Original Research Article

Pharmacokinetics Studies of some Diaryl Pyrimidinamine Derivatives as Anti-Cancer Agent: In-Silico Drug Design and Molecular Docking

Sagiru Hamza Abdullahi* (ID)**, Adamu Uzairu** (ID)**, Gideon Adamu Shallangwa, Sani Uba, Abdullahi Bello Umar** (ID)

Department of Chemistry, Faculty of Physical Sciences, Ahmadu Bello University, P.M.B.1045, Zaria, Kaduna State, Nigeria

ARTICLE INFO

ABSTRACT

The cost and duration of novel drug discovery and synthesis have been the significant drawbacks to the chemotherapeutic treatment of breast cancer. To combat these challenges, a validated QSAR model was developed to predict the inhibitive capacities of diaryl-pyridinamine analogs against the MCF-7 breast cancer cell line and to design novel derivatives with better activities. Compound 7, with the highest activity (pIC_{50} = 5.347) and low residual value (0.013), was embraced as the design template. Compared to the template, the designed compounds revealed better activities ranging from pIC50 = 6.06 to 7.14. The results of molecular docking studies demonstrated that the designed compounds exhibit higher binding affinities ranging from -155.9 to -181.4 cal/mol compared to the control drug: Tamoxifen (-155.2 cal/mol). The designed compounds exhibit drug-likeness and promising ADMET properties, as revealed from pharmacokinetics studies. Therefore, the aftermaths of this research could be significant in discovering new and improved anti-breast cancer agents.

GRAPHICAL ABSTRACT
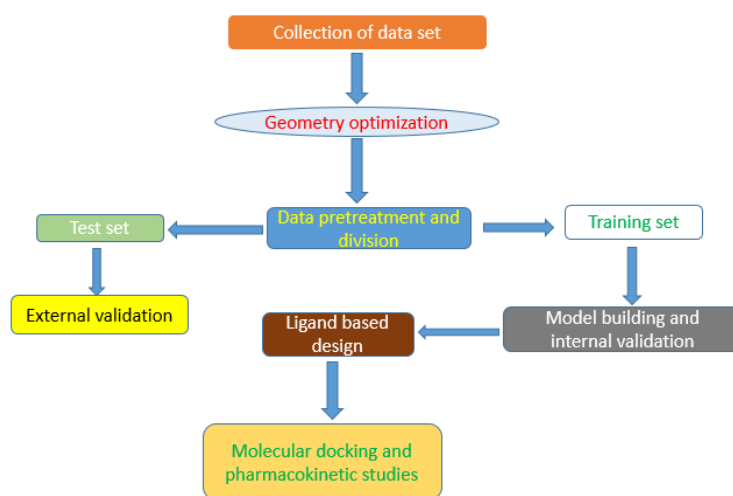
* Corresponding author: Abdullahi, Sagiru Hamza

✉ E-mail: sagirwasai@gmail.com

Introduction

Cancer is still one of the most complicated disorders and is the second most lethal malady after coronary diseases. The incidence and death statistics indicate that it is skyrocketing in developed and economically developing countries [1, 2]. Cancer-related deaths in women are mainly due to breast carcinoma. It is responsible for 23% of all cancer cases, 14% resulting in death, and the incidence rises significantly with age [3]. Breast cancer is a group of diversified tumors that emerge from the epithelial cells. It is a lethal disorder that gave rise to a severe challenge in medicine and immunology [4]. The emergence of resistivity to the approved anti-cancer drugs and numerous side effects has limited their applications. Therefore, alternatives are strongly needed to thoroughly test breast cancer treatment options. The best anti-cancer drugs are supposed to annihilate cancer cells without causing damage to normal tissues [5, 6]. Unfortunately, currently, there are no available drugs that satisfy these criteria. Therefore, these limitations make it necessary to find new drugs with different chemical structure as potential anti-cancer drugs [7]. The mathematical framework that interfaces molecules' molecular structures (molecular descriptors) to their biological activities is called quantitative structure-activity relationship (QSAR).

Moreover, QSAR strategies can save resources and accelerate the development of new molecules for use as drugs, materials, and additives or for whatever purposes [8, 9]. In addition, another critical aspect of drug development is understanding the mechanism of the ligand/receptor interactions by employing the molecular docking simulation method. A molecular docking simulation is a computational approach to anticipating the binding proficiency of the active site residues to specific receptor groups and revealing the strength of interaction

[10-12]. Abdullahi *et al.* [13] have successfully developed a robust QSAR model based on some selected quinazoline derivatives [13]. A molecular docking simulation was executed to determine the lead compound based on binding affinities from which more potent compounds were designed. A well-validated QSAR model was developed in another similar research using a series of chromen-2-one derivatives [14]. The ligand-based drug design approach was utilized to design more potent compounds which were further subjected to molecular docking studies to elucidate their interactions with the Estrogen receptor (ER⁺) binding site residues. In-silico design of new and improved diaryl pyridinamine derivatives through QSAR modeling, molecular docking, and pharmacokinetic profile studies are the main aims of this research work.

Materials and Methods

Collection of data set

A series of diaryl pyrimidinamine analogs and their bioactivities (IC_{50}) in μM unit were retrieved from the literature [15]. 2D structure drawings of the compounds were done cautiously via Chem draw 12.0 software and then changed to a 3D layout with Spartan 14.0 software. The inhibitive capacities of the compounds were transformed to a logarithmic scale using equation 1.0 to lessen the data skewness. 2D structures of the compounds are presented in the supplementary file Table 1.

