Ind. J. Pharm. Edu. Res., 2023; 57(3):838-853.
https://www.ijper.org

Original Article

2D-QSAR, Molecular Docking, and *in silico* Pharmacokinetics Analysis on N-substituted Urea and Thiourea Derivatives as Tankyrase Inhibitors for Implication in Cancer

Hafiz A. Makeen[1,*], Mohammed Albratty[2]

[1]Department of Pharmacy Practice, Pharmacy Practice Research Unit, College of Pharmacy, Jazan University, Jazan, SAUDI ARABIA.
[2]Department of Pharmaceutical Chemistry and Pharmacognosy, Faculty of Pharmacy, Jazan University, Jazan, SAUDI ARABIA.

ABSTRACT

Objectives: Cancer is among four major Non-Communicable Diseases (NCD). Based on the World Health Organization (WHO), it is the biggest cause of mortality globally, claiming about 10 million lives in recent years. Tankyrases is a poly polymerase (ADP-ribose) family enzyme, inhibiting its enzymatic processes plays a crucial part in cancer etiology. **Materials and Methods:** The Algorithm of Kennard-Stone was utilized to develop QSAR models of thirty-four cytotoxic compounds of N-naphthoyl thioureas and N-aryl-N′-benzylurea using the multiple linear regression approach. 2D QSAR best models were developed and the finest model was selected using statistical reliability (R^2) of 0.9253, (R^2_{adj}) of 0.9045, (Q^2_{cv}) cross-validation coefficient of 0.8767, and (R^2_{test}) of 0.6015. **Results:** The R^2 value of 0.9253 shows the model is promising by indicating 92.53% of the residual deviation, and this model was not over-fitted, as seen by how near Q^2_{cv} is to internal R^2. The molecular docking studies were conducted between some selected compounds (based on activity) and Tankyrases protein receptors to investigate the binding modalities and ADMET was also conducted to determine their oral bioavailability. Three compounds C18, C25, and C33 showed favourable interaction and good binding energy and were consistent with drug-likeness parameters. Furthermore, the crystal structure of the compounds bound to Tankyrases protein receptors has yet to be confirmed experimentally. **Conclusion:** Thus this article provides insight into the residual interaction between these compounds and Tankyrases protein receptors.

Keywords: Cancer, QSAR, Docking, Tankyrase inhibitors.

Correspondence:
Dr. Hafiz A. Makeen
Department of Pharmacy Practice, Pharmacy Practice Research Unit, College of Pharmacy, Jazan University, Jazan, SAUDI ARABIA.
Email: hafiz@jazanu.edu.sa

Received: 15-02-2023;
Revised: 10-04-2023;
Accepted: 16-06-2023.

INTRODUCTION

Cancer is life-threatening to the world even though treatment methods such as chemotherapy, immunotherapy, and radiotherapy are less effective.[1,2] Cancer treatments are covered with many problems such as pharmaceutical high costs, resistance to the drug, and lack of drug selectivity.[3,4] As such, there is a need for new effective drugs that is safer and more selective than chemotherapeutic agents.[5] Cancer of Colon is currently among one of those with highest mortality rate. The rate of mortality of this type of cancer is not from the primary tumor itself, but rather from the cellular mortality increase, extravasation, and multiple-step process for extracellular changes of matrix invasion support called metastasis nature of the tumor.[6] Tankyrases belong to the versatile polymerase poly (ADP-ribose) family that control a varity of biological activities. Inhibiting their metabolic activity reduces Wnt/-catenin activity, which is critical in cancer etiology. Several investigations have found tankyrase inhibitors to be important and useful in cancer treatments.[7-9] Tankyrases add numerous ADP-ribose groups (PARsylate) to a variety of target proteins, particularly AXIN, a WNT signaling pathway negative regulator. PARsylated AXIN is degraded in the proteasomes by ubiquitination, leading to the activation of the WNT signaling pathway and β-catenin translocation to the nucleus. Inhibiting tankyrase decreases Wnt transmission and cancer progression in APC-mutant colorectal cancers.[10-12] A Chemotherapy conventionaly combined with inhibition tankyrase often has synergistic anti-cancer benefits. As a result, it is envisaged that more sophisticated and superior tankyrase inhibitors would be created, allowing for an effective approach for innovative cancer therapy techniques.[13]

This investigation focus on creating a robust and predictive model for anti-cancer property of substituted Urea and thiourea derivatives using two-dimensional QSAR utilizing a multiple linear regression approach, moreover, to dock the compounds

DOI: 10.5530/ijper.57.3.102

Copyright Information :
Copyright Author (s) 2023 Distributed under Creative Commons CC-BY 4.0

Publishing Partner : EManuscript Tech. [www.emanuscript.in]

for possible inhibition of Tankyrases protein receptor which has not yet been reported. Furthermore, compounds with good interaction with the receptor are chosen and evaluated using ADMET to assess their drug-likeness.

A flexible method for swiftly constructing (thio)urea pharmacophoric derivatives such as thiazolidines, triazole, and tetrazole was used to produce novel lead compounds for the synthesis of a variety of active heterocycles.[14] Thus, series of mono- and bis-thioureas were gathered and their cytotoxic behaviours were analysed to obtain other possible influence on activity of other substituents irrespective of Urea or thiourea moieties. The thiourea cyclization to thiazolidine derivative as reported slightly reduces the anticancer activity. Likewise, presence of arylidene chlorophyll motif showed anticancer improvement behaviour. A broad range of pharmaceutical applications due to drugs and proteins' important interactions are possible. This interaction can either be ADMET or enzyme catalysis, as well as protein-ligand interactions in which hydrogen bonds are a vital process.[15] Functional groups present in Urea which are responsible for donors and acceptors hydrogen bonds are reported to have excellent cytotoxic application.[16] The scaffold of Thiourea and that of its urea counterpart display an important role in enzymes and bio receptors binding ability as well as tyrosine kinases, topoisomerase, and NADH oxidase enzyme inhibition capability.[17,18] Urea and derivatives showcase a wide range of applications which includes antimicrobial, anticancer activities, apoptosis induction, multiple cancer cell lines high activities, cytotoxicity, and inhibition of histone deacetylase and EGFR kinase enzyme.[19-23] Furthermore, multiple interesting types of research had been used to develop more potent chemotherapeutic agents by using urea and Thiourea as multipurpose bioactive molecules.[24] The commercially available drug-based thiourea and urea moieties reported are as follows; methylthiouracil, carbimazole, propylthiouracil,[25] sorafenib, and regorafenib.[26]

QSAR is a predictive model that use scientific metrics to relate metabolic processes, such as physico-chemical properties, pharmacological effects, and certain side effects, to descriptor traits or molecular structure.[27] Numerous variables affect the accuracy of a QSAR model, which finally leads in good predictive predictions of new compound responses that are reliable and accurate. This includes the input data quality, the attributes employed, and the numerical modelling and validation methods.[28] The use of molecular docking to help with new therapeutic techniques has also emerged. This is being used in conjunction between several modelling methods such as QSAR inside planned strategies.[29] Currently, the three-dimensional shape of molecules has paved the way for new ways to enhance drugs activity by adjusting some characteristics such as ADME variables, solubility and lipophilicity.[30,31] Advanced technology is increasing awareness in fragment-based regenerative medicine and important characteristics, with a focus on chemical and biological exploration and improved synthetic availability.[31,32]

MATERIALS AND METHODS

QSAR modelling and compounds Molecular Structure

Compounds were collected by data mining from two published articles and synthesized using the following scheme of reaction. 2-naphthoyl chloride with KSCN and cyclohexylamine in acetone to give aroyl thiourea derivative. The designed compounds showed the inclusion of the chlorophenyl arylidene motif, 2-naphthoyl chloride and other lipophilic chloro-substituent improved cytotoxic effects. This resulted in building of biologically active promising thioureas and enhancement of association with hydrophilic binding sites. As previously reported in the literature that ammonium thiocyanate (NH_4SCN) or Potassium Thiocyanate (KSCN) are versatile in N-aroyl thioureas synthetic routes.[33-35] Also, another novel series of urea derived compounds were synthesized using derivatives of pyridine hydrochloride and different intermediate isocyanate.[26]

Collection of Data

Thirty-four (34) anti-cancer compounds are gathered from the existing literature for this investigation.[26,36] To obtain a well-defined range, these applications are transformed using Eq. (1) to a logarithmic scale. This process minimized the skew in the activities also normalized the compounds' activity.[32,37-39]

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

Descriptors calculation, sketching of compound and optimization

ChemDraw software Ultra version 12.0 was used to create the compounds (2D) structures, which were then uploaded to Spartan software version 14 V.1.1.4 to determine the best (3D) structures at DFT level utilizing basis set of B3LYP 6-31G*.[40,41] The improved format of Spartan chemicals were then converted into SD files and thereafter imported into the program (PaDEL) to derive model parameters for proposed study.[40]

