Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 16 (2006) 4931–4935

Synthesis of cell-permeable peptide nucleic acids and characterization of their hybridization and uptake properties

Peng Zhou,[a] Anca Dragulescu-Andrasi,[a] Birendra Bhattacharya,[b] Heather O'Keefe,[b] Paolo Vatta,[b] Jens J. Hyldig-Nielsen[b] and Danith H. Ly[a,*]

[a]*Department of Chemistry, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA*
[b]*Applied Biosystems, 850 Lincoln Center Drive, Foster City, CA 94404, USA*

Received 16 May 2006; revised 13 June 2006; accepted 14 June 2006
Available online 30 June 2006

Abstract—Guanidine-based peptide nucleic acid (GPNA) monomers and oligomers containing all four natural (adenine (A), cytosine (C), guanine (G), and thymine (T)) and two unnatural (2-thiouracil (^{s}U) and 2,6-diaminopurine (D)) nucleobases have been synthesized. Thermal denaturation study showed that GPNA oligomers containing alternate D-backbone configuration bind sequence-specifically to DNA and, when incubated with mammalian cells, localized specifically to the endoplasmic reticulum (ER).

Peptide nucleic acid (PNA) is a synthetic analogue of DNA and RNA, developed more than a decade ago in which the naturally occurring sugar phosphate backbone has been replaced by the *N*-(2-aminoethyl) glycine units.[1] PNA can hybridize to complementary DNA or RNA strand through Watson–Crick base-pairing;[2] but unlike DNA or RNA which is susceptible to enzymatic degradation, PNA is not easily degraded by proteases or nucleases.[3] These properties make PNA an attractive reagent for biotechnology applications.

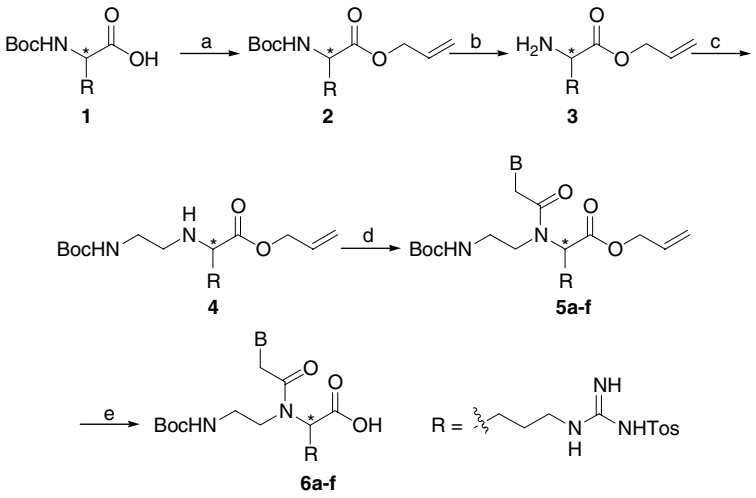
While several PNA-based applications have been successfully developed,[4–6] the full potential of PNA for in vivo applications has not yet been realized—much of which could be attributed to its poor cellular uptake property. Considerable effort has been vested in the last decade trying to develop means to transport PNA into cells. Several strategies,[7,8] including microinjection,[9] electroporation,[10] DNA-mediated transduction,[11] and covalent attachment to cell-penetrating peptides (CPPs),[12,13] have been developed, but for the most part they are limited to small-scale experimental setups. The latter approach, although promising since no delivery system is required, is presently limited by uptake efficiency and intracellular distribution. Most PNAs delivered this way are trapped in the endocytic vesicles.[7] In addition, since these PNA–CPP constructs are amphipathic in nature, they could cause cytotoxic effects since most amphipathic peptides found in nature, such as magainin and mellitin, are known to be cytotoxic.[14]

Recently, we have shown that arginine-derived, polythymine PNA oligomers (commonly known as GPNAs) were readily taken-up by mammalian cells.[15] Herein, we report the synthesis of GPNA monomers for all four natural (A, C, G, and T) and two unnatural (^{s}U and D) nucleobases and the corresponding oligomers, along with their hybridization and cellular uptake properties. The synthetic strategy reported herein has two distinct advantages over the conventional design, where the transduction domains are covalently linked either to the C- or N-terminus of PNA. First, it minimizes the amphipathic character of the resulting complex, thereby reducing its cytotoxic effects. And second, it eliminates the extra synthetic steps required for appending molecular transporters to the termini of PNAs. GPNAs containing alternate *N*-(2-aminoethyl) D-arginine and *N*-(2-aminoethyl) glycine backbone units bind sequence-specifically to DNA (and RNA),[16] and are taken-up by mammalian cells and localized specifically to the endoplasmic reticulum (ER).

