

Dual sensing of hairpin and quadruplex DNA structures using multi-colored Peptide Nucleic Acid (PNA) fluorescent probes

Girish Koripelly,[†] Kamel Meguellati[†] and Sylvain Ladame^{*‡}

*To whom correspondence should be addressed :

Institut de Science et d'Ingénierie Supramoléculaires (ISIS),

Université de Strasbourg, CNRS UMR 7006

8, allée Gaspard Monge, 67083 Strasbourg Cédex, France

Fax: (+33) 368855115

E-mail: s.ladame@isis.u-strasbg.fr

[†] Both authors contributed equally to this work

[‡] Current address: Imperial College London, Department of Bioengineering, Royal School of Mines building, South Kensington Campus, London SW7 2AZ, U.K.

E-mail: sladame@imperial.ac.uk, Phone: (+44) 2075945308

Running Title: Enlightening DNA secondary structures

Abstract: Synthesis of water-soluble 5-mer Peptide Nucleic Acids (PNAs) functionalized at their 5' and 3'-ends with two original precursors of pentamethine cyanine dye synthesis is reported. The successful use of these PNA probes for sensing DNA hairpin structures in vitro was also demonstrated where specific hairpin formation was associated with the appearance of a characteristic fluorescence signal at 660 nm. A comparative study between three different strategies where PNAs were targeting either the stem or the loop of the hairpin was carried out. Best sensitivity was obtained using PNA sequences complementary to the loop sequence and directing both functional moieties toward the base of loop. Unprecedented proof-of-concept for the simultaneous sensing of hairpin and quadruplex DNAs with a non-overlapping two-color system (C3 and C5) is also demonstrated.

Introduction: Although the most abundant form of DNA in vivo is the double-stranded helical conformation reported for the first time in 1953 by Watson and Crick (1), it has been known for more than 50 years that nucleic acids can also adopt alternative (i.e. non B-DNA) secondary structures under physiological conditions. Such structures can be either single-stranded (e.g. hairpins, intramolecular G-quadruplexes, pseudoknots) or double-stranded (e.g. cruciforms, bimolecular G-quadruplexes) and have received increasing attention because of the role they were proposed to play in transcription regulatory mechanisms. (2) For instance, it has been known for several decades that DNA sequences containing a high density of clustered guanines were able to adopt four-stranded secondary structures named guanine (G)-quadruplexes or tetraplexes in the presence of physiological cations, notably K^+ and Na^+ (Figure 1). (3-7) More recently, in silico studies have revealed the high prevalence of such G-rich DNA sequences throughout the human genome, particularly in gene promoter regions. (8-10) Hairpin formation has also been demonstrated both in vitro and in vivo. (11) Stem-loop hairpin structures are obtained by folding of a partially self-complementary or palindromic DNA sequences (Figure 1). Their presence in gene promoters (12) or near gene replication origins (13) for instance strongly suggests that DNA hairpins could play a functional role in naturally occurring regulatory mechanisms of gene expression. For this reason, non-Watson-Crick DNA secondary structures represent attractive therapeutic targets to regulate the expression of specific genes via chemical or biochemical intervention. Successful examples of gene regulation with small molecules that bind to and stabilize

hairpin or quadruplex DNA (or RNA) have been recently reported in the literature. (14-18) Although these studies represent converging evidences that such structures do form in the context of the cell, there is still an urgent need for chemical probes capable of sensing the formation of nucleic acid structures both in vitro and in vivo.

There is currently a growing interest in the development of new fluorescent chemosensors based on supramolecular systems which have the ability to recognize and bind to a specific metabolite with high affinity. As a result of this interaction, the system re-equilibrates thus leading to a change of the physical properties (i.e. changes in fluorescence intensity and/or fluorescence emission wavelength) of the medium. Since the 1960s, engineered DNA molecules have been used extensively as highly specific fluorescent probes in order to explore various biological processes (e.g. DNA transcription, DNA recombination) and more recently for genotyping applications. Among them, molecular beacons (MBs) have received particular attention because of their broad applications in modern bioresearch. (19) Typically, MBs are DNA hairpin structures functionalized at their 5' and 3' end with a fluorescence-donor and a fluorescence-acceptor. Binding of the MB to a DNA or RNA target, through sequence-specific hybridization, results in the opening of the MB, thus separating the fluorescence donor from its neighboring fluorescence acceptor. Other examples include the use of modified (e.g. fluorescently labeled) oligonucleotides or oligonucleotide mimics in oligonucleotide-templated reactions. In such systems, a specific reaction of dye formation occurs only upon hybridization of the probes to a complementary nucleic acid target. A characteristic fluorescent signal will then be emitted which can be directly linked to the detection of the targeted sequence and/or structure. (20) Most recent reports of DNA-templated reactions applied to oligonucleotide sensing are based on the Staudinger reaction, (21-25) aldol-type, (26-27) organomercury-activated (28) or S_NAr (29) reactions.

Because peptide nucleic acids (30-31) offer the advantage of a higher stability, they have been recently used as an alternative to DNA for the design of highly specific nucleic acid probes. (32-36) We recently reported the use of fluorogenic PNA probes for sensing G-quadruplex formation. Upon binding of both probes to the single-stranded flanking arms of a quadruplex, formation of a trimethine cyanine dye from two non-fluorescent precursors was templated. (27) Herein we describe the synthesis of a novel fluorescent Fisher's base aldehyde probe-head, its incorporation in a PNA strand, and the applications of the resulting fluorescently labeled probe for sensing hairpin DNA secondary structures. Reaction between this fluorescent Fisher's base aldehyde and a non-fluorescent indolenine molecule leads to the

formation of a pentamethine cyanine dye (Cy5) absorbing and emitting at a significantly longer wavelength than its precursor. The system has been designed such that the Cy5 product forms solely upon simultaneous hybridization of both probes to either the loop or the stem of a unique hairpin DNA target. Therefore, hairpin formation will be directly linked to the appearance of a fluorescent signal at long wavelength ($\lambda_{em} > 650$ nm), well above the residual fluorescence of biological media, and above the fluorescence emission wavelength of most fluorescent products formed in oligonucleotide-templated reactions, (21-29) including that previously reported by us (27). Simultaneous sensing of various DNA structures via a two-color (Cy3-Cy5) system will also be presented.

Experimental Procedures:

Ethyl 2-(3,3-dimethyl-2-methyleneindolin-1-yl)-acetate (1). A mixture of 2,3,3-trimethyl-3H-indole (1.00 g, 6.28 mmol) and ethyl iodoacetate (1.61 g, 7.53 mmol) were heated at 80°C for 5 h. After cooling the reaction mixture to room temperature, the red solid obtained was triturated with ether, filtered and dried under vacuum. To the solid, 1 N NaOH (25 mL) was added and stirred for 1 h at RT. The reaction mixture was then extracted with CH₂Cl₂ (3 x 30 mL) and the combined organic layers were dried over Na₂SO₄, filtered and concentrated in vacuo. The residue was purified by column chromatography on SiO₂ (PE/Ether, 3:1; R_f = 0.48) to give the title compound (1.28 g, 83%) as a yellow oil. ¹H NMR (400 MHz, DMSO): δ = 7.16 (dd, J = 7.3, 0.9 Hz, 1H), 7.06 (td, J = 7.7, 1.2 Hz, 1H), 6.74 (td, J = 7.5, 0.6 Hz, 1H), 6.67 (d, J = 7.9 Hz, 1H), 4.39 (s, 2H), 4.10 (q, J = 7.2 Hz, 2H), 3.87 (d, J = 2.2 Hz, 1H), 3.81 (d, J = 2.0 Hz, 1H), 1.29 (s, 6H), 1.17 (t, J = 7.2 Hz, 3H) ppm. ¹³C NMR (100 MHz, DMSO): δ = 168.6, 161.0, 145.4, 136.7, 127.3, 121.8, 118.8, 105.7, 74.8, 60.4, 43.6, 43.5, 29.6 (x 2) 14.1 ppm. HRMS (ESI): m/z: calcd for C₁₅H₁₉NO₂ [M+H]⁺ 246.149; found 246.150.

