

Designing of Trifluoromethyl Substituted Pyrimidine Pharmacophore for Antiprostate Activity through a Collective Computational Approach

Kalyani Asgaonkar*, Sushruti Tanksali, Kshitija Abhang, Krishna Shevate, Shital Patil, Trupti Chitre

Department of Pharmaceutical Chemistry, AISSMS College of Pharmacy, Kennedy Road, Pune, Maharashtra, INDIA.

ABSTRACT

Background: Prostate cancer is one of the leading causes of death. Though many drugs are being used as effective anticancer agents resistance and side effects necessitates development of target selective and safe anticancer agents. **Objectives:** The objectives of our work were to identify the important pharmacophoric features and correlate structure of Trifluoro substituted pyrimidine and predict their anticancer activity using QSAR, pharmacophore modelling, and docking studies. **Materials and Methods:** In this research investigation molecular modelling approach was adopted to develop efficacious Trifluoro substituted pyrimidine derivatives as anticancer agents. **Results:** Statistically significant models were generated using 2D-MLR method (correlation coefficient (r^2) of 0.9207 and a significant Leave-one-out cross-validated correlation coefficient (q^2) of 0.8187) and 3D QSAR studies by SA-kNN method ($q^2 = 0.723$). 2D QSAR models indicated the important of SssNHcount and H- bond acceptor groups to augment the activity. Similarly, 3D QSAR results indicated the requirement of less electronegative and less steric substituents. Essential features such as aromatic and hydrogen bond acceptor features of compounds to interact with the target were highlighted using pharmacophore mapping. These studies prompted us to design new molecules using lead grow tool. Designed compounds were screened for their drug likeness, activity and toxicity using SWISS ADME, Osiris and PASS respectively. Finally docking and binding affinities of designed molecules were studied on EGFR 1M17. **Conclusion:** Thus, these studies collectively have helped to establish relationship between pyrimidine nucleus and anticancer activity.

Keywords: Prostate cancer, QSAR, Combilib, ADME, PASS, OSIRIS, Docking.

Correspondence:

Mrs. Kalyani Dharendra Asgaonkar

Assistant Professor, Department of
Pharmaceutical Chemistry, AISSMS
College of Pharmacy, Kennedy Road,
Pune, Maharashtra, INDIA.
Email: kalyani_a@aissmscop.com

Received: 08-08-2022;

Revised: 22-12-2022;

Accepted: 31-03-2023.

INTRODUCTION

Cancer is a generic term that defines the rapid multiplication and proliferation of abnormal cells in the body beyond their usual levels. According to WHO, it accounts for about 10 million deaths in 2020, of which prostate cancer was one leading type, accounting for about 1.41 million cases.^{1,2}

Prostate Cancer (PCa) is most frequently diagnosed cancer among men in the United States and Europe and the third leading cause of death from cancer. Although different treatments are available, Progression towards Castration Resistance (CRPC) is a major task in the treatment of PCa. CRPC ultimately results in development of extreme resistance to androgen deprivation therapy and chemotherapy. There is a report that patients with advanced PCa may be at a greater risk of death due to COVID-19.³⁻⁶

The Epidermal Growth Factor Receptor (EGFR) belongs to the class of Tyrosine kinase activity that ensures specificity and affinity in their binding sites due to the presence of an extracellular transmembrane and intracellular domain. EGFR- a multi-functional membrane glycoprotein serves as an important target for therapeutic intervention and its expression is found to be increased while the progression of prostate cancer is observed.⁷⁻⁹

In the recent trends of treatment many drugs are used as effective therapy for the treatment of Prostate Cancer eg: Relugolix a GnRh antagonist, Gefitinib-anti-EGFR drug, Darolutamide- a second generation Androgen Receptor Inhibitors (ARI's) and Enzalutamide.¹⁰⁻¹⁵

Pyrimidine is a nitrogen containing heterocycle, with a plethora of pharmacological activities. The potential of pyrimidine scaffolds against prostate cancer has also been proved in the literature.¹⁶⁻¹⁹ Pyrimidineheterocycle containing drugs are being used in the treatment of Prostate Cancer such as Alisertib—an Aurora B kinase Inhibitor and Relugolix—a GnRHR Antagonist, Brigatinib- ALK-EGFR inhibitor and Vecaburtinib- a BTK inhibitor.²⁰



DOI: 10.5530/ijper.57.3.101

Copyright Information :

Copyright Author (s) 2023 Distributed under
Creative Commons CC-BY 4.0

Publishing Partner : EManuscript Tech. [www.emanuscript.in]

Literature study has drawn attention towards some key groups such as presence of sulfhydryl group on the 2nd and trifluoromethyl group on 6th position of the pyrimidine ring responsible for good bioactivity. Similarly, Liu *et al.* have explored trifluoromethyl-substituted pyrimidine derivatives for Human prostate cancer activity. Their studies have revealed some crucial structural requirements which lead to better bioactivity than the standard drug. Likewise, presence of acrylamide group, benzene ring substituted with electron withdrawing groups especially at para or meta position have shown promising results.²¹

Quantitative Structural Activity Relationship (QSAR) is an important strategy for the development of leads that is used to identify the contributions of the chemical structures to the targeted pharmacological activities. A large quantum of Computational work has been carried out on many Pyrimidine derivatives as evident from many research articles in search of lead molecules for anticancer activity.²²⁻²⁶

As apparent from different observations, treatment of PCa is complex and necessitates efforts to develop more target specific therapy that can be used to overcome drug-resistant in CRPC patients. Hencein the present investigation we have explored trifluoromethyl-substituted pyrimidine scaffold for prostate anti-cancer activity using molecular modelling studies. QSAR and pharmacophore mapping studies were used to optimize the pharmacophore and a library of New Chemical Entities (NCE's) was generated using Leadgrow tool. NCE's were screened for their absorption, distribution, metabolism, excretion, toxicity, activity prediction spectra followed by docking studies.

MATERIALS AND METHODS

QSAR

Data set

Considering the chemical and biological variations, a data set comprising of 47 novel trifluoromethyl-substituted pyrimidine derivatives bioactive against human tumour cells reported by Liu *et al.*²¹ was considered for the QSAR study. Log-transformed IC₅₀ were used for the QSAR models. Chemical structures of the compounds and their biological activity are shown in Table S1.

Structure optimisation

QSAR studies were performed using V-Life sciences MDS Version 4.4. 2D structures of the molecules were sketched using ChemDraw tool and converted to 3D structures using the V-Life software. Geometry optimisation of the structures was carried out using the standard Merck Molecular Force Field (MMFF) and the initial conformations were minimised using the Powell method by using a distance-dependent dielectric constant of 1.0 and the convergence criterion of 0.0001 kcal/mol.

Molecular descriptor pruning

Molecular descriptors are calculated as they represent key structural properties of molecules. QSAR worksheet was generated by using the antitumor activity expressed as pIC₅₀ values as dependent variables while the molecular descriptors served as the independent variables. Descriptors which did not contribute to the biological activity as represented by invariable columns in the worksheet were removed before further processing. In 2D QSAR physicochemical, structural, topological, electro-topological and alignment-based descriptors were generated. Where as in 3D QSAR steric, hydrophobic and electrostatic interaction energies were computed at the lattice points of the grid. The examination of the correlation matrices confirmed the absence of collinearity problems between the descriptors. Predominant descriptors i.e., those showing highest influence on antitumor activity were selected for generation of the QSAR model.

