

NAR Breakthrough Article

5'-fluoro(di)phosphate-labeled oligonucleotides are versatile molecular probes for studying nucleic acid secondary structure and interactions by ^{19}F NMR

Marek R. Baranowski ¹, Marcin Warminski ¹, Jacek Jemielity ² and Joanna Kowalska ^{1,*}

¹Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Ludwika Pasteura 5, 02-093 Warsaw, Poland and ²Centre of New Technologies, University of Warsaw, Stefana Banacha 2c, 02-097 Warsaw, Poland

Received April 29, 2020; Revised May 19, 2020; Editorial Decision May 20, 2020; Accepted May 22, 2020

ABSTRACT

The high sensitivity of ^{19}F nucleus to changes in the chemical environment has promoted the use of fluorine-labeled molecular probes to study structure and interactions of nucleic acids by ^{19}F NMR. So far, most efforts have focused on incorporating the fluorine atom into nucleobase and ribose moieties using either monomer building blocks for solid-phase synthesis, or nucleoside triphosphates for enzymatic synthesis. Here, we report a simple and efficient synthesis of 5'-fluoromonophosphorylated and 5'-fluorodiphosphorylated oligodeoxyribonucleotides, which combines solid-phase and in-solution synthesis methods and requires only commercially available nucleoside phosphoramidites, followed by their evaluation as ^{19}F NMR probes. We confirmed that the fluorine atom at the oligonucleotide 5' end did not alter the secondary structure of DNA fragments. Moreover, at the same time, it enabled real-time ^{19}F NMR monitoring of various DNA-related biophysical processes, such as oligonucleotide hybridization (including mismatch identification), G-quadruplex folding/unfolding and its interactions with thrombin, as well as formation of an *i*-motif structure and its interaction with small-molecule ligands.

INTRODUCTION

The development of robust methods enabling monitoring of oligonucleotide secondary and tertiary structure transformations, as well as detection of oligonucleotide interactions with (bio)molecules is crucial for improving our

understanding of nucleic acid function and the discovery of small molecules targeting nucleic acids. Various spectroscopic methods, including UV-VIS (1), circular dichroism (CD) (2,3), surface plasmon resonance (SPR) (4–6) and fluorescence-based techniques (7–9), have been used to achieve this goal. These methods have found numerous applications, but are nevertheless limited by the need to introduce bulky reporter groups (SPR and fluorescence), narrow scope, and the inability to provide detailed structural information. Nuclear magnetic resonance (NMR) spectroscopy, despite being less sensitive than the above methods, can provide more details about nucleic acid conformational equilibria, conversion mechanisms and interaction sites (10–13). NMR methods also enable the discovery and quantification of intermolecular interactions in the micromolar to millimolar K_D range (14,15). As a result, NMR techniques have become increasingly popular for the study of nucleic acids. To this end, the ^{19}F nucleus in fluorine-labeled nucleic acids exhibits several advantageous NMR-related properties, such as nuclear spin of $1/2$, high-magnetogyric coefficient and 100% abundance, which mean that ^{19}F NMR sensitivity amounts to 83% of ^1H NMR (16). The absence of fluorine in natural compounds makes typical ^{19}F NMR spectra simple and easy to analyze. Finally, fluorine nuclei are sensitive to changes in the local chemical environment, so even structurally similar entities can be distinguished by ^{19}F NMR. As a result, these features have transformed ^{19}F NMR studies on fluorine-labeled nucleic acids into a valuable biophysical tool (17,18).

Several strategies have been developed to obtain fluorine-labeled oligonucleotides for ^{19}F NMR experiments. Most of them encompass internal labeling of the oligonucleotide sequence with single or multiple fluorinated nucleotide-derived building blocks. The latter can be incorporated either by (i) solid-supported chemical synthesis or (ii)

*To whom correspondence should be addressed. Tel: +48 225 543 774; Fax: +48 225 543 771; Email: jkowalska@fuw.edu.pl

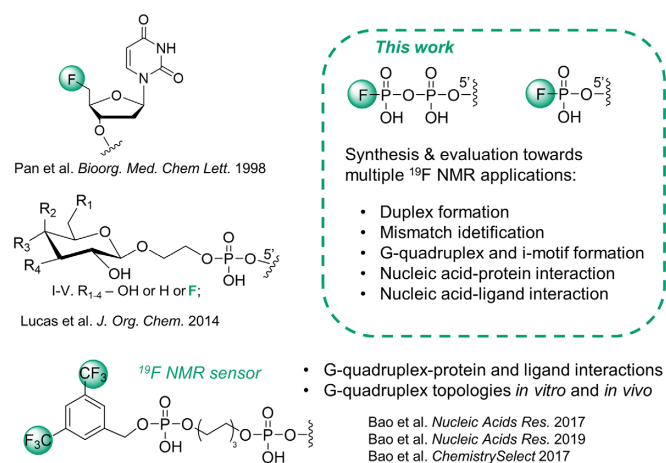


Figure 1. Structures of 5' end fluorine-labeled oligonucleotides with one example used in ^{19}F NMR studies.

polymerase-catalyzed enzymatic synthesis in the presence of fluorinated nucleoside triphosphates (19). In the chemical approach, a fluorine atom or a fluorinated substituent is introduced in the ribose or nucleobase moiety of an appropriate phosphoramidite building block through multiple synthetic steps. The fluorinated phosphoramidite is subsequently used in solid-phase synthesis of DNA or RNA. Various modified nucleotides have been synthesized following this methodology (20–24). Virta *et al.* used the click chemistry approach to incorporate the trifluoromethyl (CF_3) group at the 4' position of thymidine- and at the 2'-*O* position of cytidine-derived building blocks to study DNA duplex/triplex formation and DNA/RNA secondary structure stabilization by neomycin (25,26). Micura *et al.* reported a 2'-deoxy-2'-trifluoromethylthio (2'-SCF₃) uridine building block for RNA solid-phase synthesis and prepared 2'-SCF₃-labeled RNA probes to study RNA–protein and RNA–ligand interactions, conformational transitions and to monitor RNA folding in *Escherichia coli* cell lysates at low probe concentrations. This was achieved despite the fact that 2'-ribose labeling destabilized the RNA double helix by 15°C (1.9 kcal/mol) (27). The same team also reported the introduction of 2'-SCF₃ into pyrimidine nucleosides to study RNA stability and conformational equilibria of hairpins by ^{19}F NMR (28). Enzymatic synthesis of fluorinated DNA probes was reported by Hocek *et al.*, who incorporated trifluoroacetophenone-linked nucleoside triphosphates (uridine, cytidine, 7-deazaadenosine, and 7-deazaguanosine) into DNA to monitor DNA duplex formation and nucleic acid–protein interaction by ^{19}F NMR (29).

The incorporation of fluorine-containing tags at the 5' end of oligonucleotides (30,31) increases synthetic accessibility owing to the possibility of using post-synthetic functionalization (32,33). However, only a few examples of combining 5'-terminal fluorine-labeling and ^{19}F NMR applications have been reported (Figure 1) (34–37). Xu *et al.* introduced 3,5-bis(trifluoromethyl)phenyl phosphoramidite at the 5' end of oligonucleotides during the final cycle of solid-phase synthesis to functionalize the 5'

end of both RNA and DNA G-quadruplex-forming sequences (Figure 1). The resulting probes were used to study the telomere RNA sequence, detect different G-quadruplex assemblies *in vitro* and in living cells (single-stranded, dimeric and high-order structures) and monitor G-quadruplex–ligand interactions (34,35). More recently, the same 3,5-bis(trifluoromethyl)phenyl moiety was introduced at the 5' end of the telomere DNA sequence and was applied to study different G-quadruplex topologies in HeLa cells (37).

Despite the above examples of synthesis and ^{19}F NMR application of fluorine-labeled oligonucleotide probes, to the best of our knowledge, a universal labeling strategy compatible with multiple applications has not been proposed. We hypothesized that the most effective way to develop such a versatile and synthetically accessible labeling strategy that ensures minimal perturbation of the studied phenomena, was to rely on the 5'-terminal functionalization of oligonucleotides with a structurally compact fluorinated label. One such candidate is the fluorophosphate moiety, which we previously used to develop small nucleotide-derived ^{19}F NMR probes to monitor multiple enzymatic activities and protein binding events (36). In a proof of concept experiment, we also demonstrated that short DNA duplex formation could be monitored by ^{19}F NMR. Unlike other halogenophosphates or fluorophosphate diesters, fluorophosphate monoesters and phosphoanhydrides are chemically stable in aqueous solutions under a wide pH range. Moreover, the δ_{F} chemical shifts of fluorophosphates are insensitive to changes in pH or presence of divalent metal ions, but sensitive to intramolecular chemical transformations, which makes them good candidates for ^{19}F NMR probes (36). Hence, in the present study, we aimed to systematically investigate the properties, application scope and limitations of oligonucleotides carrying 5'-fluoromonophosphate (FP) or 5'-fluorodiphosphate (FPP) moieties as ^{19}F NMR probes. We synthesized several FP- and FPP-functionalized oligodeoxyribonucleotides using straightforward methods that avoided commercially unavailable phosphoramidite building blocks. The ^{19}F -labeled oligonucleotides were evaluated as molecular tools to: (i) probe duplex formation and distinguish 5'-terminal mismatches; (ii) monitor G-quadruplex formation; (iii) detect DNA–protein interactions; and (iv) monitor *i*-motif formation and its interactions with small-molecule ligands. We found that the FP and FPP moieties acts as universal ^{19}F reporter tags enabling monitoring of a variety of nucleic acids-related phenomena.

MATERIALS AND METHODS

Oligonucleotide synthesis

Deoxyribonucleotides (ON1, ON3-ON5, ON9, ON11, ON20-ON25, hTeloC and PON11) were synthesized at a 2–5 μmol scale using an automated DNA/RNA synthesizer (ÄKTA oligopilot 10 plus, GE Healthcare) equipped with a column filled with either PrimerSupport™ 200 T 80s (82 $\mu\text{mol/g}$), dA 80s (79 $\mu\text{mol/g}$), dG 80s (78 $\mu\text{mol/g}$) and dC 80s (84 $\mu\text{mol/g}$), or PrimerSupport™ 5G dC 350 (360 $\mu\text{mol/g}$). In the coupling step, 10 equivalents of 0.1 M phos-

phoramidite (DMT-dG(dmf), DMT-dC(N-acetyl), DMT-dA(N-bz) and DMT-T purchased from ChemGenes) in acetonitrile and 0.3 M BTT in acetonitrile were recirculated through the column for 5 min. If required, the oligodeoxyribonucleotides were 5'-phosphorylated on solid-phase, using 0.1 M bis(2-cyanoethyl)-*N,N*-diisopropylphosphoramidite (Sigma-Aldrich) in acetonitrile and 0.3 M BTT. A solution of 3% (v/v) dichloroacetic acid in toluene was used as a detritylation reagent and 0.05 M iodine in pyridine was used for oxidation, 20% (v/v) *N*-methylimidazole in acetonitrile served as Cap A, while a mixture of 40% (v/v) acetic anhydride and 40% (v/v) pyridine in acetonitrile served as Cap B. After the last cycle of synthesis, DNAs still attached to the solid support were treated with 20% (v/v) diethylamine in acetonitrile to remove 2-cyanoethyl protecting groups. Finally, the solid support was washed with acetonitrile and dried with argon, followed by cleavage and deprotection using a 1:1 (v/v) mixture of 30% ammonia and 40% aqueous methylamine at 40°C for 2 h. Subsequently, the solution was filtered and the support was washed with water several times. Combined filtrates were subjected to evaporation and lyophilization (Supplementary Table S1).

Synthesis of FP analogs (FPON1, FPON3 and FP-hTeloC) proceeded as follows: the 5'-OH group of the oligonucleotide still attached to the solid-phase support (5–20 μ mol scale) was reacted with a solution of diphenyl *H*-phosphonate (0.1–0.2 M) in anhydrous pyridine (4 ml) for 2 h; then, it was washed with acetonitrile and finally passed with a solution of 0.1 M triethylammonium bicarbonate buffer (pH 8.0) to obtain stable 5'-*H*-phosphonate. In the next step, the dry solid support was washed with a solution of *N,O*-bis(trimethylsilyl)acetamide (0.4 ml, 1.64 mmol), imidazole (150 mg, 2 mmol), CCl₄ (0.75 ml), acetonitrile (0.75 ml) and triethylamine (0.1 ml) for 4 h, and subsequently washed with acetonitrile, and dried under argon to obtain an imidazole-activated 5'-phosphate group (38). Finally, the imidazolide was reacted for 1–2 h with a 1.0 M solution of tetrabutylammonium fluoride (TBAF) in tetrahydrofuran (THF) (0.2 ml, 0.2 mmol) and diluted with a dimethylsulfoxide (DMSO; 3 ml) solution of ZnCl₂ (10 mg, 0.07 mmol). Next, the solid support was washed with ethylenediaminetetraacetic acid (EDTA) (10 mg/ml)/NaHCO₃ (5 mg/ml) aqueous solution followed by washing with acetonitrile, dried under argon and cleaved from the solid support as described above. After cleavage of FPON1 and FP-hTeloC, an additional step involving nucleophilic substitution by TBAF was performed by dissolving the oligonucleotide in a solution of TBAF in THF and diluting it with ZnCl₂ in DMSO as above. After overnight stirring, the EDTA aqueous solution (see above) was added to the reaction mixture.

