QSAR study of diethyl p-nitrophenyl phosphate derivatives for paraoxonase 1

Mustafa Lawar1, Ismaıl Yalçın2, Murat Alper Başaran3, Hakan Sezgin Sayiner4 and Fatma Kandemirli5

1Kastamonu University, Faculty of Engineering & Architecture, Department of Genetic & Bioengineering, 37000 Kastamonu, Turkey
2Ankara University, Faculty of pharmacy, Department of chemical pharmacy., Ankara, Turkey
3Alanya Alaaddin Keykubat University, Faculty of Engineering, Department of Management Engineering, 07425, Antalya, Turkey
4Department of Infectious Diseases, Faculty of Medicine, Adiyaman University, Adiyaman, Turkey
5Biomedical Engineering Department, Faculty of Engineering and Architecture, Kastamonu University, Kastamonu, Turkey

Abstract

It is perhaps noteworthy in the recent literature of biomedical engineering that there have been a plethora of studies concerned with the activation of human paraoxonase1 (PON1) to reduce the high concentrations of homocysteine in the human serum. Some of diethyl p-nitrophenyl phosphate derivatives were studied at the B3LYP/6-311g(d,p), B3LYP/6-311++g(2d,2p) and B3LYP/6-311++g(3df,3dp) basis set level through the relationship between their molecular and electronic structure, as well as via their relations with paraoxonase1 activity. In an attempt to shed light on which descriptors may contribute to the activation of PON1, this paper investigates the nature of the relationship between the activity of the enzymes and the descriptors of 15 molecules, namely HOMO, LUMO, Energy gap, hardness, softness, Electronegativity, chemical potential, electrophilicity index(ω), Nucleofugality, Electrofugality. Two data analysis methods, CART Decision Tree and Artificial Neural Networks (ANNs) were used for the linear part and non-linear part of the data set, respectively. The results of the study show correlations between the activity of the enzyme and the studied descriptors. Moreover, the results of the study reveal that by using the CART method not only do we know the significant descriptors but also we have their critical values and orders.

Introduction

Paraoxonase is a group of enzymes that catalyzes the organophosphates and lactones. This group contains three genotypic forms coded as PON set, which are located in the long arm of chromosome 7 of human beings [1-3]. The three types of paraoxonase are: PON1 that is synthesized in the liver and functions as an antioxidant, PON2 that is expressed as intracellular protein that can protect cells against oxidative damage, and PON3 that is similar to PON1 in activity but differs from it in substrate specificity [4,5]. Both PON1 and PON3 have clinical significance in that they are involved in lowering the risk of developing atherosclerosis and coronary artery disease [2,6]. PON1 or serum paraoxonase, known also as serum aryldialkylphosphatase 1 and homocysteine thiolactonase, is an enzyme that is found in many mammalian species and in humans and is encoded by the PON1gene. Although PON1 hydrolyzes aromatic carboxylic acid esters, toxic organophosphate compounds, and lactones, its natural substrates and physiological function(s) are far from settled [3-5]. Human PON1 is a glycoprotein composed of 354 amino acids, which associates with high-density lipoprotein (HDL, a cholesterol carrier in the circulation) [1,3,4,7,8]. Serum PON1 is secreted mainly by the liver, local synthesis occurs in several tissues though. The structure contains 2 calcium ions which are essential for catalytic activity and enzyme stability [3-5,8,9].

PON1 is one of three members of mammalian family enzymes that contain PON2 and PON3. These enzymes were originally discovered through their involvement in the hydrolysis of organophosphate. Members of paraoxonase exhibit a wide range of physiologically important hydrolytic activities, including drug metabolism and detoxification of nerve agents. Due to the similarity between PON1 and PON3, both of these enzymes come to be associated with high-density lipoprotein (HDL) [5].

The "natural" substrates for PON1 are lactones [2]. However, PON1 has evolved to be a highly promiscuous enzyme that is capable of hydrolyzing a wide variety of substrates, such as lactones (including a number of important pharmaceutical agents viz., statins), glucuronide drugs, thiolactones, arylesters, cyclic carbonates, organophosphorus pesticides, let alone nerve gases such as sarin, soman and VX, estrogen esters and lipid-peroxides (oxidized lipids). PON1 was first discovered through its ability to hydrolyze and therefore detoxify organophosphorus compounds which are widely used as pesticides and nerve gases. Despite decades of research, it is only now becoming clear that PON1 protects humans from the acute and chronic harmful effects of these compounds [10,11].

