Supporting Information

Synthesis of C4-linked C₀- and C₂-imidazole 2'-deoxyribonucleoside phosphoramidites and imidazole base pairing effects on DNA

Shinya Harusawa,*a Hiroki Yoneyama,a Yoshihide Usami,a Daisuke Yamamotob and Zheng-yun Zaoc

a Laboratory of Pharmaceutical Organic Chemistry, Osaka University of Pharmaceutical Sciences, 4-20-1 Nasahara, Takatsuki, Osaka 569-1094, Japan.
Fax: +81(72)6901086; E-mail: harusawa@gly.oups.ac.jp
b Biomedical Computation Center, Osaka Medical College, 2-7 Daigakacho, Takatsuki, Osaka 569-8686, Japan
c Centre for Plant Science & Bio-polymer Research, Faculty of Engineering, Computing and Creative Industries, Edinburgh Napier University, 10 Colinton Road, Edinburgh EH10 5DT, U.K.

Table of Contents:

Title and table of contents S1

Synthesis of 2'-deoxyribonucleoside phosphoramidites

31P NMR Spectra of Imz-C₀-2'-deoxyribonucleoside PAs 3a and 3b S2

1H and 13C NMR Spectra of Imz-C₂-2'-deoxyribonucleosides 9-13, 3b S3–S10

DNA synthesis

General Experimental Information S11

Oligo Sequences, Concentrations and Yields S12

Mass Spectrometry S13–S19

HPLC Analytical S-20–S25
$^{31}\text{P NMR}$ Spectrum of compound 3a in CDCl$_3$

$^{31}\text{P NMR}$ Spectrum of compound 3b in CDCl$_3$
$^1$H NMR Spectrum of carboxylate compound in CDCl$_3$

$^{13}$C NMR Spectrum of carboxylate compound in CDCl$_3$
$^1$H NMR Spectrum of compound 9 in CDCl$_3$

$^{13}$C NMR Spectrum of compound 1v in CDCl$_3$
$^1$H NMR Spectrum of compound (E)-11 in CDCl$_3$

$^{13}$C NMR Spectrum of compound (E)-11 in CDCl$_3$
$^1$H NMR Spectrum of compound (Z)-11 in CDCl$_3$

$^{13}$C NMR Spectrum of compound (Z)-11 in CDCl$_3$
$^1$H NMR Spectrum of compound ($E$)-**12** in CDCl$_3$

$^{13}$C NMR Spectrum of compound ($E$)-**12** in CDCl$_3$
$^1$H NMR Spectrum of compound (Z)-12 in CDCl$_3$

$^{13}$C NMR Spectrum of compound (Z)-12 in CDCl$_3$
$^1$H NMR Spectrum of compound 13 in CD$_3$OD

$^{13}$C NMR Spectrum of compound 13 in CD$_3$OD
$^1$H NMR Spectrum of compound 3b in CDCl$_3$

$^{13}$C NMR Spectrum of compound 3b in CDCl$_3$
General Experimental Information

Unless otherwise stated, solvents and reagents were purchased as highest grade from commercial suppliers and used without further purification. HPLC-grade water was used throughout DNA work.

Automated DNA synthesis of all sequences was performed at the University of Birmingham on an Applied Biosystems ABI 394 synthesizer (Foster city, California).

HPLC purification was carried out at the University of Birmingham using a Dionex system with Summit P580 pump and Summit UVD 170s UV/VIS Multi-Channel Detector with prep flow cell. Chromeleon 6.8 was used as a software to control the program and for data collection.

The concentrations of oligonucleotides were determined by absorbance at 260nm using a UV-1800 Shimadzu UV spectrophotometer. The oligonucleotides were kept at -20°C for future use.

DNA melting temperatures were determined on a Varian Cary 5000 with a peltier heating accessory on a range of 15 to 85 °C with a heating rate of 0.5 °C /min. The cycle has two dissociation and one annealing processes. The absorbances at 260nm were recorded as a function of temperature. Tm values were reported from the average first derivative of the two heating dissociation curve using Varian software, with an error less than 0.5 °C.
<table>
<thead>
<tr>
<th>Oligo No. and name</th>
<th>Sequences (5' to 3&quot;)</th>
<th>Concentration (μM)</th>
<th>Yield (%)</th>
<th>Modifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>978</td>
<td>TGGACTC5CTCAATG</td>
<td>175</td>
<td>26.2</td>
<td>5=IM</td>
</tr>
<tr>
<td>979</td>
<td>CATTGAG5GAGTCCA</td>
<td>243</td>
<td>36.4</td>
<td>5=IM</td>
</tr>
<tr>
<td>980</td>
<td>TGGAC5CTC5CAATG</td>
<td>87.3</td>
<td>13.0</td>
<td>5=IM</td>
</tr>
<tr>
<td>981</td>
<td>TGGACTC5CTCAATG</td>
<td>254</td>
<td>38.1</td>
<td>5=IM2</td>
</tr>
<tr>
<td>982</td>
<td>CATTGAG5GAGTCCA</td>
<td>267</td>
<td>40.0</td>
<td>5=IM2</td>
</tr>
<tr>
<td>983</td>
<td>TGGAC5CTC5CAATG</td>
<td>111</td>
<td>16.6</td>
<td>5=IM2</td>
</tr>
<tr>
<td>984</td>
<td>CATTGGGAGG GTCCA</td>
<td>241</td>
<td>36.1</td>
<td></td>
</tr>
<tr>
<td>985</td>
<td>CATTGCGAGCGTCCA</td>
<td>283</td>
<td>42.4</td>
<td></td>
</tr>
<tr>
<td>986</td>
<td>CATTTGTGAGTGTCGA</td>
<td>286</td>
<td>42.8</td>
<td></td>
</tr>
<tr>
<td>987</td>
<td>CATTG5GAG5GTCCA</td>
<td>295</td>
<td>44.2</td>
<td>5=Ab</td>
</tr>
<tr>
<td>988</td>
<td>TGGACTC5CTCAATG</td>
<td>230</td>
<td>34.4</td>
<td>5=Ab</td>
</tr>
</tbody>
</table>
Mass Spectrometry

