DOI: 10.1002/cbic.200800441

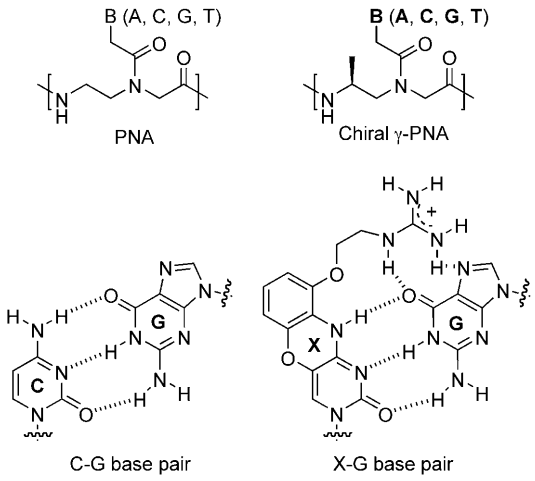
A Simple Cytosine to G-Clamp Nucleobase Substitution Enables Chiral γ-PNAs to Invade Mixed-Sequence Double-Helical B-form DNA

Venugopal Chenna,[a, b] Srinivas Rapireddy,[a] Bichismita Sahu,[a] Cristina Ausin,[c] Enrique Pedroso,[c] and Danith H. Ly*[a]

Nature uses Watson–Crick base pairings as a means to store and transmit genetic information because of their high fidelity. These specific A–T (or A–U) and G–C nucleobase interactions, in turn, provide chemists and biologists with a general paradigm for designing molecules to bind to DNA and RNA. With knowledge of the sequence information, one can design oligonucleotides to bind to just about any part of these biopolymeric targets simply by choosing the corresponding nucleobase sequence according to these digital base-pairing rules. Although conceptually simple, such principles in general can only be applied to the recognition of single-stranded DNA or RNA, but not the double-stranded form. The reason is that in double-helical DNA (or RNA) not only are the Watson–Crick faces of the nucleobases already occupied, they are buried within the double helix.[1] Such molecular encapsulation imposes a steep energetic barrier on the designer molecules. To establish binding, not only would they need to be able to gain access to the designated nucleobase targets, which are blocked by the existing base pairs, they would also need to be able to compete with the complementary DNA strand to prevent it from re-annealing with its partner—a task that has rarely been accomplished by any class of molecules.

To circumvent this challenge, most of the research effort to date has been focused on establishing principles for recognizing chemical groups in the minor and major groove instead because they are more readily accessible and energetically less demanding.[2] While impressive progress has been made on this front, especially in the development of triplex-forming oligonucleotides,[3–5] polyamides,[6–8] and zinc-finger-binding peptides,[9–13] the issues of sequence selection, specificity and/or target length still remain unresolved for many of these approaches.[13–16] In the last decade, however, several studies have shown that peptide nucleic acid (PNA), a particular class of nucleic acid mimics that are comprised of pseudopeptide backbone (Scheme 1), could invade double-stranded DNA (dsDNA).[17–21] This finding is significant because it demonstrates that the same Watson–Crick base-pairing principles that guide the recognition of single-stranded DNA and RNA can also be applied to dsDNA. Aside from the simplicity, this recognition strategy is general and could potentially be applied to any sequence or target length, just as in the recognition of single-stranded DNA or RNA. The downside to this approach, however, is that PNA can only recognize homopurine and homopyrimidine targets. Mixed-sequence PNAs do not have sufficient binding free energy to invade double-helical B-DNA. Though a double-duplex invasion strategy has been developed to try to overcome this energetic barrier,[21] the issue of sequence selection still remains due to the unresolved issue with self-quenching.[22] In this Communication, we show that a simple nucleobase substitution, that is, replacing cytosine with a G-clamp (Scheme 1) provides the necessary energetics for chiral γ-PNAs to invade mixed-sequence B-DNA. Unlike the double-duplex invasion strategy, which requires two strands of PNAs, only a single strand of γ-PNA is required for binding to B-DNA.

