5-Fluoro pyrimidines: labels to probe DNA and RNA secondary structures by 1D 19F NMR spectroscopy

Barbara Puffer1, Christoph Kreutz1, Ulrike Rieder1, Marc-Olivier Ebert2, Robert Konrat3 and Ronald Micura1,*

1Institute of Organic Chemistry, Center for Molecular Biosciences (CMBI), University of Innsbruck, 6020 Innsbruck, Austria, 2Laboratory of Organic Chemistry, ETH Zürich, 8093 Zürich, Switzerland and 3Max Perutz Laboratories, Vienna Biocenter, University of Vienna, 1030 Vienna, Austria

Received June 1, 2009; Accepted September 25, 2009

ABSTRACT

19F NMR spectroscopy has proved to be a valuable tool to monitor functionally important conformational transitions of nucleic acids. Here, we present a systematic investigation on the application of 5-fluoro pyrimidines to probe DNA and RNA secondary structures. Oligonucleotides with the propensity to adapt secondary structure equilibria were chosen as model systems and analyzed by 1D 19F and 1H NMR spectroscopy. A comparison with the unmodified analogs revealed that the equilibrium characteristics of the bistable DNA and RNA oligonucleotides were hardly affected upon fluorine substitution at C5 of pyrimidines. This observation was in accordance with UV spectroscopic melting experiments which demonstrated that single 5-fluoro substitutions in double helices lead to comparable thermodynamic stabilities. Thus, 5-fluoro pyrimidine labeling of DNA and RNA can be reliably applied for NMR based nucleic acid secondary structure evaluation. Furthermore, we developed a facile synthetic route towards 5-fluoro cytidine phosphoramidites that enables their convenient site-specific incorporation into oligonucleotides by solid-phase synthesis.

INTRODUCTION

The use of fluorine labels in context with NMR spectroscopy represents an emerging tool to study structure and function of nucleic acids (1–3). In contrast to 1H NMR spectroscopy, the problem of resonance degeneracy is practically absent for 19F NMR spectroscopy. Together with the option to site-specifically incorporate fluorine labels into DNA and RNA and thus to avoid time-consuming signal assignment procedures, these advantages can be exploited to monitor conformational changes and folding of nucleic acids or to investigate binding processes in small molecule/nucleic acid or enzyme/nucleic acid complexes. This was demonstrated with short oligonucleotides (4–17) but also with transfer RNAs (18,19), ribozymes (20), RNA aptamers (21), or with substrates of DNA methyltransferases (22), to mention a few examples.

Concerning the position of labeling, fluorine modifications at both the ribose and the nucleobase unit come into consideration (Figure 1). The majority of 19F NMR studies on nucleic acids have involved 5-fluoro uracil so far which is easily accessible as nucleoside phosphoramidite either for solid phase synthesis of DNA or RNA (6,23), or as nucleoside triphosphate derivative for enzymatic RNA synthesis by T7 polymerase (18,19,24,25). Significantly less studies deal with 5-fluoro cytosine in nucleic acids (22,24–27). A comprehensive and comparative study on the usage of both 5-fluoro uracil and 5-fluoro cytosine as potential labels for NMR based DNA and RNA secondary structure probing (28) is lacking to date and represents a major goal here. This includes a critical assessment of the impact of these fluorine modifications on DNA and RNA base pairing properties. For such an undertaking, we considered short bistable RNA and DNA oligonucleotides to be the most appropriate systems (29–34). Bistable nucleic acids are highly sensitive with respect to single nucleotide mutations or chemical modifications since they directly reflect the interference by a change of the secondary structure equilibrium position. We previously took advantage of this sensitivity to explore the influence of nucleobase methylations and to control switching of secondary structures for a series of RNAs (30,31). Others extended the latter concept by using photolabile chemical modifications to study photochemically induced RNA folding by real time NMR spectroscopy (33,35,36) and

*To whom correspondence should be addressed. Tel: +43 512 507 5210; Fax: +43 512 507 2892; Email: ronald.micura@uibk.ac.at
other biophysical methods (37–41). In this sense, the use of short bistable oligonucleotides will directly reveal the potential influence of 5-fluoro pyrimidine labeling on secondary structure formation—if there is any.

**MATERIALS AND METHODS**

**Synthesis of 5-fluoro cytidine phosphoramidites (5 and 9)**

**General.** $^1$H, $^{13}$C and $^{31}$P NMR spectra were recorded on a Bruker DRX 300 MHz, Avance II + 600 MHz or Varian Unity 500 MHz instrument. The chemical shifts are referenced to the residual proton signal of the deuterated solvents: CDCl$_3$ (7.26 ppm), d$_6$-DMSO (2.49 ppm) for $^1$H NMR spectra; CDCl$_3$ (77.0 ppm) or dimethoxytritylchloride (143 mg, 0.421 mmol) in three portions over a period of 1 h. The reaction mixture was stirred for 14 h at room temperature, evaporated and coevaporated with toluene and dichloromethane. The crude product was then suspended in 5 ml diethyl ether, filtered and washed thoroughly with ether. Yield: 144 mg of 2 as white powder (60%). TLC (CH$_2$Cl$_2$/CH$_3$OH, 99/1–97/3 v/v): $R_f$ = 0.50; $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 2.64 (s, 3H, COCH$_3$); 3.34 [dd, $J = 2.4, 11.1$ Hz, 1H, H1–C(5’)]; 3.40 [dd, $J = 2.4, 11.1$ Hz, 1H, H2–C(5’)]; 3.58 [s, br, 1H, HO–C(3’)]; 3.78, 3.79 (2s, 6H, 2x OCH$_3$); 4.40 [m, 1H, H–C(4’)]; 4.48 [m, 1H, H–C(4’)]; 4.53 [m, 1H, H–C(2’)]; 5.62 [s, br, 1H, HO–C(2’)]; 5.79 [dd, $J = 3.9$ Hz, 1H, H–C(1’)]; 6.82 [m, 4H, H–C(ar)]; 7.25 [m, 9H, H–C(ar)]; 8.02 [s, br, 1H, HN–C(4’)]; 8.11 [dd, $J = 5.7$ Hz, 1H, H–C(6’)] ppm; $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 31.53 (2$^1$H$_3$); 60.00 [C(5’)]; 68.81, 75.11 [C(3’), C(2’)]; 84.67 [C(4’)]; 90.92 [C(1’)]; 130.39; 135.99; 139.21; 153.03; 153.91; 170.12 (COCH$_3$) ppm; UV/Vis (MeOH): $\lambda$ (\textit{e}) = 260 (5100 nm (mol$^{-1}$ dm$^{-3}$ cm$^{-1}$); ESI-MS (m/z): [M + H]$^+$ calculated for C$_{11}$H$_{14}$FN$_2$O$_6$, 304.25; found 304.09.

