Arabian Journal of Chemical and Environmental Research

Vol. 13 Issue 1 (2026) 18–40

DFT - QSAR Model Generation of Pyrimidocarbazole Derivatives as Breast Cancer Inhibitors by the Genetic Algorithm and Multiple Linear Regression (GA- MLR) Method

Abdulrahman Ibrahim Kubo[1*], Hassan Garba Wafi[1], Chika Attama[1], Anthony E. Aiwonegbe[2], Janet Hadi John[1] and Saminu M.A.[1]

[1]*Department of Pure and Applied Chemistry, Adamawa State University, Mubi, Nigeria*
[2]*Department of Chemistry, Faculty of Physical Sciences, University of Benin, Benin City, Nigeria*
Corresponding Author Email: abdulrahmankuboibrahim@gmail.com, Phone number: +2348066225336

Received 18 Sept 2025, Revised 21 Oct 2025, Accepted 22 Oct 2025

Cited as: *Kubo A.I., Wafi H.G., Attama C., Aiwonegbe A.E., John J.H. and Saminu M.A. (2026) DFT- QSAR Model Generation of Pyrimidocarbazole Derivatives as Breast Cancer Inhibitors by the Genetic Algorithm and Multiple Linear Regression (GA- MLR) method, Arab. J. Chem. Environ. Res. 13(1), 18-40*

Abstract: Cancer contributes to roughly 13% of global deaths each year, ranking among the top causes of mortality worldwide. In high-income countries, it is responsible for over 20% of all deaths, underscoring its significant impact. Among women, breast cancer is one of the most common long-term cancers and accounts for approximately 24% of all female cancer cases, making it a leading cause of cancer-related deaths. This study developed quantitative structure–activity relationship (QSAR) models using the Multiple Linear Regression combined with Genetic Function Approximation (MLR-GFA) approach. Among the models generated, Model 1 showed superior performance and was chosen for further analysis, displaying strong statistical metrics: $R^2 = 0.9806$, adjusted $R^2 = 0.9767$, SEE = 0.0548, MAE = 0.0511, Q^2 (LOO) = 0.9714 and CCC = 0.9384. Molecular docking studies revealed high binding affinities between the compounds and the target receptor, ranging from (-28.715 to -30.308 kcal/mol). Notably, Compound 16 demonstrated the strongest binding affinity (-30.308 kcal/mol), followed by Compounds 30 (-29.721 kcal/mol), 21 (-29.648 kcal/mol), 6 (-29.475 kcal/mol), and 27 (-28.715 kcal/mol). Drug-likeness and toxicity assessments using SwissADME and ProTox-3.0 confirmed the compounds favorable oral bioavailability profiles. These findings suggest that, with careful evaluation of their potential carcinogenic risks, the identified compounds hold promise as candidates for further clinical development.

Keywords: QSAR, Breastcancer, ProTox- 3.0, Drug- likeness, Docking study

*Corresponding author. Abdulrahman Ibrahim Kubo
E-mail address:abdulrahmankuboibrahim@gmail.com

1. Introduction

Cancer is responsible for approximately 13% of global deaths annually, making it one of the leading causes of mortality worldwide (Mahdy *et al.,* 2019). In high income countries, it accounts for over 20% of all deaths, highlighting its significant relative impact. Targeted therapies for cancers such as breast cancer have been approved, offering more precise treatment options. These therapies work by interfering with specific molecules involved in tumor growth and progression, thereby limiting the adverse effects of traditional non selective chemotherapy and overcoming resistance often seen with existing anticancer drugs (Garber 2008). Cancer treatment encompasses several strategies, including the inhibition of angiogenesis an essential process for tumor development making it a promising therapeutic target (Kerbel, 2008, Titi *et al.,* 2020). Breast cancer is one of the most prevalent long-term cancers among women and remains a leading cause of cancer related deaths (Raja *et al.,* 2009), accounting for about 24% of all female cancer cases (Xiao *et al.,* 2018). Common risk factors among breast cancer patients include older age, minimal or no breastfeeding, weight gain, late age at first childbirth, and sedentary lifestyle (Liu *et al.,* 2018). Advancements in understanding cancers molecular mechanisms and pathways have significantly contributed to the discovery of new anticancer agents (Bhaumik *et al.,* 2019; Bouslamti *et al.,* 2023; Zafar *et al.,* 2025). Personalized and targeted cancer therapies have emerged; however, these methods are often costly and not always effective, indicating an urgent need for alternative treatment approaches (Nagireddy *et al.,* 2019; Bouammali *et al.,* 2024). To ensure the effectiveness of new drugs, their binding affinity to a specific therapeutic target must be thoroughly evaluated. DFT, QSAR and Docking analysis is a powerful tool in this regard, as it helps determine the drug' ability to reach its target, produce the desired effect, and be metabolized within a suitable timeframe. This analysis also helps minimize the risk of failure in later stages of drug development (Kavallaris 2010; Faris *et al.,* 2023; Salih *et al.,* 2023; Alruwaili *et al.,* 2025; Kadda *et al.,* 2025; Pasha and Manojmouli *et al.,* 2025).

This study aims to explore the novel derivatives of pyrimidocarbazole with vigorous anticancer activity. It involves building a Quantitative Structure - Activity Relationship (QSAR) model to predict the anti-proliferative effects of these compounds, and performing molecular docking studies with the (3ERT) protein receptor to analyze drug target interactions. The ultimate goal is to discover effective and less toxic treatment options for breast cancer.

2. Materials and methods

2.1. Selection of compounds

In this study, a dataset comprising 43 pyrimidocarbazole derivatives and their reported anticancer activities against breast cancer was obtained from the research conducted by Mohareb R.M and

colleagues (Mohreb *et al.,* 2017). The anticancer activity data, originally expressed in IC_{50} (μM), were converted to pIC_{50} values using equation one. These compounds, along with their corresponding activities, are presented in **Table 1**;

$$pIC_{50} = -log_{10}(IC_{50} \times 10^{-6}) \quad (1)$$

2.2. Computational resources

Several software tools were employed throughout this computational analysis. ChemDraw v12.1, Spartan 14 v1.1.0 and ICMpro Docking Software were used for molecular design, quantum calculations, and docking studies, respectively. The receptor-ligand interactions were further analyzed and visualized in both 2D and 3D using Discovery Studio software. The entire research was performed on an HP EliteBook laptop featuring a dual-core Intel processor at 2.5 GHz with 4 GB of RAM, running on Windows 8. Additionally, online platforms such as SwissADME and ProTox-3 were used to predict drug-likeness and evaluate the ADMET profiles of the studied compounds (Isa *et al.,* 2024; Abdulrahman & Ibrahim, 2024).

2.3. Density functional theory

For quantum chemical calculations, Density Functional Theory (DFT) was utilized using Spartan version 14. The molecular geometries of all 43 compounds were optimized using the (B3LYP) Becke's three-parameter hybrid functional with the Lee-Yang-Parr correlation functional together with the 6-31G* basis set. This approach was used to determine the most stable conformations by identifying global minima on the potential energy surface (Yunusa *et al.,* 2021).

2.4. Molecular docking

The docking simulations were conducted using ICMpro Docking Software, with the protein receptor (PDB ID: 3ERT) selected for assessing the binding poses of the small molecules under investigation.

2.5. Drug-likeness and ADMET

Accessible online tools such as SwissADMEand ProTox- 3.0 were also used to evaluate the drug-likeness and toxicity characteristics of the compounds.