2-(3,3-dimethyl-2-methyleneindolin-1-yl)acetic acid (2). To a solution of Ethyl 2-(3,3-dimethyl-2-methyleneindolin-1-yl)acetate (1.30 g, 5.30 mmol) in ethanol (40 mL) was added 1N NaOH solution (30 mL) at RT and the mixture was stirred at the same temperature for overnight. Later the solvent was removed in vacuo and the residue was dissolved in minimum amount of water (15 mL) and then acidified to pH=3 with 1N HCl solution. The solution was concentrated and the residue was dissolved in ethanol. The suspension was filtered and this process was repeated at least for two times. Finally the filtrate was

concentrated and the residue was purified by column chromatography on SiO₂ (CH₂Cl₂/MeOH 8:2; R_f = 0.20) to give the title compound (860 mg, 75 %) as a thick brown oil. ¹H NMR (400 MHz, DMSO): δ = 7.15 (dd, *J* = 7.3, 0.7 Hz, 1H), 7.06 (td, *J* = 7.7, 1.2 Hz, 1H), 6.73 (t, *J* = 7.2 Hz, 1H), 6.65 (d, *J* = 7.9 Hz, 1H), 4.26 (s, 2H), 3.85 (d, *J* = 2.0 Hz, 1H), 3.81 (d, *J* = 2.0 Hz, 1H), 1.28 (s, 6H) ppm, CO₂H signal not visible. ¹³C NMR (100 MHz, DMSO): δ = 170.1, 160.8, 145.5, 136.7, 127.3, 121.7, 118.6, 105.6, 74.4, 43.6, 43.5, 29.7 (x 2) ppm. HRMS (ESI): *m/z*: calcd for C₁₃H₁₅NO₂ [M+H]⁺ 218.116; found 218.119.

Ethyl 2-((2E)-2-((E)-3-formylallylidene)-3,3-dimethylindolin-1-yl)acetate (3). Acetic anhydride (12.5 mL) was added dropwise to dimethylaminoacrolein (1.76 g, 1.78 mL, 17.8 mmol) and the mixture was stirred at RT for 15 minutes. Then a solution of Ethyl 2-(3,3-dimethyl-2-methyleneindolin-1-yl)acetate (1.25 g, 5.09 mmol) in CH₂Cl₂ (5 mL) was added dropwise. After stirring the reaction mixture for 2 d at RT, the solvent was removed in vacuo. The residue was dissolved in THF (30 mL) and a saturated sol. of K₂CO₃ (60 mL) was added carefully. After stirring the mixture for 30 min at RT, the layers were separated and the aqueous layer was extracted with EtOAc (3 x 20 mL). The combined organic layers were dried over Na₂SO₄, filtered and concentrated in vacuo. The residue was purified by column chromatography on SiO₂ (CH₂Cl₂/EtOAc 9.5:0.5; R_f = 0.20) to give the title compound (1.32 g, 86%) as a yellow oil. ¹H NMR (400 MHz, DMSO): δ = 9.46 (d, *J* = 8.3 Hz, 1H), 7.88-7.82 (m, 1H), 7.37 (d, *J* = 7.3 Hz, 1H), 7.19 (td, *J* = 7.8, 1.1 Hz, 1H), 6.98-6.93 (m, 2H), 5.81 (dd, *J* = 14.2, 8.3 Hz, 1H), 5.62 (d, *J* = 12.6 Hz, 1H), 4.70 (s, 2H), 4.14 (q, *J* = 7.3 Hz, 2H), 1.59 (s, 6H), 1.19 (t, *J* = 7.3 Hz, 3H) ppm. ¹³C NMR (100 MHz, DMSO): δ = 192.4, 167.8, 164.6, 150.2, 143.4, 138.5, 127.8, 123.6, 121.9, 121.4, 107.8, 95.3, 61.0, 46.3, 43.7, 27.9 (x 2), 14.0 ppm. HRMS (ESI, positive mode): *m/z*: calcd for C₁₈H₂₁NO₃ [M+H]⁺ 300.159; found 300.160.

2-((2E)-2-((E)-3-formylallylidene)-3,3-dimethylindolin-1-yl)acetic acid (4). To a solution of Ethyl 2-((2E)-2-((E)-3-formylallylidene)-3,3-dimethylindolin-1-yl)acetate (300 mg, 1.00 mmol) in EtOH (2 mL) was added 1N NaOH (1.5 mL) at RT and the mixture was stirred at the same temperature for 1.5 h. Later the solvent was removed in vacuo and the residue was dissolved in minimum amount of water (3 mL) and then acidified to pH=3 with 1N HCl solution. The aqueous layer was extracted with EtOAc (4 x 10 mL). The combined organic layers were dried over Na₂SO₄, filtered and concentrated in vacuo to give the title compound

(260 mg, 96%) as a blue-green solid. The crude product was directly used for PNA coupling without further purification. ^1H NMR (400 MHz, DMSO): δ = 13.1 (br. s, 1H), 9.45 (d, J = 8.3 Hz, 1H), 7.88-7.81 (m, 1H), 7.36 (d, J = 7.0 Hz, 1H), 7.19 (td, J = 7.8, 1.1 Hz, 1H), 6.98-6.93 (m, 2H), 5.81 (dd, J = 14.2, 8.3 Hz, 1H), 5.62 (d, J = 12.6 Hz, 1H), 4.58 (s, 2H), 1.59 (s, 6H) ppm. ^{13}C NMR (100 MHz, DMSO): δ = 192.3, 169.1, 164.8, 150.4, 143.6, 138.6, 127.8, 123.3, 121.8, 121.3, 107.9, 95.1, 46.4, 43.7, 27.9 (x 2) ppm. HRMS (ESI, negative mode): m/z : calcd for $\text{C}_{16}\text{H}_{17}\text{NO}_3$ $[\text{M}-\text{H}]^-$ 270.112; found 270.110.

PNA synthesis: PNAs were synthesized on a solid support using Fmoc chemistry. The PNA strand was cleaved from the resin with trifluoroacetic acid/ triisopropylsilane/ H_2O (95:2.5:2.5) solution and purified by HPLC. The final compounds were characterized by MALDI-TOF: **Pna1** m/z : $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{87}\text{H}_{116}\text{N}_{38}\text{O}_{19}^+$ 1997.928, found 1997.718; **Pna2** m/z : $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{88}\text{H}_{124}\text{N}_{39}\text{O}_{19}^+$ 2030.993, found 2030.923; **Pna3** m/z : $[\text{M}+\text{H}]^+$ calcd for $\text{C}_{84}\text{H}_{115}\text{N}_{36}\text{O}_{19}^+$ 1931.913, found 1932.082

DNA folding: Hairpin and quadruplex DNA folding was carried out by heating a DNA solution (5 μM) in potassium phosphate buffer (10 mM, pH 7.4 also containing 100 mM KCl) at 95°C for 10 min and then letting it slowly cool down to room temperature over 6 hours.