Model Development of QSAR and Validation

The program's Kennard-Stone Algorithm was used to arrange the data set into two subgroups, training and test. The models of QSAR were created through using training data set, that comprised 70% (24 compounds) of all data, as well as the test data set, consisting 30% (10 compounds) collectively; the latter was employed to assess prediction skills of the model.[40,41] To determine the pIC activity levels of all compounds under assessment, the produced QSAR model was used. To increase the models' accuracy, using the Genetic Function Algorithm (GFA) in order to choose acceptable descriptors.[40] Multiple

Linear Regression (MLR) was used to analyse the relationship between the independent and dependent variables (pIC_{50}), X and Y respectively using the training set as a starting point (molecular descriptors). The dependent variables contingents mean (pIC_{50}) Y was calculated by studying (molecular descriptors) X, based on regression theory. The optimum was selected based on QSAR model validation metrics which include the coefficient of correlation (R^2), modified coefficient of correlation (R^2_{adj}), Q^2_{CV}, and coefficient of correlation for an external projection set (R^2_{pred}) [Supplied in formulations (2, 3, 4, 5) beneath].

$$R^2 = 1\frac{\Sigma(Y_{exp} - Y_{pred})2}{\Sigma(Y_{exp} - Ym_{training})2} (2)$$

$$R^2 pred = 1\frac{\Sigma(Y_{pred} - Y_{exp})2}{\Sigma(Y_{exp} - Ym_{training})2} (3)$$

$$Q^2 cv = 1\frac{\Sigma(Y_{exp} - Y_{pred})2}{\Sigma(YY_{exp} - Ym_{training})2} (4)$$

$$R^2 adj = \frac{(N-1)R^2 - P}{N - P + 1} (5)$$

The representative sample is N in the model, and the independent variables number is P. Also Y_{exp} denotes experimental activity function, Y_{pred} denotes projected activity function, and $Ym_{training}$ is the average experimental compounds' performance throughout the QSAR model set.[40]

The QSAR application Domain

A QSAR model's Application Domain (AD) is the chemical theoretical space that includes its relevant variables in addition to the response predicted. In this domain, based on the chemical data set utilized in the model's building, the degree of uncertainty in recognizing a specific molecule may be estimated. The AD is also employed in the analysis to define the training set's X-outliers and to detect molecules lying beyond the test set given AD, according to the essential notion of the standardization method. QSAR models' AD has been defined in a variety of ways[42] Gramatica is the most popular approach.[40] It made use of data set leverage. The leveraging method enables analysis of a new molecules' location in the model (QSAR).[41] Therefore, the leverage strategy is employed, as shown in Eq. (6):

$$h_i = X_{i(}X^T X) - 1_{xi} (6)$$

Small x is the compound under investigation's descriptor vector, and Capital X illustrate the descriptor matrix, as created in the training set descriptors values. The warning leverage (h*) was calculated using Eq. (7):

$$h^* = \frac{3(P+1)}{N} (7)$$

The number of the training compound is N, and model number's descriptors are P.

Y – Randomization

To boost the model's efficacy, a new model was developed utilizing an external test validation, which is known as Y-randomization. A good model requires a coefficient of randomized squared correlation (cR2p) greater than 0.5, which is given as Eq. (8)

$$c R_p^2 = R[R^2(R_r)^2]^2 (8)$$

For random models 'R, the Y-randomization coefficient is indicated as cR2p, and the average is indicated as Rr.[41]

Mean Effect (ME) and Variance Inflation Factor (VIF)

VIF was used to examine the linearity of the descriptors' Model. ME was utilized to illustrate importance of Model's descriptions. A score of 1 in VIF implies the absence of linearity between the descriptors, whereas greater than 10 values suggest that the model is defective. The ME and VIF are calculated using Eq. (9) and (10) respectively.

$$ME = \frac{B_j \Sigma_i^n D_j}{\Sigma_j^m (B_j \Sigma_i^n D_j)} (9)$$

Dj is the descriptor's value throughout the matrix on every training set data point, Bj is coefficient of the descriptor j in the model. The quantity of descriptors inside the model and the quantity of molecules in each training set are denoted by m and n, respectively.

$$VIF = \frac{1}{1 - R^2} (10)$$

Where R^2 denotes the multivariate regression coefficient of correlation among model's parameters presented in the literature.[41]

Molecular docking Analysis (Protein-ligand preparation)

Spartan software version 14 was employed to optimize and organize the chosen compounds based on their activity, as well as doxorubicin as a reference drug for docking studies into Protein Data Bank files (PDB). The receptors Tankyrase (TNKS2) crystal structure. (PDB id 6KRO)[43,44] was generated from (www.rcsb.org). The Tankyrase enzyme PDB file was created in Discovery studio by eliminating the superfluous molecules of water inside the X-ray structure and hydrogen molecule optimization, along with eliminating the associated ligand from the substrate prior to docking. Using Pyrex's software with embedded AutodockVina docking tool, all of the identified compounds were successfully docked well into the active Tankyrase, with a few changes, as published in the literature.[44] The box was placed in the middle of the co-crystallized ligands to calculate grid energy. Default docking parameters were used for each chemical to generate 9 docked conformations. The compounds and receptor were saved after a thorough analysis of ligand-receptor interactions.

Prediction of ADMEt

SwissADME submission website was used (http://www.swissadme.ch), where compounds was be submitted for chemistry-friendly medicinal characteristics, pharmacokinetics, physicochemistry, drug-likeness, and ADME, to predict ADME attributes. Individually stored and improved compounds were loaded and then transformed to smileys using the same interface's key, the calculation was performed, and the results are displayed under the structures.

RESULTS AND DISCUSSION

QSAR modelling and compounds molecular Structure

In this work, novel compounds were reported to be synthesized from[36] and.[26] Subsequently, another series of urea scaffold were designed based on the strategy of molecular hybridization technique via Williamson reaction and through to the different intermediate isocyanate reported see.[26]

Data collection

The structural effect of anticancer drugs was examined. Table 1 displays the structures of 34 anti-cancer compounds of N-naphthoyl thioureas and N-aryl-N'-benzylurea36 and[26] which mathematically IC_{50}, represented as $pIC_{50} = - Log_{10} (IC_{50})$.

Two-dimensional QSAR models were created for this investigation.

The better model was picked using statistical criteria that included the squared coefficient of correlation (R^2>0.6) observed, which is a measure of relative fit quality. Both internally and externally validations were used to examine the importance of developed models and their accuracy in prediction.[45,46] The (Q^2) high value, which represents the coefficient of correlation for the leave one out squared cross-validated, indicated the quality of the model (QSAR), as did the value of >0.3, which indicates the disparity between Q^2 and R^2.[47] As a result, model 1 was selected carefully to predict the activity of the chemical compounds.

This is due to its statistical quality and the highest (R^2) correlation coefficient of 0.9253, (R^2_{adj}) adjusted coefficient of correlation at 0.9045, (R^2_{test}) external confirmation of 0.6015, and (Q^2_{cv}) cross-validation coefficient of 0.8767. Model 1 was examined and validated using parameters parameters both externally and internally to meet the minimal condition of a reliable QSAR model as stated in previous literature Standard Validation parameters.[47] The equations shown below were generated using the DTC-QSAR program.

The model 1 parameters were chosen to calculate the anti-cancer application of the training set and test set outcomes. Table 2 displays the generated results. The correlation coefficient (R^2) was near to 1.0, indicating descriptors of model 1 are enough for a valid model (QSAR) prediction for freshly created compounds. The number of 0.9253 indicates that the model accounts for 92.53% of the residual variance, indicating that it is promising Table 3. Furthermore, the model wasn't over-fitted, as evidenced by the closeness of Q^2_{cv} to internal R^2.

