Supporting information

Synthesis and biophysical analysis of modified thymine-containing DNA oligonucleotides

Fumiko Kawasaki, [a] Pierre Murat, [a] Zhe Li, [a] Tobias Santner, [a] and Shankar Balasubramanian [a, b, c]

a Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

b Cancer Research UK Cambridge Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, U.K.

c School of Clinical Medicine, University of Cambridge, Cambridge CB2 0SP, U.K.

Index

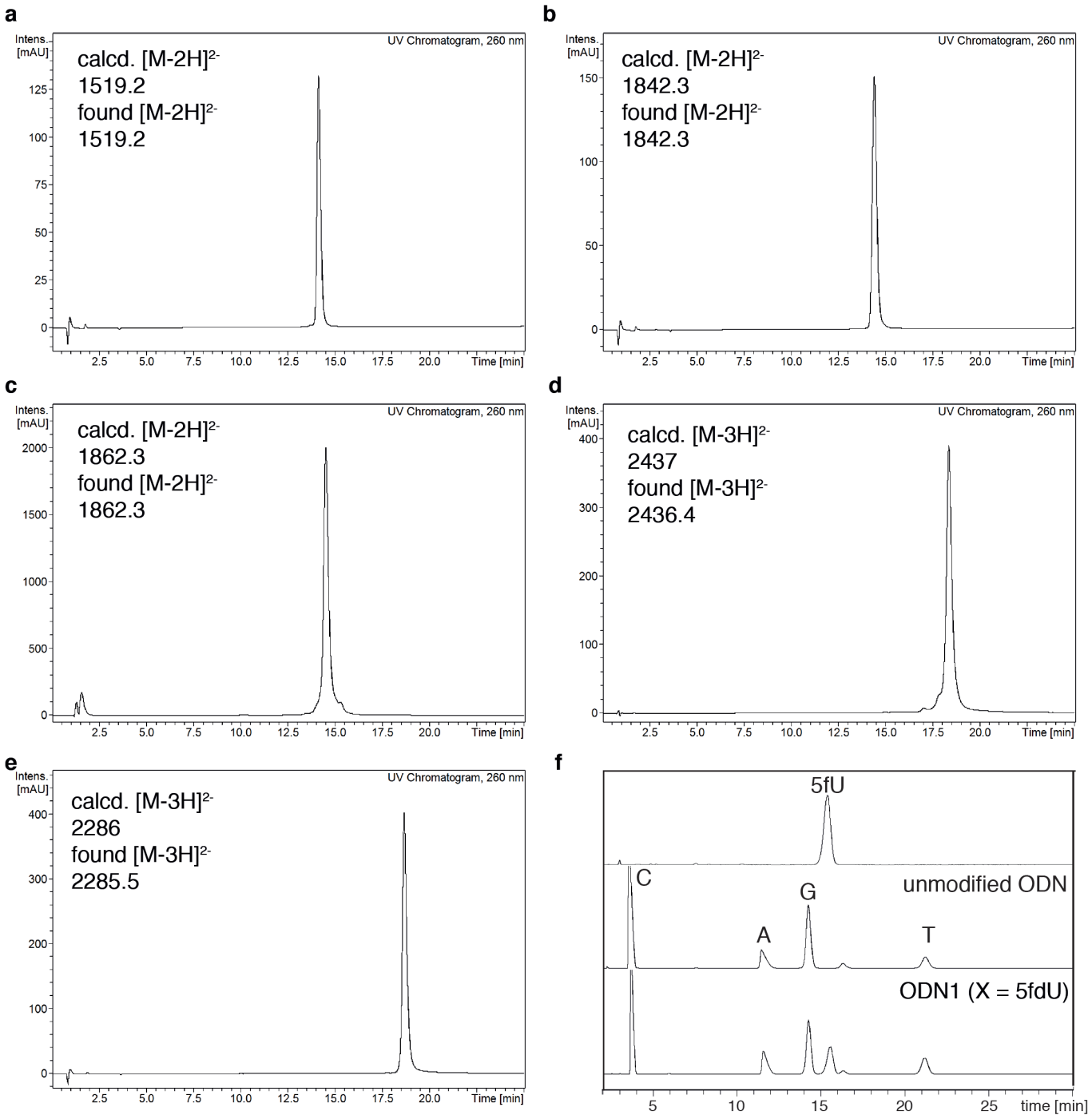


Fig S1. LC-MS chromatogram of ODNs (X = 5fU) after purification and mononucleoside composition analysis of ODN1 (X = 5fU). LC-MS chromatogram for 5fU modified ODN1 (a), ODN2 (b), ODN3 (c), ODN6 (d), ODN7 (e). A LC trace of digested ODN1 recorded at 280 nm together with LC traces of a digested non-modified ODN, as well as 5fU mononucleoside (f). The presented MS signals are the ESI signal of the indicated peak tops. For experimental details including LC conditions, see page S10.

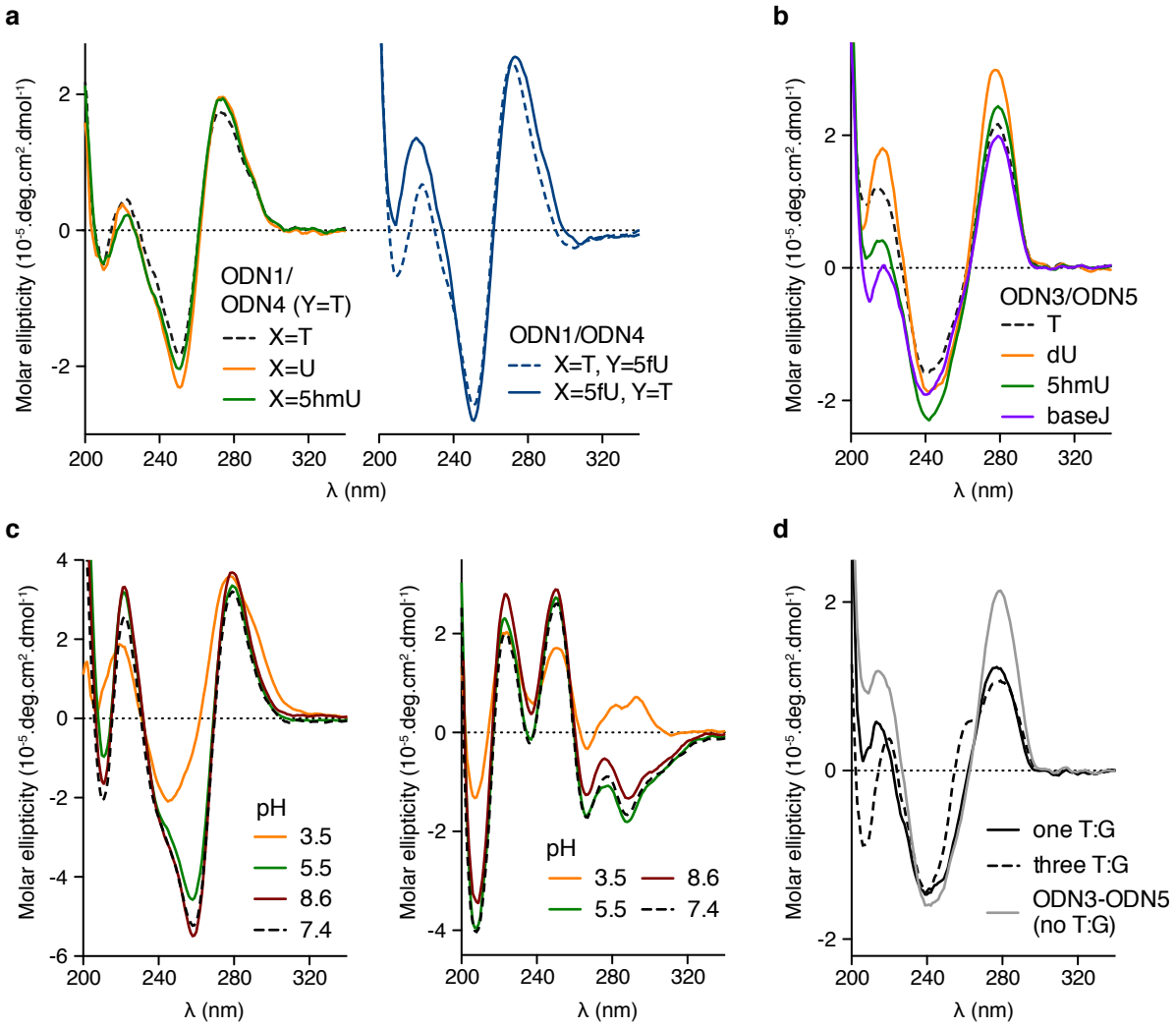


Fig S2. CD spectra of ODNs. (a) CD spectra of ODN1 (X=U, 5hmU, or 5fU)–ODN4 (Y=T or 5fU) duplexes. A copy of CD spectrum for non-modified control (Fig 2a) is shown in a dashed line; (b) ODN3–ODN5 (X=U, 5hmU and base J) duplexes. A copy of CD spectrum for non-modified control (Fig 2b) is shown in a dashed line; (c) non-modified (left) and 5fU-containing (right) self-complementary ODN2 duplex at different pH. Copies of data at pH 7.2 (Fig 2c) are shown in dashed lines. (d) non-modified ODN3–ODN5 duplex with one (annealed to 5'-CTCAGGCAGTTC-3'), two (annealed to 5'-CTCGGACGGTTC-3'), and three (annealed to 5'-CTCGGGCGGTTC-3') T:G pairs.

