QSAR studies on anti HIV-1 N-substituted betulinic acid amides

Arun K Srivastava, Mohd Shakeel & Arbab A Khan
Department of Chemistry, University of Allahabad,
Allahabad-211002, India.
Received 19 July 1999, accepted (revised) 22 January 2001

QSAR studies, using structural parameters such as equalized electronegativity ‘Xeq’, molecular connectivity $\chi^b$ and hydrophobicity 'logP' on a novel series of $\omega$-amino alkanolic acid derivatives of betulinic acid, have been discussed. Anti HIV-1 activity of these compounds both in CEM-4 and MT-4 cell cultures is found to correlate well with steric $\chi^b$ as well as electronic Xeq parameters. The presence of an amide function is found to be important for activity.

HIV has drawn global attention of both medical practitioners and medicinal chemists. Glycyrrhetinic acids have shown limited activity against a wide spectrum of viruses including HIV-1 type-I. Salaspermic acid 2 and suberol 3 are known to inhibit HIV-1 in H9 cells. Bile acid derivatives 4 were found slightly active (at 10$^{-6}$ M) against HIV in MT-4 cells. Francoise Soler and coworkers 5,6 have shown that the derivatives of betulinic acid are potent inhibitors of the cytopathicity of HIV-1 both in CEM-4 and MT-4 cells. For the synthesis of these derivatives, betulinic acid (Figure 1) has been used as the common starting material. This new class of anti HIV-1 compounds seems to block a post binding event involved in the virus-cell fusion.

For the purpose of present study the betulinic acid derivatives have been divided into two series:
(a) $\omega$-amino alkanolic acid amides of betulinic acids (Table I).
(b) (Betulinyl amino)- octanoic acid amides of $\alpha,\beta$ and $\gamma$ amino acids (Table II).

Figure 1 – Betulinic acid

The electronic parameter i.e. equalized electronegativity ‘Xeq’ was evaluated using Pauling 7 formula, the steric parameter first order molecular connectivity $\chi^b$ was calculated as defined by Kupchik 8,9,10 and the partition coefficient 'logP' was calculated by the fragmental method developed by Nys and Rekker 11.

Results and Discussion

First QSAR studies on $\omega$-amino alkanolic acid amides of betulinic acid derivatives are discussed.

All the eighteen analogues in this series are listed in Table I along with their physicochemical data and biological activity. In this Table I log(1/IC$_{50}$) i.e. p(1) represents the anti HIV-1 activity against CEM-4 cells while p(2) represents activity against MT-4 cells. IC$_{50}$ (50% cell culture inhibitory concentrations) are defined as those which inhibit HIV-1 induced cytopathicity by 50%.

Table III contains the correlation matrix with p(1)activity showing correlation among the structural parameters in this series of compounds and it is clear that auto correlation exists between Xeq and logP.

Regression analysis of p(1) with Xeq, $\chi^b$ and logP gave the following significant correlation,

$$p(1) = 0.587 (\pm 0.178) \chi^b - 8.819$$

$$n = 18, \text{EV} = 40.4%, \text{SEE} = 0.674, F(1,16) = 10.840$$

where 'n' represents the number of data points, 'r' the correlation coefficient, 'F' the variance ratio, 'SEE' the standard error of estimate and data within the parentheses represents the confidence interval of regression constant at 95% level.

Introduction of an indicator parameter ‘Ind’ whose value was taken 1 for the presence of -NHCO group at Y and zero otherwise, resulted in the following equations,

$$p(1) = 0.493 (\pm 0.162) \chi^b + 0.918 (\pm 0.387) \text{Ind} - 7.197$$

$$n = 18, \text{EV} = 56.6%, \text{SEE} = 0.593, F(2,15) = 9.80$$

$$p(1) = 0.456 (\pm 0.158) \chi^b + 0.169 (\pm 0.119) \text{logP} + 1.011 (\pm 0.380) \text{Ind} - 7.221$$

$$n = 18, \text{EV} = 62.1%, \text{SEE} = 0.574, F(2,15) = 7.656$$

Correlation of p(1), with $\chi^b$ and logP, gave a statistically acceptable regression equation (3) with 62.1%
Table I—Biological activities and physico-chemical data for o-amino alkanoic acid amide of betulinic acid derivatives