Recently, we showed that randomly folded, single-stranded PNAs can be preorganized into a right-handed helix simply by installing an L-alanine-derived, *S* chiral center at the γ-position of the *N*-(2-aminoethyl)glycine backbone unit.[23] These helical γ-PNAs exhibit strong binding affinity and sequence selectivity for DNA and RNA and are capable of invading mixed-sequence



Sequences of PNA (PNA1) and γ-PNA (**PNA2-4**) oligomers
PNA1: H-^{L}Lys-GACCACAGAT-^{L}Lys-NH_2
PNA2: H-^{L}Lys-**GACCACAGAT**-^{L}Lys-NH_2
PNA3: H-^{L}Lys-**GAXCACAGAT**-^{L}Lys-NH_2
PNA4: H-^{L}Lys-**GAXCAXAGAT**-^{L}Lys-NH_2

Scheme 1. Chemical structures of PNA, γ-PNA, C–G and X–G base pairs along with the sequences of the oligomers used in this study. Bold letters indicate γ-backbone modifications.

[a] *Dr. V. Chenna, Dr. S. Rapireddy, Dr. B. Sahu, Prof. Dr. D. H. Ly*
Department of Chemistry, Carnegie Mellon University
4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213 (USA)
Fax: (+1) 412–268–5579
E-mail: dly@andrew.cmu.edu

[b] *Dr. V. Chenna*
Present address: Department of Oncology
Johns Hopkins Medical Institutes, Baltimore, Maryland 21231 (USA)

[c] *Dr. C. Ausin, Prof. Dr. E. Pedroso*
Departament de Quimica Organica IBUB, Universitat de Barcelona
Marti I Franques 1–11, 08028 Barcelona (Spain)

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

dsDNA as demonstrated by in situ footprinting (data not shown), however, the invasion complex is not sufficiently stable under prolonged electrophoresis, a necessary condition to separate the bound from the unbound DNA target.[24] This result indicates that the binding free energies of these γ-PNAs are already within the invasion threshold of B-DNA. The additional binding free energies that are needed to stabilize the invasion complex, therefore, might not be much more and as such, could potentially be attained by replacing a cytosine nucleobase with 9-(2-guanidinoethoxy) phenoxazine (G-clamp, X)—a cytosine analogue that can form five H-bonds with guanine in addition to providing extra base-stacking as a result of the expanded phenoxazine ring system.[25] Prior studies showed that a single C→X nucleobase substitution can enhance the melting transition of a PNA–DNA duplex by as much as 23 °C.[26,27] This level of stabilization might be sufficient for γ-PNAs to invade and stabilize the B-DNA invasion complex. To test this hypothesis, we synthesized a series of dodecameric γ-PNA oligomers (**PNA2** through **4**) as shown in Scheme 1, in which the cytosine nucleobase was systematically replaced with G-clamp, and characterized their conformation, thermal stability and DNA-strand-invasion capabilities by using a combination of spectroscopic and biochemical techniques.

To confirm that these γ-PNA oligomers adopted helical structures, we measured the CD spectra of **PNA2** through **4** and compared it to that of the unmodified PNA (PNA1). Consistent with our earlier result,[24] we did not observe noticeable CD signals for PNA1 in the 220–300 nm nucleobase absorption regions (Figure 1); this indicates a lack of helical structure. We ruled out the possibility of PNA existing with an equal proportion of a left- and right-handed helix, as previously suggested by MD simulations,[28] on the basis of multidimensional NMR spectroscopic analyses.[23] In the case of **PNA2** through **4**, however, we observed pronounced CD signals, with biphasic exciton coupling patterns that are characteristic of a right-handed PNA–DNA double helix.[29] Substituting C with X did not appear to have a significant effect on the overall conformation of the oligomers, as judged from the similarities in the CD profiles of **PNA3** and **PNA4** to that of **PNA2**. A small difference in the amplitudes in the 200–250 nm regions could be attributed to the difference in the absorption strength of the cytosine and G-clamp nucleobase, and/or variations in the backbone conformation as the result of these oligomers trying to accommodate the more sterically hindered G-clamp nucleobase. In addition to CD measurements, we have also attempted to record the melting transitions (T_m values) of the corresponding PNA–DNA hybrid duplexes for this particular series of oligomers, but to no avail because the T_m values were too high to be measured accurately by UV spectroscopy. The T_m of the unsubstituted **PNA2** with its complementary (dodecameric) DNA strand alone, at 5 μM strand concentration each in 10 mM NaPi buffer, is already in excess of 90 °C.[24] Because the binding affinities of these γ-PNAs are already exceptionally high, replacing C with X should enable them to bind to DNA with an even greater affinity—perhaps sufficient to invade and form a stable complex with double-helical B-DNA.

