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 G-rich nucleic acid sequences have a propensity to fold into highly stable four-stranded structures in vitro in the presence of physiological cations, notably K^+ and Na^+.[1] Such structures, termed quadruplexes, have had their biological significance demonstrated for a number of processes. For example, it has been shown that 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] Amongst 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 exist increasing evidences 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 upstream 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 like DNA replication, transcription or even recombination, thus allowing the folding of each DNA strand into alternative (i.e. non B-DNA) structures.[13]

Herein, we were interested in designing sensitive fluorescent biosensors that would be highly specific for unique G-quadruplexes in the genome. The general strategy consists in targeting simultaneously the quadruplex structure itself but 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 functionalised with two non-fluorescent components **A** and **B** of a fluorogenic reaction (i.e. reaction between non-fluorescent 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 through 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, while they will be kept separated if the DNA remains single-stranded (Figure 1).

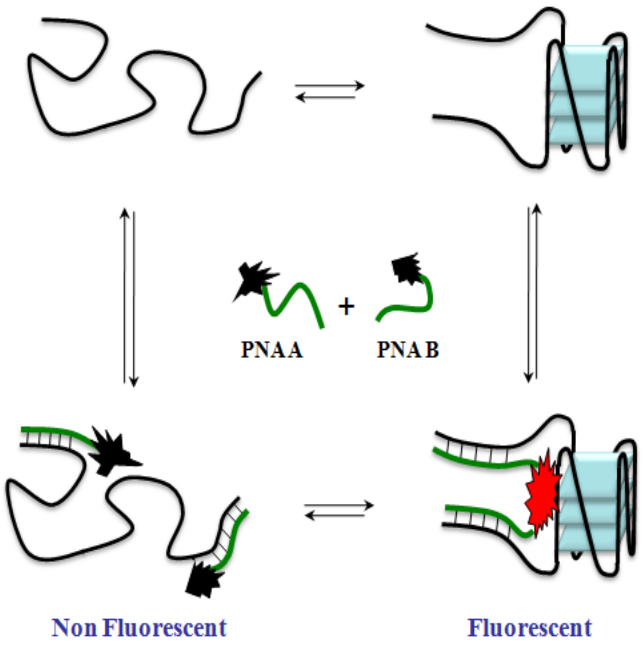


Figure 1. Schematic representation of a quadruplex-templated fluorogenic reaction via hybridisation of two labelled (and non-fluorescent) PNAs with the single-stranded flanking arms of a G-quadruplex.

Oligonucleotide-templated reactions that can be monitored with high sensitivity by the appearance/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 extension.[15] To date, and 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

[*] Kamel Meguellati,[+] Dr Girish Koripelly, [+] Dr Sylvain Ladame
ISIS Université de Strasbourg
8, Allée Gaspard Monge, BP 70028, 67083 Strasbourg, France.
Fax: (+33) 368855115
E-mail: s.ladame@isis.u-strasbg.fr
[+] These authors contributed equally to this paper

[**] SL 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://www.angewandte.org or from the author.

SNAr[19] reactions. They were all developed for detecting oligonucleotide sequences with potential applications as single nucleotide polymorphism (SNP) probes or RNA sensors in cells. Briefly, two modified oligonucleotides (or oligonucleotide analogues) are designed so that (i) they can hybridize specifically to a unique nucleic acid template through Watson-Crick base-pairing and (ii) 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 via an aldolisation-elimination reaction between two non-fluorescent precursors (i.e. a Fisher's base aldehyde and either an *N*-alkyl-2,3,3-trimethylindolenine or an *N*-alkyl-benz[e]trimethylindolenine) was applied for sensing G-quadruplex formation in vitro. Two PNAs were designed that can each hybridize in a sequence-specific manner with 5 nucleobases upstream and 5 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 a Fisher's base aldehyde or an *N*-alkyl-indolenine moiety (Figure 2). 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]

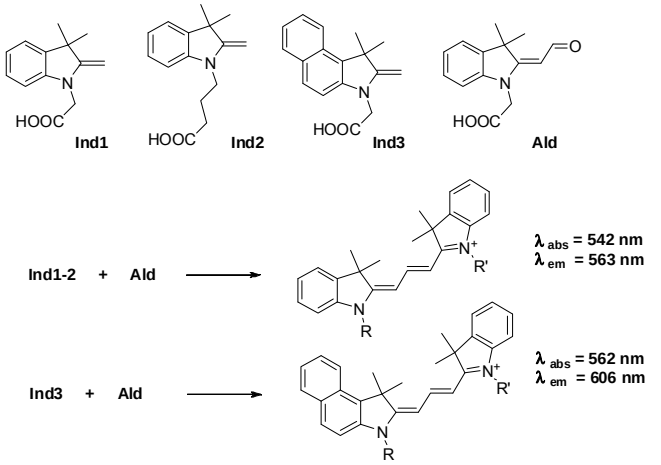


Figure 2. Structures of non-fluorescent 2-methylene-indolenines (**Ind1**, **Ind2** and **Ind3**) and of Fisher's base aldehyde (**Ald**). The structures and spectroscopic properties (taken in water) of the cyanine dyes formed upon reaction of **Ind1-3** with **Ald** are also given.

