

RESEARCH

Open Access



# Theoretical activity prediction, structure-based design, molecular docking and pharmacokinetic studies of some maleimides against *Leishmania donovani* for the treatment of leishmaniasis

Fabian Audu Ugbe\* , Gideon Adamu Shallangwa, Adamu Uzairu and Ibrahim Abdulkadir

## Abstract

**Background:** Leishmaniasis is a neglected tropical disease caused by a group of protozoan of the genus *Leishmania* and transmitted to humans majorly through the bite of the female sand fly. It is prevalent in the tropical regions of the world especially in Africa and estimated to affect a population of about 12 million people annually. This theoretical study was therefore conducted in support of the search for more effective drug candidates for the treatment of leishmaniasis. This study focuses on the in silico activity prediction of twenty-eight (28) maleimides, structure-based design, molecular docking study and pharmacokinetics analysis of the newly designed maleimides. All the studied compounds were drawn using ChemDraw Ultra and optimized by the density functional theory (DFT) approach using B3LYP with 6-31G/ basis set.

**Results:** The built QSAR model was found to satisfy the requirement of both internal and external validation tests for an acceptable QSAR model with  $R^2 = 0.801$ ,  $R^2_{adj} = 0.748$ ,  $Q^2_{cv} = 0.710$ ,  $R^2_{test} = 0.892$  and  $cR_p^2 = 0.664$  and has shown excellent prediction of the studied compounds. Among the five (5) protein receptors utilized for the virtual docking screening, pyridoxal kinase (PdxK) receptor (Pdb id = 6k91) showed the strongest binding interactions with compounds **14**, **21** and **24** with the highest binding affinities of  $-7.7$ ,  $-7.7$  and  $-7.8$  kcal/mol, respectively. The selected templates (**14**, **21** and **24**) were used to design twelve (12) new compounds (N1–N12) with higher docking scores than the templates. **N7** (affinity =  $-8.9$  kcal/mol) and **N12** ( $-8.5$  kcal/mol) showed higher binding scores than the reference drug pentamidine ( $-8.10$  kcal/mol), while **N3** and **N7–N12** showed higher predicted  $pIC_{50}$  than the templates. Also, the pharmacokinetics properties prediction revealed that the newly designed compounds, obeyed the Lipinski's rule for oral bio-availability, showed high human intestinal absorption (HIA), low synthetic accessibility score, CNS and BBB permeability and were pharmacologically active.

**Conclusions:** The activities of the various maleimides were predicted excellently by the built QSAR model. Based on the pharmacokinetics and molecular docking studies therefore, the newly designed compounds are suggested for further practical evaluation and/or validation as potential drug candidates for the treatment of leishmaniasis.

**Keywords:** DFT, Leishmaniasis, *L. Donovanii*, Molecular docking, Pharmacokinetics, QSAR

\*Correspondence: [ugbefabianaudu@gmail.com](mailto:ugbefabianaudu@gmail.com)  
Department of Chemistry, Faculty of Physical Sciences, Ahmadu Bello University, P.M.B. 1044, Zaria, Kaduna State, Nigeria

## Background

Leishmaniasis is among those diseases which are classified as neglected tropical diseases. It was reported as the second most prevalent parasite infection after malaria (Tonelli et al. 2018). Leishmaniasis is known to affect a population of about 350 million people majorly living in the tropical impoverished regions of the world, with approximately 12 million people been infected annually (Baquedano et al. 2016). It is caused by a group of protozoan parasites of the genus *leishmania*, which are responsible for the three common clinical types: cutaneous, mucocutaneous and visceral leishmaniasis (VL) among which VL is the most fatal if left untreated (Keurulainen et al. 2018). VL is majorly caused by *Leishmania Donovanii* (*L. donovani*) and *L. infantum*, inherited through the bite of the female sand fly (Abdelhameed 2018). The challenge remains that only few drugs are available for the treatment of leishmaniasis including Pentostam (contains heavy metal antimony), pentamidine, amphotericin B, paromomycin and miltefosine with varying shortcomings such as high cost, toxicity and less efficacy among others (Fan et al. 2018). For example, fever, chill, nephrotoxicity, hypokalemia and myocarditis are the major adverse effects associated with amphotericin B (Ghorbani and Farhoudi 2018). Miltefosine has been linked to contraindication in pregnancy and mandatory contraception for women in child bearing age (Ghorbani and Farhoudi 2018). The major safety concern associated with pentamidine is induction of insulin-dependent diabetes mellitus, while ototoxicity and nephrotoxicity are associated with paromomycin (Seifert 2011). Also, the parasites which are increasingly becoming drug-resistant pose a great threat to the treatment of leishmanial infections (Fan et al. 2018), and more so that it has received only a little global attention compared to other infections such as malaria, tuberculosis and cancer. Therefore, the development of new medicines with clinical attributes that overcome all of these drawbacks is highly necessary.

In this study, a computer-aided drug design approach was utilized to propose some set of new compounds for further evaluation as leishmania inhibitors. In the process of new drug development, computational tools play a critical role because it saves time and cost and has proven to be more effective than the crude traditional methods (Lawal et al. 2021). Notable computational tools include quantitative structure activity relationship (QSAR), molecular docking, molecular dynamics, pharmacokinetics and homology modeling among others (Adeniji et al. 2019; Ibrahim et al. 2020, 2021). The knowledge of QSAR helps in establishing a relationship between various molecular structures of molecules and their experimental activities (Adeniji et al. 2019). Molecular docking simulation is a computer-aided virtual screening method

which probes the binding of ligands in the active sites of the protein target using a valid docking tool (Ibrahim et al. 2020). Pharmacokinetics analysis is important in the preclinical study of new drug compounds in order to ascertain how such drug compounds affect the living organism when administered. Some of the most important pharmacokinetics properties to be determined during preclinical testing include absorption, distribution, metabolism, excretion and toxicity (ADMET) (Lawal et al. 2021; Ibrahim et al. 2021). Physicochemical properties such as molecular weight, topological polar surface area (TPSA), lipophilicity, water solubility, hydrogen bond donors, and hydrogen bond acceptors are necessary to predict drug's likelihood of being orally bioavailable (Lipinski et al. 2001). The choices of molecules for oral bioavailability have been guided by several rules such as the Lipinski's 'rule of 5' (RO5), Veber rule, Ghose rule, Egan and Muegge (Sun et al. 2020).

Some therapeutic protein targets of *Leishmania* spp available in the protein data bank for docking with drug ligands include Ribokinase from *L. donovani* (5ZWY) (Gatreddi et al. 2019), *L. major* mitochondrial ribosome (7ANE) (Soufari et al. 2020), pyridoxal kinase from *L. donovani* in complex with ADP and pyridoxine (6K91) (Are et al. 2020) and *L. donovani* UMP synthase (3QW4) (French et al. 2011). Pyridoxal kinase (PdxK) from *L. donovani* is an interesting drug target; it was reported to catalyze the phosphorylation of the 5' hydroxyl group of pyridoxal to form pyridoxal-5'-phosphate, the biologically active form of vitamin B6 (Are et al. 2020). PdxK was in a recent study shown to be significant for parasite growth and key to host infection (Kumar et al. 2018). Known anti-malaria medicines like chloroquine and primaquine had earlier shown inhibition against PdxK (Kimura et al. 2014). Therefore, PdxK represents a promising target for designing novel anti-leishmanial agents. Structure-based drug design (SBDD) otherwise referred to as the direct drug design is a critical part of industrial drug discovery processes, and a useful tool in academic researches. This is because it deals with the knowledge of the 3D structures of the target enzymes and the related diseases at the molecular level (Batool et al. 2019). The method comprises of steps such as obtaining useful information on the 3D structure of the target protein determined via X-ray crystallography; NMR spectroscopy or homology modeling; retrieval and preparation of such enzyme target; and the actual design of novel compounds (Abdullahi et al. 2022).

Maleimides have been reported as antimicrobial compounds (Chen et al. 2015; Li et al. 2012; Shen et al. 2013) and also have shown excellent enzyme inhibitory activities (Silvia and Maria 2005; Slavica et al. 2007). Additionally, maleimides had earlier been described as

anti-inflammatory (Nara et al. 2010), anti-cancer (Khan et al. 2004) and anti-anxiety (Jerzy 2003). In order to exploit the anti-leishmanial effect of maleimides therefore, Fan et al. (2018) synthesized a series of maleimides and reported their biological activities against *L. donovani*. Consequently, this study was undertaken to develop a QSAR model for predicting with precision the activities of some maleimides as potent anti-leishmanial agents, conduct a virtual docking screening with different protein receptors, carry out a structure-based design of new maleimides and subjecting same to pharmacokinetics and molecular docking studies in order to evaluate their drug-likeness properties and binding interaction pattern, respectively.

## Methods

### Data collection

A series of 28 maleimides with reported anti-leishmanial activities ( $IC_{50}$  in  $\mu\text{g/mL}$ ) were obtained from the literature (Fan et al. 2018). The bioactivities ( $IC_{50}$ ) were first converted from g/L to Molar (M) using Eq. (1) and thereafter to logarithmic scale ( $pIC_{50}$ ) using Eq. (2) (Wang et al. 2020). The molecular structures of the various maleimides with their observed  $IC_{50}$  and  $pIC_{50}$  values are shown in Table 1.

$$IC_{50}(\text{molL}^{-1}) = \frac{IC_{50}(\text{gL}^{-1})}{\text{Molar mass}(\text{gmol}^{-1})}. \quad (1)$$

$$pIC_{50} = -\log_{10}(IC_{50} \times 10^{-9}) \quad (2)$$

### Molecular geometry optimization

The two-dimensional (2-D) structures of the various compounds were drawn using ChemDraw Ultra (version 12.0.2) and saved as MDL mol file format. The resulting files were then fed separately onto the Spartan'14 (version 1.1.4) graphical user interface in three-dimensional (3-D) structural form. The 3-D structures were first subjected to energy minimization and molecular mechanics force field (MMFF) in order to minimize their chemical structures and to remove tension energy of the molecules' conformation. The main optimization process was then conducted using density functional theory (DFT) with Becke's three-parameter read-Yang-Parr hybrid (B3LYP) option and utilizing the 6-31G basis set. The thoroughly optimized structures were then saved in SD file and PDB formats for use in descriptor calculation and molecular docking, respectively (Wang et al. 2020; Li et al. 2004).

### Generation of molecular descriptors

The resulting data in SD file format obtained earlier from the optimization process were imported into the Pharmaceutical Data Exploration Laboratory (PaDEL)-Descriptor software (version 2.20) in order to calculate the molecular descriptors for all twenty-eight (28) compounds (Lawal et al. 2021).

### Dataset pretreatment and division into training and test sets

The Drug Theoretical and Chem-informatics Laboratory (DTC Lab)-based software GUI 1.2 was used to remove non-informative descriptors from the generated descriptor pool (Adeniji et al. 2020). The pretreated data were then divided into the modeling train set data and external evaluation test set data in the ratio of 70:30, respectively, with the help of DTC Lab-derived software which utilizes the Kennard Stone method for dataset division (Kennard and Stone 1969). The splitting of dataset into training and test sets was based on the closeness of the representative points of the test set to the representative points of the training set in the multidimensional descriptor space (Ugbe et al. 2021).

### Model building

The genetic function approximation (GFA) as a statistical technique in the Material Studio software (version 8.0) was used to generate the models based on multi-linear regression (MLR) approach. GFA was used to obtain the optimum descriptor combinations constituting the QSAR models, while MLR helps to establish the relationship between the biological activities,  $pIC_{50}$  (dependent variable) and the molecular descriptors (independent variables) (Arthur et al. 2020). The multi-linear regression equation assumes the following form (Eq. 3) (Adawara et al. 2020):

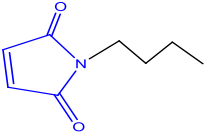
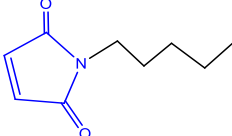
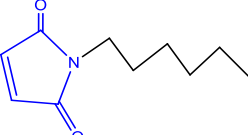


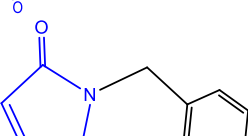
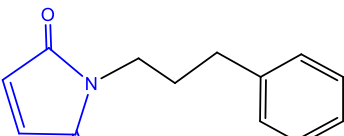
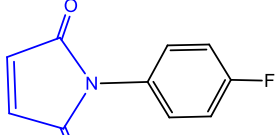
$$Y = k_1x_1 + k_2x_2 + k_3x_3 + \dots C \quad (3)$$

where  $Y$  represents the dependent variable; ' $k$ 's and ' $x$ 's represent, respectively, regression coefficients and independent variables; ' $C$ ' equals intercept or regression constant.

