



## Original Research Article

# Theoretical Investigation and Design of Novel Anti-proliferative Agents against Hepatocellular Carcinoma from Benzimidazole-Chalcone derivatives

John Philip Ameji <sup>1</sup> \* , Ameh Ohiga Ebune <sup>1</sup>, Ibraheem Wasiu Aderemi <sup>2</sup>, Abdullahi Moyosore <sup>3</sup>, Godwin Idah <sup>1</sup>

<sup>1</sup> Department of Chemistry, Federal University Lokoja, P.M.B. 1154, Lokoja Nigeria

<sup>2</sup> Department of Science Laboratory Technology, Osun State College of Technology, Esa-Oke, Nigeria

<sup>3</sup> Department of Chemistry, Federal College of Education Kastina, P.M.B. 2041, Kastina Nigeria

## ARTICLE INFO

## Article history

Submitted: 14 October 2022

Revised: 16 December 2022

Accepted: 10 January 2023

Available online: 16 January 2023

Manuscript ID: [AJCA-2210-1336](#)

Checked for Plagiarism: Yes

DOI: [10.22034/AJCA.2023.365669.1336](#)

## KEYWORDS

Hepatocellular carcinoma

EGFR

Doxorubicin

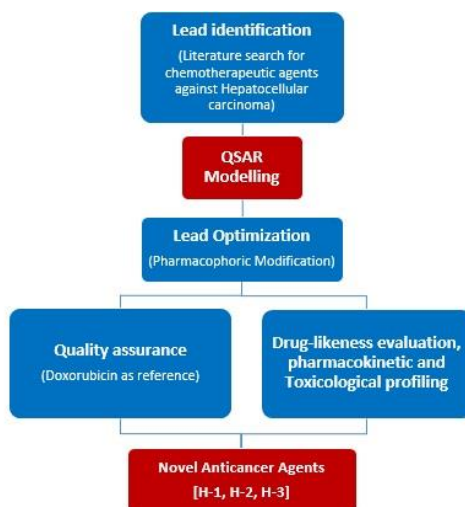
HepG2

Benzimidazole-chalcone

## ABSTRACT

Hepatocellular carcinoma is recognized as a major cause of cancer-related deaths worldwide and has been a healthcare burden globally. In the search for novel chemotherapeutic agents for treating this disease, a set of benzimidazole-chalcone derivatives with proven inhibitory activities against HepG2 cancer cell line were subjected to QSAR modeling. Structural optimization of the most active derivatives guided by the dominant descriptors in the validated model ( $R^2 = 0.93$ ,  $R^2_{Adj} = 0.92$ ,  $Q^2_{L00} = 0.91$ ,  $R^2_{Pred} = 0.81$ ) led to the design of more potent analogs; H-1, H-2, and H-3, with predicted  $IC_{50}$  values of  $26.07\mu M$ ,  $43.38\mu M$ , and  $178.69\mu M$ , respectively. Molecular docking simulations of these novel ligands against the active sites of epidermal growth factor receptor (EGFR) protein target recorded binding energy values of -8.4, -8.9, and -8.1 for H-1, H-2, and H-3, respectively. These values are better compared to -7.6 kcal/mol recorded for Doxorubicin, a standard drug for managing liver cancer used herein for quality assurance. Also, ADMET evaluation of the designed drug candidates reveals that they possess excellent pharmacokinetic and toxicity profiles. It is envisaged that the wealth of information obtainable from this work could help to discover and develop novel chemotherapeutic agents for treating systemic and early-stage hepatocellular carcinoma.

## GRAPHICAL ABSTRACT



\* Corresponding author: Ameji, John Philip

✉ E-mail: [ameji4real55@gmail.com](mailto:ameji4real55@gmail.com)

© 2023 by SPC (Sami Publishing Company)

## Introduction

Hepatocellular carcinoma (HCC) is a liver cancer that arises from abnormal proliferation of the hepatocytes and accounts for 80% of primary liver cancer cases [1]. HCC is recognized as a significant cause of cancer-related deaths worldwide and has been a healthcare burden globally [2]. The major causes of the infection include endemic hepatitis B and C viruses, alcohol, non-alcoholic fatty liver diseases, and exposure to aflatoxin [3-5]. The early stage of HCC is often managed via surgery, radiofrequency ablation, and liver transplant. However, the asymptomatic nature of the disease at its early stage and the dearth of effective systemic treatments in the advanced stage have led to unfavorable treatment outcomes [6-8].

The emergence of chemotherapeutic agents targeting specifically the signaling pathways provide a new era of selective chemotherapeutic hits that could violently attack cancerous cells and the tumor microenvironment that, causes tumor proliferation with greater efficacy and low side effects on the normal cells [9,10]. An essential signaling pathway is the tyrosine kinase (T.K.) Epidermal Growth Factor Receptor (EGFR) regulates cell growth, survival, multiplication, differentiation, and apoptosis [11]. EGFRs are over-expressed in many human cancer cells, HCC inclusive [12-14], making this macromolecule a promising target in the rational search for novel chemotherapeutic agents.

The significant challenges militating against the discovery and development of novel drugs include the enormous time and resources expended from the discovery stage to the preclinical and clinical levels. However, applying Computer Aided Drug Design techniques such as quantitative structure-activity relationship (QSAR) modeling, molecular docking, drug-likeness evaluation, and ADMET prediction has helped to mitigate these challenges recently. QSAR modeling provides the mathematical link

between the biological properties of molecules and their descriptors. The information garnered from the model could then be used to optimize the potencies of the molecules. Also, Molecular docking is a technique that helps to ascertain the binding affinities of therapeutic compounds to the active sites of a protein target [15,16]. ADME/T, on the other hand, ADME/T defines the fates of bioactive compounds in the biological system with specific reference to their absorption, distribution, metabolism, excretion, and toxicity. The success of a drug in clinical trials is greatly dependent on these properties. Hence, computer-aided ADME/T profiling of drug candidates is essential to modern drug design. Early prediction of these properties in drug candidates helps minimize the attrition rate in drug development [17-23].