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

Geometry optimization

Geometry optimization of 3D molecular structures of the compounds was achieved by using Spartan version 14 graphic user interface software utilizing the density functional theory (DFT) calculations. B3LYP/631G* basis set was selected for the calculation of equilibrium geometries because it yields more accurate

results which lead to the achievement of more realistic conformers of the molecules [16, 17]. The geometrically optimized 3D structures were kept in Spatial Document (SD) file and then imported to the PaDEL descriptor tool kit to compute at least 2000 molecular descriptors for all the optimized structures [18].

Data pretreatment and partitioning

The computed descriptors had been pretreated manually and then with the Kennard-Stone data pretreatment software to eliminate redundant and non-relevant descriptors such as those that are constant and highly inter-correlated [19]. Also, the pretreated data was then sliced into training and test sets using the Kennard-Stone data division software [20, 21].

QSAR Model development and Validation

Eighteen (18) compounds from the training data set had been utilized to generate a robust QSAR model using Material studio software version 8.0, and eight (8) test set compounds were used to examine the predictive strength of the model generated [22]. Selection of the most relevant descriptors was accomplished by the genetic function algorithm (GFA). Meanwhile, the dependent (pIC$_{50}$) and independent variables (molecular descriptors) were linearly correlated by Multi-linear regression (MLR) during the evolution process [23]. Statistical parameters adopted in validating a QSAR model include the correlation coefficient (R^2), adjusted R^2 (R$^2_{adj}$), cross-validation coefficient (Q$^2_{CV}$), and correlation coefficient for an external prediction set (R$^2_{pred}$). They are calculated using Equations 2, 3, 4, and 5.

$$R^2 = 1 - \frac{\sum(Y_{exp} - Y_{pred})^2}{\sum(Y_{exp} - Y_{mtraining})^2} \tag{2}$$

$$R^2_{adj} = 1 - (1 - R^2)\frac{N-1}{N-P-1} = \frac{(N-1)R^2 - P}{N-P+1} \tag{3}$$

$$Q^2_{cv} = 1 - \frac{\sum(Y_{pred} - Y_{exp})^2}{\sum(Y_{exp} - Y_{mtraining})^2} \tag{4}$$

p is the number of molecular descriptors found in the QSAR model, N is the number of the training set data, Y$_{exp}$, Y$_{pred}$, and Y$_{mtraining}$ are experimental, predicted, and mean experimental pIC$_{50}$ of the modeling data set molecules [22].

Applicability domain

The applicability domain enables the detection of influential and outlier compounds in the entire data set. Many techniques were utilized to define the domain. However, the leverage (*h$_1$*) approach is the most frequently used within a ±3 standardized residual boundaries for the detection of outlier molecules and the cut-off leverage (*h**) boundary for checking the influential compounds [24]. The leverage and its cut-off values are calculated using equations 5 and 6.

$$h_i = x_i (X^T X)^{-1} X_i^T \tag{5}$$

Where x denotes the descriptor vector of a referred molecule and X refers to the descriptor matrix originating from the modeling (training) set descriptor values. The cut-off leverage (h*) was computed using Equation 6 below:

$$h^* = \frac{3(Q+1)}{N} \tag{6}$$

Where Q represents the number of independent variables in the developed model, N is the number of training molecules set.

**Molecular Docking Studies**

Target protein retrieval and preparation

The crystal structure of the ER$^+$ target protein (pdb id: 3ERT) in complex with 4-hydroxy Tamoxifen was downloaded from the Protein data bank (https://www.rcsb.org/). The crystal structure of the receptor was prepared using Molegro virtual docker (MVD) software by eliminating solvent molecules and co-crystallized ligand enclosed in its crystal structure. Additionally, amino acid residues with structural

errors were repaired and rebuilt and the prepared ER+ receptor was kept in MVD recommended file format (pdb). The 3D

structure of the prepared ER+ receptor is displayed in Figure 1.
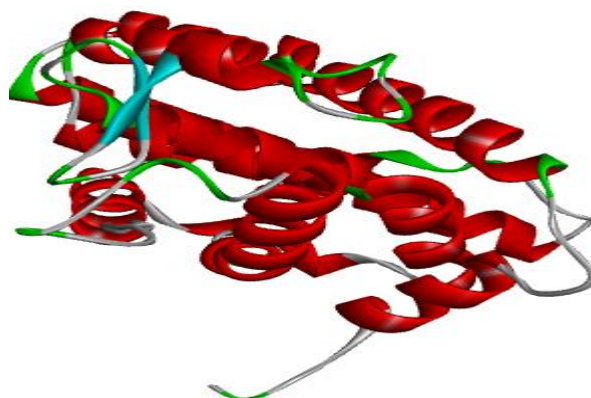
Figure 1. 3D structure of the prepared ER+ receptor

Binding site prediction and ligand-protein docking

The binding site of the ER+ receptor was predicted and set inside a constrained sphere having X, Y, Z coordinates of 24.12, 3.48, and 20.11 Å. The ligands were imported onto the MVD surface, and docking was performed using 0.30Å grid resolution. The Root Mean Square Deviation (RMSD) threshold was set as 1.00 Å for the multiple clusters poses with 100.00 energy penalty values. The docking algorithm was set for a maximum of 1500 iterations with a maximum population size of 50. The docking simulation was run for at least 50 rounds for the 10 poses, and the MolDock score was utilized as the scoring function [25]. Discovery studio visualizer software was utilized to view and interpret of the ligand-ER+ docked complexes.

ADMET and Drug Likeness properties of the designed compounds

SwissADME(http://www.swissadme.ch/index.php)andpkCSM(http://structure.bioc.cam.ac.uk/pkcsm) online tools were utilized to predict the drug-likeness and central ADMET properties of the designed entities to ascertain their viability as drug candidates [26, 27].