To determine whether these γ-PNA oligomers can invade dsDNA, we performed an electrophoretic mobility shift assay. A 171 bp PCR fragment that contained an internal binding site (Scheme S1 in the Supporting Information) was incubated with different concentrations of γ-PNA oligomers in 10 mM sodium phosphate buffer (pH 7.4) at 37 °C for 2 h. The mixtures were then separated on nondenaturing polyacrylamide gel (PAGE) and stained with SYBR-Gold for visualization. Consistent with the earlier finding,[24] our result showed no evidence of binding for **PNA2** (Figure 2, lanes 2 and 3). In the case of **PNA3** and **PNA4**, however, we noticed a distinct slow-moving band the intensity of which gradually increased with increasing oligomer concentrations (Figure 2, compare lanes 5 with 4, and lanes 7 with 6). Formation of this complex appeared to be complete at a PNA/DNA ratio of 3:1 for **PNA4**, as judged from the nearly complete disappearance of the unbound DNA target. Under identical conditions, only ~20% binding was complete for **PNA3**. In addition to the binding efficiency, the mobility of the shifted complex also appeared to be different for the two oligomers; it was lower for **PNA4** than for **PNA3**. The difference in mobility could be attributed to the difference in the overall

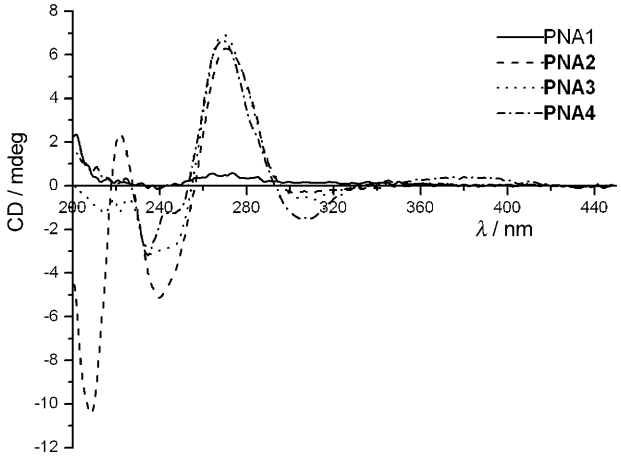


Figure 1. CD spectra of single-stranded PNA (PNA1) and γ-PNA (**PNA2** through **4**) oligomers at 5 μM strand concentration each in 10 mM NaPi buffer (pH 7.4), recorded at room temperature.

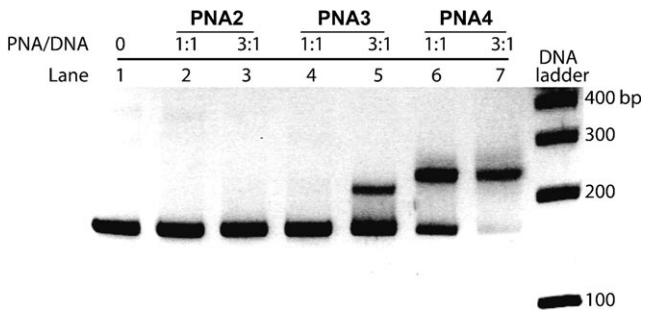


Figure 2. Gel-shift assay of a 171 bp DNA fragment that contained a perfectly matched target with various concentrations of γ-PNA oligomers. DNA (0.4 μM duplex concentration) was incubated with 0 (lane 1), 0.4 (lanes 2, 4 and 6) and 1.2 μM (lanes 3, 5 and 7) of each respective γ-PNA oligomer in 10 mM NaPi buffer (pH 7.4) at 37 °C for 2 h. The mixtures were separated on 10% nondenaturing PAGE for 3 h at 5 $V cm^{-1}$ and then stained with SYBR-Gold.

size and charge density of the complex, and reflects the difference in the number of the G-clamps that are present in the oligomers (two for **PNA4** and one for **PNA3**). Formation of the shifted band was only observed in the presence of a perfectly-matched target (Figure 3, lanes 1–3). Introduction of an inverted, single-base mismatch (Scheme S1) completely abolished the binding (Figure 3, lanes 4–6). This result shows that **PNA4** binding occurred in a sequence-specific manner.