This was shown to be consistent with previously reported modelling studies.[37]

Model 1

$$pIC_{50} = 1.011(+/-1.0635) + 0.6242(+/-0.1838)\ piPC7 + 0.09(+/-0.0354)\ L2u + 0.0704(+/-0.0275)\ RDF25p - 0.0719(+/-0.0262)\ RDF25m$$

Model 2

$$pIC_{50} = 4.3449(+/-0.2966) + 0.0833(+/-0.0097)\ RDF140v + 1.4881(+/-0.4994)\ GATS1s + 0.0534(+/-0.0079)\ RDF125m - 0.0002(+/-0.0001)\ ATSC5m + 0.2386(+/-0.0169)\ gmin$$

Model 3

$$pIC_{50} = 5.0712(+/-0.086) + 0.0281(+/-0.006)\ RDF140e + 0.2771(+/-0.0375)\ gmin - 0.0553(+/-0.0296)\ minHBint6 - 0.0149(+/-0.0041)\ WNSA\text{-}3$$

Table 2 displayed a visualization of the pIC_{50} experimental and sets validation versus the anticipated pIC_{50} for the anti-Cancer values model. The computed activities (pIC_{50}) values normalized matched with the test set, indicating that the model was error-free.[37]

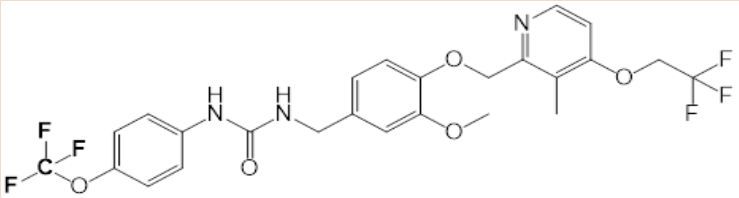
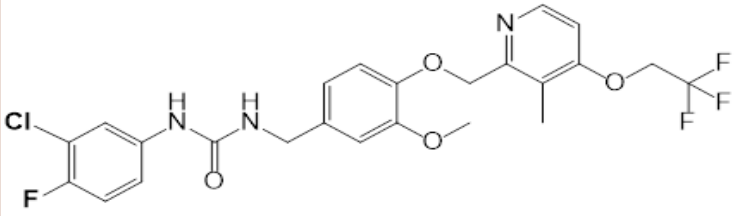
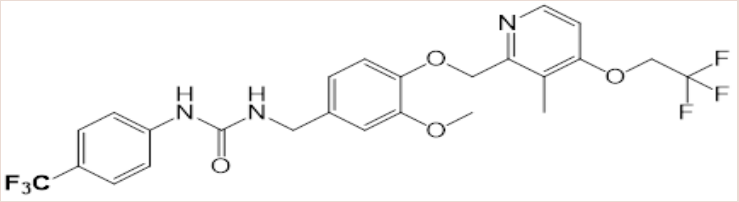
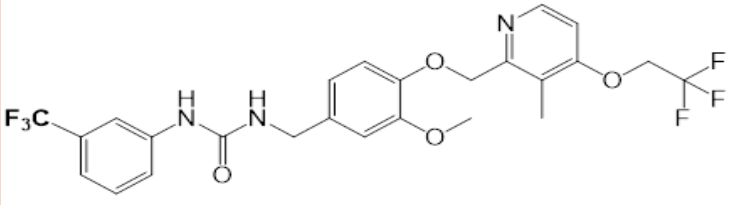
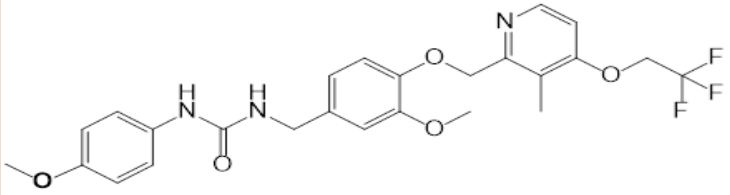
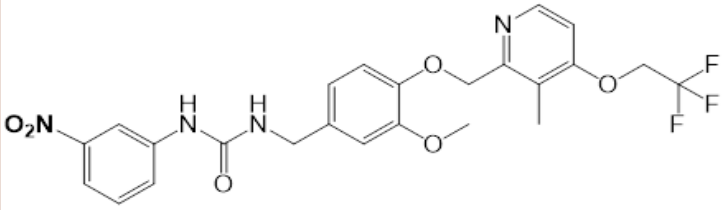
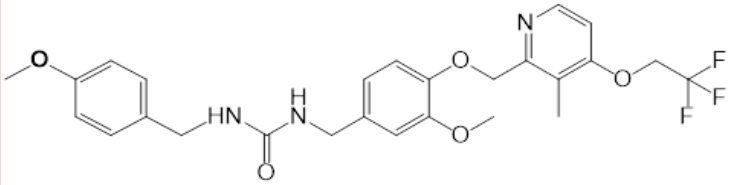
Y – Randomisation

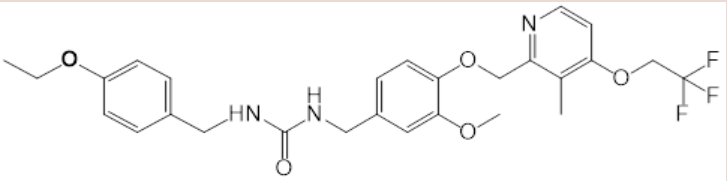
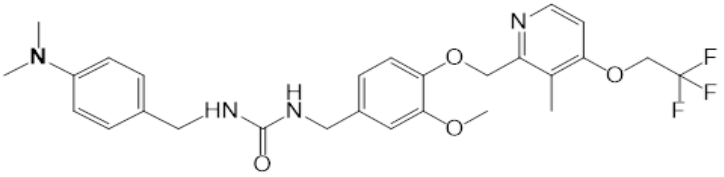
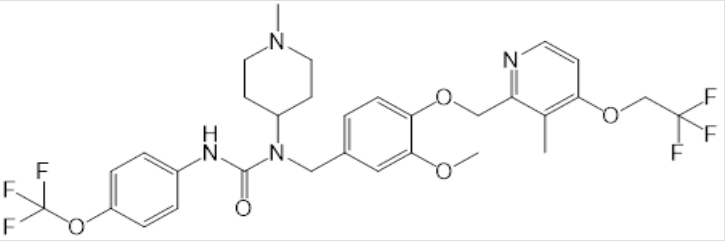
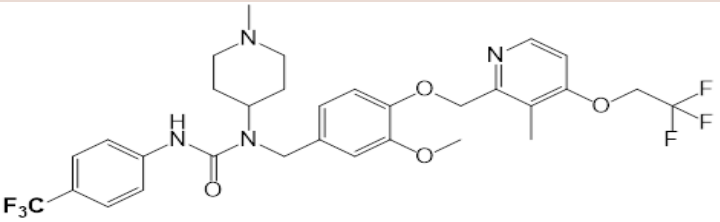
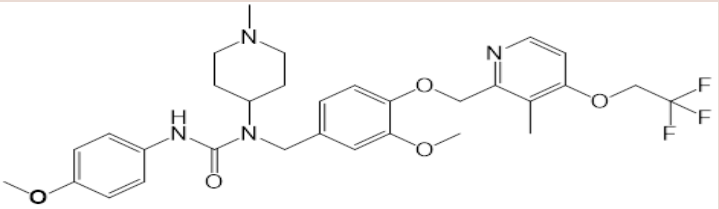
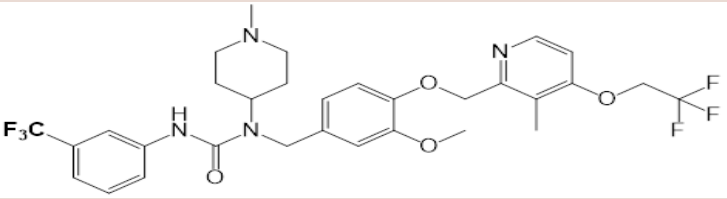
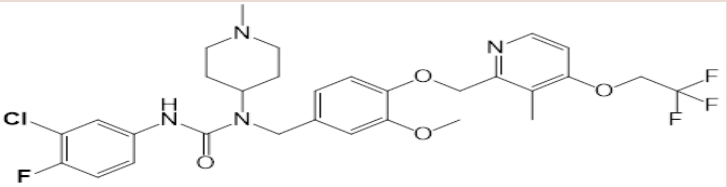
A randomization test was utilized to investigate the robustness of the established models[48] in order to substantially confirm descriptors association and the activities are not by coincidence. The observations are mixed 10 times at randomly to achieve this, with the pIC_{50} result column varying at random however the description columns staying constant. As a result, 10 models with R^2_{cv} and an R^2 average of 0.6763 and 0.21678, respectively, were adopted in Table 4. This output demonstrates that the obtained models were not acquired by coincidence.

