Synthesis and applicability of base-discriminating DNA triplex-forming ^{19}F NMR probes

Naresh Bhuma,[a] Ville Tähtinen,*[a] and Pasi Virta,*[a]

Abstract: Phosphoramidite building blocks of CF_3-modified *N*^{4}(6-amino-2-pyridinyl)deoxycytidines were synthesized, incorporated into triplex-forming 2'-deoxyoligonucleotide strands and the applicability of the probes to recognize nucleobase content in the pyrimidine-rich strand of double helical DNA targets was evaluated. As expected, the obtained ^{19}F NMR resonances were sensitive to the base content and unique ^{19}F NMR-spectral fingerprints could be obtained.

Introduction

Most of the DNA probe assays ultimately rely upon sequence recognition events associated with hybridization energies. The differences in energies are, however, small when only a single nucleobase mismatch in an extended oligodeoxynucleotide is involved. This may lead to a modest base discrimination. Fluorescence-labeled DNA probes have the potential to simplify DNA probe assays if the fluorescence label exhibits a drastic change in fluorescence intensity between the hybridized and non-hybridized probe.[1-4] Due to this fluorescence change, the bases on the complementary strands can be fluorometrically read out. Recently, ^{19}F NMR -based sensors, because of their sensitivity to local environments, have also been utilized in this context.[5-8] Tanabe et al. described ^{19}F NMR probes for the detection of mismatches and bulges, in which shift discrimination between different nucleotide content could be provided (**1**, Figure 1).[5] Hocek et al. used biaryl-substituted nucleotides as ^{19}F-NMR/fluorescent-dual labels that were capable of base-discrimination, detection of deletions and mismatches and sensing of hairpins (**2**, Figure 1).[6] Like the described fluorescent sensors, the detection in these two examples required duplex formation between the target and the sensor-labelled strand. Sakamoto et al. developed a 3,5-bis(trifluoromethyl)benzene-modified bisbenzimide H33258 probe that binds to double-helical DNAs having an AATT sequence.[7] Based on the ^{19}F NMR chemical shift, this external probe was able to discriminate the sequence neighboring the AATT binding site, and also discriminate single-nucleotide polymorphisms (SNP) in bulge structures close to the binding site.[8]

The Hoogsteen-face recognition (i.e. triplex formation) serves another option to recognize double helices directly without unwinding the duplex. Fluorescent sensors based on this recognition motif at purine-rich stretches have been described.[2-4] On the face of it, there would appear to be no simple choice for the recognition of pyrimidine-rich regions on the double helices, although nucleobase-modifications that bind to pyrimidine bases from the Hoogsteen face have been reported.[9]

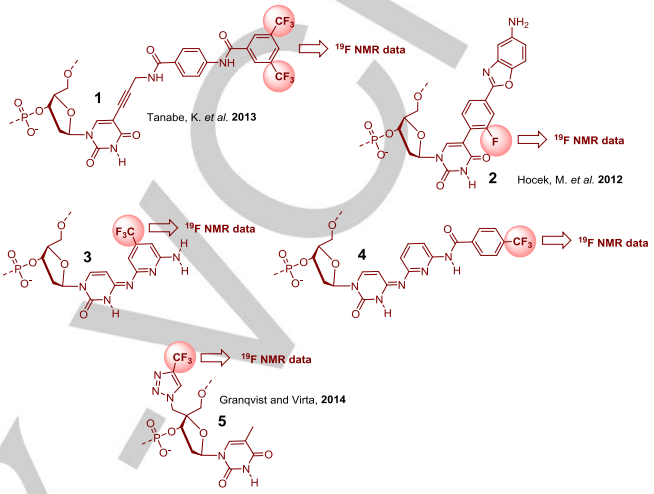


Figure 1. Nucleoside-derived ^{19}F NMR sensors (**3** and **4** described in the present study).

In the present study, phosphoramidite building blocks of appropriate ^{19}F NMR sensors (**3** and **4**, i.e. 2'-deoxy-*N*^{4}-(6-aminopyridin-2-yl)cytidine[10] derivatives, Figure 1) were synthesized, incorporated to a triplex-forming oligonucleotide (TFO), and the applicability of the building blocks as sensors to recognize single nucleotide content on the pyrimidine rich strand of a double helical DNA target was evaluated. The ^{19}F NMR shift reflection of the sensors (**3** and **4**), together with that of previously described 4'-(CF_3-triazolylmethyl)thymidine[11] (**5**), to the nucleotide content was evaluated in a 15 nucleotide long triple helical 2'-deoxyoligonucleotide model[12] (Figure 2). The plausible ^{19}F NMR shift discrimination by **5** may be expected to be based on changes in the groove environment only, whereas **3** and **4** may provide more specific recognition of the whole base-pair.[9,10] The orientation of the CF_3-group in the base modifications of **3** and **4** (i.e. CF_3-analogs of a previously reported triplet-forming nucleobase-modification[10]), together with plausible existence or absence of the extra hydrogen bonding, intercalation in the triplet, and potential changes in the steric environment, may expectedly increase the nucleobase-dependent ^{19}F NMR-shift discrimination of the sensors (Figure 2).

[a] Dr. N. Bhuma, M. Sc. V. Tähtinen, Prof. P. Virta
Department of Chemistry, University of Turku, Turku 20014 (Finland)
E-mail: pamavi@utu.fi, http://bioorganic.utu.fi/

Supporting information for this article is given via a link at the end of the document.

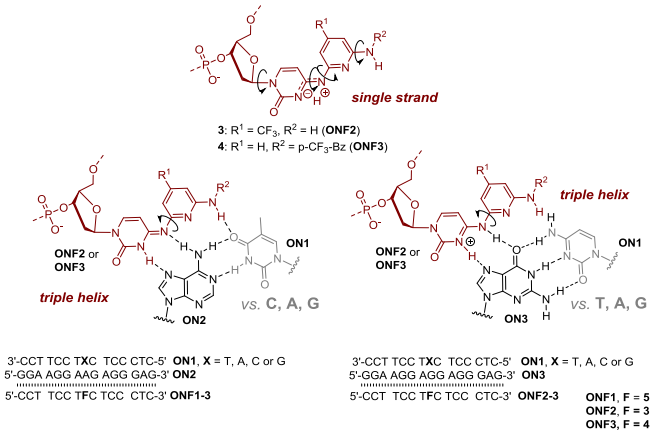
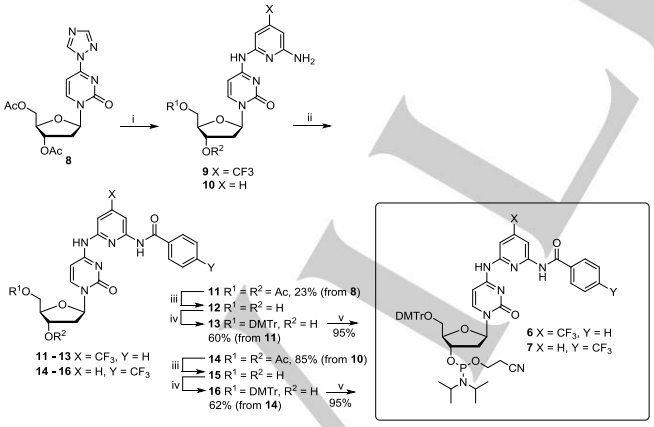