FPP analogues were synthesized by reacting a lyophilized 5'-phosphate oligonucleotide with fluorophosphate imidazolide lithium salt, which was synthesized as described previously (36). The 5'-phosphorylated oligonucleotide was moisturized with water (5–10 μ l), followed by addition of DMSO (0.2 ml), fluorophosphate imidazolide lithium salt (20–30 equivalents), and MgCl₂ (20–30 equivalents). The solution was stirred for 6–24 h. Reaction progress was monitored by reversed-phase high-performance liquid chromatography (RP-HPLC) and mass spectrometry with

electrospray ionization. After adequate time, the reaction was diluted with an aqueous mixture of EDTA (10 mg/ml).

Unmodified, FP and FPP oligodeoxyribonucleotides were purified by semi-preparative RP-HPLC using a Discovery RP Amide C-16 HPLC column (25 \times 2.12 cm, 5 μ m, flow rate 5.0 ml/min) or Vision-HT RP C-18 column (25 \times 2 cm, 10 μ m, flow rate 5.0 ml/min) with detection at 254 nm using a linear gradient of acetonitrile in 0.05 M ammonium acetate buffer (pH 5.9). The identity of all products was confirmed by high-resolution mass spectrometry (HRMS), whereas their homogeneity and purity were determined by RP-HPLC and ¹⁹F NMR spectroscopy.

Oligonucleotide sample preparation for NMR and UV-VIS studies

Stock solutions were prepared by dissolving oligonucleotides in 100–200 μ l of ultrapure water at ~2–10 mM. The exact concentrations of oligonucleotides (FPON1, FPON3, FPPON3, FP-hTeloC, ON1, ON3-ON5, ON9, ON11, ON20-ON25, hTeloC and PON11) were determined spectrophotometrically at 260 nm using theoretical extinction coefficients in phosphate buffer (pH 7.0). The appropriate oligonucleotide was then diluted in a selected buffer based on the end application: duplex formation studies (20 mM sodium phosphate, 0.15 M NaCl, 0.2 mM disodium EDTA, pH 7.0, 10% D₂O); *i*-motif studies (50 mM sodium citrate, pH 4.2, 10% D₂O); G-quadruplex studies (10 mM KH₂PO₄, 70 mM KCl, 0.2 mM disodium EDTA, pH 7.0, 10% D₂O).

¹H NMR spectroscopy studies

¹H NMR spectra were recorded on a Bruker Avance III HD 500 MHz spectrometer equipped with a 5-mm PABBO BB/19F-1H/D Z-GRD probe at a frequency of 500.24 MHz. If needed, water suppression was achieved using a 3-9-19 pulse sequence with gradients (WATERGATE) (39,40). The ¹H NMR chemical shifts were compared to sodium 3-(trimethylsilyl)-2,2',3,3'-tetradeuteriopropionate (TSP) (δ_{H} 0 ppm) as an internal standard.

¹⁹F NMR spectroscopy studies

¹⁹F NMR spectra were recorded as above at a frequency of 470.67 MHz in 5-mm NMR samples. Typical experimental parameters were as follows: ¹⁹F excitation pulse, 15.0 μ s; acquisition time, 1.19 s; prescan delay, 6.0 μ s; relaxation delay, 1.0 s; usual number of scans, 512–4096 or up to 8000 in the case of G-quadruplex studies. The ¹⁹F NMR chemical shifts were compared to either 10 mM NaF in D₂O (δ_{F} = –121.5 ppm) or 0.5 mM 2,2,2-trifluoroethanol (δ_{F} = –76.0 ppm) used as external standards.

Duplex formation and mismatch identification. Fluorophosphorylated oligonucleotides FPON3 and FPPON3 in 10 mM aqueous stock solutions were diluted in phosphate buffer (20 mM sodium phosphate, 0.15 M NaCl, 0.2 mM EDTA, pH 7.0, 10% D₂O) to a final concentration of 0.1 and 0.2 mM, respectively. Titrations were performed by

adding aliquots of complementary oligonucleotide from the stock solutions. Before acquiring each NMR spectrum, the sample was incubated at 90°C for 3 min and then cooled to room temperature on ice. NMR measurements were carried out at 23°C.

***i*-Motif studies.** Fluorophosphorylated oligonucleotides FPON1, FPPON1 and FP-hTeloC were prepared as 5–8 mM stock solutions in ultra-pure water and diluted to a final concentration of 0.1–0.5 mM in citrate buffer (50 mM sodium citrate, pH 4.2, 10% D₂O) or cacodylate buffer (10 mM sodium cacodylate, 100 mM NaCl, pH 5.5, 10% D₂O). If needed, pH was determined in a 1.5-ml tube by a calibrated pH-meter equipped with a glass electrode specifically adapted for measurements in a small volume. The pH was adjusted by dropwise additions of 0.1 M HCl or 0.1 M NaOH/KOH. In the case of temperature-dependent studies, the samples were incubated inside the magnet for 10 min before starting data acquisition. For pH-dependent studies the temperature was set to 25°C. Mitoxantrone dihydrochloride and cationic porphyrin TMPyP4 tetratosylate were used as titrants in the studies with FP-hTeloC and FPON1, respectively; they were purchased from Sigma-Aldrich as lyophilized powders and dissolved in water at a concentration of 10 mM. Adenosine 5'-fluoromonophosphate (AMPF) was used as negative control in interaction studies between *i*-motifs and ligands; it was synthesized as previously described (36) and dissolved in ultra-pure water at a concentration of 10 mM. To calculate the dissociation constant (K_d), a one-site-specific binding model was fitted to the data using GraphPad Prism 8.3 software (Equation 1):

$$\Delta\delta_{ppm} = \frac{\Delta\delta_{ppm}^{\max} * c_M}{K_d + c_M} \quad (1)$$

where $\Delta\delta_{ppm}$ represents the observed chemical shift change relative to free oligonucleotide, c_M is the concentration of mitoxantrone (in μ M), $\Delta\delta_{ppm}^{\max}$ is the chemical shift change value extrapolated to saturating concentrations of mitoxantrone and K_d is the equilibrium dissociation constant.

***G*-quadruplex formation and protein interaction studies.** The fluorophosphorylated oligonucleotide FPPON11 in 2.50 mM aqueous stock solution was diluted to a final concentration of 70 μ M in phosphate buffer (10 mM KH₂PO₄, 70 mM KCl, 0.2 mM disodium EDTA, pH 7.0, 10% D₂O). Prior to NMR measurements, the sample was heated to 90°C and cooled on the bench to room temperature (21–23°C). Thrombin from bovine plasma served as a titrant in nucleic acid–protein interaction studies; it was purchased from Sigma-Aldrich (T4648-1KU) as a lyophilized powder and dissolved in water to a concentration of 10 mg/ml (137 μ M).

UV-VIS measurements

For duplex formation studies, oligonucleotides FPON3, FPPON3, ON3-ON5, ON9 and ON20-ON25 were diluted with sodium phosphate buffer to a final concentration of 2 μ M (20 mM sodium phosphate, 0.15 M NaCl, 0.2 mM

EDTA, pH 7.0, 10% D₂O) and mixed together in a 1:1 ratio (to a final concentration of 2 μ M each). For G-quadruplex formation studies, oligonucleotides FPPON11, PON11 and ON11 were diluted in potassium phosphate buffer to a final concentration of 4 μ M (10 mM KH₂PO₄, 70 mM KCl, 0.2 mM disodium EDTA, pH 7.0, 10% D₂O). For *i*-motif formation studies, oligonucleotides ON1, FPON1, FPPON1 and hTeloC were diluted in sodium citrate buffer to a final concentration of 12 μ M (50 mM sodium citrate, 10% D₂O, pH 4.2 or 5.5). The measurements were performed in 1 × 1 cm cuvettes with a sample volume of 2 ml. Melting temperature determination was performed on a Thermo Scientific Evolution 350 UV-VIS spectrometer, equipped with an 8 Cell Peltier Thermostatted System and Temperature Probe, at a temperature ranging from 10°C to 90°C (ramp rate 1.0°C/min). Duplex formation was monitored at 260 nm, whereas G-quadruplex and *i*-motif formation were monitored at 295 and 260 nm. All experiments were performed in triplicate.

Melting temperature and transition pH determination

¹⁹F NMR and UV-VIS data were analyzed by normalizing the integrated ¹⁹F NMR signal or absorbance in GraphPad Prism 8.3 software: 0% was defined as the smallest mean, whereas 100% was defined as the largest mean value in each dataset. To determine the melting temperatures (T_m) or transitional pH (pH_T) the following equation was fitted:

$$A \text{ or } I = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{1 + e^{\frac{V_{50} - X}{\text{slope}}}} \quad (2)$$

where normalized absorbance (A) or normalized ¹⁹F NMR signal integration (I) vary from BOTTOM to TOP, and V_{50} is the value of T_m or pH_T .

RESULTS

Synthesis of 5'-fluorophosphate oligonucleotides

To explore the use of FPPONs and FPONs as molecular probes for ¹⁹F NMR studies, we synthesized a set of 18 oligonucleotide probes containing these moieties (Table 1). To obtain FPPONs, a previously developed approach based on an imidazole-activated fluorophosphate subunit (FpIm) as an electrophilic fluorophosphorylating agent was adapted (Figure 2A) (36). FPONs were generated by reacting imidazole-activated oligonucleotide 5'-monophosphates with TBAF (Figure 2B).

FPPONs were synthesized in solution starting from oligonucleotide 5'-monophosphates prepared by conventional phosphoramidite solid-phase synthesis (41). After cleavage and deprotection, the oligonucleotide 5'-monophosphate methylamine salt was converted into 5'-fluorodiphosphate in a one-step reaction, with an excess of activated FpIm (30 equivalents) and MgCl₂ (30 equivalents) as a reaction mediator (Figure 2A). Using this approach, 15 oligonucleotides of varying length (from 6 to 25 nt) (Table 1) were synthesized. The set included ten oligonucleotides of various lengths designed for duplex formation studies (FPPON2–FPPON10 and FPPON14), a 15-nt