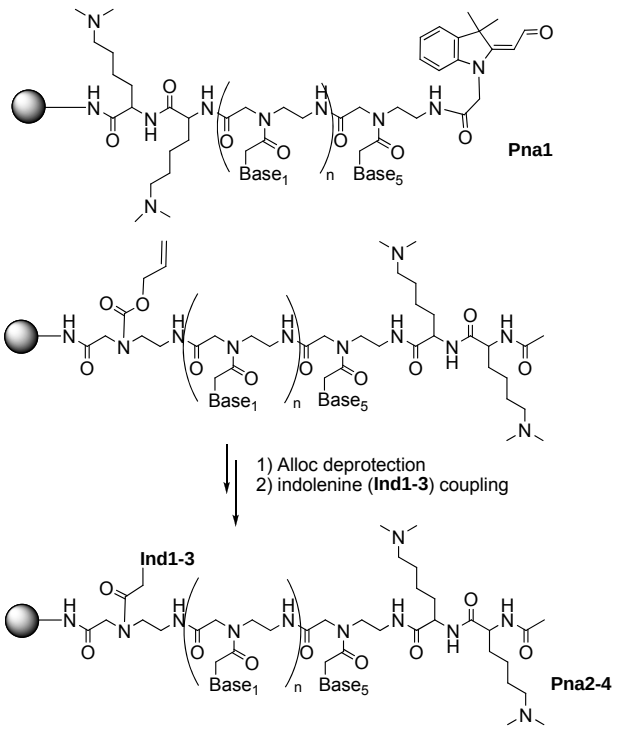
PNAs were synthesized on a rink amide resin (Merck Biosciences, loading 0.67 mmol/g) using standard solid-phase Fmoc chemistry (Fmoc = 9-fluoromethoxycarbonyl). Fisher's base aldehyde (**Ald**) was introduced at the N-terminus of the PNA via amide coupling on solid support (**Pna1**, Table 1). For the introduction of an indolenine moiety (**Ind1**, **Ind2** or **Ind3**) at the C-terminus of the PNA, a versatile synthetic strategy was chosen which involves the use of an *N*-Fmoc-[2-(*N*-Alloc)aminoethyl]glycine PNA monomer (Scheme 1). Briefly, a PNA strand was synthesized on solid support starting with the *N*-Fmoc-[2-(*N*-Alloc)aminoethyl]glycine monomer. After incorporation of the final residue, the resin was treated with tetrakis-(triphenylphosphine)palladium(0) in the presence of dimethylamine–borane complex under strictly anhydrous and anaerobic conditions in order to remove selectively the Alloc protecting group.[21] The indolenine moiety (**Ind1**, **Ind2** or **Ind3**) was finally coupled via amide bond formation using HATU as coupling agent (Scheme 1). This original approach allows convenient functionalization of any immobilised PNA sequence at their C-terminal end just prior cleavage from the solid support. Herein, this approach was used to introduce various indolenine or benz[e]indolenine moieties at the C-terminal end of a unique PNA sequence (**Pna2-4**, Table 1).

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**G*A*G**C**G*A*G**CGCGA**G*A*G**A**G*A*G**TTCGGC[a]
Quad3	*TTCGT*C**GGG**C**GGG**CGCGA**GGG**A**GGG**T*TAAGT*[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.

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



Scheme 1. Solid-phase synthesis of fluorogenic PNAs (**Pna1-4**) on a rink amide resin.

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 in order 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 which are not complementary to the fluorogenic PNAs **Pna1-4**. Each DNA (200 µM) was folded in potassium phosphate buffer (10 mM, pH 7.4) also containing 100 mM KCl.[22] Under such conditions, **Quad1** and **Quad3** formed a parallel-stranded quadruplex (see FigS1 supporting information) while **Quad2** remained single-stranded. Briefly, a stoichiometric mixture of aldehyde (**Pna1**) and indolenine (**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 indolenine (**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 hours (Figure 3), when equilibrium was finally reached (Figure S2 supporting information). At equilibrium, a 45-fold increase in fluorescence intensity was observed when compared to the quadruplex-free experiment.

