Quadruplex Sensors

DOI: 10.1002/anie.201000291

DNA-Templated Synthesis of Trimethine Cyanine Dyes: A Versatile Fluorogenic Reaction for Sensing G-Quadruplex Formation**

*Kamel Meguellati, Girish Koripelly, and Sylvain Ladame**

It has been known for several decades that guanine-rich nucleic acid sequences have a propensity to fold into highly stable four-stranded structures in vitro in the presence of physiological cations, notably potassium and sodium.[1] Such structures, termed quadruplexes, have had their biological significance demonstrated for a number of processes. For example, the single-stranded 3′-end of telomeric DNA could adopt a quadruplex conformation under near physiological conditions, which has implications on telomere maintenance mechanisms.[2] More recently, a number of DNA G-quadruplex sequences have been identified in the promoter region of genes that have been proposed to act as regulatory elements for gene expression at the transcriptional level.[3] Among the 43% of human genes that contain a putative quadruplex-forming sequence in their promoter, specific oncogenes have received particular attention. These include the *c-myc*,[4] *bcl-2*,[5] *K-ras*,[6] and *c-kit*[7] genes. Although there is an increasing amount of evidence for the formation of G-quadruplexes at telomere ends in vivo,[8] the possible existence of promoter quadruplexes in vivo is still subject to debate. Recent studies using small-molecule approaches have demonstrated that quadruplex formation within the nuclease hypersensitive element of the *c-myc* gene or within the promoter of the *c-kit* gene were coupled to a significant inhibition of *c-myc*[9] and *c-kit*[10] expression at the transcriptional level in various cell lines. However, whilst the 3′-overhang of telomeric DNA is single-stranded, and therefore is free to adopt any stable secondary structure, quadruplex formation within a promoter would require at least a local and temporary opening of the DNA double helix, despite the high stability of Watson–Crick G-C base pairs. Recent studies using fluorescence resonance energy transfer (FRET)[11] or fluorescent probes[12] have demonstrated that quadruplexes could potentially form, even when in competition with a thermodynamically more stable duplex form. Moreover, it is well established that double-stranded DNA transiently becomes single-stranded during key biological processes, such as DNA replication, transcription or even recombination, thus allowing the folding of each DNA strand into alternative (that is, non-B-DNA) structures.[13]

We are interested in designing sensitive fluorescent biosensors that would be highly specific for unique G-quadruplexes in the genome. The general strategy consists in simultaneously targeting the quadruplex structure itself and also its two flanking regions in a sequence specific manner. Briefly, two short peptide nucleic acids (PNAs)[14] complementary to both quadruplex flanking regions are functionalized with two nonfluorescent components **A** and **B** of a fluorogenic reaction (that is, the reaction between nonfluorescent derivatives **A** with **B** leads to the formation of fluorescent entity **C**). The system can be designed in such a way that, upon hybridization of the PNA probes to their complementary DNA sequences by Watson–Crick base pairing, **A** and **B** will be in close enough proximity to react with each other when the DNA sequence between both PNAs is folded into a quadruplex structure only, whereas they will be kept separated if the DNA remains single-stranded (Figure 1).

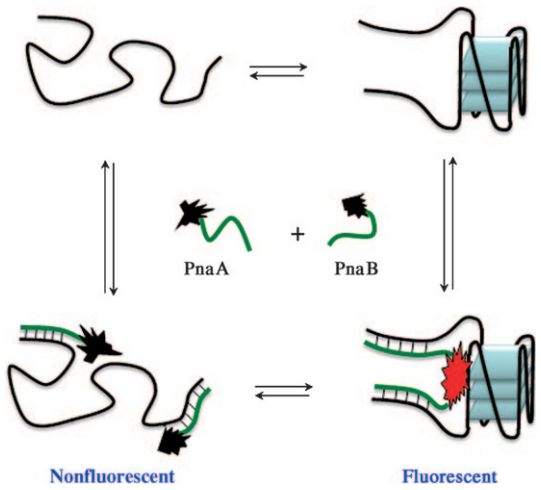


Figure 1. Quadruplex-templated fluorogenic reaction by hybridization of two labeled and nonfluorescent peptide nucleic acids PnaA and PnaB with the single-stranded flanking arms of a G-quadruplex.