Fluorescence studies: Fluorescence emission spectra were recorded in quartz cells (500 μL with a 5 mm path length) at 20 °C on a Jobin Yvon Fluorolog 3.22 instrument. The excitation and emission bandwidths were fixed to 5 nm and 5 nm, respectively.

Ligation/Sensing experiments: In a typical experiment, 500 μL of a mixture of **Pna1** (500 nM), **Pna2** (500 nM) and pre-folded hairpin DNA (500 nM) in a potassium phosphate buffer (10 mM, pH 7.4 also containing 100 mM KCl) was prepared freshly in a fluorescence Quartz cuvette (5 mm path length) and kept sealed at 20 °C for the duration of experiment. Formation of the pentamethine cyanine dye product was then monitored by fluorescence spectroscopy (at t_0 and then every hour over 10 hours).

Dual sensing experiments: After a screening of various concentrations of both DNA and PNA sequences, the following ratio 1:4:4:2:4 **Pna1/Pna2/Pna3/Hairp1a/Quad1** was found to give the optimal sensitivity and selectivity. Working with stoichiometric amounts of all

PNA and DNA strands resulted in good sensitivity but lower selectivity (data not shown). Ligation experiments in the presence of both quadruplex and hairpin DNA were carried out in 500 μ L quartz cuvettes by incubating a freshly prepared mixture of three PNA probes (**Pna1-3**) and two DNA targets (**Hairp1a** and **Quad1**) in potassium phosphate buffer.

Results and Discussion:

Design and synthesis of the PNA probes: Two 5-mer PNAs (**Pna1** and **Pna2**) were synthesized on solid support, which also contained two ϵ -*N,N*-dimethyl Lysine residues to ensure solubility of both PNAs in water at near-physiological pH. These PNA probes were functionalized at their C- or N-terminus with two precursors for the synthesis of pentamethine cyanine dyes: *N*-alkyl-2,3,3-trimethylindolenine **2** and aldehyde **4**, respectively. Both precursors were synthesized from commercially available 2,3,3-trimethylindoline (Scheme 1).

Reaction of 2,3,3-trimethyl-3H-indole with ethyl iodoacetate gave the corresponding quarternized salt which was subsequently isomerized to 2-methylene-indolenin **1** via treatment with 1N NaOH solution. Condensation of **1** with dimethylaminoacrolein in the presence of acetic anhydride provided the activated hemicyanine intermediate which was then hydrolyzed into aldehyde **3** with a saturated K_2CO_3 solution. Final hydrolysis of the ethyl ester of compounds **1** and **3** using 1N NaOH afforded both C5 precursors **2** and **4**, respectively. While indolenine **2** is non-fluorescent, aldehyde **4** emits maximally at 476 nm when excited at 410 nm. Reaction between compounds **2** and **4** leads to the formation of a symmetrical pentamethine cyanine dye (C5) characterized by its unique fluorescent properties ($\lambda_{exc} = 625$ nm, $\lambda_{em} = 660$ nm), very different from those of its two precursors. Although highly unfavorable in water, this aldolisation-elimination reaction can be templated when carried out in a favorable chemical environment. **Pna1** and **Pna2** (see Figure 2 and Table 1 for structures and sequences) were synthesized on rink amide MBHA resin (Merck Biosciences, loading 0.56 mmol/g) using standard solid-phase Fmoc chemistry. They were subsequently purified by reversed-phase HPLC and characterized by MALDI.

5-mers PNA was chosen as the minimal length necessary for a full and sequence-specific PNA-DNA hybridization under near-physiological conditions. Based on the abundant literature on PNAs (37), it is anticipated that our short PNA/DNA duplexes will have melting temperatures above 40°C, indicating that almost all the PNA probes will be

hybridized to the DNA target in the conditions of our experiments. It is also noteworthy that the presence of two Lysine residues is likely to increase even further the stability of the heteroduplex formed. In case of unusually unstable (AT rich) sequences, however, stability of the PNA-DNA heteroduplex would need to be increased. For instance, this could be achieved by increasing the length of the PNA strands, without interfering with the reaction of cyanine dye formation which occurs at the other end of the PNA.

Hairpin DNA sensing: Reaction between stoichiometric amounts of **Pna1** and **Pna2** (500 nM each) was investigated in a 10 mM potassium phosphate solution (pH=7.4) also containing 100 mM potassium chloride, in the presence and in the absence of various nucleic acid templates. In each case, reaction efficiency was determined by monitoring the formation of the pentamethine cyanine (product of the reaction between indolenine **2** and aldehyde **4**) by fluorescence spectroscopy over a ten hour period. In the absence of any template, and at a 500 nM concentration of both PNAs, only traces of the fluorescent cyanine dye product were ever detectable. We then investigated the possibility to use our PNA probes for sensing hairpin DNA structures. Three different strategies were developed that consisted in targeting either the hairpin single-stranded flanking arms or the hairpin loop (Figure 2). In the former approach, a hairpin structure was designed which contained two single-stranded flanking arms at its 3' and 5' ends, the sequences of which were complementary to those of **Pna1** and **Pna2**. This system (sequences **Hairp3a-c**, table1) was designed so that (i) **Pna1** and **Pna2** could hybridize in a sequence-specific manner to these two single-stranded arms and that (ii) the reactive functionalities of both PNAs would be brought in close enough proximity to react with each other only when the hairpin was formed, while kept separated in the absence of any DNA folding.

In the latter approach, a hairpin DNA structure was chosen that also contained a large central loop, the sequence of which included both **Pna1** and **Pna2** complementary sequences (Figure 2). Once again, only upon simultaneous hybridization of both PNA probes to their complementary DNA sequences located in the loop region, the aldehyde will be able to react with the 2-methylene-indolenine. Two systems were designed in which the cyanine dye was expected to form either at the centre of the loop (sequences **Hairp2a-c**, table1) or at the intercept between the loop and the stem (sequences **Hairp1a-d**, table1). It is noteworthy that (i) all DNA sequences **Hairp1-3** were shown to form stable hairpin structures under the conditions of our experiment and in the absence of PNA probes, and (ii) DNA hairpins

Hairp1 and **Hairp2** were designed which contained larger loops than **Hairp3** in order to allow complete hybridization of both 5-mer PNAs to the hairpin loop.

