3D QSAR Studies of Some Substituted Imidazolinones Derivatives Angiotensin II Receptor Antagonists

Mukesh Chandra Sharma and Dharm Veer Kohli

Department of Pharmaceutical Sciences, Drug Research Laboratory, Dr. H.S. Gour University, Sagar (M.P.) 470 003, India

Abstract: A quantitative structure activity relationship study was performed on a series of imidazolinones as nonpeptide angiotensin II receptor antagonists to for establishing quantitative relationship between biological activity and their physicochemical / structural properties. The k-Nearest Neighbor Molecular Field Analysis (kNN-MFA), a three dimensional quantitative structure activity relationship (3D-QSAR) method has been used in the present case to study the correlation between the molecular properties and the angiotensin II receptor antagonists activities on a series of imidazolinones derivatives using SW variable selection method. kNN-MFA calculations for electrostatic, steric and hydrophobic field were carried out. The master grid maps derived from the best model has been used to display the contribution of electrostatic, steric and hydrophobic. The statistical results by SW variable selection method has shown significant correlation coefficient r² (q²) of 0.7143, r² for external test set (pred r²) 0.7618, coefficient of correlation of predicted data set (pred r²se) of 0.5387 and k nearest neighbor of 3. Set of molecular descriptors that can signify the antihypertensive new design molecules.

Key words: Angiotensin II receptor antagonists • Imidazolinones • Antihypertensive

INTRODUCTION

Renin-angiotensin system is a cascade of proteolytic enzymes (renin and angiotensin converting enzyme (ACE)) that result in the production of the systemic hormone angiotensin II (Ang II). The blockade of RAS with inhibitors of ACE has demonstrated the effectiveness of the reduction of levels of Ang II on cardiovascular and kidney hemodynamics, aldosterone production and release and the absorption of sodium. The therapeutic availability is less for the peptidic Ang II antagonist due to their poor bioavailability; short plasma half-life and partial agonist activity but the nonpeptidic Ang II receptors antagonist lacks the defect of peptidic antagonist [1]. The octapeptide angiotensin II (Ang II, Asp-Arg-Val-Tyr-Ile-His-Pro-Phe) mediates its biological actions by activating at least two distinct receptor subtypes, designated AT₁ and AT₂. Both receptors are seven transmembrane G-protein coupled receptors with 32-34% sequence homology [2,3]. Most of the more well-known physiological effects of Ang II, including vasoconstriction, aldosterone release, stimulation of sympathetic transmission and cellular growth, are generally attributed to AT₁ receptor activation [4-6]. Quantitative structure activity relationship (QSAR) modeling results in a quantitative correlation between chemical structure and biological activity. The QSAR approach helps to correlate the specific biological activities or physical properties of a series of compounds with the measured or computed molecular properties of the compounds, in terms of descriptors [7]. A number of quantitative structure-activity relationship (QSAR) studies related to design of angiotensin II receptor antagonists drugs have also been reported [8-11]. Many different approaches to QSAR have been developed over the years. The rapid increase in the three-dimensional structural information (3D) of bioorganic molecules coupled with the development of fast methods for 3D structure alignment such as active analogue approach [12-13]. We report here the development of a new method (kNN-MFA) that adopts a k-nearest neighbor principle for generating relationships of molecular fields with the experimentally reported activity to provide further insight into the key structural features required to design potential drug.

Corresponding Author: Dr. Mukesh Chandra Sharma, Department of Pharmaceutical Sciences, Drug Research Laboratory, Dr. H.S. Gour University, Sagar (M.P.) 470 003, India, E-mail: mukeshsharma@yahoo.com.
MATERIALS AND METHODS

A data set of fifteen compounds of Imidazolines for angiotensin II receptor antagonists were taken from the literature and used for kNN-MFA analysis [14]. Selection of test set molecules was made by considering the fact that test set molecules represent structural features similar to compounds in the training set. The biological activity values [IC₅₀ (nM)] for angiotensin II receptor antagonists reported in literature were converted to their molar units and then further to negative logarithmic scale (-logIC₅₀) and subsequently used as the dependent variable for the QSAR analysis. The molecular modeling calculations were performed using the molecular design suite (MDS) 3.5 (www.vlifesciences.com). We report here the development of a new method (kNN-MFA) that adopts a k-nearest neighbor principle for generating relationships of molecular fields with the experimentally reported activity. Training set (11 compounds) and the test set (04 compounds) were selected by considering the fact that the test set compounds represents structural diversity and a range of biological activities similar to that of training set. The descriptors selected for modeling activity of the derivatives are summarized in (Table 2). We considered the most active compound as a template for the alignment. The compound moiety of the bioactive molecule was used as a substructure and the rest of the molecules were aligned on it using database alignment method. We hereby report the models, as generated by kNN-MFA in conjunction with stepwise (SW) forward-backward variable selection methods. In the kNN-MFA method, several models were generated for the selected members of training and test sets and the corresponding best models are reported herein. Compounds in test set allowed us to use one test compounds per three training compounds thus resulting in more rigorous validation of the training model. In addition, a wide range of structural diversity of compounds in the test set permit us to evaluate the extrapolative accuracy of the QSAR models. 3D QSAR methods, k-nearest neighbor molecular field analysis (k-NN MFA) requires suitable alignment of given set of molecules. This is followed by generation of a common rectangular grid around the molecules. The steric and electrostatic interaction energies are computed at the lattice points of the grid using a methyl probe of charge +1. These interaction energy values are considered for relationship generation and utilized as