<table>
<thead>
<tr>
<th>Compd</th>
<th>m</th>
<th>n</th>
<th>Y</th>
<th>Xeq</th>
<th>$\chi^b$</th>
<th>logP</th>
<th>p((l_1))</th>
<th>p((l_2))</th>
<th>p((l_2)) Calculated</th>
<th>p((l_2)) Calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>5</td>
<td>0</td>
<td>—</td>
<td>2.240</td>
<td>16.900</td>
<td>2.700</td>
<td>1.639</td>
<td>0.620</td>
<td>1.140</td>
<td>1.088</td>
</tr>
<tr>
<td>A2</td>
<td>6</td>
<td>0</td>
<td>—</td>
<td>2.239</td>
<td>17.400</td>
<td>3.240</td>
<td>0.921</td>
<td>0.367</td>
<td>1.136</td>
<td>1.412</td>
</tr>
<tr>
<td>A3</td>
<td>7</td>
<td>0</td>
<td>—</td>
<td>2.239</td>
<td>17.900</td>
<td>3.780</td>
<td>2.125</td>
<td>1.333</td>
<td>1.633</td>
<td>1.736</td>
</tr>
<tr>
<td>A4</td>
<td>8</td>
<td>0</td>
<td>—</td>
<td>2.238</td>
<td>18.400</td>
<td>4.320</td>
<td>2.377</td>
<td>0.469</td>
<td>1.880</td>
<td>2.059</td>
</tr>
<tr>
<td>A5</td>
<td>9</td>
<td>0</td>
<td>—</td>
<td>2.237</td>
<td>18.900</td>
<td>4.860</td>
<td>2.222</td>
<td>1.377</td>
<td>2.126</td>
<td>2.383</td>
</tr>
<tr>
<td>A6</td>
<td>10</td>
<td>0</td>
<td>—</td>
<td>2.237</td>
<td>19.400</td>
<td>5.400</td>
<td>2.639</td>
<td>1.354</td>
<td>2.373</td>
<td>2.707</td>
</tr>
<tr>
<td>A7</td>
<td>11</td>
<td>0</td>
<td>—</td>
<td>2.236</td>
<td>19.900</td>
<td>5.940</td>
<td>2.260</td>
<td>1.603</td>
<td>2.620</td>
<td>3.031</td>
</tr>
<tr>
<td>A8</td>
<td>3</td>
<td>5</td>
<td>CONH</td>
<td>2.281</td>
<td>19.003</td>
<td>1.976</td>
<td>1.260</td>
<td></td>
<td>2.177</td>
<td>1.942</td>
</tr>
<tr>
<td>A9</td>
<td>4</td>
<td>4</td>
<td>CONH</td>
<td>2.281</td>
<td>19.003</td>
<td>1.490</td>
<td>1.347</td>
<td></td>
<td>2.177</td>
<td>1.859</td>
</tr>
<tr>
<td>A10</td>
<td>5</td>
<td>3</td>
<td>CONH</td>
<td>2.281</td>
<td>19.003</td>
<td>1.872</td>
<td>1.523</td>
<td></td>
<td>2.177</td>
<td>1.924</td>
</tr>
<tr>
<td>A11</td>
<td>6</td>
<td>2</td>
<td>CONH</td>
<td>2.281</td>
<td>19.003</td>
<td>2.483</td>
<td>1.347</td>
<td></td>
<td>2.177</td>
<td>2.028</td>
</tr>
<tr>
<td>A12</td>
<td>7</td>
<td>1</td>
<td>CONH</td>
<td>2.281</td>
<td>19.003</td>
<td>3.094</td>
<td>2.284</td>
<td>1.807</td>
<td>2.177</td>
<td>2.132</td>
</tr>
<tr>
<td>A13</td>
<td>7</td>
<td>2</td>
<td>CONH</td>
<td>2.280</td>
<td>19.503</td>
<td>3.023</td>
<td>3.000</td>
<td>2.207</td>
<td>2.424</td>
<td>2.352</td>
</tr>
<tr>
<td>A14</td>
<td>7</td>
<td>3</td>
<td>CONH</td>
<td>2.278</td>
<td>20.003</td>
<td>2.952</td>
<td>3.357</td>
<td>1.890</td>
<td>2.671</td>
<td>2.572</td>
</tr>
<tr>
<td>A15</td>
<td>7</td>
<td>4</td>
<td>CONH</td>
<td>2.277</td>
<td>20.503</td>
<td>3.110</td>
<td>3.215</td>
<td></td>
<td>2.917</td>
<td>2.831</td>
</tr>
<tr>
<td>A16</td>
<td>7</td>
<td>1</td>
<td>NHC O</td>
<td>2.281</td>
<td>19.003</td>
<td>3.094</td>
<td>3.302</td>
<td>2.469</td>
<td>3.095</td>
<td>3.143</td>
</tr>
<tr>
<td>A17</td>
<td>7</td>
<td>2</td>
<td>NHC O</td>
<td>2.280</td>
<td>19.503</td>
<td>3.023</td>
<td>3.023</td>
<td>2.276</td>
<td>3.341</td>
<td>3.363</td>
</tr>
<tr>
<td>A18</td>
<td>7</td>
<td>3</td>
<td>NHC O</td>
<td>2.278</td>
<td>20.003</td>
<td>2.570</td>
<td>3.699</td>
<td></td>
<td>3.588</td>
<td>3.518</td>
</tr>
</tbody>
</table>