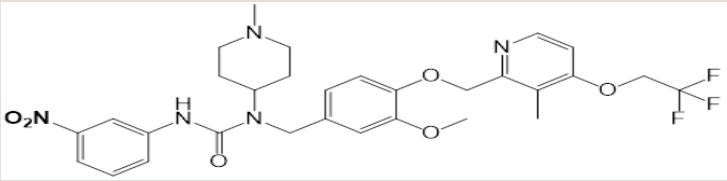
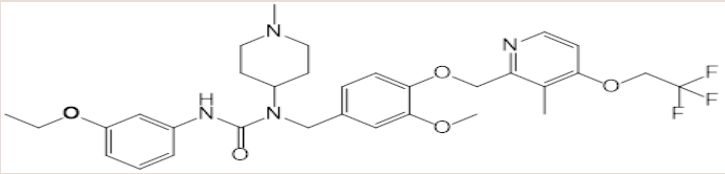
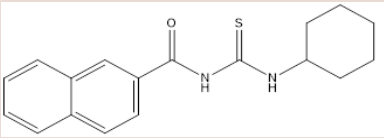
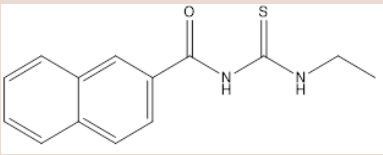
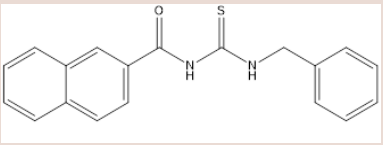
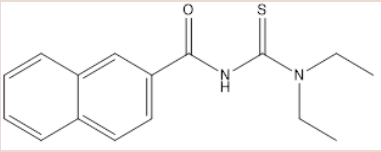
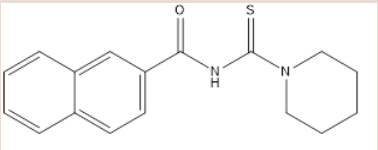
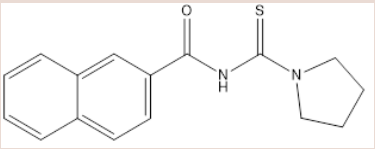
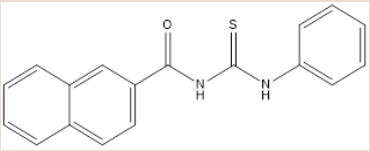
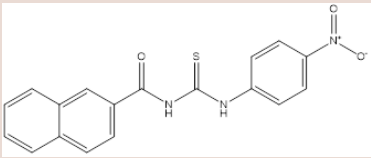
Mean Effect (ME) and Variance Inflation Factor (VIF)

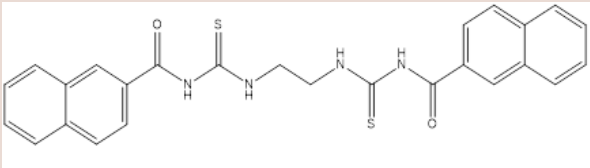
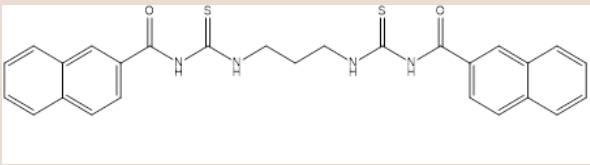
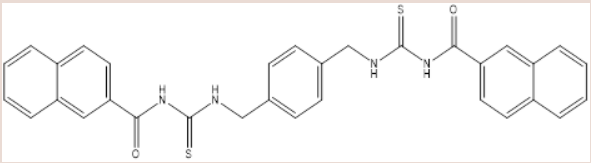
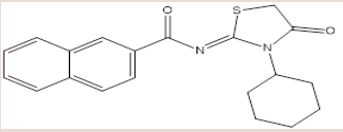
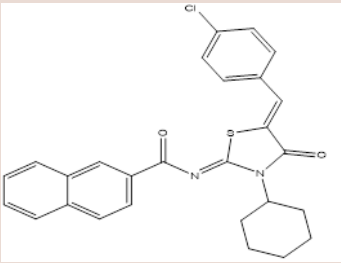
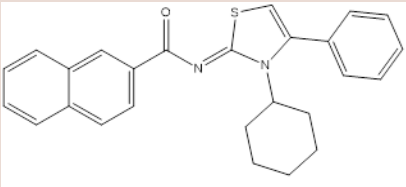
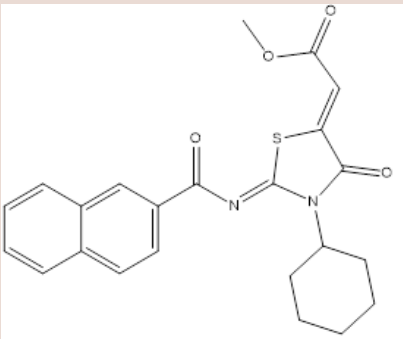
The generated values of 0.407848, 1.008611, 0.393494, -0.06787 and -0.06787 for RDF140v, GATS1s, RDF125m, and ATSC5m, respectively, are less than 10, indicating that a reasonable model was statistically indicated and adequately orthogonal descriptors were formed Table 4.[37,49] Each model description generates an impact value to evaluate their roles and contributions. It also provides crucial descriptors model information developed for determining chemical activity and the model's power.[38] According to the results, RDF125m is the most significant descriptor with the best potential mean effect value; hence, the relevance of the

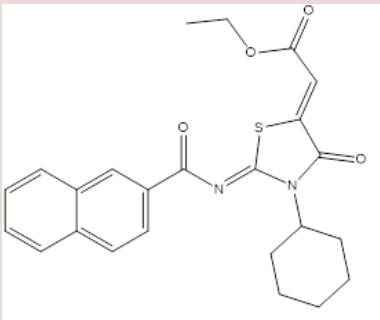
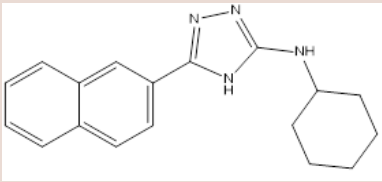
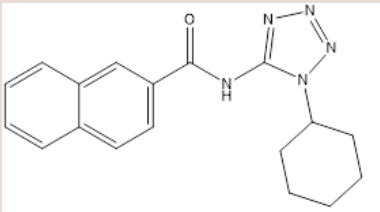
Table 1: 2D structure of Dataset Compound, names, IC_{50} and their pIC_{50}

Sl. No	Structure of Compounds	IC_{50}	pIC_{50}
C1		7.25	5.139662
C2		9.33	5.030118
C3		11.38	4.943858
C4		9.44	5.025028
C5		28.81	4.540457
C6		13.35	4.874519
C7		7.84	5.105684

Sl. No	Structure of Compounds	IC_{50}	pIC_{50}
C8		23.31	4.632458
C9		10.88	4.963371
C10		4.41	5.355561
C11		2.9	5.537602
C12		12.9	4.88941
C13		2.56	5.59176
C14		5.82	5.235077

Sl. No	Structure of Compounds	IC_{50}	pIC_{50}
C15		5.18	5.28567
C16		2.93	5.533132
C17		4.62	5.335358
C18		4.14	5.383
C19		4.55	5.341989
C20		5.07	5.294992
C21		9.12	5.040005
C22		6.22	5.20621
C23		5.77	5.238824
C24		6.09	5.215383

Sl. No	Structure of Compounds	IC_{50}	pIC_{50}
C25		1.31	5.882729
C26		1.23	5.910095
C27		1.35	5.869666
C28		9.12	5.040005
C29		9.28	5.032452
C30		19.22	4.716247
C31		12.22	4.912929

Sl. No	Structure of Compounds	IC_{50}	pIC_{50}
C32		11.38	4.943858
C33		2.53	5.596879
C34		3.06	5.514279

compounds' pIC_{50} values will be exceptional. The descriptors are listed in the order of their influence on pIC_{50} values GATS1s> RDF140v> RDF125m>ATSC5m. As a result, the mean impact value with the greatest positive value of RDF140v and lowest negative value of GATS1s suggested that descriptors may enhance the compound's anti-cancer activity, and anticipated activity is proportional inversely to the values.

Application Domain Chemistry

The determination of applicability domains is expressly requested in validation methods, making it crucial at the OECD level.[50] A model created with this set does not cover the full chemical space, yet compound predictions are considered valid. The applicability domain is the chemical space region comprising the chemicals in the model learning set. In this study, the "Leverage" technique is utilized to examine the applicability domain. Because the "Leverages" obtained are greater or less than the value in the threshold, there are no outliers and all data have normalized residues. This model is highly predictive, and it generates very exciting data and details on the nature of the molecules (Figure 1), which might be utilized to drive future anti-cancer drug research. C18 and C19 Compounds, on the contrary hand, were discovered to be relatively mildly active, which could be due to the existence of both carbonylinamine (HN–C=O) and thiolidine (HN–C=S) groups present in the backbone structure of the compounds.

Molecular structure Docking Studies

Interactions of Compounds C18, C25, and C33 with Tankyrase (TNKS2) in the Crystal Structure. (PDB code 6KRO) indicate the ligands moieties binding site. Docking experiments revealed that drugs attached to the receptor's active site with a favorable association and strong affinity affinity. Table 6 and Figures 2-4. Compound C18: careful observation presents Six hydrophobic pi – pi stacked interactions of A: TYR1009 – N: 6KRO and N: 6KRO – N: 6KRO from Pi – orbitals to Pi – orbitals interactions. Additional hydrophobic pi – pi T – shaped interaction of A: TRP1006 – N:6KRO from Pi – orbitals to Pi – orbitals. Also conventional three hydrogen bonds of A: VAL1000:HN – N:

Figure 1: Parity Graph of predicted training activity versus actual activity.