Figure 2. The triple helical models studied. Proposed hydrogen bonding, including tautomers and rotamers of **3** and **4** in the triple helices are described.

Results and Discussion

Synthesis of the phosphoramidite building blocks of the sensors 3 and 4

The synthesis of the phosphoramidites **6** and **7** is described in Scheme 1. 4-triazolyl-2'-deoxyuridine (**8**) was synthesized as previously reported[10] and exposed to transamination with 2,6-diaminopyridines. The subsequent protection with 4-trifluoromethylbenzoic acid and benzoyl chloride gave base-modified nucleosides **11** and **12**. The 3'- and 5'-O-acetyl groups were selectively removed with a gentle treatment with ammonia and the 5'-OH group was 4,4'-dimethoxytritylated to give **13** and **16**. Phosphitylation of the 3'-OH group with 2-cyanoethyl *N*,*N*-diisopropylphosphoramidochloridite gave the desired phosphoramidites **6** and **7**.



Scheme 1. Synthesis of the phosphoramidite building blocks (**6** and **7**) of the sensors **3** and **4**. Conditions: i) 4-trifluoromethyl-2,6-diaminopyridine in pyridine for **9** and 2,6-diaminopyridine in pyridine for **10**; ii) Benzoyl chloride in pyridine for **11** and 4-trifluoromethylbenzoic acid, PyBOP, DIEA, DMF for **14**; iii) Ammonia in methanol; iv) DMTrCl in pyridine; v) 2-cyanoethyl *N*,*N*-diisopropylphosphoramidochloridite, Et_3N, DCM.

Synthesis of the ^{19}F-labelled triplex-forming oligonucleotides

Oligonucleotides (**ONF1-3**) were synthesized on a 1.0 µmol scale using an automatic DNA-synthesizer (**ONF1** synthesized as previously described[12]). **ONF2** was synthesized on a standard dC-3'-*O*-yl succinate loaded and **ONF3** on a dC-3'-*O*-oyl hydroquinone *O*,*O*'-diacetate ('Q-linker')[13] loaded LCAA-CPG supports. Benzylthiotetrazol was used as an activator. Coupling times of 20 and 600s were used for the commercially available DNA building blocks (N^4-Ac-dC- and T-phosphoramidites) and for the phosphoramidites **6** and **7**, respectively. According to the DMTr assay and RP HPLC analysis of the released products, nearly quantitative couplings of **6** and **7** could be provided by this phosphoramidite cycle. After the chain assembly, the solid supported **ONF2** and **ONF3** were released by concentrated ammonia and a mixture of 0.05 M K_2CO_3 in methanol, respectively. The gentle cleavage from the Q-linker remained the 4-trifluoromethylbenzoyl group of **ONF3** intact. The benzoyl protection of the 4-trifluoromethyl-2,6-diaminopyridine moiety proved, however, surprisingly stable. After cleavage with concentrated ammonia, **ONF2** was further exposed to a mixture of 40% aqueous methylamine and concentrated ammonia (1:1, *v/v*, 5h at 55°C) to complete the protecting group removal. **ONF2** and **ONF3** were purified by RP HPLC and their authenticity was verified by MS (ESI-TOF) spectroscopy (Table S1). According to UV-absorbance at 260 nm, isolated yields of **ONF2** and **ONF3** were 18 and 25%, respectively.

UV-melting profile analysis of the triple helices

The effect of the sensors **3**, **4** and **5** on the stability of triple helix models (**ON1**/**ON2**/**ONF1-3** and **ON1**/**ON3**/**ONF2-3**) were first evaluated by UV-melting profile analysis. It may be worth noting that N^4-(6-amino-2-pyridinyl)deoxycytidine (i.e. pattern analogue for **3** and **4**) has originally been designed to interact primarily with a **C**-G base pair interruption of the purine target tract, but interaction with **A**-T- and to a lesser extent with **G**-C-base pairs have also been reported[10] (bold letters indicate the Hoogsteen-face targeted bases). Sensors **3** and **4** may be expected to interact with **A**-T and **G**-C-base pairs in similar manner. On the **ON1**/**ON2**-duplex: the triplex-forming probes **ONF1-3** were targeted to the Hoogsteen-face of the purine rich strand (**ON2**) and the sensors **3**, **4** and **5** were aimed to bind to the adenine residue (cf. Figure 2). The nucleobase content (X = T, A, C or G) at the Watson-Crick face of the adenine residue (**ON1**) was varied. Depending on the tautomers, sensors **3** and **4** may also bind to guanine base. The stabilities of **ON1**/**ON3**/**ONF2**- and **ON1**/**ON3**/**ONF3**-models, with the variable base-content at **ON1**, were hence also studied (cf. Figure 2). The measurements were carried out at 260 (data not shown) and 295 nm using 2 µmol L^{-1} of each 2'-deoxyoligonucleotide in a mixture of 10 mmol L^{-1} sodium cacodylate (pH 5.5) and 0.1 mol L^{-1} NaCl. The obtained melting temperatures (T_m[3]) are shown in Table 1. The impact of the sensors (**3**-**5**) on the triplex stability is emphasized as ΔT_m[3]-values (parentheses) in comparison to the T_m[3]-values of unmodified triple helices **ON1**/**ON2**/**ON4** and **ON1**/**ON2**/**ON5**.

