Functionalization of 2′-Amino-LNA with additional nucleobases

Tadashi Umemotoa, Jesper Wengel a and Andreas Stahl Madsena

aNucleic Acid Center, Department of Physics and Chemistry, University of Southern Denmark, 5230, Odense M, Denmark.

Supporting Information

General experimental ........................................................................................................................................ S2
MALDI-MS of synthesized ONs (Table S1) ............................................................................................ S3
Full NMR assignment of 3a-3c........................................................................................................ S4
Spectra of 3a (1H NMR, 13C NMR) .................................................................................................. S7
Spectra of 3b (1H NMR, 13C NMR) .................................................................................................. S9
Spectra of 3c (1H NMR, 13C NMR) .................................................................................................. S11
Spectrum of 4a (31P NMR) ............................................................................................................. S13
Spectrum of 4b (31P NMR) ............................................................................................................. S14
Spectrum of 4c (31P NMR) ............................................................................................................. S15
General experimental
All reagents and solvents were of analytical grade and obtained from commercial suppliers and used without further purification except for dichloromethane, which was distilled prior to use. Petroleum ether of the distillation range 60-80 °C was used. Acetonitrile was dried through storage over activated 3Å molecular sieves. Anhydrous dichloromethane, 1,2-dichloroethane, and N,N’-diisopropylethylamine were dried through storage over activated 4Å molecular sieves. All reactions were conducted under an atmosphere of argon, and were monitored by thin-layer chromatography (TLC) using silica gel coated plates with fluorescence indicator (SiO₂-60, F-254) which were visualized a) under UV light, and b) by dipping in a solution of molybdate-phosphoric acid (12.5 g/L) and cerium(IV)sulfate (5 g/L) in 3% conc. sulfuric acid in water (v/v) followed by heating. Silica gel column chromatography was performed with Silica gel 60 (particle size 0.040–0.063 mm, Merck) using moderate pressure (pressure ball). Silica gel columns were generally built with an initial starting eluent containing 1% (v/v) of triethylamine or pyridine. Evaporation of solvents was carried out under reduced pressure with a temperature not exceeding 50 °C. After column chromatography, appropriate fractions were pooled, evaporated and dried at high vacuum for at least 12 h to give obtained products in high purity (>95%), unless stated otherwise. ¹H NMR, ¹³C NMR and/or ³¹P NMR ascertained sample purity. Chemical shifts are reported in parts per million (ppm) relative to tetramethylsilane or deuterated solvent as the internal standard (δH: DMSO-d₆ 2.50 ppm; δC: DMSO-d₆ 39.51 ppm) or an external standard (δP: 85% H₃PO₄ 0.00 ppm). Exchangeable (ex) protons were detected by disappearance of peaks on D₂O addition. Assignments of NMR spectra are based on 2D spectra (COSY, HSQC, and HMBC) and follow standard carbohydrate/nucleoside nomenclature. The carbon atom of C4’ substituents is numbered C-5” in nucleoside derivatives. Similar conventions apply for the corresponding hydrogen atoms.
Numbering of nucleobases of N2′-functionalities is given with triple primes (e.g. H6‴). MALDI-HRMS were recorded in positive ion mode on an IonSpec Fourier transform mass spectrometer.