GPNA monomers were synthesized according to the procedure outlined in Scheme 1. A similar procedure has been reported previously,[17,18] but this particular protocol employed dicarboxybenzyl (diCbz)-protection for the arginine side chain and was limited to thymine

Keywords: Cell-permeable; Guanidine-based peptide nucleic acid; GPNA; Monomers; Oligomers; Synthesis; Endoplasmic reticulum; ER.
* Corresponding author. E-mail: dly@andrew.cmu.edu

0960-894X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2006.06.052

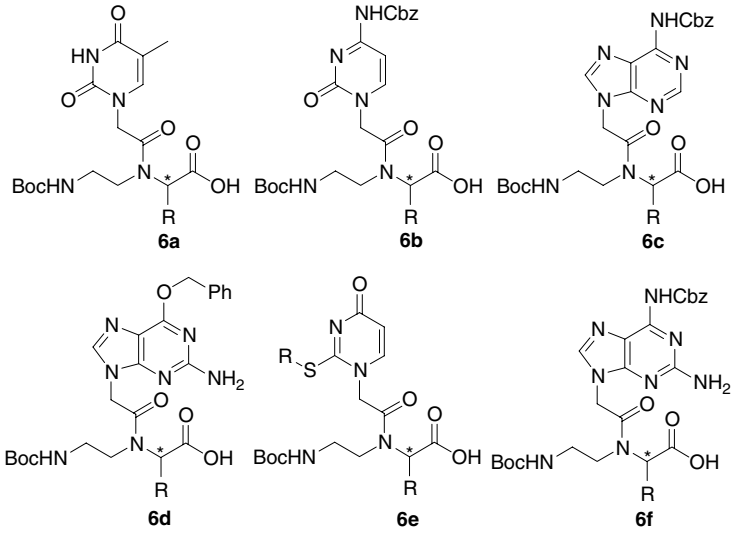


Scheme 1. Synthesis of GPNA monomers. Reagents and conditions: (a) CH_2CHCH_2Br, $NaHCO_3$, DMF, rt, overnight (>95% yield); (b) *p*-toluene sulfonic acid, benzene, reflux, 2 h (>95% yield); (c) Boc-$NHCH_2CHO$, MeOH, 4 °C, 4 h, then acetic acid, $NaBH_3CN$, 30 min (65% yield); (d) carboxymethylnucleobase, DCC, DhbtOH, DMF, 40 °C, 24 h (50–62% yield); (e) $Pd(PPh_3)_4$, morpholine, THF, rt, 2 h (90% yield).

monomer. Attempt to prepare the remaining three monomers (adenine, cytosine, and guanine) using this protocol was met with only limited success due to the poor solubility of the final monomers in DMF and NMP—solvents commonly used in solid-phase synthesis. By switching to tosyl (Tos) group, we were able to synthesize monomers for all four natural and two unnatural nucleobases that are soluble in DMF and NMP. These monomers showed comparable coupling efficiency to that of the unmodified PNA monomers on solid-support.[19]

The starting material **1** (either L- or D-arginine amino acid residue) was first protected with allylic ester, followed by removal of the Boc-protecting group. The resulting amine **3** was then coupled with 2-aminoacetaldehyde via reductive amination to afford the backbone intermediate **4**. The final monomers (**6a**–**f**, Scheme 2) were prepared by coupling the backbone intermediate with each respective carboxymethylnucleobase, which was prepared according to established procedures,[19] followed by the final deprotection step with $Pd(PPh_3)_4$.

GPNA oligomers were synthesized according to standard Boc-protected PNA chemistry,[20] using cross-linked polystyrene beads functionalized with 4-methylbenzhydrylamine (MBHA) as supporting resin. Chain extension followed a three-step cycle: (i) removal of the Boc-protecting group from the terminal amine with TFA, (ii) coupling of the next monomer onto the N-terminus of the growing chain with HBTU, and (iii) capping of the unreacted amines with acetic anhydride. In an attempt to minimize the monomer consumption, only a fourth of the recommended amount of monomer was used in each coupling step. This ratio provided sufficient amount of oligomers required for most hybridization



Scheme 2. GPNA monomers containing natural and unnatural nucleobases synthesized. R = $CH_2C_6H_6$–*p*-OMe.

