MOLECULAR MODELLING AND DYNAMIC SIMULATION OF CORROSION INHIBITORS ON STEEL IN ACIDIC MEDIUM

Supported by tetfund TERTIARY EDUCATION TRUST FUND

A. U Bello*, A. Uzairu and G. A. Shallangwa
Department of Chemistry, Ahmadu Bello University, Zaria, Nigeria
*Corresponding author: abdallahbum@yahoo.com

Received: December 10, 2017 **Accepted:** July o6, 2018

Abstract: Experimental corrosion study was often high-cost and time-consuming since large-scale trial experiments were carried out. An In-silico method was used to study the inhibition performance of twenty-five amino acids and related compounds. Density Functional Theory (B3LYP/6-31G*) quantum chemical calculation method was used to find the optimized geometry of the studied inhibitors. Additionally, a linear quantitative structure-activity relationship (QSAR) model was built by Genetic Function Approximation (GFA) method to run the regression analysis and establish correlations between different types of descriptors and the measured corrosion inhibition efficiencies which was used to predict the corrosion inhibition efficiencies of the studied inhibitors. The prediction of corrosion efficiencies of these inhibitors nicely matched the experimental measurements. The correlation parameters obtained are $R^2_{train} = 0.98, R^2_{adjusted} = 0.98, Q^2_{LOO} = 0.97,$ R^2_{test} =0.86. This indicates that the model was excellent on verifying with internal and external validation parameters. The affection of acidic solution was considered in molecular dynamics simulation and the calculated adsorption energies for most of the inhibitors is >100 kcal mol^{-1} suggesting chemisorptive interactions.

Keywords: Amino acids, DFT (B3LYP/6-31G*), QSAR, GFA, Molecular dynamics simulation

Introduction

Corrosion is an undesirable process that affects several areas of industrial activity, especially the oil industry, resulting in huge economic losses (Szklarska-Smialowska, 2005). It is a common problem for steel and directly impacts its cost and safety. The corrosion of iron can cause structural damage and lead to changes in the mechanical and chemical properties of plants, vessels, pipes, and other processing equipment. Several countries have attempted to relate the cost of corrosion to their gross national product. The annual cost of corrosion worldwide is estimated at $ 2.2 Trillion (2010), which is about 3% of the world's gross domestic product of $ 73.33 Trillion (United States in 2012-G2MT Labs, Al Hashem, 2011). Preventing the corrosion of steel has played an important role in various industries, especially in the chemical and petrochemical processing industries that employ the use of steel (Singh *et al.*, 2016). Although it is not possible to completely avoid the corrosion process but there are several ways to prevent it or slow down the corrosion rate (Nwankwo, 2016). Several organic compounds especially those that contain N, O, S, and P heteroatoms, as well as π-electron systems have been previously used as corrosion inhibitors for metals in aqueous solutions (Murulana *et al*., 2012; Arslan *et al*., 2009; Praveen and Venkatesha, 2009). Although many heterocyclic compounds have been successfully used as corrosion inhibitors in several metallic systems, most of them were toxic and non-biodegradable (Eddy and Mamza, 2009).

With the current advancement of environmental safety, researchers were focused on the environmental friendly corrosion inhibitors (Yadav *et al.*, 2012; Lyon, 2004). Amino acids which were non-toxic, easily available and completely soluble in aqueous media were considered to be the most promising green inhibitors (Lyon, 2004). Various experimental and theoretical techniques have been used to study the corrosion in an acidic solution such as electrochemical, weight loss, quantum chemical and surface morphology (Singh *et al*., 2016; Nan *et al*., 2015; Khaled, 2012; Hassan *et al*., 2007; Yurt *et al*., 2005). Though the experimental measures are the conventional methods but were often expensive and time-consuming since large-scale trial experiments were often carried out (Zhang *et al*., 2011; Muster *et al*., 2009). In-silico techniques, which can overcome these challenges, have attracted researchers' great attention in recent years (Beese, 2013; Yuan, 2013; Venzlaff, 2013). Quantum chemical methods have already proven to be very useful in determining the molecular structure as well as elucidating the electronic and reactivity centers of a compound (Kraka and Cremer, 2000).

Recently, quantitative structure-activity relationship (QSAR) and Molecular dynamic simulation has aroused many researchers' interest in the studies of corrosion inhibitors (Tong *et al*., 2005). The success of the QSAR approach can be explained by the insights offered for the structural determination of chemical properties and the possibility of estimating the properties of the new chemical compounds without any need for them to be synthesized and tested (He and Jurs, 2005). However, the success of any QSAR model depends on the accuracy of input data, selection of the appropriate descriptors, statistical tools, and most important validation of the developed models (He and Jurs, 2005; Ghafourian and Cronin, 2005; Tropsha *et al*., 2003). It has been proved to be very helpful for predicting the inhibition efficiencies of novel corrosion inhibitors (Zhang *et al*., 2005; Zhao *et al*., 2014).