Result and Discussions

In this research work, the Kennard-Stone algorithm in Material studio 8.0 software was utilized to partition the data set into the training and the external validation (test) sets. The GFA selected AATSC0m, AATS3m, MATS3s and SpMin7_Bhm as the most relevant descriptors. MLR equation was developed using the training set compounds by adopting these descriptors as the response variables and their biological activities as the dependent variables [28]. In an additional study, it was employed to evaluate the developed QSAR model's predictive potential to predict the pIC_{50} values of the data set, and the results are placed in Table 1. The predicted pIC_{50} values are in good agreement with the experimental values obtained from the literature, and the low residual values and other validation assessments verified the reliability of the reported model. Internal and external validation results of the reported model presented in Table 2 are in good agreement with the recommended validation parameters for robust and stable QSAR models [29]. Hence, the developed QSAR model can be utilized in predicting the inhibitive activities of the diaryl pyridinamine derivatives.

Developed model:

pIC$_{50}$ = - 2.110465633* (**AATSC0m**) - 0.484136308* (**AATSC3m**) + 6.012137349* (**MATS3m**) + 9.020055372* (**SpMin7_Bhm**) + 62.82077

Table 1. Training and test set descriptors used in generating the model.

ID	Exp pIC$_{50}$	Descriptors				Pred pIC$_{50}$	Residuals
		AATSC0m	AATSC3m	MATS3s	SpMin7_Bhm		
Training set							
2	5.47	34.25175	-6.95667	-0.14101	1.389819	5.59	-0.12
3	4.40	34.08369	-5.56008	-0.09939	1.24963	4.25	0.15
4	4.31	34.23231	-2.70653	-0.07243	1.478692	4.79	-0.48
7	5.35	34.06552	-4.23608	-0.05983	1.409742	5.34	0.01
9	4.58	33.91264	-1.90791	-0.00771	1.372874	4.51	0.07
8	4.71	33.89971	-2.9542	-0.02328	1.378786	5.00	-0.29
10	4.48	34.53762	-4.97977	-0.1041	1.418354	4.51	-0.03
11	4.72	34.63302	-5.60664	-0.10586	1.408235	4.51	0.21
12	5.74	34.20166	-5.42138	-0.07563	1.418356	5.60	0.14
14	5.28	34.15875	-4.9514	-0.04372	1.372874	5.25	0.03
16	5.44	34.10365	-5.0325	-0.09399	1.406143	5.40	0.04
17	4.71	34.20949	-7.55718	-0.1408	1.24963	4.71	0.00
18	5.30	33.99074	-0.75184	0.011356	1.512106	5.16	0.14
19	5.29	34.25631	-3.09418	-0.07846	1.477931	4.88	0.41
21	5.34	33.91264	-2.67139	-0.00729	1.409316	5.21	0.13
22	4.77	33.89691	-1.65406	0.010101	1.406112	4.83	-0.06
23	4.02	34.51602	-3.95781	-0.06321	1.408496	4.22	-0.19
26	5.28	34.30282	-6.45322	-0.08053	1.372874	5.45	-0.17
Test set							
1	5.60	34.2563	-6.3495	-0.145	1.40332	5.38	0.22
5	4.57	34.0894	-1.4426	-0.0347	1.5115	4.99	-0.43
6	5.66	34.0657	-3.876	-0.0526	1.4234	5.33	0.33
13	5.39	34.2941	-6.0248	-0.0785	1.40892	5.59	-0.21
15	5.44	34.0894	-4.4924	-0.0884	1.41694	5.30	0.14
20	5.59	33.8969	-2.3583	0.02273	1.4234	5.40	0.18
24	5.38	34.0906	-4.1661	-0.0171	1.4234	5.63	-0.25
25	5.32	34.1347	-4.5658	-0.034	1.40872	5.49	-0.17

Table 2. Statistical parameters for the developed model MLR technique

Name	Threshold	Built Model	Comment	Reference
Friedman Lack of Fit (LOF)	Low value	0.249552	Passed	[29]
Correlation coefficient (R^2)	≥0.6	0.827	Passed	[29]
Cross validation coefficient (Q^2)	≥0.5	0.649	Passed	[29]
Adjusted R-squared (R^2_{adj})	> 0.6	0.773583	Passed	[29]
Critical SOR F-value (95%)	> 2.09	3.226561	Passed	[29]
Difference between R^2 and Q^2 (R^2-Q^2)	≤ 0.3	0.178	Passed	[29]
Minimal number of test set	≥5	8.000	Passed	[29]
Correlation coefficient for the test set (R^2_{test})	≥0.6	0.755	Passed	[29]

Variance Inflation Factor (VIF)

VIF values were computed to study the degree of correlation between the modeled descriptors. A model can be accepted if the descriptor's VIF values fall between 1 and 5, and a value higher than 10 discloses that the model is no longer acknowledgeable and therefore needs to be reexamined [30]. VIF values of a descriptor are calculated using Equation 7 below:

$$VIF = \frac{1}{1-R^2} \qquad (7)$$

R^2 is the correlation coefficient for each descriptor pair.

The VIF values of all the molecular descriptors in the modeling set are calculated and found to be less than 10. This shows the fitness and applicability of the reported model and that the descriptors are independent of each other. Pearson's correlation statistics were performed to further examine the descriptors that appear in the model and check whether the inter-correlation between the descriptors exists [30]. The correlation coefficient between each descriptor was found to be less than ±0.8. These values confirm the non-existence of multicollinearity between the descriptors. Additionally, one-way analysis of variance (ANOVA) at 95% confidence level probability value of each descriptor was found to be less than 0.05. This confirmed the existence of a

correlation between the bioactivity of each compound and the molecular descriptors [30]. VIF values of the modeled descriptors are presented in Table 3.

