Supplementary Information

Structural Basis of Interstrand Cross-link Repair by O^6-Alkylguanine DNA Alkyltransferase

Alexey Y. Denisov[a†], Francis P. McManus[ab†], Derek K. O'Flaherty[ac], Anne M. Noronha[a], Christopher J. Wilds[a]*

[a]Department of Chemistry and Biochemistry, Concordia University, 7141 Sherbrooke St. West, Montréal, QC, Canada H4B 1R6
[b] Current address: Institute for Research in Immunology and Cancer, Université de Montréal, 2950 Chemin de Polytechnique, Montréal, Québec, Canada.
[c] Current address: Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, Massachusetts, United States.

† These authors contributed equally to the paper as first authors.

Contents

Supplementary Methods

Chemical Synthesis of Modified Nucleosides. The solution and solid-phase synthesis of **XLGG7** has been described previously.[1] 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-4-triazolyl-2'-deoxyuridine, 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-O^4-(hydroxybutyl)-2'-deoxyuridine and 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-O^4-(hydroxyheptyl)-2'-deoxyuridine were prepared according to published procedures. [2,3]

1- {O^4-[3'-O-(t-butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-4-{O^4-[3'-O-(t-butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-butane (**1a**)

To an ice cold solution of 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-O^4-(hydroxybutyl)-2'-deoxyuridine (0.500 g, 0.697 mmol) in acetonitrile (6 mL) was added 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-4-triazolyl-2'-deoxyuridine (0.582 g, 0.837 mmol) followed by the dropwise addition of DBU (0.313 mL, 2.09 mmol). The reaction was allowed to progress on ice for 5 min and then warmed up to room temperature. After 4 days, the crude product was taken up in CH_2Cl_2 (100 mL), the solution washed with two portions of sodium bicarbonate (2 x 100 mL) followed by one portion of brine (100 mL). The organic layer was dried over sodium sulphate and concentrated to produce a yellow gum. The crude product was purified by flash column chromatography using a hexanes : ethyl acetate (2 : 3) solvent system to give 0.790 g (84%) of product as a colorless foam. R_f (SiO_2 TLC): 0.29 hexane / ethyl acetate (2 : 3). $\lambda_{max(MeCN)}$ = 277 nm. ^{1}H NMR (500MHz, $CDCl_3$, ppm): 8.27 (d, 2H, H6), 7.28-7.44 (m, 20H, Ar), 6.87-6.90 (m, 6H, Ar), 6.31 (dd, 2H, H1', J = 4 Hz), 5.60 (d, 2H, H5), 4.52 (m, 2H, H3'), 4.44 (m, 4H, $ArOCH_2$), 3.99 (m, 2H, H4'), 3.84 (s, 12H, OCH_3), 3.57 (m, 2H, H5'), 3.36 (m, 2H, H5'), 2.55 (m, 2H, H2'), 2.25 (m, 2H, H2'), 1.88 (m, 4H, $(CH_2)_2$), 0.85 (s, 18H, $SiC(CH_3)_3$), 0.04 (s, 12H, $Si(CH_3)_2$). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): 171.3, 158.7, 155.8, 144.3, 142.8, 135.4, 135.3, 130.1, 130.1, 128.2, 127.9, 127.1, 113.21, 113.19, 95.3, 86.7, 86.4, 86.3, 70.3, 66.6, 61.8, 55.2, 42.1, 25.7, 25.2, 17.9, -4.62, -5.01. IR (thin film); ν_{max} (cm^{-1}) = 2954, 2930, 2856, 2362. 2238, 1670, 1607, 1540, 1509, 1460, 1300, 1251, 1177, 1108, 1079, 1035, 835, 783 731. HRMS (*m/z*): Observed 1343.6376, calculated for $C_{76}H_{95}N_4O_{14}Si_2^+$ 1343.6383 $[M+H]^+$

1- {O^4-[3'-O-(t-butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-7-{O^4-[3'-O-(t-butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-heptane (**1b**)

To an ice cold solution of 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-O^4-(hydroxyheptyl)-2'-deoxyuridine (0.510 g, 0.672 mmol) in acetonitrile (6.0 mL) was added 3'-*O*-(*t*-butyldimethylsilyl)-5'-O-(4,4'-dimethoxytrityl)-4-triazolyl-2'-deoxyuridine (0.582 g, 0.837 mmol) followed by the dropwise addition of DBU (0.313 mL, 2.09 mmol). The reaction allowed to progress on ice for 5 min and then warmed up to room temperature. After 4 days, the crude product was taken up in CH_2Cl_2 (100 mL), the solution washed with two portions of sodium bicarbonate (2 x 100 mL) followed by one portion of brine (100 mL). The organic layer was dried over sodium sulphate and concentrated to produce a yellow gum. The crude product was purified by flash column chromatography using a hexanes : ethyl acetate (1 : 1) solvent system to afford 0.776 g (83%) of pure product as a colorless foam. R_f (SiO_2 TLC): 0.47 hexane / ethyl acetate (2 : 3). $\lambda_{max(MeCN)}$ = 276 nm. ^{1}H NMR (500MHz, $CDCl_3$, ppm): 8.22 (d, 2H, H6), 7.22-7.41 (m, 20H, Ar), 6.83-6.86 (m, 6H, Ar), 6.28 (dd, 2H, H1', J= 4.0 Hz), 5.57 (d, 2H, H5), 4.48 (m, 2H, H3'), 4.35 (m, 4H, $ArOCH_2$), 3.96 (m, 2H, H4'), 3.80 (s, 12H, OCH_3), 3.54 (m, 2H, H5'), 3.32 (m, 2H, H5'), 2.51 (m, 2H, H2'), 2.23 (m, 2H, H2'), 1.73 (m, 4H, $(CH_2)_2$), 1.39 (m, 6H, $(CH_2)_3$), 0.81 (s, 18H, $SiC(CH_3)_3$), 0.01 (s, 12H, $Si(CH_3)_2$). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): 171.5, 158.6, 155.8, 144.3, 142.6, 135.4, 135.3, 130.12, 130.09, 128.2, 127.9, 127.1, 113.20, 113.18, 95.4, 86.7, 86.4, 86.3, 70.3, 67.2, 61.8, 55.2, 42.1, 29.0, 28.5, 25.8, 25.7, 17.9, -4.63, -5.01. IR (thin film); ν_{max} (cm^{-1}) = 2955, 2928, 2855, 2362, 2335, 1675, 1628, 1608, 1540, 1507, 1472, 1301, 1251, 1177, 1109, 1078, 1034, 835, 782, 726. HRMS (*m/z*): Observed 1385.6847, calculated for $C_{79}H_{101}N_4O_{14}Si_2^+$ 1385.6853 $[M+H]^+$

1-{O^4-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-4-{O^4-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-butane (**2a**)

To a solution of **1a** (0.767 g, 0.570 mmol) in THF (7.0 mL) was added dropwise TBAF (1 M in THF, 2.57 mL, 2.57 mmol). After 10 min the solvent was evaporated in vacuo and the crude product taken up in CH_2Cl_2 (50 mL). The solution was washed with sodium bicarbonate (2 x 150 mL) followed by a back extraction of the sodium bicarbonate washes with CH_2Cl_2 (2 x 5 mL). The organic layer was dried over sodium sulphate and concentrated to produce a yellow gum. The crude product was purified by flash column chromatography using a CH_2Cl_2 / methanol solvent system (49 : 1 → 9 : 1) to afford 0.634 g (> 99 %) of product as a colorless foam. R_f (SiO_2 TLC): 0.24 CH_2Cl_2 / methanol (19 : 1). $\lambda_{max(MeCN)}$ = 276 nm.^{1}H NMR (500MHz, $CDCl_3$, ppm): 8.06 (d, 2H, H6), 7.19-7.37 (m, 20H, Ar), 6.80-6.82 (m, 6H, Ar), 6.24 (d, 2H, H1'), 5.60 (d, 2H, H5), 4.47 (m, 2H, H3'), 4.37 (m, 4H, $ArOCH_2$), 4.03 (m, 2H, H4'), 3.78 (s, 12H, OCH_3), 3.50 (m, 2H, H5'), 3.41 (m, 2H, H5'), 2.59 (m, 2H, H2'), 2.23 (m, 2H, H2'), 2.00 (m, 2H, OH), 1.83 (m, 4H, $(CH_2)_2$). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): 171.4, 158.6, 155.8, 144.4, 142.7, 135.4, 135.3, 130.0, 128.1, 128.0, 127.1, 113.3, 95.5, 86.9, 86.5, 85.9, 71.0, 66.6, 62.7, 55.2, 41.9, 25.2. IR (thin film); ν_{max} (cm^{-1}) = 3381, 3056, 2956, 2836, 2362, 2335, 1652, 1616, 1608, 1540, 1508, 1474, 1302, 1250, 1220, 1177, 1035, 828, 790, 735. HRMS (*m/z*): Observed 1115.4658, calculated for $C_{64}H_{67}N_4O_{14}^+$ 1115.4654 $[M+H]^+$

1-{O^4-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-7-{O^4-[5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-heptane (**2b**)

To a solution of **1b** (0.575 g, 0.415 mmol) in THF (4.0 mL) was added dropwise TBAF (1 M in THF, 1.87 mL, 1.87 mmol). After 10 min the solvent was evaporated in vacuo and the crude product taken up in CH_2Cl_2 (50 mL). The solution was washed with sodium bicarbonate (150 mL) followed by brine (150 mL). The aqueous washes were back extracted with CH_2Cl_2 (2 x 5 mL). The organic layer was dried over sodium sulphate and concentrated to produce a yellow gum. The crude product was purified by flash column chromatography using a CH_2Cl_2 / methanol solvent system (97 : 3 → 19 : 1) to afford 0.476 g (> 99 %) of product as a colorless foam. R_f (SiO_2 TLC): 0.29 CH_2Cl_2 / methanol (19 : 1). $\lambda_{max(MeCN)}$ = 276 nm. ^{1}H NMR (500MHz, $CDCl_3$, ppm): 8.06 (d, 2H, H6), 7.20-7.37 (m, 20H, Ar), 6.79-6.82 (m, 6H, Ar), 6.25 (dd, 2H, H1', J = 5.5 Hz), 5.60 (d, 2H, H5), 4.48 (m, 2H, H3'), 4.32 (m, 4H, $ArOCH_2$), 4.03 (m, 2H, H4'), 3.77 (s, 12H, OCH_3), 3.48 (m, 2H, H5'), 3.40 (m, 2H, H5'), 2.60 (m, 2H, H2'), 2.29 (d, 2H, OH), 2.25 (m, 2H, H2'), 1.71 (m, 4H, $(CH_2)_2$), 1.38 (m, 6H, $(CH_2)_3$). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): 171.5, 158.6, 156.0, 144.4, 142.6, 135.44, 135.35, 130.0, 128.1, 127.9, 127.0, 113.2, 95.6, 86.9, 86.5, 85.9, 70.9, 67.2, 62.7, 55.2, 41.9, 28.8, 28.4, 25.7. IR (thin film); ν_{max} (cm^{-1}) = 3378, 3056, 2932, 2836, 2363, 2335, 1653, 1607, 1540, 1508, 1456, 1305, 1250, 1220, 1177, 1099, 1034, 960, 828, 735. HRMS (*m/z*): Observed 1157.5120, calculated for $C_{67}H_{73}N_4O_{14}^+$ 1157.5123 $[M+H]^+$

1-{O^4-[3'-O-(β-cyanoethyl-N,N'-diisopropyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-4-{O^4-[3'-O-(β-cyanoethyl-N,N'-diisopropyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-butane (**3a**)

To a solution of **2a** (0.614 g, 0.551 mmol) in THF (3.0 mL) was added diisopropylethylamine (288 μL, 1.653 mmol) followed by *N*,*N*-diisopropylamino cyanoethyl phosphonamidic chloride (308 μL, 1.381 mmol) dropwise. After 30 min, the solvent was evaporated in vacuo and the crude product taken up in ethyl acetate (50 mL). The solution was then washed with two portions of sodium bicarbonate (2 x 100 mL) and once with brine (100 mL). The organic layer was dried over sodium sulphate and concentrated. The product, a colorless powder, was dissolved in 3 mL of ethyl acetate and precipitated from 1 L of hexanes at -78 °C (0.757 g, 91 %). R_f (SiO_2 TLC): 0.24, 0.44, 0.58 ethyl acetate. $\lambda_{max(MeCN)}$ = 276 nm. ^{1}H NMR (500MHz, $CDCl_3$, ppm): 8.13-8.08 (m, 2H, 2 x H6), 7.50-7.47 (m, 4H, Ar), 7.38-7.31 (m, 12H, Ar), 7.27-7.23 (m, 2H, Ar), 6.93-