3. Results and discussion

3.1. QSAR Results

FINAL GA-MLR MODEL RESULTS:

MLR equation:

pIC50 = 14.33, (*SpMin2_Bhs) =-2.0132,(*JGI10) = -229.5961,(*SpMin6_Bhs) = +0.3874, (*ATSC4m) = +0.0003, (*JGI3) = -102.4387

Internal Validation metrics using training set:

Number of Training set datapoints: 31

R^2 = 0.9806, R^2(Adjusted) = 0.9767, Standard Error of Estimation (SEE) = 0.0548, Q^2(LOO) = 0.9714, SDEP(LOO) = 0.0598, Scaled average Rm^2(LOO) = 0.9602, Scaled delta Rm^2 (LOO) = 0.0209, Mean Absolute Error(MAE) = 0.0511

External Validation metrics using a Test set:

Number of Test set datapoints: 12

Q^2(F1)Test = 0.8924, Q^2(F2)Test = 0.8855, Scaled average Rm^2(Test) = 0.8445, Scaled delta Rm^2 (Test) = 0.0784, CCC (Test) = 0.9384, Mean Absolute Error(MAE, Test) = 0.106.

In this study, the Kennard-Stone algorithm was employed to divide the dataset into training (modeling) and test (validation) sets in order to assess the robustness and statistical reliability of the developed QSAR model. A total of five models were generated using the Multiple Linear Regression- Genetic Function Approximation (MLR-GFA) method. Among these, Model 1 demonstrated the best performance and was therefore selected for detailed evaluation. The statistical metrics for internal validation of this model include: coefficient of determination (R^2) = 0.9806, adjusted R^2 = 0.9767, standard error of estimation (SEE) = 0.0548, leave-one-out (Q^2_LOO) = 0.9714, standard deviation of prediction error (SDEP_LOO) = 0.0598, scaled average Rm^2 (LOO) = 0.9602, scaled delta Rm^2 (LOO) = 0.0209, and mean absolute error (MAE) = 0.0511. The high R^2 value of 0.9806 indicates that approximately 98% of the variability in the observed data is captured by the model, highlighting its strong predictive potential. The close agreement between R^2 and adjusted R^2 further confirms that the model accurately reflects the influence of the selected molecular descriptors on the pIC_{50} values, affirming the models explanatory power. External validation was also carried out to confirm the models predictive reliability. The following metrics were obtained: Q^2 (F1) = 0.8924, Q^2 (F2) = 0.8855, scaled average Rm^2 (Test) = 0.8445, scaled delta Rm^2 (Test) = 0.0784, concordance correlation coefficient (CCC) = 0.9384, and mean absolute error for the test set (MAE Test) = 0.106. These results collectively demonstrate that the model is both statistically robust and capable of making accurate predictions.

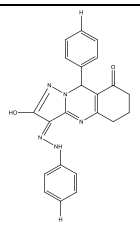
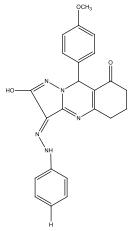
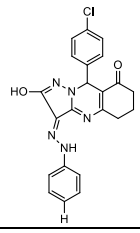
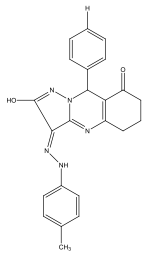
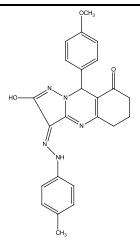
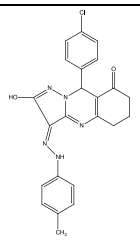
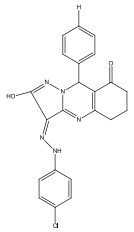
The positive coefficients of the descriptors SpMin6_Bhs and ATSC4m in the build QSAR model indicate that these variables contribute positively to the inhibitory activity of the pyrimidocarbazole compounds. This suggests that increasing the presence or influence of these descriptors within the molecular structure may enhance the potency of the compounds against their biological target, the

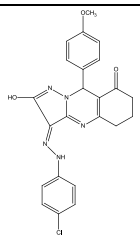
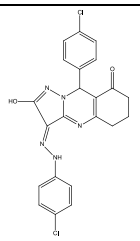
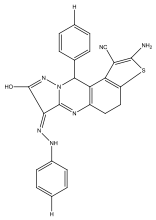
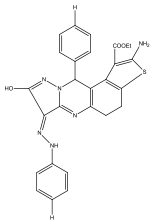
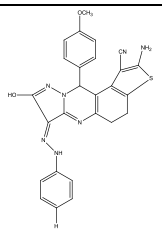
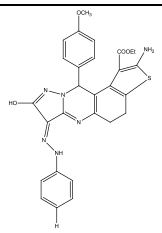
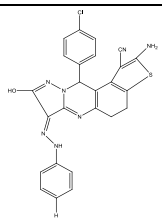
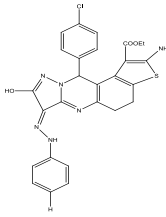
3ERT protein receptor. Conversely, the descriptors SpMin2_Bhs**,** JGI10**,** and JG13 were associated with negative coefficients, implying that they negatively impact the compounds inhibitory activity. Therefore, reducing the values of these descriptors in the molecular design could potentially improve the compounds efficacy toward the target receptor (3ERT). **Table 1** presents the 2D molecular structures and associated descriptor values of the pyrimidocarbazole derivatives. **Table 2** provides detailed results, including experimental and predicted pIC_{50} values, residuals, docking scores, and hydrogen bonding interactions for all the studied compounds. The close agreement between experimental and predicted pIC_{50} values, reflected in the low residuals, confirms the strong predictive performance and reliability of the developed model.

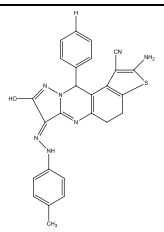
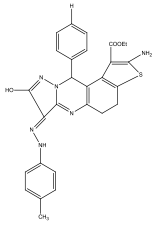
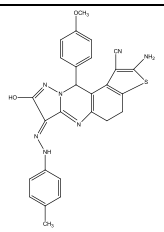
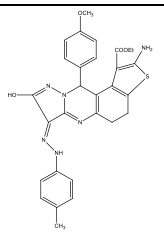
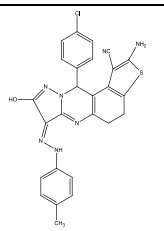
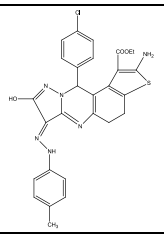
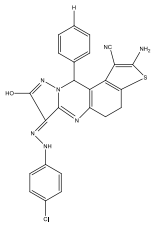
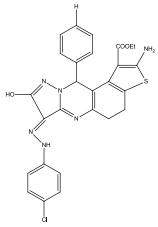
The robustness of the model was further validated through Y-scrambling tests, as shown in **Table 3** The low R^2 and Q^2 values obtained from this test indicate that the model was not developed by chance. **Table 4** contains the definitions of all descriptors included in the final QSAR model, while **Table 5** presents statistical analyses, including correlation between descriptors and their individual impacts on the model. The Variance Inflation Factor (VIF) was used to evaluate multicollinearity among descriptors and for a model to be considered valid, VIF values should fall between 1 and 10. A VIF below 1 suggests no correlation between descriptors, while a value above 10 indicates problematic multicollinearity. In this study, all VIF values were found to be between 1 and 6, as reported in **Table 5**, confirming that multicollinearity was not an issue in the dataset. This reinforces the reliability and statistical soundness of the QSAR model developed.

