# Supporting Information

# Site-Specific Generation of Deoxyribonolactone Lesions in DNA Oligonucleotides

Hamilton J. Lenox, Christopher P. McCoy and Terry L. Sheppard*

*Contribution from the Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3113*

**General Procedures**. All glassware was oven-dried (140 °C). Flash column chromatography was performed using silica gel (40-63 μm, EM Science). Thin layer chromatography (TLC) utilized glass plates with 0.250 mm of silica 60 F254 (EM Science), and spots were visualized using a short-wave UV lamp or by staining with anisaldehyde. All reagents were supplied by Aldrich and were used without further purification. Anhydrous reagents were supplied by Aldrich in Sure-Seal™ bottles. $N_2$ was used as an inert atmosphere. All reactions involving compounds **7**, **13**, **14**, **15**, and **16** were performed under inert atmosphere and in the absence of light. $^1$H- and $^{13}$C-NMR data were recorded at 500 and 125 MHz, respectively, on a Varian 500 spectrometer, unless noted otherwise. Values are given in part per million (ppm) downfield from the internal standard tetramethylsilane in the following format: chemical shift (multiplicity, integration, coupling constant in hertz, proton assignment). Proton assignments in $^1$H-NMR spectra were determined using $^1$H-COSY data. $^{31}$P-NMR data were recorded at 121 and 162 MHz on a Varian 300 or 400 spectrometer, respectively, and were referenced to an external standard of 85% aqueous phosphoric acid. Mass spectral data for synthetic intermediates were obtained by FAB, unless stated otherwise, using a Micromass 70SE-4F spectrometer, using glycerol, 3-nitrobenzyl alcohol or 3:1 dithiothreitol: dithioerythritol as matrices. T4 polynucleotide kinase was obtained from US Biochemical and [$\gamma$-$^{32}$P] ATP (7000 Ci/mmol) was purchased from ICN. Radioactive bands in polyacrylamide gels were visualized using a Molecular Dynamics Storm Phosphorimager and quantitated using Molecular Dynamics ImageQuant software.

**Cyano 3,5-*O*-(di-4-chlorobenzoyl)-2-deoxy-D-ribofuranose (11).** Compound **10** (see paper, reference 15) (9.62 g, 22.4 mmol) was dissolved in 30 mL anhydrous THF,

and the resulting solution was cooled in an ice water bath. The solution was protected from light and 35.0 mL $Et_2AlCN$ (30.2 g, 35.0 mmol) was added dropwise. The ice water bath was removed at 45 min, and the reaction was stirred for a total of 4 h. The reaction then was cooled in an ice water bath and quenched by dropwise addition of 23.0 mL methanol. The solution was concentrated, and the resulting solid was redissolved in 45 mL $CH_2Cl_2$ and applied to a 150 mL silica plug (silica covered with a 5 mm layer of anhydrous $Na_2SO_4$) in a 600 mL fritted funnel. The silica plug was washed under vacuum with 500 mL $CH_2Cl_2$, the filtrate was concentrated, and the residue was dried under vacuum for 7 h. The solid residue was dissolved in 60 mL ethyl acetate, and 500 mL hexanes was added to induce crystallization. The solution was stored at -20°C for 2 h to complete the crystallization. The crystals then were separated by vacuum filtration and dried under vacuum for 12 h to provide 7.06 g of **11** (75%) as an anomeric mixture. $^1$H NMR ($CDCl_3$): δ 8.06-8.01 (m, 2H, Ar-H), 7.95 (m, 2H, Ar-H), 7.45 (m, 4H, Ar-H), 5.61 (d, 1H, $J$ = 5.5 Hz, H-3), 5.06 (m, 1H, H-1), 4.93 (m, 1H, H-1), 4.64 (m, 1H, H-4), 4.55 (m, 2H, H-5), 2.69 (m, 2H, H-2); $^{13}$C NMR ($CDCl_3$): δ 165.49, 165.35, 165.06, 140.58, 140.53, 140.17, 131.55, 131.32, 131.27, 129.23, 129.17, 128.06, 127.65, 127.58, 118.62 & 118.10 (*C*N), 84.27 & 83.82 (C-4), 75.86 & 75.39 (C-3), 66.76 & 66.11 (C-1), 64.20 & 64.07 (C-5), 37.99 & 37.89 (C-2). HRMS-APCI (*m*/*z*) (Micromass Quatro II LC/MS/MS spectrometer): $[M + H]^+$ calcd for $C_{20}H_{15}Cl_2NO_5$+H, 420.0406; found, 420.0412 (anomeric mix).