Table 1. Summary of fluoro(di)phosphorylated oligonucleotide probes

Probe name	Length (nt)	5'-terminal group ^[a]	Sequence (5'→3')	HPLC conversion	Isolated yield ^[b]	Molecular formula of [M-nH] ⁿ⁻ ion	Calc. (m/z)	Actual (m/z)
FPON1 ^[c]	6	Fp	TCC CCC	80%	20%	C ₅₅ H ₇₂ FN ₁₇ O ₃₇ P ₆ ²⁻	883.63478	883.63531
FPPON1 ^[c]	6	Fpp	TCC CCC	61%	35%	C ₅₅ H ₇₃ FN ₁₇ O ₄₀ P ₇ ²⁻	923.61795	923.61885
FPPON2	7	Fpp	GTC AAT G	86%	45%	C ₆₉ H ₈₆ FN ₂₇ O ₄₅ P ₈ ²⁻	1139.65835	1139.65965
FPON3 ^[c]	10	Fp	AGA CAT TGA C	60%	15%	C ₉₈ H ₁₂₀ FN ₄₀ O ₃₈ P ₁₀ ³⁻	1037.83489	1037.83648
FPPON3 ^[c]	10	Fpp	AGA CAT TGA C	86%	30%	C ₉₈ H ₁₂₁ FN ₄₀ O ₆₁ P ₁₁ ³⁻	1064.49033	1064.49194
FPPON4	10	Fpp	TGA CAT TGA C	93%	49%	C ₉₈ H ₁₂₂ FN ₃₇ O ₆₃ P ₁₁ ³⁻	1061.48648	1061.48788
FPPON5	10	Fpp	GTC AAT GTC C	82%	41%	C ₉₇ H ₁₂₂ FN ₃₅ O ₆₄ P ₁₁ ³⁻	1053.48273	1053.48426
FPPON6	10	Fpp	GGA CAT TGA C	99%*	48%	C ₉₈ H ₁₂₁ FN ₄₀ O ₆₂ P ₁₁ ³⁻	1069.82197	1069.82293
FPPON7	10	Fpp	ATC AAT GTC G	85%	51%	C ₉₈ H ₁₂₂ FN ₃₇ O ₆₃ P ₁₁ ³⁻	1061.48648	1061.48737
FPPON8	10	Fpp	CGA CAT TGA T	99%*	58%	C ₉₈ H ₁₂₂ FN ₃₇ O ₆₃ P ₁₁ ³⁻	1061.48648	1061.48751
FPPON9	10	Fpp	GTC AAT GTC T	88%	48%	C ₉₈ H ₁₂₃ FN ₃₄ O ₆₅ P ₁₁ ³⁻	1058.48262	1058.48364
FPPON10	13	Fpp	GTC AAT GTC AGC G	80%	41%	C ₁₂₇ H ₁₅₈ FN ₅₀ O ₈₁ P ₁₄ ³⁻	1377.20362	1377.20421
FPPON11 ^[d]	15	Fpp	GGT TGG TGT GGT TGG	82%	42%	C ₁₅₀ H ₁₈₄ FN ₅₇ O ₉₉ P ₁₆ ⁴⁻	1220.42309	1220.42932
FPPON12	16	Fpp	GGA TAC TTT TGT ATC C	78%	35%	C ₁₅₇ H ₁₉₄ FN ₆₂ O ₉₈ P ₁₇ ⁴⁻	1254.43380	1254.43400
FPPON13	16	Fpp	GGA TAT TTT TAT ATC C	97%*	19%	C ₁₅₈ H ₁₉₈ FN ₅₂ O ₁₀₃ P ₁₇ ⁴⁻	1254.18499	1254.18523
FPPON14	16	Fpp	GTC AAT GTC AGC GAT A	85%	30%	C ₁₅₇ H ₁₉₅ FN ₆₂ O ₉₈ P ₁₇ ³⁻	1687.25737	1687.25876
FPPON15	25	Fpp	CCT GGG GGA GTA TTG CGG AGG AAG G	95%*	20%	C ₂₄₇ H ₃₀₁ FN ₁₀₇ O ₁₅₂ P ₂₆ ⁵⁻	1604.24604	1604.24765
FP-hTeloC ^[b]	24	Fp	TAA CCC TAA CCC TAA CCC TAA CCC	76%	24%	C ₂₂₈ H ₂₈₉ FN ₈₄ O ₁₄₀ P ₂₄ ⁴⁻	1801.30192	1801.29517

^[a] Fpp, fluorodiphosphate; Fp, fluorophosphate; ^[b] For FP overall yield including solid phase synthesis is reported; For FPP the yield for the final step (reaction with FpIm) is reported; ^[c] Oligonucleotide used as ¹⁹F molecular probe in monitoring of *i*-motif structure changes; ^[d] Oligonucleotide used as ¹⁹F molecular probe in monitoring of G-quadruplex structure changes and nucleic acid-protein interactions; *approximate value due to HPLC signal overlap.

Oligonucleotides are listed by length.

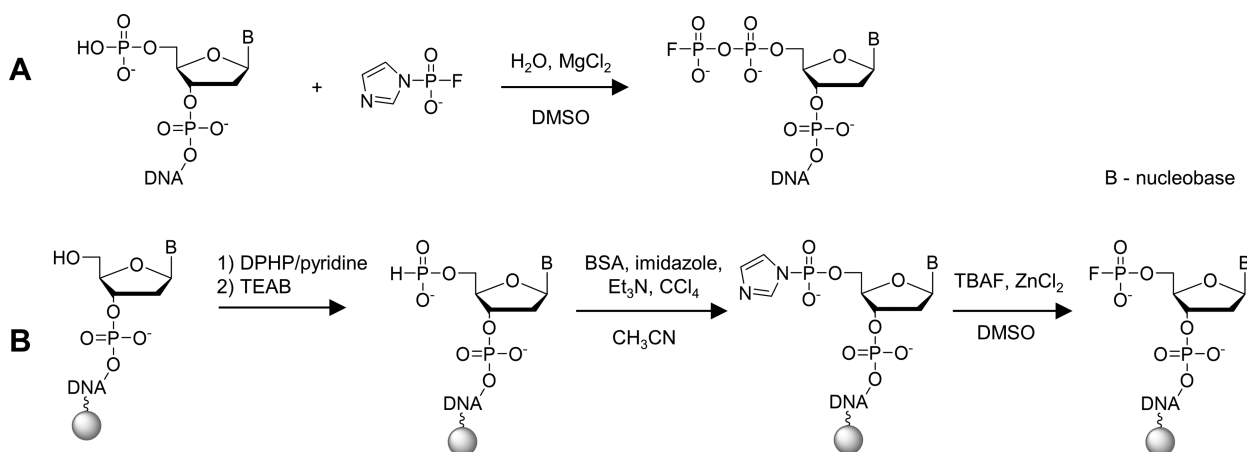


Figure 2. Synthesis of 5'-fluoro(mono)- and 5'-fluoro(di)phosphate-labeled oligonucleotides (FPONs and FPPONs). (A) FPPONs obtained by reacting solid-phase synthesized oligonucleotide 5'-phosphates with an activated fluorophosphate subunit (FpIm). (B) FPONs obtained by activation on a solid support and reaction with TBAF. DPHP: Diphenyl *H*-phosphonate; BSA: *N,O*-Bis(trimethylsilyl)acetamide.

G-quadruplex-forming sequence (FPPON11), an *i*-motif-forming sequence (FPPON1) and three longer sequences of 16 nt (FPPON12 and FPPON13) or 25 nt (FPPON15). The isolated yields for fluorophosphorylation varied from 19 to 58%, but usually reached 30%. The final products were purified by RP-HPLC.

To obtain FPONs on solid support, 5'-deprotected oligonucleotides synthesized by the conventional phosphoramidite method were modified while still attached to the support. The 5'-OH group of the oligonucleotide was converted to *H*-phosphonate monoester by treatment with diphenyl *H*-phosphonate in pyridine, followed by hydrolysis with triethylammonium bicarbonate buffer. The resulting 5'-*H*-phosphonate was oxidized to *P*-imidazolide by a reac-

tion with carbon tetrachloride in the presence of imidazole and *N,O*-bis(trimethylsilyl)acetamide (BSA) (38). Finally, a nucleophilic substitution of imidazole with fluoride (using TBAF as a source) was carried out in the presence of excess ZnCl₂, followed by cleavage from the support with a mixture of 30% ammonia and 40% aqueous methylamine (AMA) to yield the desired FPON (Figure 2B) (36). Notably, no side-reactions associated with the P-F bond cleavage were observed under these conditions. The approach was used to synthesize a molecular probe designed for DNA duplex formation monitoring (FPON3) with a 15% yield, and two cytosine-rich sequences, FPON1 and FP-hTeloC, designed to study *i*-motif structure changes and stabilization, with yields of 20 and 24%, respectively (Table 1). All

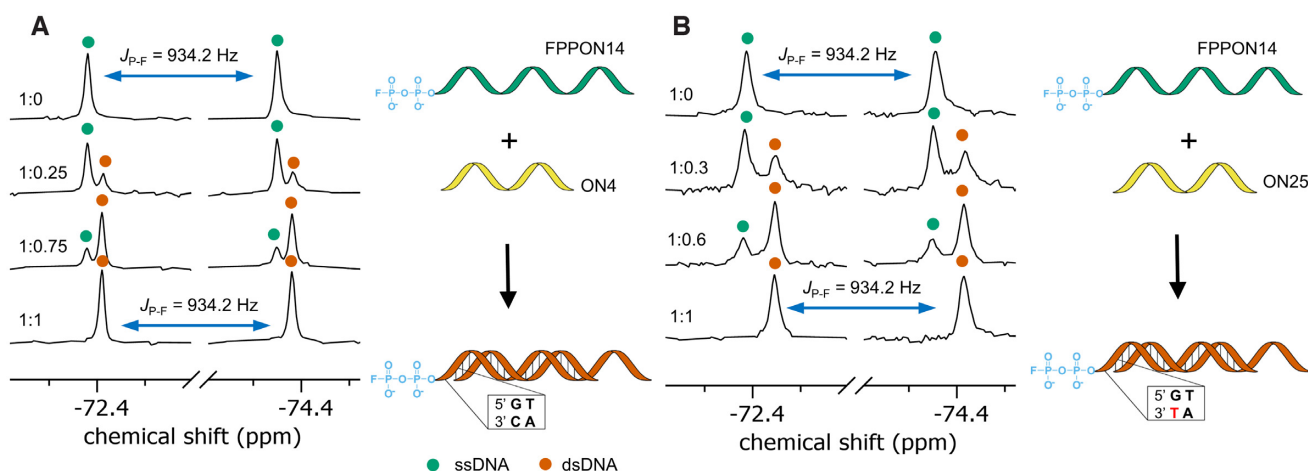


Figure 3. ^{19}F NMR (470.67 MHz) spectra of fluorine-labeled oligonucleotides in single-stranded (ss) and double-stranded (ds) form. (A) ^{19}F NMR duplex formation monitoring between FPPON14 and ON4. (B) ^{19}F NMR duplex formation monitoring between FPPON14 and ON25; Green dots indicate ssDNA, orange dots indicate dsDNA. All spectra were recorded at 23°C in 20 mM sodium phosphate, 0.15 M NaCl, 0.2 mM disodium EDTA, pH 7.0 and 10% D_2O .

synthesized oligonucleotides were purified by RP-HPLC and characterized by HRMS as well as ^{19}F NMR spectroscopy (Supplementary Data).

To verify whether the fluoro(di)phosphate oligonucleotide analogues were suitable for ^{19}F NMR studies, we evaluated them as molecular probes to monitor different phenomena including DNA duplex formation, G-quadruplex formation and interaction with proteins, as well as *i*-motif assembly and interaction with small molecules (Supplementary Table S2).

Duplex formation and mismatch identification

We first explored the suitability of fluorophosphorylated oligonucleotides (FPON3, FPPON3 and FPPON14) to monitor duplex formation and identify 3'-terminal mismatches in a complementary strand by ^{19}F NMR. Oligonucleotides used in these experiments differed in sequence, length (10–16 nt) and 5'-terminal tags (either FP or FPP). For initial studies, we chose the 16-nt FPP oligonucleotide FPPON14. The ^{19}F NMR spectrum of FPPON14 showed a single doublet at $-\delta_{\text{F}} = -73.38$ ppm and $J_{\text{p-F}} = 934.2$ Hz (Figure 3A and B, top spectra). We next verified whether it was possible to monitor duplex formation between FPPON14 and either the fully complementary 10-nt ON4 or the single-mismatched 10-nt ON25 (Supplementary Table S1). The addition of 0.25 equivalents of ON4 to FPPON14 resulted in the appearance of a new signal at $\delta_{\text{F}} = -73.40$ ppm ($J_{\text{p-F}} = 934.2$ Hz), whose intensity increased with further addition of complementary oligonucleotide (up to one equivalent), while the original signal of the ssDNA disappeared completely (Figure 3A). In an analogous experiment, the addition of one equivalent of ON25 to FPPON14 resulted in the appearance of a new signal at $\delta_{\text{F}} = -73.42$ ppm and disappearance of the ssDNA signal (Figure 3B). Duplex formation under the above conditions was independently confirmed using UV-VIS spectroscopy (Supplementary Figure S1). By comparing the changes in chemical shifts between

ssDNA and dsDNA, a chemical shift of 0.02 ppm was detected between duplexes with ON4 and ON25. This observation suggested that our probes could be used for 5'/3' terminal mismatch identification by ^{19}F NMR.

