

Prediction of HemO Inhibitors Based on Iminoguanidine using QSAR, 3D-QSAR Study, Molecular Docking, Molecular Dynamic Simulation and ADMET

Emmanuel Israel Edache*, Adamu Uzairu, Paul Aliteri P Mamza and Gideon Adamu Shallengwa

*Department of Chemistry, Ahmadu Bello University, Zaria, Nigeria.

Received April 16, 2020; Accepted April 24, 2020; Published July 30, 2020

ABSTRACT

The computational proficiencies put on in this screening are QSAR, with fraction of variance $r^2 = 0.8499$ and LOO-CV $q^2 = 0.5008$. 3D-QSAR (CoMFA and CoMSIA) and structure-based approach (molecular dynamics simulations assisted molecular docking study) were applied for the identification of important features of iminoguanidine-based, responsible for *Pseudomonas aeruginosa* inhibition. Alignments were tried to obtain best 3D-QSAR model, of which, compound 5 was found to be the best method, as it gave good statistical results. In CoMFA, Leave One Out (LOO) cross validated coefficients (q^2), and conventional coefficient (r^2) values were found to be 0.5177 and 0.968, respectively. Similarly, in CoMSIA, q^2 , and r^2 values were found to be 0.576 and 0.985, respectively. Multi-parameter optimization (MPO), golden triangle rule, structure activity/property relationships (SAR/SPR), Drug-likeness properties, and lipophilicity indices are reported and discussed in terms of the biological activity of iminoguanidine-based inhibitors of HemO. These techniques offer the ability to guide drug design and selection to quickly identify the compounds with desirable drug-like attributes.

Keywords: Hemeoxygenase (HemO), *Pseudomonas aeruginosa*, Iminoguanidine derivatives, QSAR, CoMFA, docking, MD Simulations, ADME, Lipophilicity indices, Golden Triangle

INTRODUCTION

Heme is chemical compounds composed of iron and an organic part called protoporphyrin IX, relating to or serving as a prosthetic radical of many enzymes which facilitates a wide variety of processes like electron transport (cytochrome P450 family), enzyme catalysis (e.g. peroxidase, catalase, cyclooxygenase, nitric oxide synthase), reversible binding of gases (hemoglobin, myoglobin, guanylate cyclase) [1] and various degrees of antioxidant, anti-inflammatory, anti-apoptotic, anti-proliferative, and immunomodulatory effects, most of which may play an important part in the defense against atherosclerotic lesion formation [2]. Besides, it involved in the regulation of protein synthesis and cell differentiation [1] and protective genetic factor against proatherogenic effects produced by environmental factors such as cigarette smoking or exposure to air pollution [2]. Heme which is an author of iron (a vital micronutrient essential for the survival and virulence of most bacterial pathogens) for *Pseudomonas aeruginosa* on infection of the host. The flexibility of heme into the cell is driven by the catalytic action of hemeoxygenase (HemO) and controlled by the heme binding protein (PhuS). Despite advances in structural characterization of bacterial heme uptake proteins,

the mechanism of heme transfer is poorly defined [3]. The opportunistic pathogen *Pseudomonas aeruginosa* encodes two distinct heme uptake systems: The *Pseudomonas* heme uptake (phu) and the heme assimilation system (has) [4]. It was recently determined that the Phu system is the primary iron acquisition system in *P. aeruginosa* chronic infection of the cystic fibrosis lung and up-regulated coincident with a decrease in the levels of the major secreted siderophore pyoverdine [5,6]. This increase is in large part caused by point mutations within the promoter region of the phu operon leading to increased expression of PhuR [5]. Therefore, understanding the molecular mechanisms by which *P. aeruginosa* acquires and uses heme may provide a therapeutic

Corresponding author: Emmanuel Israel Edache, Department of Chemistry, Ahmadu Bello University, Zaria, Nigeria, Tel: +2348066776802; E-mail: edacheson2004@gmail.com

Citation: Edache EI. (2020) Prediction of HemO Inhibitors Based on Iminoguanidine using QSAR, 3D-QSAR Study, Molecular Docking, Molecular Dynamic Simulation and ADMET. J Drug Design Discov Res, 1(2): 36-52.

Copyright: ©2020 Edache EI. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

strategy in the treatment of *P. aeruginosa* chronic infections. To combat the decline in viable antibiotics and growing antimicrobial resistance, an alternative hypothesis suggests targeting virulence factors that are not required for survival but are essential for pathogenesis within the host [3].

A new cure aims to fight multidrug-resistant infections, such as the iron-regulated hemeoxygenase (HemO) of *Pseudomonas aeruginosa*, due to connections between iron and virulence and a requirement on heme as an iron source during infection. In this study, we determined structural elements QSAR study for modeling the inhibitory effects of 25 synthesized iminoguanidine derivatives [https://pubchem.ncbi.nlm.nih.gov/bioassay/1315712]. Multiple linear regression and genetic algorithm (MLR-GA) methods were used for modeling the relationship between pMIC50 and their structural descriptors. 3D-QSAR models and molecular docking analysis of all the compounds were predicted. Molecular dynamic simulations to explore the binding modes of these compounds and identify the vital moieties for biological activity. Lastly, ADMET and golden triangle methods were used to assess drug load and identify compounds with good stability of the many physicochemical and biological properties necessary to become a successful, effective, and harmless drug.

MATERIALS & METHODS

Experimental data

Synthesis and activity of a series of iminoguanidine-based inhibitors of HemO were taken from the literature reported by PubChem

[https://pubchem.ncbi.nlm.nih.gov/bioassay/1315712]. The experimental MIC50 values of all compounds in $\mu\text{g.mL}^{-1}$ where converted in pMIC50 by taking $-\text{Log}(1/\text{MIC50})$ and were used as the dependent variable. There was a total of 25 iminoguanidine derivatives which are then split into a training set of 16 compounds for generating QSAR and 3D-QSAR modeling which were carried out through molecular interaction fields (MIFs) analysis, and a test set of 7 compounds for validating the quality of the models. Remaining 2 compounds that failed to optimize were removed. The compounds in the test set were selected from the original pool of structures based on the Kennard-Stone method (DTC-Lab v1.2) with a ratio of 70:30 percent respectively (the test set is marked by *). All the structures (Figure 1) and inhibitory activities are listed in Table 1.

COMPUTATIONAL DETAILS AND DESCRIPTOR CALCULATIONS

All the 2D structures were generated by Marvin View software package and transferred to Spartan'14 version 14.1.0 software to create the 3D structure, pre-optimized using semi-empirical PM3 minimization. Then a more precise optimization was performed to make the conformations with the least potential energy using Merck molecular force field

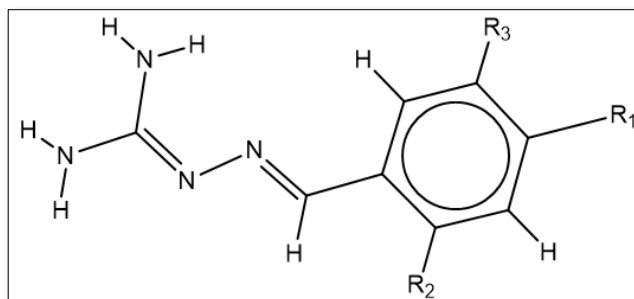


Figure 1. 2D chemical structure of iminoguanidine-based inhibitors of HemO.

(MMFF) with density functional theory (DFT) with Becke's three-parameter hybrid functional [7] using LYP correlation functional [8]. The standard Pople's 6-311+G** basis set was used. Descriptors were calculated using Spartan'14 v1.1.4, PaDEL-Descriptor v2.20, and MedChem Designer v3.1 which include: Quantum-chemical, molecular orbital, thermodynamics, charged partial surface area, electrostatic geometrical, and topological descriptors.

