Original Research Paper

**QSAR Modeling of Thirty Active Compounds for the Inhibition of the Acetylcholinesterase Enzyme**

1Nour El Houda Hammoudi, 1Yacine Benguerba and 2Widad Sobhi

1Laboratoire des Matériaux Polymères Multiphasiques, LMPMP, Université Ferhat ABBAS Sétif-1, 19000 Sétif, Algeria
2Laboratoire de Biochimie Appliquée, Université Ferhat ABBAS Sétif-1, 19000 Sétif, Algeria

**Abstract:** This work aims at developing a reliable and predictive QSAR model which allows, on one hand, an exploration of the main molecular descriptors responsible for the inhibitory activity towards the Acetylcholinesterase enzyme and, on the other hand, predict the inhibitory activity of new compounds before testing them experimentally. This study involves a series of DL0410 and its 29 DL0410 derivatives. The Multiple Linear Regression (MLR) analysis is carried out to derive the QSAR model. The results indicate that the QSAR model is robust and possesses a high predictive capacity.

**Keywords:** QSAR Model, Acetylcholinesterase, Alzheimer Disease, MLR Model

**Introduction**

Cholinesterase inhibitors (Acetylcholinesterase, Butyrylcholinesterase) are currently the most established treatment strategy in Alzheimer’s Disease (AD) (Cummings et al., 2018; Vellas et al., 2007). In order to reach a better treatment of AD, much focus was put on the development of cholinesterase inhibitor drugs (Deb et al., 2012).

Nowadays, the development of any particular drug for human consumption takes an average of 10 to 15 years before being allowed to enter the marketplace, in addition to the costly process. For this reason, the use of alternative methods such as QSAR/QSPR is cheaper, faster and indispensable (Ridzuan et al., 2012; Slater and Kontoyianni 2019). In the present work, a QSAR study was carried out on a series of 30 compounds DL0410 and its 29 derivatives which have an inhibitory effect against the Acetylcholinesterase enzyme (AChE). The objective of this application is to develop a reliable and predictive QSAR model for the determination of the inhibitory activities of new drug candidates and to examine factors other than volume that can control the inhibitory concentration of these compounds (Patel et al., 2014).

**Quantitative Structure Activity Relationship (QSAR)**

QSAR: is a mathematical model that links the structural features of the compounds (i.e., molecular descriptors) to their quantity showing specific biological activity (Rauf et al., 2019).

**Methodology**

**Database and Calculation Methods**

The QSAR model was generated using 30 compounds DL0410 and its 29 derivatives, 20 molecules used as a training set and 10 molecules as test set. The structures and the experimental values of the biological activity were taken from the published literature (Pang et al., 2017).

The quantum chemical calculations were carried out assuming the Generalized Gradient Approximation (GGA) within the framework of the Density Functional Theory (DFT). This task was achieved with the software package DMol3 in Materials Studio in order to optimize the molecule geometry and to obtain the quantum chemical parameters. The geometric optimization was made by the Triple Numerical with Polarization (TNP) basis set and the functional exchange-correlation (BP) (Musa et al., 2012a; 2012b). The quantum chemical parameters calculated were HOMO, LUMO, Molecular sizes (area and volume) and sigma profiles of all the compounds. The other descriptors were obtained from the Swiss ADME website.

**MLR Model Elaboration**

The Multiple Linear Regression (MLR) analysis using stepwise regression was carried out to derive...
QSAR model coefficients. The best obtained QSAR model for acetylcholinesterase (ES: 3.1.1.7) inhibition activity is given as follows:

\[ \text{pIC}_{50} = 45.33 + 0.013 \times \text{Volumle} - 7.09 \times \text{Gap} - 0.19 \times \text{DMM} + 0.99 \times \text{Oint} - 2.2 \times \text{Nint} - 0.48 \times \text{Cyp} + 1.05 \times \text{NAR} + 0.33 \times \text{MR} + 8.63 \times \text{LUMO} \]

From the statistical analysis, the calculated coefficient of determination \( R^2 \) was equal to 0.94, the root-mean-square deviation (RMSE) was equal to 0.260 and the Adjusted \( R^2 \) was equal to 0.88. The MLR model provides an accurate fit of the experimental data set and it is characterized by a high predictivity. The volume, Intracycle Oxygen, Non-Aromatic Ring, Rotatable Bonds, Molar Refractivity and LUMO are preceded by positive sign. Consequently, these parameters have an increasing effect on the dependent variable value PIC50. As for the Gap, Intracycle Nitrogen and cyclopropanes are preceded by a negative sign. Consequently, these parameters have a diminishing effect on the dependent variable PIC50.

