Inhibition of Hill Reaction by 2-Azido-s-triazine Derivatives: QSAR Study with Molecular Connectivity Indices

Milan Soškic
Department of Chemistry, Faculty of Agricultural Sciences, University of Zagreb, 41000 Zagreb, Croatia, Yugoslavia
Aleksandar Sabljic
Institute Ruder Bošković, P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia
Z. Naturforsch. 44c, 255–261 (1989); received November 29, 1988

2-Azido-s-Triazines, Connectivity Indices. Inhibition of Hill Reaction, Photosystem II. QSAR Modelling

This study was undertaken to find a simple and accurate structural parameters for the quantitative description of inhibitory potency of 2-azido-s-triazines in Hill reaction and to gain more information about the mechanism of inhibition on molecular level. A very good correlation \( r = 0.946 \) was obtained between the \( p_{10} \) values (the negative logarithm of the molar concentration that causes 50% inhibition) and the valence zero-order and the difference between the second-order and the valence second-order molecular connectivity indices. This model, when compared with the empirical models based on the 1-octanol/water partition coefficients and the chromatographic retention data, shows superior performance in accuracy and range of applicability. In addition, the direct correspondence between molecular structure and above connectivity indices makes it possible to locate structural features responsible for the inhibitory potency of 2-azido-s-triazines in Hill reaction. From our QSAR analysis, the interaction between the chloroplast receptor site and 2-azido-s-triazines, which causes inhibition of Hill reaction, is primarily influenced by the size of alkylamino substituents and it accounts for the most variation in the \( p_{10} \) data. The structural features of secondary importance that control the magnitude of \( p_{10} \)'s are the polarity of alkylamino chains and the degree of branching on alpha carbon atom of \( R_1 \) alkylamino substituent. Compared with the main factor, the size of alkylamino substituents, they can be viewed as a fine tuning elements for the inhibitory potency of 2-azido-s-triazines.

Introduction

The majority of commercial herbicides act as inhibitors of photosynthesis in chloroplast. Numerous biochemical and biophysical experiments, mainly on isolated chloroplast, indicated [1–5] that those structurally different classes of herbicides (phenylureas, s-triazines, as-triazinones, uracils, pyridazinones, benzimidazoles, phenols, etc.) all bind competitively to a protein with a molecular weight of 32,000–34,000 within the thylakoid membrane and inhibit the electron flow on the reducing side of photosystem II [6]. Photosystem II inhibitors can be differentiated by their mode of action into two chemically different classes that are referred to as the diuron/triazine type and the phenol type. The \textit{in vitro} inhibitory potency of those herbicides is determined from their ability to inhibit the Hill reaction [7]. In our recent article [8], a very successful QSAR (Quantitative Structure Activity Relationship) model [9], based on the molecular connectivity indices, was formulated for the inhibition of the Hill reaction by 3-alcoxyuracils (the diuron/triazine type inhibitors).

In this study we will extend our investigation on s-triazines and their ability to inhibit the Hill reaction. We will try to create the QSAR model for 2-azido-s-triazines with various alkylamino substituents on positions 4 and 6 (see structure 1).

\[
\begin{align*}
\text{R}^1\text{HN} & \quad \text{N} \quad \text{N} \quad \text{R}^2 \\
\text{NH} & \quad \text{N} \\
\end{align*}
\]

This class of s-triazines is selected for our study because of its structural homogeneity. The molecular connectivity indices [10–13] will be used as the quantitative descriptors for the molecular structure of 2-azido-s-triazines. They have been demonstrated to be very successful in creating numerous QSAR models with physico-chemical properties [10–14], bio-

Reprint requests to Dr. A. Sabljic.
Verlag der Zeitschrift für Naturforschung, D-7400 Tübingen
0341 – 0382/89/0300–0255 © 01,30/0
logical activities [10—13, 15, 16], and environmental behavior [17—27] of chemicals. In addition, these nonempirical structural descriptors can be obtained very fast, with high precision and this process is inexpensive.