Table S1. Sequences and melting temperature of additional ODNs

X	Y	T_m (°C)[a] (ΔT_m based on T control)
ODN1: 5'-ATCGCAXGTA-3' ODN4: 3'-TAGCGYACAT-5'		
5fU	T	44.0 ± 0.1 (0.5)[b]
T	5fU	44.5. ± 0.1 (1.0)[b]
U	T	44.8. ± 0.8 (1.3)[b]
5hmU	T	43.5 ± 0.9 (0)[b]
ODN6: 5'-AGGGXTAGGGXTAGGGXTAGGGT-3' ODN7: 3'-TCCCAAYCCCAAYCCCAAYCCCA-5'		
U	U	68.3 ± 0.2 (–0.5)[b]
5hmU	5hmU	66.0 ± 0.3 (–2.8)[b]
base J	base J	68.0 ± 0.3 (–0.8)[b]
5fU	T	65.8 ± 0.8 (–3.0)[b]
T	5fU	68.5 ± 0.4 (–0.3)[b]
ODN6 (G-quadruplex)		
U		65.3 ± 0.0 (2.1)[c]
5hmU		64.9 ± 0.5 (1.7)[c]
base J		65.0 ± 0.1 (1.8)[c]

a. Mean ± sd values for three measurements are shown. b. T_m values were measured by UV_{260} absorbance at 5 µM of each ODN in PBS. c. T_m values were measured by UV_{295} absorbance at 10 µM in 10 mM phosphate buffer (pH 7.4) and 70 mM KCl for G-quadruplexes.

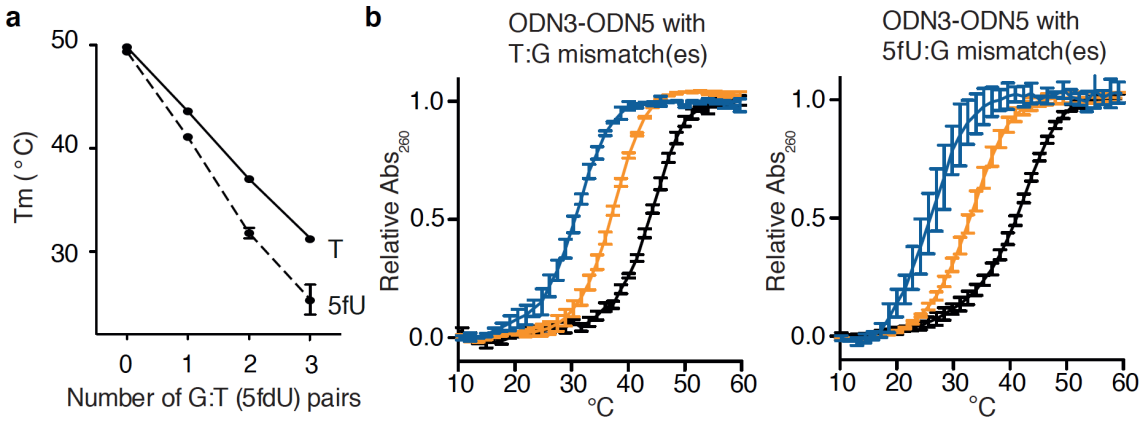


Fig S3. Tm of ODN3 (X=T or 5fU) when annealed with ODN5 (no mismatch), 5'-CTCAGGCAGTTC-3' (one mismatch), 5'-CTCGGACGGTTC-3' (two mismatches), 5'-CTCGGGCGGTTC-3' (three mismatches). (a) Tm values plotted as mean ± sd of three measurements. (b) Melting curves shown as mean ± sd of three measurements (black: one mismatch; orange: two mismatches; blue: three mismatches).

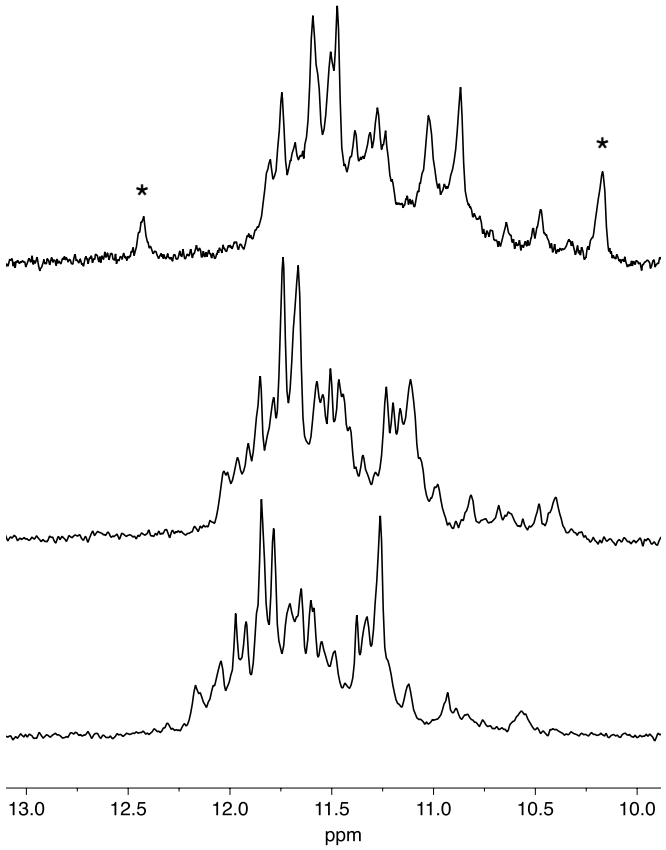


Fig S4. ^{1}H NMR spectra of OND6 (X=5fU) G-quadruplexes at 5 °C (top), 25 °C (middle), and 40 °C (bottom). Signals marked by asterisks were broadened (or were unobservable) at the elevated temperatures. Measurements were carried out at 500 MHz in 10% D_2O in water.

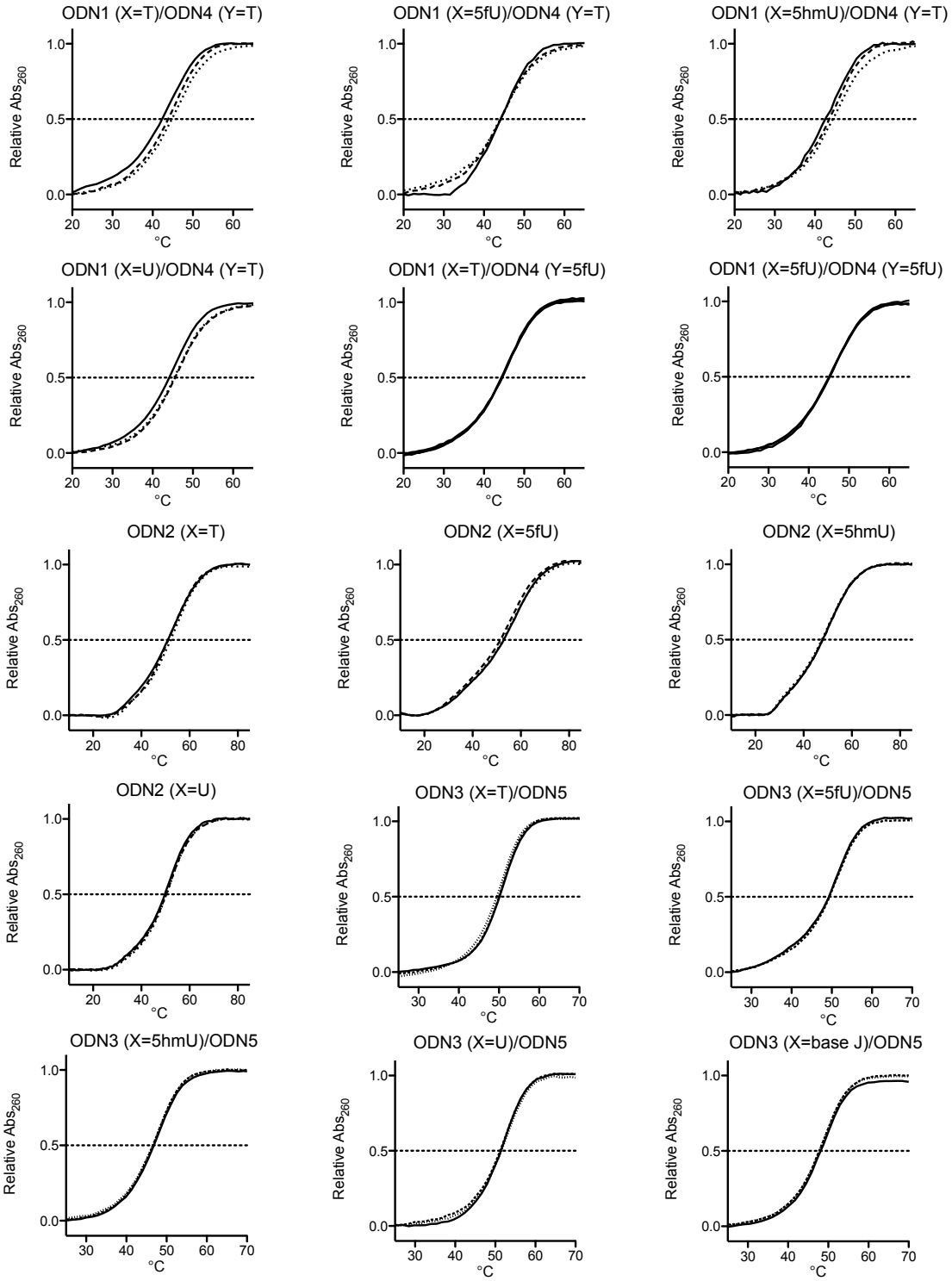


Fig S5. Normalised melting curves corresponding to T_m values in Table 1 and Table S1. Curves obtained from three measurements are shown in solid, dashed, and dotted lines.