Mean Effect (ME)

The contribution of each descriptor in the reported model was appraised by computing its mean effect value (ME) using the equation below [25].

$$MF_j = \frac{B_j \sum_{j=1}^{i=n} d_{ij}}{\sum_j^m B_j \sum_i^n d_{ij}} \qquad (8)$$

where MF_j is the mean effect of a descriptor j in a model, β_j depicts the coefficient of the descriptor J in the model and d_{ij} is the value of the descriptor in the data matrix per sample in the model building set, m is the descriptor's numbers that turn up in the model and n is the number of samples in the modeling set. The magnitude and signal of the descriptors indicate the various directions of either increase or decrease in the biological activity of a molecule [22]. Hence, AATSC0m and MATS3s with positive ME values improve the bioactivity of the compounds by increasing their values. In contrast, AATSC3m and SpMin7_Bhm with negative ME values decreased the bioactivity of the compounds when their values increased numerically. ME values of the descriptors are presented in Table 3.

**Table 3**. Pearson's correlation matrix, P-value, VIF and ME of the developed model

	AATSC0m	*AATSC3m*	*MATS3s*	*SpMin7_Bhm*	*P-value*	*VIF*	*ME*
AATSC0m	1				0.000209	2.597	1.246
AATSC3m	-0.52613	1			0.000112	4.029	-0.036
MATS3s	-0.65625	0.786047	1		3.8E-05	3.411	0.0068
SpMin7_Bhm	0.064633	0.605655	0.404267	1	0.033804	2.385	-0.218

Applicability Domain

William's plot of the developed QSAR model is demonstrated in Figure 2. The cut-off leverage (h*) was found to be 0.833. According to their

leverage values, only two compounds from the external validation set (5 and 24) lie beyond the defined range of applicability domain of the developed QSAR model ($h_i > h^*$). Therefore, they are recognized as structurally influential

compounds. These compounds impact the model capability but are not outliers as their standardized residual values are within ±3 range [31].
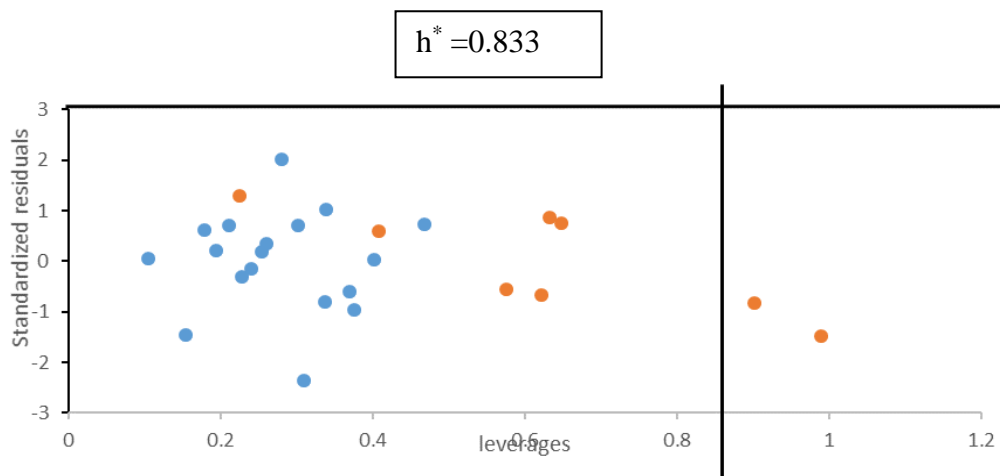

Figure 2. Williams plot of the selected model

Ligand-based drug design

Based on the reported QSAR model, an in-silico screening approach was employed to design more effective diaryl pyrimidinamine analogs. Compound 7 from the training set was opted as the design template due to the high pIC_{50} (5.35) and the least residual value (0.064). The structure of compound 7 and the template used for the design are depicted in Figures 3 and 4. It is modified by introducing various fragments at positions A, B and C as presented in Fig. 4. Five (5) novel compounds with better inhibitive capacities compared to the template and the standard drug: Tamoxifen (pIC_{50} = 4.71), were designed (Table 4). Higher predicted activities of the designed compounds might be due to the addition of electron-donating imidazolidine, dimethyl amine, amino, and isopropyl groups which push electrons to the ring systems of the 2-pyridinamine pharmacophores through +I inductive effect, consequently increasing its electron density and hence increasing its basic character. Designed compound N3 exhibits the highest bioactivity (pIC_{50} = 7.1365). Consequently, it is the most potent compound.
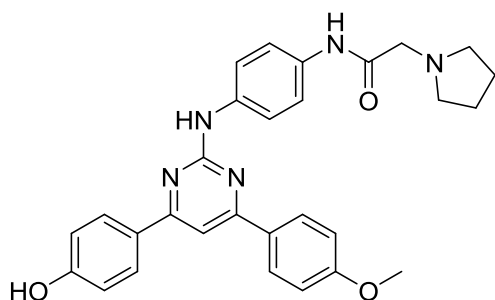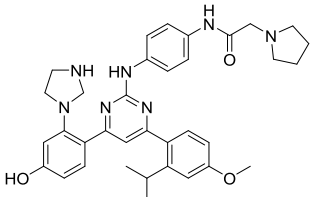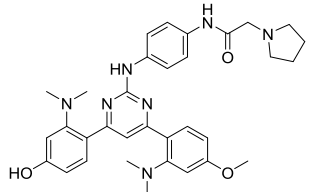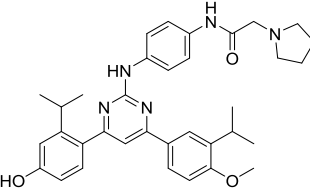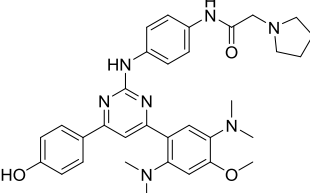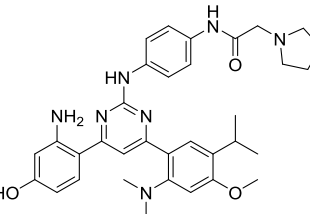