In a similar vein, Diethyl Phosphate belongs to a group of compounds called organophosphates. Organophosphate (OP) or phosphate ester is the general name for esters of phosphoric acid. Many of the most important biochemicals are organophosphates, including DNA and RNA, as well as many of the cofactors essential for...
Material and method

The 15 descriptors mentioned above have been carried out at the B3LYP level of theory using Gaussian-03 series of program package [12,13]. The calculations were based on 6-31G (d,p) basis set. This method has been widely implemented to study the relationship between corrosion inhibition efficiency of the molecules and their electronic properties [13] In order to set up correlation between experimental data and structural and electronic characteristics of the investigated activators, the geometry of the molecules were optimized by the density functional theory (DFT) [13], with the Becke’s three parameter exchange functional[1] along with the Lee–Yang–Parr correlation functional theory (B3LYP) [14].

Results and discussion

Tables 1 and 2 delineates the calculated parameters of the diethyl p-nitrophenyl phosphate derivative molecules using B3LYP/6-311G(d,p), B3LYP/6-311++G(2d,2p) methods.

As HOMO is often associated with the electron donating ability of a molecule, high value of HOMO is likely to indicate the tendency of the molecule to donate electrons to appropriate acceptor molecules with lower energy MO [15-17]. HOMO and LUMO orbitals of the 15 molecules were obtained from the quantum chemical calculation by the DFT using B3LYP/6-311G(d,p), B3LYP/6-311+G(2d,2p) bases sets as shown in Figure 1 below [18,19].

Both HOMO and LUMO of the 15 molecules were centered on the phenyl group, which illustrated the reactive orbitals of these molecules. The highest $E_{HOMO}$ among the 15 molecules were -6.682 eV, -6.745 eV, -6.685 eV, -6.922 eV, -6.931 eV, -6.967 eV, -7.055 eV, -7.087 eV, -7.091 eV, -7.123 eV, -7.19 eV, -7.254 eV, -7.305 eV, -7.423 eV, and -7.479 eV. These results were observed in the molecules 15, 7, 14, 12, 10, 8, 13, 9, 6, 3, 4, 11, 5, 1, and 2 respectively. The $E_{LUMO}$ of these molecules were -0.849 eV, -1.208 eV, -1.453 eV, -2.024 eV, -2.106 eV, -2.133 eV, -2.144 eV, -2.245 eV, -2.495 eV, -2.617 eV, -3.007 eV, -5.058 eV, -5.092 eV, -5.476 eV, and -5.66 eV for the molecules 15, 14, 7, 9, 13, 12, 3, 10, 8, 11, 2, 4, 6, 1, and 5 respectively. The energy gap results were 1.645 eV, 1.948 eV, 1.999 eV, 2.133 eV, 4.472 eV, 4.636 eV, 4.686 eV, 4.742 eV, 4.789 eV, 4.949 eV, 4.979 eV, 5.064 eV, 5.292 eV, 5.677 eV, and 5.833 eV correspondingly. In line with the results of energy gap calculated by B3LYP/6-311G(d,p) method and the interpretation of Figures (2-4), we can understand that the molecules with smaller $E_{HOMO}$-$E_{LUMO}$ energy gap lead to lower kinetic stability and higher chemical reactivity [15]. Thus, the molecules with the highest reactivity are 5, 1, 6, 4, 8, 11, 10, 2, 12, 13, 3, 9, 7, 14, and 15.

The negative charge as shown in the above ESP charts is around O32 of the aldehyde group that is attached to the phenyl ring, O31 of the acetophenone of the molecule (10), and N31 of the cyanophenyl of the molecule (13).

With respect to Nucleofugality (defined as the propensity of an atom or group of them to depart bearing the bonding electron pair in a heterolytic cleavage process [20,21], the highest nucleofugality of the 15 molecules was 89.62, 73.007, 62.583, 58.62, 14.828, 12.303, 11.677, 9.964, 9.263, 9.224, 9.034, 8.549, 5.826, 4.865 and 3.651 for the molecules 5, 1, 6, 4, 2, 11, 8, 10, 12, 3, 13, 9, 7, 14 and 15 respectively. According to these results, the molecules that can activate the PON1 enzyme of the aforementioned group of molecules. The Table 3 below shows the correlations between the calculated parameters and the activity of the enzyme. Because of there are some parameters show a weak relationship with the activity of the enzyme that measured in the reference (Khrersonsky O, Tawfik DS. 2005). We know that the bond length is: the distance between two bonded atoms at their minimum potential energy, or the average distance between two bonded atoms. The correlation between the bond length and reactivity is that the longer the bond length is the more reactive it will be. The short one is more stable, however [22]. The longer bonds length includes those bonds within the phosphate group attached to the phenyl ring, mainly the bond between the phosphor atom and oxygen atom 4 (P1-O4: 1.6383Å), (P1-O2: 1.5868 Å), and (P1-O3: 1.5782 Å) that bonded to the phenyl ring, whereas the shorter bonds were (O4-C20: 1.3615 Å), (O2-C6: 1.4592 Å) and (P1-O5: 1.4609 Å) [23].