6.88 (m, 8H, Ar), 6.27-6.22 (m, 2H, H1'), 5.70-5.69 (m, 2H, 2 x H5), 4.77-4.69 (m, 2H, H3'), 4.35-4.33 (m, 4H, $ArOCH_2$), 4.25-4.19 (m, 2H, H4'), 3.91-3.60 (m, 20H, 2 x CH_2OP, 4 x CHN & 4 x OCH_3), 3.49-3.48 (m, 2H, 2 x H5'), 3.46 (m, 2H, 2 x H5''), 2.78-2.75 (m, 2H, CH_2CN), 2.67-2.58 (m, 4H, 2 x H2' & CH_2CN), 2.39-2.33 (m, 2H, 2 x H2''), 1.86-1.84 (m, 4H, CH_2CH_2), 1.22-1.18 (m, 18H, 6 x CH_3), 1.12-1.11 (m, 6H, 2 x CH_3). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): 171.12, 171.10, 170.0, 158.8, 154.6, 144.92, 144.90, 143.33, 143.25, 135.60, 135.58, 135.47, 135.43, 130.17, 130.15, 130.14, 128.14, 128.11, 127.82, 127.81, 126.84, 126.82, 118.1, 117.9, 113.1, 94.4, 94.3, 86.61, 86.58, 86.40, 86.35, 85.49, 85.46, 85.3, 85.2, 73.1, 73.0, 72.6, 72.4, 66.0, 62.8, 62.6, 59.6, 58.74, 58.67, 58.6, 58.5, 54.68, 54.66, 43.1, 43.0, 40.43, 40.40, 40.3, 40.2, 25.07, 24.05, 24.02, 23.99, 23.96, 23.9, 22.3, 22.2, 19.94, 19.89, 19.85, 19.83, 19.80, 13.6. ^{31}P NMR (202.3 MHz, d_6-acetone, ppm): 148.3 and 148.2. HRMS (*m/z*): Observed 1515.6805, calculated for $C_{82}H_{101}N_8O_{16}P_2^+$ 1515.6811 $[M+H]^+$

1- {O^4-[3'-O-(β-cyanoethyl-N,N'-diisopropyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridinyl]}-7-{O^4-[3'-O-(β-cyanoethyl-N,N'-diisopropyl)-5'-O-(4,4'-dimethoxytrityl)- 2'-deoxyuridinyl]}-heptane (**3b**)

To a solution of **2b** (0.456 g, 0.394 mmol) in THF (3.0 mL) was added diisopropylethylamine (203 µL, 1.165 mmol) followed by *N*,*N*-diisopropylamino cyanoethyl phosphonamidic chloride (218 µL, 0.977 mmol) dropwise. After 30 min, the solvent was evaporated in vacuo and the crude product taken up in ethyl acetate (50 mL). The solution was then washed with two portions of sodium bicarbonate (2 x 100 mL) and once with brine (100 mL). The organic layer was dried over sodium sulphate and concentrated. The product, a colorless powder, was dissolved in 3 mL of ethyl acetate and precipitated from 1 L of hexanes at -78 °C (0.488 g, 80 %). R_f (SiO_2 TLC): 0.30, 0.44, 0.55 hexanes / ethyl acetate (2 : 8). $\lambda_{max(MeCN)}$ = 276 nm. ^1H NMR (500MHz, $CDCl_3$, ppm): ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): 8.12-8.07 (m, 2H, 2 x H6), 7.50-7.46 (m, 4H, Ar), 7.37-7.30 (m, 12H, Ar), 7.27-7.22 (m, 2H, Ar), 6.93-6.88 (m, 8H, Ar), 6.27-6.21 (m, 2H, H1'), 5.68-5.67 (m, 2H, 2 x H5), 4.77-4.69 (m, 2H, H3'), 4.30-4.27 (m, 4H, $ArOCH_2$), 4.25-4.18 (m, 2H, H4'), 3.92-3.59 (m, 20H, 2 x CH_2OP, 4 x CHN & 4 x OCH_3), 3.48-3.47 (m, 2H, 2 x H5'), 3.46-3.45 (m, 2H, 2 x H5''), 2.78-2.76 (m, 2H, CH_2CN), 2.67-2.57 (m, 4H, 2 x H2' & CH_2CN), 2.38-2.32 (m, 2H, 2 x H2''), 1.75-1.71 (m, 4H, 2 x CH_2), 1.46-1.43 (m, 6H, $CH_2CH_2CH_2$), 1.21-1.18 (m, 18H, 6 x CH_3), 1.12-1.11 (m, 6H, 2 x CH_3). ^{13}C NMR (125.7 MHz, $CDCl_3$, ppm): ^{13}C NMR (126 MHz, acetone) δ 171.19, 171.17, 158.8, 154.7, 144.92, 144.90, 143.3, 143.2, 135.60, 135.58, 135.48, 135.44, 130.16, 130.14, 130.13, 128.14, 128.11, 127.81, 127.79, 126.83, 126.80, 118.1, 117.9, 113.1, 94.43, 94.36, 86.61, 86.57, 86.4, 86.3, 85.5, 85.4, 85.24, 85.19, 73.1, 73.0, 72.6, 72.4, 66.4, 62.8, 62.6, 58.73, 58.66, 58.58, 58.51, 54.66, 54.64, 43.1, 43.0, 40.4, 40.2, 28.3, 25.6, 24.0, 24.00, 23.97, 23.94, 23.90, 19.87, 19.84, 19.82, 19.78. ^{31}P NMR (202.3 MHz, d_6-acetone, ppm): 148.3 and 148.2. HRMS (*m/z*): Observed 1557.7260, calculated for $C_{85}H_{107}N_8O_{16}P_2^+$ 1557.7280 $[M+H]^+$

Synthesis of Modified Oligonucleotides. Solid-phase synthesis of **XLUU4** and **XLUU7** was performed on an Applied Biosystems Model 3400 synthesizer on a 1 (**XLUU4**) or 2 µmol (**XLUU7**) scale employing standard β-cyanoethylphosphoramidite cycles supplied by the manufacturer with modifications to coupling times that are indicated below. Fast deprotecting 2'-deoxynucleoside-3′-*O*-phosphoramidites and the modified bis-2'-deoxynucleoside-3'-*O*-phosphoramidites (**3a** or **3b**) were dissolved in anhydrous acetonitrile to a concentration of 0.1 and 0.05 M, respectively. Sequence assembly was initiated by detritylation (3% TCA in CH_2Cl_2), followed by phosphoramidite coupling: Commercial 2'-deoxynucleoside-3′-*O*-phosphoramidites for 2 min and modified bis-2'-deoxynucleoside-3'-*O*-phosphoramidite (**3a** or **3b**) for 30 min; capping with phenoxyacetic anhydride/pyridine/tetrahydrofuran (1:1:8, v/v/v; solution A, and 1-methyl-1H-imidazole/tetrahydrofuran 16:84 w/v; solution B) and oxidation (0.02 M iodine in tetrahydrofuran/water/pyridine 2.5:2:1). Final removal of the 5'-terminal trityl group was carried out on the synthesizer.

The solid support was removed from the column, placed into a microfuge tube and 500 μL of a 10% DBU in *n*-propanol solution added per μmole of support bound oligonucleotide. Deprotection and cleavage of the oligonucleotide proceeded for 5 days in the dark at room temperature. This was followed by neutralization of the DBU with an equimolar amount of acetic acid and the oligonucleotides solubilized in acetonitrile prior to being transferred into clean vials to separate the solid support. The crude oligonucleotide mixture was dried down in a speed vacuum and then desalted using C-18 SEP PAK cartridges (Waters Inc.) prior to purification.

Purification of the desired oligonucleotide species was achieved by loading the crude material on a 20% 7 M urea denaturing polyacrylamide gels (19 : 1) using 1X TBE [89 mM Tris-HCl, 89 mM boric acid, 2 mM EDTA (pH 8.0)] as running buffer. The desired bands, observed by UV shadowing, were excised from the gel and the oligonucleotide extracted with 0.1 M sodium acetate. The extracted oligomers were then desalted using C-18 SEP PAK cartridges as previously described. [4]

ESI-MS and Nuclease Digestion of Modified Oligonucleotides. The identities of the purified oligonucleotides, **XLUU4** and **XLUU7**, were verified by mass spectrometry and nuclease digest.

ESI mass spectra for these oligonucleotides were obtained at the Concordia University Centre for Biological Applications of Mass Spectrometry on a LTQ Orbitrap Velos-ETD (Thermo Scientific) in full scan, negative ion mode.

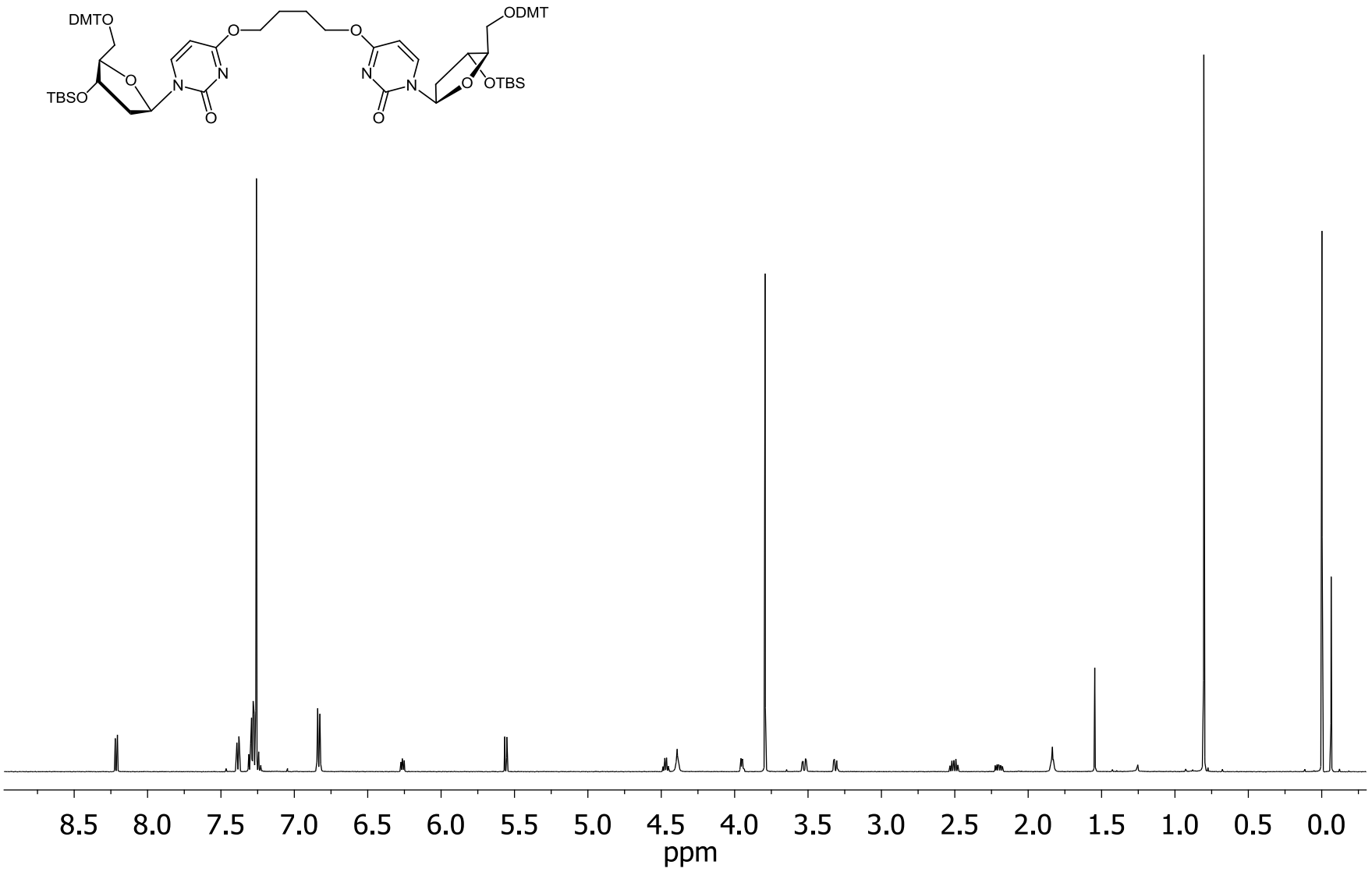
0.1 A_{260} units of **XLUU4** and **XLUU7** and a single-strand control DNA **U** (5'-CGAAAUTTTCG) were analyzed by exonuclease digestion (snake venom phosphodiesterase: 0.28 units and calf intestinal phosphatase: 5 units, in 10 mM Tris, pH 8.1 and 2 mM magnesium chloride) for 3 days at 37 °C. The resulting nucleosides were analyzed by reverse phase HPLC on a Symmetry® C-18 5μm column (0.46 x 15 cm) employing a linear gradient of 0-70% buffer B over 30 min (buffer A, 50 mM sodium phosphate, pH 5.8, 2% acetonitrile and buffer B, 50 mM sodium phosphate, pH 5.8, 50% acetonitrile). The retention times of the eluted peaks were compared to the standard nucleotides obtained which eluted at the following times: dC (4.3 min), dU (5.6 min), dG (6.8 min), dT (7.7 min), dA (9.0 min), O^4-2'-deoxyuridine-butylene-O^4-2'-deoxyuridine (17.8 min) and O^4-2'-deoxyuridine-heptylene-O^4-2'-deoxyuridine (26.5 min) and the ratio of nucleosides determined (see Supporting Information for the RP HPLC chromatographs).