**Table S1. MALDI-MS of synthesized ONs.**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Found m/z [M-H]</th>
<th>Calc. m/z [M-H]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'-GTG AAT$_T^L$ AGC C</td>
<td>3251</td>
<td>3245</td>
</tr>
<tr>
<td>5'-GTG AAT$_A^L$ AGC C</td>
<td>3256</td>
<td>3254</td>
</tr>
<tr>
<td>5'-GTG AAT$_{ph}^L$ AGC C</td>
<td>3199</td>
<td>3197</td>
</tr>
<tr>
<td>5'-GGC TAT$_T^L$ TCA C</td>
<td>3194</td>
<td>3196</td>
</tr>
<tr>
<td>5'-GGC TAT$_A^L$ TCA C</td>
<td>3203</td>
<td>3205</td>
</tr>
<tr>
<td>5'-GGC TAT$_{ph}^L$ TCA C</td>
<td>3150</td>
<td>3148</td>
</tr>
<tr>
<td>5'-GGC T$_T^L$ AT TCA C</td>
<td>3193</td>
<td>3196</td>
</tr>
<tr>
<td>5'-GGC T$_A^L$ AT TCA C</td>
<td>3202</td>
<td>3205</td>
</tr>
<tr>
<td>5'-GGC T$_{ph}^L$ AT TCA C</td>
<td>3148</td>
<td>3148</td>
</tr>
<tr>
<td>5'-GTG AAT$_T^L$ T$_T^L$GC C</td>
<td>3426</td>
<td>3429</td>
</tr>
<tr>
<td>5'-GAC GT$_T^L$T$_A^L$ T$_A^L$T$_T^L$T$_A^L$ T$_T^L$T$_T^L$T$_A^L$ GCA C</td>
<td>6437</td>
<td>6434</td>
</tr>
</tbody>
</table>

---

a For the structure of the monomers see Fig. 1.
Full NMR assignment

1-(2-Amino-2-deoxy-5-O-4,4′-dimethoxytrityl-2-N,4-C-methylene-2-N-(thymin-1-ylacetyl)-β-D-ribofuranosyl)thymine (3a).

$^1$H NMR (400 MHz, DMSO-$d_6$): $\delta$ 11.50 (br s, 1.5H, ex, N3-H$_A$), 11.39 (br s, 1H, ex, N3-H$_B$), 11.31 (br s, 2.5H, ex, N3‴-H$_{A+B}$), 7.54 (s, 1H, H6$_B$), 7.53 (s, 1.5H, H6$_A$), 7.22-7.47 (m, 25H, H6‴$′$=H$_{A+B}$, H2$_{DMT,A+B}$, H6$_{DMT,A+B}$, H2′$_{DMT,A+B}$, H6′$_{DMT,A+B}$, H2″$′$=H$_{A+B}$, H6″$′$=H$_{A+B}$, H5″$′$=H$_{A+B}$), 6.90-6.95 (m, 10H, H3$_{DMT,A+B}$, H5′$_{DMT,A+B}$, H3″$′$=H$_{A+B}$, H5″$′$=H$_{A+B}$), 6.09 (d, $J = 4.2$ Hz, 1.5H, ex, 3′-OH$_A$), 6.04 (d, $J = 4.3$ Hz, 1H, ex, 3′-OH$_B$), 5.63 (s, 1.5H, H1′$'_A$), 5.44 (s, 1H, H1′$'_B$), 4.76 (d, $J = 16.7$, 1.5H, COCH$_2$-$A$, 4.69 (s, 1H, H2′$_B$), 4.57 (d, $J = 16.7$, 1.5H, COCH$_2$-$A$), 4.51 (s, 1.5H, H2′$_A$), 4.49 (s, 2H, COCH$_2$-$B$), 4.29 (d, $J = 4.2$ Hz, 1.5H, H3′$'_A$), 4.26 (d, $J = 4.3$ Hz, 1H, H3′$'_B$), 3.75 (s, 15H, OCH$_3$), 3.57 (s, 2H, COCH$_3$), 3.48 (d, $J = 11.1$ Hz, 1H, H5′$'_B$), 3.47 (d, $J = 11.0$ Hz, 1.5H, H5′$'_A$), 3.32-3.42 (m, 5H, H5″$′$=H$_{A+B}$), 1.76 (s, 7.5H, 5‴-CH$_{3,A+B}$), 1.51 (s, 4.5H, 5-CH$_3$-$A$), 1.49 (s, 3H, 5-CH$_3$-$B$).

