A Novel Peptide Nucleic Acid Monomer for Recognition of Thymine in Triple-Helix Structures

Anne B. Eldrup,[†] Otto Dahl,[†] and Peter E. Nielsen*[,‡]

Center for Biomolecular Recognition
Department of Chemistry, University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø, Denmark
Department of Biochemistry B
The Panum Institute, Blegdamsvej 3
DK-2200 Copenhagen N, Denmark

Received May 28, 1997

Triplex targeting of double-stranded DNA using natural nucleobases as introduced ten years ago[1] is still essentially limited to homopurine stretches. The problem lies in the asymmetric nature of the triple-helix structure, in which only the purine part of the base pair is recognized by T•A−T, C$^+$•G−C, A•A−T, or G•G−C triplet formation.[2−4] Attempts to circumvent this problem by targeting homopurine tracts on alternate strands[5] has expanded the sequence recognition repertoire somewhat, but naturally has not resulted in a general solution. It is widely agreed that novel "nucleobases" that are able to sequence specifically recognize T(−A) and C(−G) base pairs are required for this task. Several nucleobases that recognize cytosine in an oligonucleotide context have been described,[6] whereas successful recognition of thymine is still lacking. Duplex-invading homopyrimidine peptide nucleic acids (PNAs)[7,8] also rely on triplex formation for recognition.[9] However, in this case a PNA•DNA−PNA triplex is formed, and thus T(or C) recognition by the Hoogsteen PNA strand may be less stringent, since part of the recognition occurs *via* Watson−Crick base pairing of the other PNA strand. In an approach to expand the triplex recognition repertoire of PNAs beyond that of homopurine targets, we now report the synthesis and recognition properties of a novel "nucleobase", 3-oxo-2,3-dihydropyridazine, designed to recognize the T(−A) base(pair). In order to recognize a T−A base pair from the major groove,

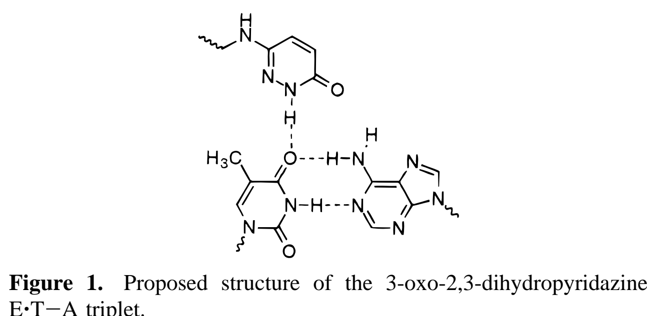

Figure 1. Proposed structure of the 3-oxo-2,3-dihydropyridazine E•T−A triplet.

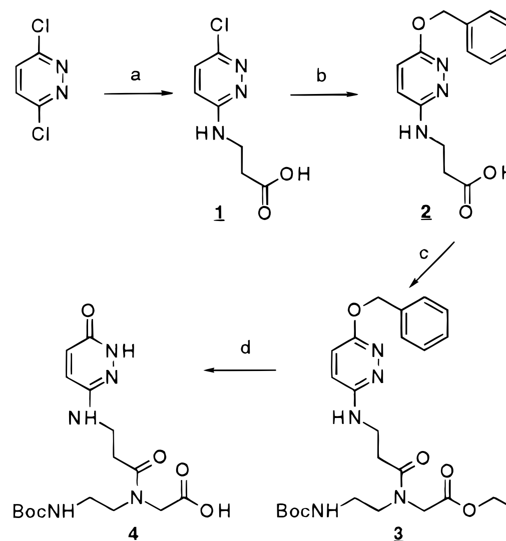

Figure 2. Synthesis of the PNA monomer **4**: (a) β-alanine, K_2CO_3 in absolute EtOH, reflux overnight, 50%; (b) benzyl alcohol and NaH, 3 h at 165 °C, 57%; (c) ethyl *N*-(2-Boc-aminoethyl)glycinate, HODhbt and dicyclohexylcarbodiimide in DMF, overnight at rt, 78%; (d) 1 M LiOH (aqueous) in THF, 30 min at rt, then 2 M HCl (aqueous), H_2, Pd/C in absolute EtOH, 2 h at 0 °C, 80%.

a heterocycle should be connected to the PNA backbone with a linker long enough to circumvent the 5-methyl group of thymine and have a hydrogen donor positioned to bind the 4-oxo group of thymine. If possible, a hydrogen acceptor to form a hydrogen bond to one of the N-6 hydrogen atoms of adenine would also be advantageous. Computer model building indicated that 3-oxo-2,3-dihydropyridazine (E), connected to the PNA backbone via a β-alanine linker from the 6-position, would fulfill these requirements (Figure 1). It should also be noted that the lack of a hydrogen atom on N-1 greatly reduces any steric interference with the thymine 5-methyl group.