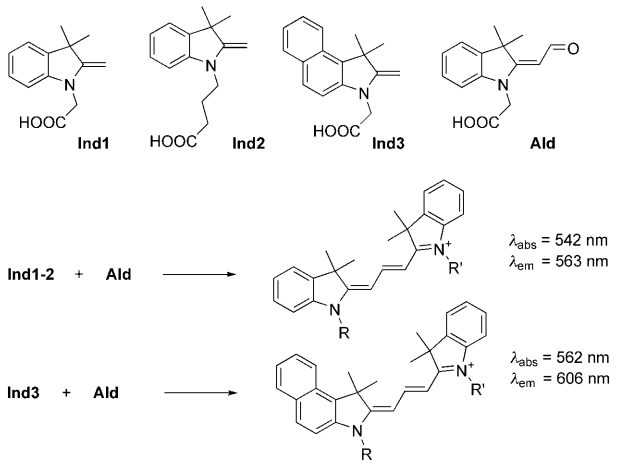
Oligonucleotide-templated reactions that can be monitored with high sensitivity by the appearance or disappearance of a fluorescent signal upon binding to the oligonucleotide target have recently received particular attention. Representative examples of such technologies include the use of fluorogenic probes (e.g. molecular beacons), or rely on fluorogenic reactions of chemical ligation or primer exten-

[*] K. Meguellati,[+] Dr. G. Koripelly,[+] Dr. S. Ladame
ISIS Université de Strasbourg
8, Allée Gaspard Monge, BP 70028, 67083 Strasbourg (France)
Fax: (+33) 3-6885-5115
E-mail: s.ladame@isis.u-strasbg.fr
[+] These authors contributed equally to this paper.

[**] S.L. thanks the International Centre for Frontier Research in Chemistry (icFRC) for financial support.
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201000291.

sion.[15] To date, although such systems offer the advantage of a very high signal-to-noise (S/N) ratio, there has been only few reports of DNA-templated fluorogenic reactions applied to oligonucleotide sensing. Most recent reports are based on the Staudinger reaction,[16] aldol-type,[17] organomercury-activated,[18] or S_NAr[19] reactions. They were all developed for detecting oligonucleotide sequences with potential applications as single-nucleotide polymorphism (SNP) probes or RNA sensors in cells. Two modified oligonucleotides (or oligonucleotide analogues) are designed so that 1) they can hybridize specifically to a unique nucleic acid template through Watson–Crick base-pairing and 2) their hybridization to the complementary template only brings both reactive groups in close enough proximity to react with each other.

Herein, the fluorogenic synthesis of a symmetrical or unsymmetrical trimethine cyanine dye by an aldolization-elimination reaction between two nonfluorescent precursors was applied for sensing G-quadruplex formation in vitro. Two PNAs were designed that can each hybridize in a sequence-specific manner with five nucleobases upstream and five nucleobases downstream of the parallel-stranded ckit21T quadruplex[7b,c] chosen as a model system. They were functionalized at their C-terminal or N-terminal end with either an *N*-alkyl-2-methyleneindoline (Ind1–3) or an *N*-alkyl-2-(3,3-dimethylindolin-2-ylidene)acetaldehyde (Ald; Scheme 1). Two ε-*N*,*N*-dimethyl lysine residues per PNA strand were also added to ensure solubility of both PNAs in water at near-physiological pH.[20]



Scheme 1. Structures of nonfluorescent 2-methyleneindolines (Ind1, Ind2, and Ind3) and of an aldehyde derivative (Ald). Absorption and emission wavelengths of the cyanine dyes formed upon reaction of Ind1–3 with Ald are also given.