### Assessment of model quality

Internal validation assessment of the built model was carried out on Material studio by GFA approach using the Friedman formula, correlation coefficient ( $R^2$ ) and cross-validation coefficient ( $Q^2_{cv}$ ), and also by the Y-randomization test in order to show how well the model predict the activity values of the model building compounds (Adeniji et al. 2018; Adawara et al. 2020). The model's predictive power was also assessed externally to show if the model could predict the activity values of the test

**Table 1** Molecular structures of maleimides and their anti-leishmanial activities ( $IC_{50}$  and  $pIC_{50}$ )

| Comp ID | Molecular structure   | $IC_{50}$ ( $\mu\text{g/mL}$ ) | Mm (g/mol) | $IC_{50}$ ( $\mu\text{M}$ ) | $pIC_{50}$ |
|---------|---|--------------------------------|------------|-----------------------------|------------|
| 1       |    | 0.08                           | 153.181    | 0.522                       | 6.2821     |
| 2       |    | 0.36                           | 167.208    | 2.150                       | 5.6670     |
| 3       |    | 0.11                           | 181.235    | 0.607                       | 6.2168     |
| 4       |    | 0.27                           | 209.289    | 1.290                       | 5.8894     |
| 5       |   | 2.21                           | 265.397    | 8.330                       | 5.0795     |
| 6       |  | 0.08                           | 187.198    | 0.427                       | 6.3692     |
| 7       |  | 0.11                           | 215.252    | 0.511                       | 6.2916     |
| 8       |  | 0.32                           | 191.161    | 1.670                       | 5.7762     |

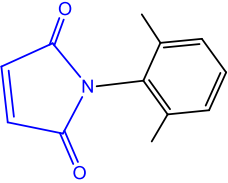
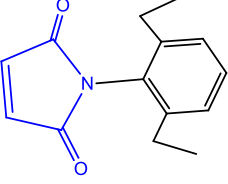
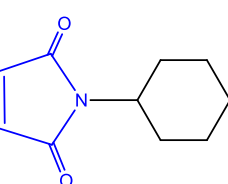
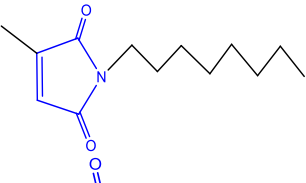
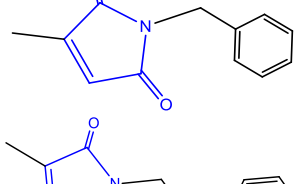
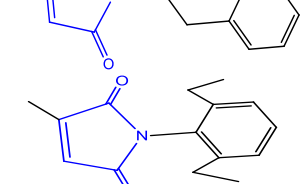
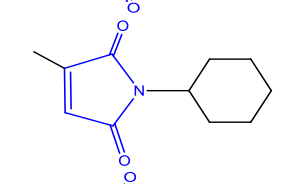
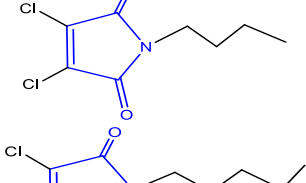
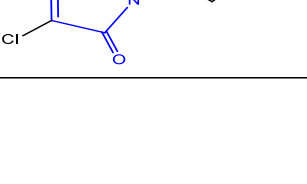

set compounds. This involves calculating the correlation coefficient ( $R^2$  test) for the test set, upon which the predictive strength of the model depends (Isyaku et al. 2020). Furthermore, the data were subjected to the Golbraikh and Tropsha acceptable model criteria using the MLR-plusValidation tool (version 1.3) (Roy et al. 2013; Edache et al. 2020). The various equations and parameters used for the model validation are presented in Table 2.

#### Evaluation of descriptors' mean effects

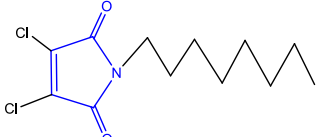
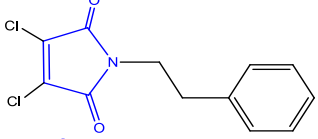
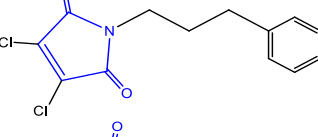
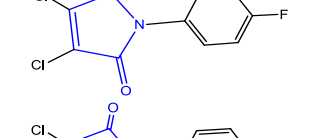
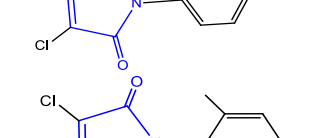
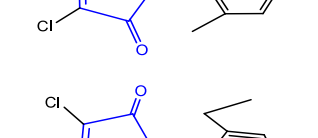
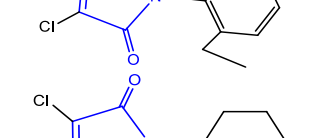
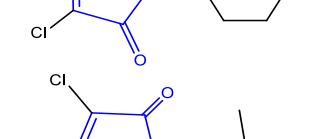
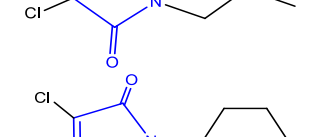
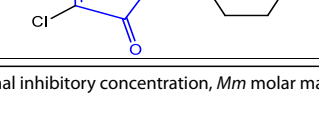
The mean effect (ME) value shows the relative contribution of each descriptor in a model, defined as (Eq. 10):

$$ME = \frac{B_j \sum_i^n D_j}{\sum_j^m (B_j \sum_i^n D_j)} \quad (10)$$

**Table 1** (continued)

|    |   |      |         |       |        |
|----|---|------|---------|-------|--------|
| 9  |    | 0.10 | 201.225 | 0.497 | 6.3037 |
| 10 |    | 0.08 | 229.279 | 0.349 | 6.4573 |
| 11 |    | 0.26 | 179.219 | 1.450 | 5.8384 |
| 12 |   | 0.96 | 223.316 | 4.300 | 5.3666 |
| 13 |  | 0.58 | 201.225 | 2.880 | 5.5403 |
| 14 |  | 1.22 | 215.252 | 5.670 | 5.2466 |
| 15 |  | 0.55 | 243.306 | 2.260 | 5.6458 |
| 16 |  | 0.89 | 193.246 | 4.610 | 5.3367 |
| 17 |  | 0.47 | 222.071 | 2.120 | 5.6744 |
| 18 |  | 0.65 | 236.098 | 2.750 | 5.5602 |

**Table 1** (continued)

|    |   |      |         |       |        |
|----|---|------|---------|-------|--------|
| 19 |    | 0.48 | 278.179 | 1.730 | 5.7631 |
| 20 |    | 1.64 | 270.115 | 6.070 | 5.2167 |
| 21 |    | 0.40 | 284.142 | 1.410 | 5.8515 |
| 22 |    | 0.36 | 260.051 | 1.380 | 5.8588 |
| 23 |   | 0.59 | 276.506 | 2.130 | 5.6709 |
| 24 |  | 0.45 | 270.115 | 1.670 | 5.7783 |
| 25 |  | 0.13 | 298.169 | 0.436 | 6.3605 |
| 26 |  | 0.13 | 248.109 | 0.524 | 6.2807 |
| 27 |  | 0.43 | 222.071 | 1.940 | 5.7130 |
| 28 |  | 0.26 | 248.109 | 1.050 | 5.9797 |

$IC_{50}$  half-maximal inhibitory concentration,  $Mm$  molar mass;  $pIC_{50}$ -negative log of  $IC_{50}$

**Table 2** Some equations and parameters used for the model validation

| Parameter  | Equation   | Eq | Significance  | Threshold value                                |
|--|--|----|---|--|
| <i>Internal validation</i>                                       |  |    |   |  |
| Friedman Lack-Of-Fit (LOF)                                       | $LOF = \frac{SEE}{\left(1 - \frac{c+d \times p}{M}\right)^2}$  | 4  | Allows for the best fitness score to be obtained  | -  |
|  | $SEE = \sqrt{\frac{(Y_{exp} - Y_{pred})^2}{N - p - 1}}$  |    |   |  |
| Correlation coefficient ( $R^2$ )                                | $R^2 = 1 - \left[ \frac{\sum (Y_{exp} - Y_{pred})^2}{\sum (Y_{exp} - \bar{Y}_{training})^2} \right]$               | 5  | Measures the degree of fitness of the regression equation                                 | $\geq 0.6$                                     |
| Adjusted $R^2$   | $R^2_{adj} = \frac{R^2 - p(n-1)}{n - p + 1}$   | 6  | Ensures model's stability and reliability   | $\geq 0.5$                                     |
| Cross-validation regression coefficient ( $Q^2_{cv}$ )           | $Q^2_{cv} = 1 - \left[ \frac{\sum (Y_{pred} - Y_{exp})^2}{\sum (Y_{exp} - \bar{Y}_{training})^2} \right]$          | 7  | Indicates a high internal predictive power  | $\geq 0.5$                                     |
| The coefficient of determination ( $cR^2_p$ ) of Y-Randomization | $cR^2_p = RX[R^2 - (R_r)^2]^2$   | 8  | This is for a confirmation that the QSAR model built is strong and not created by chance  | $cR^2_p > 0.50$                                |
| <i>External validation</i>                                       |  |    |   |  |
| Predicted $R^2$ ( $R^2_{test}$ )                                 | $R^2_{test} = 1 - \frac{\sum (Y_{pred_{test}} - Y_{exp_{test}})^2}{\sum (Y_{pred_{test}} - \bar{Y}_{training})^2}$ | 9  | Measures the ability of the model to predict activity values of external set of compounds | $\geq 0.6$                                     |
| Golbraikh and Tropsha acceptable model criteria                  | $\left  r^2_o - r'^2_o \right $<br>$\left  r^2 - \frac{r'^2}{r^2} \right $<br>$k' \text{ (threshold value)}$       | -  | Assess the robustness and stability of the model  | $< 0.3$<br>$< 0.1$<br>$0.85 \leq k' \leq 1.15$ |

SEE Standard error of estimation;  $c$  number of terms in the model;  $d$  user-defined smoothing parameter,  $p$  total number of descriptors in the model,  $M$  number of data in the training set,  $\bar{Y}_{training}$  = mean experimental activity of the training set,  $Y_{exp}$  experimental activity in the training set,  $Y_{pred}$  predicted activity in the training set,  $n$  number of compounds in the training set,  $cR^2_p$  Y-randomization coefficient,  $R$  correlation coefficient for Y-Randomization,  $R_r$  average 'R' of random models,  $Y_{pred_{test}}$  predicted activity of test set,  $Y_{exp_{test}}$  experimental activity of test set,  $r^2$  square correlation coefficients of the plot of experimental activity versus predicted activity values,  $r^2_o$  square correlation coefficients of the plot of experimental activity versus predicted activity values at zero intercept,  $r'^2_o$  square correlation coefficients of the plot of predicted activity versus experimental activity at zero intercept,  $k'$  slope of the plot of predicted activity against experimental activity at zero intercept

where  $\beta_j$  is the coefficient of the descriptor  $j$  in the model,  $D_j$  is the value of each descriptor in the data matrix for each molecule in the training set,  $m$  is the number of the descriptor that appears in the model,  $n$  is the number of molecules in the training set (Abdullahi et al. 2019).

#### Variance inflation factor

The degree of multi-co-linearity or correspondence between the descriptors is measured by the variance inflation factor (VIF), usually defined as (Eq. 11):

$$VIF = \frac{1}{(1 - R^2)} \quad (11)$$

where  $R^2$  is the correlation coefficient of the multiple regressions between the variables within the model. VIF values of 1 indicate no inter-correlation exists for each variable; for VIF in the range of 1–5, the related model is acceptable; and if VIF is greater than 10, the related model is unstable and unacceptable (Abdullahi et al. 2019).

#### Evaluation of the model's applicability domain

Evaluating the applicability domain (AD) of a QSAR model is important to ascertain the reliability and robustness of the built QSAR model. AD provides one the chance to estimate the uncertainty in the prediction of compounds based on their similarity with the training set compounds, used in the model building (Tropsha et al. 2003). The leverage approach was used to describe the AD of the developed model. The leverage ( $h$ ) of a particular chemical compound is defined thus (Eq. 12):

$$h = X(X^T X)^{-1} X^T \quad (12)$$

where  $X = m \times k$  descriptor matrix of the training set compound,  $X^T$  = transpose matrix of  $X$ .

The warning leverage ( $h^*$ ) which is the range of values used to check for influential molecule or outlier is defined below (Eq. 13):

$$h^* = 3 \frac{(j + 1)}{m} \quad (13)$$

where  $m$  number of training set compounds,  $j$  number of descriptors in the model.

A plot of the standardized residuals against leverages otherwise called the William's plot was used to evaluate the significant area in the model's chemical space. As a rule, compounds which fall within this area on the plot are the approved predicted compounds (Adeniji et al. 2020; Veerasamy et al. 2011).

### Virtual docking screening

Molecular docking investigation was conducted separately between five (5) different protein receptors of *L. donovani* and all 28 maleimides using the Auto Dock Vina of PyRx v software tool. To revalidate the docking results, the reference drug (pentamidine) was equally docked onto the same binding pockets of the 5 receptors. This screening was carried out in order to identify the best therapeutic protein target for the maleimide molecules. The various receptors in 3-D form were retrieved from the protein data bank in the PDB file format and fully prepared on the Biovia Discovery Studio Visualizer by excluding water molecules and co-crystallized ligand enclosed within the protein structure. The 3D structures of all the ligands (maleimides) were saved in PDB file format after the optimization process using Spartan 14 (Adeniji et al. 2019; Adawara et al. 2020). The various receptors used in the virtual docking screening are described in Table 3.

### Structured-based drug design

In this study, twelve (12) compounds (N1–N12) were designed via SBDD using the template compounds (14, 21 and 24) basically by addition of varying substituent groups at the meta, para and ortho positions of the phenyl ring system in the various template structures. N1–N3 were designed out of compound 14, N4–N6 out of 21 and N7–N12 via compound 24. The molecular structures of the newly designed compounds were drawn using the ChemDraw Ultra and subjected to molecular geometry optimization according to the procedures earlier reported herein for molecular geometry optimization. The

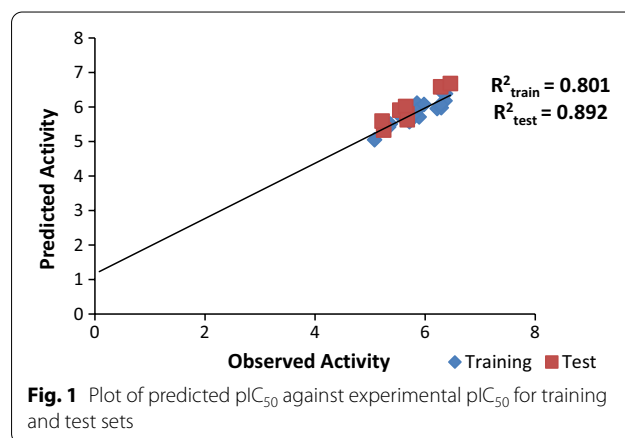
optimized structures were then saved in SD and PDB files format for further analyses involving molecular descriptor calculations and molecular docking study (Abdullahi et al. 2022). Standard docking was performed between the newly designed maleimides and the preferred target receptor (PdxK), using the Auto Dock Vina of PyRx v software tool, while the Biovia Discovery Studio tool was used to visualize the resulting pharmacological interactions as earlier reported herein for virtual docking screening.