**Bromo 1-cyano-3,5-*O*-(di-4-chlorobenzoyl)-2-deoxy-D-ribofuranose (12).** To a stirred solution of **11** (1.50 g, 3.58 mmol) in 22.0 mL anhydrous $CCl_4$, *N*-bromosuccinimide (0.866 g, 4.87 mmol), and benzoyl peroxide (0.186 g, 0.767 mmol)

were added. The reaction was refluxed (85-90 °C) for 2 h. The reaction solution then was filtered while still warm, and the filtered solid material was washed with 80 mL cold anhydrous $CCl_4$. The combined filtrate and washings were evaporated and applied to a 250 mL silica plug in a 600 mL fritted funnel (covered with ~1 cm sand) as a solution in 20:1 toluene-ether. Product that crystallized on the sand was removed, dissolved in $CH_2Cl_2$, filtered, and added to the filtrate from the silica plug. The plug was washed with ~600 mL toluene-ether. The combined filtrate was evaporated, and the product **12** was precipitated from 250 mL hexanes-ether (9:1) (1.46 g, 82%), and stored under inert atmosphere at –20 °C. NMR data for single anomer only. $^1$H NMR ($CDCl_3$): δ 8.00 (m, 4H, Ar-H), 7.45 (m, 4H, Ar-H), 5.77 (m, 1H, H-3), 4.84 (dd, 1H, $J_{4,5a}$ = 9.00 Hz, $J_{4,5b}$ = 4.50 Hz, H-4), 4.71 (dd, 1H, $J_{5a,4}$ = 12.50 Hz, $J_{5a,5b}$ = 4.50 Hz, H-$5_a$), 4.65 (dd, 1H, $J_{5b,4}$ = 12.50 Hz, $J_{5a,5b}$ = 4.50 Hz, H-$5_b$), 3.50 (dd, 1H, $J_{2a,3}$ = 14.50 Hz, $J_{2a,2b}$ = 6.50 Hz, H-$2_a$), 3.27 (dd, 1H, $J_{2b,3}$ = 14.50 Hz, $J_{2b,2a}$= 4.50 Hz, H-$2_b$); $^{13}$C NMR ($CDCl_3$): δ 165.30, 164.98, 140.94, 140.37, 131.46, 129.36, 129.18, 127.80, 127.02, 116.08 (*C*N), 86.71 (C-4), 75.46 (C-1), 73.92 (C-3), 63.05 (C-5), 50.96 (C-2). HRMS-FAB (*m*/*z*): $[M + H]^+$ calcd for $C_{20}H_{14}BrCl_2NO_5$+H, 497.9511; found 497.9515 (anomeric mix).

**2-Nitrobenzyl 1′-cyano-3′, 5′-*O*-(di-4-chlorobenzoyl)-2′-deoxy-D-ribofuranoside (13).** To a stirred solution of **12** (0.574 g, 1.15 mmol) in 11.5 mL anhydrous $CH_2Cl_2$, were added 2-nitrobenzyl alcohol (0.494 g, 3.22 mmol), 2,6-lutidine (0.49 mL, 0.45g, 4.2 mmol), and silver triflate (0.706 g, 2.75 mmol). The reaction was stirred for 4 h at 24 °C, after which, it was filtered through 1 cm of Celite. The filtrate solution was evaporated, and the residue was dissolved in 20 mL $CH_2Cl_2$. The organic phase was washed with 20 mL $H_2O$, then dried over anhydrous $MgSO_4$ and concentrated