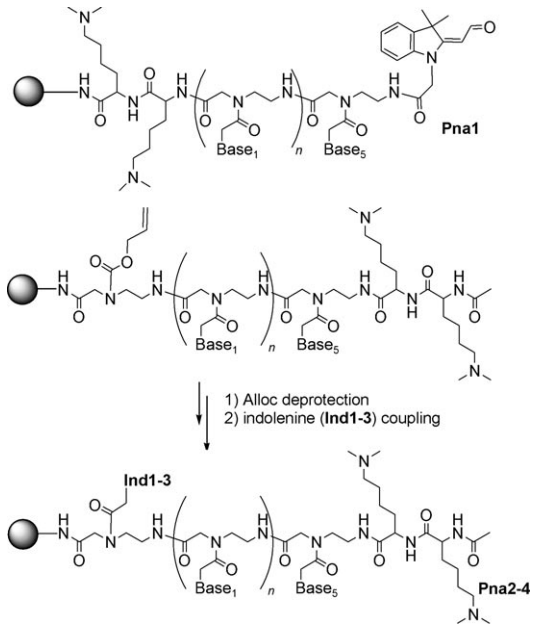
PNAs were synthesized on a rink amide resin (Merck Biosciences, loading 0.67 mmol g^{-1}) using standard solid-phase Fmoc chemistry (Fmoc = 9-fluoromethoxycarbonyl). Ald was introduced at the N-terminus of the PNA by amide coupling on solid support (Pna1; Table 1). For the introduction of an indoline moiety (Ind1, Ind2, or Ind3) at the C-terminus of the PNA, a versatile synthetic strategy was chosen that involves the use of an *N*-Fmoc-[2-(*N*-Alloc)aminoethyl]-glycine PNA monomer (Alloc = allyloxycarbonyl; Scheme 2).

Table 1: Oligonucleotide and PNA sequences.

	DNA[a] or PNA[b] sequences
Quad1	GCATCC**GGG**C**GGG**CGCGA**GGG**A**GGG**TTCGGC[a]
Quad2	GCATCC**GAG**C**GAG**CGCGA**GAG**A**GAG**TTCGGC[a]
Quad3	TTCGTC**GGG**C**GGG**CGCGA**GGG**A**GGG**TTAAGT[a]
Pna1	Lys(NMe_2)-Lys(NMe_2)-CGTAG-Ald[b]
Pna2	Ind1-AGCCG-Lys(NMe_2)-Lys(NMe_2)[b]
Pna3	Ind2-AGCCG-Lys(NMe_2)-Lys(NMe_2)[b]
Pna4	Ind3-AGCCG-Lys(NMe_2)-Lys(NMe_2)[b]

[a] DNA sequences are given from the 5′ to 3′ end. [b] PNA sequences are given from the C-terminal to N-terminal end.

A PNA strand was synthesized on a solid support starting with the *N*-Fmoc-[2-(*N*-Alloc)aminoethyl]glycine monomer. After incorporation of the final residue, the resin was treated with



Scheme 2. Solid-phase synthesis of fluorogenic PNAs Pna1–4 on a rink amide resin (gray sphere). $Base_1$, $base_5$ = A, T, C, or G.

$[Pd(PPh_3)_4]$ in the presence of dimethylamine–borane under strictly anhydrous and anaerobic conditions to remove selectively the Alloc protecting group.[21] The indoline moiety (Ind1, Ind2, or Ind3) was finally coupled by amide bond formation using 2-(7-aza-1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) as coupling agent (Scheme 2). This original approach allows convenient functionalization of any immobilized PNA sequence at their C-terminal end just prior to cleavage from the solid support. Herein, this approach was used to introduce various indoline or benz[e]indoline moieties at the C-terminal end of a unique PNA sequence (Pna2–4, Table 1).

Functionalized PNAs were finally cleaved from their solid support by treatment with a solution of trifluoroacetic acid/triisopropylsilane/H_2O (95:2.5:2.5) and the desired PNA was isolated by HPLC and characterized by MALDI.

