Journal of Biomedical and Pharmaceutical Research
Available Online at www.jbpr.in
CODEN: - JBPRAU (Source: - American Chemical Society)
NLM (National Library of Medicine): ID: (101671502)
Index Copernicus Value 2020: 76.36
Volume 14, Issue 4, 2025, 114-152

ISSN (Online): 2279-0594
ISSN (Print): 2589-8752

Research Article

Molecular Docking and QSAR Studies of Pyrimidine Analogues as BCL2 Proteins Inhibitors and EGFR Tyrosine Kinase Inhibitors

Sulekha Mandal[1], Dilip Agrawal[2]

[1]Professor, Mahatma Gandhi College of Pharmaceutical Sciences, Sitapura, Jaipur
[2]Professor and Principal, Mahatma Gandhi College of Pharmaceutical Sciences, Jaipur

Article Info: Received: 25-06-2024 / Revised: 23-07-2025 / Accepted: 26-08-2025
Address for Correspondence: Dr. Sulekha Mandal
DOI: https://doi.org/10.32553/jbpr.v14i4.1353
Conflict of interest statement: No conflict of interest

Abstract:
The CoMFA models MMFF94 were generated from training set of 45 molecules with pIC_{50} value ranging from 3.4661 to 5.2749 using leave-one-out PLS analysis with an optimized component of 1 to give a good cross-validated correlation coefficient q^2 of 0.787, which suggest that the model should be reasonable tool for predicting the IC_{50} values. The CoMSIA model, consisting steric (S), electrostatic (E), hydrophobic (H), Donor (D) and acceptor (A) fields, can be generated using these fields in different combinations. The results of CoMSIA analysis with different combination on different charge. The CoMSIA model with a combination of steric, electrostatic, hydrophobic, donor and acceptor fields gave a good cross-validated correlation coefficient q^2 of 0.805 with an optimized component of 1. A good non-cross-validated correlation coefficient r^2 of 0.833 was attained. All the designed compounds as per QSAR, docking were subjected to evaluate from previously generated QSAR model their activity were predicted .While predicting the activity, it was found that out of thirty nine designed compounds, seven compounds were supposed to having a better activity
Keywords: CoMSIA, CoMFA, HQSAR, Docking, Pyrimidine, BCL2 proteins inhibitors, EGFR

Introduction:

The use of computational methods in the field of chemistry has increased significantly in recent years. When the properly developed mathematical method is automated on computers for the usein chemistry, then term computational chemistry is employed. Different applications of computational chemistry are possible [1]. Experimental ways may not be the easiest form of obtaining some of the properties of a molecule. Computational chemistry gives details of molecular bonding and it is not possible to get these details by experimental methods [2]. Modeling a molecular system before synthesizing that molecule is one of the computational chemistry‟s best applications. Modeling is helpful as it results in the lesser time requirement as compared to Synthesize the molecule and also reduces toxic waste which is obtained as byproduct [3]. Also, ithelps in understanding the problem in depth. Due to the advantages of computational chemistry many researchers are now using computational modeling to attain better knowledge and understanding [4]. With the implementation of computational chemistry, transformation took place in almost every field of chemistry which

has opened up various innovations in the research and development area of modern pharmaceutical chemistry [5]. Quantitative structure-activity and relationships, often simply known as QSAR, is an analytical application that can be used to interpret the quantitative relationship between the biological activities of a particular molecule and its structure [6]. It is considered a major method of chemical researching all over the world today and is frequently used in agricultural, biological, environmental, medicinal, and physical organic studies [7]. QSAR is a powerful tool for explaining the relationships between chemical structure and experimental observations. Key elements of the method are the numerical descriptors used to translate a chemical structure into mathematical variables, the quality of the observed data and the statistical methods used [8].

Materials and Methods

The biological dataset were retrieved from four series of Bcl2 protein inhibitor synthesized by Farid M. Sroor et al. (2023) and consisted of total 23 (twenty three) Bcl2 proteins inhibitors. The Bcl2 inhibitory IC_{50} values (μM; performed against P53 mutant Y220C) were converted into negative logarithmic units (pIC_{50}). The structure of compounds was listed in and doxorubicin was used as a reference, since its crystallographic structure bound to Bcl2 is available under PDB ID: 5O1H. The biological dataset were retrieved from four series EGFR-TK inhibitors synthesized by Aya I. Hassaballah et al. (2024) [9] and consisted of total 16 (Sixteen) EGFR-TK inhibitors. The structure of compounds was listed in and doxorubicin was used as a reference. The crystallographic structure bound to Bcl2 is available under PDB ID: 4ZAU [10]. Since the inhibitorsare structurally related, the training set and test set were assigned by diversity method as followed. Using this approach, the test set are spread over the whole range of activity values, and there is, at least, one test set from each range of biological activity [11]. The CoMFA and CoMSIA models were developed using 17 Compounds as training set, and externally validated using 6 as test set from the Farid M. Sroor et al. (2023) and 12 as training set and 04 as test set retrieved from Aya I. Hassaballah et al. (2024) and 16 molecule from newly designed in training set and 7 as test set were using in computational studies [12].

Structural alignment:

The molecular modelling studies were performed using SYBYL X2.0 software (Tripos) running on a core-2 duo Intel processor workstation. The molecules to be analysed were aligned on an appropriate template, which is considered to be common substructure.

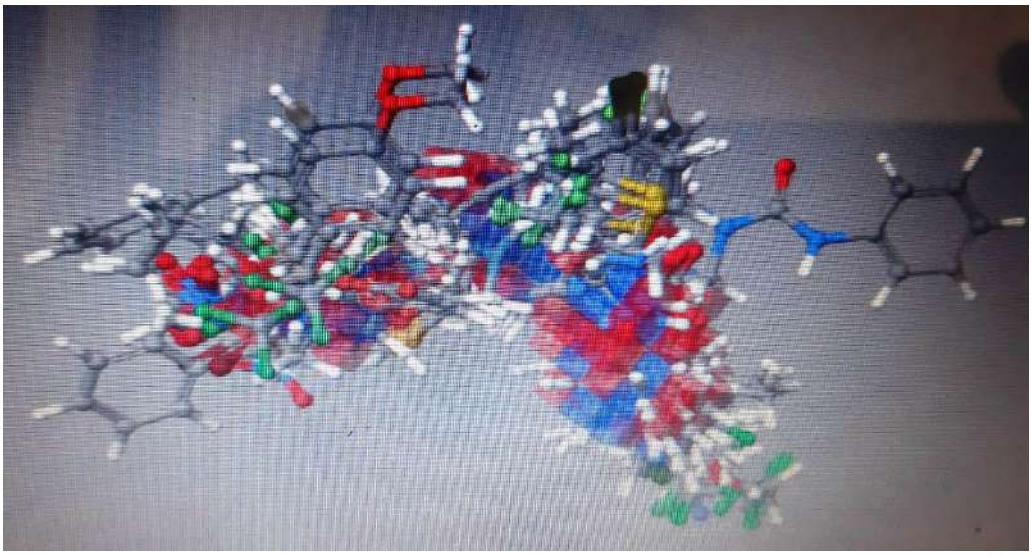


Figure 1: Alignment of all selected molecules

CoMFA

The aligned sets of molecules were positioned inside four grids boxes with grid spacing values of 1.5,2.0,2.5 and 3.0Aȋn all Cartesian directions and CoMFA fields were calculated using the QSAR modules of SYBYL. The interaction energies for each molecule were calculated at each grid point using two probe atoms: an sp^3 hybridised carbon atom with van der Waals radius of 1.52 Å and a +1.0 charge (default probe) and an sp^3 hybridised oxygen atom with a vdW radius of 1.38Å and a -1.0 charge [13]. The steric (vdW interaction) and electrostatic (Coulombic values with a 1/r distance-dependent dielectric function) fields were calculated at each intersection on the regularly spaced grid. In order to reduce noise and improved efficiency, column filtering (minimum sigma) was set to 2.0 kcal mol^{-1}, excluding from the analysis those column (lattice points) whose energy variance is less than 2.0 kcal mol^{-1}. Cuttoff for both steric and electrostatic fields were set to tentative values of 30 (defaults cuttoff), 20 and 15 kcal mol^{-1} [14].

CoMSIA

CoMSIA similarity index descriptors were derived using the same lattice boxes as those used in CoMFA calculations. Five properties, i.e., steric (S), electrostatic (E), hydrophobic (H), hydrogen bond donor (D) and hydrogen bond acceptor (A), were evaluated using a probe atom of 1.0 Å radius and +1.0 charge [15].

In CoMSIA, the steric indices are related to the third power of the atomic radii, the electrostatic descriptors are derived from atomic partial charges, the hydrophobic fields are derived from atom – based parameters developed by Vishwanath and co-workers, and the hydrogen bond donor and acceptor indices are obtained from a rule-based method derived from experimental values.

A Gussian function is used to evaluate the mutual distance between the probe atom and each molecule atom [16]. Because of the different shape of the Gussian function, the similarity indices (A) can be calculated at all grid points of the molecular surface according to equation (1).

$$A_{F.k}^{q}(j) = \sum W_{probe.k}\, W_{ik}\, e^{ar2_{iq}}$$

In Equation (1), A is the similarity index ar grid point q, summed over all atom i of the molecule j; $W_{probe.k}$ is the probe atom with radius 1.0Å, charge +1.0, hydrophobicity +1, HBA and HBD +1; W_{ik} is the actual value of the physicochemical property k of atom I; r_{iq} is the mutual distance between the probe atom at grid point q and atom I of the test molecule. Alfa (α) is the attenuation factor, with a default value of 0.3, and an optimal value normally ranging from 0.2-0.4[14, 16]

HQSAR

HQSAR is a new 2D-QSAR technique which employs specialized fragment fingerprints as predictive variables of biological activity. HQSAR does not require 3D alignment for model generation and is sensitive to three parameters concerning hologram generation, including hologram length, fragment size, and fragment distinction. The fragment distinct are atoms (A), bonds (B), connections (C), hydrogen atom (H), chirality (Ch), and donor (D). Initially, various models were developed by using the default fragment size of 4-7 and different component, then based on the different fragment distinction determined by the first step, the models were developed using different sizes [17]. The models with better results were applied to different fragment size and component number (Jae et al., 2008). The better statistical results were obtained in fragment size 4-7 and 2-6, A/B/C and A/B/C/Ch distinct and number of component 6 [18].

Partial least square and predictive r^2 analysis.

In order to generate statistically significant 3D-QSAR models, PLS regression was used to analyse the training set by correlating the variation in their pIC_{50} values (the dependent variable) with variations in their CoMFA/CoMSIA interaction fields (the independent variable) [19].

Docking Analysis

Molecular docking studies were carried out using the Schrödinger Maestro version 2016. The protein structure of pdb name along with their inhibitors was retrieved from RCSB Protein Data Bank (PDB entry code: 5O1H and 4ZAU). The protein structures were subjected to energy minimization and charge calculation (MMFF94). After that the known complex protein structure was used to investigate and validate the docking protocol.

All ligand and water molecules were removed. The bloat values was set as 1 and the threshold values as 0.5 for generation of protomol and position was considered to be the active sites for potential receptor's binding sites [20].

Pharmacophore Mapping

Genetic algorithm with linear assignment of hypermolecular alignment of datasets (GALAHAD) was used to generate the pharmacophore models. All the s in the training set were prepared by the following procedures; the structures were checked for bond orders, hydrogen atoms were added and minimization procedures was implemented using the MMFF94, force-field GALAHAD was run for 60 generation with a population size of 100. The rest of the parameters were set as default values [21]. The generated models were evaluated by a test database; several parameters were employed for model evaluation.

Designing of Compounds: On the basis of reported structure activity relationship of pyrimidine analogues as an α – oxidosidase inhibitor ,QSAR studies using CoMFA, CoMSIA, HQSAR, and Molecular modelling (Docking) studies , one hundred and two compounds were designed. On the designed compounds, further Computational QSAR studies CoMFA, CoMSIA, HQSAR and Molecular modelling Docking was done in order to select the best compounds for synthesis [22].

