ORGANIC LETTERS

2006
Vol. 8, No. 9
1897−1900

Hybridization of Pyrrolidinyl Peptide Nucleic Acids and DNA: Selectivity, Base-Pairing Specificity, and Direction of Binding

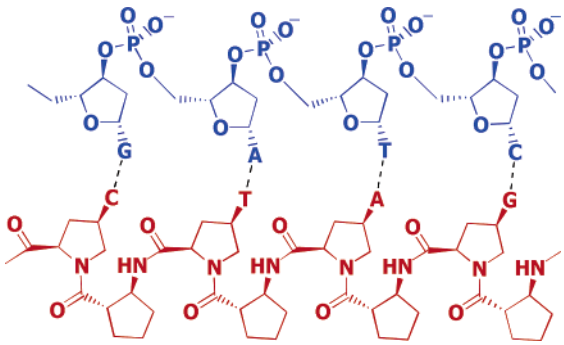
Tirayut Vilaivan* and Choladda Srisuwannaket

Organic Synthesis Research Unit, Department of Chemistry, Faculty of Science, Chulalongkorn University, Phayathai Road, Patumwan, Bangkok 10330, Thailand

vtirayut@chula.ac.th

Received February 21, 2006

ABSTRACT



A mixed-base, β-amino acid containing, pyrrolidinyl peptide nucleic acid (PNA) binds strongly and selectively to complementary DNA in an exclusively antiparallel fashion. The PNA–DNA binding specificity strictly follows the Watson–Crick base-pairing rules.

Peptide nucleic acid (PNA) is a class of DNA mimic systems consisting of nucleobases attached to a peptidelike backbone. The pioneering work of Nielsen on aminoethylglycyl PNA (aegPNA, **1**) revealed that it can hybridize with DNA and RNA to form extraordinary stable hybrids in a highly sequence-specific fashion.[1] The high stability was ascribed to the lack of electrostatic repulsion between the neutral PNA backbone and the negatively charged phosphate groups of nucleic acids.[2] PNA and its hybrids with nucleic acids are not substrates for most nucleic acid processing enzymes. These properties are not only interesting academically but also lead to many direct applications, mainly in the field of diagnostics and therapeutics.[3] Since the first report on **1** in 1991, a number of new PNA systems have been developed with the aim to gain further insight into the nature of strong binding affinity and specificity and to obtain new PNA systems with improved binding properties such as directional preference (parallel/antiparallel) and selectivity between DNA and RNA.[4] Most of these PNA designs are based on modification of the original aeg backbone of **1** by incorporating some chiral elements or structural rigidification at a few specific residues along the PNA strand.[5,6] To date, very few

(1) (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497−1500. (b) Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Nordén, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566−568.

(2) Tamac, S.; Sarkar, M.; Ratilainen, T.; Wittung, P.; Nielsen, P. E.; Nordén, B.; Gräslund, A. *J. Am. Chem. Soc.* **1996**, *118*, 5544−5552.

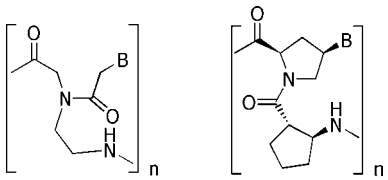
(3) Reviews: (a) Demidov, V. V. *Drugs Discovery Today* **2002**, *7*, 153−154. (b) Uhlman, E.; Peyman, A.; Breipohl, G.; Will, D. W. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2796−2823. (c) Nielsen, P. E. *Curr. Opin. Biotechnol.* **2001**, *12*, 16−20.

(4) Some recent examples: (a) Govindaraju, T.; Kumar, V. A.; Ganesh, K. N. *J. Am. Chem. Soc.* **2005**, *127*, 4144−4145. (b) Pokorski, J. K.; Witschi, M. A.; Purnell, B. L.; Appella, D. H. *J. Am. Chem. Soc.* **2004**, *126*, 15067−15073. (c) Ahn, D.-R.; Mosimann, M.; Leumann, C. J. *J. Org. Chem.* **2003**, *68*, 7693−7699. (d) Huang, Y.; Dey, S.; Zhang, X.; Sönnichsen, F.; Garner, P. *J. Am. Chem. Soc.* **2004**, *126*, 4626−4640. (e) Samuel, T. H.; Hickman, D. T.; Morral, J.; Beadham, I. G.; Micklefield, J. *Chem. Commun.* **2004**, 516−517.