Data set distribution

To avoid statistical leakage and to examine predictive ability of the generated models, Sphere exclusion method was used to divide the entire data set into training (80% = 38 molecules) and external/test sets (20% = 9 molecules) respectively with the appropriate composition for 2D and 3D QSAR. Correctness of selection criteria for training and test set molecules was gauged by Uni-Column statistics.

2D and 3D QSAR model generation

QSAR was computed via different statistical models like Partial Least Square (PLS), Multiple Linear Regression (MLR) and Principal Component Regression (PCR) for 2D and k Nearest Neighbour Molecular Field Analysis (kNN MFA) using Simulated Annealing (SA) variable selection for 3D QSAR. Generation of conformers was done using Monte Carlo conformation search method. In 3D QSAR a common template was prepared to align the optimised molecules. This was followed by generation of common rectangular grid around the molecules where in steric, hydrophobic and electrostatic interaction energies were computed at the lattice points of the grid to build a relationship with the biological activity. Stepwise Forward-Backward variable selection method with its corresponding parameters was chosen to develop 3D QSAR models.

Statistical testing and validation of the QSAR models

Internal and external validation of generated QSAR models was performed by computing statistical parameters. Internal validation included correlation coefficient R^2 , cross-validated correlation coefficient (Q^2) using standard Leave One-Out (LOO) method and for external validation (pred_r2) was calculated. r²m were considered to check the robustness and predictive ability of the models.²⁷⁻³³

Pharmacophore mapping

The MolSign module in V-LifeMDS was used for aligning the data set molecules and analysing them for five Pharmacophoric features viz; Aromatic (Aro), Aliphatic (Ala) and Positive (+ve) and Negative (-ve) Charge center, H-bond Acceptor (HAc) and H-bond Donor (HDr). The larger tessellated spheres are indicative of the common pharmacophores identified in the molecules, the smaller solid features are of the individual molecules. Most active compound was used as reference molecule and the entire data set was aligned on the reference molecule. Five pharmacophoric features were selected from the molecular set.³⁴

Combinatorial library generation and screening of ADME properties

A combinatorial library of 225 compounds was generated using Lead Growmodule of V Life MDS taking into consideration the results obtained from QSAR studies. Drug-likeness property is one of the main criteria for selection of a newer molecule at an initial stage of drug discovery. The online tool SwissADME (www.swiss-adme.ch/) was used to calculate the properties of ADME and drug-likeness parameters. The structures of the compounds were drawn using Chem Draw Ultra 8 software and the Simplified Molecular Input Line Entry System (SMILES) was generated for each compound. These SMILES were inserted in SwissADME to generate the results. The library of compounds was screened using Lipinski's Filter and compounds with Lipinski Score 6 were selected for further studies. Lipinski's Filter screens the compounds based on various pharmacokinetic profiles such as-Number of hydrogen bond acceptors (A) ≤ 10 , Number of hydrogen bond donors (D) ≤ 5 , Number of Rotatable bonds (R) ≤ 10 , partition coefficient $X \log P (X) \leq 5$, molecular weight (W) ≤ 500 g/mol, polar surface area (S) ≤ 140 A⁰.

Toxicity prediction studies

In silico studies of drug-like properties serves as an important tool in identifying a suitable candidate for its drug development and product capacity. OSIRIS Data Warrior software was used to evaluate the various toxicity parameters of the compounds screened using the Lipinski's Screen. The drug likeliness, mutagenicity, tumorigenicity, irritant reproductive effect and parameter of Partition coefficient (CLogP) values were studied to evaluate the potential of the drug candidate. The drug score (ds) value is of a contribution calculated by the respective ClogP, Solubility (ClogS), molecular weight, drug-likeness and toxicity risk within one practical value. The drug score was calculated using the Osiris Property explorer.^{35,36}

PASS

Prediction of Activity Spectra of Substances (PASS) tool was used to evaluate the general biological potential of the designed compounds for their anti-neoplastic activity. The activities

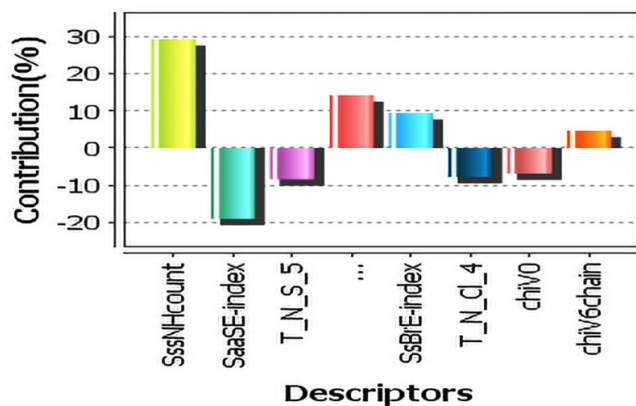
were defined in the terms of Probability to be active (Pa) and Probability to be inactive (Pi) values.^{37,38}

Docking

In order to understand interaction between designed ligands and protein, molecular docking studies were performed using AutoDock Vina. Ligands were drawn and optimised using the Merck Molecular Force Field (MMFF) so as to ensure that ligands adopt a binding mode of lowest free energy within the protein-binding site. Co-crystallised form of the Epidermal Growth Factor Receptor tyrosine kinase domain with 4-anilinoquinazoline inhibitor Erlotinib (1m17) was obtained from protein data bank. This protein structure was pre-processed and optimized using Merck Molecular Force Field. In the protein specific cavity with the key amino acids involved in the interaction with co-crystallised ligand was selected. The low energy stable conformers were selected and docked within the cavity. Post docking process Hydrogen bond, charged and pi-pi interactions were analysed, Dock score and Binding energy were computed for protein-ligand complexes and compared with standard drug Dasatinib, Gefatinib and 5-Fluorouracil. The 2D structures of the ligands were drawn using ChemDraw Ultra 8.0. Ligand preparation was carried out using the conversion of 2D structures of ligand molecules to 3D followed by optimization that ensures the lowest energy conformer was involved in the binding of the receptor. The python scrips and modified parameter files were used to carry out the docking. The protein crystal structure downloaded from RCSB was visualised using BIOVIA Discovery Studio Visualizer where the co-crystallized ligand AQ4 was highlighted and the atoms of the active site were defined. The sphere generated using these parameters allowed the marking of the active site coordinates of the protein. The Ramachandran plot was analysed for the percentage of residues of protein in the allowed region and disallowed region. Protein optimization was carried out using AutoDock Vina. The protein structure (PDB: 1m17) was obtained from Protein Data Bank. Water molecules were deleted and heteroatoms as well as co-crystallised ligands were removed. Polar hydrogen atoms and missing atoms were added. Kollman charges were computed and AD4 type was assigned. A grid box was generated in AutoDock Vina with the help of active site coordinates previously marked using BIOVIA Discovery Studio Visualizer. Active site was identified by generating the grid box using AutoDock. The grid box generation was achieved with the size of $23 \times 0 \times 56$ A^o (x, y and z) and the co-ordinates of centre $82 \times 54 \times 54$ A^o (x, y, and z). The protein was saved in PDBQT format. Molecular docking was employed to study the ligand and protein interactions by identifying the most suitable active sites for the ligand and protein to bind with each other with the best geometry. Docking was performed with all molecules treated as rigid. The exhaustiveness value used was eight with the energy range of four. The pre-processed protein

Table 1: Unicolumn statistics for training and test set.