UV Thermal Denaturation of the DNA Duplexes. Molar extinction coefficients for all oligonucleotides were determined by the nearest-neighbor method, using the thymidine coefficients for the O^4 modified 2'-deoxyuridines. For the control duplex consisting of **U** and its complementary strand (5'-CGAAAATTTCG), an equimolar amount of both were combined, lyophilized to dryness and resuspended in 1 mL of buffer composed of 90 mM sodium chloride, 10 mM sodium phosphate, 1 mM EDTA buffer (pH 7.0) to a final duplex concentration of 4 μM. For the cross-linked oligonucleotides, **XLUU4** and **XLUU7**, extinction coefficients for double stranded DNA were determined to prepare solutions containing 4 μM of duplex in the same buffer as the control sample. The samples were heated to 90 °C and held for 10 min before being cooled back to room temperature and finally placed at 4 °C overnight. Prior to analysis the samples were degassed for 2 min using a speed-vac. Thermal denaturation experiments were conducted at a heating rate of 0.5 °C min^{-1}, using a Varian CARY Model 3E spectrophotometer fitted with a 6-sample thermostated cell block and a temperature controller while monitoring absorbance at 260 nm. The data were processed as described by Puglisi and Tinoco. [5]

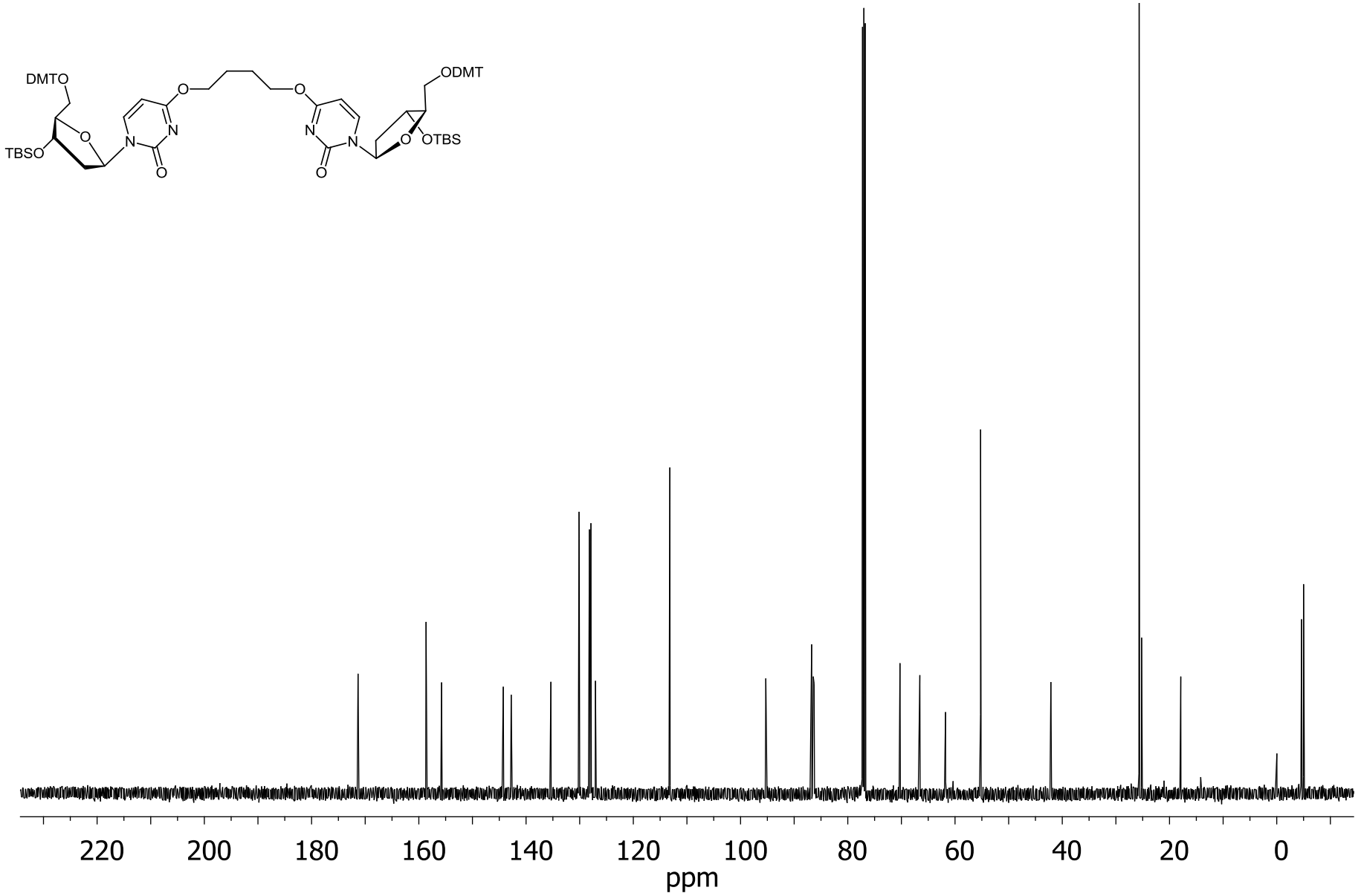
Circular Dichroism (CD) Spectroscopy of the DNA Duplexes. Samples for circular dichroism spectroscopy were prepared in the same manner as described for the UV thermal denaturation studies. Experiments were carried out on a Jasco J-815 spectropolarimeter equipped with a Julaba F25 circulating bath using fused quartz cells (Starna 29-Q-10). The spectra were an average of 5 scans run at a rate of 20 nm min^{-1}, a bandwidth of 1 nm and a sampling wavelength of 0.2 nm. Scans were performed between 320 and 220 nm at 10 °C.

AGT Preparation and Purification. All proteins were expressed through a T5 promoter of the pQE30 vector, except for the C146S Ada-C variant, which was expressed by the pET15b T7 promoter. [6] [7] [8] Cells containing a plasmid coding for AGT were grown in 1L of LB broth supplemented with the proper antibiotic(s) until an OD_{600} 0.6-0.8 was observed. Protein induction was achieved with 0.3 mM IPTG, the cells incubated for 4 h at 37 °C with shaking at 225 rpm and harvested by centrifugation at 6000 x g at 4 °C for 20 min. The cell pellets were resuspended in a buffer consisting of 20 mM Tris-HCl (pH 8.0), 250 mM NaCl, 20 mM β-mercaptoethanol supplemented with Complete, Mini, EDTA-free Protease Inhibitor Cocktail Tablets at 5 mL of buffer per gram of wet pellet. Cell lyses was achieved by two passes through a French press. The lysate was clarified by centrifugation at 17000 x g for 45 min at 4 °C. The clarified lysate was applied to a Ni-NTA column pre-equilibrated with 20 mM Tris-HCl (pH 8.0), 250 mM NaCl, 20 mM β-mercaptoethanol. The column was washed with the equilibration buffer supplemented with 10 mM imidazole and the protein eluted with the same buffer supplemented with 200 mM imidazole. Fractions containing protein were pooled and dialyzed against 50 mM Tris-HCl (pH 7.6), 250 mM NaCl, 5 mM DTT and 0.1 mM EDTA using 8000 Da cutoff dialysis tubing.

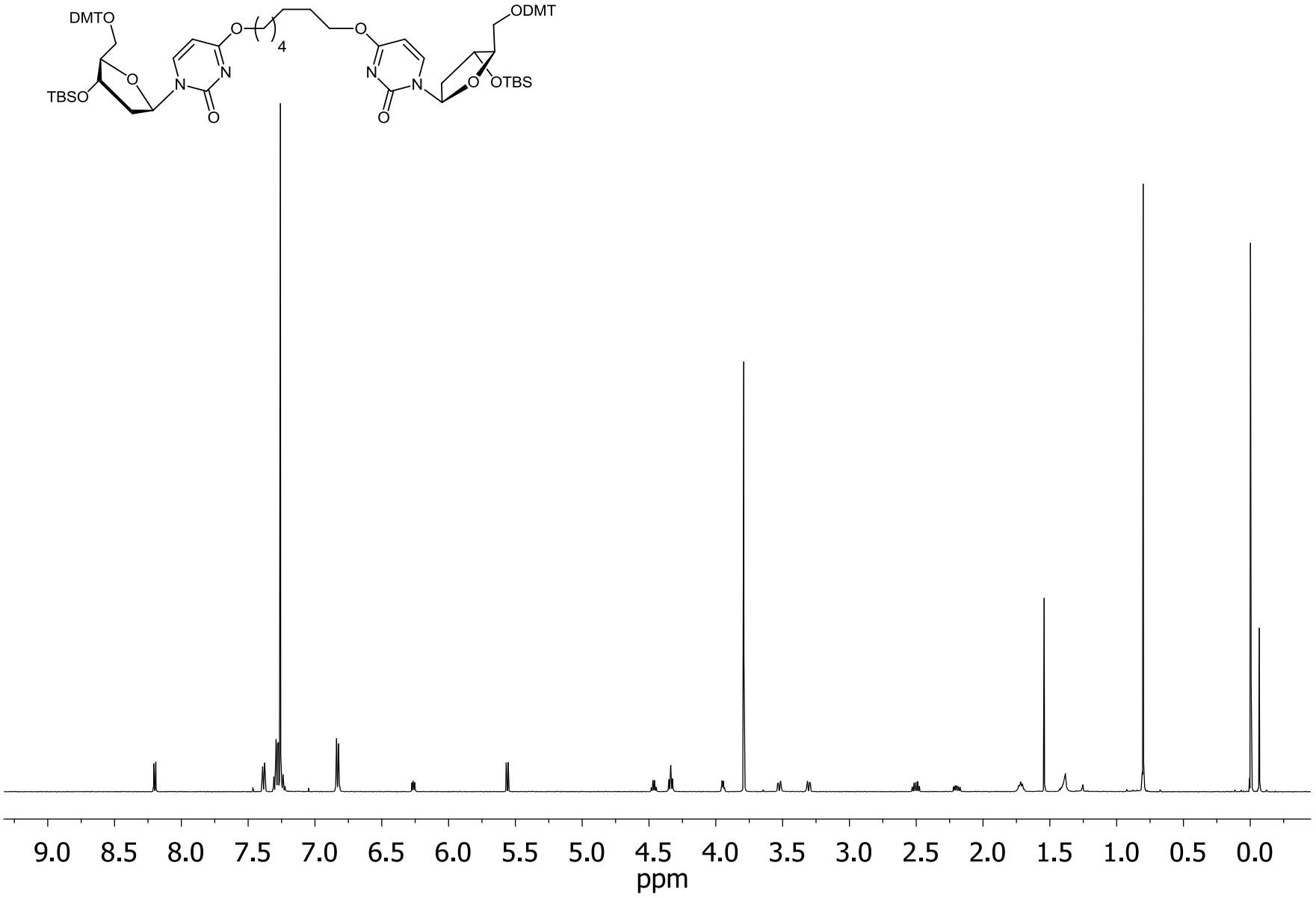
Supplementary Figure S1 - 500 MHz ^{1}H NMR spectrum of compound (**1a**) (in $CDCl_3$)



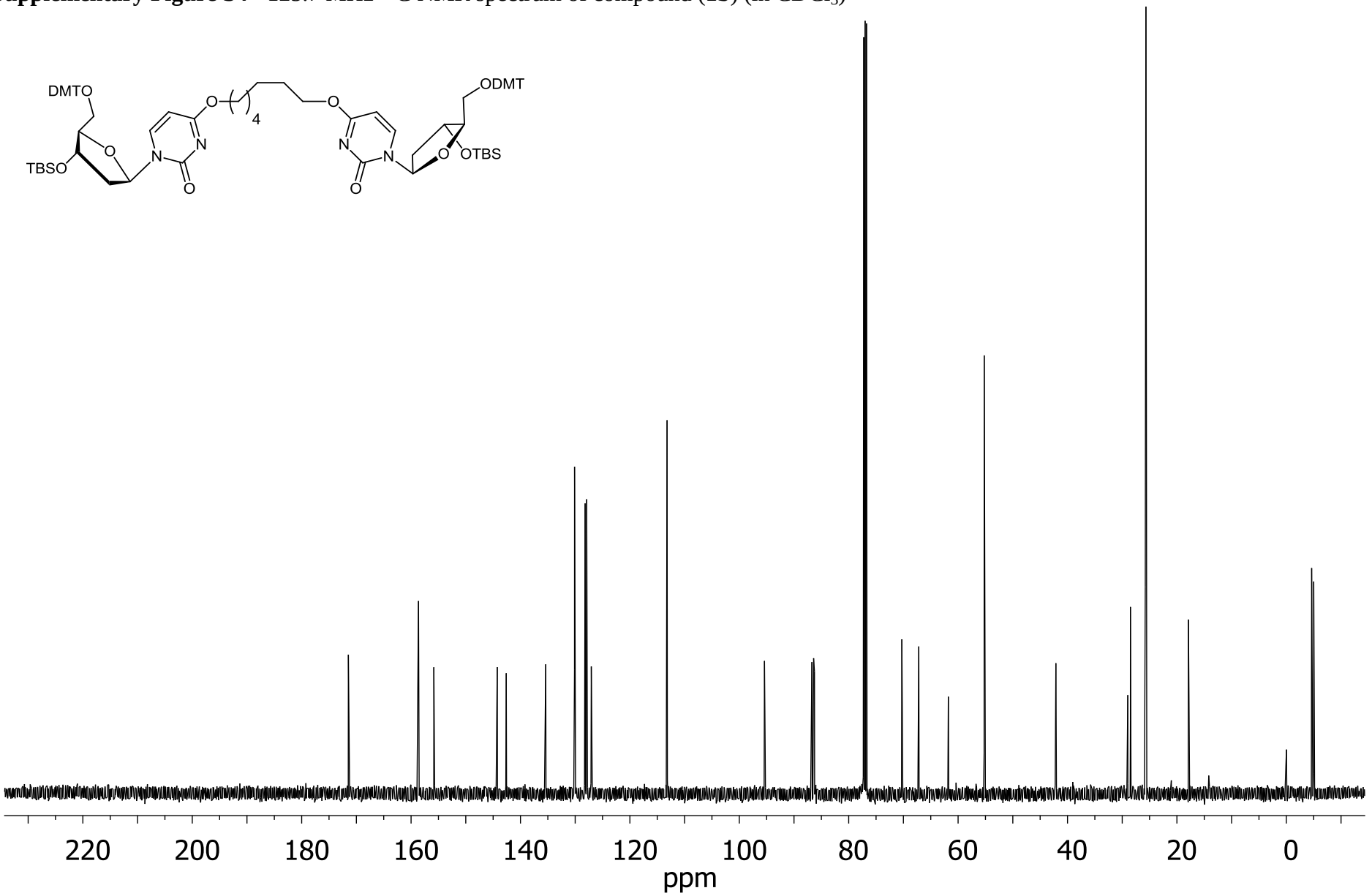
Supplementary Figure S2 - 125.7 MHz ^{13}C NMR spectrum of compound (**1a**) (in $CDCl_3$)