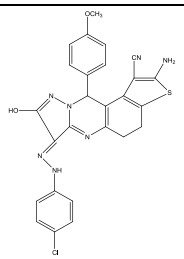
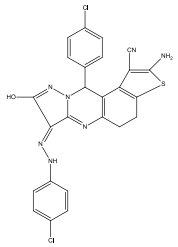
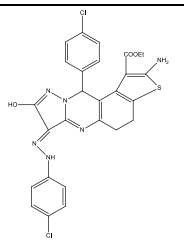
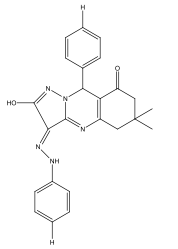
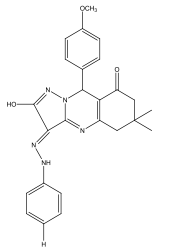
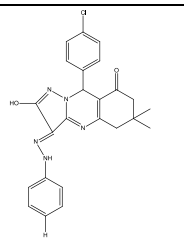
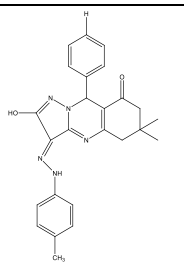
The Mean Effect (ME) values of all descriptors were calculated to evaluate their individual influence and relative contribution within the selected QSAR model. These values are shown in **Table 5**. The sign of each descriptor coefficient indicates whether it is positively or negatively impacts the compounds potency. Among all the descriptors, SpMin6_Bhs emerged as the most influential, as it had the highest (ME) value, signifying a strong positive effect on the pIC_{50} values. The descriptors were ranked according to their contributions to the biological activity pIC_{50} of the compounds, in the following descending order: SpMin6_Bhs > ATSC4m > JGI10 > SpMin2_Bhs > JGI3. **Figure 1** illustrates a scatter plot of predicted pIC_{50} versus actual pIC_{50} values for both training and test sets. The close alignment of data points along the diagonal line confirms the strong predictive accuracy of the model. **Figure 2** presents a residual plot, where residuals for both data sets are distributed evenly above and below the zero line, indicating the absence of systematic errors in the model.

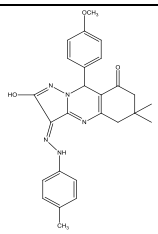
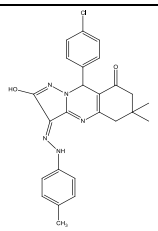
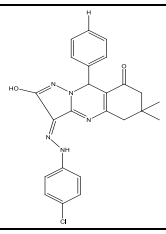
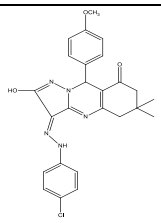
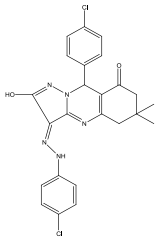
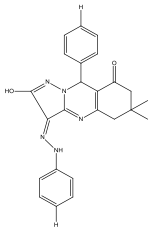
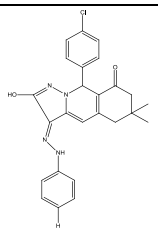
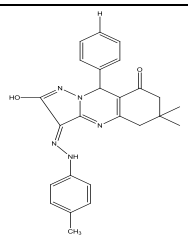
Table 1: 2D structure and descriptors of pyrimidocarbazole derivatives

S/N	Compounds	SpMin2_Bhs	JGI10	SpMin6_Bhs	ATSC4m	JGI3
1.		1.786	0.003	1.041	228.1	0.047
2.		1.784	0.003	1.081	-103.8	0.045
3.		1.792	0.003	1.036	-200.2	0.047
4.		1.836	0.004	1.149	251.2	0.047
5.		1.833	0.004	1.149	-76.375	0.044
6.		1.827	0.004	1.142	-140.2	0.046
7.		1.810	0.004	1.305	-200.2	0.047

8.		1.808	0.004	1.074	-553.0	0.044
9.		1.802	0.004	1.030	-653.8	0.046
10.		1.805	0.004	1.270	-11.744	0.050
11.		1.787	0.005	1.380	242.5	0.050
12.		1.809	0.005	1.274	-358.3	0.048
13.		1.765	0.006	1.377	-1.04.1	0.048
14.		1.821	0.004	1.273	-528.8	0.049
15.		1.803	0.005	1.378	-332.6	0.050

16.		1.850	0.005	1.280	73.437	0.049
17.		1.831	0.005	1.381	326.6	0.050
18.		1.847	0.006	1.284	-272.1	0.047
19.		1.829	0.006	1.380	-6.324	0.048
20.		1.843	0.005	1.283	-401.8	0.049
21.		1.824	0.005	1.377	-208.7	0.049
22.		1.826	0.005	1.264	-528.8	0.049
23.		1.807	0.005	1.374	-332.6	0.050

24.		1.824	0.006	1.269	-888.2	0.047
25.		1.891	0.005	1.267	-1065	0.049
26.		1.800	0.005	1.370	-927.2	0.042
27.		1.831	0.006	1.104	1446	0.055
28.		1.828	0.005	1.064	1769	0.058
29.		1.840	0.005	1.062	1588	0.057
30.		1.840	0.006	1.154	1723	0.057

31.		1.838	0.006	1.153	1409	0.054
32.		1.838	0.005	1.147	1556	0.056
33.		1.804	0.006	1.099	1295	0.054
34.		1.820	0.006	1.057	1588	0.057
35.		1.833	0.005	1.055	1410	0.056
36.		1.828	0.005	1.064	1769	0.058
37.		1.845	0.005	1.070	1399	0.057
38.		1.838	0.006	1.153	1409	0.054

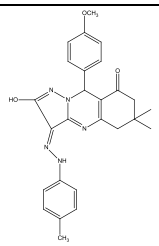
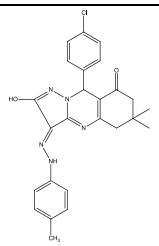
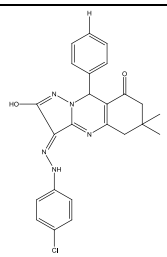
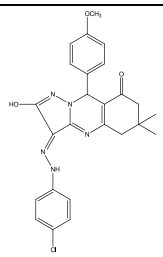
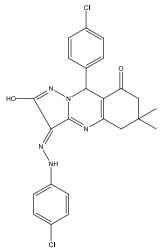
39.		1.845	0.005	1.070	1399	0.057
40.		1.840	0.006	1.154	1723	0.057
41.		1.838	0.006	1.153	1409	0.054
42.		1.838	0.005	1.147	1556	0.056
43.		1.820	0.006	1.057	1588	0.057

Table 2: Experimental pIC_{50}, Predicted pIC_{50}, Residual, Docking score and H- Bond of pyrimidocarbazole derivatives

S/NO	Experimental pIC_{50}	Predicted pIC_{50}	Residue	Docking score	H-Bond
*1.	6	5.607	-0.393	-21.459	-1.571
2.	5.698	5.683	-0.015	-24.846	-2.954
3.	5.522	5.461	-0.061	-22.000	-1.623
4.	5.397	5.342	-0.055	-28.610	-2.979
*5.	5.221	5.249	0.028	-26.323	-2.963

6.	5.301	5.441	0.14	-29.475	-3.016
7.	5.154	5.199	0.045	-23.644	-1.656
*8.	5.096	5.304	0.208	-26.098	-1.172
9.	5.045	5.086	0.041	-26.623	-1.172
*10.	5	5.007	0.007	-25.400	-3.040
11.	4.958	4.981	0.023	-22.965	-1.255
12.	4.920	4.875	-0.045	-21.419	-3.240
13.	4.886	4.904	0.018	-25.657	-3.025
*14.	4.853	4.909	0.056	-18.268	-1.579
15.	4.823	4.788	-0.044	-23.693	-1.726
16.	4.795	4.805	0.01	-30.308	-2.902
17.	4.769	4.828	0.059	-26.481	-1.264
*18.	4.744	4.741	-0.003	-26.210	-2.214
19.	4.721	4.749	0.028	-24.803	-1.243
*20.	4.698	4.780	0.082	-19.724	-0.367
21.	4.677	4.708	0.031	-29.684	-2.377
*22.	4.657	4.647	-0.01	-21.515	-3.162
23.	4.638	4.656	0.018	-26.527	-3.167
24.	4.619	4.579	-0.04	-20.232	-0.584
25.	4.585	4.603	0.018	-23.473	-1.253
26.	4.568	4.517	-0.051	-23.071	-0.534
27.	4.552	4.468	-0.084	-28.715	-1.480
*28.	4.537	4.504	-0.033	-18.068	-0.925
29.	4.522	4.547	0.025	-19.810	-1.673
30.	4.508	4.363	-0.145	-29.721	-5.168