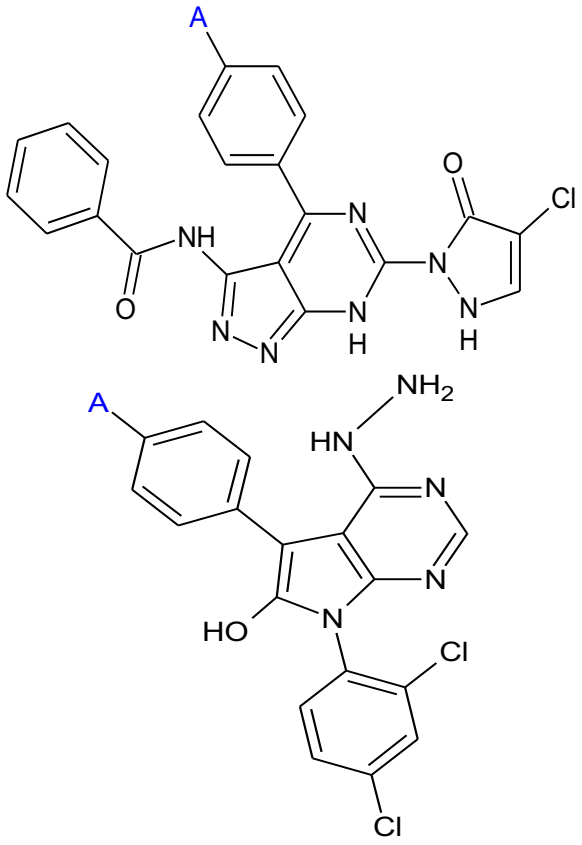
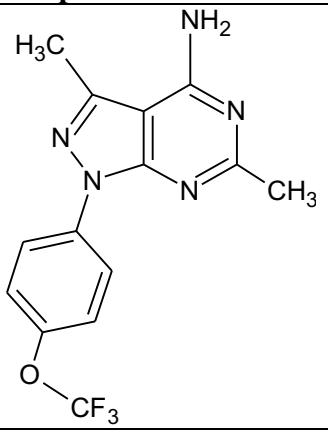
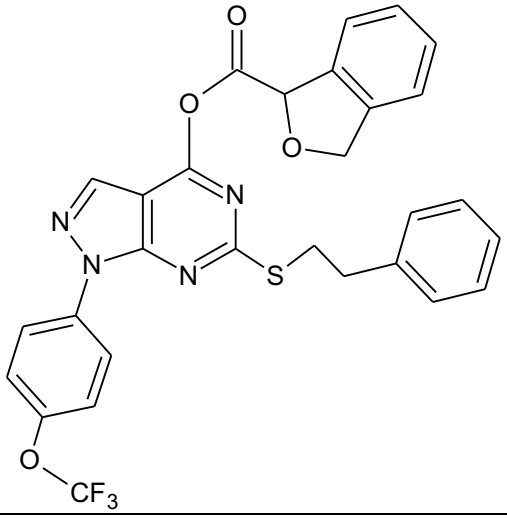
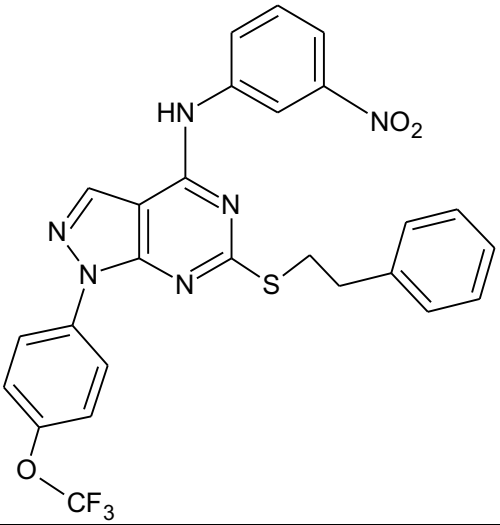
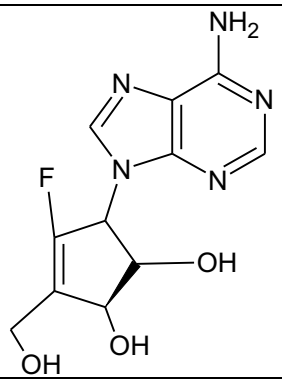
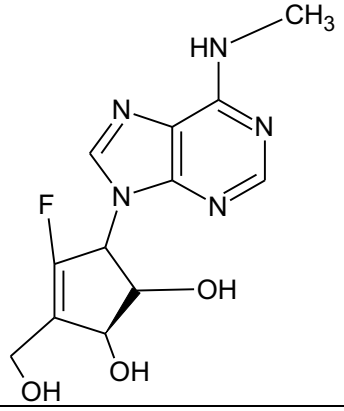
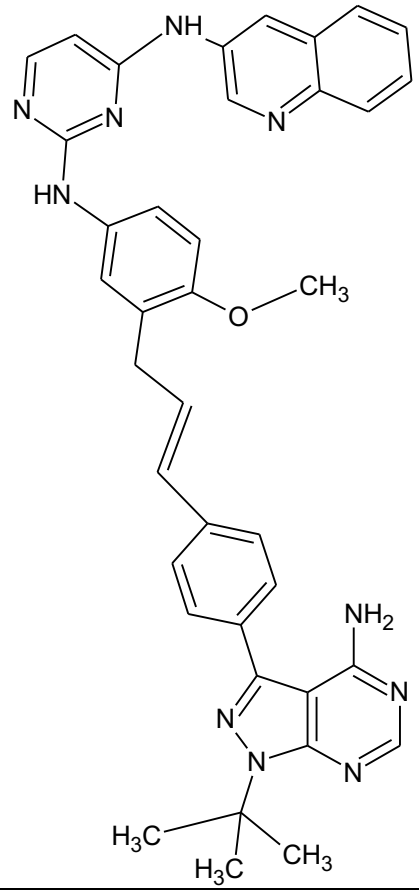
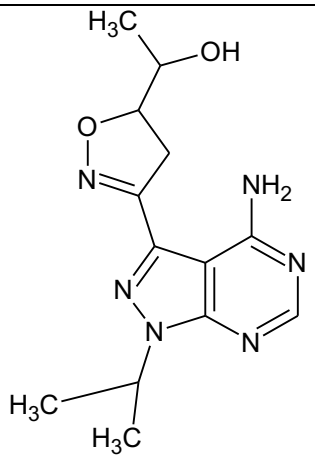
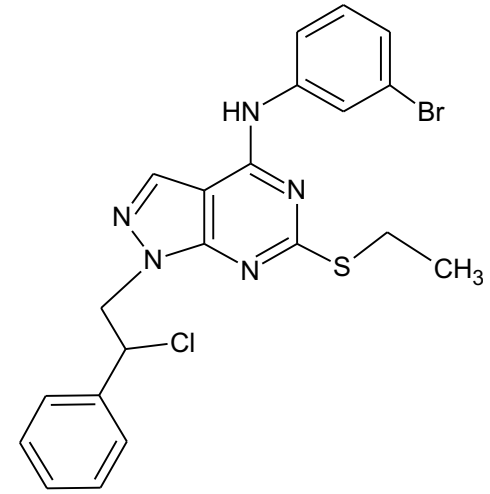
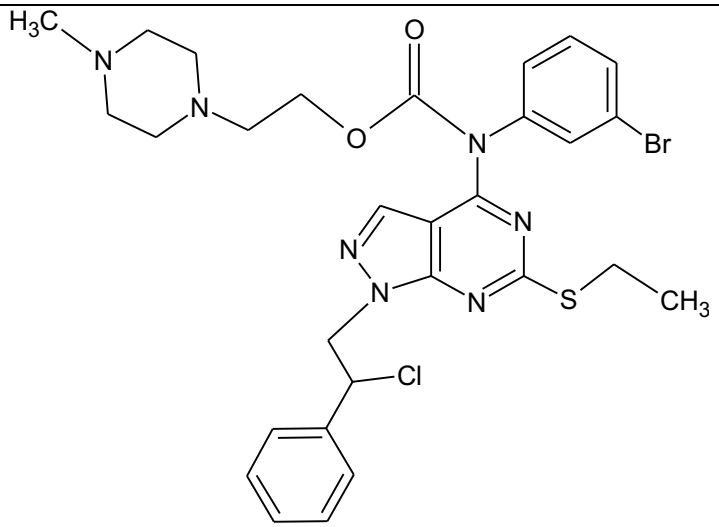


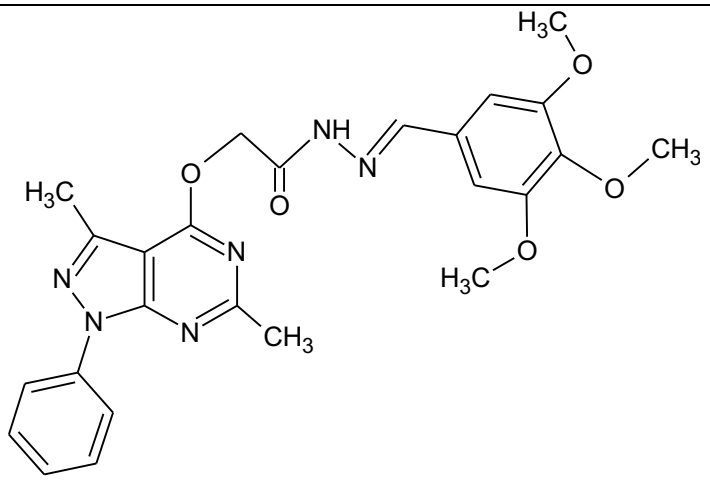
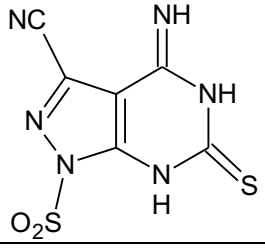
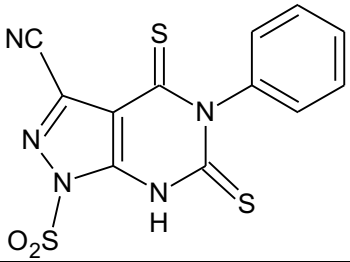
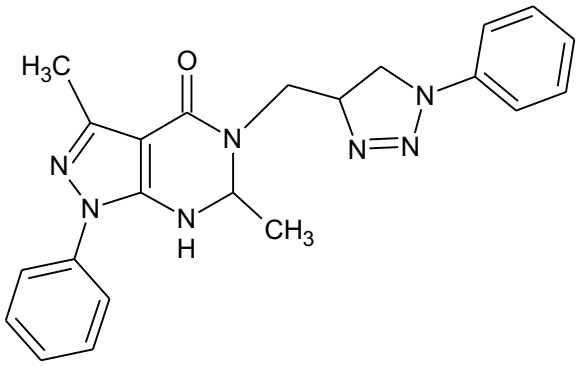
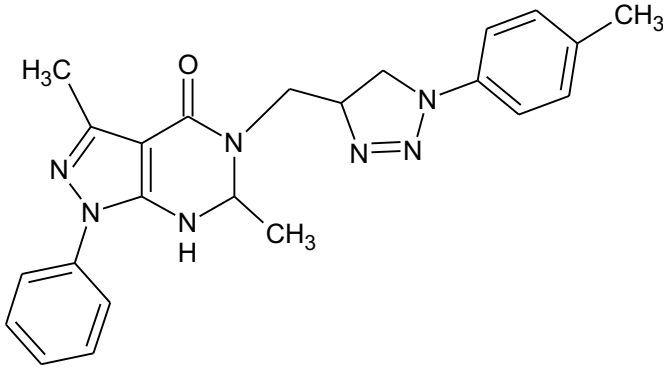
Figure 2: Target compounds Structure

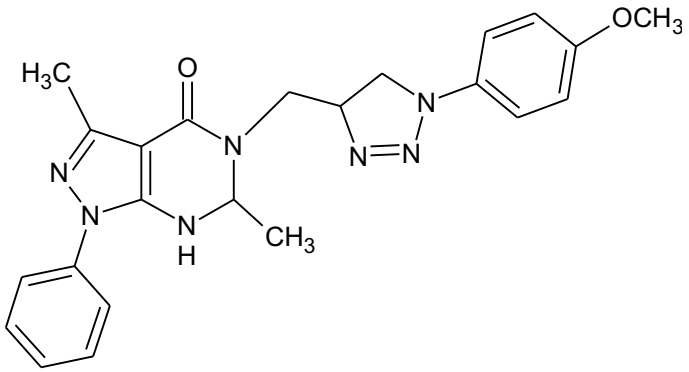
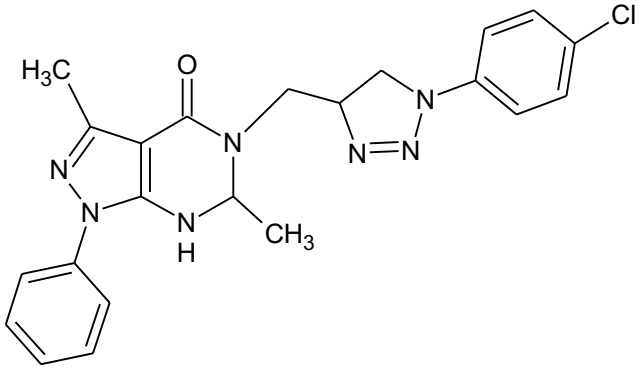
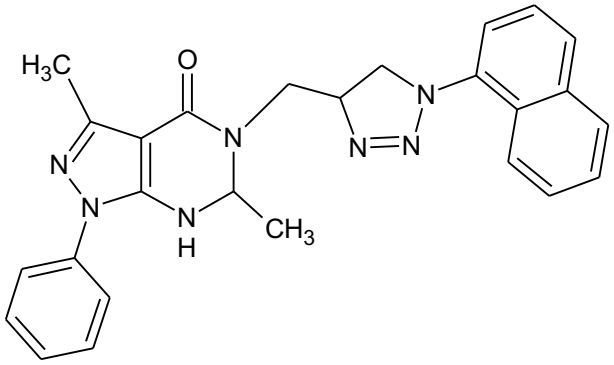
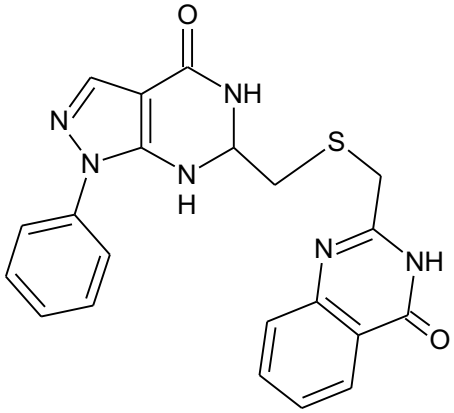
Table 1: Structure of Designed Pyrimidine analogues

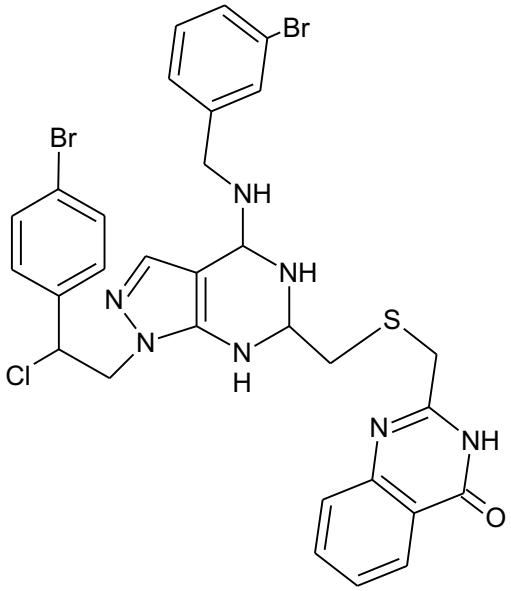
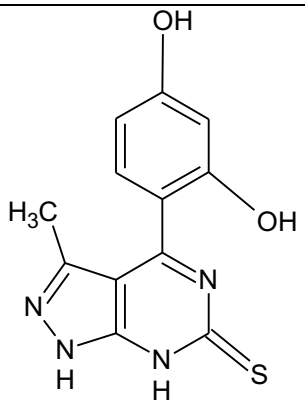
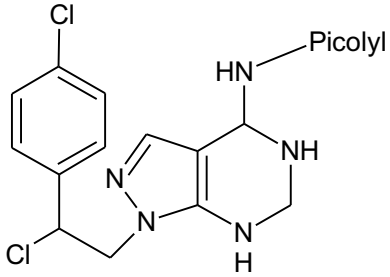
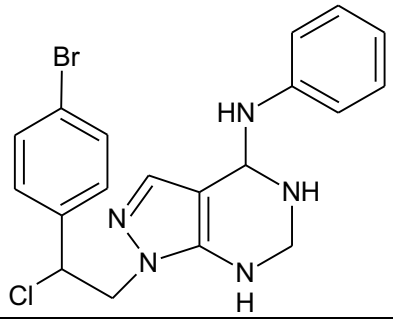
S. No.	Compound Structure
1	
2	
3	

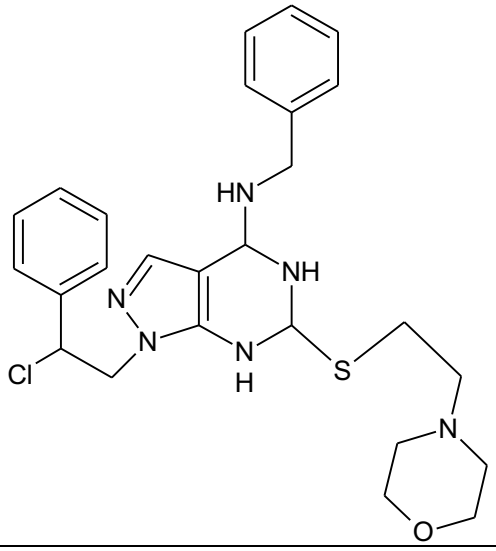
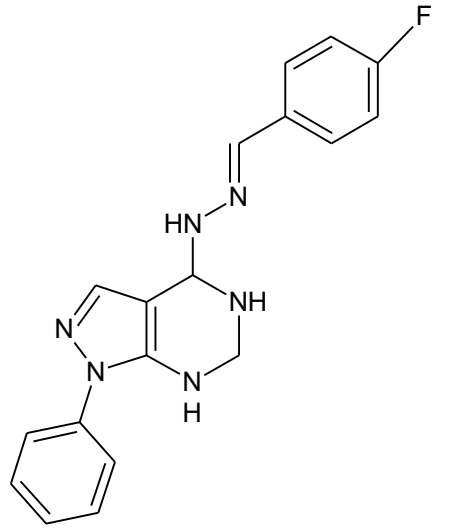
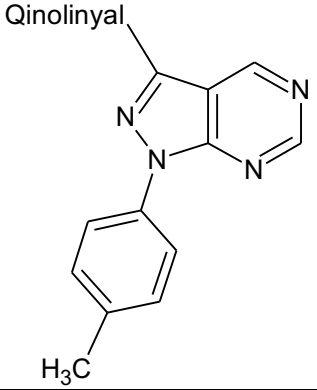
4	
5	
6	

