Novel Semisynthetic Derivatives of Bile Acids as Effective Tyrosyl-DNA Phosphodiesterase 1 Inhibitors

Oksana V. Salomatina1, Irina I. Popadyuk1, Alexandra L. Zakharenko2, Olga D. Zakharova2, Dmitriy S. Fadeev1, Nina I. Komarova1, Jóhannes Reynisson3, H. John Arabshahi3, Raina Chand3, Konstantin P. Volcho1,4, Nariman F. Salakhutdinov1,4, Olga I. Lavrik2,4

1 N.N. Vorozhtsov Novosibirsk Institute of Organic Chemistry, SB RAS, 630090, Russia, Novosibirsk, acad. Lavrentjev ave. 9; ana@nioch.nsc.ru (O.V.S.), popadyuk@nioch.nsc.ru (I.I.P.), dsf@nioch.nsc.ru (D.S.F.), komar@nioch.nsc.ru (N.I.K)

2 Novosibirsk Institute of Chemical Biology and Fundamental Medicine, SB RAS, 630090, Russia, Novosibirsk, acad. Lavrentjev ave. 8; sasha@niboch.nsc.ru (A.L.Z), isar@niboch.nsc.ru (O.D.Z.)

3 School of Chemical Sciences, University of Auckland, New Zealand, j.reynisson@auckland.ac.nz (J.R.), j.arabshahi@auckland.ac.nz (H.J.A.), rcha387@aucklanduni.ac.nz (R.Ch.)

4 Novosibirsk State University, 630090, Russia, Novosibirsk, Pirogova str. 2; anvar@nioch.nsc.ru (N.F.S.)

lavrik@niboch.nsc.ru (O.I.L.)

* Correspondence: volcho@nioch.nsc.ru; Tel.: +7 383 3308870

ELECTRONIC SUPPORTING INFORMATION

Virtual screening & molecular modeling data………………………………………………………………………………………………………2
NMR 1H and 13C spectra of new compounds………………………………………………………………………………………………………5
Table S1. The scores predicted by the four scoring functions for compounds 1a,b, 2a,b, 3a-d, 4a,b-8a,b and 9a-11a docked in Tdp1.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Scoring Function</th>
<th>IC₅₀</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLP</td>
<td>GS</td>
</tr>
<tr>
<td>1a</td>
<td>71.8</td>
<td>56.5</td>
</tr>
<tr>
<td>2a</td>
<td>70.4</td>
<td>58.6</td>
</tr>
<tr>
<td>3a</td>
<td>80.8</td>
<td>60.8</td>
</tr>
<tr>
<td>3c</td>
<td>71.9</td>
<td>49.7</td>
</tr>
<tr>
<td>Fur</td>
<td>45.3</td>
<td>47.6</td>
</tr>
<tr>
<td>1b</td>
<td>75.2</td>
<td>44.6</td>
</tr>
<tr>
<td>2b</td>
<td>74.8</td>
<td>59.5</td>
</tr>
<tr>
<td>3b</td>
<td>80.0</td>
<td>63.5</td>
</tr>
<tr>
<td>3d</td>
<td>72.7</td>
<td>67.9</td>
</tr>
<tr>
<td>4a</td>
<td>41.5</td>
<td>28.2</td>
</tr>
<tr>
<td>4b</td>
<td>65.9</td>
<td>41.9</td>
</tr>
<tr>
<td>5a</td>
<td>62.1</td>
<td>53.9</td>
</tr>
<tr>
<td>5b</td>
<td>66.4</td>
<td>49.7</td>
</tr>
<tr>
<td>6a</td>
<td>64.7</td>
<td>40.3</td>
</tr>
<tr>
<td>6b</td>
<td>66.5</td>
<td>45.8</td>
</tr>
<tr>
<td>7a</td>
<td>63.3</td>
<td>59.4</td>
</tr>
<tr>
<td>7b</td>
<td>63.4</td>
<td>50.1</td>
</tr>
<tr>
<td>8a</td>
<td>64.5</td>
<td>60.2</td>
</tr>
<tr>
<td>8b</td>
<td>63.5</td>
<td>47.3</td>
</tr>
<tr>
<td>9a</td>
<td>69.4</td>
<td>59.1</td>
</tr>
<tr>
<td>10a</td>
<td>56.9</td>
<td>56.2</td>
</tr>
<tr>
<td>11a</td>
<td>56.5</td>
<td>61.9</td>
</tr>
</tbody>
</table>
Table S2. The calculated molecular descriptors for the bile acid derivatives.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Molecular Descriptor</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mol. weight</td>
<td>Log P</td>
<td>TPSA</td>
<td>H bond donor</td>
<td>H bond acceptor</td>
<td>Rotatable bond</td>
</tr>
<tr>
<td>1a</td>
<td>618.9</td>
<td>5.5</td>
<td>97.5</td>
<td>2</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>2a</td>
<td>618.9</td>
<td>5.5</td>
<td>97.5</td>
<td>2</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>3a</td>
<td>618.9</td>
<td>5.3</td>
<td>97.5</td>
<td>2</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>3c</td>
<td>576.8</td>
<td>5.2</td>
<td>91.4</td>
<td>3</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Fur</td>
<td>304.4</td>
<td>2.9</td>
<td>112.9</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>1b</td>
<td>534.8</td>
<td>5.0</td>
<td>85.4</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>2b</td>
<td>534.8</td>
<td>5.0</td>
<td>85.4</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>3b</td>
<td>534.8</td>
<td>5.0</td>
<td>85.4</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>3d</td>
<td>576.8</td>
<td>5.2</td>
<td>91.4</td>
<td>3</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>4a</td>
<td>551.8</td>
<td>5.7</td>
<td>81.7</td>
<td>1</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>4b</td>
<td>467.7</td>
<td>5.4</td>
<td>69.6</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>5a</td>
<td>630.6</td>
<td>6.5</td>
<td>81.7</td>
<td>1</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>5b</td>
<td>546.6</td>
<td>6.2</td>
<td>69.5</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>6a</td>
<td>565.8</td>
<td>6.1</td>
<td>81.7</td>
<td>1</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>6b</td>
<td>481.7</td>
<td>5.9</td>
<td>69.6</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>7a</td>
<td>552.8</td>
<td>4.4</td>
<td>94.6</td>
<td>1</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>7b</td>
<td>468.7</td>
<td>4.1</td>
<td>82.5</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8a</td>
<td>609.9</td>
<td>5.7</td>
<td>81.7</td>
<td>1</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>8b</td>
<td>525.8</td>
<td>5.4</td>
<td>69.6</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>9a</td>
<td>708.0</td>
<td>9.0</td>
<td>101.9</td>
<td>2</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>10a</td>
<td>588.9</td>
<td>4.7</td>
<td>84.9</td>
<td>1</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>11a</td>
<td>589.8</td>
<td>4.0</td>
<td>94.2</td>
<td>1</td>
<td>5</td>
<td>12</td>
</tr>
</tbody>
</table>
Table S3. Criteria of lead-like, drug-like and known drug space (KDS) in terms of molecular descriptors.