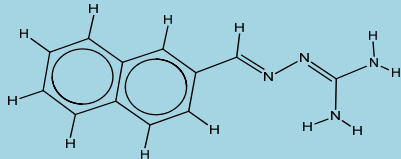
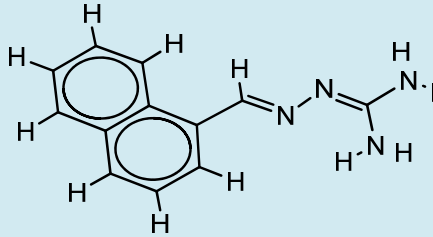
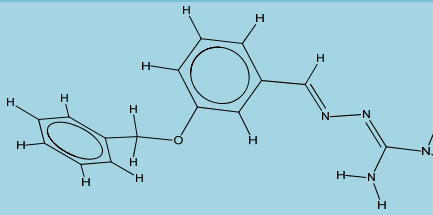
ALIGNMENT AND COMFA ANALYSIS

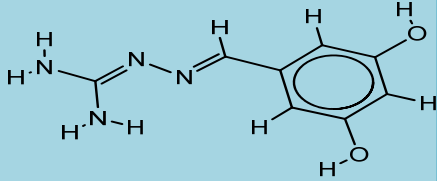
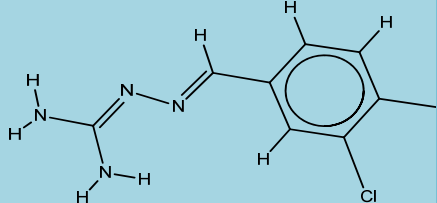
Proper alignment of the chemical structure relative to one another is one of the most important steps in the 3D-QSAR analysis for obtaining an effective molecular interaction field (MIF) model. The molecules were superimposed using the atom-based and pharmacophore-based alignment by the Open3DALIGN tools v2.3 [9]. The energy minimized and optimized structures were aligned by the template-based method, and compound 5 was selected as the template to construct other compounds because of its high representative chemical structure and the alignment was completed by open3DALIGN software. The alignment corresponding to the highest cumulative O3A score was selected for further analysis. Figure 2 shows the best alignment in which compound 5 was selected as the template. CoMFA is a versatile method to describe 3D-QSAR quantitatively. Open3DQSAR software (version 2.3) is open-source software available for high-throughput chemometric analysis of molecular interaction fields (MIFs) [10]. This study used Open 3DQSAR to perform CoMFA analysis.

1. The best alignment with compound 5 as a template as shown in Figure 2 is placed in a 3D cubic lattice with 2 Å grid size and a 5.0 Å out gap.
2. The steric fields were computed using an SP_3 hybridized carbon atom probe with a +1 charge. Similarly, electro-static fields were computed using a volume-less probe.
3. These steric and electrostatic interaction energies were considered as independent variables (CoMFA descriptors).

Before creating of CoMFA model following pretreatment operations were carried out to reduce the noise hidden in PLS matrix and hence reduced the computational time:

Table 1. Structures and structural formulae of compounds and their pMIC₅₀ values.

Compound	R ₁	R ₂	R ₃	ugmL ⁻¹	pMIC ₅₀
1*	-Cl	-H	-H	49	1.6902
2	-NC ₂ H ₆	-H	-H	121	2.0828
3	-H	-OH	-H	123	2.0899
4	-H	-OCH ₃	-H	99.5	1.9978
5	-NO ₂	-H	-H	72.8	1.8621
6	-H	-F	-H	131.2	2.1179
7				48.4	1.6848
8	-H	-H	-Cl	52.4	1.7193
9	-H	-Cl	-H	47.3	1.6749
10	-CH(CH ₃) CH ₃	-H	-H	52.3	1.7185
11*	-OCH ₃	-H	-H	113	2.0531
12		-	-	32.1	1.5065
13	-F	-H	-H	162	2.2095
14*	-Br	-H	-H	52.5	1.7202
15 ^R				23.2	1.3655
16*	-H	-Br	-H	50.1	1.6998
17	-H	-OH	-Br	80.5	1.9058
18 ^R	-C ₆ H ₅	H	H	59.1	1.7716

19*		-	-	1140	3.0569
20*	-OH	-H	-H	408	2.6107
21	-H	-H	-OH	750	2.8751
22*	-H	-H	-OCH ₃	134	2.1271
23	-H	-H	-F	66.6	1.8235
24	-H	-H	-Br	51.3	1.7101
25		-	-	32.3	1.5092

*Test set; ^RCompound removed from the model

4. The minimum and maximum energy values of steric and electrostatic were set to a cutoff value -30.0 and +30.0 kcal/mol, respectively. This pretreatment avoids the infinity of energy values inside the molecule.

5. Low energy values (<0.05 kcal/mol) were set to zero in both fields.

6. Standard deviation set to <0.1 to improve the signal-to-noise ratio.

7. N-level variables that are variables that assume only N values across the training set were removed, most of which distributed on a small number of objects. This process avoids overweighting the importance of particular substituents present in a single molecule. Otherwise, it might negatively affect the whole model.

8. The whole block of X or Y variables scaled by block unscaled weighting (BUW) technique.

Prediction of the CoMFA model can be significantly improved by suitable variable clustering and selection procedures specified as smart region definition (SRD) and fractional factorial design (FFD). These variable selection techniques selectively eliminate noisy variables with no certainty. The SRD procedure carries out variable grouping based on their closeness in the 3D area to reduce the redundancy arising from the existence of multiple nearby descriptors which mainly encrypt the identical form of content [11]. FFD aims at selecting the variables which significantly increase the predictive ability (using the LOO, LTO, or LMO models), and can operate on both single

variables or groups identified by a previous SRD run, thereby removing unhelpful variables groups as performed in GOLPE [12].

PLS analysis carried out in Open3DQSAR was engaged to obtain a correlation between the descriptors derived by CoMFA (independent variables) and pMIC50 values (dependent variable). Open3DQSAR produces a PLS model through the Non-linear iterative partial least squares (NIPALS) algorithm [13]. The statistical parameters like the coefficient of determination (R²), Standard Deviation Error in Calculation (SDEC), Standard Deviation Error in Predictivity (SEDP), and F-ratio test were computed the overall significance of the model. Moreover, the CoMFA color contour maps are derived for the steric and electrostatic fields.

MOLECULAR DOCKING (MD) AND MOLECULAR DYNAMICS SIMULATIONS (MDSS)

The docking methodology can be used to perfect the interaction between a small molecule (drugs) and a protein (target) at the atomic level, which allow us to illustrate the behavior of drugs in the binding site of target proteins as well as to explain vital biochemical processes [14]. Docking of iminoguanidine-based inhibitors of HemO (25 compounds) was checked by AutoDock4.2 [15] program to detect their binding site, the best path, and the binding energy [16]. The crystal structure (target) of HemO (PDB entry code: 4mgf) was extracted from the Brook haven Protein Database (PDB <http://www.rcsb.org/pdb>). The X-ray crystal structure (target)



Figure 2. Alignment of 25 iminoguanidine-based inhibitors of HemO.

PDB: 4mgf was used to dock. At the beginning of docking, all the water molecules, Hetatm and B chain were removed, hydrogen atoms (polar only), and Kollman charges were added to the protein, and the file saved as pdbqt. A grid box was created with a grid point spacing of 0.375 Å and 40 × 40 × 40 points. Preparation of grid parameter file (.gpf) and a docking parameter file (.dpf), the Lamarckian genetic algorithm, and default procedures for docking a flexible ligand to a rigid protein were followed. The optimized receptor was then saved as a PDB file and used for docking simulation and molecular dynamics simulation (NAMD) [17].

RESULTS AND DISCUSSION

PMIC50 = 1.43985(+/-1.46294) +0.52064(+/-0.29982)
EHOMO -0.9114(+/-0.20999) ELOMO -0.02651(+/-
0.00365) PSA -0.00004(+/-0.00005) Ho -----Model 1

Internal Validation Parameters

SEE = 0.1538; r^2 = 0.8499; r^2 adjusted = 0.7953; F = 15.5722 (DF: 4, 11)

Leave-One-Out (LOO) Result

Q^2 = 0.5008; PRESS = 0.8651; SDEP = 0.2325

Without scaling	After scaling	
rm^2(Loo)	: 0.47126	0.48689
rm^2'(Loo)	: 0.23983	0.22357
average rm^2(LOO)	: 0.35555	0.35523
delta rm^2(LOO)	: 0.23142	0.26332

External Validation Parameters (Without Scaling)

r^2 = 0.91153; r_0^2 = 0.8677; reverse r_0^2 = 0.7660;
 $rm^2(test)$ = 0.7207; reverse $rm^2(test)$ = 0.5638; average
 $rm^2(test)$ = 0.6423; delta $rm^2(test)$ = 0.1568; $rmsep$ =
0.2417; $rpred^2$:0.7806; Q^2_{f1} = 0.7806; Q^2_{f2} = 0.7776

Some External Validation Parameters (After Scaling):
 $rm^2(test)$ = 0.6084; reverse $rm^2(test)$ = 0.3574; average
 $rm^2(test)$ = 0.4829; delta $rm^2(test)$ = 0.25096