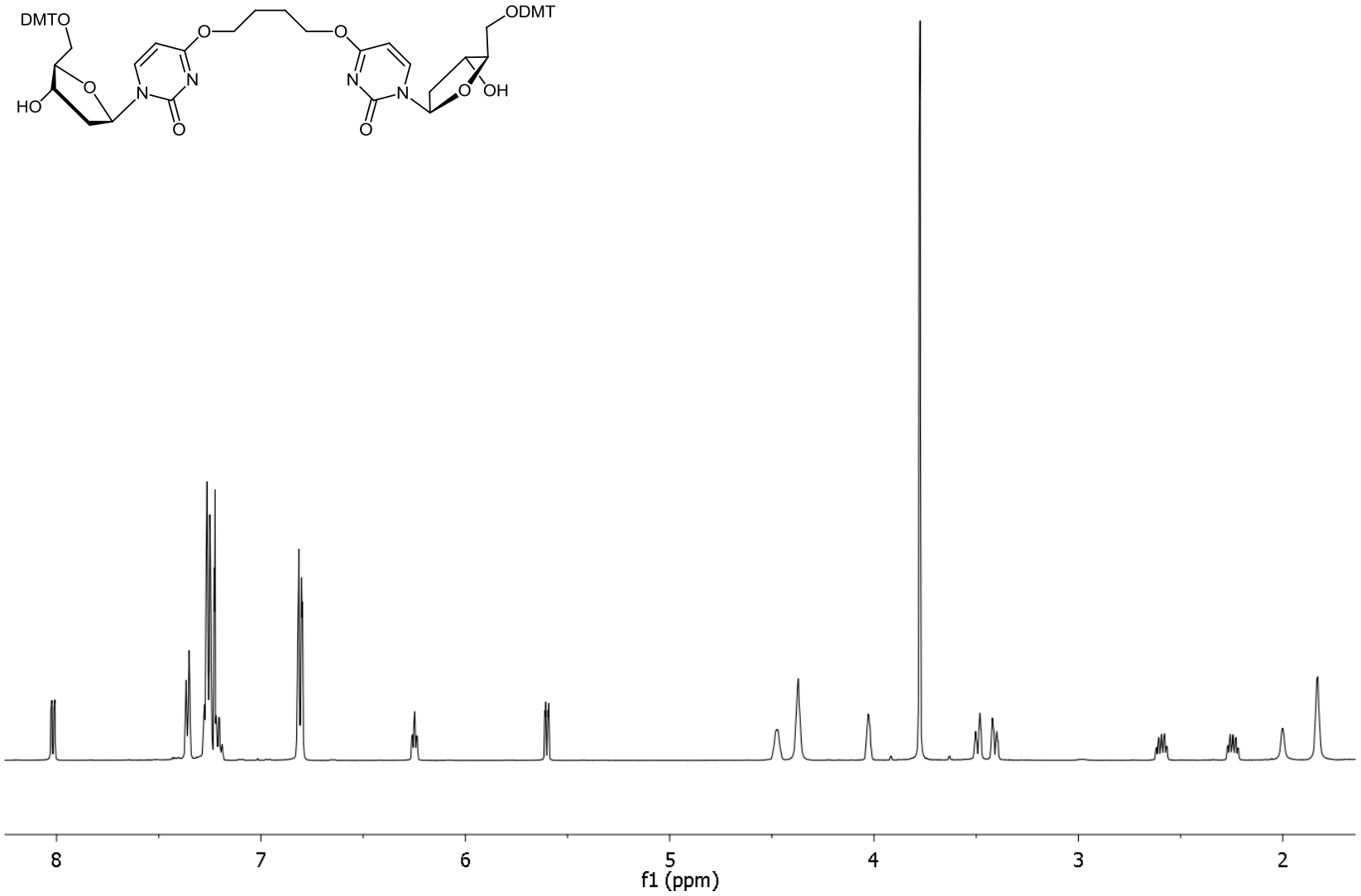
Supplementary Figure S3 - 500 MHz ^{1}H NMR spectrum of compound (**1b**) (in $CDCl_3$)



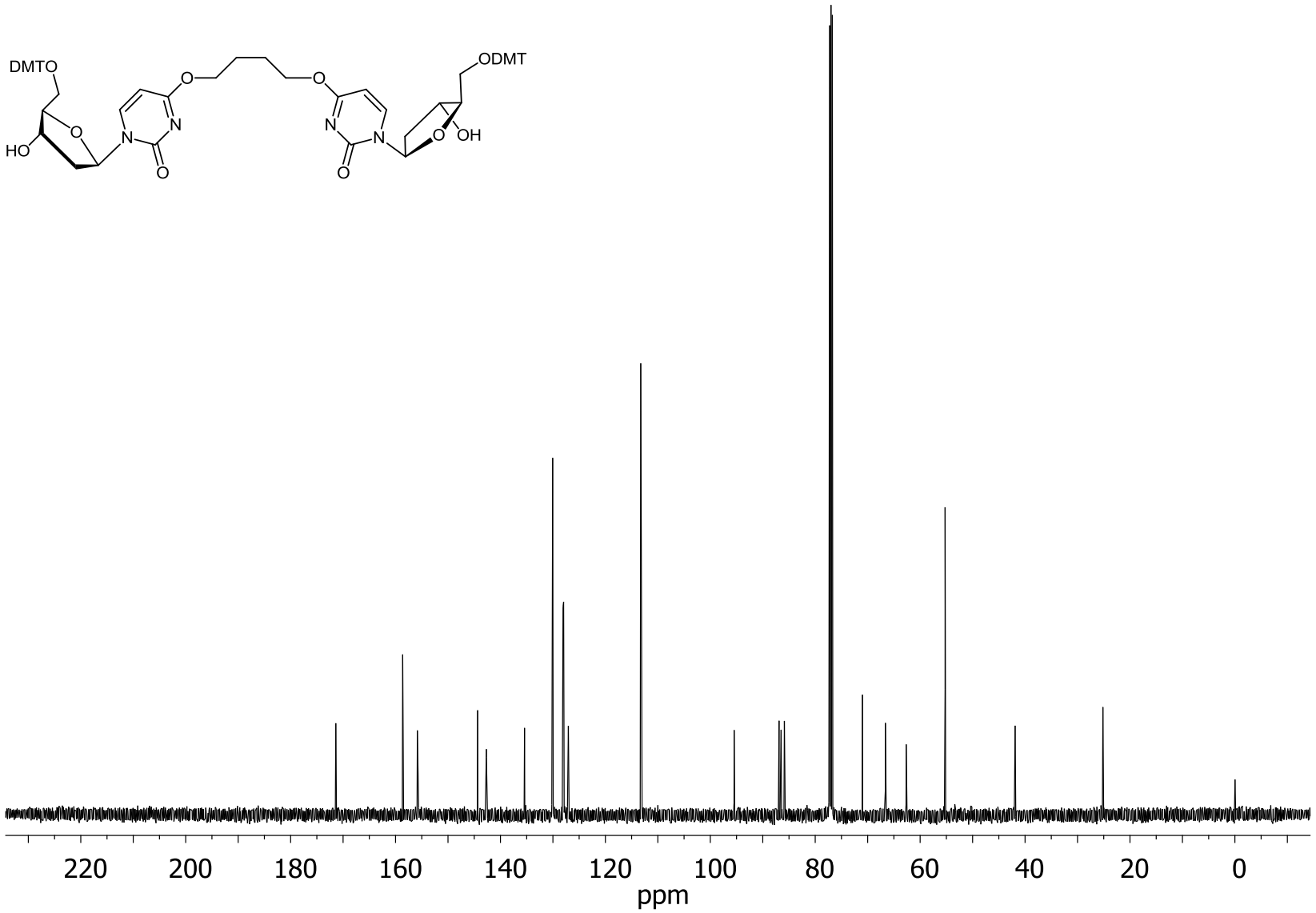
Supplementary Figure S4 - 125.7 MHz ^{13}C NMR spectrum of compound (**1b**) (in $CDCl_3$)



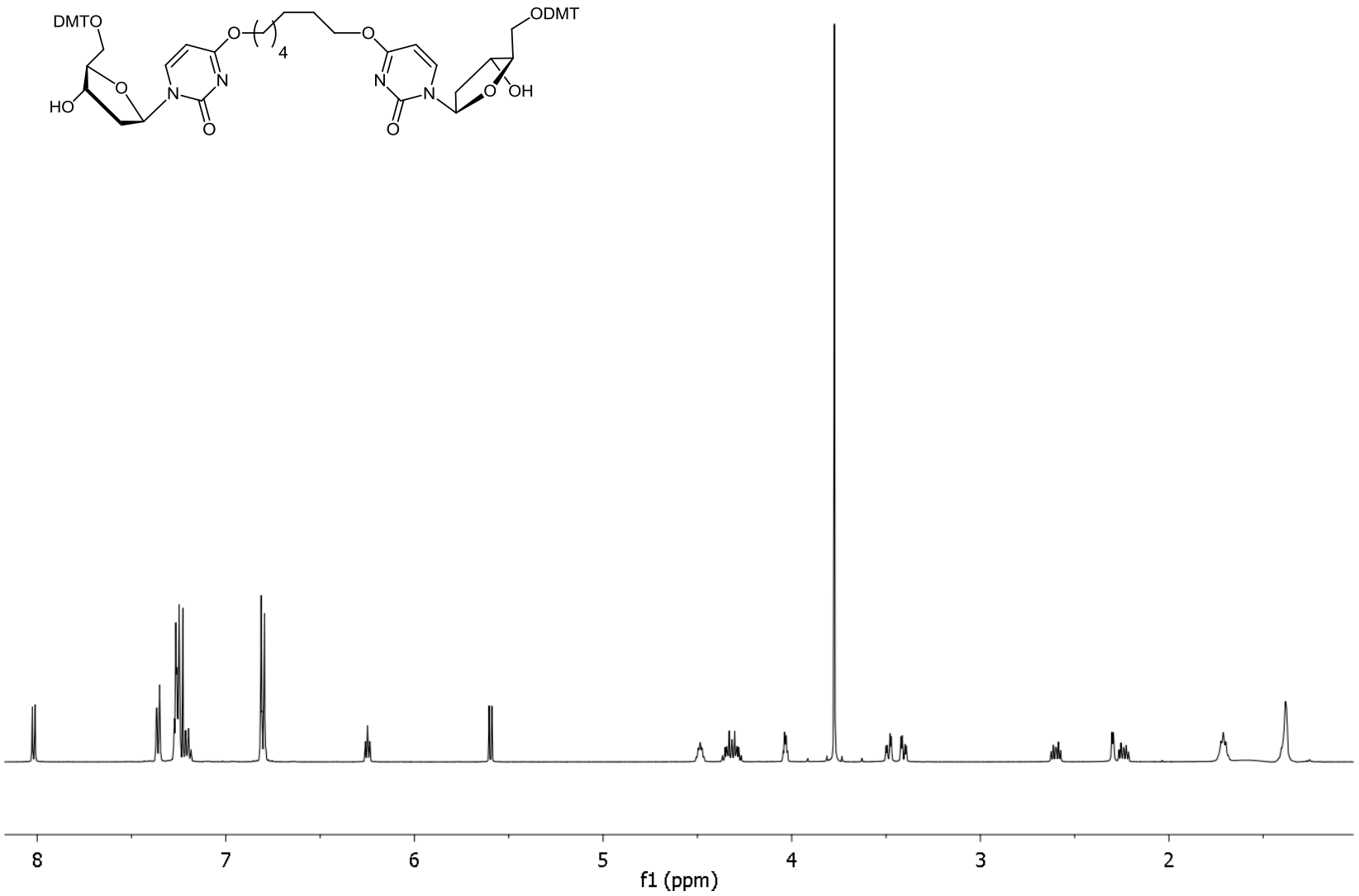
Supplementary Figure S5 - 500 MHz ^{1}H NMR spectrum of compound (**2a**) (in $CDCl_3$)