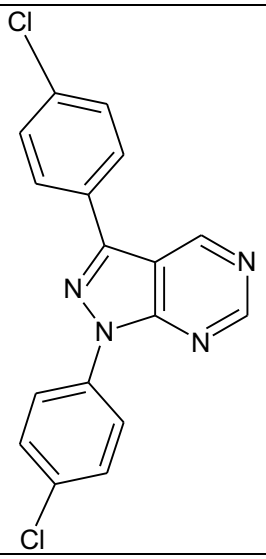
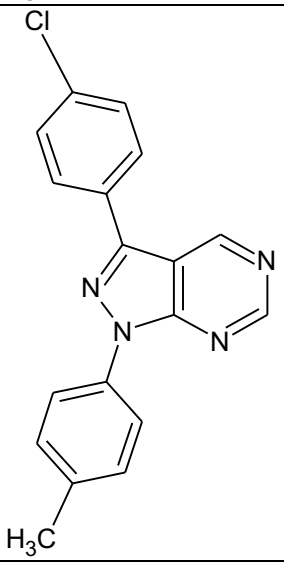
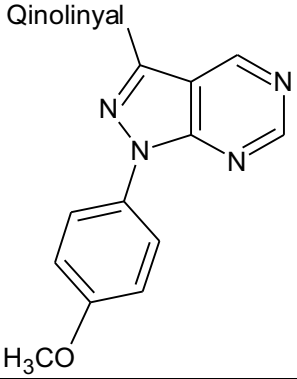
7	
8	
9	

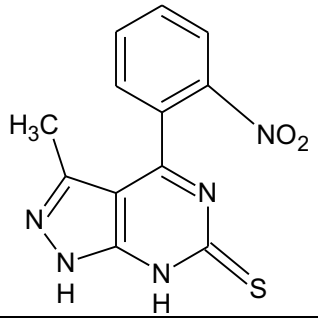
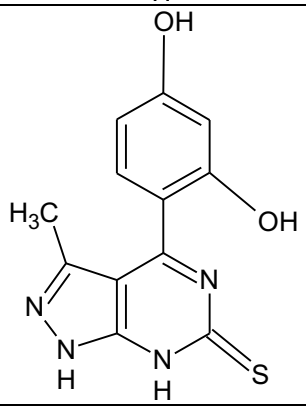
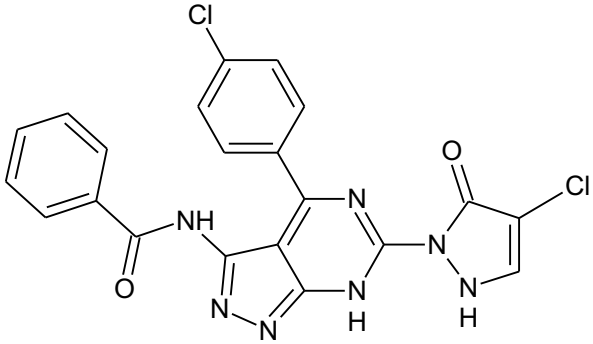
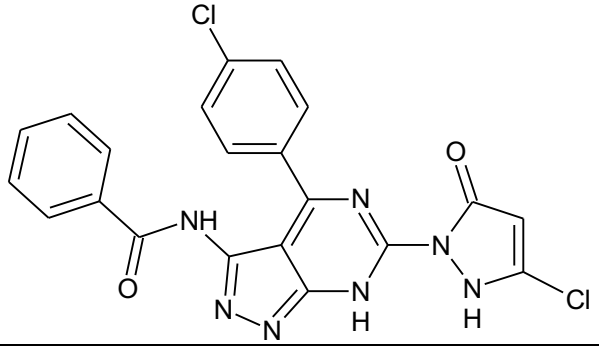
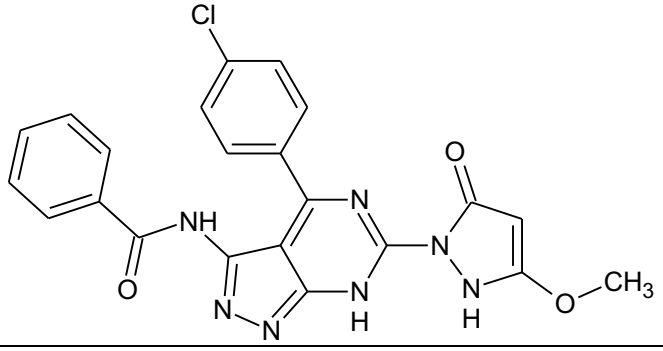
10	
11	
12	
13	
14	

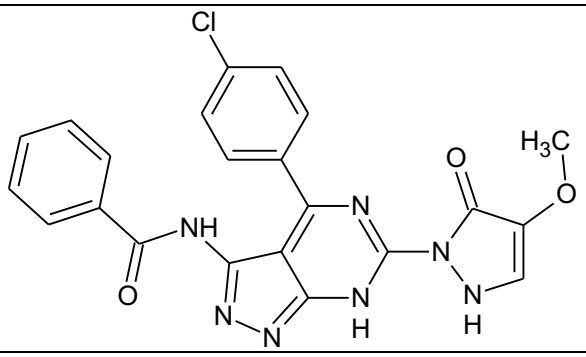
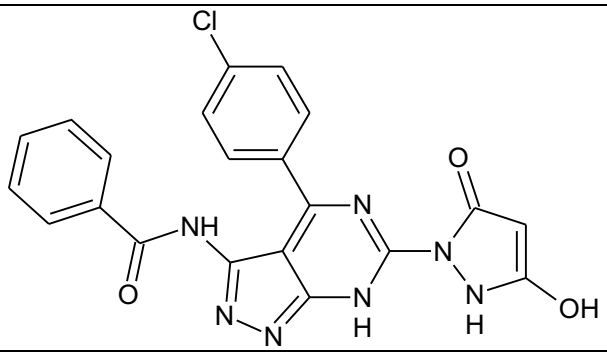
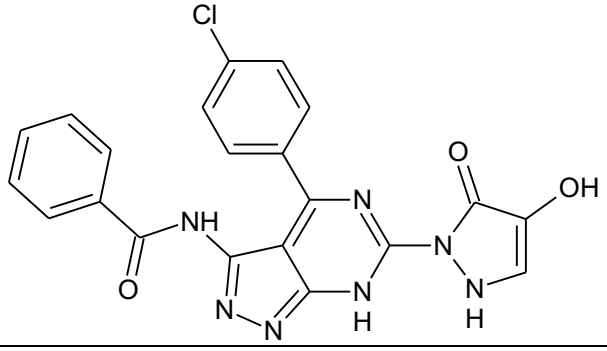
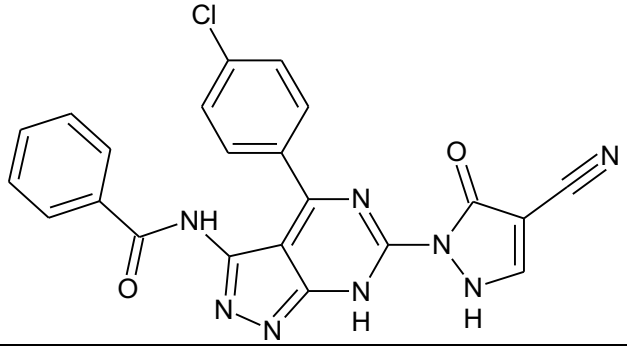
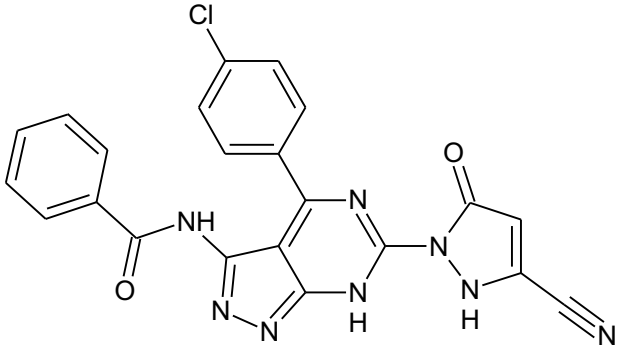
15	
16	
17	
18	

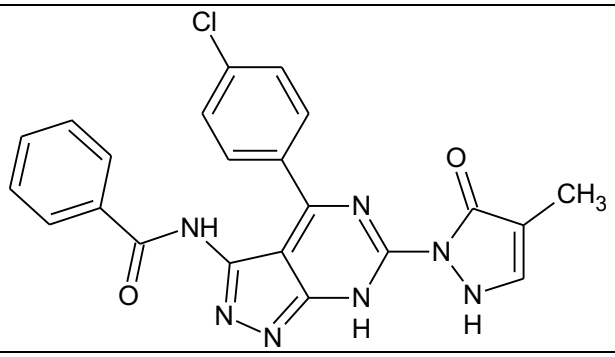
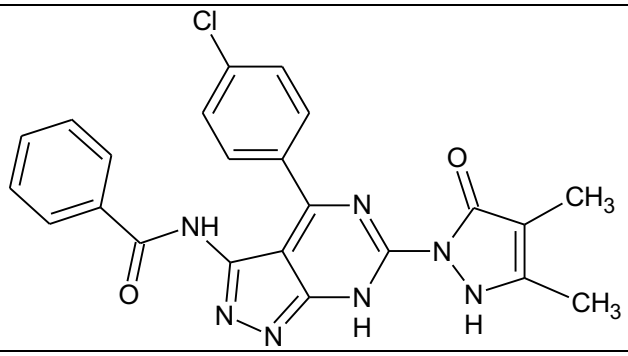
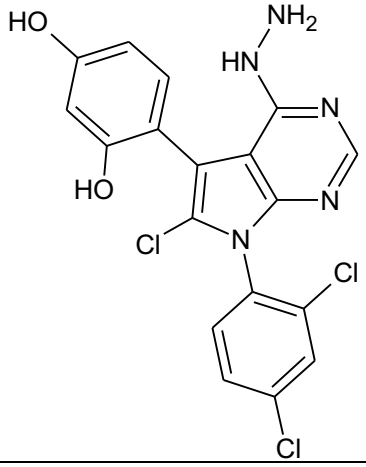
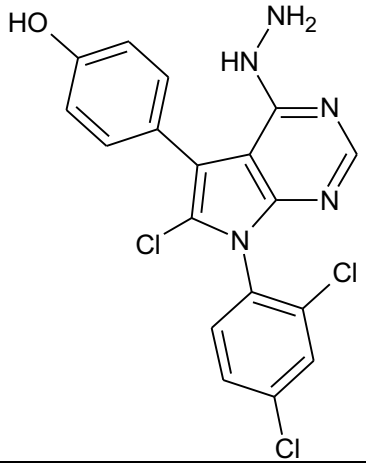
19	
20	 CH_4
21	
22	

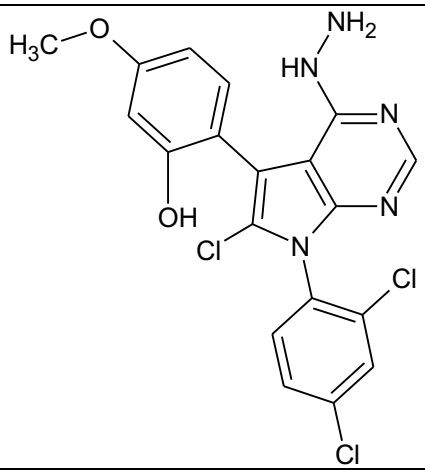
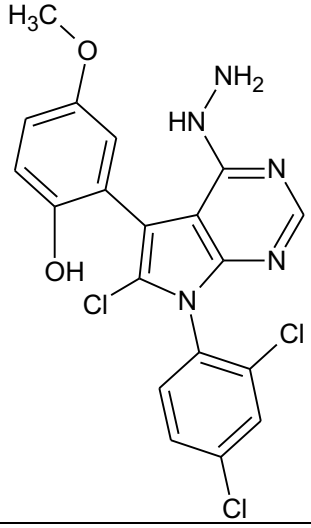
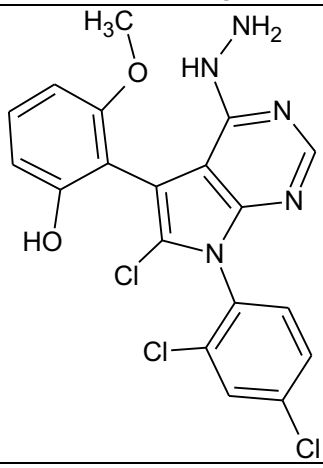
23	
24	
25	

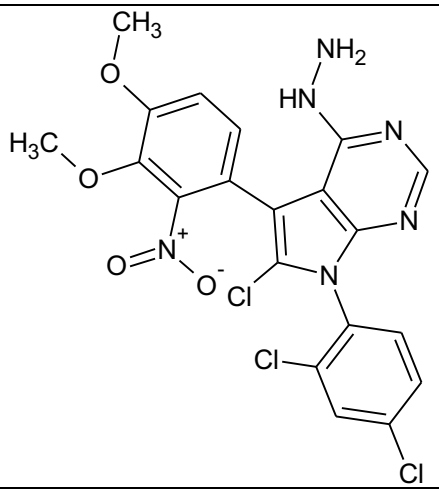
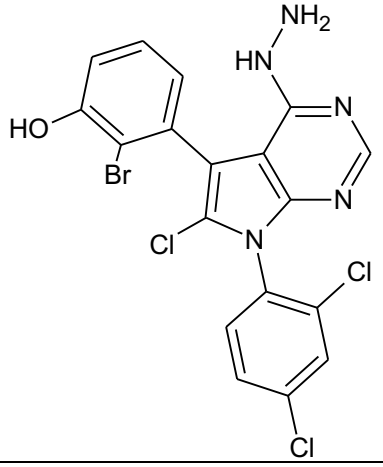
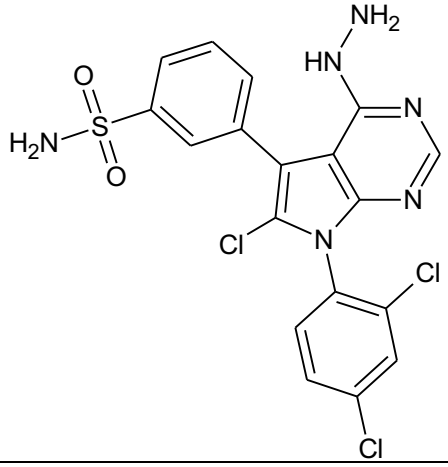
26	
27	
28	