$^{N^4}$-acetetyl-5’-O-(4,4’-dimethoxytrityl)-5-fluoro cytidine (3). To a suspension of $^{N^4}$-acetetyl-5-fluoro cytidine 2 (121 mg, 0.383 mmol) in pyridine (1.5 ml) was added 4,4’-dimethoxytritylchloride (143 mg, 0.421 mmol) and silver nitrate (285 mg, 1.679 mmol) was stirred in THF (3.9 ml) and pyridine (0.4 ml) overnight at room temperature. Then, tert.-butyldimethylchlorosilane (253 mg, 1.679 mmol) was added and stirring was continued for 1 h. After evaporation of the solvent, the product was dissolved in dichloromethane and precipitated with ether to give a white precipitate. Yield: 146.1 mg of 4a. TLC (CH$_2$Cl$_2$/CH$_3$OH, 99/1–97/3 v/v): $R_f$ = 0.50; $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 2.64 (s, 3H, COCH$_3$); 3.34 [dd, $J = 2.4, 11.1$ Hz, 1H, H1–C(5’)]; 3.40 [dd, $J = 2.4, 11.1$ Hz, 1H, H2–C(5’)]; 3.58 [s, br, 1H, HO–C(3’)]; 3.78, 3.79 (2s, 6H, 2x OCH$_3$); 4.40 [m, 1H, H–C(4’)]; 4.48 [m, 1H, H–C(4’)]; 4.53 [m, 1H, H–C(2’)]; 5.62 [s, br, 1H, HO–C(2’)]; 5.79 [dd, $J = 3.9$ Hz, 1H, H–C(1’)]; 6.82 [m, 4H, H–C(ar)]; 7.25 [m, 9H, H–C(ar)]; 8.02 [s, br, 1H, HN–C(4’)]; 8.11 [dd, $J = 5.7$ Hz, 1H, H–C(6’)] ppm; $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 31.53 (2$^1$H$_3$); 60.00 [C(5’)]; 68.81, 75.11 [C(3’), C(2’)]; 84.67 [C(4’)]; 90.92 [C(1’)]; 130.39; 135.99; 139.21; 153.03; 153.91; 170.12 (COCH$_3$) ppm; UV/Vis (MeOH): $\lambda$ (\textit{e}) = 260 (5100 nm (mol$^{-1}$ dm$^{-3}$ cm$^{-1}$); ESI-MS (m/z): [M + H]$^+$ calculated for C$_{11}$H$_{14}$FN$_2$O$_6$, 304.25; found 304.09.

$^{N^4}$-acetetyl-2’-O-(tert.-butyldimethylsilyl)-5’-O-(4,4’-dimethoxytrityl)-5-fluoro cytidine (4b). Compound 3 (565 mg, 0.933 mmol) and silver nitrate (285 mg, 1.679 mmol) was stirred in THF (3.9 ml) and pyridine (0.4 ml) overnight at room temperature. Then, tert.-butyldimethylchlorosilane (253 mg, 1.679 mmol) was added and stirring was continued for 1 h. After evaporation of the solvent, the product was dissolved in dichloromethane and precipitated with ether to give a white precipitate. Yield: 146.1 mg of 4b. TLC (CH$_2$Cl$_2$/CH$_3$OH, 99/1–97/3 v/v): $R_f$ = 0.50; $^1$H NMR (300 MHz, CDCl$_3$): $\delta$ 2.64 (s, 3H, COCH$_3$); 3.34 [dd, $J = 2.4, 11.1$ Hz, 1H, H1–C(5’)]; 3.40 [dd, $J = 2.4, 11.1$ Hz, 1H, H2–C(5’)]; 3.58 [s, br, 1H, HO–C(3’)]; 3.78, 3.79 (2s, 6H, 2x OCH$_3$); 4.40 [m, 1H, H–C(4’)]; 4.48 [m, 1H, H–C(4’)]; 4.53 [m, 1H, H–C(2’)]; 5.62 [s, br, 1H, HO–C(2’)]; 5.79 [dd, $J = 3.9$ Hz, 1H, H–C(1’)]; 6.82 [m, 4H, H–C(ar)]; 7.25 [m, 9H, H–C(ar)]; 8.02 [s, br, 1H, HN–C(4’)]; 8.11 [dd, $J = 5.7$ Hz, 1H, H–C(6’)] ppm; $^{13}$C NMR (75 MHz, CDCl$_3$): $\delta$ 31.53 (2$^1$H$_3$); 60.00 [C(5’)]; 68.81, 75.11 [C(3’), C(2’)]; 84.67 [C(4’)]; 90.92 [C(1’)]; 130.39; 135.99; 139.21; 153.03; 153.91; 170.12 (COCH$_3$) ppm; UV/Vis (MeOH): $\lambda$ (\textit{e}) = 260 (5100 nm (mol$^{-1}$ dm$^{-3}$ cm$^{-1}$); ESI-MS (m/z): [M + H]$^+$ calculated for C$_{11}$H$_{14}$FN$_2$O$_6$, 304.25; found 304.09.
1 h. The resulting white suspension was filtered, the solvents were purified by column chromatography on SiO₂ (hexane/ethylacetate, 6/4 – 5/5 v/v). Yield: 325 mg of 4a (48%) and 95 mg of 4b (14%) as white foam. TLC (hexane/ethylacetate, 3/7): Rf (4a) = 0.50, Rf (4b) = 0.24.