Figure 3. Structure of compound 7
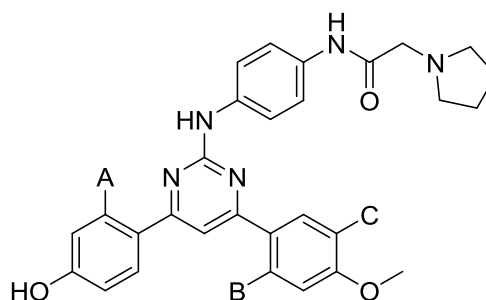
Figure 4. Structure of the template used for design

Table 4. Designed compounds, calculated descriptors and predicted pIC_{50} values

S/No	Structure	AATSC0m	AATSC3m	MATS3s	SpMin7_Bhm	Pred pIC_{50}
N1		34.15311	-2.51836	-0.04862	1.613863	6.23
N2		34.35413	-4.43944	-0.08609	1.582556	6.22
N3		33.79221	-1.94965	-0.04367	1.657602	7.14
N4		34.35413	-4.16222	-0.0599	1.628355	6.66
N5		34.26396	-2.90886	-0.06661	1.612364	6.06

Key: N1: N-(4-((4-(4-hydroxy-2-(imidazolidin-1-yl)phenyl)-6-(2-isopropyl-4-methoxyphenyl)pyrimidin-2-yl)amino)phenyl)-2-(pyrrolidin-1-yl)acetamide
N2: N-(4-((4-(2-(dimethylamino)-4-hydroxyphenyl)-6-(2-(dimethylamino)-4-methoxyphenyl)pyrimidin-2-yl)amino)phenyl)-2-(pyrrolidin-1-yl)acetamido
N3: N-(4-((4-(4-hydroxy-2-isopropylphenyl)-6-(3-isopropyl-4-methoxyphenyl)pyrimidin-2-yl)amino)phenyl)-2-(pyrrolidin-1-yl)acetamide
N4: N-(4-((4-(2,5-bis(dimethylamino)-4-methoxyphenyl)-6-(4-hydroxyphenyl)pyrimidin-2-yl)amino)phenyl)-2-(pyrrolidin-1-yl)acetamide
N5: N-(4-((4-(2-amino-4-hydroxyphenyl)-6-(2-(dimethylamino)-5-isopropyl-4-methoxyphenyl)pyrimidin-2-yl)amino)phenyl)-2-(pyrrolidin-1- yl)acetamide

Molecular docking studies of the designed compounds

The designed compounds were subjected to molecular docking studies with the active site of the ER+ receptor (pdb id = 3ERT) to study their interactions with the active site residues of the target receptor. The docking studies' results revealed that the designed compound's inhibitory activities were correlated with their corresponding docking scores ranging from -155.9 to -181.4 cal/mol. Meanwhile, designed compounds N1 and N3 displayed the highest binding scores of -181.4 and -179.4 cal/mol with the target receptor compared to the standard

drug: Tamoxifen (-155.2 cal/mol) and the rest of the designed compounds. Docking scores and interactions between ER+ receptors and designed compounds are presented in the supplementary file Table 2.

Interpretation of Ligand ‑ Receptor Complex of designed compounds N1, N3 and Tamoxifen

The interaction pattern of the docked complexes of designed compounds N1 and N3 was viewed using discovery studio visualizer software. Their interactions with the target receptor "ER+" was presented in Figs. 5 and 6. Designed compound N1 was observed to have interacted with the ER+ receptor via two (2) conventional hydrogen bonds, a single carbon-hydrogen bond, Pi-anion, Pi-Sigma, Pi-Sulfur, and several hydrophobic Alky and Pi-Alkyl interactions. The hydrogen atom of the hydroxyl group attached to the benzene ring forms a conventional Hydrogen bond with VAL533 at a distance of 2.37Å. The other one is formed between LEU525 with a Hydrogen atom attached to the Nitrogen of the imidazoline ring at a distance 2.055 Å. A carbon-hydrogen bond is observed between LEU346 and the hydrogen atom of the pyrrolidine ring at a distance of 2.63Å. A single Pi-anion electrostatic interaction is found between ASP351 and phenyl ring at a 3.039Å distance. Pi-Sigma interaction is formed between TRP383 and the hydrogen atom of the isopropyl group at a distance of 2.88Å. Pi-Sulfur interaction is found between the Benzene ring and MET343 at a distance of 5.71 Å. Amino acid residues LEU346, LEU349, ALA350, LEU354, and LEU536 forms alkyl interactions while TRP383, VAL533, LEU536, LEU534, LEU539, ALA350, and LEU525 form a hydrophobic Pi-Alkyl interaction.
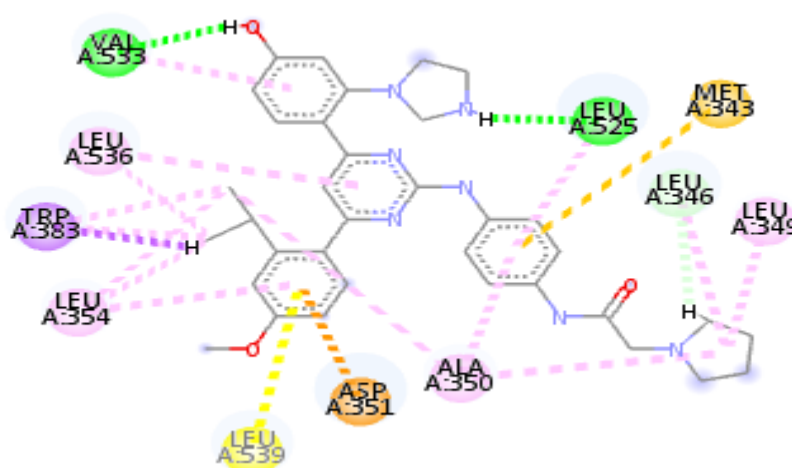
Figure 5. interactions of designed compound N1 with the active site of the ER+ receptor