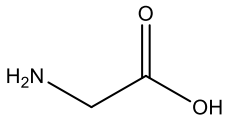
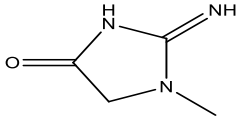
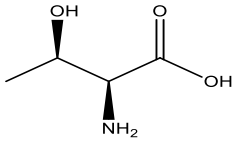
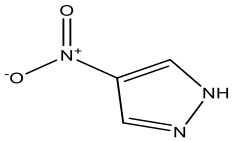
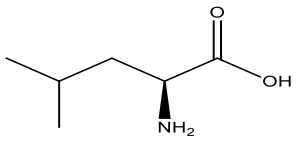
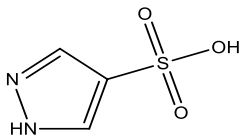
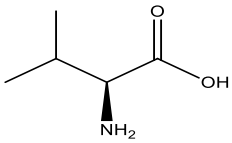
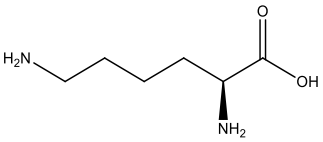
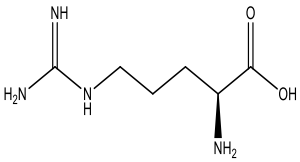
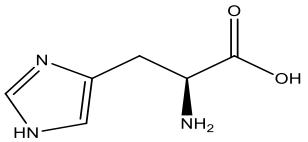
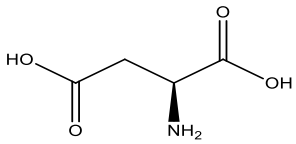
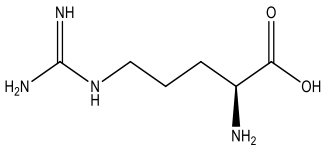
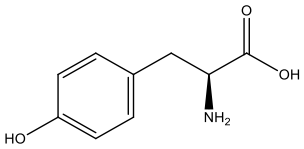
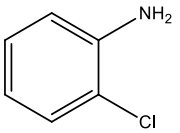
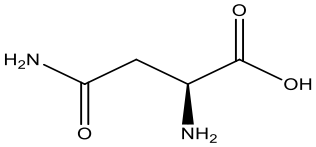
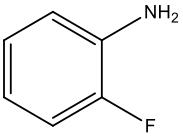
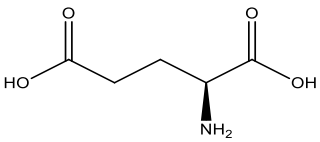
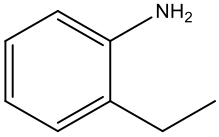
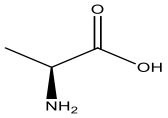
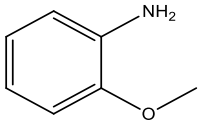
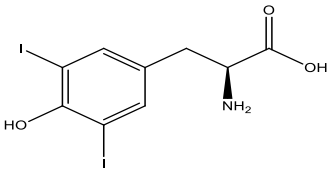
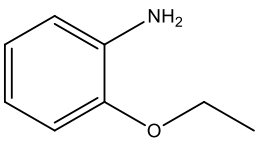
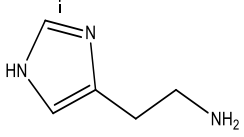
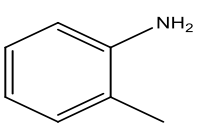
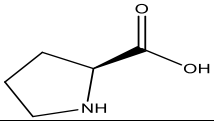
Moreover, to study the adsorption behavior of amino acids on a metal surface, molecular dynamics simulation was used to study the adsorption configuration and adsorption strength of amino acids on a metal surface (Khaled, 2010; Fu *et al.*, 2010; Amin *et al*., 2010; Deng *et al*., 2012; Hu and Dai, 2012). Also various factors, such as the adsorption of the solvent molecules and the affection of the acidic solution, which would influence the adsorption behaviors of the amino acid compounds greatly, should also be considered in the molecular dynamics simulation. In this paper, Molecular modeling and dynamics simulations were carried out to provide theoretical based explanations for the inhibitory behavior of some amino acids derivatives.

Materials and Methods

Data sets

Twenty-five amino acids and their related compounds used as steel corrosion inhibitors were collected from the literature (Khaled and Hackerman, 2003; Hluchan *et al*., 1988; Babić-Samardžija *et al*., 2003) and used for this present study. Their molecular structures and inhibition efficiencies were shown in Table 1.

Table 1: Inhibition efficiencies and molecular structures of the studied inhibitor series

S/N	Compound	%IE	S/N	Compound	%IE
1		50	14		34
2		59	15		43
3		63	16		77.4
4		47	17		75.1
5		80	18		41
6		52	19		71
7		39	20		63.62
8		73	21		71.79
9		53	22		63.24
10		51	23		66.83
11		87	24		49.88
12		75	25		60.09
13		67			

Geometry optimization and QSAR descriptors calculations

Geometry optimization

The chemical structure of each compound in the data sets was drawn with ChemDraw ultra V12.0. Optimization were performed using the SPARTAN'14 V1.1.0 Wave Function programming Package on Dell Intel(R)Core(TM)i7-5500U CPU), 16.00GB RAM @ 2.400GHz 2.400GHz processor on Windows 8.1 Pro 64-bit Operating system, ×64-based processor. The computational method invoked for calculating geometries in the present study is Density Functional Theory (DFT) method in combination with the B3LYP functional with 6-311+G (d,p) basis set (Larif, 2013). The B3LYP hybrid functional of DFT method uses Becke's three-parameter functional (B3) and incorporates a blend of HF with DFT exchange terms associated with the gradient-corrected correlation functional of Lee, Yang, and Parr (LYP) (Becke, 1993). Full geometry optimizations were carried out on neutral molecules in vacuum.