and cellular uptake experiments. Treating the solid resin with *m*-cresol/thioanisole/TFMSA/TFA (1:1:2:6) mixture resulted in simultaneous removal of the protecting groups and cleavage of the oligomers from the resin. (Note that HF is not required to remove the Tos group.) The crude oligomers were then precipitated with ethyl ether and purified by HPLC (preparative Waters column at 40 °C) using 0.1% TFA/acetonitrile/water mixture and characterized by MALDI-TOF mass spectrometry. Figure 1 shows a representative HPLC trace of the crude oligomer, along with the corresponding MALDI-TOF spectrum of the purified product. The peak labeled GPNA8* indicates the presence of a second conformer, detected only at relatively high concentration.

To determine the effects of arginine side chains, backbone configurations, and spacing on the hybridization properties of PNA, we measured the melting transitions (T_ms) of various GPNA–DNA hybrid duplexes using UV-spectroscopic technique and compared to that of PNA–DNA. The results are tabulated in Table 1. Similar to the observations made with RNA,[16] the T_ms of the GPNA–DNA hybrid duplexes with consecutive backbone modifications decreased with increasing number of modified units—to a larger extent for the L- (GPNA1 to GPNA3) as compared to the D- (GPNA4 to GPNA6) series. The extent of destabilization for the D-series, however, could be recovered by spacing the arginine-modified at every other position with the unmodified backbone units (compare the T_m of GPNA6 to that of GPNA7). The two oligomers (GPNA6 and GPNA7) are identical to each other in every respect with the exception of backbone spacing. The T_m of the alternate GPNA (GPNA7) is 4 °C higher than that of the unmodified PNA (PNA1), as compared to the consecutive GPNA (GPNA6) which is 5 °C lower than the unmodified PNA. The T_m of the hybrid duplex further increased with increasing number of D-arginine-derived backbone units (see the T_m of GPNA8)—consistent with the observations made with RNA.[16] The detrimental effects observed with the consecutive backbone modifications could be attributed to intrastrand steric and/or electrostatic repulsion, which could be minimized by placing the modified backbone units further apart as demonstrated with GPNA7 and GPNA8. This result is in agreement with the previous findings with lysine-derived chiral PNAs.[21] The enhancement in the T_m of GPNA8 could be due to increased electrostatic interaction between the guanidinium groups of GPNA and the phosphate groups of DNA.

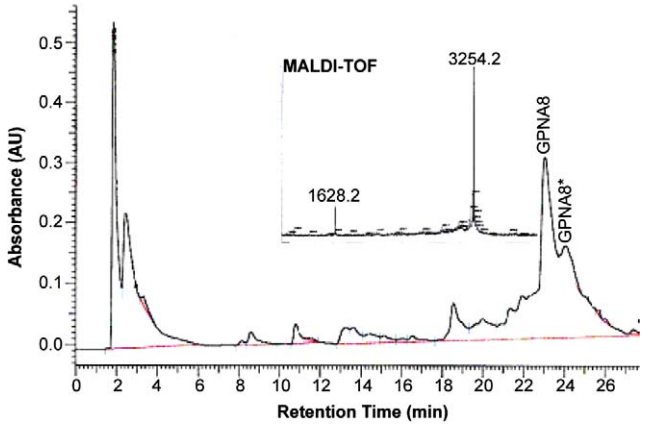


Figure 1. An HPLC trace of the crude GPNA8 oligomer. Inset: MALDI-TOF spectrum of the purified GPNA8. GPNA8* indicates the existence of a second conformer based on the following observations: (i) GPNA8* has the same *m*/*z* as GPNA8, (ii) the ratio of GPNA8:GPNA8* varied with concentration (GPNA8* only showed up at relatively high concentration), (iii) GPNA8 and GPNA8* were present at relatively high concentration regardless of whether the samples were collected from the first or second peak.