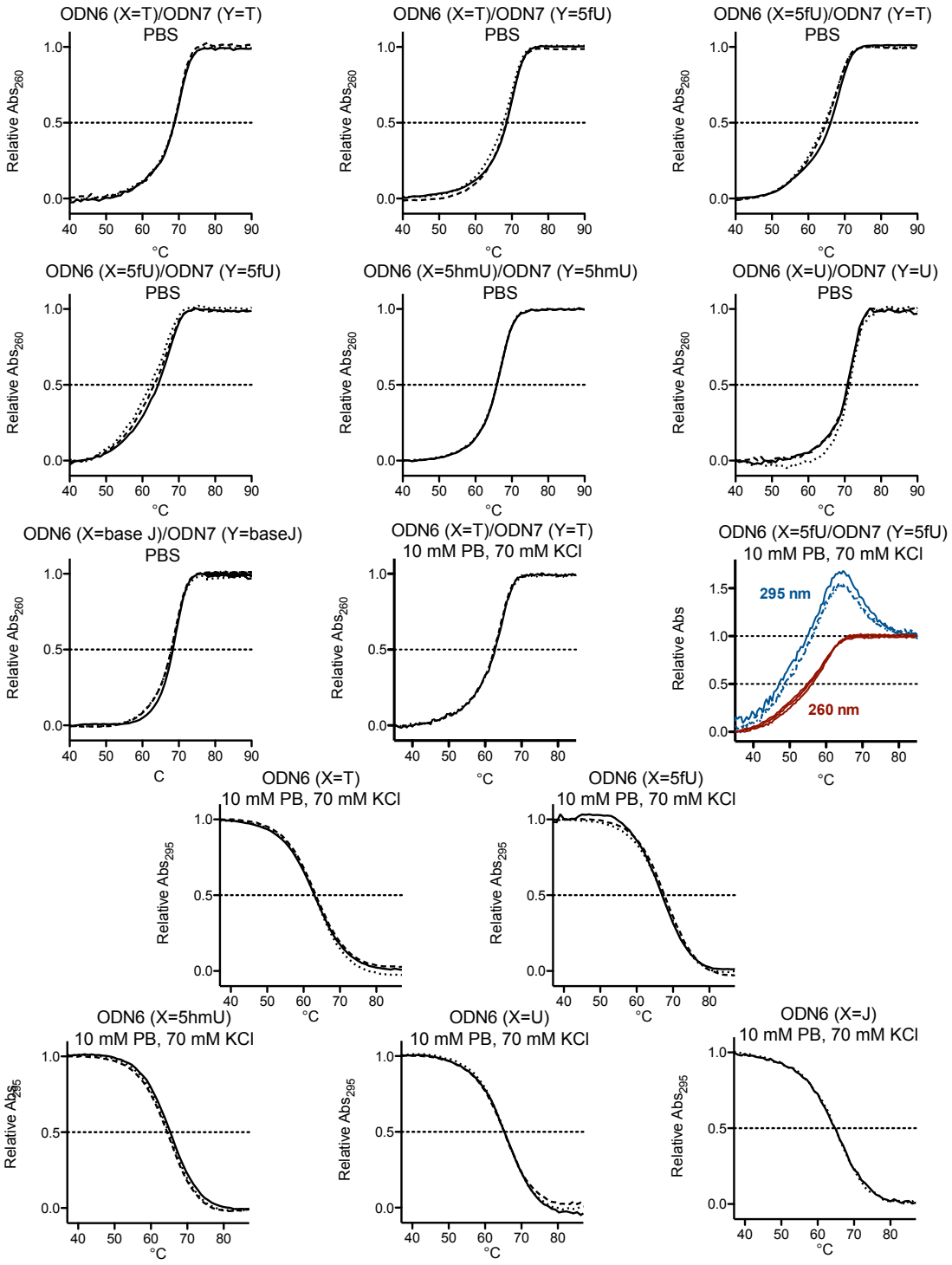


Fig S6. Normalised melting curves corresponding to T_m values in Table 2, and Table S1. Curves obtained from three measurements are shown in solid, dashed, and dotted lines. For ODN6 (X=5fU) in the KCl supplemented buffer, a curve obtained at UV_{295} was also shown. The absorbance peak at UV_{295} indicates a presence of G-quadruplexes in competition to the duplex at elevated temperatures.

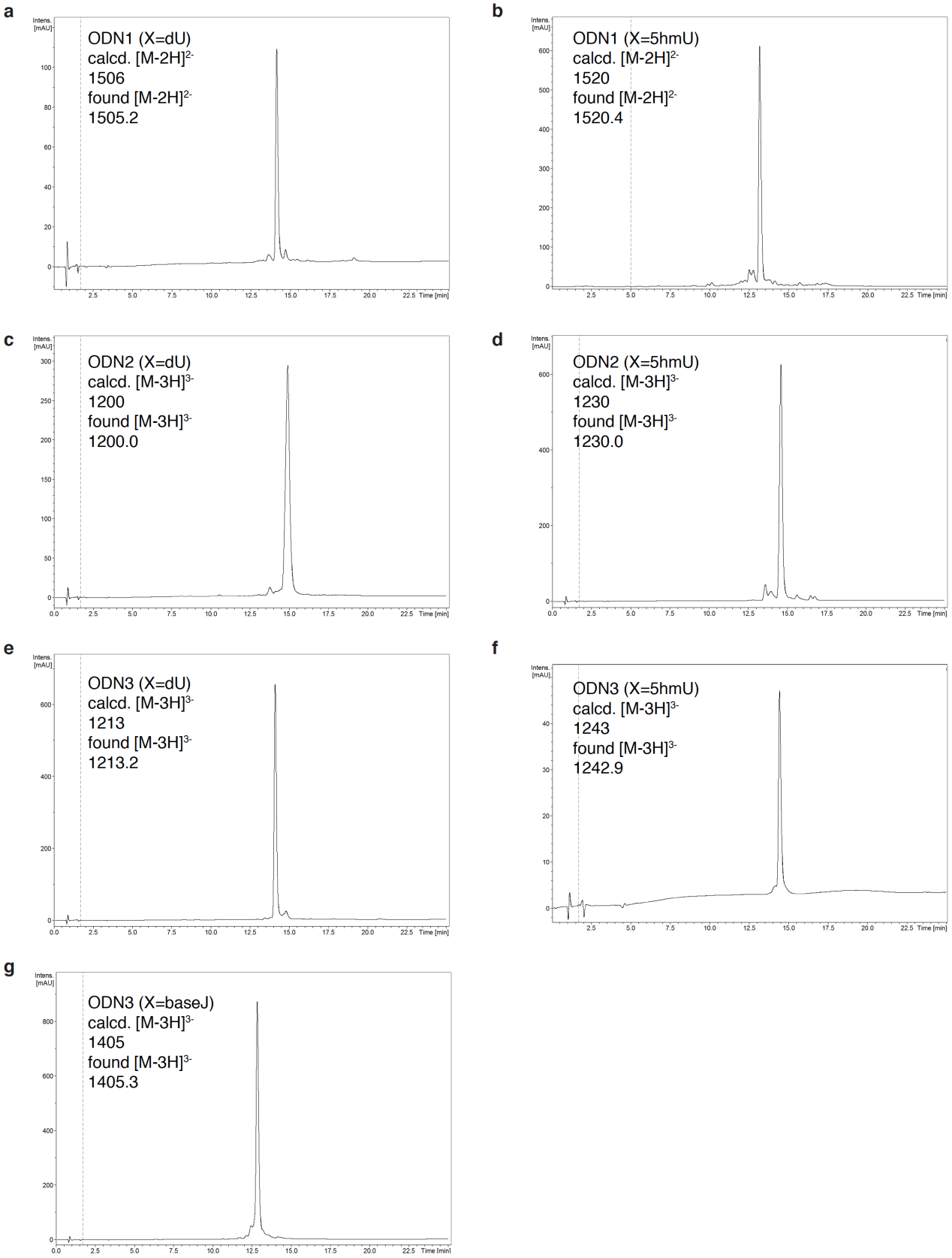


Fig S7. LC-MS chromatogram of ODN1–3 (X = dU, 5hmU, and base J)