Organic compounds, especially those containing benzimidazole moiety, have been identified as an essential pharmacophore in the discovery of novel anti-tumor agents [24-28]. This study aims to apply computer-aided techniques to design highly potent and less toxic benzimidazole-chalcone-based chemotherapeutic agents against HCC by targeting the EGFR receptor. This enzyme governs the growth and dissemination of tumor cells. In achieving the stated aim, a literature search was conducted for benzimidazole-chalcone analogs with proven anti-cancer activities against HCC. The bioactive compounds were subsequently subjected to QSAR modeling, molecular docking simulation, design, drug-likeness, and ADMET profiling.

## Experimental

### Methods

A dataset of twenty four (24) benzimidazole-chalcone derivatives with proven anti-cancer activities against the HepG2 cancer line was accessed from reported work [29]. Meanwhile, the reported activities of these series of compounds were transformed from  $IC_{50}$  ( $\mu$ M) to

pIC<sub>50</sub> using the expression; pIC<sub>50</sub> = -log IC<sub>50</sub> to make their biological activity values fall within statistically normal distribution [30,31]. The bioactive compounds' 2D chemical structures, pIC<sub>50</sub> values, and IUPAC nomenclatures are presented in Table S1 as a supplementary file. The geometries of all the molecules were optimized to determine their stable conformations using the DFT (B3LYP/6-31G\*) method of Spartan 14 software. PaDEL Descriptor tool was used to calculate nearly 2045 pool descriptors for all the optimized structures. Filtration of the calculated descriptors was carried out with V-WSP data pretreatment tool 1.2v to eliminate descriptors with zero and redundant values. Subsequently, the filtered descriptors were partitioned into a 70% training set for model development and a 30% test set for external validation of the model by using Kennard and Stone's algorithm of the DTC laboratory. The QSAR model was built using the Genetic Function Approximation (GFA) tool of Material Studio v8.0 by setting mutation probability to 0.3 [32,33].

#### Model Validation

The generated QSAR model was validated using the following internal validation parameters; Friedman Lack of fit (LOF), square of correlation coefficient (R-squared), Adjusted R-squared, Cross validated R-squared (Q<sup>2</sup><sub>cv</sub>). Likewise, external validation of the QSAR model is crucial to obtain models with more reliable predictions. The optimum QSAR model was externally validated using the test set of eight molecules by using the model to predict the pIC<sub>50</sub> of the compounds [34].

#### Model Applicability Domain

Applicability Domain (A.D.) refers to the physicochemical, structural, or biological space in which the training set of the model has been developed. The resulting model can be reliably

applicable only to those compounds which are inside this domain. It helps to estimate the uncertainty in predicting the bioactivities of a congeneric set of molecules. The AD of the optimum model was defined using the standardization approach in A.D. executable jar file of the DTC laboratory at <http://poi.apache.org> [35].

#### Design

The primary essence of this is to develop analogs of benzimidazole-chalcones with enhanced potencies and excellent ADMET properties. In achieving this feat, compounds 9 and 13 were selected as templates for designing new analogues because they fall within the model's AD and display the best inhibitory activity against the HepG2 cell line. Optimization of the potencies of the templates was achieved via their structural modifications guided by the dominant descriptors in the model, giving rise to newly designed ligands. The IC<sub>50</sub> values of these new bioactive analogs were predicted using the optimum QSAR model.

#### Molecular Docking Simulation

To substantiate the potencies of the designed ligands and the templates against the HepG2 cell line, they were subjected to molecular docking simulation against EFGR protease to examine their binding affinities and mechanisms of interaction with the active sites of the protein target. The geometry-optimized structures of the compounds were prepared using the AutoDock Vina interface and saved as pdbqt files [36]. The 3-D crystal structure of EGFR kinase in complex with allosteric inhibitor (PDB Code: 7JXQ) was retrieved from the RCSB protein data bank at [www.rcsb.org/pdb](http://www.rcsb.org/pdb). The protein was prepared on the Discovery Studio interface, where water molecules and heteroatoms were removed. Using the AutoDock Vina interface, missing atoms were checked and repaired, partial atomic charges

were assigned, and all necessary valency tests and H-atom additions were performed on the prepared protein. EasyDock, a graphical user interface of AutoDock Vina virtual screening software, was used to carry out the docking and the grid box revealed the following binding locations in EGFR; center-x = 48.2036; center-y = 0.0159; center-z = 70.1600; Dimension-x = 59.5890; Dimension-y = 54.6382; Dimension-z = 88.8233. The Discovery Studio Visualizer v16.1.0.15350 was deployed to analyze the kind of interactions in the stable protein-ligand complexes formed. Likewise, the chemical structure of Doxorubicin, a validated anti-cancer drug used as a reference for comparison in this study, was geometry-optimized, prepared, and docked with EGFR protein target to examine its binding affinity to the target [37,38].

#### *Drug-Likeness, Pharmacokinetics, and Toxicity Assessment*

Drug-likeness evaluation was performed on the designed compounds to ascertain their oral bioavailability. The Lipinski's and Veber's rules were used to evaluate this critical parameter. The former rule states that for a drug candidate to be orally bioavailable, it must not violate more than one of the following indices; molecular weight (Mw) < 500, number of hydrogen bond donors (HBD) < 5, octanol/ water partition coefficient (Log P) < 5 and number of hydrogen bond acceptors (HBA) < 10 [39]. The Veber's rule, on the hand, states that for a drug candidate to be orally bioavailable, its topological polar surface area (TPSA) and the number of rotatable bonds (NRB) must be less than 140 Å<sup>2</sup> and 10, respectively. The aforementioned physicochemical properties of the molecules were calculated with the aid of the SwissADME

([www.swissadme.ch/](http://www.swissadme.ch/)) online tool and the DataWarrior chemoinformatic program.