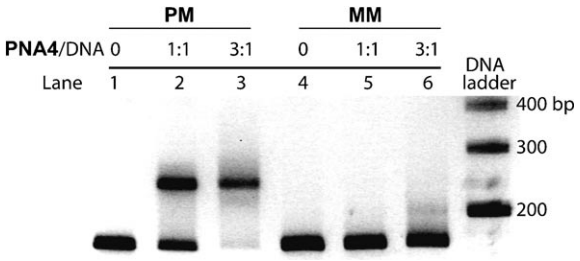


Figure 3. Gel-shift assay of **PNA4** with perfect-match (**PM**) and single-base-mismatch (**MM**) DNA targets. Samples were prepared by incubating 0.4 μM of 171 bp DNA fragment that contained **PM** (lanes 1–3) and **MM** (lanes 4–6) targets with 0 (lanes 1 and 4), 0.4 (lanes 2 and 5) and 1.2 μM (lanes 3 and 6) of **PNA4** in 10 mM NaPi buffer (pH 7.4) at 37 °C for 2 h. The samples were separated and stained under identical condition as before.

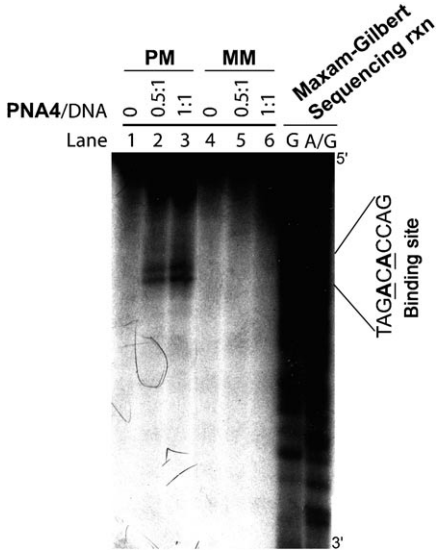


Figure 4. DEPC treatment after incubation of **PNA4** with **PM** (lanes 1–3) and **MM** (lanes 4–6) DNA targets. The homologous DNA strand was 3′-labeled with P-32. The samples were prepared by incubating 10000 cpm of the labeled, and 0.4 μM of the cold (unlabeled) DNA with 0, 0.2, and 0.4 μM of **PNA4** in 10 mM NaPi buffer at 37 °C for 2 h, followed by DEPC treatment. The samples were separated on 10% denaturing PAGE, and the cleavage patterns were visualized by autoradiography.

To further demonstrate that the observed binding occurred through a strand invasion mechanism, we performed a diethyl pyrocarbonate (DEPC) probing assay. DEPC is an acylating reagent that is known to react selectively with adenines, and to a lesser extent with guanines within a single-stranded or a perturbed region of a DNA duplex, which can be revealed in the form of strand cleavage following piperidine treatment.[30,31] Sequence-specific strand invasion of dsDNA by **PNA4** is expected to result in a local displacement of the homologous DNA strand, which can be readily detected by the DEPC assay. Consistent with the strand invasion mechanism, DEPC treatment of the bound DNA complex, in which the 3′-end of the homologous DNA strand was labeled with P-32, revealed selective cleavage at the adenine residues on the homologous DNA strand across from the binding site (Figure 4). Strand cleavage was observed even at a PNA/DNA ratio of 0.5:1, but was less intense than at a 1:1 ratio (Figure 4, lanes 2 and 3). Although there are four adenines within the expected locally displaced DNA strand, only the two middle adenine residues showed significant cleavage pattern, but very little, if any, was observed for the two residues localized near the termini. The difference in the cleavage intensity could be attributed to the difference in the degree of interaction between these adenine residues and their neighboring nucleobases. Because they are located within the melted junction of the DNA duplex, the two flanking adenine residues could still be interacting with the adjacent nucleobases, or at least to a greater extent than those that reside in the middle of the looped-out strand. This might explain why only the middle adenines are more susceptible to DEPC treatment than those near the termini. Similar to the data obtained from the gel-shift assay, strand cleavage was only observed in the presence of the perfect-match target (Figure 4, compare lanes 1–3 with lanes 4–6). These results, taken together, show that **PNA4** binding occurred in a sequence-specific manner through a strand-invasion mechanism.