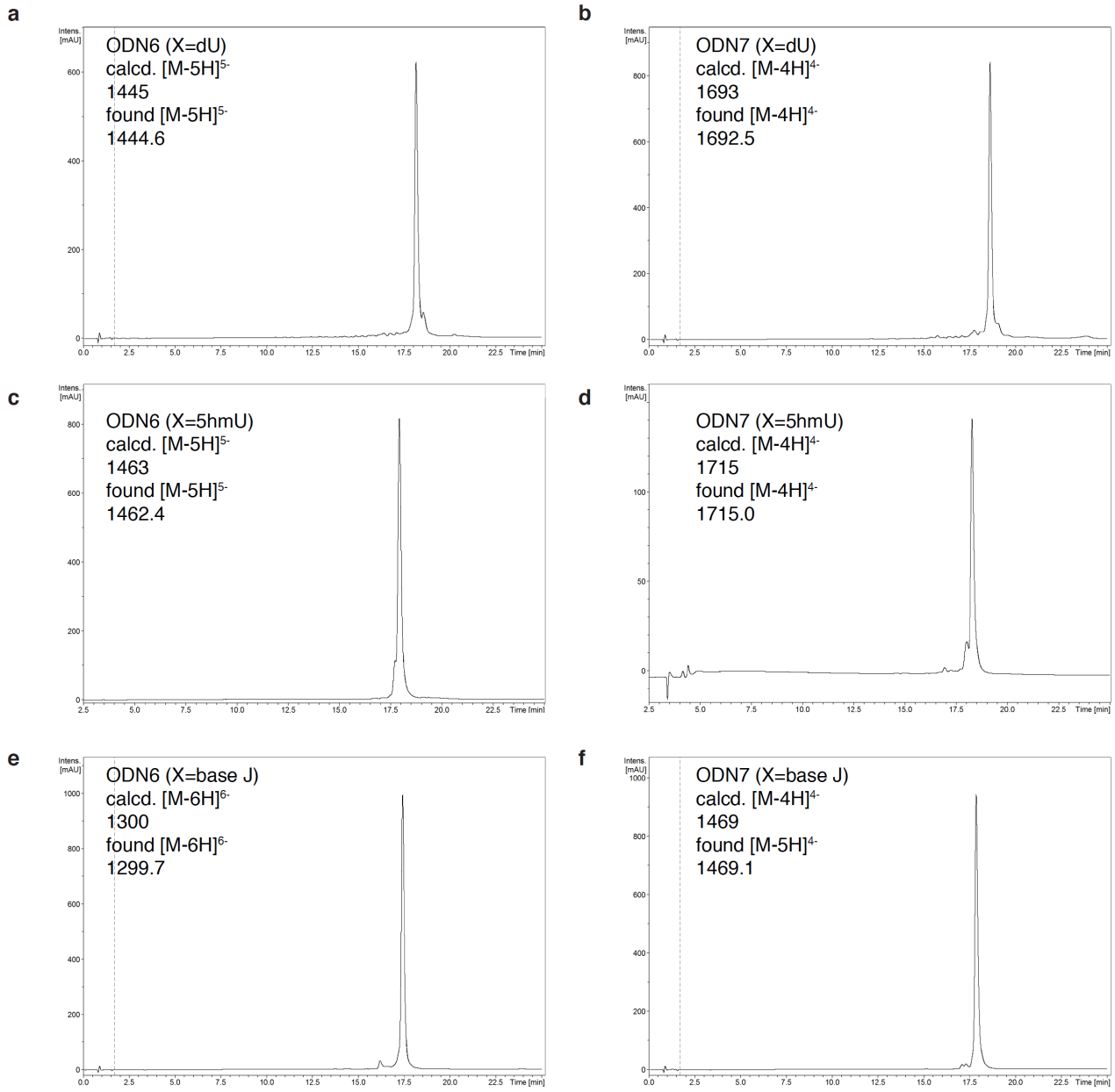


Fig S8. LC-MS chromatogram of ODN6 and ODN7 (X = dU, 5hmU, and base J)

Experimental Procedures

General

NMR spectra were recorded on a AVANCEIII HD Smart Probe Spectrometer (Bruker). ^{1}H NMR spectra were obtained at 400 MHz with $CHCl_3$ (δ 7.26) or with CD_3CN (δ 1.94) as internal standards. ^{13}C NMR spectra were obtained at 100 MHz with $CDCl_3$ (δ 77.0) in $CDCl_3$ or with CD_3CN (δ 1.4) as internal standards. ^{31}P NMR spectra were obtained at 161 MHz with external calibration. HRMS spectra were recorded on Micromass® Q-Tof (ESI) spectrometer (Waters). LC-ESIMS spectra were recorded on an Amazon ESI-MS (Bruker) connected to Ultimate 3000 LC (Dionnex). All the

reactions were conducted under an inert atmosphere unless otherwise noted. Silica gel column chromatography was performed on CombiFlash Rf+ (TELEDYNE ISCO) using with pre-packed cartridges unless otherwise noted. Dry organic solvents were prepared prior to use. The other organic solvents were reagent grade and used as received. Water was in ultrapure water quality and produced by Synergy® UV Remote Water Purification System (Merck Millipore). CD spectra were recorded on Chirascan Plus spectropolarimeter (Applied Photophysics, UK). UV spectra were recorded on Cary 100 UV-Vis Spectrophotometer (Agilent Technologies, California, USA).

Chemical synthesis of phosphoramidite 5

3',5'-*O*-Di-tert-butylsilyl-5-(1,3-dioxan-2-yl)-2'-deoxyuridine (2)

To a solution of 3',5'-*O*-di-tert-butylsilyl-2'-deoxyuridine[1] (396 mg, 1 mmol) in anhydrous CH2Cl2 (20 mL) were added 1,3-propanediol (300 uL, 4.2 mmol) and triethylorthoformate (200 uL, 1.2 mmol) at 0 °C. To the mixture, $TiCl_4$ (1 M in anhydrous CH_2Cl_2, 420 uL, 0.42 mmol) was added dropwise at 0 °C, and the reaction mixture was stirred at 0 °C for 1 h. The reaction was quenched by addition of sat. $NaHCO_3$ aq. (20 mL) and the organic layer was separated. The aqueous layer was extracted with CH_2Cl_2 (20 mL x 2). Organic layers were combined, filtered over a cotton plug, and concentrated. The crude product was purified by silica gel column chromatography (CH_2Cl_2–MeOH, 1:0 to 9:1, v/v) to give **2** as a colorless foam (365 mg, 80%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.26 (br, 1H), 7.59 (s, 1H), 6.14 (dd, 1H *J* = 3.0 Hz, 7.1 Hz), 5.51 (s, 1H), 4.43 (dd, 1 H, *J* = 5.1 Hz, 9.4 Hz) 4.26 (app. q, 1H, *J* = 9.3 Hz), 4.17 (m, 2H), 4.02 (app. t, 1H, *J* = 9.5 Hz), 3.94 (m, 2H), 3.68 (td, 1H, *J* = 5.1 Hz, 9.4 Hz), 2.43–2.30 (m, 2H), 2.19–2.04 (m, 1H), 1.40 (m, 1H), 1.04 (s, 9H), 1.00 (s, 9H). ^{13}C NMR (100 MHz, $CDCl_3$) δ 161.5 (C), 149.6 (C), 138.8 (CH), 112.7 (C), 94.9 (CH), 84.6 (CH), 78.2 (CH), 74.3 (CH), 67.4 (CH2), 67.3 (CH2), 38.6 (CH2), 27.4 (CH3), 27.1 (CH3), 25.6 (CH2), 22.6 (C), 20.0 (C). HRMS calcd. $[M+H]^+$ 455.2214, found $[M+H]^+$ 455.2216.

5-(1,3-Dioxan-2-yl)-2'-deoxyuridine (3)

To a solution of **2** (365 mg, 0.80 mmol) in anhydrous THF−pyridine (100:1.5, v/v, 10 mL), was added HF·pyridine (ca. 70% HF, 50 uL) at 0 °C. The reaction mixture was stirred at 0 °C for 10 min then rt for 1 h, followed by addition of sat. $NaHCO_3$ aq. (5 mL). The mixture was concentrated and absorbed to silica gel. The crude product was purified by silica gel column chromatography (CH_2Cl_2–MeOH, 1:0 to 4:1, v/v) to give **3** (260 mg, 99%) as a colorless foam. ^{1}H NMR (400 MHz, $CDCl_3$–CD_3CN) δ 9.09 (br, 1H), 8.02 (s, 1H), 6.15 (t, 1H *J* = 6.6 Hz), 5.40 (s, 1H), 4.37–4.31 (m, 1H),

4.13–4.06 (m, 2H), 3.92–3.82 (m, 3H), 3.76–3.62 (m, 2H), 3.36–3.26 (m, 1H, 3'-OH), 3.15–3.10 (m, 1H, 5'-OH), 2.25–2.15 (m, 2H), 2.05–1.96 (m, 1H), 1.38 (m, 1H). ^{13}C NMR (100 MHz, $CDCl_3$–CD_3CN) δ 162.2 (C), 150.7 (C), 140.2 (CH), 113.0 (C), 96.0 (CH), 88.1 (CH), 86.3 (CH), 71.6 (CH), 67.9 (CH2), 67.9 (CH2), 41.0 (CH2), 26.1 (CH2). HRMS calcd. $[M+H]^+$ 315.1192, found $[M+H]^+$ 315.1187.

