The Journal of Engineering and Exact Sciences – jCEC, Vol. 06 N. 02 (2020)
journal homepage: https://periodicos.ufv.br/ojs/jcec
doi: 10.18540/jcecvl6iss2pp0107-0128
OPEN ACCESS – ISSN: 2527-1075

THEORETICAL AND INSILCO PHARMACOKINETIC INVESTIGATIONS ON SOME PHENYL PIPERIDINE DERIVATIVES AS NOVEL ANTIDEPRESSANT AGENTS

S. B. OLASUPO[1,*], A. UZAIRU[2], G. A SHALLANGWA[2] and S. UBA[2]

[1]National Agency for Food and Drug Administration and Control (NAFDAC)
[2]Department of Chemistry, Ahmadu Bello University Zaria, Nigeria

[*]*Corresponding author. National Agency for Food and Drug Administration and Control (NAFDAC), Abuja, Nigeria*
e-mail addressl olasabit@yahoo.com (S. B. OLASUPO).

A R T I C L E I N F O

Article history:
Received 2019-04-04
Accepted 2020-06-17
Available online 2020-06-17

k e y w o r d s
Serotonin
Antipsychotic
QSAR
Descriptors
Model

A B S T R A C T

A theoretical and Insilco pharmacokinetic studies were carried out on some Phenyl piperidine derivatives using Density Functional Theory (DFT/B3LYP/6-31G) with Spartan 14 V1.1.4 software to investigate the antipsychotic activity of the compounds. PaDEL-Descriptor software 2.20 version was utilized to generate molecular descriptors while Genetic Function Algorithm (GFA) was used for variable selections to develop Penta-parametric Multi-linear regression models. The statistical parameters of the best model (R^2_{Train}= 0.8572, R^2_{adj} = 0.8274, R^2_{Test} = 0.678, Q2cv (LOO) = 0.7664, □^{2}= 0.0036, r^2_m (LOO)= 0.694, cR^2_p= 0.763, RMSE= 0.168 and Delta r^2_m (LOO)= 0.0051) revealed that the model was predictive, robust and possessed good quality. Similarly, the descriptors (AATS8v, GATS1e, SpMAD_Dzs, SP-7 and RDF135v) were found to influence the inhibitory activity of the compounds. Likewise, descriptors SpMAD_Dzs (38.94%) with positive correlation and SP-7 (33.17%) with negative correlation showed predominant influences on the observed activity of the compounds evidenced by their highest percentage contributions. The model proved to be reliable, stable and could be accepted because it satisfied the general requirements for QSAR model development. More so, Insilco Pharmacokinetics and ADMET Risk screenings showed that four compounds (1,2,5 and 6) possessed exceptional good distribution profiles with low ADMET Risk. Consequently, the obtained results are envisaged to provide a rationale blueprint for the structural requirements for the development of novel Phenyl piperidine analogues as potent antidepressant agents.*

1. INTRODUCTION

The inability to instinctively sustain, withhold, or modify adaptive behavior in response to varying situational demands could be referred to as cognitive inflexibility. This is highly correlated with diverse psychiatric disorders such as depression, schizophrenia and obsessive-compulsive disorders (Lin *et al.*, 2017; Boulougouris, Glennon and Robbins, 2008).

Cognitive dysfunction appears to be an independent and core domain of depression which may lead to reduce life quality of depression patients (Poyurovsky et al., 2003). Depression is a serious and common disorder including symptoms like a feeling of sadness, hopelessness, weight loss or gain, tiredness, changes in sleeping routine and thinking of suicide(Kaya et al., 2017). Depression is a major public health problem and the fourth cause of the global burden of disease(Kaya *et al.*, 2017;Khattab *et al.*, 2015).

Depression constitutes a serious threat to the health of approximately 121 million people and thus is among the top five primary causes of disease and disability burden all over the world and also, cognitive dysfunction has been recognized as a leading cause of morbidity and mortality throughout the world (Lin et al., 2017)

The serotoninergic system is connected in the control of various physiological and behavioral activities; thus, it is known to regulate emotion, mood and appetite. Reduced serotoninergic neurotransmission has been proposed to play a key role in the etiology of depression. Drugs inhibiting serotonin transport have been very useful for the treatment of depression because, the concentration of synaptic serotonin is regulated by its reuptake into the presynaptic terminal. Highly specific serotonin reuptake inhibitors such as Ritanserin, YM992, M100907, LY367265, paroxetine, and nefazodone have been developed in the treatment of psychotic disorder and are increasingly prescribed for depressed patients(Lin *et al.*, 2017; Barr *et al.*, 2004). Despite many developments in the

field of antidepressants, the clinical use of currently used drugs was restricted as a result of various adverse effects and a response in less than 50% of patients (Fishback, Robson, Xu, & Matsumoto, 2010). Thus, the search for new class of more effective and safer antidepressants has become a sine qua non. With advances in computer technology, chemical biology, and molecular biology, computer simulation technology plays a prominent role in the growth of new agents (LIN *et al.*, 2016;Gao, Han and Ren, 2016). Computer-aided drug design (CADD) can greatly raise the efficiency of developing and designing novel drugs, and thus has been used tremendously in the present pharmaceutical industry(Wang, Yang, Li, & Wang, 2016). In essence, CADD approaches, such as the 3D-QSAR and molecular docking have been vastly conducted in the optimization and development of inhibitors(Yang et al., 2016) in which 3D-QSAR modeling has proven its efficiency in exploring the pharmacological properties of the studied molecules in modern drug discovery(Wu et al., 2014). The 3D-QSAR methods have been successfully utilized to obtain insights into the structural requirements that affect their biological activity for many series of molecules(Yang et al., 2016).

In the present study, bioinformatic investigation via the Quantitative Structure-Activity Relationship (QSAR) approach and Insilco Pharmacokinetics studies using ADMET PredictorTM software (SimulationPlus Inc., USA) and MedChem DesignerTM software were procured to harness the crucial structural features of some Phenyl piperidine derivatives to develop a mathematical model that could quantitatively define the relationship between the structures of these compounds and their molecular descriptors and evaluate the Insilco Pharmacokinetics properties of the compounds respectively. The results of this research are hoped to guide rational structural modification and design of novel and more potent Phenyl piperidine analogues that would offer some reference for further experimental study.

2. MATERIALS AND COMPUTATIONAL METHODS

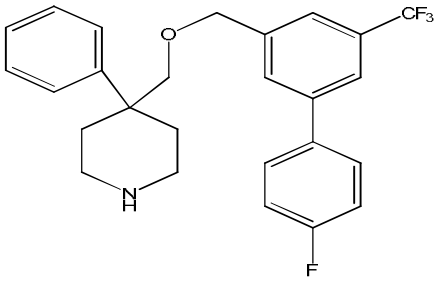
2.1. Preparation and Division of Experimental Data set

A total of 43 experimental data set of Phenyl piperidine derivatives with reported inhibitory activity (IC50 values) against the serotonin transporter (SERT) sourced from the literature were used in this present study(Zare, Fereidoonnezhad, Afshar, & Ramezani, 2017). The inhibitory activity (IC_{50} values) of the experimental data set compounds were converted into consistent pIC_{50} (-logIC50) values for improving the normal distribution of the experimental data points defined as the dependent variables in the QSAR modeling. The chemical names, 2D chemical structures and inhibitory activity of the compounds (Phenyl piperidine derivatives) is presented in Supplementary Table S1.

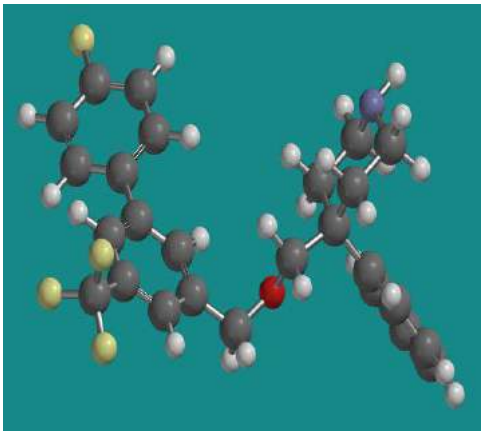
The experimental data set (Phenyl piperidine derivatives) was split into 70% training set (30 compounds) and 30% test set (13 compounds) in line with the optimum splitting pattern of a data set in QSAR study(Patil, 2012) using ñDataset Division GUI 1.2ò software based on Kennard stone algorithm technique(Kennard & Stone, 1969). The training set was used for the QSAR model development while the test set compounds were employed for the external validation of the model(S. B. Olasupo, Uzairu, Shallangwa, & Uba, 2019).

2.2 Molecule standardization and optimization

The ChemDraw software ultra-version 12.0 used to draw 2D structures of the compounds while the Spartan 14 V1.1.4 wavefunction software package was employed to optimize molecular geometries of the molecules (Hehre & Huang, 1995). The 2D (two-dimensional) structures were converted into 3D (three-dimensional) structures by importing it into Spartan 14 V1.1.4 wavefunction software for energy minimization in two steps. In the first phase, the converted 3D structures were minimized by Molecular Mechanics Force Field (MMFF) optimization to remove strain energy. In the second step, the minimized MMFF molecules were re-optimized via the Spartan wavefunction software package for complete geometry optimization with the aid of Density Functional Theory (DFT) by the B3LYP/6-31G* basis set(Bauernschmitt & Ahlrichs, 1996) as to compute the molecular descriptors of the molecules. Figure 1 depicts 2D and 3D chemical structures of compound 28.



(a)



(b)

Figure 1- (a) 2D and (b) 3D chemical structures of compound 28.

2.3 Descriptors calculations

The molecular descriptors which include physicochemical, atom type count used in constructing the models were calculated for each molecule by using the PaDEL-Descriptor software 2.20 version tool kit (Yap, 2011) and Spartan 14 software to generate thousands of descriptors comprising of 0D, 1D, 2D, and 3D types for each molecule. The highly correlated descriptors were removed in other to select the best subset of descriptors. Pearsonôs correlation matrix was used to select the suitable descriptors for Genetic

Function Approximation (GFA) analysis based on the correlation coefficients. The details of the descriptors and correlation matrix for the developed QSAR model are presented in Table 2 and Table 3 respectively.

