Supporting Information

**Chimeric (α-amino acid + nucleoside-β-amino acid)_n peptide oligomers show sequence specific DNA/RNA recognition**

Khirud Gogoi* and Vaijayanti A Kumar*

Division of Organic Chemistry, National Chemical Laboratory, Pune-411008, India
E-mail: va.kumar@ncl.res.in

S2-S3: Experimental procedures
S4-S6: ¹H NMR spectra of 7 and 8, ¹³C and DEPT NMR spectra of 8 and Mass spectra of compound 8
S6-S7: Oligomer synthesis, Solid phase synthesis scheme and structure of ONs synthesized.
S8-S12: Table of ONs synthesized and their RP-HPLC and MALDI-TOF mass spectral data.
S13-S16: UV-melting experiments and UV-Job’s plots.
S16-S17: CD graphs.
Experimental Procedure:

Melting points of samples were determined in open capillary tubes using Buchi Melting point B-540 apparatus and are uncorrected. IR spectra were recorded on an infrared Fourier Transform spectrometer. Column chromatographic separations were performed using silica gel 60-120 and 230-400 mesh, Ethyl acetate, Petroleum ether, Dichloromethane and Methanol as the solvent system. $^1$H and $^{13}$C spectra were obtained using Bruker AC 200 (200MHz) NMR spectrometers.

Synthesis of 3’-azidothymidine-5’-carboxylic acid 7$^1$:

![Chemical structure](image)

BAIB (3.9g, 12.4 mmol) and TEMPO (180mg, 1.12mmol) and 3’-azido-thymidine 6 (1.5g, 5.6 mmol) were combined in a reaction vessel and to the mixture was added 12 ml of 1:1 acetonitrile-water solution. The reaction mixture was stirred for 3h. Solvents were removed in vacuo and the resulting residue was triturated sequentially with diethyl ether and acetone, filtered and dried in vacuo. Yield: 1.22g, 78%.

M.P. 174-177 ºC FT-IR (Nujol) 3440, 2110, 1748, 1721, 1694, 1610 cm$^{-1}$

$^1$H NMR (DMSO-d$_6$) $\delta$ 1.78(s, 3H), 2.2(m, 2H), 4.35(d, 1H), 4.64(m, 1H), 6.2(t, 1H), 8.36(s, 1H).

Synthesis of 3’-N-Fmoc-3’-aminothymidine-5’-carboxylic acid 8:

![Chemical structure](image)

3'-azidothymidine-5'-carboxylic acid 7 (1.2 g, 4.26 mmol) was dissolved in 10 ml methanol and to it was added 120 mg (10%) Pd-C catalyst. The mixture was subjected to hydrogenation at 40 Psi of hydrogen pressure for 3.5 h. The catalyst was filtered over celite and concentration of the filtrate in vacuo gave 1.1g, 97% of 3'-deoxy-3'-amino thymidine-5'-carboxylic acid.

The amino acid 3'-deoxy-3'-aminothymidine-5'-carboxylic acid (1.1g, 4.3 mmol) was treated with Fmoc-succinimide (1.74g, 5.17 mmol) and NaHCO₃ (1.81g, 21.5 mmol) in 10 ml 1:1 acetone-water. After stirring for five hours, acetone was evaporated and the reaction mixture was diluted with 5 ml water. The water layer was extracted with ethyl acetate to remove all the nonpolar unwanted impurities. The water layer was neutralized by adding dilute HCl to bring the pH=7.0. The water layer was then extracted with ethyl acetate. The organic layer was washed with water (1x10ml) and saturated NaCl solution (1x10ml), dried over Na₂SO₄ and evaporated to dryness to afford 1.25g of 8 (61%) as pure white amorphous solid. M.p.147-150°C.

¹H NMR (CDCl₃ + 1 drop DMSO-d₆, 200 MHz) 1.87(s, 3H), 2.0-2.2(m, 1H), 2.3-2.4(m, 1H), 4.1-4.5(m, 4H), 6.46(t, 3H), 7.2-7.4(m, 5H), 7.6-7.73(m, 5H), 8.1(s, 1H), 10.4(s, 1H).

¹³C NMR (DMSO-d₆, 50 MHz) 12.4, 36.2, 46.6, 54.5, 65.6, 81.5, 85.1, 109.3, 120.1, 125.1, 127.1, 127.6, 136.4, 140.7, 143.8, 150.5, 155.5, 163.7, 172.2.