### Prediction of pharmacokinetic properties

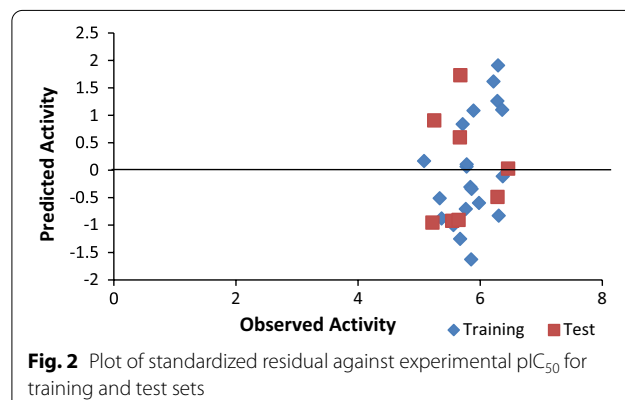
Pharmacokinetics properties prediction constitute an absolutely necessary stage in drug discovery's early phase because only molecules with good drug-likeness properties and excellent ADMET profiles advance into the preclinical research phase (Lawal et al. 2021). Therefore, all the newly designed compounds were investigated for their drug-likeness and ADMET properties using the online web servers; <http://www.swissadme.ch/index.php> and <http://biosig.unimelb.edu.au/pkcsdm>, respectively. The Lipinski's RO5 is a widely used criterion for oral bioavailability. Hence, the tested compounds would be assessed for oral bioavailability using the RO5 criteria (Lipinski et al. 2001).

**Table 3** Description of the various enzymes used for the target fishing

| S. no | Enzyme                                      | Organism           | PDB ID | Resolution (Å) |
|-------|---|--------------------|--------|----------------|
| 1     | Ribokinase                                  | <i>L. donovani</i> | 5ZWY   | 1.95           |
| 2     | Pteridine reductase (PTR 1)                 | <i>L. donovani</i> | 2XOX   | 2.50           |
| 3     | O-acetyl serine sulfhydrylase               | <i>L. donovani</i> | 3SPX   | 1.79           |
| 4     | Pyridoxal kinase/ADP and pyridoxine complex | <i>L. donovani</i> | 6K91   | 2.00           |
| 5     | UMP synthase                                | <i>L. donovani</i> | 3QW4   | 3.00           |

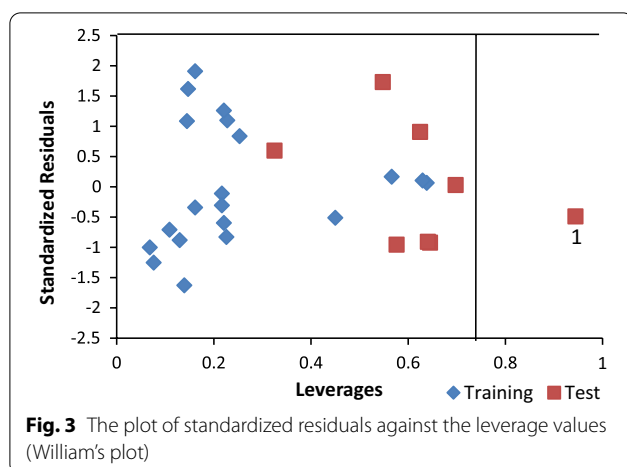


**Fig. 1** Plot of predicted  $pIC_{50}$  against experimental  $pIC_{50}$  for training and test sets



**Fig. 2** Plot of standardized residual against experimental  $pIC_{50}$  for training and test sets





## Results

### QSAR study

The results of theoretical study (QSAR) conducted on the twenty-eight (28) maleimides are presented in Figs. 1, 2, 3 and Tables 4, 5, 6, 7, 8. The built QSAR model (Eq. 14) is composed of four (4) descriptors described clearly in Table 4. In order to ascertain the stability, robustness, reliability and predictive power of the built QSAR model, it was subjected to both internal and external validation tests, and the results are presented in Table 5. The calculated descriptors, experimental activities ( $pIC_{50}$ ) of the various compounds, together with those predicted by the model as well as their residuals, are presented in Table 6. Also, Fig. 1 shows the plot of predicted activity values against those of experimental activity for both training and test sets. A further plot of standardized residuals against experimental activities was obtained and is presented in Fig. 2.

**Table 4** Selected descriptors used in the QSAR model

| S/no | Descriptor symbol | Description   | Class |
|------|-------------------|---|-------|
| 1    | ATSC6v            | Centred Broto–Moreau autocorrelation of lag 6 weighted by van der Waals volume                | 2D    |
| 2    | GAT55c            | Geary autocorrelation of lag 5 weighted by gasteiger charge                                   | 2D    |
| 3    | SpMin8_Bhs        | The smallest absolute eigenvalue of Burden modified matrix – n 8/weighted by relative I-state | 2D    |
| 4    | LOBMIN            | The L/B ratio for the rotation that results in the minimum area                               | 3D    |

**Table 5** Validated parameters of the QSAR model

| Validation parameters                            | Model     | Threshold                | Remarks |
|--|-----------|--------------------------|---------|
| <i>Training set</i>                              |           |                          |         |
| Friedman LOF                                     | 0.151244  | –                        | –       |
| R-squared ( $R^2$ )                              | 0.801298  | $\geq 0.6$               | Passed  |
| Adjusted R-squared ( $R^2_{adj}$ )               | 0.748311  | $\geq 0.5$               | Passed  |
| Cross-validated R-squared ( $Q^2_{cv}$ )         | 0.710284  | $\geq 0.5$               | Passed  |
| $R^2 - Q^2_{cv}$                                 | 0.091014  | $\leq 0.3$               | Passed  |
| Significant regression                           | Yes       | –                        | –       |
| Significance-of-regression F-value               | 15.122516 | –                        | –       |
| Critical SOR F-value (95%)                       | 3.103976  | –                        | –       |
| Replicate points                                 | 1         | –                        | –       |
| Computed experimental error                      | 0.21286   | –                        | –       |
| Lack-of-fit points                               | 14        | –                        | –       |
| Min expt. error for nonsignificant LOF (95%)     | 0.000000  | –                        | –       |
| <i>Test set</i>                                  |           |                          |         |
| R-squared ( $R^2_{test}$ ) i.e. $r^2$            | 0.892     | $\geq 0.6$               | Passed  |
| Number of test set compounds ( $N_{test\ set}$ ) | 8         | $\geq 5$                 | Passed  |
| $ r^2_o - r'^2_o $                               | 0.00599   | $< 0.3$                  | Passed  |
| $ r^2 - r'^2_o $                                 | 0.00093   | $< 0.1$                  | Passed  |
| $k'$   | 1.03956   | $0.85 \leq k' \leq 1.15$ | Passed  |

**Table 6** Calculated descriptors, experimental  $pIC_{50}$ , predicted  $pIC_{50}$  and residuals of the maleimides

| Comp ID | ATSC6v   | GATS5c   | SpMin8_Bhs | LOBMIN    | Exp. $pIC_{50}$ | Pred. $pIC_{50}$ | Residuals |
|---------|----------|----------|------------|-----------|-----------------|------------------|-----------|
| *1      | -0.87061 | 0.807177 | 0.487764   | 2.99E-302 | 6.2821          | 6.5850           | -0.3029   |
| *2      | 68.11814 | 0.850080 | 0.247823   | 2.956504  | 5.6670          | 5.8008           | -0.1338   |
| 3       | 3.519155 | 0.802467 | 0.125168   | 3.229268  | 6.2168          | 5.9548           | 0.26203   |
| 4       | 14.31615 | 0.743074 | 0.317457   | 3.923050  | 5.8894          | 5.7134           | 0.17596   |
| 5       | 20.85557 | 0.695707 | 0.702430   | 5.417251  | 5.0795          | 5.0522           | 0.02731   |
| 6       | -281.902 | 0.789288 | 0.174618   | 2.131400  | 6.3692          | 6.3870           | -0.0178   |
| 7       | 335.0083 | 0.749431 | 0.246926   | 2.679580  | 6.2916          | 5.9821           | 0.30943   |
| 8       | -1.64535 | 1.087869 | 0.744944   | 0.000000  | 5.7762          | 5.7657           | 0.01060   |
| 9       | -523.742 | 0.935578 | 0.138658   | 1.345829  | 6.3037          | 6.4380           | -0.1343   |
| *10     | -1824.44 | 1.039158 | 0.371490   | 1.032339  | 6.4573          | 6.6796           | -0.2224   |
| 11      | -472.640 | 1.186918 | 0.022460   | 1.637375  | 5.8384          | 5.8878           | -0.0494   |
| 12      | 17.12542 | 0.945125 | 0.485891   | 2.646233  | 5.3666          | 5.5092           | -0.1425   |
| *13     | 155.7281 | 1.022659 | 0.090320   | 1.631637  | 5.5403          | 5.9109           | -0.3706   |
| *14     | 602.2820 | 1.051403 | 0.315048   | 2.086024  | 5.2466          | 5.3322           | -0.0856   |
| *15     | -1714.91 | 1.217162 | 0.612817   | 1.142319  | 5.6458          | 6.0140           | -0.3682   |
| 16      | -835.324 | 1.390670 | 0.180784   | 1.827930  | 5.3367          | 5.4195           | -0.0828   |
| *17     | -652.101 | 0.995548 | 0.564835   | 2.629568  | 5.6744          | 5.6318           | 0.04262   |
| 18      | -101.100 | 0.875543 | 0.360434   | 2.967507  | 5.5602          | 5.7221           | -0.1619   |
| 19      | -61.4039 | 0.756412 | 0.294463   | 3.442068  | 5.7631          | 5.8777           | -0.1146   |
| *20     | 199.6291 | 0.852694 | 0.115018   | 3.836036  | 5.2167          | 5.5924           | -0.3757   |
| 21      | 144.0129 | 0.767797 | 0.229505   | 2.425906  | 5.8515          | 6.1149           | -0.2634   |
| 22      | -267.907 | 0.769120 | 0.455067   | 3.071717  | 5.8588          | 5.9138           | -0.0550   |
| 23      | -180.781 | 0.849008 | 0.25338    | 3.07202   | 5.6709          | 5.8734           | -0.2026   |
| 24      | 779.8424 | 0.988572 | 0.172124   | 1.225518  | 5.7783          | 5.7612           | 0.01713   |
| 25      | -728.677 | 1.038058 | 0.352118   | 1.159004  | 6.3605          | 6.1823           | 0.17821   |
| 26      | -956.98  | 1.178622 | 0.052329   | 1.69209   | 6.2807          | 6.0763           | 0.20442   |
| 27      | -664.409 | 1.183124 | 0.53361    | 1.5443    | 5.713           | 5.5772           | 0.13585   |
| 28      | -956.98  | 1.178622 | 0.052329   | 1.69209   | 5.9797          | 6.0763           | -0.0966   |

\* Test set compound

**Table 7** Pearson's correlation and statistical analyses of descriptors used in the QSAR model

| Descriptors | Inter-correlation |         |            |        | Statistical parameters |        |         |          |
|-------------|-------------------|---------|------------|--------|------------------------|--------|---------|----------|
|             | ATSC6v            | GATS5c  | SpMin8_Bhs | LOBMIN | VIF                    | ME     | t Stat  | p-value  |
| ATSC6v      | 1                 |         |            |        | 1.8938                 | -0.036 | -3.4079 | 0.00389  |
| GATS5c      | -0.6434           | 1       |            |        | 3.178                  | 0.7093 | -5.735  | 3.95E-05 |
| SpMin8_Bhs  | 0.2589            | -0.2312 | 1          |        | 1.0884                 | 0.0865 | -3.9339 | 0.00133  |
| LOBMIN      | 0.2894            | -0.6848 | 0.1889     | 1      | 2.0482                 | 0.2402 | -5.8204 | 3.37E-05 |

VIF variance inflation factor, ME mean effect

$$\begin{aligned}
 pIC_{50} = & -0.000437131 * ATSC6v - 2.149874608 * GATS5c \\
 & - 0.840892310 * SpMin8\_Bhs \\
 & - 0.292116894 * LOBMIN + 8.730133535
 \end{aligned}
 \tag{14}$$

Additionally, Pearson's correlation statistical analyses were performed on the values of all four descriptors in

the built QSAR model, and the results are reported in Table 7. Another significant validation test is the Y-randomization test, which was also performed in order to confirm that the QSAR model is strong and not created by chance. The resulting parameters are presented in Table 8. Also, the scatter plot of the standardized residuals versus the leverages (William's Plot) obtained to

**Table 8** Y-randomization test parameters

| Model                           | R        | R <sup>2</sup> | Q <sup>2</sup> |
|---------------------------------|----------|----------------|----------------|
| Original                        | 0.895153 | 0.801298       | 0.710284       |
| Random 1                        | 0.351112 | 0.12328        | -0.69927       |
| Random 2                        | 0.484525 | 0.234764       | -0.13213       |
| Random 3                        | 0.266796 | 0.07118        | -0.88374       |
| Random 4                        | 0.440654 | 0.194176       | -0.28691       |
| Random 5                        | 0.647923 | 0.419804       | 0.035498       |
| Random 6                        | 0.52495  | 0.275572       | -0.19051       |
| Random 7                        | 0.6118   | 0.374299       | -0.02487       |
| Random 8                        | 0.440392 | 0.193945       | -0.67219       |
| Random 9                        | 0.758172 | 0.574825       | 0.01589        |
| Random 10                       | 0.478221 | 0.228695       | -0.61115       |
| <i>Random models parameters</i> |          |                |                |
| Average r:                      | 0.500454 |                |                |
| Average r <sup>2</sup> :        | 0.269054 |                |                |
| Average Q <sup>2</sup> :        | -0.34494 |                |                |
| cR <sub>p</sub> <sup>2</sup>    | 0.664372 |                |                |

cR<sub>p</sub><sup>2</sup> Y-randomization coefficient, R correlation coefficient for Y-randomization  
 R<sup>2</sup>-correlation coefficient, Q<sup>2</sup>-cross-validated R<sup>2</sup>

ascertain the model's applicability domain is shown in Fig. 3.