Designed compound N3 was observed to have interacted with the active site of the ER+ receptor via two (2) conventional hydrogen bonds, VAL533 with the hydrogen atom of the hydroxyl group attached to the benzene ring at a distance of 1.8827Å and THR347 with a hydrogen atom attached to the amido group nitrogen at a distance of 2.653Å. Single Pi-sigma hydrophobic interaction with LYS529 with the pyrimidine scaffold at a distance of 1.99Å. TYR526 formed single Pi-Pi T-shaped hydrophobic interaction with phenyl ring moiety at a distance of 5.055Å. Weak alkyl interactions are also found with LEU346, LEU525, VAL533, and LEU536. Other weak Pi-Alkyl interactions are formed with TYR526, LYS529, VAL533, PRO535, MET522, and LEU525 amino acid residues.
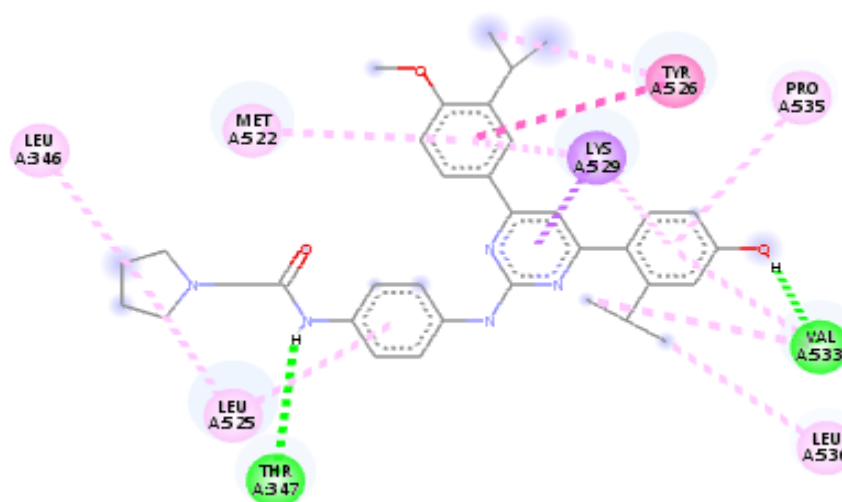
Figure 6. Interactions of designed compound N3 with the active site of the ER+ receptor

The interaction mode of the standard drug: Tamoxifen, was through a single conventional hydrogen bond, five (5) carbon-hydrogen, bonds and numerous hydrophobic Alkyl and Pi-Alkyl interactions. GLY420 forms a conventional hydrogen bond with the –OH group hydrogen at a 1.79Å distance, GLY421 forms a carbon-hydrogen bond with the –OH group oxygen atom at a distance of 2.37 Å, THR347 and ASP351 form

three other carbon-hydrogen bonds at 2.89, 2.20, and 2.62 Å, ALA 350 forms another with trimethyl amine hydrogen at 2.79 Å. PHE404, MET421, LEU346, LEU387, LEU 349, and LEU525 form hydrophobic Alkyl and Pi-Alkyl interactions. 2D interaction modes of Tamoxifen with the active site of the 3ERT receptor are depicted in Figure 7.
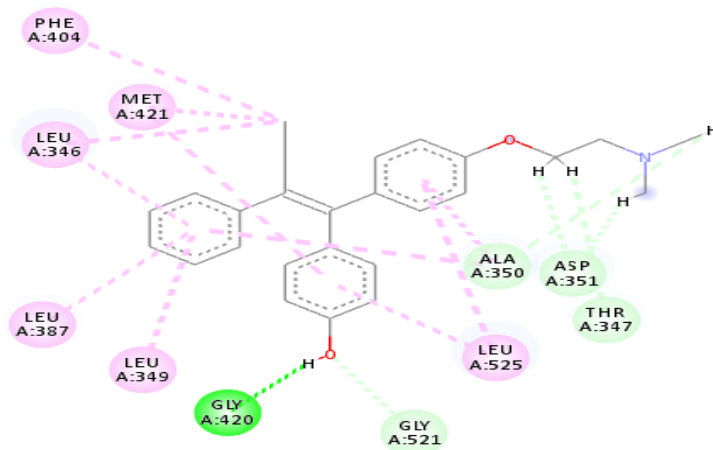
Figure 7. 2D interaction modes of Tamoxifen with the active site of the ER+ receptor.

Apart from hydrogen bonds, and hydrophobic interactions observed for N1, N3, and Tamoxifen, the designed compounds formed additional interactions such as electrostatic Pi-anion, Pi-sigma, Pi-Pi T-shaped, and Pi-Sulfur interactions.

These additional interactions provide a tangible suggestion to support the assertion that the designed compounds N1 and N3 bind efficiently with the ER+ binding pocket compared to the standard drug "Tamoxifen".