31.	4.494	4.430	-0.064	-27.012	-3.078
32.	4.481	4.484	0.003	-24.746	-3.052
33.	4.468	4.321	-0.147	-23.746	-1.583
*34.	4.455	4.439	-0.016	-20.513	0
35.	4.443	4.410	-0.033	-20.735	-1.338
36.	4.431	4.468	0.037	-19.011	0
37.	4.408	4.477	0.069	-24.687	-3.053
38.	4.387	4.430	0.043	-19.777	-1.671
39.	4.397	4.363	-0.034	-21.601	-1.340
*40.	4.376	4.484	0.108	-27.073	-3.078
41	4.366	4.321	-0.045	-21.179	-1.669
*42	4.356	4.439	0.083	-22.292	-1.332
43.	4.346	4.410	0.064	-26.709	-3.061

*Denote test set

Table 3: Y- Scramble

MODEL TYPE	R^2	Q^2-LOO
Original	0.9806	0.9714
Random 1	0.0394	-0.4269
Random 2	0.1635	-0.3774
Random 3	0.0592	-0.4816
Random 4	0.0916	-0.3437
Random 5	0.2025	-0.1995
Random 6	0.1731	-0.5007
Random 7	0.1743	-0.2524
Random 8	0.1107	-0.4497
Random 9	0.2133	-0.2697
Random 10	0.1570	-0.2846
Summary:		
R^2	Original Model	0.9806
Q^2-LOO	Original Model	0.9714
Average R^2	10 Random Models	0.1384
Average Q^2-LOO	10 Random Models	-0.3586

Table 4: Definition of descriptors and their class for model 1

Descriptors	Definition	Class
SpMin2_Bhs	Smallest absolute eigenvalue of Burden modified matrix - n 2 / weighted by relative I-state	2D
JGI10	Mean topological charge index of order 10	2D
SpMin6_Bhs	Smallest absolute eigenvalue of Burden modified matrix - n 6 / weighted by relative I-state	2D
ATSC4m	CenteredBroto-Moreau autocorrelation - lag 4 / weighted by mass	2D
JGI3	Mean topological charge index of order 3	2D

Table 5: Statistical analysis of model 1 parameters

	SpMin2_Bhs	*JGI10*	*SpMin6_Bhs*	*ATSC4m*	*JGI3*	*VIF*	*M/E*
SpMin2_Bhs	1	0.272475	-0.22825	0.489965	0.498242	1.3471	0.2458
JGI10	0.272475	1	0.258818	0.416633	0.534294	2.2261	0.2512
SpMin6_Bhs	-0.22825	0.258818	1	-0.54233	-0.3848	2.3777	0.2574
ATSC4m	0.489965	0.416633	-0.54233	1	0.9026	7.3573	0.2544
JGI3	0.498242	0.534294	-0.3848	0.9026	1	6.5690	-6.1977

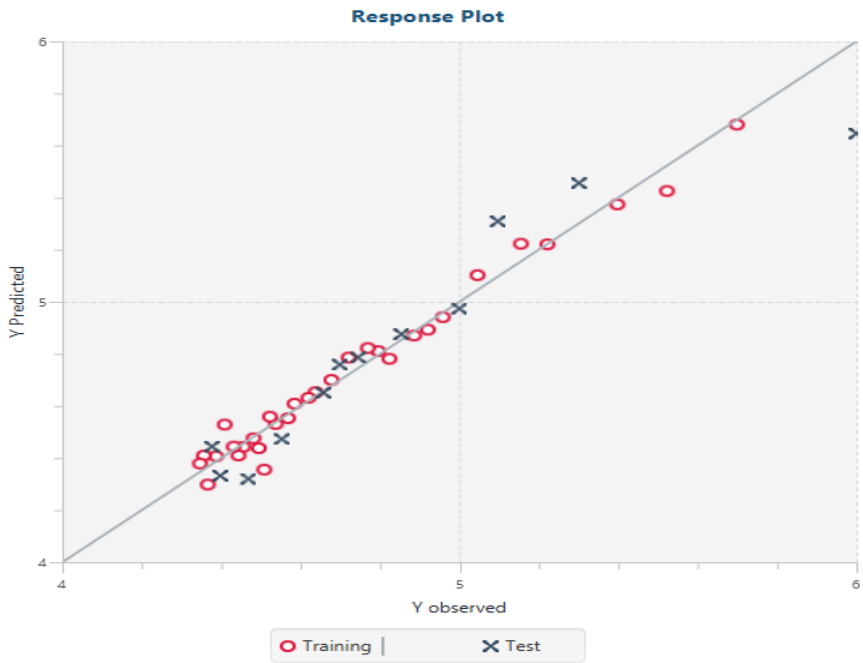


Figure 1: Plot of the predicted pIC_{50} versus actual pIC_{50} for the test and training sets compounds

3.2 Docking results

Molecular docking analysis is used to understand how two or more molecules interact such as a drug binding to a protein receptor. Docking tools are widely used in drug discovery, especially for virtual screening, which helps identify promising compounds for further study from a large molecular database. This requires computational tools that are both efficient and reliable (Yang *et al.,* 2011). To visualize the molecular interactions, including hydrogen bonds, salt- bridged, alkyl

related and Pi-related interactions, Discovery Studio software was used. The ligand- receptor interactions between the pyrimidocarbazole derivatives and the target receptor were analyzed and are displayed in **Table 6**. The docking results in this study showed binding affinities ranging from (-28.715 and -30.308 kcal/mol), indicating strong interactions between the compounds and the receptor. Specifically, Compound 16 exhibited the highest binding affinity **(**-30.308 kcal/mol**),** followed by Compound 30 **(**-29.721 kcal/mol**),** Compound 21 (**-**29.648 kcal/mol**),** Compound 6 (-29.475 kcal/mol) and Compound 27 (-28.715 kcal/mol) respectively**..**

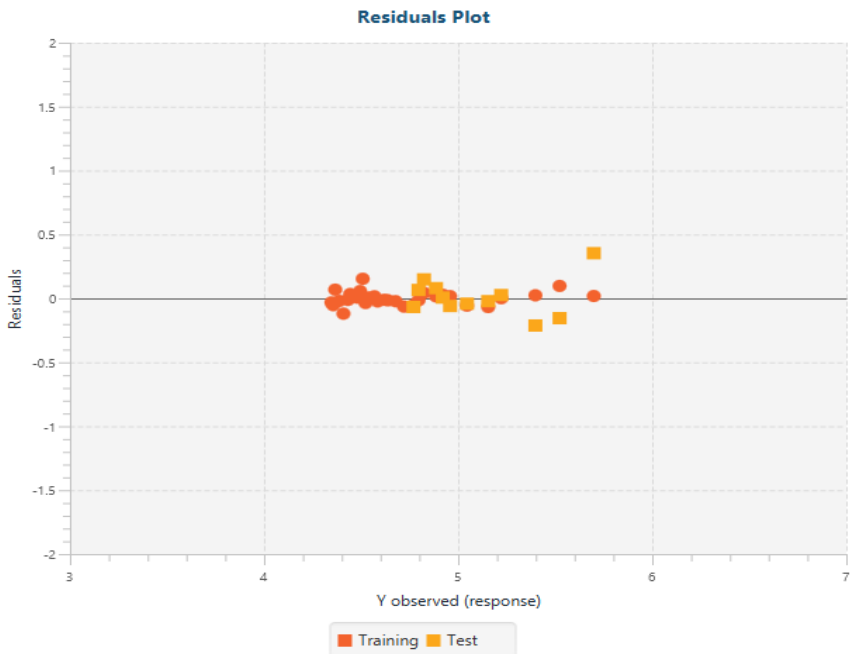


Figure 2: Scatter plot of the residuals against actual pIC_{50}

Compound 16 has the highest docking score (-30.308 kcal/mol) formed one conventional hydrogen bonds with LYS530 at a distance of (3.93 Å), Pi- sulfur with MET343 at a distance of (7.56 Å), MET522 at a distance of (5.65 Å). Additionally, it interacted with TRY526 at a distance of (4.76 Å) and CYS530 at a distance of (4.76 Å). It formed unfavourable donor- donor with CYS530 at a distance of (3.71 Å), it also interacted with Pi- Pi T- shaped with TYR528 at a distance of (4.76 Å), Alkyl with VAL533 at a distance of (5.57 Å), LEU346 at a distance of (4.75 Å), and Pi- alkyl with MET522 at a distance of (4.57 Å), LEU525 at a distance of (4.27 Å), ALA350 at a distance of (4.75 Å) and (6.59 Å) respectively. The following amino acid residue LYS, MET, TRY, CYS, VAL, LEU and ALA might be the reason why compound 16 has higher binding affinity. Both 3D and 2D visual representations of compound 16 bound to the receptor are shown in **Figure 3**.

