Efficient pH-independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA

Michael Egholm+, Leif Christensen, Kim L. Dueholm, Ole Buchardt†, James Coull1 and Peter E. Nielsen2,*

Center for Biomolecular Recognition, Department of Organic Chemistry, The H.C. Ørsted Institute, Universitetsparken 5, DK 2100 Ø Copenhagen, Denmark, 1Biosearch, Core R & D, 75 A Wiggins Institute, Bedford, MA 01730, USA and 2Center for Biomolecular Recognition, Department of Medical Biochemistry and Genetics, Biochemistry Laboratory B, The Panum Institute, Blegdamsvej 3c, DK-2200 N Copenhagen, Denmark

Received November 1, 1994; Revised and Accepted December 16, 1994

ABSTRACT

The synthesis and DNA binding properties of bis-PNA (peptide nucleic acid) are reported. Two PNA segments each of seven nucleobases in length were connected in a continuous synthesis via a flexible linker composed of three 8-amino-3,6-dioxaoctanoic acid units. The sequence of the first strand was TCTCTTT (C- to N-terminal), while the second strand was TTTCTCT or TTTJTJT, where J is pseudoisocytosine. These bis-PNAs form triple-stranded complexes of somewhat higher thermal stability than monomeric PNA with complementary oligonucleotides and the thermal melting transition shows very little hysteresis. When the J base is placed in the strand parallel to the DNA complement ('Hoogsteen strand'), the DNA binding was pH independent. The bis-PNAs were also superior to monomeric PNAs for targeting double-stranded DNA by strand invasion.

INTRODUCTION

Homopyrimidine peptide nucleic acids (PNA) (Fig. 1) bind to complementary DNA or RNA oligonucleotides forming (PNA)2/DNA (RNA) triplexes of very high thermal stability (1–3). Due to the high stability of such (PNA)2/DNA triplexes, binding of PNA to homopurine targets in double-stranded DNA takes place via strand displacement (1,4,5), instead of conventional triple helix formation as observed with natural oligonucleotides (6,7). While PNA/DNA duplexes are most stable in the antiparallel configuration (8), PNAs are oriented from the N- to the C-terminal, thus if the N-end of the PNA faces the 5'-end of the oligonucleotide the complex is termed parallel), the parallel orientation seems to be preferred for (PNA)2/DNA triplexes if only one species of PNA is employed. As would be predicted from these results, the most stable (PNA)2/DNA triplexes are formed when the Watson–Crick base pairing PNA strand is in the anti-parallel orientation relative to the DNA strand (8) and the Hoogsteen strand is in the parallel orientation relative to the DNA strand (i.e. the two PNA strands are anti-parallel relative to each other) (Nielsen et al, in preparation).

As a logical consequence of this, dimeric PNA molecules (bis-PNAs) with one strand designed for Watson–Crick recognition of DNA (RNA) and the other strand designed for Hoogsteen recognition of a PNA-DNA duplex should be optimal for (PNA)2/DNA triplex stability and thus enhance strand displacement binding to double-stranded DNA. We describe here such bis-PNAs which furthermore contain pseudoisocytosine instead of cytosine in the ‘Hoogsteen strand’ to avoid pH sensitivity of the binding.

MATERIALS AND METHODS

Synthesis of N-(2-boc-aminoethyl)-N-(N2-(benzoylcarbonyl)isocytosin-5-ylacetyl)glycine, the pseudoisocytosine PNA monomer

Methylα-formylsuccinate. This procedure is a modification of the one published by Fissekis and Sweet (9). Sodium methoxide (40.5 g, 0.75 mol) was suspended in dry ether (500 ml) and stirred under nitrogen at 0°C. A mixture of dimethyl succinate (65.4 ml, 0.50 mol) and methylformate (123 ml, 2.00 mol) was added dropwise over 30 min. The reaction mixture was stirred at 0°C for 2 h and then at room temperature overnight. Subsequently, the reaction mixture was evaporated to a viscous brown residue which was washed once with petroleum ether and then dissolved in 3 M hydrochloric acid (160 ml). This solution was made weakly acidic with concentrated hydrochloric acid and then extracted with dichloromethane (4×250 ml). The organic phase was dried (MgSO4), filtered and evaporated under reduced

---

*To whom correspondence should be addressed
+Present address: Biosearch, Core R & D, 75 A Wiggins Institute, Bedford, MA 01730, USA
†Deceased
room temperature overnight. Subsequently, the reaction mixture was refluxed under nitrogen for 3 h and then stirred at room temperature. A solution of methyl α-formylsuccinate (30.0 g, 0.17 mol) and guanidine hydrochloride (49.4 g, 0.52 mol) was added. The solution was filtered, and the filter cake was washed once with methanol. The collected filtrate and washing were evaporated under reduced pressure. The resulting residue was dissolved in water (80 ml) and the solution was acidified with concentrated hydrochloric acid to pH 4.2. After having been stirred at 0°C the mixture was filtered, the precipitate washed once with water and then freeze-dried leaving 28.29 g (97%) of the title compound as a white solid.