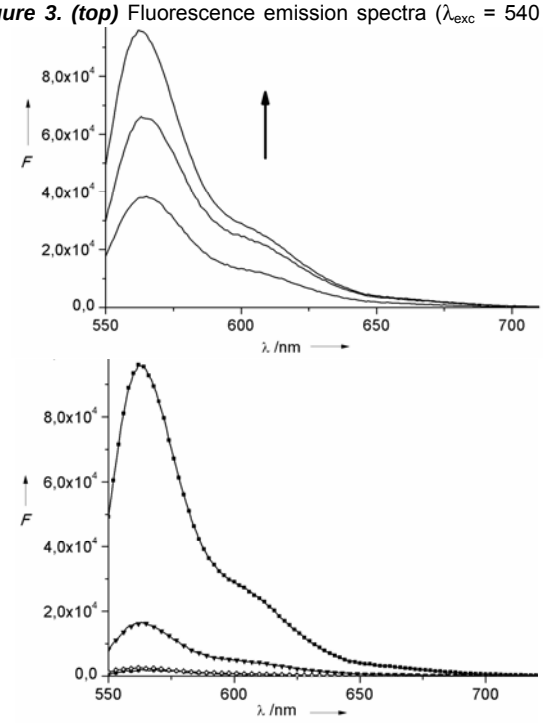


Figure 3. (top) Fluorescence emission spectra (λ_{exc} = 540 nm) of a mixture of **Pna1**, **Pna2** and **Quad1** (500 nM each) in potassium phosphate buffer (10 mM, pH 7.4) + 100 mM KCl after 10 min, 1 h and 2 h (bottom to top) at rt.; ***(bottom)*** Fluorescence emission spectra (λ_{exc} = 540 nm) 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 hours.

A similar trend, although of weaker intensity, was also observed when decreasing the PNA and DNA concentrations down to 200 nM (see Figure S3 supporting information). In order 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 when 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.

Next, the influence of the linker between the PNA and the indolenine on the efficiency of the fluorogenic reaction was 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 proved significantly weaker, thus suggesting the influence of the linker (e.g. flexibility) on the reaction efficiency (see Figure S4 supporting information).

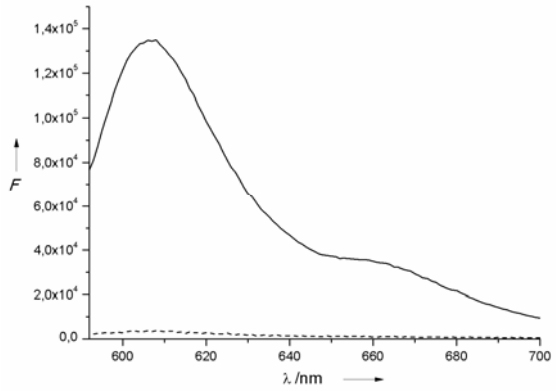


Figure 4. Fluorescence emission spectra (λ_{exc} = 562 nm) of a mixture of **Pna1** and **Pna 4** (20 µM each) in the absence (dotted line) or presence (plain line) of **Quad1** (20 µM) in potassium phosphate buffer (10 mM, pH 7.4) + 100 mM KCl.

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. In order to shift our quadruplex-specific fluorescent biosensor toward longer wavelengths **Pna4** was synthesized which differs from **Pna2** by the substitution of the indolenine moiety by a benz[e]indolenine (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** (Figure 2).[23] While no reaction was observed when reacting **Pna1** and **Pna4** at 20 µM, a strong quadruplex-templating effect (c.a. 40-fold increase in fluorescence) was observed at this concentration resulting in the time-dependant appearance of a characteristic fluorescence signal (λ_{em} = 606 nm, see Figure 4). This study demonstrates the possibility for 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

since a c.a.15-fold loss of sensitivity was observed when replacing **Ind1** by **Ind3**. This is 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 µg/mL).

We reported here 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 non-fluorescent aldehyde and indolenine 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 quadruplex is formed. Although a DNA-programmed synthesis of hemicyanine dyes that proceeds via a similar aldol-type reaction had already been reported by Huang and Coull,[17] our system offers the advantage of being more biocompatible since working at physiological pH and requiring no amine additive. This original fluorescent biosensor enables the specific detection of a unique quadruplex in vitro, i.e., that 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 colours are currently underway in our laboratory.

Received: ((will be filled in by the editorial staff))
Published online on ((will be filled in by the editorial staff))

[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 Discov.* **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. Gonzalez, 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*, 12929-12927.

[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. Int. Ed.* **2009**, *48*, 9362-9365.

[13] For a recent review, see H.J. Lipps, D. Rhodes, *Trends in Cell Biology* **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. 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 the Fischer's base aldehyde.

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

[22] **Quad1-3** were folded into a 10 mM potassium phosphate buffer + 100 mM KCl solution at pH 7.4 after heating the DNA solution (200 µM) at 95 °C for 5 min and letting it slow cool to rt over 8 hours.

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