10.1021/ol060448q CCC: $33.50
Published on Web 04/04/2006

PNAs not deriving from such partial modification of **1** have been evaluated for base-pairing specificity in detail.[7] Furthermore, none of these appeared to exhibit a comparable level of binding affinity and specificity toward DNA and RNA as **1**.

Proline is often used as a basic structural motif in the design of new PNA because of the possibility of conformational restriction of the PNA through a limited flexibility of the pyrrolidine ring.[6] We have recently reported a series of new β-pyrrolidinyl PNA (**2**) consisting of an alternating sequence of D-proline and a cyclic β-amino acid spacer (Figure 1).[8] The homothymine decamer of **2** with a (1*S*,2*S*)-



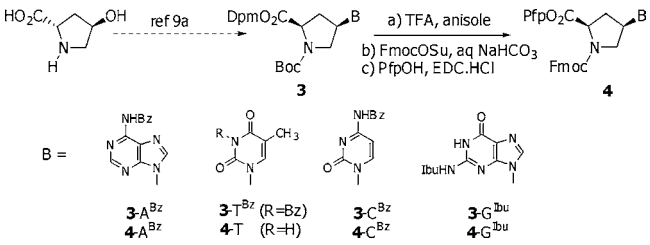
1 **2**

Figure 1. Structure of PNA **1** and **2**.

2-amino-1-cyclopentanecarboxylic acid (*SS*-ACPC) spacer was shown to bind with its complementary DNA to form a 1:1 hybrid with high affinity and sequence specificity.[8c] In this paper, we demonstrate that **2** incorporating all four DNA nucleobases (A, T, C, and G) exhibits a remarkable fidelity in recognition of DNA following Watson−Crick base-pairing rules (A·T, C·G). Furthermore, the PNA **2** with asymmetric sequences also shows a strong preference for binding to DNA in only one specific orientation. Finally, we disclose a preliminary result on an unusual preference of **2** in binding to DNA over RNA and over self pairing.

The four pentafluorophenyl (Pfp)-activated, Fmoc-protected proline nucleotide equivalents **4** were synthesized from *N*-Boc *C*-diphenylmethyl (Dpm)-protected monomers **3**[9] (Scheme 1). The protection scheme for nucleobases was chosen in accordance with standard oligonucleotide synthesis, i.e., none for thymine, benzoyl (Bz) for adenine (A) and cytosine (C), and isobutyryl (Ibu) for guanine (G). The mixed-sequence PNA was synthesized by stepwise coupling

Scheme 1. Synthesis and Structures of PNA Monomers[a]



[a] Abbreviations: Boc, *tert*-butoxycarbonyl; Bz, benzoyl; Dpm, diphenylmethyl; EDC, *N*-ethyl-*N*′′-(3-dimethylaminopropyl)carbodiimide; Ibu, isobutyryl; Fmoc, fluoren-9-ylmethoxycarbonyl; Pfp, pentafluorophenyl; Su, succinimidyl.

of the monomers (**4**) and Pfp-activated ACPC spacer (**5**)[10] using a standard Fmoc solid-phase peptide synthesis protocol.[8,9b] After the desired sequence was obtained, the N-terminal amino group was capped by acetylation. A lysine residue was also incorporated into the PNA sequence, usually at the C-terminus, to improve solubility and reduce nonspecific aggregation.[1a] The PNA was obtained after on-resin nucleobase deprotection (1:1 aq NH_3−dioxane, 55 °C), cleavage from the resin (TFA, rt), and reverse-phase HPLC purification. They were characterized by MALDI-TOF mass spectrometry. The sequences of PNAs synthesized and used in this study are as shown in Figure 2.