The potential of **Hairp3a-c**, **Hairp2a-c** and **Hairp1a-c** to serve as a DNA template for the reaction of cyanine dye formation was determined. Briefly, a stoichiometric mixture of **Pna1**, **Pna2** and one hairpin DNA was reacted in potassium phosphate buffer at physiological pH and the appearance of the cyanine dye product was monitored by fluorescence spectroscopy ($\lambda_{\text{exc}} = 625 \text{ nm}$). As a control experiment, a reaction was carried out under the same conditions but lacking the DNA template. Interestingly, every hairpin DNA tested proved capable to significantly increase the reaction efficiency, although at different levels depending on whether the PNA probes were targeting the loop or the stem of the hairpin (Figure 3). Moreover, as previously observed by us with other fluorogenic systems (27 and unpublished data) and although no detectable fluorescence is observed in the absence of template ($F_0 < 350$), DNA-catalyzed reactions are all characterized by a fast and almost instantaneous appearance of fluorescence ($F_0 > 6000$) followed by a slower increase over 10-12 hours. The kinetics of cyanine dye formation proved generally slow, the fluorescence intensity still rising even after 10 hours, even in the presence of a template (Figure 3). However, the facts that (i) only residual amounts of fluorescent product form in the absence of DNA template (i.e. very low background signal) and (ii) a very fast initial phase of cyanine dye formation is observed in the presence of template make this chemistry suitable for the detection of DNA secondary structures. Indeed, a signal-to-noise ratio of ≥ 17 was obtained for all hairpin DNA-templated reactions after less than 10 minutes reaction.

Least effective templating effect was obtained when targeting the 3' and 5' single-stranded flanking ends of the hairpin (**Hairp3a-c**). This effect also proved almost independent of the number of nucleotides located between the PNA complementary sequences and the first nucleotide involved in the stem formation. Indeed, comparable fluorescence intensities were obtained from **Hairp3a**, **Hairp3b** and **Hairp3c** which contain a one-, two- and three-nucleotide gap between the PNA/DNA duplex regions and the hairpin stem, respectively (Figure 3a).

Sensing hairpin formation proved however more efficient when both PNAs were directed against the hairpin loop, the amount of cyanine dye formed after 10 h reaction being for instance twice higher with **Hairp1a** than with **Hairp3a**. Two slightly different approaches were considered for targeting hairpin loops, the reaction of cyanine dye formation taking place either at the centre of the loop or at its base, in the vicinity of the stem. Both approaches

proved equally effective (same fluorescence intensity at the maximum emission wavelength obtained after 10 h with **Hairp1a** and **Hairp2b**) although being highly dependent on the overall design. While a four-nucleotide gap between the functional ends of **Pna1** and **Pna2** led to the optimal reaction efficiency within the **Hairp2** series (**Hairp2b**, Figure 3b), hybridization of both PNAs starting from the first nucleotides constitutive of the loop was found twice more effective than systems where **Pna1** and **Pna2** were separated further (**Hairp1a** versus **Hairp1b-c**, Figure 3c).

Whilst **Hairp1a** and **Hairp2b** proved equally good templates for the aldolisation-elimination reaction, they are based on two different strategies that can find distinct applications in modern bioresearch, in particular for sensing hairpin DNA formation in vitro and maybe also in vivo. In order to demonstrate the strong link between hairpin formation and pentamethine cyanine dye formation, an analogue of **Hairp1a** was designed which was missing six nucleotides at its 3'-end in order to prevent hairpin formation. The ability of this new single-stranded DNA (so-called **Hairp1d**) to act as a template was investigated by fluorescence spectroscopy (grey curve, Figure 3c). Unsurprisingly, the fluorescence intensity obtained with **Hairp1d** was c.a. 10 times lower than that obtained with **Hairp1a**, thus demonstrating that the templating effect observed with the latter was indeed structure-specific.

Whether the hairpin remains “hairpin-like” or opens up upon hybridization of both PNAs to the DNA template remains to be demonstrated. **Hairp2** is likely to open after PNA ligation, as it is commonly observed with MBs (thus forming a highly stable 10-mer PNA/DNA heteroduplex with single-stranded DNA flanking arms). In the case of **Hairp1** and **Hairp3**, however, the complex formed between the ligation product and the DNA template can be defined as a PNA:DNA hybrid 3-way junction and could therefore be significantly more stable than the “open” structure. Additional structural studies will be required to further address this question.

Dual Hairpin+Quadruplex DNA sensing using a dual-color system: We recently demonstrated that two fluorogenic PNA probes (**Pna2** and **Pna3**, Figure 4a) could be used for sensing the formation of G-quadruplex DNA in vitro. The system was designed so that simultaneous and sequence-specific hybridization of both PNAs to a unique G-quadruplex-containing oligonucleotide were the only conditions inducing the formation of a fluorescent trimethine cyanine dye ($\lambda_{em} = 563$ nm). (27) Since the fluorescent properties of this

trimethine cyanine dye were significantly different compared to those of **Pna1** aldehyde ($\lambda_{em} = 476$ nm) and to those of the pentamethine analogue formed when reacting **Pna1** and **Pna2** ($\lambda_{em} = 660$ nm), we reasoned that it should be possible to simultaneously detect the formation of both hairpin and quadruplex DNAs using a non-overlapping, two-color “Cy3-Cy5” system. The system was designed so that a unique nucleophilic PNA, **Pna2**, could potentially react with two electrophilic PNA aldehydes, **Pna1** and **Pna3**, to form a pentamethine ($\lambda_{em} = 660$ nm) and a trimethine ($\lambda_{em} = 563$ nm) cyanine dye, respectively. While **Pna1** and **Pna2** are directed against hairpin DNA **Hairp1a**, **Pna2** and **Pna3** are directed against quadruplex DNA **Quad1**. Hence, it was our expectation to be able to specifically correlate quadruplex and hairpin DNA formation with the appearance of characteristic fluorescence signals at $\lambda_{em} = 563$ nm and $\lambda_{em} = 660$ nm, respectively. In order to demonstrate the concept, a mixture of **Pna1** (250 nM), **Pna2** (1 μ M) and **Pna3** (1 μ M) was incubated under near-physiological conditions and in the presence of various DNA templates: (i) **Hairp1a** (500 nM), (ii) **Quad1** (1 μ M) and (iii) a mixture of **Hairp1a** and **Quad1** (500 nM and 1 μ M, respectively). In each case, the formation of the trimethine and pentamethine cyanine dyes were monitored by fluorescence spectroscopy ($\lambda_{exc} = 540$ and 625 nm, respectively) (Figure 4a). Under our experimental conditions, detection of hairpin DNA proved highly specific, even in the presence of an alternative DNA secondary structure. Indeed, comparable fluorescence intensities were obtained when using as a template either **Hairp1a** alone or a mixture of **Hair1a** and **Quad1**, while only residual fluorescence was detectable when using the quadruplex **Quad1** alone (Figure 4b).

Detection of quadruplex DNA in the presence of hairpin DNA also proved possible, the amount of trimethine cyanine dye formed decreasing by only <5% upon addition of competitor hairpin DNA (Figure 4c). However, significant background fluorescence (c.a. 50%) was also obtained under our conditions which could be interpreted as the result of a non-specific formation of trimethine cyanine dye templated by hairpin DNA. So far, attempts to combine on a single oligonucleotide the quadruplex and hairpin forming sequences (both structures being separated by a polyT linker) in order to simultaneously detect both structures were unsuccessful due to high background signal which is likely due to the formation of complex tertiary structures resulting in non-specific reaction of dye (C3 and C5) formation.