As seen in Table 1, sensor **5** (**ONF1**) reduced the triplex stability to a small extent (ΔT_m^3 = -2.4 – -3.0°C compared to **ON1**/**ON2**/**ON4**-triplex), whereas sensor **3** reduced the stability remarkably (ΔT_m^3 = -5.8 – -14.3°C compared to **ON1**/**ON2**/**ON4**-triplex and ΔT_m^3 = -9.6 – -16.8°C compared to **ON1**/**ON3**/**ON5**-triplex). Sensor **4** also reduced the triplex stability in most cases (-3.6 – -7.2°C compared to **ON1**/**ON2**/**ON4**-triplex and ΔT_m^3 = -1.8 – -8.1°C compared to **ON1**/**ON3**/**ON5**-triplex), but in the case of **ON1**/**ON2**/**ONF3**-triplex, with a **4*****A**-A triad (X = A in **ON1**), a small triplex stabilization was observed (ΔT_m^3 = +1.2°C). It may be reasonable to conclude that sensors **3** and **4** probably cannot find the optimal triple helical constructs proposed in Figure 2, but a deviation on the triple helices is resulted. The 4-CF_3-benzoyl moiety of **4** seemed to compensate this reduced stability by plausible intercalation. Sensors **3** and **4** did not show notable selectivity on the binding between adenine (**ON2**) and guanine (**ON3**) bases. This may support the idea of the alternating tautomers of **3** and **4**, albeit the triplex stability decreased in both cases. It may be worth noting that **ONF2** and **ONF3** single strands may exist as a set of tautomers and rotamers (cf. Figure 2), but the structural mobility is reduced in triple helices. Rotation around the N^4(Cyt)-C^2(Py) σ-bond may, however, be relatively facile (preventing e.g. the described hydrogen bonding to O^4(Thy), Figure 2). Protonation of the constructs, particularly that of the pyridine residue, may certainly also play important role in the binding. The UV thermal melting temperatures of the **ON1**/**ON2** and **ON1**/**ON3** duplexes are listed in Table S2.

Table 1. UV melting temperatures of the triplexes.

duplex	ON4	ON5	ONF1	ONF2	ONF3
			T_m^3 / °C		
ON1/ON2					
X = T	40.8 ± 0.4		38.4 ± 0.2 (–2.4)	26.5 ± 0.4 (–14.3)	33.6 ± 1.2 (–7.2)
X = A	36.3 ± 0.2		33.3 ± 0.3 (–3.0)	30.5 ± 0.3 (–5.8)	37.5 ± 0.5 (+1.2)
X = C	37.9 ± 0.1		34.9 ± 0.1 (–3.0)	25.2 ± 0.4 (–12.7)	32.3 ± 0.7 (–5.6)
X = G	36.9 ± 0.2		33.9 ± 0.3 (–3.0)	23.3 ± 0.6 (–13.6)	34.3 ± 1.2 (–3.6)
ON1/ON3					
X = T		40.1 ± 0.5		29.6 ± 0.3 (–10.5)	36.3 ± 0.6 (–3.8)
X = A		40.3 ± 0.1		29.7 ± 0.2 (–10.6)	33.9 ± 0.5 (–6.4)
X = C		42.5 ± 0.4		25.7 ± 0.4 (–16.8)	34.4 ± 0.6 (–8.1)
X = G		38.8 ± 0.2		29.2 ± 0.2 (–9.6)	37.0 ± 0.6 (–1.8)

Conditions and notes: The measurements were carried out at 295 nm using mixtures of 2.0 µmol L^{-1} each ON, 10 mmol L^{-1} sodium cacodylate (pH 5.5), 0.1 mol L^{-1} NaCl in H_2O. ΔT_m-values in parentheses compared to T_m-values of unmodified triple helices (**ON1**/**ON2**/**ON4** and **ON1**/**ON3**/**ON5**).

^{19}F NMR analysis

The ^{19}F NMR shift response of the sensors **3**, **4** and **5** to recognize variable base content in the pyrimidine-rich sequence (**ON1**, X = T, A, C or G) was next evaluated (Figures 3 and 4). The measurements were carried out using 10 µmol L^{-1} of each oligonucleotide, 10 mmol L^{-1} sodium cacodylate (pH 5.5) and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v*/*v*) at 20°C. First, the ability of sensor **5** (**ONF1**) to recognize variable base pairs in the target duplex was evaluated. The plausible ^{19}F NMR shift discrimination by **5** was expected to be based on changes in the local environment of the Crick-Hoogsteen groove[14]. **ONF1** alone gave a relatively sharp ^{19}F resonance signal at -62.73 ppm (Figure 3A). When 1 equivalent of the purine sequence **ON2** was added, various scattered signals were observed, indicating unspecific interactions between **ONF1** and **ON2**. Adding 1 equivalent of **ON1** resulted in one, relatively sharp signal. In the case of the complementary **ON1**/**ON2**/**ONF1** triplex (**5*****A**-T triad, X = T in **ON1**), the chemical shift of the signal was -61.87 ppm. The non-matching **5*****A**-C triad (X = C in **ON1**) gave a signal with exactly the same chemical shift. The **5*****A**-A triad (X = A in **ON1**) gave a slightly different chemical shift (-61.83 ppm), whereas in the case of the **5*****A**-G triad (X = G in **ON1**), the change in the chemical shift was notable (-61.54 ppm). Taken together, sensor **5** was able to differentiate non-matching adenine-purine base pairs (**A**-A and **A**-G) from adenine-pyrimidine base pairs (**A**-T and **A**-C), although the discrimination between the **A**-A base pair and adenine-pyrimidine base pairs was modest.

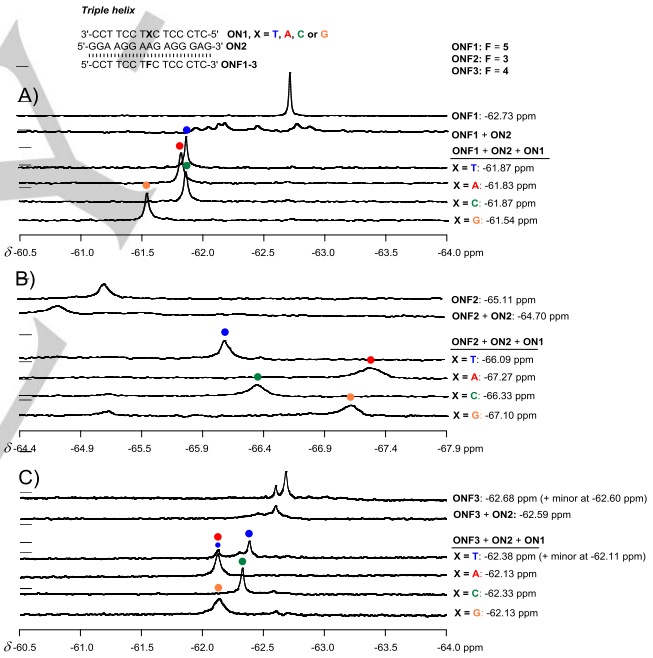


Figure 3. ^{19}F NMR spectra of the triple helices (**F*A-**X-triads, X = T, A, C or G). Conditions: 10 µmol L^{-1} of each ON, 10 mmol L^{-1} sodium cacodylate (pH 5.5) and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v*/*v*) at 20°C.