5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-5-(1,3-dioxan-2-yl)-2'-deoxyuridine (4)

To a solution of **3** (245 mg, 0.78 mmol) in anhydrous pyridine (10 mL) were added Et_3N (0.2 mL, 1.5 mmol) and 4,4'-dimethoxytriphenylmehtylchloride (401 mg, 1.2 mmol) at rt. The reaction mixture was stirred at rt for 4 h then concentrated. The crude product was purified by silica gel column chromatography (CH_2Cl_2–MeOH–Et_3N, 100:0:1 to 90:10:1, v/v/v) to give **4** (444 mg, 92%) as a colorless foam. ^{1}H NMR (400 MHz, CD_3CN) δ 8.97 (br), 7.80 (s, 1H), 7.51–7.44 (m, 2H), 7.40–7.27 (m, 6H), 7.26–7.20 (m, 1H), 6.89–6.86 (m, 4H), 6.15 (dd, 1H *J* = 5.9, 7.6 Hz), 5.32 (s, 1H), 4.32–4.26 (m, 1H), 3.95–3.89 (m, 1H), 3.85–3.66 (m, 10H), 3.30 (ABX, 1H, *J* = 3.2, 10 Hz), 3.22 (ABX, 1H, *J* = 4.3, 10 Hz), 2.27–2.07 (m, 2H), 1.64–1.50 (m, 1H), 1.14 (m, 1H). ^{13}C NMR (100 MHz, CD_3CN) δ 161.3 (C), 158.7 (C), 149.9 (C), 145.1 (C), 138.3 (CH), 136.0 (C), 136.0 (C), 130.2 (CH), 130.2 (CH), 128.2 (CH), 127.9 (CH), 126.9 (CH), 113.1 (CH), 112.7 (C), 95.3 (CH), 86.4 (C), 86.1 (CH), 84.9 (CH), 71.2 (CH), 67.0 (CH2), 66.9 (CH2), 63.5 (CH2), 55.0 (CH3), 40.2 (CH2), 25.1 (CH2). HRMS calcd. $[M+H]^+$ 617.2499, found $[M+H]^+$ 617.2513.

5'-*O*-(4,4'-Dimethoxytriphenylmethyl)- 5-(1,3-dioxan-2-yl)-2'-deoxyuridine 3'-*O*-(2-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite (5).

To a solution of **4** (425 mg, 0.68 mmol) in anhydrous CH_2Cl_2 (10 mL), were added Et_3N (0.60 mL, 4.4 mmol) and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (200 uL, 0.9 mmol) at 0 °C. The reaction mixture was stirred at 0 °C for 2 h, followed by the addition of sat. $NaHCO_3$ aq. The organic layer was separated, and the aqueous layer was extracted with CH_2Cl_2 (20 mL x2). The organic layers were combined, filtered over a cotton plug, and concentrated. The crude product was purified by silica gel column chromatography (hexane–ethylacetate–Et_3N, 70:30:1 to 0:100:1 v/v/v). All the fractions containing the product were concentrated and purified by silica gel column chromatography (ethylacetate–Et_3N, 100:1 v/v) to give **5** as a mixture of diastereomers (316 mg, 57%). ^{1}H NMR (400 MHz, CD_3CN) δ 9.28 (br), 7.87 (S), 7.84 (s), 7.50–7.20 (m, 9H), 6.92–6.83 (m, 4H), 6.22–6.14 (m, 1H), 5.35 (s, 1H), 4.53–4.44 (m, 1H), 4.14–4.02 (m, 1H), 3.85–3.49 (m), 3.44–3.33 (m, 1H), 3.29–3.21 (m, 1H), 2.63 (t, 1H, *J* = 5.9 Hz), 2.50 (t, 1H, *J*

= 5.9 Hz), 2.47–2.32 (m, 1H), 2.28–2.16 (m, 1H), 1.64–1.47 (m, 1H), 1.12-1.10 (m), 1.03 (d, 3H, J = 6.8). ^{13}C NMR (100 MHz, CD_3CN) δ 161.5 (C), 161.5 (C), 158.7 (C), 150.0 (C), 145.0 (C), 145.0 (C), 138.3 (CH). 138.3 (CH), 135.9 (C), 135.8 (C), 135.8 (C), 135.7 (C), 130.1 (CH), 130.1 (CH), 128.1 (CH), 128.0 (CH), 127.9 (CH), 126.8 (CH), 113.1 (CH), 112.8 (C), 112.8 (C), 95.1 (CH), 95.1 (CH), 86.5 (C), 86.5 (C), 85.3 (CH, J_{c-p} = 4.0 Hz), 85.1 (CH, J_{c-p} = 5.6 Hz), 84.9 (CH), 84.9 (CH), 73.6 (CH, J_{c-p} = 17 Hz), 73.2 (CH, J_{c-p} = 17 Hz), 66.9 (CH2), 66.9 (CH2), 63.2 (CH2), 63.0 (CH2), 58.6 (CH2, J_{c-p} = 8.0 Hz), 58.4 (CH2, J_{c-p} = 8.0 Hz), 54.9 (CH3), 43.1 (CH, J_{c-p} = 1.2 Hz), 43.0 (CH, J_{c-p} = 1.2 Hz), 39.5 (CH2, J_{c-p} = 3.9 Hz), 39.4 (CH2, J_{c-p} = 4.7 Hz), 25.0 (CH2), 24.0 (CH3), 23.9 (CH3), 23.9 (CH3), 23.8 (CH3), 20.1 (CH2, J_{c-p} = 7.3 Hz), 20.0 (CH2, J_{c-p} = 7.3 Hz). ^{31}P NMR (CD_3CN) δ 148.0, 147.9. HRMS calcd. $[M+H]^+$ 815.3572, found $[M+H]^+$ 815.3532.

Synthesis of ODNs containing 5fU

The 1,3-propanediol-protected ODN1–3, 6, 7 were synthesized by ATDBio Ltd (Southampton, UK) using a standard solid-support oligonucleotide synthesis protocol. LC-MS analysis indicated the formation of desired products without any major side product formation or truncated ODNs. The products were desalted (ODN1-3) or HPLC purified (ODN4, 6 and 7), and incubated with 5% (2% for ODN1) (v/v) aqueous AcOH at rt for 90 min followed by neutralization by addition of 1 M NaOH. The obtained mixture was desalted by PD-10 column (GE Healthcare, UK), lyophilized, and used for biophysics study without further purification.

Synthesis of other ODNs

ODN1–3 bearing 5hmdU and base J were synthesized by ATDBio Ltd using 5'-*O*-(acetoxymethyl)-2'-deoxyuridine 3'-O-(2-cyanoethyl-*N*,*N*'-diisopropyl) phosphoramidite and 5'-*O*-(2,4,4,6-tetra-*O*-benzoyl-ß-glucopyranosyl)hydroxymethyl-2'-deoxyuridine 3'-*O*-(2-cyanoethyl-*N*,*N*'-diisopropyl) phosphoramidite. ODN 6, 7 bearing 5hmU were purchased from Eurogentec (Seraing, Belgium). The rest ODNs were purchased from Life Technologies (CA, USA). All above ODNs were obtained as HPLC purified samples.

LC-MS analysis of ODNs

ODNs were analyzed by LC-MS using XBridge BEH C18 XP (30Å, 2.5 µm, 3 mm x 50 mm, Waters) or XTerra MS C18 Column (125Å, 2.5 µm, 2.1 mm x 50 mm, Waters). The elution solvents were 10 mM triethylamine and 100 mM

hexafluoroisopropanol (Acros Organics) in water (solvent A) and methanol (solvent B). After equilibrating at 5% B, ODNs were eluted with 2 min at 5% B, 20 min at 5–30% B, 3 min at 30% B) at 0.2 ml/min.

LC analysis of digested ODNs

ODNs (2.5 μg) were treated with DNA Degradase Plus™ (0.1 U/μL as final concentration, ZYMO RESEARCH) at 37 ºC for 4 h. The solution was subjected to HPLC analysis using Pursuit C18 (5 μm, 4.6 mm x 250 mm, Agilent) and eluted with 0.1% (v/v) formic acid in water at 1.5 mL/min.