### Virtual docking screening

The results (binding affinities) of the docking simulation conducted between five (5) protein receptors of *L. donovani* and the studied maleimides, as well as the reference drug (pentamidine) are reported in Table 9. Also, the prepared 3-D structure of PdxK is shown in Fig. 4, while the pharmacological interactions between the receptor (PdxK)'s amino acid residues and the selected templates (14, 21 and 24), as well as the reference drug (pentamidine) as compiled through Discovery Studio Visualizer, are presented in Table 10. Additionally, the 2D and 3D views of the binding interactions of compounds 14, 21, 24, and pentamidine with the binding site of PdxK receptor as adapted from the Discovery Studio Visualizer, are shown in Figs. 5, 6, 7 and 8, respectively.

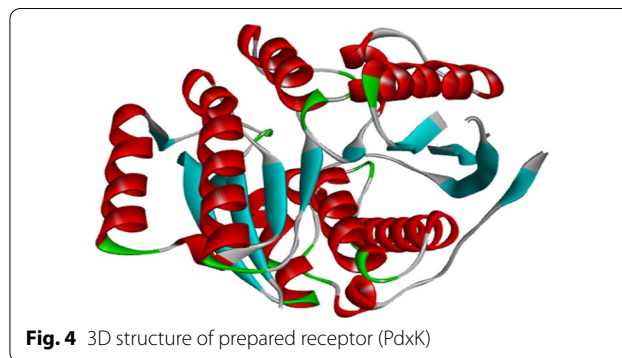
### Structured-based drug design

The 2D molecular structures, predicted pIC<sub>50</sub> and binding energies of the newly designed compounds as well as the reference drug (pentamidine) are reported in Table 11. Also, the binding interaction profiles of the target receptor (PdxK) with the newly designed compounds are presented in Table 12, while the 2D and 3D views of the binding interactions of N7 and N12 with the receptor's amino acid residues are shown in Figs. 9 and 10, respectively.

**Table 9** Summary of binding affinities of interactions between maleimides and different *L. donovani* receptors used for the target fishing

| COMP ID     | Protein–ligand binding affinities (kcal/mol) |      |      |      |      |
|-------------|--|------|------|------|------|
|             | 5ZWY   | 2XOX | 3SPX | 6K91 | 3QW4 |
| 1           | -5.4   | -5.1 | -5.2 | -5.5 | -4.8 |
| 2           | -5.3   | -5.1 | -5.6 | -6.0 | -4.9 |
| 3           | -5.4   | -5.2 | -5.3 | -6.0 | -4.8 |
| 4           | -5.4   | -5.4 | -5.7 | -6.4 | -5.3 |
| 5           | -4.9   | -5.2 | -5.9 | -6.9 | -5.0 |
| 6           | -6.2   | -6.0 | -6.7 | -7.1 | -5.6 |
| 7           | -6.4   | -6.5 | -6.7 | -7.6 | -6.3 |
| 8           | -6.4   | -6.7 | -6.6 | -7.0 | -6.5 |
| 9           | -6.2   | -6.4 | -6.6 | -6.7 | -5.9 |
| 10          | -6.2   | -6.7 | -7.0 | -5.4 | -5.8 |
| 11          | -6.1   | -6.1 | -6.1 | -6.3 | -5.9 |
| 12          | -5.8   | -6.0 | -6.8 | -6.8 | -5.3 |
| 13          | -6.8   | -6.3 | -7.1 | -7.3 | -6.0 |
| 14          | -7.0   | -6.8 | -7.1 | -7.7 | -6.3 |
| 15          | -6.5   | -6.9 | -7.3 | -6.7 | -6.3 |
| 16          | -6.7   | -6.4 | -6.5 | -6.8 | -6.2 |
| 17          | -5.6   | -5.8 | -5.9 | -5.9 | -5.1 |
| 18          | -5.7   | -5.9 | -5.7 | -6.1 | -5.5 |
| 19          | -5.5   | -5.5 | -6.1 | -6.3 | -5.8 |
| 20          | -6.3   | -6.4 | -7.2 | -7.5 | -6.5 |
| 21          | -6.4   | -6.6 | -7.1 | -7.7 | -6.2 |
| 22          | -6.7   | -7.1 | -7.1 | -7.2 | -6.8 |
| 23          | -6.4   | -7.0 | -6.8 | -6.2 | -6.5 |
| 24          | -6.6   | -6.5 | -7.0 | -7.8 | -6.2 |
| 25          | -6.7   | -6.8 | -7.4 | -7.4 | -5.9 |
| 26          | -6.1   | -6.0 | -6.3 | -7.0 | -6.3 |
| 27          | -5.7   | -6.0 | -5.9 | -6.0 | -5.2 |
| 28          | -6.1   | -6.2 | -6.3 | -7.0 | -6.3 |
| Pentamidine | -6.4   | -6.8 | -7.4 | -8.1 | -7.0 |

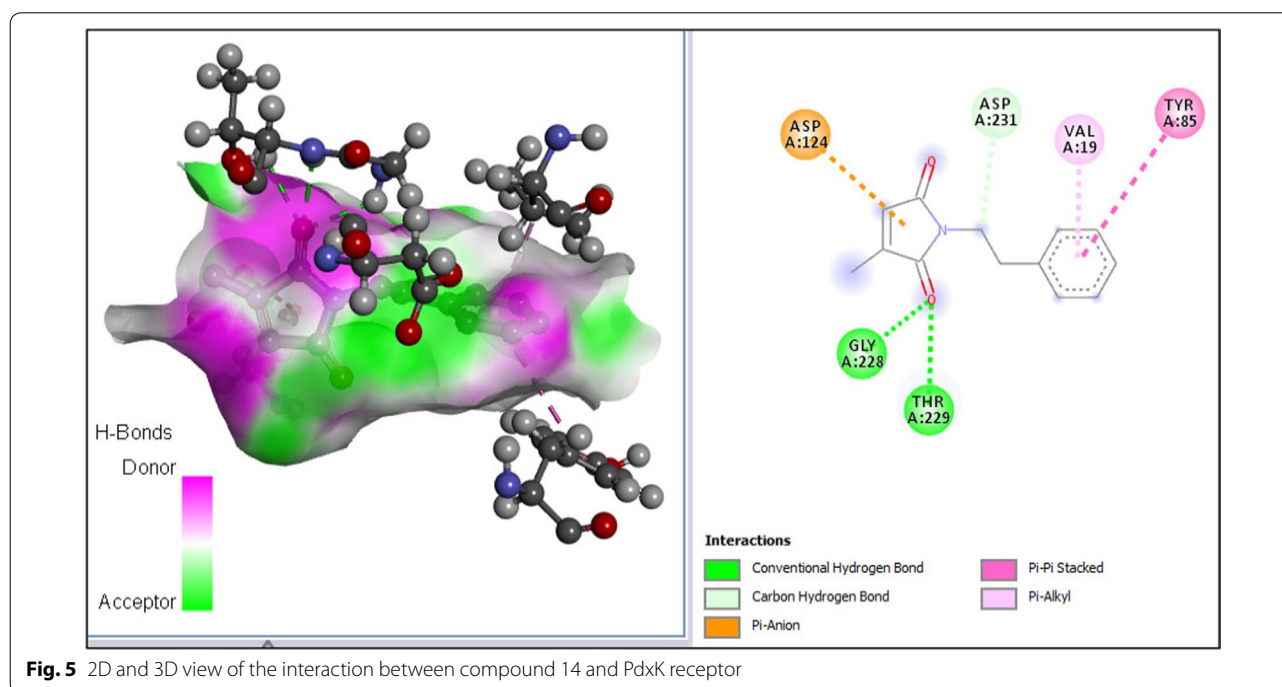
PDB ID—5ZWY, 2XOX, 3SPX, 6K91, 3QW4

**Fig. 4** 3D structure of prepared receptor (PdxK)

**Table 10** Predicted binding interaction profile of selected template molecules with PdxK receptor

| ID | Hydrogen bond interactions |                 |              | Electrostatic and hydrophobic interactions  |      |
|----|----------------------------|-----------------|--------------|---|------|
|    | Amino acid                 | Type            | Distance (Å) |   |      |
| 14 | THR-229                    | Conventional    | 2.41, 2.48   | ASP-124 ( $\pi$ -anion), TYR-85 ( $\pi$ - $\pi$ ), VAL-19 ( $\pi$ -alkyl)   |      |
|    | GLY-228                    | Conventional    | 2.37         |   |      |
|    | ASP-231                    | Carbon-hydrogen | 3.44         |   |      |
| 21 | ASN-151                    | Conventional    | 2.31         | VAL-121 (alkyl), ASP-124 ( $\pi$ -anion), LYS-187 ( $\pi$ -alkyl), LEU-257 ( $\pi$ -alkyl)                                    |      |
|    | GLY-228                    | Conventional    | 2.60         |   |      |
|    | THR-229                    | Conventional    | 2.15, 2.47   |   |      |
|    | LYS-187                    | Conventional    | 2.28         |   |      |
|    | THR-227                    | Carbon-hydrogen | 2.67         |   |      |
| 24 | –                          | –               | –            | VAL-219 (alkyl), LEU-198 (alkyl), ILE-261 ( $\pi$ -alkyl), TYR-226 (alkyl), PHE-233 (alkyl), LYS-187 (alkyl), LEU-257 (alkyl) |      |
|    | Pentamidine                | ASN-151         | Conventional |   | 1.95 |
|    |                            | SER-12          | Conventional |   | 2.28 |
|    | ASP-231                    | Carbon-hydrogen | 3.40         |   |      |

ASN asparagine, ASP aspartic acid, GLY glycine, ILE isoleucine, LEU leucine, LYS lysine, PHE phenylalanine, SER serine, THR threonine, TYR tyrosine, VAL valine

**Fig. 5** 2D and 3D view of the interaction between compound 14 and PdxK receptor

### Pharmacokinetics properties prediction

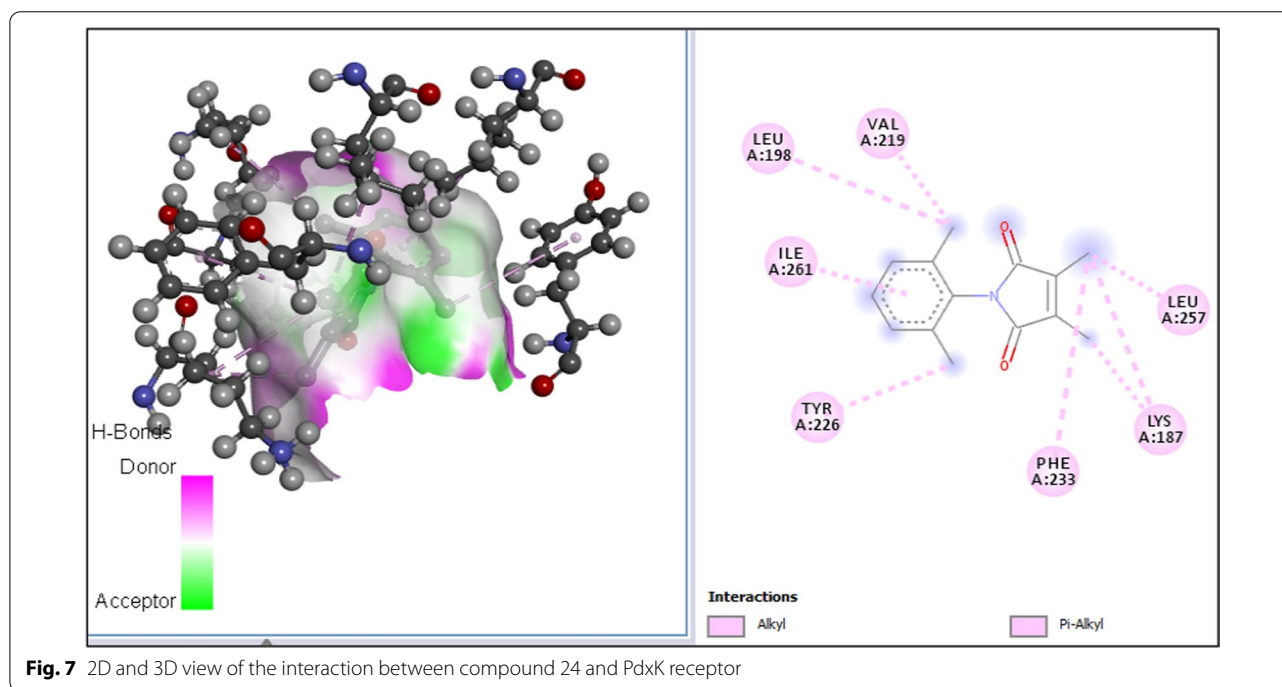
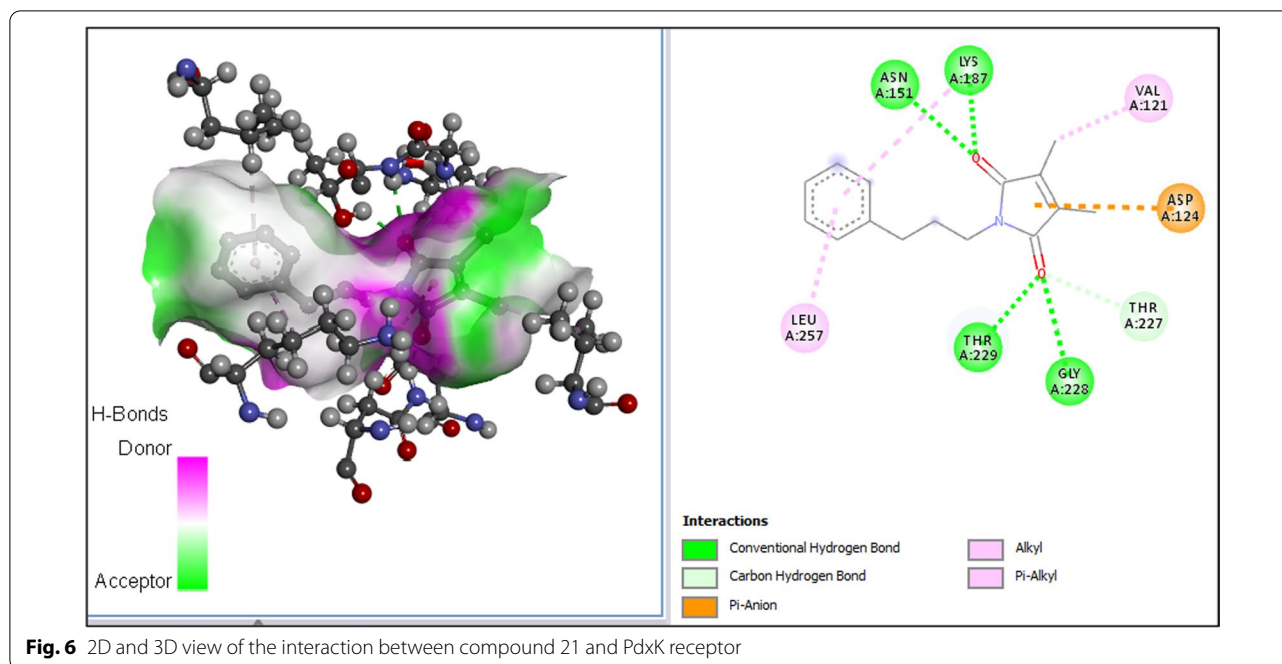
Results of the pharmacokinetics investigation conducted on the newly designed compounds are presented in Tables 13, 14, while Fig. 11 shows their Boiled Egg's representation.