QSAR descriptor calculation

The molecular descriptors used in this QSAR modeling were calculated using Padel descriptor toolkit (Yap, 2011), Spartan 14 software (Wavefunction Inc., 2013) and Material studio software (Accelrys Inc., 2007). Over 500 descriptors ranging from 0-3D were used for this work. The quantum chemical descriptors were the parameters obtained using DFT computation with Spartan 14 software such as the energy of the highest occupied molecular orbital (E-HOMO), the energy of the lowest unoccupied molecular orbital (E-LUMO), Dipole Moment, Polarizability etc. The Spartan files of all the optimized molecules were then saved in SD file format, which is the recommended input format in PaDEL-Descriptor software V2.20 (Yap, 2011). Molecular descriptors such as atom-type topological state descriptors, 2D-Autocorrelations, WHIM, Petitjean shape index, count of chemical substructures and binary fingerprints of chemical substructures were calculated using the paDEL program (PaDEL-Descriptor, 2014). In addition, some structural and topological descriptors were generated from Material studio 8.0 and incorporated into the analysis.

QSAR model building

In order to obtain validated QSAR models, the descriptors generated were divided into training and test sets. The training set was used to generate the model, while the test set was used for the external validation of the model (Kennard and Stone, 1969). The correlation between inhibition efficiency of the molecules against corrosion of steel and the calculated descriptors were obtained via correlation analysis using the Microsoft excel package in Microsoft office 2013. Pearson's correlation matrix was used as a model, in order to select the suitable descriptors for each regression analysis. The selected descriptors were subjected to regression analysis with the experimentally determined inhibition efficiencies as the dependent variable and the selected descriptors as the independent variables using Genetic function approximation (GFA) method in Material studio software.

The number of descriptors in the regression equation was set to be 5, and Population and Generation were set to be 1000 and 1000, respectively. The number of top equations returned was 4 out of which the best one was selected based on statistical significance. Mutation probability was set to be 0.1, and the smoothing parameter was 0.5. The models were scored based on Friedman's Lack of Fit (LOF). In GFA algorithm, an individual or model was represented as a one-dimensional string of bits. The GFA algorithm approach has a number of important advantages over other standard regression analysis techniques. It builds multiple models rather than a single model (Accelrys Inc., 2007). It automatically selects which features are to be used in the models and is better at discovering combinations of features that take advantage of correlations between multiple features (Khaled, 2008). GFA incorporates Friedman's lack-of-fit (LOF) error measure, which estimates the most appropriate number of features, resists over fitting, provides information not available from the standard regression analysis and allows control over the smoothness of fit. Also, it can use a larger variety of equation term types in the construction of its models.

Evaluation of the QSAR model

The predictive ability of the best model was evaluated by internal and external validation parameters. The validation parameters were compared with the minimum recommended standards for a generally acceptable QSAR model (Table 2) (Abdulfatai *et al.*, 2017).

Table 2: Validation metrics for a generally acceptable QSAR model

Symbol	Name	Threshold	Source
R^2	Coefficient of determination	≥ 0.6	Ravichandran *et al.* (2011)
Q^2	LOO cross validation coefficient	> 0.5	Ravichandran *et al.* (2011)
$R^2_{pred.}$	External test set's coefficient of determination	≥ 0.6	Ravichandran *et al.* (2011)
R^2 - Q^2	Difference between R^2 and Q^2	≤ 0.3	Ravichandran *et al.* (2011)
F value	Variation ratio	High	Ravichandran *et al.* (2011)
$P_{95\%}$	Confidence interval at 95% confidence level.	< 0.05	Jaiswal *et al.* (2004) And Shapiro & Bernhard (1988)

The various internal and external validation parameters used in this study are presented as:

(a) R^2 (the square of the correlation coefficient): Describes the fraction of the total variation attributed to the model. The closer the value of R^2 is to 1.0, the better the regression equation explains the Y variable. R^2 is the most commonly used internal validation indicator and is expressed in equation 1:

$$R^2 = 1 - \frac{\sum(Yobs - Ypred)^2}{\sum(Yobs - Ytraining)^2} \qquad 1$$

Where: Yobs; Ypred; Ytraining are the experimental property, the predicted property and the mean experimental property of the samples in the training set, respectively (Wu *et al.*, 2015).

(b) Adjusted R^2 (R^2_{adj}): R^2 value varies directly with the increase in number of regressors i.e. descriptors, thus, R^2 cannot be a useful measure for the goodness of model fit. Therefore, R^2 is adjusted for the number of explanatory variables in the model. The adjusted R^2 is defined as equation 2:

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

Where: p = number of independent variables in the model and N = sample size (Brandon and Aline, 2015).

(c) Q^2 (Leave one out cross validation coefficient): The LOO cross validated coefficient (Q^2) is given in equation 3;

$$Q^2 = 1 - \frac{\sum(Yp - Y)^2}{\sum(Y - Ym)^2} \qquad 3$$

Where: Yp and Y represent the predicted and observed activity respectively of the training set and Y_m the mean activity value of the training set (Brandon and Aline, 2015).