$^{13}$C NMR (101 MHz, DMSO-$d_6$): $\delta$ 165.9 (COCH$_2$-$A$), 165.6 (COCH$_2$-$B$), 164.3 (C4‴$″$=C$″$-$A+B$), 163.8 (C4$″$-$A$), 163.7 (C4$″$-$B$), 158.2 (C4$_{DMT,A+B}$, C4$′$$_{DMT,A+B}$), 158.1 (C4$″$$_{DMT,A+B}$, C4$″′$$_{DMT,A+B}$), 150.9 (C2‴$″$=C$″$-$A+B$), 149.9 (C2$″$-$A$), 149.7 (C2$″$-$B$), 144.53 (C1″$′$-$DMT,A$), 144.47 (C1″$′$-$DMT,B$), 142.1 (C6‴$″$-$A$), 142.0 (C6‴$″$-$B$), 135.2 (C1$″′$-$DMT,A$), 135.1 (C1$″′$-$DMT,B$), 135.0 (C1$″′$-$DMT,A$), 134.9 (C1$″′$-$DMT,B$), 134.3 (C6$″$-$A$), 134.0 (C6$″$-$B$), 129.74 (C2$_{DMT,A+B}$/C6$_{DMT,A+B}$/C2′″$′$-$DMT,A+B$/C6″$′$-$DMT,A+B$), 129.66 (C2″$′$-$DMT,A+B$/C6″$′$-$DMT,A+B$), 127.9 (C3‴$″$=C$″$-$A+B$), 127.6 (C2″$′$-$DMT,A+B$/C6″$′$-$DMT,A+B$), 126.8 (C4″$″$-$DMT,A+B$), 113.22 (C3″$″$-$DMT,A+B$/C5″$″$-$DMT,A+B$/C5″$″$-$DMT,A+B$), 113.16 (C3″$″$-$DMT,A+B$/C5″$″$-$DMT,A+B$/C3″$″$-$DMT,A+B$/C5″$″$-$DMT,A+B$), 108.6 (C5$″$-$A$), 108.5 (C5$″$-$B$), 108.1 (C5″$″$-$A$), 108.0 (C5″$″$-$B$), 87.8 (C4$″$-$B$), 86.9 (C4$″$-$A$), 86.2 (C1$″$-$A$), 86.0 (C1$″$-$B$), 85.9 (Ar$_3$CO$_B$), 85.7 (Ar$_3$CO$_A$), 69.4 (C3‴$″$=$A$), 68.2 (C3‴$″$=$B$), 62.2 (C2‴$″$=$A$), 61.1 (C2‴$″$=$B$), 59.2 (C5‴$″$=$A$), 55.0 (OCH$_3$-$A$+$B$), 51.5 (C5‴$″$=$A$), 50.7 (C5‴$″$=$B$), 48.3 (COCH$_2$-$A$), 47.9 (COCH$_2$-$B$), 12.21 (5-CH$_3$-$A$), 12.18 (5-CH$_3$-$B$), 11.81 (5‴-CH$_3$-$B$), 11.77 (5‴-CH$_3$-$A$).
1-(2-Amino-2-N-(6-N-benzoyladenin-9-ylacetyl)-2-deoxy-5-O-4,4′-dimethoxytrityl-2-N,4-C-methylene-β-D-ribofuranosyl)thymine (3b).

