Journal of Applied Pharmaceutical Research
Volume 13 Issue 2, Year of Publication 2025, Page 149 – 163
DOI: https://doi.org/10.69857/joapr.v13i2.951

Check for updates

Research Article

JOURNAL OF APPLIED PHARMACEUTICAL RESEARCH | *JOAPR*
www.japtronline.com *ISSN: 2348 – 0335*

IDENTIFICATION OF NOVEL POTENTIAL BENZIMIDAZOLE DERIVATIVES BY PHARMACOPHORE GENERATION, 3D-QSAR, VIRTUAL SCREENING, MOLECULAR DOCKING AND ADME/ TOX ANALYSIS AGAINST BREAST CANCER AS TARGETED ESTROGEN ALPHA RECEPTOR

Aastha Sharma, Nitish Banga, Rakesh Kumar Marwaha*, Balasubramanian Narasimhan

Article Information
Received: 30th December 2024
Revised: 27th March 2025
Accepted: 17th April 2025
Published: 30th April 2025

Keywords
Benzimidazole, Cancer, Docking, Pharmacophore, Virtual screening, 3D-QSAR, ADME/Tox.

ABSTRACT

Background: The estrogen alpha receptor (ERα) is critical in breast carcinogenesis. Although selective estrogen receptor modulators like tamoxifen are clinically used, their adverse effects highlight the need for safer alternatives. The study uses computational methods to identify potential ERα inhibitors within a benzimidazole scaffold. **Methodology:** This study employed computational approaches, including pharmacophore generation, 3D-QSAR, virtual screening, molecular docking, and in silico ADME/Tox analysis. The best pharmacophore model (DDRRR_1) identified two hydrogen donors and three aromatic rings as critical features. Moreover, a rigorous external validation was used on decoy databases with optimized metrics (ROC, BEDROC, AUROC). A subsequent atom-based 3D-QSAR model with a high correlation coefficient ($R^2 = 0.9$), cross-validated coefficient ($Q^2 = 0.8$), and Fisher ratio ($F = 80.1$) was developed. Benzimidazole scaffolds from PubChem were screened, followed by docking against ERα (PDB ID: 3ERT) and ADMET profiling. **Results and Discussion:** The pharmacophore model validated the importance of the identified features. The 3D-QSAR model effectively screened benzimidazole scaffolds, with five component PLS factors, supporting the pharmacophore findings. This model effectively screened benzimidazole scaffolds obtained from the PubChem database, followed by molecular docking against the targeted protein ERα (PDB ID: 3ERT) and identified five promising compounds. ADME/Tox profiling revealed PubChem ID 3074802 (2-[2-(1H-indol-3-yl) ethyl]1H-benzimidazole) has favourable pharmacokinetics and a low toxicity profile. **Conclusion:** These findings indicate that PubChem ID 3074802 is a promising candidate for further therapeutic drug development in breast cancer treatment. It demonstrates the highest binding affinity (-9.842 kcal/mol) compared to the standard drug Tamoxifen (-5.357 kcal/mol) and exhibits a favorable ADME/Tox profile.

*Department of Pharmaceutical Sciences, Maharshi Dayanand University, Rohtak, Haryana, 124001, India

***For Correspondence:** **rkmarwaha.mdu@gmail.com**

INTRODUCTION

Cancer is becoming a global health issue, becoming the second leading cause of death in the United States, with an expected count of 2,001,140 new cancer cases, with an estimated mortality rate of 611,720 in the year 2024. Breast cancer has been identified to cause the highest morbidity and the second-highest mortality in females, among all cancer types. It alone counts for 310,720 new cancer cases (32% of all cancer cases) and 42,250 deaths (15% of all cancer cases) in the year 2024. It may be presumed that by the year 2030, nearly 2 million cases of breast cancer will be reported all over the globe [1]. As the treatment and diagnosis were delayed due to the COVID-19 pandemic, it is necessary to expedite the treatment and diagnosis processes with a good healthcare service [2, 3].

In India, as per the ICMR report, "one in every nine persons has a risk of cancer" [4]. Breast cancer is the most common cancer in Indian women, also accounting for 14% of malignancies. According to statistics, a woman in India receives a breast cancer diagnosis every four minutes. It is equally endemic in rural and urban India. Breast cancer survival has become more and more challenging day by day, and more than half of Indian women diagnosed with breast cancer are in stage 3 or 4. After treatment, the survival rate for Indian women is 60%, whereas it is 80% in the United States [5]. The increase in cancer rate was due to unhealthy lifestyles, urbanization, and air pollution, which is highly prevalent in socio-economic countries like India [6].

Estrogen receptors (ER) are the major marker for predicting breast cancer. It is of two subtypes, i.e., ERα and Erβ, having different binding affinities to produce their response. Among these, the disturbance in ERαsignalling pathways is remarkably responsible for postnatal mammary gland development, breast carcinogenesis and its progression, and is responsible for hormone-type breast cancer. Selective Estrogen Receptor Modulators (SERMs) such as tamoxifen and raloxifene block estrogen signals by binding to estrogen receptors in breast tissue. More than 50% of breast cancers exhibit overexpression of estrogen receptors (ERα), and approximately 70% of these cases respond to anti-estrogen therapies such as tamoxifen [7]. However, they can cause side effects because they also exert estrogenic effects on healthy cells. Moreover, tamoxifen and its active metabolite (4-OHT) have also faced limitations by producing resistance to treatment in 86.7% of ERα-positive cases. These results highlight the necessity of developing drugs that can enhance the treatment for breast cancer with minimum side effects [8, 9]. To date, ERα-positive breast cancer remains a primary focus for clinical therapy due to its central role in regulating cell division, differentiation, apoptosis, and migration [10]. Nitrogen-containing heterocyclic rings have high significance in medicinal chemistry due to an electron-rich nitrogen atom, which favors the formation of hydrogen bonds or dipole-dipole interactions. These interactions help improve biological activity by providing high binding affinity with the biological targets or diversity of enzymes [11]. Benzimidazole, a nitrogen-containing heterocyclic moiety, contains a benzene ring fused with an imidazole ring as given in Figure 1. Its amphoteric nature plays an important role in developing chemical interactions with a wide range of therapeutic targets, thereby exhibiting various pharmacological effects, and thus has generated a great scientific interest nowadays [12].

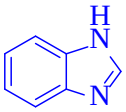


Figure 1: Chemical structure of benzimidazole nucleus

The benzimidazole structure enables it to bind with DNA, which helps in apoptosis induction, disruption of the microtubule network, inhibition of DNA synthesis, and prevents the development of cancer cells [13, 14]. There are many marketed formulations containing benzimidazole moiety in their parent structures and have a role in cancer treatment like liarozole (retinoic acid metabolism blocking agent (RAMBA)) and pracinostat (histone deacetylase inhibitor), Veliparib (PARP-1/2 inhibitor, potentiate DNA damaging agent), Carbendazim (inhibits tumor cell proliferation), Nocodazole (microtubule depolymerizing agent) and Bendamustine (alkylating agents). Still, all these drugs have faced some side effects like nausea, headache, pruritus, and epistaxis [13]. To improve this complication, there is a great need to find new therapeutic molecules with minimal side effects and better potential to deal with breast cancer.

In the past, drugs have been discovered by experimental and random screening techniques, which require a lot of time (about 10-20 years) and money (avg. cost 1 to 2 million USD) because the molecules pass through various phases before they can get market access. Nowadays, several computational approaches are available that can accelerate the efficiency of the drug design and development process, also called rational drug design. The

Computational approaches, such as pharmacophore generation, 3D-QSAR analysis, molecular docking, etc, help in examining diverse molecules in a shorter time and evaluating their binding affinity with the target of interest and pharmacokinetic profile before experimental or in vitro testing, and pave the way for medicinal chemists [15].

Therefore, the present research aimed to identify selective and potent molecules as ERα inhibitors for breast cancer treatment using a comprehensive approach that combined ligand and structure-based drug design strategies. This was followed by evaluating the identified compounds' in silico pharmacokinetic and toxicity profiles. This study addresses a critical gap in breast cancer research by leveraging advanced computational techniques, focusing on discovering novel benzimidazole derivatives as ERα inhibitors. The findings lay a strong foundation for developing safer and more effective therapeutic options, addressing the limitations of current SERMs and paving the way for future experimental validation and clinical applications.