(d) R^2_{pred}: R^2_{pred} is termed the predictive R^2 of a development model and is an important parameter that is used to test the external predictive ability of a QSAR model. The predicted R^2 value is calculated using equation 4;

$$Pred.R^2 = 1 - \frac{\sum[Y_{(te)} - Y_{Pred\,(te)}]^2}{\sum[Y_{(te)} - Y_{m\,(tr)}]^2} \qquad 4$$

Ypred.(test) and Y(test) indicate predicted and observed activity values, respectively of the test set compounds and Ym(tr) indicates mean activity value of the training set (Ravichandran *et al.*, 2011).

Applicability domain

The applicability domain (AD) of QSAR model was used to verify the prediction reliability, identify the problematic compounds and predict the compounds with an acceptable activity that fall within the domain. The most common methods used for determination of the AD of QSAR models have been described by Grammatica *et al.* (2007) that used the leverage values for each compound. The leverage approach allows the determination of the position of a new chemical in the QSAR model, i.e. whether a new chemical will lie within the structural model domain or outside of it. The leverage approach along with the Williams Plot is used to determine the applicability domain in all QSAR models. To construct the William Plot, the leverage hi for each chemical compound in which QSAR model was used to predict its property was calculated according to the equation 5:

$$h_i = x_i(X^TX)^{-1}x_i^T \qquad 5$$

Where, *x* refers to the descriptor vector of the considered compound and *X* Represents the descriptor matrix derived from the training set descriptor values. The warning leverage (h^*) was determined using equation 6:

$$h^* = \frac{3(p+1)}{N} \qquad 6$$

Where N is the number of training compounds and p is the number of descriptors in the model.

Molecular dynamics simulation

The molecular dynamics simulation was carried out to describe the interaction between the inhibitor molecules and the metallic surface. The adsorption locator module implemented in the Materials studio 8.0 software from Accelrys was used for the simulation. The inhibitor molecules were modeled and optimized using the Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies (COMPASS) force field. COMPASS is a robust and well-developed force field that was derived based on fitting against a wide range of experimental data for organic and inorganic compounds (Wymyslowski *et al.*, 2008). This informs its suitability for treating metal and non-metal containing systems. The whole system was performed at 298K controlled by the Andersen thermostat with fixed number-volume-energy (microcanonical) ensemble, with a time step of 1.0 fs, simulation time of 500ps, using the COMPASS force field. The MD simulation was carried out in a simulation box (24.82 Å ×24.82 Å ×45.27 Å) with periodic boundary conditions. The box includes a Fe slab, an acid solution layer, and an inhibitor molecule. For the iron surface, Fe (110) surface was selected as the studied surface for that Fe (110) was density packed surface and was the most stable (Khaled, 2008). The iron crystal contained ten layers and seven layers near the bottom were frozen. The density of the acid solution layer was set at 1.0 g/cm^{-3}. The adsorption energy in solution was calculated using equation 7 (Zhao *et al.*, 2014)

$$E_{adsorption} = E_{Total} - \left(E_{Fe_{surface}+solution} + E_{Inhibitor+solution}\right) + E_{Solution} \qquad 7$$

Results and Discussion

Quantitative structure–activity relationship (QSAR)

Pearson's correlation matrix was used on the training set to select the significant descriptors and it was found that among all the computed descriptors Electrophilicity index (ω), SpMin8_Bhe, SpMin2_Bhp, AVP-2 and SpMAD_D construct the best model and show strong correlation (Table 3).

Table 3: Correlation matrix for the studied variable

	%IE	ω	SpMin8_Bhe	SpMin2_Bhp	AVP-2	SpMAD_D
%IE	1					
ω	0.327205	1				
SpMin8_Bhe	-0.13403	-0.20681	1			
SpMin2_Bhp	0.438851	-0.16755	-0.4891	1		
AVP-2	0.553372	0.176477	-0.67728	0.215628	1	
SpMAD_D	0.193828	0.144248	-0.75398	0.74464	0.407443	1

The selected descriptors were subjected to regression analysis with the experimentally determined inhibition efficiencies as the dependent variable and the selected descriptors as the independent variables using Genetic function approximation (GFA) method in Material studio software, a new GFA equation (Equation 8) was developed on the basis of the training set:

$$\%\mathbf{IE} = 1.422315151 * \omega + 23.122686929 * \text{SpMin8\_Bhe} + 89.168074872 * \text{SpMin2\_Bhp} + 377.048667573 * \text{AVP} - 2 - 6.890458835 * \text{SpMAD\_D} - 153.771948048 \qquad 8$$

$N_{train} = 18, R^2_{train} = 0.98, R^2_{adjusted} = 0.98, Q^2_{LOO} = 0.97, N_{test} = 7\ R^2_{test} = 0.87$

In the further study, the constructed model from the training set was used to evaluate the predictive ability of the produced model by predicting the %IE values in the prediction set. The results are given in Table 4. The predicted inhibition efficiency were plotted against the experimental inhibition efficiency for the training and test sets in Fig. 1.

