

RESEARCH

Open Access



Computational modeling, ligand-based drug design, drug-likeness and ADMET properties studies of series of chromen-2-ones analogues as anti-cancer agents

Sagiru Hamza Abdullahi*, Adamu Uzairu, Gideon Adamu Shallangwa, Sani Uba and Abdullahi Bello Umar

Abstract

Background: In spite of the significant escalation in the depth of our conception and regulation of breast cancer over the past decades, the malady is still a serious community health challenge globally and poses a substantial tasks. Selective estrogen modulators (SERMs) such as Tamoxifen are approved for the therapy of this illness but developed drug resistance and unwanted side effects such as endometrial cancer caused by the long-term Tamoxifen chemotherapy limit their therapeutic applicability. Hence, developing new ER⁺ drugs with better therapeutic effect is strongly needed. In an attempt to overcome this challenge, this research is aimed at designing novel chromen-2-one analogues with better inhibition capacity against MCF-7 breast cancer cell line via structural modification of the reference compound and predict their activities using a developed QSAR model.

Results: Four models were developed, and the first was selected for the design as it has the highest statistical parameters such as: coefficient of determination ($R^2 = 0.950$), cross-validation coefficient ($Q_{cv}^2 = 0.912$), adjusted R^2 ($R_{adj}^2 = 0.935$), and external validation R^2 ($R_{pred}^2 = 0.7485$). Twelve (12) new novel chromen-2-one analogs were designed through structural modification of the reference compound. Their activities was predicted using the selected model, and their pIC_{50} was found to be better than that of the reference compound and standard drug (Tamoxifen) used in the research. Results of pharmacokinetic study of the designed compounds revealed that they possess drug-likeness properties as none of them violated the Lipinski's rule of five while ADMET studies confirmed designed compounds 6, 8, 11 and 12 as orally safe and non-toxic. Furthermore, molecular docking analysis was performed between these orally safe designed compounds and the active site of the ER⁺ receptor and the result showed that they have higher binding affinities than the reference compound and the standard drug used for this research.

Conclusion: Hence, designed compounds 6, 8, 11 and 12 can be used as novel ER⁺ breast cancer drug candidates after performing in vivo and in vitro studies.

Keywords: Breast cancer, QSAR, Molecular docking, ADMET studies, Chromen-2-ones, Density functional theory, Estrogen receptor

Background

The most frequently diagnosed form of cancer amongst female worldwide is breast cancer (BC) (Torre et al. 2015). In 2012, a 1.67 million estimated new cases of breast cancer were identified worldwide, and the figure is projected to rise to 1.7 million cases by 2020

*Correspondence: sagirwasai@gmail.com

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

according to a recent report (Forouzanfar et al. 2011). Estrogen receptor- α (ER- α) is one of the enormous superfamily of nuclear receptors, and overexpression of these receptors is seriously involved in at least 70% breast cancer patients (Sommer and Fuqua 2001). Estrogen receptors (ER⁺) support estrogenic actions in various significant biological progressions and play vibrant role in the discovery of therapeutic agents for the management of breast cancer (Traboulsi et al. 2017; Feitelson et al. 2015). In spite of the significant escalation in the depth of our conception and regulation of breast cancer over the past decades, the disorder is still a serious community health challenge over the world and poses a substantial tasks (Amir et al. 2010). It is a renowned fact that almost 70% of human breast cancers are hormone-dependent and ER α ⁺ (Maurer et al. 2017). Endocrine therapy is considered as a promising treatment option as it aims in the blockage of the ER transcription effect. Thus, ER α has provided an ultimate pharmaceutical target and several ER α ligands were established as antagonists against ER α positive breast cancer (Traboulsi et al. 2017). A distinctive group of ligands that serve as antagonist in breast tissue but agonist in other tissues such as bone and cardiovascular system (Maruyama et al. 2013; Bai and Gust 2009). Due to their outstanding mode of action, SERMs are still essential anti-breast cancer agents with benefits in cardiovascular system and bone density maintenance in comparison with aromatase inhibitors and pure anti-estrogen (Kaur et al. 2014; Jordan 2007). The first SERM approved for the treatment of breast cancer is Tamoxifen that has triphenyl ethylene skeleton and a basic side chain (Fig. 1a), it provides indispensable therapies for a number of patients (Jordan et al. 2014). 4-hydroxy [4-OH] Tamoxifen, which is an active metabolite of Tamoxifen (Fig. 1b), displays better binding affinity for the ER α ⁺. Several SERMs having numerous frames that

mimics Tamoxifen were produced to treat and inhibit breast cancer growth (Wang et al. 2009).

Even though SERMs have enhanced the treatment outcome for ER α ⁺ breast cancer patients, unwanted side effects limit their therapeutic applicability relentlessly. For example, long-term Tamoxifen chemotherapy intensifies the occurrence of endometrial cancer due to their partial estrogenic activity on the endometrium (Chen et al. 2014). Another frequent deficiency that restricts their use is inherent and developed drug resistance, in which breast tumors become refractory to endocrine therapies and relapse (Garcia-Becerra et al. 2012). Hence, developing new ER drugs with better therapeutic effect is still needed.

Coumarine analogs are imperative class of pharmacologically active skeletons that possess a wide biological activities such as anti-cancer (Luo et al. 2017a, b, c), anti-HIV (Olomola et al. 2013), anti-microbial (Arshad et al. 2011), and anti-inflammatory activities (Chen et al. 2017). Coumarin's therapeutic applicability depends on the substitution pattern, and as a result of their varied pharmacological effects in recent years they have attracted extreme attention (Bisi et al. 2017). Their anti-cancer properties were the most extensively studied among all their properties (Emami and Dadashpour 2015; Thakur et al. 2015). Anti-angiogenesis and independent induction of apoptosis are the major mechanisms of the anti-cancerous properties of Coumarin's as revealed by several studies (Sinha et al. 2016). Computational methods of drug discovery were established to accelerate the drug discovery process, as it reduces the time, resources and facilitates the assessment of properties of new compounds such as effectiveness and poisonousness prior to their synthesis. A mathematical frameworks that interfaces the quantitative relationship between the activities of a compounds and their molecular structures in an equation format is referred to Quantitative Structure–Activity Relationships

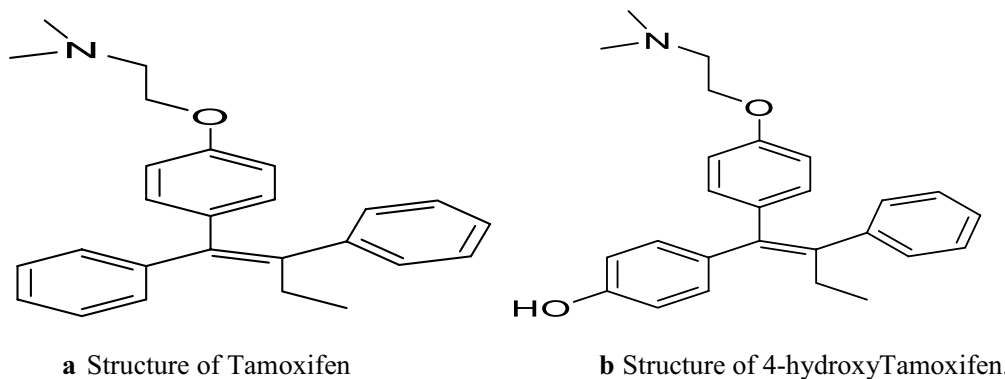


Fig. 1 **a** Structure of Tamoxifen, **b** structure of 4-hydroxyTamoxifen

(QSAR) (Abdullahi et al. 2021). Safety and efficacy of a drug to the body system are the major causes that results in the failure of drug candidate. Consequently, it is essential to discover effective compounds with better ADMET and drug-likeness properties (Abdullahi et al. 2022a, b).

This research is aimed at developing a robust QSAR models for the prediction of the anti-breast cancer activities of novel compounds with chromen-2-ones scaffold against MCF-7 cell line, design new compounds, perform ADMET and drug-likeness predictions and lastly perform molecular docking between the newly designed compounds and the active site of the ER⁺ receptor on the designed compounds to evaluate their drug-likeness characteristics.

Methods

Softwares and online tools employed for this study

This research was performed on HP laptop furnished with a dual-core Intel (R) PENTIUM (R) B940 CPU processor running at 2.0 GHz and 4.0 GB of RAM running on Windows 8. The following softwares were used in this research: Chemdraw 19.1, SPARTAN "14 v 1.1.0, Molegro Virtual Docker (MVD) and Discovery Studio. SwissADME and pkCSM online web tools are utilized in assessing the pharmacokinetics and ADMET properties of the molecules.

Chromen-2-ones data set retrieval and activity normalization

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

Drawing of two-dimensional structures of the compounds and geometry optimization

Two-dimensional structures of the chromen-2-ones were drawn with the aid of PerkinElmer ChemDraw software utilizing ACS-1996 document settings and then converted to three-dimensional format using Spartan 14.0 software program. Geometry optimization of the analogs was performed on Spartan 14.0 interface by using density functional theory calculations with B3LYP/6-31G* basis set. 2D structures of all the chromen-2-ones analogues are presented in Table 1.

Molecular descriptors generation

The optimized structures in sdf format were imported to pharmaceutical data exploratory (PADEL) software package to compute the molecular descriptors that are responsible for the biological activities of the compounds (Yap 2011).

Table 1 2D structures of the Chromen-2-one derivatives and their experimental pIC₅₀

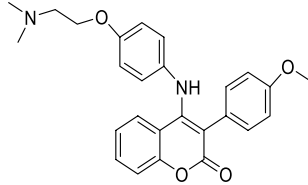
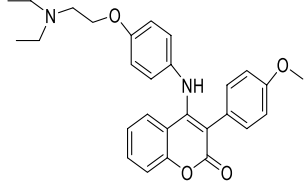
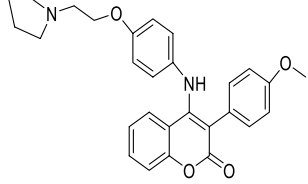
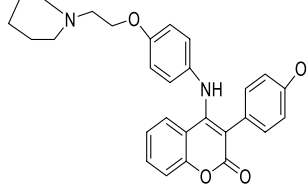
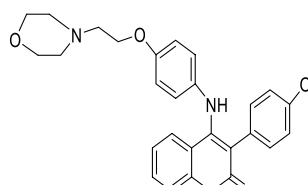
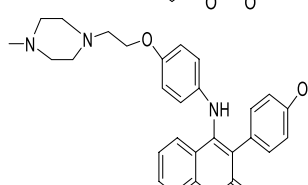
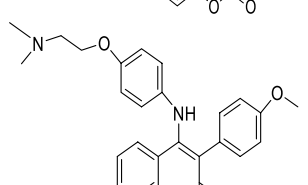
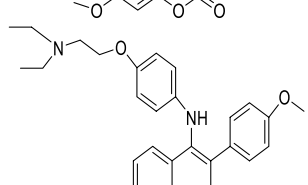
S/no	STRUCTURE	Exp pIC ₅₀
1 [†]		5.06
2		5.12
3		5.20
4		5.23
5		4.32
6		4.806
7		5.08
8		4.99

Table 1 (continued)

S/no	STRUCTURE	Exp pIC ₅₀
9		5.34
10		4.91
11		4.55
12 ^T		4.56
13		4.69
14		5.02
15		5.07
16 ^T		5.14

Table 1 (continued)

S/no	STRUCTURE	Exp pIC ₅₀
17		4.75
18		4.49
19		4.58
20 ^T		4.58
21		4.95
22 ^T		4.50
23 ^T		4.58
24		4.67

Data pretreatment and division

The result obtained from PADEL in excel worksheet was pretreated using Kennard-Stone data pretreatment software to eliminate redundant and non-relevant descriptors. Data division software was then utilized to partition

Table 1 (continued)

S/no	STRUCTURE	Exp pIC ₅₀
25		4.90
26		4.71
27 ^T		4.76
28		4.86
29 ^T		5.15
30		4.75
31		4.81
32		4.84
33 ^T		5.02

the data set into training and test set. This data set was separated into 73% training set and 27% test set. This partitioning confirms that an interrelated principle can be utilized to estimate the biological activities of the test set (Kennard and Stone 1969).

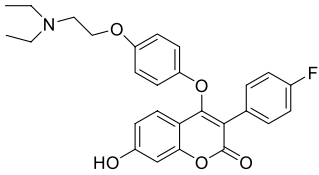
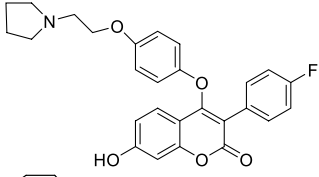
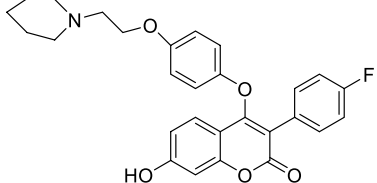
Table 1 (continued)

S/no	STRUCTURE	Exp pIC ₅₀
34		4.82
35		4.78
36		4.84
37 ^T		5.03
38		4.38
39 ^T		4.62
40 ^T		4.71
41		4.78

QSAR model building

Regression study was achieved through Genetic Function Algorithm (GFA) in material studio 8.0 software package in which the biological activities (pIC₅₀) are the

Table 1 (continued)

S/no	STRUCTURE	Exp pIC ₅₀
42		4.67
43		4.78
44		5.02

† Test set

dependent variable and the physicochemical properties (descriptors) are the independent variables (Khaled 2011). The length of the regression equation was set to 6, and population and maximum generation were set to 1000 and 1500, respectively. The number of top regression models developed was set as 4. Mutation probability was 0.1, and the user defined smoothing parameter was 0.5

Internal validation of the developed QSAR model

The models generated were evaluated by means of Friedman's Lack of Fit (LOF) which is used to measure the capability of a model. The revised formula for the Friedman's lack of fit is represented by Eq. 2 below:

$$\text{LOF} = \frac{\text{SEE}}{\left(1 - \frac{a+bq}{N}\right)^2} \quad (2)$$

SEE is the standard error of estimation, q is the total number of physicochemical parameters (descriptors) in the model, b is a user-defined smoothing parameter, a is the number of terms in the model, and N is the number of molecules in the modeling (training) set. SEE is the standard error of estimation which is the same as the standard deviation of the model. A good model has lower SEE value. Equation 3 is used to compute SEE values of a model

$$\text{SEE} = \frac{(Y_{\text{exp}} - Y_{\text{pred}})^{\frac{1}{2}}}{(N - q - 1)^{\frac{1}{2}}} \quad (3)$$

Structure of a typical regression equation is of the form:

$$A = b_1\delta_1 + b_2\delta_2 + b_3\delta_3 + \dots + c \quad (4)$$

where A is the biological activity (pIC_{50}), b 's represents the coefficient of regression for the corresponding δ 's which are the independent variables that represents the molecular descriptors of a compound and c is the mathematical constant of regression.