Table 1: Structures and Biological activity data and structures of the compounds in the series

<table>
<thead>
<tr>
<th>Comp</th>
<th>R₁</th>
<th>R₂</th>
<th>R₃</th>
<th>IC₅₀(nM)</th>
<th>-logIC₅₀</th>
<th>Predict activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Phenyl</td>
<td>Phenyl</td>
<td>phenyl</td>
<td>1400.0</td>
<td>3.1460</td>
<td>3.3540</td>
</tr>
<tr>
<td>2*</td>
<td>Methyl</td>
<td>Phenyl</td>
<td>phenyl</td>
<td>100.0</td>
<td>2.0000</td>
<td>1.8320</td>
</tr>
<tr>
<td>3</td>
<td>n-Propyl</td>
<td>Phenyl</td>
<td>phenyl</td>
<td>300.0</td>
<td>2.4770</td>
<td>2.6760</td>
</tr>
<tr>
<td>4</td>
<td>n-Propyl</td>
<td>CF₃</td>
<td>CF₃</td>
<td>20.0</td>
<td>1.3010</td>
<td>1.5630</td>
</tr>
<tr>
<td>5</td>
<td>n-Propyl</td>
<td>CH₃</td>
<td>CH₃</td>
<td>4.0</td>
<td>0.6021</td>
<td>0.8230</td>
</tr>
<tr>
<td>6*</td>
<td>n-Propyl</td>
<td>-(CH₂)₃</td>
<td>-</td>
<td>3.0</td>
<td>0.4771</td>
<td>0.3210</td>
</tr>
<tr>
<td>7</td>
<td>n-Propyl</td>
<td>-(CH₂)₃</td>
<td>-</td>
<td>0.9</td>
<td>-0.0457</td>
<td>-0.0216</td>
</tr>
<tr>
<td>8</td>
<td>n-butyl</td>
<td>-(CH₂)₃</td>
<td>-</td>
<td>0.9</td>
<td>-0.0457</td>
<td>-0.1553</td>
</tr>
<tr>
<td>9*</td>
<td>n-Propyl</td>
<td>-(CH₂)₃</td>
<td>-</td>
<td>3.0</td>
<td>0.4771</td>
<td>0.6230</td>
</tr>
<tr>
<td>10</td>
<td>n-Propyl</td>
<td>-(CH₂)₃-S-(CH₂)₃</td>
<td>-</td>
<td>2.0</td>
<td>0.3012</td>
<td>0.2989</td>
</tr>
<tr>
<td>11</td>
<td>n-Propyl</td>
<td>-(CH₂)₃-O-(CH₂)₃</td>
<td>-</td>
<td>8.0</td>
<td>0.9031</td>
<td>0.9650</td>
</tr>
<tr>
<td>12*</td>
<td>n-Propyl</td>
<td>CH₃</td>
<td>CH₃</td>
<td>3.0</td>
<td>0.4771</td>
<td>0.5480</td>
</tr>
<tr>
<td>13</td>
<td>n-Propyl</td>
<td>-(CH₂)₃-S-(CH₂)₃</td>
<td>-</td>
<td>1.0</td>
<td>0.0000</td>
<td>0.1320</td>
</tr>
<tr>
<td>14</td>
<td>n-Propyl</td>
<td>-(CH₂)₃</td>
<td>-</td>
<td>9.0</td>
<td>0.9542</td>
<td>0.9320</td>
</tr>
<tr>
<td>15</td>
<td>n-butyl</td>
<td>-(CH₂)₃</td>
<td>-</td>
<td>1.0</td>
<td>0.0000</td>
<td>0.1130</td>
</tr>
</tbody>
</table>

