Molecular Design and Dynamic Simulations of Some Novel Antioxidant Lubricant Additives

Abdulfatai Usman 1,*, Adamu Uzairu, Sani Uba, Gideon Adamu Shallangwa

Department of Chemistry, Ahmadu Bello University, Zaria, Nigeria

Abstract

A quantitative structure-property relationship (QSPR) in-silico study was performed to develop a mathematical model that correlate 2D and 3D descriptors of 37 antioxidant lubricant additives (compounds) with their properties. A molecular dynamics simulation study was also carried out to access the binding strength of these additives on the diamond like carbon (DLC) and steel crystal surfaces. Five novel antioxidant lubricant additives were designed from the information derived from the QSPR mathematical model’s high coefficient molecular descriptors. All the novel lubricant additive’s antioxidant properties were found to be better than our previous study, with the lubricant additive (Z)-3-(4-(5-amino-1-phenyl-1H-pyrazol-3-yl)-3,5-dimethylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one found to possessed excellent antioxidant properties of 0.850281 total acid values (T.A.V 0.1g/L) than its co-additives. Moreover, from our dynamic simulation study, all the designed additives were found to dynamically bind to steel crystal surface excellently than the DLC crystal surface. The molecular dynamics simulations results were found to be better than the one reported by our previous study. This investigation will help to synthesize novel and excellent anti-oxidant lubricant additives that will hinder the base oil from undergoing complete circle of oxidation and they meet environmental requirements as these novel additives do not contain Zinc and Phosphorus which often rendered exhaust pipe catalytic converter inactive thereby increasing environmental pollution.

Keywords: QSPR; Novel lubricant additive; DLC: DFT: Molecular dynamic simulations.

Introduction

The growing concern among scientists and other researchers in the lubricating oil industry to explore other alternatives to replace the long time use of zinc dialkyl dithiophosphates (ZDDPs) has continue to generate momentum. Studies shows that higher concentrations (≥1800 ppm) of ZDDP (a multifunctional lubricant additive) was reported to be responsible for less corrosion inhibitors efficiency [1], while less concentration of this additive have been reported to be responsible for engine failure [2]. Moreover, less efficiency of catalytic converters and the violation of the environmental protection regulations have been traced to the use of ZDDP as base oil additive [3].

At an elevated dynamic temperature in the alloy coated internal combustion engines, the metals such as iron, zinc, copper and their alloys were reported to act as catalyst for auto-oxidation of lubricating oil whenever the oil is in contact with the atmospheric oxygen. Also, the auto-

1 Corresponding author: uabdulfatai@abu.edu.ng
oxidized lube oil often resulted in unexpected rise in viscosity, sludge and its volume generally varnishes [4-8]. Antioxidant additives are lubricating oil additives designed to hinder the base oil from undergoing complete circle of oxidation. These additives also allow automobile lubricating oil to operate at unexpected higher dynamic temperatures. Base oil antioxidants such as Aromatic amines, hindered phenols, alkyl and aromatic sulfides has been identified as a reliable free radical tappers [4-8]. Moreover, nitrogen or sulfur heterocyclic lubricant additives have been reported also to be effective antioxidant lubricant additives and they meet environmental requirements as they do not contain Zinc and Phosphorus which often rendered exhaust pipe catalytic converter inactive thereby increasing environmental pollution [9-11]. Therefore, there is urgent need to sustain the momentum of pursuing very active, reliable and environmental friendly antioxidant lubricant additive.