Our goal is to develop a quantitative model, based on molecular connectivity indices, that will predict the inhibitory potency of 2-azido-s-triazines in Hill reaction within experimental error. To check the quality of such QSAR model it will be compared with the empirical QSAR models [28, 29] based on Hansch approach [30]. This analysis will provide important information about the performance and range of applicability of molecular topology in predicting inhibitory potencies of herbicides. In addition, we hope to gain more information about structural features that are important for the high inhibitory potency of s-triazines and to learn more about the mechanism of inhibition on molecular level. These results on s-triazines, combined with our earlier findings for 3-alcoxyuracils [8], will also help us to formulate the global model of inhibition for diuron/triazine type inhibitors of photosystem II, which is our ultimate goal.

**Method of calculation and experimental data**

Several extensive reviews of the theory and method of calculation of molecular connectivity indices have been published recently [10—13]. Thus only a brief description of the calculation of the molecular connectivity indices used in the nonempirical models discussed in this study is given here.

The zero-order valence molecular connectivity indices ($\chi^0$) are calculated from the non-hydrogen part of the molecule. Each non-hydrogen atom is described by its atomic $\delta^v$ value, which is calculated from the following equation:

$$\delta^v = (Z^v - h)/(Z - Z^v - 1)$$

where $Z$ is its atomic number, $Z^v$ is the number of valence electrons in the atom and $h$ is the number of hydrogen atoms bound to the same atom. The $\chi^0$ indices are then calculated from the atomic $\delta^v$ values by Eqn. (2),

$$\chi^0 = \Sigma (\delta^v)^{-0.5}$$

and summation is over all non-hydrogen atoms in a molecule.

The second-order molecular connectivity indices ($\chi^2$) are also calculated from the non-hydrogen part of the molecule and corresponding $\delta^v$ values (the number of adjacent non-hydrogen atoms) by Eqn. (3),

$$\chi^2 = \Sigma (\delta^v \cdot \delta^v \cdot \delta^v)^{-0.5}$$

where, i, j, and k correspond to three consecutive non-hydrogen atoms and summation is over all pairs of adjacent bonds between non-hydrogen atoms. The second-order valence molecular connectivity indices ($\chi^v$) are calculated from the atomic $\delta^v$ values by Eqn. (4).

$$\chi^v = \Sigma (\delta^v \cdot \delta^v \cdot \delta^v)^{-0.5}$$

Molecular connectivity indices were calculated by the GRAPH III computer program on an IBM PC/XT personal computer [20]. Minimum hardware and software requirements for this program are IBM PC or compatible computer, 256 KB memory, 1 double sided/double density disk drive, and PC-DOS or MS-DOS operating system version 2.1 or higher. The use of mathematical coprocessor is highly recommended. In its present version, GRAPH III can calculate the molecular connectivity indices up to the tenth order for molecules with 35 non-hydrogen atoms or less. It is possible to extend the program to handle larger molecules if sufficient memory is available.

Regression analysis was carried out using the statistical analysis system (SYSTAT) on the personal computer described above. To test the quality of the regression equations the following statistical parameters were used: the correlation coefficient ($r$), the standard error of the estimate ($s$), a test of the null hypothesis (F-test), and the amount of explained variance (EV).

The inhibitory activities of 2-azido-s-triazines are taken from the study of Gabbott [28]. They are expressed as pI50, the negative logarithm of concentration causing 50% inhibition in Hill reaction. Their experimental errors are in the range of 1—5%.

**Results and Discussion**

The 2-azido-s-triazine derivatives examined in the present QSAR study are shown in Table I together with their inhibitory potencies (pI50) and molecular connectivity indices used as structural descriptors.

From correlation diagrams it was easy to conclude that the exponential relation is apparent between the molecular connectivity indices and inhibitory potencies of 2-azido-s-triazines. Thus, the single variable models (quadratic, hyperbolic, and exponential)
Table I. The zero-order valence molecular connectivity indices, the difference between the simple and valence second-order molecular connectivity indices, Taft's polar substituent $\sigma^*$ constants, plus observed [28] and calculated (Eqs. (6) and (7)) inhibitory potencies ($p_{10}^5$) in Hill reactions of 18 2-azido-s-triazines. 2-Azido-s-triazines used in this study are described by structure 1 and only substituents $R^1$ and $R^2$ are indicated in this table.