in vacuo. The residue was purified by flash chromatography using 2% ethyl acetate in $CHCl_3$ to provide 0.414 g of **13** (63%, anomeric mixture). $^1$H NMR ($CDCl_3$): δ 8.10-7.86 (m, 10H, Ar-H), 7.71 (m, 1H, Ar-H), 7.59-7.38 (m, 12H, Ar-H), 7.30 (m, 1H, Ar-H), 5.67 (m, 1H, H-3′), 5.56 (m, 1H, H-3′), 5.41 (d, 1H, $J$ = 13.50 Hz, Ar-C$H_2$), 5.26 (m, 2H, Ar-C$H_2$), 5.15 (d, 1H, Ar-C$H_2$), 4.78 (m, 1H, H-4′), 4.67 (m, 1H, H-4′), 4.50 (m, 4H, H-5′), 2.95 (m, 4H, H-2′); $^{13}$C NMR ($CDCl_3$): δ 165.20, 165.02, 140.50-139.80 (4C's), 133.76, 133.69, 131.29, 131.21, 131.10, 131.01, 129.46, 129.08, 129.04, 129.00, 128.93, 128.83, 125.08, 124.96, 115.78 & 115.21 (*C*N), 100.95 & 100.27 (C-1′), 84.98 & 84.76 (C-4′), 74.48 & 74.44 (C-3′), 65.57 & 64.90 (Ar-*C*H$_2$), 63.94 & 63.72 (C-5′), 45.33 & 45.16 (C-2′). HRMS-FAB (*m*/*z*): $[M + H]^+$ calcd for $C_{27}H_{20}Cl_2N_2O_8$+H, 571.0675; found, 571.0675 (anomeric mix). The byproduct, 3,5-*O*-(di-4-chlorobenzoyl)-2-deoxy-D-ribonolactone, which eluted after **13**, was produced in a 20% yield.

**2-Nitrobenzyl 1′-cyano-2′-deoxy-D-ribofuranoside (14).** To a stirred solution of **13** (2.42 g, 4.24 mmol) in 4 mL of anhydrous THF was added 50 mL saturated methanolic ammonia. The reaction was stirred for 3 h at 40 °C, evaporated, and purified by flash chromatography using 100% ethyl acetate to afford 0.997 g of **14** (80%). $^1$H NMR ($CDCl_3$): δ 8.05 (m, 2H, Ar-H), 7.72 (m, 2H, Ar-H), 7.66 (m, 2H, Ar-H), 7.50 (m, 2H, Ar-H), 5.15 (m, 4H, Ar-C$H_2$), 4.49 (m, 2H, H-3′), 4.22 (m, 2H, H-4′), 3.73 (m, 4H, H-5′), 3.19-2.49 (m, 6H, O*H* & H-2′); $^{13}$C NMR ($CDCl_3$): δ 147.44, 147.39, 134.31, 134.20, 132.53, 132.33, 129.85, 129.64, 129.20, 129.10, 125.10, 116.63 & 116.15 (*C*N), 100.56 & 100.33 (C-1′), 89.46 & 89.22 (C-4′), 71.39 & 71.11 (C-3′), 65.20 & 64.98 (Ar-*C*H$_2$), 62.70 & 61.95 (C-5′), 47.88 & 47.46 (C-2′). HRMS-FAB (*m*/*z*): $[M + Na]^+$ calcd for $C_{13}H_{14}N_2O_6$+Na, 317.0750; found, 317.0751 (anomeric mix).