MATERIALS AND METHODS

Dataset

The dataset comprised 60 benzimidazole molecules with reported *in vitro* activity against the MCF-7 cancer cell line, prominently expressing the ERα receptor. IC_{50} values, converted to pIC_{50} using the formula pIC_{50}= -log IC_{50}, were the dependent variables for 3D-QSAR modelling. The molecules, retrieved from recent literature, were chosen for their diverse pharmacophoric features and consistent biological activity measured via the MTT assay method [16, 17]. Each molecule was generated in ChemDraw (version 15.0), optimized in the LigPrep module of Maestro (version 13.1), where tautomeric states and conformers were not generated, and charged species were neutralized. The OPLS2005 force field was utilized to ensure that only low-energy conformations were considered [18].

Pharmacophore Model Generation

A pharmacophore is a molecular framework consisting of specific features within a molecule that are recognized at a receptor site, influencing biological activity by either stimulating or inhibiting it [19]. Pharmacophore models were generated when the structure of a known active ligand was available, while the macromolecular target structure was absent, and is known as Ligand-based pharmacophore modeling. It is becoming a fundamental computational approach in facilitating the drug discovery process. In general, generating a pharmacophore from multiple ligands involves the following steps. Firstly, the ligands' conformational space is explored, all the ligands are aligned, and later, the essential common chemical features are determined to construct pharmacophore models [20]. In this study, the pharmacophore modeling was conducted using the PHASE module in Maestro 13.1 of Schrodinger software [21]. PHASE provides a set of six pharmacophoric features, including hydrogen bond acceptors or donors, anionic, cationic, hydrophobic, and aromatic groups [22, 23]. Finally, conformations of selective active molecules were aligned to generate common pharmacophoric features [24].

In this approach for modeling pharmacophores, only the active chemicals are often considered. The “Develop pharmacophore hypothesis” function of the Phase module was used to generate a pharmacophore from multiple ligand entries. For creating a common pharmacophore, the pIC50 values ranged from 8.276 to 4.622 are defined as active, with pIC50>5.41, and inactive, with pIC50<5.40, because active molecules possess crucial structural features essential for rational drug design or binding to the receptor binding site, thereby influencing potency and selectivity in therapy [24]. Three to five pharmacophoric key sites are utilized for pharmacophore generation to create an optimal combination of features shared among highly active compounds. Subsequently, the resulting pharmacophores are classified based on their scoring algorithm, such as number of matches, volume, vector, site score, and survival score [25]. Scoring was used to identify the most promising hypothesis by the comprehensive ranking of all hypotheses based on their performance score (survival score, site score, vector score, and volume score). High parameter values express a high-quality pharmacophore model.

Pharmacophore Validation

To ensure the effectiveness of the pharmacophore model in guiding virtual screening, it was validated through enrichment studies. This process assesses the accuracy and specificity of the model in selecting active molecules from a large pool of candidates. Validation employed an external test set containing over 1000 decoy molecules obtained from the Directory of Useful Decoys (DUD) (https://dud.docking.org/). Decoy molecules are inactive compounds with similar chemical

properties to active molecules, and including them helps test the ability of the model to differentiate true actives from random structures [26]. Additionally, 15 confirmed active molecules were included to evaluate the model's predictive ability for identifying known actives. Subsequently, the active and decoy datasets underwent preprocessing using the LigPrep module. Notably, ionization states and tautomers were not generated, assuming a neutral pH of 7.0 for the screening process. The "hypothesis validation tool" within the Phase module was utilized for the validation process. This tool inputs the pharmacophore hypothesis file and the active and decoy datasets. It evaluates the performance using various metrics such as Phase Hypo Score, Enrichment Factors (EF1%), BEDROC160.9, receiver operating characteristics (ROC), and Area Under Accumulation Curve (AUAC) [27, 28].

3D-QSAR

An external validation process was conducted to evaluate the robustness of the developed pharmacophore model. This involved predicting the biological activity of molecules from a separate test set. The original dataset of 60 molecules was randomly divided into two subsets using the automated random selection function within the Phase software. The larger portion (75%) was designated as the training set to build the model. The remaining 25% of molecules formed the test set for evaluating the model's reliability (Supplementary Table **1**). 'Phase' offers two primary alignment methods for the 3D structures of molecules, including pharmacophore-based alignment, which aligns molecules based on their functional features. In contrast, atom-based alignment aligns the molecules based on individual atom positions. The choice of alignment method can influence the performance of the model [21]. This study employed an atom-based Quantitative Structure-Activity Relationship (QSAR) model as they have the same molecular framework. QSAR is a computational technique that helps us understand the relationship between the molecule's structure and its biological activity. The atom-based approach offers a more detailed picture than other methods by representing each molecule as a collection of overlapping spheres, where each sphere corresponds to a single atom in the molecule [21, 29]. A set of rules based on atomic properties classifies each atom into one of six categories. These categories capture key chemical features: hydrogen bond donors (D), hydrophobic/nonpolar regions (H), negative and positive charges (N and P), electron-attracting atoms (W), and any remaining atom types (X). This classification system allows the model to account for the individual contributions of different atomic functionalities to the overall biological activity of the molecule [30].

To ensure the robustness of the pharmacophore model, an atom-based 3D-QSAR analysis was performed using Partial Least Squares (PLS) regression. This technique identifies latent factors that explain the relationship between the molecular structure and biological activity. Each model comprised five or more PLS factors, up to a maximum of N/3 PLS factors (where N is the number of ligands in the training set), and was used in the PLS regression. In this study, the optimal number of PLS factors was determined to be five. Including more factors resulted in diminishing returns, with a decrease in statistical significance and predictive ability [31]. The study employed several statistical metrics to assess the performance of the final PLS QSAR model. These metrics included the R^2 (regression coefficient), Q^2CV (cross-validation coefficient), F-statistic (variance), confidence interval (P), RMSE (mean squared error), and Pearson's correlation coefficient (r). These parameters provide insights into how well the model fits the data, and offer statistically reliable predictions (F-statistic, P-value). Additionally, RMSE measures the closeness of predictions to actual values, while Pearson's correlation coefficient indicates the strength and direction of the relationship between predicted and observed activity [32]. The insights from the QSAR studies are invaluable for designing and synthesizing novel, promising molecules. A key advantage of the 3D-QSAR technique is its ability to generate contour maps around the studied molecules. By analysing these coloured cubes, researchers can identify the key molecular features favourable or unfavourable for ligand interaction with the target receptor. Blue cubes depict regions favourable for activity, while red cubes depict unfavourable regions [33].

High throughput virtual screening (HTVS) and molecular docking

During the initial phase of drug design, an extensive database of compounds is evaluated using HTVS. This process helps to identify potential lead compounds that may have the desired biological activity and to undergo more in-depth analysis of a drug candidate with improved efficacy and safety profiles. The PubChem database contained 7,133 potential benzimidazole-containing molecules having druglike properties or following the Lipinski rule of five. In the PubChem database, several other

parameters, like rotatable bond (<10) and polar surface area (<140), were also applied. Further, all the potential molecules were screened from the DDRRR_1 hypothesis model by using the Phase Ligand Screening module to identify only those hit molecules with similar pharmacophoric group matches. Structure-based HTVS involves a molecular docking approach that predicts the binding mode of a database of compounds against a protein target, facilitating efficient identification of potential drug candidates. Here, ERα was selected as a targeted protein for evaluating the binding affinity of the screened compounds. The targeted protein structure (PDB ID: 3ERT) was downloaded from the Protein Data Bank (PDB) (https://www.rcsb.org/ structure/3ERT), consisting of a single ligand and protein chain with a resolution of 1.90 Å and no mutations. The preparation and optimization of the protein were done using the 'protein preparation wizard' module of GLIDE [34]. Further, the co-crystallized ligand in the targeted receptor has been used for grid generation and identification of the binding pocket. The molecules passing pharmacophore screening were then docked into the 3ERT crystal structure using Glide XP to estimate their binding affinities. Glide XP docking scores served as a preliminary assessment. Subsequently, the bound conformations were used to calculate the binding free energy or a related affinity measure for the most promising candidates. These compounds were then selected for further ADME/Tox analysis.