This bonds length was among the entire molecules, as shown in figure 6 below.

Statistical analysis

The total number of 15 molecules based on descriptors, namely, HOMO, LUMO, Energy Gap, Hardness, Softness, Electronegativity, Chemical Potential, Electrophilicity Index, Nucleofugality, Hyperpolarisibility, Fz, Alpha, Delta and dependent variable $K_{CAT}$ were investigated using CART Decision Tree and Artificial Neural Networks (ANNs) in order to examine which factors may have impact on the dependent variable, $K_{CAT}$. The data set included both linear and nonlinear calculations. While CART Decision Tree was conducted in order to pin down which descriptors have impact on the dependent variable, $K_{CAT}$ using SALFORD Predictive Modeler 8.0 for linear part, Artificial Neural Networks (ANNs) was utilized for the same purpose using SPSS 20.0 for non-linear part. CART Decision Tree tests revealed that Chemical Potential and Homo descriptors were statistically significant variables by conducting 5-fold cross-validation with 0.544 coefficient of determination.

Of the variables that have significant results that are in sync with the objective of the study are Chemical Potential and Homo. The model below delineates the output of this relationship and pinpoints these significant variables. The importance level for Chemical potential is higher since the separation starts with that descriptor. Besides, its critical value is -10.31. If the molecules display critical value less than and equal to -10.31, there will be no other significant descriptors available except Chemical Potential. By contrast, if the Chemical Potential is greater than -10.31, there will be other significant descriptors adding extra information to the model, which is called Homo. The Descirptors’ contribution to the model comes with a single separating value at -6.90.
Lawar M (2017) QSAR study of diethyl p-nitrophenyl phosphate derivatives for paraoxonase 1

Tables 1. The calculated parameters by B3LYP/6-31G(d,p) : HOMO, LUMO, Energy gap, Hardness, Softness, Electronegativity, Chemical potential, Electrophilicity index(ω), Nucleofugality, & Electrofugality.

<table>
<thead>
<tr>
<th></th>
<th>HOMO</th>
<th>LUMO</th>
<th>Energy gap</th>
<th>hardness</th>
<th>softness</th>
<th>Electronegativity</th>
<th>chemical potential</th>
<th>electrophilicity index(ω)</th>
<th>Nucleofugality</th>
<th>Electrofugality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-7.423</td>
<td>-5.476</td>
<td>1.948</td>
<td>0.974</td>
<td>0.513</td>
<td>12.899</td>
<td>-12.899</td>
<td>85.419</td>
<td>73.007</td>
<td>98.80504</td>
</tr>
<tr>
<td>3</td>
<td>-7.123</td>
<td>-2.144</td>
<td>4.979</td>
<td>2.49</td>
<td>0.201</td>
<td>9.267</td>
<td>-9.267</td>
<td>17.246</td>
<td>9.224</td>
<td>27.75778</td>
</tr>
<tr>
<td>4</td>
<td>-7.19</td>
<td>-5.058</td>
<td>2.133</td>
<td>1.066</td>
<td>0.469</td>
<td>12.248</td>
<td>-12.248</td>
<td>70.334</td>
<td>58.62</td>
<td>83.1154</td>
</tr>
<tr>
<td>5</td>
<td>-7.305</td>
<td>-5.66</td>
<td>1.645</td>
<td>0.823</td>
<td>0.608</td>
<td>12.965</td>
<td>-12.965</td>
<td>102.174</td>
<td>89.62</td>
<td>115.5502</td>
</tr>
<tr>
<td>6</td>
<td>-7.091</td>
<td>-5.092</td>
<td>1.999</td>
<td>0.999</td>
<td>0.5</td>
<td>12.183</td>
<td>-12.183</td>
<td>74.267</td>
<td>62.583</td>
<td>86.94992</td>
</tr>
<tr>
<td>7</td>
<td>-6.745</td>
<td>-1.453</td>
<td>5.292</td>
<td>2.646</td>
<td>0.189</td>
<td>8.199</td>
<td>-8.199</td>
<td>12.702</td>
<td>5.826</td>
<td>22.22404</td>
</tr>
<tr>
<td>14</td>
<td>-6.885</td>
<td>-1.208</td>
<td>5.677</td>
<td>2.838</td>
<td>0.176</td>
<td>8.094</td>
<td>-8.094</td>
<td>11.54</td>
<td>4.865</td>
<td>21.0528</td>
</tr>
<tr>
<td>15</td>
<td>-6.682</td>
<td>-0.849</td>
<td>5.833</td>
<td>2.917</td>
<td>0.171</td>
<td>7.532</td>
<td>-7.532</td>
<td>9.725</td>
<td>3.651</td>
<td>18.71468</td>
</tr>
</tbody>
</table>