Column Name	Average	Max	Min	Std Dev	Sum
Training	4.8100	5.2400	4.3200	0.2683	187.5900
test	4.8788	5.1700	4.4500	0.2747	39.0300

**Figure 1:** Contribution plot in 2D QSAR.

and ligands in PDBQT format were used for docking. The best conformer for each ligand was selected based on the data.³⁹⁻⁴²

RESULTS

2D QSAR

Among various 2D-QSAR generated models, one of the best models generated by MLR method was selected and discussed herein.

Unicolumn statistics

Feasible distribution of the data series into training and test for development of models was indicated by unicolumn statistics. The results indicated that the test set is derived within the min-max range of the train set. Maximum of test value of pMIC was less than maximum of training set and minimum of test value of pMIC was greater than that of training set. Mean of the test set value of pMIC was higher than the training set representing the presence of relatively more active molecules as compared to the inactive ones. Standard deviation was similar in both the sets indicating a comparable spread of molecules in both the sets (Table 1).

Radar Plot

Radar plot indicated the closeness between experimental and predicted activity (of training and test set (Figure S1-Supplementary material)).

The statistical validation matrices

To evaluate the quality, robustness and predictability of generated models' different statistical validation matrices were considered. The statistical validation metrics were found to be in good

agreement. The best QSAR model was generated through the MLR method were obtained with a $r^2 = 0.9207$, $q^2 = 0.8187$ and predicted $r^2 = 0.7078$, $\text{pred}_r^2\text{se} = 0.2531$, $r^2\text{m} = 0.756$. Following equation was generated -

$$\text{(Equation. 1)- pMIC} = 0.2504 (\text{SssNHcount}) - 0.4384 (\text{SaaSE-index}) - 0.1843 (\text{T_N_S_5}) + 0.1110 (\text{H-Acceptor Count}) + 0.5400 (\text{SsBrE-index}) - 0.3751 (\text{T_N_Cl_4}) - 0.0514 (\text{chiV0}) + 1.3469 (\text{chiV6chain}).$$

Contribution Plot

The contribution plot indicates the extent of influence of a particular descriptor on the bioactivity (Figure 1).

Interpretation of 2D QSAR

The total variance of more than 92% in the training set is accounted for the antitumor activity as indicated by $r^2 = 0.92$. The predictive potential of the model is also greater than 70%. Statistical fitness was proved by low value of standard error. From the results of 2D QSAR, the positively contributing descriptors such as SssNHcount defines the total number of -NH group connected with two single bonds, a positive number of H- bond acceptor groups (H-Acceptor Count), whereas SsBrE-index represents the Electrotopological state indices for number of bromines connected with one single bond and chiV6chain signifies the atomic valence connectivity index for six membered rings. These descriptors are responsible for the good activity of the compounds. The negatively contributing descriptors like SaaSE-index represents Electrotopological state indices for number of sulphur atom connected with two aromatic bonds. The negatively contributing, T_N_S_5 and T_N_Cl_4 indicate the count of number of Nitrogen atoms (single double or triple bonded) separated from any other Sulphur (single double or triple bonded) and Chlorine atom (single double or triple bonded) by 5 and 4 bonds in a molecule respectively. The negative chiV_0 descriptor signifies atomic valence connectivity index. These 2D descriptors and SAR predicted were used to design the template for the generation of new chemical entities and designing the template for the combinatorial library generation. (Table 2).

Interpretation of 3D QSAR

Compounds of the data set were aligned using a common template. Most active compound was used as a reference molecule to generate the 3D grid. The value of $q^2 = 0.723$, $3334t5q^2\text{se} = 0.1492$, $\text{pred}_r^2 = 0.6883$, $\text{pred}_r^2\text{se} = 0.1952$, $r^2\text{m} = 0.68$ reveals the reliability and predicting ability of QSAR models. Electrostatic

and steric field points around the Pyrimidine pharmacophore were computed and their values depended on the dissimilarity of the field values by applying the most active molecule and its nearest neighbour set. Positive electrostatic value E₈₀₀ (0.6097, 0.7212) indicated the requirement of less electronegative substituent's at that position. Whereas negative value of steric descriptor S₂₄₈ (-0.1332 -0.1028) indicated the requirement of sterically less bulky groups for better anti-tumour activity Table 3 and Figure 2.

Pharmacophore mapping

To establish the chemical groups and their arrangement that are essential for biological activity, pharmacophore mapping study was carried out on a set of 47 pyrimidine derivatives. Five-point biophores with two aromatic ring (AroC, yellow coloured), two hydrogen bond acceptors (HAc, blue coloured), one hydrophobic region, (orange coloured) as pharmacophoric features were developed (Figure 3). These observations indicate the importance of two nitrogen atoms present in the pyrimidine ring, sulphur

atom and trifluoromethyl as key substituents necessary for anti-prostate cancer activity. The bond distance between the two pharmacophoric features are reported in Table 4.

The colourschemes for the various chemical features are as follow: Hydrogen bond donor: greencolour, Hydrogen bond acceptor: Bluecolour, Hydrophobic: Orange color, Aliphatic: Buff colour, Negative ionizable: Dark Red colour, Positive ionizable: Dark Green color The larger tessellated spheres are indicative of the common pharmacophores identified in the molecules, the smaller solid features are of the individual molecules.

ADME

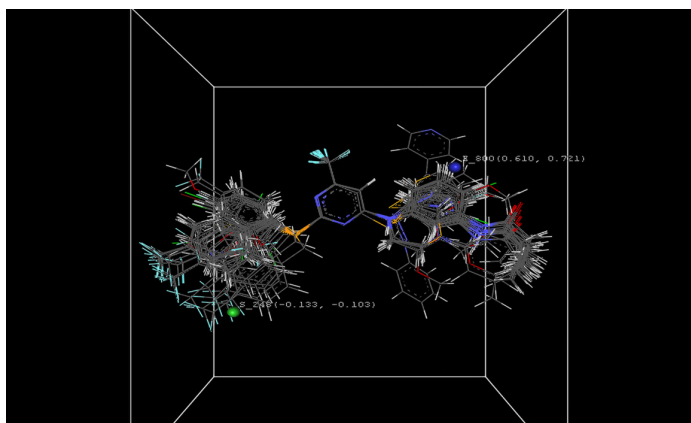
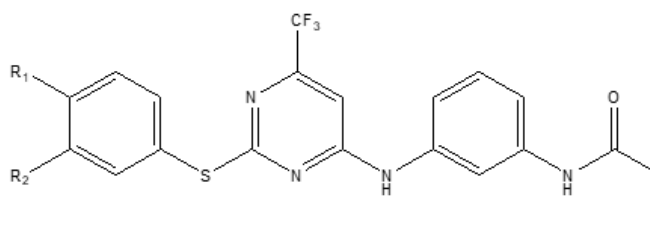
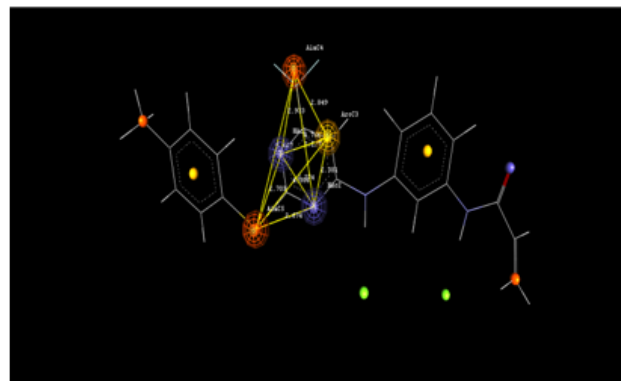
Considering the results of QSAR and Molsign, pharmacophore was optimised and a library of 250 New chemical entities (NCE's) were generated used Lead grow tool. These NCE's were subject to Lipinski filter to gauge their drug likeliness. The molecular descriptive properties are as given in Table 5. Out of the entire library 33 NCE's were selected based on their Lipinski score of 6 for further investigations.