Conclusion: Detection of nucleic acid secondary structures remains a key challenge for chemists due to their large abundance and structural diversity within the genome. Of

particular interest are structures located in regions essential for DNA replication or transcription (e.g. in gene promoters) which could serve as therapeutic targets for regulating gene expression. Whether a DNA strand can potentially form into a hairpin or a quadruplex conformation based on the DNA primary sequence can now be easily predicted, and with reasonably good accuracy. However, whether these structures actually form (or not) depending on the context (in vitro or in vivo) or the experimental conditions remains much more challenging. Therefore, fluorescent probes capable of detecting the folding of specific nucleic acid sequences would represent an important way to validate the existence of such structures in vivo and also maybe to identify new targets. Herein, we reported three different strategies for sensing hairpin DNA formation by targeting either the loop or the stem with short PNAs functionalized at their 3' or 5' end with two precursors of a pentamethine cyanine dye. When working at low PNA concentrations (500 nM each), the aldolisation-elimination reaction is highly unfavorable, thus resulting in no significant appearance of fluorescence. However, a significant increase (>50-fold) in fluorescence intensity can be obtained upon sequence-specific hybridization of both PNA to either the loop or the stem of the hairpin target. Therefore, this PNA-based approach represents an original strategy for sensing hairpin DNA formation in vitro, the best sensitivity being obtained when targeting the hairpin loop with the cyanine dye forming in the vicinity of the hairpin stem (**Hairp1a**). Compared to other DNA-templated reactions described in the literature that also involve the formation of a fluorescent product, our system offers the significant advantages of (i) absorbing and emitting in the far-red part of the spectrum ($\lambda_{em} = 660 \text{ nm}$), thus making it suitable for in vivo applications, and (ii) requiring no catalysis, thus making it biocompatible. It is also noteworthy that this unprecedented DNA-templated C5 synthesis can be orthogonal to the DNA-templated C3 synthesis recently reported by us, (27) thus allowing simultaneous detection of at least two different structures with a two-color (C3-C5) system. Although such approach still requires optimization, we demonstrated that it was possible to use a combination of three PNA probes for sensing, within a mixture of hairpin and quadruplex DNA, either or both secondary structures. The development of PNA probes of increased specificity (e.g. longer PNAs) and higher sensitivity for in vivo applications is currently underway in our group.

Acknowledgements: The authors are very thankful to the Centre National de la Recherche Scientifique (CNRS) and the International Centre for Frontier Research in Chemistry (icFRC) for financial support.

References:

- (1) Watson, J. D., and Crick, F. H. C. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 737-738.
- (2) Wadkins, R. M. (2000) Targeting DNA Secondary Structures. *Current Med. Chem.* 7, 1-15.
- (3) G.N. Parkinson Quadruplex Nucleic Acids; S. Neidle, S. Balasubramanian Eds; RSC Biomolecular Sciences Publishing, Cambridge, UK, 2006, pp. 1-30.
- (4) Huppert, J. L., and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* 35, 406-413.
- (5) Qin, Y., and Hurley, L. H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie* 90, 1149-1171.
- (6) Eddy, J., and Maizels, N. (2008) Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nucleic Acids Res.* 36, 1321-1333.
- (7) Verma, A., Yadav, V. K., Basundra, R., Kumar, A., and Chowdhury, S. (2009) Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. *Nucleic Acids Res.* 37, 4194-4204.
- (8) Wong, H. M., and Huppert, J. L. (2009) Stable G-quadruplexes are found outside nucleosome-bound regions. *Mol. BioSyst.* 5, 1713-1719.
- (9) Panayotatos, N., and Fontaine, A. (1987) A native cruciform DNA structure probed in bacteria by recombinant T7 endonuclease. *J. Biol. Chem.* 262, 11364-11368.
- (10) Varani, G. (1995) Exceptionally stable nucleic acid hairpins. *Annu. Rev. Biophys. Biomol. Struct.* 24, 379-404.

- (11) Baker, E. S., Dupuis, N. F., and Bowers, M. T. (2009) DNA hairpin, pseudoknot, and cruciform stability in a solvent-free environment *J. Phys. Chem. B* *113*, 1722-1727.
- (12) Dai, X., Greizerstein, M. B., Nadas-Chinni, K., and Rothman-Denes, L. B. (1997) Supercoil-induced extrusion of a regulatory DNA hairpin. *Proc. Natl Acad. Sci. USA* *94*, 2174-2179.
- (13) Willwand, K., Mumtsidu, E., Kuntz-Simon, G., and Rommelaere, J. (1998) Initiation of DNA replication at palindromic telomeres is mediated by a duplex-to-hairpin transition induced by the minute virus of mice nonstructural protein NS1. *J. Biol. Chem.* *273*, 1165-1174.
- (14) Dervan, P. B., Poulin-Kerstien, A. T., Fechter, E. J., and Edelson, B. S. (2005) Regulation of Gene Expression by Synthetic DNA-Binding Ligands. DNA Binders and Related Subjects. *Topics in current chemistry* *253*, 1-31.
- (15) Caldarelli, S. A., Mehiri, M., Di Giorgio, A., Martin, A., Hantz, O., Zoulim, F., Terreux, R., Condom, R., and Patino, N. (2005) A cyclic PNA-based compound targeting domain IV of HCV IRES RNA inhibits in vitro IRES-dependent translation. *Bioorg. Med. Chem.* *13*, 5700-5709.
- (16) Bejugam, M., Sewitz, S., Shirude, P. S., Rodriguez, R., Shahid, R., and Balasubramanian, S. (2007) Trisubstituted isoalloxazines as a new class of G-quadruplex binding ligands: small molecule regulation of c-kit oncogene expression. *J. Am. Chem. Soc.* *129*, 12926-12927.
- (17) Waller, Z. A., Sewitz, S. A., Hsu, S. T., and Balasubramanian, S. (2009) A small molecule that disrupts G-quadruplex DNA structure and enhances gene expression. *J. Am. Chem. Soc.* *131*, 12628-12633.
- (18) Wu, P., Ma, D. L., Leung, C. H., Yan, S. C., Zhu, N., Abagyan, R., and Che, C. M. (2009) Stabilization of G-quadruplex DNA with platinum(II) Schiff base complexes:

luminescent probe and down-regulation of c-myc oncogene expression. *Chem. Eur. J.* *15*, 13008-13021.

(19) Wang, K., Tang, Z., Yang, C. J., Kim, Y., Fang, X., Li, W., Wu, Y., Medley, C. D., Cao, Z., Li, J., Colon, P., Lin, H., and Tan, W. (2009) Molecular Engineering of DNA: Molecular Beacons. *Angew. Chem. Int. Ed.* *48*, 856-870.

(20) Ranasinghe, R. T., and Brown, T. (2005) Fluorescence based strategies for genetic analysis. *Chem. Commun.*, 5487-5502.

(21) Cai, J., Li, X., Yue, X., and Taylor, J. S. (2004) Nucleic acid-triggered fluorescent probe activation by the Staudinger reaction. *J. Am. Chem. Soc.* *126*, 16324-16325.

(22) Pianowski, Z. L., and Winssinger, N. (2007) Fluorescence-based detection of single nucleotide permutation in DNA via catalytically templated reaction. *Chem. Commun.*, 3820-3822.

(23) Franzini, R. M., and Kool, E. T. (2008) 7-Azidomethoxy-coumarins as profluorophores for templated nucleic acid detection. *ChemBioChem* *9*, 2981-2988.