2.4 MLR-Based Model Development

Multiple Linear Regression (MLR) method with a statistical analysis of the Genetic Function Approximation (GFA) approach via a Material studio software 8.0 version was utilized to construct the QSAR models. The MLR-based QSAR models were developed by using the training sets with the experimentally determined inhibitory activity on the logarithmic scale (pIC50) as the dependent variable and the descriptors as the independent variable. The MLR-GFA method has a good attribute to produce a population of model equations compare to other statistical methods that yield only a singular model (S. Olasupo, Uzairu, & Sagagi, 2016). Also, unlike the stepwise regression technique, the GFA technique generates better models and selects the basic function genetically(Olasupo *et al.*, 2019). Out of the three statistically significant developed GFA models, the best model (Model 1) was selected based on the one with the smallest Friedmanôs Lack of Fit (LOF) score. The use of Friedmanôs lack-of-fit (LOF) measure is because it has several advantages over the regular least square error measure and it is estimated measured using a slight variation of the original Friedman formula in the Materials Studio software (Ameji, Uzairu, & Idris, 2015). LOF is computed via this revised mathematical formula:

$$LOF = \frac{SSE}{(1-\frac{c-dp}{M})^2} \qquad (1)$$

SSE denotes the sum of squares of errors, c is the number of terms (basic functions) in the model, other than the constant term, d represents user-defined smoothing parameter, p represents the total number of descriptors made up in the model and M is the number of samples in the training data set..

2.5 Statistical evaluation and validation of the Model

A rigorous validation is an important, integral component of QSAR model development (Veerasamy et al., 2011). Generally, validation of a QSAR model does not only rely on the evaluation of statistical fitness and predictiveness of a model using cross-validation technique alone but also more significantly is the assessment of data quality (a well-defined End-point and unambiguous algorithm), domain of applicability, mechanistic interpretability and appropriate statistical evaluations (using the training set and the test set for internal and external validations respectively) of the models. This is a conceptual framework that constitutes the principles of the Organization for Economic Co-operation and Development (OECD) for validating a QSAR model(Leonard and Roy, 2006;OECD, 2007)

In search to build a stable, predictive, reliable, robust and acceptable QSAR model that is also satisfied OECD principles of model validations, different types of validations techniques were procured and the best model was selected/ affirmed by applying these various statistical parameters. The procedures employed for both the internal and external validations techniques include; R^2_{Pred} (R^2 for external test set), R^2 (squared of correlation coefýcient for training set), Q2 (cross-validated correlation coefýcient), root-mean-squared error (RMSE), chi-squared, degree of freedom (DF), F test (Fischerôs value) for statistical relevance; CR2P (coefficient of determination for Y-Randomization/scrambling), Domain of applicability in the chemical space (AD) and the evaluation of mechanistic associations between the descriptors that revealed in the model and the predicted endpoint (interpretation of the model).

2.6 Procured Validations Procedures

Some of the procedures and procured validation techniques employed in this study are shown via the Equations 2-7 and the statistical validations parameters alongside with the threshold values (Veerasamy et al., 2011) for the assessment of a QSAR model as a reliable screening tool in practical applications are presented in Table 1.

Least squares fit (R^2): R^2 (squared correlation coefficient) for the comparison between the predicted and experimental activities. A value of R^2 closes to 1.0 indicates the goodness of fit and R^2 is expressed mathematically as,

$$R^2 = 1 - \left[\frac{\sum(Y_{exp}-Y_{pred})^2}{\sum(Y_{exp}-Y_{training})^2}\right] \qquad (2)$$

Ypred, Yexp and training indicate the predicted, experimental, and mean values of experimental activity of training set.

Cross-validation coefficient (Q^2): The Leave-one-out (LOO) cross-validated coefficient (Q^2) is given by this formula;

$$Q^2 = 1 - \frac{\sum(A-B)^2}{\sum(B-C)^2} \qquad (3)$$

A and B indicate the predicted and experimental activity respectively of the training set and C is the mean activity value of the training set.

Adjusted R^2 (R^2_{adj}):

$$R^2_{adj} = \frac{R^2-D(M-1)}{M-D+1} \qquad (4)$$

D represents the number of descriptors and M indicates the number of molecules in the training set.

Chi-squared and Root-mean squared error (RMSE)

$$X^2 = \sum_{i=1}^{n}\left(\frac{(y_i-\hat{y}_i)^2}{\hat{y}_i}\right) \qquad (5)$$

$$RMSE = \sqrt{\left(\sum_{i=1}^{n}\frac{(\hat{y}_i-y_m)^2}{n-1}\right)} \qquad (6)$$

y and R̂ indicate the experimental and predicted activity for each compound in the training set, ym is the mean of the experimental activities, and n is the number of molecules of the data set

Predicted R^2_{Ext}: The predicted R^2_{Ext} for the external validation could be computed by using this equation

$$R^2_{Ext} = 1 - \frac{\sum[W-T]^2}{\sum(T-X)^2} \quad (7)$$

W and T represent predicted and experimental activity values of the test set molecules and X indicates the mean activity value of the training dataset.

Variance inflation factors (VIF): To verify multi-collinearity (orthogonality) among the descriptor, VIF is determined via this mathematical expression;

$$VIF = \frac{1}{1-R^2} \quad (8)$$

R^2 is the correlation coefficient of the multiple regression between the variables in the model.

MLR Y-Randomization: Y-randomization (Scrambling test) was performed on the training set data (Table 4) by permuting the activity values to the selected descriptors matrix via ñMLR Y-Randomization Test 1.2 ò software sourced from DTC Lab software (Myers, 1990). The Coefficient of determination (cR^2_p) when computed should have a value greater than 0.5 to pass the Y-randomization test. Coefficient of determination (cR^2_p) is defined by this equation formula;

$$C\square^2_{\square} = \square * (\square^2 - (\square\square\square\square\square\square\square\square)^2)^{1/2} \quad (9)$$

where c is Coefficient of determination, Average Rr is average óRô of random models

2.6.1 Defining Domain of Applicability of the Model

It is important to note that before a QSAR model is used for screening of chemicals, its domain of applicability must be correctly defined and predictions for only molecules that lie within the domain could be considered as reliable chemical candidates, no matter how statistically significant, robust and validated is the model (Tropsha, Gramatica, & Gombar, 2003). To evaluate and define the domain of application of the model, the extent of extrapolation (leverage/ standardization approach) and similarity distance (Euclidean Based) methodologies of Applicability Domain (Figure 5 and Supplementary Table S2a &b) were employed in this research.

The leverage hi is defined as;

$$H_i = x_i(X^T X)^{-1} X_i^T \quad (10)$$

Here the descriptor row is the vector of the considered compound i, hi is the n x k descriptor matrix of the training set for building the model.

The warning leverage (h*) is computed as;

$$h^* = \frac{3(p+1)}{n} \quad (11)$$

n is the number of training compounds and P depicts the number of predictor variables in the model. A model is adjudged to be reliably predicted if the leverage hi < h* for the investigated molecules and leverage greater than the warning leverage implies that the predicted response is due to significant extrapolation of the model and such model may not be reliable for practical use (Tropsha et al., 2003). With the aid of Williamôs plot (Figure 5), the significant area of the model in terms of chemical space can be evaluated and visualized.

The similarity distance (Euclidean Based) Applicability is given as:

$$\square\square\square = [\Sigma wk\,(x_k^i - x_k^j)^2]^{1/2} \quad (12)$$

where difference $x_k^i - x_k^j$ = distance of the test set compounds from the training set compounds wk = weighted vector corresponding to the importance of the k^{th} descriptor in the model calculated using auto-scaled descriptors x_k^i and x_k^j represent molecules from both the test set and training set (Netzeva et al., 2005).

2.7 In silico pharmacokinetics evaluation, Drug-likeness prediction and ADMET Risk screening of the compounds

Lipinskiôs rule of five was employed to assess the drug likeness or pharmacological active of the studied compounds in humans(Lipinski, 2016; Lipinski *et al.*, 1997). Also certain physicochemical properties were computed (Supplementary Table S3) for Insilco evaluation of pharmacokinetics of the compounds such as the quantitative estimation of drug-like properties, lipophilicity (S+logP), distribution (S+logD), bioavailability (%Fraction bioavailable), absorption (%Fraction absorbed), bloodï brain-barrier penetration (LogBB), Renal clearance, ADMET Risk, Toxicity Risk as well as human ether-a-go-go related gene (hERG_pIC50) by using the ADMET PredictorTM software (SimulationPlus Inc., USA) and MedChem DesignerTM software(Tareq Hassan Khan, 2010).

Table 1 - Chemometric validation parameters with their threshold values for accepting QSAR models as a reliable screening tool for practical applications.