Next, we investigated the effect of ionic strengths on **PNA4** binding because salt has been shown to have a profound effect on the efficiency of DNA strand invasion.[19,32,33] To assess this effect, we incubated **PNA4** with the perfect-match DNA target at a 3:1 (PNA/DNA) ratio in buffers that contained 10 mM NaPi and various concentrations of $MgCl_2$, including a simulated physiological salt concentration (150 mM KCl and 2 mM $MgCl_2$).[34] The mixtures were separated by non-denaturing PAGE and stained with SYBR-Gold as before. Inspection of Figure 5 reveals that whereas the binding efficiency of **PNA4** with its target gradually decreased with increasing Mg^{2+} concentrations, a significant fraction (~60%) of the DNA target

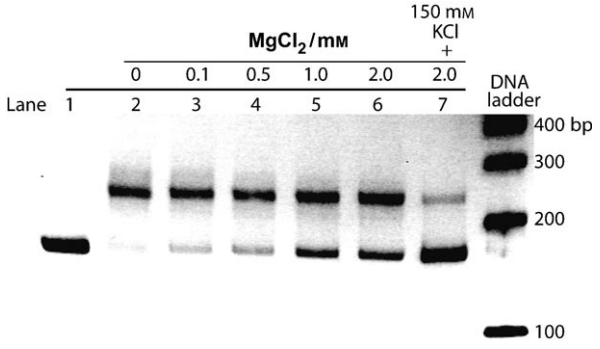


Figure 5. Binding of **PNA4** with **PM** DNA target at a 3:1 ratio in buffer that contained 10 mM NaPi and various concentrations of $MgCl_2$ (lanes 3–6), along with a simulated physiological salt concentration (lane 7). The samples were incubated at 37 °C for 2 h prior to separation and staining as before.

was found in the bound state after 2 h of incubation at 37 °C in 10 mM NaPi and 2 mM $MgCl_2$—this corresponds to less than a twofold decrease in the binding efficiency as compared to that in the original buffer with just 10 mM NaPi (Figure 5, compare lanes 6 with 2). Although the amount of the complex that was formed under physiological salt concentration was significantly less (only ~10%) than that in the 10 mM NaPi baseline, strand invasion still took place under this condition as evident from the presence of the retarded band (Figure 5, compare lanes 7 with 2). This result indicates that it is feasible to invade double-helical B-DNA under physiological temperature and ionic strength given that sufficient binding free energy is provided.

In summary, we have shown that γ-PNAs that contain G-clamp in place of cytosine nucleobase can invade mixed-sequence B-DNA. In this case, only a single strand of γ-PNA is required, and the invasion occurs through direct Watson–Crick base pairings. In 10 mM NaPi buffer and at 37 °C, the invasion of **PNA4** into DNA was complete within 2 h of incubation at a PNA/DNA ratio of 3:1. We attributed these relatively fast invasion kinetics to the preorganized structure of γ-PNAs and the stability of the invasion complex to the enhanced binding affinity of the X–G base pair. Although the efficiency was lower, strand invasion still took place at elevated ionic strengths. This study shows that it is feasible to target mixed-sequence B-DNA with γ-PNAs not only at relatively low ionic strengths, but also at physiological salt concentration. Presently however, it is not clear whether the reduction in the invasion efficiency is predominantly the result of a slower rate of base-pair "breathing", or an increase in the thermodynamic stability of the dsDNA because both are interrelated and significantly affected by ionic strengths. A prior study with homopyrimidine PNA seems to suggest that it is the former.[35] Whereas the initial searching steps for the two systems may be similar, the overall invasion efficiency is likely to be different, that is, faster for the mixed-sequence γ-PNA because only a single strand is required for binding and stabilization, as compared to two strands for homopyrimidine PNA. Additional studies will be required to tease out these various contributions. Our ability to fine-tune the energetics of γ-PNAs without the need to change the nucleobase sequence or the oligomer size, by replacing C with X (or other nucleobases with their more thermodynamically stable synthetic analogues), should enable us to address this question in the future, which is an important first step toward developing a more effective oligonucleotide platform for targeting double-stranded B-DNA on the basis of Watson–Crick pairing.

Acknowledgements

Financial supports for this work were provided in part by the National Institutes of Health (GM076251-01) to D.H.L., and Ministerio de Educacion y Ciencia (CTQ2007-680-C02-01) and Generalitat de Catalunya (2005SGR-693 and XRB) to E.P.