ADME/T Prediction

The development of in silico methods in recent times for developing new molecules has been approved by the USFDA. However, there is still a high chance of late-stage attrition or drug applicability in clinical trials, causing failure of the drug molecule. The main reason behind this problem was drug safety and toxicity profile (ADME/T), which plays a keen role in drug development. Therefore, it is necessary to find a molecule with an optimized ADME/T profile [35]. The top five hits identified through pharmacophore screening and binding affinity calculations were then subjected to ADME/T analysis using ADMET Lab 2.0, a free web tool. This in silico approach helps evaluate the physicochemical properties of the compounds early in the drug design process, potentially reducing the risk of ADME/T issues later in development. The obtained data included several pharmacokinetic characteristics, including blood-brain barrier permeability, rule of five, octanol/water coefficient, Caco cell permeability, and T1/2. There are various physicochemical parameters, as well as toxicity parameters, that are applied to eliminate undesirable and toxic compounds.

Table 1: Score of different parameters of the generated Pharmacophore hypothesis.

Sl No	HypoID	Survival Score	Site Score	Vector Score	Volume Score
1	DDRRR_1	6.118	0.976	0.985	0.879
2	ADDRR_1	5.840	0.974	0.985	0.879
3	DDRR_1	5.670	0.974	0.98	0.879
4	DRRR_1	5.65	0.989	0.984	0.871
5	DDRR_2	5.633	0.975	0.983	0.876
6	DRRR_2	5.618	0.968	0.97	0.867
7	DDRR_3	5.588	0.970	0.984	0.844
8	ADRR_1	5.457	0.961	0.977	0.868
9	ADRR_2	5.457	0.963	0.979	0.866
10	ADDR_1	5.345	0.971	0.986	0.848
11	ADDR_2	5.338	0.962	0.972	0.858
12	ADRR_3	5.045	0.765	0.842	0.690
13	DRR_1	5.323	0.994	0.98	0.871
14	DRR_2	5.297	0.996	0.981	0.871
15	DRR_3	5.284	0.999	1	0.862
16	DRR_4	5.284	0.997	0.994	0.857
17	RRR_1	5.278	0.994	0.976	0.873
18	DDR_1	5.27	0.991	0.983	0.832
19	DRR_5	5.268	0.969	0.972	0.870
20	DDR_2	5.241	0.982	0.980	0.834
21	ARR_1	5.199	0.975	0.965	0.857
22	ADR_1	5.151	0.999	1	0.858

RESULTS AND DISCUSSION

Generation of Pharmacophore Model

A ligand-based pharmacophore model was developed to search for novel ERα inhibitors. This approach utilizes the structural features (e.g., hydrogen bond donors, hydrophobic regions) of previously reported inhibitors to identify potential new candidates. Various hypotheses were generated, each representing a combination of these pharmacophoric features (acceptor [A], donor [D], hydrophobic [H], negative ionic [N], positive ionic [P], and aromatic ring [R]) found in the training set molecules. Table 1 summarizes all the generated models and their combinations of shared active compound features. A metric called the survival score was used to identify the best hypothesis. This score ranged from 6.1185 (highest for model DDRRR_1) to 5.1512 (lowest for model ADR_1). Hypotheses with only three or four features were excluded due to their low survival

scores, suggesting an inability to capture the complete range of binding interactions within the training set. Among the remaining models with five features, pharmacophore (DDRRR_1) was chosen for further analysis due to its superior survival score [25].

Pharmacophore Model Validation

The pharmacophore model was validated to evaluate the predictive potential of the pharmacophore model and authenticate the pharmacophore design [36]. Before database screening, the structure-based pharmacophore model was validated to distinguish active compounds from inactive ones (decoy molecules). Enrichment studies were employed to validate the chosen pharmacophore model (DDRRR_1). These studies assessed the model's ability to enrich for truly active compounds. The validation process involved an initial screening where the model was challenged to differentiate between 15 well-established active molecules and a database of decoy molecules. So, an extensive collection of over 1,000 decoys was retrieved from the DUD-E database, a directory of decoy databases. These decoys share similar properties with the active molecules but lack the desired inhibitory activity. Statistical parameters such as Enrichment Factor (EF), Robust Initial Enhancement (RIE), Receiver Operating Characteristic (ROC), and Boltzmann-Enhanced Discrimination of ROC (BEDROC) were assessed. These parameters were used to benchmark the model's reliability and accurately rank compounds, as reported in Table 2. Generally, a model with a higher AUC value should exhibit better predictability and provide a summary of model performance. The AUC value ranges between 0 and 1, where 1 indicates a model with perfect prediction accuracy. Five featured hypotheses showed astonishing results in our validation process, with the highest phase hypo score of 1.36 observed for ADDRR_1. Therefore, pharmacophore DDRRR_1 was selected (Table 2). The ROC plots, depicting sensitivity versus specificity, demonstrate the precision in identifying true positives, while the Enrichment Factor reveals the ability of the pharmacophore to differentiate actives from decoys. ROC analysis visually represents the sensitivity and specificity ratio validation process, showing how effectively the model distinguishes active from inactive compounds (Figure 2). A steeper ROC curve, reaching a plateau at the end, indicates a model that excels at identifying more active compounds. Higher AUC values (closer to 1) suggest better model performance, while lower values (closer to 0) indicate a less reliable model. The results for DDRRR_1 indicate that the model successfully distinguishes true actives from decoy compounds and can identify inhibitors based on their potency.

Table 2. Validation parameters of the generated pharmacophore.

Hypothesis	Phase Hypo Score	EF1%	BEDROC160.9	ROC	AUAC	Matches
DDRRR_1	1.36	100.15	1	1	1	4 of 5
ADDRR_1	1.34	100.15	1	1	1	4 of 5

3D-QSAR

The top model with good predictive power, denoted as DDRRR_1 (two hydrogen bond donors and three aromatic rings), was identified based on a five-point hypothesis (Fig. 3a). This model features two donor groups (DD, blue) and three aromatic rings (RRR, brown). Rings R6 and R7 correspond to the benzimidazole ring, while one donor group maps to the NH of the benzimidazole group, and the other donor represents the substituted NH group. The interatomic distances in Å are shown in Figure 3b, while Figure 3c and 3d depict the mapping of active and inactive compounds onto the pharmacophore, respectively. These figures show that active ligands exhibit better alignment with the DDRRR_1 pharmacophore, suggesting their potential for higher activity. This highlights the importance of good alignment for accurate activity prediction in the 3D-QSAR model. To design theoretical models and assess their robustness, 60 benzimidazole scaffolds were used. All compounds exhibit biological activity within a three-unit logarithmic range.

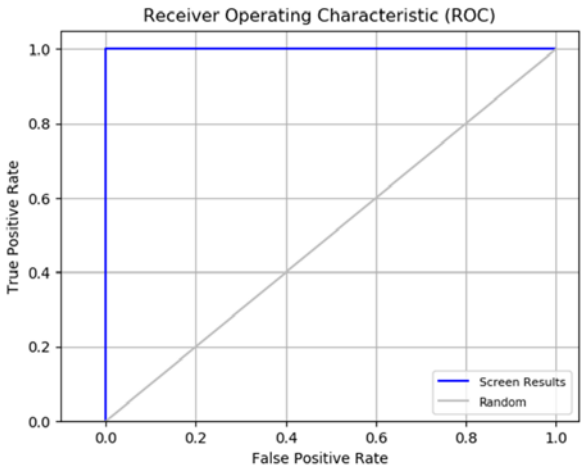


Figure 2. ROC curve of a generated best pharmacophore model (DDRRR1)

The developed atom-based 3D-QSAR model involved generating multiple factors using the partial least squares (PLS) technique. Among these factors, the 7th partial least-squares factor emerged as the most suitable model due to its good statistical significance and strong predictive power. The selection of the 5th factor was based on the observed gradual improvement in model performance up to this point. The combined effects of H-bond donors, hydrophobic/non-polar character, negative ionic character, positive ionic character, and electron-withdrawing groups were evaluated, and all statistical parameter values are presented in Table 3.