To further explore this possibility, we synthesized a 10-nt model sequence, ON3, carrying either a 5'-monofluorophosphate (FPON3) or 5'-difluorophosphate (FPPON3) and a library of oligonucleotides (ON5, ON9 and ON20–ON25; Supplementary Table S1) that were either fully complementary to the ON3 sequence, or complementary except for a single mismatch at the last (X) or second to last (Y) position from their 3' ends. To assess the influence of FPP and FP moieties on DNA duplex stability, we first studied the stability of these fully or partially complementary duplexes with respect to an identical but unmodified (5'-OH) ON3 sequence (Supplementary Figures S2–4). T_{m} was determined via UV-VIS spectroscopy by fitting a theoretical model to the determined temperature-dependent sigmoidal curves (Table 2) (42). The fully complementary duplexes of FPON3, FPPON3 and ON3 with ON9 showed similar stability, with T_{m} values of $40.8 \pm 0.1^\circ\text{C}$, $41.0 \pm 0.1^\circ\text{C}$ and $41.6 \pm 0.1^\circ\text{C}$, respectively (Table 2, line 1). Duplexes with sequences carrying a single mismatch (ON5, ON20–ON21) at the 3' end (i.e. at the 5' end of the fluorophosphorylated probe) resulted in T_{m} values of 39–41°C, with the 5'-modified duplexes showing lesser stability (0.6–1.5°C lower T_{m}) than the unmodified counterparts (Table 2, lines 2–4). The sequences with single mismatch (ON22–ON24) at the second to last position from the 3' end resulted in T_{m} values of 29–31°C (Table 2, lines 5–7), with the 5'-terminal modifications having an additional slightly destabilizing influence. Overall, these results revealed that the complementarity of the last base pair had little influence on duplex stability; therefore, these oligos are good model sequences to study mismatches and other modifications that do not affect duplex stability significantly.

We subsequently tested whether it was possible to detect duplex formation for these sequences and identify 3'-

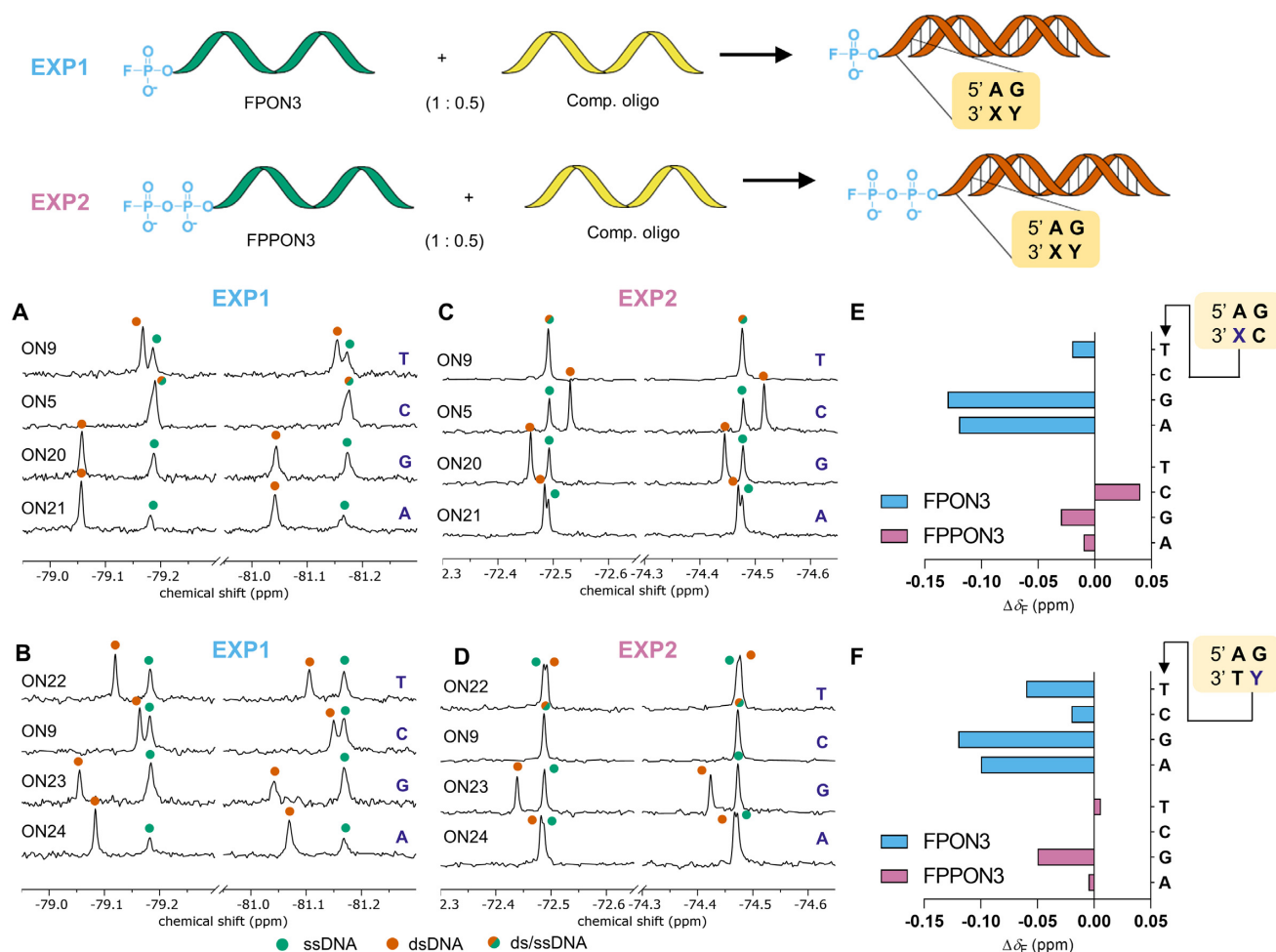


Figure 4. DNA hybridization monitoring and mismatch identification by ^{19}F NMR spectroscopy. FPN3 (EXP1) and FPPON3 (EXP2) sequences were used to monitor dsDNA formation and mismatch identification. Fluorophosphorylated ONs were mixed with 0.5 equivalents of different complementary or partially complementary (single mismatch at X or Y position from the 3' end) oligonucleotides (ON5, ON9, ON20–ON24) and ^{19}F NMR spectra were recorded. Color-coding was used to indicate the resonances for ssDNA (green dots), dsDNA (orange dots) and overlapping dsDNA and ssDNA signals (green-orange dots). Downfield or upfield shifted components of each doublet were analyzed to calculate $\Delta\delta_{\text{F}}$ differences between ssDNA and dsDNA. Measurements were performed in 20 mM sodium phosphate, 0.15 M NaCl, 0.2 mM disodium EDTA, pH 7.0, 10% D_2O buffer at 23°C.

quadruplexes may serve as gene expression modulators and drugs in anti-cancer therapy (48). Hence, developing methods to study these DNA structures is of broad interest.

To verify the usefulness of FP labeling for studying G-quadruplexes by ^{19}F NMR, we introduced the fluorodiphosphate moiety at the 5' end of the thrombin-binding aptamer (TBA) sequence (ON11). TBA is an artificial 15-nt deoxyribooligonucleotide that forms a potassium-stabilized G-quadruplex consisting of two G-tetrads and a TGT loop (49). The aptamer, discovered in 1992 by systematic evolution of ligands by exponential enrichment, interacts with thrombin with nanomolar K_{d} and inhibits fibrin-clot formation (50,51). Using ^1H NMR, we confirmed that FPPON11 formed a G-quadruplex structure at 25°C in the presence of K^+ , similar to the unmodified TBA sequence (52). The imino resonances typical of G-quadruplexes (δ_{H} 11.8–12.5 ppm) were clearly observed at room temperature, but not at 78°C (Supplementary Figure S19). Next, we tested whether temperature-induced unfolding of this

TBA G-quadruplex could be observed by ^{19}F chemical shift changes in temperature-dependent spectra of FPPON11. ^{19}F NMR spectra of FPPON11 at 25°C showed a doublet at -72.43 ppm ($J_{\text{P-F}} = 933$ Hz). A stepwise temperature increase (up to 78°C) resulted in the appearance of a new signal (-72.51 ppm, $J_{\text{P-F}} = 929$ Hz, $\Delta\delta_{\text{F}} = 0.016$ ppm at 40°C), while the original signal decreased in intensity until it disappeared completely at 55°C (Figure 5A). The ^{19}F NMR signals at different temperatures were integrated, normalized, and the relative intensities were plotted as a function of temperature, followed by fitting a theoretical curve to determine T_{m} . The T_{m} value was $37.8 \pm 0.5^\circ\text{C}$ if G-quadruplex signal was used for the analysis and $37.8 \pm 0.7^\circ\text{C}$ if ssDNA signal was analyzed (Figure 5B). The results were independently verified under the same conditions by UV-VIS temperature-dependent measurements. T_{m} values for unmodified TBA (ON11), 5'-phosphate TBA (PON11) and FPPON11 were determined (Figure 5C). The 5'-phosphate (P) group lowered the melting temperature by 7°C ($T_{\text{m}} = 41.1 \pm 0.1^\circ\text{C}$)

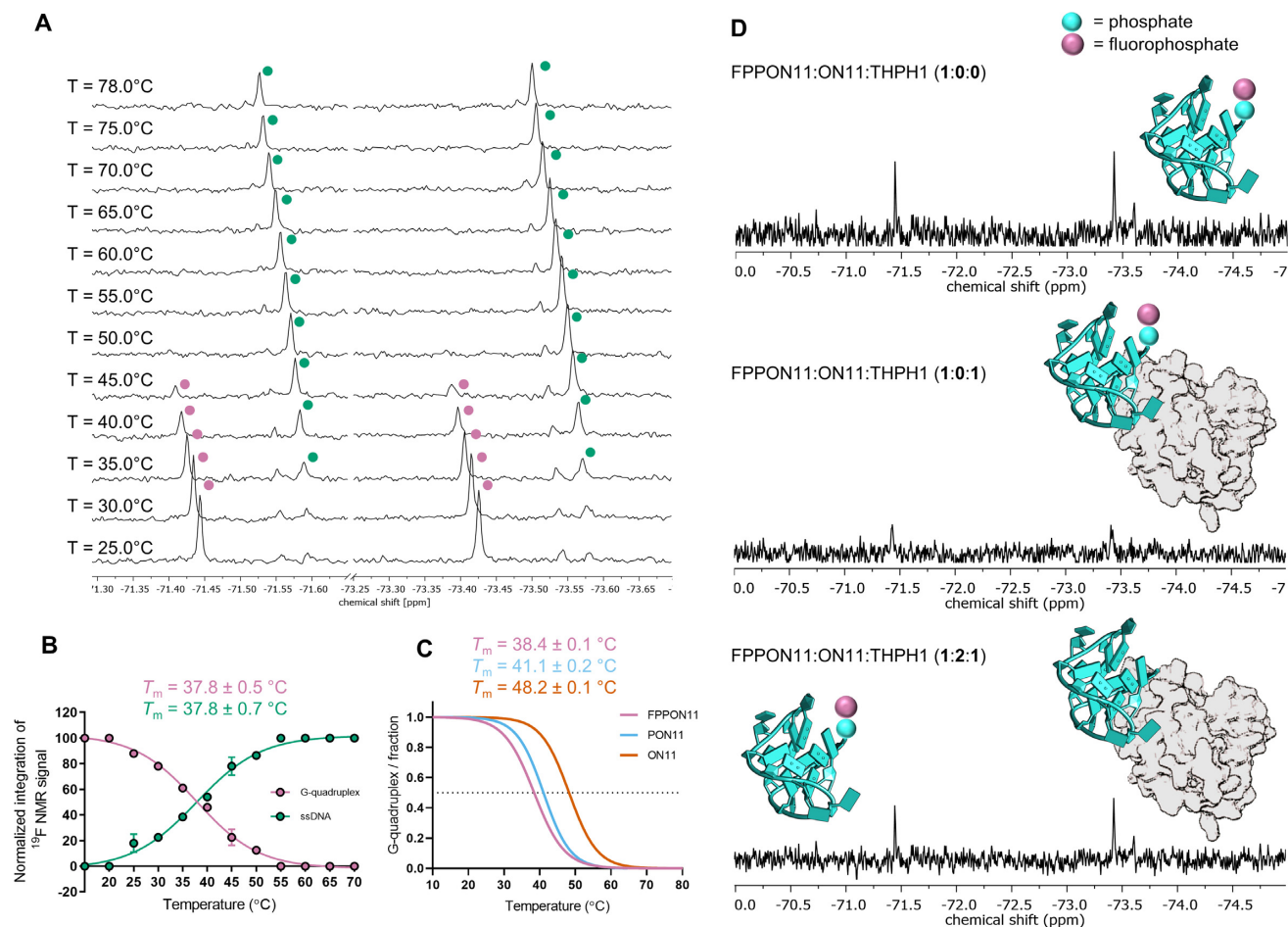


Figure 5. ^{19}F NMR studies on a 5'-fluorodiphosphorylated TBA sequence forming a G-quadruplex structure (FPPON11). (A) ^{19}F NMR spectra of FPPON11 as a function of temperature. (B) Melting temperature determination by analysis of integrated ^{19}F NMR signals. (C) UV/VIS melting profile of FPPON11, ON11 and PON11 ($A_{295\text{nm}}$). (D) Protein–nucleic acid interaction and ligand displacement experiment. Top: ^{19}F NMR spectrum of FPPON11; middle: ^{19}F NMR spectrum of FPPON11 in complex with thrombin; bottom: ^{19}F NMR spectrum recorded after addition of two equivalents of TBA sequence (ON11). All signals were referenced to $\text{CF}_3\text{CH}_2\text{OH}$ (-76.00 ppm). The protein structure was adapted from PDB entry: 4DII (73), the scale has not been preserved.