*test compound
 descriptors to decide nearness between molecules. The term descriptor is utilized in the following discussion to indicate field values at the lattice points. The optimal training and test sets are generated using the sphere exclusion algorithm. This algorithm allows the construction of training sets covering descriptor space occupied by representative points. Once the training and test sets are generated, k-NN methodology is applied to the descriptors generated over the grid [15]. Once the training and test sets were generated, kNN methodology was applied to the descriptors generated over the grid. In the present kNN-MFA study, (-11.7809 to14.2145) × (-10.3215 to16.4321) × (-13.4354 to17.3215) Å Grid at the interval of 2.00 was generated around the aligned compounds. For calculation of field descriptor values, both electrostatic and steric, hydrophobic field type with cutoff values of 10.0 and 30.0 Kcal/mole respectively were selected and charge type was selected as Gasteiger-Marsili. Dielectric constant was set to 1.0 considering the distance dependent dielectric function. The KNN-MFA models were generated using the variable selection methods, viz. stepwise (SW) forward-backward method.

K-Nearest Neighbor (kNN) Method: The kNN methodology relies on a simple distance learning approach whereby an unknown member is classified according to the majority of its k-nearest neighbors in the training set. The nearness is measured by an appropriate distance metric (e.g., a molecular similarity measure calculated using field intersections of molecular structures) [16]. This method employs the kNN classification principle combined with the stepwise variable selection procedure for optimization of (i) the number of nearest neighbours (k) used to estimate the activity of each compound and optimization of (ii) selection of variable from the original pool of all molecular descriptors (steric and electrostatic fields at the lattice points) that are used to calculate similarities between compounds.

Stepwise (SW) Method: The kNN-MFA model for all the Ang II activities was developed using stepwise forward backward method with cross correlation limit set to 0.5 and term selection criteria as q2. The method resulted in selection of compounds as test set and training set. F-test ‘in’ was set to 4.0 and F-test ‘out’ to 3.99. As some additional parameters, variance cut-off was set as 2 Kcal/mol Å and auto scaling, additionally the K-nearest Neighbor parameter setting was done within the range of 2-5 and prediction method was selected as distance base weighted average.

Cross-Validation: This is done to test the internal stability and predictive ability of the QSAR models. Developed QSAR models were validated by the following procedure:

Internal and External Validation: A molecule in the training set was eliminated and its biological activity was predicted as the weighted average activity of the k most similar molecules [17]. The similarities were evaluated as the inverse of Euclidean distances between molecules using only the subset of descriptors corresponding to the current trial solution.

\[ q^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - y_{\text{mean}})^2} \]

- Step 1 was repeated until every molecule in the training set has been eliminated and its activity predicted once.
- The cross-validated \( r^2 \) (\( q^2 \)) value was calculated, where \( y_i \) and \( \hat{y}_i \) are the actual and predicted activities of the \( i \)th molecule, respectively and \( y_{\text{mean}} \) is the average k-Nearest neighbor activity of all molecules in the training set. Both summations are over all molecules in the training set. Since the calculation of the pairwise molecular similarities and hence the predictions, were based upon the current trial solution, the \( q^2 \) obtained is indicative of the predictive power of the current kNN-MFA model.

External Validation: The predicted \( r^2 \) (\( \text{pred}_r^2 \)) value was calculated using, where \( y_i \) and \( \hat{y}_i \) are the actual and predicted activities of the \( i \)th molecule in test set, respectively and \( y_{\text{mean}} \) is the average activity of all molecules in the training set. Both summations are over all molecules in the test set. The \( \text{pred}_r^2 \) value is indicative of the predictive power of the current kNN-MFA model for external test set.

\[ \text{pred}_r^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - y_{\text{mean}})^2} \]

Randomization Test: To evaluate the statistical significance of the QSAR model for an actual data set, we have employed a one-tail hypothesis testing.
The robustness of the QSAR models for experimental training sets was examined by comparing these models to those derived for random data sets. Random sets were generated by rearranging biological activities of the training set molecules. The significance of the models hence obtained was derived based on calculated Zscore [18].

\[
Z\text{score} = \frac{(h - \mu)}{\sigma}
\]

Where \( h \) is the \( q^2 \) value calculated for the actual dataset, \( \mu \) the average \( q^2 \) and \( \sigma \) is its standard deviation calculated for various iterations using models build by different random data sets. The probability (\( \alpha \)) of significance of randomization test is derived by comparing \( Z \) score value.