$^1$H NMR (DMSO-$d_6$): $\delta$ 11.56 (br s, 1H, ex, NH$_A$/6‴″-NH$_A$), 11.39 (br s, 1H, ex, NH$_B$/6‴″-NH$_B$), 11.17 (br s, 2.3H, ex, NH$_A$/B/6‴″-NH$_A$/B), 8.72 (s, 2.3H, H$_2$″/A+B), 8.47 (s, 1.3H, H8‴″/A), 8.38 (s, 1H, H8‴″/B), 8.07 (d, $J = 7.5$ Hz, 4.6H, H$_2$/Bz,A+B, H$_6$/Bz,A+B), 7.62-7.67 (m, 2.3H, H$_4$/Bz,A+B), 7.53-7.59 (m, 6.9H, H$_3$/Bz,A+B, H$_5$/Bz,A+B, H$_6$/A+B), 7.24-7.52 (m, 20.7H, H$_2$/DMT,A+B, H$_6$/DMT,A+B, H$_2′$/DMT,A+B, H$_6′$/DMT,A+B, H$_2′$/DMT,A+B, H$_3′$/DMT,A+B, H$_4′$/DMT,A+B, H$_5′$/DMT,A+B, H$_6′$/DMT,A+B), 6.91-6.96 (m, 9.2H, H$_3$/DMT,A+B, H$_5$/DMT,A+B, H$_3′$/DMT,A+B, H$_5′$/DMT,A+B), 6.18 (d, $J = 4.3$ Hz, 1.3H, ex, 3′-OH$_A$), 6.08 (d, $J = 4.3$ Hz, 1H, ex, 3′-OH$_B$), 5.75 (s, 1.3H, H1‴/A), 5.53 (s, 1H, H1‴/B), 5.48 (d, $J = 17.0$ Hz, 1.3H, COCH$_2$/A,a), 5.30 (d, $J = 17.0$ Hz, 1.3H, COCH$_2$/A,b), 5.28 (d, $J = 17.2$ Hz, 1H, COCH$_2$/B,b), 5.20 (d, $J = 17.2$ Hz, 1H, COCH$_2$/B,a), 4.73 (s, 1.3H, H2‴/A), 4.72 (s, 1H, H2‴/B), 4.36 (d, $J = 4.3$ Hz, 1.3H, H3‴/A), 4.28 (d, $J = 4.3$ Hz, 1H, H3‴/B), 3.74-3.77 (m, 15.8H, OCH$_3$/A+B, H5‴″/B), 3.36-3.55 (m, 7.2H, H5‴/A+B, H5‴″/A), 1.54 (s, 3.9H, 5-CH$_3$/A), 1.51 (s, 3H, 5-CH$_3$/B).

$^{13}$C NMR (DMSO-$d_6$): $\delta$ 165.5 (COPh$_{A+B}$), 165.2 (COCH$_2$/A), 165.1 (COCH$_2$/B), 163.9 (C4‴/A), 163.7 (C4‴/B), 158.24 (C4′/DMT,A+B, C4‴/DMT,A+B), 158.21 (C4′/DMT,A+B, C4‴/DMT,A+B), 158.1 (C4′/DMT,A+B, C4‴/DMT,A+B), 152.8 (C4‴″/A,B), 151.5 (C2‴″/A,B), 151.4 (C2‴″/A,B), 150.1 (C2‴/C2‴/C6‴″/A,A,B), 150.0 (C2‴/C2‴/C6‴″/A,A,B), 149.8 (C2‴/C2‴/C6‴″/A,B), 145.6 (C8‴″/A), 145.5 (C8‴″/B), 144.61 (C1‴/DMT,A), 144.56 (C1‴/DMT,B), 135.3 (C1‴/DMT,A/C1‴/DMT,A), 135.2 (C1‴/DMT,B/C1‴/DMT,B), 135.02 (C1‴/DMT,A/C1‴/DMT,B), 134.99 (C1‴/DMT,B/C1‴/DMT,B), 134.3 (C6‴/B), 134.1 (C6‴/A), 133.4 (C1‴/Bz,A+B), 132.3 (C4‴/Bz,A+B), 129.8 (C2‴/DMT,A+B/C6‴/DMT,A+B/C2‴/DMT,A+B/C6‴/DMT,A+B), 129.7 (C2‴/DMT,A+B/C6‴/DMT,A+B/C2‴/DMT,A+B/C6‴/DMT,A+B), 128.4 (C2‴/Bz,A+B, C3‴/Bz,A+B, C5‴/Bz,A+B, C6‴/Bz,A+B), 128.0 (C3‴/DMT,A+B, C5‴″/DMT,A+B), 127.7 (C2‴″/DMT,A,B, C6‴″/DMT,A,B), 126.9 (C4‴/DMT,A,B), 124.9 (C5‴″/A,B), 113.3 (C3‴/DMT,A+B/C5‴/DMT,A+B/C3‴/DMT,A+B/C5‴″/DMT,A,B), 113.2 (C3‴/DMT,A+B/C5‴/DMT,A+B/C3‴/DMT,A+B/C5‴″/DMT,A,B), 108.7 (C5‴/A), 108.6 (C5‴/B), 88.0 (C4‴/B), 87.1 (C4‴/A), 86.4 (C1‴/A), 86.0 (C1‴/B), 85.9 (Ar3/COB), 85.8 (Ar3/COA), 69.5 (C3‴/A), 68.3 (C3‴/B), 62.5 (C2‴/A), 61.3 (C2‴/B), 59.4 (C5‴/A), 59.2 (C5‴/B), 55.1 (OCH$_3$/A+B), 51.6 (C5‴/A), 51.0 (C5‴/B), 44.6 (COCH$_2$/A), 44.3 (COCH$_2$/B), 12.29 (5-CH$_3$/A), 12.26 (5-CH$_3$/B).
1-(2-Amino-2-deoxy-5-O-4,4′-dimethoxytrityl-2-N-phenylacetyl-β-D-ribofuranosyl)thymine (3c).