S/N	*Statistical Parameter*	*Definition*	*Validation Type*	*Threshold value*	*Computed value*	*Implication*
1.	R^2 Training set	Co-efýcient of determination	Internal	Ó0.6	0.857	goodness-of-fit
2.	R^2_{Ext} (Test set)	Co-efýcient of determination of external and test set	External	Ó0.5	0.678	good predictivity
3.	R^2_{adj}	Adjusted R-squared	Internal	>0.6	0.827	goodness-of-fit

4.	Q^2_{cv}	Cross-Validation Co-efýcient	Internal	>0.5	0.766	pass
5.	R^2-Q^2_{cv}	Difference between R^2 and Q	Internal	Ò0.3	0.091	pass
6.	LOF	Friedman Lack of fit score	Internal	Minimal		statistically significance
7.	X^2	Chi-squared	Internal	<0.5	0.0036	good predictivity
8.	RMSE	Root-mean squared error	Internal	0.3	0.168	good predictivity
9.	$_cR^2p$	Coefficient of determination for ⍰-randomization	Random model	>0.5	0.763	Robustness
10.	VIF	Variance Inflation Factor		1ÒVIF Ò10		orthogonal and statistical significance.
11.	r^2_m(LOO)	predictability of the selected model	External	Ó0.5	0.694	good external prediction
12.	Delta r^2_m(LOO)	predictability of the selected model	External	<0.2	0.0051	good external prediction

Table 2 - Notations, symbols and definition of the descriptors in the developed model

S/N	Descriptor Notation	Descriptor Symbols	Definition
1	t	AATS8v	Broto-Moreau autocorrelation weighted by van der Waals volumes
2	w	GATS1e	Geary autocorrelation weighted by Sanderson electronegativities
3	x	SpMAD_Dzs	Spectral mean absolute deviation from Barysz matrix weighted by I-state
4	y	SP-7	Simple path, order 7
5	z	RDF135v	Radial distribution function weighted by relative mass

Table 3 - Pearson's correlation matrix and Statistical quality parameters of the Model

	AATS8v	*GATS1e*	*SpMAD_Dzs*	*SP-7*	*RDF135v*	*VIF*	t-statistics	p value	*% Contribution*
AATS8v	1					1.21	5.666	7.79E-06	9.68
GATS1e	-0.188	1				1.05	-6.797	4.97E-07	13.63
SpMAD_Dzs	0.436	-0.131	1			1.48	9.618	1.05E-09	38.94
SP-7	0.647	-0.386	0.727	1		2.09	-10.158	3.62E-10	33.17
RDF135v	0.357	-0.055	0.376	0.383	1	1.04	-3.783	0.0009	4.61

Table 4 - MLR Y-Randomization Test

Model	R	R^2	Q^2
Original	0.926	0.857	0.766
Random 1	0.509	0.259	-0.125
Random 2	0.582	0.339	-0.048
Random 3	0.575	0.331	0.0083
Random 4	0.333	0.111	-0.461
Random 5	0.340	0.116	-0.349
Random 6	0.324	0.105	-0.633
Random 7	0.355	0.126	-0.344
Random 8	0.371	0.138	-0.522
Random 9	0.314	0.099	-0.406
Random 10	0.506	0.256	-0.183

Random Models Parameters	
Average r :	0.421
Average r^2 :	0.188
Average Q^2 :	-0.306
cRp^2 :	0.763

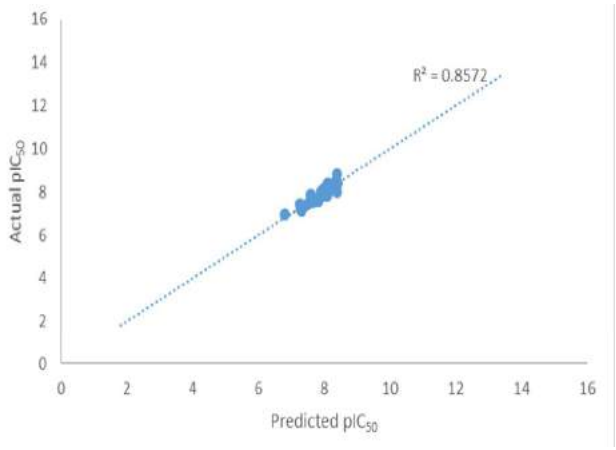


Figure 2 - Plot of (Actual) Experimental pIC_{50} against the predicted pIC_{50} of training set.

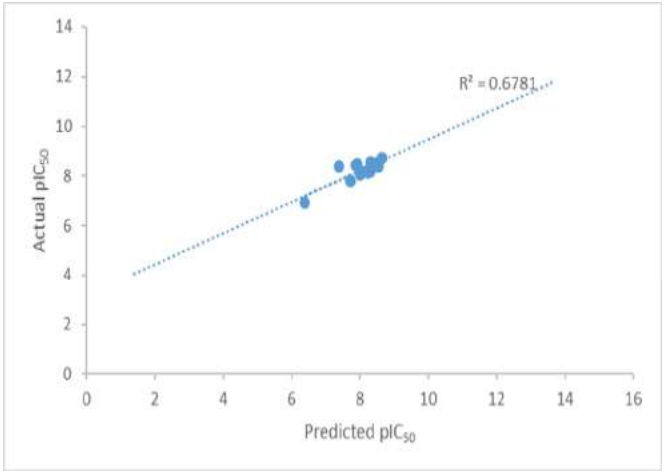


Figure 3 - Plot of (Actual) Experimental pIC_{50} against predicted pIC_{50} of test set.

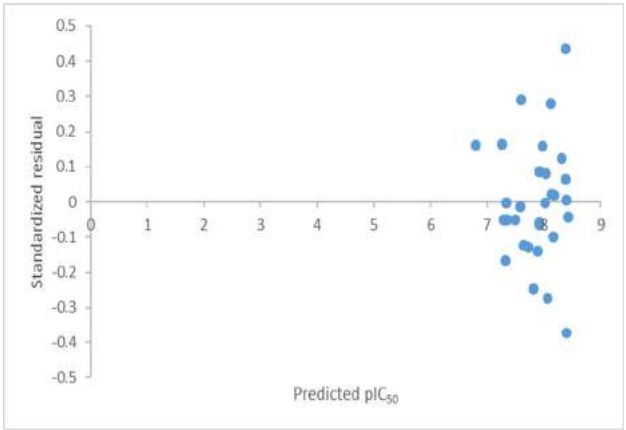


Figure 4 - Plot of standardized residual against predicted pIC_{50} of the compounds.

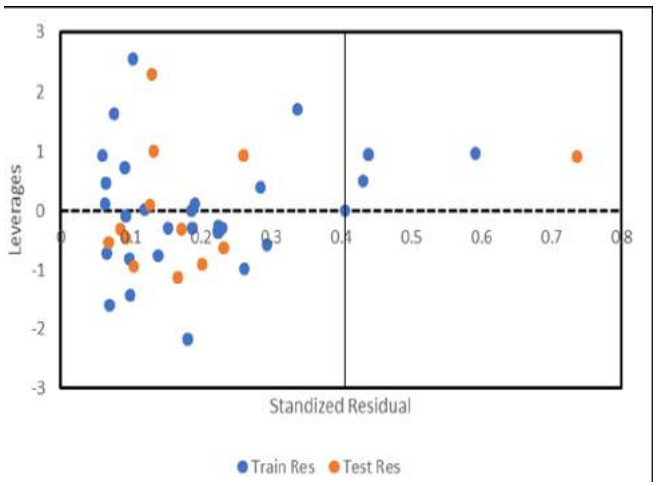


Figure 5 - William's plot of Model 1.

3. RESULTS AND DISCUSSION

3.1 Bioinformatic modeling and statistical analysis

Bioinformatic modeling was performed on 43 compounds of Phenyl piperidine derivatives reported being inhibitors of the serotonin transporter (SERT) to investigate and evaluate their antipsychotic properties for mood disorder. This led to the development of a Penta-parametric equation of three GFA models. Model 1 with the most statistically significant parameters (R^2_{Train}= 0.8572, R^2_{adj} = 0.8274, Q^2_{cv} = 0.7664, R^2_{Test} = 0.678, F-value = 28.81) was picked as the best model (Equation 13). The descriptor notations, symbols and their definition as it appears in the model are presented in Table 2.

Model 1:

pIC_{50} = 1.176655838t - 1.594327634w + 4.721202317x - 3.966993752y-0.554324221z+8.199016114 (13)

n = 43, Friedman LOF = 0.1578, R^2 = 0.8572, $R^2_{adj.}$ = 0.8274, Q^2 = 0.7664, F-value = 28.811.

Based on the significance of cheminformatic and statistical parameters for a reliable model, the selected (Model 1) gives the best GFA derived QSAR model for predicting the antipsychotic activity of the studied compounds vis-a-viz the validation parameters of the model. The reported cheminformatic parameters of the model (R^2_{Train}= 0.8572, R^2_{adj} = 0.8274, Q^2_{cv} = 0.7664, R^2_{Test} = 0.678, R^2-Q^2_{cv} = 0.091, ^{2}= 0.0036, r^2_m (LOO)= 0.694 and Delta r^2_m (LOO)= 0.0051) are all in agreement with acceptable validation parameters presented in Table 1 for a reliable QSAR model(Golbraikh & Tropsha, 2002). This suggests that the developed model is very predictive and reliable because the R^2_{Train}>0.5, , R^2_{Test} > 0.6, , Q^2_{cv} > 0.5 and r^2_m> 0.5(S. B. Olasupo et al., 2019). That is, the internal (R^2_{Train}) and external (R^2_{Test}) predictive ability of the model was 85.6% and 67.8% respectively. A plot of predicted pIC_{50} against experimental pIC_{50} gives an insight into how well the model was trained and how well it predicts the