Due to the poor solubility properties of the product it was further characterized as its sodium salt. 7 (0.42 g, 2.5 mmol) and sodium bicarbonate were dissolved in boiling water (35 ml). The solution was cooled and evaporated. The residue was dissolved in water (6 ml) and ethanol (4 ml) and isopropanol (8 ml) were added. The sodium salt of 7 was collected by filtration, washed with absolute ethanol and petroleum ether and dried to yield 0.31 g (65%) as white crystals.

'HNMR (D₂O/TMS): δ = 3.05 (s, 2H, CH₂COO); 7.40 (s, 1H, H-6). 13C-NMR (DMSO-d₆/TMS): δ = 44.5 (CH₂COO); 121.5 (C-5); 153.8 (C-2); 156.2 (C-6); 168.6 (C-4); 179.3 (CO/OH). MS (FAB+) m/z (%): 253 (100, M+H).

Methyl isocytosin-5-ylacetate (8). Thionyl chloride (3.6 ml, 50 mmol) was added to stirred methanol (210 ml) at -40°C under nitrogen. Isocytosin-5-ylactic acid (7, 7.0 g, 41 mmol) was added and the reaction mixture was stirred at room temperature for 1 h, at 60°C for 3 h and overnight at room temperature. The reaction mixture was evaporated to dryness and the residue was dissolved in saturated aqueous sodium bicarbonate (80 ml) giving a foamy precipitate. The precipitate was collected by filtration, washed with water, recrystallized from water and freeze-dried yielding 4.66 g (62%) of methyl isocytosin-5-ylacetate as white crystals.

'H-NMR (D₂O/TMS): δ = 2.85 (s, 2H, CH₂COO); 6.27 (br s, 2H, NH₂); 7.52 (s, 1H, H-6). 13C-NMR (DMSO-d₆/TMS): δ = 32.0 (CH₂COO); 51.5 (COOMe); 108.4 (C-5); 153.3 (C-2); 156.4 (C-6); 164.0 (C-4); 171.8 (CH₂COO). MS (FAB+) m/z (%): 184 (100, M+H).

Due to the poor solubility properties of the product it was further characterized as its sodium salt. 7 (0.42 g, 2.5 mmol) and sodium bicarbonate were dissolved in boiling water (35 ml). The solution was acidified with concentrated hydrochloric acid to pH 4.2. After having been stirred at 0°C the mixture was filtered, the precipitate washed once with water and then freeze-dried leaving 28.29 g (97%) of the title compound as a white solid.


Methyl N²-(benzoxycarbonyl)isocytosin-5-ylacetate (9). Methyl isocytosin-5-ylacetate (8, 9.5 g, 52 mmol) was dissolved in dry DMF (95 ml) and the solution was stirred at 0°C under nitrogen. N-Benzoxycarbonyl-N'-methylimidazolium triflate (37.99 g, 104 mmol) was added slowly and the reaction mixture was stirred for 30 min at 0°C and then overnight at room temperature. Dichloromethane (800 ml) was added and the resultant mixture was washed with half-saturated aqueous sodium bicarbonate (2x400 ml), half-saturated aqueous potassium hydrogen sulfate (2x400 ml) and with brine (1x400 ml). The organic phase was dried (MgSO₄), filtered and evaporated under reduced pressure. The residue was recrystallized from methanol affording 13.32 g (81%) of the title compound as white crystals.