The most frequently used internal assessment parameter for the QSAR model is correlation coefficient of the training set (R^2). It provides explanation of the fragment of the total deviation of the model. The more closer the R^2 value is to unity the better the model generated. R^2 is calculated using Eq. (5) below:

$$R^2 = 1 - \frac{\sum(Y_{\text{exp}} - Y_{\text{pred}})^2}{\sum(Y_{\text{exp}} - Y_{\text{mtraining}})^2} \quad (5)$$

where Y_{exp} , Y_{pred} , and $Y_{\text{mtraining}}$ are the actual experimental activity, the predicted activity and the mean experimental activity of the training set.

Adjusted R^2 (R_{adj}^2) value varies directly with a rise in the number of molecular descriptors; value of R^2 alone is insufficient for assessing the stability of a model. R^2 is adjusted to get a stable and reliable model. The adjusted R^2 is defined by Eq. (6) below:

$$R_{\text{adj}}^2 = 1 - \left(1 - R^2\right) \frac{N - 1}{N - P - 1} = \frac{(N - 1)R^2 - q}{N - P + 1} \quad (6)$$

where N is the number of samples in the training set, q =number of descriptors in the model (Abdullahi et al. 2021). The ability of a QSAR model to predict the activity of new molecules is determined using the cross-validation coefficient (Q_{cv}^2), and it is calculated using Eq. (7) below:

$$Q_{\text{cv}}^2 = 1 - \frac{\sum(Y_{\text{pred}} - Y_{\text{exp}})^2}{\sum(Y_{\text{exp}} - Y_{\text{mtraining}})^2} \quad (7)$$

External validation of a QSAR model

The external validation of the built QSAR models is evaluated based on the value of R_{test}^2 as defined in Eq. (8):

$$R_{\text{test}}^2 = 1 - \frac{\sum(Y_{\text{pred}} - Y_{\text{exp}})^2}{\sum(Y_{\text{exp}} - Y_{\text{mtraining}})^2} \quad (8)$$

Applicability domain

The determination of influential and outlier compounds used to develop a QSAR model is performed by studying its applicability domain (AD). The robustness and reliability of a model can as well be affirmed using the domain (Tropsha et al. 2003). Technique used to evaluate the AD of a QSAR model is the leverage approach, for a molecule the leverage h_i is defined by Eq. 9:

$$h_i = y_i \left(Y^T Y \right)^{-1} y_i^T \quad (9)$$

where y indicates the vector descriptor of the referred sample and Y signifies the matrix of the descriptor obtained from the training set descriptor values. The threshold leverage (h^*) was computed using Eq. (10) below:

$$h^* = \frac{3(Q + 1)}{N} \quad (10)$$

N is the number of training set data and Q is the number of independent variables (descriptors) used in building the model. William's plot is a plot of standardized residual values against the leverage values of molecules. It facilitates the viewing of AD of a QSAR model. When a leverage of a compound exceeds the threshold value (h^*), it is alleged to have influence the performance of the model and the compound may be eliminated from the domain, compounds having residual values within ± 3 regions are not tagged as outliers since points lying within this region cover 99% of the normally distributed data (Abdullahi et al. 2022a, b). Hence, the leverage together with standardized residuals was jointly used to characterize and determine the applicability domain.

Quality assurance of the model

Internal and external validations parameters are used to assess the reliability and predictive ability of a QSAR model. Table 2 gives the general minimum requirement values for the assessment of a QSAR model.

ADMET and drug-likeness studies

Available online web sites such as SwissADME (<http://www.swissadme.ch/index.php>) and pkCSM (<http://structure.bioc.cam.ac.uk/pkcsml>) are applied to explore the drug-likeness and ADMET properties of the studied molecules. These sites allow researchers to discover an innovative drug candidate, to minimize the number of empirical experimentations and to upgrade the success rate (Abdullahi et al. 2021). Lipinski's rule of five (ROF) was employed as the principal screening step for the drug-likeness properties trailed by computing the central ADMET properties which are measures of the pharmacokinetics of the molecules under research.

Molecular docking studies

Ligand–protein docking studies was performed to study the nature of binding interactions between the newly designed chromen-2-one analogs and the binding pocket of the ER⁺ receptor and to gain insight into the amino acid residues that are responsible for the ligand–protein interaction (Abdullahi et al. 2022a, b). The simulation studies was carried out on HP workstation furnished with a dual-core Intel (R) PENTIUM (R) B940 CPU processor running at 2.0 GHz and 4.0 GB of RAM running on Windows 8. X-ray crystallized structure of the ER⁺ protein was downloaded from protein data bank (pdb id=3ERT) and was prepared using Molegro virtual docker by the elimination of excess water molecules and co-crystallized ligand enveloped in its crystal structure. Residues with structural errors were repaired and rebuilt. The active site of the 3ERT receptor was predicted and set inside a restricted sphere having X, Y, Z coordinates of 24.12, 3.48 and 20.11 Å, respectively. Ligand preparation was performed by optimization using DFT calculations with B3LYP/631G* basis set and then saved in pdb format. The docking algorithm were scored based on MolDock and Rerank scoring functions. Molegro virtual docker was utilized for the docking studies due to its ability to offer better and accurate results in comparison with other docking softwares. Discovery studio software was utilized to view the various ligand/protein interactions in the docked complexes.

Table 2 Minimum recommended values for acceptable QSAR model

Symbol	Name	Threshold limit	Model 1	Model 2	Model 3	Model 4
LOF	Friedman's lack of fit	Low value	0.016488	0.01648900	0.01683500	0.01705300
R^2	Co-efficient of determination of the training set	≥ 0.6	0.94794	0.94793800	0.94684500	0.94615500
R_{adj}^2	Adjusted R-squared	≥ 0.6	0.935	0.93544300	0.93408800	0.93323200
Q^2	Cross-validation coefficient	≥ 0.5	0.912	0.91105100	0.90975700	0.90712400
$R^2 - Q^2$	Difference between R^2 and Q^2	< 0.3	0.012494	0.036887	0.037088	0.039031

Results

Four QSAR models were developed from the training set data using genetic function algorithm (GFA) coupled with multi linear regression (MLR), and their expressions are presented below:

Model 1

$$Y = 0.342327907 * \text{apol} + 0.002006877 * \text{ATSC8m} + 0.021947183 * \text{ATSC7s} - 2.110146447 * \text{SM1_Dzm} - 0.027702443 * \text{SpAbs_Dzs} + 0.122940438 * \text{ZMIC4} - 9.882891756.$$

Model 2

$$Y = 0.333966562 * \text{apol} + 0.001909583 * \text{ATSC8m} + 0.019049122 * \text{ATSC7s} - 2.079324191 * \text{SM1_DzZ} - 0.027112784 * \text{SpAbs_Dzs} + 0.119956742 * \text{ZMIC4} - 9.456381109.$$

Model 3

$$Y = 0.342932868 * \text{apol} + 0.002004154 * \text{ATSC8m} + 0.021734174 * \text{ATSC7s} - 2.067273713 * \text{SM1_Dzm} - 0.027821824 * \text{SpAD_Dzs} + 0.122642435 * \text{ZMIC4} - 9.889860694.$$

Model 4

$$Y = 0.334437304 * \text{apol} + 0.001906343 * \text{ATSC8m} + 0.018877487 * \text{ATSC7s} - 2.033875692 * \text{SM1_DzZ} - 0.027216796 * \text{SpAD_Dzs} + 0.119277728 * \text{ZMIC4} - 9.454115852.$$

Discussion

Internal and external validations of the developed models

Internal and external validation of the developed QSAR models was performed to reveal their robustness, stability, reliability and predictability, and the results are presented in Table 2. As observed from the table, all the models passed the minimum requirements for an acceptable QSAR model with model 1 having the best statistical parameters. Hence, it was selected for further examination in this study. Values of different types of molecular descriptors that have appeared in model 1 together with experimental predicted pIC_{50} and residuals for the training set data are presented Table 3.

Moreover, external validation of the selected model was performed to determine its predictive ability. The value of the external validation regression coefficient (R_{ext}^2) was found to be 0.745, and this value surpasses the minimum recommended value ($R_{ext}^2 \geq 0.6$). This illustrates that the selected model is capable of providing a valid predictions

of the activities of new compounds. Step by step calculation of the external prediction correlation coefficient (R_{ext}^2) is shown in Table 4.

Figure 2 represents a plot of the predicted pIC_{50} for the model building (training) and test sets against the experimental pIC_{50} values. Furthermore, the residual values of all the data sets were plotted against the experimental activities as illustrated in Fig. 3. The predicted activities of the compounds strongly agree with their corresponding experimental activities as observed from Fig. 2, and this affirms the predictive ability of the selected model. Additionally, as observed from Fig. 3 the residual values of the compounds reside on both sides of zero, and this suggests that the selected model did not demonstrate any relative and systematic error (Abdullahi et al. 2021).

William's plot

Williams's plot of the selected model is portrayed in Fig. 4. It can be observed that only six (1, 13, 19, 25, 36 and 40) compounds from the test set data are found to be beyond the defined domain of applicability, i.e., they have leverage values greater than the threshold value ($h^* = 0.656$). These compounds are termed as influential compounds, and their high leverage values might be related to their differences structurally from the other compounds in the data set.

Mean effect of the relevant molecular descriptors

The individual impact and role of the relevant descriptors is designated by their values of mean effects. Key information's on the effect of the molecular descriptors on a built QSAR model is offered by the mean effect of the descriptors. The magnitude and sign of the molecular descriptors coupled with their mean effect values illustrates their powerfulness in influencing the biological activity of a compound (Abdullahi et al. 2021). Positive mean effect value of a descriptor indicates that biological activity of a molecule rises with the raise in the value of the descriptor, while negative descriptor value suggests that the biological activity of a molecule rises with the decrease in the descriptor's value. This implies that apol, ZMIC4 and ATSC7s descriptors increase the pIC_{50} of the compounds when their values are increased, while increase in the values of SpAbs_Dzs, SM1_Dzm and ATSC8m descriptors will lessen the pIC_{50} of the compounds (Abdullahi et al. 2021).

Mean effect value of descriptors is calculated using Eq. (11) below.

$$MF_j = \frac{\alpha_j \sum_{j-1}^{i-n} \beta_{ij}}{\sum_j^m \alpha_j \sum_i^n \beta_{ij}} \quad (11)$$

Table 3 Values of different types of molecular descriptors that have appeared in model 1, experimental and predicted pIC_{50}

Name	Apol	ATSC8m	ATSC7s	SM1_Dzm	SpAbs_Dzs	ZMIC4	Exp pIC_{50}	Pred pIC_{50}	Residual values
2	76.02538	-511.394	7.008239	1.282063	387.2488	26.52987	5.12	5.098739	0.022165
3	74.69179	-387.898	6.193359	1.282063	374.857	26.21461	5.206908	5.176698	0.030211
4	76.45179	-461.806	4.403051	1.282063	392.8177	26.59491	5.232102	5.140775	0.091327
5	75.49379	-296.667	5.573471	1.531328	402.4475	26.65645	4.323764	4.384738	-0.06097
6	78.21858	-291.019	7.702415	1.424563	429.1153	26.93659	4.806041	4.896536	-0.0905
7	73.73379	-219.21	0.728265	1.531328	364.528	27.52942	5.079877	4.989133	0.090744
8	79.92096	-629.4	3.951389	1.531328	415.1461	28.48906	4.990974	5.070446	-0.07947
10	81.68096	-525.919	2.964508	1.531328	430.3128	28.13309	5.344862	5.395041	-0.05018
11	82.11417	-369.884	7.882883	1.673828	463.1924	28.97157	4.906929	4.855971	0.050958
12	66.74462	-322.919	20.87111	1.282063	315.6919	26.03773	4.547447	4.525972	0.021475
13	71.5982	-637.728	22.96139	1.282063	351.8555	26.80669	4.68508	4.694322	-0.00924
14	70.6402	278.127	-10.5874	1.531328	346.1908	25.91132	5.020907	4.988915	0.031992
15	75.49379	-36.6476	-7.9154	1.531328	382.9678	26.65645	5.074688	5.150156	-0.07547
17	79.02058	-24.4477	-1.44281	1.673828	440.2504	27.35784	4.754734	4.722681	0.032053
18	65.83282	-106.726	33.4814	1.400586	322.3998	26.51099	4.488651	4.546728	-0.05808
19	72.02	-573.958	36.53421	1.400586	370.4859	27.55149	4.578725	4.58991	-0.01119
21	73.78	-464.979	35.60511	1.400586	384.7591	27.18907	4.946154	4.950763	-0.00461
24	74.582	-589.541	31.62424	1.649852	378.6411	27.83961	4.666956	4.611436	0.05552
25	76.342	-483.953	37.0656	1.649852	403.2157	28.20248	4.900319	4.909093	-0.00877
26	66.63482	262.9249	16.48696	1.649852	323.6971	26.99087	4.714668	4.687209	0.027458
28	71.48841	-81.1025	19.51191	1.649852	359.6619	27.71358	4.855737	4.817233	0.038504
30	71.43541	-225.952	1.505883	1.638094	335.6486	28.43851	4.748605	4.892375	-0.14377
31	77.62258	-654.296	5.409569	1.638094	393.3604	29.35224	4.813892	4.750034	0.063857
32	76.289	-549.968	4.533031	1.638094	381.4404	29.09015	4.842241	4.781636	0.060605
34	67.43003	-105.825	37.02106	1.756617	315.2689	27.96859	4.816445	4.798453	0.017992
35	73.6172	-585.876	41.43313	1.756617	364.8873	28.90328	4.778847	4.790285	-0.01144
36	72.28362	-461.299	40.51983	1.756617	353.5163	28.64644	4.836839	4.847156	-0.01032
38	65.24824	-10.4331	11.79608	1.638094	297.1394	27.62792	4.38321	4.399839	-0.01663
41	64.33645	259.7639	27.13186	1.756617	296.974	27.85831	4.77963	4.749339	0.030291
42	70.52362	-255.052	31.53991	1.756617	345.6058	28.80669	4.671009	4.700326	-0.02932
43	69.19003	-115.882	30.64934	1.756617	334.3837	28.55617	4.775208	4.783637	-0.00843
44	72.28362	-155.826	31.37194	1.756617	361.2802	28.42581	5.020452	5.017228	0.003223

where MF_j is the mean effect of molecular descriptor j in a model, α_j denotes the coefficient of the descriptor J in the model and β_{ij} is the value of the descriptor in the data matrix for each compound in the training set, m demonstrates the number of descriptors found in the model and n is the number of samples in the training set (Abdullahi et al. 2022a, b). Mean effect values of the relevant descriptors that appeared in the selected model are placed in Table 5.