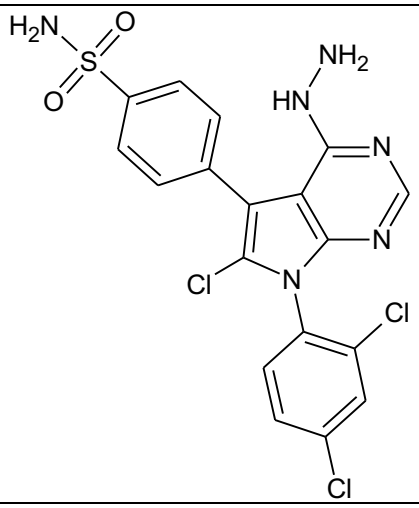
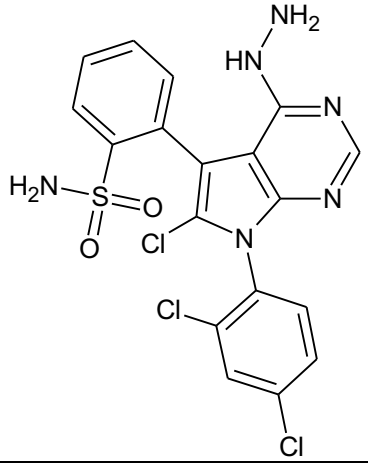
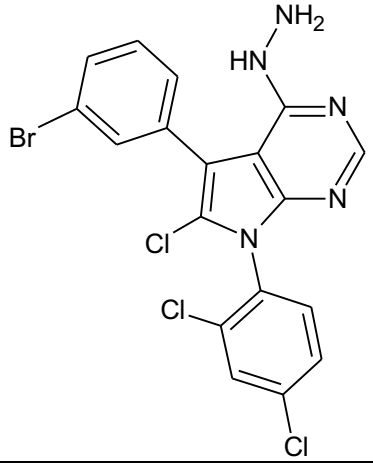
29	
30	
31	
32	
33	

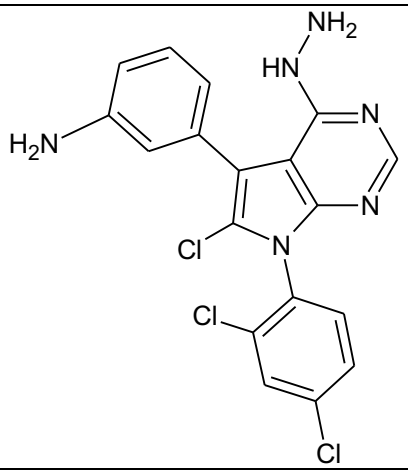
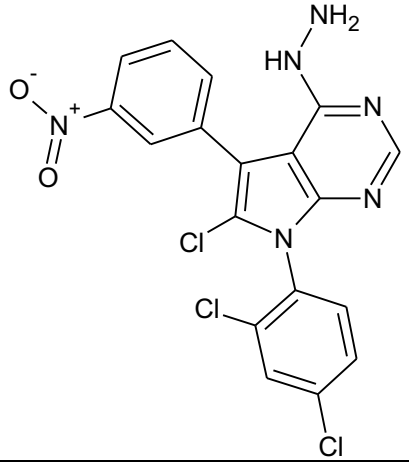
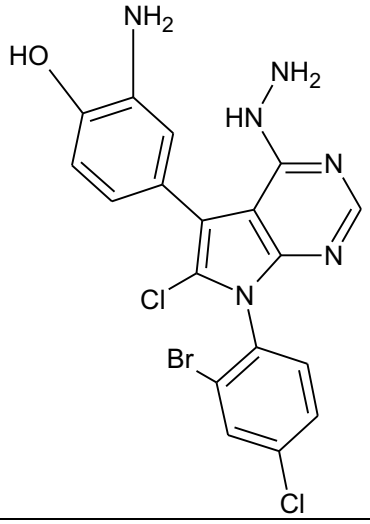
34	
35	
36	
37	
38	

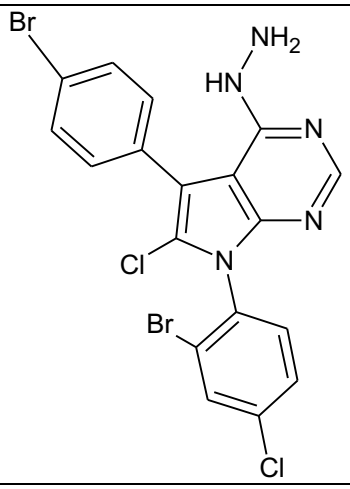
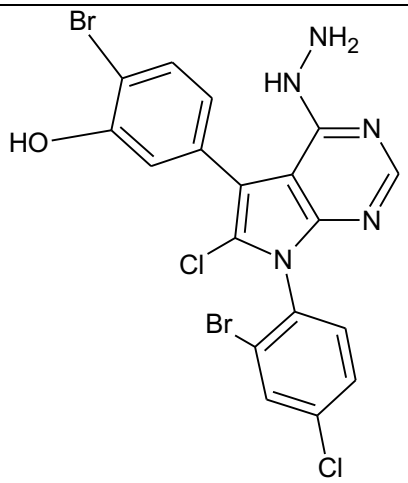
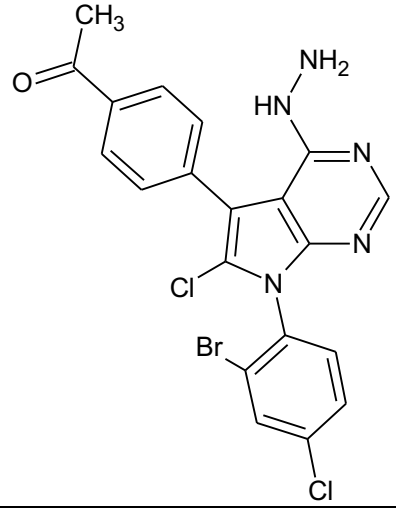
39	
40	
41	
42	

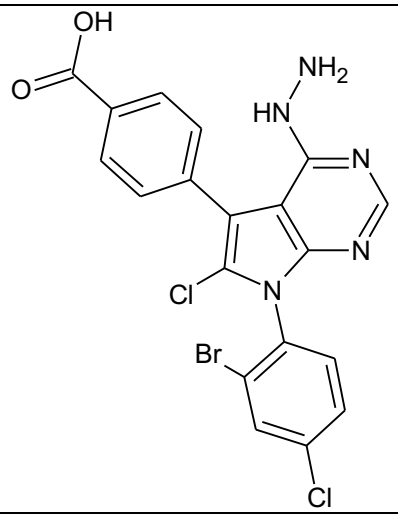
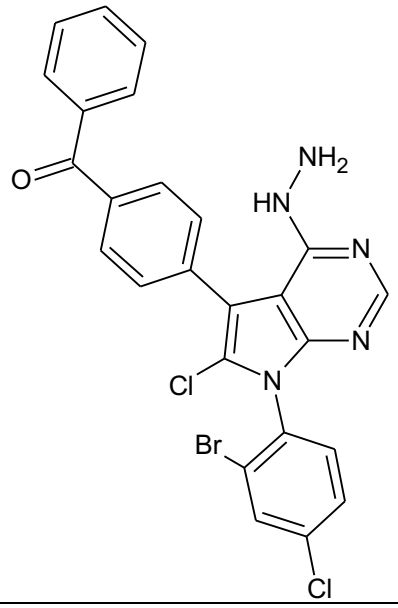
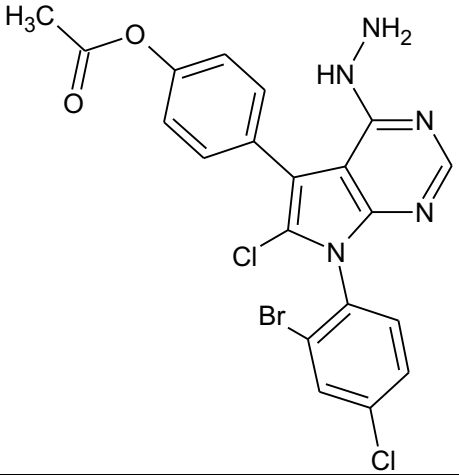
43	
44	
45	

46	
47	
48	

49	
50	
51	

52	
53	
54	

55	
56	
57	

58	
59	
60	

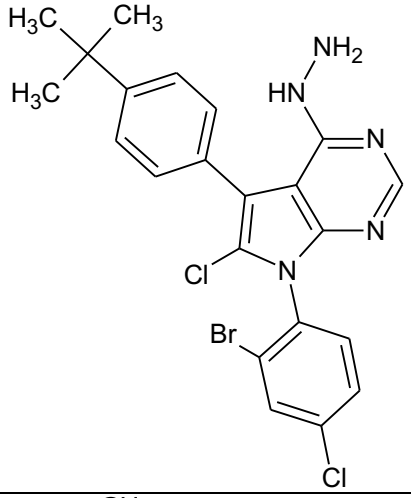
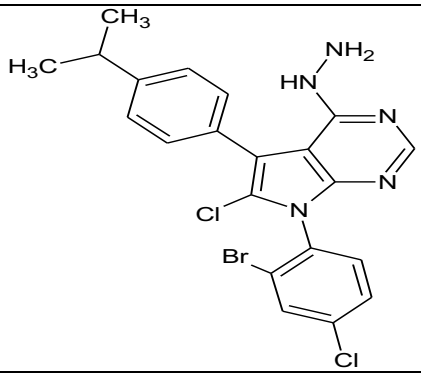
61	
62	

Table 2: Designed Pyrimidine analogues on the basis of computational studies with their predicted data

Compound	Pred pIC_{50}			
	CoMFA	**CoMSIA**	**HQSAR**	**Docking Score**
1	4.3521	4.4758	4.282	4.5033
2	4.3484	4.4751	5.03	3.8241
3	4.3438	4.4782	4.328	3.6918
4	4.3534	4.4803	3.836	5.3139
5	4.3477	4.4731	4.696	5.4296
6	4.3456	4.4747	4.915	3.0611
7	4.3488	4.4753	4.578	4.3919
8	4.3493	4.4782	4.447	4.9245
9	4.3486	4.4808	4.425	2.9629
10	4.3502	4.4784	4.756	3.8114
11	4.3511	4.4776	4.503	2.3645
12	4.3450	4.4755	4.518	5.3367
13	4.3493	4.4778	4.341	5.9438
14	4.3443	4.4748	4.594	4.8035
15	4.3536	4.4984	5.123	6.0693
16	4.3456	4.4717	4.597	2.7969
17	4.3462	4.4732	4.725	4.5086
18	4.3469	4.4785	4.691	4.7630
19	4.3495	4.4753	4.823	3.3900

20	4.3480	4.4792	4.941	3.0896
21	4.3461	4.4752	4.647	2.8599
22	4.3492	4.4802	4.518	4.9646
23	4.3465	4.4702	4.866	6.6412
24	4.3452	4.4737	4.826	4.7953
25	4.3467	4.4783	4.742	5.2871
26	4.3453	4.4697	4.884	6.3892
27	4.3514	4.791	4.373	8.2615
28	4.3473	4.4783	4.293	5.4088
29	4.3470	4.4761	4.939	6.3040
30	4.3431	4.4754	4.775	4.8325
31	4.3454	4.4758	4.555	4.9260
32	4.3501	4.4743	4.859	6.5207
33	4.3520	4.4804	4.182	2.8665
34	4.3474	4.4771	4.624	1.5708
35	4.3503	4.4781	4.378	5.4090
36	4.3459	4.4740	4.876	5.8507
37	4.3541	4.4793	4.745	2.9364
38	4.3471	4.4747	4.796	6.6970
39	4.3451	4.4759	5.016	3.3043
40	4.3460	4.4796	4.654	3.7888
41	4.3508	4.4800	4.947	5.4434
42	4.3467	4.4778	4.883	5.0204
43	4.3460	4.4735	4.876	2.9344
44	4.3436	4.4780	4.596	5.1306
45	4.3466	4.4739	4.721	7.1182
46	4.3453	4.4722	4.774	5.8317
47	4.3496	4.4784	4.341	4.7399
48	4.3457	4.4786	4.138	6.8212
49	4.3492	4.4780	4.884	8.3985
50	4.3506	4.4769	4.617	8.8636
51	4.3534	4.4794	4.872	5.3914
52	4.3452	4.4825	4.118	4.6358
53	4.3407	4.4767	4.721	4.1250
54	4.3337	4.4715	4.019	4.3931
55	4.3493	4.4810	3.527	4.3762
56	4.3333	4.4741	4.387	3.6550
57	4.3369	4.4748	4.606	3.5770
58	4.3357	4.4788	4.269	4.6806
59	4.3387	4.4799	4.137	5.0128
60	4.3391	4.4841	4.116	3.0643
61	4.3340	4.4757	4.447	3.5384
62	4.3432	4.4760	4.193	4.3641

***39 compound from data base and #23 compounds (compounds number 40 to 62) as newly substituted as per literature**

Result and Discussion

CoMFA and CoMSIA Results:

The results of CoMFA analysis with combination of steric and electrostatic on different charge are summarized. The statistical parameters corresponding to the CoMFA model are listed. The CoMFA models MMFF94 were generated from training set of 45 molecules with pIC_{50} value ranging from 3.4661 to 5.2749 using leave-one-out PLS analysis with an optimized component of 1 to give a good cross-validated correlation coefficient q^2 of 0.787, which suggest that the model should be reasonable tool for predicting the IC_{50} values. A high non-cross-validated correlation coefficient r^2 of 0.819 with a low standard error estimation (SEE) of 0.041 was obtained as well as an F value of 1316.074 and predictive correlation coefficient r^2_{pred} of 0.996. Contributions of steric and electrostatic fields were 0.507 and 0.493, respectively. The actual and predicted pIC_{50} values of the training and test set by the model 7 (MMFF94 charge) are listed in table. The graph of actual versus predicted pIC_{50} of the training and test set of model 29. The CoMSIA model, consisting steric (S), electrostatic (E), hydrophobic (H), Donor (D) and acceptor (A) fields, can be generated using these fields in different combinations. Among the combination models, steric, electrostatic, acceptor, donor and hydrophobic fields played an essential roles for the present series of compounds. Removal of any descriptors results in significant reduction in r^2, q^2 and r^2_{pred}, which implies that all descriptors all descriptors are crucial and the steric, electrostatic, hydrophobic, donor and acceptor functional groups were of extreme significance for the inhibitory activity. In conclusion, the combination of steric, electrostatic, hydrophobic, donor and acceptor was selected as the best model. The CoMSIA model with a combination of steric, electrostatic, hydrophobic, donor and acceptor fields gave a good cross-validated correlation coefficient q^2 of 0.805 with an optimized component of 1. A good non-cross-validated correlation coefficient r^2 of 0.833 was attained, as well as very low standard error estimation SEE of 0.065, F value of 520.302 and predicted correlation coefficient r^2_{pred} of 0.990. Contribution of steric, electrostatic, hydrophobic, donor and acceptor fields were 0.151, 0.268, 0.223, 0.234, 0.124 respectively. The actual and predicted pIC_{50} values and residual value for the training and test set compounds of model 29 (MMFF94). The association between actual and predicted pIC_{50} of training and test set compounds of model 29.