2a AAA AAA AAA
2b-e TTT TXT TTT
2b X=T **2c** X=A **2d** X=G **2e** X=C
2f GTA GAT CAC T
2g TGT ACG TCA CAA CTA

Figure 2. Sequences and abbreviations of PNA synthesized and used in this study. The base sequences are written from the N- to C-terminus. All PNA were N-acetylated and modified with L-LysNH_2 at C-termini except **2g** where N^{α}-acetyl-L-lysine was included at the N-terminus to facilitate purification by HPLC.

To investigate the base-pairing specificity between **2** and DNA, each of the four PNAs **2b**−**e** with one nucleobase variation at the central position was hybridized with one of the four DNA probes AAAA*Y*AAAA (Y = A, T, C, G). The highest four T_m values out of the possible sixteen pairs of PNA−DNA hybrids corresponded to the Watson−Crick-type base pairing (A·T, C·G) (Figure 3). The remaining mismatched hybrids had a T_m between 15 and 48 °C lower. Percent hyperchromicity for the mismatched hybrids was also significantly decreased (about 25−37% for a perfect match and 7−27% for a single mismatch) indicating a less favorable base stacking. Because **2** was shown to form only a 1:1 hybrid with DNA even with a thymine-rich sequence,[8c]

(5) (a) Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1939−1942. (b) Sforza, S.; Corradini, R.; Chirardi, S.; Dossena, A.; Marchelli, R. *Eur. J. Org. Chem.* **2000**, 2905−2913. (c) Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 6878−6879.

(6) (a) D'Costa, M.; Kumar, V.; Ganesh, K. N. *Org. Lett.* **2001**, *3*, 1281−1284. (b) Kumar, V. A. *Eur. J. Org. Chem.* **2002**, 2021−2032. (c) Kumar, V. A.; Ganesh, K. N. *Acc. Chem. Res.* **2005**, *38*, 404−412.

(7) (a) Li, Y.; Jin, T.; Liu, K. *Nucleosides, Nucleotides Nucleic Acids* **2001**, *20*, 1705−1721. (b) Kuwahara, M.; Arimitsu, M.; Shigeyasu, M.; Saeki, N.; Sisido, M. *J. Am. Chem. Soc.* **2001**, *123*, 4654−4658.

(8) (a) Vilaivan, T.; Suparpprom, C.; Harnyuttanakorn, P.; Lowe, G. *Tetrahedron Lett.* **2001**, *42*, 5533−5536. (b) Vilaivan, T.; Lowe, G. *J. Am. Chem. Soc.* **2002**, *124*, 9326−9327. (c) Suparpprom, C.; Srisuwannaket, C.; Sangvanich, P.; Vilaivan, T. *Tetrahedron Lett.* **2005**, *46*, 2833−2837.

(9) (a) Lowe, G.; Vilaivan, T. *J. Chem. Soc., Perkin Trans. 1* **1997**, 547−554. (b) Lowe, G.; Vilaivan, T. *J. Chem. Soc., Perkin Trans. 1* **1997**, 555−560.

(10) LePlae, P. R.; Umezawa, N.; Lee, H.-S.; Gellman, S. H. *J. Org. Chem.* **2001**, *66*, 5629−5632.

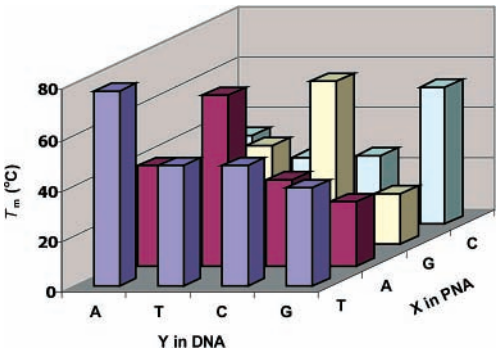


Figure 3. T_m of hybrids between PNA **2b**−**e** (TTTT*X*TTTT) and DNA (AAAA*Y*AAAA). Conditions: 10 mM sodium phosphate buffer, pH 7.0; PNA = DNA = 1.0 μM. For exact T_m figures, see Table S2 in the Supporting Information.