Likewise, pharmacokinetics deals with the study of the fates and distribution of pharmacological compounds in the biological system with particular reference to their absorption (A), distribution (D), metabolism (M), and excretion (E). Early ADME and toxicity (T) assays of potential drug candidates are necessary to minimize their failure rates due to weak pharmacokinetic and toxicity profiles during clinical trials. The newly designed ligands were analyzed for their ADMET properties with the aid of Swiss ADME server (<http://www.swissadme.ch/index.php>) and DataWarrior chemoinformatic program.

## **Results**

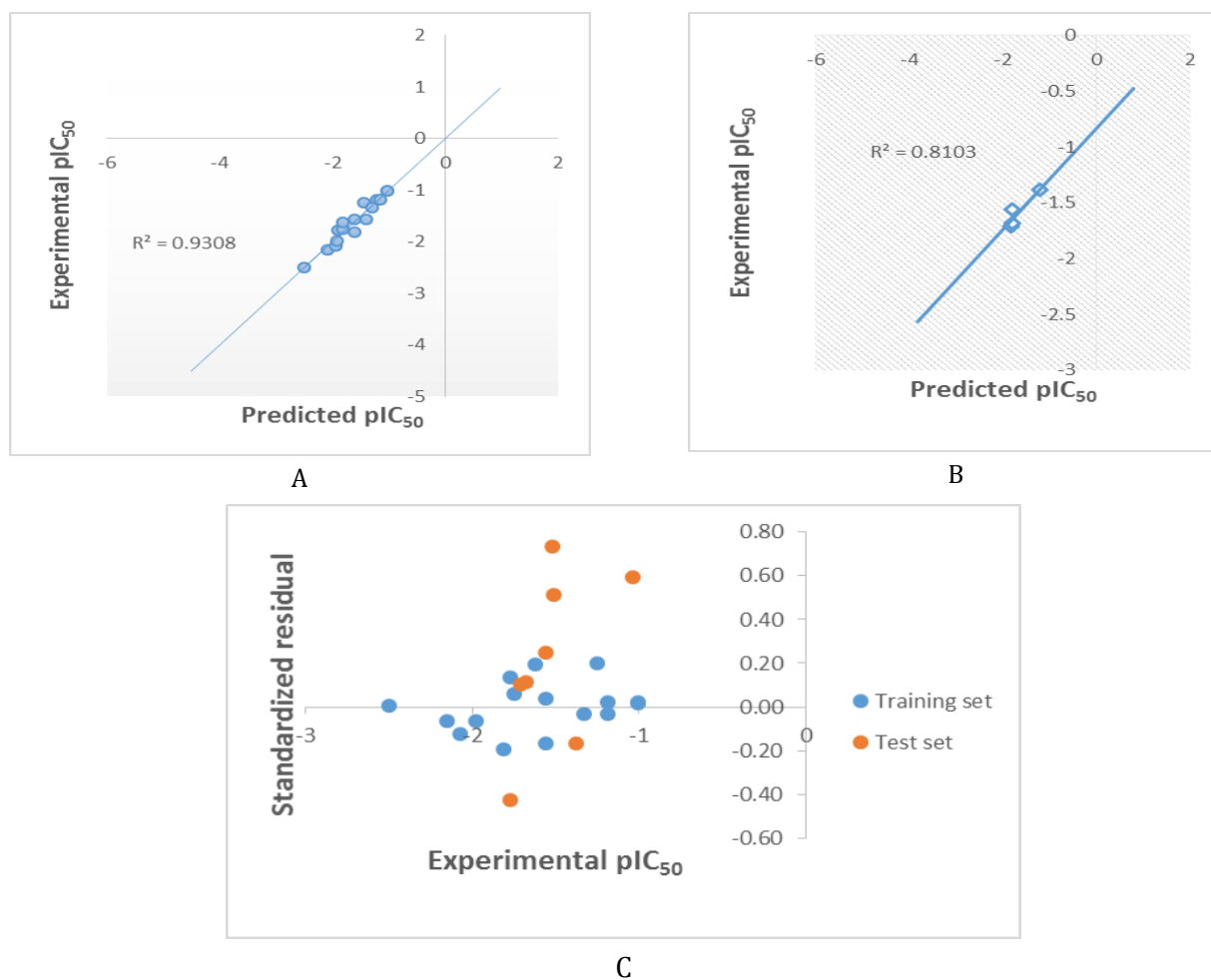
### *GFA-QSAR Model and Validation*

It is an established paradigm in Chemistry that whatever property a molecule exhibit is a function of the descriptors of its structure. QSAR model links the descriptors of a molecule to its biological properties using a regression equation. The GFA-derived QSAR model that connects the anti-cancer properties of the studied bioactive compounds to their molecular descriptors is shown by equation 1. The validation parameters for the regression model and the definitions of the descriptors in the model are presented in [Table 1](#) and [Table 2](#), respectively. [Table 3](#) presents the correlation matrix for all the descriptors. Furthermore, [Figure 1](#) presents the plot of experimental pIC<sub>50</sub> versus predicted pIC<sub>50</sub> for training and test set molecules, as well as the residual plot of the QSAR model.

$$\text{pIC}_{50} = 0.1188 * \text{ZMIC2} - 0.2277 * \text{RDF120m} - 5.4171 \quad (1)$$