Table 2: Interpretation of 2D QSAR descriptors.

Sl. No.	Descriptor	Description	Interpretation
1	SssNHcount	Defines the total number of -NH group connected with two single bonds.	The descriptor encodes the total number of -NH groups attached by a single bond to the native chain. Its positive contribution in the chain contributes to the better activity.
2	SaaSE-index	Electrotopological state indices for number of sulphur atom connected with two aromatic bonds.	The presence of Sulphur group in the series of compounds negatively contributes when connected with two aromatic bonds i.e., when present in a heterocycle and it is absent in the series.
3	T_N_S_5	Count of number of Nitrogen atoms (single double or triple bonded) separated from any other Sulphur by 5 bonds.	The topological descriptor contributes negatively to the presence of Nitrogen atoms separated from Sulphur atom by a bond distance of 5. The series of compounds shows no compounds with this description and thus favors the compound.
4	H-Acceptor Count	H- bond acceptor groups.	The positive contribution of Hydrogen bond acceptor count encodes the presence of such substitutions contributing to better activity of the compounds.
5	SsBrE-index	Electrotopological state indices for number of bromines connected with one single bond.	The positive presence of bromine atom present as a substitution in compounds potentially increases the activity.
6	T_N_Cl_4	Nitrogen atoms (single double or triple bonded) separated from any Chlorine atom (single double or triple bonded) by 4 bonds.	While this alignment independent descriptor contributes negatively, the presence of chlorine atom is particularly present as a substitution to the linked aryl group in the compound with respect to the Nitrogen atom in the compound present as -NH linkage or N-Ar (Pyrimidine) ring. This in turn satisfies the criteria in the compound.
7	chiV0	Atomic valence connectivity index	This descriptor contributes negatively to the significance of atomic valence connectivity index (order 0).
8	chiV6chain	Signifies the atomic valence connectivity index for six membered rings.	The contribution of this descriptor positively influences towards the presence of six membered rings, it contributes well towards the structure and better activity.

Table 3: Interpretation of 3D QSAR descriptors.

Sl. No.	Descriptor	Description	Interpretation
1	Electrostatic	E_800 (0.6097, 0.7212)	Positive range indicates that positive electrostatic potential is favourable for increase in the activity and hence a less electronegative substituent group is preferred in that region.
2	Steric	S_248 (-0.1332 -0.1028)	Negative range indicates that negative steric potential is favourable for increase in the activity and hence less bulky substituent group is preferred in that region.

**Figure 2:** Field generated around the aligned molecules in 3D QSAR.**Figure 3:** Pharmacophore mapping.**Table 4: Bond distance between two pharmacophoric features.**

Sl. No	Pharmacophoric features	Distance (Å°)
1.	HAc1-AroC3	2.228
2.	HAc1-AlaC4	2.933
3.	HAc1- HAc2	2.374
4.	HAc1-AlaC5	2.703
5.	HAc2-AroC3	2.849
6.	HAc2- AlaC4	4.786
7.	HAc2- AlaC5	2.674
8.	AroC3- AlaC4	2.849
9.	AroC3- AlaC5	4.386
10.	AlaC4- AlaC5	5.607

range of 0.5-0.7 greater than their respective P_i values, indicating good anti-neoplastic activity of all the designed compounds. The % absorbance of all the compounds were calculated using formula (%AB = 109 - 0.345 TPSA.) And the compounds with highest absorbance were selected for further studies. (Table 6).

Toxicity prediction

ClogP is the logarithm of the partition coefficient between n-octanol and water $\log(c_{\text{octanol}}/c_{\text{water}})$ a measure of a compound's hydrophilicity. Low hydrophilicities indicate high logP values that cause poor absorption or permeation. For a compound to have a reasonable probability of absorption their logP value must not be more than 5.0. The ClogP values of the compounds were obtained in the range of 4.0 to 5.1. Compounds with logP value less than 5 were considered. The drug likeliness property of the compounds was assessed and it was found that the designed compounds have very fewer resembling fragments to the commercial marketed drugs as the scores of drug-likeliness were obtained as negative values between -21.5 to -15.1. The other toxicity risk assessment studies indicated high mutagenicity, whereas only some compounds showed high tumorigenicity and reproductive effect. The drug score values of 1.0, 0.8, and 0.6 associate with no risk, medium and high risk. The values of the compounds were obtained in the range of 0.07- 0.12. (Table 6).

PASS

The P_a values should be greater than the respective P_i values obtained for those respective compounds. The P_a values of antineoplastic activity of various compounds were obtained in the

Table 5: Results of Combilib study.

Sl. No.	Comp. no.	R ₁	R ₂	H-Acceptor	H-Donor	Rotatable Bond Count	Mol wt.	Polar SA	Lipinski Screen Score
1	4	-H	-COOH	9	3	10	463.46	129.51	6
2	6	-H	-COCH ₃	8	2	10	461.48	109.28	6
3	9	-H	-COOCH ₃	9	2	9	462.45	132.34	6
4	10	-H	-NH ₂	8	3	9	342.44	119.85	6
5	11	-H	-CHO	8	2	9	446.45	109.28	6
6	12	-H	-OCH ₃	8	3	10	449.47	92.21	6
7	14	-H	-CN	8	2	9	444.46	116.00	6
8	24	-CH ₃	-COOCH ₃	9	2	10	476.47	132.34	6
9	25	-CH ₃	-NH ₂	8	3	10	446.47	119.85	6
10	26	-CH ₃	-CHO	8	2	10	460.48	109.28	6
11	29	-CH ₃	-CN	8	2	10	458.48	116.00	6
12	46	-COOH	-H	9	3	10	463.46	129.51	6
13	61	-COCH ₃	-H	8	2	10	461.48	109.28	6
14	106	-CHO	-H	8	2	9	446.45	109.28	6
15	107	-CHO	-CH ₃	8	2	10	460.48	109.28	6
16	114	-CHO	-COOH	10	2	10	492.47	149.41	6
17	115	-CHO	-NH ₂	9	3	10	462.47	136.92	6
18	116	-CHO	-CHO	9	2	10	476.47	126.35	6
19	119	-CHO	-CN	9	2	10	474.48	133.07	6
20	121	-OCH ₃	-H	8	3	10	449.47	92.21	6
21	136	-NH ₂	H	8	3	9	432.49	119.85	6
22	137	-NH ₂	-CH ₃	8	3	10	446.47	119.85	6
23	145	-NH ₂	-NH ₂	9	4	10	448.47	147.49	6
24	146	-NH ₂	-CHO	9	3	10	462.47	136.92	6
25	149	-NH ₂	-CN	9	3	10	460.48	143.64	6
26	196	-COOH	-H	9	2	9	462.45	132.34	6
27	197	-COOH	-CH ₃	9	2	10	476.47	132.34	6
28	206	-COOH	-CHO	10	2	10	492.47	149.41	6
29	211	-CN	-H	8	2	9	444.46	116.00	6
30	212	-CN	-CH ₃	8	2	10	458.48	116.00	6
31	220	-CN	-NH ₂	9	3	10	460.48	143.64	6
32	221	-CN	-CHO	9	2	10	474.48	133.07	6
33	224	-CN	-CN	9	2	10	472.49	139.79	6

Docking

On the basis of the results obtained from the toxicity prediction studies and the scores of activities, the compounds with drug likeliness -15.1 and drug score 0.08 and above with low tumorigenicity, irritant and reproductive effect were selected. The ClogP values of the compounds with less than 5 value were chosen. On the basis of these results, 10 compounds i.e. 114, 26, 29, 137, 46, 24, 196, 115, 119, 121, 221 were selected for the

further studies. The co-crystallized ligand Erlotinib (AQ4), was extracted and re-docked to obtain the interactions. The binding affinity and the interactions of the standard along with the 10 compounds were studied and the best docked compounds were screened. (Table 7, Figure 4).