### Discussion

#### QSAR study

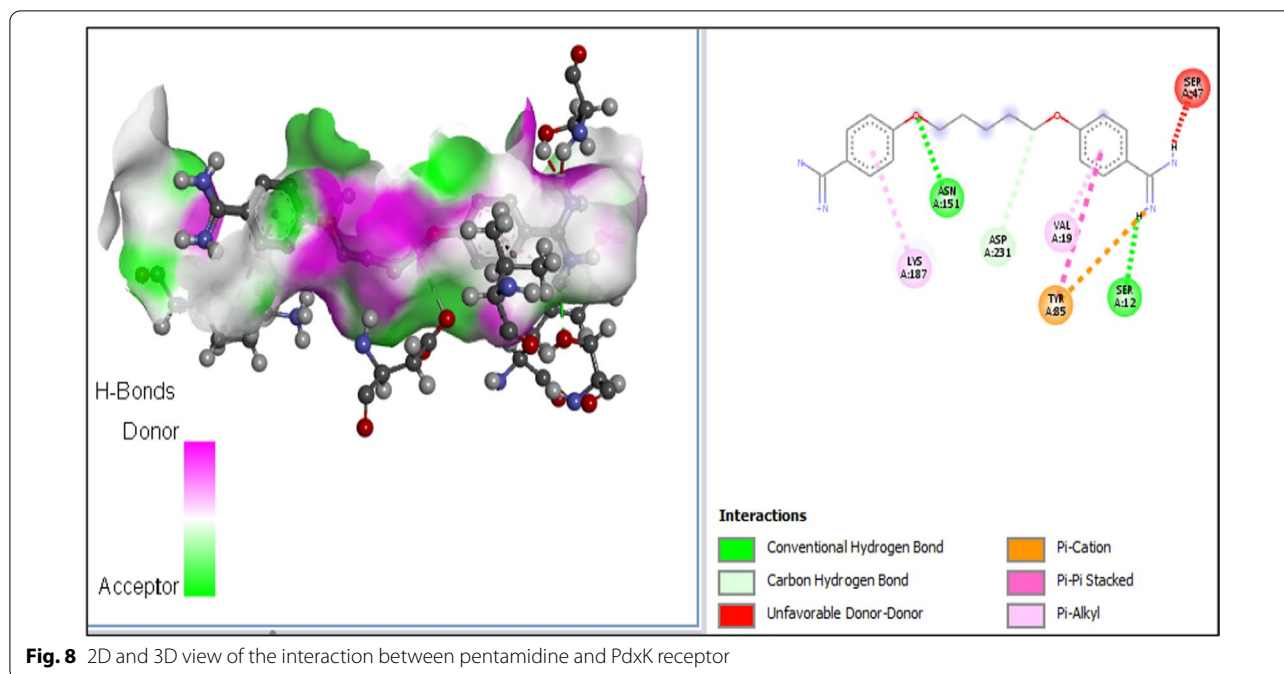
A combined GFA and MLR approach led to the selection of four (4) descriptors in the generation of the QSAR

model. The built model (Eq. 14) was found to excellently satisfy the requirement for a reliable QSAR model. The low residual values between the experimental and predicted activities as shown in Table 6 indicate a high predictive strength for the QSAR model. The  $R^2$  values of 0.801 and 0.892 for training set and test set, respectively, as obtained from the plot of exp.  $pIC_{50}$  against pred.  $pIC_{50}$  in Fig. 1 compare perfectly well with those obtained from GFA (0.8013 and 0.892) and MLRplusValidation analysis (0.8013 and 0.892) as reported in Table 5.



The close grouping of points along the line of best fit in Fig. 1 shows a very good correlation between the experimental and predicted activity values, indicating that the built model is reliable and robust. The random spread of standardized residuals on both sides of the zero mark in Fig. 2 is an indication that the built model is free of any systematic error.

Furthermore, the low correlation coefficients (less than 0.50) which exist between each pair of the descriptor in the built model (Table 7) indicate no inter-correlation between each descriptor. Similar result was also obtained elsewhere by Adeniji et al. (2018) and Abdullahi et al. (2022). The VIF values ranging from 1 to 5 for all 4 descriptors as reported in Table 7 showed that



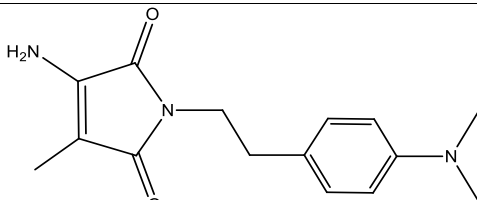
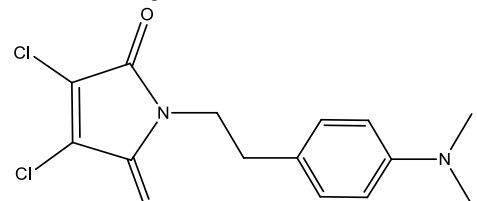
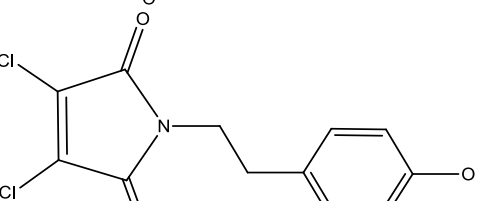
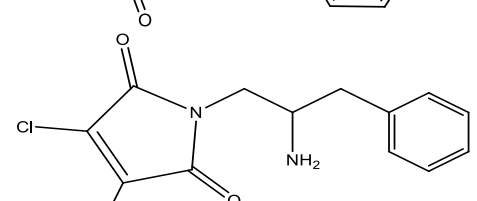
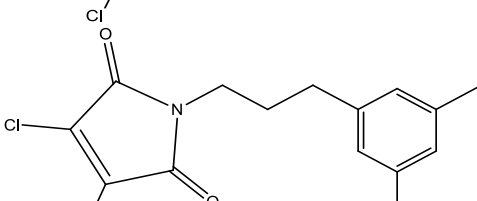
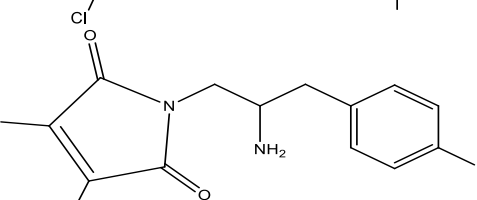
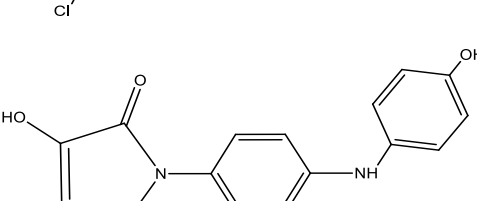
the descriptors were statistical orthogonal and the built model was statistically substantial, an indication of the stability and acceptability of the built QSAR model. Similar observation was reported by Adeniji et al. (2019). The values of the absolute *t*-statistics greater than 2 for each descriptor show that the selected descriptors were good (Adeniji et al. 2018). Also, the evaluated *p*-values for the various descriptors in the model at 95% confidence level were less than 0.05 as shown in Table 7. Therefore, the alternative hypothesis which asserts that a relationship exists between the descriptors used in generating the model and the compounds' inhibitory activities at  $p < 0.05$  holds. Additionally, the values of the mean effect (ME) reported in Table 7 provide vital information on the effect and degree of each descriptor's contributions in the built model. The magnitudes and signs of ME values signify their individual strength and direction on the molecules' inhibitory activities. All the descriptors except ATSC6v have positive ME, meaning that increase or decrease in their values will lead to an increase or decrease in the anti-proliferative activities, respectively. Increasing the values of ATSC6v will lead to a decrease in the inhibitory activities because of its negative ME value. GATS5c with the highest ME value has the greatest influence on the molecules' inhibitory activities. GATS5c is Geary autocorrelation of lag 5 weighted by gasteiger charge, which has a positive ME suggested to contribute positively to anti-leishmanial activity. The gasteiger charge is a physicochemical property calculated for every atom in the molecule and is the Geary coefficient (Mahmud et al. 2020).

The low values of  $R^2$  and  $Q^2$  obtained from the random reshuffling (Table 8) inferred that the built model is stable, robust and reliable. The value of coefficient for Y-randomization,  $cR^2_p$  (0.664372) greater than 0.50, supports the claim that the built model is powerful and not inferred by chance. The William's plot (Fig. 3) clearly shows that all the compounds fall within the square area  $\pm 2.5$  of standardized cross-validated residual. It can therefore be inferred that no outlier is present in the dataset. However, compound 1 was found with leverage value greater than the calculated warning leverage ( $h^* = 0.75$ ) and was said to be an influential molecule.

#### Virtual docking screening

Binding energies of the protein–ligand (drug) interactions are important to describe how well the drug binds to the target macromolecule. The negative value of the binding energy change shows the spontaneity of the binding process and how well ligands can fit into the target protein pocket to form the most energetically stable drug receptor (Ugbe et al. 2021). Among the studied receptors, pyridoxal kinase (PdxK) receptor (pdb id: 6K91) had relatively shown the strongest interaction with the various compounds as shown by the higher binding energy values associated with this receptor (Table 9). Consequently, PdxK was selected as the target receptor of interest in this study and subsequent discussions pertaining protein–ligand interactions shall be based on it. Also, among the 28 maleimides studied, compounds 14, 21 and 24 bind more strongly to PdxK with the highest

**Table 11** Molecular structures, predicted  $pIC_{50}$  and binding energies of the designed compounds

| Comp ID | Molecular structures  | Pred. $pIC_{50}$ | Binding energy |
|---------|---|------------------|----------------|
| N1      |    | 5.1374           | -8.0           |
| N2      |    | 5.2139           | -7.8           |
| N3      |    | 6.4988           | -7.9           |
| N4      |   | 5.8218           | -7.9           |
| N5      |  | 5.6513           | -8.0           |
| N6      |  | 5.7616           | -7.9           |
| N7      |  | 6.6091           | -8.9           |

**Table 11** (continued)

|             |  |        |      |
|-------------|--|--------|------|
| N8          |  | 5.8313 | -7.9 |
| N9          |  | 6.793  | -7.9 |
| N10         |  | 5.8653 | -8.0 |
| N11         |  | 6.6075 | -7.8 |
| N12         |  | 6.2275 | -8.5 |
| Pentamidine |  | -      | -8.1 |

reported binding energies of  $-7.7$ ,  $-7.7$  and  $-7.8$  kcal/mol, respectively. These compounds were equally well predicted by the built QSAR model with low residual values and contained within the model's applicability domain (Table 6 and Fig. 3). Therefore, **14**, **21** and **24** were selected as template molecules for designing new compounds with improved binding scores and pharmacological properties.