'H-NMR (D₂O/TMS): δ = 3.43 (s, 2H, CH₂COO); 3.67 (s, 3H, COOMe); 5.30 (s, 2H, PhCH₂); 7.43-7.52 (m, 5H, PhCH₂); 7.77 (s, 1H, H-6). 13C-NMR (DMSO-d₆/TMS): δ = 31.9 (CH₂COO); 51.6 (COOMe); 67.0 (PhCH₂); 128.1-128.5 (m, PhCH₂); 135.7 (PhCH₂); 150.7 (Z-CO); 170.8 (COO). MS (FAB+) m/z (%): 318 (3.5, M+H). Calculated for C₁₃H₁₅N₅O₇: C, 56.78; H, 4.76; N, 13.24. Found: C, 56.68; H, 4.79; N, 13.28.
**RESULTS AND DISCUSSION**

By preparing the bis-PNAs in a continuous synthesis the two PNA moieties will form an antiparallel clamp. A flexible, hydrophilic linker based on 8-amino-3,6-dioxaoctanoic acid (egl) was used to connect the two PNA segments, and computer model building (using the Biosys software InsightII and Discover) indicated that three linker units would be suitable. The bis-PNAs were synthesized by standard solid phase f-Boc-peptide chemistry as previously reported (2,3), purified by reversed phase HPLC, and characterized by FAB mass spectrometry. 

The hybridization results presented in Table 1 show that a small but significant increase in $T_m$ is gained by linking the two PNAs together (I versus 4 + 5). As expected, only a small difference is ascribed to the effect of the linker which is positioned at the 3'- and 5'-end of the oligonucleotide, respectively. 

The slight difference is ascribed to the effect of the linker which is positioned at the 3'- and 5'-end of the oligonucleotide, respectively.
It has previously been shown that changing cytosine to pseudoisocytosine in the Hoogsteen strand (Fig. 1) abolishes the pH sensitivity of the DNA triplex (12,13) without seriously affecting the stability of the triplex. Likewise, 8-oxoadenine has been employed to recognize guanine in the Hoogsteen mode (14–17). However, synthesis of the PNA pseudoisocytosine monomer seemed more straightforward (Scheme 1). We therefore synthesized bis-PNAs with thymines and cytosines in the one strand and thymines and pseudoisocytosines in the other PNA strand (2 and 3). The complexes of these PNAs with complementary oligonucleotides showed thermal stabilities at acidic pH (pH 5) comparable to that of the bis-PNA 1 complex. However, in the complexes where the cytosine PNA strand is anti-parallel to the DNA target (and thus the pseudoisocytosine strand is parallel), almost no pH dependence of the \( T_m \) is observed (complexes 2/1 and 3/1), indicating that the orientation directs the complex formation in a way that orients the Watson–Crick strand in the anti-parallel configuration and thus the Hoogsteen strand in the parallel configuration. Concordently, the complexes 2/II and 3/II show pH dependence of the \( T_m \), indicating that the cytosine PNA strand in these cases is involved in Hoogsteen hydrogen bonding. Finally, we note that the pronounced hysteresis of the melting behavior usually observed with (PNA)\(_2\)/DNA triplexes (\(-30^\circ\text{C}\) for PNA 4), indicating a very slow rate of triplex formation, is dramatically reduced in the case of bis-PNAs (2–3°C for PNAs 1–2). This is ascribed to the high local concentration of the now covalently linked second PNA strand. Thus significantly higher DNA affinity, due to a faster on-rate, would be predicted for triplex-forming bis-PNAs as compared to monomeric PNAs.

As shown in Table 2, the sequence discrimination of the bis-PNA 3, as judged from thermal stability measurements, is better than that found for monomeric PNAs and the very high cost in stability (\(-30^\circ\text{C}\)) for a base mismatch reflects the 2-fold recognition process involving both PNA strands.

**Table 1. Thermal stabilities (\( T_m \)) of bis-PNA–oligonucleotide complexes**

<table>
<thead>
<tr>
<th>PNA</th>
<th>pH</th>
<th>Oligonucleotide 1(^b)</th>
<th>Oligonucleotide II(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-TCTCT(_3)-(egl)(_3)-TCTCT-LysNH(_2) (1)</td>
<td>5</td>
<td>69.0</td>
<td>68.5</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>49.0</td>
<td>52.0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>38.5</td>
<td>41.0</td>
</tr>
<tr>
<td>H-TJTTJ(_3)-(egl)(_3)-TCTCT-LysNH(_2) (2)</td>
<td>5</td>
<td>67.0</td>
<td>65.0</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>64.0</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>60.5</td>
<td>39.0</td>
</tr>
<tr>
<td>H-TCTCT(_3)-(egl)(_3)-TJTT-LysNH(_2) (3)</td>
<td>5</td>
<td>66.0</td>
<td>61.5</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>47.0</td>
<td>60.0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>37.5</td>
<td>59.0</td>
</tr>
<tr>
<td>H-TCTCT(_3)-LysNH(_2) (4)</td>
<td>5</td>
<td>50.0</td>
<td>55.5</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>40.0</td>
<td>39.0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>–</td>
<td>(-23.5)</td>
</tr>
<tr>
<td>4 + H-T(_3)CTCT-LysNH(_2) (5)</td>
<td>5</td>
<td>62.0</td>
<td>65.5</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>45.0</td>
<td>50.0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>34.5</td>
<td>39.5</td>
</tr>
<tr>
<td>4 + H-T(_3)JTTT-LysNH(_2) (6)</td>
<td>5</td>
<td>46.0</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>36.0</td>
<td>44.0</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>37.0</td>
<td>42.5</td>
</tr>
</tbody>
</table>