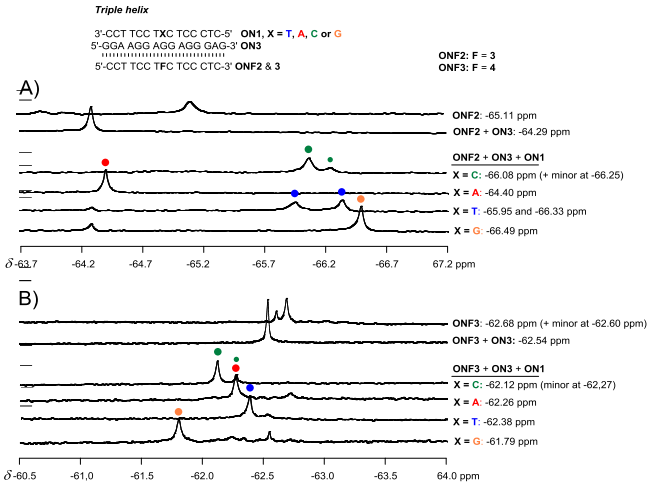


Figure 4. ^{19}F NMR spectra of the triple helices (**F*G-X**-triads, **X** = C, A, T or G). Conditions: 10 µmol L^{-1} of each ON, 10 mmol L^{-1} sodium cacodylate (pH 5.5) and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*) at 20°C.

Next, the same measurements were carried out with sensors **3** and **4** (**ONF2** and **ONF3**), which may be expected to give improved ^{19}F NMR shift reflection to the base content. **ONF2** alone gave a ^{19}F resonance signal at -65.11 ppm (Figure 3B). This signal was notably broader than that of **ONF1**. Addition of **ON2** resulted in one broad signal at -64.70 ppm. When **ON1** with variable base content was added, a signal with a distinct chemical shift was observed for each **ON1**/**ON2**/**ONF2** triplex: -66.09 ppm with the **3*A-**T triad, -67.27 ppm with the **3*A-**A triad, -66.33 ppm with the **3*A-**C triad and -67.10 ppm with the **3*A-**G triad. All the signals were relatively broad. However, all the signals could be distinguished from a mixture containing all the four **ON1**/**ON2**/**ONF2** triplexes (X = T, A, C or G in **ON1**, see Figure 5A), albeit the signals resulting from **3*A-**A and **3*A-**G overlapped partially.

ONF3 single strand gave two distinct ^{19}F resonance signals: major one at -62.68 ppm and minor one at -62.60 ppm, which was most likely originated from rotamers of **4** (Figure 3C). Addition of the duplex (**ON2** + **ON1**) gave distinct signals for the **4*A-**T triad (62.38 ppm + minor signal at -62.11 ppm), for the **4*A-**A triad (-62.13 ppm) and for the **4*A-**C triad (-62.33 ppm), but the **4*A-**A triad and the **4*A-**G triad were observed with the same chemical shift. In other words, sensor **4** was able to distinguish purine bases from pyrimidine bases in **ON1**, but not purine bases from each other. The ^{19}F NMR spectrum of a mixture containing all the four **ON1**/**ON2**/**ONF3** triplexes is shown in Figure 5B.

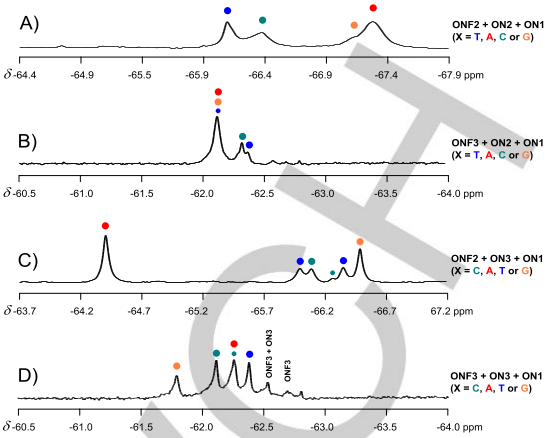


Figure 5. ^{19}F NMR spectra of the mixtures of **ON1**/**ON2**/**ONF2**, **ON1**/**ON2**/**ONF3**, **ON1**/**ON3**/**ONF2** and **ON1**/**ON3**/**ONF3** triplexes. Conditions: 40 µmol L^{-1} of **ONF2**, **ONF3**, **ON2** and **ON3**, 10 µmol L^{-1} of each **ON1**, 10 mmol L^{-1} sodium cacodylate (pH 5.5) and 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*) at 20°C.

Depending on the tautomers, sensors **3** and **4** may also bind to guanine base. Therefore, the ability of **ONF2** and **ONF3** to recognize variable base content in **ON1**/**ON3** duplex (**ON1**, X = C, A, T or G) was next evaluated (Figure 4). Interestingly, **ONF2** and **ONF3** both gave a relatively sharp signal after addition of **ON3**. When **ON1** was added to the mixture of **ONF2** and **ON3**, signals with unique chemical shifts were observed for each **ON1**/**ON3**/**ONF2** triplex (Figure 4A): -66.08 ppm and -66.25 ppm (minor) with the **3*G-**C triad, -64.40 ppm with the **3*G-**A triad, -65.95 ppm and -66.33 ppm with the **3*G-**T triad and -66.49 ppm with the **3*G-**G triad. As seen, resonances with the pyrimidine bases (C and T) split up into two signals, while resonances with purine bases gave only one relatively sharp signal. Sensor **3** (**ONF2**) targeted to AC and AT base pairs may hence adopt two relatively stable conformers resulting in different ^{19}F resonance shifts. This also disturbed the ^{19}F NMR analysis of a mixture containing all the four **ON1**/**ON3**/**ONF3** triplexes (shown in Figure 5C).