The co-crystallized ligand (AQ4 i.e., Erlotinib) showed prominent hydrogen bonding interactions with Met761, Lys721, Asp831; electrostatic interactions with Asp831 along with prominent

Table 6: OSIRIS and PASS data of the NCE's.

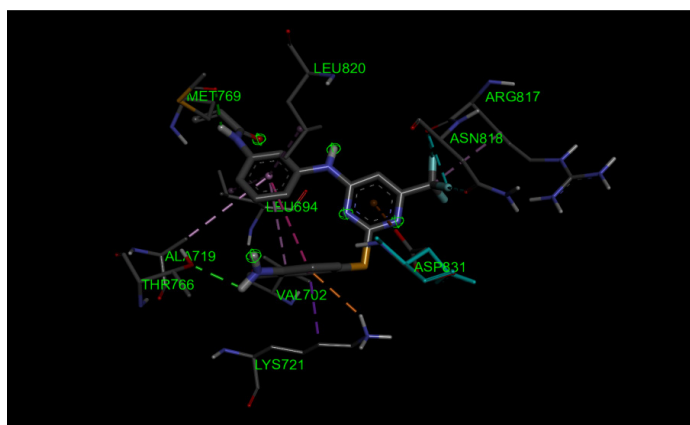
Sl. No	Comp. No	Drug Likelihood	Drug Score	Tumorigenicity	Muta-genicity	Reproductive effect	cLog P	Pa	Pi	% Absorbance
1	4	-15.108	0.10	Low	High	Low	4.639	0.517	0.067	64.31
2	6	-15.299	0.10	Low	High	Low	4.787	0.560	0.054	71.29
3	9	-17.454	0.12	Low	High	Low	4.829	0.657	0.034	68.11
4	10	-15.136	0.11	Low	High	Low	4.239	0.664	0.032	68.21
5	11	-17.258	0.10	Low	High	Low	4.849	0.668	0.032	71.29
6	12	-15.081	0.10	Low	High	Low	4.846	0.657	0.034	74.00
7	14	-19.388	0.10	Low	High	Low	4.751	0.608	0.043	68.98
8	24	-17.477	0.09	Low	High	Low	5.173	0.506	0.070	68.11
9	25	-15.193	0.08	High	High	Low	4.582	0.616	0.042	68.21
10	26	-17.323	0.09	Low	High	Low	5.193	0.675	0.030	71.29
11	29	-19.453	0.09	Low	High	Low	5.095	0.610	0.043	68.98
12	46	-17.258	0.10	Low	High	Low	4.849	0.555	0.055	71.29
13	61	-15.081	0.10	Low	High	Low	4.846	0.667	0.032	74.00
14	106	-17.258	0.07	Low	High	Low	4.849	0.702	0.026	71.29
15	107	-17.323	0.09	Low	High	Low	5.193	0.675	0.030	71.29
16	114	-17.381	0.10	Low	High	Low	4.334	0.583	0.049	58.42
17	115	-17.286	0.08	High	High	Low	4.172	0.611	0.043	62.32
18	116	-17.258	0.09	Low	High	Low	4.782	0.703	0.025	65.40
19	119	-21.538	0.10	Low	High	Low	4.685	0.703	0.025	63.09
20	121	-15.136	0.10	Low	High	Low	4.239	0.667	0.032	74.00
21	136	-15.193	0.11	Low	High	Low	4.582	0.679	0.030	68.21
22	137	-15.193	0.08	Low	High	Low	4.582	0.606	0.044	68.21
23	145	-15.136	0.09	High	High	Low	3.561	0.718	0.023	59.23
24	146	-17.286	0.11	Low	High	Low	4.172	0.601	0.045	62.32
25	149	-19.416	0.11	Low	High	Low	4.074	0.595	0.046	60.00
26	196	-15.231	0.10	Low	High	Low	4.401	0.555	0.055	64.31
27	197	-15.268	0.09	Low	High	Low	4.745	0.474	0.007	64.31
28	206	-17.381	0.10	Low	High	Low	4.334	0.583	0.049	58.42
29	211	-19.388	0.10	Low	High	Low	4.751	0.646	0.036	68.98
30	212	-19.453	0.09	Low	High	Low	5.095	0.610	0.043	68.98
31	220	-19.416	0.08	High	High	Low	4.074	0.605	0.044	60.00
32	221	-21.538	0.10	Low	High	Low	4.685	0.703	0.025	63.09
33	224	-19.388	0.10	Low	High	Low	4.587	0.638	0.037	60.77

hydrophobic interactions with key amino acids like Leu820, Leu694, Leu768, Lys721, Asp831, Glu748, Thr830, Thr766, Pro770, Gly772, Gln767, Phe771, Ala719 present in the active site. Hydrophobic interactions of the standard drugs were almost similar to the interactions with the co-crystallized ligand i.e., with amino acids Leu820 and Ala719. With reference to the above interactions of the co-crystallized ligand and standard drugs, the docked compounds were compared. The compounds with the highest number of interactions that resemble with the interactions of the co-crystallized ligand and standard drugs along with

the values of binding affinity were studied. The compound 137 showed the most interactions resembling with the co-crystallized ligand and with of a dock score of -9.2 that was found out to be more than the co-crystallized ligand and the standard drugs. The binding energy of all the ligand molecules were found to be better than the standard drugs and co-crystallized ligand. Structurally, the standard drugs as well as the co-crystallized ligand were the derivatives of the pyrimidine moiety with the presence of -N, -O, and -CH₃ as substituents to the ring in their structure. The ligand molecules with the best interactions had -NH₂, -CHO, -CN,

Table 7: Docking study results.

Compound No.	Binding Affinity (kcal/mol)	No of Hydrogen bonds	No of hydrophobic bonds	No of electrostatic interactions
24	-9.1	2	6	2
26	-9.4	1	6	2
29	-9.3	3	5	1
46	-9.1	3	4	2
114	-9.6	3	4	2
115	-9.0	3	6	0
119	-8.9	1	6	0
121	-8.8	2	4	1
137	-9.2	2	6	1
196	-9.1	2	6	0
221	-8.9	3	2	1
Co-Crystallized Ligand - Erlotinib	-7.2	4	4	1

**Figure 4:** Interactions of compound 137 with 1m17.

-COOH, -CH₃, and -COOCH₃ substitutions that were mainly responsible for their better binding and interactions with the protein molecule.