The interaction between Pna1 and Pna2–4 was next investigated by fluorescence spectroscopy in the presence and in the absence of various types of DNA. Three different DNA sequences (Quad1–3; Table 1) were tested for their capacity to template the fluorogenic reaction of cyanine dye formation. Quad1 corresponds to the previously reported ckit21T[7b–c] quadruplex-forming sequence with two additional single-stranded flanking arms, located downstream and upstream of the quadruplex, and complementary to Pna1 and Pna2–4, respectively. Quad2 differs from Quad1 by only four G→A mutations to prevent G-quadruplex formation. Quad3 contains the same ckit21T sequence as Quad1 but the quadruplex-forming motif is now flanked with randomized single-stranded arms that are not complementary to the fluorogenic PNAs Pna1–4. Each DNA (200 μM) was folded in a potassium phosphate buffer (10 mM, pH 7.4) that also contained 100 mM KCl.[22] Under such conditions, Quad1 and Quad3 formed a parallel-stranded quadruplex (see the Supporting Information, Figure S1) whereas Quad2 remained single-stranded. Briefly, a stoichiometric mixture of aldehyde (Pna1) and indoline (Pna2, Pna3, or Pna4) in potassium-containing buffer was incubated at room temperature in the presence or in the absence of an equimolar amount of folded DNA. The reaction of cyanine dye formation was then monitored by fluorescence spectroscopy at different time points. First, the reaction between the PNA aldehyde (Pna1) and the PNA indoline (Pna2) was investigated. Interestingly, only very moderate fluorescence was detectable when working at a PNA strand concentration up to 500 nM. However, when adding a stoichiometric amount (500 nM) of folded quadruplex Quad1 to the previous mixture, a strong fluorescence signal instantaneously appeared which increased up to after 2 h (Figure 2), at which time equilibrium was finally reached (Supporting Information, Figure S2). At equilibrium, a 45-fold increase in fluorescence intensity was observed compared to the quadruplex-free experiment.

A similar trend, although of weaker intensity, was also observed when decreasing the PNA and DNA concentrations down to 200 nM (Supporting Information, Figure S3). To demonstrate that the efficiency of the fluorogenic reaction was indeed linked to quadruplex formation, the same stoichiometric mixture of Pna1 and Pna2 (500 nM each) was reacted in potassium phosphate buffer and in the presence of either Quad2 or Quad3. Key mutations of the ckit21T sequence to prevent quadruplex formation resulted in a complete inhibition of the fluorogenic reaction. Randomization of the quadruplex flanking sequences to prevent PNA:DNA hybridization also led to a significant inhibition compared to the reaction templated by Quad1. These results are consistent with the proposed model suggesting that hybridization of the both aldehyde and indole PNAs to the quadruplex flanking regions associated with folding of the central DNA sequence into a quadruplex conformation are the only conditions that bring both reactive groups in close enough proximity to form the fluorescent cyanine dye. If only one of those requirements is satisfied, no or little reaction will take place.

The influence of the linker between the PNA and the indoline on the efficiency of the fluorogenic reaction was then investigated. Pna3 differs from Pna2 by two extra methylene groups between the heterocycle and the PNA scaffold. Although a specific quadruplex-templating effect was observed when mixing Pna1 and Pna3 in the presence of Quad1 which was similar to that obtained with Pna2, it was significantly weaker, thus suggesting the influence of the linker (for example flexibility) on the reaction efficiency (Supporting Information, Figure S4).