Overall Parameters

Without scaling	After scaling
rm^2(overall) :	0.63481 0.59588
reverse rm^2(overall) :	0.34942 0.29109
average rm^2(overall) :	0.49212 0.44349
delta rm^2(overall) :	0.28538 0.30479

Golbraikh and Tropsha acceptable model criterias

1. Q^2 0.5007 Passed (Threshold value $Q^2 > 0.5$)
2. r^2 0.9115 Passed (Threshold value $r^2 > 0.6$)
3. $|r_0^2 - r^2|$ 0.1017 Passed (Threshold value $|r_0^2 - r^2| < 0.3$)
4. $k = 1.08014$; $[(r^2 - r_0^2)/r^2]$ 0.0481 OR* $k' = 0.91837$;
 $[(r^2 - r_0^2)/r^2] = 0.15963$ Passed (Threshold value:
 $[0.85 < k < 1.15$ and $((r^2 - r_0^2)/r^2) < 0.1]$ OR* $[0.85 < k' < 1.15$
and $((r^2 - r_0^2)/r^2) < 0.1]$)

It can be seen that the performance of Model 1 built with four Spartan'14 descriptors, the standard errors of regression coefficients are given. Four variables EHOMO, ELUMO, PSA, and Ho in Model (1) could explain 85% of the variance (adjusted coefficient of variation) of the activity. The leave-one-out predicted variance was found to be 50%. The difference between $|r_0^2 - r^2|$ is less than 0.3 signifying the robustness of the model [18]. While Model (1) was applied for the prediction of test set compounds, the predictive R^2 value for the test set was found to be 0.7806. As r^2 and r_0^2 values are not much different, an acceptable value of $rm^2(test)$ (0.7207) was obtained [19]. The mean effect (MF) value indicates the relative importance of a descriptor, compared with the other descriptors in the model. Its sign indicates the variation direction in the values of the activities as a result of the increase (or reduction) of the descriptor values. The mean effect values are given in **Table 2**. A negative effect of these descriptors illustrates that the

biological activity increases with decreasing the value of Energy of the lowest unoccupied molecular orbital (ELOMO) and standard enthalpy (H_o). The Energy of the highest occupied molecular orbital (EHOMO) and polar surface area (PSA) mean effect has a positive sign, suggest that the biological activity is directly related to these descriptors. The leverage values can be calculated for every compound and plotted vs. standardized residuals, and it allows a graphical detection of both the outliers and the influential compound in a model. **Figure 3** shows the Williams plot, the

applicability domain is established inside a squared area within ± 3 bound for residuals and a leverage threshold h^* ($h^* = 3(k+1)/n$, where k is the number of model parameters and n is the number of compounds [20]). It demonstrates that some of the compounds of the training and test set are inside of this square area. There are influential outlier compounds with standard residuals $>3d$ for the training sets. Besides, all the chemicals have leverage lower than the warning h^* value of 0.60 except compounds 2, 5, and 19.

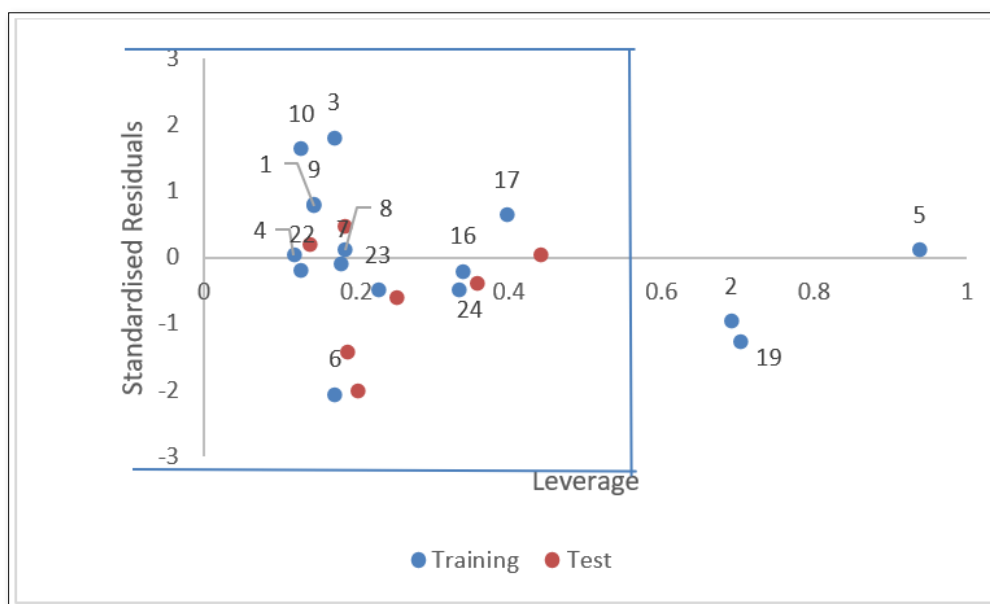


Figure 3. The Williams plot, the plot of the standardized residuals vs the leverage value.

Table 2. The calculated quantum chemical descriptors in this study.

Descriptor	Description	Mean effect
EHOMO	The energy of the highest occupied molecular orbital	0.8772
ELOMO	The energy of the lowest unoccupied molecular orbital	-0.4298
PSA	Polar Surface Area	0.5663
H^o	Standard Enthalpy	-0.0137

Since the QSAR model satisfied the Roy and Paul [19], Golbraikh and Tropsha [21] criteria, showed that the QSAR model is considered predictive.

Investigation of the 3D-QSAR structure-activity quantitative relationship

To assess the predictive ability of the 3D-QSAR models, the following statistical parameters were used: the coefficient of determination (R^2), the mean-square error of the model (SDEC), the F-statistic value, the correlation coefficient of the leave-one-out (LOO) method (Q2LOO), the correlation coefficient of the leave-two-out LTO method (Q2LTO), the

correlation coefficient of the LMO method (Q2LMO), and the mean-square error of prediction (SDEP). The statistical results of MIFs studies are summarized in **Table 3**. The predicted values of the iminoguanidine-based activity are shown in **Table 4**. The calculated values of the model have the best match with the experimental data. The level of correlation between the activities predicted by the CoMFA model and the experimental values is also shown in **Figure 3**.

As can be seen from **Table 3** (fractional factorial design (ffdsel)), the model has a high predictive ability. High coefficients of determination ($R^2 = 0.968$) and LOO cross-

validation ($Q^2 = 0.52$) indicate the statistical significance of the obtained model. The scatter plot of observed pMIC50 vs. predicted pMIC50 of both the training and test set of CoMFA models is shown in **Figure 3**. Molecular interaction field (MIF) contour maps were obtained to visualize the information about the 3D-QSAR models. The CoMFA steric

(88.56%) and electrostatic (11.44%) contour maps were shown in **Figure 4**. The steric field is represented by green and yellow contours in which green contours indicate regions where the bulky group would be favorable while the yellow contours represent regions where the bulky group would decrease the activity.

Table 3. Statistical parameters of optimal CoMFA model.

STA	R ²	SDEC	F-test	Q ² _{LOO}	SDEP _{LOO}	Q ² _{LTO}	SDEP _{LTO}	Q ² _{LMO}	SDEP _{LMO}
CoMFA (ffdsel)	0.968	0.0574	61.935	0.5177	0.2254	0.5012	0.2293	0.4731	0.0139
CoMSIA (uvepls)	0.985	0.0405	126.7	0.576	0.2114	0.5628	0.2146	0.51	0.0191

Ffdsel: fractional factorial design (FFD)