DISCUSSION

2D and 3D QSAR

The results of 2D and 3D QSAR were derived using the MLR and SA-kNN method. The conclusion from the regression equations and the validation parameters ensured the correctness of the models. The 3D QSAR was done using the aligned molecules and the 3D grid was generated. The obtained descriptors were interpreted and the points were used to generate a new set of compounds with better expectancy of activity profiles.

Pharmacophore mapping

The method of pharmacophore mapping was used to evaluate the positions of the substituent groups for better activity. The

listed groups and their bond distance values contributed to the development of the NCE's.

Combinatorial library generation and screening of ADME properties

The 2D and 3D QSAR, Pharmacophore Mapping were used to develop a template and various substituent groups for the development of NCE's. The good ADME parameters and drug-likeness scores in turn were helpful in concluding that the NCE's would prove good Leads for future study. Better drug score values and % absorption scores prove the necessity of development of these compounds.

Toxicity prediction

Toxicity prediction is an important parameter for the generation of any compounds. Compounds that show no or minimum toxicity can be essentially useful in reducing the associated side-effects and co-morbidities caused due to the same. The NCE's were obtained in the safe margin index when tested for toxicity.

PASS

In silico testing of activity of compounds ensured that the target and prediction of the study could be obtained. The percentage of activity a compound shows can aggravate the chances of development of potential drug candidates.

Docking

While screening out potential candidates, the NCE's actual binding potential with the target protein is necessarily important while developing the drug molecule. The binding affinity scores of the screened molecules showed up compound number 137 to be a potential lead for carrying out future studies.

CONCLUSION

We have attempted to attend to the global issue of prostate cancer, by conducting extensive *in silico* studies to develop a library of trifluoromethyl substituted Pyrimidine against EGFR. Following outcome has been derived from the entire investigation.

Statistically reliable QSAR models were generated indicating a positive role of SssNHcount, H-Acceptor Count descriptors, electropositive groups and sterically less bulky groups in modulating the biological activity.

Pharmacophore mapping revealed the importance of trifluoromethyl Pyrimidine nucleus through two Aromatic ring (AroC), two Hydrogen bond Acceptors (HAc) and one hydrophobic region.

Optimised of pharmacophore was the basis for designing new molecules and these molecules were found to obey the drug-likeness rule when screened for their ADME properties.

Docking studies have shown promising interactions with the key amino acids present in the active site.

Overall outcome of the present investigation proves the ability of drug design techniques in anti-cancer drug development and efficacy of compound 137 which deserves further *in vitro* and *in vivo* investigation. Hence in continuation with this work we intend to synthesize these compounds and since compound 137 has shown higher binding affinity and binding energy than the standard drugs it will promisingly provide a lead for future drug discovery.

ACKNOWLEDGEMENT

Authors are thankful to Dr. Ashwini R. Madgulkar, Principal AISSMS College of Pharmacy, Pune for her constant motivation and support.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ABBREVIATIONS

PCa: Prostate cancer; **CRPC:** Castration resistance; **EGFR:** Epidermal Growth Factor Receptor; **ARI's:** Androgen receptor inhibitors; **QSAR:** Quantitative Structural Activity Relationship; **NCE's:** New chemical entities; **MMFF:** Merck Molecular Force Field; **PLS:** Partial least square; **MLR:** Multiple linear regression; **PCR:** Principal component regression; **kNN MFA:** k Nearest neighbour molecular field analysis; **SA:** Simulated annealing; **LOO:** Leave one-out; **HAcH:** Hydrogen bond Acceptor; **HDr:** Hydrogen bond Donor; **PASS:** Prediction of Activity Spectra of Substance; **MMFF:** Merck Molecular Force Field.

SUMMARY

Though many drugs are being used as effective anticancer agents, resistance and side effects necessitates development of target selective and safe anticancer agents. The objectives of our work was to identify the important pharmacophoric features and correlate structure of Trifluoro substituted pyrimidine to predict their anticancer activity using *in silico* studies. The study was carried out by utilizing the possible *in silico* tools for model development like QSAR, Pharmacophore mapping, ADME studies, toxicity scores to finally docking for the development of Trifluoro substituted pyrimidine derivatives as anticancer agents. The 2D and 3D QSAR models were generated to design the template and various substituents contributing to the better activity of the compounds. Pharmacophore mapping essentially quoted the exact presence and distance of various chemical groups and their positioning that alleviated the biological activity profile. The NCE's were designed using the lead grow tools and the molecules were segregated using the ADME, drug-likeness, toxicity and prediction of activity spectrum type of filters. PDB 1M17 targeting prostate cancer of EGFR type was used to study the binding of the screened compounds. Compound code 137 proved to be a potential lead by passing all the filters and effectively interacting in the binding pocket of the protein.

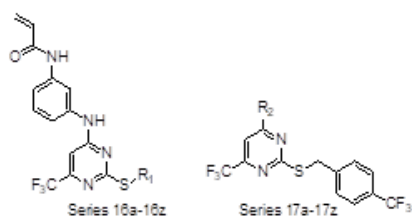
REFERENCES

1. WHO cancer report [citedMar82022]. Available from: <https://www.who.int/news-room/fact-sheets/detail/cancer>.
2. Ferlay J, Laversanne M, Ervik M, Lam F, Colombet M, Mery L, et al. Global cancer observatory: Cancer tomorrow. Lyon, France: International Agency for Research on Cancer; 2020 [citedMar82022] Available from: <https://gco.iarc.fr/tomorrow>.
3. Garin O, Suárez JF, Guedea F, Pont À, Pardo Y, Goñi A, et al. Comparative effectiveness research in localized prostate cancer: A 10-year follow-up cohort study. *Int J Radiat Oncol Biol Phys.* 2021;110(3):718-26. doi: 10.1016/j.ijrobp.2020.12.032, PMID 33388360.
4. Zamora V, Garin O, Pardo Y, Pont À, Gutiérrez C, Cabrera P, et al. Mapping the patient-oriented prostate utility scale from the expanded prostate cancer index composite and the short-form health surveys. *Value Health.* 2021;24(11):1676-85. doi: 10.1016/j.jval.2021.03.021, PMID 34711369.
5. Turco F, Armstrong A, Attard G, Beer TM, Beltran H, Bjartell A, et al. What experts think about prostate cancer management during the COVID-19 pandemic: report from the advanced prostate cancer consensus conference 2021. *Eur Urol.* 2022;82(1):6-11.
6. Zhao R, Ma X, Bai L, Li X, Mamouni K, Yang Y, et al. Overcoming prostate cancer drug resistance with a novel organosilicon small molecule. *Neoplasia.* 2021;23(12):1261-74. doi: 10.1016/j.neo.2021.11.006, PMID 34781084.
7. Rezaeiamesh A, Majidi J, Baradaran B, Movasaghpour A, Nakhband A, Barar J, et al. Impacts of anti-EGFR monoclonal antibody in prostate cancer PC3 cells. *Hum Antibodies.* 2010;19(2-3):63-70. doi: 10.3233/HAB-2010-0229, PMID 20826931.
8. Zhang H, Zhao X, Liu H, Jin H, Ji Y. Trichostatin A inhibits proliferation of PC3 prostate cancer cells by disrupting the EGFR pathway. *Oncol Lett.* 2019;18(1):687-93. doi: 10.3892/ol.2019.10384, PMID 31289542.
9. Kumar CBP, Raghu MS, Prathibha BS, Prashanth MK, Kanthimathi G, Kumar KY, et al. Discovery of a novel series of substituted quinolines acting as anticancer agents and selective EGFR blocker: molecular docking study. *Bioorg Med Chem Lett.* 2021;44:128118. doi: 10.1016/j.bmcl.2021.128118, PMID 34015505.
10. Liu YF, Fu SQ, Yan YC, Gong BB, Xie WJ, Yang XR, et al. Progress in clinical research on gonadotropin-releasing hormone receptor antagonists for the treatment of prostate cancer. *Drug Des Dev Ther.* 2021;15:639-49. doi: 10.2147/DDDT.S291369, PMID 33623372.
11. Saad F, Bögemann M, Suzuki K, Shore N. Treatment of nonmetastatic castration-resistant prostate cancer: focus on second-generation androgen receptor inhibitors. *Prostate Cancer Prostatic Dis.* 2021;24(2):323-34. doi: 10.1038/s41391-020-00310-3, PMID 33558665.
12. Ayati A, Moghimi S, Toolabi M, Foroumadi A. Pyrimidine-based EGFR TKI inhibitors in targeted cancer therapy. *Eur J Med Chem.* 2021;221:113523. doi: 10.1016/j.ejmech.2021.113523, PMID 33992931.