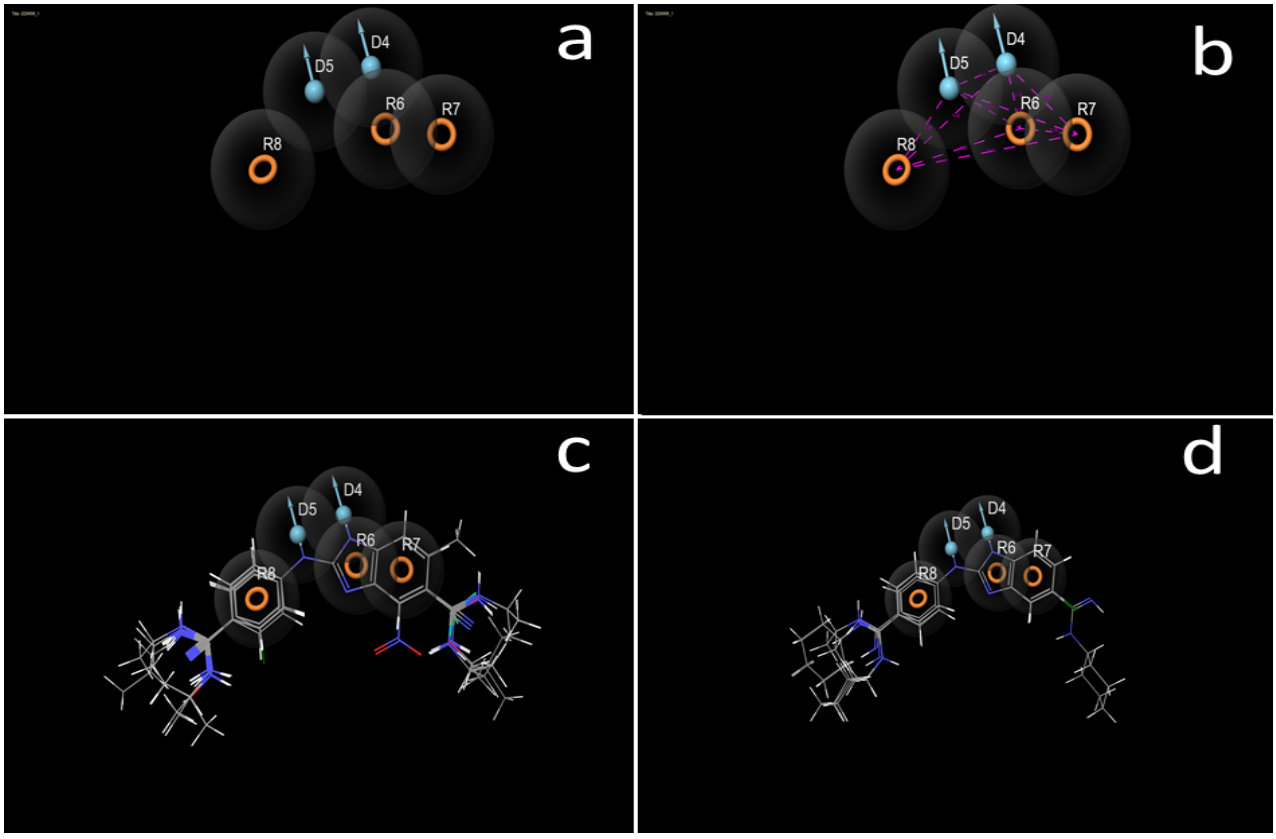


Figure 3: (a) Pharmacophoric features of validated pharmacophore model (DDRRR_1) (b) Spatial arrangement of DDRRR_1 feature (c) DDRRR_1 model alignment of active molecules. In contrast, (d) DDRRR_1 model alignment of inactive molecules.

The model with the highest Q^2 value exceeding 0.5 was chosen based on its superior predictive ability. Additionally, a comprehensive external validation was performed to assess model robustness. This validation involved calculating several metrics for both the test set (RMSE, Q^2, and Pearson correlation coefficient) and the training set (R^2, R^2CV, and Fisher's (F) ratio). A Y-randomization test yielded even better results, further supporting the model's validity. In general, the R^2 value lies between 0.7 and 0.9. Stability should be closer to 1, the F value must be maximized as much as possible, to define the model is not a false positive, the P value should be low, the RMSE value lies between 0.4 and 0.5, the Q^2 value must be closer to 1, and and the Pearson's value should be closer to 1. Internal validation indicated the robustness and predictive ability of the model. The internal validation results demonstrated excellent performance for the atom-based 3D-QSAR model. The training set, consisting of 43 compounds, exhibited a strong correlation coefficient (R^2 = 0.94), indicating a close relationship between the predicted and actual biological activity. Additionally, a high Fisher ratio (F = 80.1) further supported the statistical significance of the model. The predictive capability was assessed using a separate test set of 17 compounds. The results were highly promising, with a high cross-validated correlation coefficient (Q^2 = 0.85) and Pearson's R (0.94). These metrics indicate a strong correlation between the predicted and observed activity in the unseen test set. The large F-value observed in the training set (F = 80.1) suggests a statistically robust model, and the small significance level of the variance ratio (P) further strengthens this confidence. A comprehensive summary of these results is presented in Table 3. Scatter plots were generated to visually represent the model's

performance (Figure 4a and 4b), depicting the observed versus predicted biological activity for both the training and test set molecules. These plots demonstrate a good distribution of data points, particularly in the test set, which aligns well with the best-fit line (y = 0.83x + 0.80, R^2 = 0.89). This visual confirmation further validates the ability of the model to accurately predict the biological activity of unseen compounds [37].

Table 3. PLS statistical parameters of the selected 3D-QSAR model (DDRRR1).

PLS FACTOR	SD	R^2	R^2CV	F	P	RMSE	Q^2	PEARSON-R
1	0.6927	0.6689	0.5984	82.8	2.16e-11	0.91	0.6002	0.8555
2	0.6192	0.7420	0.5676	57.5	1.71e-12	0.91	0.6002	0.8600
3	0.5751	0.7830	0.4999	46.9	5.22e-14	0.88	0.6228	0.8722
4	0.5025	0.8385	0.3356	49.3	1.52e-14	0.73	0.7398	0.9160
5	0.4246	0.8878	0.1748	58.5	1.49e-16	0.59	0.8305	0.9423
6	0.3550	0.9236	0.1284	72.6	1.27e-18	0.57	0.8427	0.9376
7	0.3158	0.9413	0.1494	80.1	1.21e-19	0.55	0.8553	0.9442

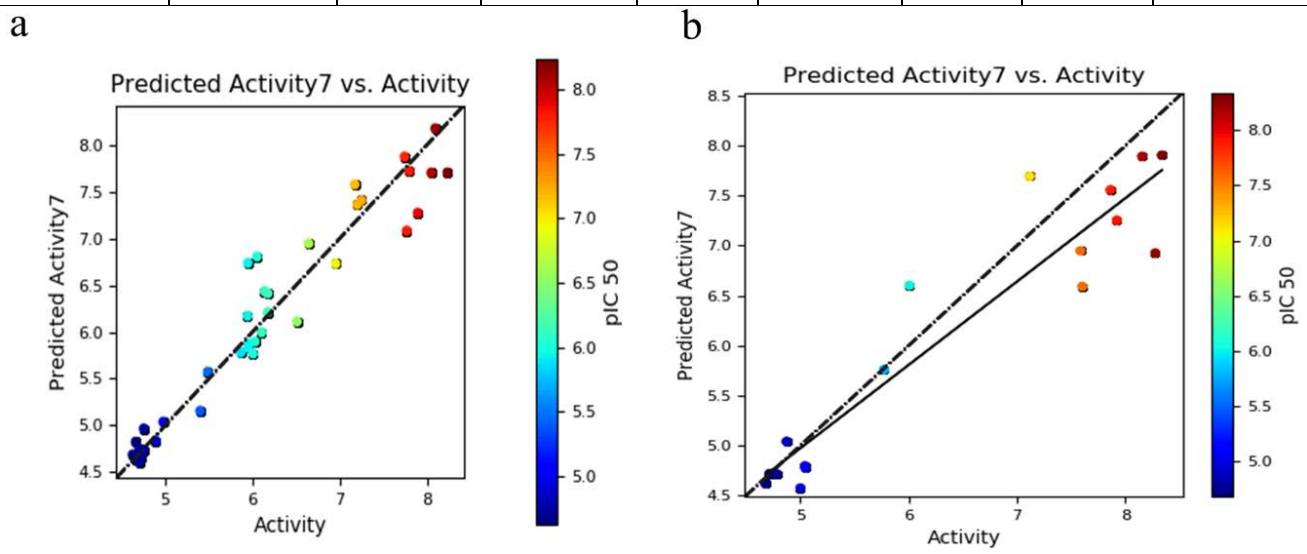


Figure 4. The scatter plots depict the correlation between the experimental and predicted biological activity of the benzimidazole inhibitors in the (a) training set and (b) test set. The solid line represents the best-fit line with the equation y = 0.83x + 0.80 (R^2 = 0.89) for the test set.