Ligand-based drug designation

Ligand-based drug design of new novel chromen-2-ones was achieved through virtual screening technique based on the chosen QSAR model. Compound 10 from the training set data was utilized as the reference compound

for the design as it has the highest pIC_{50} (5.344) values and was excellently predicted by the selected model with low residual value (-0.051), which is within the defined domain of applicability of the model. Twelve (12) new novel chromen-2-one analogs were designed through structural adjustment of the reference compound, and their 3D structures were geometrically optimized on Spartan 14.0 interface using DFT calculations with B3LYP/6-31G* basis set. Their pIC_{50} was predicted using the selected model, and they were found to have improved pIC_{50} which ranges from (5.472–8.584) compared to the reference compound and Tamoxifen ($pIC_{50}=4.843$) used in the research. The structure of the reference compound and the template used for the design is shown in Figs. 5 and 6; also, the structure of the

Table 4 Calculations of the external prediction correlation coefficient (R_{ext}^2) of the test set data

Name	Apol	ATSC8m	ATSC7s	SM1_Dzm	SpAbs_Dzs	ZMIC4	Y_{exp}	Y_{pred}	$(Y_{exp} - Y_{pred})$	$(Y_{exp} - Y_{mtrn})$	$(Y_{exp} - Y_{pred})^2$	$(Y_{exp} - Y_{mtrn})^2$
1	69.8382	-85.778	4.200617	1.282063	337.8195	25.44033	5.057	5.009	0.048895	0.222726	0.002391	0.049607
12	72.93179	-768.015	23.75772	1.282063	363.9795	27.10589	4.560	4.607	-0.0474	-0.27442	0.002246	0.075305
16	78.58738	-55.6998	-7.88985	1.531328	410.3691	26.61358	5.136	5.407	-0.27107	0.301313	0.073478	0.090789
20	70.68641	-437.171	35.74556	1.400586	358.2108	27.26468	4.583	4.695	-0.11219	-0.25158	0.012587	0.063291
22	69.72841	-264.142	28.89506	1.649852	341.7769	27.13952	4.500	4.478	0.022508	-0.33404	0.000507	0.111586
23	74.582	-584.449	37.55379	1.649852	387.9617	28.57307	4.579	4.584	-0.00388	-0.25489	1.51E-05	0.064969
27	72.822	-221.898	20.36558	1.649852	372.1313	27.99117	4.758	4.699	0.059326	-0.07681	0.00352	0.0059
29	74.582	-114.504	19.85069	1.649852	386.4599	27.625	5.151	5.063	0.087798	0.316425	0.007709	0.100125
33	79.38258	-558.235	4.732507	1.638094	402.6387	28.98203	5.017	5.228	-0.21109	0.182055	0.044557	0.033144
37	75.3772	-488.134	41.10432	1.756617	380.9881	28.52711	5.031	5.089	-0.05886	0.195814	0.003465	0.038343
39	70.10182	-368.441	15.09024	1.638094	334.5215	28.33253	4.622	4.466	0.155737	-0.2128	0.024254	0.045286
40	73.19541	-399.582	15.59836	1.638094	361.4098	28.21521	4.706	4.715	-0.00863	-0.1288	7.45E-05	0.016588
										Sum	0.174802	0.694934

$$R_{test}^2 = 1 - \frac{\sum(Y_{pred} - Y_{exp})^2}{\sum(Y_{exp} - Y_{mtrn})^2} = 0.174802 / 0.694934$$

$$R_{pred}^2 = 0.7485$$

where Y_{exp} and Y_{pred} are the experimental and predicted pIC_{50} of the test set compounds

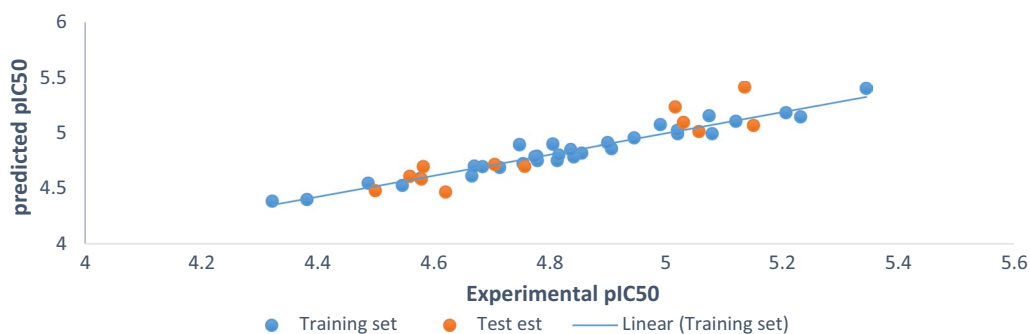


Fig. 2 Plot of predicted against experimental $P_{IC_{50}}$ of the training and test set

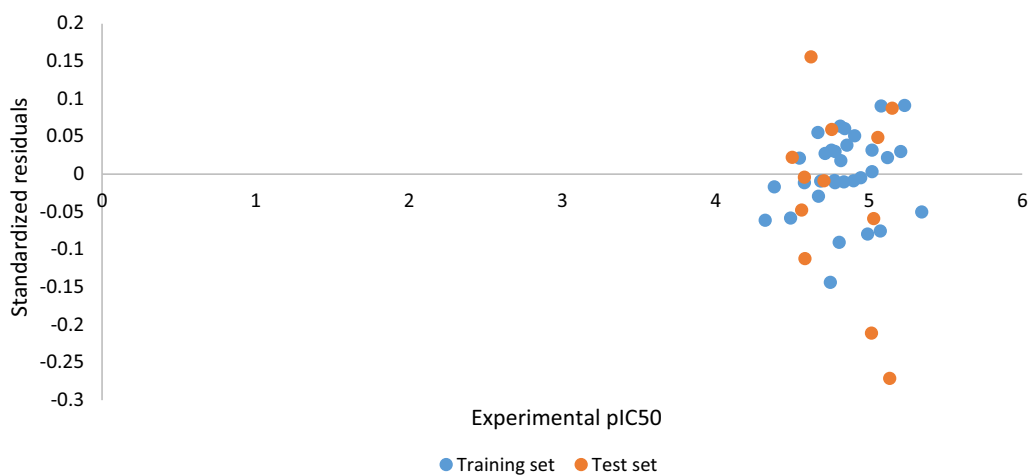


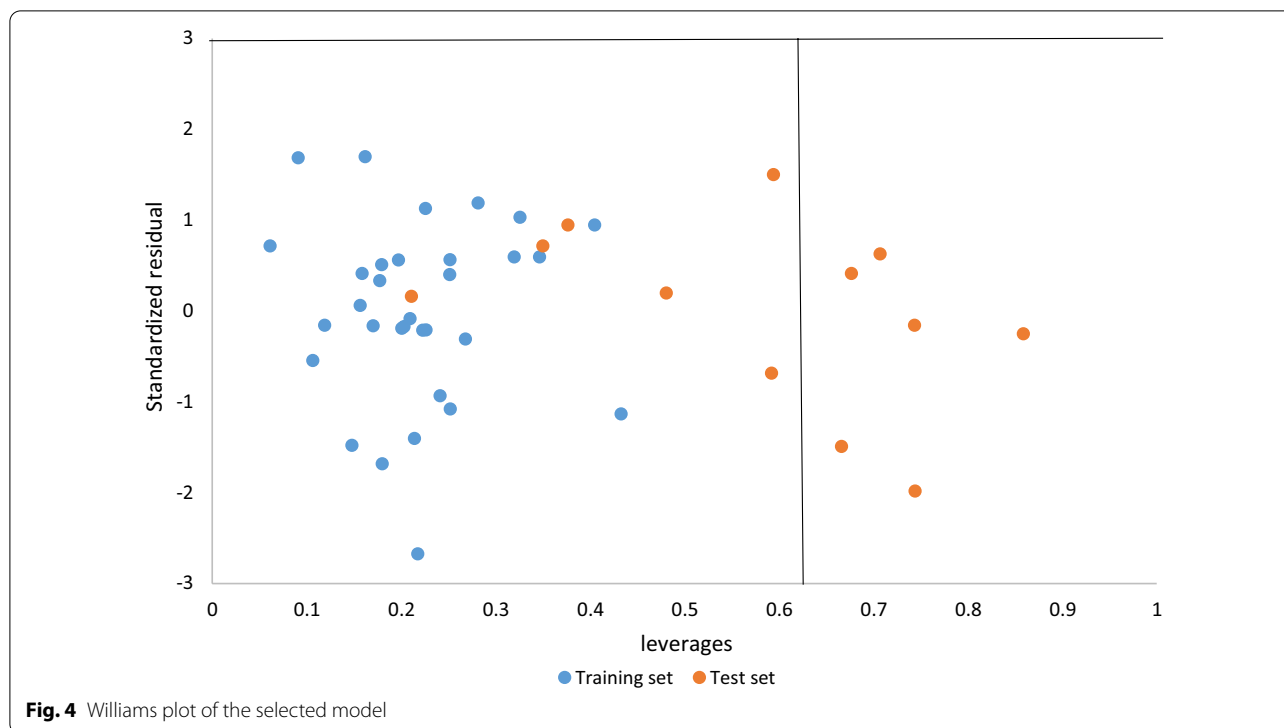
Fig. 3 Plot of standardized residuals against the experimental activities of the training and test set

newly designed compounds with their predicted pIC_{50} is shown in Table 6.

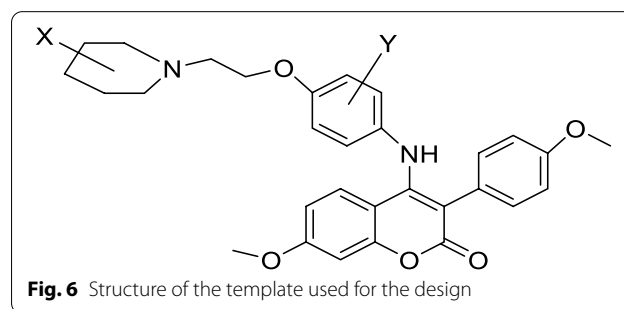
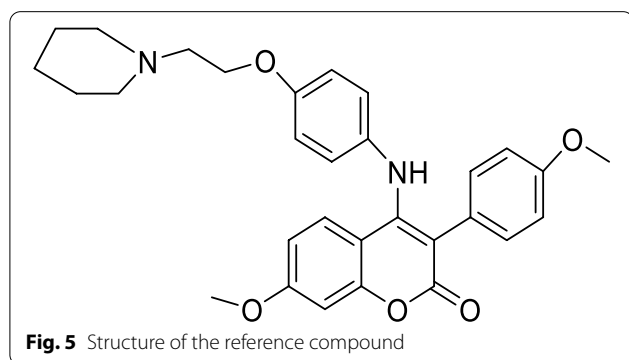
Drug-likeness properties of the designed compounds

The standards used in the screening of drug candidates at the early phase of the drug discovery process are the assessment of their drug-likeness parameters. This is accomplished by correlating the physicochemical properties of a given molecule with its bio-pharmaceutical properties in human body, mostly, its impact on oral bioavailability (Bickerton et al. 2012). To affirm that the designed chromen-2-ones analogues are the feasible drugs, their ADMET and pharmacokinetic properties were evaluated using SwissADME. The most inventive and detailed analysis of drug-likeness properties was performed by Lipinski (Lipinski et al. 1997); it results to the prevalent “rule of five,” which propose that a molecule possess drug-likeness properties only when its molecular weight (mol. wt.) < 500, its number of hydrogen bond donors (HBD) < 5, its

number of hydrogen bond acceptors (HBA) < 10, and its partition coefficient octanol/ water $\log P$ < 5. Compounds that do not violate more than two (2) of the criteria are deemed to possess drug likeness properties. Results of the drug-likeness properties of the designed compounds are presented in Table 7. All the designed chromen-2-ones possess drug-likeness properties since they violated only one of the Lipinski’s rule of five criteria (Molecular weight > 500). They have optimum profile of permeability and bioavailability as indicated by their bioavailability score of 0.55 (Martin 2005). Furthermore, the synthetic accessibility values of the designed chromen-2-ones were evaluated, based on a scale ranging from 1 (easy to synthesize) and 10 (not easily synthesize). Their predicted synthetic accessibility values range from 4.51 to 5.24 (Table 7), and these values suggested that the designed compounds can be easily synthesized.