Keywords: chiral gamma-PNA · DNA recognition · DNA strand invasion · helical PNA · nucleic acids

[1] H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 2179–2183.
[2] P. E. Nielsen, *Bioconjugate Chem.* **1991**, *2*, 1–12.
[3] H. E. Moser, P. B. Dervan, *Science* **1987**, *238*, 645–650.
[4] J. S. Sun, C. Helene, *Curr. Opin. Struct. Biol.* **1993**, *3*, 345–356.
[5] C. Giovannangeli, C. Helene, *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 413–421.
[6] J. M. Gottesfeld, L. Neely, J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1997**, *387*, 202–205.
[7] S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature* **1998**, *391*, 468–471.
[8] C. C. O'Hare, D. Mack, M. Tandon, S. K. Sharma, J. W. Lown, M. L. Kopka, R. E. Dickerson, J. A. Hartley, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 72–77.
[9] H. A. Greisman, C. O. Pabo, *Science* **1997**, *275*, 657–661.
[10] M. Moore, A. Klug, Y. Choo, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 1437–1441.
[11] C. W. Garvie, C. Wolberger, *Molecules* **2001**, *8*, 937–946.
[12] R. R. Beerli, C. F. I. Barbas, *Nat. Biotechnol.* **2002**, *20*, 135–141.
[13] D. Jantz, B. T. Amann, G. J. Gatto, J. M. Berg, *Chem. Rev.* **2004**, *104*, 789–800.
[14] D. M. Gowers, K. R. Fox, *Prog. Nucleic Acid Res. Mol. Biol.* **1999**, *27*, 1569–1577.
[15] J. J. Kelly, E. E. Baird, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6981–6985.
[16] C. L. Kielkopf, E. E. Baird, P. D. Dervan, D. C. Rees, *Nat. Struct. Biol.* **1998**, *5*, 104–109.
[17] P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497–1500.
[18] J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G. Au, S. G. Carter, D. A. Ruckenstein, A. L. Boyd, S. A. Noble, L. E. Babiss, *Science* **1992**, *258*, 1481–1485.
[19] N. J. Peffer, J. C. Hanvey, J. E. Bisi, S. A. Thomson, C. F. Hassman, S. A. Noble, L. E. Babiss, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10648–10652.
[20] T. Ishihara, D. R. Corey, *J. Am. Chem. Soc.* **1999**, *121*, 2012–2020.
[21] P. E. Nielsen, *Acc. Chem. Res.* **1999**, *32*, 624–630.
[22] V. V. Demidov, E. Protozanova, K. I. Izvolsky, C. Price, P. E. Nielsen, M. D. Frank-Kamenetskii, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5953–5958.
[23] A. Dragulescu-Andrasi, S. Rapireddy, B. M. Frezza, C. Gayathri, R. R. Gil, D. H. Ly, *J. Am. Chem. Soc.* **2006**, *128*, 10258–10267.
[24] S. Rapireddy, G. He, S. Roy, B. A. Armitage, D. H. Ly, *J. Am. Chem. Soc.* **2007**, *129*, 15596–15600.
[25] C. J. Wilds, M. A. Maier, V. Tereshko, M. Manoharan, M. Egli, *Angew. Chem.* **2002**, *114*, 123–125; *Angew. Chem. Int. Ed.* **2002**, *41*, 115–117.
[26] K. G. Rajeev, M. A. Maier, E. A. Lesnik, M. Manoharan, *Org. Lett.* **2002**, *4*, 4395–4398.
[27] J.-A. Ortega, J. R. Blas, M. Orozco, A. Grandas, E. Pedroso, J. Robles, *Org. Lett.* **2007**, *9*, 4503–4506.
[28] S. Sen, L. Nilsson, *J. Am. Chem. Soc.* **2001**, *123*, 7414–7422.
[29] M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, P. E. Nielsen, *Nature* **1993**, *365*, 566–568.
[30] P. M. Scholten, A. Nordheim, *Nucleic Acids Res.* **1986**, *14*, 3981–3993.
[31] J. C. Furlong, D. M. J. Lilley, *Nucleic Acids Res.* **1986**, *14*, 3995–4007.
[32] D. Y. Cherny, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, M. Egholm, O. Buchardt, R. H. Berg, P. E. Nielsen, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1667–1670.
[33] A. Kurakin, H. J. Larsen, P. E. Nielsen, *Chem. Biol.* **1998**, *5*, 81–89.
[34] S. W. Santoro, G. F. Joyce, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4262–4266.
[35] V. V. Demidov, M. V. Yavnilovich, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, P. E. Nielsen, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2637–2641.

Received: June 30, 2008
Published online on September 24, 2008