For the synthesis of monomer **4** containing the novel nucleobase E (Figure 2), commercially available 3,6-dichloropyridazine was converted to *N*-(3-chloropyridazin-6-yl)-3-aminopropionic acid (**1**) by nucleophilic aromatic substitution with β-alanine in the presence of potassium carbonate. Another nucleophilic substitution to yield *N*-(3-(benzyloxy)pyridazin-6-yl)-3-aminopropionic acid (**2**) was performed at elevated temperature using benzyl alcoholate in benzyl alcohol. Compound **2** was condensed with ethyl *N*-(2-Boc-aminoethyl)glycinate[10] using the previously published procedure.[8] Basic hydrolysis of the resulting benzyl-protected monomer ester **3**, followed by catalytic hydrogenation of the masked oxo functionality in the 3-position, gave the monomer **4**.

In order to assess the triplex recognition properties of the E-base, monomer **4** was incorporated into the Hoogsteen strand of a bis-decamer-linked bis-PNA[11,12] at two positions facing

† University of Copenhagen.
‡ The Panum Institute.

(1) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645−650. (b) Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J. L.; Thuong, N. T.; Lhomme J.; Hélène, C. *Nucleic Acids Res.* **1987**, *15*, 7749−7760.
(2) Mergny, J.-L.; Sun, J.-S.; Rougeé, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Hélène, C. *Biochemistry* **1991**, *30*, 9791−9798.
(3) Greenberg, W. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 5016−5022.
(4) Best, G. C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 1187−1193.
(5) Horne, D. A.; Dervan, P. G. *J. Am. Chem. Soc.* **1990**, *112*, 2435−2437.
(6) (a) Huang, C.-Y.; Bi, G.; Miller, P. S. *Nucleic Acids Res.* **1996**, *24*, 2606−2613. (b) Zimmerman, S. C.; Schmitt, P. *J. Am. Chem. Soc.* **1995**, *117*, 10769−10770. (c) Durland, R. H.; Rao, T. S.; Bodepudi, V.; Seth, D. M.; Jayaraman, K.; Revankar, G. R. *Nucleic Acids Res.* **1995**, *23*, 647−653. (d) Sasaki, S.; Nakashima, S.; Nagtasugi, F.; Yoshitsugu, T.; Hisatome, M.; Maeda, M. *Tetrahedron Lett.* **1995**, *36*, 9521−9524. (e) Koshlap, K. M.; Gillespie, P.; Dervan, P. B.; Feigon, J. *J. Am. Chem. Soc.* **1993**, *115*, 7908−7909. (f) Stilz, H. U.; Dervan, P. B. *Biochemistry* **1993**, *32*, 2177−2185. Huang, Y.-C.; Cushman, C. D.; Miller, P. S. *J. Org.Chem.* **1993**, *58*, 5048−5049. (g) Griffin, L. G.; Kiessling, L. L.; Beal, P. A.; Gillespie, P.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 7976−7982.
(7) (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497−1500. (b) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 1895−1897. (c) Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992**, *114*, 9677−9678. (d) Hyrup, B.; Nielsen, P. E. *Bioorg. Med. Chem.* **1996**, *4*, 5−23.
(8) Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767−5773.
(9) (a) Cherny, D. Y.; Belotserkovskii, B. P.; Frank-Kamenetskii, M. D.; Egholm, M.; Bucharrdt, O.; Berg, R. H.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1667−1670. (b) Nielsen, P. E.; Egholm, M.; Buchardt, O. *J. Mol. Recognit.* **1994**, *7*, 165−170.
(10) Dueholm, K.; Egholm, M.; Buchardt, O. *Org. Prep. Proc. Int.* **1993**, *25*, 457−461.

S0002-7863(97)01742-3 CCC: $14.00 © 1997 American Chemical Society

Table 1. Thermal Stabilities of PNA−DNA Complexes

bis-PNA[a]	T_m (°C)[b]
H-TJT**E**TJ**E**TTT-eg1-eg1-eg1-TTTACTATCT-NH_2[c]	57.0
H-TJT-**aeg**(**Ac**)-TJ-**aeg**(**Ac**)-TTT-eg1-eg1-eg1-TTTACTATCT-NH_2[d]	47.5
H-TJT**G**TJ**G**TTT-eg1-eg1-eg1-TTTACTATCT-NH_2	46.0
H-TTT-eg1-eg1-eg1-TTTACTATCT-NH_2	37.0
H-TTTACTATCT-NH_2	27.0
H-TJT**E**TJ**E**TTT-eg1-eg1-eg1-TTTGCTGTCT-NH_2[c]	42.0