analysis of the results is much more straightforward than **1** which may also form triple helical (**1**)$_2$·DNA hybrids.[11]

We next moved on to PNA with random sequences, **2f** and **2g**,[1b] which allow the test for directional preference in binding (Table 1). Pleasingly, **2f** and **2g** formed stable

Table 1. T_m of PNA **2f** and **2g** with DNA[a]

PNA **2f**		PNA **2g**	
DNA (5′→3′)	T_m (°C)[b]	DNA (5′→3′)	T_m (°C)[c]
CATCTAGTGA[d]	<20	ACATGCAGTGTTGAT[d]	<20 (56)
AGTG*A*TCTAC[e]	57.6	TAGTTGT*G*ACGTACA[e]	78.6, 73.7 (69)
AGTG*T*TCTAC[f]	33.8	TAGTTGT*T*ACGTACA[f]	63.1, 59.1 (51)
AGTG*C*TCTAC[f]	31.0	TAGTTGT*A*ACGTACA[f]	65.0, 61.0 (49)
AGTG*G*TCTAC[f]	28.9	TAGTTGT*C*ACGTACA[f]	65.2, 61.1 (50)

[a] Conditions: [PNA] = [DNA] = 1.0 μM; 10 mM sodium phosphate buffer, pH 7.0. T_m values are accurate within ±0.5 °C. [b] Measured at [NaCl] = 0 mM. [c] Measured at [NaCl] = 0, 100 mM; values in parentheses are the corresponding T_m of PNA **1** at [NaCl] = 100 mM taken from ref 1b. [d] Perfect match parallel. [e] Perfect match antiparallel. [f] Single mismatch antiparallel.

hybrids with complementary DNA only in an antiparallel fashion.[12] The T_m's of both antiparallel perfectly matched **2**·DNA hybrids were slightly higher than the corresponding **1**·DNA hybrids at the same salt concentration. The mixtures between **2f** and **2g** with parallel complementary DNAs showed no observable melting. The antiparallel **2f**·DNA hybrid showed a distinct induced CD signal, whereas the parallel hybrid exhibited no spectral change, confirming the lack of parallel hybrid formation (Figure 4). The high directional preference is probably a direct consequence of the stereoregular nature of **2**, which does not exist in **1**.[13]

(11) Egholm, M.; Christensen, L.; Behrens, C.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Chem. Soc., Chem. Commun.* **1993**, 800−801.

(12) The term antiparallel used here means the pairing between PNA and DNA in such a way that the N-terminus of PNA faces the 3′-terminus of DNA and the C-terminus of the PNA faces the 5′-terminus of the DNA. The opposite orientation of binding is therefore parallel.

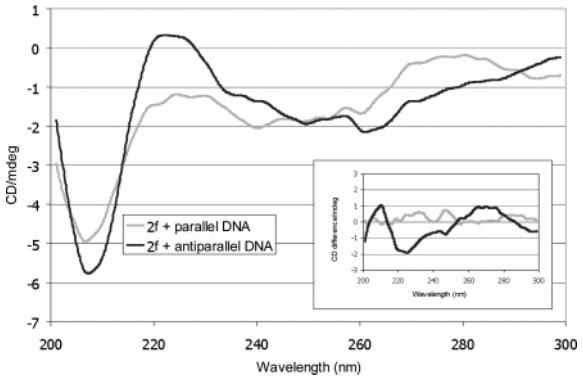


Figure 4. CD spectra of 1:1 mixtures of **2f** with parallel and antiparallel complementary DNA. The inset shows the difference CD spectra (observed − sum of CD spectra of all components). Conditions: [PNA] = 1 μM; 10 mM sodium phosphate buffer, pH 7.0; 25 °C.

Introduction of a mismatched base to the DNA sequence caused a reduction of T_m of the **2f**·DNA hybrid in a position and type-of-mismatch dependent fashion. As expected, the effect was strongest at the middle of the sequence (data not shown). From Table 1, it is clear that the high Watson−Crick base-pairing specificity was still retained in this PNA with random sequences. A significant drop in T_m was observed for singly mismatched hybrids (**2f**: 23.8−28.7 °C; **2g**: 12.6−16.5 °C). This indicates that, despite high affinity, the specificity was not compromised. The large difference in stability of perfectly matched and mismatched **2**·DNA duplexes suggests that **2** could be useful as a DNA sequence probe.