RESULTS AND DISCUSSION

The kNN-MFA technique was used to derive 3D-QSAR model for Substituted Imidazolinones Angiotensin II Receptor Antagonists which inhibits Antihypertensive activity. The in vitro inhibitory activity (IC\textsubscript{50} values) in nM, were converted to pIC\textsubscript{50} was used as dependant variable. Relative alignment of all the energy minimized molecules was then carried out by using two techniques namely atom and template based for better results and better assessment between both.

This method utilizes the active analogue principle that lies at the foundation of medicinal chemistry [19-21]. Selecting training and test set by spherical exclusion method. The 3D QSAR models were evaluated using following statistical measures, \( n \), number of observations (molecules); \( V_n \), number of descriptors; \( k \), number of nearest neighbors; \( q^2 \), cross-validated \( r^2 \) (by the leave one out method); \( \text{pred}_r^2 \), predicted \( r^2 \) for the external test set. Hence all the molecules were constructed using the standard geometry with 3D molecular module of Molecular Design Suite. The calculation of 3D descriptors and partial least square analysis was performed on Pentium IV workstation using Molecular Design Suite.

\[
pIC_{50} = E_{-735}(-0.8619-0.4873) - E_{-801}(-0.92460.1268) + H_{-333} (0.29620.5219) - S_{176} (-0.0194-0.0159),
\]

\( N = 16 \), Optimum Components = 3, DF = 20, \( r^2 = 0.7864, q^2 = 0.7143, F_{test} = 27.675, r^2_{se} = 0.3176, q^2_{se} = 0.6954, \text{pred}_r^2 = 0.7618, \text{pred}_{r^2e} = 0.5387, Z\text{Score} Q^2 = 1.321, \) Best Rand \( Q^2 = 0.7643 \).

The descriptors E_{-735}, E_{-801}, H_{-333} and S_{176} are the steric and electrostatic field energy of interactions between probe (CH\textsubscript{3}) and compounds at their corresponding spatial grid points of 735, 801, 333 and 176. The plot of observed versus predicted activity for the training and test sets of compounds in both the cases and contribution chart of selected descriptors are represented in Figure (1). The respective relative contribution of steric, electrostatic and hydrophobic fields indicates that electrostatic field is more predominant. It is evident that the predicted activities of all the compounds in the test set are in good agreement with their corresponding experimental activities and optimal fit is obtained (Table 1). The contribution plot
of steric electrostatic and hydrophobic field interactions indicates relative regions of the local fields (steric electrostatic and hydrophobic) around the aligned molecules, leading to activity variation in the model. The green-coloured balls specify the positions of the steric descriptors and the descriptors with positive or negative coefficients show a region where bulky substituent is favored or disfavored, respectively. Electrostatic field descriptors (blue-coloured balls) with positive coefficients represent regions where electropositive (electron-withdrawing) groups are favorable, whereas negative coefficient indicates that electronnegative (electron-rich or electron-donating) groups are favorable in this region [22]. From 3D-QSAR model it is observed that electrostatic descriptors like E_735 (-13.65%) and E_801 (-24.56%) with negative coefficient are from the R_1 and R_3 position of the imidazolones ring. This indicates that electronnegative groups are favorable on this site and presence of electronnegative groups increases the activity of imidazolones compounds [7, 13]. Most of the compounds (compounds 3, 6-10, 11-14, etc.) with higher activity having electropositive substitution at the R_1, R_2 and R_3 position of imidazolones ring strongly support the above statement. The presence of steric descriptors S_176 (-14.19%) with negative coefficients are near from the R_1 position of the imidazolones ring which indicates that less bulky groups are favorable on this site and presence of bulky groups increases the antihypertensive activity of imidazolones compounds. H_333 descriptors to Hydrogen group nearer to R_1 and R_2 respectively indicates that positive hydrophobic field is favorable for increasing the activity. Hence less hydrophobic or more hydrophilic substituent groups near R_1 and R_2 is preferred. The above results are in close agreement with the experimental observations where compounds 5, 12 and 14 with side chain at the R_1 position produce high activity values.

CONCLUSION

Among combination, SW-based PLS method provides the best results in 3D QSAR study. The 3D results reveal that a less bulky substituent, electronnegative groups increases the activity at of imidazolones ring and electropositive groups at R_1, R_2 and R_3 position of imidazolones ring, are required for potent antihypertensive molecules. Furthermore, we hope that the current study provides better insight into the designing of more potent Ang II inhibitors as antihypertensive agent in the future before their synthesis.

ACKNOWLEDGMENTS

The author wishes to express gratitude to V-life Science Technologies Pvt. Ltd for providing the software for the study and Head, School of Pharmacy, Devi Ahilya Vishwanidyalaya for providing facilities to carry out the work.

REFERENCES