**Table 1.** Validation Parameters of the Model

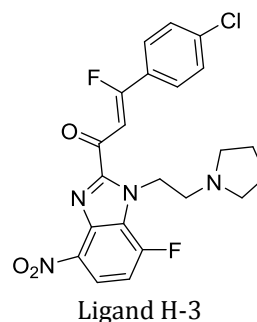
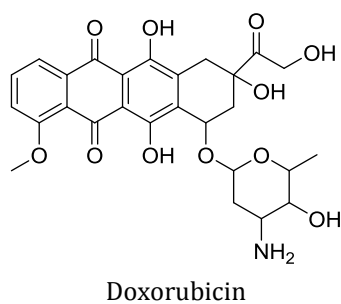
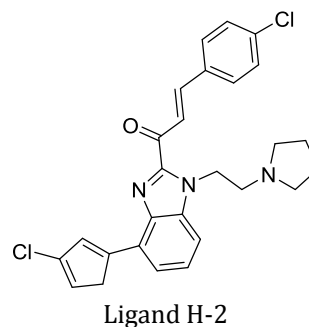
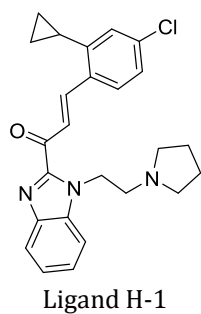
S/n	Parameter	Threshold	Model value	Statement
1	Square of Coefficient of determination ( $R^2$ )	$\geq 0.6$	0.93	Excellent
2	Adjusted R-squared ( $R^2_{Adj.}$ )	$\geq 0.6$	0.92	Stable
3	Cross validated R-squared ( $Q^2_{L00}$ )	$\geq 0.5$	0.91	Reliable
4	Predictive R-squared ( $R^2_{pred}$ )	$\geq 0.5$	0.81	Robust
5	$R^2 - Q^2_{L00}$	$\leq 0.3$	0.02	Stable
6	Friedman LOF	low	0.03	Stable

**Figure 1.** The plot of experimental  $pIC_{50}$  verses predicted  $pIC_{50}$  for the training set (A), test set (B), and the residual plot (C) of the model**Table 2.** Symbols, definitions, and classes of the descriptors in equation 1

S/n	Descriptor	Definition	Class
1	ZMIC2	Z-modified information content index (neighborhood symmetry of 2-order)	2D
2	RDF120m	Radial distribution function - 120 / weighted by relative mass	3D

**Table 3.** Correlation Matrix of Descriptors

	pIC <sub>50</sub>	ZMIC2	RDF120m
PIC50	1		
ZMIC2	0.771824	1	
RDF120m	-0.45902	0.147131	1



**Figure 2.** 2D chemical structures of the designed ligands and standard Doxorubicin ligand (**H-1**: (*E*)-3-(4-chloro-2-cyclopropylphenyl)-1-(1-(2-(pyrrolidin-1-yl)ethyl)-1*H*-benzo[*d*]imidazol-2-yl)prop-2-en-1-one, **H-2**: (*E*)-1-(4-(3-chlorocyclopenta-1,3-dien-1-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1*H*-benzo[*d*]imidazol-2-yl)-3-(4-chlorophenyl)prop-2-en-1-one, **H-3**: (*Z*)-3-(4-chlorophenyl)-3-fluoro-1-(7-fluoro-4-nitro-1-(2-(pyrrolidin-1-yl)ethyl)-1*H*-benzo[*d*]imidazol-2-yl)prop-2-en-1-one)

#### Molecular Docking Simulations of the designed compounds

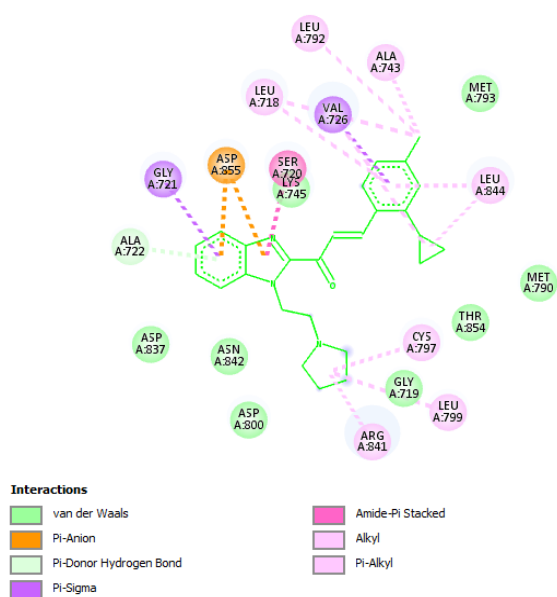
The binding affinities ( $\Delta G$ ) of the novel ligands and Doxorubicin against the active sites of the EGFR are presented in Table 4. Their 2D diagrams of interactions with the protein target are displayed in Figure 3.

#### In-silico Drug-likeness and ADMET Assay

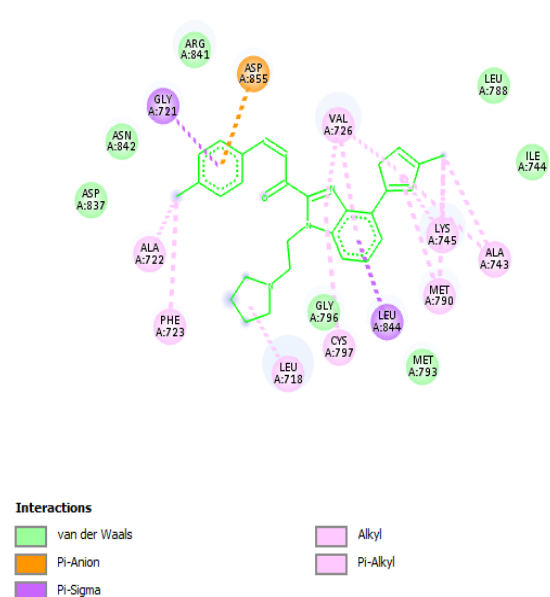
The key physicochemical descriptors of the novel ligands used to assess their drug-likeness and ADMET profiles are presented in Tables 5 and 6, respectively.