<table>
<thead>
<tr>
<th>No.</th>
<th>Compound</th>
<th>$R^1$</th>
<th>$R^2$</th>
<th>$\chi$</th>
<th>$\chi - 2\chi'$</th>
<th>$\sigma^*$</th>
<th>$p_{10}^5$ (exp.)</th>
<th>$p_{10}^5$ (calc.) Eqn. (6) Eqn. (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Methyl Ethyl</td>
<td>7.890</td>
<td>3.039</td>
<td>-1.100</td>
<td>5.26</td>
<td>5.03</td>
<td>5.02</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Methyl n-Propyl</td>
<td>8.597</td>
<td>2.995</td>
<td>-1.15</td>
<td>5.50</td>
<td>5.66</td>
<td>5.70</td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Methyl i-Propyl</td>
<td>8.761</td>
<td>3.111</td>
<td>-0.90</td>
<td>5.82</td>
<td>6.04</td>
<td>6.03</td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Methyl i-Butyl</td>
<td>9.468</td>
<td>2.976</td>
<td>-1.125</td>
<td>6.11</td>
<td>6.18</td>
<td>6.17</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Methyl t-Butyl</td>
<td>9.683</td>
<td>3.170</td>
<td>-0.300</td>
<td>6.75</td>
<td>6.69</td>
<td>6.74</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Methyl t-Amyl</td>
<td>10.39</td>
<td>3.139</td>
<td>-0.315</td>
<td>6.70</td>
<td>6.75</td>
<td>6.80</td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Ethyl Ethyl</td>
<td>8.597</td>
<td>3.138</td>
<td>-0.200</td>
<td>5.99</td>
<td>5.96</td>
<td>5.94</td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Ethyl i-Propyl</td>
<td>9.468</td>
<td>3.211</td>
<td>-0.290</td>
<td>6.54</td>
<td>6.69</td>
<td>6.66</td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Ethyl n-Butyl</td>
<td>10.01</td>
<td>3.095</td>
<td>-0.230</td>
<td>6.80</td>
<td>6.61</td>
<td>6.57</td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Ethyl i-Butyl</td>
<td>10.17</td>
<td>3.077</td>
<td>-0.225</td>
<td>6.64</td>
<td>6.60</td>
<td>6.56</td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>Ethyl s-Butyl</td>
<td>10.17</td>
<td>3.176</td>
<td>-0.310</td>
<td>7.04</td>
<td>6.81</td>
<td>6.81</td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td>Ethyl t-Butyl</td>
<td>10.39</td>
<td>3.270</td>
<td>-0.400</td>
<td>7.46</td>
<td>7.03</td>
<td>7.05</td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td>Ethyl t-Amyl</td>
<td>11.10</td>
<td>3.239</td>
<td>-0.415</td>
<td>6.77</td>
<td>6.85</td>
<td>6.87</td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td>i-Propyl i-Propyl</td>
<td>10.34</td>
<td>3.283</td>
<td>-0.380</td>
<td>6.87</td>
<td>7.05</td>
<td>7.00</td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td>i-Propyl t-Butyl</td>
<td>11.26</td>
<td>3.342</td>
<td>-0.490</td>
<td>6.85</td>
<td>7.01</td>
<td>7.01</td>
<td></td>
</tr>
</tbody>
</table>