$^1$H NMR (DMSO-$d_6$): δ 11.50 (br s, 1.5H, ex, NH$_A$), 11.39 (br s, 1H, ex, NH$_B$), 7.57 (s, 1H, H$_6B$), 7.54 (s, 1.5H, H$_6A$), 7.39-7.46 (m, 5H, H$_2''$DMT,A+B, H$_6''$DMT,A+B), 7.20-7.37 (m, 30H, H$_2$DMT,A+B, H$_6$DMT,A+B, H$_2''$DMT,A+B, H$_6''$DMT,A+B, H$_3''$DMT,A+B, H$_4$DMT,A+B, H$_5''$DMT,A+B, H$_2$Bz,A+B, H$_3$Bz,A+B, H$_5$Bz,A+B, H$_6$Bz,A+B), 6.88-6.94 (m, 10H, H$_3$DMT, H$_5$DMT, H$_3''$DMT, H$_5''$DMT), 5.99 (d, $J = 4.2$ Hz, 1.5H, ex, 3′-OH$_B$), 5.96 (d, $J = 4.1$ Hz, 1H, ex, 3′-OH$_B$), 5.53 (s, 1.5H, H$_1'A$), 5.40 (s, 1H, H$_1'B$), 4.72 (s, 1H, H$_2'B$), 4.59 (s, 1.5H, H$_2'A$), 4.26 (d, $J = 4.2$ Hz, 1.5H, H$_3'A$), 4.24 (d, $J = 4.1$ Hz, 1H, H$_3'B$), 3.77 (s, 3H, COCH$_2A$), 3.74 (s, 15H, OCH$_3$A+B), 3.63 (d, $J = 15.8$ Hz, 1H, COCH$_2B,a$), 3.59 (d, $J = 15.8$ Hz, 1H, COCH$_2B,b$), 3.57 (d, $J = 9.6$ Hz, 1.5H, H$_5''A,a$), 3.51 (d, $J = 9.6$ Hz, 1.5H, H$_5''A,b$), 3.32-3.48 (m, 7H, H$_5'A+B$, H$_5''B$) 1.51 (s, 4.5H, 5-CH$_3A$), 1.47 (s, 3H, 5-CH$_3B$).