As seen from Table 10 and Figs. 5, 6, 7, 8, compound **14** was observed to have interacted well with the binding

site of the PdxK receptor through three (3) conventional hydrogen bonds, one (1) carbon-hydrogen bond, one (1)  $\pi$ -anion, one (1)  $\pi$ - $\pi$  stacked and one (1) pi-alkyl interactions. One of the carbonyl oxygen atoms on the pyrrole ring system formed 3 conventional hydrogen bonds with 2 THR-229 at distances of 2.41 Å and 2.48 Å, and GLY-228 at a distance of 2.37 Å. It also formed a carbon-hydrogen bond with ASP-231 at a distance of 3.44 Å. Others include  $\pi$ -anion between the pyrrole ring system and ASP-124, pi-pi stacking between the phenyl



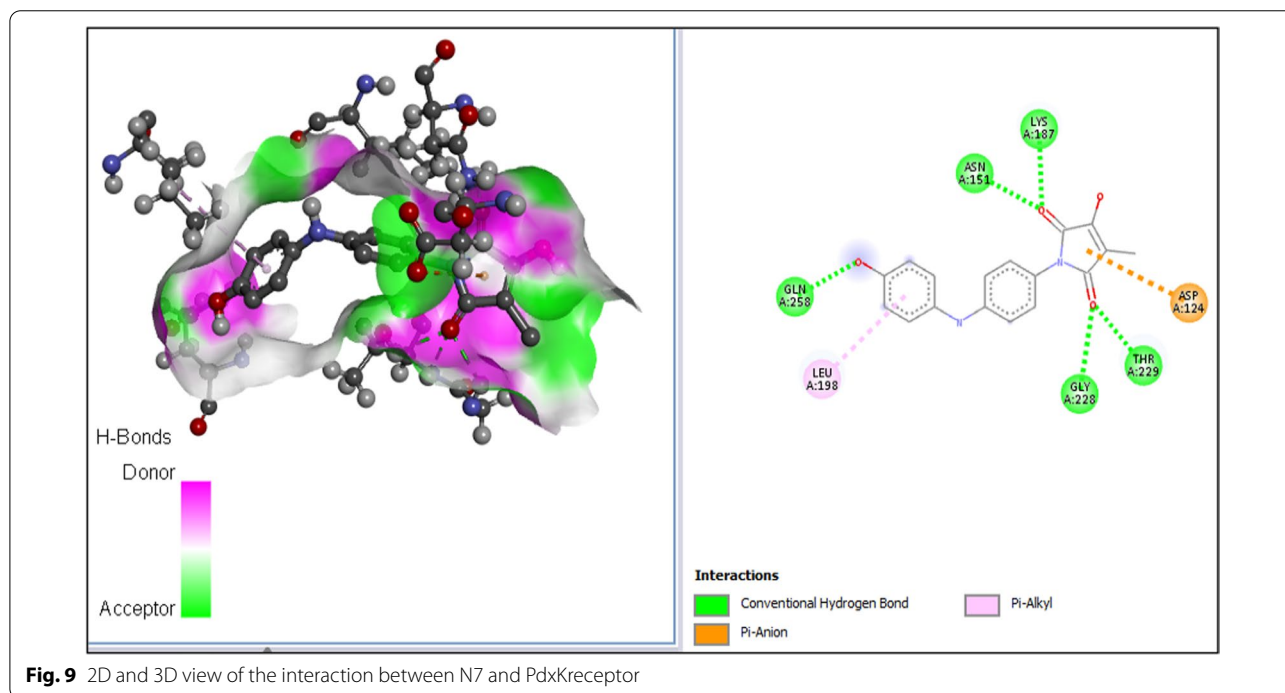
**Table 12** Predicted binding interaction profile of newly designed molecules with PdxK receptor

| ID | Hydrogen bond Interactions |                 |                  | Electrostatic and hydrophobic interactions                    |
|----|----------------------------|-----------------|------------------|---|
|    | Amino acid                 | Type            | Distance (Å)     |   |
| N1 | TYR-129                    | Conventional    | 2.93             | TYR-85, VAL-19, VAL-121                                       |
|    | THR-229                    | Conventional    | 2.55             |   |
|    | GLU-154                    | Conventional    | 2.38, 2.73       |   |
|    | GLY-228                    | Conventional    | 2.39             |   |
|    | GLY-230                    | Conventional    | 2.07             |   |
| N2 | ASP-231                    | Carbon-hydrogen | 3.47             | VAL-121, ASP-124, LYS-187                                     |
|    | ASN-151                    | Conventional    | 2.28             |   |
|    | GLY-228                    | Conventional    | 2.40             |   |
|    | THR-229                    | Conventional    | 2.17             |   |
|    | LYS-187                    | Conventional    | 2.31             |   |
| N3 | THR-227                    | Carbon-hydrogen | 2.59             | VAL-121, ASP-124, VAL-19, TYR-85                              |
|    | THR-229                    | Carbon-hydrogen | 3.28             |   |
|    | LYS-187                    | Carbon-hydrogen | 3.79             |   |
|    | GLY-228                    | Conventional    | 2.28             |   |
|    | SER-12                     | Conventional    | 2.09             |   |
| N4 | THR-229                    | Conventional    | 2.49             | TYR-129, LEU-257, VAL-121, ASP-124, LYS-187                   |
|    | ASP-231                    | Carbon-hydrogen | 3.72             |   |
|    | THR-227                    | Carbon-hydrogen | 2.88             |   |
|    | ASP-124                    | Conventional    | 2.75             |   |
|    | GLY-228                    | Conventional    | 2.55             |   |
| N5 | LYS-187                    | Conventional    | 2.06             | VAL-121, ASP-124, LYS-187, MET-254, LEU-257, ILE-261, PHE-233 |
|    | THR-229                    | Conventional    | 2.27, 2.30, 2.41 |   |
|    | THR-229                    | Carbon-hydrogen | 3.39             |   |
|    | ASN-151                    | Conventional    | 2.40             |   |
|    | GLY-228                    | Conventional    | 2.53             |   |
| N6 | LYS-187                    | Conventional    | 2.12             | VAL-121, LYS-187, ASP-124                                     |
|    | THR-229                    | Conventional    | 2.28, 2.41       |   |
|    | THR-229                    | Carbon-hydrogen | 3.39             |   |
|    | ASN-151                    | Conventional    | 2.39             |   |
|    | ASP-124                    | Conventional    | 2.83             |   |
| N7 | GLN-258                    | Conventional    | 2.69             | LEU-198 ( $\pi$ -alkyl), ASP-124 ( $\pi$ -anion)              |
|    | GLY-228                    | Conventional    | 2.57             |   |
|    | LYS-187                    | Conventional    | 2.06             |   |
|    | THR-229                    | Conventional    | 2.22, 2.38       |   |
|    | THR-229                    | Carbon-hydrogen | 3.32             |   |
| N8 | ASN-151                    | Conventional    | 2.32             | VAL-121, THR-85, TYR-129, ASP-231, VAL-19, LEU-43             |
|    | GLN-258                    | Conventional    | 1.98             |   |
|    | GLY-228                    | Conventional    | 2.40             |   |
|    | LYS-187                    | Conventional    | 2.11             |   |
|    | THR-229                    | Conventional    | 2.38, 2.40       |   |
| N9 | –                          |                 |                  | VAL-121, ASP-231, VAL-19, TYR-85                              |
|    | GLY-228                    | Conventional    | 2.22             |   |
|    | SER-47                     | Conventional    | 2.09, 2.32       |   |
|    | HIS-46                     | Carbon-hydrogen | 2.86             |   |
|    | THR-227                    | Carbon-hydrogen | 2.83             |   |
|    | SER-47                     | Carbon-hydrogen | 2.84             |   |

**Table 12** (continued)

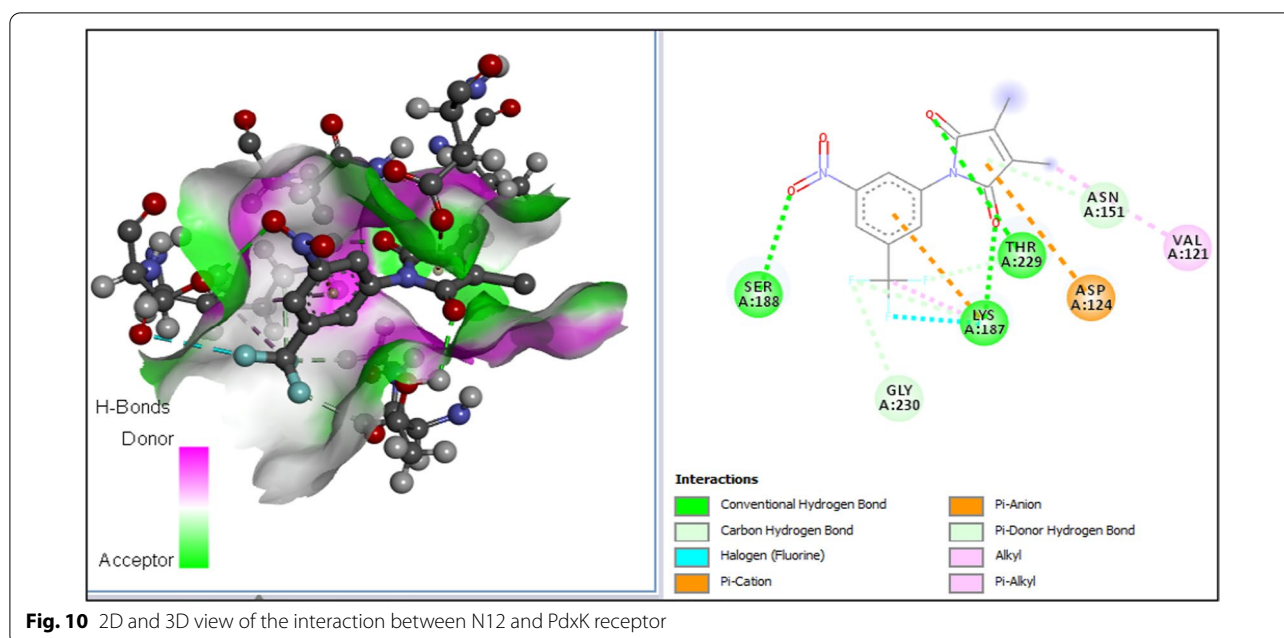
| ID  | Hydrogen bond Interactions |                 |              | Electrostatic and hydrophobic interactions  |
|-----|----------------------------|-----------------|--------------|---|
|     | Amino acid                 | Type            | Distance (Å) |   |
| N10 | LYS-187                    | Conventional    | 2.06         | ASP-124, VAL-121  |
|     | SER-188                    | Conventional    | 2.68         |   |
|     | THR-229                    | Conventional    | 2.39         |   |
|     | ASN-151                    | Carbon-hydrogen | 3.23         |   |
|     | GLY-230                    | Carbon-hydrogen | 2.75         |   |
| N11 | LYS-187                    | Conventional    | 2.05         | ASP-124, VAL-121, LYS-187   |
|     | THR-229                    | Conventional    | 2.40         |   |
|     | ASN-151                    | Carbon-hydrogen | 3.17         |   |
|     | GLY-230                    | Carbon-hydrogen | 2.68         |   |
|     | THR-229                    | Carbon-hydrogen | 2.35         |   |
| N12 | LYS-187                    | Conventional    | 2.06         | VAL-121 (alkyl), ASP-124 ( $\pi$ -anion), LYS-187 (fluorine), LYS-187 (alkyl), LYS-187 ( $\pi$ -cation) |
|     | SER-188                    | Conventional    | 2.64         |   |
|     | THR-229                    | Conventional    | 2.46         |   |
|     | ASN-151                    | Pi-donor H-bond | 3.19         |   |
|     | GLY-230                    | Carbon-hydrogen | 2.71         |   |
|     | LYS-187                    | Carbon-hydrogen | 3.04         |   |
|     | THR-229                    | Carbon-hydrogen | 2.38         |   |

ASN asparagine, ASP aspartic acid, GLN glutamine, GLU glutamic acid, GLY glycine, HIS histidine, ILE isoleucine, LEU leucine, LYS lysine, MET methionine, PHE phenylalanine, SER serine, THR threonine, TYR tyrosine, VAL valine



ring and aromatic ring of TYR-85, and  $\pi$ -alkyl interaction between the phenyl ring and VAL-19. Similarly, Compound **21** binds well into the binding pockets of the

PdxK receptor via five (5) conventional hydrogen bonds, a carbon-hydrogen bond, one  $\pi$ -anion, two  $\pi$ -alkyl, and one alkyl interactions. One of the carbonyl groups oxygen

**Table 13** Predicted drug-likeness properties of the newly designed compounds

| Comp ID | MW (g/mol) | TPSA (Å <sup>2</sup> ) | WLOGP | MLOGP | Log S (ESOL) | HBD | HBA | SA   | RO5 violation | Drug-likeness |
|---------|------------|------------------------|-------|-------|--------------|-----|-----|------|---------------|---------------|
| N1      | 273.33     | 66.64                  | 0.52  | 0.80  | -2.41        | 1   | 2   | 2.52 | 0             | YES           |
| N2      | 313.18     | 40.62                  | 1.97  | 1.88  | -3.78        | 0   | 2   | 2.32 | 0             | YES           |
| N3      | 286.11     | 57.61                  | 1.61  | 1.37  | -3.4         | 1   | 3   | 2.11 | 0             | YES           |
| N4      | 299.15     | 63.40                  | 1.23  | 1.36  | -3.08        | 1   | 3   | 2.73 | 0             | YES           |
| N5      | 312.19     | 37.38                  | 2.91  | 2.70  | -4.38        | 0   | 2   | 2.32 | 0             | YES           |
| N6      | 344.15     | 109.22                 | 1.14  | 0.39  | -3.15        | 1   | 5   | 2.88 | 0             | YES           |
| N7      | 330.72     | 89.87                  | 2.64  | 1.45  | -3.59        | 3   | 4   | 2.75 | 0             | YES           |
| N8      | 298.16     | 37.38                  | 3.17  | 2.73  | -4.23        | 0   | 2   | 2.29 | 0             | YES           |
| N9      | 287.06     | 83.20                  | 1.78  | 1.48  | -3.01        | 0   | 4   | 2.25 | 0             | YES           |
| N10     | 332.05     | 129.02                 | 1.68  | 0.59  | -3.08        | 0   | 6   | 2.63 | 0             | YES           |
| N11     | 310.06     | 37.38                  | 4.04  | 2.62  | -3.8         | 0   | 5   | 2.19 | 0             | YES           |
| N12     | 355.05     | 83.20                  | 3.95  | 2.41  | -4.21        | 0   | 7   | 2.62 | 0             | YES           |

MW molecular weight, TPSA topological polar surface area, ESOL estimated solubility, HBD hydrogen bond donors, HBA hydrogen bond acceptors, RO5 Lipinski rule of five, SA synthetic accessibility score