Compound 30 with docking score (-29.721 kcal/mol) formed two conventional hydrogen bond with CYS530 at a distance of (3.48 Å), MET522 at a distance of (3.18 Å) and (6.25 Å). It also formed

carbon hydrogen bond with TYR526 at a distance of (4.43 Å), LYS529 at a distance of (4.99 Å). Pi-sulfur was observed to interact with CYS381 at a distance of (6.85 Å), Pi- Pi T-shaped interacted with TYR526 at a distance of (4.77 Å). It formed alkyl with VAL533 at a distance of (4.59 Å), LYS529 at a distance of (4.91 Å) and Pi- alkyl with LEU525 at a distance of (5.03 Å) and (4.91 Å), MET522 at a distance of (4.81 Å) and (6.25 Å) respectively. Both 3D and 2D visual representations of compound 30 bound to the receptor are shown in **Figure 4**.

Compound 21 with docking score (-21.684 kcal/mol) formed two conventional hydrogen bond with MET343 at a distance of (4.27 Å) and VAL534 at a distance of (5.03 Å). It formed sulfux-x with MET343 at a distance of (7.08 Å), (8.24 Å) and (5.03 Å) respectively. Amide Pi- stacked is interacted with LEU346 at a distance of (5.57 Å), alkyl with LEU391 at a distance of (5.99 Å) and ALA350 at a distance of (4.77 Å). Pi- alkyl is observed to interact with TRP383 at a distance of (5.21 Å), LEU536 at a distance of (5.85 Å) and (3.91 Å), VAL533 at a distance of (5.11 Å) and (6.68 Å), CYS530 at a distance of (5.64 Å), LEU 525 at a distance of (4.58 Å) and (4.91 Å). Both 3D and 2D visual representation of compound 21 bound to the receptor are shown in **Figure 5**.

Compound 6 with docking score (-29.475 kcal/mol) formed one conventional hydrogen bond with CYS530 at a distance of (3.81 Å), unfavourable donor- donor with CYS530 at a distance of (3.99 Å). It formed Pi- sulfur with CYS530 at a distance of (4.65 Å), MET343 at a distance of (7.83 Å), Pi- alkyl with LYS529 at a distance of (7.31 Å), VAL533 at a distance of (4.70 Å) and (4.25 Å). It also formed PRO535 at a distance of (3.79 Å), LEU523 at a distance of (4.72 Å) and (4.86 Å), ALA350 at a distance of (6.50 Å) and (4.86 Å) and alkyl with LEU346 at a distance of (4.93 Å) and MET522 at a distance of (5.23 Å) respectively. Both 3D and 2D visual representation of compound 6 bound to the receptor are shown in **Figure 6**.

Compound 27 with docking score (-28.715 kcal/mol) formed conventional hydrogen bond with CYP530 at a distance of (3.75 Å), carbon hydrogen bond with LYS529 at a distance of (4.49 Å). It formed Pi- sulfur with TYR526 at a distance of (4.57 Å), MET343 at a distance of (6.87 Å) and CYP530 at a distance of (5.58 Å). Pi- amide interacted with LEU525 at a distance of (6.87 Å), LEU346 at a distance of (6.95 Å), alkyl with LEU346 at a distance of (4.06 Å), LYS529 at a distance of (6.44 Å), LEU539 at a distance of (4.96 Å) and MET528 at a distance of (5.71 Å). It also formed Pi- alkyl with ALA350 at a distance of (5.13 Å) and (5.39 Å), LEU525 at a distance of (4.81 Å), (5.24 Å) and (4.02 Å), VAL533 at a distance of (5.32 Å) and (5.62 Å) respectively. Both 3D and 2D visual representation of compound 27 bound to the receptor are shown in **Figure 7.**

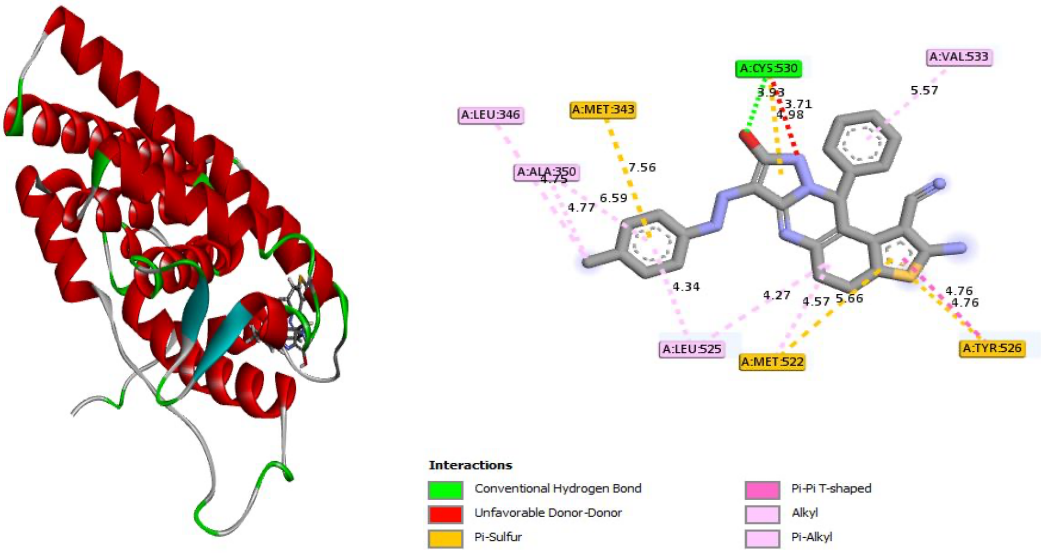


Figure 3: 3D and 2D representations of compound 16 in the active site of the 3ert receptor

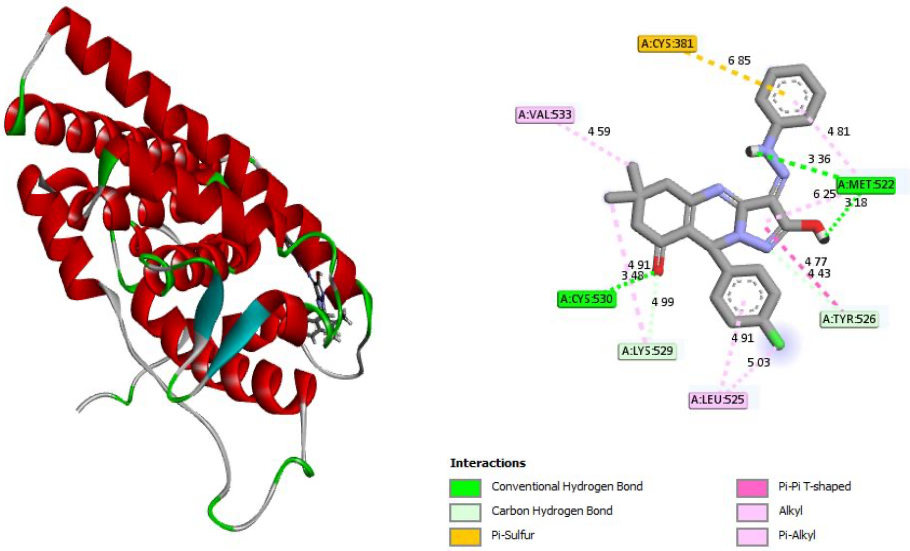


Figure 4: 3D and 2D representations of compound 30 in the active site of the 3ert receptor