Among the other experiments, **ONF3** (sensor **4**) with **ON1**/**ON3** duplex gave the most promising ^{19}F NMR data. Distinct chemical shifts for each **ON1**/**ON3**/**ONF3** triplex (Figure 4B) could be observed: -62.12 ppm and -62.27 ppm (minor) with the **4*G-**C triad, -62.26 ppm with the **4*G-**A triad, -62.38 ppm with the **4*G-**T triad and -61.79 ppm with the **4*G-**G triad. The signals could also be distinguished from mixtures containing all the four **ON1**/**ON3**/**ONF3** triplexes (Figure 5D). When sensors **3** and **4** are compared to each other, the 4-trifluoromethylbenzoyl group (**4**) seems to give a beneficial effect for the detection. The 2,5-diaminopyridinemoiety of **3** on the triple helix (**ON1**/**ON2**/**ONF2**) is probably too prone to adopt alternative conformers (cf. Figure 2 and Figure 5C, particularly, when X = T) that disturbs the obtained ^{19}F NMR resonances.

To ascertain the temperature dependence of the chemical shifts, ^{19}F NMR spectra of the **ON1**/**ON3**/**ONF3** system were additionally recorded in different temperatures (at 20,

25, 30, 35 and 40°C, Figure S27). Passive temperature-dependent shift of the signals and gradual conversion of the triplexes eventually to the single strand (**ONF3**) may be observed in increasing temperature, but the signals could be well distinguished from each other.

Conclusions

The applicability of triplex-forming ^{19}F NMR probes to recognize nucleobase content in the pyrimidine-rich strand of DNA double helices has been evaluated. For that purpose, phosphoramidite building blocks (**6** and **7**), based on CF_3-modified *N*^{4}-(6-amino-2-pyridinyl)deoxycytidine cores, were synthesized and incorporated into triplex-forming 2'-deoxyoligonucleotide strands by an automated DNA-synthesizer. The incorporated sensors **3** and **4** on triplex-forming strands (**ONF2** and **ONF3**) reduced the stability of the triple helices, but the obtained ^{19}F NMR shift responses were, as expected, sensitive to the base-content in the pyrimidine-rich strand of the target duplexes. Unique fingerprints for the base-content were observed, particularly with **F*G-**X-triads, in which **F** = **4** and X = T, A, C or G. The described proof of concept of the triplex-forming ^{19}F NMR probes is now demonstrated with GC-rich duplexes, the targeting of which via Hoogsteen face is limited to acidic conditions (pH 5.5 used). Incorporation of appropriate base-modifications (e.g. aminopyridine residues) into TFOs may be used to increase the triplex stability in physiological pH.[15] In summary, the idea of the base-discriminating DNA triplex-forming ^{19}F NMR probes proved promising. In comparison to fluorescent-based detection methods[1-4], the sensitivity of the described ^{19}F NMR-based detection is modest, limiting its applicability e.g. for SNP, but it may find applications (and become a superior detection method) to studies, in which the goal is to simultaneously detect changes of local environments (e.g. upon binding of nucleobase-specific ligands) in structurally resembling double helical stretches.

Experimental Section

3',5'-di-*O*-acetyl-*N*^{4}-(*N*^{6}-benzoyl-4-trifluoromethyl-6-aminopyridin-2-yl)-2'-deoxycytidine (11). A mixture of 3',5'-di-*O*-acetyl-4-(1,2,4-triazol-1-yl)-2'-deoxyuridine (**8**, 0.90 g, 2.5 mmol) and 2,4-diamino-4-trifluoromethylpyridine (1.0 g, 5.7 mmol) in dry pyridine (5.0 mL) was refluxed for three days. Pyridine was removed by evaporation, the residue was dissolved in chloroform and washed with water. The organic layer was dried over Na_2SO_4, evaporated to dryness and the residue was filtered through a short silica gel column (5% Et_3N, 10% MeOH in DCM). The fractions of the transamination product (**9**) were evaporated to dryness and co-evaporated with dry pyridine. The residue (**9**, 0.35 g, contained some impurities) was dissolved in dry pyridine (3.0 mL) and benzoyl chloride (0.10 mL, 0.88 mmol) was added to the mixture at 0°C. The mixture was allowed to warm up to room temperature, stirred for 2 hours and evaporated to dryness. The residue was dissolved in chloroform and washed with water. The organic layer was dried over Na_2SO_4, filtered and evaporated to dryness. The residue was purified by silica gel chromatography (20% hexane in EtOAc) to give 0.33 g (23% overall yield from **8**) of the product **11** as yellowish foam. ^{1}H NMR (400 MHz, DMSO-*d*$_6$): *δ* 8.25 (b, 1H), 8.14 (b, 1H), 7.92 (m, 2H), 7.83 (b, 1H), 7.54 (m, 1H), 7.46 (m, 2H), 6.40 (b, 1H), 6.16 (b, 1H), 5.20 (b, 1H), 4.35 – 4.22 (m, 3H), 2.59 (m, 1H), 2.19 (m, 1H), 2.06 (b, 6H); ^{13}C NMR (100 MHz, DMSO-*d*$_6$): *δ* 170.8, 170.6, 167.0, 162.5, 156.0, 151.5, 151.4, 141.6 (q, *J* = 32.1 Hz), 141.1, 133.8, 132.1, 128.3, 127.5, 123.0 (q, *J* = 272 Hz), 105.8, 104.4, 96.6, 86.8, 82.7, 74.6, 63.6, 37.7, 19.5, 19.4; HRMS (ESI-TOF): *m/z* calcd for $C_{26}H_{25}F_3N_5NaO_7$: 598.1520 $[M+Na]^+$; found 598.1539.