Most of the inhibitory concentration points (Fig. 1) are concentrated around the middle of the fit line, meaning that the considered compound has a similar inhibitory action. In general, this dispersion shows that the provided data were well matched by the MLR model.

**Model Validation**

**Internal Validation**

Cross-validation statistical procedure was used to evaluate the predictive power of QSAR model. The coefficient that describes this validation is given by the equation below (Katritzky et al., 2010):

\[ R_{CV}^2 = 1 - \frac{\sum (y_i^{\text{pred}} - y_i^{\text{obs}})^2}{\sum (y_i^{\text{obs}} - \bar{y}_i)^2} \]

In this case, the obtained \( R_{CV}^2 \) is 0.60.

**External Validation**

After getting the QSAR model, a series of molecules was tested by the retrieved model and compared with the experimental values. The results are shown in Table 1. where the experimental and the predicted values are close to one another.

**Applicability Domain**

The Applicability Domain (AD) is a specific physico-chemical, structural or biological space (Sheridan et al., 2004), which allows to define the area in which a compound can be confidently predicted (Katritzky et al., 2010). In this study, the leverage method was carried out in order to determine the AD of the obtained model. The result is shown in Fig. 2.

**Table 1:** Tested compounds pIC50

<table>
<thead>
<tr>
<th>Compound</th>
<th>pIC50 (experimental)</th>
<th>pIC50 (predicted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>5.4788619</td>
<td>5.2123648</td>
</tr>
<tr>
<td>3.5</td>
<td>5.3106911</td>
<td>5.4499948</td>
</tr>
<tr>
<td>4.1</td>
<td>4.2640819</td>
<td>3.7478612</td>
</tr>
<tr>
<td>4.2</td>
<td>5.5128616</td>
<td>5.7782385</td>
</tr>
<tr>
<td>4.3</td>
<td>4.7594508</td>
<td>4.2963982</td>
</tr>
<tr>
<td>4.4</td>
<td>4.804931</td>
<td>4.0373946</td>
</tr>
<tr>
<td>5.1</td>
<td>6.05061</td>
<td>5.0325662</td>
</tr>
<tr>
<td>5.2</td>
<td>5.4100504</td>
<td>4.4260751</td>
</tr>
<tr>
<td>5.3</td>
<td>4.8513973</td>
<td>2.7336582</td>
</tr>
<tr>
<td>6.1</td>
<td>7.2218487</td>
<td>6.1383127</td>
</tr>
</tbody>
</table>

**Fig. 1:** Parity diagram (observed vs predicted) pIC50 values using MLR model
The blue circles represent the training set molecules and the green circles represent the test set molecules. All these compounds have a residual and leverage that does not exceed the threshold $h^* = 3 \frac{p}{n}$. If an external molecule is outside the defined space of the model ($h^* > 3 \frac{p}{n}$), it is considered outside of the model’s applicability domain and will not have a reliable prediction (Oluwaseye et al., 2018).

$p$: Number of descriptors + 1,
$n$: Number of molecules (training set)

**Conclusion**

In this study the resulted QSAR model proved to be robust and possesses a high predictive capacity. The selected descriptors directly explain the structural features of the compounds responsible for the inhibitory activity of the Acetylcholinesterase. Therefore it can be concluded that the most active predicted compounds are characterized by a Gap, Intracycle Nitrogen and cyclopropanes which should not be elevated and the higher the number of, Intracycle Oxygen, Non-Aromatic Ring, Rotatable Bonds, Molar Refractivity, volume and LUMO the higher the $pIC_{50_{ACHE}}$ value.

**Acknowledgement**

The authors gratefully acknowledge the General Directorate of Scientific Research and Technological Development (DGRST).

**Author’s Contributions**

**Nour El Houda Hammoudi**: Work realization (calculation and redaction).

**Yacine Benguerba**: Orientation and correction of the molecular modeling part.

**Widad Sobhi**: Orientation and correction of the biological part

**Ethics**

This research was subjected to ethical clearance from the faculty of technology, Frehat ABBAS university setif-1, setif/Algeria

**References**