(24) Z. Pianowski, Z., Gorska, K., Oswald, L., Merten, C. A., and Winssinger, N. (2009) Imaging of mRNA in live cells using nucleic acid-templated reduction of azidorhodamine probes. *J. Am. Chem. Soc.* *131*, 6492-6497.

(25) Furukawa, K., Abe, H., Hibino, K., Sako, Y., Tsuneda, S., and Ito, Y. (2009) Reduction-triggered fluorescent amplification probe for the detection of endogenous RNAs in living human cells. *Bioconjug. Chem.* *20*, 1026-1036.

(26) Huang, Y., and Coull, J. M. (2008) Diamine catalyzed hemicyanine dye formation from nonfluorescent precursors through DNA programmed chemistry. *J. Am. Chem. Soc.* *130*, 3238-3239.

- (27) Meguellati, K., Koripelly, G., and Ladame, S. (2010) DNA-templated synthesis of trimethine cyanine dyes: a versatile fluorogenic reaction for sensing G-quadruplex formation. *Angew. Chem. Int. Ed.* 49, 2738-2742.
- (28) Franzini, R. M., and Kool, E. T. (2008) Organometallic activation of a fluorogen for templated nucleic acid detection. *Org. Lett.* 10, 2935-2938.
- (29) Shibata, A., Abe, H., Ito, M., Kondo, Y., Shimizu, S., Aikawa, K., and Ito, Y. (2009) DNA templated nucleophilic aromatic substitution reactions for fluorogenic sensing of oligonucleotides. *Chem. Commun.*, 6586-6588.
- (30) Nielsen, P. E., Egholm, M., Berg, R. H., and Buchardt, O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* 254, 1497-1500.
- (31) Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B., and Nielsen, P. E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* 365, 566-568.
- (32) Grossmann, T. N., and Seitz, O. (2009) Nucleic acid templated reactions: consequences of probe reactivity and readout strategy for amplified signaling and sequence selectivity. *Chem. Eur. J.* 15, 6723-6730.
- (33) Hansen, M. E., Bentin, T., and Nielsen, P. E. (2009) High-affinity triplex targeting of double stranded DNA using chemically modified peptide nucleic acid oligomers. *Nucleic Acids Res.* 37, 4498-4507.
- (34) Wojciechowski, F., and Hudson, R. H. (2009) Peptide nucleic acid containing a meta-substituted phenylpyrrolocytosine exhibits a fluorescence response and increased binding affinity toward RNA. *Org. Lett.* 11, 4878-4881.

(35) Socher, E., Jarikote, D. V., Knoll, A., Röglin, L., Burmeister, J., and Seitz, O. (2008)

FIT probes: peptide nucleic acid probes with a fluorescent base surrogate enable real-time

DNA quantification and single nucleotide polymorphism discovery. *Anal. Biochem.* 375,

318-330.

(36) Oquare, B. Y., and Taylor, J. S. (2008) Synthesis of peptide nucleic acid FRET probes

via an orthogonally protected building block for post-synthetic labeling of peptide nucleic

acids at the 5-position of uracil. *Bioconjug. Chem.* 19, 2196-2204.

(37) Sugimoto, N., Satoh, N., Yasuda, K., and Nakano, S. (2001) Stabilization factors

affecting duplex formation of peptide nucleic acid with DNA. *Biochemistry* 40, 8444-8451.

Tables

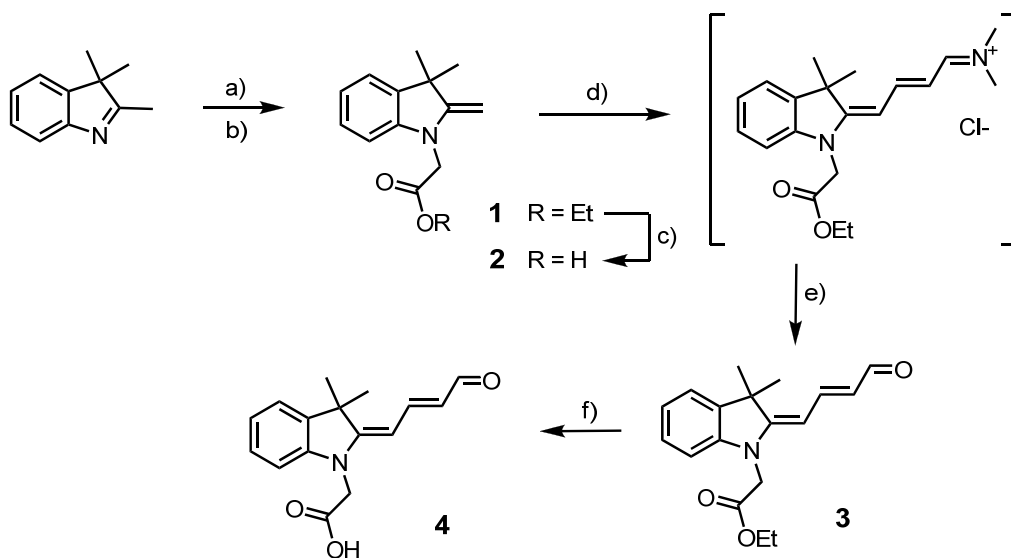
Table 1. Oligonucleotide and PNA sequences

	DNA ^[a] and PNA ^[b] sequences
Pna1	Lys(NMe ₂)-Lys(NMe ₂)- <u>GAGTG</u> -Ald4 ^[b]
Pna2	Ind2- <u>AGCCG</u> -Lys(NMe ₂)-Lys(NMe ₂) ^[b]
Hairp1a	<u>CCTAGCTCGGCTTTTTCTCACGCTAGG</u>
Hairp1b	<u>CCTAGCTTCGGCTTTTCTCACTGCTAGG</u>
Hairp1c	<u>CCTAGCTTTCGGCTCTCACTTGCTAGG</u>
Hairp1d	<u>CCTAGCTCGGCTTTTTCTCAC</u>
Hairp2a	<u>CCTAGCTTTTCTCACTTTTTCGGCTTTGCTAGG</u>
Hairp2b	<u>CCTAGCTTCTCACTTTTTCGGCTTTGCTAGG</u>
Hairp2c	<u>CCTAGCTTTTCTCACTTTCGGCTTTGCTAGG</u>
Hairp3a	<u>CTCACTTGCGATACTCATCGCATTTCGGC</u>
Hairp3b	<u>CTCACTTTGCGATACTCATCGCATTTCGGC</u>
Hairp3c	<u>CTCACTTTTTCGGATACTCATCGCATTTCGGC</u>
Quad1	<u>GCATCCGGGCGGGCGCGAGGGAGGGTTTCGGC</u>

^[a] DNA sequences are given from the 5' to 3' end.

^[b] PNA sequences are given from the C-terminal to N-terminal end.

Schemes and legends



Scheme 1. Synthesis of 2-methylene-Indolenin derivative and Fischer's base Aldehyde.

Reagents and conditions: a) ethyl iodoacetate, 80°C, 5 h; b) 1N NaOH, EtOH, 25°C, 1 h (83% over two steps), c) 1N NaOH, EtOH, 25°C, 16 h (75%); d) dimethylaminoacrolein, AcOH, CH₂Cl₂, 48 h; e) Sat. aq. K₂CO₃, THF, 25°C, 1 h, (86% over two steps); f) 1N NaOH, EtOH, 25°C, 1.5 h, (96%).