antipsychotic activity of the test compounds. The correlation coefficient R^2 value for the plots of predicted against experimental activity for both training set and test set data is greater than 0.6 (Figures 2 and 3), this shows that the QSAR model produced a high antipsychotic activity-descriptors relationship accuracy (R^2_{Train}= 0.8572) and a good activity prediction accuracy (R^2_{Test} = 0.678), a good indication that there is a linear relationship between the experimental and predicted activity of the compounds. The computed difference between the R^2 and R^2_{adj} value (R^2- R^2_{adj}= 0.0298) is less than 0.3, this suggests that the number of descriptors involved in the developed model is good and acceptable (Veerasamy et al., 2011). The plot of standardized residual against predicted pIC_{50} of the compounds (Figure 4) shows that the model lack systematic error as the propagation of residuals was observed on both sides of zero (Adedirin, Uzairu, Shallangwa, & Abechi, 2018). The other statistical validation parameters (Table 1, 3, 4) like t-statistics (t-test >2), Chi-squared (X^2= 0.0036) and Root-mean squared error (RMSE= 0.168) for error checking, Variance Inflation Factor (VIF<10) to check for possible multi-collinearity among the descriptors and coefficient of determination for ?-randomization (cR^2_p= 0.763) to evaluate the robustness of the model are within the threshold limits and also statistically significant (Alam & Khan, 2018). The estimated value of VIF is less than 10 for each of the descriptors, the maximum correlation coefficient between a pair of descriptors is less than 0.7 and the coefficient of determination for ?-randomization is greater than 0.5 with low R^2 and Q^2 values. All these implied that the QSAR model lacked multi-co-linearity effect and reasonably orthogonal, the descriptors contributed significantly to the developed model and the model was not only robust but very stable (Tropsha et al., 2003). To ascertained the applicability and reliability of the developed model, the applicability domain of the model was further evaluated on the studied compounds (both the training set and test set) using the Euclidean based method and Leverage approach (Williamôs Plot) procedures to detect which of the compounds of training set and test set are within or outside the applicability domain and their outliers. The Williamôs Plot (a plot of standardized residuals against the leverage values) as shown in Figure 5 revealed that all compounds are within the applicability domain with warning leverage h*= 0.42 except four compounds, three from the training set (1,12 and 20) and one from the test set (13). Hence, those compounds (1,12, 13 and 20) are influential compounds that could be due to the obvious differences in their structures compare with the rest of the dataset. More so, most of the compounds had leverage lower than the warning h* value of 0.4, a good indication of an appreciably high applicability domain of the model. Furthermore, the computed Euclidean normalized mean distance scores for both the training set and test set compounds (Supplementary Table S2a and b) are found to be within the threshold boundaries of 0-1, showing that all the compounds (training set and test set) were fell within the acceptable domain of applicability (Alam & Khan, 2018). The results reported so far showed that the derived QSAR model displayed a good quality assurance for bioinformatic application of the model as a screen tool (Tropsha et al., 2003).

3.1 Mechanistic evaluation and Significant of the Descriptors in the model

The Penta-parametric equation derived model (Equation 13) revealed that five molecular descriptors (AATS8v, GATS1e, SpMAD_Dzs, SP-7 and RDF135v) were significantly correlated with the antipsychotic activity of the studied compounds. From the equation 13, two descriptors (AATS8v and SpMAD_Dzs) were found to be positively correlated with the antipsychotic activity of the compounds. This means that the antipsychotic activity of the compounds increases, whenever the value of any of this descriptor increases. Conversely, the descriptor GATS1e, SP-7 and RDF135v were negatively correlated with the antipsychotic activity of the compounds, a good implication that the antipsychotic activity decreases as the value of the descriptors increases. The descriptors were subjected to Pearsonôs correlation matrix and the percentage contribution of each descriptor was also computed as reported in Table 4. For each pair of descriptors, the value obtained for the Pearson correlation coefficients was less than 0.5, an indication of insignificant inter-correlation among the descriptors. Also, from the results (Table 4), the absolute t-statistics values for each descriptor is greater than 2 and p-values of all descriptors in the derived model are less than 0.05. These showed that the selected descriptors were good and that there is a significant connection between the descriptors and the inhibitory activities of the compounds (Tareq Hassan Khan, 2010). Furthermore, the computed percentage contribution for each descriptor to the observed inhibitory activities of the compounds are; AATS8v (9.68%), GATS1e (13.63%), SpMAD_Dzs (38.94%), SP-7 (33.17%), RDF135v (4.61%). The definition of the descriptors is presented in Table 2. Descriptor SpMAD_Dzs and SP-7 have demonstrated a pronounced influence on the observed antipsychotic bio-activity of the compounds.

The descriptor SpMAD_Dzs is defined as Spectral mean absolute deviation from Barysz matrix weighted by the ionization state of the molecule. It is a descriptor of the rate at which a neutral molecular is converted to electrically charged chemical species. Its positive correlation with pIC_{50} as shown in the model reveals that the higher its value in a molecule, the better its inhibitory activity of the compound against serotonin transporter. Thus, it could be inferred that the presence of easily ionizable substituents in a molecule would enhance its antipsychotic bioactivity. Also, the descriptor, SP-7 is defined as Simple path, order 7. Its negative correlation with pIC_{50} as shown in the model implies that the lower its value in a molecule, the higher the inhibitory activity against serotonin transporter. It is a descriptor of molecular connectivity. Structural features such as size and branching are encoded in it. Its calculation is based on the representation of molecular structures as graphs, where atoms are represented by vertices and covalent chemical bonds by edges. Hence, to enhanced antipsychotic bioactivity, larger substituents in the compounds could be substituted with smaller ones so as to minimize the size of the compounds.

3.2 Insilco Pharmacokinetics evaluation, Drug-likeness Assessment and ADMET Risk screening Results

Physicochemical property is an important attribute that influences efficacy, safety, metabolism and pharmacological activeness of a compound and it could be evaluated by applying Lipinskiôs rule of five and other Insilco methods. The result (Supplementary Table S3) shows that all the compounds have no more than one violation of Lipinskiôs rule of five except compound 39 with two violations (i.e. Molecular weight= 522.47 and LogP= 6.51). This implies that most of the

compounds are pharmacological activity and exhibit properties that would make them orally active drug in human (Lipinski et al., 1997). Also, to evaluate lipophilicity of a substance and determine effective lipophilicity via the distribution of a molecule within the body, Insilco estimation of LogP and LogD are very significant (Triggle & Taylor, 2006). Log P value greater than 2 and Log D value between 1 and 3 are mostly considered for a drug-like molecule to cross the blood ï brain barrier (BBB) including CNS drugs(Waring, 2010). From our findings (Supplementary Table S3), all the compounds have LogP greater 2 but only four compounds (1,2,5 and 6) have the computed values for LogD less than 3, ADMET_Risk less than 7.0 and hERG toxicity endpoints (pIC50) less than 6 including Brexpiprazole (Standard Drug) with toxicity endpoints slightly above 6 (pIC50= 6.82). This suggests that these four compounds (1,2,5 and 6) possessed exceptional good distribution profiles, excellent potential drug candidates and unlikely to exhibit some hERG toxicity property(Lagorce *et al.*, 2017;Gabbert and Weikard, 2013).

4. CONCLUSION

This study has offered a profound information via the computational modeling and Pharmacokinetic investigation. The developed model revealed that the antipsychotic properties of the compounds were influenced by AATS8v, GATS1e, SpMAD_Dzs, SP-7 and RDF135v molecular descriptors with the descriptors SpMAD_Dzs and SP-7 having a pronounced contributions on the observed bio-activity of the compounds evidenced by their highest percentage contributions and the direction of their influences. Likewise, the statistical diagnostic and the model validations showed that the derived QSAR model was good, predictive, robust, reliable, stable and could be accepted because it fulfilled the general requirements and the OECD Principles for model development. Similarly, the Insilco Pharmacokinetic evaluation, Drug-likeness Assessment and ADMET Risk screening portend that four of the compounds (1,2,5 and 6) possessed good quality assurance, high distribution profiles, pharmacological activeness and none of the four the compounds exhibit hERG toxicity. Hence, the obtained results would serve as a reliable blueprint for the structural requirements and physicochemical parameters needed to develop and design novel inhibitors of serotonin transporter as antidepressant agents with improved inhibitory properties.

Acknowledgment

We thank the theoretical and physical chemistry team of the chemistry department, Ahmadu Bello University Zaria. We sincerely appreciate David Arthur, Abdulateef Jimoh, Abdulfatai Usman and Jonh Philip Ameji for their technical support and advice in the course of this study.

SUPPLEMENTARY MATERIAL

Free access to the Supplementary Material at jCEC website.

REFERENCES

Adedirin, O., Uzairu, A., Shallangwa, G. A., & Abechi, S. E. (2018). Computational studies on Ŭ-aminoacetamide derivatives with anticonvulsant activities. *Beni-Suef University Journal of Basic and Applied Sciences*, *7*(4), 709ï 718.

Alam, S., & Khan, F. (2018). QSAR, docking, ADMET, and system pharmacology studies on tormentic acid derivatives for anticancer activity. *Journal of Biomolecular Structure and Dynamics*, *36*(9), 2373ï 2390.

Ameji, J. P., Uzairu, A., & Idris, S. O. (2015). Quantum modeling of the toxicity of selected Anti-Candida albicans Schiff bases and their Nickel (II) complexes. *Journal of Computational Methods in Molecular Design Scholars Research Library*, *5*, 91ï 103.

Barr, A. M., Lehmann-Masten, V., Paulus, M., Gainetdinov, R. R., Caron, M. G., & Geyer, M. A. (2004). The selective serotonin-2A receptor antagonist M100907 reverses behavioral deficits in dopamine transporter knockout mice. *Neuropsychopharmacology*, *29*(2), 221ï 228.

Bauernschmitt, R., & Ahlrichs, R. (1996). Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chemical Physics Letters*, *256*(4ï 5), 454ï 464.

Boulougouris, V., Glennon, J. C., & Robbins, T. W. (2008). Dissociable effects of selective 5-HT 2A and 5-HT 2C receptor antagonists on serial spatial reversal learning in rats. *Neuropsychopharmacology*, *33*(8), 2007ï 2019.

Fishback, J. A., Robson, M. J., Xu, Y.-T., & Matsumoto, R. R. (2010). Sigma receptors: potential targets for a new class of antidepressant drug. *Pharmacology & Therapeutics*, *127*(3), 271ï 282.

Gabbert, S., & Weikard, H.-P. (2013). Sequential testing of chemicals when costs matter: A value of information approach. *Human and Ecological Risk Assessment: An International Journal*, *19*(4), 1067ï 1088.

Gao, X., Han, L., & Ren, Y. (2016). In silico exploration of 1, 7-diazacarbazole analogs as checkpoint kinase 1 inhibitors by using 3D QSAR, molecular docking study, and molecular dynamics simulations. *Molecules*, *21*(5), 591.

Golbraikh, A., & Tropsha, A. (2002). Beware of q2! *Journal of Molecular Graphics and Modelling*, *20*(4), 269ï 276.