**2-Nitrobenzyl 1′-cyano-5′-*O*-(4-monomethoxytrityl)-2′-deoxy-D-ribofuranoside (15).** Compound **14** (0.473 g, 1.61 mmol) was dissolved in 3.2 mL anhydrous pyridine and MMTCl (0.766 g, 2.48 mmol) and DMAP (0.036 g, 0.29 mmol) were added. The reaction was stirred for 18.5 h at 24 °C, and then 9 mL saturated aqueous $NaHCO_3$ was added. A water-ethyl acetate solution (1:1, 80 mL) was added to dissolve the precipitate, and the organic phase was separated. The $H_2O$ layer was back-extracted with 20 mL ethyl acetate, and the combined organic phases were washed with 30 mL saturated aqueous NaCl. The organic phase was dried over $Na_2SO_4$, concentrated, and the residue was purified by flash chromatography (2:3 ethyl acetate-hexanes) to provide an anomeric mixture of **15** (0.729 g, 80%). The anomeric compounds **15a** (alpha stereochemistry at C-1′) and **15b** (beta stereochemistry) were separated by flash chromatography (1% acetone in $CH_2Cl_2$).

Compound **15a** (alpha anomer) eluted first (35% isolated yield after second column). $^1$H NMR ($CD_3CN$): δ 8.09 (d, 1H, $J$ = 8.00 Hz, H-6), 7.87 (d, 1H, $J$ = 7.50 Hz, H-3), 7.73 (t, 1H, $J$ = 7.50 Hz, H-4), 7.55 (t, 1H, $J$ = 8.50 Hz, H-5), 7.46 (s, 2H, Ar-H), 7.45 (s, 2H, Ar-H), 7.32 (m, 6H, Ar-H), 7.26 (m, 2H, Ar-H), 6.90 (s, 1H, Ar-H), 6.88 (s, 1H, Ar-H), 5.26 (d, 1H, $J$ = 14.00 Hz, Ar-C$H_2$), 5.14 (d, 1H, $J$ = 14.00 Hz, Ar-C$H_2$), 4.35 (m, 1H, H-3′), 4.16 (m, 1H, H-4′), 3.78 (s, 3H, C$H_3$), 3.35 (d, 1H, $J$ = 5.00 Hz, O$H$), 3.28 (dd, 1H, $J_{5'a,4'}$ = 3.00 Hz, $J_{5'a,5'b}$ = 10.50 Hz, H-5′$_a$), 3.11 (dd, 1H, $J_{5'b,4'}$ = 4.50 Hz, $J_{5'b,5'a}$ = 10.50 Hz, H-5′$_b$), 2.91 (dd, 1H, $J_{2'a,3'}$ = 7.50 Hz, $J_{2'a,2'b}$ = 14.50 Hz, H-2′$_a$), 2.45 (dd, 1H, $J_{2'b,3'}$ = 2.50 Hz, $J_{2'b,2'a}$ = 14.00 Hz, H-2′$_b$); $^{13}$C NMR ($CD_3CN$): δ 159.87, 148.57, 145.52, 145.40, 136.25, 134.94, 133.80, 131.37, 130.53, 129.86, 129.31, 128.96, 128.11, 125.77, 117.41, 114.17 (*C*N), 101.74 (C-1′), 89.69 (C-4′), 87.48 (*C*-Ar$_3$), 72.18 (C-3′),

65.28 (C-5′), 64.44 (Ar-*C*$H_2$), 55.98 (*C*$H_3$), 48.60 (C-2′). HRMS-FAB (*m*/*z*): $[M]^+$ calcd for $C_{33}H_{30}N_2O_7$, 566.2053; found, 566.2050.