An interesting intrinsic property of cyanine dyes is the possibility to tune their spectroscopic properties by varying either the nature of the nitrogen-containing heterocycles or the length of the polymethine chain between them. To shift our quadruplex-specific fluorescent biosensor toward longer wavelengths, Pna4 was synthesized, which differs from Pna2 by the substitution of the indoline moiety by a benz[e]indoline (Table 1). Reaction of Pna4 with Pna1 was then expected to generate an unsymmetrical cyanine dye absorbing and emitting at significantly longer wavelengths than the symmetrical dye formed upon reaction between Pna1 and Pna2 (Scheme 1).[23] Although no reaction was observed when

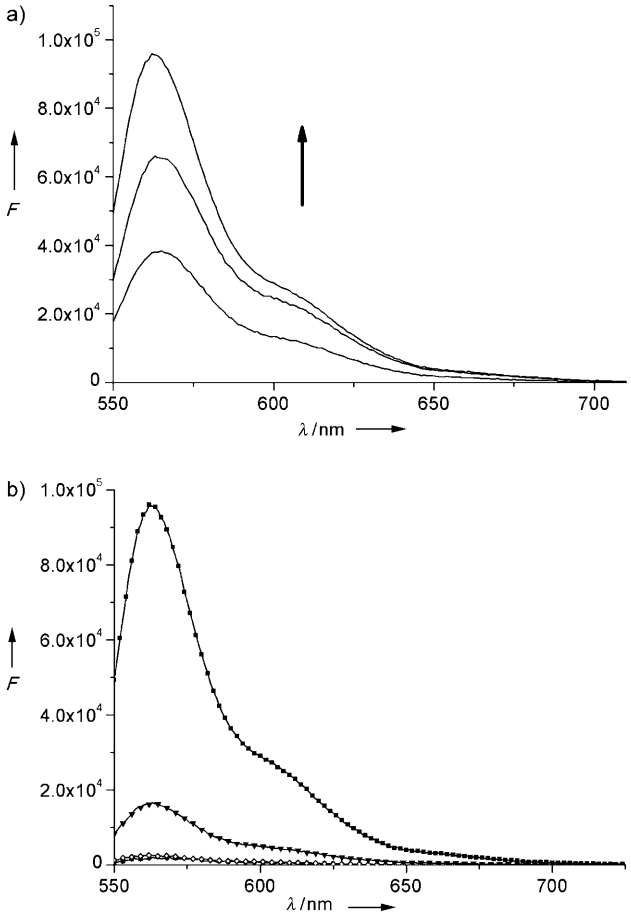


Figure 2. a) Fluorescence emission spectra (λ_{exc} = 540 nm) of a mixture of Pna1, Pna2, and Quad1 (500 nM each) in potassium phosphate buffer (10 mM, pH 7.4) and 100 mM KCl after 10 min, 1 h, and 2 h (bottom to top) at RT. b) Fluorescence emission spectra (λ_{exc} = 540 nm) of a mixture of Pna1 and Pna2 (500 nM each) in potassium phosphate buffer and in the absence (●) or in the presence of 500 nM of Quad1 (■), Quad2 (◇), or Quad3 (▼). Fluorescence spectra were recorded after 2 h.

reacting Pna1 and Pna4 at 20 μM, a strong quadruplex-templating effect (circa 40-fold increase in fluorescence) was observed at this concentration, resulting in the time-dependent appearance of a characteristic fluorescence signal (λ_{em} = 606 nm; see Figure 3). This study demonstrates the possibility

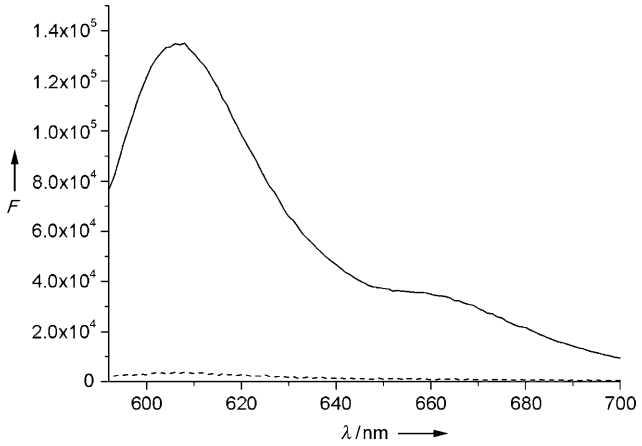


Figure 3. Fluorescence emission spectra (λ_{exc} = 562 nm) of a mixture of Pna1 and Pna4 (20 μM each) in the absence (-----) or presence (——) of Quad1 (20 μM) in potassium phosphate buffer (10 mM, pH 7.4) and 100 mM KCl.