Hehre, W. J., & Huang, W. W. (1995). *Chemistry with computation: an introduction to SPARTAN*. Wavefunction, Inc.

Kaya, B., Yurttaķ, L., Saĵ lik, B. N., Levent, S., Özkay, Y., & Kaplancikli, Z. A. (2017). Novel 1-(2-pyrimidin-2-yl) piperazine derivatives as selective monoamine oxidase (MAO)-A inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *32*(1), 193ï 202.

Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics*, *11*(1), 137ï 148.

Khattab, S. N., Moneim, S. A. H. A., Bekhit, A. A., El Massry, A. M., Hassan, S. Y., El-Faham, A., é Amer, A. (2015).

Exploring new selective 3-benzylquinoxaline-based MAO-A inhibitors: design, synthesis, biological evaluation and docking studies. *European Journal of Medicinal Chemistry*, *93*, 308ï 320.

Lagorce, D., Douguet, D., Miteva, M. A., & Villoutreix, B. O. (2017). Computational analysis of calculated physicochemical and ADMET properties of protein-protein interaction inhibitors. *Scientific Reports*, *7*, 46277.

Leonard, J. T., & Roy, K. (2006). On selection of training and test sets for the development of predictive QSAR models. *QSAR & Combinatorial Science*, *25*(3), 235ï 251.

LIN, F., FU, X.-M., WANG, C., JIANG, S.-Y., WANG, J.-H., ZHANG, S.-W., é LI, Y. (2016). QSAR, Molecular Docking and Molecular Dynamics of 3C-like Protease Inhibitors. *Acta Physico-Chimica Sinica*, *32*(11), 2693ï 2708.

Lin, F., Li, F., Wang, C., Wang, J., Yang, Y., Yang, L., & Li, Y. (2017). Mechanism exploration of arylpiperazine derivatives targeting the 5-HT2A receptor by in silico methods. *Molecules*, *22*(7), 1064.

Lipinski, C. A. (2016). Rule of five in 2015 and beyond: Target and ligand structural limitations, ligand chemistry structure and drug discovery project decisions. *Advanced Drug Delivery Reviews*, *101*, 34ï 41.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *23*(1ï 3), 3ï 25.

Myers, R. H. (1990). Classical and modern regression with applications, Raymond H. Myers. *Duxbury Classic Series.*

Netzeva, T. I., Worth, A. P., Aldenberg, T., Benigni, R., Cronin, M. T. D., Gramatica, P., é Marchant, C. A. (2005). Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships: The report and recommendations of ecvam workshop 52. *Alternatives to Laboratory Animals*, *33*(2), 155ï 173.

OECD, D. E. (2007). Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q) SAR] Models. *Organisation for Economic Co-Operation and Development, Paris, France.*

Olasupo, S. B., Uzairu, A., Shallangwa, G., & Uba, S. (2019). QSAR analysis and molecular docking simulation of norepinephrine transporter (NET) inhibitors as anti-psychotic therapeutic agents. *Heliyon*, *5*(10), e02640.

Olasupo, S. B., Uzairu, A., Shallangwa, G., & UBA, S. (n.d.). Quantitative Structure-Activity Relationship (QSAR) Studies and Molecular docking Simulation of Norepinephrine Transporter (NET) Inhibitors as Anti-psychotic Therapeutic Agents. *Journal of the Turkish Chemical Society Section A: Chemistry*, *7*(1), 179ï 196.

Olasupo, S., Uzairu, A., & Sagagi, B. (2016). Quantitative Structure-Activity relationship (QSAR) models for predicting Toxicity of Dioxin compounds. *Journal of Computational Methods in Molecular Design*, *6*(2), 1ï 12. Retrieved from https://pdfs.semanticscholar.org/84fb/76731ef2955f734e6a95ded1922d8ad9f950.pdf

Patil, S. S. (2012). A Least Square Approach ToAnlayze Usage Data For Effective Web Personalization.

Poyurovsky, M., Koren, D., Gonopolsky, I., Schneidman, M., Fuchs, C., Weizman, A., & Weizman, R. (2003). Effect of the 5-HT2 antagonist mianserin on cognitive dysfunction in chronic schizophrenia patients: an add-on, double-blind placebo-controlled study. *European Neuropsychopharmacology*, *13*(2), 123ï 128.

Tareq Hassan Khan, M. (2010). Predictions of the ADMET properties of candidate drug molecules utilizing different QSAR/QSPR modelling approaches. *Current Drug Metabolism*, *11*(4), 285ï 295.

Triggle, D. J., & Taylor, J. B. (2006). *Comprehensive Medicinal Chemistry II* (Vol. 7). Elsevier.

Tropsha, A., Gramatica, P., & Gombar, V. K. (2003). The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR & Combinatorial Science*, *22*(1), 69ï 77.

Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Varghese, C. P., & Agrawal, R. K. (2011). Validation of QSAR models-strategies and importance. *International Journal of Drug Design & Discovery*, *3*, 511ï 519.

Wang, J., Yang, Y., Li, Y., & Wang, Y. (2016). Computational study exploring the interaction mechanism of benzimidazole derivatives as potent cattle bovine viral diarrhea virus inhibitors. *Journal of Agricultural and Food Chemistry*, *64*(29), 5941ï 5950.

Waring, M. J. (2010). Lipophilicity in drug discovery. *Expert Opinion on Drug Discovery*, *5*(3), 235ï 248.

Wu, M., Li, Y., Fu, X., Wang, J., Zhang, S., & Yang, L. (2014). Profiling the interaction mechanism of quinoline/quinazoline derivatives as MCHR1 antagonists: an in silico method. *International Journal of Molecular Sciences*, *15*(9), 15475ï 15502.

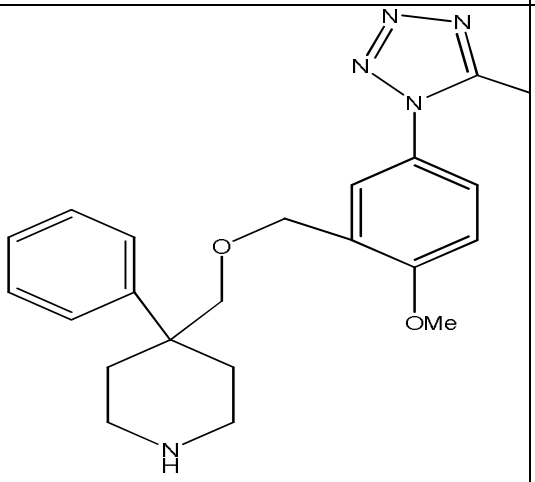
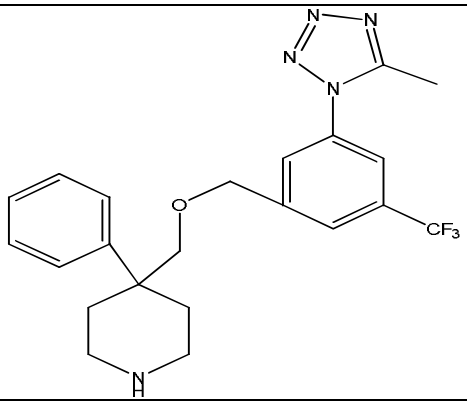
Yang, Y., Li, Y., Pan, Y., Wang, J., Lin, F., Wang, C., é Yang, L. (2016). Computational analysis of structure-based interactions for novel H1-antihistamines. *International Journal of Molecular Sciences*, *17*(1), 129.

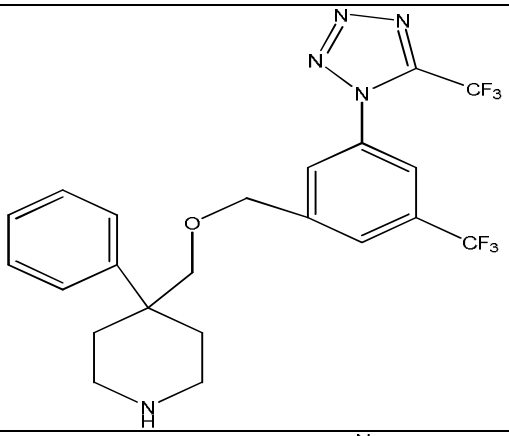
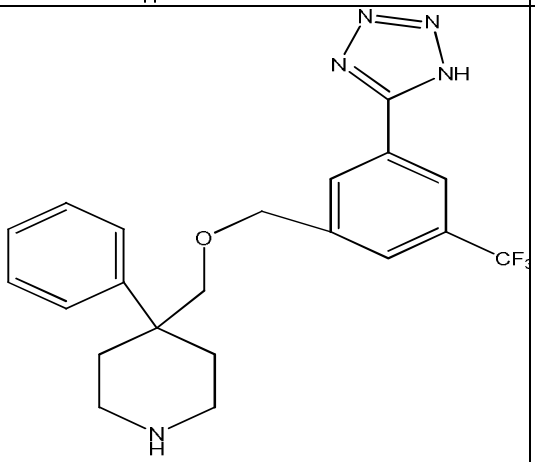
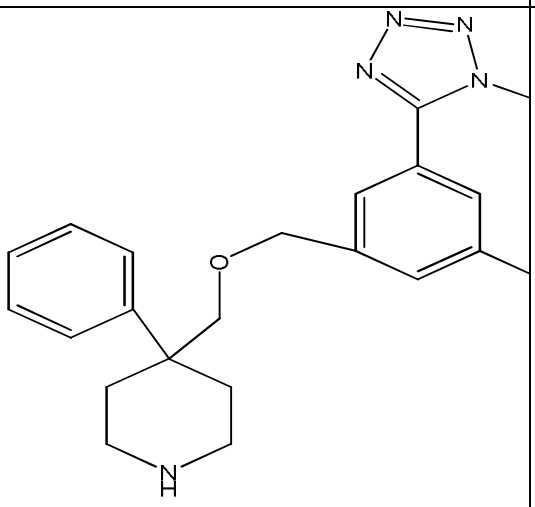
Yap, C. W. (2011). PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*, *32*(7), 1466ï 1474.