**Table 4.** Predicted IC<sub>50</sub> and binding affinity values of the designed ligands

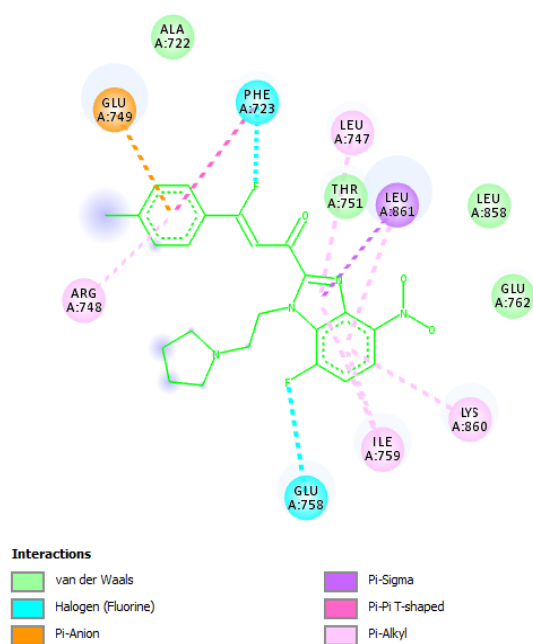
S/n	ZMIC2	RDF120m	pIC <sub>50</sub>	IC <sub>50</sub> ( $\mu$ M)	$\Delta G$ (kcal/mol)
H-1	36.1295	1.27872	-1.41608	26.07	-8.4
H-2	41.16739	4.878857	-1.63733	43.38	-8.9
H-3	38.69496	6.288769	-2.25209	178.69	-8.1
9	-	-	-	-	-7.9
13	-	-	-	-	-7.9
Doxorubicin	-	-	-	-	-7.6



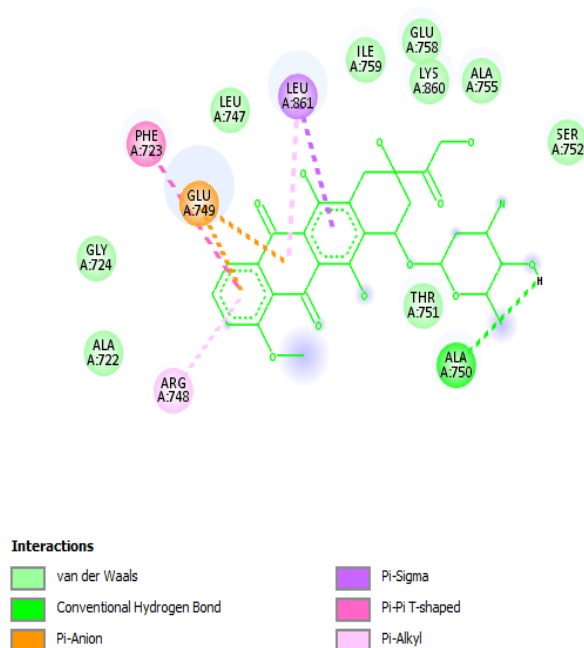
Ligand H-1 with the active sites of EGFR



Ligand H-2 with the active sites of EGFR



Ligand H-3 with the active sites of EGFR



Doxorubicin with the active sites of EGFR

**Figure 3.** 2D diagram of interactions of the designed ligands and Doxorubicin with the active sites of EGFR protease



**Table 5.** Physicochemical properties of the compounds.

Compound Rule	H-1	H-2	H-3
Lipinski's	yes	yes	yes
HBA	3	3	7
HBD	0	0	0
Mw (gmol <sup>-1</sup> )	419.95	478.41	460.86
MLogP <sub>(o/w)</sub>	3.92	4.37	3.63
Veber's	yes	yes	yes
NRB	7	7	7
TPSA (Å <sup>2</sup> )	38.13	38.13	83.95

HBA: number of hydrogen bond acceptor, HBD: number of hydrogen bond donor, Mw: molecular weight, NRB: number of the rotatable bond, TPSA: topological polar surface area

**Table 6.** Pharmacokinetics and Toxicity profiles of the Compounds

Cpd	T	M	I	R	CYP45 substrate	GIA	BBB	Pg-P substrate
H-1	No	No	low	No	No	Yes	Yes	Yes
H-2	No	No	low	low	Yes	Yes	No	Yes
H-3	No	No	low	low	Yes	Yes	No	No

T; Tumorigenic, M; Mutagenic, I; Irritant, R; Reproductive effect, GIA: gastrointestinal absorption, BBB: blood-brain barrier penetration, Pg-P: P-glycoprotein

## Discussions

### QSAR Model

Equation 1 gives the GFA-derived QSAR model, linking the dominant descriptors of the investigated benzimidazole-chalcones to their anti-cancer properties. The model's statistical parameters comply with the standard validation metrics shown in Table 1, confirming its stability, robustness, and reliability [40,41]. Likewise, the high linearity of the plot of experimental pIC<sub>50</sub> against predicted pIC<sub>50</sub> for the training set of molecules (Figure 1) further substantiates the high extrapolative power of the model.

As a way of externally validating the model, it was used to predict the pIC<sub>50</sub> of the test set of molecules that were not used during the model building. The high linearity of the plot of their experimental pIC<sub>50</sub> values against predicted pIC<sub>50</sub> values ( $R^2_{\text{pred}} = 0.81$ ), as displayed in Figure 1 implies that the model can accurately predict a new set of chemicals within its applicability

domain. Furthermore, possible biases that may arise in the model-building processes were checked via the plot of standardized residuals against the experimental pIC<sub>50</sub> values (residual plot). The dispersal of residuals on both sides of the zero line (Figure 1) implies that there is no systematic error in the course of building the model [42,43].