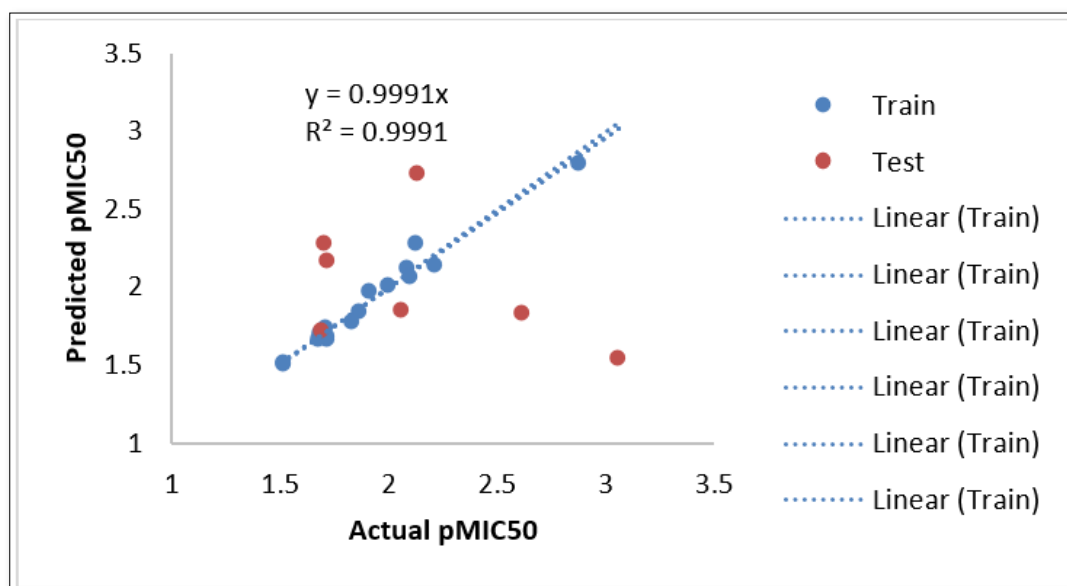


Figure 4. Activity plots of actual pMIC50 vs. predicted pMIC50 of training and test set of iminoguanidine-based inhibitors of HemO by CoMFA models.

As shown in **Figure 4**, green contour near the R_s position specified bulky groups at this position would increase the potency. The electrostatic field is indicated by blue and red contours, which demonstrate the regions where an electron-donating group and an electron-withdrawing group would be favorable, respectively. Great attention should be paid to the red region which indicates that the presence of a negatively charged group would increase the activity. The small yellow contours at the tail portions of the R₂ residue show that too long alkyl chains might have a negative influence on its activity. A large yellow contour indicates that the -OH group is preferable at this position as compared to OCH₃. In an electrostatic field, red and blue contours are mostly distributed near to the core of the iminoguanidine. So, the electronegative and electropositive substituents in these

regions are likely to boost biological activity. The red contours surrounded the R₁ and R₃ position suggests that the electron-rich substituents in this region are likely to enhance biological activity. It indicates R₁ group substituent is preferable than R₃ group substituent because R₃ group substituent is more electron-rich and it allows the electrons to be donated easily but not capable of ionization. The blue electrostatic contour near the R₃ position indicates the presence of an electron-deficient group is favorable at this position.

The CoMFA model (uvepls) for iminoguanidine-based inhibitor of HemO yielded a strong PLS correlation with $R^2 = 0.985$, $Q^2_{LOO} = 0.576$, $Q^2_{LTO} = 0.563$, and $Q^2_{LMO} = 0.51$ respectively, indicating an excellent internal prediction

Table 4. The experimental and predicted pMIC₅₀ values of the training and test set.

Name	pMIC ₅₀	ffdsel	Residual values	uvepls	Residual values
2	2.0828	2.1155	-0.0327	1.9769	0.1059
3	2.0899	2.0729	0.017	2.1038	-0.0139
4	1.9978	2.0067	-0.0089	1.9096	0.0882
5	1.8621	1.8473	0.0148	1.8692	-0.0071
6	2.1179	2.2826	-0.1647	2.3303	-0.2124
7	1.6848	1.6996	-0.0148	1.6322	0.0526
8	1.7193	1.6659	0.0534	1.6644	0.0549
9	1.6749	1.6659	0.009	1.6644	0.0105
10	1.7185	1.6768	0.0417	1.8176	-0.0991
12	1.5065	1.5076	-0.0011	1.5521	-0.0456
13	2.2095	2.1418	0.0677	2.1771	0.0324
17	1.9058	1.9704	-0.0646	1.8483	0.0575
21	2.8751	2.7919	0.0832	2.8051	0.07
23	1.8235	1.7789	0.0446	1.8262	-0.0027
24	1.7101	1.7466	-0.0365	1.7621	-0.052
25	1.5092	1.5175	-0.0083	1.5484	-0.0392
1*	1.6902	1.7236	-0.0334	1.7129	-0.0227
11*	2.0531	1.8593	0.1938	1.8144	0.2387
14*	1.7202	2.1649	-0.4447	2.1906	-0.4704
16*	1.6998	2.2818	-0.582	2.3389	-0.6391
19*	3.0569	1.5399	1.517	1.8087	1.2482
20*	2.6107	1.8275	0.7832	1.7982	0.8125
22*	2.1271	2.7255	-0.5984	2.73	-0.6029

power of the established model. The conventional R₂, F-test, and SDEC for the CoMFA model were found reasonable. The values of experimental and predicted activities along with the residual values of the training set and test set molecules are summarized in **Table 4**. The scatter plot of the observed vs. predicted values of pMIC₅₀ of both the training and test set of CoMFA models is shown in **Figure 5**. This data shows that the observed and predicted activities of inhibitors are very close to each other. Most of the molecules show residual values less than 0.4. This graphical representation conforms to the good predictive power of the established model and also indicate that the developed CoMFA model is reliable and could be used in designing new inhibitors (**Figures 6 and 7**).

The Green and yellow contours represent the steric fields (79.54%). In detail, the green region in the steric contour maps indicates an area where the bulky groups are favored for activity while the yellow contours represent regions where the bulky groups are not favored for the activity. The red and blue contour represents electrostatic contour maps (20.46%). The blue contour defines a region of space where positively charged substituent increases activity, whereas the red contour defines a region of space where negatively charged substituent increases activity. This sterically crowded R₂ group substituent may bring a hydrophobic nature to the parent iminoguanidine thereby enhances the activity.

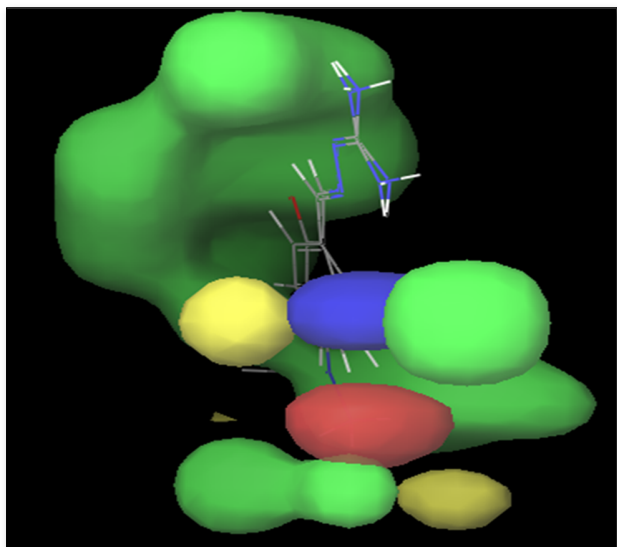


Figure 5. (FFDSEL): Steric and electrostatic contour maps representing the comparative molecular field analysis (CoMFA) model for iminoguanidine-based inhibitors of HemO. Green regions and yellow region indicate areas where steric interactions increase and decrease activity. Blue denote enhancing and detrimental electrostatic effects with the positively charged probe, respectively.

MOLECULAR DOCKING RESULTS: INTERACTION MECHANISM BASED ON THE BINDING FREE ENERGY ANALYSIS

The docking results confirm the QSAR and 3D-QSAR analysis and experimental values (**Table 1**). The docking results of selected compounds are shown in **Table 5**. Compound 12 (**Figure 8**), which is the best-performing molecule in the data set, has the most binding energy (-6.54kcal/mol) and inhibition constant (K_i) as 16.01 μ M. The steric (vdw_lib_desolve_energy) of -6.88 and the electrostatic_energy of 0.03 with ligand efficiency of -0.41. The docked conformations showed that all ligands bind to the active residues in the predefined hydrophobic binding pocket. The key residues in the binding pocket were Ala263, Glu325, Leu326, Asn79, Gly261, Ile327, Val260, Glu329, Phe259, Val328, Asn262, and Ile266.