In recent times, the rapid improvement of computer speed and programs has lead to the development of in-silico studies. Computational in-silico simulations act as a boundary between microscopic level, time scales and the macroscopic world of real laboratory experiment [11]. Quantitative Structure-Properties Relationship (QSPR) which is a theoretical in-silico study was developed to effectively predict and design active leading compounds. QSPR does that by relating chemical properties of a group of compounds with its molecular structures. This in-silico study has recorded tremendous success in predicting and designing lead compounds with the improve activities and properties [12, 13]. Also, molecular dynamic (MD) simulation is another in-silico method that could predict accurately the dynamic interaction energies in close contact between any potent compounds and crystal surface [13]. In this research, efforts were geared towards designing new active antioxidant lubricant additives and to investigate their dynamic binding energies on Diamond Like-Carbon and steel crystal surfaces through QSPR and MD simulation approach.

Materials and Methods

Data set and Molecular Descriptors Generation

The set of data used for this research was made up of 37 (Table 1) heterocyclic derived antioxidant lube oil additives obtained from literatures [14-16] with their antioxidant expressed in term of total acid values (0.1g/L), p(T.A.V 0.1g/L). All the selected heterocyclic derived lubricant additives 2D structure were drawn with ChemDraw ultra V12.0 module and transferred to Spartan’14 version 1.1.2 [17] where pre-energy minimization were carried out by molecular mechanics (MM) and followed by complete geometry optimization by B3LYP/6-311++ G** using Spartan’14 version 1.1.2 software. From the computational mathematical point of view, this density functional theory (DFT/B3LYP/6-311++ G**) method can provide reliable results compared to other in-silico methods [12, 13]. The geometrically optimized additives were saved in sdf format and transferred to Dragon 6.0 software toolkits and padel software [18] where about 4234 generated descriptors were reduced to 3500 after removing all the irrelevant descriptors. Moreover, Using random selection method, the lubricant additive dataset was divided into training (24) and test sets (13).The 24 training set were used for the building and development of the QSPR models while the leftover 13 test sets were used to evaluate the predictive power of the built models [19].

QSPR model building and development
Using 24 data of the training sets, the genetic function algorithm (GFA) method in material studio chemical model building software version 8.0 was employed to perform the correlation analysis between the anti-oxidant properties of the lubricant additives and the calculated molecular descriptors. The GFA has the following advantages over other methods; substituting regression calculations with the GFA allows the construction of models competitive with those produced by standard techniques and makes available additional information not provided by other techniques. GFA as a result provides multiple reliable QSPR/QSAR models, and the population of the model is created by evolving random initial model using a genetic algorithm [12, 13].

Assurance of the QSPR models and descriptors variance inflation factor (VIF) Determination

The reliability, robustness, and predictive ability of the constructed QSPR models were accessed by internal and external validation parameters and compared with the standard recommended values for a generally acceptable QSPR model [20] in Table 2. The internal validation parameters such as the square correlation coefficient ($R^2$) (2.1), the adjusted square correlation coefficient ($R^2_{adj}$) (2.2) which increases the number of descriptors in the model, and the Leave one out cross validation coefficient ($Q^2_{cv}$) (2.3) and all these statistical values must be less than 0.5 but greater than 1.0 for a reliable, predictive and guaranteed QSPR model [20].

$$R^2 = 1 - \frac{\sum(Y_{observed} - Y_{predicted})^2}{\sum(Y_{observed} - Y_{training})^2}$$  

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

$$Q^2 = 1 - \frac{\sum(y_p - Y)^2}{\sum(y - Y_m)^2}$$

where $p$ = number of independent variables in the model, $N$ = sample size [21] where $Y_p$ and $Y$ represent the predicted and observed activity respectively of the training set and $Y_m$ the mean activity value of the training set [21]. The models were externally validated by testing the previously excluded compounds which form the test set. The value of $R^2_{pred}$ which gives an indication of the predictive power of a model was calculated using equation (3.4) [20]

$$R^2_{pred} = 1 - \frac{\sum(y_{pred}(test) - Y_{observed}(test))^2}{\sum(Y_{observed}(test) - Y_{mean}(training))^2}$$  