Table 4: Experimental and predicted %IE with Residuals

S/N	Exp. %IE	Pred. %IE	Residuals	S/N	Exp. %IE	Pred. %IE	Residuals
1a	50	50.27	-0.27	14a	34	36.39	-2.39
2b	59	58.07	0.93	15a	43	42.96	0.04
3a	63	62.1	0.9	16a	77.4	76.54	0.86
4a	47	46.55	0.45	17b	75.1	75.23	-0.13
5b	80	81.3	-1.3	18a	41	38.5	2.5
6a	52	55.49	-3.49	19a	71	70.84	0.16
7a	39	38.85	0.15	20a	63.62	62.77	0.85
8b	73	74.48	-1.48	21a	71.79	72.66	-0.87
9a	53	52.16	0.84	22b	63.24	64.64	-1.4
10a	51	51.87	-0.87	23a	66.83	66.53	0.3
11a	87	84.61	2.39	24a	49.88	48.05	1.83
12a	75	78.28	-3.28	25b	60.09	58.34	1.75
13b	67	66.5	0.5				

a = training set, b = test set

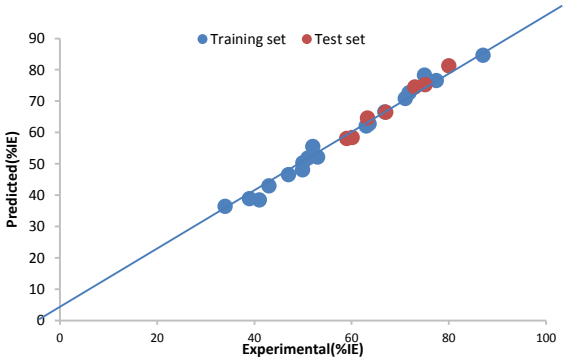


Fig. 1: Plot of predicted versus actual inhibition efficiency (%IE) value for the model

The result of the GFA QSAR model is in conformity with the standard shown in Table 3 as seen in equation-3. The closeness of coefficient of determination (R^2) to its absolute value of 1.0 is an indication that the model explained a very high percentage of the response variable (descriptor) variation, high enough for a robust QSAR model. The high adjusted R^2 (R^2_{adj}) value and its closeness in value to the value of R^2 implies that the model has excellent explanatory power to the descriptors in it. More so, the high and closeness of Q^2 value to R^2 revealed that the model was not over fitted. The high R^2_{pred} is an indication that the model is capable of providing valid predictions for new molecules that fall within its applicability domain.

Furthermore, the equation contains five descriptors and each descriptor has a positive or negative coefficient attached to it. These coefficients along with the value of descriptor have a significant role in deciding the overall inhibition efficiency of the inhibitor molecules. Examination of equation-13 shows that coefficients of each descriptor played an important role in deriving the inhibition efficiency. From the point of view of inhibition of the molecules in terms of %IE values, the weight of positive coefficient is very significant because it contributes towards the increased value of %IE. So the descriptors with high weight positive coefficients are most important followed by descriptors with the low weight negative coefficient and lastly the descriptors with high weight negative coefficients. On the basis of values of the coefficients on the model, the associated descriptors are arranged in a sequence pertaining to their contribution towards overall inhibition efficiency of the inhibitors, in the following increasing order of inhibition efficiency of inhibitors towards corrosion of steel.

$$\mathbf{AVP-2 > SpMin2\_Bhp > SpMin8\_Bhe > \omega > SpMAD\_D}$$

Table 5: Specification of entered descriptors in genetic algorithm of the developed model

S/N	Symbol	Description	Class/Group
1	Ω	Electrophilicity index	3D/Electronic Descriptor
2	SpMin8_Bhe	Smallest absolute eigenvalue of Burden modified matrix - n 8 / weighted by relative Sanderson electronegativities	2D/Burden Modified Eigen values Descriptor
3	SpMin2_Bhp	Smallest absolute eigenvalue of Burden modified matrix - n 2 / weighted by relative polarizabilities	2D/Burden Modified Eigenvalues Descriptor
4	AVP-2	Average valence path, order 2	2D/PaDELChiPath Descriptor
5	SpMAD_D	Spectral mean absolute deviation from topological distance matrix	2D/TopologicalDistanceMatrixDescriptor

The leverages for every compound in the dataset were plotted against their standardized residuals, leading to discovery of outliers and influential chemicals in the models. Figure 2 shows the Williams plot of standardized residuals against calculated leverages for both the training and test set.

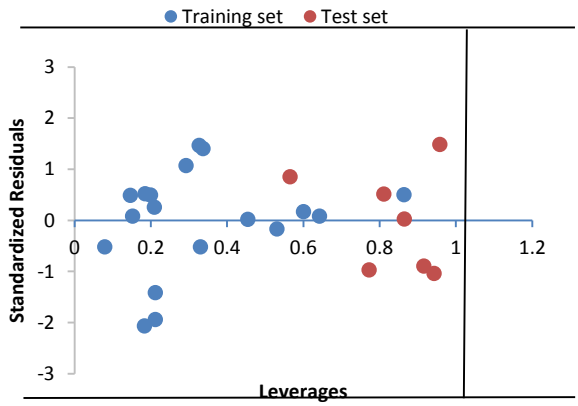


Fig. 3: The Williams plot, the plot of standardized Residuals versus the leverage value for all the data set

The Williams plot for the QSAR is illustrated in Fig. 3. The warning leverage (h^*), was found to be1.0 (N = 18 and p=5) for the developed QSAR model (Roy, 2015). The chemicals that had a standardized residual more than the standard deviation units were considered to be outliers while chemicals with a leverage value higher than h* were considered to be influential or high leverage chemicals. Based on the leverages none of the compounds were found to be outside of the defined AD (Fig. 2) of the QSAR model. In addition, no outlier compounds with standardized residuals > ±3d for the dataset were identified.

