Supporting Information

DNA-Templated Polymerization of Side-Chain-Functionalized Peptide Nucleic Acid Aldehydes
Ralph E. Kleiner, Yevgeny Brudno, Michael E. Birnbaum, and David R. Liu*
Howard Hughes Medical Institute and the Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Sample preparation and MALDI-TOF analysis of polymerization products
The DNA-templated polymerization reactions of unfunctionalized H$_2$N-ggatt-CHO shown in lanes 2 and 3 of Figure 3b were prepared using 250 pmol template. The DNA-templated polymerization reactions of multiply-functionalized H$_2$N-gg*at*t-CHO and H$_2$N-gg*a*t*t-CHO shown in lanes 10-13 of Figure 3a were prepared using 500 pmol template. Each of these samples was loaded onto a preparative 15% denaturing polyacrylamide TBE/urea gel, and the product bands (for lane 13 in Figure 3a, the fastest migrating band was excised; the major bands were excised for the other samples) were visualized by UV shadowing and excised. The excised polyacrylamide was homogenized, soaked in TE buffer overnight, and filtered. The resulting solution was concentrated, passed through a sephadex gel filtration column (Princeton Separation), and evaporated to dryness. Samples were resuspended in 19 µL 1x S1 nuclease reaction buffer (Promega) and 1 µL S1 nuclease (Promega) was added. Following incubation at 37 °C overnight, an additional 1 µL S1 nuclease (Promega) was added, and the samples were heated at 65 °C for two hours. The reactions were purified using a ZipTip$_{µ}$C$_{18}$ (Millipore) and analyzed by MALDI-TOF mass spectrometry using a sinapic acid matrix with 25 mM ammonium citrate. In all cases, the observed masses corresponded to PNA polymers linked to 5’-amino deoxythymidine. For the matched template, the observed mass corresponded to eight PNA pentamers linked to 5’-amino deoxythymidine (Table S2), and for the mismatched template, the observed mass corresponded to four PNA pentamers linked to 5’-amino deoxythymidine (Table S2).

Modeling studies of glutamine-modified PNA-DNA heteroduplex
Starting with the solution structure of PNA-DNA heteroduplex gctagtgc•d(GACATAGC) (Eriksson, M.; Nielsen, P.E. Nature Structural Biology 1996, 3, 410), a glutamine side-chain was built on the PNA-thymine monomer at position five (gctat*gtc). The modified structure was then enclosed in a periodic box of water with dimensions x = 46 Å, y = 41 Å, and z = 56 Å. The original atoms of the unmodified structure were fixed, and a geometry optimization was performed using molecular mechanics with the Amber99 force field and the Polak-Ribiere minimization algorithm. The single point energy of the solvated system was then calculated. This approach was utilized for all four glutamine regio- and stereochemistries. For the α-functionalized structures, $E_{α glut} = -52002.42$ kcal/mol and $E_{α D glut} = -52061.06$ kcal/mol. For the γ-functionalized structures, $E_{γ glut} = -52025.91$ kcal/mol and $E_{γ D glut} = -51915.28$ kcal/mol. All structure optimization and energy calculations were performed in HyperChem 8.0. Optimized structures (with the water removed) were rendered using PyMol.
Table S1. ESI-MS (m/z) characterization of PNA pentamer aldehydes

<table>
<thead>
<tr>
<th>PNA Pentamer Aldehyde</th>
<th>Expected Mass [M + 2H]^{2+}</th>
<th>Observed Mass [M + 2H]^{2+}</th>
</tr>
</thead>
<tbody>
<tr>
<td>ggtt</td>
<td>559.23</td>
<td>559.23</td>
</tr>
<tr>
<td>ggtt [\gamma-(L)-lys]</td>
<td>594.76</td>
<td>594.77</td>
</tr>
<tr>
<td>ggtt [\gamma-(L)-lys]</td>
<td>607.27</td>
<td>607.27</td>
</tr>
<tr>
<td>ggatt</td>
<td>696.78</td>
<td>696.78</td>
</tr>
<tr>
<td>ggtt [\gamma-(L)-lys]</td>
<td>732.32</td>
<td>732.32</td>
</tr>
<tr>
<td>ggatt [\gamma-(L)-lys]</td>
<td>732.33</td>
<td>732.33</td>
</tr>
<tr>
<td>ggatt [\gamma-(L)-lys]</td>
<td>767.85</td>
<td>767.86</td>
</tr>
<tr>
<td>ggatt [\gamma-(L)-lys]</td>
<td>803.39</td>
<td>803.40</td>
</tr>
<tr>
<td>ggtt [\gamma-(L)-lys]</td>
<td>559.62 [M + 3H]^{3+}</td>
<td>559.63 [M + 3H]^{3+}</td>
</tr>
</tbody>
</table>

Table S2. MALDI-TOF analysis of polymerization products.