Table 2: Training and test dataset residual values.

Data Set	pIC_{50}	Predicted pIC_{50}	Residual
C1	5.139662	5.095476	0.044186
C2	5.030118	4.893803	0.136315
C3	4.943858	4.897403	0.046455
C4	5.025028	5.079503	-0.05448
C5	4.540457	4.68693	-0.14647
C6	4.874519	4.840572	0.033947
C7	5.105684	5.013496	0.092188
C8	4.632458	4.793146	-0.16069
C9	4.963371	4.927634	0.035737
C10	5.355561	5.602023	-0.24646
C11	5.537602	5.510439	0.027163
C12	4.88941	4.860784	0.028627
C13	5.59176	5.460565	0.131195
C14	5.235077	5.264861	-0.02978
C15	5.28567	5.206753	0.078917
C16	5.533132	5.501856	0.031276
C17	5.335358	5.226108	0.10925
C18	5.383	5.278277	0.104723
C19	5.341989	5.278277	0.063712
C20	5.294992	5.285564	0.009428
C21	5.040005	5.124086	-0.08408
C22	5.20621	5.188652	0.017557
C23	5.238824	5.352009	-0.11319
C24	5.215383	5.231038	-0.01566
C25	5.882729	5.911193	-0.02846
C26	5.910095	5.93054	-0.02045
C27	5.869666	5.857222	0.012444
C28	5.040005	4.740321	0.299684
C29	5.032452	5.055163	-0.02271
C30	4.716247	5.042231	-0.32598
C31	4.912929	4.683267	0.229662
C32	4.943858	4.723065	0.220792
C33	5.596879	5.569356	0.027523
C34	5.514279	4.9731	0.541179

Table 3: Validations Values of QSAR Model.

Validation parameter	Model 1	Model 2	Model 3
R^2	0.9253	0.8345	0.7951
R^2_{adj}	0.9045	0.7996	0.7496
Q^2_{CV}	0.8767	0.725	0.7003
$R^2 - Q^2_{CV}$	0.068375	0.1083	0.0717
$N_{ext.\ test\ set}$	10	10	11
R^2_{test}	0.6015	-	-
$P_{95\%}$	0.004663	-	-

Table 4: Y- Randomization.

Model	R	R^2	Q^2
ORIGINAL	0.961919	0.925287846	0.876677
RANDOMIZED	0.410683	0.168660573	-0.5817
RANDOMIZED	0.587991	0.345733489	-0.8104
RANDOMIZED	0.553723	0.306609311	-1.08242
RANDOMIZED	0.433338	0.187781439	-0.94205
RANDOMIZED	0.557297	0.310580032	-0.16716
RANDOMIZED	0.267849	0.071742876	-0.49356
RANDOMIZED	0.463739	0.215053594	-0.34992
RANDOMIZED	0.508031	0.258095924	-0.2471
RANDOMIZED	0.607019	0.36847242	0.023866
RANDOMIZED	0.416926	-0.93739	0.645698
Average Randomized model			
R^2	(Original Model)	0.92528	
Average R^2	(50 Random Models)	0.21678	
Average Q^2	(50 Random Models)	-0.6763	

Table 5: QSAR model Descriptors' VIF Mean Effect (ME) and correlation matrix

Descriptors	Correlation matrix					*ME*	*VIF*
	RDF140v	GATS1s	RDF125m	ATSC5m	gmin		
RDF140v	1					2.351027	0.407848
GATS1s	-0.03504	1				1.040534	1.008611
RDF125m	0.637556	0.009448	1			3.070099	0.393494
ATSC5m	0.577613	0.063015	0.605345	1		1.99204	-0.06787
gmin	-0.64831	-0.07556	-0.72458	-0.34178	1	2.754561	-0.74208

Table 6: Major important interactions between the Tankyrase (TNKS2) receptor and Ligands.

Compounds binding interactions	C18	C25	C33
Hydrophobic pi – pi stacked interactions.	A:TYR1009 - N:6KRO and N:6KRO - N:6KRO	A:TRP1006 - N:6KRO	-
Hydrophobic pi – pi T – shaped interactions.	A:TRP1006 - N:6KRO		-
Hydrophobic Pi – Sigma interactions.	-	-	A:ILE997:CG1 - N:6KRO and A:ILE997:CG2 - N:6KRO
Hydrophobic alkyl interactions.	-	A:VAL1000 - N:6KRO	-
Hydrophobic Pi – Alkyl interactions.	-	N:6KRO - A:ARG1100	N:6KRO - A:ILE994 and N:6KRO - A:MET972
Conventional hydrogen bonds.	A:VAL1000:HN - N:6KRO:O, A:ASN1002:HN - N:6KRO:O and A:ARG1100:HH22 - N:6KRO:S	N:6KRO:H - A:PHE1098:O	A:PHE1098:HN - N:6KRO:O and N:6KRO:H - A:GLN998:O
Electrostatic pi – cation interaction.	-	A:LYS999:NZ - N:6KRO	-

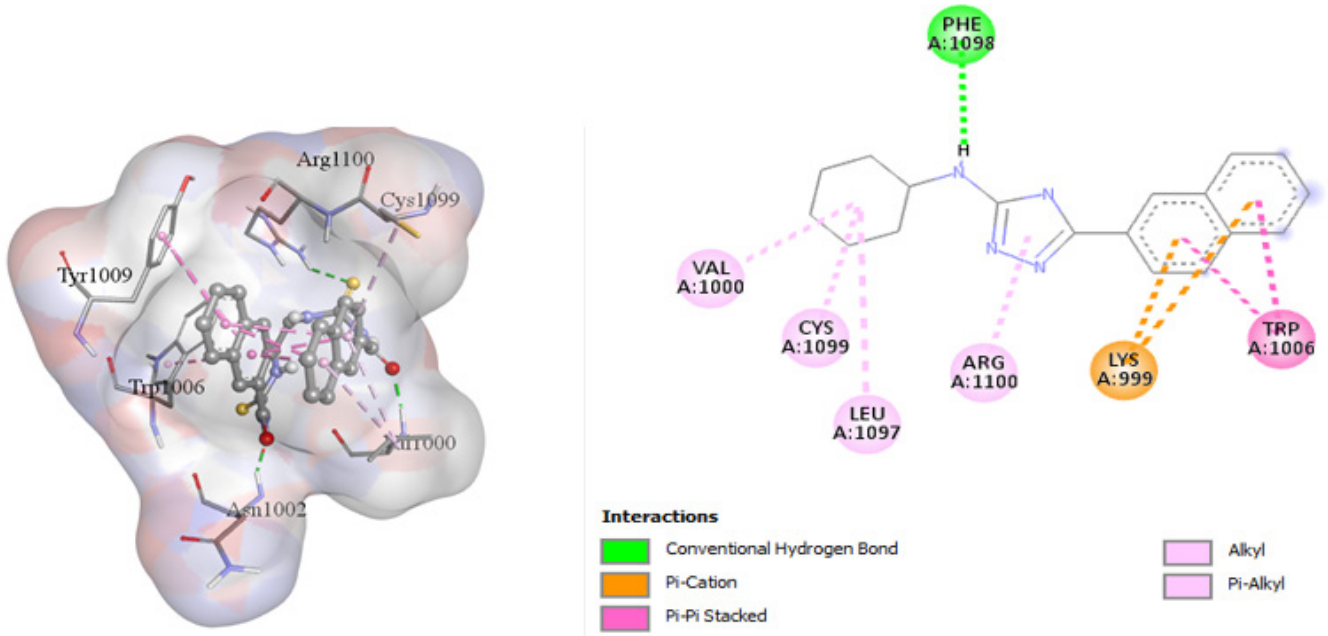


Figure 2: Representation of compound C18 interaction with Tankyrases enzyme (pdb code: 6KRO).

6KRO:O, A:ASN1002:HN – N: 6KRO:O, and A:ARG1100:HH22 – N: 6KRO:S from H – Donor to H – Acceptor Figure 2.

Compound C25: observed four hydrophobic pi – pi stacked interactions of A:TRP1006 – N:6KRO from Pi – orbitals to Pi – orbitals. Three hydrophobic alkyl interactions of A:VAL1000 – N:6KRO from Alkyl – Alkyl. Additional one hydrophobic Pi – Alkyl interactions of N:6KRO – A:ARG1100 from Pi – orbital – Alkyl. Two electrostatic pi – cation interactions of A:LYS999:NZ – N:6KRO from Positive – Pi – orbitals. One conventional hydrogen bond of N:6KRO:H – A:PHE1098:O from H – Donor to H – Acceptor (Figure 3).