$Y_{pred}$(test) and $Y_{observed}$(test) indicate predicted and observed activity values respectively of the test set compounds and $Y_{mean}$ (training) indicates mean activity value of the training. The best regression model was generated by considering all the possible combination of descriptors. Variance inflation factor (VIF) [22] was used to identifying the multi-collinearity among variables. The VIF for the regression coefficient is expressed as:

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

where $R^2$ is the correlation coefficient of the multiple regression between the variables within the model. If VIF equals to 1, then no inter-correlation exists for each variable; if VIF falls into the range of 1–5, the related model is acceptable; and if VIF is larger than 10, the related model is unacceptable.
**QSPR Model’s Applicability domain (AD)**

The AD of a QSPR model is defined as a theoretical region in chemical space, defined by the model descriptors and modeled response, and thus by the nature of the chemicals in the training set, as represented in each model by specific molecular descriptors. Thus only the predictions for chemicals falling within this domain can be considered reliable and not model extrapolations [23]. Without the restriction placed by the applicability domain, QSPR models can predict the activity of any compound even if such a compound is structurally different from those included in the training set. This would lead to unreasonable extrapolation of the model in chemistry space and therefore heighten the chances of inaccurate predictions. Thus for a QSPR model to give reliable outcome, its applicability domain must be defined. The leverage is defined as a compound’s distance from the centroid of $X$. Mathematically, the leverage ($h_{ii}$) of a given compound in the multidimensional descriptor space, can be calculated as (Eq. 2.6)

$$h_{ii} = x_i^T(X^TX)^{-1}x_i$$

where $x_i$ the descriptor row matrix of the compound under consideration and $X$ is the multidimensional matrix carrying the structural information (calculated molecular descriptors) for each training set compound. The model predictions should be referred as unreliable for those compounds for which $h_{ii}$ diagonal elements are greater than the cut-off leverage value ($h^*$). These compounds are located far from the structural centroid of the model, and therefore could be referred as structurally-influential outliers [24]. The cut-off leverage value ($h^*$) is usually defined by Equation (2.7)

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

where $p$ is the total number of descriptors used for developing of the QSPR model, while $n$ is the total number of the training set compounds. Moreover, the compounds for which the calculated standardized residual values are greater than standard deviation units could be considered as response outliers [24].

**Virtual Screening Method; Template Based**

QSPR is an in-silico computational screening technique which design or predict the properties of un-synthesis chemical compounds with the desired properties by relating molecular properties of compound with its molecular structure. This virtual screening method could determine whether the generated model could predict as much structures as possible as those used for the training and validation sets and to identify which structural modifications could be permitted using the domain of applicability to design better novel chemical compound of desirable property. In this current research, a template based chemical compound design method was used to design novel antioxidant lubricant additives with improved properties. In the pool of experimental lubricant antioxidant additives, a template chemical compound with a good leverage as well as excellent experimental properties was chosen as a useful scaffold toward the base for structural modification with the help of interpretations of molecular descriptors generated from the developed QSPR best model.

**Molecular dynamics simulation studies**

Chemical Compounds with anti-oxidant properties were reported to prevent oxidation by acting as radical trapper [14-16]. Therefore, molecular dynamic simulation study was carried out in order to calculate which of the designed compounds have good dynamic binding energy against
diamond-like carbon (DLC) surface. The dynamic module implemented in the Materials studio 8.0 software from Accelrys was used for the simulation calculations.

Hydrogen-containing DLC (a-C: H) crystals and of the 3D lubricant additives Preparation

The 3D structure of hydrogen-containing DLC crystal surface which was reported in many literatures to be better allotropy of carbon in term of wearing resistance ability and antioxidant in the sliding interface [25-30] was constructed from the carbon (C) model using Materials studio 8.0 simulation software. This was done by cleaving the carbon surface at point 1.1.0 (h k l). Top (1.006 Å) and Thickness (24.121 Å) into the crystal unit. Moreover, the repeated units (super cell) of the carbon crystal units were formed at U (9) and V (8) while hydrogenation of the super cell, vacuum slab and geometric optimization were performed. The 2D structures of the lubricant antioxidant additives were drawn with Chemdraw software and were then converted to 3D structures by Materials studio 8.0 simulations Software optimized and saved by the same software.