Table 3: CoMFA on different charge

S.no	Model	q^2	r^2	SE	NC
1	Model 1 Gastegier	0.775	0.816	0.265	1
2	Model 2 G-H	**0.786**	**0.819**	**0.262**	**1**
3	Model 3 Delre	0.771	0.814	0.266	1
4	Model 4 Pullman	0.780	0.814	0.266	1
5	Model 5 Formal Charge	0.779	0.812	0.266	1
6	Model 6 MMFF94	**0.787**	**0.819**	**0.262**	**1**

Table 4: CoMFA on MMFF94 charge

Sno	Name	q^2	r^2	SE	NC
1	Model 7 clogP	0.777	0.834	0.255	2
2	Model 8 CMR	0.773	0.826	0.261	2
3	Model 9 CPSA	0.775	0.823	0.264	2
4	Model 10 D-M	0.769	0.836	0.254	2
5	Model 11 MP-Area	0.773	0.825	0.262	2
6	Model 12 MP-PSA	0.778	0.826	0.261	2
7	Model 13 MP-PV	0.773	0.825	0.262	2

8	Model 14 MP-vol	0.771	0.826	0.261	2
9	Model15 Mol-Wt	0.777	0.829	0.259	2
10	Model 16 Acceptor	0.775	0.828	0.260	2
11	Model 17 Atom-Count	0.776	0.825	0.262	2
12	Model 18 Bond-Count	0.777	0.825	0.262	2
13	Model 19 Chiral	0.787	0.819	0.262	1
14	Model 20 Donor	0.779	0.826	0.261	2
15	Model 21 Hydrophobe	0.777	0.828	0.259	2
16	Model 22 Ring Count	0.774	0.834	0.255	2
17	Model 23 Rot-Bond	0.754	0.829	0.259	2

Table 5: Residual values of Training set of molecules of the CoMFA model 7.

S No.	IC_{50}	pIC_{50}	Predicted pIC50	Residual Valu0e
1*	6.73	3.6579	3.8016	-0.1437
2	5.21	4.3639	3.8789	0.485
3	4.11	3.6213	3.8078	-0.1865
4*	3.23	3.4117	4.0432	-0.6315
5	8.32	3.6873	3.8829	-0.1956
6	5.63	3.7563	3.9741	-0.2178
7	8.23	4.4347	3.8897	0.545
8*	7.23	4.7392	4.0403	0.6989
9	6.52	4.1642	3.9498	0.2144
10	3.91	3.4661	3.9013	-0.4352
11	5.40	3.7534	3.9112	-0.1578
12*	4.41	3.3722	3.7974	-0.4252
13	9.70	3.7173	3.8383	-0.121
14	1.81	3.8362	3.8785	-0.0423
15	3.55	3.7132	3.8534	-0.1402
16*	3.16	4.0037	4.016	-0.0123
17	6.32	3.909	3.9062	0.0028
18	5.43	4.2563	3.8841	0.3722
19	6.40	3.6453	3.8988	-0.2535
20	7.72	3.7552	3.9677	-0.2125
21	4.39	4.3336	4.0258	0.3078
22	6.50	5.1871	4.9646	0.2225
23	8.23	4.9901	4.9235	0.0666
24	6.29	4.9473	5.0454	-0.0981
25	8.48	5.0716	4.9601	0.1115
26	9.22	4.95	5.0041	-0.0541
27	6.97	5.1568	5.027	0.1298
28	5.55	5.2557	5.0047	0.251
29*	5.75	4.8945	4.9155	-0.021
30	7.09	4.8213	4.9923	-0.171
31	4.58	5.2534	5.0292	0.2242
32*	3.38	4.5787	4.983	-0.4043
33	7.12	5.1475	4.9941	0.1534

34	3.17	4.7913	5.0014	-0.2101
35	8.05	5.0942	5.0036	0.0906
36*	8.02	4.5525	4.9661	-0.4136
37	7.33	4.7368	5.0323	-0.2955
38	6.37	5.0773	5.0028	0.0745
39	8.07	5.0931	4.9718	0.1213
40	5.31	5.2749	4.968	0.3069
41	2.09	4.9551	4.9889	-0.0338
42	9.12	5.0357	4.9485	0.0872
43	5.34	4.2729	5.0297	-0.7568
44*	6.8	4.3487	4.9767	-0.628
45	8.85	4.9263	5.0492	-0.1229

Table 6: CoMSIA on training set at different charges at MMFF94 charge

Sno	Name	q^2	r^2	SE	NC
1	Model 24 S	0.784	0.813	0.266	1
2	Model 25 SE	0.794	0.825	0.258	1
3	Model 26 SHE	0.805	0.832	0.252	1
4	Model 27 SEHD	0.803	0.830	0.254	1
5	**Model 28 SEHDA**	**0.805**	**0.831**	**0.253**	**1**

Table 7: CoMSIA with MMFF94 Charge

Sno	Model	q^2	r^2	SE	NC
1	Model 29 clogP	0.800	0.854	0.239	2
2	Model 30 CMR	0.793	0.845	0.247	2
3	Model 31 CPSA	0.792	0.838	0.252	2
4	Model 32 DM	0.779	0.855	0.238	2
5	Model 33 MP_Area	0.791	0.843	0.248	2
6	Model 34 MP_PSA	0.800	0.846	0.245	2
7	Model 35 MP_PV	0.794	0.844	0.247	2
8	Model 36 MP_Vol	0.790	0.845	0.247	2
9	Model 37 Mol-Wt	0.795	0.848	0.244	2
10	Model 38 Atom Count	0.788	0.844	0.247	2
11	Model 39 Bond Count	0.792	0.843	0.248	2
12	Model 40 Chiral	0.805	0.831	0.253	1
13	Model 41 Ring Count	0.796	0.853	0.240	2
14	Model 42 RotBonds	0.783	0.848	0.244	2

Table 8: Residual values of Training set of molecules of the CoMFA model 29.

S. No.	IC_{50}	pIC_{50}	Predicted pIC_{50}	Residual Value
1*	6.73	3.6579	4.3401	-0.6822
2	5.21	4.3639	3.9408	0.4231
3	4.11	3.6213	3.8592	-0.2379
4*	3.23	3.4117	4.365	-0.9533
5	8.32	3.6873	3.8546	-0.1673
6	5.63	3.7563	3.9542	-0.1979

7	8.23	4.4347	3.8808	0.5539
8*	7.23	4.7392	4.3639	0.3753
9	6.52	4.1642	3.9218	0.2424
10	3.91	3.4661	3.8709	-0.4048
11	5.40	3.7534	3.9132	-0.1598
12*	4.41	3.3722	4.34	-0.9678
13	9.70	3.7173	3.8452	-0.1279
14	1.81	3.8362	3.8613	-0.0251
15	3.55	3.7132	3.8469	-0.1337
16*	3.16	4.0037	4.3554	-0.3517
17	6.32	3.909	3.8919	0.0171
18	5.43	4.2563	3.8721	0.3842
19	6.40	3.6453	3.9065	-0.2612
20	7.72	3.7552	3.9649	-0.2097
21	4.39	4.3336	3.9883	0.3453
22	6.50	5.1871	4.9881	0.199
23	8.23	4.9901	4.9627	0.0274
24	6.29	4.9473	5.0087	-0.0614
25	8.48	5.0716	4.9353	0.1363
26	9.22	4.95	5.0478	-0.0978
27	6.97	5.1568	5.0269	0.1299
28	5.55	5.2557	5.0385	0.2172
29*	5.75	4.8945	4.6344	0.2601
30	7.09	4.8213	5.0305	-0.2092
31	4.58	5.2534	5.0674	0.186
32*	3.38	4.5787	4.6372	-0.0585
33	7.12	5.1475	4.9695	0.178
34	3.17	4.7913	4.9662	-0.1749
35	8.05	5.0942	4.9759	0.1183
36*	8.02	4.5525	4.6396	-0.0871
37	7.33	4.7368	4.9692	-0.2324
38	6.37	5.0773	5.064	0.0133
39	8.07	5.0931	4.9936	0.0995
40	5.31	5.2749	4.9747	0.3002
41	2.09	4.9551	4.9681	-0.013
42	9.12	5.0357	4.9865	0.0492
43	5.34	4.2729	4.9916	-0.7187
44	6.8	4.3487	4.6592	-0.3105
45	8.85	4.9263	5.0445	-0.1182

Table 9: Results of CoMFA and CoMSIA Models

PLS Statistics	**CoMFA**	**CoMSIA**
q^2	0.787	0.805
r^2	0.819	0.831
SEE	0.041	0.065
F value	1316.074	520.302

r^2_{pred}	0.996	0.990
NC	1	1
	Field contribution	
Steric	0.507	0.151
Electrostatic	0.493	0.268
Hydrophobic	-	0.223
Donor	-	0.234
Acceptor	-	0.124

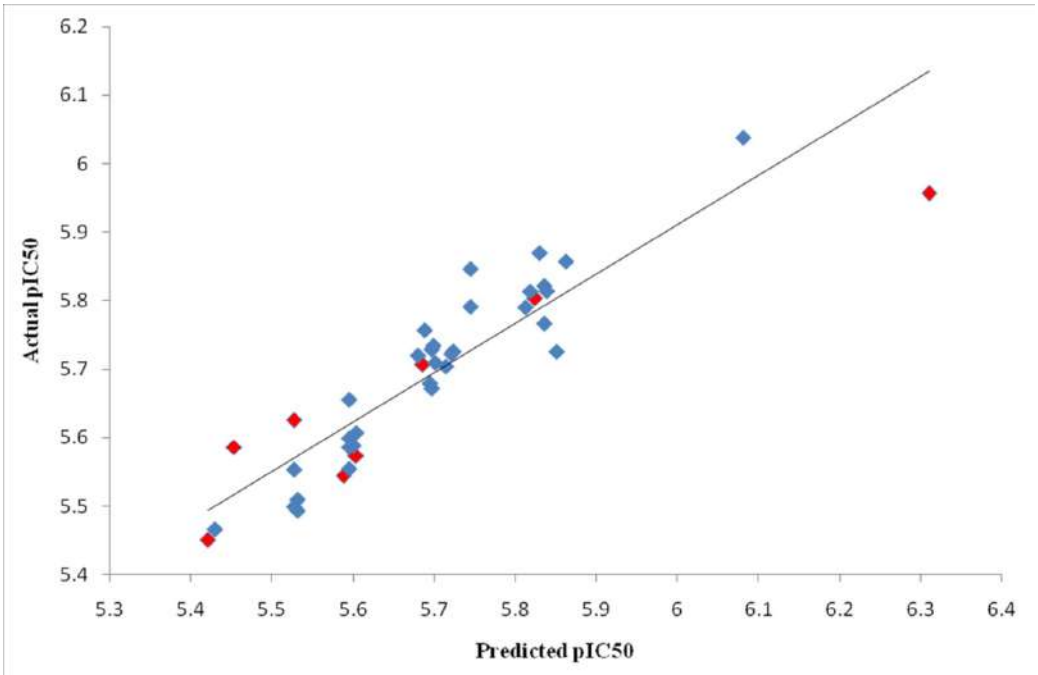


Figure 3: Graph of actual versus predicted pIC_{50} values of the training set and the test set molecules of Model 7 (MMFF94) using the CoMFA model.

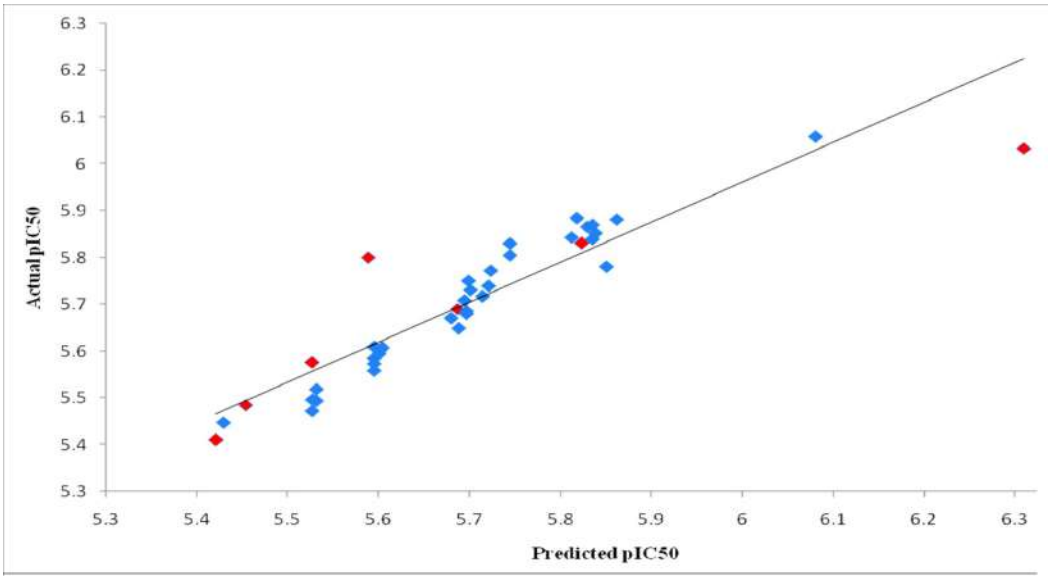


Figure 4: Graph of actual versus predicted pIC_{50} values of the training set and the test set molecules of Model 29 (MMFF94) using the CoMSIA model.

CoMFA contour Maps:

To view the information of the resultant 3D-QSAR model, CoMFA contour maps were generated to rationalize the region in 3D space around the molecules where changes in the steric and electrostatic fields were predicted to enhance or lessen the activity of the compound.

The steric fields is characterised by green and yellow contour, in which yellow contour is unfavourable for the introduction of bulky group, while the green contour is favourable for the introduction of the bulky groups. Compounds 40 and 28 were selected as a reference structure. The 4th position of primidine ring, over hydrazine hydarate group, 3rd position of phenyl ring attached to hydroxyl and methoxy substituted primidine ring and phenyl ring at R_1 position was surrounded by green contour indicates the need of bulky group for increasing the Bcl2 inhibitors activity. which exhibit the regions where electron-donating group and electron-accepting groups would be favourable respectively. Compound 40 and 28 was selected as a reference molecule. In the CoMFA electrostatic fields, blue contour at 2nd position of imidazole ring and 2nd and 4rd position of phenyl ring attached to substituted imidazole ring shows the electron-donating substitution was essential for the inhibitory activity.