**Table 5** Mean effect values of the most relevant descriptors

Descriptor	Definition	Class	Mean effect
apol	Sum of the atomic polarizabilities	2D	1.703
ATSC8m	Centered Broto-Moreau autocorrelation—lag 8/weighted by mass	2D	-0.039
ATSC7s	Centered Broto-Moreau autocorrelation—lag 7/weighted by I-state	2D	0.025
SM1_Dzm	Spectral moment of order 1 from Barysz matrix/weighted by mass	2D	-0.224
SpAbs_Dzs	Graph energy from Barysz matrix/weighted by I-state	2D	-0.695
ZMIC4	Z-modified information content index (neighborhood symmetry of 4-order)	2D	0.231



ADMET properties of the designed compounds

ADMET properties of a molecule deals with its absorption, distribution, metabolism, excretion, and toxicity, in

and through the human body. These properties constitute the pharmacokinetic profile of a drug molecule and is very essential in evaluating its pharmacodynamics activities. Many online tools and offline software programs are widely available for the prediction of ADMET properties

Table 6 Structure and predicted PiC_{50} of the newly designed compounds pIC_{50}

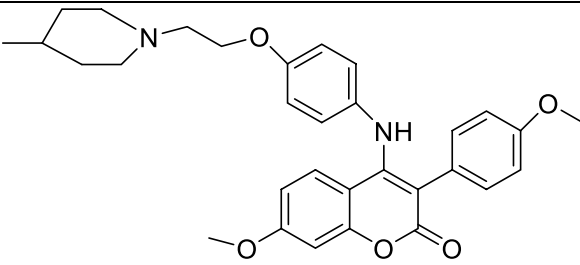
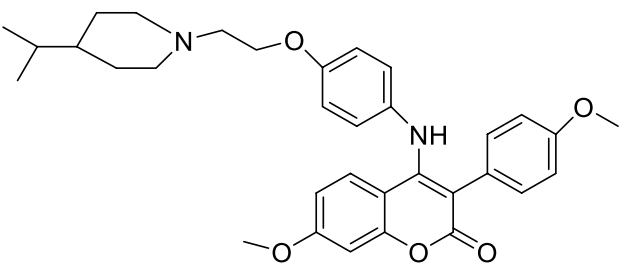
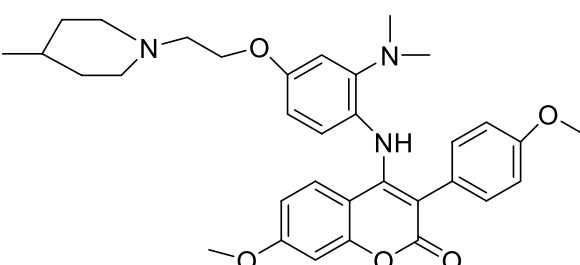
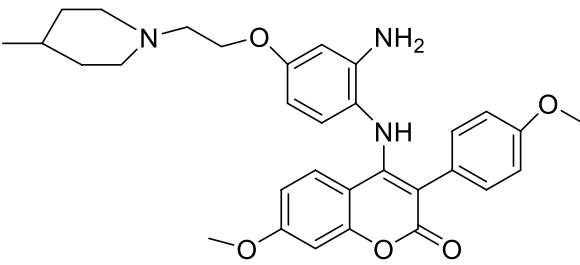
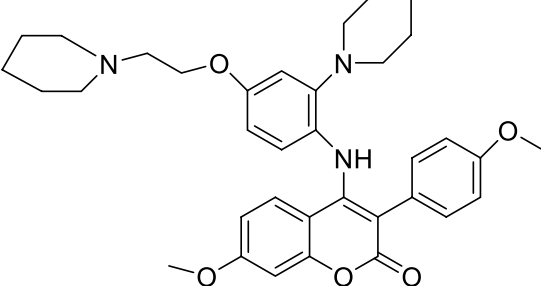
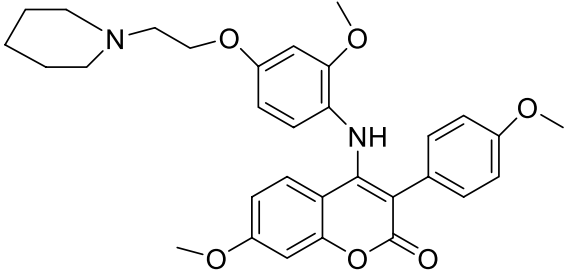
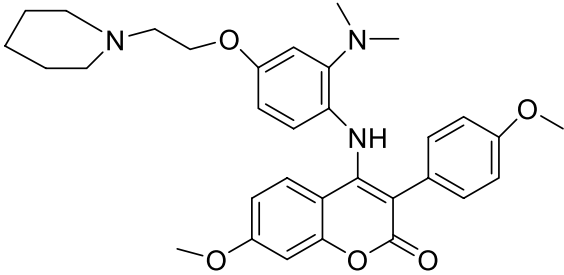
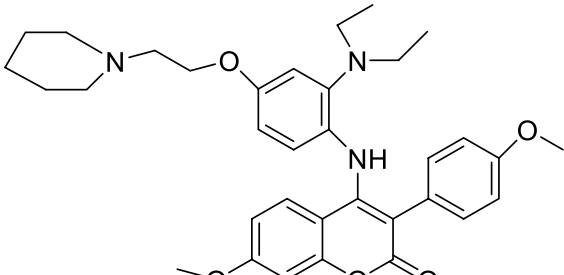
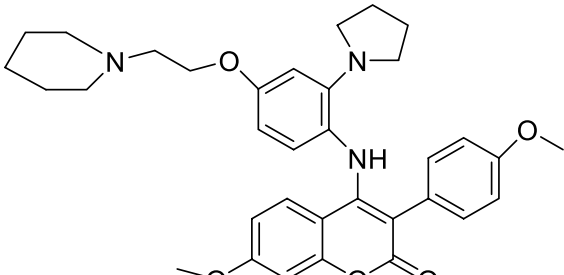
S/no	Structure	Predicted pIC_{50}
1		5.472
2		5.919
3		6.772
4		5.595
5		8.584

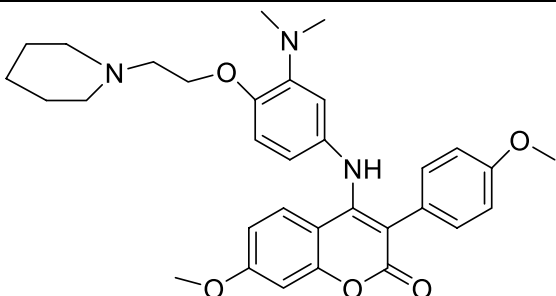
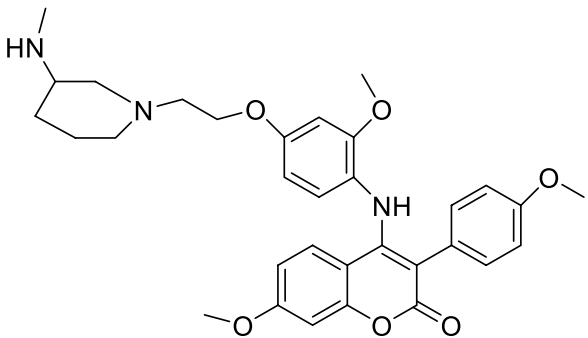
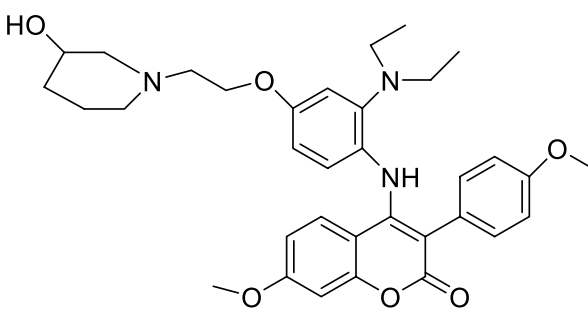
Table 6 (continued)

S/no	Structure	Predicted pIC ₅₀
6		5.987
7		6.477
8		7.958
9		8.007

of a molecule. In this study, pkCSM online server was used for this purpose. The results of predicted ADMET properties of the designed compounds are presented in Table 6. All the designed chromen-2-one analogues exhibit excellent human intestinal absorption between 92.520 to 100%, and these values exceed the minimum recommended percentage of absorption (30%), as such

they are well absorbed by the human intestine. The allowed range of Blood–brain barrier (BBB) and the central nervous system (CNS) permeability is >0.3 to <-1 log BB and >-2 to <-3 log PS; thus, the designed compounds have a high possibility of crossing the blood–brain barrier (BBB) and central nervous system (CNS) as their log BB and log PS are within the acceptable range

Table 6 (continued)

S/no	Structure	Predicted pIC_{50}
10		7.430
11		5.587
12		6.902
Tamoxifen		4.843

(Umar et al. 2019). The biotransformation of a drug in a body is demonstrated by its enzymatic metabolism; hence, it is very essential to consider the drug's metabolism. A category of super enzymes, cytochrome P450, plays a critical role in drug's metabolism. CYP families accountable for the drug's metabolism include 1A2, 2C9, 2C19, 2D6, and 3A4, among which the most essential is the 3A4 enzyme. All the designed chromen-2-ones are the substrate as well as inhibitors of the 3A4 enzyme. The relationship between the elimination rate of a drug and its concentration is explained by a parameter called total clearance; all the designed analogues possess high values of this parameter which are within the acceptable range

of a drug candidate in a human body. Additionally, it is necessary to evaluate the toxicity and adverse side effects of drug candidate at the preclinical and clinical phase as the safety of the drug one the most important issue. Results from Table 8 show that compounds 6, 8, 11 and 12 are non-toxic. These compounds exhibit promising pharmacokinetics and ADMET properties; thus, they can be recommended as ER⁺ inhibitors and breast cancer drug candidates for further analysis.

Moreover, to further affirm drug-likeness properties of the orally safe compounds (6, 8, 11 and 12), their bioavailability radar and boiled egg plot were analyzed. The Bioavailability Radar allows a first glance at

Table 7 Drug-likeness properties of the designed chromen-2-ones

S/nos.	MW	HBA	HBD	mlogP	SA	ABS	Lipinski violation	Drug likeness
1.	514.61	6	1	3.49	4.83	0.55	1	Yes
2.	542.67	6	1	3.86	5.07	0.55	1	Yes
3.	557.68	6	1	3.34	5.24	0.55	1	Yes
4.	529.63	6	2	2.97	4.99	0.55	1	Yes
5.	583.72	6	1	3.7	4.96	0.55	1	Yes
6.	530.61	7	1	2.97	4.51	0.55	1	Yes
7.	543.65	6	1	3.15	4.66	0.55	1	Yes
8.	571.71	6	1	3.52	4.91	0.55	1	Yes
9.	569.69	6	1	3.52	4.83	0.55	1	Yes
10.	543.65	6	1	3.15	4.7	0.55	1	Yes
11.	559.65	8	2	2.38	5.15	0.55	1	Yes
12.	587.71	7	2	2.74	5.4	0.55	1	Yes

MW, molecular weight; HBA, hydrogen bond acceptors; HBD, hydrogen bond donors; SA, synthetic accessibility; ABS, Bioavailability Score