Table 1. The sequence of PNA and GPNA oligomers, along with the corresponding T_ms upon hybridizing to perfectly matched, antiparallel complementary DNA strand

Oligomers	Sequence	T_m (°C)	ΔT_m (°C)
PNA1	H-GCATGTTTG A-^{L}Lys-NH_2	43	
GPNA1	H-GCATG^{L}**T**TTG A-NH_2	39	−4
GPNA2	H-GCATG^{L}**T**^{L}**T**TG A-NH_2	37	−6
GPNA3	H-GCATG^{L}**T**^{L}**T**^{L}**T**G A-NH_2	35	−8
GPNA4	H-GCATG^{D}**T**TTG A-NH_2	41	−2
GPNA5	H-GCATG^{D}**T**^{D}**T**TG A-NH_2	39	−4
GPNA6	H-GCATG^{D}**T**^{D}**T**^{D}**T**G A-NH_2	38	−5
GPNA7	H-GCA^{D}**T**G^{D}**T**T^{D}**T**G A-NH_2	47	+4
GPNA8	H-G^{D}**C**A^{D}**T**G^{D}**T**T^{D}**T**G^{D}**A**-NH_2	50	+7

Experimental conditions: PNA–DNA and GPNA–DNA duplexes were hybridized by mixing stoichiometric amount of each oligomer strand (2 μM) in sodium phosphate buffer (10 mM NaPi, 100 mM NaCl, and 0.1 mM EDTA, pH 7) and incubated at 90 °C for 5 min, followed by a gradual cooling to room temperature. UV-absorption was monitored at 260 nm as the function of temperatures from 20 to 90 °C at the rate of 1 °C/min. Both the heating and cooling runs were recorded in order to assess the reversibility of the hybridization process. Bold letters indicate modified backbone units and superscripts (L or D) indicate the configuration of the amino acids from which the backbone units were derived.

Next we determined the sequence-specificity of GPNA8 by comparing the T_ms of the hybrid duplex containing perfectly matched with those containing single-base mismatched sequences. Figure 2 shows the melting profiles of GPNA8–DNA containing fully match (GPNA8–DNA) and single-base T ↔ C (GPNA8–DNA-C) and T ↔ G (GPNA8-DNA-G) mismatches. The melting profile of T ↔ T (data not shown) did not show noticeable transition, and therefore was omitted from Figure 1. On average, a single-base mismatch lowered the T_m of the hybrid duplex by ~17 °C, similar to the observation made with PNA.[22] This result is important because it shows that GPNA maintains the sequence specificity of PNA, despite the positively charged side chains.

In addition to the natural nucleobases, we have also examined the effects of arginine side chains on the hybridization properties of the modified nucleobases ^{s}U and D. These modified nucleobases have been exploited in the development of pseudocomplementary PNAs (psPNAs) for targeting double helical DNA, in which both strands of the double helical DNA are targeted simultaneously.[23,24] Nielsen and co-workers[25]

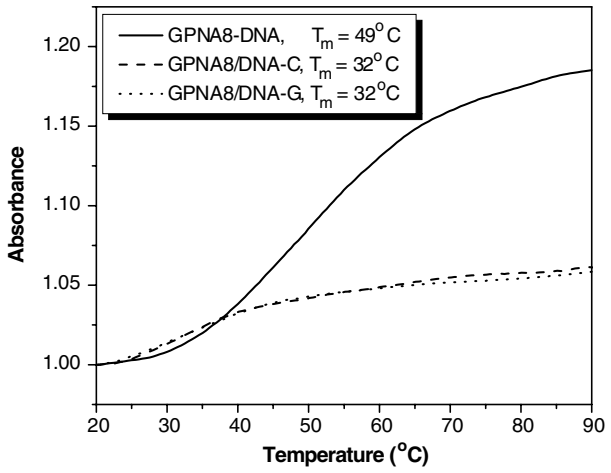


Figure 2. UV-melting curves of GPNA8–DNA duplex containing perfectly matched and single-base mismatched sequences. GPNA8/DNA-C and GPNA8/DNA-G contained T ↔ C and T ↔ G single-base mismatch, respectively.

have shown that the A → D substitution enhanced the stability of the PNA–DNA hybrid duplex by 2–6 °C, while the T → ^{s}U substitution destabilized the duplex by 2–3 °C. Table 2 shows the T_ms of the GPNA oligomers with different combinations of backbone and nucleobase modifications. The effect of the A → D substitution is enhancing (compare the T_ms of PNA3, GPNA9, and GPNA10 to that of PNA2, Table 2). On the other hand, a slight destabilization, but still within the expected range, was observed for the T → ^{s}U substitution (compare the T_ms of PNA5, GPNA11, and GPNA12 to that of PNA4, Table 2). The cause of this destabilization is not clear, but could be due to steric repulsion between the relatively large sulfur atom of the thiouracil and the arginine side chains of GPNA. This finding is significant because it suggests the possibility of using arginine side chains as a means to further minimize PNA self-quenching in double-duplex invasion, while at the same time improve cellular uptake—which are important for in vivo applications.