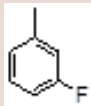
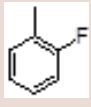
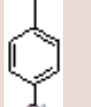
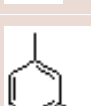
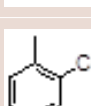

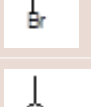
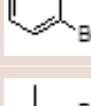

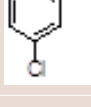
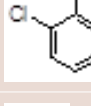

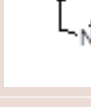
13. Abbas N, Swamy PMG, Dhiwar P, Patel S, Giles D. Development of fused and substituted pyrimidinederivatives as potent anticancer agents (A review). *Pharm Chem J*. 2021;54(12):1215-26. doi: 10.1007/s11094-021-02346-8.
14. Mahapatra A, Prasad T, Sharma T. Pyrimidine: a review on anticancer activity with key emphasis on SAR. *Future J Pharm Sci*. 2021;7(1):123. doi: 10.1186/s43094-021-00274-8.
15. Arias-Gómez A, Godoy A, Portilla J. Functional pyrazolo[1,5-a]pyrimidines: Current Approaches in Synthetic Transformations and Uses As an Antitumor Scaffold. *Molecules*. 2021;26(9):2708. doi: 10.3390/molecules26092708, PMID 34063043.
16. Kiss MA, Peřina M, Bazgier V, May NV, Baji A, Jorda R, *et al.* Synthesis of dihydrotestosterone derivatives modified in the A-ring with (hetero)arylidene, pyrazolo[1,5-a]pyrimidine and triazolo[1,5-a]pyrimidine moieties and their targeting of the androgen receptor in prostate cancer. *J Steroid Biochem Mol Biol*. 2021;211:105904. doi: 10.1016/j.jsbmb.2021.105904, PMID 33933576.
17. Suh H, Choi KW, Lee J, Ryou C, Rhee H, Lee CH. Effects of a novel carbocyclic analog of pyrrolo[2,3-d]pyrimidine nucleoside on pleiotropic induction of cell death in prostate cancer cells with different androgen responsiveness. *Bioorg Med Chem Lett*. 2016;26(4):1130-5. doi: 10.1016/j.bmcl.2016.01.057, PMID 26832220.
18. Angelucci A, Schenone S, Gravina GL, Muzi P, Festuccia C, Vicentini C, *et al.* Pyrazolo[3,4-d]pyrimidines C-Src inhibitors reduce epidermal growth factor-induced migration in prostate cancer cells. *Eur J Cancer*. 2006;42(16):2838-45. doi: 10.1016/j.ejca.2006.06.024, PMID 16973347.
19. Park SJ, Kim HT, Lee DH, Kim KP, Kim SW, Suh C, *et al.* Efficacy of epidermal growth factor receptor tyrosine kinase inhibitors for brain metastasis in non-small cell lung cancer patients harboring either exon 19 or 21 mutation. *Lung Cancer*. 2012;77(3):556-60. doi: 10.1016/j.lungcan.2012.05.092, PMID 22677429.
20. Kerru N, Gummidi L, Maddila S, Gangu KK, Jonnalagadda SB. A review on recent advances in nitrogen-containing molecules and their biological applications. *Molecules*. 2020;25(8):1909. doi: 10.3390/molecules25081909, PMID 32326131.
21. Liu L, Wang Z, Gao C, Dai H, Si X, Zhang Y, *et al.* Design, synthesis and antitumor activity evaluation of trifluoromethyl-substituted pyrimidine derivatives. *Bioorg Med Chem Lett*. 2021;51:128268. doi: 10.1016/j.bmcl.2021.128268, PMID 34302974.
22. Gelardi T, Caputo R, Damiano V, Daniele G, Pepe S, Ciardiello F, *et al.* Enzastaurin inhibits tumours sensitive and resistant to anti-EGFR drugs. *Br J Cancer*. 2008;99(3):473-80. doi: 10.1038/sj.bjc.6604493, PMID 18665191.
23. Goudzal A, Hadaji E, Bouachrine M, El Hamdani H, Ouammou A. QSAR and docking molecular models to predict anticancer activity on a series of azacalix (2). *Mater Today Proc*. 2022;51:1831-7. doi: 10.1016/j.matpr.2020.08.002.
24. Oyewole RO, Oyebamiji AK, Semire B. Theoretical calculations of molecular descriptors for anticancer activities of 1, 2, 3-triazole-pyrimidine derivatives against gastric cancer cell line (MGC-803): DFT, QSAR and docking approaches. *Heliyon*. 2020;6(5):e03926. doi: 10.1016/j.heliyon.2020.e03926, PMID 32462084.
25. Sun C-c, Feng L-j, Sun X-h, Yu R-l, Chu Y-y, Kang C-m. Study on the interactions of pyrimidine derivatives with FAK by 3D-QSAR, molecular docking and molecular dynamics simulation. *New J Chem*. 2020;44(45):19499-507. doi: 10.1039/D0NJ02136A.
26. Li G, Fu L, He Q, Hu Y, Su X, Liu H, *et al.* Structural optimization for pyrimidine analogues inhibitors against MAP kinase interacting serine/threonine kinase 1(MNK1) based on molecular simulation. *J Mol Struct*. 2021;1243:130688. doi: 10.1016/j.molstruc.2021.130688.
27. V-lifeMDS;2004. Molecular design suite. Version 4.3.
28. Pratim Roy PP, Paul S, Mitra I, Roy K. On two novel parameters for validation of predictive QSAR models. *Molecules*. 2009;14(5):1660-701. doi: 10.3390/molecules14051660, PMID 19471190.
29. Chitre TS, Asgaonkar KD, Patil SM, Kumar S, Khedkar VM, Garud DR. QSAR, docking studies of 1,3-thiazinan-3-yl isonicotinamide derivatives for antitubercular activity. *Comput Biol Chem*. 2017;68:211-8. doi: 10.1016/j.compbiolchem.2017.03.015, PMID 28411471.
30. Bhadoriya KS, Kumawat NK, Bhavthankar SV, Avchar MH, Dhupal DM, Patil SD, *et al.* Exploring 2D and 3D QSARs of benzimidazole derivatives as transient receptor potential melastatin 8 (TRPM8) antagonists using MLR and kNN-MFA methodology. *J Saudi Chem Soc*. 2016;20:S256-70. doi: 10.1016/j.jscs.2012.11.001.
31. Jawarkar RD, Bakal RL, Zaki MEA, Al-Hussain S, Ghosh A, Gandhi A, *et al.* QSAR based virtual screening derived identification of a novel hit as a SARS CoV-229E 3CLpro Inhibitor: GA-MLRQSAR modeling supported by molecular Docking, molecular dynamics simulation and MMGBSA calculation approaches. *Arab J Chem*. 2022;15(1):103499. doi: 10.1016/j.arabjc.2021.103499, PMID 34909066.
32. Daoui O, Elkhatabi S, Chtita S, Elkhlabi R, Zgou H, Benjelloun AT. QSAR, molecular docking and ADMET properties *in silico* studies of novel 4,5,6,7-tetrahydrobenzo[D]thiazol-2-yl derivatives derived from dimedone as potent anti-tumor agents through inhibition of C-Met receptor tyrosine kinase. *Heliyon*. 2021;7(7):e07463. doi: 10.1016/j.heliyon.2021.e07463, PMID 34296007.
33. Mishra M, Agarwal S, Dixit A, Mishra VK, Kashaw V, Agrawal RK, *et al.* Integrated computational investigation to develop molecular design of quinazoline scaffold as promising inhibitors of plasmodium lactate dehydrogenase. *J Mol Struct*. 2020;1207:127808. doi: 10.1016/j.molstruc.2020.127808.
34. Wolber G, Dornhofer AA, Langer T. Efficient overlay of small organic molecules using 3D pharmacophores. *J Comput Aid Mol Des*. 2006;20(12):773-88. doi: 10.1007/s10822-006-9078-7, PMID 17051340.
35. Ajay Kumar TV, Kabilan S, Parthasarathy V. Screening and toxicity risk assessment of selected compounds to target cancer using QSAR and pharmacophore modelling. *Int J PharmTech Res*. 2017;10(4):219-24. doi: 10.20902/IJPTR.2017.10428.
36. Filimonov DA, Lagunin AA, Glorizova TA, Rudik AV, Druzhilovskii DS, Pogodin PV, *et al.* Prediction of the biological activity spectra of organic compounds using the PASS online web resource. *Chem Heterocycl Comp*. 2014;50(3):444-57. doi: 10.1007/s10593-014-1496-1.
37. Dmitriev AV, Filimonov DA, Rudik AV, Pogodin PV, Karasev DA, Lagunin AA, *et al.* Drug-drug interaction prediction using pass. *SAR QSAR Environ Res*. 2019;30(9):655-64. doi: 10.1080/1062936X.2019.1653966, PMID 31482727.
38. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comp Chem*. 2010;31(2):455-61. doi: 10.1002/jcc.21334, PMID 19499576.
39. BIOVIA Dassault Syst, [Discovery Studio visualizer]. v21.1.0.20298. San Diego: Dassault Systèmes;2020.
40. Stamos J, Sliwowski MX, Eigenbrot C. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J Biol Chem*. 2002;277(48):46265-72. doi: 10.1074/jbc.M207135200, PMID 12196540.
41. Seeliger D, de Groot BL. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aid Mol Des*. 2010;24(5):417-22. doi: 10.1007/s10822-010-9352-6, PMID 20401516.
42. Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc*. 2016;11(5):905-19. doi: 10.1038/nprot.2016.051, PMID 27077332.