Table 2. The other calculated parameters by B3LYP/6-31G(d,p): Hyperpolarisibility, Pz valance, Alpha, Delta (alpha) & the K_{cat}/K_{m} (activity).

<table>
<thead>
<tr>
<th></th>
<th>Hyperpolarisibility</th>
<th>Pz (elect)</th>
<th>Alpha(au)</th>
<th>Delta (alpha) esu 10-24</th>
<th>K_{cat}/K_{m} (activity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.68</td>
<td>0.014</td>
<td>266.983</td>
<td>39.56693</td>
<td>130</td>
</tr>
<tr>
<td>2</td>
<td>8.30</td>
<td>-0.051</td>
<td>234.831</td>
<td>34.802</td>
<td>4040</td>
</tr>
<tr>
<td>3</td>
<td>0.73</td>
<td>-0.258</td>
<td>202.744</td>
<td>30.04666</td>
<td>122</td>
</tr>
<tr>
<td>4</td>
<td>7.89</td>
<td>-0.280</td>
<td>236.182</td>
<td>35.00222</td>
<td>8300</td>
</tr>
<tr>
<td>5</td>
<td>1.12</td>
<td>0.014</td>
<td>265.589</td>
<td>39.36024</td>
<td>5600</td>
</tr>
<tr>
<td>6</td>
<td>9.23</td>
<td>-0.665</td>
<td>235.636</td>
<td>34.92126</td>
<td>5800</td>
</tr>
<tr>
<td>7</td>
<td>2.49</td>
<td>-0.147</td>
<td>198.966</td>
<td>29.48676</td>
<td>48</td>
</tr>
<tr>
<td>8</td>
<td>7.92</td>
<td>-0.763</td>
<td>229.294</td>
<td>33.98132</td>
<td>400</td>
</tr>
<tr>
<td>9</td>
<td>3.88</td>
<td>-0.743</td>
<td>229.322</td>
<td>33.98552</td>
<td>220</td>
</tr>
<tr>
<td>10</td>
<td>8.21</td>
<td>-0.761</td>
<td>242.487</td>
<td>35.93657</td>
<td>285</td>
</tr>
<tr>
<td>11</td>
<td>3.66</td>
<td>-0.534</td>
<td>232.672</td>
<td>34.48199</td>
<td>58</td>
</tr>
<tr>
<td>12</td>
<td>7.39</td>
<td>-0.789</td>
<td>243.966</td>
<td>36.15576</td>
<td>130</td>
</tr>
<tr>
<td>13</td>
<td>2.35</td>
<td>-0.673</td>
<td>227.786</td>
<td>33.75784</td>
<td>88</td>
</tr>
<tr>
<td>14</td>
<td>1.70</td>
<td>-0.228</td>
<td>198.971</td>
<td>29.48755</td>
<td>0.57</td>
</tr>
<tr>
<td>15</td>
<td>0.44</td>
<td>-0.748</td>
<td>214.728</td>
<td>31.82269</td>
<td>0.403</td>
</tr>
</tbody>
</table>

1                        1-HOMO                  1-LUMO                      1-ESP
2                            -HOMO                      2-LUMO                    2-ESP
3                         3-HOMO                  3-LUMO                        3-ESP
4                                 4-HOMO                    4-LUMO                       4-ESP
5                            5-HOMO                    5-LUMO                     5-ESP
6                         6-HOMO                     6-LUMO                     6-ESP
7                           7-HOMO                   7-LUMO                    7-ESP
8                            8-HOMO                       8-LUMO               8-ESP
9                            9-HOMO                   9-LUMO              9-ESP
10                             10-HOMO                    10-LUMO                  10-ESP
11                        11-HOMO                  11-LUMO                   11-ESP
12                          12-HOMO                 12-LUMO                12-ESP
13                            13-HOMO              13-LUMO                13-ESP
14                          14-HOMO                 14-LUMO               14-ESP
15                              15-HOMO             15-LUMO                15-ESP
Lawar M (2017) QSAR study of diethyl p-nitrophenyl phosphate derivatives for paraoxonase 1