Thermal denaturation experiments of the hybrids of homothymine (**2b**) as well as mixed-sequence PNA (**2f**,**g**) and their complementary nucleic acids were also carried out to investigate the binding selectivity among different types of nucleic acids (Table 2). The affinity of the homothymine

Table 2. T_m of PNA **2** with Complementary Nucleic Acids[a]

entry	PNA	DNA, RNA, or PNA	T_m (°C)
1	**2b**	poly(dA)	>80
2	**2b**	poly(rA)	37.7
3	**2g**	d(TAGTTGTGACGTACA)	78.6
4	**2g**	r(UAGUUGUGACGUACA)	63.2
5	**2b**	**2a**	<20
6	**2f**	**2**-(Ac-AGTGATCTAC-LysNH_2)	<20
7	**2f**	**2**-(Ac-CATCTAGTGA-LysNH_2)	<20

[a] Conditions: PNA = 1.0 μM; PNA/nucleic acids = 1:1 (as nucleotides); 10 mM sodium phosphate buffer, pH 7.0. T_m values are accurate within ±0.5 °C.

oligomer **2b** toward complementary DNA was much higher than RNA (entries 1 and 2). This higher affinity for DNA

(13) An exclusive antiparallel hybrid formation was observed in aegPNA **1** with its backbone partially modified by a few D-lysine "chiralbox" residues (ref 5b).

appears to be general because a significant T_m difference (15.4 °C) between hybrids of the mixed oligomer **2g** with complementary DNA and RNA was also observed (entries 3 and 4). The lack of binding between the two complementary strands of **2** (entry 5) was confirmed by both T_m and CD spectroscopy. Upon addition of a complementary DNA probe (dA_9 or dT_9) to the mixture between **2a** and **2b**, a CD spectral change was evidenced. This suggested that both strands of **2** were free to hybridize with DNA even in the presence of their complementary strands.[14] The mixed oligomer **2f** also failed to bind with its complementary PNA sequences in both antiparallel and parallel orientations (entries 6 and 7) according to T_m measurements.

The preference for binding of **2** to DNA over RNA is in sharp contrast to other PNA systems, which usually prefer binding to RNA.[1b,4a,e] The same phenomenon was observed in a PNA related to **2** with a D-aminopyrrolidinecarboxylic acid as the spacer.[8b] The reason for a large difference between T_m (**2**·DNA) and T_m (**2**·RNA) in the homooligomer compared to the mixed sequences is not fully understood at present, but we have confirmed by UV titration that in all cases only duplexes were formed. The lack of self pairing between two complementary strands of **2** is rather unexpected and, to the best of our knowledge, has not been reported in other PNA systems. The basis of this unique selectivity will be the subject of future investigations.

In conclusion, we have synthesized a PNA **2** with all four nucleobases incorporated and demonstrated several interesting features of this PNA system. Apart from forming very stable PNA·DNA hybrids with a stringent Watson−Crick base-pairing specificity, **2** shows a strong preference for binding to DNA over RNA and over self pairing in an antiparallel orientation. These unique and desirable properties render **2** as a potential candidate for diagnostic, and possibly therapeutic, applications.

Acknowledgment. Financial support of this work comes from The Thailand Research Fund and The Commission on Higher Education (RSA/04/2545 and RMU4880030, both to T.V.) and from Chulalongkorn University (Rachadapisek Sompoj Endowment to OSRU).

Supporting Information Available: Experimental details, NMR spectra of PNA monomers, and UV and CD spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

OL060448Q

(14) The following T_m values were obtained at 1 μM PNA and DNA, 10 mM sodium phosphate buffer, pH 7.0: **2a** + dT_9 = 79.9 °C, **2a** + **2b** + dT_9 = 78.7 °C, **2b** + dA_9 = 77.3 °C, **2a** + **2b** + dA_9 = 76.9 °C.