Table 8 Predicted ADMET properties of the designed chromen-2-ones

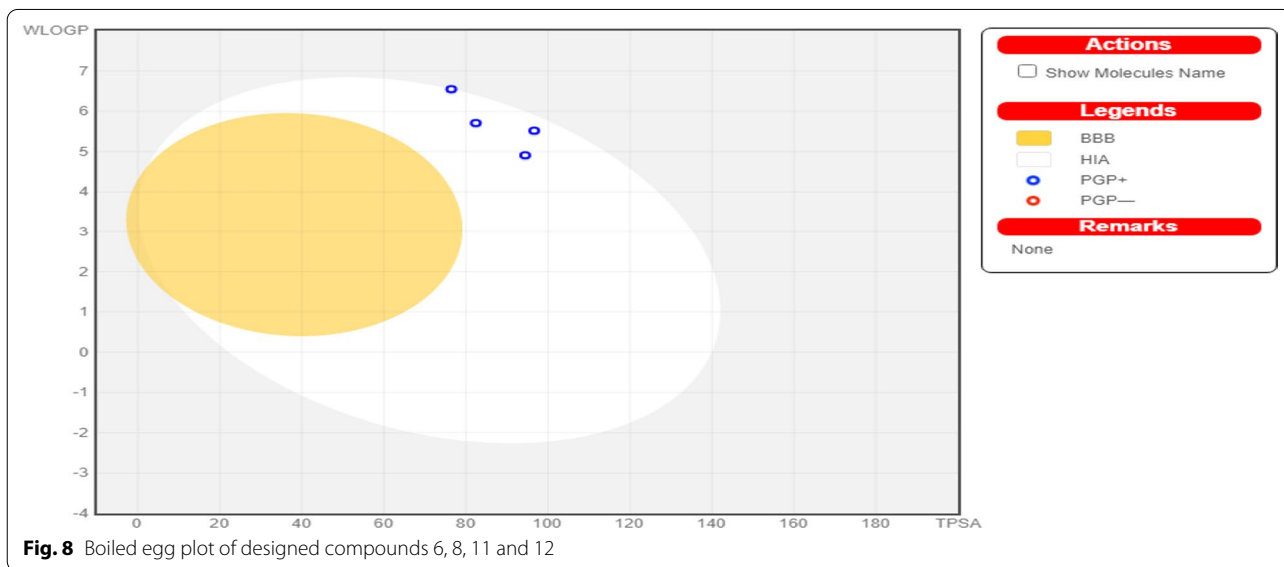
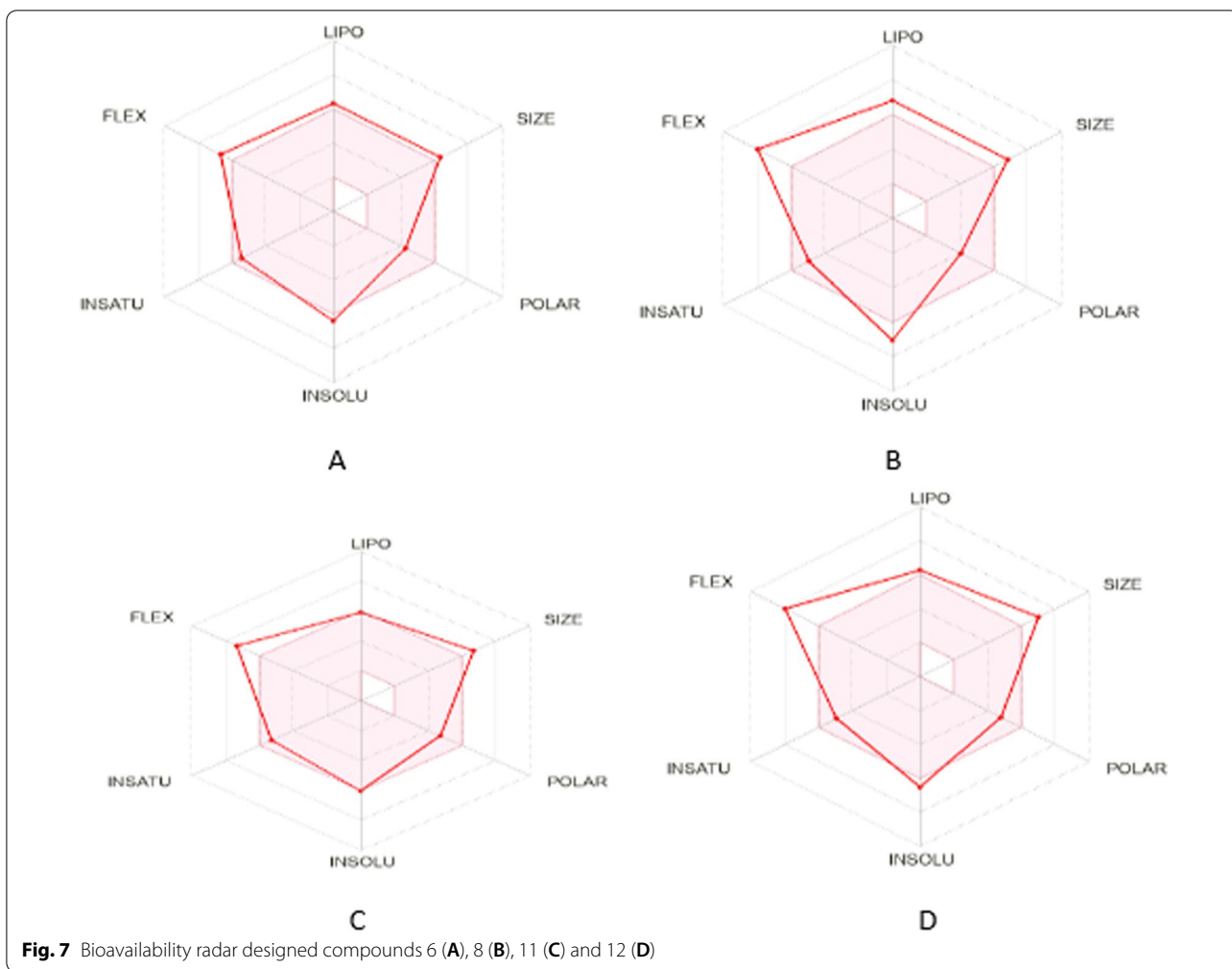
S/nos	Intestinal (human) absorption	Distribution		Metabolism						Total clearance	AMES toxicity	
		log BB	logPS	Substrate			CYP	Inhibitors				
				2D6	3A4	1A2	2C19	2C9	2D6			3A4
1.	92.714	-0.199	-1.989	No	Yes	No	Yes	Yes	Yes	Yes	0.669	Yes
2.	92.52	-0.087	-1.886	No	Yes	No	Yes	No	Yes	Yes	0.547	Yes
3.	93.818	-0.307	-2.062	No	Yes	No	No	No	No	Yes	0.618	Yes
4.	94.727	-1.311	-2.108	Yes	Yes	No	Yes	Yes	No	Yes	0.817	Yes
5.	92.949	-0.286	-1.962	No	Yes	No	No	No	No	Yes	0.829	Yes
6.	93.876	-0.442	-2.279	No	Yes	No	No	Yes	Yes	Yes	0.727	No
7.	93.605	-0.359	-2.149	No	Yes	No	No	No	No	Yes	0.708	Yes
8.	95.354	-0.457	-2.220	No	Yes	No	Yes	No	No	Yes	0.882	No
9.	93.338	-0.308	-2.053	No	Yes	No	No	No	No	Yes	0.855	Yes
10.	94.033	-0.340	-2.161	No	Yes	No	No	Yes	No	Yes	0.661	Yes
11.	95.698	-1.309	-2.692	No	Yes	No	Yes	Yes	No	Yes	0.472	No
12.	97.519	-1.511	-2.564	No	Yes	No	Yes	No	No	Yes	0.861	No

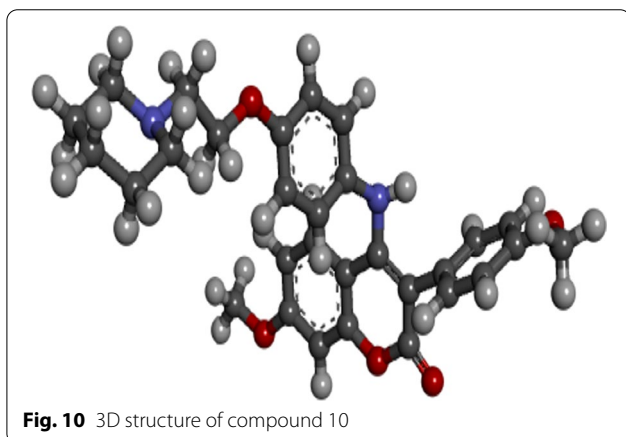
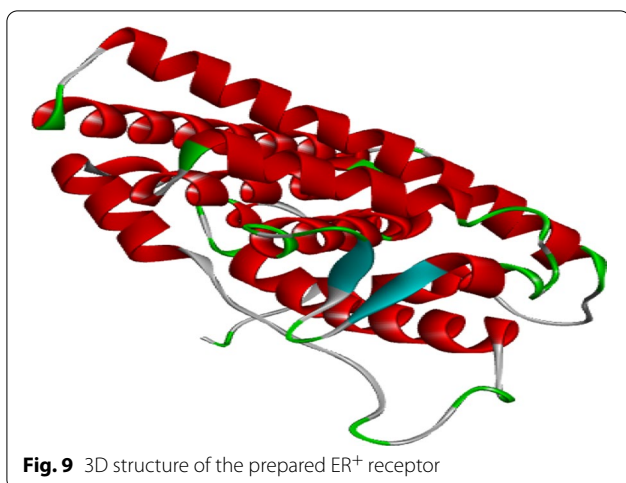
the drug-likeness of a molecule. The pink area signifies the optimum limit for each properties (lipophilicity: XLOGP3 between -0.7 and $+5.0$, size: MW between 150 and 500 g/mol, polarity: TPSA between 20 and 130 Å², solubility: $\log S \leq 6$, saturation: fraction of carbons in the sp³ hybridization ≥ 0.25 , and flexibility: ≤ 9 rotatable bonds (Daina et al. 2017). Designed compounds 6 and 11 are the most orally bioavailable (Fig. 7), since most of their predicted properties are within the pink region while 8 and 12 having most of their predicted properties placed outside the pink region are deemed to be not orally bioavailable. Similarly, as illustrated from their boiled egg plot (Fig. 8), they possess high gastrointestinal values as they are all located within the

white area of the plot and are all P-gp substrate as indicated by their blue colors.

Molecular docking studies

Designed chromen-2-ones 6, 8, 11, and 12 being orally safe were docked into the active site of the ER⁺ receptor using Molegro virtual docker to reveal the nature of interactions with the amino acid residues at the active pocket of the receptor (pdb id=3ERT). They have better docking scores compared to the reference compound, and the standard drug (Tamoxifen) utilized in this research. 3D structures of the prepared ER⁺ receptor and compound 10 are shown in Figs. 9 and 10 while the Docking scores and various kinds of interactions between the docked





chromen-2-ones and the active site of the ER⁺ receptor are presented in Table 9, respectively.

Interpretation of the docking result

Designed compound 6 (MolDock score = -142.117, Rerank score = -99.5654) interacted with the active site of the ER⁺ receptor via seven (7) Carbon-Hydrogen bonds and six (6) Pi-Alkyl hydrophobic Pi-Alkyl interactions. Carbon-Hydrogen bonds are between TYR526 and the Oxygen atom attached directly to the ethylpiperidine group at distance 2.83 Å, LYS529, CYS530, LYS531 and VAL 533 with Oxygen and Hydrogen atoms of the methoxy group attached to the chromen-2-one scaffold at distances 2.84 Å, 2.455 Å, 2.64 Å, and 2.839 Å. MET522 forms the other Carbon-Hydrogen bonds with Hydrogen atoms of the ethylpiperidine group at distances 1.586 Å and 2.287 Å. TYR526, LEU354, LEU536, VAL533 and MET522 forms Hydrophobic Pi-Alkyl interactions. 3D and 2D interactions of designed compound 6 with the active site of the ER⁺ receptor is shown in Fig. 11.

Designed compound 8 (MolDock score = -166.475 Rerank score = -102.019) interacted with the active site of the ER⁺ receptor through a conventional Hydrogen bond, three Carbon-Hydrogen bond, double Pi-Sulfur interactions, Pi-Pi stacked and amide-Pi stacked interactions, and several alkyl as well as Pi-Alkyl hydrophobic interactions. Para methoxy oxygen atom forms a conventional Hydrogen bond with LEU536 at distance 2.147 Å, VAL534 forms two carbon-hydrogen bonds with the methoxy group Hydrogen atoms at 3.04 and 2.82 Å, and the last Carbon-Hydrogen bond is between GLU380 and Hydrogen atom of the piperidine group at 3.011 Å. MET343 and MET522 form Pi-Sulfur interactions, and TRP383 and LEU525 form hydrophobic Pi-Pi stacked and Amide-Pi stacked interactions. Finally, LYS529, CYS530, VAL533, MET528, ALA350, LEU525, and LEU536 residues formed Alkyl and Pi-Alkyl Hydrophobic interactions. 3D and 2D interactions of designed compound 8 with the active site of the ER⁺ receptor are shown in Fig. 12.

Designed compound 11 (MolDock score = -154.303 Rerank score = -104.610) is found to have interacted with the binding pocket of the ER⁺ receptor through two (2) conventional Hydrogen bonds, five (5) Carbon-Hydrogen bonds, Pi-Sulfur, Amide-Pi stacked, Alkyl and Pi-Alkyl hydrophobic interactions. Carbonyl Oxygen atom and Hydrogen atom of the dimethyl amine groups form two conventional Hydrogen bonds with TRP383 and MET343 at distances 2.25 and 2.42 Å. THR347 forms two Carbon-Hydrogen bonds with Hydrogen atom of the piperidine and dimethyl amine group Hydrogen atom at 2.83 and 2.62 Å, methoxy group Hydrogen atoms attached to the chromen-2-ones scaffold forms two additional Carbon-Hydrogen bonds with ASN591 at 3.04 and 2.877 Å, VAL534 forms the last Carbon-Hydrogen bond with para methoxy group Hydrogen atom at distance 2.86 Å. MET522 forms a pair of Pi-Sulfur interactions, LEU525 forms an Amide-Pi stacked hydrophobic interactions, while ALA350, TRP383, MET522, LEU536, LEU525 and LYS529 residues formed Alkyl and Pi-Alkyl hydrophobic interactions. 3D and 2D interactions of designed compound 11 with the binding site of the ER⁺ receptor is portrayed in Fig. 13.