Multi-collinearity occurs in a model building if the descriptors are significantly inter-correlated. This scenario weakens the statistical power and stability of a model. The presence of possible multi-collinearity among the descriptors in the regression model was checked via the correlation matrix in Table 3. The low inter-correlation ( $R^2 < 0.5$ ) among the two descriptors confirms the high statistical power of the model.

### Importance of the descriptors in the Model

One of the fundamental functions of QSAR modeling is to harness the dominant descriptors of a set of molecules responsible for their



observed bioactivities. The optimum model in Equation 1 reveals the dominant influences of *ZMIC2* and *RDF120m* descriptors on the observed chemotherapeutic properties of benzimidazole chalcones against the HepG2 cancer cell line. The positive coefficient of the *ZMIC2* descriptor, as shown in Equation 1 implies that the anti-cancer properties of the studied compounds vary directly with the value of this descriptor in the molecules. Thus, for enhanced potency, the structures of the bioactive compounds ought to be optimized to have a higher value of the *ZMIC2* descriptor. Also crucial to the observed anti-tumor properties of the studied compounds is the *RDF120m* descriptor. Its negative coefficient in the model reveals the inverse relationship between this descriptor and the anti-cancer properties of the studied benzimidazole-chalcone compounds. Consequently, for enhanced anti-cancer bioactivity of the compounds, this descriptor's value has to be as low as possible.

#### Molecular docking simulation

EGFR is a target of many chemotherapeutic agents because it is over-expressed in a considerable number of human cancer cells, HCC inclusive. Here, molecular docking analysis of the designed ligands was performed to examine their binding patterns with the active sites of EGFR protease and compare them with the template molecules and the standard Doxorubicin inhibitor. The result of the docking simulation displayed in Table 4 reveals that the newly designed compounds (Figure 2) have minimum binding energy ranging from -8.1 to -8.9 kcal/mol, with the best result achieved with H-2 ( $\Delta G = -8.9$  kcal/mol). All the compounds displayed better binding efficacies compared to the template molecules ( $\Delta G = -7.9$  kcal/mol each) and Doxorubicin ( $\Delta G = -7.9$  kcal/mol).

The 2D diagrams of the interaction of the designed ligands and the standard drug within the binding pocket of EGFR shown in Figure 3

reveal that H-1 binds to the active sites of the protease through pi-sigma interactions with GLY721 and VAL726, pi-anion interactions with ASP855, alkyl interactions with CYS797, ARG841, LEU799, LEU844, LEU792, LEU718, and ALA743, a pi-alkyl interaction with LEU844, an amide-pi stacked interaction with SER720, and pi-donor hydrogen bond with ALA722. Also, the following interactions were found in the complex of H-2 with EGFR; pi-sigma with GLY721 and LEU844, alkyl with ALA722, ALA743, PHE723, LEU718, and MET790, pi-alkyl with VAL726, CYS797, MET790, LYS745, and ALA743, and a pi anion with ASP855. Likewise, H-3 interacted with the active sites of EGFR through pi-alkyl with ARG748, LEU747, LEU861, ILE759, and LYS860; a pi-anion with GLU749; halogen type with PHE 723, and GLU753; and a pi-pi T-shaped with PHE723. Furthermore, the standard ligand, Doxorubicin, binds to the active pocket of the macromolecule via pi-alkyl interactions with ARG748 and LEU861; a pi-pi T-shaped interaction with PHE723; a pi-sigma interaction with LEU861, and pi-anion interactions with GLU749. A conventional hydrogen bond stabilized these interactions with the ALA750 amino acid residue of the protease. By binding to EGFR active sites via pi-sigma, pi-anion, pi-alkyl, and pi-pi T-shaped interaction types, the standard drug exhibits similar mechanisms with the designed ligands. All the designed ligands showed better binding affinities when compared with the control (Doxorubicin). These results are similar to the findings of Al-Anazi *et al.* [44], wherein the authors carried out molecular docking investigations on a series of designed chalcone compounds against the active binding sites of EGFR, and docking scores of some of the designed ligands were found to be better than that of TAK 285 used as control.

#### Drug-likeness Assessment

Drug likeness is the quantitative description of the oral bioavailability of drugs. The evaluation of the oral bioavailability of drug candidates is sacrosanct because a majority of drugs are administered via the oral route. This examination was carried out on the designed ligands using the Lipinski's and Verber's rule. The results (Table 5) reveal that the compounds violate none of the rules and, as such, could be inferred that they possess excellent oral bioavailability.

Furthermore, poor pharmacokinetics and high toxicity account for the humongous attrition rates in drug development. Thus, the need to profile bioactive compounds concerning these parameters is a rational drug design strategy. The fates of the designed ligands in the biological system concerning their absorption, distribution, metabolism, excretion, and toxicity were assessed via *in silico* pharmacokinetic and toxicity profiling. The results (Table 6) show that they possess good gastrointestinal absorption. Also, the blood vessels that vascularize the central nervous system possess a unique feature known as the blood-brain barrier (BBB) that regulates the movement of ions, molecules, and cells between the blood and the brain [45]. The BBB protects the Central Nervous System (CNS) from toxins and xenobiotics. The results (Table 6), as it concerns BBB penetration potentials of the ligands, reveal that only H-1 ligand could penetrate the BBB, unlike H-2 and H-3. It implies that H-1 can affect the CNS if administered as a drug, unlike H-2 and H-3, which pose no significant threat to the CNS. Furthermore, P-glycoprotein (P-gp), also known as multidrug-resistant protein, is one of the proteins expressed naturally in many human body tissues that functions as a biological barrier by extruding toxins and xenobiotics from the system [46]. Substrates of P-gp have reduced serum concentrations due to the efflux action of this protein. The newly designed ligands except for H-3 (Table 6), were found to be substrates of P-gp. It implies that the efficacies of H-1 and H-2 may

be threatened if used as anti-cancer agents due to possible reduction in their serum concentrations by the efflux action of P-gp.