[a] Pseudoisocytosine (J) was used instead of protonated cytosine in the Hoogsteen strand of the bis-PNA. Three units of 8-amino-3,6-dioxaoctanoic acid (eg1) connect the two antiparallel 10-mer PNA strands. [b] T_m measurements were performed in 100 mM NaCl, 1 mM EDTA, 10 mM Na_2HPO_4, pH 7.0 on a Gilford Response spectrophotometer at a heating rate of 0.5 °C/min. [c] E is the 3-oxo-2,3-dihydropyridazine unit derived from monomer **4**. [d] aeg(Ac) is *N*-acetyl-*N*-(2−aminoethyl)glycine, a no-base PNA unit.

an adenine in the Watson−Crick strand (Table 1). Thus, thymines in the DNA target should be recognized at these two positions. All other positions consisted of T,T and J,C[11] pairs for standard recognition of adenine and guanine in the DNA target, respectively. In the same system we also substituted the E-base for a no-base[13] position or a guanine.[14,15] In an analogous system we placed a guanine facing the E-base to study the recognition of cytosine in the DNA target. The thermal stabilities of complexes of these PNAs with their respective DNA targets are given in Table 1. These results clearly demonstrate that the E-base binds stronger to thymine than to either guanine or a no-base ($\Delta T_m \approx 5$ °C per unit) and also show that the E-base recognizes thymine over cytosine ($\Delta T_m \approx 7$ °C per unit).

Table 2. Effect of Mismatches on Bis-PNA Complex Stabilities

oligonucleotide target sequence[a]	T_m (°C)[b]
5′-dCGCAGA**T**AG**T**AAACGC-3′	57.0
5′-dCGCAGA**A**AG**T**AAACGC-3′	45.0
5′-dCGCAGA**G**AG**T**AAACGC-3′	40.5
5′-dCGCAGA**C**AG**T**AAACGC-3′	40.5
5′-dCGCAGA**U**AG**U**AAACGC-3′	63.5
5′-dCGCAGA**A**AG**U**AAACGC-3′	49.5
5′-dCGCAGA**G**AG**U**AAACGC-3′	44.5
5′-dCGCAGA**C**AG**U**AAACGC-3′	44.5

[a] The bis-PNA; H-TJTETJETTT-eg1-eg1-eg1-TTTACTATCT-NH_2, was used. [b] T_m measurements were performed in 100 mM NaCl, 1 mM EDTA, 10 mM Na_2HPO_4, pH 7.0, on a Gilford Response spectrophotometer at a heating rate of 0.5 °C/min.

The sequence discrimination of this system was studied by introducing a mismatch at one of the E,A positions. The results presented in Table 2 show a ΔT_m discrimination in the order of 12−16 °C. Finally, we introduced uracil in the target in place of thymine in order to assess the interference with the 5-methyl group. The thermal stability results (Table 2) indicate that the design of the E-base has not completely solved this steric clash problem since the E-base PNAs bind somewhat stronger to the uracil containing targets than to the thymine containing targets ($\Delta T_m \approx 3$ °C per unit). In conclusion, our results prove the E•T−A triplet to be considerably more stable than the G•T−A or the no-base triplets, although significantly less stable than the canonical C^+•G−C and T•A−T triplets, since complexes of homopyrimidine decameric bis-PNAs with homopurine DNA targets typically show thermal stabilities around 70−80 °C. The results presented herein could direct a novel method for the future design of heterocycles designed for specific interaction with T−A base pairs in DNA•DNA−DNA triplexes.

Acknowledgment. This work was supported by the Danish National Research Foundation. We thank Henrik F. Hansen for fruitful discussions.

Supporting Information Available: Experimental details and charactarization data for all compounds (3 pages). See any current masthead page for ordering and Internet access instructions.

JA9717424

(11) Egholm, M.; Christensen, L.; Dueholm, K.; Buchardt, O.; Coull, J.; Nielsen P. E. *Nucleic Acids Res.* **1995**, *23*, 217−222.

(12) PNA oligomerisation was performed as previously described: Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J.; Berg R. H. *J. Pept. Sci.* **1995**, *3*, 175−183.

(13) L. Christensen and P. E. Nielsen, unpublished results.

(14) Guanine has been shown to interact with thymine to form fairly stable G•T−A triplets in DNA•DNA−DNA triplets[2,3] while surprisingly G•U−A triplets have been reported to be substantially less stable.[15]

(15) Miller, P. S.; Cushman, C. D. *Biochemistry* **1993**, *32*, 2999−3004.