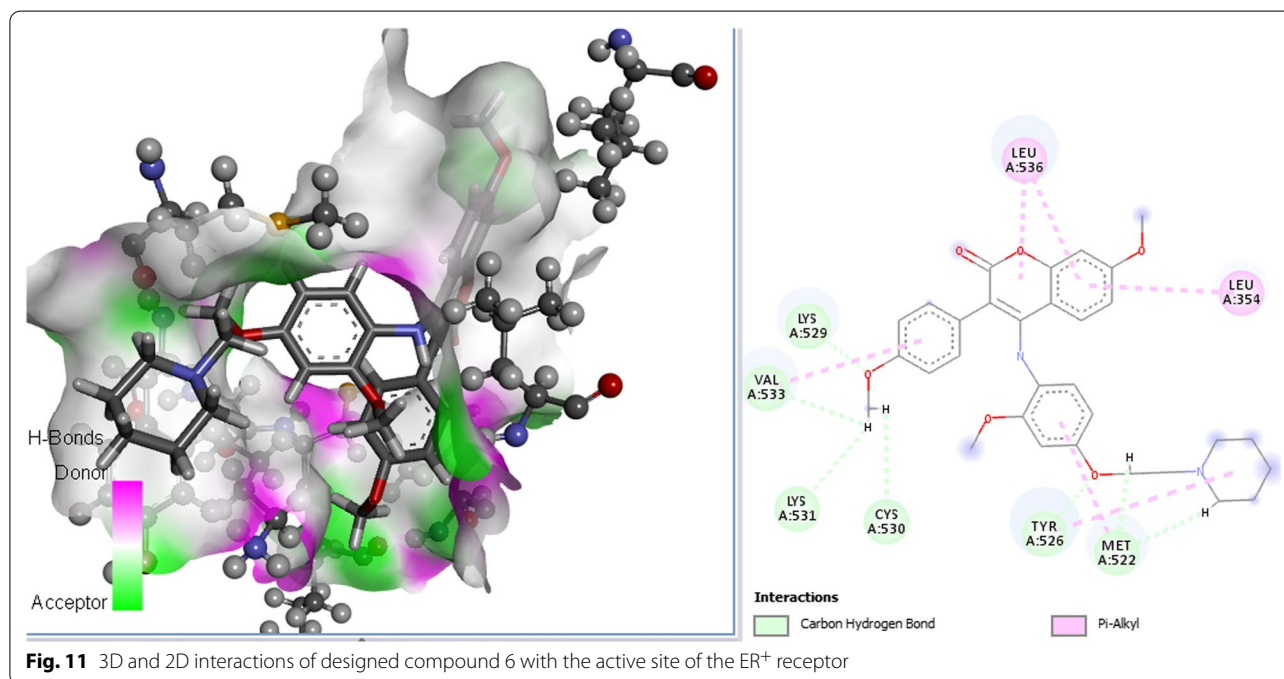
The binding mode of designed compound 12 (MolDock score = -170.357 Rerank score = -171.531) with the active site of the ER⁺ receptor is through a single conventional Hydrogen bond, seven (7) Carbon-Hydrogen bonds, Pi-Sulfur and Amide-Pi stacked interactions, and several alkyl as well as Pi-Alkyl hydrophobic interactions. Methoxy group oxygen atom attached to the chromen-2-one frame forms a conventional Hydrogen bond with HIS524 at 2.34 Å distance, GLY520 forms a pair of Carbon-Hydrogen bonds with Oxygen and Hydrogen

Table 9 Docking scores and various kinds of interactions of the designed chromen-2-ones and the active site of the ER⁺ receptor

S/nos	MolDock score	Rerank score	Amino acid residues	Categories	Interaction type
6	- 142.117	- 99.5654	TYR526	Hydrogen bond	Carbon-hydrogen
			LYS529	Hydrogen bond	Carbon-hydrogen
			CYS530	Hydrogen bond	Carbon-hydrogen
			LYS531	Hydrogen bond	Carbon-hydrogen
			VAL533	Hydrogen bond	Carbon-hydrogen
			MET522	Hydrogen bond	Carbon-hydrogen
			TYR526	Hydrophobic	Pi-alkyl
			LEU354	Hydrophobic	Pi-alkyl
			LEU536	Hydrophobic	Pi-alkyl
			VAL533	Hydrophobic	Pi-alkyl
8	- 166.475	- 102.019	MET522	Hydrophobic	Pi-alkyl
			LEU536	Hydrogen bond	Conventional
			VAL534	Hydrogen bond	Carbon-hydrogen
			GLU380	Hydrogen bond	Carbon-hydrogen
			MET343	Other	Pi-sulfur
			MET522	Other	Pi-sulfur
			TRP383	Hydrophobic	Pi-Pi stacked
			LEU525	Hydrophobic	Amide-Pi stacked
			TYR526	Hydrophobic	Amide-Pi stacked
			LYS529	Hydrophobic	Alkyl
			CYS530	Hydrophobic	Alkyl
			VAL533	Hydrophobic	Alkyl
			MET528	Hydrophobic	Alkyl
			ALA350	Hydrophobic	Pi-alkyl
LEU525	Hydrophobic	Pi-alkyl			
LEU536	Hydrophobic	Pi-alkyl			
11	- 154.303	- 104.61	TRP383	Hydrogen bond	Conventional
			MET343	Hydrogen bond	Conventional
			THR347	Hydrogen bond	Carbon-hydrogen
			VAL534	Hydrogen bond	Carbon-hydrogen
			THR347	Hydrogen bond	Carbon-hydrogen
			ASN519	Hydrogen bond	Carbon-hydrogen
			MET522	Other	Pi-sulfur
			LEU525	Hydrophobic	Amide-Pi stacked
			TYR526	Hydrophobic	Amide-Pi stacked
			ALA350	Hydrophobic	Alkyl
			TRP383	Hydrophobic	Pi-alkyl
			MET522	Hydrophobic	Pi-alkyl
			LEU536	Hydrophobic	Pi-alkyl
			LEU525	Hydrophobic	Pi-alkyl
			LYS529	Hydrophobic	Pi-alkyl
			12	- 170.357	- 117.531
GLY420 -GLU353	Hydrogen bond	Carbon-hydrogen			
ASP351	Hydrogen bond	Carbon-hydrogen			
MET343	Hydrogen bond	Carbon-hydrogen			
LEU346	Other	Pi-sulfur			
THR347	Hydrophobic	Amide-Pi stacked			
ALA350	Hydrophobic	AmidePi-stacked			
LEU354	Hydrophobic	Alkyl			
LEU536	Hydrophobic	Alkyl			
LEU387	Hydrophobic	Alkyl			
MET388	Hydrophobic	Alkyl			
LEU346	Hydrophobic	Alkyl			
MET421	Hydrophobic	Alkyl			
LEU525	Hydrophobic	Pi-alkyl			
LEU346	Hydrophobic	Pi-alkyl			
LEU391	Hydrophobic	Pi-alkyl			
ALA350	Hydrophobic	Pi-alkyl			

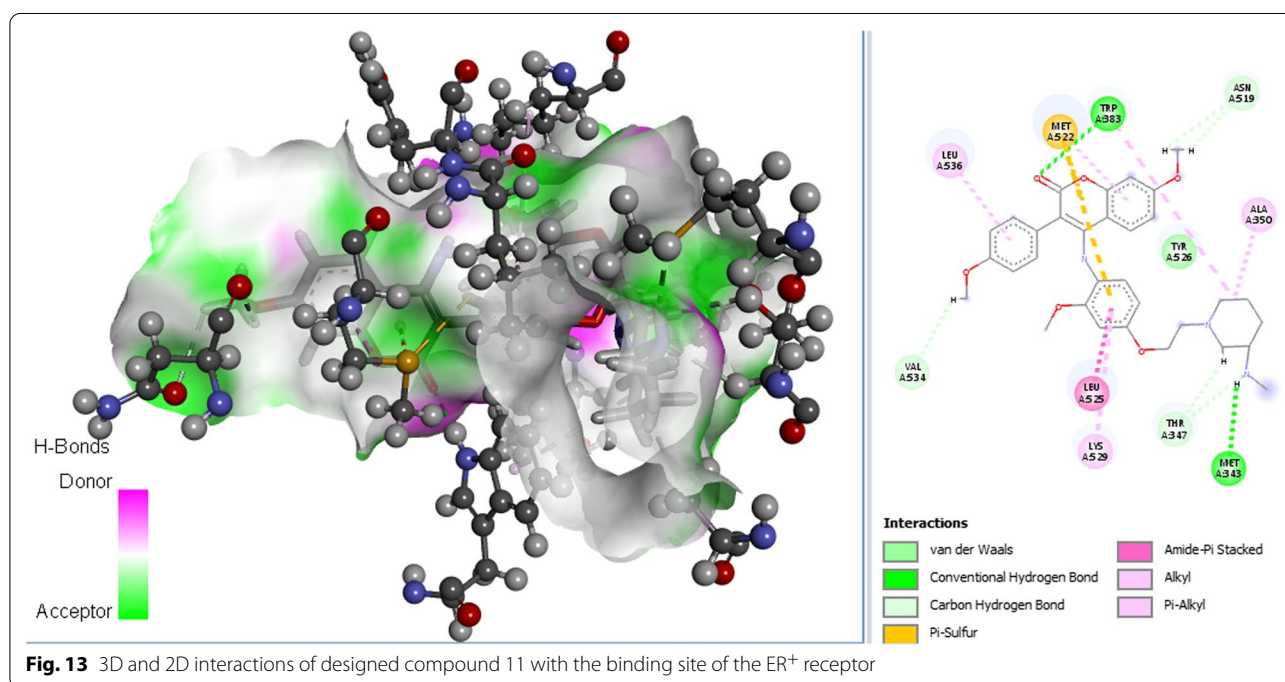
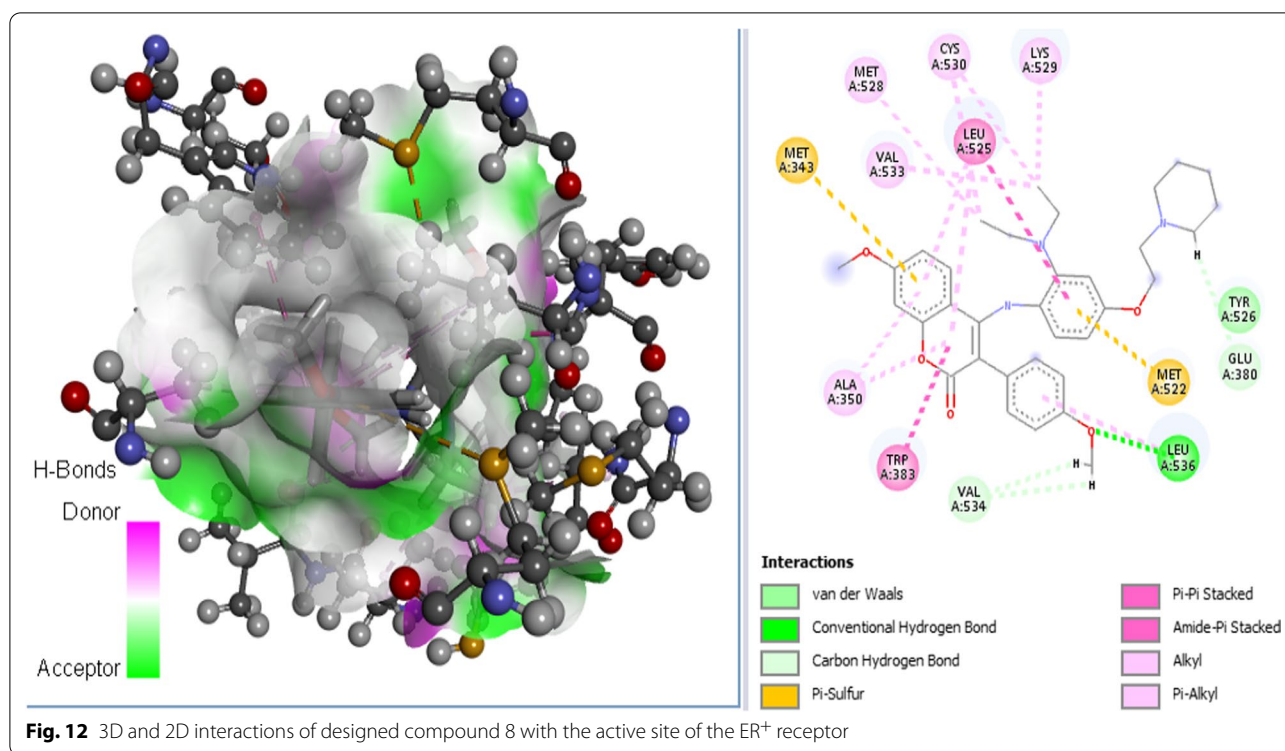
Table 9 (continued)

S/nos	MolDock score	Rerank score	Amino acid residues	Categories	Interaction type
Reference compound (10)	− 142.022	− 108.524	CYS530	Hydrogen bond	Conventional
			LEU525	Hydrogen bond	Conventional
			THR347	Hydrogen bond	Carbon–hydrogen
			LEU346	Hydrophobic	Alkyl
			ALA350	Hydrophobic	Alkyl
			LEU525	Hydrophobic	Alkyl
			LYS529	Hydrophobic	Pi-alkyl
			CYS530	Hydrophobic	Pi-alkyl
			LEU525	Hydrophobic	Pi-alkyl
			Original Tamoxifen		
ASP351	Hydrogen bond	Conventional			
MET343	Hydrogen bond	Carbon–hydrogen			
LEU346	Other	Pi-sulfur			
LEU346	Hydrophobic	Amide-Pi stacked			
ALA350 LEU387	Hydrophobic	Pi-alkyl			
LEU346 LEU525 MET421	Hydrophobic	Pi-alkyl			
	Hydrophobic	Pi-alkyl			
	Hydrophobic	Pi-alkyl			
	Hydrophobic	Pi-alkyl			
Redocked Tamoxifen	− 155.184	− 114.415	ASP351	Hydrogen bond	Carbon–hydrogen
			MET388 MET421 ILE424	Hydrophobic	Alkyl
			LEU428	Hydrophobic	Alkyl
			ALA350 LEU525 MET421	Hydrophobic	Alkyl
			LEU346 LEU387	Hydrophobic	Pi-alkyl
				Hydrophobic	Pi-alkyl
				Hydrophobic	Pi-alkyl
				Hydrophobic	Pi-alkyl
				Hydrophobic	Pi-alkyl
				Hydrophobic	Pi-alkyl



atoms of the methoxy group attached to the chromen-2-one scaffold at 2.71 Å and 1.54 Å, GLU533 residues forms another pair of Carbon–Hydrogen bonds with