MOLECULAR DYNAMICS SIMULATION RESULTS

The MD simulations were performed using the NAMD package [17] and was performed with CHARMM 27 force field. The parameter file par_all27_prot_lipid.inp was used. First the protein-ligand complex was solvate in water box as shown in **Figure 9**. Minimization was performed to optimize the initial structure of protein-ligand complex. After that, the temperature of system was gradually heated up from 0K to 310K in 50ps. Subsequently, the system was equilibrated at 310K for 100ps with NVT ensemble. Results of this simulation study revealed that the protein get stabilized after

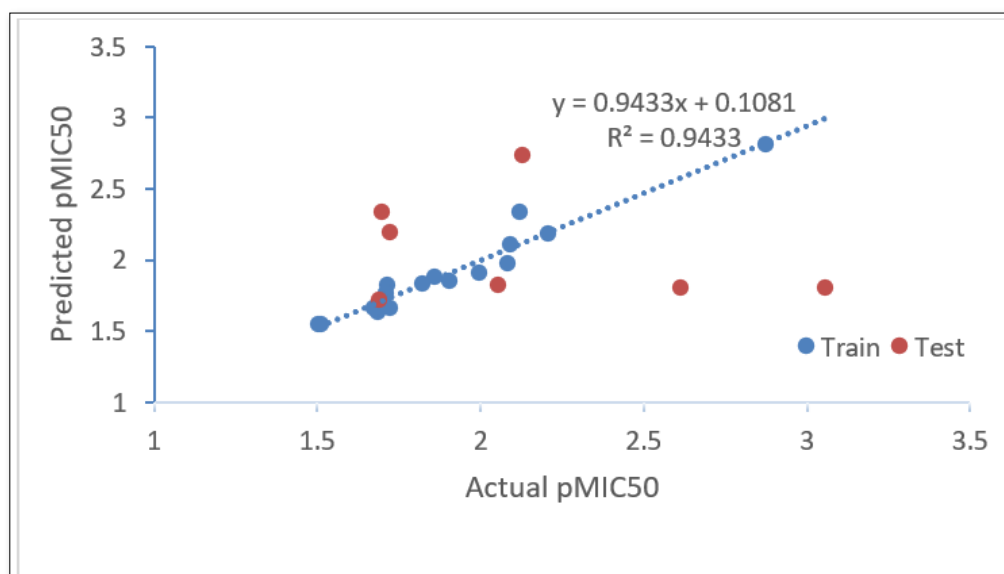


Figure 6. Activity plots of actual pMIC50 vs. predicted pMIC50 of training and test set of iminoguanidine-based inhibitors of HemO by CoMFA models.

500fs and it stayed stabilized up to 1000fs with respect to kinetic energy, temperature, total energy and RMSD was also determined as shown in **Figure 10**, respectively.

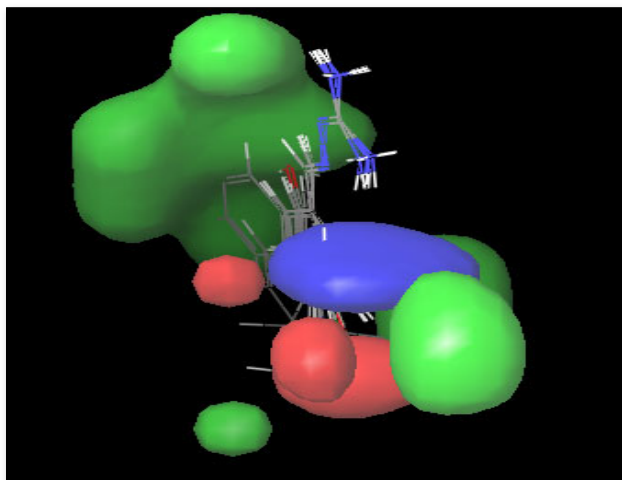


Figure 7. (UVEPLS): MIF contour maps for the iminoguanidine: Steric and electrostatic contour maps representing the comparative molecular field analysis (CoMFA) model for iminoguanidine-based inhibitors of HemO. Green and yellow contours indicate regions where bulky groups increase or decrease the activity. Blue and red contours indicate regions where electron-donating or electron-withdrawing groups increase the potency.

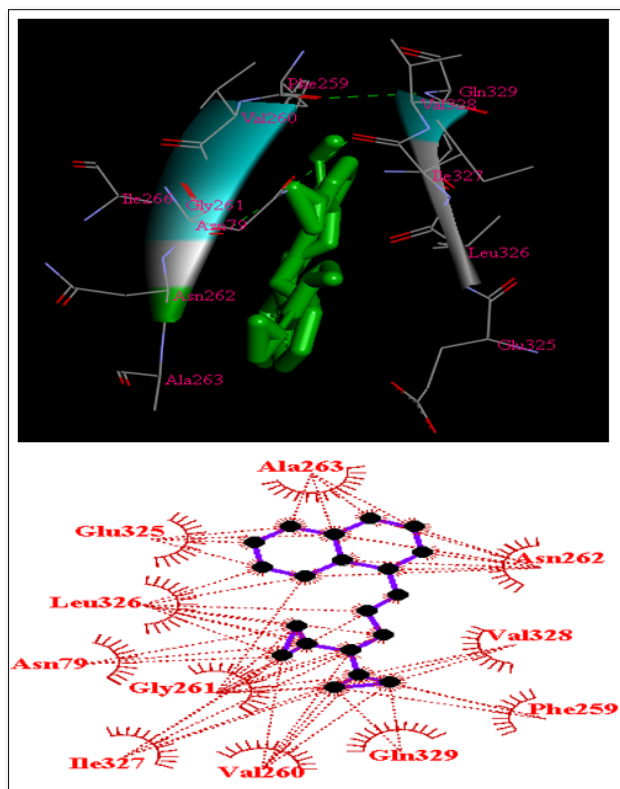


Figure 8. Molecular docking interactions of compound 12 with unsimulated protein structure.

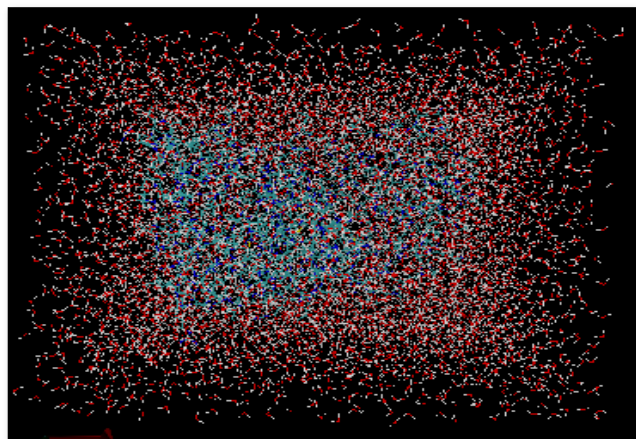


Figure 9. Solvate protein (pdb entry code: 4mgf) in a water box.

This result of molecular dynamics simulation proved that the docking of compound 12 was done correctly in the active site of HemO. It also gave the idea that the compound 12 formed a stable complex with protein when protein was simulated up to or beyond 500fs concerning temperature (at or above 298 K), kinetic energy (at or above 23895.44 kJ/mol) and total energy (at or above -57629.53 kJ/mol). To check the consistency of simulated and docked protein structure, they were arrayed over each other as shown in **Figure 11**. Exact alignment of docked and simulated protein structure with RMSD value 0 proved that the sequence of amino acid remained the same before and after the simulation and hence there was no change in the binding cavity of protein. From the results, the increase in the hydrophobic bond interaction with improved docking score was due to the stable conformation of simulated protein not due to change in the sequence of amino acids.

DRUG-LIKE PROPERTIES

The investigation of the World Drug Index (WDI), leads to Lipinski's 'rule-of five' [22-24]- that identified several critical properties that should be considered for compounds with oral delivery in mind. These properties, are usually viewed more as guidelines rather than absolute cutoffs, and they are:

Molecular weight, MW < 500 Da

Lipophilicity, log P or the calculate of 1-octanol-water partition coefficient, ClogP < 5 (or MlogP < 4.15)

Number of hydrogen-bond donors, OH plus NH count, <5

Number of hydrogen-bond acceptors, O plus N atoms, <10

The Ghose's "rule of four"[25, 26] are:

Calculated logP is between -0.4 and 5.6 or AlogP between 1.3 and 4.1

Molecular weight, between 160 and 480

Molar refractivity, between 40 and 130

Total number of atoms, between 20 and 70

Table 5. Docking score and PSA of iminoguanidine-based inhibitors of HemO.