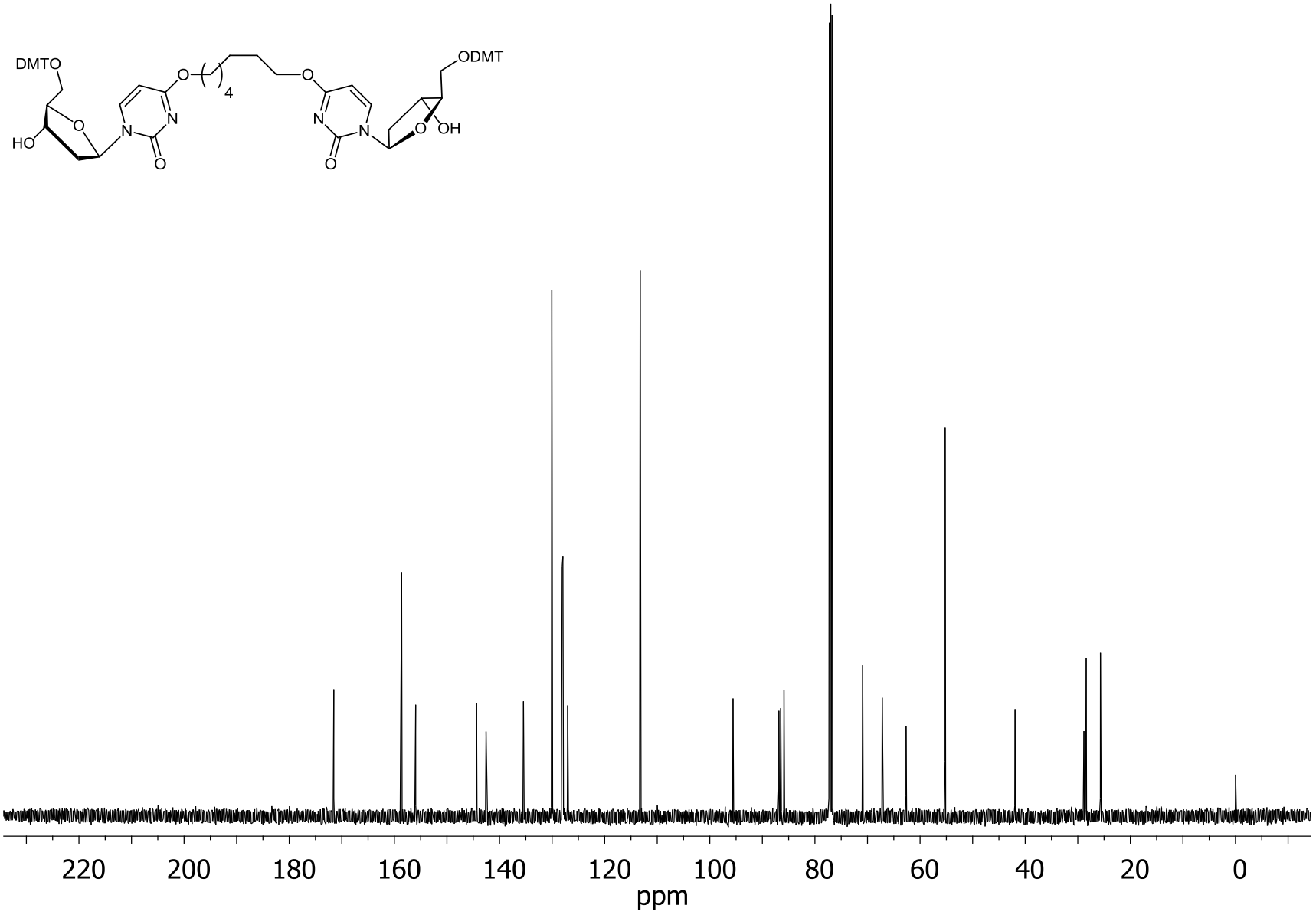
Supplementary Figure S6 - 125.7 MHz ^{13}C NMR spectrum of compound (**2a**) (in $CDCl_3$)



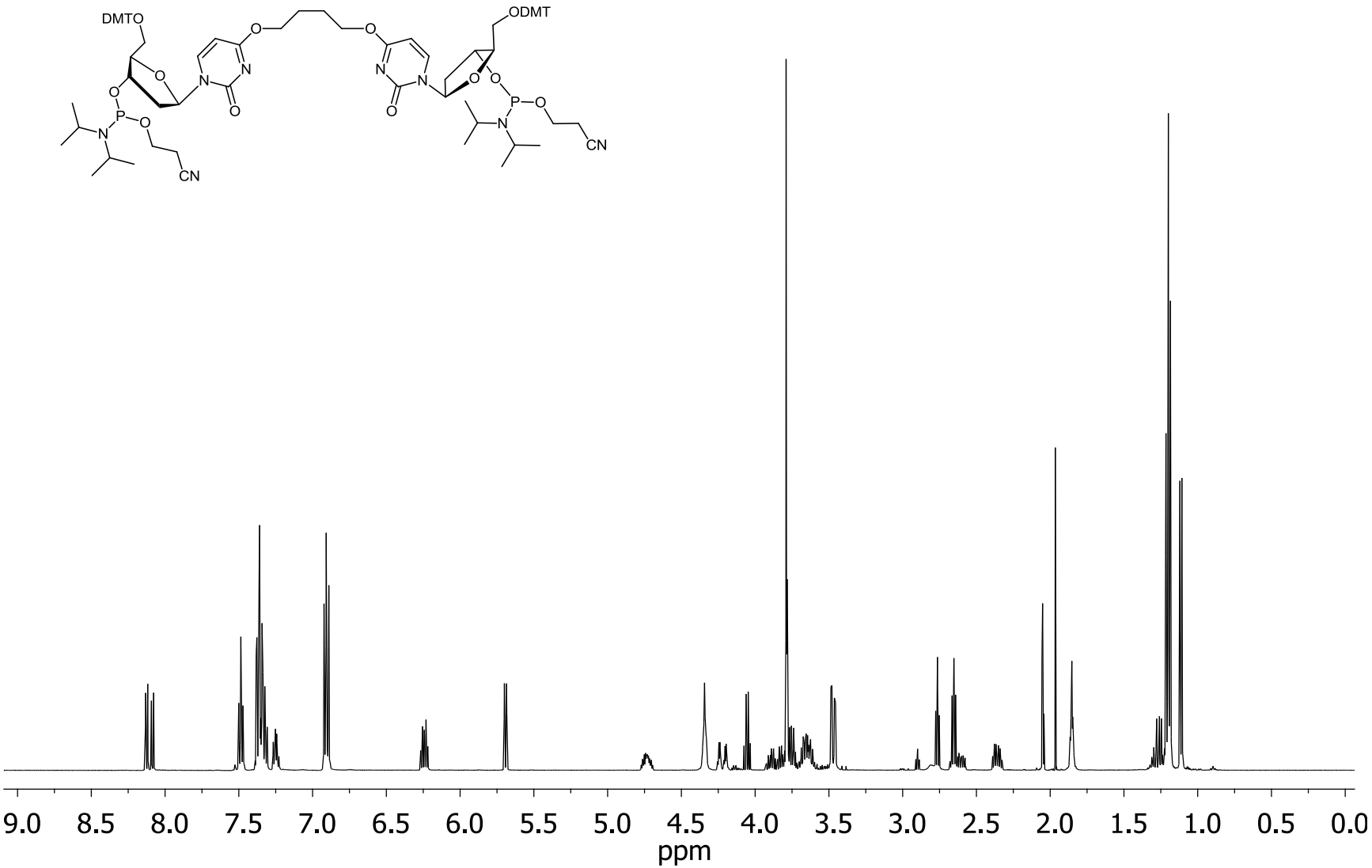
Supplementary Figure S7 - 500 MHz ^{1}H NMR spectrum of compound (**2b**) (in $CDCl_3$)



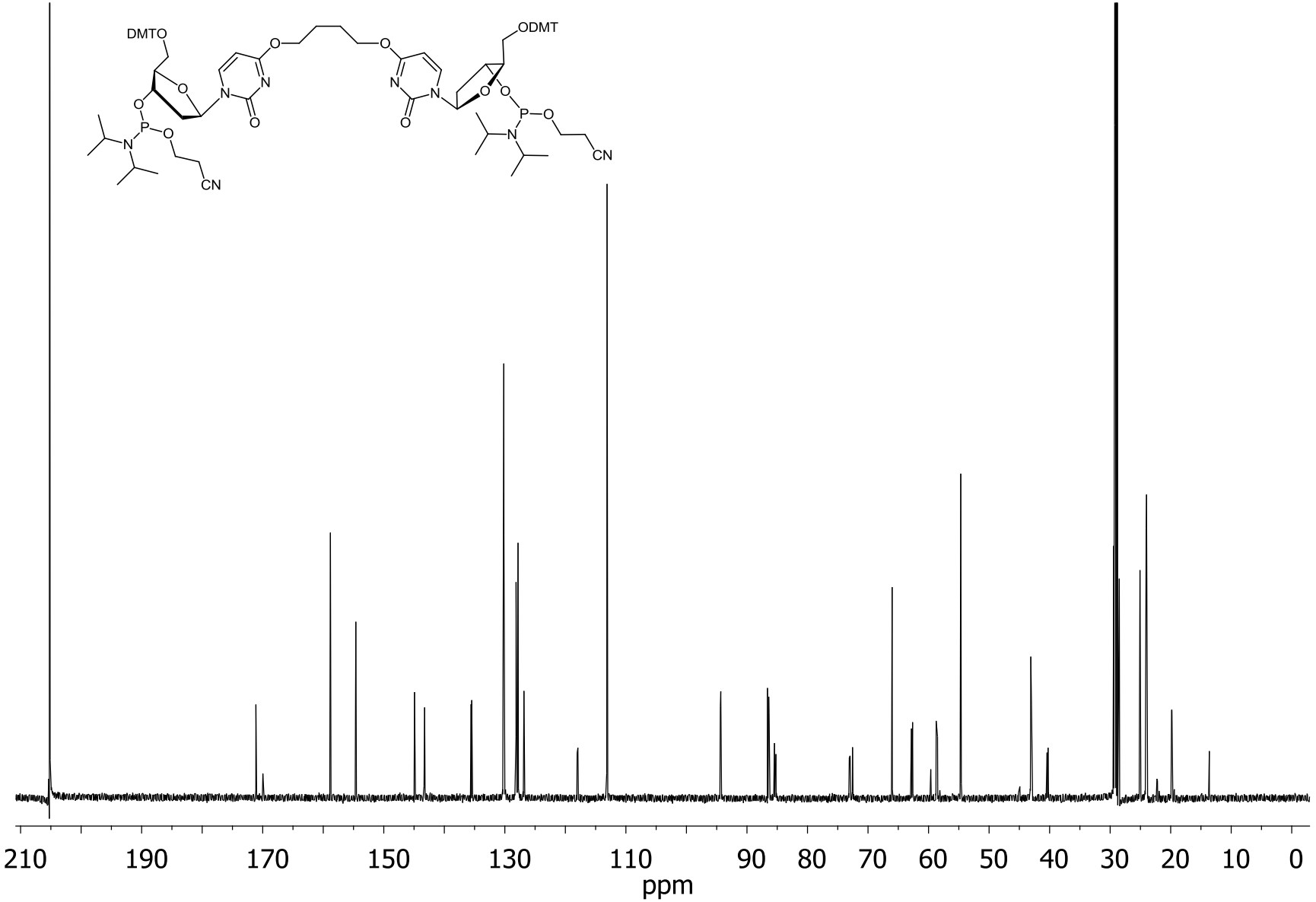
Supplementary Figure S8 - 125.7 MHz ^{13}C NMR spectrum of compound (**2b**) (in $CDCl_3$)



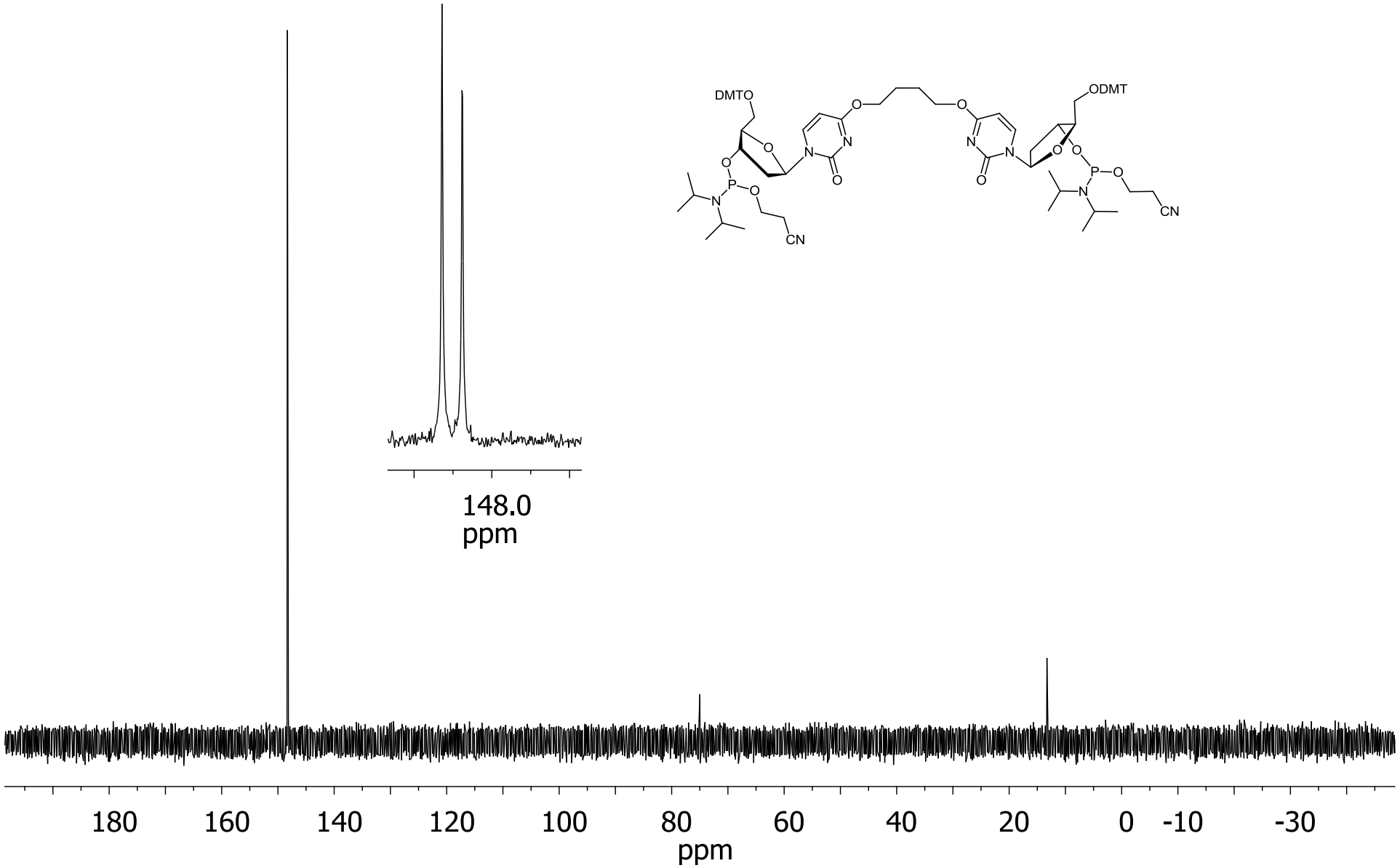
Supplementary Figure S9 - 500 MHz ^{1}H NMR spectrum of compound (**3a**) (in $CDCl_3$)