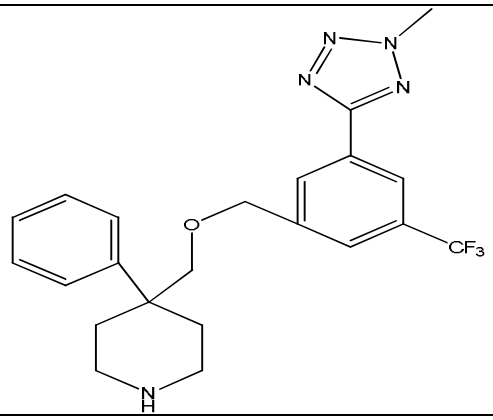
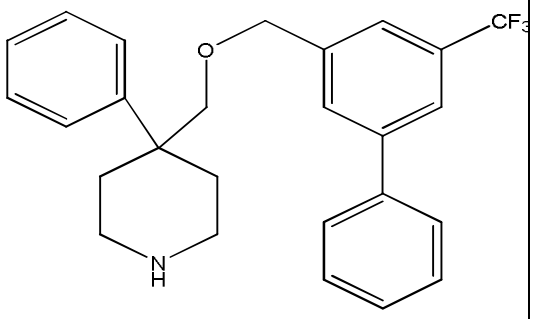
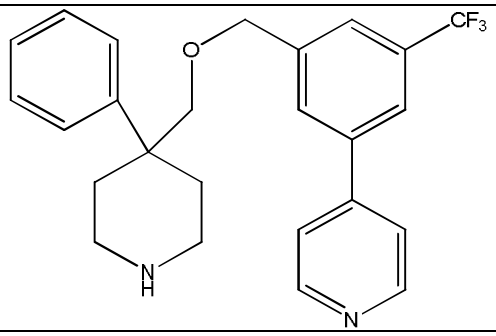
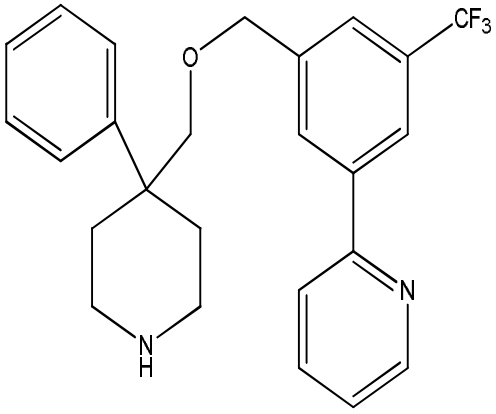
Zare, S., Fereidoonnezhad, M., Afshar, D., & Ramezani, Z. (2017). A comparative QSAR analysis and molecular docking studies of phenyl piperidine derivatives as potent dual NK1R antagonists/serotonin transporter (SERT) inhibitors. *Computational Biology and Chemistry*, *67*, 22ï 37.

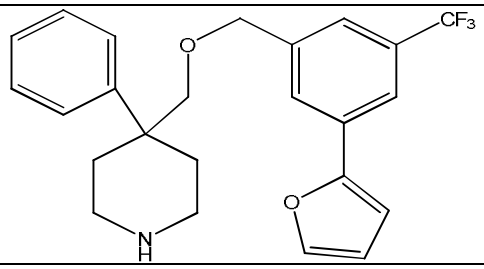
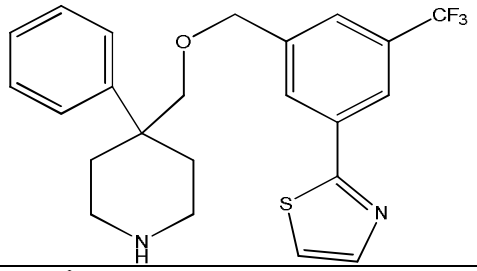
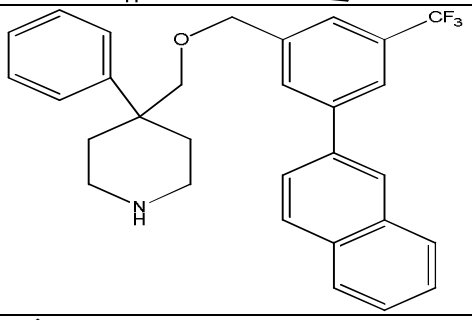
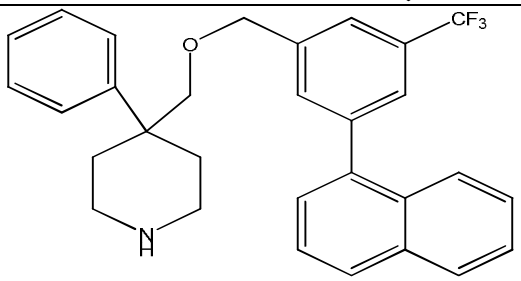
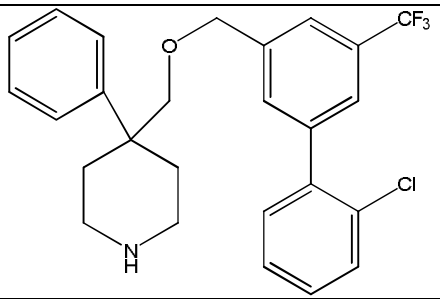
SUPPLEMENTARY MATERIALS

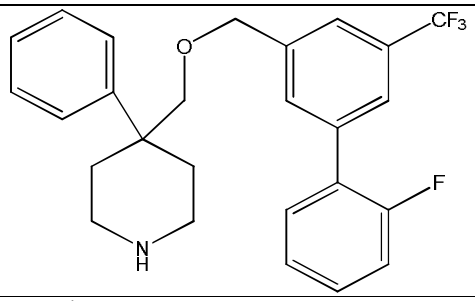
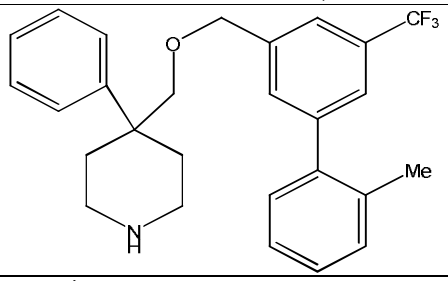
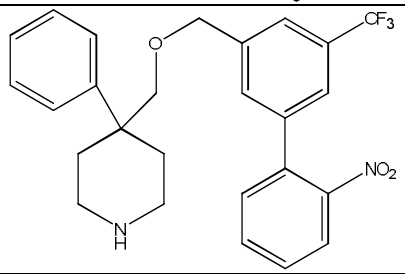
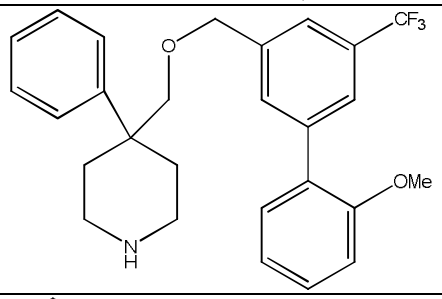
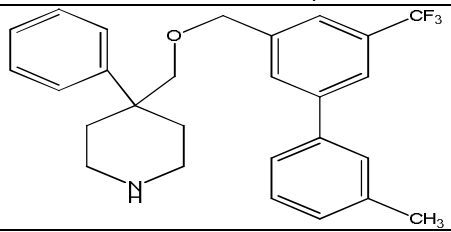
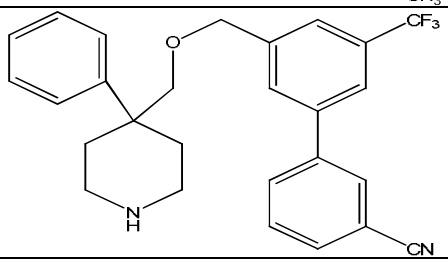
Supplementary Table S1 - Chemical Names, Chemical Structures and Experimental pIC_{50} values of Phenyl piperidine (Data set)

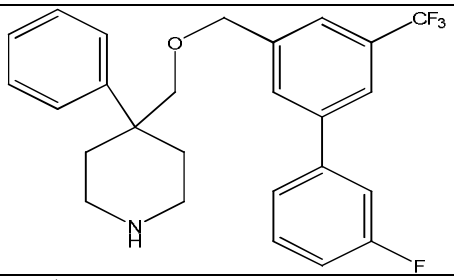
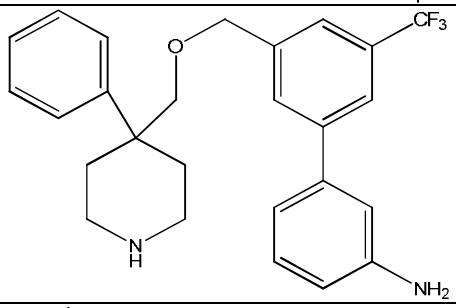
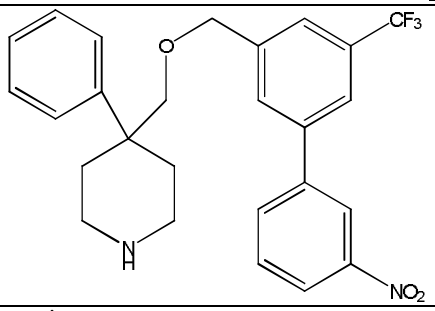
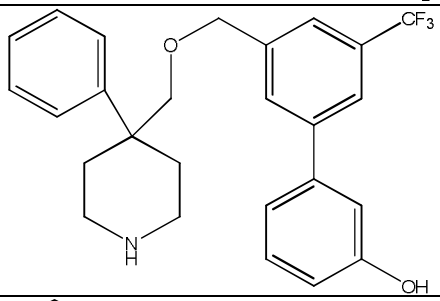
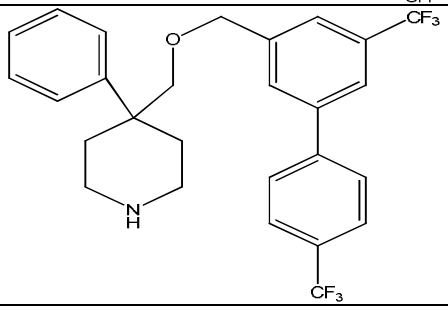
S/N	Chemical Name	Chemical Structure	Experimental pIC_{50}
1.	4-(((2-methoxy-5-(5-methyl-1H-tetrazol-1-yl)benzyl)oxy)methyl)-4-phenylpiperidine		7.432
2.	4-(((3-(5-methyl-1H-tetrazol-1-yl)-5-(trifluoromethyl)benzyl)oxy)methyl)-4-phenylpiperidine		7.155