Molecular dynamic simulation study

Corrosion inhibition by organic compounds is mainly due to their ability to adsorb onto a metal surface to form a protective film. The adsorption of organic inhibitors at the metal/solution interface takes place through the replacement of water molecules by organic inhibitor molecules according to the following process;

$$\boldsymbol{Org_{(sol)} + xH_2O_{(ads)} \leftrightarrow Org_{(ads)} + xH_2O_{(sol)}}$$

Where Org(sol) and Org(ads) are organic molecules in the solution and adsorbed on the metal surface, respectively, x is the number of water molecules replaced by the organic molecules.

It is essential to know the mode of adsorption and the adsorption energy that can give valuable information on the interaction of inhibitor and metal surface. The adsorption of the studied inhibitor molecules on steel surface was simulated by modeling the interactions between the inhibitor molecules and Fe(110) crystal surface. The equilibrium configurations of the simulated system for inhibitor-11 are shown in Fig. 3; all different systems for the inhibitors were studied similarly. The

adsorption energies in kcal/mol of the systems are listed in Table 6.

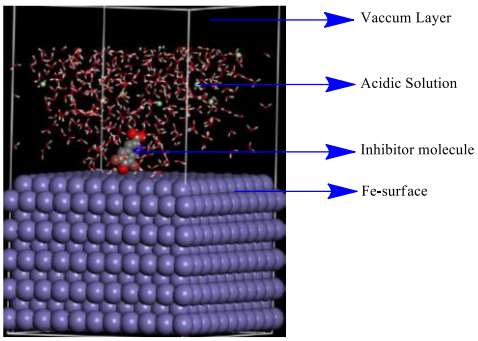


Fig. 3: Molecular dynamic system of Fe, inhibitor molecule and acidic layer

Table 6: Adsorption energies and inhibition efficiencies (%IE) of Amino acids derivatives

S/N	Adsorption Energy (kcal/mol)	Exp. %IE	S/N	Adsorption Energy (kcal/mol)	Exp. %IE
1	-42.82	50	14	-181.67	34
2	-43.16	59	15	-197.00	43
3	-256.72	63	16	-280.41	77.4
4	-214.82	47	17	-259.21	75.1
5	-399.12	80	18	-128.24	41
6	-229.23	52	19	-227.37	71
7	-244.21	39	20	-285.05	63.62
8	-297.5	73	21	-126.58	71.79
9	-234.58	53	22	-467.95	63.24
10	-227.71	51	23	-253.35	66.83
11	-651.7	87	24	-380.73	49.88
12	-290.94	75	25	-346.24	60.09
13	-257.21	67			

It is apparent from the molecular structures of examined inhibitors; these molecules contain various lone pair electrons on N, and O atoms as well as π-aromatic frameworks. Therefore, giving the lone pair electrons on heteroatoms to the empty d orbitals of iron. It can be noticed from Fig. 3; the inhibitor is adsorbed nearly parallel to the Fe (110) surface with the assistance of the donation of the lone pair of electrons of the N in the amino functionality and is also adsorbed through the carboxylic functional group.

It was accounted for in many investigations that the primary mechanism of the interaction between corrosion inhibitors and metallic iron is by adsorption. In this way, the adsorption energies calculated via molecular dynamics simulations approach to give us an immediate understanding to compare the anticorrosive performances of inhibitor molecules. It is seen from the Table 1 that the calculated adsorption energies of the examined inhibitors on the iron surface are generally negative values and most of them are greater than 100 kcal/mol suggesting chemisorptive interaction (Akalezi *et al*., 2012). Only two of the inhibitors 1 and 2 were found to be <100 kcal mol−1 which fall within the range of physisorption interactions (Akalezi *et al*., 2012). These negative values denote that the adsorption happening amongst metal and inhibitors could happen spontaneously. The largest negative adsorption energy represents that the system is most stable and adsorption is very strong. This implies that the corrosion inhibitor binds to Fe (110) surface more easily and firmly (Obot *et al*., 2016). The values of the adsorption energies obtained via molecular dynamic simulation approach are in agreement with experimentally observed corrosion inhibition efficiencies of the studied inhibitors.

Conclusion

The inhibitive performance of twenty-five amino acids and related compounds on steel corrosion in 1 M HCl were investigated using QSAR analysis and Molecular dynamic simulations. The following conclusions can be drawn from the results:

1. The developed GFA model shows good statistical significance. The prediction of corrosion efficiencies nicely matched the experimental measurements.
2. The studied inhibitor molecules inhibit steel corrosion in 1 M HCl by adsorbing on the steel surface and form the protective film. Their adsorption occurs mainly via Chemisorption.
3. Molecular dynamics simulations revealed that the studied molecules have strong interactions with Fe surface.

Acknowledgement

We acknowledge the Department of Chemistry, Ahmadu Bello University, Zaria, Nigeria for providing the softwares used and Abdulfatai Usman for his positive suggestions which anonymously helped at improving the quality of the manuscript. The authors received no direct funding for this research.

References

Abdulfatai U, Uzairu A & Uba S 2017. Quantitative structure-activity relationship and molecular docking studies of a series of quinazolinonyl analogues as inhibitors of gamma amino butyric acid aminotransferase. *J. Adv. Res*., 8: 33–43.

Akalezi OC, Enenebaku CK & Oguzie EE 2012. Application of aqueous extracts of coffee senna for control of mild steel corrosion in acidic environments. *Int. J. Ind. Chem.*, 3: 13.

Al Hashem A., 2011. "Corrosion in the Gulf Cooperation Council (GCC) states: Statistics and Figures," in Proceedings of the Corrosion UAE, Abu Dhabi, UAE.