1H NMR (300 MHz, CDCl₃): δ = 0.23 (s, 3H, SiCH₃); 0.34 (s, 3H, SiCH₃); 0.96 [s, 9H, Si(CH₃)₃]; 2.41 [d, J = 9.3 Hz, 1H, HO–C(3)]; 2.68 (s, br, 3H, COCH₃); 3.57 [m, 2H, H₂–C(5)]; 3.80 [s, 6H, 2 × OCH₃]; 4.12 [m, 1H, H–C(4)]; 4.40 [m, 1H, H–C(2)]; 4.45 [m, 1H, H–C(3)]; 5.77 [s, 1H, H–C(1)]; 6.85 [d, J = 8.7 Hz, 4H, H–C(2); H–C(3)]; 7.23–7.44 [m, 9H, H–C(ar)]; 7.81 [s, br, 1H, HN–C(4)]; 8.38 [d, J = 4.5 Hz, 1H, H–C(6)] ppm; 4b: δ = 0.00 (s, 3H, SiCH₃); 0.09 (s, 3H, SiCH₃); 0.87 [s, 9H, Si(CH₃)₃]; 2.71 (s, br, 3H, COCH₃); 3.13 [s, br, 1H, HO–C(2)]; 3.33 [dd, J = 3.0, 9.0 Hz, 1H, H1–C(5)]; 3.58 [m, 1H, H–C(2)]; 3.82 (s, 6H, 2 × OCH₃); 4.18 [m, 1H, H–C(4)]; 4.22 [m, 1H, H–C(2)]; 4.37 [m, 1H, H–C(1)]; 5.96 [d, J = 3.0 Hz, 1H, H1–C(5)]; 6.86 [d, J = 9.0 Hz, 4H, H–C(ar)]; 7.25–7.40 [m, 9H, H–C(ar)]; 7.58 [s, br, 1H, HN–C(4)]; 8.25 [m, 1H, H–C(6)] ppm; 13C NMR (75 MHz, CDCl₃): 4a: δ = 5.23, -4.19 (2 × SiCH₃); 18.22 [Si(CH₃)₂]; 25.98 [Si(CH₃)₂]; 26.43 (COCH₃); 55.37 (2 × OCH₃); 61.32 [C(5)]; 69.22 [C(3)]; 67.57 [C(2)]; 83.45 [C(4)]; 87.32, 91.09 [C(1)]; 113.47, 133.51 [C(ar)]; 127.21 [C(ar)]; 128.12, 128.16 [C(ar)]; 128.46, 130.13, 130.19 [C(ar)]; 135.23, 135.48, 144.43 [C(ar)]; 152.66, 152.81; 158.83; 158.87; 171.64 (COCH₃) ppm; UV/Vis (MeOH): λ (ε) = 260 (5500 nm) (mol⁻¹ dm³ cm⁻¹); ESI-MS (m/z): [M + H]⁺ calculated for C₁₅H₁₄F₃N₃O₆Si, 366.32; found 366.03.

N⁴-benzoyl-5-fluoro cytidine (6). A suspension of 5-fluoro cytidine 1 (100 mg, 0.383 mmol) in anhydrous ethanol (4 mL) was treated with benzoic anhydride (1.04 g, 4.594 mmol) and stirred for 14 h at room temperature. After evaporation of the solvent, the product was coevaporated with diethyl ether and dried under high vacuum. The crude product was then suspended in 2 mL diethyl ether, filtered and washed thoroughly with ether. Yield: 134 mg of 6 as white powder (99%). TLC (CH₂Cl₂/CH₃OH, 90/10): Rf = 0.40; 1H NMR (300 MHz, DMSO): δ = 3.63 [m, 1H, H1–C(5)]; 3.74 [m, 1H, H2–C(5)]; 3.91 [m, 1H, H–C(4)]; 4.04 [m, 2H, H–C(3) + H–C(2)]; 5.05 [m, 1H, HO–C(2) or HO–C(3)]; 5.32 [m, 1H, HO–C(5)]; 5.51 [m, 1H, HO–C(3) or HO–C(2)]; 5.74 [s, 1H, H–C(1)]; 7.51 [m, 2H, H–C(ar)]; 7.62 [m, 1H, H–C(ar)]; 7.99 [m, 2H, H–C(ar)]; 8.61 [m, 1H, H–C(6)]; 11–12.5 [s, br, 1H, HN–C(4)] ppm; 13C NMR (75 MHz, DMSO): δ = 60.30 [C(5)]; 69.21, 74.90 [C(3), C(2)]; 85.00 [C(4)]; 90.37 [C(1)]; 129.11–129.85 [C(ar), C(6)]; 133.47 [C(ar)]; 137.59: 140.73 ppm; UV/Vis (MeOH): λ (ε) = 260 (1300 nm) (mol⁻¹ dm³ cm⁻¹); ESI-MS (m/z): [M + H]⁺ calculated for C₁₅H₁₄F₃N₃O₆Si, 366.32; found 366.03.

N⁴-benzoyl-5′-O-(4,4′-dimethoxytrityl)-5-fluoro cytidine (7). To a suspension of N⁴-benzoyl-5-fluoro cytidine 6 (247 mg, 0.676 mmol) in pyridine (2.6 mL) was added 4,4′-dimethoxytritylchloride (252 mg, 0.744 mmol) in three portions over a period of 1 h. The reaction mixture was stirred for 14 h at room temperature, evaporated and coevaporated with toluene and dichloromethane. The crude product was purified by column chromatography on SiO₂ (CH₂Cl₂/CH₃OH, 99/1 – 97/3 v/v). Yield: 251 mg of 7 as white foam (56%). TLC (CH₂Cl₂/CH₃OH, 9:1): Rf = 0.60; 1H NMR (300 MHz, CDCl₃): δ = 1.85 [s, br, 1H, HO–C(2)]; 3.42 [m, 2H, H₂–C(5)]; 3.61 [s, br, 1H, HO–C(3)]; 3.76, 3.77 [2s, 6H, 2 × OCH₃]; 4.37 [m, 1H, H–C(4)]; 4.48 [m, 1H, H–C(3)]; 4.57 [m, 1H, H–C(2)]; 5.85 [d, J = 3.6 Hz, 1H, H1–C(5)]; 6.84 [m, 4H, H–C(ar)]; 7.19–7.43 [m, 9H, H–C(ar)]; 7.48 [m, 2H, H–C(ar)]; 7.57 [m, 1H, H–C(ar)]; 8.07 [m, 1H, H–C(6)]; 8.16 [m, 2H, H–C(ar)] ppm; 13C NMR (75 MHz, CDCl₃): δ = 55.20 (2 × OCH₃); 62.89 [C(5)]; 71.76 [C(3)]; 76.68 [C(2)]; 85.71 [C(4)]; 87.40; 92.39 [C(1)]; 113.33, 113.36 [C(ar)]; 127.08 [C(ar)]; 127.88, 128.03 [C(ar)]; 128.48 [C(ar)]; 129.83, 129.98 [C(ar)]; 133.13 [C(ar)]; 135.02, 135.37 [C(ar)]; 144.08 [C(ar)]; 158.67, 158.70 [C(ar)] ppm; UV/Vis (MeOH): λ (ε) = 260 (2200 nm) (mol⁻¹ dm³ cm⁻¹); ESI-MS (m/z): [M + H]⁺ calculated for C₃₇H₃₄F₃N₃O₈, 668.69; found 668.15.