of individual G-quadruplexes to template the formation of various trimethine cyanine dyes that absorb and emit at different wavelengths. However, it is noteworthy that structural modifications of the fluorogenic probes are also accompanied with changes in sensitivity, as a 15-fold loss of sensitivity was observed upon replacing Ind1 by Ind3. This is most likely due to a significantly lower reactivity of Ind3 when compared with Ind1. The specificity of this quadruplex fluorescent biosensor was finally assessed by reacting Pna1 with Pna2 in the presence of various amounts of double-stranded Calf Thymus DNA. Interestingly, no fluorescence was observed when working at low PNA concentration (500 nM each) and high CT concentration (10 $\mu g\,mL^{-1}$).

In conclusion, we reported the first example of fluorogenic synthesis of a trimethine cyanine dye that can be templated by a parallel-stranded G-quadruplex DNA in a "sequence + structure"-specific manner. By attaching two nonfluorescent aldehyde and indoline building blocks at the end of two PNA strands complementary to both single-stranded flanking regions of a DNA quadruplex, the fluorogenic reaction occurs only when a quadruplex is formed. Although a DNA-programmed synthesis of hemicyanine dyes that proceeds by a similar aldol-type reaction had already been reported by Huang and Coull,[17] our system offers the advantage of being more biocompatible as it involves working at physiological pH and requiring no amine additive. This fluorescent biosensor enables the specific detection of a unique quadruplex in vitro that is located between both PNAs complementary sequences. Considering the versatility of the PNA functionalization and the broad spectral range covered by cyanine dyes, tunable quadruplex fluorosensors based on this principle can potentially be designed that emit at different but specific wavelengths. Attempts for simultaneously sensing various quadruplex sequences and/or folds with different colors are currently underway in our laboratory.

Received: January 18, 2010
Published online: March 12, 2010

Keywords: biosensors · cyanines · fluorescent probes · G-quadruplexes · peptide nucleic acids