Compound	Docking Score	LE	Inh_Con	vdw	Electr	PSA
1	-4.95	-0.38	234.54	-5.37	0.0	64.537
2	-5.35	-0.36	119.93	-5.91	-0.08	65.687
3	-5.9	-0.45	47.33	-6.18	-0.03	83.266
4	-5.35	-0.38	119.76	-5.8	-0.02	70.39
5	-5.39	-0.36	112.56	-5.18	-0.9	103.31
6	-5.44	-0.42	103.53	-5.77	-0.1	63.978
7	-5.96	-0.37	43.03	-6.25	-0.11	64.496
8	-5.56	-0.43	84.24	-5.96	-0.01	64.461
9	-5.72	-0.44	64.1	-6.00	-0.1	63.712
10	-5.33	-0.36	122.88	-5.87	-0.02	64.552
11	-5.35	-0.38	119.82	-5.82	-0.16	71.511
12	-6.54	-0.41	16.01	-6.88	0.03	64.416
13	-5.29	-0.41	133.14	-5.66	-0.01	64.551
14	-5.01	-0.39	213.11	-5.48	-0.01	64.535
15	-6.17	-0.31	29.78	-6.95	-0.1	71.013
16	-5.66	-0.44	71.39	-5.95	-0.07	63.444
17	-5.36	-0.38	118.62	-6.06	-0.13	77.871
18	-5.85	-	-	5.05	0.08	64.527
19	-5.34	-0.38	120.86	-6.2	-0.13	103.854
20	-5.29	-0.41	133.28	-5.87	-0.13	84.311
21	-5.61	-0.43	77.5	-6.28	-0.05	84.224
22	-5.45	-0.39	100.87	-6.11	0.02	71.416
23	-5.53	-0.43	88.97	-5.82	-0.11	64.501
24	-5.44	-0.47	102.35	-5.85	-0.04	64.452
25	-5.4	-0.39	110.8	-5.86	-0.04	64.451

LE=ligand_efficiency; Inh_con=inhibition constant; vdw=vdw_lib_desolve_energy; `Electr=electrostatic_energy

The Veber's "rule of two" [27] are:

The number of rotating bonds (nrotb) < 10

Polar surface area (PSA) equal to or less than 140 Å²

The above-mentioned parameters were calculated for all the series of the iminoguanidine derivatives (**Table 3**). Compound is probs to be dynamic orally if it follows the Lipinski and Veber rule [22-24,27]; however, it is deserving indicating some oral drugs which do not respect the rule. The partition coefficient (Log P) of Octanol/water is widely used for an estimation of membrane penetration and permeability, including gastrointestinal absorption, brain-brain crossing

(BBC), and correlations with pharmacokinetic properties [28,29]. In other words, Low molecular weight drug molecules (<500) are easily transported, diffused and absorbed in relation to heavy molecules. The amounts of hydrogen bond acceptors (O and N atoms) and hydrogen-bonding donors (NH and OH) are found to be critical in drug development because they have an influence on absorption and permeation [30]. Screening process with Lipinski Rule of Five showed that all the compounds meet the criteria of drug-likeness assessment except compound 19. However, compound 19 were rejected with one violation (**Table 6**). According to the screening process with Ghose rules showed

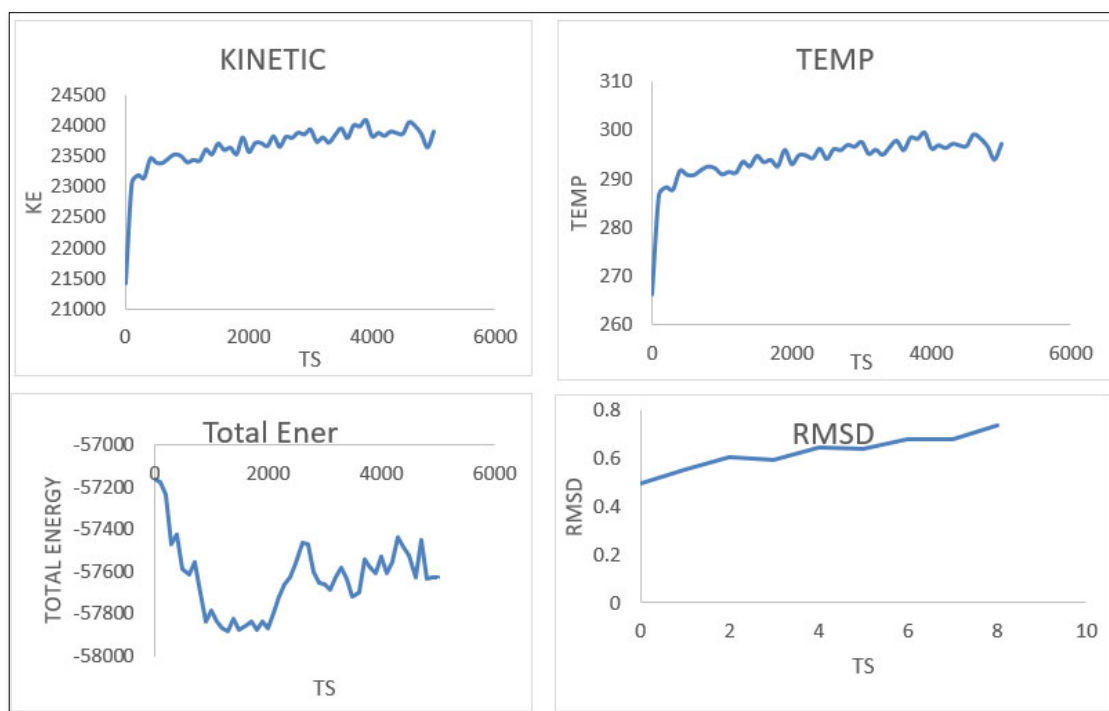


Figure 10. Molecular dynamics simulations study of A-chain of protein (PDB ID: 4mgf) concerning (A) time vs Kinetic; (B) time vs temperature; (C) time vs total energy; (D) time vs RMSD.

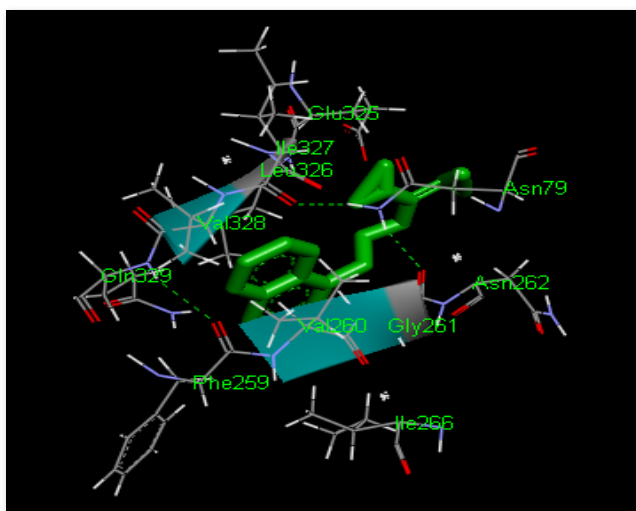


Figure 11. Docking interactions of compound 12 with simulated protein structure.

that all compounds meet the criteria except compound 19 were their AlogP is less than 1.3 (**Table 7**) [25-26]. Polar surface area is defined as a sum of surfaces of polar atoms (usually oxygens, nitrogens and attached hydrogens) in a molecule. This parameter has been shown to correlate very well with the human intestinal absorption, Caco-2 monolayer's permeability, and blood-brain barrier penetration. Molecules with PSA values of 140 Å² or more

are expected to exhibit poor intestinal absorption. PSA values using Spartan'14 software (**Table 5**) and MedChem Designer (**Table 6**) are below 140 Å². Number of rotatable bonds (nRotBt) is a simple topological parameter that measures molecular flexibility and is considered to be a good descriptor of oral bioavailability of drugs [31]. All the screened compounds using PaDEL software v2.0 were flexible, all the compounds 1 to 25 have rotatable bonds less than 8 (**Table 7**). However, the screening process with Veber rules [27], all the 25 compounds meet the criteria of drug likeness assessment, suggesting that these compounds theoretically have ideal oral bioavailability. These physicochemical parameters are associated with acceptable aqueous solubility and intestinal permeability that are the first steps in oral bioavailability.

The ideas such as ligand efficiency (LE), which correlates potency and molecular weight, and ligand-lipophilicity efficiency (LLE), which correlates potency and lipophilicity, have been employed by medicinal chemists to all together evaluate compound potency to properties [32]. Ligand efficiency (LE) is an important new concept, which estimates the efficiency of a binding interaction concerning the magnitude of ligand physical properties, most notably molecular weight [33]. The ligand efficiency (LE) is defined as followed:

$$LE = \frac{pMIC_{50}}{\text{number of heavy atoms}}$$

Table 6. Calculation of electronic parameters of drug likeness of the iminoguanidine-based by using MedChem Designer.