Compound C33: indicated two hydrophobic Pi – Sigma bonds of A:ILE997:CG1 – N:6KRO and A:ILE997:CG2 – N:6KRO from CH to Pi – orbitals. Additional two hydrophobic Pi – alkyl interactions of N:6KRO – A:ILE994 and N:6KRO – A:MET972 from Pi – orbitals to Alkyl. Also two conventional hydrogen bonds of A:PHE1098:HN – N:6KRO:O and N:6KRO:H – A:GLN998:O from H – Donor to H – acceptor (Figure 4).

Figures 2-4 showed the three-dimensional (3D) and two-dimensional (2D) structures of docked molecules and interactions. Furthermore, the compounds have high overall clearance values are within the range of permissible drug in the body.

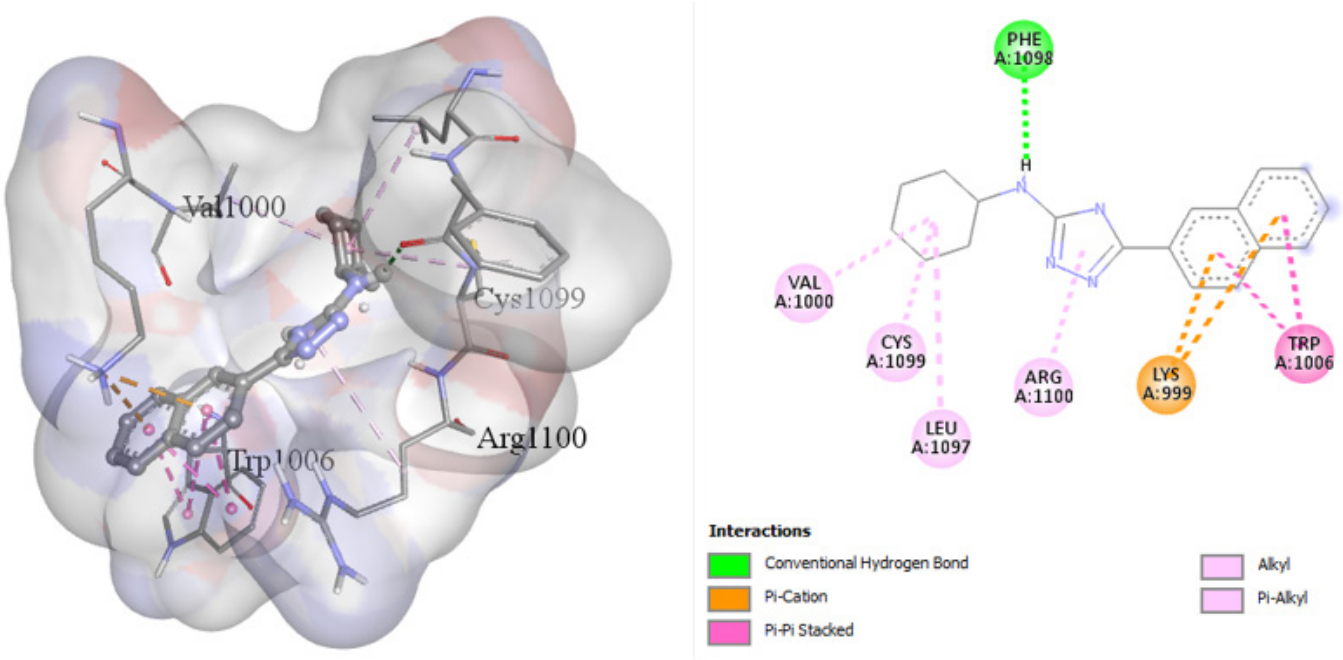


Figure 3: Representation of compound C25 interaction with Tankyrases enzyme (pdb code: 6KRO).

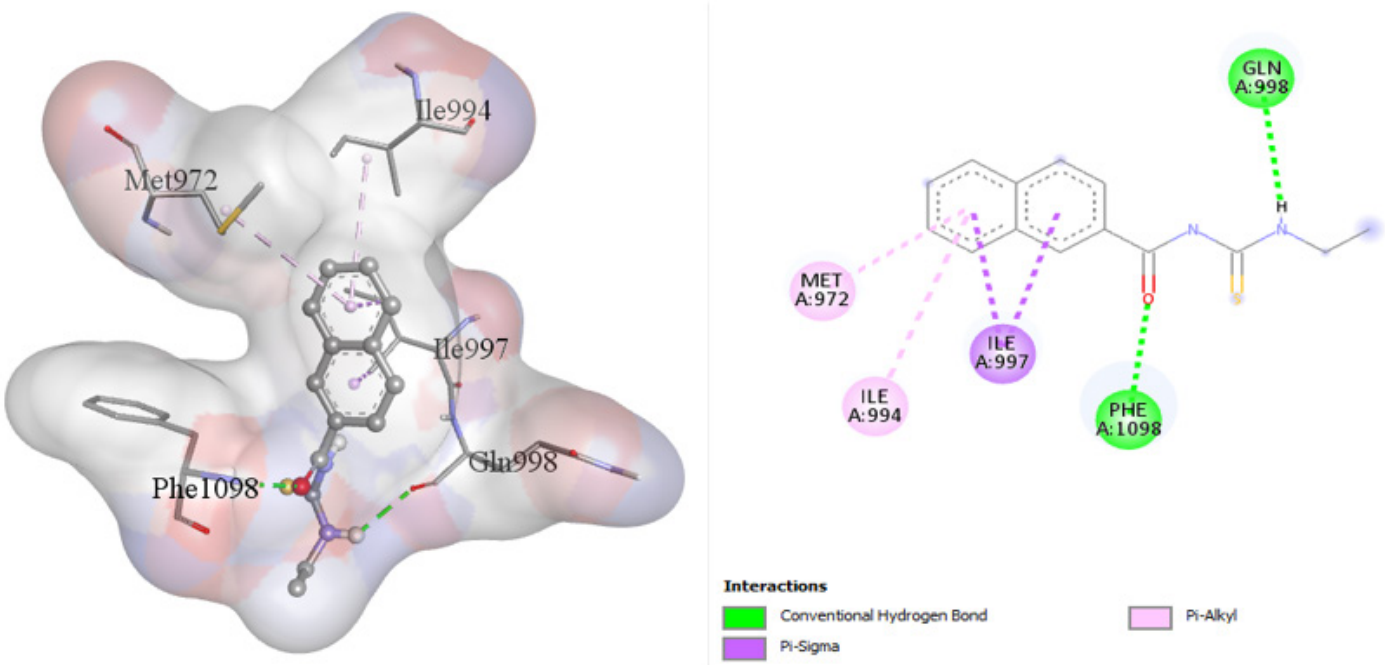


Figure 4: Representation of compound C33 interaction with Tankyrases enzyme (pdb code: 6KRO).

ADMET properties

The value of Egan and Lipinski's guidelines for the selected chemicals was calculated using the criteria employed in the published literature.[33] As a result, the projected values of the chosen compounds C13, C16, C18, C25, C26, C27, C33, and C34 are in conformity with properties of drug-likeness (Figure 5 and Table 7). Compounds C13, C16, C26, and C27 violated one of Egan and Lipinski's rules' parameters for having molecular weights more than 500. Similarly, the compounds pro'osed ma'

Table 7: Predicted Lipinski's and Egan's rule based on selected compounds.

Compound	MW	nHBD	nHBA	TPSA	WLOGP	Lipinski violations	Egan violations
C13	640.62	1	12	76.16	8.60	1	1
C16	616.67	1	10	85.39	6.83	1	1
C18	258.34	2	1	73.22	2.46	0	0
C25	486.61	4	2	146.44	3.90	0	1
C26	500.64	4	2	146.44	4.29	1	1
C27	562.7	4	2	146.44	5.30	1	1
C33	292.38	2	2	53.60	4.18	0	0
C34	321.38	1	4	72.70	3.39	0	0
Sorafenib	464.82	3	7	92.35	6.32	0	1
Doxorubucin	543.52	6	12	206.07	-0.32	3	1

Figure 5: Drug-likeness Radar map parameter diagram of C18, C33, C34, and Doxorubicin.

be viable therapeutic options since they meet Lipinski's criterion. Furthermore, LogP values smaller than 5 are an important descriptor. A logP of less than five indicates better ligand bioavailability, according to Lipinski's criteria.

Furthermore, the compounds have a high human gastrointestinal absorption. The findings demonstrate that no compounds penetrate the BBB. Subsequently, when taken orally, these chemicals are said to have low toxicity, good absorption, and to be permeable and bioavailable.[48] Since enzyme metabolism is used to explain biotransformation of drugs in the body, considering the metabolism of drugs became critical. The QSAR, docking, and drug similarity properties of the structures extend the options for developing, manufacturing, and testing new cytotoxic compounds.

