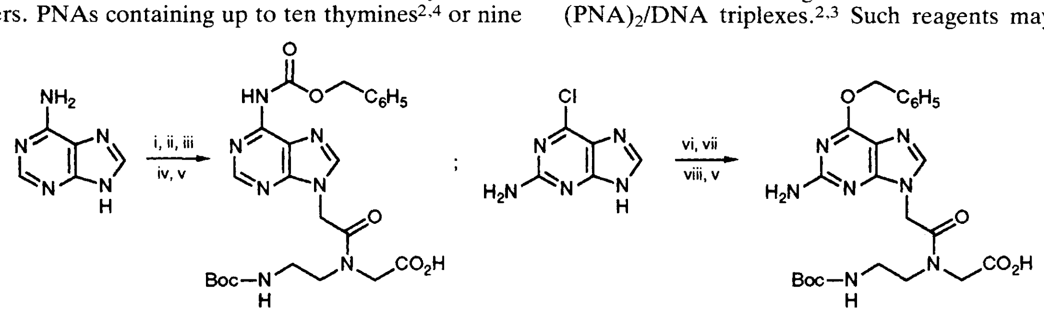
Peptide Nucleic Acids containing Adenine or Guanine recognize Thymine and Cytosine in Complementary DNA Sequences

Michael Egholm,[a] Carsten Behrens,[a] Leif Christensen,[a] Rolf H. Berg,[b] Peter E. Nielsen[c] and Ole Buchardt*[a]

[a] *Research Center for Medical Biotechnology, Chemical Laboratory II, The H. C. Ørsted Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark*
[b] *Polymer Group, Materials Department, Risø National Laboratory, DK-4000 Roskilde, Denmark*
[c] *Department of Biochemistry B, The Panum Institute, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark*

Peptide nucleic acid (PNA) monomer building blocks for the introduction of G and A are prepared and used to synthesise H-T_4XT_5-Lys-NH_2 (X = G or A), which are shown by T_m measurements to recognize their complementary DNA sequences in both the parallel (N-terminal PNA/5′-DNA) and the anti-parallel mode; the stoichiometry in each case is $(PNA)_2$/DNA.

Peptide nucleic acids (PNA)[1–5] are DNA analogues in which the deoxyribose-phosphate backbone has been replaced by a peptide oligomer consisting of, *e.g.* (2-aminoethyl)glycine units to which the nucleobases are attached *via* methylenecarbonyl linkers. PNAs containing up to ten thymines[2,4] or nine thymines and one cytosine[3,4] have previously been shown to recognize their complementary DNA with high fidelity and to bind with much higher binding strength than that of DNA to DNA or RNA owing to the formation of very stable $(PNA)_2$/DNA triplexes.[2,3] Such reagents may be of use as



1 2

Scheme 1 i, $BrCH_2CO_2Et/K_2CO_3$/DMF; ii, $C_6H_5CH_2OC(O)$-imidazolium-N-Et,BF_4^-/DMF/CH_2Cl_2;[7] iii, LiOH/THF/H_2O; iv, Boc-NHCH_2CH_2NHCH_2CO_2Et/DCC/DhbtOH/DMF/CH_2Cl_2; v, NaOH/MeOH/H_2O; vi, $BrCH_2CO_2H$, K_2CO_3, DMF; vii, C_6H_5–CH_2ONa/DMF; viii, Boc-NHCH_2CH_2NHCH_2CO_2Me/PyBroP/DMF

Table 1. Melting temperatures of hybrids (°C)[a] between PNA (H-T_4XT_5-Lys-NH_2 and 5′-d($CGCA_4YA_5CGC$)

		DNA Y			
PNA	X	C	T	A	G
3	G[b]	*63*	50	48	51
4	A[c]	51	*55*	48	48
5	T	66	65	*76*	63
6	C	49	44	46	*79*

[a] The melting temperatures were determined as previously described with the exception that the buffer contained 100 mmol dm^{-3} NaCl, 10 mmol dm^{-3} phosphate, 0.1 mmol dm^{-3} EDTA, pH 7. [b] FAB MS (M + 1), Calc. (Found), 2832.15 (2832.25). [c] FAB MS (M + 1), Calc. (Found), 2816.16 (2815.95).