Thermal denaturation study of DNA duplex

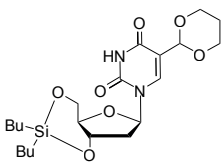
Each ODN (5 μM) was annealed with its complementary ODN (5 μM, see Table 1 for sequences) in PBS buffer (10 mM sodium phosphate salt, 137 mM NaCl and 2.7 mM KCl, pH 7.2). Samples in quartz cuvettes with 1 cm path length were heated from 5 °C (20 °C for ODN6–ODN7 duplex) to 70 °C (90 °C for ODN6–ODN7 duplex) at the ramping rate of 1 °C /min, then cooled from 70 °C (90 °C for ODN6–ODN7 duplex) to 5 °C (20 °C for ODN6–ODN7 duplex) at the rate of 1, 2, or 4 °C /min. The heating-cooling cycle was repeated for three times, with data collected at every 1 °C, and UV absorbance at 260 nm was recorded with Cary 100 UV-Vis Spectrophotometer (Agilent). Tm values were defined as the temperature where half of the duplex is dissociated in the heating process.

T_m analysis of single stranded ODN6

ODN6 (10 μM) was annealed in 75 mM KCl in 10 mM sodium phosphate buffer (pH 7.2). Samples were heated from 20 °C to 90 °C at the rate of 1 °C /min, then cooled from 90 °C to 20 °C at the rate of 2 °C per min for three cycles and absorbance at UV290 was measured.

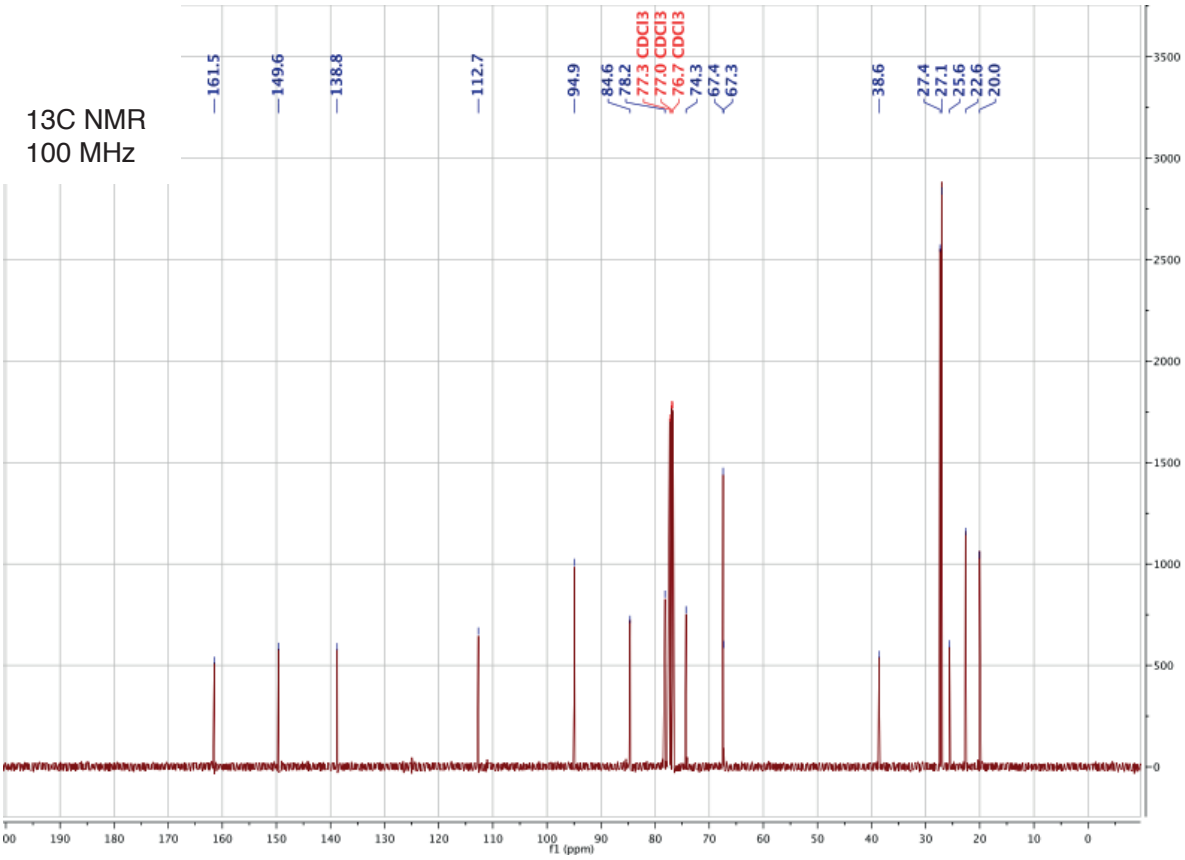
CD measurement

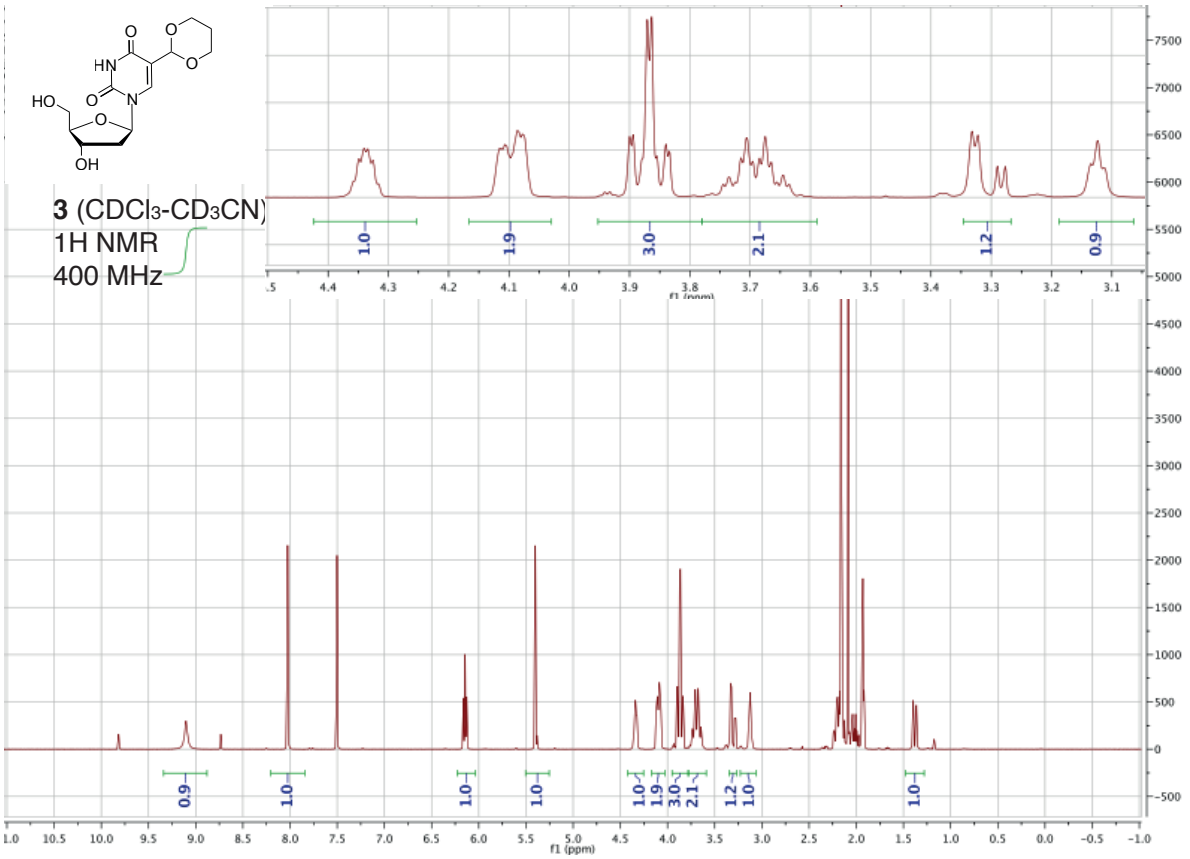
Each DNA duplex (10 μM) in PBS, or single-stranded ODN6 (10 μM) in 10 mM sodium phosphate buffer (pH 7.2) in presence or absence of 75 mM KCl was annealed by heating at 95 °C for 5 minutes, and gradually cooling down to 4 °C. CD spectra were measured with Chirascan Plus spectropolarimeter (Applied Photophysics, UK) in quartz cuvettes with 0.1 cm path length at 20 °C. Results from three replicates were averaged, smoothed, and corrected with the spectra of the buffer and the molar ellipticity value at 350 nm.