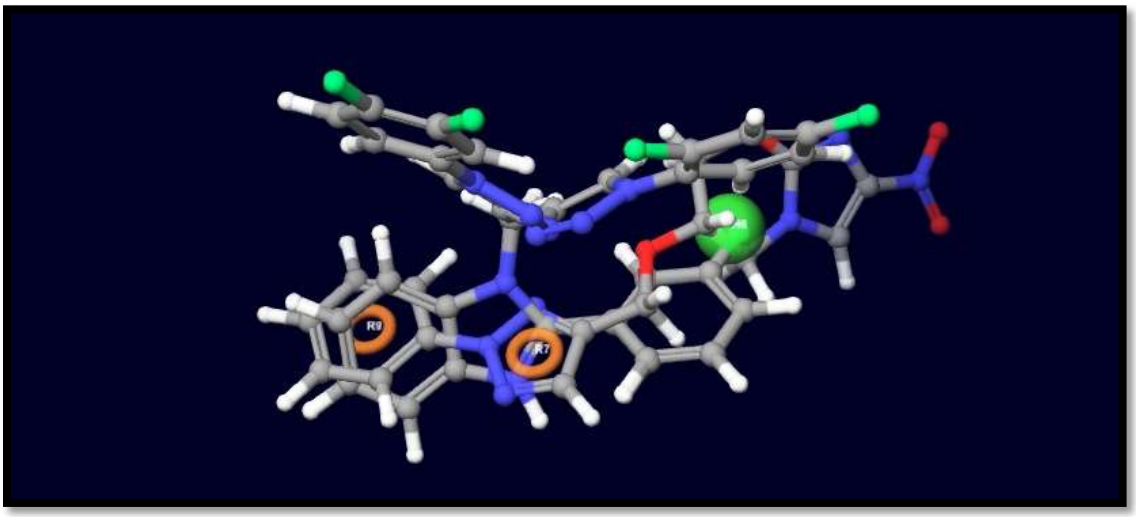


Figure 5: Contour map of Compound 36

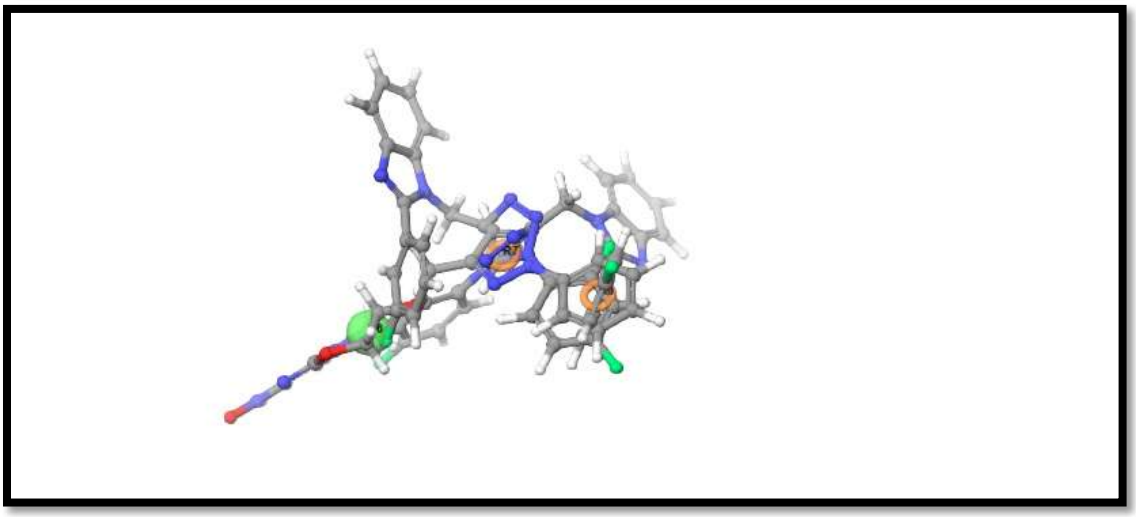


Figure 6: Contour map of Compound 13

Figure 6: Std* coeff contour maps of CoMFA analysis with 2Å grid spacing in combination with compound 40 and 28. 4 and 4 Steric fields; green contour indicates regions where bulky groups increases the activity; yellow contour indicates regions where bulky groups decreases the activity and Electrostatic fields; blue contour represent regions where electron-donating groups increases the activity; red contour represent the regions where electron-withdrawing groups increases the activity.

Table 10: The determination of statistical parameters for the models of the series based on different distinct with default size 4-7.

S no	Fragment Distinct	q^2	r^2	q^2 SE	r^2 SE	Ensemble	Best length	NC
1	A/B	0.792	0.946	0.286	0.150	0.948	151	6
2	A/B/H	0.771	0.955	0.300	0.141	0.939	353	6
3	**A/B/C**	**0.800**	**0.943**	**0.276**	**0.160**	**0.933**	**257**	**6**
4	A/B/Ch	0.787	0.951	0.290	0.147	0.947	151	6
5	A/B/C/H	0.794	0.950	0.285	0.149	0.937	151	6
6	A/B/DA	0.797	0.941	0.278	0.163	0.937	257	6
7	A/B/C/DA	0.781	0.947	0.289	0.154	0.935	353	6
8	A/B/C/Ch	0.800	0.943	0.276	0.160	0.933	257	6
9	A/B/H/DA	0.793	0.932	0.285	0.174	0.922	257	6
10	A/B/H/Ch	0.785	0.939	0.291	0.164	0.930	257	6
11	A/B/Ch/DA	0.796	0.937	0.276	0.168	0.933	353	6
12	A/B/C/H/DA	0.773	0.961	0.298	0.132	0.946	353	6
13	A/C/H/DA	0.783	0.951	0.292	0.148	0.939	257	6
14	A/C/H/Ch/DA	0.782	0.944	0.293	0.158	0.935	307	6

Table: 11: For Test:

Pred R^2	SE	Bond Length	NC
0.938	0.166	307	6

Table 12: Residual value of molecules of the HQSAR model

S. No.	Actual	Predicted pIC50	Residual	Sno.	Actual	Predicted pIC50	Residual
1*	3.6579	3.84	-0.1821	24	4.9473	5.03886	-0.09156
2	4.3639	4.31899	0.04491	25	5.0716	5.18102	-0.10942
3	3.6213	3.70718	-0.08588	26	4.95	4.98164	-0.03164
4*	3.4117	3.897	-0.4853	27	5.1568	5.05202	0.10478
5	3.6873	3.65625	0.03105	28	5.2557	5.17937	0.07633
6	3.7563	3.72401	0.03229	29*	4.8945	5.034	-0.1395
7	4.4347	4.40343	0.03357	30	4.8213	4.82304	-0.00174
8*	4.7392	3.581	1.1582	31	5.2534	5.17144	0.08196
9	4.1642	3.97353	0.19067	32*	4.5787	4.882	-0.3033
10	3.4661	3.85838	-0.39228	33	5.1475	5.02115	0.12635
11	3.7534	3.68661	0.06679	34	4.7913	4.63785	0.15345
12*	3.3722	4.281	-0.9088	35	5.0942	5.05461	0.03959
13	3.7173	3.60906	0.10824	36*	4.5525	5.275	-0.7225
14	3.8362	3.78962	0.04658	37	4.7368	4.9274	-0.1906

15	3.7132	3.85739	-0.14419	38	5.0773	5.31175	-0.23445
16*	4.0037	4.665	-0.6613	39	5.0931	5.13309	-0.03999
17	3.909	3.93422	-0.02522	40	5.2749	4.91999	0.35491
18	4.2563	4.10691	0.14939	41	4.9551	4.85028	0.10482
19	3.6453	3.99175	-0.34645	42	5.0357	4.96971	0.06599
20	3.7552	3.81998	-0.06478	43	4.2729	4.65657	-0.38367
21	4.3336	3.97848	0.35512	44	4.3487	4.588	-0.2393
22	5.1871	4.98552	0.20158	45	4.9263	5.10494	-0.17864
23	4.9901	5.09693	-0.10683	46	4.9539	4.89264	0.06126

Table 13: Summary of the statistical parameters of HQSAR studies:

S. No	Statistical parameters	Model (A/B/C)	Model (A/B/C/Ch)
1	Fragment size	2-6	2-6
2	q^2	0.801	0.801
3	r^2	0.920	0.920
4	Ensemble	0917	0.917
5	SE	0.189	0.189
6	NC	6	6
7	Best Length	151	151

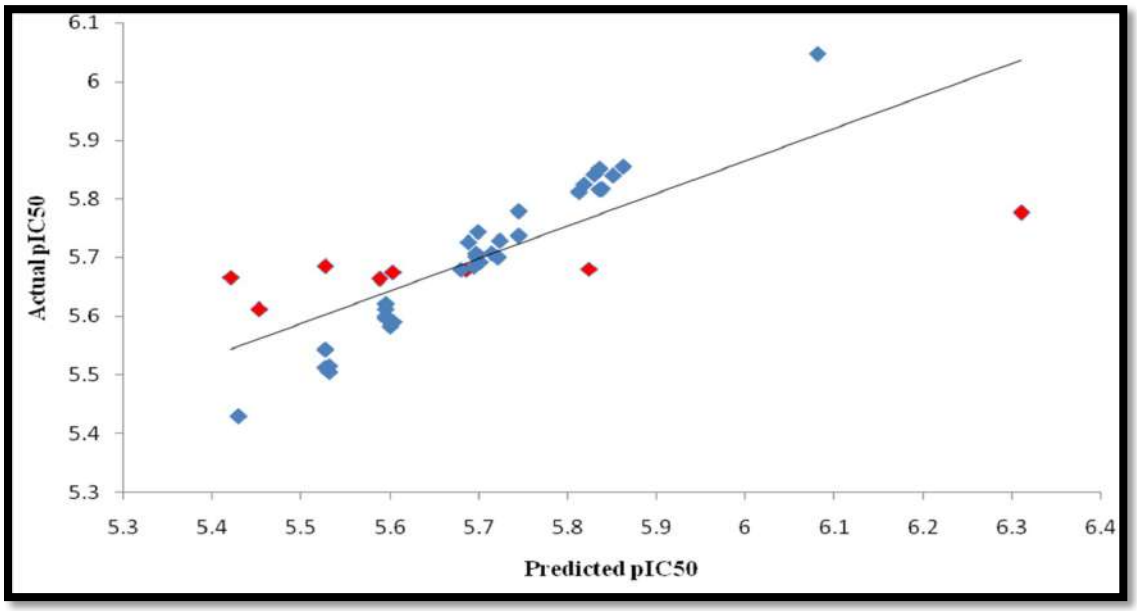


Figure 7: Graph of actual versus predicted pIC50 values of the training set and the test set molecules of Model A/B/C at 2-6 fragment size using the HQSAR.

Pharmacophore Modelling:

Ten GALAHAD models were generated by using training set compounds. Model 8 and 10 had high energy which is considered to be due to steric clashes, leading to their exclusion from the analysis. The other 20 models were generated and evaluated successively by the test database constructed previously. Table 5.13 and 5.14 shows the predictable results for each model. Model 8 with the highest value was considered to be the best model.

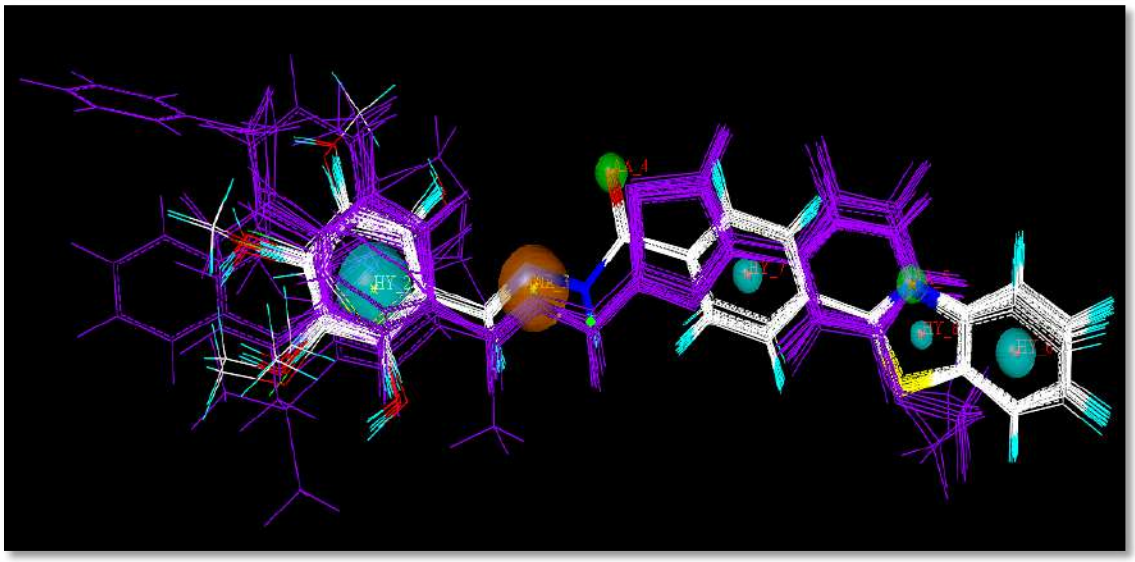


Figure 8: Pharmacophore model 8 and molecular alignment of the compound

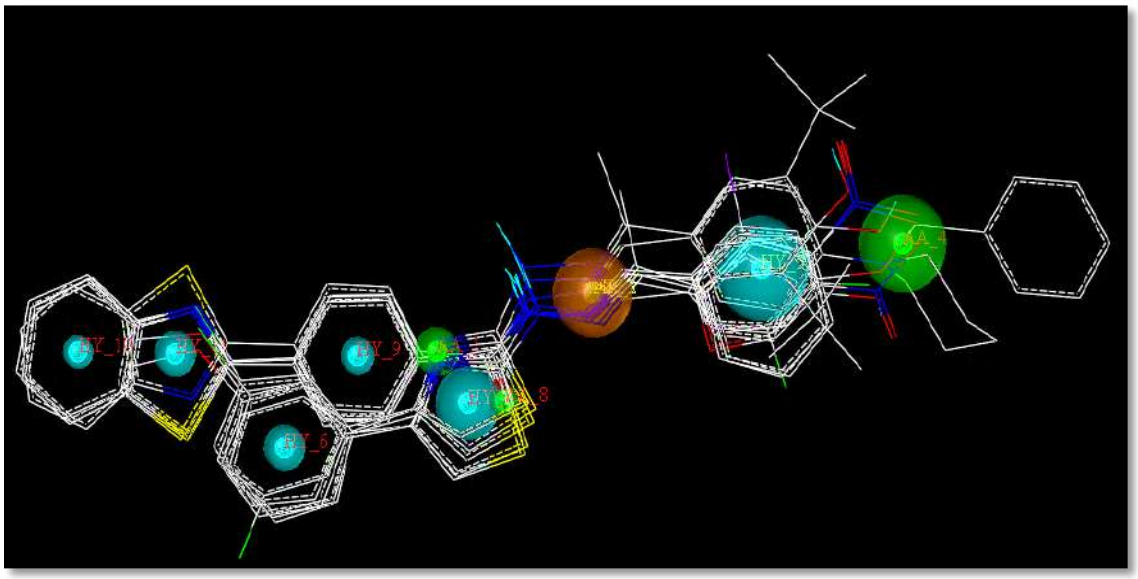


Figure 9: Alignment of all test set compounds using pharmacophore modelling.