Contour Map Analysis

Contour plot analysis was performed at specific locations of the molecule to elucidate the essential pharmacophoric regions. The 3D-QSAR models make it easier to see the contours of the ligand-receptor interactions and pinpoint the positive and negative activity coefficients that responsible for the biological activity like hydrogen bond donor, hydrophobic/ non-polar, positive, and negative ionic and electron withdrawing properties. The blue cubes represent the positive contribution of different groups while the red cubes indicate negative contribution. For visualizing the generated 3D-QSAR model and exploring its correlation with inhibitory activity, one or more ligands from the series have diverse inhibitory activities, were selected for analysis. For this, the most active compound 28 (pIC_{50}= 8.337) and least active compound 56(pIC_{50}= 4.622) were taken for visualization purpose and shown in Figure 5a, 5b, 5c, 5d, and 5e. The effect of hydrogen bond donor, the blue cubes at amino group (C-1 and C-2) of benzimidazole favours the activity indicated the preference of hydrogen bond donor group at D5 and D4 (pharmacophoric features) position for activity of compound 28 as shown in Figure 5a. Compound 56 has the lowest activity (pIC_{50} 4.6) which might be due to absence of 'NH' group near the benzimidazole moiety (D7 pharmacophoric features) that means the presence of 'NH' favourably contributed to activity. Presence of blue cubes nearer to benzimidazole ring favours the activity while presence of red cubes at terminal end of substituted benzimidazole ring indicated unfavourable regions for hydrophobic interactions which high inhibitory activity of compound 28 (Figure 5b). While in compound 56, the attachment of substituted alkyl chain at

benzimidazole ring indicated unfavourable for inhibitory activity in Figure 5c. In the contour plot of electron withdrawing features, compound 28 contain 'NH' of benzimidazole ring favours the activity and red cubes at terminal end indicates unfavourable regions for electron withdrawing regions (Figure 5d). Compound 56 the attachment of long chains to C-2 position of benzimidazole moiety was unfavourable for inhibitory activity (Figure 5e).

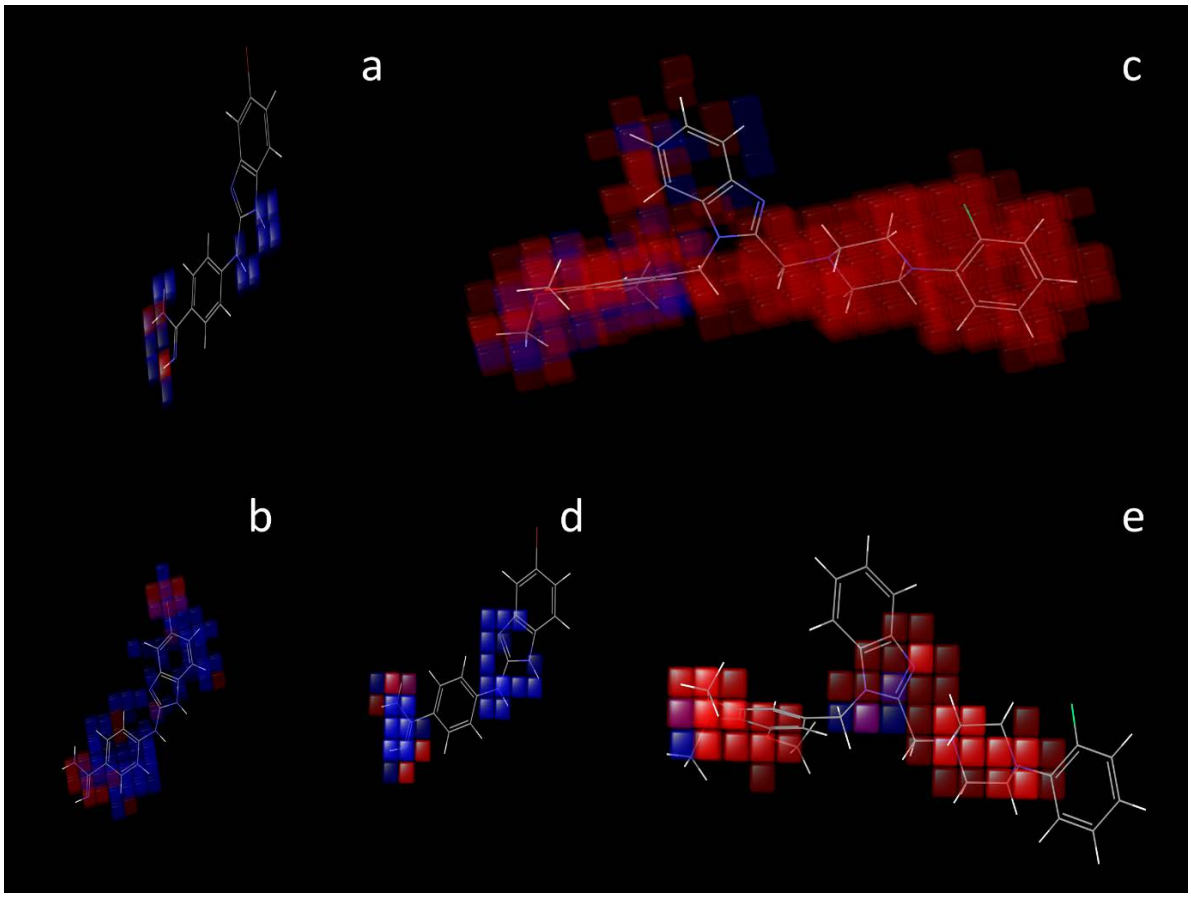


Figure 5. Atom-based 3D-QSAR based contour maps in the context of favourable and unfavourable regions with blue and red effects. Compound 28 (highest activity) and Compound 56 (least activity) were used as the template and were shown in ball-stick model (a) hydrogen bond donor (b) hydrophobic/ non-polar (c) Electron withdrawing

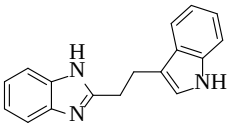
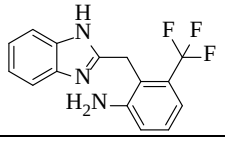
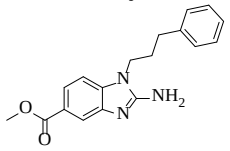
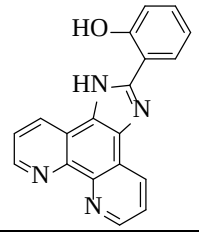
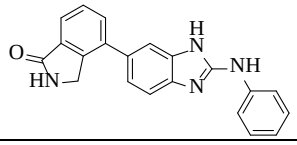
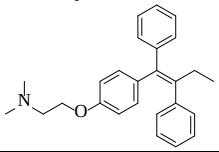
Pharmacophore-based Virtual Screening for identifying novel benzimidazole inhibitors

The compounds obtained from the PubChem database after employing filters like the Lipinski rule of 5, no. of rotatable bonds, polar surface area (PSA), and drug likeness filter have been used to obtain potential molecules, which could be identified as a potential inhibitor for ERα. A virtual screening process was employed to identify potential new inhibitors with novel scaffolds. Initially, a search of the PubChem substructure database identified 7,133 compounds containing the benzimidazole moiety. To refine this large set, a validated pharmacophore model (DDRRR_1) was used as a 3D query for filtering. This approach effectively reduced the number of candidate inhibitors to 1143, demonstrating the efficiency of the pharmacophore model in eliminating unsuitable structures based on their 3D arrangement of key features. The resultant compounds were subjected to Molecular Docking against the targeted receptor, i.e., ERα (PDB ID: 3ERT) [38]. Molecular docking for the selected compounds was executed using the GLIDE program provided by Schrödinger 2023.

Before docking, the compounds' 3D structures were prepared using the Maestro LigPrep module. This module generated multiple conformations for each ligand, considering different ionization states within the pH range of 7.0 ± 2.0. This comprehensive approach facilitates the exploration of ligand-protein interactions and supports the rational design of potential therapeutic agents. The molecular docking screening process was performed on 1143 ligands obtained after pharmacophore-based screening.

The top five best docked molecules were selected to evaluate more in-depth parameters like docking score, glide energy, and essential amino acid interaction. All the compounds were also compared with the standard drug Tamoxifen, SERM, has a docking score of -5.796kcal/mol, as given in Table 4, while 2D and 3D interaction images of all five top docked complexes were given in Figure 6. Thus, the results indicated that all the top 5 protein-ligand complexes show better binding affinity, so they are considered for further ADME/Tox profile to identify the best molecule with the least toxicity profile and as a better therapeutic agent.

Table 4. The selected hits obtained by Pharmacophore-based virtual screening of compounds from the PubChem Database with their binding affinity using the GLIDE module.