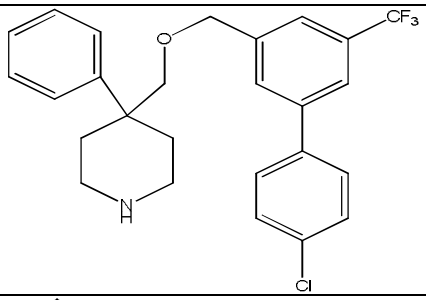
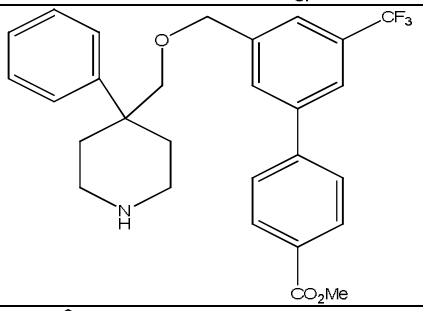
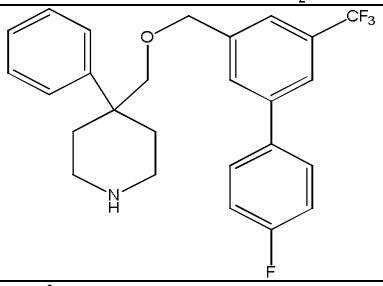
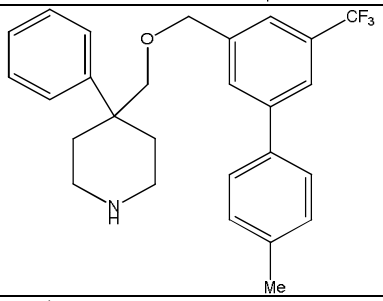
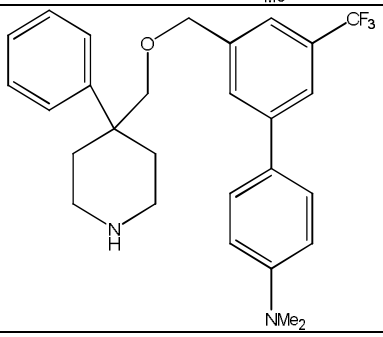
3.	4-phenyl-4-(((3-(trifluoromethyl)-5-(5-(trifluoromethyl)-1H-tetrazol-1-yl)benzyl)oxy)methyl)piperidine		7.244
4.	4-(((3-(1H-tetrazol-5-yl)-5-(trifluoromethyl)benzyl)oxy)methyl)-4-phenylpiperidine		8.066
5.	4-(((3-(1-methyl-1H-tetrazol-5-yl)-5-(trifluoromethyl)benzyl)oxy)methyl)-4-phenylpiperidine		7.301

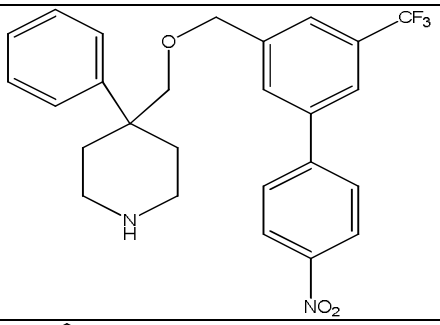
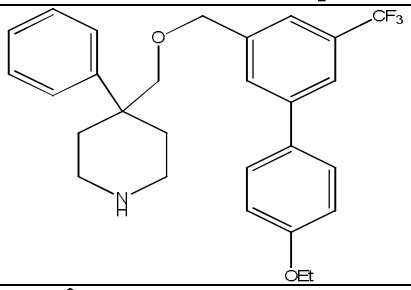
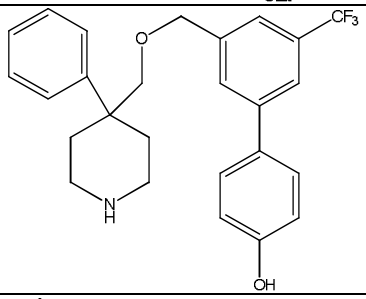
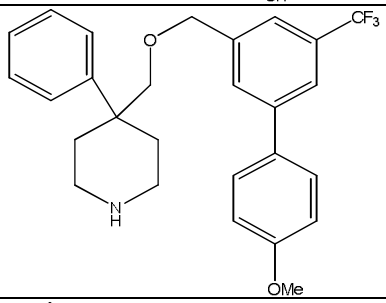
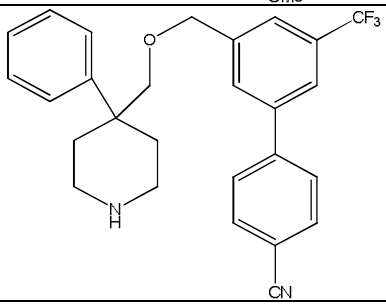
6.	4-(((3-(2-methyl-2H-tetrazol-5-yl)-5-(trifluoromethyl)benzyl)oxy)methyl)-4-phenylpiperidine		8.018
7.	4-phenyl-4-(((5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)piperidine		8.432
8.	4-(3-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5-(trifluoromethyl)phenyl)pyridine		7.796
9.	2-(3-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5-(trifluoromethyl)phenyl)pyridine		8.131

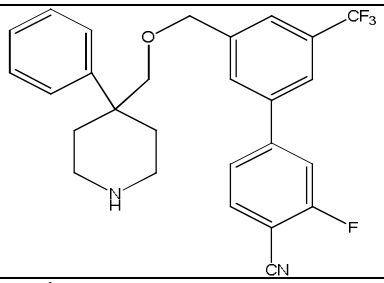
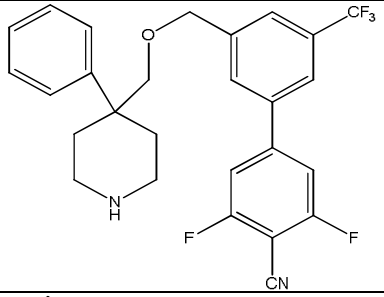
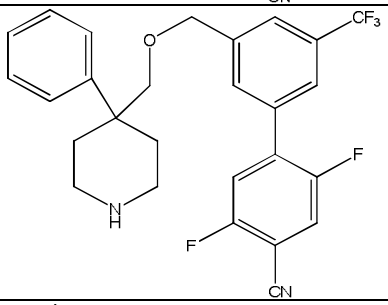
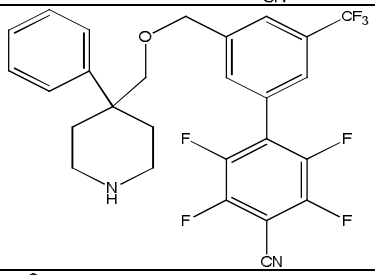
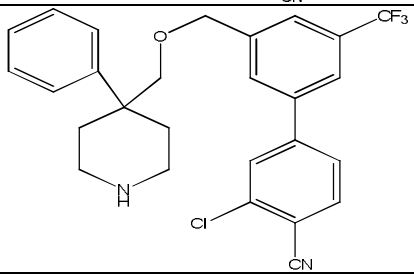
10.	4-(((3-(furan-2-yl)-5-(trifluoromethyl)benzyl)oxy)methyl)-4-phenylpiperidine		7.745
11.	2-(3-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5-(trifluoromethyl)phenyl)thiazole		8.387
12.	4-(((3-(naphthalen-2-yl)-5-(trifluoromethyl)benzyl)oxy)methyl)-4-phenylpiperidine		6.959
13.	4-(((3-(naphthalen-1-yl)-5-(trifluoromethyl)benzyl)oxy)methyl)-4-phenylpiperidine		6.921
14.	4-(((2'-chloro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.194

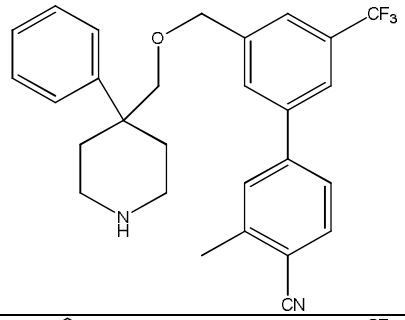
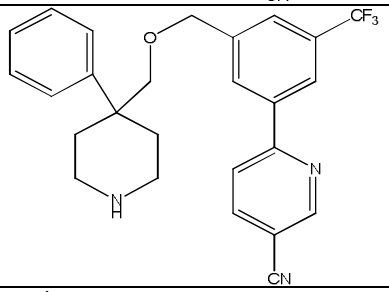
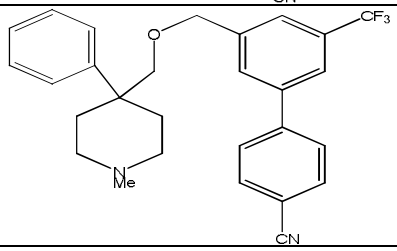
15.	4-(((2'-fluoro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.119
16.	4-(((2'-methyl-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		7.886
17.	4-(((2'-nitro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.456
18.	4-(((2'-methoxy-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.377
19.	4-(((3'-methyl-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		7.854
20.	3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-3-carbonitrile		8.000

21.	4-(((3'-fluoro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.201
22.	3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-3-amine		7.523
23.	4-(((3'-nitro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.377
24.	3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-3-ol		7.569
25.	4-(((4',5-bis(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		7.444