Table 3. Comparison of melting temperatures of thrombin-binding aptamers containing different chemical groups at the 5' end

Oligo.	5'-terminal group	5' GGT TGG TGT GGT TGG 3'	
		T_m [°C]	method
ON11	OH	48.2 ± 0.1	UV/VIS
PON11	P	41.1 ± 0.2	UV/VIS
FPPON11	FPP	38.4 ± 0.1	UV/VIS
		37.8 ± 0.5^a	^{19}F NMR
		37.8 ± 0.7^b	^{19}F NMR

^abased on G-quadruplex ^{19}F NMR signal, ^b based on ssDNA ^{19}F NMR signal.

compared to unmodified ON11 ($T_m = 48.2 \pm 0.1^\circ\text{C}$). Further destabilization by 2.7°C was observed for FPPON11, resulting in a T_m of $38.4 \pm 0.1^\circ\text{C}$ (Table 3).

After confirming that the fluorophosphate moiety enabled monitoring of G-quadruplex formation by ^{19}F NMR, we next tested whether a similar approach could be used

to monitor the TBA–thrombin interaction. To this end, thrombin (THPH1) was titrated into a solution of FPPON11 and the corresponding ^{19}F NMR spectrum was recorded after each addition (Figure 5D). A significant widening of the ^{19}F NMR signal was observed following THPH1 addition, in line with a significantly increased molecular weight and consequently slower tumbling of oligo in the complex with THPH1 (53). The signal was restored after addition of a competitive TBA binder (ON11), in accordance with displacement of the fluorinated probe from the complex (Figure 5D, bottom and Supplementary Figure S20).

i-Motif formation monitoring

i-Motifs are another class of non-canonical secondary nucleic acid structures, which are formed by cytosine-rich sequences. *i*-Motifs were initially found to be stable only under acidic conditions and at low temperatures; hence, the potential biological significance of these structures has not been fully appreciated. In recent years, the initial views on *i*-motifs have been challenged as it has been revealed that

these structures can be stable even at neutral pH (54,55) and are additionally stabilized under molecular crowding conditions (56). Finally, *i*-motif structures have been detected in the nuclei of human cells, suggesting that they may play regulatory roles in the genome, which has sparked renewed interest (57). Interactions between ligands and *i*-motifs have been investigated with the aim of identifying molecules capable of modulating *i*-motif stability (58–60).

Here, we aimed to verify whether fluorophosphate-labeled oligonucleotides could be used to monitor pH- and temperature-dependent *i*-motif (un) folding. To this end, we first generated a TCCCC sequence labeled with FP (FPON1) or FPP (FPPON1) moieties and their unmodified counterpart (ON1). This short DNA sequence consisting of a single thymine and five cytidines forms a tetrameric *i*-motif and is a simple, well-known model molecule for studying *i*-motif stability, as well as the effects of different chemical modifications on *i*-motif structure (61). Using ^1H NMR, we confirmed that all three variants of the ON1 sequence formed *i*-motifs at 10°C in an aqueous buffer at pH 4.2 (10% D_2O). As expected, the presence of imino signals (δ_{H} 15–16 ppm) was clearly visible in the ^1H NMR spectra, which was consistent with the presence of protonated C-CH⁺ base pairs characteristic of the *i*-motif structure (Supplementary Figure S21) (62). Increasing the temperature above 50°C resulted in the disappearance of imino resonances. ^{19}F NMR spectra of FPON1 and FPPON1 recorded under conditions expected to favor *i*-motif formation (pH 4.2, 25°C) revealed presence of a doublet for each form (δ_{F} –79.86 ppm and –73.06 ppm, respectively; Supplementary Figure S22). Both FPON1 and FPPON1 were used in a series of temperature- and pH-dependent ^{19}F NMR experiments, in which we attempted to monitor stepwise *i*-motif unfolding (Figure 6A). The temperature-dependent ^{19}F NMR spectra of FPPON1 and FPON1 were recorded in the range of 25 to 70°C (step of 5°C) at pH 4.20 (Supplementary Figures S23 and 24). For FPPON1 at 25°C, a single doublet (δ_{F} –73.06 ppm, $J_{\text{P-F}}$ = 934.7 Hz) was observed, which disappeared with increasing temperature along with the appearance of a new doublet (δ_{F} –72.24 ppm at 70°C, $J_{\text{P-F}}$ = 929.8 Hz) (Figure 6B). At intermediate temperatures (40–60°C), varying proportions of both signals were clearly observed. Similar changes in ^{19}F NMR spectra were observed for FPON1. A signal visible at 25°C (δ_{F} –79.86 ppm, $J_{\text{P-F}}$ = 936.4 Hz) disappeared at 70°C, while a new doublet appeared (δ_{F} –78.68 ppm, $J_{\text{P-F}}$ = 930.0 Hz) (Figure 6C). The influence of pH changes (in the range of ~3.5–7) on ^{19}F NMR spectra at a constant temperature (25°C) was also studied. At low pH (3.56), one doublet was observed for FPPON1 (δ_{F} –73.07 ppm, $J_{\text{P-F}}$ = 934.6 Hz); whereas increasing the pH resulted in the appearance of a new doublet shifted upfield (δ_{F} = –73.15 ppm, $J_{\text{P-F}}$ = 933.9 Hz, $\Delta\delta_{\text{F}}$ = 0.08 ppm, pH 6.72; Figure 6D). Similarly, for FPON1, one major signal was observed at pH 3.86 (δ_{F} –79.86 ppm, $J_{\text{P-F}}$ = 935.9 Hz), but it disappeared with increasing pH and a new downfield-shifted doublet emerged (δ_{F} –79.77 ppm, $J_{\text{P-F}}$ = 934.2 Hz, $\Delta\delta_{\text{F}}$ –0.09 ppm, pH 6.65; Figure 6E). In both pH- and temperature-dependent experiments, the appearance of new spectral signals suggested unfolding of the *i*-motif structure and appearance of the single-stranded form. Thus, the intensity of normalized ^{19}F NMR signals

from each experimental set was plotted as a function of temperature or pH to determine T_{m} or pH_{T} , respectively (Figure 6). The T_{m} and pH_{T} values were calculated by fitting a theoretical curve to the data (Table 4). The T_{m} values were also independently determined by UV/VIS measurements at 260 nm (pH 4.2), and were usually in good agreement with those obtained by ^{19}F NMR (Supplementary Figure S25 and Table 4). Moreover, the determined values were in good agreement with published data, further confirming that the observed changes in ^{19}F NMR spectra were related to *i*-motif (un) folding (Supplementary Table S3).

Interestingly, while studying FPON1 and FPPON1 by ^{19}F NMR, we noted that the spectra recorded at acidic pH contained a varying number of signals, depending on the sample incubation time. Freshly prepared solutions of oligonucleotides usually exhibited more than one signal in the ^{19}F NMR spectra (Supplementary Figure S26A); whereas after 48 h of incubation at room temperature (~20–21°C) only one clearly visible signal was observed (Supplementary Figure S26B). A single signal was observed at high temperatures, confirming the presence of only one fluorine-tagged sequence in solution (Supplementary Figure S26C). Spectra recorded 15 min after lowering the temperature from 70 to 25°C showed at least four clearly visible signals, each with a different intensity. ^{19}F NMR spectra produced better signal separation for FPON1 than FPPON1, but in both cases peaks with different chemical shifts could be observed (Supplementary Figure S26D). We believe this observation indicates the presence of different *i*-motif topologies in solution. Such phenomenon was previously observed for a short C₄T sequence, whereby ^1H NMR studies revealed an equilibrium between so-called S- and R-forms (63). A more recent study utilized a longer (21-nt) C-rich sequence, which formed major (5'E) and minor (3'E) *i*-motif conformations following a change in pH from 9 to 6 (64). Therefore, we assumed that the initially visible ^{19}F NMR signals corresponded to several kinetically favored *i*-motif conformations, which rearranged to the most thermodynamically stable single form over time.

We next utilized FPON1 as a molecular probe to study *i*-motif interaction with a cationic porphyrin, TMPyP4, by ^{19}F NMR (65). When TMPyP4 was titrated into the preformed FPON1 *i*-motif at 0.1–1 equivalents (0–125 μM), we observed signal broadening (Supplementary Figure S27) with simultaneous downfield shifting of the corresponding resonance (Figure 7). To verify that the spectral changes were a result of *i*-motif interaction with TMPyP4, AMPF (100 μM) was used as an internal reference as it was not expected to interact with the ligand (Figure 7A). Addition of a second equivalent of TMPyP4 led to narrowing of the FPON1 ^{19}F NMR signal, which was likely caused by saturation of the complex in the presence of excess ligand. ^{19}F NMR data were analyzed by plotting chemical shift changes as a function of concentration of the ligand to determine the K_{d} value (0.18 ± 0.03 mM; Figure 7B).

Next, we turned our attention to a longer (24-nt) *i*-motif-forming sequence known as hTeloC (5'-TAA CCC TAA CCC TAA CCC TAA CCC-3'). The sequence was found in the human genome within C-rich stretches present in telomers (66). Small molecules that bind to this sequence are potential gene expression modulators. The interaction of hTe-

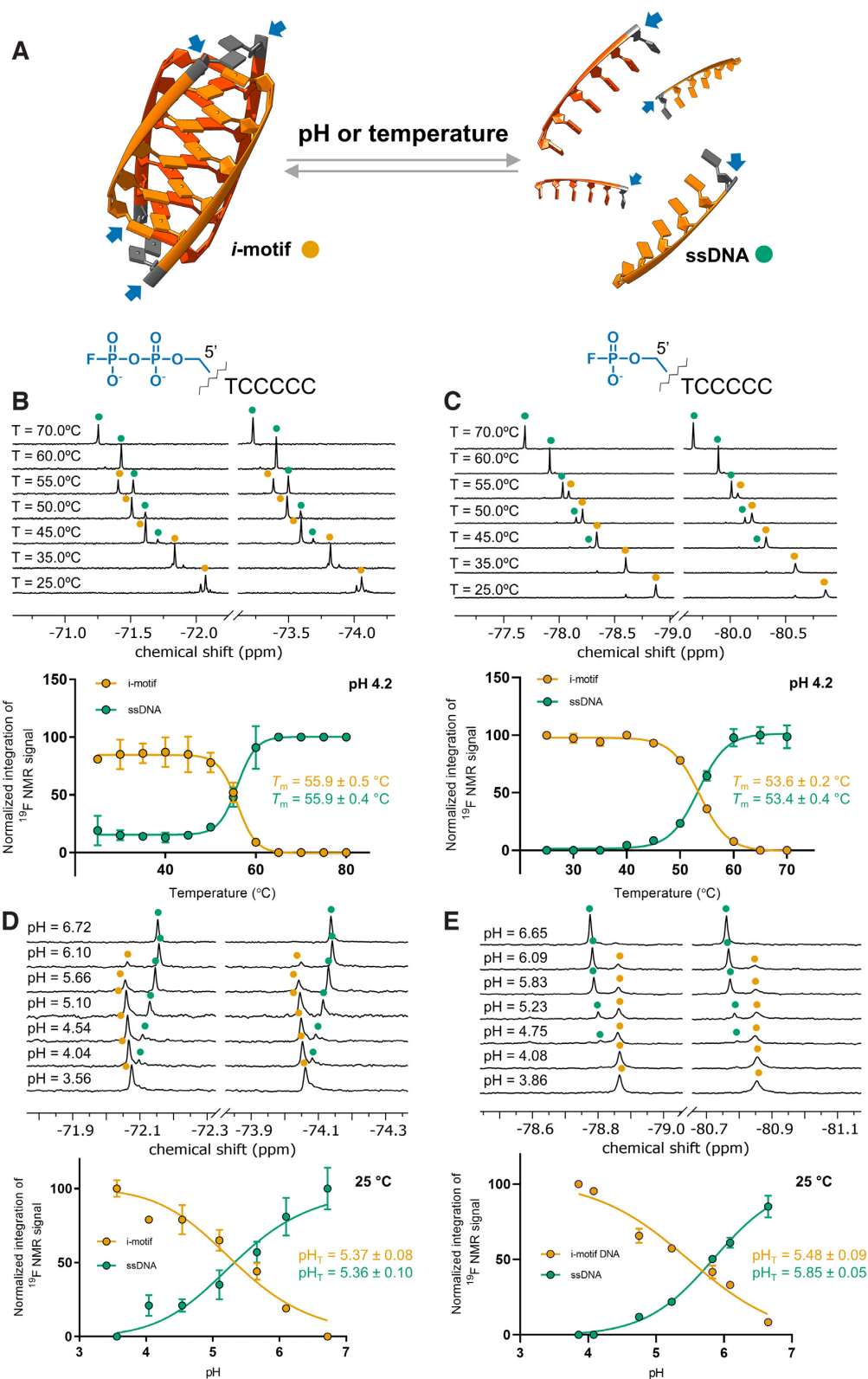


Table 4. Transitional pH (pH_T) and melting temperature (T_m) values for *i*-motif sequences