To further gain insight into the uptake properties of this particular class of oligomeric molecules, we examined the intracellular distribution of an 18-mer, alternating GPNA oligomer (labeled with FITC) designed to bind to the RNA transcript of the human survivin gene. In this case, HeLa cells were seeded on the 8-well glass-bottomed culture slides at a density of ~50,000 cells per well. The next day, cells were washed with PBS, supplemented with complete media, and then incubated with 1 μM of GPNA13 for 16 h. Cells were then thoroughly washed with PBS, fixed with 4% paraformaldehyde, and then mounted with Vectashield containing DAPI. Images were taken using Zeiss AX200 fluorescence microscope with 60× objective. Interestingly, most of the cells (>80%) incubated with GPNA13 showed specific FITC-staining in the ER compartment, while the negative control (FITC-labeled PNA) did not show any uptake. An illustration of GPNA13 localization is shown in Figure 3—an image of a tetranucleated cell specifically stained with GPNA13 (green) at the nuclear junctions, a region where the ER is known to be distributed in a tetranucleated cell. This type of staining was not observed with shorter oligomers.[16] A similar pattern was also observed with live cells (data not shown), indicating that the observed intracellular localization is not due to artifacts from cell-fixation. This result is significant from the standpoint of antisense gene regulation, since it is the ER where the ribosomes are localized and where the proteins are made. Because of its specific localization, GPNA could potentially be used to modulate gene expression in mammalian cells.

In summary, we have shown that GPNAs containing alternate, arginine-derived D-backbone configuration

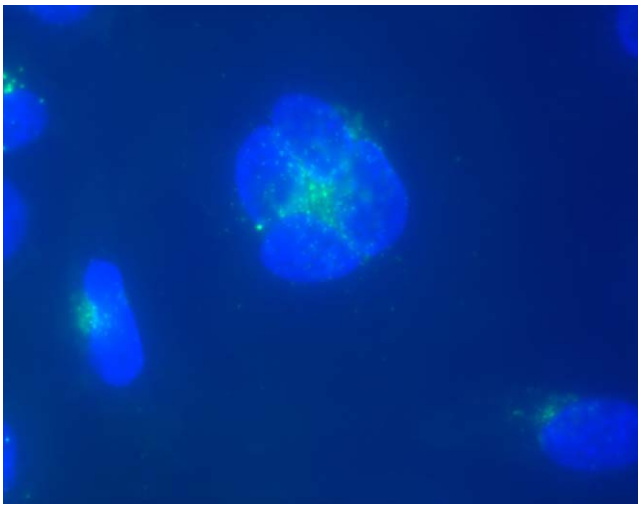


Figure 3. An image of HeLa cells incubated with 1 μM of GPNA13 for 16 h, followed by paraformaldehyde-fixing and brief staining with DAPI. Nuclei (blue), GPNA13 (green). GPNA13: FITC-^{D}**C**C-^{D}**A**C^{D}**C**T^{D}**C**T^{D}**G**C^{D}**C**A^{D}**A**C^{D}**G**G^{D}**G**T-NH_2.

Table 2. The effects of backbone and nucleobase modifications on the stability of PNA–DNA hybrid duplex

PNA	PNA sequence	T_m (°C)	ΔT_m (°C)
PNA2	H-^{L}Lys-GTAGATCACT-^{L}Lys-NH_2	50	
PNA3	H-^{L}Lys-GTDGDTCDCT-^{L}Lys-NH_2	65	+15
GPNA9	H-^{D}Lys-GT^{D}**A**G^{D}**A**TC^{D}**A**CT-^{D}Lys-NH_2	53	+3
GPNA10	H-^{D}Lys-GT^{D}**D**G^{D}**D**TC^{D}**D**CT-^{D}Lys-NH_2	68	+18
PNA4	H-^{L}Lys-AGTGATCTAC-^{L}Lys-NH_2	49	
PNA5	H-^{L}Lys-AGUGAUCUAC-^{L}Lys-NH_2	48	−2
GPNA11	H-^{D}Lys-AG^{D}**T**GA^{D}**T**C^{D}**T**AC-^{D}Lys-NH_2	51	+2
GPNA12	H-^{D}Lys-AG$^{D\text{-}S}$**U**GA$^{D\text{-}S}$**U**C$^{D\text{-}S}$**U**AC-^{D}Lys-NH_2	43	−6

bind sequence-specifically to DNA, and are taken-up by mammalian cells and localized specifically in the ER compartment. The results presented here have important implication on the future design of PNA and other related nucleic acid mimics for in vivo applications. These results are relevant to a recent finding by Janowski and co-workers,[26] which showed that PNAs could be effectively used to target chromosomal DNA if they could be delivered into cells.