CONCLUSION

To sum up, quantitative structure-activity relationship studies were performed using the multiple linear regression approach to generate models as systematic predictions to ascertain biological activity like physico-chemical characteristics, therapeutic with some adverse effects to descriptors attributes of thirty-four (34) compounds of N-naphthoyl thioureas and N-aryl-N'-benzylurea anti-colon cancer agents. The 2D QSAR models were created, and model 1 was chosen as the finest according to statistical standards (R^2>0.6).

A'high value of (Q^2), more than 0.3, suggests the quality of the QSAR model.

The R^2 value of 0.9253 indicates that the model is promising, as it accounts for 92.53% of the residual variance, and that the model has not been over-fitted, as seen by how near Q^2_{cv} is to internal R^2. The models' robustness was tested by doing a test randomization, which demonstrated that the compounds' activity and that of descriptors correlated well. This result demonstrates that the models obtained were not acquired by coincidence. The results of the docking tests revealed that compounds bind well to the active site of the receptor. Three compounds among the data set depict a good free binding energy in the receptor. These compounds pass all Lipinski´s rules of drug-likeness conducted, and can therefore be used as drug leads. Finally, model 1 was found highly effective in predicting the anticancer activity of naphthoyl thiourea and N-aryl-benzylurea derivatives prior to their synthesis.

ACKNOWLEDGEMENT

The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number ISP22-2.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ABBREVIATIONS

ADME: Absorption, distribution, metabolism, and excretion; **GFA:** Genetic function algorithm; **KSCN:** Potassium thiocyanate; **MLR:** Multiple linear regression; **NADH:** Nicotinamide adenine dinucleotide; **NCD:** Non-communicable diseases; **PDB:** Protein Data Bank files; **QSAR:** Quantitative structure-activity relationship; **WHO:** World Health Organization.

SUMMARY

Tankyrases is a poly polymerase (ADP-ribose) family enzyme, inhibiting its enzymatic processes plays a crucial part in cancer etiology. The Algorithm of Kennard-Stone was utilized to develop QSAR models of thirty-four cytotoxic compounds of N-naphthoyl thioureas and N-aryl-N'-benzylurea using the multiple linear regression approach. Three compounds C18, C25, and C33 showed favourable interaction and good binding energy and were consistent with drug-likeness parameters

REFERENCES

1. Türk S, Tok F, Erdoğan Ö, Çevik Ö, Tok TT, Koçyiğit-Kaymakçıoğlu B, *et al*. Synthesis, anticancer evaluation and *in silico* ADMET studies on urea/thiourea derivatives from gabapentin. Phosphorus Sulfur Silicon Relat Elem. 2021;196(4):382-8. Doi: 10.1080/10426507.2020.1845678.
2. Cho JH. Immunotherapy for non-small-cell lung cancer: current status and future obstacles. Immune Netw. 2017;17(6):378-91. Doi: 10.4110/in.2017.17.6.378, PMID 29302251.
3. Housman G, Byler S, Heerboth S, Lapinska K, Longacre M, Snyder N, *et al*. Drug resistance in cancer: an overview. Cancers. 2014;6(3):1769-92. Doi: 10.3390/cancers6031769, PMID 25198391.
4. Moku B, Ravindar L, Rakesh KP, Qin HL. The significance of N-methylpicolinamides in the development of anticancer therapeutics: synthesis and Structure–Activity Relationship (SAR) studies. Bioorg Chem. 2019;86:513-37. Doi: 10.1016/j.bioorg.2019.02.030, PMID 30782571.
5. Al-Harbi RAK, El-Sharief MAMS, Abbas SY. Synthesis and anticancer activity of bis-benzo [d][1, 3] dioxol-5-yl thiourea derivatives with molecular docking study. Bioorg Chem. 2019;90:103088. Doi: 10.1016/j.bioorg.2019.103088, PMID 31288134.
6. Rajput A, Dominguez San Martin I, Rose R, Beko A, Levea C, Sharratt E, *et al*. Characterization of HCT116 human colon cancer cells in an orthotopic model. J Surg Res. 2008;147(2):276-81. Doi: 10.1016/j.jss.2007.04.021, PMID 17961596.
7. Busch AM, Johnson KC, Stan RV, Sanglikar A, Ahmed Y, Dmitrovsky E, *et al*. Evidence for tankyrases as antineoplastic targets in lung cancer. BMC Cancer. 2013;13:211. Doi: 10.1186/1471-2407-13-211, PMID 23621985.
8. Arqués O, Chicote I, Puig I, Tenbaum SP, Argilés G, Dienstmann R, *et al*. Tankyrase inhibition blocks Wnt/β-catenin pathway and reverts resistance to PI3K and AKT inhibitors in the treatment of colorectal CancerTherapeutic. Clin Cancer Res. 2016;22(3):644-56. Doi: 10.1158/1078-0432.CCR-14-3081, PMID 26224873.
9. Quackenbush KS, Bagby S, Tai WM, Messersmith WA, Schreiber A, Greene J, *et al*. The novel tankyrase inhibitor (AZ1366) enhances irinotecan activity in tumors that exhibit elevated tankyrase and irinotecan resistance. Oncotarget. 2016;7(19):28273-85. Doi: 10.18632/oncotarget.8626, PMID 27070088.
10. Huang SM, Mishina YM, Liu S, Cheung A, Stegmeier F, Michaud GA, *et al*. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature. 2009;461(7264):614-20. Doi: 10.1038/nature08356, PMID 19759537.
11. Lau T, Chan E, Callow M, Waaler J, Boggs J, Blake RA, *et al*. A novel tankyrase small-molecule inhibitor suppresses APC mutation–driven colorectal tumor growth. Cancer Res. 2013;73(10):3132-44. Doi: 10.1158/0008-5472.CAN-12-4562, PMID 23539443.
12. Waaler J, Machon O, Tumova L, Dinh H, Korinek V, Wilson SR, *et al*. A novel tankyrase inhibitor decreases canonical Wnt signaling in colon carcinoma cells and reduces tumor growth in conditional APC mutant MiceJW55. Cancer Res. 2012;72(11):2822-32. Doi: 10.1158/0008-5472.CAN-11-3336, PMID 22440753.
13. Kim MK. Novel insight into the function of tankyrase. Oncol Lett. 2018;16(6):6895-902. Doi: 10.3892/ol.2018.9551, PMID 30546421.
14. Ronchetti R, Moroni G, Carotti A, Gioiello A, Camaioni E. Recent advances in urea- and thiourea-containing compounds: focus on innovative approaches in medicinal chemistry and organic synthesis. RSC Med Chem. 2021;12(7):1046-64. Doi: 10.1039/d1md00058f, PMID 34355177.
15. Chen D, Oezguen N, Urvil P, Ferguson C, Dann SM, Savidge TC. Regulation of protein-ligand binding affinity by hydrogen bond pairing. Sci Adv. 2016;2(3):e1501240. Doi: 10.1126/sciadv.1501240, PMID 27051863.
16. Okpareke OC, Henderson W, Lane JR, Okafor SN. Synthesis, structure, computational and molecular docking studies of asymmetrically di-substituted ureas containing carboxyl and phosphoryl hydrogen bond acceptor functional groups. J Mol Struct. 2020;1203:127360. Doi: 10.1016/j.molstruc.2019.127360.
17. Zarei O, Azimian F, Hamzeh-Mivehroud M, Shahbazi Mojarrad J, Hemmati S, Dastmalchi S. Design, synthesis, and biological evaluation of novel benzo [b] thiophene-diaryl urea derivatives as potential anticancer agents. Med Chem Res. 2020;29(8):1438-48. Doi: 10.1007/s00044-020-02559-8.
18. Kumar V, Chimni SS. Recent developments on thiourea based anticancer chemotherapeutics. Anti Cancer Agents Med Chem. 2015;15(2):163-75. Doi: 10.2174/1871520614666140407123526, PMID 24712324.
19. Vega-Pérez JM, Periñán I, Argandoña M, Vega-Holm M, Palo-Nieto C, Burgos-Morón E, *et al*. Isoprenyl-thiourea and urea derivatives as new farnesyl diphosphate analogues: synthesis and *in vitro* antimicrobial and cytotoxic activities. Eur J Med Chem. 2012;58:591-612. Doi: 10.1016/j.ejmech.2012.10.042, PMID 23174318.
20. Lan J, Huang L, Lou H, Chen C, Liu T, Hu S, *et al*. Design and synthesis of novel C14-urea-tetrandrine derivatives with potent anti-cancer activity. Eur J Med Chem. 2018;143:1968-80. Doi: 10.1016/j.ejmech.2017.11.007, PMID 29133049.
21. Ma LY, Pang LP, Wang B, Zhang M, Hu B, Xue DQ, *et al*. Design and synthesis of novel 1, 2, 3-triazole-pyrimidine hybrids as potential anti-cancer agents. Eur J Med Chem. 2014;86:368-80. Doi: 10.1016/j.ejmech.2014.08.010, PMID 25180925.
22. Ning C, Bi Y, He Y, Huang W, Liu L, Li Y, *et al*. Design, synthesis and biological evaluation of di-substituted cinnamic hydroxamic acids bearing urea/thiourea unit as potent histone deacetylase inhibitors. Bioorg Med Chem Lett. 2013;23(23):6432-5. Doi: 10.1016/j.bmcl.2013.09.051, PMID 24119555.
23. Li HQ, Yan T, Yang Y, Shi L, Zhou CF, Zhu HL. Synthesis and structure–activity relationships of N-benzyl-N-(X-2-hydroxybenzyl)-N'-phenylureas and thioureas as antitumor agents. Bioorg Med Chem. 2010;18(1):305-13. Doi: 10.1016/j.bmc.2009.10.054, PMID 19914837.
24. Chandrasekhar M, Syam Prasad G, Venkataramaiah C, Umapriya K, Raju CN, Seshaiah K, *et al*. *In silico* and *in vitro* antioxidant and anticancer activity profiles of urea and thiourea derivatives of 2, 3-dihydro-1 H-inden-1-amine. J Recept Signal Transduct Res. 2020;40(1):34-41. Doi: 10.1080/10799893.2019.1710848, PMID 31910703.
25. Ansari MM, Deshmukh SP, Khan R, Musaddiq M. Synthesis antimicrobial and anticancer evaluation of 1-Aryl-5-(o-methoxyphenyl)-2-S-benzyl isothiobiurets. Int J Med Chem. 2014;2014:352626. Doi: 10.1155/2014/352626, PMID 25505990.
26. Hou S, Liang S, Zhang C, Han Y, Liang J, Hu H, *et al*. Design, Synthesis and Anticancer Activity of a New Series of N-aryl-N'-[4-(odellin-2-ylmethoxy) benzyl] urea Derivatives. Molecules. 2021;26(12):3496. Doi: 10.3390/molecules26123496, PMID 34201326.
27. Tong W, Hong H, Xie Q, Shi L, Fang H, Perkins R. Assessing QSAR limitations-A regulatory perspective. Curr Comput Aid Drug Des. 2005;1(2):195-205. Doi: 10.2174/1573409053585663.
28. Dearden JC. *In silico* prediction of drug toxicity. J Comput Aid Mol Des. 2003;17(2-4):119-27. Doi: 10.1023/a:1025361621494, PMID 13677480.
29. Pinzi L, Rastelli G. Molecular docking: shifting paradigms in drug discovery. Int J Mol Sci. 2019;20(18):4331. Doi: 10.3390/ijms20184331, PMID 31487867.
30. Li Petri G, Raimondi MV, Spanò V, Holl R, Barraja P, Montalbano A. Pyrrolidine in drug discovery: a versatile scaffold for novel biologically active compounds. Top Curr Chem (Cham). 2021;379(5):34. Doi: 10.1007/s41061-021-00347-5, PMID 34373963.
31. Aldeghi M, Malhotra S, Selwood DL, Chan AWE. Two- and three-dimensional rings in drugs. Chem Biol Drug Des. 2014;83(4):450-61. Doi: 10.1111/cbdd.12260, PMID 24472495.
32. Edache EI, Uzairu A, Shallangwa GA, Mamza PA. Virtual screening, pharmacokinetics, and molecular dynamics simulations studies to identify potent approved drugs for Chlamydia trachomatis treatment. Future J Pharm Sci. 2021;7(1):1-22.
33. Limban C, Chifiriuc MC, Caproiu MT, Dumitrascu F, Ferbinteanu M, Pintilie L, *et al*. New substituted benzoylthiourea derivatives: from design to antimicrobial applications. Molecules. 2020;25(7):1478. Doi: 10.3390/molecules25071478, PMID 32218209.
34. Gonçalves IL, Da Rosa RR, Eifler-Lima VL, Merlo AA. The use of odellingne and isoxazole scaffolding in the design of novel thiourea and amide liquid-crystalline compounds. Beilstein J Org Chem. 2020;16:175-84. Doi: 10.3762/bjoc.16.20, PMID 32117474.
35. Özgeriş B. Design, synthesis, characterization, and biological evaluation of nicotinoyl thioureas as antimicrobial and antioxidant agents. J Antibiot (Tokyo). 2021;74(4):233-43. Doi: 10.1038/s41429-020-00399-7, PMID 33441970.
36. Arafa WAA, Ghoneim AA, Mourad AK. N-naphthoyl thiourea derivatives: an efficient ultrasonic-assisted synthesis, reaction, and *in vitro* anticancer evaluations. ACS Omega. 2022;7(7):6210-22. Doi: 10.1021/acsomega.1c06718, PMID 35224384.
37. Abdullahi SH, Uzairu A, Ibrahim MT, Umar AB. Chemo-informatics activity prediction, ligand based drug design, Molecular docking and pharmacokinetics studies of some series of 4, 6-diaryl-2-pyrimidinamine derivatives as anticancer agents. Bull Natl Res Cent. 2021;45:1-22.