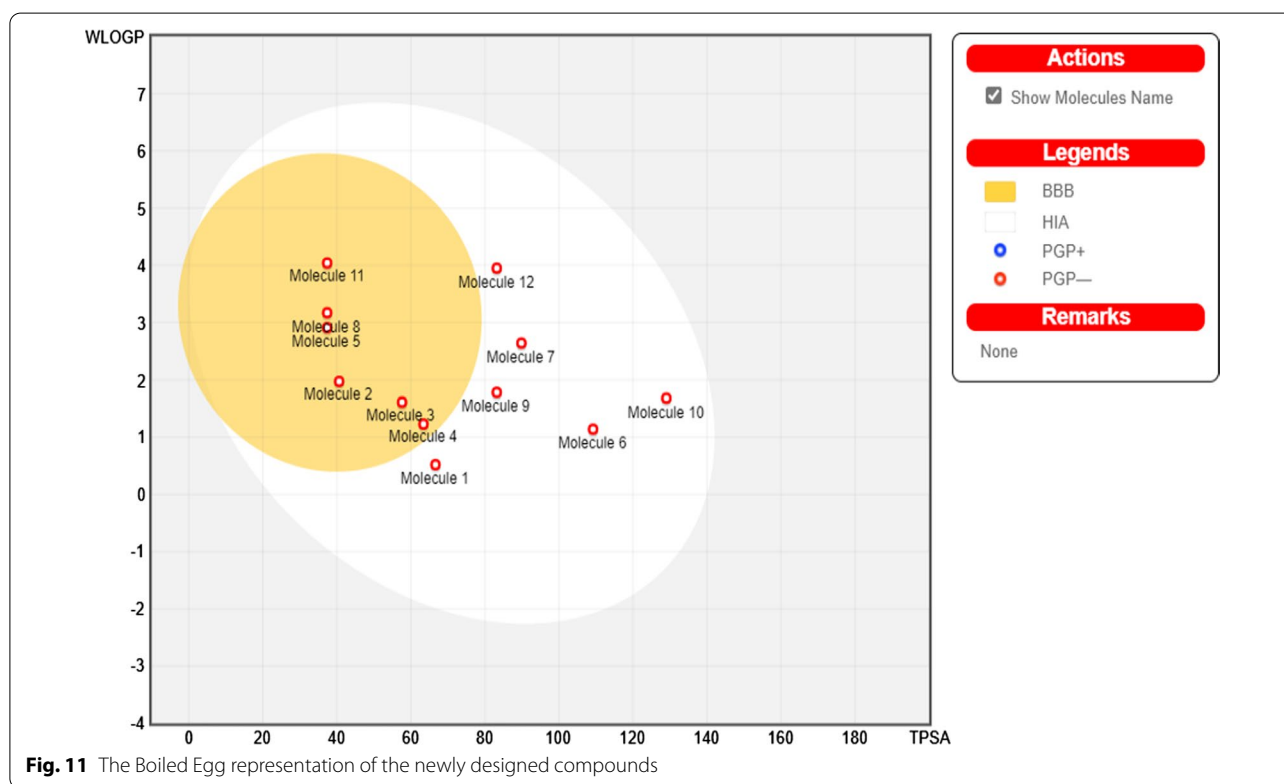
formed conventional hydrogen bonds with GLY-228 and 2 THR-229 at a distance of 2.6 Å, 2.15 Å and 2.47 Å, respectively. The other carbonyl oxygen interacted with ASN-151 and LYS-187 to form conventional hydrogen bond at a distance of 2.31 Å and 2.28 Å, respectively. The ligand also formed a carbon-hydrogen bond with THR-227 at a distance of 2.67 Å. Others include alkyl interaction with VAL-121,  $\pi$ -anion with ASP-124, and  $\pi$ -alkyl with LYS-187 and LEU-257. Distinguishably, the binding interactions of Compound **24** with the receptor were characterized only by electrostatic and hydrophobic

interactions, and without hydrogen bond interactions. This may be attributed to the steric hindrances posed by the side chain methyl groups on the phenyl ring and the chloro groups on the pyrrole ring which shields the carbonyl groups from interacting with amino acid residues to form hydrogen bonds. The observed interactions were dominated by the hydrophobic interactions (alkyl type) with the exception of ILE-261 which binds by electrostatic interaction to the  $\pi$  electron system of the benzene ring of compound **24** via its alkyl group. Lastly, the reference drug (pentamidine) was equally docked onto the

**Table 14** Predicted ADMET properties of the newly designed compounds

| Comp ID | Absorption<br>HIA (%) | Distribution |        | Metabolism |           | Excretion<br>Total clearance | Toxicity<br>MRTD |
|---------|-----------------------|--------------|--------|------------|-----------|------------------------------|------------------|
|         |                       | Log BB       | Log PS | CYP-34A    |           |                              |                  |
|         |                       |              |        | Substrate  | Inhibitor |                              |                  |
| N1      | 75.73                 | -0.131       | -2.912 | Yes        | No        | 0.477                        | 0.308            |
| N2      | 94.20                 | 0.228        | -2.09  | Yes        | No        | 0.092                        | 0.314            |
| N3      | 90.67                 | 0.088        | -2.415 | Yes        | No        | 0.436                        | 0.363            |
| N4      | 92.07                 | 0.051        | -2.55  | Yes        | No        | 0.811                        | -0.13            |
| N5      | 93.47                 | 0.265        | -1.85  | Yes        | No        | -0.028                       | 0.249            |
| N6      | 75.83                 | -0.784       | -2.757 | Yes        | No        | 0.568                        | -0.644           |
| N7      | 89.35                 | -1.016       | -2.349 | No         | No        | -0.336                       | 0.072            |
| N8      | 95.14                 | 0.274        | -1.541 | Yes        | No        | -0.062                       | 0.225            |
| N9      | 94.76                 | -0.788       | -2.249 | Yes        | No        | 0.071                        | 0.21             |
| N10     | 87.67                 | -1.31        | -2.46  | Yes        | No        | 0.292                        | -0.202           |
| N11     | 92.75                 | 0.302        | -2.205 | No         | No        | 0.041                        | 0.298            |
| N12     | 91.33                 | -1.221       | -2.092 | Yes        | No        | 0.182                        | -0.031           |

MRTD Maximum recommended tolerated dose, HIA Human intestinal absorption, LogBB logarithmic ratio of brain to plasma drug concentration, LogPS blood-brain permeability-surface area product, CYP-34A cytochrome p450 isoform



binding pocket of the pyridoxal kinase receptor in order to provide insight into their binding interaction mode and for validation purpose. Pentamidine interacted with PdxK via two (2) conventional hydrogen bonds with ASN-151 and SER-12 at a distance of 1.95 Å and 2.28 Å, respectively. It equally formed a carbon-hydrogen bond

with ASP-231 at a distance of 3.40 Å. Others are electrostatic and hydrophobic interactions including  $\pi$ -alkyl interactions with VAL-19 and LYS-187,  $\pi$ -cation with TYR-85 and  $\pi$ - $\pi$  stacked with TYR-85.

### Structured-based drug design

As reported in Table 11, the predicted  $pIC_{50}$  values of most of the designed compounds (N3, N7–N12) were greater than those of their corresponding template molecules, an evidence of being more biologically active molecules than their templates. Also, all the newly designed compounds showed higher docking scores compared to the template molecules, while only N7 and N12 with binding energy values of  $-8.9$  and  $-8.5$  kcal/mol showed better docking scores than that of the reference drug (pentamidine). This is a clear indication of how well these new maleimides would interact with the target enzyme. Because N7 and N12 bind more strongly to PdxK than pentamidine, their interactions with the binding pocket of the protein target shall be discussed further.

N7 was observed from Fig. 9 to bind excellently with the active site of PdxK via six (6) conventional hydrogen bonds, one (1)  $\pi$ -alkyl hydrophobic interaction and one (1)  $\pi$ -anion electrostatic interaction. The carbonyl groups oxygen atoms play a vital role in hydrogen bond formation just as observed with the template molecules. One of the carbonyl groups formed hydrogen bond with ASN-151 and LYS-187 via its oxygen at distance of 2.32 Å and 2.11 Å, respectively. The other carbonyl group formed a hydrogen bond with GLY-228 at a distance of 2.40 Å and 2 hydrogen bonds with THR-229 at a distance of 2.38 Å and 2.40 Å. It also formed an additional hydrogen bond with GLN-258 at a distance of 1.98 Å via a hydroxyl group at the para-position of the outer phenyl ring system. Others include  $\pi$ -alkyl hydrophobic interaction with LEU-198 via the compound's  $\pi$  electrons of the outer benzene ring, and  $\pi$ -anion electrostatic interaction with ASP-124 via the compound's  $\pi$  electrons of the pyrrole ring system. The interaction of N12 with the protein target as observed from Fig. 10 was via three (3) conventional hydrogen bonds, three (3) carbon-hydrogen bonds and one (1)  $\pi$ -donor hydrogen bond. The observed hydrophobic interactions include one each of alkyl and  $\pi$ -alkyl interactions, while the electrostatic interactions include one each of  $\pi$ -anion and  $\pi$ -cation interactions. Also, observed was a halogen interaction with the receptor. The observed conventional hydrogen bonds were formed by the interaction of its carbonyl groups oxygen with LYS-187 and THR-229 at distance of 2.06 Å and 2.46 Å, respectively, and between its nitro group oxygen and SER-188 at a distance of 2.64 Å. The carbon-hydrogen bonds were formed by its trifluoromethyl group's interaction with GLY-230, LYS-187 and THR-229 at interaction distance of 2.71 Å, 3.04 Å and 2.38 Å, respectively. Also observed was a  $\pi$ -donor hydrogen bond between the  $\pi$ -electron of the pyrrole ring system and ASN-151 at a distance of 3.19 Å. The observed hydrophobic interactions include alkyl interaction with

LYS-187 and VAL-121, and  $\pi$ -alkyl interaction with LYS-187. Furthermore, the observed electrostatic interactions comprises of  $\pi$ -cation between LYS-187 and its benzene ring  $\pi$ -system, and  $\pi$ -anion between ASP-124 and its pyrrole ring  $\pi$ -system. Therefore, these compounds have demonstrated the potentials to arrest the target receptor (PdxK) an absolutely necessary factor whose inactiveness is dangerous to the viability of the parasite (*L. donovani*).

### Pharmacokinetics properties prediction

According to the Lipinski's rule for oral bioavailability, a drug molecule is more likely to have poor absorption or permeation when it has hydrogen bond donors (HBD) of greater than 5, hydrogen bond acceptors (HBA)  $>10$ , molecular weight (MW)  $>500$  and lipophilicity (MLOGP  $>4.15$  or WLOGP  $>5$ ) (Lipinski et al. 2001). Molecules that satisfy at least three out of the four requirements are said to obey the Lipinski's rule for oral-bioavailability (Lawal et al. 2021).

As seen from Table 13, all the designed molecules perfectly obeyed the Lipinski's rule by showing no violation. Also, the reported values of topological polar surface area (TPSA) for all molecules were less than 140 Å<sup>2</sup>, beyond which a molecule may exhibit poor gastrointestinal absorption. Additionally, the synthetic accessibility scores of all tested molecules are in the easy portion ( $<5.00$ ), indicating their easy laboratory synthesis. The estimated water solubility (Log S) ranges from moderately soluble (N5, N8 and N12) to soluble (the remaining 9 compounds). The Boiled Egg representation in Fig. 11 provides for an intuitive evaluation of passive gastrointestinal absorption and brain barrier permeability as a function of the position of the molecules in the WLOGP against TPSA plot (Daina et al. 2017). As seen from Fig. 11, all the designed molecules were represented in red dots, indicating that they were predicted not to be effluated from the central nervous system by P-glycoprotein. P-glycoprotein is an enzyme which acts as a biological barrier by extruding toxins and xenobiotics, including drugs out of cells. Also, six (6) molecules (N1, N6, N7, N9, N10 and N12) were located in the Boiled Egg's yolk, meaning that these molecules were predicted to passively permeate through the blood–brain barrier (BBB), while the remaining six (6) molecules were located in the Boiled Egg's white which is an indication that they were predicted to be passively absorbed by the gastrointestinal tract.

The predicted ADMET properties in Table 14 showed that the human intestinal absorption (HIA) was high ( $>70\%$ ) for all newly designed compounds. Drug molecules are said to be poorly distributed to the brain through the blood–brain barrier (BBB) and considered as unable to penetrate the central nervous system

(CNS), when the values of the logarithmic ratio of brain to plasma drug concentration (logBB) are less than  $-1$ , and the blood–brain permeability-surface area product (logPS) are less than  $-3$ , respectively. Consequently, all the newly designed compounds were predicted to cross the BBB except **N7**, **N10** and **N12** with predicted logBB values of less than  $-1$  (Table 14). Also as predicted in Table 14, all the molecules showed CNS permeability, i.e.,  $\logPS > -3$ . Furthermore, all the studied molecules except **N7** and **N11** are substrates of the Cytochrome P450 enzyme (CYP-3A4), an important enzyme for drug metabolism in the body, with none of the compounds inhibiting the enzyme. The degree of drug elimination from the body is measured by the drug's total clearance, which is within the accepted range for these newly designed compounds. Additionally, the predicted values of maximum recommended tolerated dose (MRTD) for all molecules are included in Table 14. MRTD value of less than or equal to  $0.477 \log$  (mg/kg/day) is considered low, and high if greater than  $0.477 \log$  (mg/kg/day). The overall predicted drug-likeness and ADMET properties put these molecules on an excellent pharmacokinetics profile, and more so that they are orally bio-available. Therefore, the newly designed molecules are suggested for practical evaluation and/or validation in the laboratory as potential drug candidates for the treatment of leishmaniasis.