N⁴-benzoyl-2′-O-(tert-butyldimethylsilyl)-5′-O-(4,4′-dimethoxytrityl)-5-fluoro cytidine (8a) and N⁴-benzoyl-3′-O-(tert-butyldimethylsilyl)-5′-O-(4,4′-dimethoxytrityl)-5-fluoro cytidine (8b). Compound 7 (251 mg, 0.376 mmol) and silver nitrate (115 mg, 0.677 mmol) was stirred in THF (1.5 mL) and pyridine (0.15 mL) overnight at room
temperature. Then, tert.-butyldimethylchlorosilane (102 mg, 0.677 mmol) was added and stirring was continued for 1 h. The resulting white suspension was filtered, the solvents were evaporated and the residue was coevaporated with dichloromethane. The crude product was purified by column chromatography on SiO2 (hexane/ethylacetate, 8/2 – 7/3 v/v). Yield: 171 mg of 8a (58%) and 21 mg of 8b (7%) as white foam. TLC (hexane/ethyl acetate, 7/3): Rf (8a) = 0.50, Rf (8b) = 0.26.

1H NMR (300 MHz, CDCl3): 8a: δ 0.20 (s, 3H, SiCH3); 0.21 (s, 3H, SiCH3); 0.96 [s, 9H, Si(CH3)3]; 2.64 [m, 1H, HO–C(3)]; 3.52 [m, 2H, 2H–C(5)]; 3.82 (s, 6H, 2× OCH3); 4.18 [m, 1H, H–C(4)]; 4.39 [m, 1H, H–C(3)]; 4.46 [m, 1H, H–C(2)]; 5.98 [d, J = 2.7 Hz, 1H, H–C(1)]; 6.88 [d, J = 8.1 Hz, 4H, H–C(ar)]; 7.23 – 7.49 [m, 11H, H1–C(ar)]; 7.57 [m, 1H, H–C(ar)]; 8.11 [d, J = 5.7 Hz, 1H, H–C(6)]; 8.29 [d, J = 7.2 Hz, 2H, H–C(ar)]; 13.04 [s, 1H, HN–C(4)]; 0.09 (s, 3H, SiCH3); 0.88 [s, 9H, Si(CH3)3]; 2.84 [m, 1H, HO–C(2)]; 3.36 [dd, J = 2.7, 11.1 Hz, 1H, H1–C(5)]; 3.51 [dd, J = 2.1, 11.1 Hz, 1H, H2–C(5)]; 3.80 (s, 6H, 2× OCH3); 4.25 [m, 1H, H–C(4)]; 4.33 [m, 1H, H–C(3)]; 5.98 [d, J = 3.0 Hz, 1H, H–C(1)]; 6.86 [d, J = 8.7 Hz, 4H, H–C(ar)]; 7.25 – 7.48 [m, 11H, H1–C(ar)]; 7.55 [m, 1H, H–C(ar)]; 8.00 [d, J = 5.7 Hz, 1H, H–C(6)]; 8.28 [d, J = 7.5 Hz, 2H, H–C(ar)]; 13.01 [s, br, 1H, HN–C(4)]; ppm; 13C NMR [75 MHz, CDCl3]: 8a: δ 5.02, –4.50 (2× SiCH3); 18.17 [Si(CH3)3]; 25.82 [Si(CH3)3]; 55.41 (2× OCH3); 62.68 [C(5)]; 70.93 [C(3)]; 76.59 [C(2)]; 84.07 [C(4)]; 87.54, 89.35 [C(1)]; 113.55, 133.58 [C(ar)]; 125.68, 127.33 [C(ar)]; 127.92, 127.98, 128.09, 128.24, 128.45, 129.28, 130.18, 133.17 [C(ar)]; 135.12, 135.39; 144.36; 152.60; 152.85; 158.91, 158.93 [C(ar)]; ppm; UV/Vis (MeOH): λ (ε) = 260 (2400) nm (mol1 dm3 cm−1); ESI-MS (m/z): [M + H]+ calculated for C51H60FNSi, 773.17; found 772.12.

Solid-phase synthesis of oligonucleotides

2′-O-TOM standard nucleoside phosphoramidites were obtained from GlenResearch or ChemGenes. 2′-O-tert.-Butyldimethylsilyl-5′-O-(4,4′-dimethoxytrityl)-5-fluoro uridine phosphoramidite was purchased from ChemGenes, 5′-O-(4,4′-dimethoxytrityl)-5-fluoro 2′-deoxy-uridine phosphoramidite and 5′-O-(4,4′-dimethoxytrityl)-5-fluoro-O2′-(2,4,6-trimethylphenyl) 2′-deoxyuridine phosphoramidite were obtained from GlenResearch. All solid supports for DNA and RNA synthesis were purchased from GE Healthcare (Custom Primer Supports: riboC Ac 80, riboA 80, riboG 80, riboU 80; daA 80s, dcA 80s, dgA 80s, T 80s). All oligonucleotides were synthesized on Pharmacia instruments (Gene Assembler Plus) following DNA/RNA standard methods; detritylation (2.0 min): dichloroacetic acid/1,2-dichloroethane (4:96); coupling (3.0 min): phosphoramidites/acetanilide (0.1 M × 120 µl) were activated by benzylthiotriazole/acetanilide (0.35 M × 360 µl); capping (3 × 0.4 min): A: Ac2O/sym-collidine/acetanilide (20/30/50); B: 4-(dimethylamino)pyridine/acetonitrile (0.5 M), A:B = 1/1; oxidation (1.0 min): I2 (10 mM) in acetonitrile; ESI-MS (m/z): [M + H]+ calculated for C51H60FNSi, 773.17; found 772.12.