38. Arthur DE, Uzairu A, Mamza P, Abechi SE, Shallangwa G. *In silico* modelling of quantitative structure–activity relationship of pGI50 anticancer compounds on K-562 cell line. Cogent Chem. 2018;4(1):1432520. Doi: 10.1080/23312009.2018.1432520.
39. De Bruijn WJC, Hageman JA, Araya-Cloutier C, Gruppen H, Vincken JP. QSAR of 1, 4-benzoxazin-3-one antimicrobials and their drug design perspectives. Bioorg Med Chem. 2018;26(23-24):6105-14. Doi: 10.1016/j.bmc.2018.11.016, PMID 30471830.
40. Umar BA, Uzairu A, Shallangwa GA, Sani U. QSAR odelling 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 Model Anal Health Inform Bioinformatics. 2019;8:1-10.
41. Idris MO, Abechi SE, Shallangwa GA. Computational evaluation of some compounds as potential anti-breast cancer agents. Future J Pharm Sci. 2021;7(1):1-5.
42. Eriksson L, Jaworska J, Worth AP, Cronin MT, McDowell RM, Gramatica P. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. Environ Health Perspect. 2003;111(10):1361-75. Doi: 10.1289/ehp.5758, PMID 12896860.
43. Shirai F, Tsumura T, Yashiroda Y, Yuki H, Niwa H, Sato S, *et al*. Discovery of novel spiroindoline derivatives as selective tankyrase inhibitors. J Med Chem. 2019;62(7):3407-27. Doi: 10.1021/acs.jmedchem.8b01888, PMID 30883102.
44. Shirai F, Mizutani A, Yashiroda Y, Tsumura T, Kano Y, Muramatsu Y, *et al*. Design and discovery of an orally efficacious Spiroindolinone-based tankyrase inhibitor for the treatment of colon Cancer. J Med Chem. 2020;63(8):4183-204. Doi: 10.1021/acs.jmedchem.0c00045, PMID 32202790.
45. El Aissouq A, Toufik H, Stitou M, Ouammou A, Lamchouri F. *In silico* design of novel tetra-substituted pyridinylimidazoles derivatives as c-Jun N-terminal kinase-3 inhibitors, using 2D/3D-QSAR studies, molecular docking and ADMET prediction. Int J Pept Res Ther. 2020;26(3):1335-51. Doi: 10.1007/s10989-019-09939-8.
46. El Aissouq A, Toufik H, Lamchouri F, Stitou M, Ouammou A. QSAR study of isonicotinamides derivatives as Alzheimr's disease inhibitors using PLS-R and ANN methods International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS). IEEE Publications. 2019;2019:1-7. Doi: 10.1109/ISACS48493.2019.9068919.
47. Abdulfatai U, Uzairu A, Shallangwa GA, Uba S. *In-silico* molecular design and dynamic binding energy assessments of some multifunctional lubricating oil additives. S Afr J Chem Eng. 2020;32:78-84. Doi: 10.1016/j.sajce.2019.11.002.
48. Rücker C, Rücker G, Meringer M. y-randomization and its variants in QSPR/QSAR. J Chem Inf Model. 2007;47(6):2345-57. Doi: 10.1021/ci700157b, PMID 17880194.
49. Myers R. Classical and modern regression with application, Duxbury. Pacific Grove, CA; 1990. P. 93950.
50. Ruiz IL, Gómez-Nieto MÁ. Study of the applicability domain of the QSAR classification models by means of the rivality and modelability indexes. Molecules. 2018;23(11):2756. Doi: 10.3390/molecules23112756, PMID 30356020.

Cite this article: Makeen HA, Albratty M. 2D-QSAR, Molecular Docking, and *in silico* Pharmacokinetics Analysis on N-substituted Urea and Thiourea Derivatives as Tankyrase Inhibitors for Implication in Cancer. Indian J of Pharmaceutical Education and Research. 2023;57(3):838-53.