Name	MlogP	S+logP	S+logD	RuleOf5	RuleOf5_Code	MWt	M_NO	T_PSA	HBDH
1	2.618	1.325	0.988	0	<None>	196.64	4	76.76	4
2	2.134	0.779	-0.61	0	<None>	205.264	5	80	4
3	1.513	-0.258	-0.597	0	<None>	178.195	5	96.99	5
4	1.831	0.57	0.051	0	<None>	192.222	5	85.99	4
5	2.08	0.782	0.757	0	<None>	207.193	7	122.58	4
6	2.461	1.001	0.581	0	<None>	180.186	4	76.76	4
7	3.064	1.909	1.652	0	<None>	212.256	4	76.76	4
8	2.618	1.381	1.094	0	<None>	196.64	4	76.76	4
9	2.618	1.346	1.065	0	<None>	196.64	4	76.76	4
10	2.945	1.654	0.651	0	<None>	204.276	4	76.76	4
11	1.831	0.543	-0.367	0	<None>	192.222	5	85.99	4
12	3.064	1.874	1.645	0	<None>	212.256	4	76.76	4
13	2.461	0.916	0.23	0	<None>	180.186	4	76.76	4
14	2.771	1.401	0.772	0	<None>	241.096	4	76.76	4
15	3.258	1.978	1.166	0	<None>	268.32	5	85.99	4
16	2.771	1.41	0.9	0	<None>	241.096	4	76.76	4
17	2.252	0.058	0.004	0	<None>	257.096	5	96.99	5
18	3.522	2.228	1.479	0	<None>	238.294	4	76.76	4
19	1.009	-0.477	-0.806	1	Hb	194.194	6	117.22	6
20	1.513	-0.173	-0.893	0	<None>	178.195	5	96.99	5
21	1.513	0.098	-0.307	0	<None>	178.195	5	96.99	5
22	1.831	0.636	0.251	0	<None>	192.222	5	85.99	4
23	2.461	1.008	0.588	0	<None>	180.186	4	76.76	4
24	2.771	1.472	0.996	0	<None>	241.096	4	76.76	4
25	3.19	2.085	1.921	0	<None>	231.085	4	76.76	4

Leach and his coworkers propose lipophilicity per unit of in vitro potency, or ligand-lipophilicity efficiency [33], as a more important objective for lead generation and optimization programs. The ligand-lipophilicity efficiency (LLE/LipE) is defined as followed:

$$LLE = pMIC_{50} - S + \log D \text{ (or } \log D)$$

It is suggested to target a LipE is between 5 and 7 or even higher [34].

Ligand efficiency-dependent lipophilicity index (LELP) to combine molecular size and lipophilia into a single measure of efficacy [35]. The optimal value of LELP scores are $-10 < \text{LELP} < 10$ [36] and is defined as:

$$LELP = \frac{\log P}{LE}$$

Total polar surface area (TPSA) is a very useful parameter for prediction of drug transport properties [31]. TPSA (PaDEL software v2.0) of iminoguanidine derivatives were found in the range of 37.3457- 135.4850 and is well below the 140 Å², and we can observe obviously that all the title compounds 1 to 25 exhibited a great percentage absorption (%ABS) [37] ranging from 72.1514 to 96.1157%, indicating that these compounds should have good cellular plasmatic membrane permeability except compound 19 (**Table 7**). Some compound is situated in the suggested range $-10 < \text{LELP} < 10$. In the other hand, the compounds have LELP less than 16.5, which mean that these compounds are in the Lipinski zone (ROF-score = 4). Except the compounds 4 and 17 their LELP is 39.1730 and 476.4490 respectively are in agreement with their weak ROF-score <4 (**Table 7**).

Table 7. Structure activity, calculated metrics and Lipophilicity indice.

Compound	LELP	LLE	%ABS	TPSA	LE	nRotBt	nAtom	ALogP
1	1.7107	0.7022	88.6981	58.8460	0.1300	5	22	1.0728
2	-3.4144	2.6928	94.5916	41.7635	0.1389	7	30	0.4751
3	-3.5007	2.6869	67.0692	121.5385	0.1608	5	23	0.141
4	39.1730	1.9468	86.8864	64.0973	0.1427	6	26	0.2055
5	2.4599	1.1051	95.9328	37.8758	0.1241	6	24	0.9174
6	3.6453	1.5369	77.7409	90.6060	0.1629	5	22	0.6139
7	1.0199	0.0328	93.9921	43.5012	0.1053	4	28	1.2485
8	1.5716	0.6253	96.1157	37.3457	0.1323	5	22	1.0728
9	1.5726	0.6099	89.1249	57.6090	0.1288	5	22	1.0728
10	2.6398	1.0675	93.6771	44.4142	0.1146	7	31	1.0814
11	-5.5942	2.4201	89.2829	57.1509	0.1466	6	26	0.2055
12	0.9158	-0.1385	95.4451	39.2895	0.0942	4	28	1.2485
13	9.6067	1.9795	72.1514	106.8075	0.1700	5	22	0.6139
14	2.2282	0.9482	95.8719	38.0525	0.1323	5	22	1.1568
15	1.1711	0.1995	91.1392	51.7703	0.0683	7	36	1.3226
16	1.8887	0.7998	88.8201	58.4925	0.1308	5	22	1.1568
17	476.4490	1.9018	83.7292	73.2487	0.1361	6	23	0.5937
18	1.1978	0.2926	96.0548	37.5224	0.0984	5	32	1.7163
19	-3.7927	3.8629	33.3112	219.3878	0.2184	6	24	-0.422
20	-2.9235	3.5037	63.2197	132.6966	0.2008	5	23	0.141
21	-9.3650	3.1821	62.2577	135.4850	0.2212	5	23	0.141
22	8.4745	1.8761	89.5661	56.3301	0.1519	6	26	0.2055
23	3.1011	1.2355	72.3343	106.2774	0.1403	5	22	0.6139
24	1.7170	0.7141	96.0446	37.5518	0.1315	5	22	1.1568
25	0.7856	-0.4118	96.1157	37.3457	0.1078	6	22	1.4415

Lligand efficiency (LE); Efficacy of ligand lipophilicity efficiency LipE (LLE); Ligand efficiency-dependent lipophilicity index (LELP); Percentage of absorption (%ABS)

OPTIMIZE CLEARANCE AND ORAL ABSORPTION

S+logD and molecular weight are the key factors to determine the penetrability of drug candidates [38]. S+logD threshold values for high permeability is a lower value for ligands or compounds of lower molecular weight [39]. To achieve high permeability ligands with molecular weight less than 415 and S+logD less than 1.3 [38]. Most of our compounds are within the limits of Lipinski rules [22-24] except compound 19. The Golden Triangle [40] is a visual image built up from *in vitro*

permeability, *in vitro* authorization, and computational data designed to help medicinal chemists obtain stable, pervious, and strong drug candidates. The chance of achieving effectiveness, stability, and permeability is actualized by moving the design properties into an area with a service line of S+logD = 0.0 to S+log D = 2.5 at MWt = 200 and a peak at S+logD = 0 – 2.5 and MWt = 300, these boundaries give a triangular shape, called the golden triangle. For our series of compounds, it is seeming that all the compounds except 2 and 17, have good permeability and low clearance because are

concentrated within the golden triangle area as shown in **Figure 5**. These results should help design compounds with improved permeability.

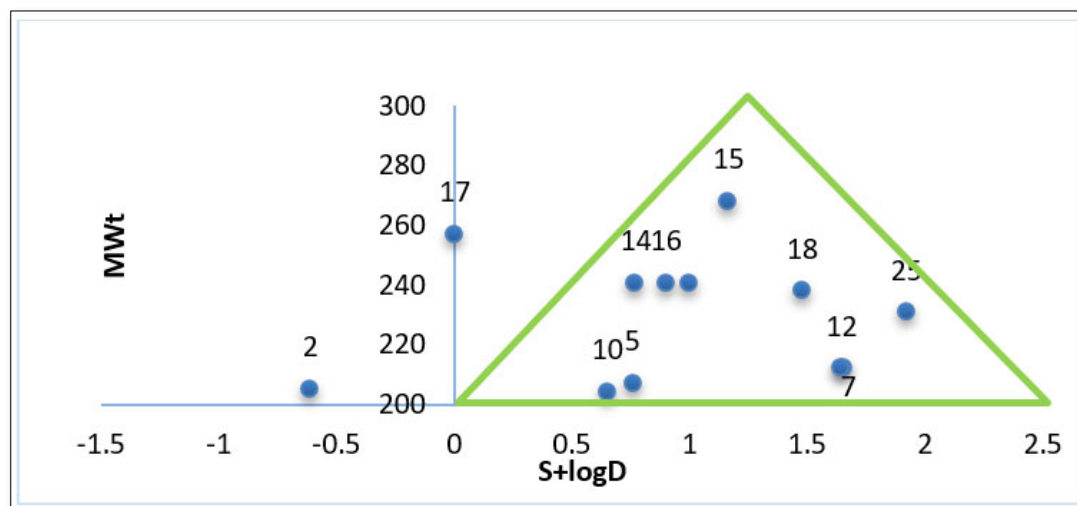


Figure 12. *In vitro* penetrability and clearance curves over MWt and S+logD.