26.	4-(((4'-chloro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.745
27.	methyl 3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carboxylate		7.602
28.	4-(((4'-fluoro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.824
29.	4-(((4'-methyl-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.194
30.	N,N-dimethyl-3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-amine		8.456

31.	4-(((4'-nitro-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		8.137
32.	4-(((4'-ethoxy-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		7.854
33.	3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-ol		7.569
34.	4-(((4'-methoxy-5-(trifluoromethyl)-[1,1'-biphenyl]-3-yl)methoxy)methyl)-4-phenylpiperidine		7.824
35.	3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		8.398

36.	3-fluoro-3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		8.051
37.	3,5-difluoro-3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		8.155
38.	2,5-difluoro-3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		8.432
39.	2,3,5,6-tetrafluoro-3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		7.328
40.	3-chloro-3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		8.509

41.	3-methyl-3'-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		8.569
42.	6-(3-(((4-phenylpiperidin-4-yl)methoxy)methyl)-5-(trifluoromethyl)phenyl)nicotinonitrile		8.398
43.	3'-(((1-methyl-4-phenylpiperidin-4-yl)methoxy)methyl)-5'-(trifluoromethyl)-[1,1'-biphenyl]-4-carbonitrile		8.022

Supplementary Table S2a - Euclidean Based Applicability Domain (Training set)

S/N	Distance Score	Mean Distance	Normalized Mean Distance
1	**24.903**	**0.830**	**0.837**
2	**17.964**	**0.599**	**0.388**
3	**12.911**	**0.430**	**0.060**
4	**15.360**	**0.512**	**0.219**
5	**16.729**	**0.558**	**0.308**
6	**13.763**	**0.459**	**0.115**
7	**12.584**	**0.419**	**0.039**
8	**13.140**	**0.438**	**0.075**
9	**12.444**	**0.415**	**0.030**
10	**15.145**	**0.505**	**0.205**
12	**25.389**	**0.846**	**0.869**
14	**11.985**	**0.400**	**0.000**
15	**12.257**	**0.409**	**0.018**

16	**19.772**	**0.659**	**0.505**
19	**19.186**	**0.640**	**0.467**
20	**27.411**	**0.914**	**1.000**
22	**12.925**	**0.431**	**0.061**
23	**17.117**	**0.571**	**0.333**
24	**13.272**	**0.442**	**0.083**
25	**17.966**	**0.599**	**0.388**
27	**15.197**	**0.507**	**0.208**
28	**12.424**	**0.414**	**0.028**
30	**14.508**	**0.484**	**0.164**
32	**19.409**	**0.647**	**0.481**
33	**13.905**	**0.463**	**0.124**
35	**13.061**	**0.435**	**0.070**
37	**17.410**	**0.580**	**0.352**
39	**20.524**	**0.684**	**0.554**
42	**12.452**	**0.415**	**0.030**
43	**13.524**	**0.451**	**0.100**

Supplementary Table S2b - Euclidean Based Applicability Domain (Test set).

S/N	Distance Score	Mean Distance	Normalized Mean Distance
11	14.002	0.467	0.131
13	20.054	0.668	0.523
17	17.336	0.578	0.347
18	13.011	0.434	0.067
21	11.931	0.398	-0.003
26	13.148	0.438	0.075
29	13.032	0.434	0.068
31	16.526	0.551	0.294
34	12.606	0.420	0.040
36	16.275	0.542	0.278
38	15.764	0.525	0.245
40	17.426	0.581	0.353
41	14.406	0.480	0.157

Table S3 - Pharmacokinetic/ ADMET properties of the studied compounds with a Standard Drug (Brexpiprazole) as a control

Compds S/N	MW	ADMET_Risk	S+logP	S+logD	LogBB	Pgp_Inh	S+CL_Renal	RO5	RO5_Code	hERG_pIC50	TOX_Risk	%Fa_hum-100.0	%Fb_hum-100.0
1	393.49	2.74	3.50	1.11	-0.30	Yes (48%)	No (99%)	0		4.85	1.25	81.37	52.78
2	431.46	5.85	4.57	2.19	0.47	Yes (61%)	No (99%)	1	LP	5.14	2.29	93.86	59.37
3	485.44	7.92	5.89	3.55	0.92	Yes (61%)	No (82%)	1	LP	5.17	2.69	97.21	55.60
4	417.44	6.24	3.43	3.43	-0.18	No (96%)	Yes (61%)	0		4.51	2.00	19.58	17.64
5	431.46	5.80	4.52	2.14	0.54	Yes (70%)	No (99%)	1	LP	5.17	2.40	94.84	58.34
6	431.46	5.32	4.69	2.32	0.62	Yes (97%)	No (99%)	1	LP	5.12	1.86	97.55	66.86
7	425.50	8.71	5.80	3.47	1.06	Yes (51%)	No (99%)	1	LP	6.45	1.00	100.00	35.64
8	426.49	6.43	5.10	2.77	0.95	Yes (59%)	No (99%)	0		6.24	1.62	99.94	44.55
9	426.49	6.29	5.26	2.92	1.05	Yes (51%)	No (99%)	1	LP	6.40	1.00	99.97	47.85
10	415.46	8.57	5.33	2.98	0.84	Yes (97%)	No (99%)	1	LP	6.06	2.30	99.97	38.93
11	432.51	8.98	5.26	2.92	1.02	Yes (97%)	No (99%)	0		6.10	4.00	99.95	49.72
12	475.56	10.25	6.74	4.46	1.14	Yes (97%)	No (99%)	1	LP	6.86	2.41	91.74	25.92
13	475.56	9.46	6.60	4.33	1.14	Yes (97%)	No (95%)	1	LP	6.73	1.73	83.08	22.14
14	459.94	9.94	6.20	3.97	1.18	Yes (97%)	No (99%)	1	LP	6.61	1.85	99.99	28.92
15	443.49	8.88	6.07	3.74	1.20	Yes (97%)	No (99%)	1	LP	6.46	1.78	99.99	29.13
16	439.52	8.24	6.04	3.70	1.00	Yes (64%)	No (89%)	1	LP	6.38	1.36	100.00	26.90
17	470.50	9.07	5.62	3.34	1.06	Yes (97%)	No (99%)	1	LP	7.06	3.00	99.94	80.33
18	455.52	7.84	5.77	3.41	1.02	Yes (97%)	No (99%)	1	LP	6.37	1.48	99.99	33.63
19	439.52	8.35	6.19	3.85	1.06	Yes (59%)	No (95%)	1	LP	6.43	1.11	100.00	26.55
20	450.51	7.76	5.46	3.19	0.73	Yes (97%)	No (92%)	1	LP	6.52	1.59	80.18	24.06
21	443.49	9.11	6.16	3.83	1.21	Yes (97%)	No (99%)	1	LP	6.55	1.84	100.00	29.98
22	440.51	8.69	5.56	3.24	0.81	Yes (97%)	No (92%)	1	LP	6.40	2.49	99.89	55.03
23	470.50	9.14	5.86	3.56	1.06	Yes (97%)	No (99%)	1	LP	7.17	2.85	99.88	79.98
24	441.50	6.63	5.36	3.29	0.88	No (69%)	No (99%)	1	LP	6.57	1.69	99.76	73.11
25	493.50	9.69	6.57	4.25	1.26	Yes (64%)	No (89%)	1	LP	6.67	2.83	97.69	19.15

26	459.94	9.43	6.42	4.19	1.23	Yes (97%)	No (99%)	1	LP	6.70	1.34	100.00	27.46
27	483.53	8.95	5.87	3.59	1.05	Yes (97%)	No (99%)	1	LP	6.32	1.51	99.97	57.05
28	443.49	9.03	6.22	3.88	1.21	Yes (97%)	No (99%)	1	LP	6.56	1.72	100.00	28.42
29	439.52	9.33	6.27	3.93	1.07	Yes (73%)	No (95%)	1	LP	6.43	1.10	100.00	26.91
30	468.57	9.87	6.52	4.11	1.18	Yes (97%)	No (99%)	1	LP	6.70	2.36	100.00	23.60
31	470.50	9.63	5.94	3.64	1.09	Yes (97%)	No (99%)	1	LP	7.14	2.99	98.42	76.14
32	469.55	10.62	6.31	3.93	1.09	Yes (97%)	No (99%)	1	LP	6.60	1.30	100.00	25.14
33	441.50	6.18	5.37	3.27	0.87	No (66%)	No (99%)	1	LP	6.51	1.18	99.76	70.61
34	455.52	9.36	5.97	3.61	1.08	Yes (97%)	No (99%)	1	LP	6.48	1.32	99.99	30.86
35	450.51	7.63	5.56	3.29	0.75	Yes (97%)	No (92%)	1	LP	6.54	1.22	77.69	23.37
36	468.50	8.75	5.92	3.64	1.01	Yes (97%)	No (92%)	1	LP	6.59	1.69	29.74	6.70
37	486.49	8.63	6.24	3.95	1.22	Yes (97%)	No (92%)	1	LP	6.62	2.00	11.15	1.71
38	486.49	9.09	6.16	3.88	1.17	Yes (97%)	No (86%)	1	LP	6.53	2.00	19.87	3.08
39	522.47	9.48	6.51	4.23	1.41	Yes (97%)	No (82%)	2	Mw; LP	6.33	2.00	5.08	0.28
40	484.95	10.06	6.19	4.00	0.99	Yes (97%)	No (99%)	1	LP	6.76	1.72	16.05	3.71
41	464.53	8.59	5.84	3.55	0.71	Yes (97%)	No (92%)	1	LP	6.52	1.51	45.29	10.54
42	451.50	6.73	5.03	2.73	0.72	Yes (97%)	No (95%)	0		6.42	1.00	99.83	31.55
43	464.53	8.72	5.81	4.48	0.81	Yes (97%)	No (95%)	1	LP	6.70	1.00	55.96	14.40
Brexpiprazole	433.57	6.583	4.725	4.396	0.22	Yes (78%)	No (99%)	0		6.82	2		

Brexpiprazole (Standard Drug) as a control