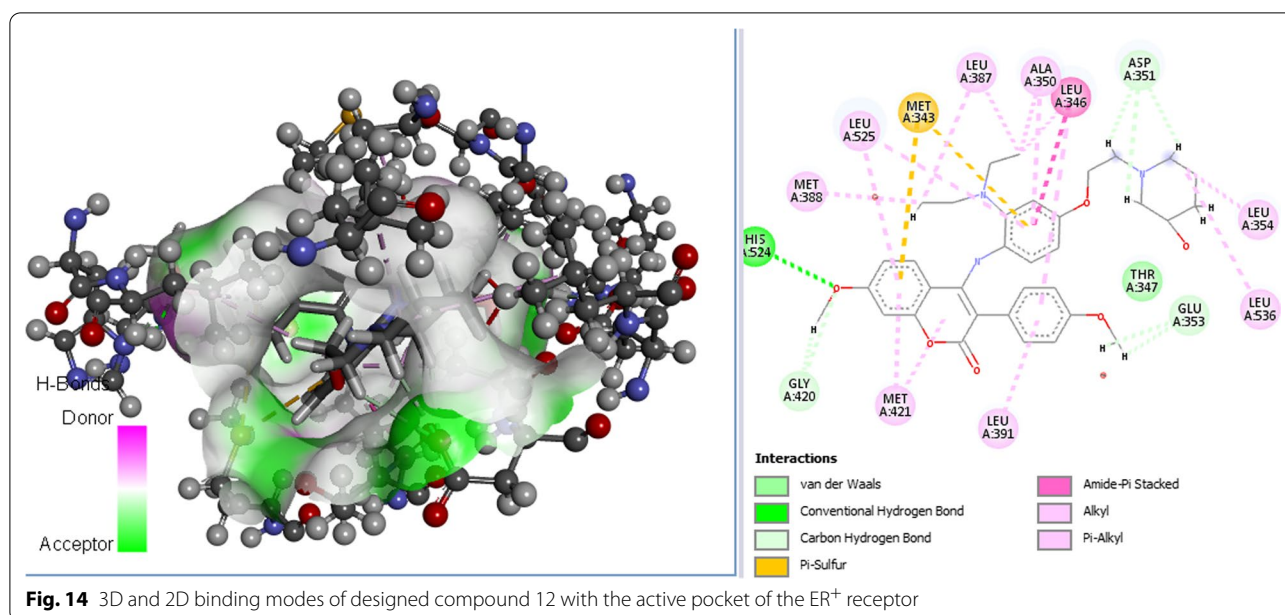
p-methoxy group Hydrogen atoms at distances 2.02 and 2.95 Å, ASP351 forms triple Carbon–Hydrogen bonds with Hydrogen atoms of the ethyl piperidine group at



distances 2.87, 2.49 and 2.78 Å, respectively. MET343 forms double Pi-Sulfur interactions, LEU346 forms an Amide-Pi stacked and ALA350, LEU354, LEU536, LEU387, MET388, LEU346, MET421, LEU525 and LEU391 residues formed Alkyl and Pi-Alkyl Hydrophobic

interaction. 3D and 2D binding modes of designed compound 12 with the active pocket of the ER⁺ receptor are shown in Fig. 14, respectively.

The reference compound (Moldock score = -142.022 Rerank score = -108.524) interacted with the binding

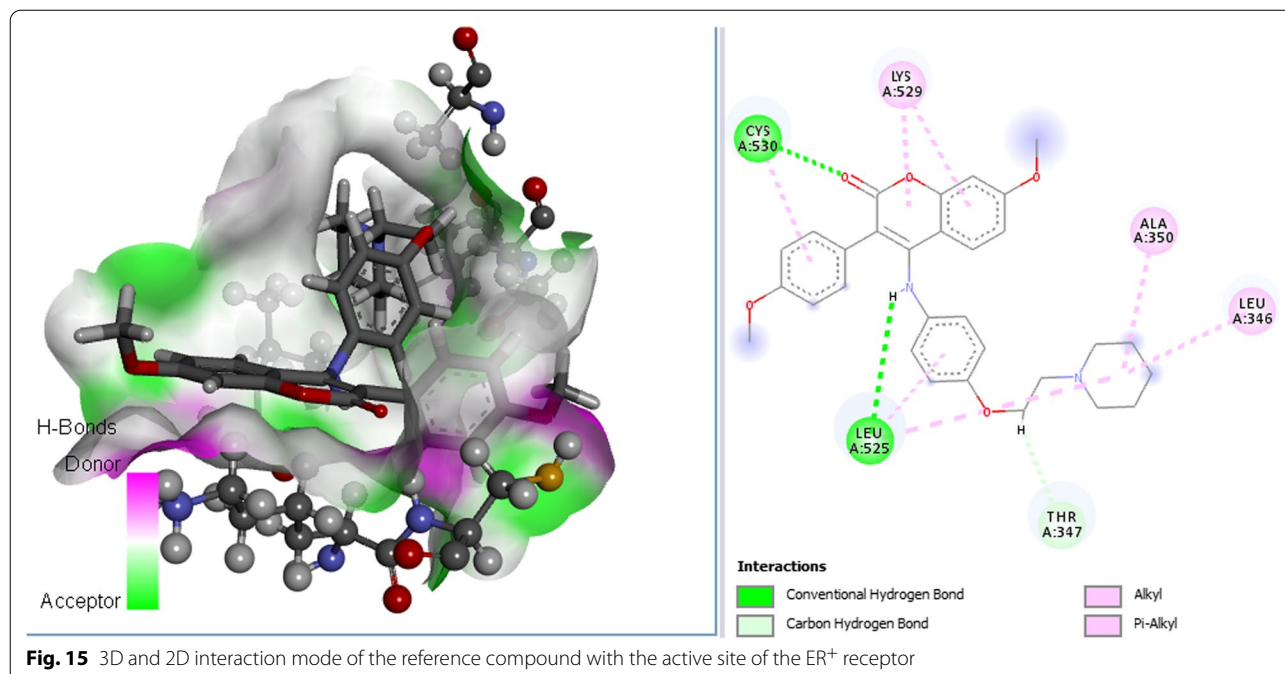


pocket of the ER⁺ receptor through two (2) conventional Hydrogen bonds, single Carbon–Hydrogen bond, and several Alkyl and Pi-Alkyl interactions. Carbonyl Oxygen atom of the chromen-2-one scaffold forms a conventional Hydrogen bond with CYS530 at distance 2.26 Å, and LEU525 forms the other conventional Hydrogen bond with Hydrogen atom of the amino group attached to the chromen-2-one frame at 2.54 Å. THR347 forms Carbon–Hydrogen bond with ethyl Hydrogen atom attached

to the piperidine group at 2.67 Å. LEU346, ALA350, LEU525, LYS529, CYS530 and LEU525 residues formed Alkyl and Pi-Alkyl hydrophobic interactions. 3D and 2D interaction mode of the reference compound with the active site of the ER⁺ receptor is shown in Fig. 15.

Validation of docking protocol

For the validation of the docking results, the reference drug (Tamoxifen) which is also the co-crystallized ligand

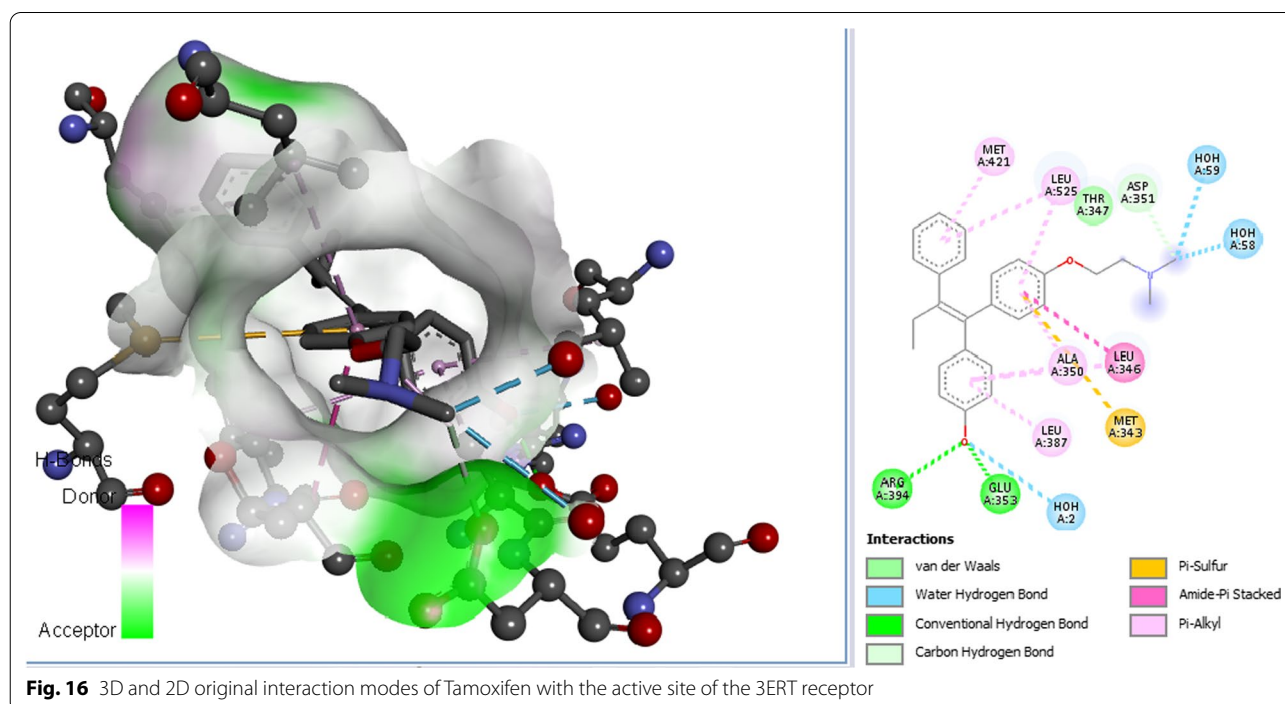


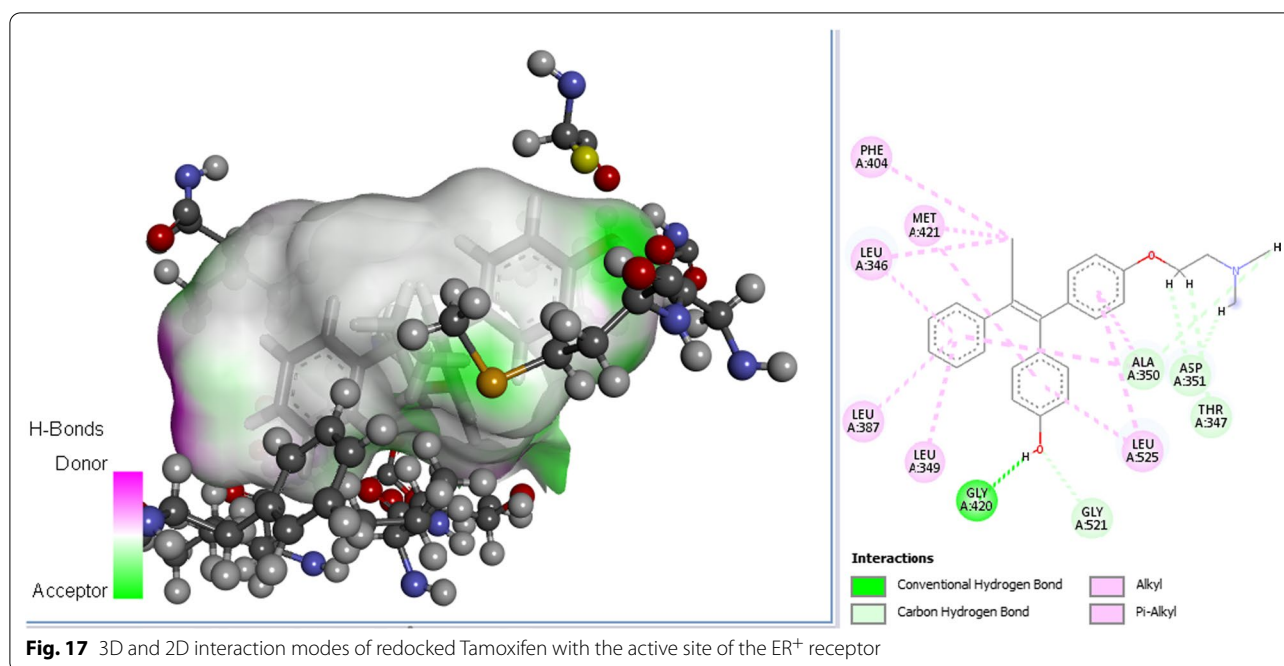
was redocked into the original binding pocket of the ER⁺ receptor. The redocked modes were compared with the original docking modes of the Tamoxifen with the binding site of the receptor. The original docking modes of interaction of Tamoxifen are through three (3) conventional Hydrogen bonds with water molecules, additional pair of conventional Hydrogen bonds, single Carbon–Hydrogen bond, Pi-Sulfur and Amide-Pi stacked interactions and several Pi-Alkyl interactions. Hydroxyl Oxygen atom and Hydrogen atoms of the trimethyl amine group form the three (3) conventional Hydrogen atoms with water molecules. Two (2) conventional Hydrogen bonds are between Hydroxyl group oxygen atom and ARG394 and GLU353 residues at 3.02 and 2.42 Å distances. ASP351 forms the Carbon–Hydrogen bonds with Nitrogen atom of the trimethyl amine group at 3.20 Å, phenyl ring intercalated in space and forms a single Pi-Sulfur interaction with MET343 and Amide-Pi stacked interaction LEU346. ALA350, LEU387, LEU346, LEU525 and MET421 residues formed Pi-Alkyl hydrophobic interactions. 3D and 2D original interaction modes of Tamoxifen with the active site of the 3ERT receptor is presented in Fig. 16.