Also, biotransformation of drugs into inactive metabolites and their subsequent excretion through the kidney and bile is a cardinal aspect of pharmacokinetic studies. This function is performed mainly by Cytochrome P450 (CYP450) enzymes mainly found in the liver and in lesser amounts in the small intestine, lungs, placenta, and kidneys [47]. All the designed anti-cancer agents, except H-1 (Table 6), were found to be substrates of CYP450 enzymes. It implies that these ligands if used as drugs, may not pose any significant toxicity threat to the biological system as they could easily be metabolized and excreted from the body. Additionally, the *in silico* toxicity assays of the designed compounds (Table 6) reveal that they exhibit little to no tendencies to cause cancer, genetic deformity, reproductive ill-health, and irritation.

## Conclusions

Hepatocellular carcinoma is one of the leading causes of cancer-related deaths globally. Lack of potent drugs for managing this disease, especially at its advanced stage, presents a serious challenge to effectively treating this infection. In the search for novel chemotherapeutic agents against hepatocellular carcinoma, a set of benzimidazole-chalcones with established inhibitory activities against HepG2 cancer cell line was subjected to GFA-QSAR modeling to harness the dominant descriptors of the molecules responsible for their anti-cancer activities. The well-validated model ( $R^2 = 0.93$ ,  $R^2_{Adj} = 0.92$ ,  $Q^2_{L00} = 0.91$ ,  $R^2_{Pred} = 0.81$ ) reveals the predominant influences of ZMIC2 and RDF120m descriptors in the observed anti-cancer bioactivities of the compounds. With the information garnered from the QSAR model, three potent drug candidates were designed using the most active molecules in the data set as

templates. The designed molecules were subjected to molecular docking simulation against the active sites of EGFR protease. They were found to be more potent than the templates and the approved chemotherapeutic agent used herein for quality assurance. The novel ligands bind to the active sites of the target EGFR protease with binding affinity values ranging from -8.1 kcal/mol to -8.9 kcal/mol as against the -7.9 kcal/mol and -7.6 kcal/mol recorded for the templates and the Doxorubicin drug, respectively. Drug-likeness and ADMET evaluation of the designed compounds reveals that they possess good oral bioavailability and favorable pharmacokinetic and toxicity profiles. It is envisaged that the wealth of information in this study could be of immense help in the search for potent and non-toxic chemotherapeutic agents against hepatocellular carcinoma.

### Acknowledgment

The authors gratefully acknowledged the technical support of Mr. Israel Emmanuel Edache of the Department of Pure and Applied Chemistry, University of Maiduguri, Maiduguri, Nigeria.

### Disclosure statement

The authors reported no potential conflict of interest.

### Orcid

J. P. Ameji : 0000-0001-5516-8802

### References

- [1] J.M. Llovet, R.K. Kelley, V. Augusto, A.G. Singal, P. Eli, R. Sasan, L. Riccardo, K. Kazuhiko, J. Zucman-Rossi, R.S. Finn, *Nat. Rev. Dis. Primers*, **2021**, 7, 6. [[CrossRef](#)], [[Google Scholar](#)]
- [2] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, *CA: Cancer J. Clin.*, **2018**, 68, 394–424. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [3] N.E. Ibrahim, W.M. Aboulthana, R.K. Sahu, *Pharm. Biosci. J.*, **2018**, 48–55 [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [4] B. Kucukcakan, Z. Hayrulai-Musliu, *Open Access Maced J. Med. Sci.*, **2015**, 3, 363–369. [[Google Scholar](#)]
- [5] K.A. McGlynn, J.L. Petrick, W.T. London, *Clin. Liver. Dis.*, **2015**, 19, 223–238. [[Google Scholar](#)], [[Publisher](#)]
- [6] L. Kulik, H.B. El-Serag, *Gastroenterology*, **2019**, 156, 477–491. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [7] A. Villanueva, J.M. Llovet, *Gastroenterology* **2011**, 140, 1410–1426. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [8] A. Vitale, M. Peck-Radosavljevic, E.G. Giannini, E. Vibert, W. Sieghart, S. Van Poucke, T.M. Pawlik, *J. Hepatol.*, **2017**, 66, 412–423. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [9] H.A. Allam, E.E. Aly, A.K. Farouk, A.M. El Kerdawy, E. Rashwan, S.E. Abbass, *Bioorg. Chem.*, **2020**, 98, 103726. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [10] S. Montoya, D. Soong, N. Nguyen, M. Affer, S.P. Munamarty, J. Taylor, *Biomedicines*, **2021**, 9, 1591. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [11] T. Al-Warhi, M.A. El Kerdawy, A.M. Said, A. Albohy, M.Z. Elsayed, N. Aljaeed, B.E. Elkaeed, M.W. Eldehna, A.H. Abdel-Aziz, A.M. Abdelmoaz, *Drug Des. Devel. Ther.*, **2022**, 16, 1457–1471 [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [12] A. Ayati, S. Moghimi, S. Salarinejad, M. Safavi, B. Pouramiri, A. Foroumadi, *Bioorg. Chem.*, **2020**, 99, 103811. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [13] G.S. Cockerill, K.E. Lackey, *Curr. Top Med. Chem.*, **2002**, 2, 1001–1010. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]