Amin MA, Khaled KF, Mohsen Q & Arida HA 2010. A study of the inhibition of iron corrosion in HCl solutions by some amino acids. *Corros. Sci*., 52: 1684– 1695.

Anonymous 2013. Wavefunction. Inc., Spartan'14, version 1.1.2. Irvine, California, USA.

Arslan T, Kandemirli F, Ebenso EE, Love I & Alemu H 2009. Quantum chemical studies on the corrosion inhibition of some sulphonamides on mild steel in acidic medium. *Corros. Sci.*, 51: 35–47, doi:10.1016/j.corsci.2008.10.016

Babić-Samardžija K, Lupu C, Hackerman N, Barron AR & Luttge A 2003. *Langmuir*, 21: 12187 – 12196.

Becke AD 1993. Density-functional thermochemistry III. The role of exact exchange. *J. Chem. Phys*., 98(7): 5648–52.

Beese P 2013. Monitoring of anaerobic microbially influenced corrosion via electrochemical frequency modulation. *Electrochim. Acta*, 105: 239–247.

Brandon K & Aline O 2015. Comprehensive R archive network (CRAN). Retrieved 03, July 2015 from http:// CRAN.R-project.org.

Cost of Corrosion to exceed $1 Trillion in the United States in 2012-G2MT Labs-The Future of Materials Conditions Assessment," http://www.g2mtlabs.com/2011/06/ nace-cost-of corrosion-study-update.

Deng Q, Shi HW, Ding NN, Chen BQ, He XP, Liu G, Tang Y, Long YT & Chen GR 2012. Novel triazolyl bis-amino

acid derivatives readily synthesized via click chemistry as potential corrosion inhibitors for mild steel in HCl. *Corros. Sci*., 57: 220–227.

Eddy NO & Mamza PAP 2009. Inhibitive and adsorption properties of ethanol extract of seeds and leaves of azadirachta indica on the corrosion of mild steel in H_2SO_4, *Electrochim. Acta*, 27: 443–456.

Fu J, Li S, Wang Y, Cao L & Lu L 2010. Computational and electrochemical studies of some amino acid compounds as corrosion inhibitors for mild steel in hydrochloric acid solution. *J. Mater. Sci*., 45: 6255–6265.

Ghafourian T & Cronin MTD 2005. The impact of variable selection on the modelling of regulatory perspective. *Curr. Comput. Aided Drug Des.*, 1: 195-205.

Gramatica P, Giani E & Papa E 2007. Statistical external validation and consensus modeling: A QSPR case study for K oc prediction. *J. Mol. Graph Model*, 25(6): 755–766.

Hassan HH, Abdelghani E & Amin MA 2007. Inhibition of mild steel corrosion in hydrochloric acid solution by triazole derivatives: Part I. Polarization and EIS studies. *Electrochim. Acta*, 52: 6359–6366, doi:10.1016/j.electacta.2007.04.046.

He L & Jurs PC 2005. Assessing the reliability of a QSAR model's predictions. *J. Mol. Graphics.*

Hluchan V, Wheeler BL & Hackerman N 1988. Amino acids as corrosion inhibitors in hydrochloric acid solutions, *Mater. Corros*., 39: 512-517.

Hu P & Dai W 2012. Computer simulation and density functional theory study on relationship between structure of amino acid and inhibition performance. *Mater. Res. Innov*., 16: 67–72.

Jaiswal M, Khadikar PV, Scozzafava A & Supuran CT 2004. Carbonic anhydrase inhibitors: The first QSAR study on inhibition of tumor-associated isoenzyme IX with aromatic and heterocyclic sulfonamides. *Bioorg. Med. Chem. Lett.*, 14: 3283–3290.

Kennard RW & Stone LA 1969. Computer aided design of experiments. *Technometrics*, 11(1): 137–48.

Khaled KF 2012. Alanine as corrosion inhibitor for iron in acid medium: A molecular level study. *Int. J. Electrochem. Sci.*, 7: 12706–12719.

Khaled KF 2008. Molecular simulation, quantum chemical calculations and electrochemical studies for inhibition of mild steel by triazoles, *Electrochim. Acta*, 53: 3484-3492.

Khaled KF 2010. Corrosion control of copper in nitric acid solutions using some amino acids a combined experimental and theoretical study, *Corros. Sci*. 52 3225–3234.

Khaled KF & Hackerman N 2003. Investigation of the inhibitive effect of orthosubstituted anilines on corrosion of iron in 1 M HCl solutions, *Electrochim. Acta*, 48: 2715– 2723.

Kraka E & Cremer D 2000. Computer design of anticancer drugs. *J. Am. Chem. Soc*., 122 8245–8264.

Larif M, Chtita S, Adad A, Hmamouchi Bouachrine R, Lakhlifi M 2013, Predicting biological activity of anticancer molecules 3- ary l-4-hydroxyquinolin-2-(1H)-one by DFT-QSAR models. *Int. J. Cancer*., 3(12).

Lyon S 2004. Materials science. A natural solution to corrosion?, *Nature*, 29 406–407.

Materials Studio 6.1 Manual (Accelrys, Inc., San Diego, CA, 2007).