The statistical parameters show that equation 5 is statistically significant above the 95% level and it accounts for 80% of the variation in the $p_{10}^5$ data. (The 95% confidence intervals are shown in parentheses.) Although a high correlation coefficient was obtained, the $\sigma^*$ index alone was not able to account for all of the variation in the $p_{10}^5$ data whose average experimental error is only 2–3%. Consequently, two-variable regression equation were screened to find a higher-order molecular connectivity index that will account for the remaining part of the variation in the $p_{10}^5$ data. The best two-variable regression model for the inhibitory potency of 2-azido-s-triazines and its statistical parameters are as follows:

$$p_{10}^5 = -25.49(\pm 5.93) + 4.90(\pm 1.06) \cdot \frac{\chi}{\chi'} - 0.24(\pm 0.06) \cdot (\frac{\chi}{\chi'})^2 + 2.14(\pm 0.69) \cdot (\chi - 2\chi') \cdot (5)$$

$N = 18 \quad r = 0.946 \quad s = 0.182 \quad F^{2.14} = 40 \quad EV = 87.2\%$.

The introduction of the second variable $(\chi - 2\chi')$ made statistically significant improvements in our model. The standard error (s) is lower by 23% and equation 6 accounts for more than 87% of the variation in the $p_{10}^5$ data. In addition, the value of F-test clearly shows that improvements are real and are not caused by the sole fact that more variables are used. The level of cross-correlation $(r^2 = 0.4)$ between the two independent variables is sufficiently low to permit their simultaneous presence in the regression model. A comparison of the observed and predicted inhibitory potency of 2-azido-s-triazines, Table I, clearly demonstrates that the molecular connectivity model (equation 6) is very accurate in predicting their $p_{10}^5$ data. The average difference between predicted and observed $p_{10}^5$'s is only 0.15 log units (factor 1.4) and only one compound is predicted outside the two standard deviations. Moreover, approximately 75% of the predicted $p_{10}^5$ data...
are within the experimental error. The high accuracy of the molecular connectivity model in predicting the inhibitory potency of 2-azido-s-triazines is also shown in Fig. 1 where the observed vs. predicted \( p_{50} \) data of test compounds from Table I are plotted.

![Graph showing observed vs. predicted inhibitory potencies of 2-azido-s-triazines](image)

Fig. 1. Correlation between the observed and the predicted inhibitory potencies (\( p_{50} \)) of 18 2-azido-s-triazines from Table I. The solid line presents their linear regression model. Predicted inhibitory potencies are calculated by Eqn. (6), which is based on molecular connectivity indices.

Although, the introduction of the second variable \((\chi - \chi')\) made very significant improvements to our model, Eqn. (6), 10% of the variation in the \( p_{50} \) data is still unaccountable. At this point, it will be logical to farther increase the complexity of our model, by increasing the number of variables. But, since the amount of experimental data is relatively low the addition of any new variable will greatly increase the possibility of chance correlation [31]. Such unfortunate situation prompted us to continue our search for the better second variable in the large pool of experimental variables [32]. The result of this search is described by Eqn. (7) and its statistical parameters.

\[
p_{50} = -18.84(\pm 5.00) + 4.90(\pm 1.06) \cdot \sigma_{R1}^{*} + 0.24(\pm 0.06) \cdot (\frac{\sigma_{R1}^{*}}{\chi}) + 2.93(\pm 0.94) \cdot (\sigma_{R1}^{*} - \sigma_{R2}^{*})
\]

\( N = 18 \quad r = 0.946 \quad s = 0.182 \quad F_{1,16} = 40 \quad EV = 87.2\% \) (7)

A brief comparison of the two-variable models, Eqn. (6) and (7), clearly shows that they are statistically identical. Thus, it is appropriate to speculate that the polar substituent constant \((\sigma_{R1}^{*} - \sigma_{R2}^{*})\) developed by Taft [33] and the difference between the simple and valence type connectivity indices describe the same physical property of studied 2-azido-s-triazines. To additionally validate this assumption we have also correlated those two variables. As expected, a very good linear correlation is obtained as shown by Eqn. (8), its statistical parameters, and Fig. 2.

\[
\sigma_{R1}^{*} + \sigma_{R2}^{*} = 2.95(\pm 0.26) - 1.02(\pm 0.08) \cdot (\frac{\chi}{\chi'})
\]

\( N = 18 \quad r = 0.952 \quad s = 0.034 \quad F_{1,16} = 155 \quad EV = 90.1\% \) (8)

Both results indicate that the information encoded in the difference between the simple and valence type connectivity indices describes electronic properties of the studied compounds. This result is consistent with the earlier findings reported by Kier and Hall [13]. They have found that such differences between connectivity indices correlate very well with various experimental and calculated descriptors of electronic structure; e.g. electronegativity, Hammett sigma values, solvent polarity, and energies of ionization.