The redocked Tamoxifen (MolDock score = −155.184 Rerank score = −144.415) interaction mode with the active pocket of the ER⁺ is through a conventional Hydrogen bond, five (5) Carbon–Hydrogen bonds and many Alkyl and Pi-Alkyl hydrophobic interactions. GLY420 forms a conventional Hydrogen bond with

Hydrogen atom of the Hydroxyl group at 1.79 Å distance, GLY421 forms a carbon–hydrogen bond with Hydroxyl group oxygen atom at distance 2.37 Å, THR347 and ASP351 form three other Carbon–Hydrogen bonds with Hydrogen atoms 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. 3D and 2D interaction modes of redocked Tamoxifen with the active site of the ER⁺ receptor are shown in Fig. 17.

Furthermore, the original and redocked Tamoxifen complexes were superimposed and aligned using Discovery studio. The root mean square deviation (RMSD) values between the superimposed proteins were found to be 1.15 Å, and this confirms the stability of the protein and that the Tamoxifen binds perfectly well to the binding pocket of the 3ERT receptor in the redocked complex (Umar et al. 2019). Thus, the docking protocol is reliably validated. The docking outcomes of the designed molecules suggested that Hydrogen bonds, electrostatic and hydrophobic (Alkyl and Pi-Alkyl) interactions are the central driving forces that regulate the binding interactions of the designed chromen-2-ones analogs and the active site residues of the ER⁺ receptor, and that the docking scores increase as the number of interactions increases. The designed compounds have higher docking scores than the reference and the standard drug (Tamoxifen), and this is related to the increase in the number of





Hydrogen bonds and other interactions due to the presence of more substituent groups in the designed compounds (Abdullahi et al. 2022a, b).

Conclusions

In this study, a predictive QSAR model capable of explaining the structural requirements accountable for the anti-cancer activities of chromen-2-one analogs was developed using genetic function algorithm (GFA). Model 1 was selected for this research based on its excellent statistical parameters. Twelve new novel Chromen-2-one analogues were designed through the structural adjustment of compound 10 adopted as reference compound, and their activities was predicted using the selected model. They have improved pIC_{50} which ranges from (5.472 to 8.584), compared to the reference compound ($pIC_{50} = 5.344$) and Tamoxifen ($pIC_{50} = 4.843$) utilized as control drug in the research. Moreover, the designed chromen-2-ones possess drug likeness properties since they do not violate the Lipinski's rule of five, but ADMET studies showed that designed compounds 6, 8, 11 and 12 are orally safe and non-toxic. These orally safe designed compounds were subjected to molecular docking studies with the active site of the ER⁺ receptor kinase, and they were found to show promising binding scores compared to the reference compound and the standard drug used in the study. Hence, these compounds can be utilized as novel ER⁺ breast cancer drug candidates after performing *in vivo* and *in vitro* studies.

Abbreviations

QSAR: Quantitative structure activity relationship; DFT: Density functional theory; PADEL: Pharmaceutical Data Exploration Laboratory; B3LYP: Bee-3-Lee Yang Par; GFA-MLR: Genetic function algorithm-multi linear regression; AD: Applicability domain; ME: Mean effect; ADMET: Absorption, distribution, metabolism, excretion and toxicity; ER⁺: Estrogen receptor⁺; SERMs: Selective estrogen modulators; RMSD: Root mean square deviation.

Acknowledgements

The authors genuinely acknowledge all the contributors of this exploratory group for their guidance and motivation during this research work and Ahmadu Bello University for supplying the softwares and favorable environment employed for this work.

Author contributions

All authors contribute accordingly: SHA executed all the computational analysis and drafted the manuscript; AU supplied all the softwares for this research and revised the manuscript to guarantee that error was minimized prior to final submission; ABU examined the results obtained from the softwares; GAS and SU evaluated the manuscript by using Plagiarism checker and organized the manuscript in accordance with the format of the journal. All authors have read and approved the manuscript.

Funding

Funds allocation was not received by the authors for this research.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable, because this article does not contain any studies with animal or human subjects.

Consent of publication

Not applicable.

Competing interest

The correspondents did not acknowledge competing of interest.

Received: 17 May 2022 Accepted: 14 June 2022

Published online: 21 June 2022

References

- Abdullahi SH, Uzairu A, Ibrahim MT et al (2021) Chemo-informatics activity prediction, ligand based drug design, Molecular docking and pharmacokinetics studies of some series of 4, 6-diaryl-2-pyrimidinamine derivatives as anti-cancer agents. *Bull Natl Res Cent* 45:167. <https://doi.org/10.1186/s42269-021-00631-w>
- Abdullahi SH, Uzairu A, Shallangwa GA et al (2022a) In-silico activity prediction, structure-based drug design, molecular docking and pharmacokinetic studies of selected quinazoline derivatives for their antiproliferative activity against triple negative breast cancer (MDA-MB231) cell line. *Bull Natl Res Cent* 46:2. <https://doi.org/10.1186/s42269-021-00690-z>
- Abdullahi SH, Uzairu A, Shallangwa GA et al (2022b) 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. <https://doi.org/10.1080/2314808X.2022.2050115>
- Amir SE, Freedman OC, Seruga B, Evans DG (2010) Assessing women at high risk of breast cancer: a review of risk assessment models. *J Natl Cancer Inst* 102:680–691. <https://doi.org/10.1093/jnci/djq088>
- Arshad A, Osman H, Bagley MC, Lam CK, Mohamad S, Zahariluddin AS (2011) Synthesis and antimicrobial properties of some new thiazolyl coumarin derivatives. *Eur J Med Chem* 46:3788–3794. <https://doi.org/10.1016/j.ejmech.2011.05.044>
- Bai Z, Gust R (2009) Breast cancer, estrogen receptor and ligands. *Arch Pharm* 342:133–149. <https://doi.org/10.1002/ardp.200800174>
- Bickerton GR, Paolini GV, Besnard J et al (2012) Quantifying the chemical beauty of drugs. *Nat Chem* 4(2):90. <https://doi.org/10.1038/nchem.1243>
- Bisi A, Cappadone C, Rampa A, Farruggia G, Sargenti A, Belluti F, Di Martino RM, Malucelli E, Meluzzi A, Iotti S, Gobbi S (2017) Coumarin derivatives as potential antitumor agents: Growth inhibition, apoptosis induction and multidrug resistance reverting activity. *Eur J Med Chem* 127:577–585. <https://doi.org/10.1016/j.ejmech.2017.01.020>
- Chen JY, Kuo SJ, Liaw YP, Avital I, Stojadinovic A, Man YG, Mannion C, Wang JL, Chou MC, Tsai HD (2014) Endometrial cancer incidence in breast cancer patients correlating with age and duration of tamoxifen use: a population based study. *J Cancer* 5:151–155. <https://doi.org/10.7150/jca.8412>
- Chen LZ, Sun WW, Bo L, Wang JQ, Xiu C, Tang WJ, Shi JB, Zhou HP, Liu XH (2017) New arylpyrazoline-coumarins: synthesis and anti-inflammatory activity. *Eur J Med Chem* 138:170–181. <https://doi.org/10.1016/j.ejmech.2017.06.044>
- Emami A, Dadashpour S (2015) Current developments of coumarin-based anti-cancer agents in medicinal chemistry. *Eur J Med Chem* 102:611–630. <https://doi.org/10.1016/j.ejmech.2015.08.033>
- Feitelson T, Arzumanyan A, Kulathinal RJ, Blain SW, Holcombe RF, Mahajna J, Marino M, Martinez-Chantar ML, Nawroth R, Sanchez-Garcia I et al (2015) Sustained proliferation in cancer: mechanisms and novel therapeutic targets. *Semin Cancer Biol* 35:25–54. <https://doi.org/10.1016/j.semcancer.2015.02.006>
- Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJL, Naghavi M (2011) Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *Lancet* 378:1461–1484. [https://doi.org/10.1016/S0140-6736\(11\)61351-2](https://doi.org/10.1016/S0140-6736(11)61351-2)
- Garcia-Becerra R, Santos N, Diaz L, Camacho J (2012) Mechanisms of resistance to endocrine therapy in breast cancer: focus on signaling pathways, miRNAs and genetically based resistance. *Int J Mol Sci* 14:108–145. [https://doi.org/10.1016/S0140-6736\(11\)61351-2](https://doi.org/10.1016/S0140-6736(11)61351-2)
- Guoshun L, Chen M, Lyu W, Zhao R, Xu Q, You Q, Xiang H (2017) Design, synthesis, biological evaluation and molecular docking studies of novel 3-aryl-4-anilino-2H-chromen-2-one derivatives targeting ER α as anti-breast cancer agents. *Bioorgan Med Chem Lett*. <https://doi.org/10.1016/j.bmcl.2017b.04.029>
- Jordan VC (2007) Chemoprevention of breast cancer with selective estrogen-receptor modulators. *Nat Rev Cancer* 7:46–53. <https://doi.org/10.1038/nrc2048>
- Jordan VC, McDaniel R, Agboke F, Maximov PY (2014) The evolution of nonsteroidal antiestrogens. *Steroids* 90:3–12. <https://doi.org/10.1016/j.steroids.2014.06.009>
- Kaur G, Mahajan MP, Pandey MK, Singh P, Ramisetty SR, Sharma AK (2014) Design, synthesis and evaluation of Ospemifene analogs as anti-breast cancer agents. *Eur J Med Chem* 86:211–218. <https://doi.org/10.1016/j.ejmech.2014.08.050>
- Kennard RW, Stone LA (1969) Computer aided design of experiments. *Technometrics* 11(1):137–148. <https://doi.org/10.1080/00401706.1969.10490666>
- Khaled KF (2011) Modeling corrosion inhibition of iron in acid medium by genetic function approximation method: a QSAR model. *Corros Sci* 53(11):3457–3465. <https://doi.org/10.1016/j.corsci.2011.01.035>
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 23(1–3):3–25. [https://doi.org/10.1016/S0169-409X\(96\)00423-1](https://doi.org/10.1016/S0169-409X(96)00423-1)
- Luo G, Li X, Zhang G, Chengzhe Wu, Tang Z, Liu L, You Q, Xiang H (2017a) Novel SERMs based on 3-aryl-4-aryloxy-2H-chromen-2-one skeleton—a possible way to dual ER α /VEGFR-2 ligands for treatment of breast cancer. *Eur J Med Chem*. <https://doi.org/10.1016/j.ejmech.2017.09.015>
- Luo GS, Muyaba M, Lyu WT, Tang ZC, Zhao RH, Xu Q, You QD, Xiang H (2017c) Design, synthesis and biological evaluation of novel 3-substituted 4-anilino-coumarin derivatives as antitumor agents. *Bioorg Med Chem Lett* 27:867–874. <https://doi.org/10.1016/j.bmcl.2017.01.013>
- Martin YC (2005) A bioavailability score. *J Med Chem* 48(9):3164–3170. <https://doi.org/10.1021/JM0492002>
- Maruyama K, Nakamura M, Tomoshige S, Sugita K, Makishima M, Hashimoto Y, M, (2013) Ishikawa Structure–activity relationships of bisphenol A analogs at estrogen receptors (ERs): discovery of an ER α -selective antagonist. *Bioorg Med Chem Lett* 23:4031–4036. <https://doi.org/10.1016/j.bmcl.2013.05.067>
- Maurer C, Martel S, Zardavas D, Ignatiadis M (2017) New agents for endocrine resistance in breast cancer. *The Breast* 34:1–11. <https://doi.org/10.1016/j.breast.2017.04.007>
- Olomola TO, Klein R, Mautsa N, Sayed Y, Kaye PT (2013) Synthesis and evaluation of coumarin derivatives as potential dual-action HIV-1 protease and reverse transcriptase inhibitors. *Bioorg Med Chem* 21:1964–1971. <https://doi.org/10.1016/j.bmcl.2013.01.025>
- Sinha S, Kumaran AP, Mishra D, Paira P (2016) Synthesis and cytotoxicity study of novel 3-(triazolyl) coumarins based fluorescent scaffolds. *Bioorg Med Chem Lett* 26:5557–5561. <https://doi.org/10.1016/j.bmcl.2016.09.078>
- Sommer S, Fuqua SA (2001) Estrogen receptor and breast cancer. *Semin Cancer Biol* 11(5):339–352. <https://doi.org/10.1006/scbi.2001.0389>
- Thakur A, Singla R, Jaitak V (2015) Coumarins as anticancer agents: a review on synthetic strategies, mechanism of action and SAR studies. *Eur J Med Chem* 101:476–495. <https://doi.org/10.1016/j.ejmech.2015.07.010>
- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. *CA Cancer J Clin* 65(2):87–108. <https://doi.org/10.3322/caac.21492>
- Traboulsi T, Ezy ME, Gleason JL, Mader S (2017) Antiestrogens: structure-activity relationships and use in breast cancer treatment. *J Mol Endocrinol* 58:15–31. <https://doi.org/10.1530/JME-16-0024>
- Tropsha A, Gramatica P, Gombar VK (2003) The importance of being earnest: validation is the absolute essential for successful application and interpretation of Qspr models. *Mol Inf* 22:69–77. <https://doi.org/10.1002/qsar.200390007>
- Umar BA, Uzairu A, Shallangwa GA, Sani U (2019) QSAR modeling for the prediction of pGI₅₀ activity of compounds on LOX IMVI cell line and ligand based design of potent compounds using in silico virtual screening. *Netw Mod Anal Health Infor Bioinform* 8(1):22. <https://doi.org/10.1016/j.heliyon.2020.e03640>
- Wang TL, You QD, Huang FSG, Xiang H (2009) Recent advances in selective estrogen receptor modulators for breast cancer. *Mini-Rev Med Chem* 9:1191–1201. <https://doi.org/10.2174/138955709789055207>
- Yap CW (2011) PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J Comput Chem* 32(7):1466–1474. <https://doi.org/10.1002/jcc.21707>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.