\(^a\)Buffer: 100 mM NaCl, 10 mM Na-phosphate, 0.1 mM EDTA. Heating rate: 0.5°C/min at 5–90°C.

\(^b\)Oligonucleotide sequences: I, 5’-d(CGCAGAGACGC); II, 5’-d(CGCA\(_3\)GAGACGC).

\(^c\)The hydrogen bond donor/acceptor pattern of pseudoisocytosine is such that dependent on the tautomer it may function both as a cytosine analog for Watson–Crick base pairing with guanine and as a protonated cytosine analog for Hoogsteen base pairing with guanine. Due to this dual recognition property we refer to the pseudoisocytosine base as J for the Greek god Janus with two faces.

**Table 2. Effect of base pair mismatches on PNA/DNA thermal stabilities (\( T_m \))**

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>PNA 2 (^a)</th>
<th>PNA 4 + 6 (^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5’-d(GCGA(_3)GAGACGC)-3’)</td>
<td>60.0</td>
<td>44.0</td>
</tr>
<tr>
<td>(5’-d(GCGA(_3)GAGACGC)-3’)</td>
<td>27.0</td>
<td>32.5</td>
</tr>
<tr>
<td>(5’-d(GCGA(_3)GAGACGC)-3’)</td>
<td>36.5</td>
<td>34.0</td>
</tr>
<tr>
<td>(5’-d(GCGA(_3)GACGCG)-3’)</td>
<td>23.0</td>
<td>33.0</td>
</tr>
<tr>
<td>(5’-d(GCGA(_3)GACGCG)-3’)</td>
<td>(\leq)11</td>
<td>(\leq)11</td>
</tr>
</tbody>
</table>

\(^a\)Measured at pH 7.0 as described in Table 1.
Scheme 1. (i) SOCl<sub>2</sub> in MeOH, overnight, 62%. (ii) N-Benzylxoycarbonyl-N'-methylmiazolium trflate in DMF, overnight, 81%. (iii) LiOH in H<sub>2</sub>O/THF, 2h, 83%. (iv) Ethyl N-(2-Boc-aminoethyl)glycinate, HBTU, N-methylmorpholine in DMF, 4h, 74%. (v) NaOH in H<sub>2</sub>O/MeOH, 75 min, 82%.

Binding to double-stranded DNA was analyzed using a <sup>32</sup>P-end labeled DNA fragment containing the PNA target and probing for strand displacement using KMnO<sub>4</sub>, since thymines in the displaced loop will react with KMnO<sub>4</sub> (1,5). The results (Fig. 2) show that all three dimeric PNAs bind efficiently to double-stranded DNA, whereas the monomeric PNA 4 does not (only a very weak thymine oxidation is seen even at the highest PNA concentration; lane A4). The observed KMnO<sub>4</sub> sensitivity of the thymines outside the PNA target is ascribed to increased DNA breathing proximal to the bound PNA. The most efficient binding is, not surprisingly, observed with PNA 3 having the cytosines in the anti-parallel (Watson-Crick) PNA strand and the pseudocytosine in the parallel (Hoogsteen) PNA strand. Furthermore, the binding did not exhibit the strong pH dependence observed with PNAs containing cytosines in the Hoogsteen strand (5).

In conclusion, we have shown that bis-PNAs with superior DNA binding properties can be constructed by incorporation of nucleobases chosen independently for optimal Watson-Crick and Hoogsteen recognition.

ACKNOWLEDGEMENTS

This work was supported by ISIS Pharmaceuticals, The Danish National Research Foundation and the Alfred Benzon Foundation (a fellowship to K.L.D.).

REFERENCES