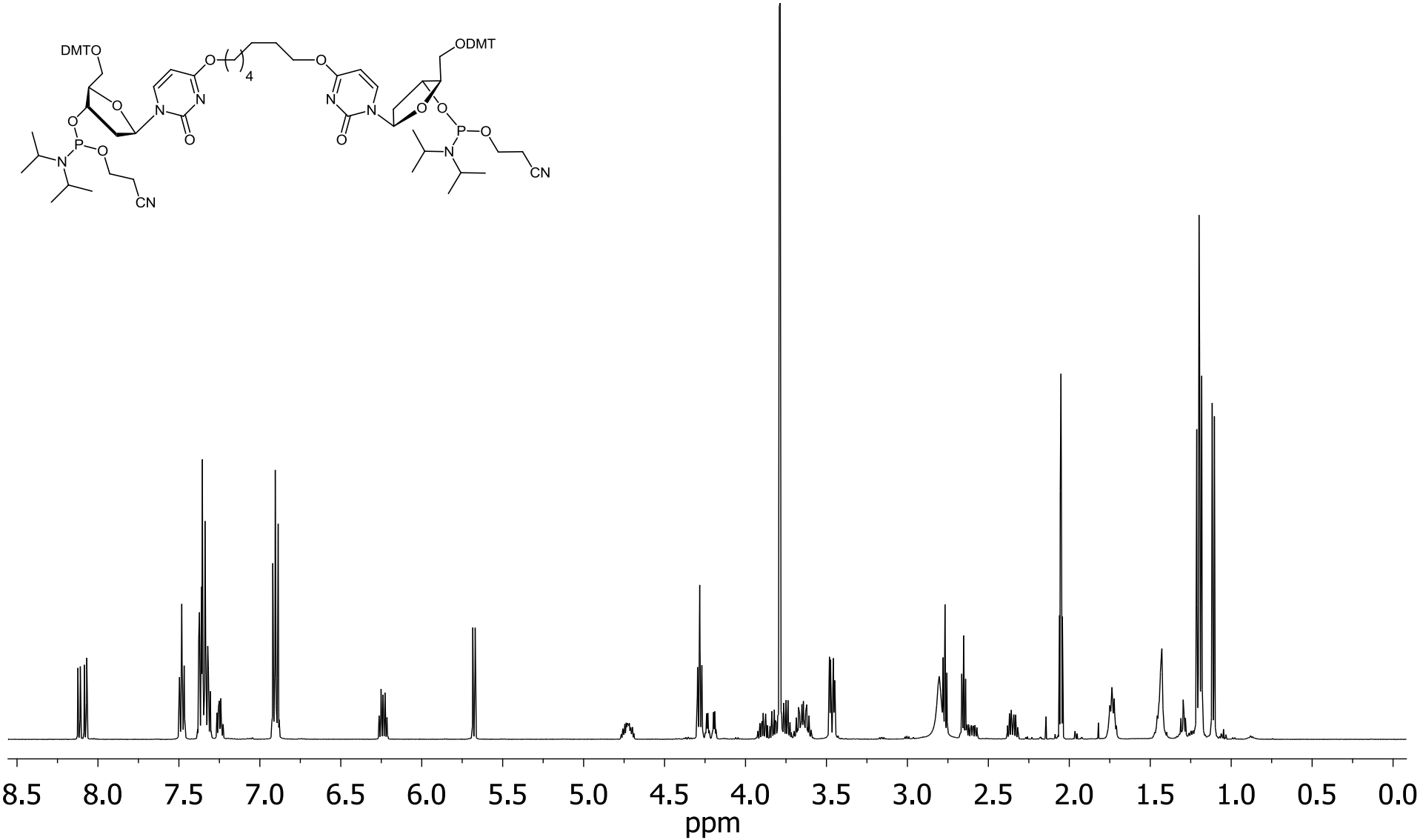
Supplementary Figure S10 - 125.7 MHz ^{13}C NMR spectrum of compound (**3a**) (in $CDCl_3$)



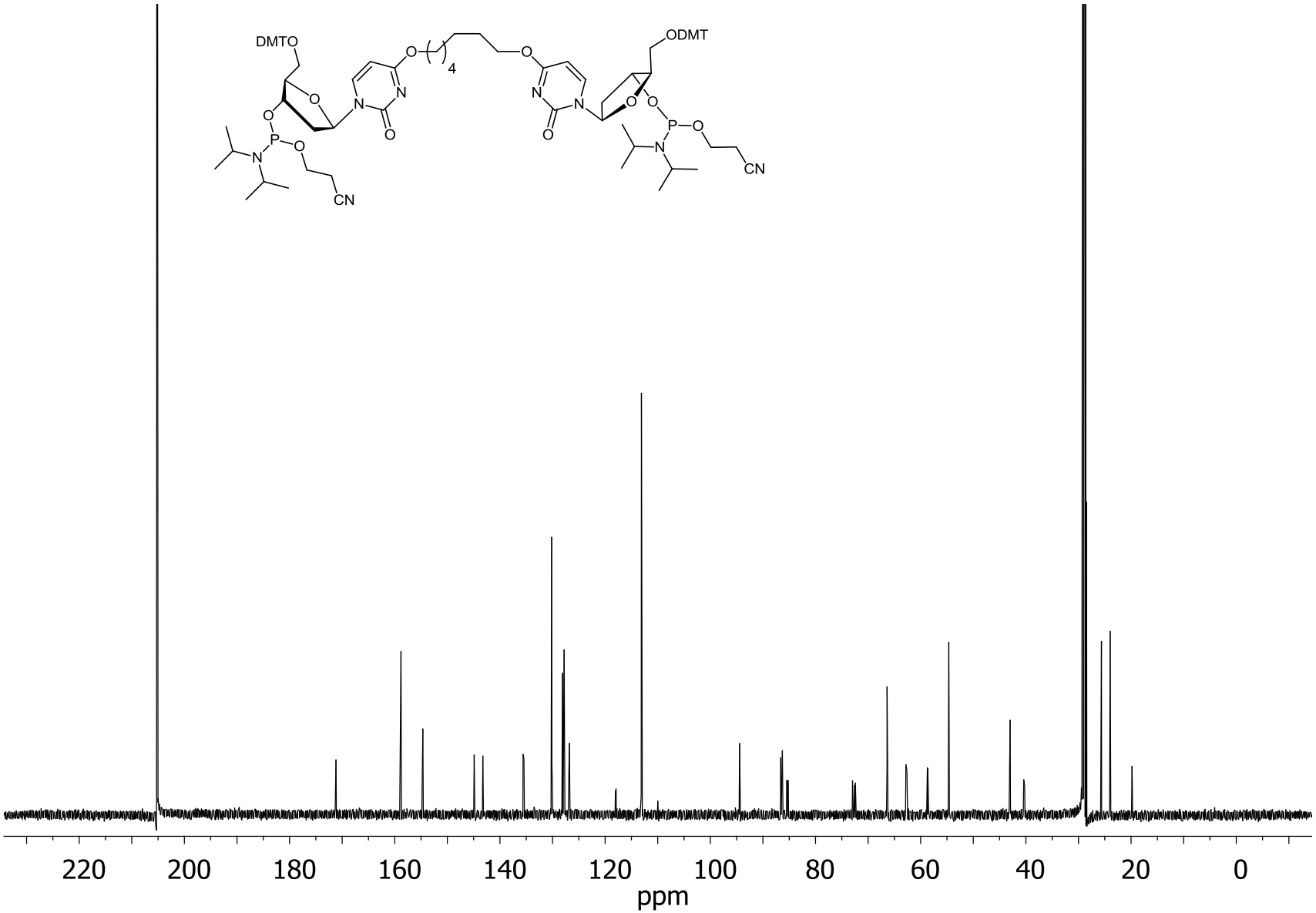
Supplementary Figure S11 - 202.3 MHz ^{31}P NMR spectrum of compound (**3a**) (in d_6-acetone)



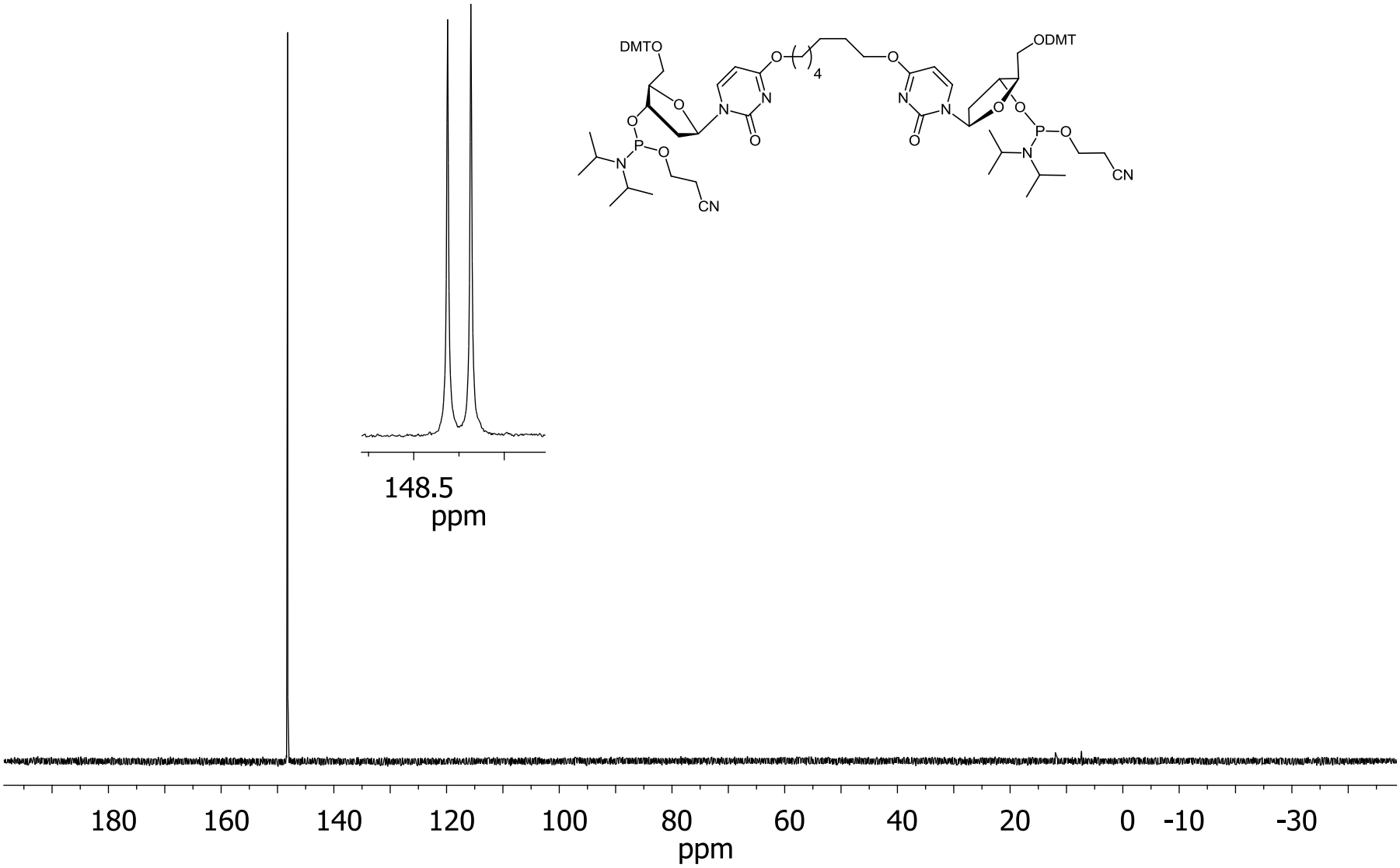
Supplementary Figure S12 - 500 MHz ^{1}H NMR spectrum of compound (**3b**) (in $CDCl_3$)



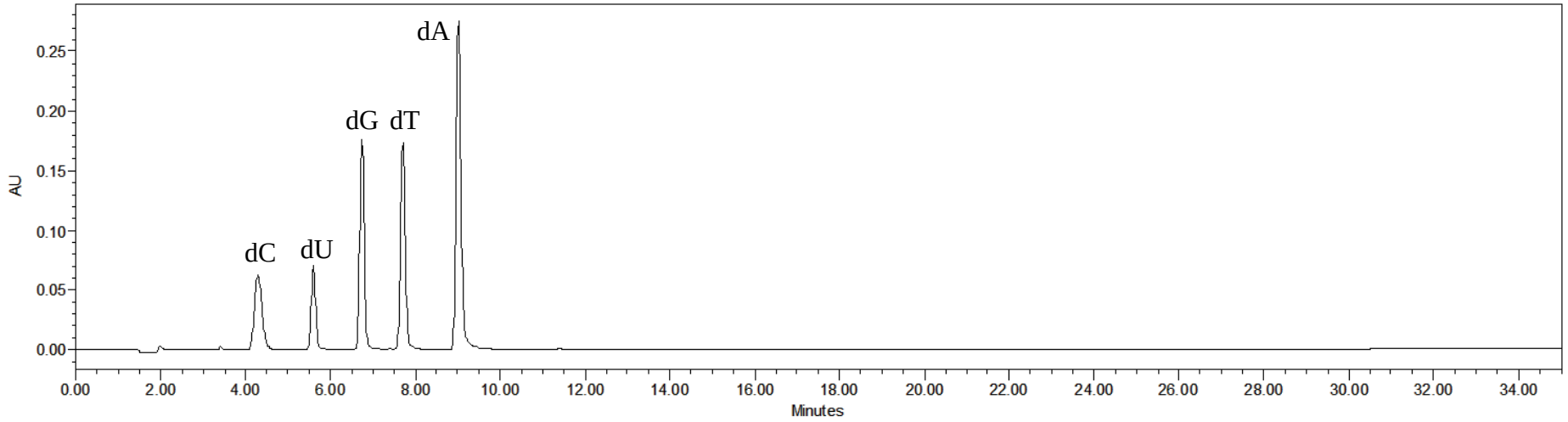
Supplementary Figure S13 - 125.7 MHz ^{13}C NMR spectrum of compound (**3b**) (in $CDCl_3$)