Molecular dynamics simulation energy calculations

The Condensed-phase Optimized Molecular Potentials for Atomistic Simulation Studies II (COMPASS II) force field method in Materials studio 8.0 Software was selected and used for the molecular dynamic simulation. COMPASS II is a robust and well-developed (than COMPASS) force field that was derived based on fitting against a wide range of experimental data for organic and inorganic compounds [31]. Using the Materials studio 8.0 Software, The dynamic simulation calculations were carried out after introducing the optimized lubricant additive compound into the simulation vacuum slab of geometrically optimized hydrogen-containing DLC crystal (24.82 Å ×24.82 Å ×45.27 Å) surface at 350.15 K and over a range of inter-surface separations. The binding energy was calculated by using equation 2.8 [32].

\[
\text{Molecular Dynamic Binding Energy} = E_{\text{total}} - (E_{\text{Lubricant Additive}} + E_{\text{DLC Surface}}) \tag{2.8}
\]

3. Results and Discussion

Some QSPR models were developed from the investigation of the structure– properties relationship of 37 synthesized antioxidant lubricant additives. In a pool of some developed QSPR models, best model with the best squared correlation coefficient (R^2), adjusted squared correlation coefficient (R^2adj), Leave one out cross validation coefficient (Q^2) and the external validation (R^2ext) of values 0.999477, 0.998823, 0.995547and 0.645 respectively were identified and selected as the model with the most predictive power [21]. This predictive model p(TAV.0.1g/L), that was expressed in term of antioxidant total acid values (TAV 0.1g/L) was further used to predict many properties of some designed antioxidant lubricant additives. Some QSPR statistical analyses were also carried out to further validate the identified model; the plot of experimental properties versus residual activities (Figure 1) indicated that there was no systemic error in the developed model as the propagation on both sides of zero in the graph was observed [32]. The plotted graph in figure 2 is in great concurrence with the validation parameters [21] in table 2, and hence the developed model did not demonstrate any relative forms of error, since the R^2 from the graph was found to be 0.7244.

QSPR Model
The QSPR model revealed that decrease in molecular descriptors such as CrippenLogP (Lipophilicity parameter/character), nAtomLAC (Longest Aliphatic Chain Descriptor), MDEN-33 (Molecular distance edge between all primary carbons) and increase in nRotBt (Rotatable Bonds Count Descriptor) and WT.eeneg (Non-directional WHIM, weighted by Mulliken atomic electronegativities) values will increase the antioxidant properties of the lubricant additive [18, 34].

**Applicability domain’s assessment of the developed model and molecular designs**

In order to be sure of the additives to use as design template for further structural modifications since the robustness of the QSPR model alone cannot be enough to predict accurately the properties of all the compounds within the chemical space [35]. In Figure 3, the William’s plot which is the applicability domain’s leverage plot was calculated and found to contain leverage danger, h² of 0.75. This plot revealed that about five compounds (14, 20, 27, 17 and 21) that go beyond the leverage danger were termed as influential antioxidant lubricant additives [24]. These influential antioxidant lubricant additives are compounds that goes beyond the danger leverage lines/chemical space and therefore, such compound cannot be used as design template [24]. The compound with serial number 15 (Table 1) has zero (0.00) residual value and was statistically found to be within the chemical space of the William’s plot with an excellent leverage value (Fig. 3) was chosen as the best template after which further modification were made.