PubChem Compound	Chemical Structure and IUPAC name	Docking score (kcal/mol)	Glide Energy (kcal/mol)	Crucial amino acid residues and H-bonding interactions
3074802 (**$C_{17}H_{15}N_3$**)	2-[2-(1H-indol-3-yl)ethyl]1H-benzimidazole 	-9.842	-37.441	GLY521, HIE524, LEU525, MET343, **LEU346 (-NH)**, THR347, LEU349, ALA350, GLU353, ARG394, LEU391, MET388, LEU387, LEU384, TRP383, LEU428, PHE404
71494439 (**$C_{15}H_{12}F_3N_3$**)	1H-benzimidazol-2-ylmethyl-[3-(trifluoromethyl)phenyl] amine 	-9.72	-35.189	ILE424, MET421, GLY420, GLY521, HIE524, LEU525, MET343, **LEU346 (-NH)**,THR346, THR347, LEU349, ALA350, ARG394, LEU391, PHE404, LEU428, MET388, LEU387, LEU384, TRP383
71682086 (**$C_{18}H_{19}N_3O_2$**)	2-amino-1-(3-phenylpropyl) benzimidazole-5-carboxylic acid methyl ester 	-9.269	-42.015	LEU391, MET388, LEU387, LEU384, TRP383, ASP351, ALA350, THR347, LEU346, MET343, LEU428, ILLE424, MET421, GLY420, GLU419, VAL418, GLY521, HIE524, LEU525
135934901 (**$C_{19}H_{12}N_4O$**)	2-(1H-imidazo[4,5-f][1,10] phenanthrolin-2-yl)phenol 	-9.221	-36.703	HIE524, LEU525, MET528, MET343, LEU346, **THR347 (-OH)**, ALA350, ASP351, GLU353, GLY420, MET421, ILE424, LEU428, ARG394, LEU391, MET388, LEU387, LEU384, TRP383, PHE404
11739311 (**$C_{21}H_{16}N_4O$**)	4-(2-anilino-3H-benzimidazol -5-yl)isoindolin-1-one 	-9.215	-44.846	LEU539, LEU536, TRP383, L3U384, LEU387, MET388, LEU391, **ARG394 (C=O)**, LEU428, LEU354, **GLU353 (-NH)**, ASP351, ALA350, LEU349, THR347, LEU346, MET343, LEU525
Tamoxifen	2-[4-[(Z)-1,2-diphenylbut-1-enyl]phenoxy]-N,N-dimethylethanamine 	-5.357	-33.220	LEU525, MET528, LYS529, CYS530, ARG394, LEU391, MET388, LEU387, LEU384, **TRP381 (-C_6H_5),** PHE404, MET343, LEU346, THR347, LEU349, ALA380, ASP 352, GLU353, LEU354

Absorption, Distribution, Metabolism, Excretion, and Toxicity predictions (ADME/Tox)

Evaluating in silico ADME/Tox prediction is a critical step in selecting lead compounds before clinical trials. This approach is widely employed to minimize late-stage attrition caused by unfavourable toxicity profiles and to provide insights into the synthetic accessibility of potential drug candidates. In this study, the top five compounds were screened using ADMET Lab 2.0, a free web tool that facilitated the assessment of key parameters such as physicochemical properties, medicinal chemistry attributes, and toxicity profiles. The results of the ADME/Tox profile of all five compounds are given in Table 5. The analysis revealed that all five compounds fell within the optimal range of pharmacokinetic parameters. However, PubChem compound 3074802 demonstrated the most favorable toxicity profile among the screened compounds, highlighting its superior safety and pharmacokinetic characteristics. The compound exhibited a low hepatotoxicity score (0.623), indicating minimal potential for liver toxicity. Additionally, its AMES test score (0.226) suggests a low likelihood of mutagenicity, while the carcinogenicity score (0.079) confirms a negligible risk of cancer-causing potential. Furthermore, the NR-AR score (0.581) indicates moderate interaction with nuclear receptors, which is acceptable for drug candidates. These toxicity parameters of PubChem compound 3074802 (2-[2-(1H-indol-3-yl) ethyl]1H-benzimidazole) underscore the suitability as a safe and effective therapeutic agent and make it a suitable candidate for therapeutic applications in breast cancer treatment. These promising properties highlight its potential for further investigation through in vitro and in vivo studies, paving the way for its development as a novel breast cancer therapeutic [39].

Table 5. ADME/Tox profile of the top five best compounds screened from Molecular Docking

Parameters	Compounds					Comment
	3074802	71494439	71682086	135934901	11739311	
Physiochemical Property						
Mol. wt.	261.13	291.1	309.15	312.1	340.13	Optimal:100-600
nHA	3	3	5	5	5	0-12
nHD	2	2	2	2	3	0-7
LogP	3.916	3.818	3.593	3.616	4.489	Optimal: 0-3 log mol/L
Lipinski	Accepted	Accepted	Accepted	Accepted	Accepted	MW≤500; logP≤5; nHA≤ 10; nHD ≤5; Good absorption
Synthetic Accessibility Score	1.975	2.007	1.954	2.442	2.346	SA<6 (easy to synthesize)
Absorption						
Caco-2 permeability	-5.099	-5.073	-4.791	-5.062	-5.251	Optimal: higher than -5.15 log cm/s
Pgp-inhibitor	0.101	0.022	0.994	0.12	0.686	0-0.3: excellent; 0.3-0.7: medium; 0.7-1.0: poor
HIA	0.012	0.005	0.004	0.025	0.049	0-0.3: excellent; 0.3-0.7: medium; 0.7-1.0: poor
MDCK Permeability	1.47E-05	1.29E-05	1.95E-05	3.05E-05	7.04E-06	$>2x10^{-6}$cm/s: excellent
Distribution						
Plasma Protein Binding	97.24%	97.50%	96.23%	98.33%	97.35%	Optimal <90% (probability of low therapeutic index)
Volume Distribution	1.772	3.3	1.569	1.012	1.619	Optimal: 0.04-20L/kg
BBB Penetration	0.336	0.406	0.948	0.388	0.439	0-0.3: excellent; 0.3-0.7: medium; 0.7-1.0: poor
Metabolism						
CYP1A2 inhibitor	0.994	0.991	0.938	0.991	0.962	Category 0: Non-substrate / Non-inhibitor Category 1: substrate / inhibitor
CYP1A2 substrate	0.821	0.875	0.885	0.16	0.156	
CYP2C19 inhibitor	0.971	0.973	0.746	0.812	0.494	
CYP2C19 substrate	0.056	0.058	0.071	0.067	0.061	
Excretion						

Clearance	4.353	4.503	9.323	3.381	3.806	≥ 5: excellent;< 5: poor
T1/2	0.807	0.404	0.378	0.525	0.301	0-0.3: excellent; 0.3-0.7: medium; 0.7-1.0: poor
Toxicity						
Hepatotoxicity	0.623	0.791	0.61	0.892	0.798	Category:0; Inactive Category:1; Active (0-0.3: excellent; 0.3-0.7: medium; 0.7-1.0: poor)
AMES	0.226	0.039	0.169	0.881	0.826	
Carcinogenicity	0.079	0.135	0.856	0.339	0.163	
NR-AR	0.581	0.369	0.008	0.007	0.566	