Mass calc. 477.42 obs. 499.62(+Na⁺)
Figure S1: $^1$H NMR of 7

Figure S2: $^1$H NMR of 8
Figure S3: $^{13}$C NMR of 8

Figure S4: DEPT NMR of 8
Figure S5: Mass spectrum of 8

Oligomer Synthesis:

The amino acids used for solid-phase oligonucleotides(ON) synthesis were Fmoc-Lys(Boc)-OH, Fmoc-Pro-OH, Fmoc-Met-OH and Fmoc-Sar-OH. All the $N^\alpha$-Fmoc L-amino acids (except Fmoc-Sar-OH) and resins were obtained from Novabiochem (Fmoc=9-fluorenyl methoxy-carbonyl). Sarcosine ($N$-methylglycine) was obtained from Lancaster chemicals and Fmoc protection was done by standard methods. The amounts of isolated ONs were determined spectrophotometrically at 260 nm wavelength. The molar absorptivities were assumed to be identical to the Deoxyribonucleosides.
Figure S6: Solid-Phase synthesis of amino acid backbone modified ONs and structure of ONs synthesized.

i) 20% piperidine/DMF
ii) Fmoc-Pro-OH, HBTU, HOBt, DIPEA, DMF

Step 1

Rink amide resin

Step 2

Repetition step 1 and step 2
7 times, then 20% piperidine/DMF

10% TFA-DCM, Triethyl silane
Table S1: Amino-acid backbone modified ONs and their HPLC and MALDI-TOF mass characterization.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Sequence</th>
<th>Mol Formula</th>
<th>HPLC $t_R$ (min)</th>
<th>Mass calculated/observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\beta$-ala-(Pro-$t$)$_8$-H 12</td>
<td>$C_{123}H_{145}N_{33}O_{42}$</td>
<td>12.7</td>
<td>2763.7/2785.6(+Na$^+$) 2801.6(+K$^+$)</td>
</tr>
<tr>
<td>2</td>
<td>$\beta$-ala-(Sar-$t$)$_8$-H 13</td>
<td>$C_{107}H_{134}N_{33}O_{42}$</td>
<td>12.3</td>
<td>2555.4/2578.8(+Na$^+$) 2594.9(+K$^+$)</td>
</tr>
<tr>
<td>3</td>
<td>(Lys-$t$)$_8$-H 14</td>
<td>$C_{128}H_{186}N_{40}O_{41}$</td>
<td>11.7</td>
<td>2941.1/2941.7</td>
</tr>
<tr>
<td>4</td>
<td>$\beta$-ala-(Met-$t$)$_8$-H 15</td>
<td>$C_{123}H_{147}N_{33}O_{42}S_8$</td>
<td>12.6</td>
<td>3036.3/3037.0</td>
</tr>
</tbody>
</table>

Figure S7: RP-HPLC of 12
Figure S8: RP-HPLC of 13

<table>
<thead>
<tr>
<th>Peak No</th>
<th>Ret. Time (min)</th>
<th>Width 1/2 (sec)</th>
<th>Peak Area (counts)</th>
<th>Result (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.208</td>
<td>7.5</td>
<td>994718</td>
<td>98.7647</td>
</tr>
<tr>
<td>2</td>
<td>13.459</td>
<td>0.0</td>
<td>70545</td>
<td>1.2353</td>
</tr>
</tbody>
</table>

Figure S9: RP-HPLC of 14

<table>
<thead>
<tr>
<th>Peak No</th>
<th>Ret. Time (min)</th>
<th>Width 1/2 (sec)</th>
<th>Peak Area (counts)</th>
<th>Result (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11.748</td>
<td>12.0</td>
<td>225517</td>
<td>99.3523</td>
</tr>
<tr>
<td>2</td>
<td>13.165</td>
<td>0.0</td>
<td>1287</td>
<td>4.6165</td>
</tr>
</tbody>
</table>

5324704 100.000
Figure S10: RP-HPLC of 15
Figure S11: MALDI-TOF mass of 15

\[ \beta\text{-Ala-(Met-t)}_8\text{-H 15} \]
Mass Calc. 3036.4
Mass Obs 3037.0

Figure S12: MALDI-TOF mass of 14

\[(\text{Lys-t})_8\text{-H 14}] \]
Mass Calc 2941.11
Mass Obs 2941.72
Figure S13: MALDI-TOF mass of 12

\[ \beta\text{-Ala-(Pro-t)}_8\text{-H 12} \]
Mass Calc. 2763.7
Mass Obs. 2785.6 (+Na\(^+\))
2801.6 (+K\(^+\))

Figure S14: MALDI-TOF mass of 13

\[ \beta\text{-Ala-(Sar-t)}_8\text{-H 13} \]
Mass Calc. 2555.4
Mass Obs 2578.8 (+Na\(^+\)),
2594.9 (+K\(^-\))
**UV-Tm studies:** The complementary DNA and RNA oligomers were synthesized on an Applied Biosystems 3900 DNA Synthesizer. The concentration was calculated on the basis of absorbance from the molar extinction coefficients of the corresponding nucleobases. The experiments were performed in 1-2 µM concentrations. The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM) and were annealed by keeping the samples at 90°C for 5 minutes followed by slow cooling to room temperature. Absorbance versus temperature profiles were obtained by monitoring at 260 nm with Perkin-Elmer Lambda 35 spectrophotometer scanning from 5 to 85°C at a ramp rate of 0.5°C per minute. The data were processed using Microcal Origin 6.1 and $T_m$ values derived from the derivative curves. All values are an average of at least 3 experiments and accurate to within ±0.5°C.