3',5'-di-*O*-acetyl-*N*^{4}-[*N*^{6}-(4-trifluoromethylbenzoyl)-6-aminopyridin-2-yl]-2'-deoxycytidine (14). 3',5'-di-*O*-acetyl-N^4-(6-aminopyridin-2-yl)-2'-deoxycytidine (**10**, 0.70 g, 1.7 mmol) was synthesized from **8** as previously described and dissolved in dry DMF (7.0 mL). 4-trifluoromethylbenzoic acid (0.99 g, 5.2 mmol), PyBOP (2.2 g, 4.2 mmol) and DIEA (1.2 mL, 7.0 mmol) were added and the mixture was stirred overnight at ambient temperature. The mixture was poured to saturated aqueous $NaHCO_3$ and the product was extracted with ethyl acetate. The organic layers were combined, dried over Na_2SO_4, filtered and evaporated to dryness. The residue was purified by silica gel chromatography (EtOAc) to give 0.85 g (85% from **10**) of the product **14** as yellowish foam. ^{1}H NMR (400 MHz, DMSO-*d*$_6$): *δ* 8.13 (d, 2H, *J* = 8.2 Hz), 7.96 (b, 1H), 7.92 (m, 1H), 7.91 (d, 1H, *J* = 8.2 Hz), 7.85 (d, 2H, *J* = 8.3 Hz), 7.80 (dd, 1H, *J* = 8.1 Hz & 8.0 Hz), 6.62 (b, 1H), 6.25 (dd, 1H, *J* = 7.9 Hz & 5.9 Hz), 5.27 (ddd, 1H, *J* = 6.5 Hz, 2.1 Hz & 1.9 Hz), 4.42 – 4.33 (m, 3H), 2.65 (ddd, 1H, *J* = 14.4 Hz, 5.8 Hz & 2.8 Hz), 2.29 (m, 1H), 2.11 (s, 3H), 2.09 (s, 3H); ^{13}C NMR (100 MHz, DMSO-*d*$_6$): *δ* 170.9, 170.7, 165.8, 162.5, 156.5, 150.6, 150.1, 141.0, 139.9, 138.2, 133.1 (q, *J* = 32.2 Hz), 128.1, 125.3, 123.9 (q, *J* = 278 Hz), 110.6, 109.3, 96.6, 87.0, 82.7, 74.6, 63.7, 37.7, 19.4, 19.3; HRMS (ESI-TOF): m/z calcd for $C_{26}H_{24}F_3N_5NaO_7$: 598,1520 $[M + Na]^+$; found 598.1500.

***N*^{4}-(*N*^{6}-benzoyl-4-trifluoromethyl-6-aminopyridin-2-yl)-5'-*O*-(4,4'-dimethoxytrityl)- -2'-deoxycytidine (13).** Compound **11** (0.15 g, 0.26 mmol) was dissolved in methanol (0.6 mL) and concentrated ammonia (0.3 mL) was carefully added. Once the acetyl removal was completed, the mixture was evaporated to dryness and the residue was dissolved in ethyl acetate. The mixture was washed with water, dried over $NaSO_4$, filtered and evaporated to dryness. The residue was vacuum-dried over phosphorus pentoxide, dissolved in dry pyridine (3.0 mL) and 4,4'-dimethoxytrityl chloride (88 mg, 0.26 mmol) was

added. The mixture was stirred over night at room temperature, concentrated, dissolved in ethyl acetate and washed with water. The organic layer was separated, dried over Na_2SO_4, filtered and evaporated to dryness. The residue was purified by silica gel chromatography (10% MeOH in $CHCl_3$) to give 0.12 g (60% from **11**) of the product **13** as yellowish oil. ^{1}H NMR (500 MHz, $CDCl_3$): (rotamers observed) δ 13.45 and 10.06 (2 × s, 1H), 8.71 and 7.69 (2 × b, 1H), 8.71 and 8.27 (2 × s, 1H), 8.08 and 7.62 (2 × d, 1H, J = 7.5 and 8.1 Hz), 7.97 and 7.81 (2 × d, 1H, J = 7.5 and 7.6 Hz), 7.54 – 7.42 (m, 3H), 7.39 – 7.15 (m, 9H), 6.85 – 6.79 (m, 4H), 6.40 and 5.65 (m and d, 1H, J = 7.8 Hz), 6.40 and 6.27 (m and dd, 1H, J = 5.6 and 5.7 Hz), 4.52 and 4.45 (2 × m, 1H), 4.07 and 4.06 (2 × m, 1H), 3.79, 3.78, 3.71 and 3.70 (4 × s, 6H), 3.47 – 3.35 (m, 2H), 2.65, 2.44 and 2.25 (3 × m, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): (rotamers observed) δ 167.2, 165.8, 161.9, 159.8, 158.7, 158.6, 158.4, 155.4, 151.8, 151.6, 150.8, 150.2, 150.1, 144.9, 144.4, 144.3, 142.7, 142.4, 136.5, 136.2, 135.5, 135.3, 135.2, 134.2, 133.6, 132.5, 132.2, 130.1, 129.1, 128.8, 128.6, 128.2, 128.1, 128.0, 128.0, 127.9, 127.6, 127.3, 127.2, 127.0, 128.8 and 126.6 (2 × q, J = 284 Hz), 113.3, 113.2, 106.4, 105.4, 105.2, 104.7, 95.4, 87.0, 86.8, 86.5, 86.3, 86.2, 85.0, 71.8, 70.7, 63.2, 62.7, 55.3, 55.2, 41.9, 40.9; HRMS (ESI-TOF): *m/z* calcd for $C_{43}H_{39}F_3N_5O_7$: 794,2796 $[M + H]^+$; found 794.2800.

5'-*O*-(4,4'-dimethoxytrityl)-*N*^{4}-[*N*^{6}-(4-trifluoromethylbenzoyl)-6-aminopyridin-2-yl]-2'-deoxycytidine (16). Compound **16** was synthesized from **14** as described for **13** from **11**. 0.2 g (0.34 mmol) of **14** gave 0.17 g (62% from **14**) of the product (**16**) as white foam. ^{1}H NMR (400 MHz, DMSO-d_6): δ 8.12 (d, 1H, J = 7.3 Hz), 8.06 (d, 2H, J = 8.1 Hz), 7.96 (b, 1H), 7.87 (d, 1H, J = 8.1 Hz), 7.76 (d, 2H, J = 8.3 Hz), 7.72 (dd, 1H, J = 8.1 Hz, both), 7.40 (m, 2H), 7.30 – 7.15 (m, 7H), 6.84 – 6.81 (m, 4H), 6.21 (b, 1H), 6.21 (dd, 1H, J = 6.0 Hz & 5.8 Hz), 4.50 (m, 1H), 4.05 (m, 1H), 3.70 (s, 3H), 3.69 (s, 3H), 3.41 – 3.31 (m, 2H), 2.52 (m, 1H), 2.28 (m, 1H); ^{13}C NMR (100 MHz, DMSO-d_6): δ 166.8, 163.7, 160.1, 157.9, 152.0, 151.3, 145.9, 143.0, 141.2, 139.4, 137.0, 136.9, 136.8, 134.5, 133.0 (q, J = 32.2 Hz), 129.4, 129.3, 128.9, 127.9, 126.6, 126.6, 126.5, 126.5, 123.9 (q, J = 270 Hz), 114.2, 111.9, 110.5, 97.5, 88.2, 88.0, 87.8, 71.6, 64.1, 55.7, 42.6; HRMS (ESI-TOF): *m/z* calcd for $C_{43}H_{39}F_3N_5O_7$: 794,2796 $[M + H]^+$; found 794.2828.