The present QSAR analysis shows that the inhibitory potencies of 2-azido-s-triazines in Hill reaction is primarily influenced by the size of alkylamino substituents, which is described the best by the \(\sigma_{R1}^{*}\) index, since its numerical value is directly proportional to the number of atoms in a molecule and it accounts
for the most variation in the pl_{50} data. The other factor that controls the magnitude of pl_{50}'s is the polarity of alkylamino substituents. The positive regression coefficients of the (\gamma - \chi') index shows that the inhibition increases with the polarity of alkylamino substituents. Compared with the main factor, the size of substituents, it can be viewed as a fine tuning element for the inhibitory potency of 2-azido-s-triazines.

The satisfactory performance of the molecular connectivity model (Eqn. 6) in predicting the inhibitory potency of 2-azido-s-triazines in Hill reaction prompted us to check the quality of its predictions by examining and comparing it with the predictive ability of QSAR models based on empirical parameters (1-octanol/water partition coefficients and chromatographic retention data) [28]. Such analysis is an important part of any modelling process. Those models are described by Eqn. (9) and (10), respectively.

\[
\begin{align*}
pl_{50} &= 0.820 - 0.667(\pi R_1 + \pi R_2)^2 + \\
& 3.69(\pi R_1 + \pi R_2) - 2.73(\sigma^s R_1 + \sigma^s R_2) \\
N &= 18 \quad r = 0.913 \quad s = 0.25 \quad EV = 79.8% \\
pl_{50} &= 1.070 - 7.94(\Delta Rm_{R1R2})^2 + \\
& 12.39(\Delta Rm_{R1R2}) - 3.01(\sigma^s R_1 + \sigma^s R_2) \\
N &= 18 \quad r = 0.947 \quad s = 0.20 \quad EV = 87.4% .
\end{align*}
\]

For the sake of clarity, the comparison of molecular connectivity model (Eqn. 6) with each empirical model will be performed separately.

A brief inspection of Eqn. (6) and (9) and their statistical parameters shows that the molecular connectivity model is superior to classical Hansch type model (Eqn. (9)). The standard error (s) of molecular connectivity model is lower by 28% and the explained variance is higher nearly 10%. Similar result is obtained in our prior study on alcoxyuracil derivatives [8]. Such consistent superior performance of molecular connectivity model is encouraging and gives confidence in correct selection of structural descriptors for modelling the inhibitory potency of Hill reaction. We also hope that the improved models will help in gaining new information about structural requirements for the potent Hill reaction inhibitors.

The comparison of Eqn. (6) and (10) clearly shows that they are statistically identical. Although, the standard error of molecular connectivity model is lower by 9%, the detailed analysis of the calculated pl_{50} data shows that their predictive abilities are very similar. However, there are some limitations that make the application of Eqn. (10) unpracticable, difficult, and theoretically questionable. The theoretical justification for using chromatographic data in QSAR analysis was that “they are free energy based constants identical with Hansch \pi values” and the high correlation between those two empirical parameters for that group of 2-azido-s-triazines [28]. However, for the very similar group of s-triazines (with a chlorine or methylthio groups replacing azido group) analogous correlation is very poor (r^2 = 0.308) [34]. In addition, a complete lack of correlation is observed between their pl_{50} data and log P values (r^2 = 0.097) [34, 35].

The second limitation in using chromatographic data for QSAR modelling is that it is not possible to make predictions for compounds that have not been synthesized and/or those whose chromatographic data have not been measured. (It has to be pointed out here that a real value of models is in their ability to correctly and efficiently handle new situations.) This makes the model based on chromatographic measurements difficult or impossible to apply for the new compounds. It is fair to conclude that in general the performance and future applications of molecular connectivity model are superior to all empirical QSAR models reported for 2-azido-s-triazines inhibition of Hill reaction.