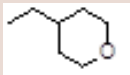


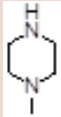
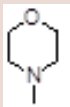
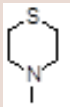
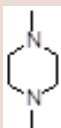
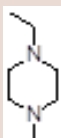
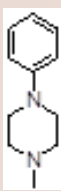
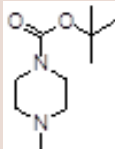
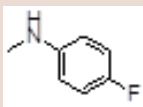
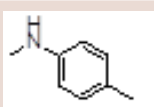
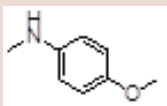
Cite this article: Asgaonkar K, Tanksali S, Abhang K, Shevate K, Patil S, Chitre T. Designing of Trifluoromethyl Substituted Pyrimidine Pharmacophore for Antiprostata Activity through a Collective Computational Approach. *Indian J of Pharmaceutical Education and Research*. 2023;57(3):827-37.

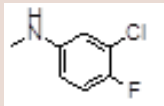
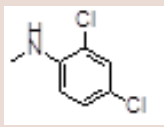
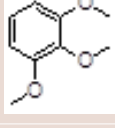
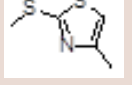
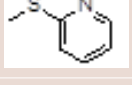
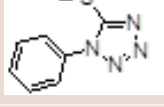
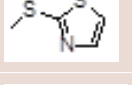
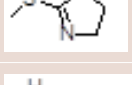
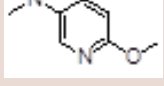
Table S1: Data set for QSAR.



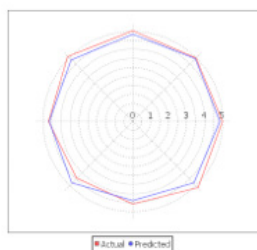
Molecule No	Compounds	R ₁ (16a-16z)/R ₂ (17a-17z)	IC ₅₀ (μM)	pIC ₅₀
1	16a		17.81	4.74
2	16b*		9.05	5.04
3	16c*		15.05	4.82
4	16d		5.65	5.24
5	16e		5.92	5.22
6	16f*		7.96	5.09
7	16g*		15.82	4.80
8	16h		11.53	4.93
9	16i		9.96	5.00
10	16j		9.22	5.03

Molecule No	Compounds	R ₁ (16a-16z)/R ₂ (17a-17z)	IC ₅₀ (μM)	pIC ₅₀
11	16k		8.57	5.06
12	16l		8.67	5.06
13	16m		9.99	5.00
14	16n*		10.52	4.97
15	16o		13.56	4.86
16	16p		6.32	5.19
17	16q		7.44	5.12
18	16r		8.17	5.08
19	16s		6.66	5.17
20	16t*		14.55	4.83
21	16u		14.16	4.84
22	16v		11.53	4.93
23	16w		12.48	4.90

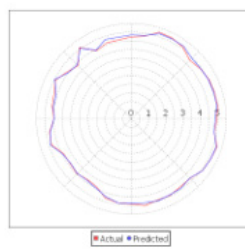
Molecule No	Compounds	R ₁ (16a-16z)/R ₂ (17a-17z)	IC ₅₀ (μM)	pIC ₅₀
24	16x		8.96	5.04
25	16y		12.87	4.89
26	16z		12.63	4.89
27	17a*		15.88	4.79
28	17b		40.84	4.38
29	17c		46.82	4.32
30	17d		34.67	4.46
31	17e		27.21	4.56
32	17f		42.74	4.36
33	17g*		40.61	4.39
34	17j		14.13	4.84
35	17k		16.46	4.78
36	17l		35.34	4.45

Molecule No	Compounds	R ₁ (16a-16z)/R ₂ (17a-17z)	IC ₅₀ (μM)	pIC ₅₀
37	17m		34.97	4.45
38	17n*		20.59	4.68
39	17o		23.15	4.63
40	17p		19.51	4.70
41	17q		11.96	4.92
42	17s		25.21	4.59
43	17u		45.26	4.34
44	17v		7.88	5.10
45	17x		40.59	4.39
46	17y		23.83	4.62
47	17z		7.35	5.13

*Test set molecules.



a)



b)

Figure S1: Radar plot for (a) Training set, (b) Test set.