<table>
<thead>
<tr>
<th></th>
<th>Lead-like Space</th>
<th>Drug-like Space</th>
<th>Known Drug Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular weight (g mol(^{-1}))</td>
<td>300</td>
<td>500</td>
<td>800</td>
</tr>
<tr>
<td>Lipophilicity (Log P)</td>
<td>3</td>
<td>5</td>
<td>6.5</td>
</tr>
<tr>
<td>Hydrogen bond donors (HD)</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Hydrogen bond acceptors (HA)</td>
<td>3</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Polar surface area (Å(^2)) (PSA)</td>
<td>60</td>
<td>140</td>
<td>180</td>
</tr>
<tr>
<td>Rotatable bonds (RB)</td>
<td>3</td>
<td>10</td>
<td>17</td>
</tr>
</tbody>
</table>

Spectra of Compound 1, $^1$H NMR, 300MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 75MHz, CDCl$_3$ (top)
Spectra of Compound 2, $^1$H NMR, 500MHz, CDCl$_3$ (bottom), $^{13}$C NMR, JMOD, 125MHz, CDCl$_3$ (top)
Spectra of Compound 3, $^1$H NMR, 300MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 75MHz, CDCl$_3$ (top)
Spectra of Compound 1a, $^1$H NMR, 400MHz, CDCl₃ (bottom); $^{13}$C NMR, JMOD, 125MHz, CDCl₃ (top)
Spectra of Compound 1b, $^1$H NMR, 400MHz, CDCl$_3$ (bottom), $^{13}$C NMR, BB, 100MHz, CDCl$_3$ (top)
Spectra of Compound 2a, $^1$H NMR, 500MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 125MHz, CDCl$_3$ (top)
Spectra of Compound 2b, $^1$H NMR, 500MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 125MHz, CDCl$_3$ (top)
Spectra of Compound 3a, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)
Spectra of Compound 3b, $^1$H NMR, 300MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 75MHz, CDCl$_3$ (top)
Spectra of Compound 3c, $^1$H NMR, 300MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 75MHz, CDCl$_3$ (top)
Spectra of Compound 3d, ^1^H NMR, 300MHz, CDCl₃ (bottom); ^1^C NMR, JMOD, 75MHz, CDCl₃ (top)
Spectra of Compound 4a, $^1$H NMR, 500MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 125MHz, CDCl$_3$ (top)
Spectra of Compound 4b, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)
Spectra of Compound 5a, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)
Spectra of Compound 5b, $^1$H NMR, 400MHz, CDCl$_3$+CD$_3$OD (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$+CD$_3$OD (top)
Spectra of Compound 6a, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)
Spectra of Compound 6b, $^1$H NMR, 300MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 75MHz, CDCl$_3$ (top)
Spectra of Compound 7a, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)
Spectra of Compound 7b, $^1$H NMR, 400MHz, CDCl$_3$+CD$_3$OD (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$+CD$_3$OD (top)
Spectra of Compound 8a, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)
Spectra of Compound 8b, $^1$H NMR, 400MHz, CDCl$_3$+CD$_3$OD (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$+CD$_3$OD (top)
Spectra of Compound 9a, $^1$H NMR, 300MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 75MHz, CDCl$_3$ (top)
Spectra of Compound 10a, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)
Spectra of Compound 11a, $^1$H NMR, 400MHz, CDCl$_3$ (bottom); $^{13}$C NMR, JMOD, 100MHz, CDCl$_3$ (top)