**Table S2:** $T_m$ (°C) values of ONs: DNA / Mismatch DNA

<table>
<thead>
<tr>
<th>Entry</th>
<th>ON Sequence</th>
<th>DNA 9</th>
<th>DNA 16</th>
<th>$\Delta T_m$ (mismatch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5’TITTTTTTTT3'</td>
<td>11</td>
<td></td>
<td>17.8</td>
</tr>
<tr>
<td>2</td>
<td>β-Ala-(Pro-t)$_8$-H</td>
<td>12</td>
<td>49.0</td>
<td>37.5</td>
</tr>
<tr>
<td>3</td>
<td>β-Ala-(Sar-t)$_8$-H</td>
<td>13</td>
<td>49.1</td>
<td>35.3</td>
</tr>
<tr>
<td>4</td>
<td>(Lys-t)$_8$-H</td>
<td>14</td>
<td>56.3</td>
<td>38.5</td>
</tr>
<tr>
<td>5</td>
<td>β-Ala-(Met-t)$_8$-H</td>
<td>15</td>
<td>47.3</td>
<td>32.9</td>
</tr>
</tbody>
</table>

DNA 9 5’ GCAAAAAAAACG 3’ DNA 16 5’ GCAAAAA[TAAACG 3’ (Single base mismatch sequence)

**Figure S15:** A. UV-melting Curves of 12, 13, 14 and 15 with complementary DNA 9 and B. Corresponding first derivative Curves.
Figure S16: C. UV-melting Curves of 12, 13, 14 and 15 with mismatch DNA 16 and D. Corresponding first derivative Curves.

Table S3: Tm (°C) values of ONs: RNA / Mismatch RNA

<table>
<thead>
<tr>
<th>Entry</th>
<th>ON Sequence</th>
<th>RNA10</th>
<th>RNA 17</th>
<th>ΔTm_{mismatch}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5’ TTTTTTTT 3’</td>
<td>11</td>
<td>15.6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>β-Ala-(Pro-t)$_8$-H 12</td>
<td>60.8</td>
<td>44.6</td>
<td>16.2</td>
</tr>
<tr>
<td>3</td>
<td>β-Ala-(Sar-t)$_8$-H 13</td>
<td>61.6</td>
<td>45.6</td>
<td>16.0</td>
</tr>
<tr>
<td>4</td>
<td>(Lys-t)$_8$-H 14</td>
<td>69.0</td>
<td>55.5</td>
<td>13.5</td>
</tr>
<tr>
<td>5</td>
<td>β-Ala-(Met-t)$_8$-H 15</td>
<td>57.3</td>
<td>41.2</td>
<td>16.1</td>
</tr>
</tbody>
</table>

RNA 10 r(5’ GCAAAAAAAACG 3’) RNA 17 r(5’ GCAAAAUAACG 3’)( single base mismatch sequence).

Figure S17: E.UV-melting Curves of 12, 13, 14 and 15 with complementary RNA 10 and F. Corresponding first derivative Curves.
Figure S18: G.UV-melting Curves of 12, 13, 14 and 15 with mismatch RNA 17 and H. Corresponding first derivative Curves.

Figure S19: Jobs plot of modified ONs 12, 13, 14 and 15 with complementary DNA 9 at 25°C in the relative molar ratios of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, and 100:0 (10mM sodium phosphate buffer pH 7.0, 100 mM NaCl, 0.1 mM EDTA).

Figure S20: Jobs plot of modified ONs 12 and 14 with complementary RNA 10 at 25°C in the relative molar ratios of 0:100, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, and 100:0 (10mM sodium phosphate buffer pH 7.0, 100 mM NaCl, 0.1 mM EDTA).


Circular dichorism studies:

The circular dichorism spectra were obtained in a JASCO J-715 spectrometer. The experiment were performed in 2µM concentration and The complexes were prepared in 10 mM sodium phosphate buffer, pH 7.0 containing NaCl (100 mM) and EDTA (0.1 mM) and were annealed by keeping the samples at 90°C for 5 minutes followed by slow cooling to room temperature. The experiments were performed at 25°C. The data were processed using Microcal Origin 6.1.
Figure S21: CD Curves of ONs 12, 13, 14 and 15 in water

Figure S23: CD curves of 12, 13, 14 and 15 with complementary RNA 10