Deprotection of oligonucleotides

DNA oligonucleotides were deprotected by treatment with ammonia in water (32%) and ethanol (final volume 1 ml, aqueous ammonia/ethanol 3/1, v/v) at 55°C for 16 h. The solution was evaporated to dryness and the crude oligonucleotide was dissolved in 1.0 ml of water. RNA oligonucleotides (including those containing N4-acetyl protected C5) were deprotected by using MeNH2 in EtOH (8 M, 0.60 ml) and MeCNH3 in H2O (40%, 0.60 ml) at room temperature for 5–6 h. After complete evaporation of the 2′-O-TOM protecting groups were removed by treatment with tetrabutylammonium fluoride trihydrate (TBAF⇌3H2O) in THF (1 M, 0.95 ml) for at least 12 h at 37°C. The reaction was quenched by addition of triethylammonium acetate (TEAA) (1 M, pH 7.0, 0.95 ml). The volume of the solution was reduced to 1 ml and the solution was loaded on a GE Healthcare HiPrep 26/10 Desalting column (2.6 × 10 cm; Sephadex G25). The crude RNA was eluted with H2O, evaporated to dryness and dissolved in 1.0 ml of water.
Analysis and purification of oligonucleotides

Analysis of crude oligonucleotides after deprotection was performed by anion-exchange chromatography on a Dionex DNAPac100 column (4 x 250 mm) at 80°C. Flow rate: 1 ml min⁻¹; eluant A: 25 mM Tris–HCl (pH 8.0), 6 M urea; eluant B: 25 mM Tris–HCl (pH 8.0), 0.5 M NaClO₄, 6 M urea; gradient: 0–40% B in A within 30 min (<20 nt) or 0–60% B in A within 45 min (>20 nt); UV-detection at 265 nm. Crude RNA products (trityl-off) were purified on a semipreparative Dionex DNAPac100 column (9 x 250 mm). Flow rate: 2 ml min⁻¹; gradient: ΔS-10% B in A within 20 min. Fractions containing oligonucleotide were loaded on a C18 SepPak cartridge (waters/millipore), washed with 0.1–0.2 M (Et₃NH)⁺HCO₃⁻ and H₂O, eluted with H₂O/CH₃CN and lyophilized to dryness. The purified oligonucleotides were characterized by mass spectrometry on a Finnigan LCQ Advantage MAX ion trap instrumentation connected to an Amersham Ettan micro LC system (negative-ion mode with a potential of −4 kV applied to the spray needle). LC: Sample (250 pmol of oligonucleotide dissolved in 20 μl of 20 mM EDTA solution; average injection volume: 10–20 μl); column (Amersham μRPC C2/C18; 2.1 x 100 mm) at 21°C; flow rate: 100 μl min⁻¹; eluant A: 8.6 mM TEA, 100 mM 1,1,1,3,3,3-hexaoxypropane, eluted with 40% CH₃CN; flow rate: 1 ml min⁻¹; eluant B: 50 mM TEA, 100 mM 1,1,1,3,3,3-hexaoxypropane, eluted with 100% CH₃CN; gradient: 0–100% B in A within 30 min; UV-detection at 254 nm.

NMR spectroscopy

19F NMR spectra with 1H-decoupling were recorded at a frequency of 564.7 MHz on a Bruker Avance II + 600 MHz NMR spectrometer equipped with a 5 mm QNP probe. Typical experimental parameters were chosen as follows: spectral width 5.647 kHz, 19F excitation pulse 12.4 μs, acquisition time 1.0 s, relaxation delay 2 s, number of scans 2 K, proton decoupling using WALTZ-16 with γB₁ = 1 kHz. Prior to Fourier transformation all time domain data was processed with an exponential window function using a line broadening factor of 2 Hz. 19F-domain data was processed with an exponential window function using a line broadening factor of 2 Hz. 19F NMR spectra with 1H-decoupling were recorded as follows: spectral width 5.647 kHz, 19F excitation pulse 12.4 μs, acquisition time 1.0 s, relaxation delay 2 s, number of scans 2 K, proton decoupling using WALTZ-16 with γB₁ = 1 kHz. Prior to Fourier transformation all time domain data was processed with an exponential window function using a line broadening factor of 2 Hz.

RESULTS AND DISCUSSION

Synthesis of 5-fluoro pyrimidine containing DNA and RNA

All oligonucleotides investigated here were chemically synthesized by solid-phase synthesis. For 5-fluoro pyrimidine (dU₅F or dC₅F) modified DNA, we used the commercially available building blocks of either 5’-O-(4,4′-dimethoxytrityl)-5-fluoro 2′-O-deoxyuridine 3’-[(2-cyanoethyl)-(N,N-diisopropyl)]phosphoramidite or 5’-O-(4,4′-dimethoxytrityl)-5-fluoro-O²-(2,4,6-trimethylphenyl) 2′-deoxyuridine 3’-[2-cyanoethyl-(N,N-diisopropyl)phenyl] phosphoramidite (26). The latter was converted into 5-fluoro 2′-deoxyctydine after oligonucleotide synthesis by ammonium hydroxide used for the routine deprotection steps. For 5-fluoro uridine (U₅F) modified RNA, we applied commercially available 2′-O-(tert.-butyldimethylsilyl)-5′-O-(4,4′-dimethoxytrityl)-5-fluoro uridine 3’-[2-cyanoethyl-(N,N-diisopropyl)]phosphoramidite (6,23). For the preparation of 5-fluoro cytidine (C₅F) modified RNAs, we elaborated a short synthetic route towards 5-fluoro cytidine phosphoramidites equipped for solid-phase synthesis (Scheme 1). Up to now, 5-fluoro cytidine containing RNAs have been provided mostly by enzymatic methods, producing uniformly 5-fluoro cytidine labeled derivatives (24,25), or by chemical solid-phase synthesis using a convertible 5-fluoro-O²-methyl uridine nucleoside (6). Ammonolysis has been required to convert the uridine precursor into the corresponding cytidine nucleoside after completed strand assembly (NH₃/MeOH, r.t., 80 h) (6). The severe limitations of ammonolysis for the deprotection of larger chemically synthesized RNA strands prompted us to develop a synthetic route for the straightforward incorporation of 5-fluoro cytidine into RNA.