Figures and Legends

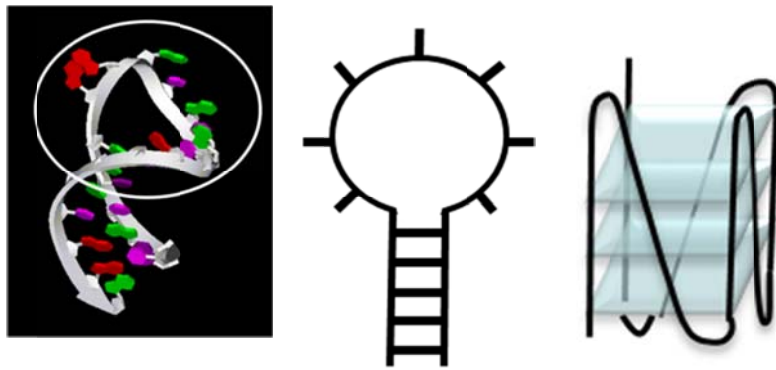


Figure 1. X-ray crystal structure and schematic representation of a DNA hairpin loop and schematic representation of a parallel-stranded intramolecular G-quadruplex (from left to right).

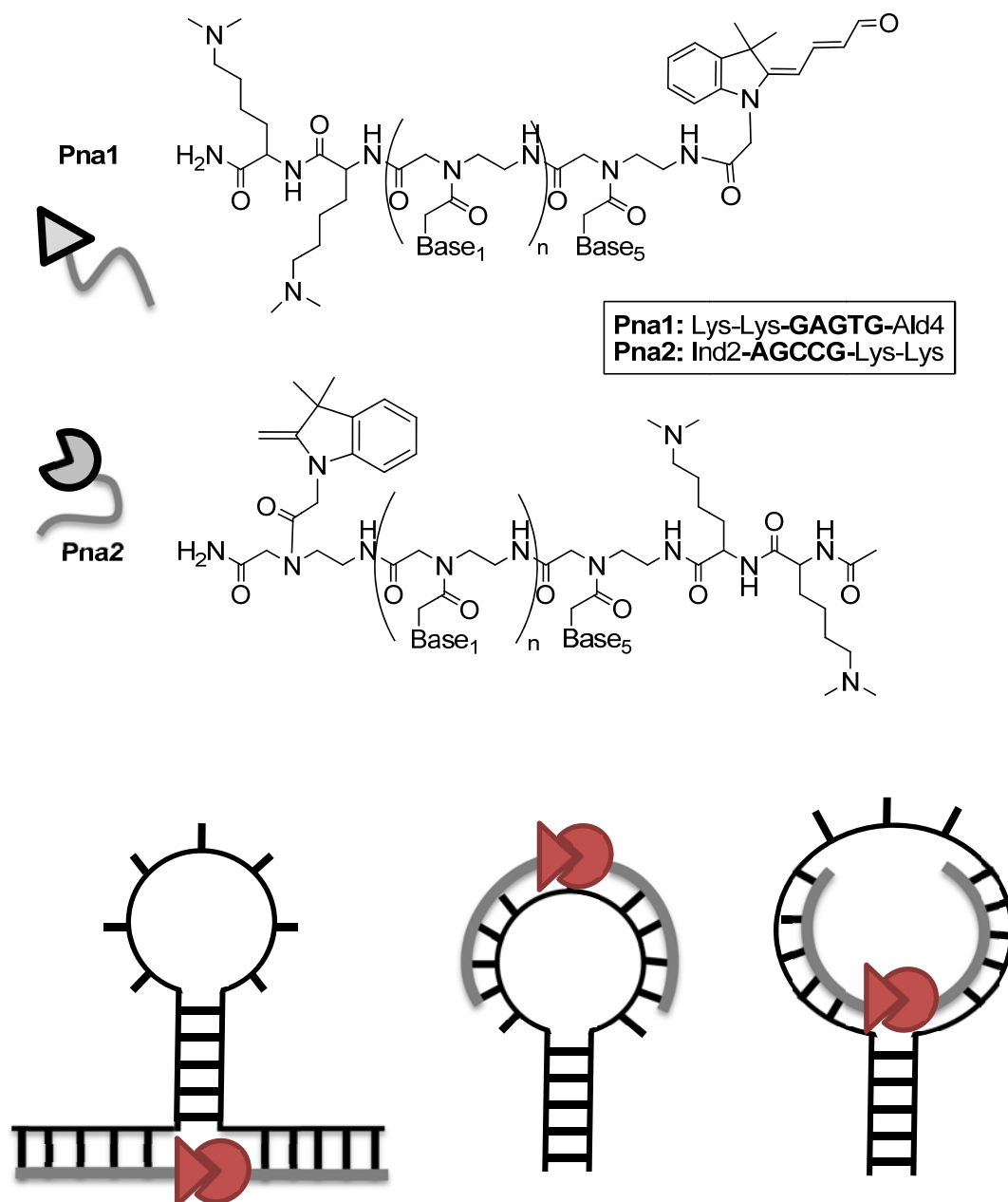


Figure 2. Structures of PNA probes **Pna1** and **Pna2** (top). Schematic representation of the hairpin DNA-templated synthesis of the pentamethine cyanine dye upon simultaneous hybridisation of both PNA probes to either the hairpin flanking regions (**Hairp3a-c**) or the hairpin loop (**Hairp2a-c** and **Hairp1a-c**) (bottom).