## Conclusions

In this study, a four-descriptor QSAR model was developed with a series of twenty-eight (28) maleimides, which was used to excellently predict their anti-leishmanial activities, and those of the newly designed compounds. The virtual docking screening conducted between the 28 maleimides and five (5) target enzymes, identified pyridoxal kinase (PdxK) receptor as a biological target of interest. Compounds **14**, **21** and **24** were used as templates for designing twelve (12) new maleimides, because of their relatively higher binding energies of interaction with PdxK. The newly designed compounds interacted more strongly with the target site than the chosen template molecules, but only **N7** and **N12** were better than the reference drug (pentamidine) in this regard. The predicted pharmacological interaction profiles of the designed compounds generally showed a good fitting into the target site cavity. Also, all the newly designed compounds were said to be orally bio-available, and showed good drug-likeness and ADMET properties. Hence, these new molecules have demonstrated the potential to arrest PdxK enzyme, which could inhibit the growth and viability of the parasite (*L. donovani*). It is therefore recommended that, these new compounds be developed and further

evaluated as potential drug candidates for the treatment of leishmaniasis.

## Abbreviations

AD: Applicability domain; ADMET: Absorption, distribution, metabolism, excretion and toxicity; ASN: Asparagine; ASP: Aspartic acid; BBB: Blood–brain barrier; LogBB: Logarithmic ratio of brain to plasma drug concentration; B3LYP: Becke's three-parameter read-Yang-Parr hybrid; CNS: Central nervous system; CYP-3A4: Cytochrome p450 isoform; DFT: Density functional theory; DTC: Drug theoretical and cheminformatics; ESOL: Estimated solubility; LOF: Friedman lack-of-fit; GFA: Genetic function approximation; GLN: Glutamine; GLU: Glutamic acid; GLY: Glycine; HIA: Human intestinal absorption; h\*: Warning leverage; HBA: Hydrogen bond acceptor; HBD: Hydrogen bond donor; HIS: Histidine; ILE: Isoleucine; IC<sub>50</sub>: Half-maximal inhibitory concentration; LEU: Leucine; LYS: Lysine; MET: Methionine; MRTD: Maximum recommended tolerated dose; ME: Mean effect; MMFF: Molecular mechanics force field; MW: Molecular weight; MLR: Multi-linear regression; NTD: Neglected tropical diseases; NMR: Nuclear magnetic resonance; PaDEL: Pharmaceutical Data Exploration Laboratory; PDB: Protein data bank; pIC<sub>50</sub>: Negative log of IC<sub>50</sub>; PdxK: Pyridoxal kinase receptor; PHE: Phenylalanine; LogPS: Blood–brain permeability-surface area product; QSAR: Quantitative structure activity relationship; RO5: Rule of five; SBDD: Structure-based drug design; SEE: Standard error of estimation; SER: Serine; TPSA: Topological polar surface area; THR: Threonine; TYR: Tyrosine; VAL: Valine; VIF: Variance inflation factor; VL: Visceral leishmaniasis.

## Acknowledgements

The authors sincerely acknowledge G.F.S. Harrison Quantum Chemistry Research Group, Ahmadu Bello University Zaria, for providing all software and suitable environment utilized for this study.

## Author contributions

GAS, AU and FAU conceived and designed the study. FAU carried out the study and drafted the manuscript. IA conducted the technical editing. All authors read and approved the final manuscript.

## Funding

No funding was received for this study.

## Availability of data and materials

All data related to this study are included herein otherwise available on request.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

Received: 22 February 2022 Accepted: 22 March 2022

Published online: 05 April 2022

## References

- Abdelhameed A (2018) The design and synthesis of cyanines and arylimidamideazole hybrids as antileishmanial agents, PhD dissertation The Ohio State University. <https://gradsch.osu.edu/calendar/events/design-and-synthesis-cyanines-and-arylimidamide-azole-hybrids-antileishmanial-agents>. Accessed 11 Feb 2022
- Abdullahi A, Shallangwa GA, Ibrahim MT, Bello AU, Arthur DE, Uzairu A, Mamza P (2019) QSAR studies on some C14-urea tetrandrine compounds as potent anti-cancer agents against Leukemia cell line (K562). *JOTCSA* 5(3):1391–1402

- Abdullahi SA, Uzairu A, Shallangwa GA, Uba S, Umar AB (2022) 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 Nat Res Centre* 46:2
- Adawara SN, Shallangwa GA, Mamza PA, Ibrahim A (2020) Molecular docking and QSAR theoretical model for prediction of phthalazinone derivatives as new class of potent dengue virus inhibitors. *Beni-Suef Univ J Basic Appl Sci* 9(50)
- Adeniji SE, Uba S, Uzairu A (2018) QSAR modeling and molecular docking analysis of some active compounds against mycobacterium tuberculosis receptor (Mtb CYP121). *J Pathogens Hindawi* 2018:24–64
- Adeniji SE, Uba S, Uzairu A (2019) Activity modeling of some potent inhibitors against mycobacterium tuberculosis using genetic function approximation approach. *Adiyaman Univ J Sci* 9(1):77–98
- Adeniji SE, Arthur DE, Abdullahi A, Ugbe FA (2020) Computer-aided modeling of triazole analogues, docking studies of the compounds on DNA gyrase enzyme and design of new hypothetical compounds with efficient activities. *J Biomol Struct Dyn*
- Are S, Gatreddi S, Jakkula P, Qureshi IA (2020) Structural attributes and substrate specificity of pyridoxal kinase from *Leishmania donovani*. *Int J Biol Macromol* 152:812–827
- Arthur DE, Uzairu A, Mamza P, Abechi SE, Shallangwa GA (2020) Activity and toxicity modeling of some NCI selected compounds against leukemia P388ADR cell line using genetic algorithm-multiple linear regressions. *J King Saud Univ Sci* 32(1):324–331
- Baquedano Y, Alcolea V, Toro MA, Gutiérrez KJ, Nguewa P, Font M, Moreno E, Espuelas S, Jiménez-Ruiz A, Palop JA, Plano D, Sanmartín C (2016) Novel heteroaryl selenocyanates and diselenides as potent antileishmanial agents. *Antimicrob Agents Chemother* 60:3802–3812
- Batool M, Ahmad B, Choi S (2019) A structure-based drug discovery paradigm. *Int J Mol Sci* 20(11):2783
- Chen XL, Zhang LJ, Li FG, Fan YX, Wang WP, Li BJ, Shen YC (2015) Synthesis and antifungal evaluation of a series of maleimides. *Pest Manag Sci* 71:433–440
- Daina A, Michielin O, Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, druglikeness and medicinal chemistry friendliness of small molecules. *Sci Rep* 7:42717
- Edache EI, Samuel H, Sulyman YI, Arinze O, Ayine OI (2020) QSAR and molecular docking analysis of substituted tetraketone and benzyl-benzoate analogs as anti-tyrosine: A novel approach to anti-tyrosine kinase drug design and discovery. *Chem Res J* 5(6):79–100
- Fan Y, Lu Y, Chen X, Tekwani B, Li X, Shen Y (2018) Anti-Leishmanial and cytotoxic activities of a series of maleimides: synthesis, biological evaluation and structure-activity relationship. *Molecules* 23:2878
- French JB, Yates PA, Soysa DR, Boitz JM, Carter NS, Chang B, Ullman B, Ealick SE (2011) The *Leishmania donovani* UMP synthase is essential for promastigote viability and has an unusual tetrameric structure that exhibits substrate-controlled oligomerization. *J Biol Chem* 286:20930–20941
- Gatreddi S, Pillalamarr V, Vasudevan D, Addlagatta A, Qureshi IA (2019) Unraveling structural insights of ribokinase from *Leishmania donovani*. *Int J Biol Macromol* 136:253–265
- Ghorbani M, Farhoudi R (2018) Leishmaniasis in humans: Drug or vaccine therapy? *Drug Des Devel Ther* 12:25–40
- Ibrahim MT, Uzairu A, Shallangwa GA, Uba S (2020) Lead identification of some anti-cancer agents with prominent activity against Non-small Cell Lung Cancer (NSCLC) and structure-based design. *Chem Afr* 3:1023–1044
- Ibrahim MT, Uzairu A, Uba S, Shallangwa GA (2021) Design of more potent quinazoline derivative EGFRWT inhibitors for the treatment of NSCLC: a computational approach. *J Pharm Sci* 7:140
- Isyaku Y, Uzairu A, Uba S, Ibrahim MT, Umar AB (2020) QSAR, molecular docking, and design of novel 4-(N, N-diarylmethyl amines) Furan-2(5H)-one derivatives as insecticides against *Aphis craccivora*. *Bull Natl Res Centre* 44:44
- Jerzy K (2003) Synthesis of new N-substituted cyclic imides with potential anxiolytic activity. XXV derivatives of halogenodibenzo(e,h)bicyclo(2.2.2) octane-2,3-dicarboximide. *Acta Pol Pharm* 60:1183–1189
- Kennard RW, Stone LA (1969) Computer aided design of experiments. *Technometrics* 11:137–148
- Keurulainen L, Siiskonen A, Nasereddin A, Kopelyanskiy D, Sacerdoti-Sierra N, Leino TO, Tammela P, Yli-Kauhaluoma J, Jaffe CL, Kiuru P (2018) Synthesis and biological evaluation of 2-arylbenzimidazoles targeting *Leishmania donovani*. *Bioorg Med Chem Lett*
- Khan MI, Baloch MK, Ashfaq M (2004) Biological aspects of new organotin (IV) compounds of 3-maleimidopropionic acid. *J Organomet Chem* 689:3370–3378
- Kimura T, Shirakawa R, Yaoita N, Hayashi T, Nagano K, Horiuchi H (2014) The antimalarial drugs chloroquine and primaquine inhibit pyridoxal kinase, an essential enzyme for vitamin B6 production. *FEBS Lett* 588:3673–3676
- Kumar V, Sharma M, Rakesh BR, Malik CK, Neelagiri S, Neerupudi KB, Garg P, Singh S (2018) Pyridoxal kinase: a vitamin B6 salvage pathway enzyme from *Leishmania donovani*. *Int J Biol Macromol* 119:320–334
- Lawal HA, Uzairu A, Uba S, (2021) QSAR, molecular docking studies, ligand-based design and pharmacokinetic analysis on Maternal Embryonic Leucine Zipper Kinase (MELK) inhibitors as potential anti-triple-negative breast cancer (MDA-MB-231 cell line) drug compounds. *Bull Natl Res Centre* 45(90)
- Li Z, Wan H, Shi Y, Ouyang P (2004) Personal experience with four kinds of chemical structure drawing software: review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. *J Chem Inf Comput Sci* 44:1886–1890
- Li W, Fan YX, Shen ZZ, Chen XL, Shen YC (2012) Antifungal activity of simple compounds with maleic anhydride or dimethylmaleimide structure against *Botrytis cinerea*. *J Pest Sci* 37:247–251
- Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46:3–26
- Mahmud AW, Shallangwa GA, Uzairu A (2020) In silico studies of 2,5-disubstituted furans as active antimalarial drug candidates. *Bull Natl Res Centre* 44:77
- Nara LM, Gislaine F, Carla S (2010) N-antipyrine-3,4-dichloromaleimide, an effective cyclic imide for the treatment of chronic pain: the role of the glutamatergic system. *Anesth Analg* 110:942–950
- Roy K, Chakraborty P, Mitra I, Ojha PK, Kar S, Das RN (2013) Some case studies on application of “rm2” metrics for judging quality of quantitative structure–activity relationship predictions: Emphasis on scaling of response data. *J Comput Chem* 34(12):1071–1082
- Seifert K (2011) Structures, targets and recent approaches in anti-leishmanial drug discovery and development. *Open Med Chem J* 5:31–39
- Shen ZZ, Fan YX, Li FG, Chen XL, Shen YC (2013) Synthesis of a series of N-substituted dimethylmaleimides and their antifungal activities against *Sclerotinia sclerotiorum*. *J Pest Sci* 86:353–360
- Silvia NL, Maria VC (2005) In vitro antifungal properties, structure-activity relationships and studies on the mode of action of N-phenyl, N-aryl, N-phenylalkyl maleimides and related compounds. *Arzneim-Forsch* 55:123–132
- Slavica A, Dib I, Nidetzky B (2007) Selective modification of surface-exposed thiol groups in *Trigonopsis variabilis* D-amino acid oxidase using poly (ethylene glycol) maleimide and its effect on activity and stability of the enzyme. *Biotechnol Bioeng* 96:9–17
- Soufari H, Waltz F, Parrot C, Durrieu-Gaillard S, Bochler A, Kuhn L, Sissler M, Hashem Y (2020) Structure of the mature kinetoplasts mitoribosome and insights into its large subunit biogenesis. *Proc Natl Acad Sci USA* 117:29851–29861
- Sun Y, Yang AW, Hung A, Lenon GB (2020) Screening for a potential therapeutic agent from the herbal formula in the 4th edition of the Chinese national guidelines for the initial-stage management of COVID-19 via molecular docking. *Evid Based Complement Altern Med* 2020:3219840
- Tonelli M, Gabriele E, Piazza F, Basilio N, Parapini S, Tasso B, Loddò R, Sparatore F, Sparatore A (2018) Benzimidazole derivatives endowed with potent antileishmanial activity. *J Enzyme Inhib Med Chem* 33(1):210–226
- 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
- Ugbe FA, Shallangwa GA, Uzairu A, Abdulkadir I (2021) Activity modeling, molecular docking and pharmacokinetic studies of some boron-pleuromutilins as anti-wolbachia agents with potential for treatment of filarial diseases. *Chem Data Collect* 36:100783
- Veerasamy R, Rajak H, Jain A, Sivadasan S, Varghese CP, Agrawal RK (2011) Validation of QSAR models-strategies and importance. *Int J Drug Des Discov* 3:511–519
- Wang X, Dong H, Qin Q (2020) QSAR models on aminopyrazole-substituted resorcylic acid compounds as Hsp90 inhibitors. *J Comput Sci Eng* 48:1146–1156

## Publisher's Note

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