***N*^{4}-(*N*^{6}-benzoyl-4-trifluoromethyl-6-aminopyridin-2-yl)-3'-*O*-[(2-cyanoethoxy)(*N,N*-diisopropylamino)phosphinyl]-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine (6).** 2-Cyanoethyl-*N,N*-diisopropylphosphoramidochloridite (62 mg, 0.26 mmol) was added to a mixture of **13** (0.16 g, 0.20 mmol) and triethylamine (0.1 mL) in dichloromethane (2.0 mL). The mixture was stirred under nitrogen for 2h and eluted through a short silica gel column (5% Et_3N, 50-95% EtOAc in hexane) to give 0.19 g (95%) of the product (**6**) as yellowish oil. ^{1}H NMR (500 MHz, CD_3CN): (a mixture of diastereomers) δ 9.15 (b, 1H), 8.83 (b, 1H), 8.21 (b, 1H), 8.19 (s, 1H), 8.02 and 7.95 (2 × d, 1H, J = 7.5 Hz, both), 7.92 – 7.89 (m, 2H), 7.59 (m, 1H), 7.50 – 7.42 (m, 4H), 7.34 – 7.18 (m, 9H), 6.86 – 6.83 (m, 4H), 6.39 (b, 1H), 6.21 and 6.18 (2 × dd, J = 6.1 and 6.0 Hz), 4.64 – 4.55 (m, 1H), 4.16 – 4.09 (m 1H), 3.78 – 3.53 (m, 4H), 3.72, 3.72 and 3.71 (3 × s, 6H), 3.41 – 3.28 (m, 2H), 2.63 – 2.51 (m, 1H), 2.62 and 2.52 (2 × dd, 2H, J = 6.0 Hz, each), 2.35 – 2.28 (m, 1H), 1.17, 1.16, 1.15, 1.14, 1.14, 1.13, 1.06 and 1.05 (8 × s, 12H); ^{13}C NMR (125 MHz, CD_3CN): (a mixture of diastereomers) δ 167.3, 163.3, 159.7, 155.7, 153.0, 152.4, 145.8, 143.4, 142.0 (q, J = 33.0 Hz), 136,7, 136.6, 134.9, 133.4, 131.1, 130.0, 129.6, 129.1, 129.0, 128.9, 128.7, 127.9, 124.1 (q, J = 271 Hz), 119.5, 119.4, 114.1, 106.7, 104.9, 96.5, 96.4, 87.6, 87.5, 86.3 (d, J = 4.2 Hz), 86.2 (d, J = 6.2 Hz), 73.8 (d, J = 16.8 Hz), 73.1 (d, J = 16.3 Hz), 63.8, 63.5, 59.5 (d, J = 14.2 Hz), 59.3 (d, J = 14.3 Hz), 55.9, 44.1, 44.1, 44.0, 44.0, 41.2, 41.0, 24.9, 24.9, 24.9, 24.8, 24.8, 24.7, 21.0, 21.0, 20.9; ^{31}P (200 MHz, CD_3CN): δ 148.1 and 148.0; ^{19}F NMR (470MHz, CD_3CN): δ -65.34; HRMS (ESI-TOF): *m/z* calcd for $C_{52}H_{56}F_3N_7O_8P$: 994,3875 $[M + H]^+$; found 994.3861.

3'-*O*-[(2-cyanoethoxy)(*N,N*-diisopropylamino)phosphinyl]-5'-*O*-(4,4'-dimethoxytrityl)-*N*^{4}-[*N*^{6}-(4-trifluoromethylbenzoyl)-6-aminopyridin-2-yl]-2'-deoxycytidine (7). Compound 7 was synthesized from **16** as described for **6** from **13**. 0.16 g (0.20 mmol) of **16** gave 0.19 g (95%) of the product **7** as white foam. ^{1}H NMR (400 MHz, $CDCl_3$): (a mixture of diastereomers) δ 8.18 – 7.55 (m, 7H), 7.46 (b, 1H), 7.41 – 7.17 (m, 9H), 6.85 – 6.78 (m, 4H), 6.36 – 6.04 (m, 2H), 4.65 (m, 1H), 4.17 (m, 1H), 3.87 – 3.37 (m, 6H), 3.74 (s, 6H), 2.77 – 2.27 (m, 4H), 1.29 – 1.05 (m, 12H); ^{13}C NMR (100 MHz, $CDCl_3$): (a mixture of diastereomers) δ 164.6, 162.3, 158.7, 155.1, 150.5, 149.8, 144.3, 142.6, 140.8, 137.5, 135.4, 133.6 (q, J = 33.0 Hz), 130.2, 128.3, 128.2, 127.9, 127.1, 125.7, 123.6 (q, J = 271 Hz), 117.4, 113.2, 111.1, 109.8, 94.5, 86.9, 86.5, 85.4, 84.9, 71.4, 71.2, 61.8, 58.3, 58.2, 55.2, 43.4, 43.2, 40.8, 24.7, 24.6, 24.5, 20.2, 20.1; ^{31}P (200 MHz, CD_3CN): δ 148.04 and 147.99; ^{19}F NMR (470MHz, CD_3CN): δ -63.41; HRMS (ESI-TOF): *m/z* calcd for $C_{52}H_{56}F_3N_7O_8P$: 994,3875 $[M + H]^+$; found 994.3873.