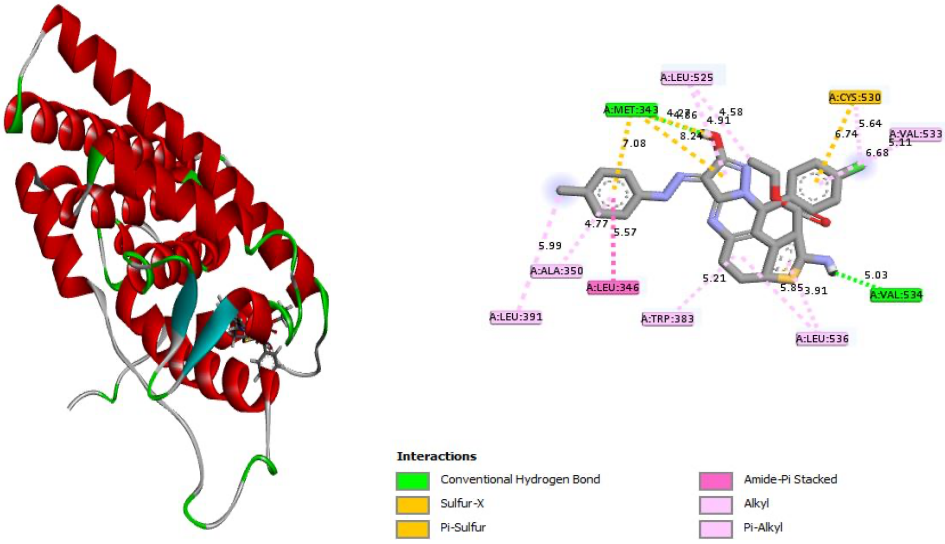


Figure 5: 3D and 2D representations of compound 21 in the active site of the 3ert receptor

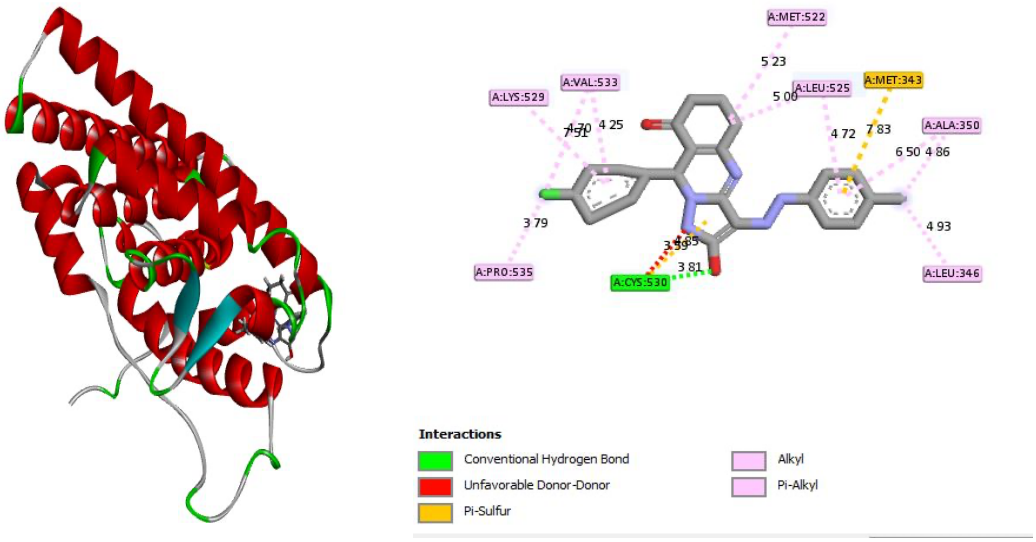


Figure 6: 3D and 2D representations of compound 6 in the active site of the 3ert receptor

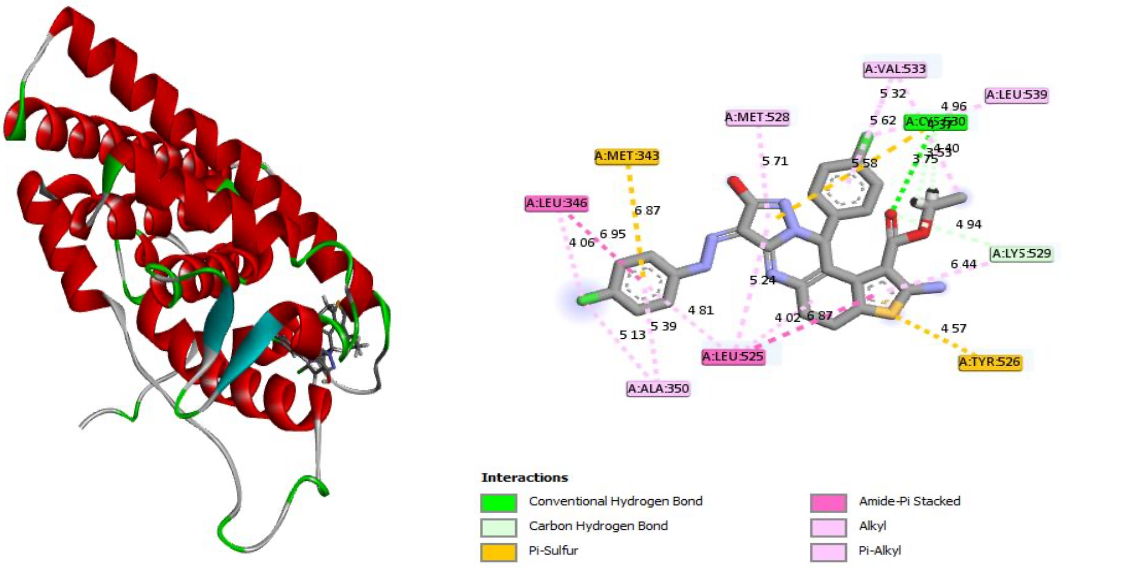


Figure 7: 3D and 2D representations of compound 27 in the active site of the 3ert receptor

3.3 Drug-likeness and Toxicity studies of the best compounds of pyrimidocarbazole

Using free online tools, SwissADME and ProTox- 3.0, various parameters were evaluated to assess the compounds drug-likeness and toxicity characteristics. The radar plots in **Figure 8** illustrate that all physico-chemical properties fall within the acceptable range defined by Lipinski's rule of five, confirming their drug-likeness and potential for acceptable oral bioavailability. The toxicity and metabolism predictions, summarized in **Table 8**, indicate that the most promising compounds exhibit moderate ADMET profiles and do not show signs of extreme toxicity. **Figure 9** and **Figure 10** indicate the molecular weight distribution for compound 16, identified as the lead candidate. In

these diagrams, the red line represents the mean molecular weight and the black line indicates the actual molecular weight, while the LD50 distribution is similarly depicted with the mean in red and the predicted median lethal dose in black. Overall, these results suggest that, with careful consideration of potential carcinogenicity, these compounds could serve as viable candidates for further clinical trials.