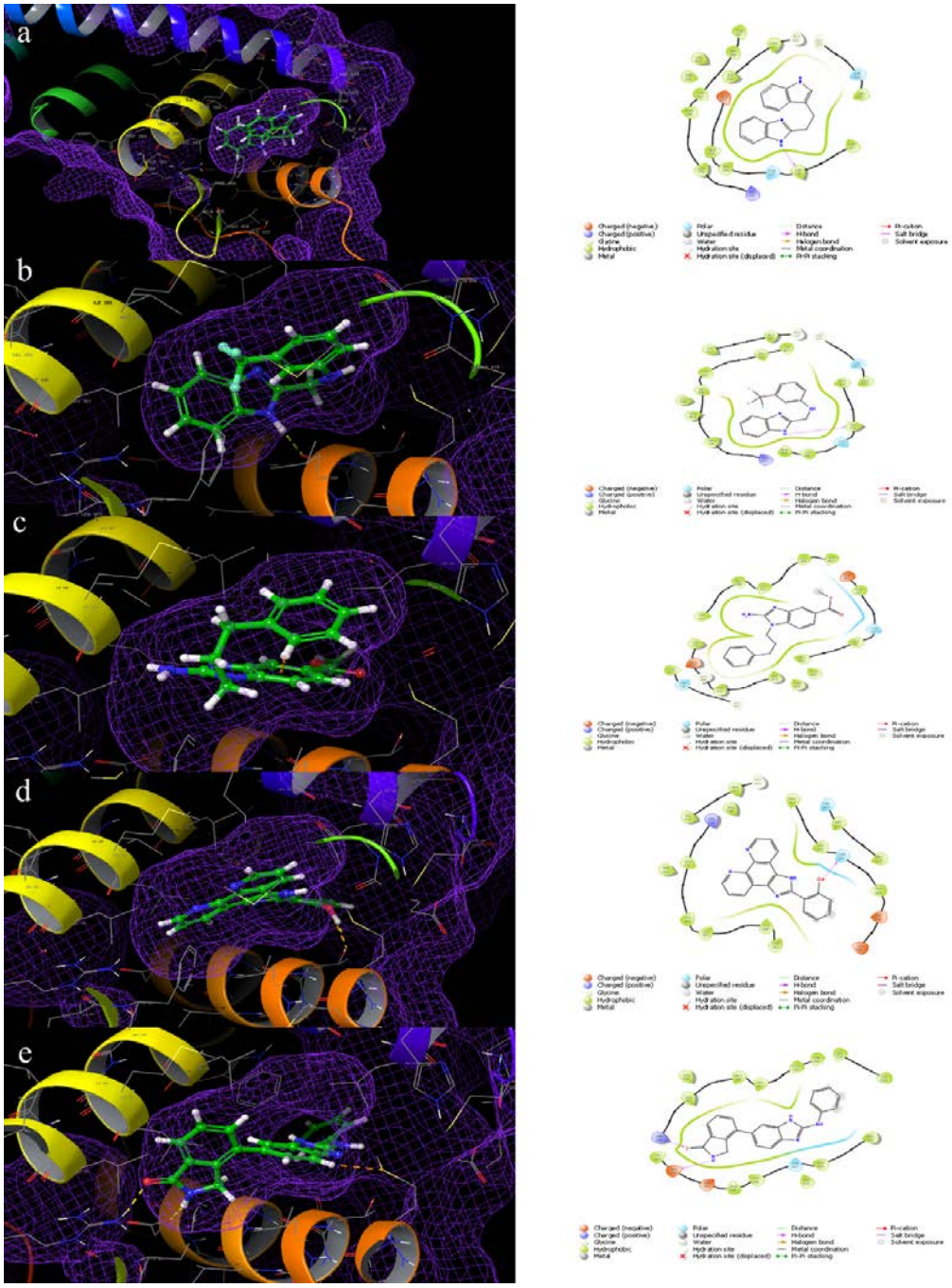


Figure 6: The 2D and 3D interaction images of Glide XP (Xtra Precision) docking of best five compounds against targeted receptor ERα (PDB ID:3ERT) (a) Compound 3074802 (b) Compound 71494439 (c) Compound 71682086 (d) Compound 135934901(e) Compound 11739311.

CONCLUSION

This computational study outputs the design of cancer medication on structural insights and elucidates the selectivity mechanism for the quinoxaline analogues in marketed formulations. The common pharmacophore hypothesis, DDRRR1, was developed, revealing the importance of two hydrogen bond donor groups and three aromatic rings for optimal activity. A robust, predictive atom-based 3D-QSAR model was successfully developed. This model, designated DDRRR_1, exhibited excellent performance metrics, with strong agreement between the experimentally observed and predicted activity for the training set (R^2 = 0.9413) and the test set (R^2 = 0.89). Furthermore, visualization of the 3D-QSAR model provided valuable insights into the structure-activity relationships within the studied molecules.

This understanding allows for identifying potential modifications to the molecular structure that could optimize binding interactions. A virtual screening approach was employed to identify potential inhibitors of the ERα receptor for breast cancer treatment. This strategy involved screening a benzimidazole compound library from PubChem using a pharmacophore model. Docking analysis was then performed on the top hits to assess their binding interaction with the ERα receptor. Moreover, ADME/T profile evaluation of the top 5 docked complexes, compound PubChem ID 3074802 (2-[2-(1H-indol-3-yl) ethyl]1H-benzimidazole) emerged as the most promising candidate.

This compound exhibited a favourable docking score (-9.842 kcal/mol), an optimal pharmacokinetic profile, and a limited toxicity profile, suggesting its potential for further development as a breast cancer therapeutic agent targeting ERα. The innovative use of advanced computational techniques in this work provides a robust framework for the rational design of safer and more effective ERα inhibitors, paving the way for future in vitro and in vivo investigations and potential clinical applications.

ACKNOWLEDGEMENTS

The authors thank the Department of Pharmaceutical Sciences, Maharshi Dayanand University, Rohtak, for assisting with software.

FINANCIAL ASSISTANCE

Nil

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTION

Aastha Sharma contributed to conceptualization, writing, methodology, investigation, and the first draft. Nitish Banga collected the data. Rakesh Kumar Marwaha and Balasubramanian Narasimhan contributed to Methodology, review, editing, and supervision resources.

ABBREVIATIONS

pIC_{50} – The predicted concentration of the compound producing 50% inhibition; NR-AR – Steroid hormone receptors; ADME/T – Absorption, Distribution, Metabolism, Excretion/ Toxicity; QSAR – Quantitative Structure Activity Relationship; nHA – No. of Hydrogen bond Acceptors; nHD – No. of Hydrogen bond Donors; HIA – Human Intestinal Absorption; Glide – Grid Based Ligand Docking with Energetics

REFERENCES

[1] DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics. *CA Cancer J. Clin.*, **64**(1), 52-62 (2014) https://doi.org/10.3322/caac.21203.

[2] Ghoshal S, Rigney G, Cheng D, Brumit R, Gee MS, Hodin RA, Lillemoe KD, Levine WC, Succi MD. Institutional surgical response and associated volume trends throughout the COVID-19 pandemic and postvaccination recovery period. *JAMA Network*, **5**(8), 1-11 (2022) https://doi.org/10.1001/jamanetworkopen.2022.27443.

[3] Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J. Clin.*, **74**(1), 12-49 https://doi.org/10.3322/caac.21820.

[4] Chaturvedi M, Sathishkumar K, Lakshminarayana SK, Nath A, Das P, Mathur P. Women cancers in India: Incidence, trends and their clinical extent from the National Cancer Registry Programme. *Cancer Epidemiol.*, **80**, 1-8 (2022) https://doi.org/10.1016/j.canep.2022.102248.

[5] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin*., **71**(3), 209-49 (2021) https://doi.org/10.3322/caac.21660.

[6] Kulothungan V, Kumar S, Leburu S, Ramamoorthy T, Stephen S, Basavarajappa D, Tomy N, Mohan R, Menon GR, Mathur P. Burden of cancers in India-estimates of cancer crude incidence, YLLs, YLDs and DALYs for 2021 and 2025 based on National Cancer Registry Program. *BMC cancer*, **22**(1), 1-12 (2022) https://doi.org/10.1186/s12885-022-09578-1.

[7] Ali S, Coombes RC. Estrogen receptor alpha in human breast cancer: occurrence and significance. *J. Mammary Gland Biol.*

Neoplasia, **5**, 271-81 (2000) https://doi.org/10.1186/s12885-022-09578-1.

[8] Muchtaridi M, Dermawan D, Yusuf M. Molecular docking, 3D structure-based pharmacophore modeling, and ADME prediction of alpha mangostin and its derivatives against estrogen receptor alpha. *J. Young Pharm.*, **10**(3), 252-9 (2018) https://doi.org/10.5530/jyp.2018.10.58.

[9] Abdullahi SH, Uzairu A, Shallangwa GA, Uba S, Umar AB. Molecular docking, ADMET and pharmacokinetic properties predictions of some di-aryl pyridinamine derivatives as estrogen receptor (Er+) kinase inhibitors. *Egypt J. Basic Appl. Sci.*, **9**(1), 180-204 (2022) https://doi.org/10.1080/2314808X.2022.2050115.

[10] Brisken C, O'Malley B. Hormone action in the mammary gland. *Cold Spring Harb. Perspect. Biol.*, **2**(12), 1-15 (2010) https://doi.org/10.1101/cshperspect.a003178.

[11] Liu L, Tang Z, Wu C, Li X, Huang A, Lu X, You Q, Xiang H. Synthesis and biological evaluation of 4, 6-diaryl-2-pyrimidinamine derivatives as anti-breast cancer agents. *Bioorg. Med. Chem. Lett*, **28**(6), 1138-42 (2018) https://doi.org/10.1016/j.bmcl.2017.12.066.