[1] a) S. Neidle, S. Balasubramanian in *Quadruplex Nucleic Acids*, Royal Society of Chemistry, Cambridge, **2006**; b) S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402 – 5415.
[2] a) E. H. Blackburn, *Nature* **1991**, *350*, 569 – 573; b) J. F. Riou, L. Guittat, P. Mailliet, A. Laoui, E. Renou, O. Petitgenet, F. Mégnin-Chanet, C. Hélène, J. L. Mergny, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2672 – 2677; c) S. Neidle, G. N. Parkinson, *Nat. Rev. Drug Discovery* **2002**, *1*, 383 – 393.
[3] For recent reviews or examples, see: a) Y. Qin, L. H. Hurley, *Biochimie* **2008**, *90*, 1149 – 1171; b) J. L. Huppert, *Biochimie* **2008**, *90*, 1140 – 1148; c) J. Eddy, N. Maizels, *Nucleic Acids Res.* **2008**, *36*, 1321 – 1333.
[4] V. González, L. H. Hurley, *Annu. Rev. Pharmacol. Toxicol.* **2009**, *50*, 111 – 129.
[5] J. Dai, T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R. A. Jones, D. Yang, *J. Am. Chem. Soc.* **2006**, *128*, 1096 – 1098.
[6] S. Cogoi, L. E. Xodo, *Nucleic Acids Res.* **2006**, *34*, 2536 – 2549.
[7] a) S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2005**, *127*, 10584 – 10589; b) H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle, S. Balasubramanian, *Biochemistry* **2006**, *45*, 7854 – 7860; c) S. T. Hsu, P. Varnai, A. Bugaut, A. P. Reszka, S. Neidle, S. Balasubramanian, *J. Am. Chem. Soc.* **2009**, *131*, 13399 – 13409.
[8] a) K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, H. J. Lipps, *Nat. Struct. Mol. Biol.* **2005**, *12*, 847 – 854; b) J. Tang, Z. Y. Kan, Y. Yao, Q. Wang, Y. H. Hao, Z. Tan, *Nucleic Acids Res.* **2008**, *36*, 1200 – 1208.
[9] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11593 – 11598.
[10] M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid, S. Balasubramanian, *J. Am. Chem. Soc.* **2007**, *129*, 12926 – 12927.
[11] a) P. S. Shirude, B. Okumus, L. Ying, T. Ha, S. Balasubramanian, *J. Am. Chem. Soc.* **2007**, *129*, 7484 – 7485; b) N. Kumar, B. Sahoo, K. A. Varun, S. Maiti, S. Maiti, *Nucleic Acids Res.* **2008**, *36*, 4433 – 4442.
[12] J. Alzeer, B. R. Vummidi, P. J. Roth, N. W. Luedtke, *Angew. Chem.* **2009**, *121*, 9526 – 9529; *Angew. Chem. Int. Ed.* **2009**, *48*, 9362 – 9365.
[13] For a recent review, see: H. J. Lipps, D. Rhodes, *Trends Cell Biol.* **2009**, *19*, 414 – 422.
[14] a) P. E. Nielsen, M. Egholm, R. H. Berg, Buchardt, *Science* **1991**, *254*, 1497 – 1500; b) 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.
[15] For recent examples, see: a) K. Wang, Z. Tang, C. J. Yang, Y. Kim, X. Fang, W. Li, Y. Wu, C. D. Medley, Z. Cao, J. Li, P. Colon, H. Lin, W. Tan, *Angew. Chem.* **2009**, *121*, 870 – 885; *Angew. Chem. Int. Ed.* **2009**, *48*, 856 – 870; b) T. N. Grossmann, O. Seitz, *Chem. Eur. J.* **2009**, *15*, 6723 – 6730.
[16] a) J. Cai, X. Li, X. Yue, J. S. Taylor, *J. Am. Chem. Soc.* **2004**, *126*, 16324 – 16325; b) Z. L. Pianowski, N. Winssinger, *Chem. Commun.* **2007**, 3820 – 3822; c) R. M. Franzini, E. T. Kool,

ChemBioChem **2008**, *9*, 2981 – 2988; d) Z. Pianowski, K. Gorska, L. Oswald, C. A. Merten, N. Winssinger, *J. Am. Chem. Soc.* **2009**, *131*, 6492 – 6497; e) K. Furukawa, H. Abe, K. Hibino, Y. Sako, S. Tsuneda, Y. Ito, *Bioconjugate Chem.* **2009**, *20*, 1026 – 1036.

[17] Y. Huang, J. M. Coull, *J. Am. Chem. Soc.* **2008**, *130*, 3238 – 3239.

[18] R. M. Franzini, E. T. Kool, *Org. Lett.* **2008**, *10*, 2935 – 2938.

[19] A. Shibata, H. Abe, M. Ito, Y. Kondo, S. Shimizu, K. Aikawa, Y. Ito, *Chem. Commun.* **2009**, 6586 – 6588.

[20] ε-*N*,*N*-dimethyl lysines were preferred to lysines to avoid the presence of nucleophilic primary amine within the PNA strand that could potentially react with Ald.

[21] D. V. Jarikote, O. Köhler, E. Socher, O. Seitz, *Eur. J. Org. Chem.* **2005**, 3187 – 3195.

[22] A DNA solution containing either Quad1, 2, or 3 (200 μM) in potassium phosphate buffer (10 mM, pH 7.4) with KCl solution (100 mM) was heated at 95 °C for 5 minutes and slowly cooled to room temperature over 8 hours.

[23] S. J. Mason, J. L. Hake, J. Nairne, W. J. Cummins, S. Balasubramanian, *J. Org. Chem.* **2005**, *70*, 2939 – 2949.