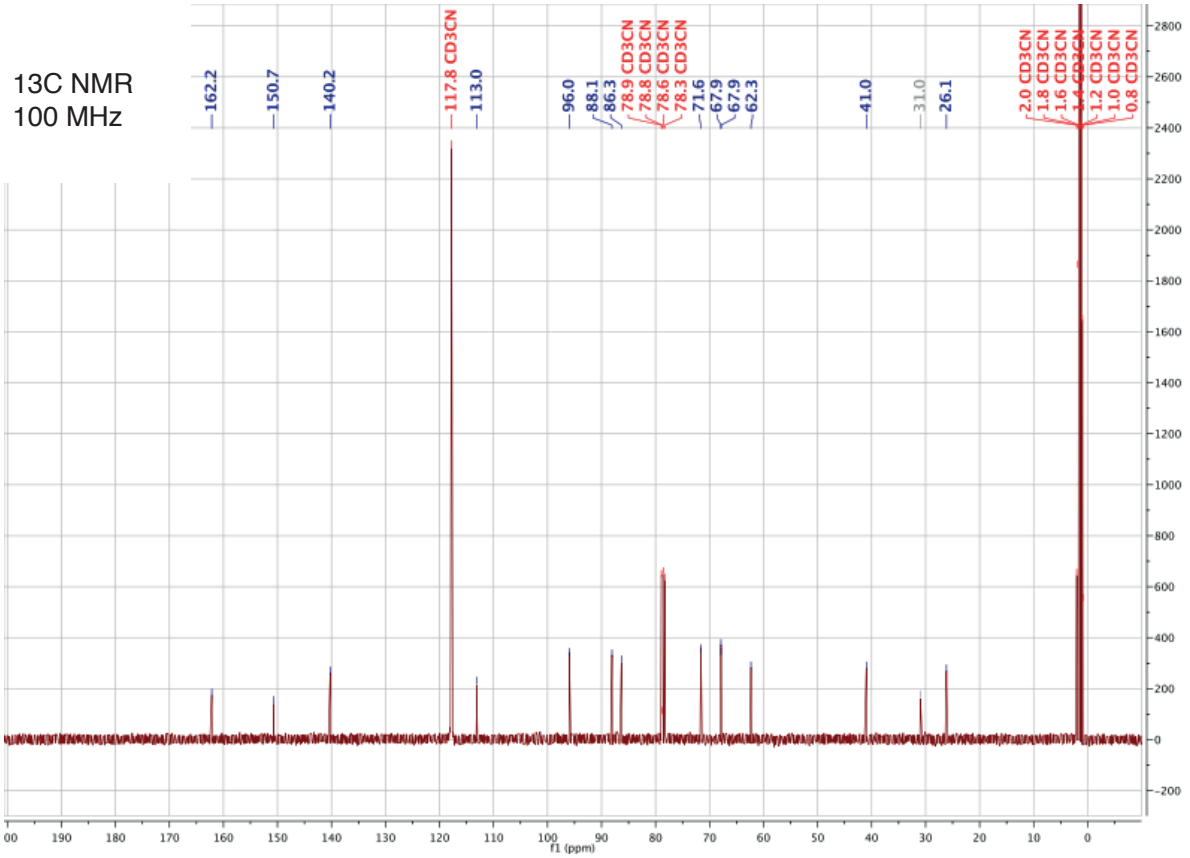


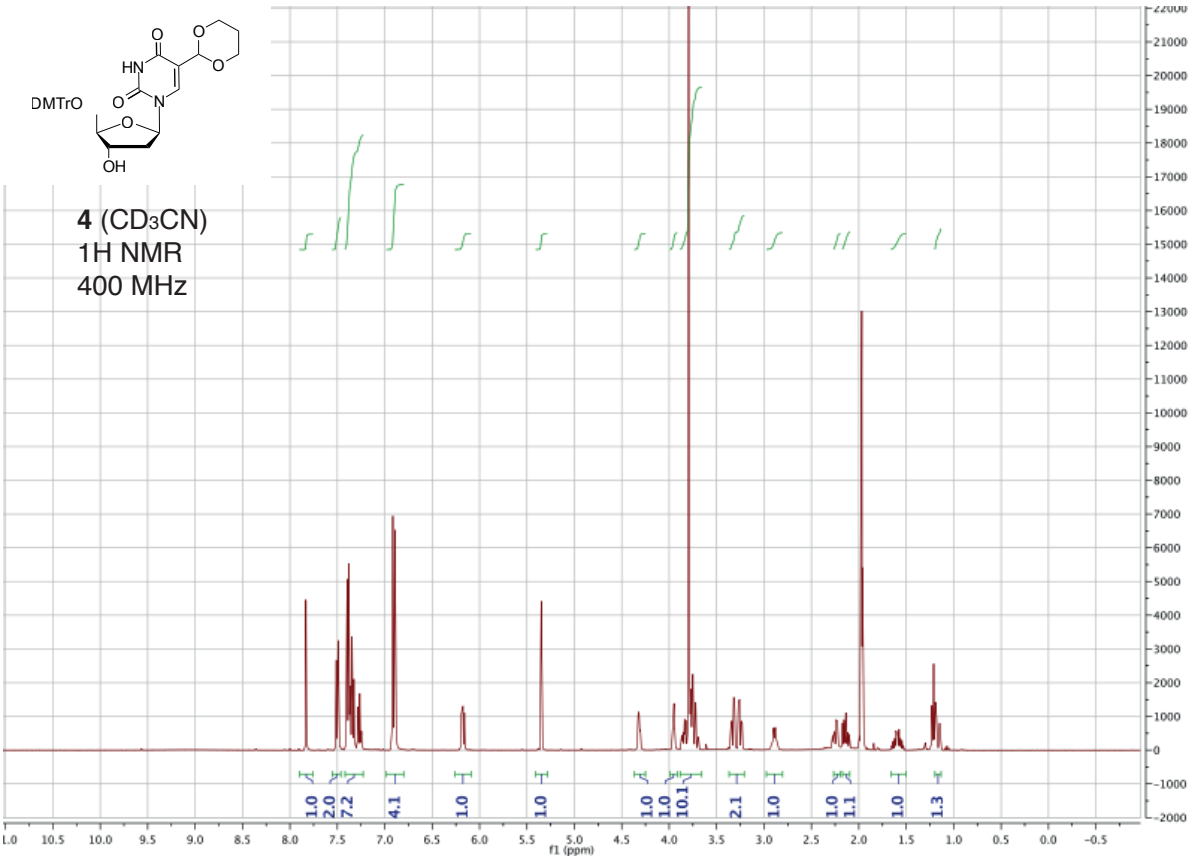
2 ($CDCl_3$)
1H NMR
400 MHz

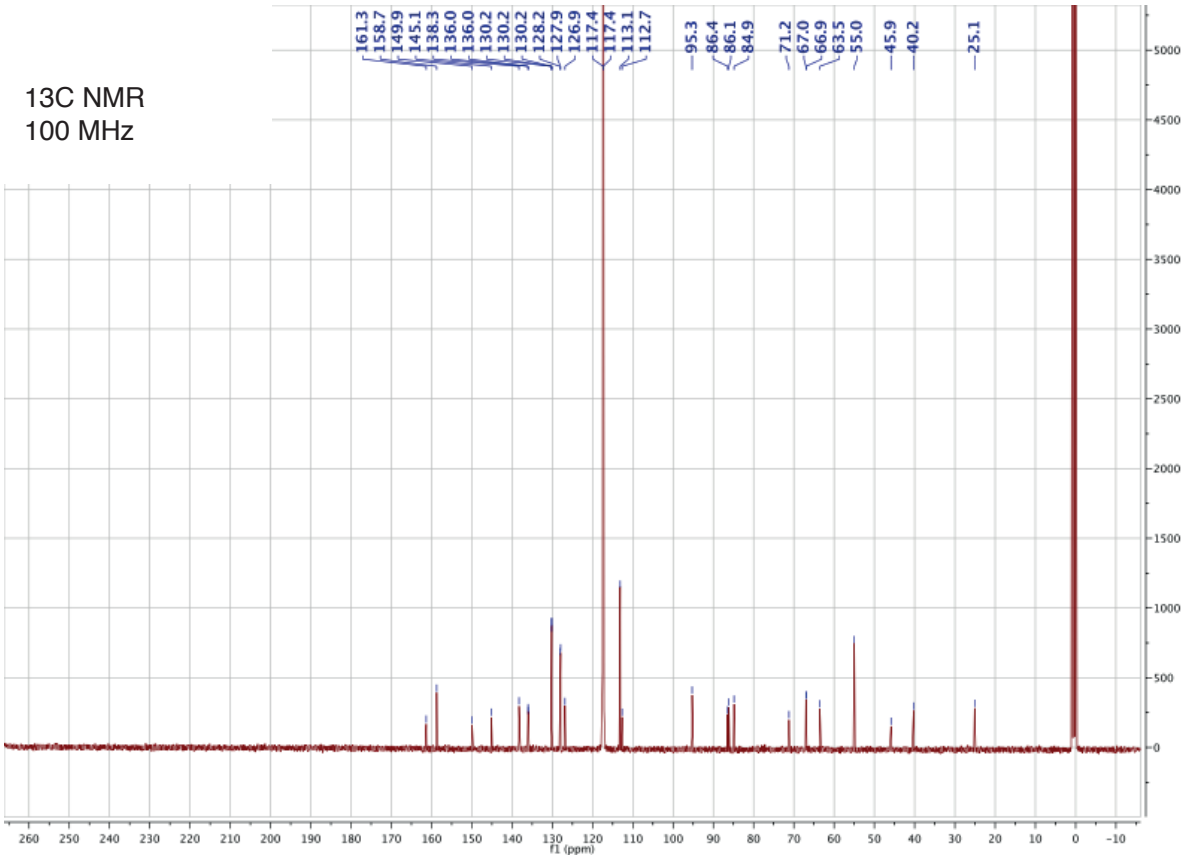
13C NMR
100 MHz

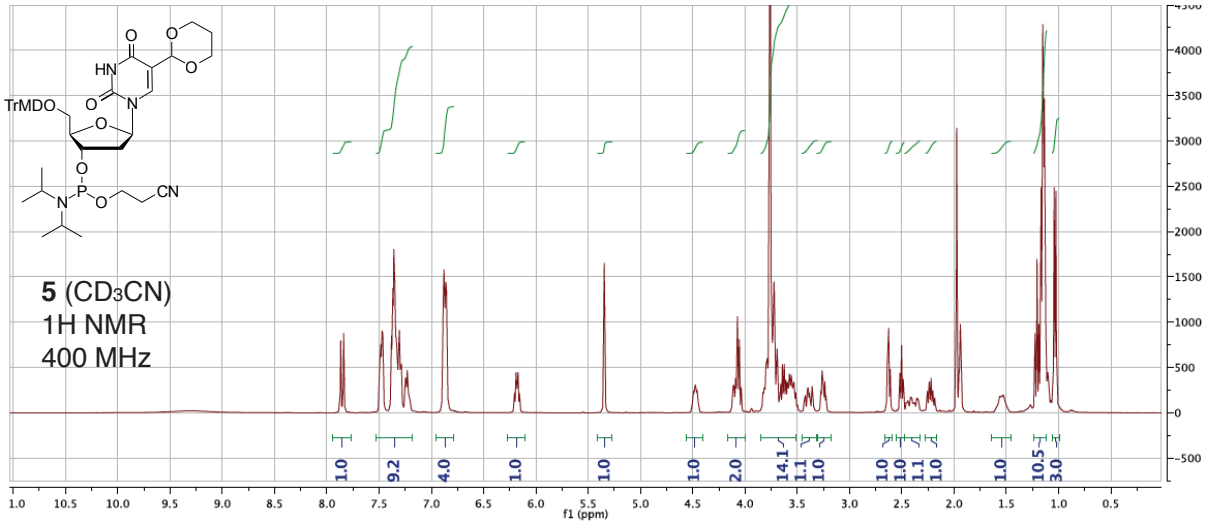


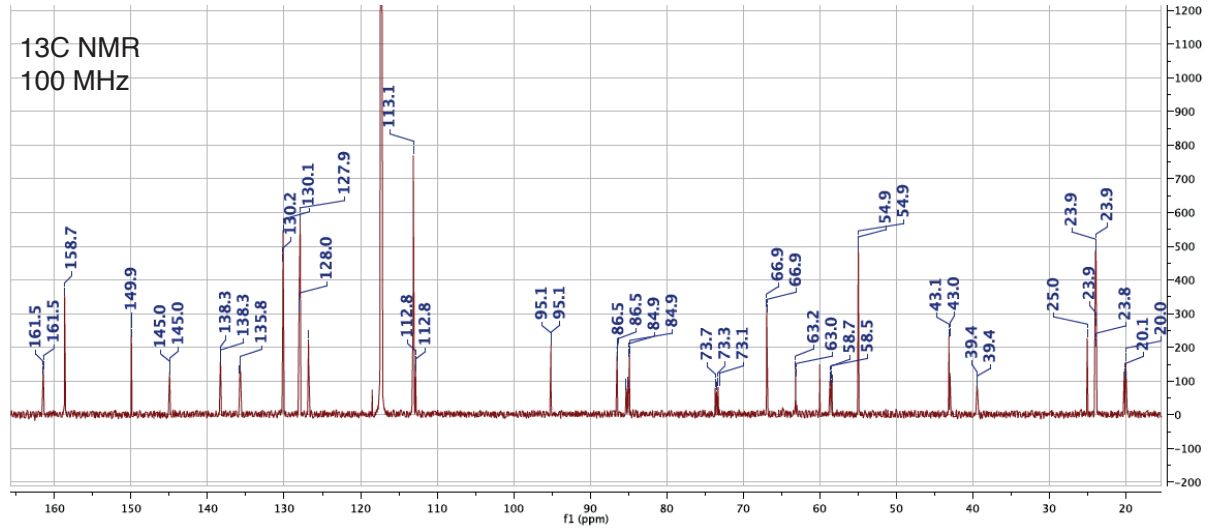

HO
HN
O
N
O
O
O
OH
3 (CDCl3-CD3CN)
1H NMR
400 MHz
1.0
1.9
3.0
2.1
1.2
0.9
0.9
1.0
1.0
1.0
1.9
3.0
2.1
1.2
1.0
1.0
f1 (ppm)

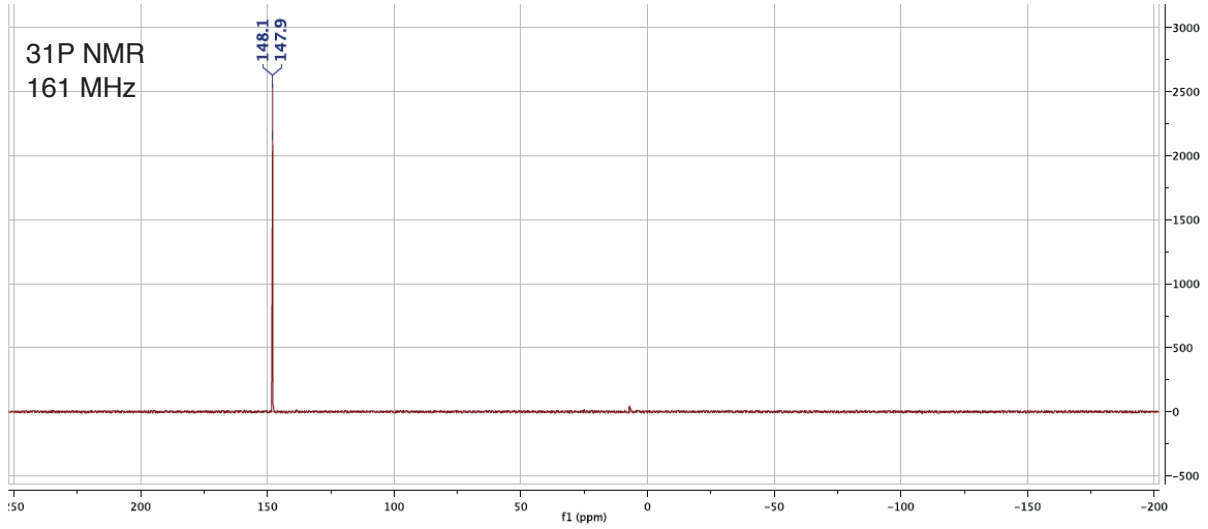

13C NMR
100 MHz
162.2
150.7
140.2
117.8 CD3CN
113.0
96.0
88.1
86.3
78.9 CD3CN
78.8 CD3CN
78.6 CD3CN
78.3 CD3CN
71.6
67.9
67.9
62.3
41.0
31.0
26.1
2.0 CD3CN
1.8 CD3CN
1.6 CD3CN
1.4 CD3CN
1.2 CD3CN
1.0 CD3CN
0.8 CD3CN
f1 (ppm)


DMTrO
HN
O
N
O
OH
4 (CD3CN)
1H NMR
400 MHz
1.0