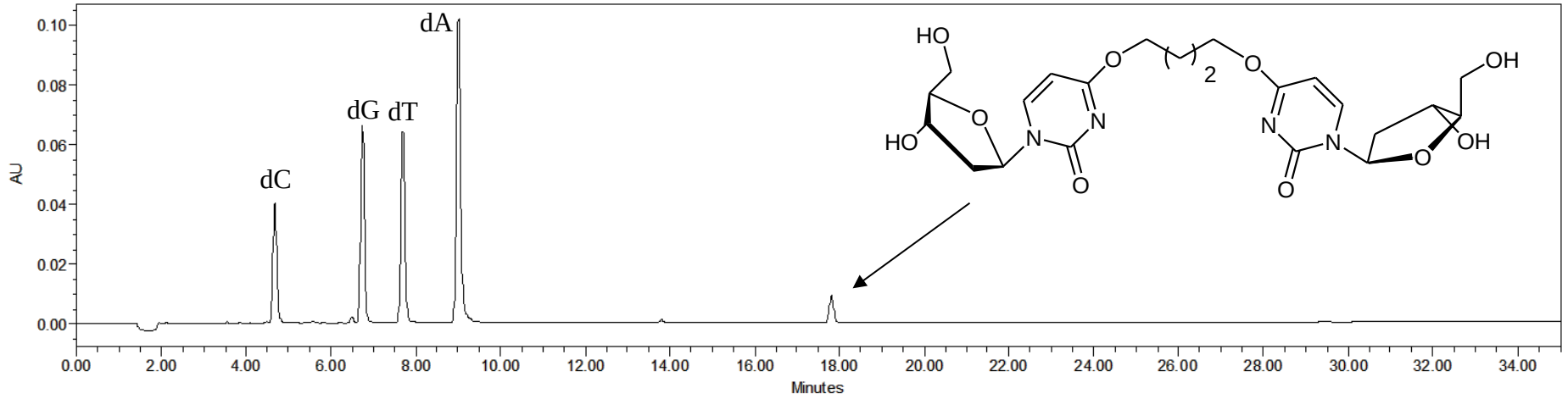
Supplementary Figure S14 - 202.3 MHz ^{31}P NMR spectrum of compound (**3b**) (in d_6-acetone)



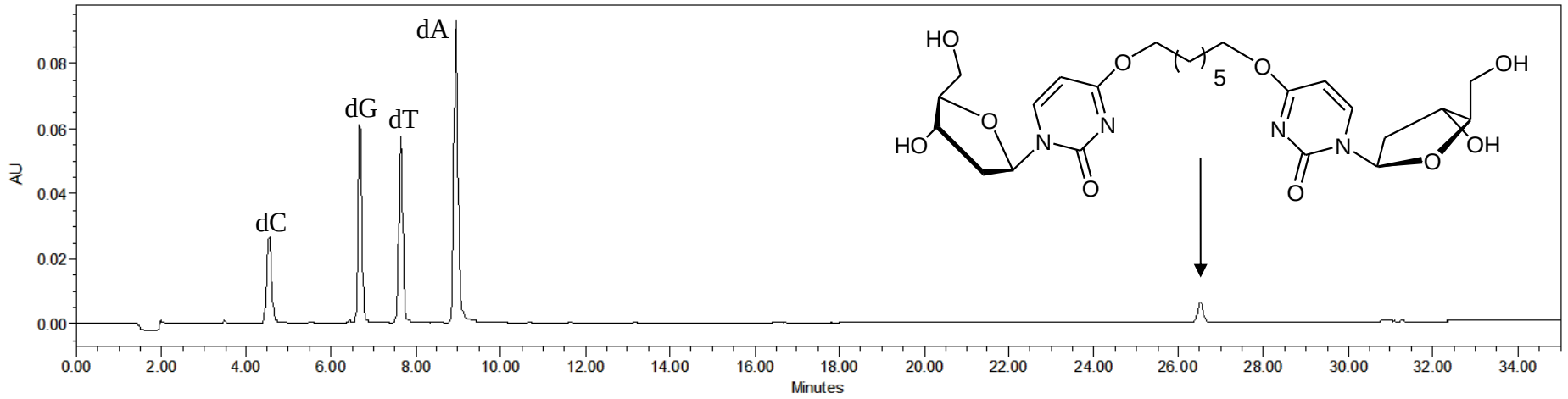
Supplementary Figure S15 - C-18 HPLC profile of digested control DNA **U** (5'-CGAAAUTTTCG). The column was eluted with a linear gradient of 0-70% buffer B over 30 min (buffer A: 50 mM sodium phosphate, pH 5.8 and buffer B: 50 mM sodium phosphate, pH 5.8, 50% acetonitrile).



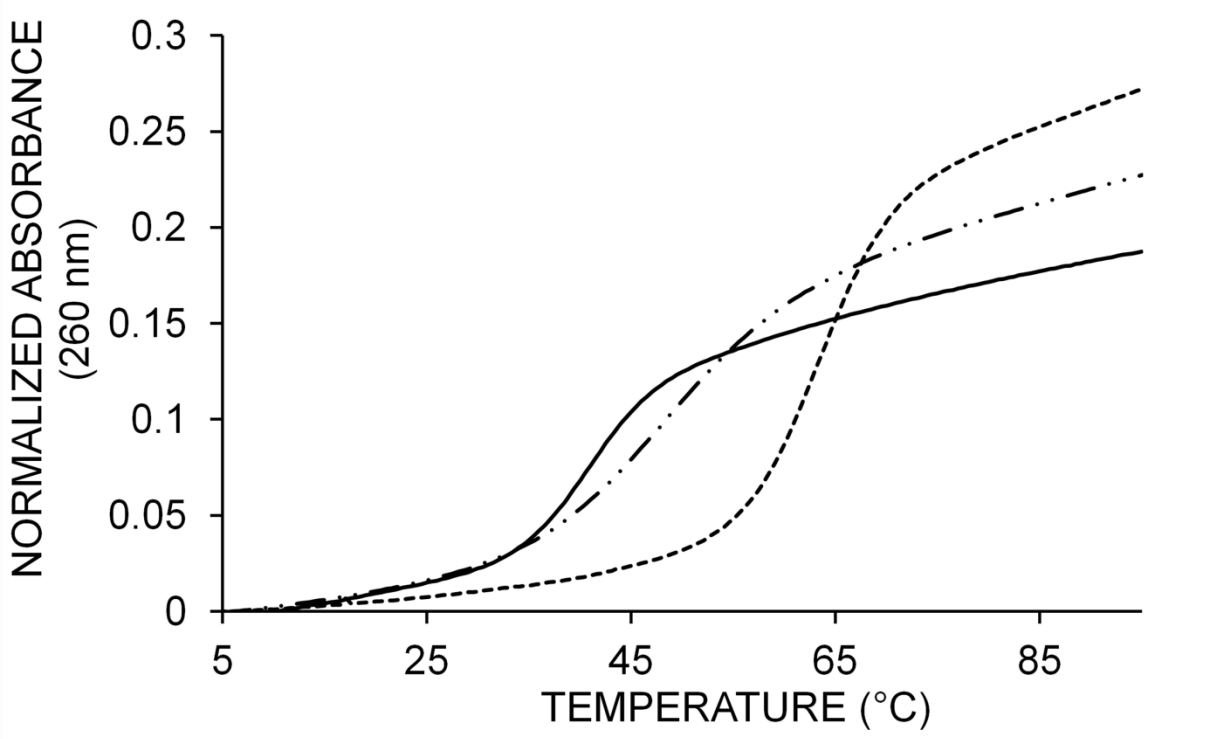
Supplementary Figure S16 - C-18 HPLC profile of digested cross-linked duplex **XLUU4**. The column was eluted with a linear gradient of 0-70% buffer B over 30 min (buffer A: 50 mM sodium phosphate, pH 5.8 and buffer B: 50 mM sodium phosphate, pH 5.8, 50% acetonitrile).



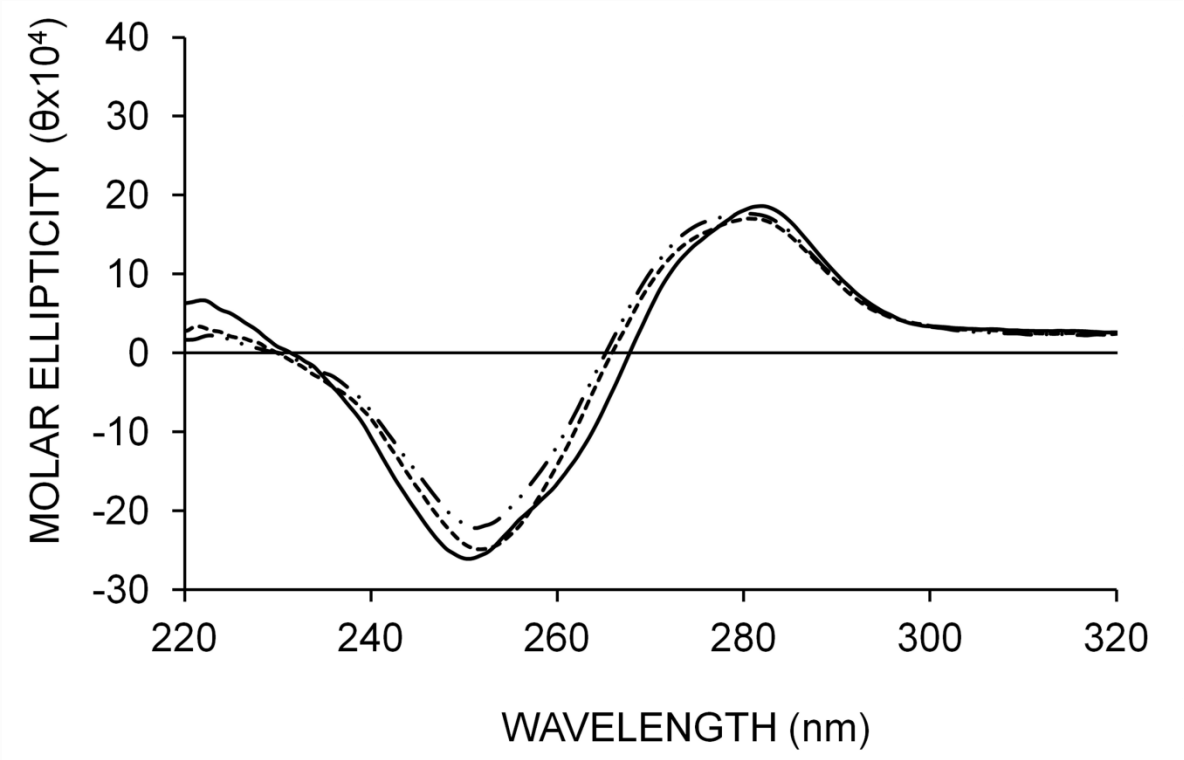
Supplementary Figure S17- C-18 HPLC profile of digested cross-linked duplex **XLUU7.** The column was eluted with a linear gradient of 0-70% buffer B over 30 min (buffer A: 50 mM sodium phosphate, pH 5.8 and buffer B: 50 mM sodium phosphate, pH 5.8, 50% acetonitrile).