Murulana LC, Singh AK, Shukla SK, Kabanda MM & Ebenso EE 2012. Experimental and quantum chemical studies of some bis (trifluoromethyl-sulfonyl) imide imidazolium-based ionic liquids as corrosion inhibitors for mild steel in hydrochloric acid solution. *Ind. Eng. Chem. Res.*, 51: 13282–13299, doi:10.1021/ie300977d

Muster TH, Hughes AE, Furman SA, Harvey T, Sherman N, Hardin S, Corrigan P, Lau D, Scholes FH, White PA, Glenn M, Mardel J, Garcia SJ & Mol JMC 2009. A rapid screening multi-electrode method for the evaluation of corrosion inhibitors. *Electrochim. Acta*, 54: 3402–3411.

Nan L, Xu D, Gu T, Song X & Yang K 2015. Microbiological influenced corrosion resistance characteristics of a 304L-Cu stainless steel against. *Escherichia coli*. *Mater. Sci. Eng. C.*, 48: 228–234, doi:10.1016/j.msec.2014.12.004.

Nwankwo HU 2016. Synthesis, characterization, antimicrobial studies and corrosion inhibition potential of 1, 8-dimethyl-1, 3, 6, 8, 10, 13-hexaazacyclotetradecane: experimental and quantum chemical studies. *Mater.*, 9: 1–19, doi:10.3390/ma9020107.

Obot IB, Kaya SC & Tuzun B 2016. Density Functional Theory (DFT) modeling and Monte Carlo simulation assessment of inhibition performance of some carbohydrazide Schiff bases for steel corrosion. Physica E: Low-dimensional Systems and Nanostructures, 80: 82–90.

Praveen BM & Venkatesha TV 2009. Metol as corrosion inhibitor for steel. *Int. J. Electrochem. Sci.*, 4: 267–275, doi:10.1007/s11030-015-9640-0.

Ravichandran V, Harish R, Abhishek J, Shalini S, Christapher PV & Ram KA 2011. *Int. J. Drug Design & Discovery*, 2: 511-519.

Roy K, Kar S & Ambure P 2015. On a simple approach for determining applicability domain of QSAR models. *Chemometr. Intell. Lab. Syst*., 145: 22–29.

Shapiro S & Bernhard G 1998. Inhibition of oral bacteria by phenolic compounds; QSAR analysis using molecular connectivity. *Molecular Informatics*, 17(4): 327-337.

Singh P, Ebenso EE, Olasunkanmi LO, Obot IB & Quraishi MA 2016. Electrochemical, theoretical, and surface morphological studies of corrosion inhibition effect of green naphthyridine derivatives on mild steel in hydrochloric acid. *J. Phys. Chem. C.*, 120: 3408–3419, doi:10.1021/acs.jpcc.5b11901.

Szklarska-Smialowska Z 2005. Pitting and Crevice Corrosion, NACE International.

Tong W, Hong H, Xie Q, Shi L, Fang H & Perkins R 2005. Assessing QSAR limitations-A oestrogenicity. *SAR QSAR Environ. Res.*, 16: 171-190.

Tropsha A, Gramatica P & Gombar VK 2003. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb. Sci*., 22: 69-77.

Venzlaff H., 2013. Accelerated cathodic reaction in microbial corrosion of iron due to direct electron uptake by sulfate-reducing bacteria. *Corros. Sci.*, 66: 88–96, doi:10.1016/j.corsci.2012.09.006.

Wu W, Zhang C, Lin W, Chen Q, Guo X & Qian Y 2015. Quantitative structure-property relationship (QSPR) modeling of drug-loaded polymeric micelles via genetic function approximation. *PLoSONE*, 10(3): 35-43.

Wymyslowski A, Iwamoto N, Yuen M & Fan H 2008. Molecular Modeling and Multiscaling Issues for Electronic Material Applications. Springer, Spring Street, New York.

Yadav DK, Quraishi MA & Maiti B 2012. Inhibition effect of some benzylidenes on mild steel in 1 M HCl: an experimental and theoretical correlation. *Corros. Sci.*, 55: 254– 266, doi:10.1016/j.corsci.2011.10.030.

Yap CW 2011. PaDEL-descriptor: Open source software to calculate molecular descriptors and fingerprints. *J Comp. Chem.*, 32(7): 1466–1474.

Yuan S, Liang B, Zhao Y & Pehkonen SO 2013. Surface chemistry and corrosion behaviour of 304 stainless steel in simulated seawater containing inorganic sulphide and sulphate reducing bacteria. *Corros. Sci.*, 74: 353–366.

Yurt A, Bereket G & Ogretir C 2005. Quantum chemical studies on inhibition effect of amino acids and hydroxy carboxylic acids on pitting corrosion of aluminium alloy 7075 in NaCl solution. *J. Mol. Struct.: THEOCHEM*, 725: 215–221.

Zhang J, Qiao G, Hu S, Yan Y, Ren Z & Yu L 2011. Theoretical evaluation of corrosion inhibition performance of imidazoline compounds with different hydrophilic groups. *Corros. Sci*., 53: 147–152.

Zhang SG, Lei W & Xia MZ 2005. Wang FY QSAR study on N-containing corrosion inhibitors: quantum chemical approach assisted by topological index. *THEOCHEM*, 732: 173–182.

Zhao H, Zhang X, Ji L, Hu H & Li Q 2014. Quantitative structure-activity relationship model for amino acids as corrosion inhibitors based on the support vector machine and molecular design. *Corr. Sci*. doi http://dx.doi.org/10.1016/j.corsci.2014.02.023