CONCLUSION

In this study, carried out 25 iminoguanidine derivatives based on DFT/QSAR prove that our calculated results are similar to experimental data taken from literature. MIFs studies, against experimented biological activities. The atom-based and pharmacophore-based alignment with varying the 3D grid spacing method was used to provide the model for MIFs analysis. These studies have recognized that the model derived through MIFs studies is quite reliable and significant. We have investigated that the PLS analysis at 2.0Å 3D grid spacing by Open 3DQSAR tools has presented quite statistical results in terms of R² and Q²LOO values and showed a high degree of agreement with the experimented inhibitory activities. Also, the molecular interaction mechanism of inhibitor binding to HemO active binding site was explained based on molecular docking binding free energy analysis and molecular dynamics simulations. Our results suggested that inhibitor can exactly bind to the active binding site of HemO to display inhibitory activity and the van der Waals interactions could be driving force for the binding of inhibitor with HemO. A series of differently substituted iminoguanidine has been considered to check its potential biological activity. It has been found that these compounds should exhibit good cell plasma membrane permeability. Likewise, they satisfy Lipinski's rule of five, Veber's rule of two and Ghose's rule of four, indicating that these compounds theoretically would not have oral bioavailability problems. In particular, compound 12 which has the highest binding energy value of the dataset could be considered as a good candidate for biological testing.

REFERENCES

1. Podkalicka P, Mucha O, Józkwicz A, Dulak J, Łoboda A (2018) Heme oxygenase inhibition in cancers: Possible tools and targets. *Contemp Oncol (Pozn)* 22: 23-32.
2. Araujo JA (2012) Nrf2 and the promotion of atherosclerosis: Lessons to be learned. *Clin Lipidol* 7: 123-116.
3. Deredge DJ, Huang W, Hui C, Matsumura H, Yue Z, et al. (2017) Ligand-induced allostery in the interaction of the *Pseudomonas aeruginosa* heme binding protein with heme oxygenase. *Proc Natl Acad Sci* 114: 3421-3426.
4. Ochsner UA, Johnson Z, Vasil ML (2000) Genetics and regulation of two distinct heme-uptake systems, phu and has, in *Pseudomonas aeruginosa*. *Microbiol* 146: 185-198.
5. Alekshun MN, Levy SB (2004) Targeting virulence to prevent infection: To kill or not to kill? *Drug Discov Today: Ther Strategies* 1: 483-489.
6. Rasko DA, Sperandio V (2010) Anti-virulence strategies to combat bacteria-mediated disease. *Nat Rev Drug Discov* 9: 117-128.
7. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 98: 5648.
8. Lee C, Yang W, Parr RG (1988) Development of the colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev* 37: 785-789.

9. Tosco P, Balle T (2011) Open3DALIGN: An open-source software aimed at unsupervised ligand alignment. *J Comput Aided Mol Des* 25: 777-783.
10. Tosco P, Balle T (2011) Open3DQSAR: A new open-source software aimed at high-throughput chemometric analysis of molecular interaction fields. *J Mol Model* 17: 201-208.
11. Pastor M, Cruciani G, Clementi S (1997) Smart region definition: a new way to improve the predictive ability and interpretability of three-dimensional quantitative structure-activity relationships. *J Med Chem* 40: 1455-1464.
12. Massimo SCB, Costantino G, Cruciani G, Riganelli D, Valigi R (1993) Generating optimal linear PLS estimations (GOLPE): An advanced chemometric tool for handling 3D-QSAR problems. *Quant Struct Relat* 12: 9-20.
13. Wold S, Sjöström M, Eriksson L (2001) PLS-regression: A basic tool of chemometrics. *Chemometr Intell Lab Syst* 58: 109-130.
14. Conkey BJM, Sobolev V, Edelman M (2002) The performance of current methods in ligand-protein docking. *Curr Sci* 83: 845-855.
15. Huey R, Morris G (2008) Using AutoDock 4 with AutoDockTools: A tutorial. The Scripps Research Institute, pp: 1-56.
16. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 30: 2785-2791.
17. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26.
18. Eriksson L, Jaworska J, Worth AP, Cronin MTD, Dowell MM, et al. (2003) Methods for reliability and uncertainty assessment and for applicability evaluations of classification-and regression-based QSARs. *Environ Health Perspect* 111.
19. Roy K, Paul S (2008) Exploring 2D and 3D QSARs of 2, 4-diphenyl-1, 3-oxazolines for ovicidal activity against *tetranychus urticae*. *QSAR Comb Sci* 28: 406-425.
20. Netzeva TI, Worth AP, Aldenberg T, Benigni R, Cronin MTD, et al. (2005) Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. *Altern Lab Anim* 33: 155-173.
21. Golbraikh A, Tropsha A (2002) Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *J Comput Aided Mol Des* 16: 357-369.
22. 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* 23: 3-25.
23. Lipinski CA (2004) Lead-and drug-like compounds: The rule-of-five revolution. *Drug Discov Today Technol* 1: 337-341.
24. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 64: 4-17.
25. Ghose AK, Crippen GM (1986) Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure-activity relationships. I. Partition coefficients as a measure of hydrophobicity. *J Comput Chem* 7: 565-577.
26. Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. A qualitative and quantitative characterization of known drug databases. *J Comb Chem* 1: 55-68.
27. Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, et al. (2002) Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 45: 2615-2623.
28. Winiwarter S, Bonham NM, Ax F, Hallberg A, Lennernäs H, et al. (1998) Correlation of human jejunal permeability (in vivo) of drugs with experimentally and theoretically derived parameters. A multivariate data analysis approach. *J Med Chem* 41: 4939-4949.
29. Waterbeemd HV, Camenisch G, Folkers G, Chretien JR, Raevsky OA (1998) Estimation of blood-brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors. *J Drug Target* 2: 151-165.
30. Bemis GW, Murcko MA (1996) The properties of known drugs. 1. Molecular frameworks. *J Med Chem* 39: 2887-2893.
31. Rouane A, Tchouar N, Kerassa A, Belaidi S, and Cinar M (2017) Structure-based virtual screening and drug-like of quercetin derivatives with anti-malaria activity. *Rev Theor* 5.
32. Wager TT, Hou X, Verhoest PR, Villalobos A (2010) Moving beyond rules: The development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS Chem Neurosci* 1: 435-449.
33. Leach AR, Hann MM, Burrows JN, Griffen EJ (2006) Fragment screening: an introduction. *Mol Biosyst* 2: 429-446.

34. Smida KB, Belaidi S, Benbrahim I, BoughdiriS (2015) Theoretical studies of structure/activity relationships applied to flavone derivate for drug discovery. *Res J Pharm Biol Chem Sci* 6: 874-885.
35. Kesserü GM, Makara GM (2009) The influence of lead discovery strategies on the properties of drug candidates. *Nat Rev Drug Discov* 8: 203-212.
36. Hopkins AL, Groom CR, Alex A (2004) Ligand efficiency: A useful metric for lead selection. *Drug Discov Today* 9: 430-431.
37. Zhao YH, Abraham MH, Le J, Hersey A, Luscombe CN, et al. (2002) Rate-limited steps of human oral absorption and QSAR studies. *Pharmacol Res* 19: 1446-1457.
38. Waring MJ (2009) Defining optimum lipophilicity and molecular weight ranges for drug candidates' molecular weight dependent lower log D limits based on permeability. *Bioorg Med Chem Lett* 19.
39. Martin YC (2005) A bioavailability score. *J Med Chem* 48.
40. Johnson TW, Dress KR, Edwards M (2009) Using the golden triangle to optimize clearance and oral absorption. *Bioorg Med Chem Lett* 19.