Pharmaco-kinetics and ADMET properties of the designed compounds

To ensure that the designed diaryl pyrimidinamine analogs are the possible drug candidates, their ADMET, and pharmacokinetic properties were predicted, and presented in supplementary file Table 3 and 4. All the designed compounds violated only one of Lipinski's rules of five (MW > 500). As such, they possess drug-likeness properties. Their bioavailability score values were 0.55, illustrating that they possess an excellent form of permeability and bioavailability [32]. Their synthetic accessibility values are less than 5, affirming that they are easily synthesized when referenced to a scale between 1 (easily synthesized) and 10 (challenging to synthesize). Additionally, they exhibit high intestinal (Human) absorption, ranging from 84.545% to 96.085%. As such, they are found to be well soaked by the human intestine since for a molecule to be poorly absorbed, and its absorbance should be at most 30% [33].

BBB and CNS penetration ratings were utilized to establish if a molecule will pass through the blood-brain barrier and central nervous system. A log BB > 0.3 suggests that a molecule can easily flow through the blood-brain barrier, whereas log BB < -1 suggests that such a molecule will be poorly distributed. Additionally, log PS > -2 illustrates that a molecule can be readily dispersed into the central nervous system, while log PS < - 3 suggests poor dispersal of such a molecule [34]. Predicted log BB and log PS of the designed compounds indicated that they have no potential of crossing the blood-brain barrier but are readily dispersed to the central nervous system, thus, indicating lower toxicity profiles.

The body's biochemical modification of a drug candidate is best described by its metabolism. As a result, drugs usually offer numerous metabolites, which might differ in pharmacological and physicochemical properties

[33]. A class of super enzymes called Cytochrome P450 (CYP450) plays a substantial part in the drug's metabolism as it is the primary liver protein system accountable for oxidation (phase-1 metabolism), as in the case of this research. In addition, cytochrome CYP3A4 inhibition is an essential phenomenon in this study [34]. This research revealed that the designed compounds are the inhibitors of 2C19, 2C9, and CYP3A4.

Clearance relates the drug level in the body to its elimination rate. Low total clearance value anticipated an eminent endurance of the drugs in the human body, and all the designed compounds showed good endurance in the body for the drug. Additionally, examining the toxicity level is fundamental as it plays a significant role in choosing the best drug candidates. This study showed that the designed compounds do not display any serious toxicity threat.

Conclusions

A genetic function algorithm coupled with multilinear regression analysis was employed to develop a robust QSAR model based on a series of diaryl pyridinamine. The developed model was found to be statistically significant both internally (R^2_{train} = 0.827, R^2_{adj} = 0.774, and Q^2_{cv} = 0.649) and externally (R^2_{test} = 0.755). Compound 7 was a template to design five (5) more potent diaryl pyrimidine analogs. The molecular docking performed between the design compounds and the ER+ active site showed better binding scores ranging from -155.962 to -181.444 cal/mol compared to the standard drug: Tamoxifen (Docking score = -145.933 cal/mol). Moreover, ADMET and drug-likeness predictions suggested that the designed compounds are viable drug candidates that are orally safe with no toxicity threat. Hence, the findings of this research could help design new and improved anti-breast cancer agents.

Acknowledgment

The authors genuinely acknowledge the physical chemistry research team under the leadership of Prof. Adamu Uzairu for their guidance and motivation during this research work.

Disclosure statement

The authors reported no potential conflict of interest.

ORCID

Sagiru H. Abdullahi 🆔: 0000-0002-2976-9636
Adamu Uzairu 🆔: 0000-0002-6973-6361
Abdullahi B. Umar 🆔: 0000-0003-0984-5969

References

[1] J. Liu, B. Ming, G.H. Gong, D. Wang, G.L.Bao, L.J. Yu, *RSC Adv.,* **2018**, *8*, 4386–4416. [CrossRef], [Google Scholar], [Publisher]

[2] E. Golipour-Chobar, F. Salimi, RG Ebrahimzadeh, *Monatsh Chem.,* **2020**, *151*, 309–318. [CrossRef], [Google Scholar], [Publisher]

[3] X. Li, C. Wu, X. Lin, X. Cai, L. Liu, G. Luo, *Eur. J. Med. Chem.,* **2019**, *161*, 445–455. [CrossRef], [Google Scholar], [Publisher]

[4] A. Mullard, *Nat. Rev. Drug Discov.*, **2011**, *11*, 91–94. [CrossRef], [Google Scholar], [Publisher]

[5] S.H. Abdullahi, A. Uzairu, M.T. Ibrahim A.B. Umar, *Bull. Natl. Res. Cent.,* **2021**, *45*, 167. [CrossRef], [Google Scholar], [Publisher]

[6] F. Kamali, G. Ebrahimzadeh-Rajaei, S. Mohajeri, A. Shamel, M. Khodadadi-Moghaddam, *Inorg. Chem. Commun.*, **2022**, *141*, 109527. [CrossRef], [Google Scholar], [Publisher]

[7] I.A. Al-Suwaidan, A.A. Abdel-Aziz, T.Z. Shawer, R.R. Ayyad, A.M. Alanazi, A.M. El-Morsy, *J. Enzyme Inhib. Med. Chem.,* **2016**, *31*, 78–89. [CrossRef], [Google Scholar], [Publisher]

[8] B.A. Umar, A. Uzairu, G.A. Shallangwa, U. Sani, *Netw. Model. Anal. Health Inform. Bioinforma.,* **2019**, *8*, 22. [CrossRef], [Google Scholar], [Publisher]