gene-targeting drugs, as diagnostics, and as molecular biological tools.[6]

We have now prepared the A and G monomers (Scheme 1) and used them for the synthesis of oligomers of the form H-T_4XT_5-Lys-NH_2 (X = G or A), in order to study the DNA recognizing properties of purine-containing PNA. The PNAs **3** (X = G) and **4** (X = A) were assembled by Merrifield synthesis, as previously reported[3] for PNA **6** (X = C) using the adenine monomer **1** and the guanine monomer **2**. The resulting crude PNAs were purified by preparative HPLC, and their identity was established using fast atom bombardment mass spectrometry (Table 1).

The thermal stability of the hybrids between these PNA oligomers, as well as H-T_4XT_5-Lys-NH_2 (X = T or C)[2,3] and hexadecameric oligonucleotides of the form 5′-d($CGCA_4YA_5CGC$), where Y = A, G, C or T, was measured (Table 1). The stoichiometry of the hybrids was determined to be $(PNA)_2$/DNA by UV titrations of selected samples, analogously to what was found for the hybrids between H-T_4XT_5-Lys-NH_2 (X = T or C) and the complementary DNA sequences 5′-d(A_4YA_5) (Y = A or G) as well as for hybrids containing one mismatch.[2,3]

Initial T_m measurements on complexes of the above PNAs with 5′-d(A_4YA_5) (Y = A, C, G or T) showed less discrimination than expected by the purines in PNA and therefore the flanking sequences [5′-d(CGC)] of the oligonucleotides were placed to ensure that no stacking or generation of staggered complexes between individual hybrids would take place.

The results (Table 1) clearly show that A and G in PNA recognize T and C in DNA, respectively. However, the thermal stability of the complexes between PNA containing A or G and their complementary DNA is lower than those of the homopyrimidine PNA/DNA complexes. We ascribe this to the fact that all these complexes are $(PNA)_2$/DNA triplexes. In the case of homopyrimidine PNA perfectly matched Hoogsteen/Watson–Crick (T-A-T or C^+-G-C) triplets can be formed, whereas in the case of PNA containing A or G, only the Watson–Crick strand matches while the second PNA strand binds with a mismatch. Thus, we infer from these results that PNA/DNA duplexes (and presumably PNA/RNA duplexes as well) will exhibit straightforward Watson–Crick base-pair recognition.†

We have shown that all four DNA bases can be incorporated into PNA and that they bind to their complementary bases in DNA, presumably through Watson–Crick base pairing. The presently examined T-rich PNA strands form $(PNA)_2$/DNA triple helices where the second PNA is believed to bind *via* Hoogsteen base pairing.

This work was supported by ISIS Pharmaceuticals and Millipore Corporation. Ms Annette W. Jørgensen, Ms Vivi Hedeboe, Ms Susanne Klitgaard and Mr Nikolaj Hansen are thanked for their skilful technical assistance.

Received, 27th January 1993; Com. 3/00540B

References

1 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497.
2 M. Egholm, O. Buchardt, P. E. Nielsen and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895.
3 M. Egholm, P. E. Nielsen, O. Buchardt and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 9677.
4 D. Y. Cherny, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, M. Egholm, O. Buchardt, R. H. Berg and P. E. Nielsen, *Proc. Natl. Acad. Sci., USA*, 1993, **90**, 1667.
5 B. H. Madsen, M. Egholm, R. H. Berg, P. E. Nielsen and O. Buchardt, *J. Chem. Soc., Chem. Commun.*, 1993, 518.
6 E. Uhlmann and A. Peyman, *Chem. Rev.*, 1990, **90**, 543.
7 B. E. Watkins, J. S. Kiely and H. Rapoport, *J. Am. Chem. Soc.*, 1982, **104**, 5702.

† All compounds showed satisfactory elemental analysis, ^1H and ^{13}C NMR spectra, and the expected molecular mass (FAB-MS).