Facile synthesis of a 5-fluoro cytidine phosphoramidite derivative and incorporation into RNA

The synthesis started with commercially available 5-fluoro cytidine 1 using acetic anhydride in anhydrous ethanol for acetylation at C(4)–NH₂ to obtain compound 2 (Scheme 1). In analogy, the usage of benzoic anhydride resulted in derivative 6. After regioselective protection of the 5′-OH with the 4,4′-dimethoxytrityl group under standard conditions to yield compounds 3 and 7, respectively, the 2′ hydroxy group was reacted with tert.-butyldimethoxysilane in the presence of silver nitrate. The resulting derivatives 4a and 8a, respectively, were finally incorporated into phosphoramidite 5 and 9 by reaction with (2-cyanoethyl)-(N,N-diisopropylethoxysilane in the presence of 1-methylimidazole and N-ethylidisopropylamine in dichloromethane. This route provided compound 5 in a 17% overall yield and compound 9 in a 14% overall yield in four steps with three chromatographic purifications. In total, 0.7 g of 5 and 0.4 g of 9 were prepared in the course of this study. These building blocks were incorporated into RNA by solid-phase synthesis with >98% coupling efficiencies. Importantly, deprotection of RNA containing N⁴-acetyl-5-fluoro cytidine was accomplished under standard conditions using methanol in H₂O/EtOH whereas deprotection of the N⁴-benzoyl protected counterpart required treatment with aqueous ammonia prior to application of methanol to circumvent potential substitution of the phenyl carboxamide moiety at C(4) by methylene. Because of this, we preferred the preparation of 5-fluoro cytidine modified RNAs using the N⁴-acetyl protection concept.

The molecular weights of all 5-fluoro substituted DNA and RNA oligonucleotides synthesized in this study were confirmed by LC-ESI mass spectrometry (Table 1).
Bistable nucleic acids

Bistable nucleic acids exist in dynamic equilibria between two defined secondary structures (29). We have previously shown this behavior for a series of rationally designed oligoribonucleotides (29–31) and we also demonstrated the applicability of 2'-fluoro ribose labeling to probe the secondary structures of bistable RNA (42). The intention of the present study is on the one hand, to introduce bistable DNA oligonucleotides and on the other hand,

Table 1. Selection of chemically synthesized 5-fluoro uracil and 5-fluoro cytosine modified DNA and RNA oligonucleotides

<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence a</th>
<th>Length</th>
<th>Isolated yield</th>
<th>Molecular weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>5'-d(TCG TAC CGG GTA CGA ACC dU5F GC)-3'</td>
<td>26 nt</td>
<td>88 OD 260 nm</td>
<td>320 nmol</td>
</tr>
<tr>
<td>S2b</td>
<td>5'-d(CCG GAA GGT ACG AAC C dU5F T CCG)-3'</td>
<td>21 nt</td>
<td>78 OD 260 nm</td>
<td>350 nmol</td>
</tr>
<tr>
<td>S4</td>
<td>5'-r(GAC CGG AAG GUC CGC U5F U CC)-3'</td>
<td>20 nt</td>
<td>68 OD 260 nm</td>
<td>320 nmol</td>
</tr>
<tr>
<td>S4b</td>
<td>5'-r(GCG AAG TCG U5F GC CC)-3'</td>
<td>17 nt</td>
<td>86 OD 260 nm</td>
<td>500 nmol</td>
</tr>
<tr>
<td>S5</td>
<td>5'-r(GAC CGG AAG GUC CGC U5F U CC)-3'</td>
<td>20 nt</td>
<td>66 OD 260 nm</td>
<td>323 nmol</td>
</tr>
<tr>
<td>S5b</td>
<td>5'-r(GCG AAG TCG U5F CC CC)-3'</td>
<td>17 nt</td>
<td>27 OD 260 nm</td>
<td>161 nmol</td>
</tr>
<tr>
<td>S6</td>
<td>5'-d(TCG TAC CGG GTA CGA ACC dC5F UC)-3'</td>
<td>26 nt</td>
<td>150 OD 260 nm</td>
<td>550 nmol</td>
</tr>
<tr>
<td>S6b</td>
<td>5'-d(C CGG AAG GTA CGA AC dC5F TTC CG)-3'</td>
<td>21 nt</td>
<td>93 OD 260 nm</td>
<td>420 nmol</td>
</tr>
<tr>
<td>S7</td>
<td>5'-r(GAC CGG AAG GUC CGC cU5F CC)-3'</td>
<td>20 nt</td>
<td>64 OD 260 nm</td>
<td>310 nmol</td>
</tr>
<tr>
<td>S7b</td>
<td>5'-r(GCG AAG TCG cU5F CC)-3'</td>
<td>17 nt</td>
<td>64 OD 260 nm</td>
<td>370 nmol</td>
</tr>
</tbody>
</table>

a dU5F, 5-fluoro 2'-deoxy uridine; dC5F, 5-fluoro 2'-deoxy cytidine; U5F, 5-fluoro uridine; C5F, 5-fluoro cytidine.

b LC-ESI MS.

Bistable nucleic acids

Bistable nucleic acids exist in dynamic equilibria between two defined secondary structures (29). We have previously shown this behavior for a series of rationally designed oligoribonucleotides (29–31) and we also demonstrated the applicability of 2'-fluoro ribose labeling to probe the secondary structures of bistable RNA (42). The intention of the present study is on the one hand, to introduce bistable DNA oligonucleotides and on the other hand,
to assess a fluoro label that can be used in equivalent manner for both RNA and DNA secondary structure probing. The previously applied 2'-fluoro ribose label does not fulfill this requirement since this modification favors ribose C3'-endo over C2'-endo conformation and therefore supports structural integrity of RNA only (11). For these reasons, we focused on the assessment of 5-fluoro pyrimidine labeling in the present comparative study.