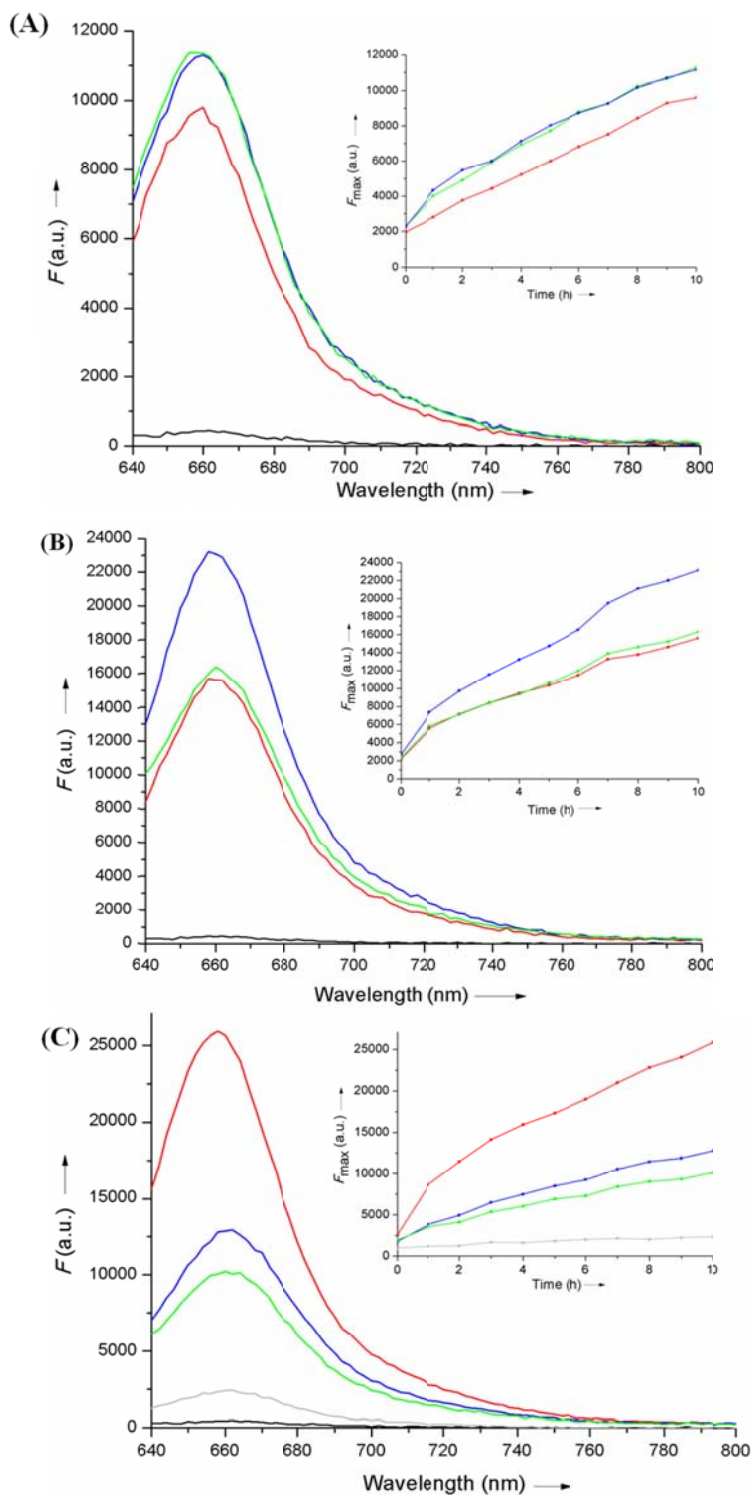


Figure 3. Fluorescence emission spectra of a stoichiometric mixture of **Pna1** and **Pna2** (500 nM each) measured after 10 h ($\lambda_{\text{exc}} = 625$ nm) in the absence (black) or in the presence of various DNA templates: (A) **Hairp3a** (red), **Hairp3b** (blue) and **Hairp3c** (green); (B) **Hairp2a** (red), **Hairp2b** (blue), **Hairp2c** (green); (C) **Hairp1a** (red), **Hairp1b** (blue), **Hairp1c** (green), **Hairp1d** (grey). Reaction time courses (maximum fluorescence intensity vs time) are also given as insets for each DNA templated reaction.

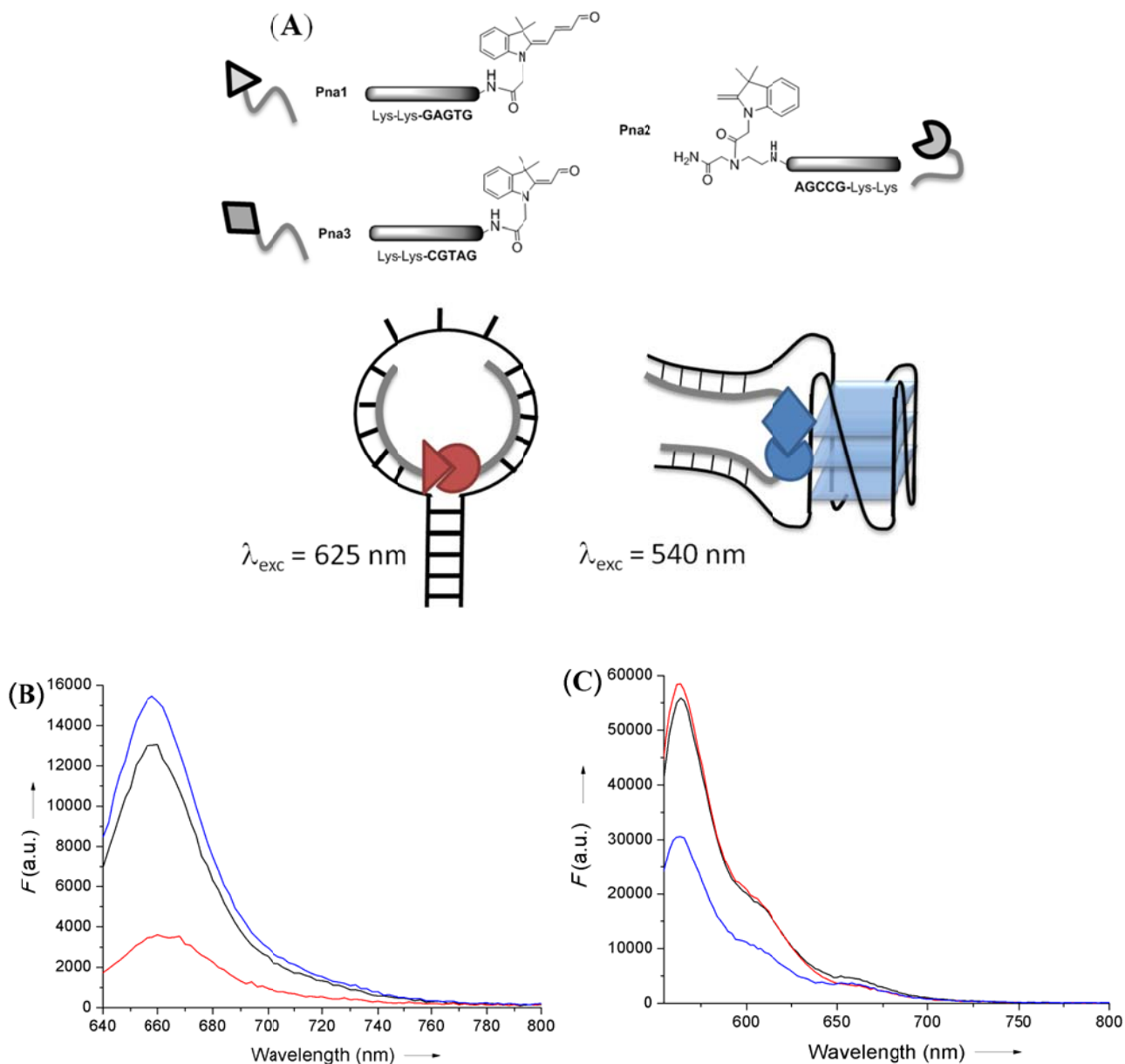


Figure 4. Schematic representation of the general system for sensing simultaneously hairpin and quadruplex DNA, based on a dual-colour detection system (A). Fluorescence emission spectra ($\lambda_{\text{exc}} = 540 \text{ nm}$) of a mixture of **Pna1** (250 nM), **Pna2** (1 μM) and **Pna3** (1 μM) in potassium phosphate buffer (10 mM, pH = 7.4 also containing 100 mM KCl) in the presence of 500 nM **Hairp1a** (blue), 1 μM **Quad1** (red), or a mixture of 500 nM **Hairp1a** and 1 μM **Quad1** (black). Excitation wavelengths were either $\lambda_{\text{exc}} = 625 \text{ nm}$ (B) or $\lambda_{\text{exc}} = 540 \text{ nm}$ (C). Emission spectra were recorded after 12 h incubation at room temperature.