$^{13}$C NMR (DMSO-$d_6$): δ 169.3 (COCH$_2A$), 169.2 (COCH$_2B$), 163.8 (C$_4A$), 163.7 (C$_4B$), 158.2 (C$_4$DMT,A+B, C$_4'$DMT,A+B), 158.1 (C$_4$DMT,A+B, C$_4'$DMT,A+B), 150.1 (C$_2A$), 149.8 (C$_2B$), 144.62 (C$_1''$DMT,A), 144.55 (C$_1''$DMT,B), 135.4 (C$_1$DMT,A/C$_1'$DMT,A/C$_1$Ph,A), 135.30 (C$_1$DMT,B/C$_1'$DMT,B/C$_1$Ph,B), 135.27 (C$_1$DMT,A/C$_1'$DMT,A/C$_1$Ph,A), 135.2 (C$_1$DMT,B/C$_1'$DMT,B/C$_1$Ph,B), 135.02 (C$_1$DMT,A/C$_1'$DMT,A/C$_1$Ph,A), 134.98 (C$_1$DMT,B/C$_1'$DMT,B/C$_1$Ph,B), 134.4 (C$_6B$), 134.1 (C$_6A$), 129.8 (C$_2$DMT,A+B/C$_6$DMT,A+B/C$_2'$DMT,A+B/C$_6'$DMT,A+B), 129.7 (C$_2$DMT,A+B/C$_6$DMT,A+B/C$_2'$DMT,A+B/C$_6'$DMT,A+B), 129.4 (C$_2$Ph,A/C$_3$Ph,A/C$_5$Ph,A/C$_6$Ph,A), 129.3 (C$_2$Ph,B/C$_3$Ph,B/C$_5$Ph,B/C$_6$Ph,B), 128.14 (C$_2$Ph,B/C$_3$Ph,B/C$_5$Ph,B/C$_6$Ph,B), 128.11 (C$_2$Ph,A/C$_3$Ph,A/C$_5$Ph,A/C$_6$Ph,A), 127.9 (C$_3''$DMT,A+B, C$_5''$DMT,A+B), 127.7 (C$_2''$DMT,A+B, C$_6''$DMT,A+B), 126.8 (C$_4$Ph,A+B), 126.4 (C$_4'$DMT,A), 126.3 (C$_4'$DMT,B), 113.3 (C$_3$DMT,A+B/C$_5$DMT,A+B/C$_3'$DMT,A+B/C$_5'$DMT,A+B), 113.2 (C$_3$DMT,A+B/C$_5$DMT,A+B/C$_3'$DMT,A+B/C$_5'$DMT,A+B), 108.6 (C$_5A$), 108.5 (C$_5B$), 87.9 (C$_4'B$), 87.3 (C$_4'A$), 86.8 (C$_1'A$), 86.2 (C$_1'B$), 85.8 (Ar$_3$CO$_B$), 85.7 (Ar$_3$CO$_A$), 69.3 (C$_3'A$), 68.1 (C$_3'B$), 63.5 (C$_2'A$), 60.8 (C$_2'B$), 59.3 (C$_5'A$), 59.2 (C$_5'B$), 55.0 (OCH$_3$A+B), 51.8 (C$_5''B$), 51.3 (C$_5''A$), 39.9 (COCH$_2A$+B, overlap with DMSO-$d_6$), 12.3 (5-CH$_3A$), 12.2 (5-CH$_3B$).
3a; $^1$H NMR
3a; $^{13}$C NMR
3b; $^1$H NMR
**3b; $^{13}$C NMR**

![Chemical Structure](image)
3c: $^1$H NMR
3c; $^{13}$C NMR

![Chemical Structure](image)

Supplementary Material (ESI) for Organic & Biomolecular Chemistry
This journal is (c) The Royal Society of Chemistry 2009
4a; $^{31}$P NMR
$^{31}$P NMR

Supplementary Material (ESI) for Organic & Biomolecular Chemistry
This journal is (c) The Royal Society of Chemistry 2009
Supplementary Material (ESI) for Organic & Biomolecular Chemistry
This journal is (c) The Royal Society of Chemistry 2009

4c; $^{31}$P NMR

![Chemical Structure](image)

![NMR Spectrogram](image)

S15