[12] Gaba M, Mohan. Development of drugs based on imidazole and benzimidazole bioactive heterocycles: recent advances and future directions. *Med. Chem. Res.*, **25**, 173-210 (2016) https://doi.org/10.1007/s00044-015-1495-5.

[13] Shrivastava N, Naim MJ, Alam MJ, Nawaz F, Ahmed S, Alam O. Benzimidazole scaffold as anticancer agent: synthetic approaches and structure–activity relationship. *Archiv der Pharmazie*, **350**(6), 1-80 (2017) https://doi.org/10.1002/ardp.201700040.

[14] Lee YT, Tan YJ, Oon CE. Benzimidazole and its derivatives as cancer therapeutics: The potential role from traditional to precision medicine. *Acta Pharm. Sin B,* **3**(2), 478-97 (2023) https://doi.org/10.1016/j.apsb.2022.09.010.

[15] Spiegel J, Senderowitz HA. Comparison between Enrichment Optimization Algorithm (EOA)-Based and Docking-Based Virtual Screening. *Int. J. Mol. Sci.*, **23**(1), 43 (2021) https://doi.org/10.3390/ijms23010043.

[16] Karaaslan C, Bakar F, Goker H. Antiproliferative activity of synthesized some new benzimidazole carboxamidines against MCF-7 breast carcinoma cells. *Z Naturforsch C. J. Biosci.*, **73**(3-4), 137-45 (2021) https://doi.org/10.1515/znc-2017-0067.

[17] Özdemir A, Turanli S, Çalişkan B, Arka M, Banoglu E. Evaluation of cytotoxic activity of new benzimidazole-piperazine hybrids against human MCF-7 and A549 cancer cells. *Pharm. Chem. J.*, **53**, 1036-46 (2020) https://doi.org/10.1007/s11094-020-02119-9.

[18] Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, Sherman W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comput.*, **6**(5), 1509-19 (2010) https://doi.org/10.1021/ct900587b.

[19] Khedkar SA, Malde AK, Coutinho EC, Srivastava S. Pharmacophore modeling in drug discovery and development: an overview. *Med. Chem.*, **3**(2), 187-97 (2007) https://doi.org/10.2174/157340607780059521.

[20] Muhammed MT, Esin AY. Pharmacophore modeling in drug discovery: methodology and current status. *J. Turk. Chem. Soc A: Chem*, **8**(3), 749-62 (2021) https://doi.org/10.18596/jotcsa.927426.

[21] Dixon SL, Smondyrev AM, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput. Aided Mol. Des.*, **20**, 647-71 (2006) https://doi.org/10.1007/s10822-006-9087-6.

[22] Qing X, Yin Lee X, De Raeymaeker J, RH Tame J, YJ Zhang K, De Maeyer M, RD Voet A. Pharmacophore modeling: advances, limitations, and current utility in drug discovery. *J. Recept. Ligand Channel Res.*, **7**, 81-92 (2014) https://doi.org/10.2147/jrlcr.S46843.

[23] Lu X, Yang, H.; Chen, Y.; Li, Q.; He, S.-y.; Jiang, X.; Feng, F.; Qu, W.; Sun, H. The development of pharmacophore modeling: Generation and recent applications in drug discovery. *Curr. Pharm. Des.*, **24**(29), 3424-39 (2018) https://doi.org/10.2174/1381612824666180810162944.

[24] Bhansali SG, Kulkarni VM. Pharmacophore generation, atom-based 3D-QSAR, docking, and virtual screening studies of p38-α mitogen activated protein kinase inhibitors: pyridopyridazin-6-ones (part 2). *Res. Rep. Med. Chem.*, **4**, 1-21 (2013) https://doi.org/10.2147/rrmc.S50738.

[25] Sanapalli BKR, Yele V, Jupudi S, Karri VVSR. Ligand-based pharmacophore modeling and molecular dynamic simulation approaches to identify putative MMP-9 inhibitors. *RSC Adv.*, **11**(43), 26820-31 (2021) https://doi.org/10.1039/d1ra03891e.

[26] Imrie F, Bradley AR, Deane CM. Generating property-matched decoy molecules using deep learning. *Bioinformatics*, **37**(15), 2134-41 (2021) https://doi.org/10.1093/bioinformatics/btab080.

[27] Dixon SL, Smondyrev AV, Knoll EH, Rao SN, Shaw DE, Friesner RA. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput. Aided Mol. Des.*, **20**, 647-71 (2006) https://doi.org/10.1007/s10822-006-9087-6.

[28] Zia K, Ashraf S, Jabeen A, Saeed M, Nur-e-Alam M, Ahmed S, Al-Rehaily AJ, Ul-Haq Z. Identification of potential TNF-α inhibitors: from in silico to in vitro studies. *Sci. Rep.*, **10**(1), 1-9 (2020) https://doi.org/10.1038/s41598-020-77750-3.

[29] Mali SN, Pandey A. Molecular modeling studies on 2, 4-disubstituted imidazopyridines as anti-malarials: Atom-based 3D-

QSAR, molecular docking, virtual screening, in-silico ADMET and theoretical analysis. *J. Comput. Biophys. Chem.*, **20**(03), 267-82 (2021) https://doi.org/10.1142/S2737416521500125.

[30] Fratev F, Gutierrez DA, Aguilera RJ, Tyagi A, Damodaran C, Sirimulla S. Discovery of new AKT1 inhibitors by combination of in silico structure based virtual screening approaches and biological evaluations. *J. Biomol. Struct. Dyn.*, **39**(1), 368-77 (2021) https://doi.org/10.1080/07391102.2020.1715835.

[31] Baroni M, Clementi S, Cruciani G, Costantino G, Riganelli D, Oberrauch E. Predictive ability of regression models. Part II: Selection of the best predictive PLS model. *J. Chemom.*, **6**(6), 347-56 (1992) https://doi.org/10.1002/cem.1180060605.

[32] Malik R, Mehta P, Srivastava S, Choudhary BS, Sharma M. Pharmacophore modeling, 3D-QSAR, and in silico ADME prediction of N-pyridyl and pyrimidine benzamides as potent antiepileptic agents. *J. Recept. Signal Transduction Res.*, **37**(3), 259-66 (2017) https://doi.org/10.1080/10799893.2016.1217883.

[33] Mellado M, González C, Mella J, Aguilar LF, Viña D, Uriarte E, Cuellar M, Matos MJ. Combined 3D-QSAR and docking analysis for the design and synthesis of chalcones as potent and selective monoamine oxidase B inhibitors. *Bioorg. Chem.*, **108**, 1-11 (2021) https://doi.org/10.1016/j.bioorg.2021.104689.

[34] Madhavi SG, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.*, **27**, 221-34 (2013) https://doi.org/10.1007/s10822-013-9644-8.

[35] Guan L, Yang H, Cai Y, Sun L, Di P, Li W, Liu G, Tang Y. ADMET-score–a comprehensive scoring function for evaluation of chemical drug-likeness. *Medchemcomm*, **10**(1), 148-57 (2019) https://doi.org/10.1039/c8md00472b.

[36] Opo FADM, Rahman MM, Ahammad F, Ahmed I, Bhuiyan MA, Asiri AM. Structure based pharmacophore modeling, virtual screening, molecular docking and ADMET approaches for identification of natural anti-cancer agents targeting XIAP protein. *Sci. Rep.*, **11**(1), 1-18 (2021) https://doi.org/10.1038/s41598-021-83626-x.

[37] Kulkarni S, Singh Y, Biharee A, Bhatia N, Monga V, Thareja S. Molecular docking, 3D-QSAR and simulation studies for identifying pharmacophoric features of indole derivatives as 17β-hydroxysteroid dehydrogenase type 5 (17β-HSD5) inhibitors. *J. Biomol. Struct. Dyn.*, **41**(22), 1-18 (2023) https://doi.org/10.1080/07391102.2023.2175265.

[38] Sheridan RP, Singh SB, Fluder EM, Kearsley SK. Protocols for bridging the peptide to nonpeptide gap in topological similarity searches. *J. Chem. Inf. Comput. Sci.*, **41**(5), 1395-1406 (2001) https://doi.org/10.1021/ci0100144.

[39] Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, Yin M, Zeng X, Wu C, Lu A. ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res.*, **49**(W1), W5-W14 (2021) https://doi.org/10.1093/nar/gkab255.