Electrospray ionization (ESI) mass spectra in negative mood for oligonucleotides sequences were obtained on a Waters micromass LCT ESI-ToF (time of flight) mass spectrometer. 20 µl HPLC purified sample, approximately 50-80 µM, was added to a mixture of acetonitrile and 1% TFA in water (50 µl, 1/1), 10 µl was injected into ESI.

Mass results obtained for synthesized oligonucleotides

<table>
<thead>
<tr>
<th>Oligo No.</th>
<th>Sequences (5’ to 3’)</th>
<th>Expected mass</th>
<th>Observed mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>978</td>
<td>TGGACTC5CTCAATG</td>
<td>(5=IM)</td>
<td>4484.8</td>
</tr>
<tr>
<td>979</td>
<td>CATTGAG5GAGTCCA</td>
<td>(5=IM)</td>
<td>4533.9</td>
</tr>
<tr>
<td>980</td>
<td>TGGAC5CTC5CAATG</td>
<td>(5=IM)</td>
<td>4426.7</td>
</tr>
<tr>
<td>981</td>
<td>TGGACTC5CTCAATG</td>
<td>(5=IM2)</td>
<td>4512.9</td>
</tr>
<tr>
<td>982</td>
<td>CATTGAG5GAGTCCA</td>
<td>(5=IM2)</td>
<td>4561.9</td>
</tr>
<tr>
<td>983</td>
<td>TGGAC5CTC5CAATG</td>
<td>(5=IM2)</td>
<td>4482.8</td>
</tr>
<tr>
<td>984</td>
<td>CATTGGGAGGGGTCCA</td>
<td></td>
<td>4633.1</td>
</tr>
<tr>
<td>985</td>
<td>CATTGCGAGCGGTCCA</td>
<td></td>
<td>4553.1</td>
</tr>
<tr>
<td>986</td>
<td>CATTGTGAGTGTCCA</td>
<td></td>
<td>4583.1</td>
</tr>
<tr>
<td>987</td>
<td>CATTG5GAG5GTCCA</td>
<td>(5=Ab)</td>
<td>4335.1</td>
</tr>
<tr>
<td>988</td>
<td>TGGACTC5CTCAATG</td>
<td>(5=Ab)</td>
<td>4419.1</td>
</tr>
</tbody>
</table>
Mass spectrum of oligo 978

Mass spectrum of oligo 979
Mass spectrum of oligo 980

Mass spectrum of oligo 981
Mass spectrum of oligo 982

Mass spectrum of oligo 983
Mass spectrum of oligo 984

Mass spectrum of oligo 9
Mass spectrum of oligo 986

Mass spectrum of oligo 987
Mass spectrum of oligo 988
**HPLC Analytical**

Analytical HPLC for each oligo was carried out using a Shimadzu UFLC on Phenomenex Clarity 5u Oligo-Rp column, 150x4.6mm. Aliquot samples (ca 70 μM, 20 μl) were loaded in an auto tray sitting in a fridge. The samples were injected into the HPLC column, which was eluted with 0.1M TEAA, pH 7.0 and acetonitrile mixtures.

**Conditions (Oligo analytical method):**

**Solvent system A:** 5% MeCN, 0.1 M TEAA pH 7.0; **Solvent system B:** 15% MeCN, 0.1 M TEAA pH 7.0; **Solvent system C:** MeCN

Gradient (linear increase): 0 - 25 mins, 30% B - 50% B; 35 - 45 mins, 0% - 100% C; 45 - 55 mins, 100% C hold; 55-60 mins, 30% B. Flow rate: 1.0 ml/min.

Analytical HPLC of oligo 978
Analytical HPLC of oligo 979

Analytical HPLC of oligo 980
Analytical HPLC of oligo 981

Analytical HPLC of oligo 982
Analytical HPLC of oligo 983

Analytical HPLC of oligo 984
Analytical HPLC of oligo 985

Analytical HPLC of oligo 986
Analytical HPLC of oligo 987

Analytical HPLC of oligo 988