Note: Cal. $p_1$ and Cal $p_{12}$ corresponding to the activities calculated from eqns.2, 3, 4 & 6 respectively.

Q SAR on the data given in the Table II gave the following regression equations,

$$p(\(l_2\)) = 25.051 (\pm 6.235) \text{Xeq} - 55.036 \quad \ldots (4)$$

$$n = 12, \text{EV} = 61.8\%, \text{SEE} = 0.448, F_{(1,10)} = 16.141$$

$$p(\(l_2\)) = 0.516 (\pm 0.151) \text{Xeq} - 8.242 \quad \ldots (5)$$

$$n = 12, \text{EV} = 53.9\%, \text{SEE} = 0.492, F_{(1,10)} = 11.707$$

Correlation level was found to improve when both Xeq and $\chi^b$ were considered together.

$$p(\(l_2\)) = 17.969 (\pm 5.613) \text{Xeq} + 0.327 (\pm 0.124) \text{Xeq}^2 - 45.217 \quad \ldots (6)$$

$$n = 12, \text{EV} = 78.5\%, \text{SEE} = 0.354, F_{(2,9)} = 16.391$$

Equations (4), (5) and (6) make it clear that the first order molecular connectivity $\text{Xeq}$ and equalized electronegativity Xeq play an important role in determining the activity. Incremental lengthening of the chain in compounds (A1-A7) also shows that significant activity was observed for the compounds positioned between A3 to A7 i.e. between the betulinylamino octanoic and (betulinylamino) dodecanoic acids.
For the purpose of performing QSAR studies on the (betulinylamino)octanoic acid amides of α-β- and γ-amino acids, the calculated physicochemical data and anti HIV-1 activity $p(I_3) = [p(\log IC_{50})$ against MT-4 cells] have been listed in Table II. The correlation matrix for the structural parameters in Table V shows the absence of any autocorrelation between $'\chi^b$, Xeq & logP.

The regression analysis gave the following statistically significant relationship between $p(I_3)$ and Xeq.

$$p(I_3) = 8.334 (\pm1.179) \text{Xeq} - 17.715 \quad \ldots (7)$$

$n = 11$, $EV = 84.7\%$, $SEE = 0.139$, $F_{(2,8)} = 49.904$

This correlation explains 84.7% variation in the activity. The standard error of estimate is low and F-ratio is fairly high. The ‘t-test’ also suggest that this relationship is significant at 95% confidence level ($|t_{cal}| > t_{0.05} = 2.262$).

When for the compounds in which CH(OH)CH$_2$COOH group is present in variant position (RNHR') of the substituent (comps. B7, B8, B10 & B11) an indicator parameter Ind1 was assigned a value 1 otherwise it was treated as 0, a highly significant regression equation was obtained.

$$p(I_3) = 7.132 (\pm0.752) \text{Xeq} + 0.234 (\pm0.055) \text{Ind}_1 - 15.078 \quad \ldots (8)$$

$n = 11$, $EV = 95.3\%$, $SEE = 0.082$, $F_{(2,8)} = 80.635$
For this equation explained variance is very high at 95.3%, standard error of estimate SEE is very low and F-ratio is significant at 95% confidence level.

It can therefore be suggested that $X_{eq}$ plays a dominant role in determining the activity against MT-4 cells. In future designing of more potent drugs in this class one may look for the compounds with higher equalized electronegativity having \(-\text{CH(OH)}\text{CH}_2\text{COOH}\) group in variant portion (RNHR') of the substituent.

Acknowledgement
This research was supported by Grant (1648)/00/EMR II awarded by CSIR, New Delhi.

References
8 Kupchick E J, Quant Struct Act Relat, 4, 1985, 123.