Sequence (5'→3')	Bases	Abbreviation	25°C	pH 4.2		pH 5.5	
			pH_T	^{19}F NMR [a]	UV-VIS	^{19}F NMR [a]	UV-VIS
TCCCCC	6	FPON1	5.48 ± 0.09	$53.6 \pm 0.2^\circ\text{C}$	$57.7 \pm 0.3^\circ\text{C}$	n.d.	n.d.
TCCCCC	6	FPPON1	5.37 ± 0.08	$55.9 \pm 0.5^\circ\text{C}$	$60.7 \pm 0.7^\circ\text{C}$	n.d.	n.d.
(TAACCC) ₄	24	FP-hTeloC	6.47 ± 0.07	$64.6 \pm 1.1^\circ\text{C}$	$54.7 \pm 2.4^\circ\text{C}$	$40.9 \pm 0.3^\circ\text{C}$	$42.2 \pm 0.1^\circ\text{C}$

Conditions: 50 mM sodium citrate, 10% D₂O; [a] ^{19}F NMR signal corresponding to *i*-motif structure was analyzed.

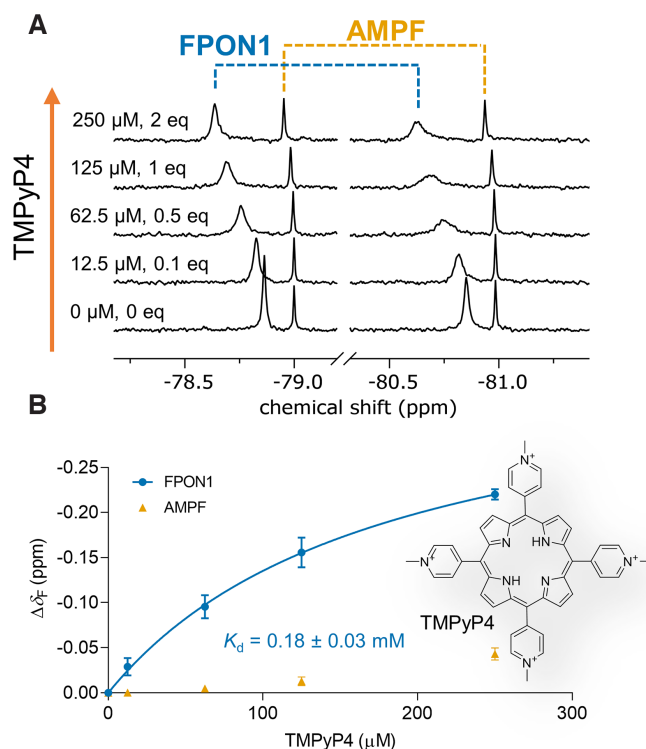


Figure 7. Interaction of *i*-motif-forming sequence, FPON1, with TMPyP4 ligand observed by ^{19}F NMR. (A) ^{19}F NMR spectra recorded during titration of FPON1 with TMPyP4 at pH 4.2 at 500 μM FPON1 (125 μM of *i*-motif structure), 100 μM AMPF, 50 mM sodium citrate buffer pH 4.2, 10% D₂O, 25°C. (B) δ_F chemical shift changes as a function of TMPyP4 concentration ($\Delta\delta_F = [\delta_F(\text{FPON1} + \text{TMPyP4}) - \delta_F(\text{FPON1})]$); data points represent mean values from duplicate experiments \pm S.E.

loC with multiple ligands has been evaluated by different biophysical methods including CD, fluorescence resonance energy transfer melting screen, UV/VIS spectroscopy and SPR (58,67). To generate an ^{19}F NMR probe for the discovery of potential binders, we synthesized an hTeloC sequence labeled with the FP moiety. The initial characterization of FP-hTeloC was performed as described for FPON1 and FPPON1. This confirmed that the sequence formed an *i*-motif structure under acidic conditions (Supplementary Figure S21D) and that pH-dependent unfolding of FP-hTeloC could be observed by ^{19}F NMR (Supplementary Figure S28) to determine the pH_T value (pH_T 6.4; Table 4). We also performed ^{19}F NMR-monitored temperature-

dependent experiments at two different pH values (pH 4.2 and 5.5; Supplementary Figures S29 and 30) to determine T_m values of 60 and 40°C, respectively (Table 4). To confirm our findings, we performed UV/VIS spectroscopy temperature-dependent experiments of FP-hTeloC and its unmodified counterpart, hTeloC, at both pH 4.2 and 5.5 (Supplementary Figures S31 and 32). The observed T_m values for the tested hTeloC sequences were 55°C at pH 4.2 and 43°C at pH 5.5 (Supplementary Table S4), which was in good agreement with published data (Supplementary Table S5), albeit there was a noticeable discrepancy between the values determined for FP-hTeloC by ^{19}F NMR and UV/VIS at pH 4.2 (Table 4). We hypothesize that this discrepancy may be a result of either contribution from some minor secondary form that affects the results of UV-VIS measurements (which form was observed by ^{19}F NMR at lower pH values, Supplementary Figure S29) or insufficient equilibration times prior to ^{19}F NMR measurements. After this preliminary characterization, we investigated the effect of mitoxantrone on the stability of FP-hTeloC by determining T_m in the presence of this ligand by both ^{19}F NMR and UV/VIS experiments. Intriguingly, ^{19}F NMR showed that addition of 0.3 equivalents of mitoxantrone had a small stabilizing effect on the *i*-motif ($\Delta T_m = +1.1^\circ\text{C}$), whereas higher concentrations had an opposite effect ($\Delta T_m = -3.2^\circ\text{C}$ in the presence of two equivalents). UV/VIS studies of FP-hTeloC and hTeloC confirmed the destabilizing effect of excess mitoxantrone on *i*-motif structure (Supplementary Figure S33). Finally, we used ^{19}F NMR to study FP-hTeloC and mitoxantrone interaction in a concentration-dependent manner at three different pH values (4.2, 5.5 and 6.4). To that end, ^{19}F NMR-monitored titration of FP-hTeloC with mitoxantrone was performed, while AMPF was used as a non-interacting reference. At pH 4.2 and 5.5, *i*-motif signals visible at -79.94 and -79.95 ppm shifted upfield in the presence of one equivalent of mitoxantrone, to -79.97 and -79.99 ppm, respectively (Figure 8A and B). ^{19}F NMR spectra at pH_T resulted in the appearance of the *i*-motif signal and single-stranded oligo signal in a 45:55 ratio. Addition of up to one equivalent of mitoxantrone induced changes to both resonances, but especially the one corresponding to *i*-motif (Figure 8C). No interaction between mitoxantrone and the reference AMPF molecule was observed. By plotting the chemical shifts as a function of concentration and fitting a 1:1 binding model, we determined the corresponding apparent dissociation constant equal 0.07 ± 0.03 mM for interaction at pH 4.2 (Figure 8D) which is in the same order of magnitude

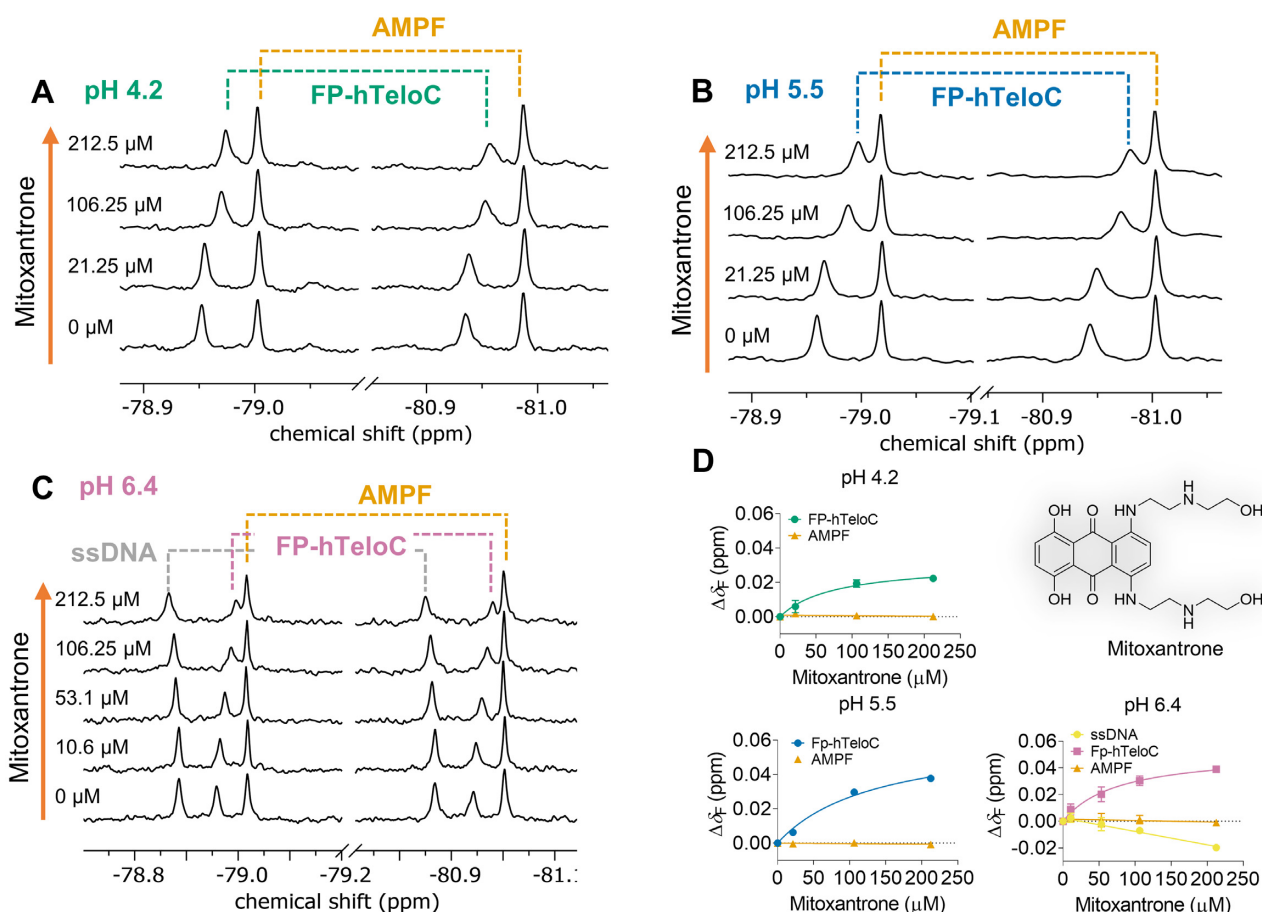


Figure 8. ^{19}F NMR study of FP-hTeloC with mitoxantrone at different pH. (A) Titration of FP-hTeloC with mitoxantrone at pH 4.2. (B) Titration of FP-hTeloC with mitoxantrone at pH 5.5. (C) Titration of FP-hTeloC with mitoxantrone at pH 6.4. (D) Chemical shift changes as a function of mitoxantrone concentration under different pH conditions ($\Delta\delta_{\text{F}} = [\delta_{\text{F}}(\text{FP-hTeloC} + \text{mitoxantrone}) - \delta_{\text{F}}(\text{FP-hTeloC})]$); data points represent mean values from duplicate experiments \pm S.E. Conditions: 212.5 μM FP-hTeloC, 80 μM AMPF, 50 mM sodium citrate buffer pH 4.2, 10% D_2O or 10 mM sodium cacodylate, 100 mM NaCl, 10% D_2O , pH 6.4. Experiments were performed in duplicate.

as the value previously determined by SPR (K_{d} 0.12 ± 0.003 mM measured at pH 5.5 in 10 mM sodium cacodylate buffer with 100 mM NaCl, 0.05% Tween-20 and 5% DMSO) (67).