Table 14: The parameter values of Training set for each pharmacophore model:

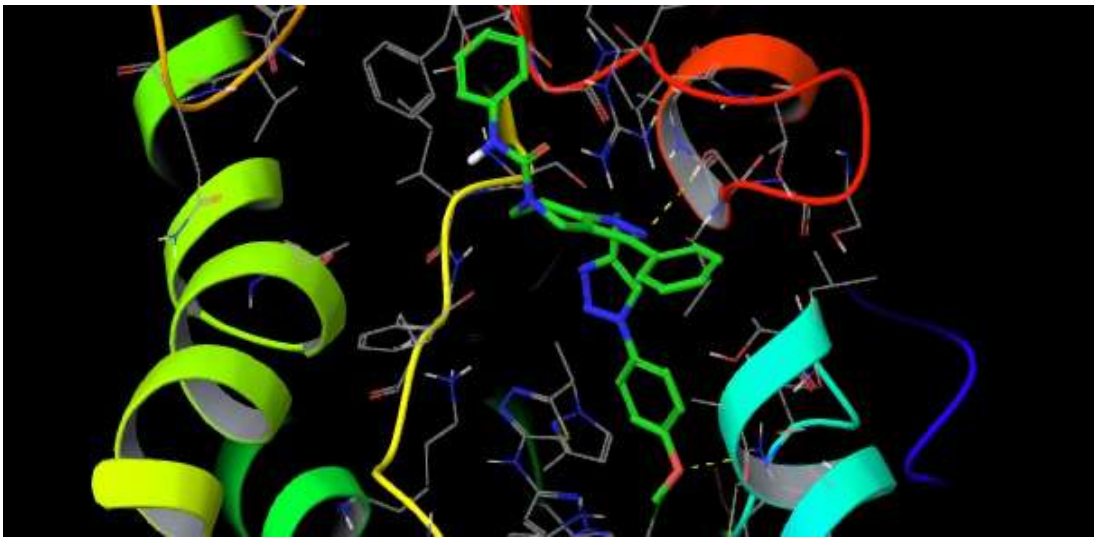
NAME	Specific.	N_HITS	FEATS	PARETO	Energy	Steric	HBOND	MOL_QRY
Model_001	3.818	-16	8	0	12.16	1344.7	328.5	102.39
Model_002	3.651	-16	9	0	11.05	1302.6	326.7	101.73
Model_003	3.812	-16	8	0	15.43	1431.7	326.1	103.38
Model_004	1.66	-16	9	0	8.05	1217.9	321.6	104.1
Model_005	3.823	-16	8	0	10.95	1338.4	320.8	104.34
Model_006	4.979	-16	8	0	17.59	1255.7	336	107.9
Model_007	3.814	-16	8	0	15.09	1308.6	325	107.49
Model_008	3.822	-16	8	0	10.95	1292.2	326.7	72.97
Model_009	3.8	-16	9	0	9.89	1340.7	322.1	66.9
Model_010	3.825	-16	8	0	8.36	1159.2	326.5	88.92

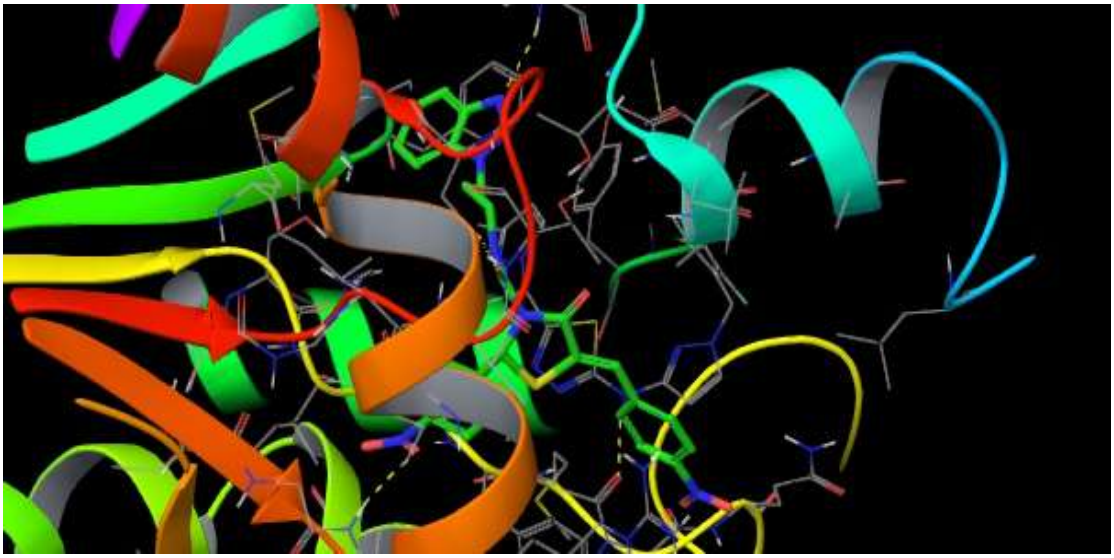
Table 15: The parameter value of Test set for each pharmacophore model:

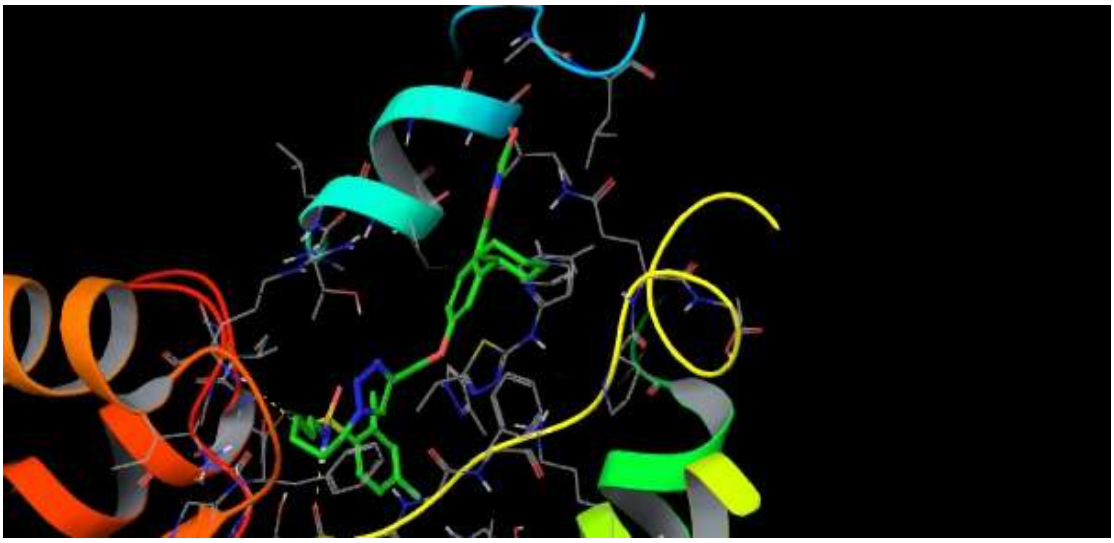
NAME	Specific.	N_HITS	FEATS	PARETO	Energ.	Steric	HBOND	MOL_QRY
Model_001	3.710	0	14	0	7.91	1363.50	279.80	119.80
Model_002	3.774	2	13	0	11.75	1144.60	287.40	119.56
Model_003	3.391	9	12	0	11.12	1394.50	250.20	96.99
Model_004	3.405	8	12	0	12.78	1487.70	277.80	90.49
Model_005	3.710	0	14	0	15.08	1232.30	286.00	107.41
Model_006	3.767	1	13	0	21.92	1191.60	287.80	112.80
Model_007	3.398	9	12	0	14.83	1035.00	288.10	112.07
Model_008	3.477	9	11	0	11.15	1288.20	281.60	58.79
Model_009	2.399	9	12	0	16.15	1233.70	282.00	95.57
Model_010	2.371	9	13	0	180.71	1170.10	292.10	97.11

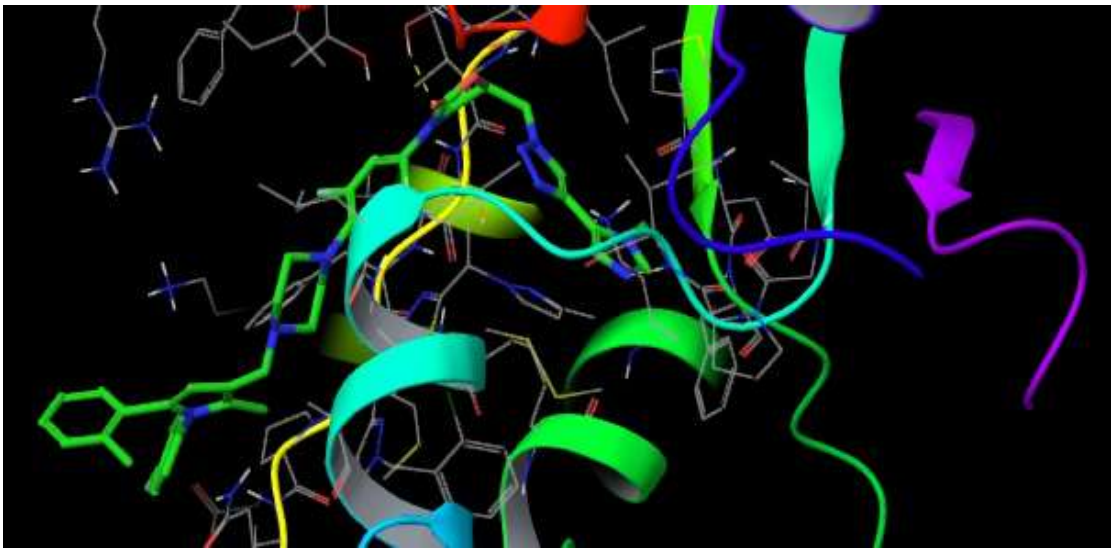
Table 16: Docking pose view of the compound 36 and 13 based on 5O1H- PDB

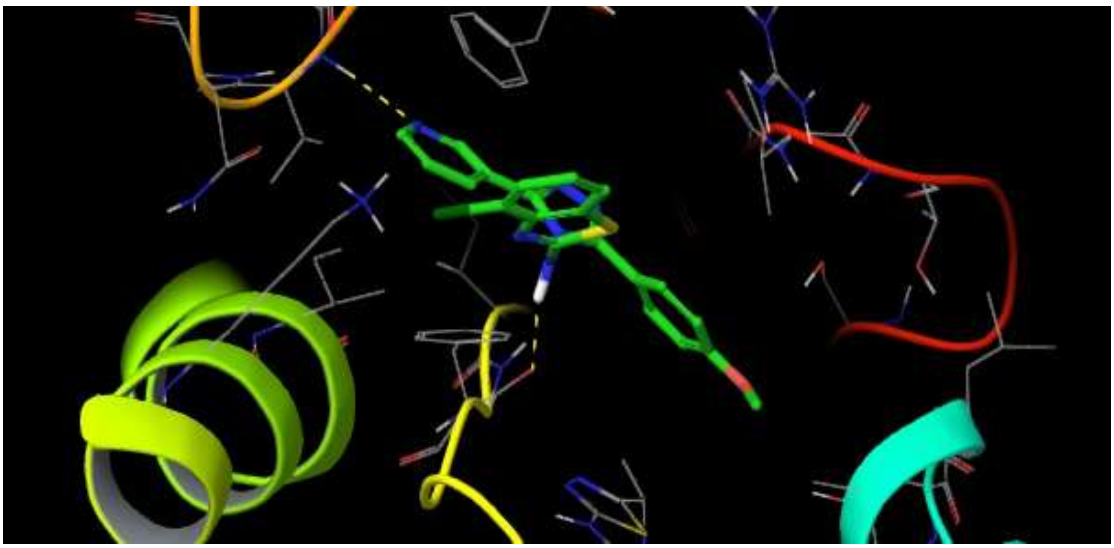
Sno	Compound	Total Score	Sno	Compound	Total Score	Sno.	Compound	Total Score
1	20	8.75	16	28	6.49	31	3	5.33
2	22	8.58	17	27	6.48	32	25	5.18
3	7	7.87	18	11	6.47	33	23	5.03
4	1	7.67	19	13	6.46	34	33	4.93
5	26	7.56	20	17	6.43	35	36	4.89
6	30	7.54	21	10	6.32	36	18	4.88
7	38	7.39	22	19	6.22	37	9	4.84
8	15	7.3	23	16	6.18	38	39	4.81
9	8	7.27	24	4	5.92	39	2	4.75
10	43	7.21	25	12	5.9	40	34	4.33
11	21	7.11	26	29	5.83	41	24	4.04
12	31	7.11	27	5	5.79	42	37	3.89
13	45	7.06	28	14	5.79	43	44	3.69
14	42	6.76	29	35	5.75	44	32	3.44
15	41	6.54	30	6	5.43	45	40	3.24