[9] T. Qidwai, *In Silico Pharmacol.,* **2017**, *5*, 6. [CrossRef], [Google Scholar], [Publisher]

[10] S.H. Abdullahi, A. Uzairu, G.A. Shallangwa, S. Uba, A.B. Umar, *Egypt. J. Basic Appl. Sci.,* **2022**, *9*, 180–204. [CrossRef], [Google Scholar], [Publisher]

[11] O. Daoui, S. Elkhattabi, S. Chtita, R. Elkhalabi, H. Zgou, A.T. Benjelloun, *Heliyon,* **2021**, *7*, e07463. [CrossRef], [Google Scholar], [Publisher]

[12] G.M. Benhander, A.A.A. Abdusalam, *Chemistry Africa,* **2022**, *5*, 57–67. [CrossRef], [Google Scholar], [Publisher]

[13] S.H. Abdullahi, A. Uzairu, G.A. Shallangwa, S. Uba, A.B. Umar, *Bull. Nat. Res. Cent.,* **2022**, *46*, 2. [CrossRef], [Google Scholar], [Publisher]

[14] S.H. Abdullahi, A. Uzairu, G.A. Shallangwa, S. Uba, A.B. Umar, *Bull. Nat. Res. Cent.,* **2022**, *46,* 177. [CrossRef], [Google Scholar], [Publisher]

[15] L. Liu, Z. Tang, C. Wu, X. Li, A. Huang, X. Lu, Q. You, H. Xiang, *Bioorg. Med. Chem. Let.,* **2018**, *28*, 1138–1142. [CrossRef], [Google Scholar], [Publisher]

[16] F. Kamali, G.E. Rajaei, S. Mohajeri, A. Shamel, M. Khodadadi-Moghaddam, *Monatsh Chem.,* **2020**, *151*, 711–720. [CrossRef], [Google Scholar], [Publisher]

[17] M.H.N. Zamanpour, G. Ebrahimzadeh, *Theor. Chem. Acc.,* **2006**, *115*, 343–347. [CrossRef], [Google Scholar], [Publisher]

[18] C.W. Yap, *J. Comput. Chem.,* **2011**, *32*, 1466–1474. [CrossRef], [Google Scholar], [Publisher]

[19] R.W. Kennard, L.A. Stone, *Technometrics,* **1969**, *11*, 137–148. [Google Scholar], [Publisher]

[20] A. Tropsha, P. Gramatica, V.K. Gombar, *Mol. Inform.,* **2003**, *22*, 69–77. [CrossRef], [Google Scholar], [Publisher]

[21] S. Khamouli, S. Belaidi, M. Ouassaf, T. Lanez, S. Belaaouad, S. Chtita, *J. Biomol. Struct. Dyn.,* **2022***, 40*, 1285–1298. [CrossRef], [Google Scholar], [Publisher]

[22] S.E. Adeniji, D.E. Arthur, M. Abdullahi, A. Abdullahi, F.A. Ugbe, *J. Biomol. Struct. Dyn.*, **2022**, *40*, 4004–4020. [CrossRef], [Google Scholar], [Publisher]

[23] M. Taha, S. Sultan, M. Herizal M.Q. Fatmi, M. Selvaraj, K. Ramasamy, S.A. Halim, S.M. Lim, F. Rahim, K. Ashraf, A. Shehzad, *J. Saudi Chem. Soc.,* **2019**, *23*, 1168–1179. [CrossRef], [Google Scholar], [Publisher]

[24] R. Thomsen, M.H.Christensen, *J. Med. Chem.,* **2006**, *49*, 3315–3321. [CrossRef], [Google Scholar], [Publisher]

[25] M. El Fadili, M. Er-Rajy, M. Kara, A. Assouguem, A. Belhassan, A. Alotaibi, N.N. Mrabti, H. Fidan, R. Ullah, S. Ercisli, S. Zarougui, M. Elhallaoui, *Pharmaceuticals,* **2022**, *15*, 670. [CrossRef], [Google Scholar], [Publisher]

[26] A. Daina, O. Michielin, V. Zoete, *Sci. Rep.,* **2017**, *7*, 42717. [CrossRef], [Google Scholar], [Publisher]

[27] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, *Adv. Drug Deliv. Rev.,* **1997**, *23*, 3–25. [CrossRef], [Publisher]

[28] H. Hadni, M. Elhallaoui, *New J. Chem.*, **2020**, *44*, 6553-6565. [CrossRef], [Google Scholar], [Publisher]

[29] A. Tropsha, P. Gramatica, V.K. Gombar, *QSAR Comb. Sci.,* **2003**, *22*, 69–77. [CrossRef], [Google Scholar], [Publisher]

[30] S.E. Adeniji, D.E. Arthur, M. Abdullahi, A. Haruna, *Chemistry Africa,* **2020**, *3*, 391–408. [CrossRef], [Google Scholar], [Publisher]

[31] P. Gramatica, E. Giani, E. Papa, *J. Mol. Graph. Model.,* **2007**, *25*, 755–766. [CrossRef], [Google Scholar], [Publisher]

[32] R. Thomsen, M.H.Christensen, *J. Med. Chem.,* **2006**, *49*, 3315–3321. [CrossRef], [Google Scholar], [Publisher]

[33] B. Shaik, S.P. Gupta, T. Zafar, K. Anita, *Lett. Drug Des. Discov.,* **2018**, *15*, 169–180. [CrossRef], [Google Scholar], [Publisher]

[34] M. Šrejber, V. Navrátilová, M. Paloncýová, V. Bazgier, K. Berka, P. Anzenbacher, M. Otyepka, *J. Inorg. Biochem.,* **2018**, *183*, 117–136. [CrossRef], [Google Scholar], [Publisher]