Notably, the interaction of mitoxantrone with the *i*-motif could be observed even at pH 6.4, whereby only a fraction of the probe actually adapted the *i*-motif fold. Moreover, under these conditions ^{19}F NMR signal of ssDNA also visibly changed its chemical shift upon mitoxantrone addition ($\Delta\delta_{\text{F}} = -0.02$ ppm), which suggested the ligand weakly interacts with FP-hTeloC sequence also in ssDNA form (Figure 8D).

DISCUSSION

In this study, we sought to develop ^{19}F NMR oligonucleotide probes that combined functional versatility with synthetic accessibility. A careful review of the literature led us to conclude that the key factors for ^{19}F NMR probe development were biocompatibility, synthetic availability, responsiveness, sensitivity and application scope. In developing our ^{19}F NMR oligonucleotide probes, we focused on FP and FPP moieties, both of which are structurally compact and isosteric with naturally occurring mono- and diphos-

phate moieties. This was advantageous, as it could ensure highly responsive and biocompatible labels. However, due to the presence of only a single fluorine atom and strong coupling with the phosphorus nucleus ($^1J_{\text{P-F}} \sim 935$ Hz) that causes signal splitting into a doublet, the sensitivity of those labels is relatively low. Such low sensitivity could be partially compensated by the synthetic accessibility of such probes, enabling their usage even at high micromolar concentrations.

The first step of our study was the synthesis of 18 deoxyribonucleotide sequences modified at the 5' end with either FP or FPP moieties (Table 1 and Figure 2). We confirmed that the probes could be synthesized at good yields by adapting methods previously developed for mononucleotides. Importantly, the synthetic protocols were based only on commercially available building blocks for solid-supported synthesis, and as such did not require laborious preparation of modified phosphoramidites, thereby increasing the robustness and scalability of the synthetic protocol.

These fluorinated probes were subsequently evaluated as molecular tools for ^{19}F NMR studies. We demonstrated that the probes enabled us to distinguish between ssDNA and dsDNA due to different chemical shift values for those

entities. Moreover, the probes could also distinguish between perfect and imperfect duplexes (with a single mismatch at the 3' end of the complementary strand) independently of their thermodynamic stability (T_m values), thus overcoming a limitation of many existing probes for mismatch detection. Importantly, a combination of ^{19}F NMR signatures from the experiments performed with FP and FPP probes for the same sequence enabled unambiguous determination of every investigated duplex. We believe this finding can be further harnessed to develop assays for different DNA 3' end-processing enzymatic activities.

We next investigated whether FP and FPP probes could be applied to monitor formation and interactions of non-canonical secondary structures, i.e. G-quadruplexes and *i*-motifs. These structures have previously been investigated by ^{19}F NMR, but to the best of our knowledge a universal approach applicable to different non-canonical structures has not been proposed. The influence of FP/FPP moieties did not prevent formation of these secondary structures, albeit it affected their stability. G-quadruplexes were destabilized by FP and FPP moieties, a phenomenon that was observed upon addition of unmodified 5'-phosphate; whereas the stability of *i*-motifs was increased in the presence of FP/FPP. Next, we demonstrated that both G-quadruplex and *i*-motif formation, and either temperature- or pH-dependent transformations, could be detected by ^{19}F NMR with the use of our FP and FPP probes. Furthermore, we found that different *i*-motif topologies might also be detected by ^{19}F NMR using our probes. ^{19}F NMR experiments revealed the formation of more than one topology of the tetrameric *i*-motif structures at 25°C and pH 4.2. Similar phenomena have been previously reported for *i*-motif sequences. As an example, a fluorinated short TC₅ human telomeric and centromeric DNA sequence incorporating 2'-F-araC has been used to study *i*-motif topology by ^{19}F NMR (68). *i*-Motifs are stabilized by acidic conditions when a hemiprotonated cytosine-cytosine⁺ interaction results in two base-paired parallel-stranded duplexes, with fully intercalated nucleobases and antiparallel orientation (69). ^{19}F NMR revealed the existence of two possible topologies for this four-stranded *i*-motif form: the R-form and the S-form, in which either the 3' or the 5' end cytidine were located outside stacked cytosine-cytosine⁺ pairs (63,70). This phenomenon has also been examined by gel electrophoresis (68).

Finally, we demonstrated that ^{19}F NMR could detect the interaction between these secondary structures and either proteins or small molecules. The interaction with proteins was demonstrated by means of a TBA sequence (ON11) and thrombin (Figure 5). This sort of experiment may be performed to verify results obtained by other biophysical methods, as well as for the development of ^{19}F NMR-based screening assays to identify small molecules targeting nucleic acid-binding proteins. The ability to monitor nucleic acid interaction with small molecules by means of FP/FPP probes was exemplified using two *i*-motif structures and their previously identified ligands: TMPyP4 and mitoxantrone (65,71). So far, the discovery of small molecules targeting G-quadruplexes and *i*-motifs has relied mostly on probes emitting fluorescent light upon intercalating these secondary structures (66,72). However, such competition

assays are prone to fluorescence-associated interference and may lead to both false positive and negative results. As demonstrated here, ^{19}F NMR experiments enable direct observation of the interaction with small molecules due to chemical shift perturbation. Therefore, ^{19}F NMR experiments with our probes could be used in the future to validate hits from fluorescent assays, characterize binding affinities for different target sequences, as well as to develop fluorescence-independent high-throughput screening experiments aimed at identifying nucleic acid-binding ligands and verifying their specificity for a particular sequence or topology.

Further studies are required to determine whether FP and FPP labels are suitable for investigation of other types of nucleic acid-related phenomena (e.g. hairpin formation, RNA-DNA duplex formation, duplex versus triplex identification) and labeling of different types of nucleic acids (e.g. RNA, PNA).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Prof. Joanna Trylska (University of Warsaw) for providing access to UV-VIS spectrophotometer and Dr Mikolaj Chrominski (University of Warsaw) for critical reading of the manuscript.

Author contributions: M.B., J.K. and J.J. designed the study. M.B. and M.W. synthesized FPP-, FP- and unmodified oligonucleotides, M.B. performed UV/VIS and ^{19}F NMR experiments. M.B. and J.K. wrote the first draft of the manuscript. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

FUNDING

National Science Centre, Poland [UMO-2018/31/B/ST5/03821 to J.K., UMO-2016/21/B/ST5/02556 to J.J., UMO-2019/32/T/ST4/00071 to M.B.]; Ministry of Science and Higher Education, Poland [0149/DIA/2014/43 to M.B.]. Funding for open access charge: National Science Centre, Poland [UMO-2018/31/B/ST5/03821].

Conflict of interest statement. None declared.

REFERENCES

- Mergny, J.L., Phan, A.T. and Lacroix, L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Lett.*, **435**, 74–78.
- Vorlickova, M., Kejnovska, I., Sagi, J., Renciuik, D., Bednarova, K., Motlova, J. and Kypr, J. (2012) Circular dichroism and guanine quadruplexes. *Methods*, **57**, 64–75.
- Ponikova, S., Antalik, M. and Hianik, T. (2008) A circular dichroism study of the stability of guanine quadruplexes of thrombin DNA aptamers at presence of K⁺ and Na⁺ ions. *Gen. Physiol. Biophys.*, **27**, 271–277.
- Calabrese, D.R., Zlotkowski, K., Alden, S., Hewitt, W.M., Connelly, C.M., Wilson, R.M., Gaikwad, S., Chen, L., Guha, R., Thomas, C.J. *et al.* (2018) Characterization of clinically used oral antiseptics as quadruplex-binding ligands. *Nucleic Acids Res.*, **46**, 2722–2732.