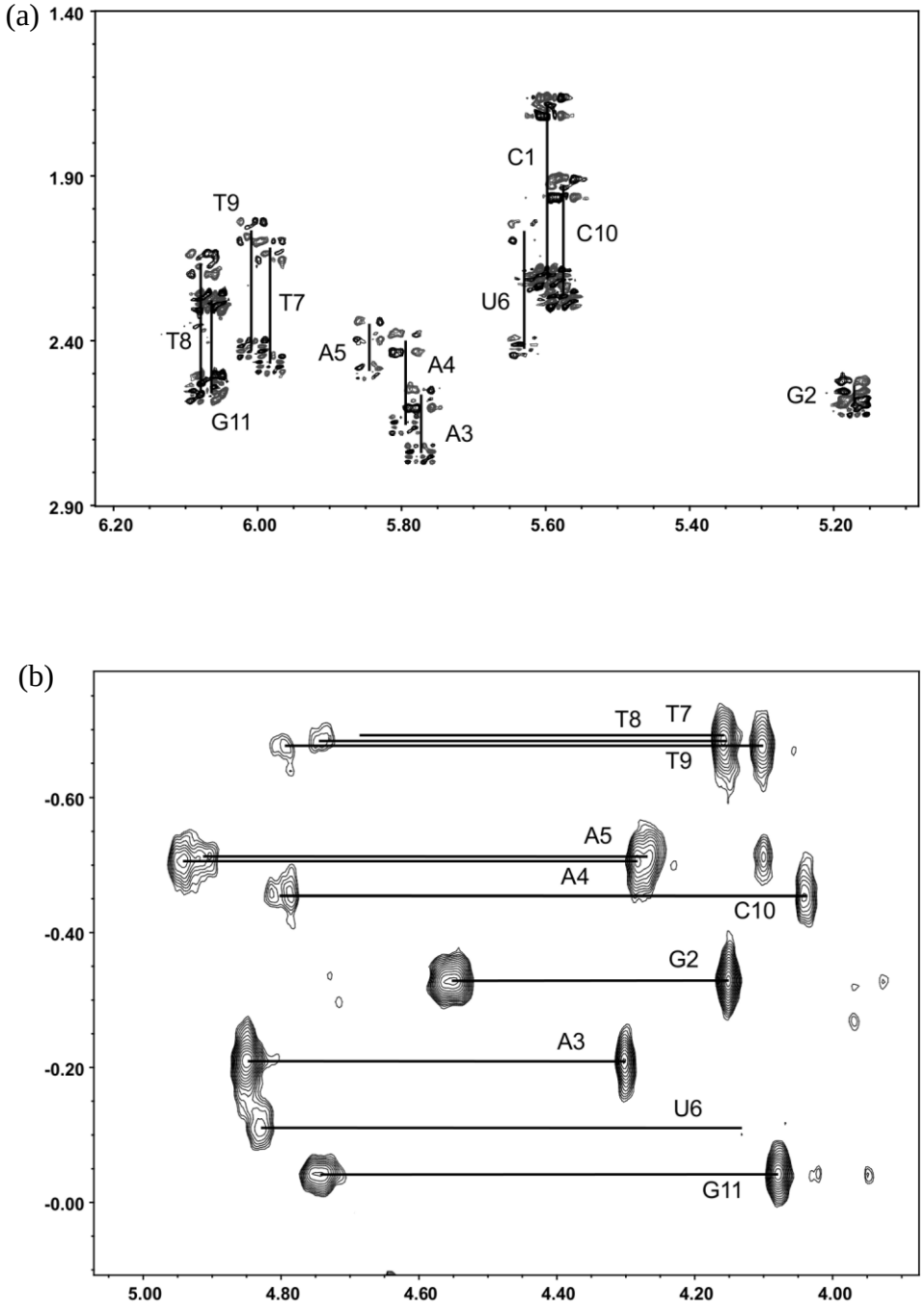
Supplementary Figure S18 - T_m Curves of **XLUU4**, **XLUU7** and control DNA duplexes. Control DNA (——), **XLUU4** (……) and **XLUU7** (- • • -).



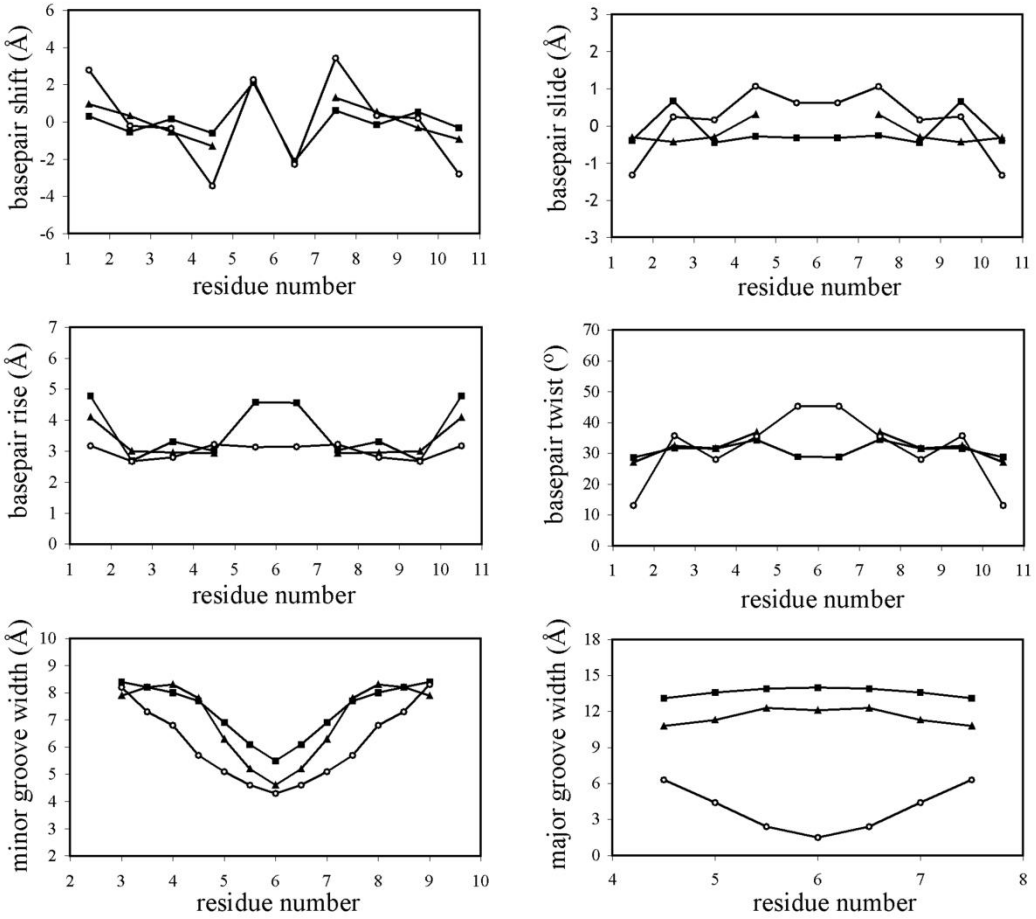
Supplementary Figure S19 - Far-UV Circular Dichroism Spectra of **XLUU4**, **XLUU7** and control DNA duplexes. Control DNA (——), **XLUU4** (……) and **XLUU7** (- • • -).



Supplementary Figure S20 - Expansions of (a) DQF-COSY and (b) H,P-HSQC spectra of **XLUU7** at 25 °C. *H1'-H2'* and *H1'-H2"* cross-peaks are connected as well as ^{31}*P-H4'* and ^{31}*P(i)-H3'(i-1)* cross-peaks.



Supplementary Figure S21. Helical parameters and groove widths of the average minimized structures of **XLUU7** (filled triangles), **XLGG7** (filled squares) and previously characterized **XLCC2** (PDB ID *1N4B*) (open circles). The helical parameters involving central nucleotides in **XLUU7** could not be used in the comparison due to the *syn*-conformation of the U6-bases and were omitted.



Supplementary Table S1 - ESI-MS results of oligonucleotide substrates synthesized

DNA	Calculated Mass (Da)	Observed Mass (Da)
U[a]	3315.58	3315.55
XLUU4[b]	6685.21	6685.14
XLUU7[b]	6727.25	6727.19

[a] Oligonucleotide sequence, 5' CGAAAUTTTCG
[b] Oligonucleotide sequence, 5' CGAAAYTTTCG / 3' GCTTTYAAAGC

Supplementary Table S2 - Nucleoside ratio of the various modified oligonucleotides as observed by snake venom digestion and RP-HPLC

DNA	Nucleoside	Retention time (min)	Nucleoside ratios	
			Expected	Observed
U[a]	dC	4.30	2.00	2.00
	dU	5.59	1.00	0.98
	dG	6.75	2.00	2.06
	dT	7.70	3.00	2.99
	dA	9.02	3.00	3.10
XLUU4[b]	dC	4.63	4.00	4.00
	dG	6.75	4.00	4.07
	dT	7.69	6.00	5.85
	dA	9.01	6.00	6.11
	dU-butylene-dU	17.82	1.00	0.96
XLUU7[b]	dC	4.61	4.00	4.00
	dG	6.67	4.00	4.08
	dT	7.64	6.00	5.86
	dA	8.93	6.00	6.25
	dU-heptylene-dU	26.54	1.00	1.03

[a] Oligonucleotide sequence, 5' CGAAAUTTTCG
[b] Oligonucleotide sequence, 5' CGAAAYTTTCG / 3' GCTTTYAAAGC

Supplementary Table S3. K_d of AGTs binding to Control and ICL DNA **XLUU4** and **XLUU7**.

DNA	OGT (μM)	Ada-C (μM)	hAGT (μM)	Chimera (μM)
Control[a]	>20	>100	>20	>20
XLUU4[b]	8.40 ± 0.18	8.73 ± 0.42	2.78 ± 0.03	4.02 ± 0.18
XLUU7[b]	7.61 ± 0.54	8.22 ± 0.60	1.74 ± 0.22	2.33 ± 0.17

[a] 5' CGAAAUTTTCG / 3' GCTTTAAAAGC

[b] 5' CGAAAYTTTCG / 3' GCTTTYAAAGC, where Y are cross-linked by a butylene (**XLUU4**) or heptylene (**XLUU7**) linker.

Supplementary Table S4 - Proton and phosphorus chemical shifts (ppm) of **XLUU7** (*1*) and **XLGG7** (*2*) duplexes at 25 °C[a].

Duplex	Residue	*H6/H8*	*H5/H2/Me*	*H1'*	*H2'*	*H2''*	*H3'*	*H4'*	*^{31}P(5')*
1	**C1**	7.48	5.80	5.60	1.68	2.20	4.55	3.93	-
2		7.48	5.79	5.59	1.69	2.20	4.55	3.93	-
1	**G2**	7.81	-	5.17	2.53	2.57	4.85	4.14	-0.33
2		7.81	-	5.17	2.54	2.57	4.85	4.15	-0.34
1	**A3**	8.04	7.29	5.77	2.57	2.73	4.95	4.30	-0.21
2		8.05	7.22	5.75	2.58	2.74	4.94	4.30	-0.24
1	**A4**	7.90	7.25	5.80	2.40	2.64	4.91	4.28	-0.50
2		7.87	7.21	5.74	2.35	2.59	4.89	4.27	-0.57
1	**A5**	7.82	7.55	5.85	2.37	2.48	4.83	4.26	-0.51
2		7.72	7.52	5.26	2.23	2.13	4.74	4.20	-0.72
1	**U6**	7.36	5.24	5.64	2.06	2.40	4.68	4.13	-0.11
2	**G6**	7.90	-	6.03	2.65	2.74	4.88	4.33	0.44
1	**T7**	7.33	1.40	5.99	2.11	2.46	4.74	4.16	-0.69
2		6.97	1.07	5.84	1.98	2.48	4.61	4.16	-0.73
1	**T8**	7.46	1.56	6.08	2.15	2.54	4.80	4.15	-0.69
2		7.47	1.51	6.10	2.15	2.55	4.82	4.19	-0.61
1	**T9**	7.38	1.60	6.01	2.06	2.43	4.80	4.10	-0.68
2		7.37	1.60	5.98	2.04	2.40	4.79	4.09	-0.66
1	**C10**	7.43	5.65	5.58	1.92	2.24	4.74	4.04	-0.45
2		7.42	5.65	5.57	1.91	2.25	4.74	4.03	-0.49
1	**G11**	7.86	-	6.06	2.51	2.27	4.58	4.07	-0.05
2		7.85	-	6.06	2.52	2.26	4.58	4.07	-0.06

[a] Proton signals for linker - CH_2(1) 3.47/3.62, CH_2(2) 1.18, CH_2(3) 0.86, CH_2(4) 0.95 ppm for **XLUU7**, and CH_2(1) 3.55/3.88, CH_2(2) 1.23/1.42, CH_2(3) 1.09, CH_2(4) 1.01 ppm for **XLGG7**.

Supplementary Table S5. Structural statistics for 10 final individual structures of **XLUU7** and **XLGG7** duplexes

Duplex	**XLUU7**	**XLGG7**
Number NOE distance restraints	131	150
Intranucleotide	77	84
Internucleotide	54	66
Torsion angle restraints	87	87
Hydrogen bond restraints	34	34
NOE violation (>0.2 Å)	0	0
Torsion angle violation (>5 °)	0	0
Heavy atoms RMSD relative to average structure:		
Full duplex (Å)	0.58	0.41
Residues 1:5 of strand A and 7:11 of strand B (Å)	0.24	0.29
Average RMSD from covalent geometry:		
Bond lengths (Å)	0.0026 ± 0.0001	0.0026 ± 0.0001
Angles (°)	0.380 ± 0.005	0.411 ± 0.011
Impropers (°)	0.315 ± 0.023	0.355 ± 0.022

References

1. Q. Fang, A. M. Noronha, S. P. Murphy, C. J. Wilds, J. L. Tubbs, J. A. Tainer, G. Chowdhury, F. P. Guengerich and A. E. Pegg, *Biochemistry*, 2008, **47**, 10892-10903.
2. F. P. McManus and C. J. Wilds, *Chembiochem*, 2014, **15**, 1966-1977.
3. D. M. Noll, A. M. Noronha and P. S. Miller, *J Am Chem Soc*, 2001, **123**, 3405-3411.
4. F. P. McManus, D. K. O'Flaherty, A. M. Noronha and C. J. Wilds, *Organic & biomolecular chemistry*, 2012, **10**, 7078-7090.
5. J. D. Puglisi and I. Tinoco, Jr., *Methods Enzymol*, 1989, **180**, 304-325.
6. M. E. Dolan and A. E. Pegg, *Carcinogenesis*, 1985, **6**, 1611-1614.
7. A. J. Doherty, S. R. Ashford, J. A. Brannigan and D. B. Wigley, *Nucleic acids research*, 1995, **23**, 2074-2075.
8. P. E. Verdemato, J. A. Brannigan, C. Damblon, F. Zuccotto, P. C. Moody and L. Y. Lian, *Nucleic acids research*, 2000, **28**, 3710-3718.