<table>
<thead>
<tr>
<th>Polymerization Template</th>
<th>Polymer Species</th>
<th>Expected [M + H]^+</th>
<th>Observed [M + H]^+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mismatched</td>
<td>4x ggatt + dT</td>
<td>5744.32</td>
<td>5739.43 ± 15</td>
</tr>
<tr>
<td>Matched</td>
<td>8x ggatt + dT</td>
<td>11246.52</td>
<td>11267.15 ± 30</td>
</tr>
<tr>
<td>Mismatched</td>
<td>4x gg*att + dT</td>
<td>6312.89</td>
<td>6321.81 ± 15</td>
</tr>
<tr>
<td>Matched</td>
<td>8x gg*att + dT</td>
<td>12383.664</td>
<td>12400.57 ± 30</td>
</tr>
<tr>
<td>Mismatched</td>
<td>4x gg<em>a</em>tt + dT</td>
<td>6597.18</td>
<td>6600.65 ± 15</td>
</tr>
<tr>
<td>Matched</td>
<td>8x gg<em>a</em>tt + dT</td>
<td>12952.25</td>
<td>12981.45 ± 30</td>
</tr>
</tbody>
</table>
Figure S1. (a) Polymerization of functionalized and unfunctionalized PNA tetramer aldehydes in the presence of matched and mismatched templates. Short black arrows indicate full-length product corresponding to the polymerization of 10 tetramers resulting in a 40mer of PNA. Reactions were performed for 30 min at the reaction temperature (40 °C for ggtt and ggt*t; 50 °C for gggt and ggg*t; temperatures were chosen to optimize efficiency and sequence specificity) in 10 mM TAPS pH 8.5 with 100 mM NaCl, using 0.4 µM DNA template and 16 µM of the indicated PNA tetramer aldehyde and then quenched with excess allyl amine. Prior to the addition of NaBH₃CN, reactions were pre-equilibrated by heating to 95 °C for 10 min and slowly
cooling to the reaction temperature over 15 min. (b) Comparison of polymerization efficiencies of functionalized and unfunctionalized PNA tetramer aldehydes. Reactions were performed as in Figure S1a. Product yield was determined by gel densitometry using denaturing PAGE. Values represent the average and standard deviation of three independent reactions.
Figure S2. Effect of PNA building block concentration on polymerization efficiency and basic sequence specificity. H$_2$N-ggatt-CHO PNA pentamer aldehyde was polymerized at varying concentrations in the presence of matched and mismatched templates. Reactions were performed for 30 min at 40°C in 10 mM TAPS pH 8.5 with 100 mM NaCl, using 0.4 µM DNA template and 6.4 µM (2 equiv), 12.8 µM (4 equiv), 19.2 µM (6 equiv), or 25.6 µM (8 equiv) of the H$_2$N-ggatt-CHO building block and then quenched with excess allyl amine. Prior to the addition of NaBH$_3$CN, reactions were pre-equilibrated to the reaction temperature as described in Figure 2.
Figure S3. UV-melting curves of side-chain functionalized H2N-ggat*tt-CHO pentamer building blocks on an eight-codon DNA-template (sequence: 5’-[AATCC]₈-3’). Thermal denaturation was performed on an Amersham Biosciences Ultrospec 3300 Pro UV/visible spectrophotometer at a 0.67 µM DNA concentration and 5.33 µM PNA building block concentration in 100 mM NaCl and 10 mM phosphate buffer pH 7.2. UV absorbance was measured at 260 nm from 20 °C to 70 °C, in 0.1 °C increments. (a) unfunctionalized and α-(L)-functionalized pentamers. (b) α-(D)-functionalized pentamers. (c) γ-(L)-functionalized pentamers. (d) γ-(D)-functionalized pentamers.
$^1$H- and $^{13}$C-NMR of 4a
$^1$H- and $^{13}$C-NMR of 4b
$^{1}H$- and $^{13}C$-NMR of 5a
$^1$H- and $^{13}$C-NMR of 5b
$^1$H- and $^{13}$C-NMR of 7a
$^1$H- and $^{13}$C-NMR of 8a

δ (ppm)

δ (ppm)
$^1$H- and $^{13}$C-NMR of 9a
$^1$H- and $^{13}$C-NMR of 9b
$^1$H- and $^{13}$C-NMR of 9c
$^1$H- and $^{13}$C-NMR of 9d
$^{1}\text{H}-\text{and}^{13}\text{C-NMR of 9e}$
$^{1}\text{H-} \text{and} ^{13}\text{C-NMR of 11a}$

\[ \delta \text{ (ppm)} \]

\[ \delta \text{ (ppm)} \]
$^{1}H$- and $^{13}C$-NMR of 12a

\begin{center}
\includegraphics[width=\textwidth]{nmr_spectrum.png}
\end{center}

\begin{center}
\includegraphics[width=\textwidth]{13c_nmr_spectrum.png}
\end{center}
$^1$H- and $^{13}$C-NMR of 13a

\[
\text{Chemical Structure Image}
\]

\[
\text{NMR Spectra Image}
\]
$^1$H- and $^{13}$C-NMR of 13b
$^{1}H$- and $^{13}C$-NMR of $13c$
$^{1}H$- and $^{13}C$-NMR of 13d
$^1$H- and $^{13}$C-NMR of 13e