Rational design of bistable oligonucleotides

Following the sequence design of two hairpins that compete for the same internal sequence partition, we demonstrated that a 26 nt DNA with the sequence of 5'-TCG TAC CGG AAG GTA CGA ACC TTC CG-3' exhibited a 70:30 (estimated accuracy ±10) structure equilibrium between fold S1' and S1'' (Figure 2A). This was verified by comparative imino proton 1H NMR spectroscopy with a set of shorter reference oligonucleotides (S1a and S1b). (C) 1H decoupled 19F NMR spectra of S2 and the corresponding reference hairpin S2b and reference single strand 5'-d[ACC(dU5F)TCCG]. Conditions: 0.3 mM oligonucleotide, 25 mM sodium arsenate buffer, H2O/D2O = 9/1, pH 7.0, 300 K.

**Figure 2.** Bistable 26 nt DNA oligonucleotide. (A) Imino proton 1H NMR spectra of bistable DNA S1 (existing in a dynamic equilibrium of secondary structures S1' and S1'') and of reference hairpins S1a and S1b. (B) Same for the 5-fluoro 2'-deoxyuridine modified bistable DNA S2 (S2' and S2'') and the corresponding reference hairpins S1a and S2b. (C) 1H decoupled 19F NMR spectra of S2 and the corresponding reference hairpin S2b and reference single strand 5'-d[ACC(dU5F)TCCG]. Conditions: 0.3 mM oligonucleotide, 25 mM sodium arsenate buffer, H2O/D2O = 9/1, pH 7.0, 300 K.

Bistable DNA and RNA with 5-fluoro uracil labels

Alternative secondary structures of the same RNA sequence can be distinguished by use of 19F NMR spectroscopy if a distinct 2'-F modified nucleoside resides within the double helix of 1-fold while it is part of a single-stranded region within the alternative fold; we demonstrated this aspect previously (42). However, since 2'-F modified nucleosides are incompatible with the structural integrity of DNA as outlined above, we considered that the chemical environment of a fluorine atom attached at C5 of 2'-deoxyuridine should also be different for the two cases of double helical versus single-stranded arrangements. We therefore expected different chemical shift values of the corresponding 19F NMR resonances. In this sense, we replaced thymidine by 5-fluoro 2'-deoxyuridine in position 22 of S1 to yield S2 and analyzed the respective modified bistable DNA (Figure 2B). Indeed, the equilibrium position of S2, analyzed by comparative imino proton 1H NMR spectroscopy exhibited a 70:30 ratio between S2' and S2'' and was well comparable to that of the unmodified DNA S1.
Also, the proton-decoupled $^{19}$F NMR spectrum displayed two individual resonances (separated by $\Delta \delta/\delta_{C24} = 0.35$ ppm) corresponding to the two secondary structures (Figure 2C). The ratio determined on the basis of the $^{19}$F NMR signals was 70:30 and again in good agreement with the value determined from the $^1$H NMR spectrum. Together, this indicated that the influence of a single 5-fluoro uracil substituent on the DNA structure equilibrium was minor and within the range of accuracy. Interestingly, the $^{19}$F NMR resonance of 5-fluoro uracil showed an increased line width when it was involved in base pairing compared to the situation when it was located in an unpaired strand region. Moreover, we stress that the $^1$H NMR spectra of the bistable DNA S2 as well as the hairpin reference S2b were lacking the N(3)-H imino proton resonances of the respective dA:dU$^{SF}$ base pairs at 300 K (compare Figure 2A, top to Figure 2B, top; and Figure 2A, bottom to Figure 2B, bottom). All other imino protons showed similar chemical shift values as the ones observed for the unmodified DNA counterparts, S1 and S1b. Importantly, we observed the same behavior for the 20 nt bistable RNA S3 when one or even two A:U base pairs were replaced by A:U$^{SF}$ to yield S4 and S5, respectively (Figure 3A). Again, at 300 K, the $^1$H NMR imino proton resonance pattern of S4 and S5 was lacking one or two signals, respectively, that were attributed to 5-fluoro substituted Watson–Crick base pairs while all other imino protons provided chemical shift values comparable to that of the unmodified RNA counterpart. Our finding fits to a report by James and coworkers who showed in a temperature dependent series of $^1$H NMR spectra for the self-complementary duplex of [d(GGAAT(dU$^{SF}$)CC)]$_2$ that the dU$^{SF}$ N(3)-H resonance is extremely broad at low temperatures and becomes non-observable at room temperature (5). Also Hennig et al., observed for a uniformly 5-fluoro uridine labeled HIV TAR hairpin significant line broadening of 5-fluoro uridine N(3)-H imino protons together with a downfield chemical shift change between 0.6 and 0.7 ppm (24). The line broadening most likely reflects an increased exchange rate of dA:dU$^{SF}$ and A:U$^{SF}$ imino protons and relates to the considerable electronegativity of the 5-fluorine atom that lowers the apparent $pK_a$ of uridine N(3)-H from 9.18 ± 0.02 (unmodified uridine) to a value of 7.55 ± 0.02 (5-fluoro uridine) (43).

With respect to the structure equilibrium of the bistable RNA S3 ($S3':S3'' = 20:80$), a single 5-fluoro uridine substitution, as in S4, did not change the equilibrium ratio of 20:80 between S4' and S4'', determined by $^1$H NMR and $^{19}$F NMR spectroscopy in independent manner (Figure 3B).
and C). The chemical shift difference between the $^{19}$F resonances assigned for the RNA folds of $S_4^0$ and $S_4^{00}$ was 3.5 ppm, and therefore remarkably larger compared to the chemical shift difference of 0.35 ppm observed for the $^{19}$F resonances of the two DNA folds $S_2^0$ and $S_2^{00}$. This most likely reflects the different stacking patterns of Watson–Crick base pairs within A- and B-form (Figure 4 and Supplementary Data) and moreover, the much more pronounced difference in the chemical environment of a 5-fluoro pyrimidine atom that resides within an RNA duplex compared to its RNA single strand than the respective environments encountered for DNA.

Taken together, 5-fluoro uracil represents a non-invasive label for structural probing by $^{19}$F NMR spectroscopy of both, bistable DNA and RNA. This rationale holds true even for double labeling; the bistable RNA $S_5$ which contained two successive 5-fluoro uridine replacements revealed only a marginally shifted equilibrium position for the two secondary structures of $S_5^0$ and $S_5^{00}$ (Figure 3C, bottom).