- [14] A.E. Wakeling, *Curr. Opin. Pharmacol.*, **2002**, 2, 382–387. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [15] P.P. Kore, M.M. Mutha, V.R. Antre, J.R. Oswal, S.S. Kshirsagar, *Open J. Med. Chem.*, **2012**, 2, 26238. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [16] S.S. Padole, J.A. Asnani, D.R. Chaple, G.S. Katre, *GSC Biol. Pharm. Sci.*, **2022**, 19, 075–083. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [17] U. Abdulfatai, S. Uba, B.A. Umar, M.T. Ibrahim, *SN Appl. Sci.*, **2019**, 1, 499. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [18] A. Bello, A. Uzairu, M. Ibrahim, Y. Gatugel, *Sci. World J.*, **2019**, 14, 32–37. [[Google Scholar](#)], [[Publisher](#)]
- [19] A. Daina, O. Michielin, V. Zoete, *Sci. Rep.*, **2017**, 7, 42717. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [20] L.L. Ferreira, A.D. Andricopulo, *Drug Discov. Today*, **2019**, 24, 1157–1165. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [21] T.M. Ibrahim, A. Uzairu, G.A. Shallangwa, A. Ibrahim, *J. King Saud. Univ. Sci.*, **2020**, 32, 423–432 [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [22] Y.P.G. Zheng, *Epigenetic technological applications*. Academic Press, **2015**. [[Google Scholar](#)]
- [23] J. Verma, V.M. Khedkar, E.C. Coutinho, *Curr. Top. Med. Chem.*, **2010**, 10, 95–115. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [24] Y. Luo, J. Yao, L. Yang, C. Feng, W. Tang, G. Wang, J. Zuo, W. Lu, *Bioorg. Med. Chem.*, **2010**, 18, 5048–5055. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [25] M. Serafini, E. Torre, S. Aprile, E. Del Grosso, A. Gesù, A. Griglio, G. Colombo, C. Travelli, S. Paiella, A. Adamo, E. Orecchini, *J. Med. Chem.* **2020**, 63, 3047–3065. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [26] M. Tuncbilek, T. Kiper, N. Altanlar, *Eur. J. Med. Chem.* **2009**, 44, 1024–1033. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [27] S. Yadav, B. Narasimhan, *Anti-Cancer Agents Med. Chem.* **2016**, 16, 1403–1425. [[Google Scholar](#)], [[Publisher](#)]
- [28] H. Zhang, G.L.V. Damu, G. Cai, C. Zhou, *Eur. J. Med. Chem.* **2013**, 64, 329–344. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [29] C.Y. Hsieh, P.W. Ko, Y.J. Chang, M. Kapoor, Y.C. Liang, H.L. Chu, H.H. Lin, J.C. Horng, M.H. Hsu, *Molecules*, **2019**, 24, 3259. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [30] Y. Shao, L.F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S.T. Brown, *Phys. Chem. Chem. Phys.*, **2006**, 8, 3172–3191. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [31] A.B. Umar, A. Uzairu, A.G. Shallangwa, S. Uba, *Heliyon*, **2020**, 6, e03640. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [32] O.M. Idris, E.S. Abechi, A.G. Shallangwa, *Future J. Pharm. Sci.*, **2021**, 7, 167. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [33] E. Pourbasheer, R. Aalizadeh, R. Ganjali. *Arab. J. Chem.*, **2019**, 12, 2141–2149. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [34] S.N. Adawara, G.A. Shallangwa, P.A. Mamza, I. Abdulkadir, *Adv. J. Chem. A*, **2022**, 5, 118–137. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [35] R. Veerasamy, H. Rajak, A. Jain, V.C.P. Sivadasan, R.K. Agrawal, *Int. J. Drug Des. Discov.*, **2011**, 3, 511–519. [[Google Scholar](#)]
- [36] Trott O, A.J. Olson AJ, **2010**, *J. Comput. Chem.*, 31, 455–461. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [37] D.E. Arthur, N.A. Samuel, S. Ejeh, S.E. Adeniji, O. Adedirin, M. Abdullahi, *Sci. Afr.*, **2020**, 10, e00612. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [38] G.C. Noble, C.C. She, A.T. Chao, P.Y. Shi, *J. Virol.*, **2012**, 86, 438–446. [[Google Scholar](#)]
- [39] A.C. Lipinski, *Drug Discov. Today Technol.*, **2004**, 1, 337–341. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [40] A. Golbraikh, A. Tropsha, *J. Mol. Graph. Model.* **2002**, 20, 269–276. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]

- [41] K. Roy, R.N. Das, P. Ambure, R.B. Aher, *Chemometr. Intell. Lab. Syst.*, **2016**, 152, 18–33. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [42] M. Jalali-Heravi, A. Kyani, *J. Chem. Inf. Comput. Sci.*, **2004**, 4, 1328–1335. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [43] I.Y. Zakari, A. Uzairu, G.A. Shallangwa, S.E. Abechi, *Heliyon*, **2021**, 7, e05924. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [44] M. Al-Anazi, M. Khairuddean, O.B. Al-Najjar, M.M. Alidmat, M.N.S.N.N. Nik Nur Syazni Nik Kamal, M. Muhamad, M. Hariono, *Sys. Rev. Pharm.*, **2021**, 12, 809–820. [[Google Scholar](#)]
- [45] F. Konig, F. Muller, *Pharmacol. Rev.*, **2013**, 65, 944–966. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [46] A. Seelig, *Front. Oncol.*, **2020**, 10, 576559. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]
- [47] R.L. Slaughter, D.J. Edwards, *Ann. Pharmacother.*, **1995**, 29, 619–624. [[CrossRef](#)], [[Google Scholar](#)], [[Publisher](#)]

### HOW TO CITE THIS ARTICLE

John Philip Ameji \*, Ameh Ohiga Ebune, Ibraheem Wasiu Aderemi, Abdullahi Moyosore, Godwin Idah. Theoretical Investigation and Design of Novel Anti-proliferative Agents against Hepatocellular Carcinoma from Benzimidazole-Chalcone derivatives. *Adv. J. Chem. A*, **2023**, 6(2), 92-104.

DOI: [10.22034/AJCA.2023.365669.1336](https://doi.org/10.22034/AJCA.2023.365669.1336)

URL: [http://www.ajchem-a.com/article\\_164975.html](http://www.ajchem-a.com/article_164975.html)