Table 6: Amino Acid Residues involved in major interaction with 1z8l and Distance (Å)

Compound	Hydrogen bond	Salt- bridge/ others	Pi- related	Alkyl related
16	LYS530 (3.93 Å)	LYS530	MET522, LEU525, ALA350, TYR526, CYS530, MET343	VAL533, LEU346
30	CYS530 (3.48 Å), MET522 (3.18 Å, 6.25 Å), TYR526 (4.43 Å), LYS529 (4.99 Å)	-	CYS381, TYR526, LEU525, MET522	VAL533, LYS529
21	MET343 (4.27 Å), VAL543 (5.03 Å)	MET343	CYS530, LEU346, TRP383, VAL533	LEU391, ALA350
6	CYS530 (3.81 Å)	CYS530	CYS530, MET343, LYS529, VAL533, PRO535, LEU523, ALA350	LEU346, MET522
27	CYP530 (3.75 Å), LYS529 (4.49 Å)	-	TYR526, MET343, CYP530, LEU525, ALA350, VAL533	LEU346, LYS529, MET528

Table 7: Drug- likeness of the inhibiting compounds under investigation

S/No	MW (g/mol)	HBA	HBD	MLOGP	P-gb Subtrate	Fraction Csp3	Bioavailability score	Synthetic accessibility	Lipinski
16	479.56	5	3	3.01	NO	0.15	0.55	5.06	YES
30	447.92	5	2	3.77	NO	0.25	0.55	4.98	YES
21	561.5	6	3	4.18	NO	0.21	0.17	5.38	NO
6	433.89	5	2	3.57	NO	0.22	0.55	4.78	YES
27	581.47	6	3	4.45	NO	0.19	0.17	5.26	NO

Table 8: Toxicity and Metabolism of the inhibiting compounds under investigation

		Ligand 16	Ligand 30	Ligand 21	Ligand 6	Ligand 27
Classification	**Target**	**Prediction**	**Prediction**	**Prediction**	**Prediction**	**Prediction**
Toxicity end point	Carcinogenicity	Inactive	Inactive	Inactive	Active	Inactive
Toxicity end point	Mutagenicity	Inactive	Inactive	Inactive	Inactive	Inactive
Toxicity end point	Cytotoxicity	Inactive	Inactive	Inactive	Inactive	Inactive
Toxicity end point	BBB- barrier	Inactive	Inactive	Inactive	Inactive	Inactive
Metabolism	Cytochrome CYP1A2	Inactive	Inactive	Inactive	Inactive	Inactive
Metabolism	Cytochrome CYP2C19	Inactive	Inactive	Inactive	Inactive	Inactive
Metabolism	Cytochrome CYP2C9	Active	Active	Active	Active	Active
Metabolism	Cytochrome CYP2D6	Inactive	Inactive	Inactive	Inactive	Inactive
Metabolism	Cytochrome CYP3A4	Active	Active	Active	Active	Active
Metabolism	Cytochrome CYP2E1	Inactive	Inactive	Inactive	Inactive	Inactive

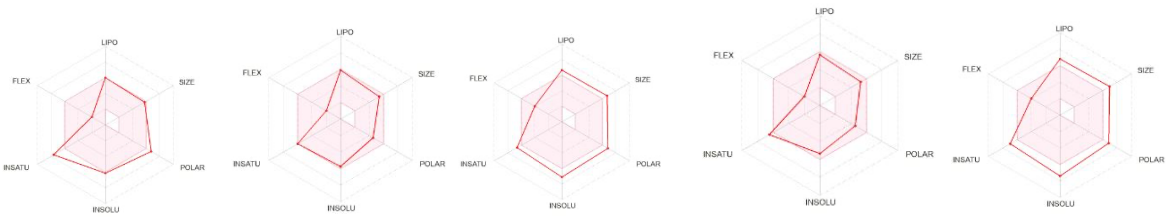


Figure 8: Radar for compound 16, 30, 21, 6 and 27

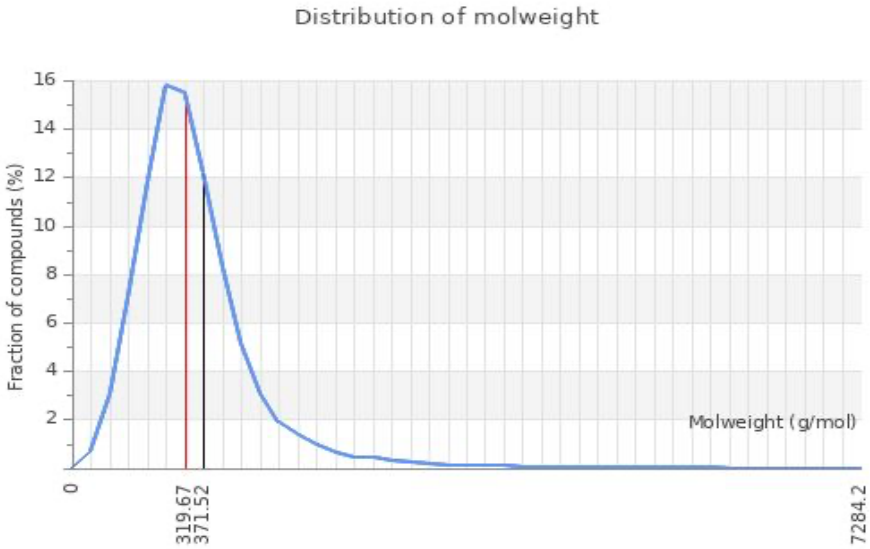


Figure 9: The molecular weight (MW) distribution of compound 16

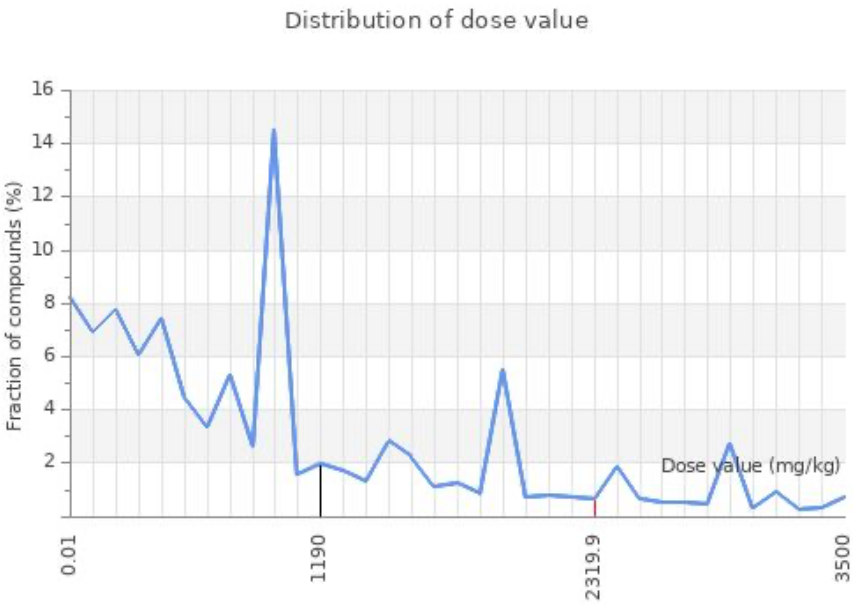


Figure 10: The distributionof LD50 valuesofcompound 16

Conclusion

This study employed a dataset of 43 pyrimidocarbazole derivatives with reported anticancer activity against breast cancer. Density Functional Theory (DFT) calculations were performed using Spartan 14 software, applying the B3LYP functional and the 6-31G* basis set to identify the most stable conformations by locating the global minima on the potential energy surface. The QSAR model developed showed a strong predictive ability, as evidenced by a high R^2 value of 0.9806, indicating that about 98% of the variation in the observed activity data is explained by the model. Among the molecular descriptors used, SpMin6_Bhs and ATSC4m had positive coefficients, suggesting that these features contribute favorably to the inhibitory activity of the pyrimidocarbazole compounds. Enhancing these descriptors in the molecular structure could potentially increase their effectiveness against the 3ERT protein receptor. On the other hand, descriptors such as SpMin2_Bhs, JGI10, and JG13 exhibited negative coefficients, indicating a detrimental effect on the compounds' inhibitory activity. Thus, minimizing the presence or influence of these descriptors in molecular design could improve the compounds' efficacy against the target receptor (3ERT).

Declaration of interest statement; Authors declare no conflict of interest.

Source of funding: This research did not receive any specific grant from funding agencies in the public, commercial or not- for- profit sectors.