5. Baltzinger, M., Sharma, K.K., Mely, Y. and Altschuh, D. (2013) Dissecting the oligonucleotide binding properties of a disordered chaperone protein using surface plasmon resonance. *Nucleic Acids Res.*, **41**, 10414–10425.
6. Milkani, E., Morais, S., Lambert, C.R. and McGimpsey, W.G. (2010) Detection of oligonucleotide systematic mismatches with a surface plasmon resonance sensor. *Biosens. Bioelectron.*, **25**, 1217–1220.
7. Sabharwal, N.C., Savikhin, V., Turek-Herman, J.R., Nicoludis, J.M., Szalai, V.A. and Yatsunyk, L.A. (2014) N-methylmesoporphyrin IX fluorescence as a reporter of strand orientation in guanine quadruplexes. *FEBS J.*, **281**, 1726–1737.
8. Lawson, C.P., Fuchtbauer, A.F., Wranne, M.S., Giraud, T., Floyd, T., Dumat, B., Andersen, N.K., El-Sagheer, A.H., Brown, T., Graden, H. et al. (2018) Synthesis, oligonucleotide incorporation and fluorescence properties in DNA of a bicyclic thymine analogue. *Sci. Rep.*, **8**, 13970.
9. Singh, D., Kumar, V. and Ganesh, K.N. (1990) Oligonucleotides, part 5+: synthesis and fluorescence studies of DNA oligomers d(AT)₅ containing adenines covalently linked at C-8 with dansyl fluorophore. *Nucleic Acids Res.*, **18**, 3339–3345.
10. Lane, A.N. (1995) Determination of fast dynamics of nucleic acids by NMR. *Methods Enzymol.*, **261**, 413–435.
11. Chirayil, S., Wu, Q., Amezcua, C. and Luebke, K.J. (2014) NMR characterization of an oligonucleotide model of the MiR-21 pre-element. *PLoS One*, **9**, e108231.
12. Barnwal, R.P., Yang, F. and Varani, G. (2017) Applications of NMR to structure determination of RNAs large and small. *Arch. Biochem. Biophys.*, **628**, 42–56.
13. Tonelli, M. and James, T.L. (1998) Insights into the dynamic nature of DNA duplex structure via analysis of nuclear overhauser effect intensities. *Biochemistry*, **37**, 11478–11487.
14. Fielding, L. (2003) NMR methods for the determination of protein-ligand dissociation constants. *Curr. Top. Med. Chem.*, **3**, 39–53.
15. Fielding, L. (2007) NMR methods for the determination of protein-ligand dissociation constants. *Prog. Nucl. Mag. Res. Sp.*, **51**, 219–242.
16. Dalvit, C., Fagerness, P.E., Hadden, D.T.A., Sarver, R.W. and Stockman, B.J. (2003) Fluorine-NMR experiments for high-throughput screening: Theoretical aspects, practical considerations, and range of applicability. *J. Am. Chem. Soc.*, **125**, 7696–7703.
17. Guo, F.M., Li, Q. and Zhou, C.Z. (2017) Synthesis and biological applications of fluoro-modified nucleic acids. *Org. Biomol. Chem.*, **15**, 9552–9565.
18. Chrominski, M., Baranowski, M.R., Chmielinski, S., Kowalska, J. and Jemielity, J. (2020) Synthesis of trifluoromethylated purine ribonucleotides and their evaluation as ¹⁹F NMR probes. *J. Org. Chem.*, **85**, 3440–3453.
19. Morihiro, K., Kasahara, Y. and Obika, S. (2017) Biological applications of xeno nucleic acids. *Mol. Biosyst.*, **13**, 235–245.
20. Tanabe, K., Sugiura, M. and Nishimoto, S.-i. (2010) Monitoring of duplex and triplex formation by ¹⁹F NMR using oligodeoxynucleotides possessing 5-fluorodeoxyuridine unit as ¹⁹F signal transmitter. *Bioorg. Med. Chem.*, **18**, 6690–6694.
21. Puffer, B., Kreutz, C., Rieder, U., Ebert, M.O., Konrat, R. and Micura, R. (2009) 5-Fluoro pyrimidines: labels to probe DNA and RNA secondary structures by 1D ¹⁹F NMR spectroscopy. *Nucleic Acids Res.*, **37**, 7728–7740.
22. Tanabe, K., Tsuda, T., Ito, T. and Nishimoto, S.-I. (2013) Probing DNA mismatched and bulged structures by using ¹⁹F NMR Spectroscopy and oligodeoxynucleotides with an ¹⁹F-labeled nucleobase. *Chemistry*, **19**, 15133–15140.
23. Bhumra, N., Tahtinen, V. and Virta, P. (2018) Synthesis and applicability of base-discriminating DNA-triplex-forming ¹⁹F NMR Probes. *Eur. J. Org.*, **2018**, 605–613.
24. Kieger, A., Wiester, M.J., Proccisi, D., Parrish, T.B., Mirkin, C.A. and Thaxton, C.S. (2011) Hybridization-Induced “off-on” ¹⁹F NMR signal probe release from DNA-functionalized gold nanoparticles. *Small*, **7**, 1977–1981.
25. Granqvist, L. and Virta, P. (2014) 4'-C-[(4-Trifluoromethyl-1H-1,2,3-triazol-1-yl)methyl]thymidine as a sensitive ¹⁹F NMR sensor for the detection of oligonucleotide secondary structures. *J. Org. Chem.*, **79**, 3529–3536.
26. Granqvist, L. and Virta, P. (2015) 2'-O-(4-CF₃-triazol-1-yl)methyl Uridine—a sensitive ¹⁹F NMR sensor for the detection of RNA secondary structures. *J. Org. Chem.*, **80**, 7961–7970.
27. Fauster, K., Kreutz, C. and Micura, R. (2012) 2'-SCF₃ uridine—a powerful label for probing structure and function of RNA by ¹⁹F NMR spectroscopy. *Angew. Chem. Int. Ed.*, **51**, 13080–13084.
28. Kosutic, M., Jud, L., Da Veiga, C., Frener, M., Fauster, K., Kreutz, C., Ennifar, E. and Micura, R. (2014) Surprising base pairing and structural properties of 2'-trifluoromethylthio-modified ribonucleic acids. *J. Am. Chem. Soc.*, **136**, 6656–6663.
29. Olszewska, A., Pohl, R. and Hocek, M. (2017) Trifluoroacetophenone-linked nucleotides and DNA for studying of DNA-protein interactions by ¹⁹F NMR spectroscopy. *J. Org. Chem.*, **82**, 11431–11439.
30. Pan, D.F., Gambhir, S.S., Toyokuni, T., Iyer, M.R., Acharya, N., Phelps, M.E. and Barrio, J.R. (1998) Rapid synthesis of a 5'-fluorinated oligodeoxy-nucleotide: a model antisense probe for use in imaging with positron emission tomography (PET). *Bioorg. Med. Chem. Letters*, **8**, 1317–1320.
31. Lucas, R., Penalver, P., Gomez-Pinto, I., Vengut-Climent, E., Mtashobya, L., Cousin, J., Maldonado, O.S., Perez, V., Reyes, V., Avino, A. et al. (2014) Effects of sugar functional groups, hydrophobicity, and fluorination on carbohydrate-DNA stacking interactions in water. *J. Org. Chem.*, **79**, 2419–2429.
32. Godeau, G., Arnion, H., Brun, C., Staedel, C. and Barthelemy, P. (2010) Fluorocarbon oligonucleotide conjugates for nucleic acids delivery. *Med. Chem. Comm.*, **1**, 76–78.
33. Metelev, V., Zhang, S.R., Zheng, S.K., Kumar, A.T.N. and Bogdanov, A. (2017) Fluorocarbons enhance intracellular delivery of short STAT3-sensors and enable specific imaging. *Theranostics*, **7**, 3354–3368.
34. Bao, H.L., Ishizuka, T., Iwanami, A., Oyoshi, T. and Xu, Y. (2017) A simple and sensitive ¹⁹F NMR approach for studying the interaction of RNA G-quadruplex with ligand molecule and protein. *ChemistrySelect*, **2**, 4170–4175.
35. Bao, H.L., Ishizuka, T., Sakamoto, T., Fujimoto, K., Uechi, T., Kenmochi, N. and Xu, Y. (2017) Characterization of human telomere RNA G-quadruplex structures in vitro and in living cells using ¹⁹F NMR spectroscopy. *Nucleic Acids Res.*, **45**, 5501–5511.
36. Baranowski, M.R., Nowicka, A., Rydzik, A.M., Warminski, M., Kasprzyk, R., Wojtczak, B.A., Wojcik, J., Claridge, T.D., Kowalska, J. and Jemielity, J. (2015) Synthesis of fluorophosphate nucleotide analogues and their characterization as tools for ¹⁹F NMR studies. *J. Org. Chem.*, **80**, 3982–3997.
37. Bao, H.L., Liu, H.S. and Xu, Y. (2019) Hybrid-type and two-tetrad antiparallel telomere DNA G-quadruplex structures in living human cells. *Nucleic Acids Res.*, **47**, 4940–4947.
38. Thillier, Y., Decroly, E., Morvan, F., Canard, B., Vasseur, J.J. and Debart, F. (2012) Synthesis of 5' cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine-N-7)-methyl transferase. *RNA*, **18**, 856–868.
39. Piotto, M., Saudek, V. and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR-spectroscopy of aqueous-solutions. *J. Biomol. NMR*, **2**, 661–665.
40. Sklenar, V., Piotto, M., Leppik, R. and Saudek, V. (1993) Gradient-tailored water suppression for ¹H-¹⁵N HSQC experiments optimized to retain full sensitivity. *J. Magn. Reson.*, **102**, 241–245.
41. Beaucage, S.L. and Caruthers, M.H. (1981) Deoxynucleoside phosphoramidites—a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron. Lett.*, **22**, 1859–1862.
42. Mergny, J.L. and Lacroix, L. (2003) Analysis of thermal melting curves. *Oligonucleotides*, **13**, 515–537.
43. Mathad, R.I. and Yang, D. (2011) G-Quadruplex structures and G-quadruplex-interactive compounds. In: Songyang, Z. (ed). *Telomeres and Telomerase. Methods in Molecular Biology (Methods and Protocols)*. Humana Press, Vol. 735.
44. Paeschke, K., Bochman, M.L., Garcia, P.D., Cejka, P., Friedman, K.L., Kowalczykowski, S.C. and Zakian, V.A. (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature*, **497**, 458–462.
45. Lopes, J., Piazza, A., Bermejo, R., Kriegsman, B., Colosio, A., Teulade-Fichou, M.P., Foiani, M. and Nicolas, A. (2011) G-quadruplex-induced instability during leading-strand replication. *EMBO J.*, **30**, 4033–4046.

46. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. and Hurley, L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 11593–11598.
47. Balasubramanian, S., Hurley, L.H. and Neidle, S. (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.*, **10**, 261–275.
48. Kim, M.Y., Vankayalapati, H., Kazuo, S., Wierzbicka, K. and Hurley, L.H. (2002) Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *J. Am. Chem. Soc.*, **124**, 2098–2099.
49. Wang, K.Y., Krawczyk, S.H., Bischofberger, N., Swaminathan, S. and Bolton, P.H. (1993) The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity. *Biochemistry*, **32**, 11285–11292.
50. Bock, L.C., Griffin, L.C., Latham, J.A., Vermaas, E.H. and Toole, J.J. (1992) Selection of single-stranded-DNA molecules that bind and inhibit human thrombin. *Nature*, **355**, 564–566.
51. Avino, A., Fabrega, C., Tintore, M. and Eritja, R. (2012) Thrombin binding aptamer, more than a simple aptamer: chemically modified derivatives and biomedical applications. *Curr. Pharm. Des.*, **18**, 2036–2047.
52. Macaya, R.F., Schultze, P., Smith, F.W., Roe, J.A. and Feigon, J. (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 3745–3749.
53. Foster, M.P., McElroy, C.A. and Amero, C.D. (2007) Solution NMR of large molecules and assemblies. *Biochemistry*, **46**, 331–340.
54. Wright, E.P., Huppert, J.L. and Waller, Z.A.E. (2017) Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.*, **45**, 2951–2959.
55. Fleming, A.M., Ding, Y., Rogers, R.A., Zhu, J., Burton, A.D., Carlisle, C.B. and Burrows, C.J. (2017) 4n-1 Is a “Sweet Spot” in DNA i-motif folding of 2'-deoxycytidine homopolymers. *J. Am. Chem. Soc.*, **139**, 4682–4689.
56. Rajendran, A., Nakano, S. and Sugimoto, N. (2010) Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.*, **46**, 1299–1301.
57. Zeraati, M., Langley, D.B., Schofield, P., Moye, A.L., Rouet, R., Hughes, W.E., Bryan, T.M., Dinger, M.E. and Christ, D. (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.*, **10**, 631–637.
58. Pagano, A., Iaccarino, N., Abdelhamid, M.A.S., Brancaccio, D., Garzarella, E.U., Di Porzio, A., Novellino, E., Waller, Z.A.E., Pagano, B., Amato, J. et al. (2018) Common G-quadruplex binding agents found to interact with i-motif-forming DNA: unexpected multi-target-directed compounds. *Front. Chem.*, **6**, 281.
59. Sheng, Q., Neaverson, J.C., Mahmoud, T., Stevenson, C.E.M., Matthews, S.E. and Waller, Z.A.E. (2017) Identification of new DNA i-motif binding ligands through a fluorescent intercalator displacement assay. *Org. Biomol. Chem.*, **15**, 5669–5673.
60. Sabale, P.M., Tanpure, A.A. and Srivatsan, S.G. (2018) Probing the competition between duplex and G-quadruplex/i-motif structures using a conformation-sensitive fluorescent nucleoside probe. *Org. Biomol. Chem.*, **16**, 4141–4150.
61. Tsvetkov, V.B., Zatsepin, T.S., Belyaev, E.S., Kostyukevich, Y.I., Shpakovski, G.V., Podgorsky, V.V., Pozmogova, G.E., Varizhuk, A.M. and Aralov, A.V. (2018) i-Clamp phenoxazine for the fine tuning of DNA i-motif stability. *Nucleic Acids Res.*, **46**, 2751–2764.
62. Fenna, C.P., Wilkinson, V.J., Arnold, J.R.P., Cosstick, R. and Fisher, J. (2008) The effect of 2'-fluorine substitutions on DNA i-motif conformation and stability. *Chem. Commun.*, 3567–3569.
63. Kanaori, K., Maeda, A., Kanehara, H., Tajima, K. and Makino, K. (1998) ¹H nuclear magnetic resonance study on equilibrium between two four-stranded solution conformations of short d(CnT). *Biochemistry*, **37**, 12979–12986.
64. Lieblein, A.L., Buck, J., Schlepckow, K., Furtig, B. and Schwalbe, H. (2012) Time-resolved NMR spectroscopic studies of DNA i-motif folding reveal kinetic partitioning. *Angew. Chem. Int. Ed.*, **51**, 250–253.
65. Fedoroff, O.Y., Rangan, A., Chemeris, V.V. and Hurley, L.H. (2000) Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. *Biochemistry*, **39**, 15083–15090.
66. Debnath, M., Fatma, K. and Dash, J. (2019) Chemical regulation of DNA i-motifs for nanobiotechnology and therapeutics. *Angew. Chem. Int. Ed.*, **58**, 2942–2957.
67. Wright, E.P., Day, H.A., Ibrahim, A.M., Kumar, J., Boswell, L.J.E., Huguin, C., Stevenson, C.E.M., Pors, K. and Waller, Z.A.E. (2016) Mitoxantrone and analogues bind and stabilize i-motif forming DNA sequences. *Sci. Rep.*, **6**, 39456.
68. Abou Assi, H., Harkness, R.W., Martin-Pintado, N., Wilds, C.J., Campos-Olivas, R., Mittermaier, A.K., Gonzalez, C. and Damha, M.J. (2016) Stabilization of i-motif structures by 2'-β-fluorination of DNA. *Nucleic Acids Res.*, **44**, 4998–5009.
69. Gehring, K., Leroy, J.L. and Gueron, M. (1993) A tetrameric DNA-structure with protonated cytosine-cytosine base-pairs. *Nature*, **363**, 561–565.
70. Day, H.A., Pavlou, P. and Waller, Z.A.E. (2014) i-Motif DNA: Structure, stability and targeting with ligands. *Bioorg. Med. Chem.*, **22**, 4407–4418.
71. Qin, T.X., Liu, K.H., Song, D., Yang, C.F. and Su, H.M. (2017) Porphyrin bound to i-motifs: intercalation versus external groove binding. *Chem. Asian J.*, **12**, 1578–1586.
72. Vummidi, B.R., Alzeer, J. and Luedtke, N.W. (2013) Fluorescent probes for G-quadruplex structures. *Chembiochem*, **14**, 540–558.
73. Russo Krauss, I., Merlino, A., Randazzo, A., Novellino, E., Mazzarella, L. and Sica, F. (2012) High-resolution structures of two complexes between thrombin and thrombin-binding aptamer shed light on the role of cations in the aptamer inhibitory activity. *Nucleic Acids Res.*, **40**, 8119–8128.