2.0
7.2
4.1
1.0
1.0
1.0
1.0
10.1
2.1
1.0
1.0
1.1
1.0
1.3
f1 (ppm)


13C NMR
100 MHz
161.3
158.7
149.9
145.1
138.3
136.0
136.0
130.2
130.2
130.2
128.2
127.9
126.9
117.4
117.4
113.1
112.7
95.3
86.4
86.1
84.9
71.2
67.0
66.9
63.5
55.0
45.9
40.2
25.1
f1 (ppm)


TrMDO
HN
O
N
O
O
O
O
O
P
N
O
CN
5 (CD3CN)
1H NMR
400 MHz
1.0
9.2
4.0
1.0
1.0
1.0
2.0
14.1
1.1
1.0
1.0
1.0
1.1
1.0
1.0
10.5
3.0
f1 (ppm)


13C NMR
100 MHz
161.5
161.5
158.7
149.9
145.0
145.0
138.3
138.3
135.8
130.2
130.1
127.9
128.0
113.1
112.8
112.8
95.1
95.1
86.5
86.5
84.9
84.9
73.7
73.3
73.1
66.9
66.9
63.2
63.0
58.7
58.5
54.9
54.9
43.1
43.0
39.4
39.4
25.0
23.9
23.9
23.9
23.8
20.1
20.0
f1 (ppm)


31P NMR
161 MHz
148.1
147.9
f1 (ppm)