Authors contribution : The authors contributed equally in drafting the manuscript

Reference

Alruwaili M, Elsaman T, Mohamed MA, Elderdery AY, Mills J, Alruwaili Y, Hamza SMA, Mekki SEI, Alotaibi HA, Alrowily MJ and Althobiti MM (2025) Molecular docking, free energy calculations, ADMETox studies, DFT analysis, and dynamic simulations highlighting a chromene glycoside as a potential inhibitor of PknG in Mycobacterium tuberculosis. *Front. Chem.* 13, 1531152. doi: 10.3389/fchem.2025.1531152

Arjmand F., Yousuf I., Ben Hadda T., Toupet L. (2014). Synthesis, crystal structure and antiproliferative activity of Cu(II) nalidixic acid–DACH conjugate: Identification of antitumor, antifungal and antibacterial pharmacophore sites, *European Journal of Medicinal Chemistry*, 81, 76-88

Bhaumik I., Pal K., Debnath U., Karmakar P., Jana K., Misra AK. (2019) Natural product inspired allicin analogs as novel anti-cancer agents. *Journal of Bioorganic Chemistry*, 86, 259- 272. https://doi.10.1016/j.bioorg.2019.01.057.

Bouammali, H., Zraibi, L., Ziani, I., Merzouki, M., Bourassi, L., Fraj, E., Challioui, A., Azzaoui, K., Sabbahi, R., *et al.* (2024). Rosemary as a Potential Source of Natural Antioxidants and Anticancer Agents: A Molecular Docking Study. *Plants,* 13, 89. https://doi.org/10.3390/plants13010089

Bouslamti M., Loukili E.H., Elrherabi A., El Moussaoui A., Chebaibi M., Bencheikh N., Nafidi H-A., *et al.* (2023). Phenolic Profile, Inhibition of α-Amylase and α-Glucosidase Enzymes, and Antioxidant Properties of Solanum elaeagnifolium Cav. (Solanaceae): In Vitro and In Silico Investigations. *Processes.* 11(5), 1384. https://doi.org/10.3390/pr11051384

Faris A., Edder Y., Louchachha I., Ait Lahcen I., Azzaoui K., *et al*. (2023) From Himachalenes to trans-Himachalol: Unveiling Bioactivity through Hemisynthesis and Molecular Docking Analysis, *Scientific reports,* 13, 17653. https://doi.org/10.1038/s41598-023-44652-z

Gerber DE. (2008) Targeted therapies: a new generation of cancer treatments. *American Family Physician.* 77(3), 311-19. PMID: 18297955.

Isa A.S., Ibrahim A.K., Mukhtar A.M. (2024). Quinazoline derivatives as cancer therapeutic agents. *Journal of Medicinal and Medical Chemistry*, 1(3), 138-155. https://doi:10.22034/medmedchem.2025.492523.1022.

Kadda S., Merzouki M., Alaloui M.M. (2025). DFT and Docking Techniques: Bibliometric analysis, *Arab. J. Chem. Environ. Res.* 12(2), 167-178

Kavallaris M. (2010) Microtubules and resistance to tubulin-binding agents. *Erratum in Nat Rev Cancer.* 10(3), 194- 204. https://doi:10.1038/nrc2803.

Kerbel RS. (2000) Tumor angiogenesis: past, present and the near future. *Carcinogenesis.* 21(3), 505-15. https://doi:10.1093/carcin/21.3.505.

Kubo A.I., Bwatanglang I.B. (2024) QSAR, Molecular docking Studies and Pharmacokinetics Properties Prediction of some thiosemicarbazone-derivatives containing Indole Fragments Targeting prostate cancer Cell. *International Research Journal of Innovations in Engineering and Technology,* 8(8), 156-166. https://doi.org/100.47001/IRJIET/2024.808017

Liu L., Tang Z., Wu C., Li X., Huang A., Lu X., You Q., Xiang H. (2018) Synthesis and biological evaluation of 4,6-diaryl-2-pyrimidinamine derivatives as anti-breast cancer agents. *Bioorg Med Chem Lett.* 28(6), 1138- 1142. https://doi:10.1016/j.bmcl.2017.12.066.

Mahdy H.M., Mohammed K. Ibrahim., Ahmed M. Metwaly., Amany Belal, Ahmed B.M. Mehany., Kamal M.A. El-Gamal., Abdou El-Sharkawy., Iet al. (2020) Design, synthesis, molecular modeling, in vivo studies and anticancer evaluation of quinazolin-4(3H)-one derivatives as potential VEGFR-2 inhibitors and apoptosis inducers. *Journal of Bioorganic Chemistry*, 94, 103422, ISSN 0045-2068. https://doi.org/10.1016/j.bioorg.2019.103422.

Mohareb R.M., Abbas N.S., Mohamed A.A. (2017) Synthesis of tetrahydropyrazoloquinazoline and tetrahydropyrazolo-pyrimidocarbazole derivatives as potential anti-prostate cancer agents and Pim-1 kinase inhibitors. *Med Chem Res.,* 26, 1073–1088. https://doi.org/10.1007/s00044-017-1811-3.

Nagireddy P. K. R., Kommalapati V.K., Siva Krishna V., Sriram D., Tangutur A.D., Kantevari S. (2019) Imidazo[2,1-*b*]thiazole-Coupled Natural Noscapine Derivatives as Anticancer Agents. *ACS Omega,* 4(21), 19382-19398. https://doi:10.1021/acsomega.9b02789.

Pasha T.Y., Manojmouli C. (2025) Pyruvate kinase enzyme: Comprehensive overview of molecular insights, clinical trials, and development of activators for the treatment of cancer and hemolytic anemia, *Mor. J. Chem.,* 13(4), 1748-1777

Salih R.H.H., Hasan A.H., Hussen N.H., Hawaiz F.E., Hadda T.B., Jamalis J., *et al.* (2023). Thiazole-pyrazoline hybrids as potential antimicrobial agent: Synthesis, biological evaluation, molecular docking, DFT studies and POM analysis *Journal of Molecular Structure* 1282, 135191

Solomon R.V., Changkun Hu., Hoyun Lee. (2009) Hybrid pharmacophore design and synthesis of isatin–benzothiazole analogs for their anti-breast cancer activity, *Bioorganic & Medicinal Chemistry*, 17(21), 7585-7592, https://doi.org/10.1016/j.bmc.2009.08068.

Titi A., Messali M., Alqurashy B.A., Touzani R., Shiga T., *et al*. (2020). Synthesis, characterization, X-Ray crystal study and bioctivities of pyrazole derivatives: Identification of antitumor, antifungal and antibacterial pharmacophore sites, *Journal of Molecular Structure, 1205*, 127625, ISSN 0022-2860, https://doi.org/10.1016/j.molstruc.2019.127625

Xiao X., Xu M., Yang C., Yao Y., Liang LN., Ed Chung P., Long Q.,*et al*. (2018) Novel racemosin B derivatives as new therapeutic agents for aggressive breast cancer. *Bioorg Med Chem.* 26(23-24), 6096-6104. https://doi:10.1016/j.bmc.2018.11.014.

Yang Y., Qin J., Liu H., Yao X. (2011) Molecular dynamics simulation, free energy calculation and structure-based 3D-QSAR studies of B-RAF kinase inhibitors. *J Chem Inf Model.,* 51(3), 680-92. https://doi:10.1021/ci100427j.

Yunusa U., Umar U., Idris S., Kubo A., Abdullahi T. (2021). Experimental and DFT computational Insights on the Adsorption of Selected Pharmaceuticals of Emerging concern from water system onto magnetically Modified Biochar. *Journal of the Turkish Chemical Society Section A Chemistry* 8(4), 1179-1196. https://doi:10.18596/jotcsa.900197.

Zafar, A., Khatoon, S., Khan, M.J. *et al.* (2025). Advancements and limitations in traditional anti-cancer therapies: a comprehensive review of surgery, chemotherapy, radiation therapy, and hormonal therapy. *Discov. Onc*. 16, 607, https://doi.org/10.1007/s12672-025-02198-8