Volume 2(2): 5-9
Figure 1. Molecular structure, HOMO, LUMO & ESP of the 15 molecules.
As for the non-linear part, ANNs was conducted in order to examine which descriptors may have impact on the dependent variable, \( KCAT \). Owing to the fact that the data set we had was small, we had the model run 100 times and then the average was taken. The most significant variables were \( \text{Pz} \), \( \text{Hyperbola} \), \( \text{Delta} \) and \( \text{Alpha} \) with 62 percent of determination of coefficient \((R^2=0.62)\). In consequence, the total number of 15 molecules consisting of 14 descriptors and one independent variable that have been already mentioned above were investigated in order to examine which descriptors would impact the dependent variable. Since the data set contained both linear and non-linear calculations, two different statistical models, CART Decision Tree and ANNs, were utilized for the linear part and non-linear part, respectively. The results revealed that Chemical Potential and Homo were significant descriptors for the linear part. Similarly, \( \text{Pz} \), \( \text{Hyperpolarizability} \), \( \text{Delta} \) and \( \text{Alpha} \) were significant descriptors for the non-linear part.

To sum up, utilizing ANNs for the non-linear part shows that the descriptors, \( \text{Pz} \), \( \text{Hyperpola.} \), \( \text{Delta} \) and \( \text{Alpha} \) do have an impact on the activity of the enzymes. Similarly, CART model used to measure the effect of the descriptors for the linear part on the activity of the enzymes reveals that the most significant descriptors are Chemical Potential and Homo. Most importantly, however, the results show that the order of the descriptors in the CART model bear significance. That is, while Chemical Potential with its value of greater than -10.31 was the first splitting variable having impact on dependent variable \( KCAT \), Homo came second in importance, with a value of less than and equal to -6.90. By utilizing the CART method, thereupon, not only do we have the significant descriptors but also we have their critical values and orders.

**Summary and Conclusion**

The major aim of this paper was to examine the descriptors of the 15 aforementioned molecules that may activate PON1. The results of the study reveal that there is a strong relationships between the activity of the enzyme and the studied descriptors. In order to examine the relationships, two data analysis methods, CART Decision Tree and Artificial Neural Networks (ANNs) were used for the linear part and non-linear part of the data set, respectively. The most significant
Figure 3. HOMO & LUMO at B3LYP/6-311++G(2d,2p).

Figure 4. HOMO & LUMO at B3LYP/6-311+G(3df,3pd).

Figure 5. The longest and shortest bond length.

Lawar M (2017) QSAR study of diethyl p-nitrophenyl phosphate derivatives for paraoxonase 1


Volume 2(2): 8-9
descriptors for the linear part were Chemical Potential and Homo. However, the order of the descriptors was determined by CART model. While Chemical Potential with its value of greater than -10.31 was the first splitting variable to have impact on dependent variable KCAT, Homo was the second variable that showed an impact on dependent variable, with its value of less than and equal to -6.90. The importance of CART method was that it not only revealed and sorted out significant descriptors but also pinpointed their critical values and orders. In result, CART Decision Tree turned out that Chemical Potential and Homo descriptors were statistically significant variables by conducting 5-fold cross-validation with 0.544 coefficient of determination. In a similar fashion, ANNs, employed for non-linear part in order to determine which descriptors have impact on dependent variable, revealed that Pz, Hyperpolarizability, Delta and Alpha were the most significant variables by conducting 3-fold validation with 0.62 percent of determination of coefficient (R² =0.62).

References

8. Yeung DT, Josse RJ, McAndrew CW, et al. (2004) Structure/ function analyses of human serum paraoxonase (HuPON1) mutants designed from a DFPase-like homology model. Biochim Biophys Acta 1702: 67-77. [Crossref]
17. Osman OI (2017) DFT Study of the Structure, Reactivity, Natural Bond Orbital and Hyperpolarizability of Thiazole Acry Dyes. Int J Mol Sci 18. [Crossref]
19. Khersonsky O1, Tawfik DS (2005) Structure-reactivity studies of serum paraoxonase PON1 suggest that its native activity is lactonase. Biochemistry 44: 6371-6382. [Crossref]

Copyright: ©2017 Lawar M. 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.