The following section is dedicated to qualitative observations on the structural requirements of chloroplast receptor site for optimal interaction (binding) with alkyl chains in Hill reaction. The conclusions made here are based on analyzing the hierarchy of effects of alkyl substituents on the resulting inhibitory potencies of 2-azido-s-triazines. Such analysis was made possible by superb study of Gabbot [28] where the size and shape of receptor site (i.e. its hydrophobic region) was systematically probed by continuous change in the size and shape of alkyl substituents. First, the asymmetric substitution is a necessary prerequisite for the strong interaction with the hydrophobic region of receptor site. Once the s-triazine ring is fixed in its position the optimal size for alkyl chains are two and four carbon atoms, respectively. Second, the branching on the alpha carbon atom of the larger alkyl chain is very important for strong binding. The higher is substitution the stronger is binding and inhibitory potency. Thus, isopropyl substituent is better than n-propyl substituent and t-butyl is the best of all butyl isomers. This means that the larger side of hydrophobic region is
not extensive and that it resembles a large cavity which can nicely accommodate bulky 1-butyl group. However, the branching on smaller alkyl chain considerably decreases activity of the resulting compound and even prevents 1-butyl group from optimal packing and interaction with receptor site. It is a pity that there is no experimental data for m-propyl substituent at R₂ position, thus it is not possible to draw final conclusion about the size and shape of “smaller” hydrophobic region in chloroplast receptor site. Finally, if the alkyl substituent at position R₁ is kept constant (methyl or ethyl), identical substituent series are obtained for increased contribution of larger substituent (R₂) to the resulting activity of 2-azido-s-triazines:

\[
\text{Me} < \text{Et} < \text{n-Pr} < \text{i-Pr} < \text{i-Bu} < \text{n-Bu} < \text{s-Bu} < \text{t-Bu}.
\]

In the closing paragraph we will compare the molecular connectivity models for alcoxyuracils [8] and azidotriazines activity in Hill reaction and discuss their common features as well as differences since both groups are classified as diuron/triazine type inhibitors of Photosystem II. In general features these models are very similar. The inhibitory potency of 3-alcoxyuracils and 2-azido-s-triazines in Hill reaction is primarily influenced by the size of alcoxy and alkylamino chains, respectively. In both model this relationship is exponential and increases until an optimal value or plateau is reached. For alcoxyuracils this is n-alkyl chain with nine carbon atoms and for 2-azido-s-triazines this is ethyl group plus highly branched 1-butyl group. The structural features of secondary importance differ for those two classes of Hill reaction inhibitors. For alcoxyuracils this is the flexibility of alcoxy chain and for 2-azido-s-triazines the degree of branching on alpha carbon of larger alkylamino substituent (R₂). This variation strongly suggests that alcoxy and alkylamino chains interact with the different sections of chloroplast hydrophobic region. The additional support for this assumption is that it is not possible to logically overlap uracil and s-triazine rings and position the alcoxy and alkylamino chains to occupy the same space.

**Conclusions**

In this investigation, we have demonstrated that a simple model, based on topological indices, can be used to accurately describe the inhibitory potency of 2-azido-s-triazines in Hill reaction. From our QSAR analysis, the size of alkylamino chains accounts for majority of quantitative differences in inhibitory potency found for the studied 2-azido-s-triazines. This relationship is exponential (quadratic) and increases until an optimal value is reached. The structural features of secondary importance for inhibitory potency are the polarity of alkylamino chains and the degree of branching on alpha carbon atom of R₂ alkylamino substituent. The inhibitory activity of 2-azido-s-triazines is directly proportional to them. Created model, when compared with the existing empirical models, shows superior performance in accuracy and future applications.

The molecular connectivity indices correlate nicely (correlation coefficient above 0.9) with the $p_{50}$ data for that small group (6 compounds) of s-triazines. The details about that correlation will be published in one of our future studies.