From the developed QSPR model, the molecular descriptor with the highest/ contributor to the development of the model were WT.eeneg (Non-directional WHIM, weighted by Mulliken atomic electronegativities). This descriptor was chosen as the best contributor due to the highest coefficient value among the co-descriptors. Therefore, interpretation of this descriptor revealed that addition of more weighted electronegative compound such as nitrogen atoms from NH₂ and other substituent like phenyl and methyl to the additive template (Fig. 4) revealed that better novel antioxidant lubricant additives were designed (Table 3). Moreover, all the five designed antioxidant lubricant additives along with their properties (total acid values) were found to be excellently better than the experimental antioxidant lubricant additives (Table 1).

**Dynamic simulations studies**

The five designed antioxidant lubricant additives such as; (Z)-3-(4-(5-amino-1-phenyl-1H-pyrazol-3-yl)-3,5-dimethylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one, (Z)-3-(4-(3-amino-4-methyl-1-phenyl-1H-pyrazol-5-yl)phenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one, (Z)-3-(4-(3-amino-4-methyl-1-phenyl-1H-pyrazol-5-yl)-3-methylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one, (Z)-3-(4-(5-amino-4-methyl-2-phenyl-2,5-dihydroisoxazol-3-yl)-3-methylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one and (Z)-3-(4-(5-amino-2-phenyl-2,5-dihydroisoxazol-3-yl)-3-methylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one were found to be dynamically bound to the DLC (-12285.2, -12281.9, -12278.8, -12286.8 and -12286.4kcal/mol) and steel (-171817, 172835, -171767, -171835 and -172870 kcal/mol) surfaces respectively (Table 3). Moreover, from table 3, it can be observed that the designed additives were found to dynamically bind to steel crystal
surface than the DLC crystal surface. This study also revealed that (Z)-3-(4-(5-amino-2-phenyl-2,5-dihydroisoxazol-3-yl)-3-methylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one was excellently bound to steel surface (-172870 kcal/mol) than its co-additives, while (Z)-3-(4-(5-amino-4-methyl-2-phenyl-2,5-dihydroisoxazol-3-yl)-3-methylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one was also excellently bound to DLC surface (-12286.8 kcal/mol) than other antioxidant lubricant additives (Table 3).

**Conclusion**

Quantitative structure-properties relationships (QSPR) which is an in-silico method was used to correlate 2D and 3D properties (descriptors) of antioxidant lubricant additives (compounds) with their properties. QSPR mathematical model along with molecular descriptors that has high coefficient value was used to design five novel antioxidant lubricant additives. All the novel additive’s antioxidant properties were found to be better than the experimental additives in table 1 and our previous study [36], with the novel lubricant additive (Z)-3-(4-(5-amino-1-phenyl-1H-pyrazol-3-yl)-3,5-dimethylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one found to possessed excellent antioxidant properties of 0.850281 total acid values (T.A.V 0.1g/L). Moreover, all the designed additives were found to dynamically bind to steel crystal surface excellently than the DLC crystal surface; this was in agreement with our previous study [36]. The additives; (Z)-3-(4-(5-amino-2-phenyl-2,5-dihydroisoxazol-3-yl)-3-methylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one and (Z)-3-(4-(5-amino-4-methyl-2-phenyl-2,5-dihydroisoxazol-3-yl)-3-methylphenyl)-2-phenyl-5-(thiophen-2-ylmethylene)-3,5-dihydro-4H-imidazol-4-one was excellently bound to steel (-172870 kcal/mol) and to DLC (-12286.8 kcal/mol) surfaces than their co-additives. The molecular dynamic simulations results were found to be better than the one reported by our previous study [36]. This investigation will help to synthesize novel and excellent anti-oxidant lubricant additives that will hinder the base oil from undergoing complete circle of oxidation and they meet environmental requirements as these novel additives do not contain Zinc and Phosphorus which often rendered exhaust pipe catalytic converter inactive thereby increasing environmental pollution [9-11].

**Conflict of Interest**
No conflict of interest.

**Funding**
The authors did not receive any direct funding for this research work.

**Compliance with Ethics Requirements**
This article does not contain any studies with human or animal subjects.
References