**Bistable DNA and RNA with 5-fluoro cytosine labels**

The bistable DNA and RNA sequences, $S_1$ and $S_3$, were labeled with single 5-fluoro cytosines and—as their uracil counterparts—turned out to be non-invasive labels. The dC$_{5F}$ at position 21 of DNA $S_6$ resulted in an equilibrium position between $S_6^0$ and $S_6^{00}$ of 70:30 (Figure 5A) and was unchanged compared with the unmodified counterpart DNA $S_1$ (Figure 2A). Also for the C$_{5F}$ modified RNA $S_7$, the determined equilibrium position of 15:85 between $S_7^0$ and $S_7^{00}$ (Figure 5B) revealed that the structure equilibrium was—within the limits of
resonances have been previously reported for small organic molecules as well as fluorine labeled biomolecules (45). In the case of 5-fluoro cytidine modified oligonucleotides, the appearance of additional signals cannot only be attributed to solvent exposure of the fluorine nucleus but additionally stems from the isotopic replacement of exchangeable protons within the exocyclic amino group of the nucleobase (Figure 6C). A potential intramolecular hydrogen bond between C(5F)–H⋯N(C4) may be the cause for this pronounced effect and would be consistent with the lack of this phenomenon in the 5-fluoro uracil series.

Influence of 5-fluoro pyrimidine modifications on thermodynamic stability

Since 5-fluoro modifications left the equilibrium position of bistable DNA and RNA unchanged, we expected their impact on thermodynamic base pairing stability to be minor. To provide evidence for this assumption, we recorded UV melting curves of DNA hairpin references S1b (unmodified), S2b (dU5F) and S6b (C5F) (Supplementary Data). In aqueous buffer system of 150 mM NaCl and 10 mM phosphate buffer, at pH 7.0 and at 2 μM DNA concentration, their Tm values varied only within 72.5 ± 1.0°C and their thermodynamic stabilities were in the small range of ΔG° = 7.1 ± 0.5 kcal mol⁻¹. The Tm values of the RNA hairpin references S3b (unmodified), S4b (U5F), S5b (U5F,U5F) and S7b (C5F) also varied only within one degree, 75.5 ± 1.0°C, and their thermodynamic stabilities were within the range of ΔG° = 7.1 ± 0.3 kcal mol⁻¹. In extrapolation, these thermal and thermodynamic data suggest that the stabilities of the two competing hairpin structures in bistable oligonucleotides are minimally affected by the 5-fluoro pyrimidine labels.

Pseudoknot formation of a riboswitch aptamer

To demonstrate the value of 5-fluoro pyrimidine labeling for biologically relevant RNAs, we monitored the formation of a riboswitch aptamer/ligand complex (Figure 7). Based on the recently published X-ray and NMR structures of a 7-aminomethyl-7-deazaguanine (preQ1) sensing mRNA aptamer domain (46-49), we expected that fluorine labeling of U32 in the single-stranded nucleotide 3′-overhang (…(C5F–U32)AG-3′) would allow to follow base pairing with four complementary nucleotides in the loop (double-helix formation). Indeed, a significant shift of the fluorine resonance from 164.4 to 166.5 ppm was observed upon ligand addition corresponding to the free versus ligand-bound RNA (Figure 7). Ongoing experiments focus on the verification of ligand-induced secondary structure rearrangements within the expression platform of the full-length preQ1 riboswitch domain containing single fluorine labels.

CONCLUSIONS

With the present study, we demonstrated the reliability of 5-fluoro labeling at pyrimidine nucleobases for secondary structure probing by 19F NMR spectroscopy. For bistable
DNA and RNAs, the impact of single 5-fluoro pyrimidine substitutions on the equilibrium position was within the accuracy of determination. Watson–Crick base pairing properties were preserved corresponding to the observation that thermodynamic stabilities remained unchanged for the modified versus non-modified double helices. 5-fluoro pyrimidine modifications hence turned out to be non-invasive spin label options for secondary structure evaluation by 19F NMR spectroscopy. This was not expected in such strict manner, since the $pK_a$ values of the N(3) imino proton of 5-fluoro uracil and of N(3) protonated 5-fluoro cytosine are decreased by the order of two. A change of $pK_a$ values could potentially affect base pair formation, in particular base pairing strength because of higher proton mobility and higher proton exchange rates with the solvent, and consequently could affect the RNA structure equilibrium. However, our study showed that this concern did not come true.

In comparison with RNA 2′-F ribose labeling that we explored previously in our laboratory (42), 5-fluoro labeling is superior in one aspect: the modification does not directly affect ribose puckering and is therefore applicable for both DNA and RNA. However, we recall that 2′-F ribose represents an ideal label if constitutionally equivalent labeling of all four standard nucleosides is required (42). We also recall that 2′-F modified pyrimidine nucleosides shifted the equilibrium position of bistable RNAs by ~25% (42) which can be regarded as disadvantage compared to 5-F pyrimidine labeling. Only 2′-F modified purine nucleosides left the RNA equilibrium position unchanged as observed here for the 5-fluoro pyrimidine series (42).

RNA with site-selective 5-F pyrimidine labels are conveniently accessible by chemical solid-phase synthesis for a size up to ~65 nucleotides. Larger RNAs with site-specific labels in nmol quantities that are required for various types of biophysical experiments are accessible by combined approaches using chemical synthesis and enzymatic ligation methods as we have recently exemplified for adenine and TPP riboswitches containing fluorescent dyes (50–53). We consider the advantage of site-specific 5-F pyrimidine labeling over fully 5-F pyrimidine labeled RNA [which are efficiently prepared biochemically (24,25)] in the avoidance of constitutional assignments of 19F resonances and also in the avoidance of signal overlaps that are likely to be encountered in the fully labeled counterparts. We are confident that the use of RNAs labeled with single fluorine probes is a requirement to assist facile verification of RNA secondary structure models and we are currently working along this line.

Taken together, we consider this study as an important step on the way to identify reliable fluoro labels for RNA structure probing by NMR spectroscopy, with the long-term aim to contribute to the elucidation of folding pathways of biologically relevant, mid-sized RNA (up to ~200 nt), such as riboswitches (49–52).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Funding by the Austrian Science Fund FWF (I317-N17 to R.M.; SFB F17 to R.K.) and the Ministry of Science and Research bmf:wf (GenAU project ‘Non-coding RNAs’, P0726-012-012 to R.M.) is acknowledged. Funding for open access charge: Austrian Science Fund FWF.

Conflict of interest statement. None declared.

REFERENCES