Synthesis of ^{19}F-labelled oligonucleotides ONF2 and ONF3

6 and **7** were dissolved in acetonitrile to provide 0.1 mol L^{-1} solutions of the phosphoramidites and loaded to reagent vessels of an automated Applied Biosystems 3400 DNA synthesizer. The commercially available building blocks (N^4-Ac-dC and T phosphoramidites) were placed to the reagent vessels as usual. **ONF2** and **ONF3** were synthesized on a 1.0 µmol scale on a standard dC-3'-*O*-yl succinate and a dC-3'-*O*-yl hydroquinone *O,O'*-diacetate loaded LCAA-CPG support, respectively. Benzylthiotetrazol was used as an activator. 20 and 600 s coupling times were used for the commercially available DNA building blocks and for **6** and

7, respectively. The chain assembly was monitored by DMTr assay that showed nearly quantitative coupling yields for **6** and **7**. After the chain assembly, the CPG supports were removed from the synthesizer and **ONF2** and **ONF3** were released using the following procedures: The CPG-support bearing **ONF2** was suspended with concentrated ammonia (for 5h at 55°C), filtered and the filtrate was evaporated to dryness. The residue was then dissolved in a mixture of 40% aqueous methylamine and concentrated ammonia (1:1, *v/v*, for 5h at 55°C) to complete the benzoyl group removal of the 4-trifluoromethyl-2,6-diaminopyridine residue (RP HPLC-monitoring of the deprotection, see Figure S26). Finally, the crude mixture of **ONF2** was evaporated to dryness and subjected to a RP HPLC purification. **ONF3** was released by suspending the CPG support to a mixture of 0.05 mol L^{-1} K_2CO_3 in methanol (for 5h at r.t.). The mixture was neutralized by addition of acetic acid, filtered, the filtrate was evaporated to dryness and the crude **ONF3** was purified by RP HPLC. The authenticity of **ONF2** and **ONF3** was verified by MS(ESI-TOF) spectroscopy (Table S1). Isolated yields, according to UV-absorbance at 260 nm, of **ONF2** and **ONF3** were 18 and 25%, respectively.

UV-melting profile analysis. The melting curves (absorbance *vs.* temperature, Table 1) were measured at 295 nm on a PerkinElmer Lambda 35 UV-Vis spectrometer equipped with a multiple cell holder and a Peltier temperature controller. The validity of the target temperature was confirmed by an internal thermometer. Temperature was changed from 10 to 90 °C at a rate of 0.5 °C/ min. The detailed conditions of the samples are described in Table 1. Each T_m value was determined as the minimum of the first derivative of the melting curve.

^{19}F NMR measurements

ONF1-ONF3 (as triethyl ammonium salts) were dissolved in 10 mmol L^{-1} sodium cacodylate (pH 5.5) containing 0.1 mol L^{-1} NaCl in D_2O-H_2O (1:9, *v/v*) to result 10 µmol L^{-1} (spectra in Figures 3 and 4) and 40 µmol L^{-1} (spectra in Figure 5) of the oligonucleotides. After the sample preparation, pH was checked, the sample was heated to 90°C, allowed to cool down to the ambient temperature and then the NMR measurement was carried out at the target temperature. Spectra were recorded at a frequency of 470.6 MHz on a Bruker Avance III HD 600 MHz spectrometer equipped with a cryogenic probe (the sensitivity ca. tenfold compared to measurements with normal probes that work at ambient temperature). The parameters were optimized to gain the signals with the longest relaxation rate.

Acknowledgements

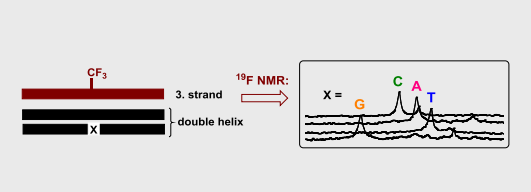
Financial support from the Academy of Finland (251539 and 256214) and from the EXPERT Sustain program is gratefully recognized.

Keywords: ^{19}F NMR spectroscopy • triplex–forming oligonucleotides • modified nucleosides

[1] A. Okamoto, K. Tainaka, Y. Ochi, K. Kanatami, I. Saito, *Mol. Biosyst*. **2006**, *2*, 122–127.
[2] M. Hausmann, G. Winkler, J. Hildenbrand, A. Finstark, A. Weisel, A. Rapp, E. Schmitt and J. C. Cremer, *Biotechniques*, 2003, **35**, 564–577.
[3] I. Van Daele, N. Bomholt, V. V. Van Calenberg and E. B. Pedersen, *ChemBioChem*, 2008, **9**, 791–801.
[4] X. Li, Y. Wang, J. Guo and X. Tang, *ChemBioChem*, 2011, **12**, 2863–2870.
[5] K. Tanabe, T. Tsuda, T. Ito and S. Nishimoto, *Chem. Eur. J.*, 2013, **19**, 15133–15140.
[6] J. Riedl, R. Pohl, L. Rulisek and M. Hocek, *J. Org. Chem.*, 2012, **77**, 1026–1044.
[7] T. Sakamoto, D. Hasegawa and K. Fujimoto, *Chem. Commun.*, 2015, **51**, 8749–8752.
[8] T. Sakamoto, D. Hasegawa and K. Fujimoto, *Analyst*, 2016, **141**, 1214–1217.
[9] Y. Hari, S. Obika and T. Imanishi, *Eur. J. Org. Chem.*, 2012, **15**, 2875–2887.
[10] C.-Y. Huang, G. Bi, P. S. Miller, *Nucleic Acids Res.* **1996**, *24*, 2606–2613.
[11] L. Granqvist and P. Virta, *J. Org. Chem.*, 2014, **79**, 3529–3536.
[12] V. Tähtinen, L. Granqvist, P. Virta, *Bioorg. Med. Chem.* **2015**, *23*, 4472–4480.
[13] R. T. Pon, S. Yu, *Nucleic Acids Res.*, 1997, **25**, 3629–3635.
[14] D. P. Arya, *Acc. Chem. Res.* **2011**, *44*, 134–146.
[15] S. Hildbrand, A. Blaser, S. P. Parel, C. J. Leumann, *J. Am. Chem. Soc.* **1997**, *119*, 5499–5511.

Entry for the Table of Contents (Please choose one layout)

FULL PAPER



N. Bhuma, V. Tähtinen, and P. Virta,**

Page No. – Page No.

Synthesis and applicability of base-discriminating DNA triplex-forming ^{19}F NMR probes

The detection of nucleobase content in DNA double helices has been demonstrated by triplex–forming ^{19}F NMR probes. For that purpose the applicability of two base-modified nucleosides, incorporated into the third strand, was evaluated. Unique ^{19}F NMR spectral finger prints, depending on the nucleobase content could be obtained, which may find application to simultaneously detect local environments of structurally resembling double helical stretches.

Key topic: ^{19}F NMR-based detection of nucleic acids

[a] Dr. N. Bhuma, M. Sc. V. Tähtinen, Prof. P. Virta
Department of Chemistry, University of Turku,
Turku 20014 (Finland)
E-mail: pamavi@utu.fi

Supporting information for this article is given via a link at the end of the document.