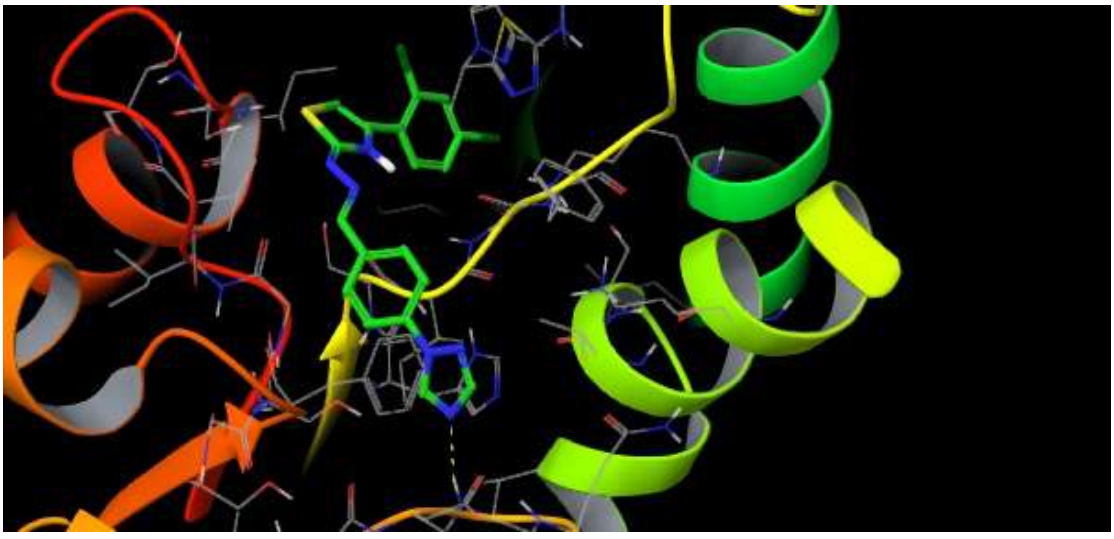


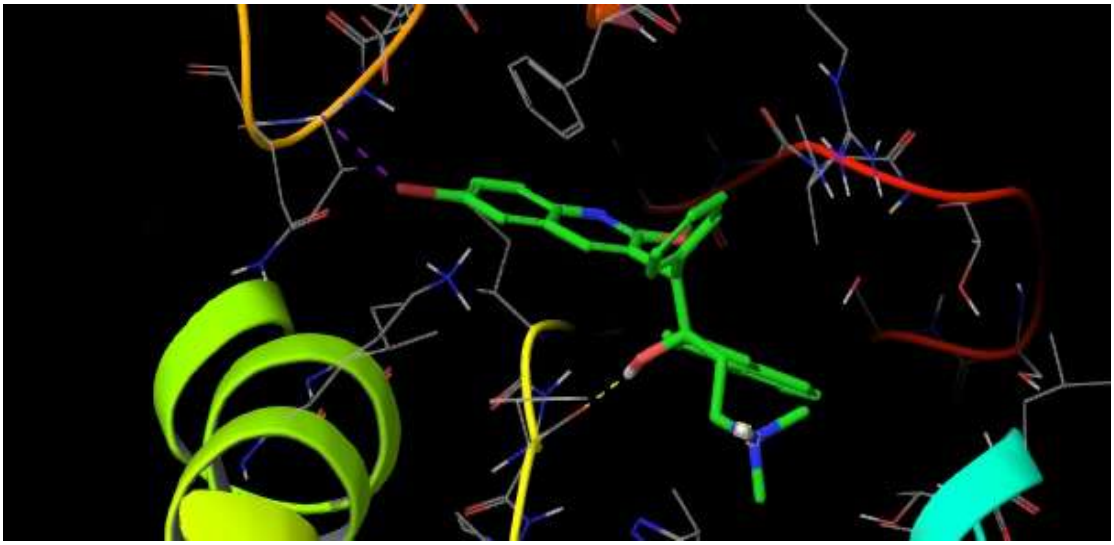


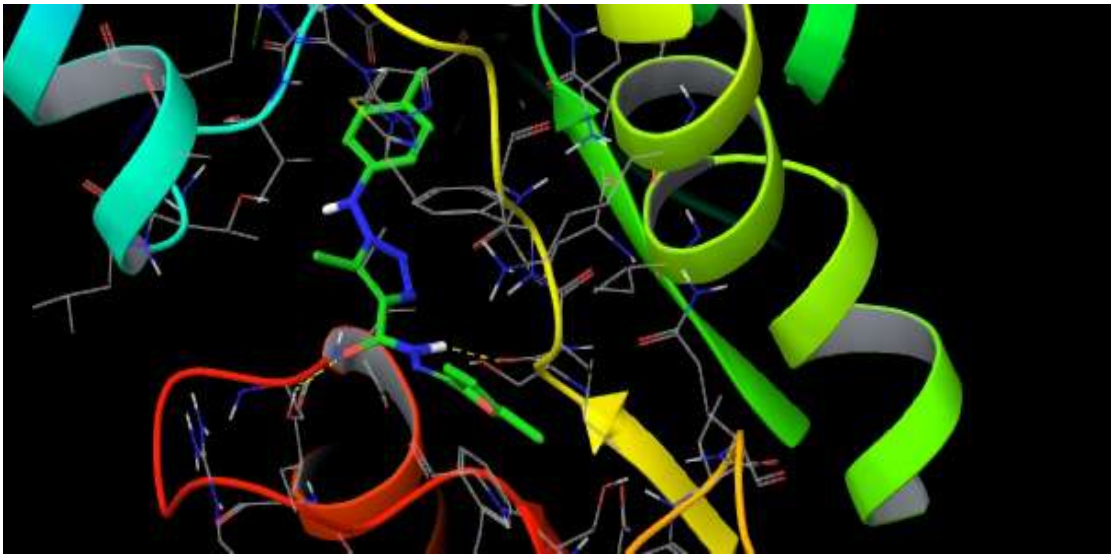


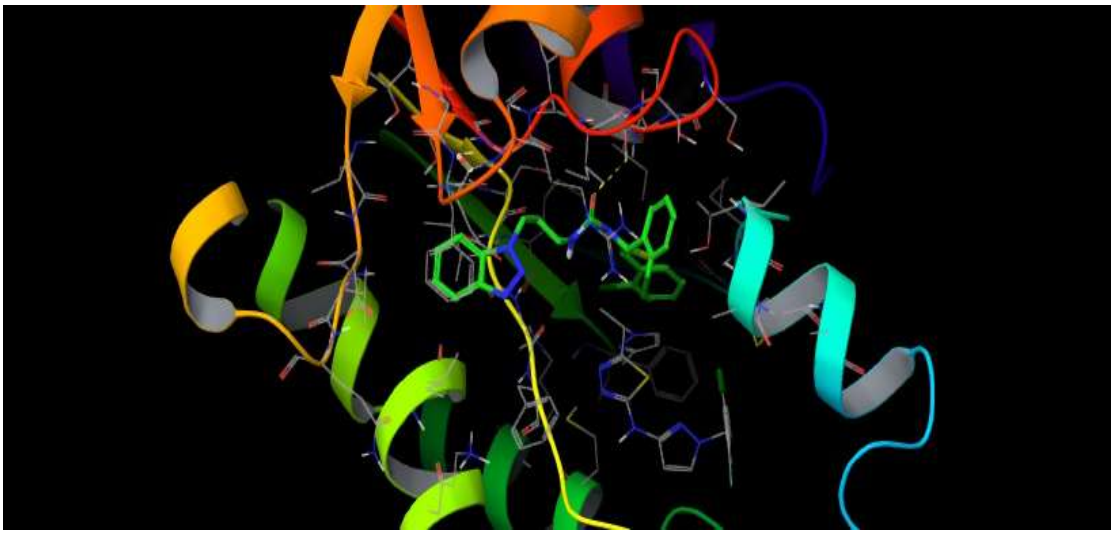












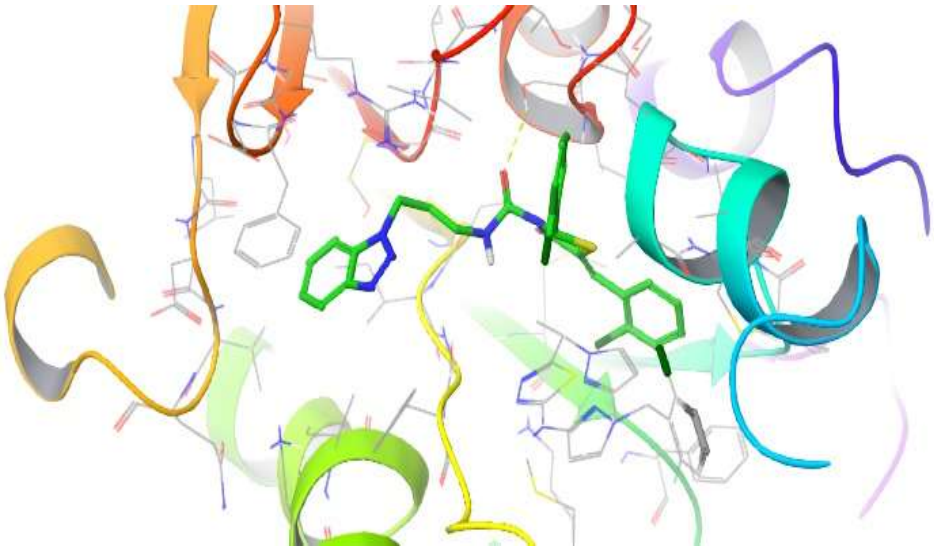
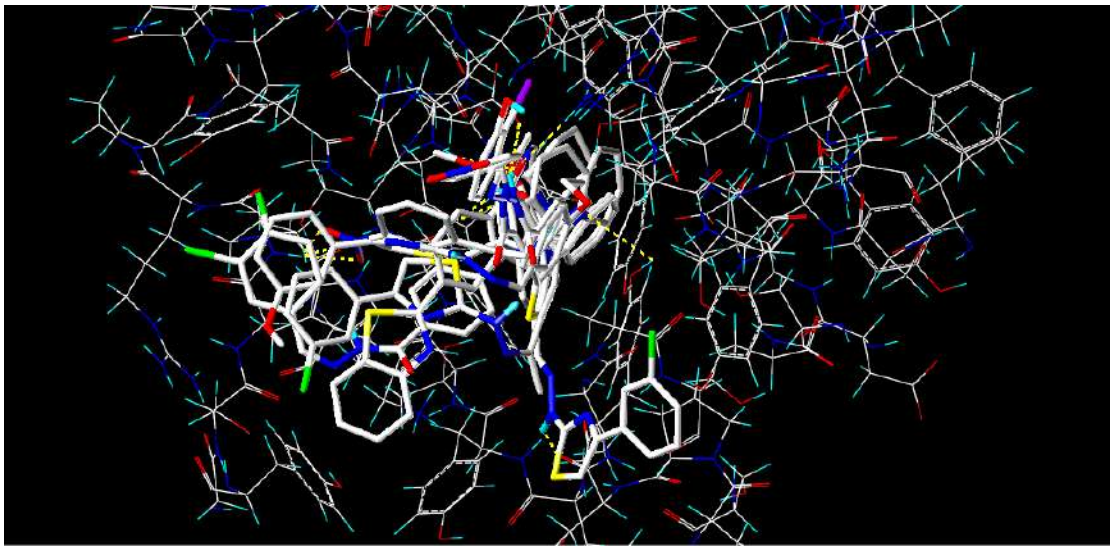


Figure 10: Full Docking view of all compounds on 5O1H -PDB:



Training Compound **Test Compound**

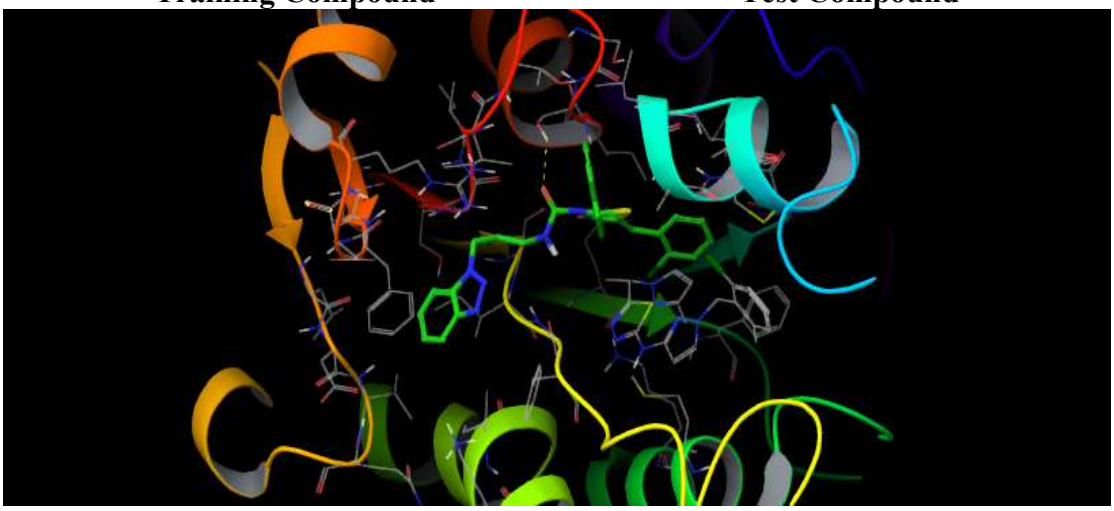


Figure 11: Interaction point of compound 36

Conclusion:

All the designed compounds as per QSAR, docking were subjected to evaluate from previously generated QSAR model their activity were predicted. While predicting the activity, it was found that out of thirty nine designed compounds, seven compounds were supposed to having a better activity in all the computational studies from the docking data of newly designed compounds, the 3 compounds having a better activity.

References:

1. Karabasanagouda T, Adhikari AV, Girisha M, "Synthesis of some pyrazoline and isoxazoles carrying 4-methylthiophenyl moiety as potential analgesic and anti-inflammatory agents" Indian J Chem. 2009, 48B, 430-7.
2. Burger, A., Hansch, C., Sammes, P.G., Taylor, J.B.,Eds., Comprehensive Medicinal Chemistry, Pergamon press, 1990, 1
3. Camille, G., Wermuth, Eds., The Practice of Medicinal Chemistry, Elsevier, 2001, 2, 29.
4. AnanthnarayanR.Panikar, C. K.J.; In Textbook of Microbiology, Orient Longman Publisher, 6th Editon, 2009, 7-23, 49-61.
5. Pelczar, M. J. Chan, E.C. S. Krieg, N. R. In Micorobiology, Tata McGraw HillPublication, 5 th. Edition, 1986, 75-99.
6. Wang, M., Han, X., Zhou, Z. New substituted benzimidazole derivatives: A patent review, Expert Opin. Ther.Pat.2015, 25, 595–612.
7. Soderlind, K.J., Gorodetsky, B. Singh, A.K., Bachur, N.R., Miller, G.G., Lown, J.W. Bis-benzimidazole anticancer agents: Targeting human tumour helicases. Anticancer Drug Des. 1999, 14, 19–36.
8. Kumar, K., Awasthi, D., Lee, S.Y., Cummings, J.E., Knudson, S.E., Slayden, R.A., Ojima, I, Benzimidazole-based antibacterial agents against Francisellatularensis. Bioorg. Med. Chem. 2013, 21, 3318–3326.
9. Ke, Y., Zhi, X., Yu, X., Ding, G., Yang, C., Xu, H. Combinatorial synthesis of benzimidazole-azo-phenol derivatives as antifungal agents.,Comb. Chem. High Throughput Screen, 2014, 17, 89–95.
10. Tonelli, M., Paglietti, G., Boido, V., Sparatore, F., Marongiu, F., Marongiu, E., La Colla, P., Loddo, R. Antiviral activity of benzimidazole derivatives. I. Antiviral activity of 1-substituted-2-[(benzotriazol-1/2-yl)methyl] benzimidazoles. Chem. Biodivers.2008, 5, 2386–2401.
11. Shingalapur, R.V., Hosamani, K.M., Keri, R.S., Hugar, M.H. Derivatives of benzimidazolepharmacophore: Synthesis, anticonvulsant, antidiabetic and DNA cleavage studies. Eur. J. Med. Chem.2010, 45, 1753–1759.
12. Siddiqui, N., Andalip; Bawa, S., Ali, R., Afzal, O., Akhtar, M.J., Azad, B., Kumar, R. Antidepressant potential of nitrogen-containing heterocyclic moieties: An updated review. J. Pharm. Bioallied Sci. 2011, 3, 194–212.
13. Datar, P.A., Limaye, S.A. Design and Synthesis of Mannich bases as Benzimidazole Derivatives as Analgesic Agents. Anti-Inflamm. Anti-Allergy Agents Med. Chem.2015, 14, 35–46.
14. Achar, K.C.S., Hosamani, K.M.; Seetharamareddy, H.R. In vivo analgesic and anti-inflammatory activities of newly synthesized benzimidazole derivatives. Eur. J. Med. Chem.2010, 45, 2048–2054.
15. Bansal, Y., Silakari, O. The therapeutic journey of benzimidazoles: A review. Bioorg. Med. Chem.2012, 20, 6208–6236.
16. Abd-Alsalam E, Hafez HN, Assay MG, Ali AK, El-Farargy AF, Abbass EM. Exploring the Antiproliferative Potency of Pyrido [2, 3-d] Pyrimidine Derivatives: Studies on Design, Synthesis, Anticancer Evaluation, SAR, Docking, and DFT Calculations. Chemistry & Biodiversity. 2024 Feb; 21(2):e202301682.
17. Park SY, Saralamma VV, Nale SD, Kim CJ, Seong JY, Baig MH, Cho J. Design, synthesis, and evaluation of purine and pyrimidine-based KRAS G12D inhibitors:

Towards potential anticancer therapy. Heliyon. 2024 Apr 2.
18. Pattabi V, Veeraboina MR, Eppakayala L, Navuluri S, Mulakayala N. Design, synthesis and biological evaluation of aryl urea derivatives of oxazole-pyrimidine as anticancer agents. Results in Chemistry. 2024 Jan 1; 7: 101442.
19. Teli G, Pal R, Maji L, Purawarga Matada GS, Sengupta S. Explanatory review on pyrimidine/fused pyrimidine derivatives as anticancer agents targeting Src kinase. Journal of Biomolecular Structure and Dynamics. 2024 Feb 11; 42(3):1582-614.
20. Myriagkou M, Papakonstantinou E, Deligiannidou GE, Patsilinakos A, Kontogiorgis C, Pontiki E. Novel Pyrimidine Derivatives as Antioxidant and Anticancer Agents: Design, Synthesis and Molecular Modeling Studies. Molecules. 2023 May 5; 28(9):3913.
21. Al-Tuwaijri HM, Al-Abdullah ES, El-Rashedy AA, Ansari SA, Almomen A, Alshibl HM, Haiba ME, Alkahtani HM. New Indazol-Pyrimidine-Based Derivatives as Selective Anticancer Agents: Design, Synthesis, and In Silico Studies. Molecules. 2023 Apr 23; 28(9):3664.
22. Cherukumalli PK, Tadiboina BR, Gulipalli KC, Bodige S, Badavath VN, Sridhar G, Gangarapu K. Design and synthesis of novel urea derivatives of pyrimidine-pyrazoles as anticancer agents. Journal of Molecular Structure. 2022 Mar 5; 1251:131937.