Acknowledgments

Financial support for this work was provided by Applied Biosystems (APPLERA-LY) and National Institutes of Health (GM077261-01).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2006.06.052.

References and notes

1. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
2. Nielsen, P. E. *Acc. Chem. Res.* **1999**, *32*, 624.
3. Gambacorti-Passerini, C.; Mologni, L.; Bertazzoli, C.; LeCoutre, P.; Marchesi, E.; Grignani, F.; Nielsen, P. E. *Blood* **1996**, *88*, 1411.
4. Buchardt, O.; Egholm, M.; Berg, R. H.; Nielsen, P. E. *Trends Biotechnol.* **1993**, *11*, 384.
5. Nielsen, P. E. *Curr. Opin. Biotechnol.* **1999**, *10*, 71.
6. Nielsen, P. E. *Curr. Opin. Biotechnol.* **2001**, *12*, 16.
7. Koppelhus, U.; Nielsen, P. E. *Adv. Drug Deliv. Rev.* **2003**, *55*, 267.
8. Marin, V. L.; Roy, S.; Armitage, B. *Exp. Opin. Biol. Ther.* **2004**, *4*, 337.
9. Hanvey, J. C.; Peffer, N. J.; Bisi, J. E.; Thomson, S. A.; Cadilla, R.; Josey, J. A.; Ricca, D. J.; Hassman, C. F.; Bonham, M. A.; Au, K. G.; Carter, S. G.; Ruckenstein, D. A.; Boyd, A. L.; Noble, S. A.; Babiss, L. E. *Science* **1992**, *258*, 1481.
10. Shammas, M. A.; Simmons, C. G.; Corey, D. R.; Reis, R. J. S. *Oncogene* **1999**, *18*, 6191.
11. Hamilton, S. E.; Simmons, C. G.; Kathiriya, I. S.; Corey, D. R. *Chem. Biol.* **1999**, *6*, 343.
12. Cutrona, G.; Carpaneto, E. M.; Ulivi, M.; Roncella, S.; Landt, O.; Ferrarini, M.; Boffa, L. C. *Nat. Biotech.* **2000**, *18*, 300.
13. Koppelhus, U.; Awasthi, S. K.; Zachar, V.; Holst, H.; Ebbesen, P.; Nielsen, P. E. *Antisense Nucleic Acids Drug Dev.* **2002**, *12*, 51.
14. Zasloff, M. *Nature* **2002**, *415*, 389.
15. Zhou, P.; Wang, M.; Du, L.; Fisher, G. W.; Waggoner, A.; Ly, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 6878.
16. Dragulescu-Andrasi, A.; Zhou, P.; He, G.; Ly, D. H. *Chem. Commun.* **2005**, *2*, 244.
17. Puschl, A.; Sforza, S.; Haaima, G.; Dahl, O.; Nielsen, P. E. *Tetrahedron Lett.* **1998**, *39*, 4707.
18. Haaima, G.; Lohse, A.; Buchardt, O.; Nielsen, P. E. *Angew. Chem., Int. Ed.* **1996**, *35*, 1939.
19. 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.
20. Christensen, L.; Fitzpatrick, R.; Gildea, B.; Petersen, K. H.; Hansen, H. F.; Koch, T.; Egholm, M.; Buchardt, O.; Nielsen, P. E.; Coull, J. *J. Pept. Sci.* **1995**, *1*, 175.
21. Sforza, S.; Corradini, R.; Ghirardi, S.; Dossena, A.; Marchelli, R. *Eur. J. Org. Chem.* **2000**, 2905.
22. Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, P. E. *Nature* **1993**, *365*, 566.
23. Lohse, J.; Dahl, O.; Nielsen, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11804.
24. Demidov, V. V.; Protozanova, E.; Izvolsky, K. I.; Price, C.; Nielsen, P. E.; Frank-Kamenetskii, M. D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5953.
25. Haaima, G.; Hansen, H. F.; Christensen, L.; Dahl, O.; Nielsen, P. E. *Nucleic Acids Res.* **1997**, *25*, 4639.
26. Janowski, B. A.; Kaihatsu, K.; Huffman, K. E.; Schwartz, J. C.; Ram, R.; Hardy, D.; Mendelson, C. R.; Corey, D. R. *Nat. Chem. Biol.* **2005**, *1*, 210.