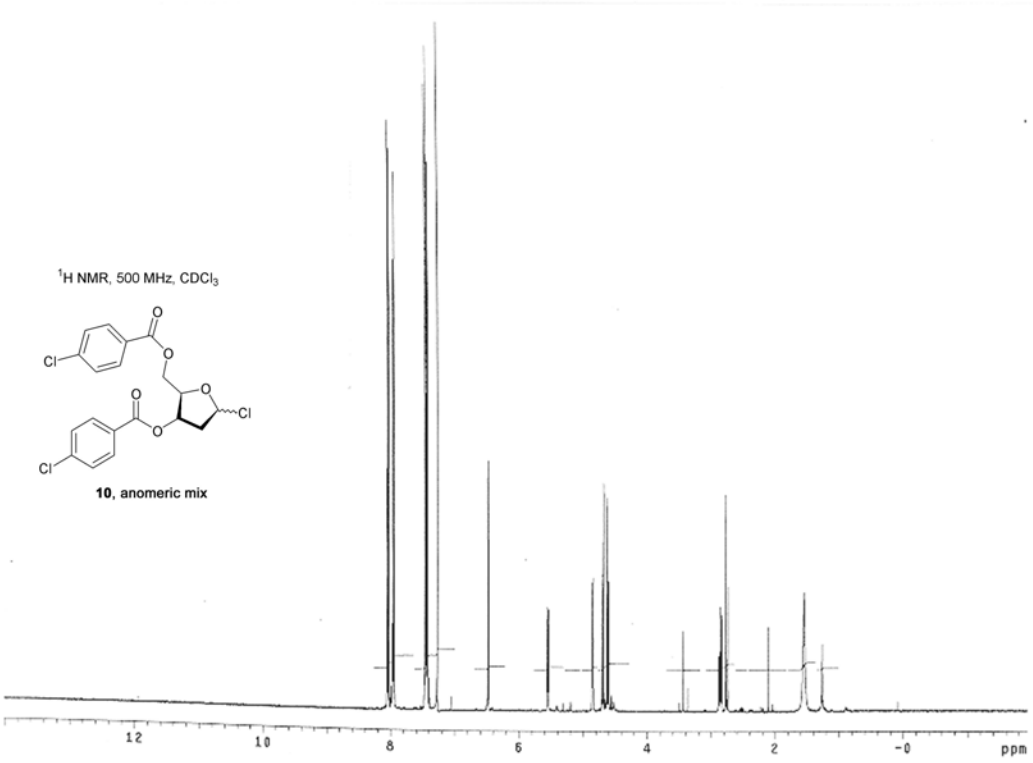
Compound **15b** (beta anomer) eluted second (24% isolated yield after second column). $^1$H NMR ($CD_3CN$): δ 8.05 (d, 1H, *J* = 8.00 Hz, H-6), 7.64 (m, 2H, H-3 & H-4), 7.51 (dt, 1H, *J* = 2.50 Hz, *J* = 8.00 Hz, H-5), 7.38 (s, 2H, Ar-H), 7.37 (s, 2H, Ar-H), 7.21 (m, 8H, Ar-H), 6.78 (s, 1H, Ar-H), 6.76 (s, 1H, Ar-H), 5.16 (d, 1H, *J* = 14.50 Hz, Ar-C*H*$_2$), 5.14 (d, 1H, *J* = 14.00 Hz, Ar-C*H*$_2$), 4.48 (m, 1H, H-3′), 4.24 (m, 1H, H-4′), 3.72 (s, 3H, C*H*$_3$), 3.55 (d, 1H, *J* = 4.50 Hz, O*H*), 3.22 (dd, 1H, $J_{5'a,4'}$ = 4.00 Hz, $J_{5'a,5'b}$ = 10.50 Hz, H-5′$_a$), 3.09 (dd, 1H, $J_{5'b,4'}$ = 5.00 Hz, $J_{5'b,5'a}$ = 10.50 Hz, H-5′$_b$), 2.75 (dd, 1H, $J_{2'a,3'}$ = 6.00 Hz, $J_{2'a,2'b}$ = 13.50 Hz, H-2′$_a$), 2.59 (dd, 1H, $J_{2'b,3'}$ = 5.5 Hz, $J_{2'b,2'a}$ = 13.50 Hz, H-2′$_b$); $^{13}$C NMR ($CD_3CN$): δ 159.77, 148.25, 145.38, 145.32, 136.26, 135.03, 133.53, 131.19, 129.97, 129.73, 129.21, 128.99, 128.87, 128.78, 128.60, 128.04, 125.84, 117.56, 114.06 (*C*N), 101.19 (C-1′), 88.99 (*C*-Ar$_3$), 87.35 (C-4′), 71.48 (C-3′), 65.30 (C-5′), 64.06 (Ar-*C*$H_2$), 55.93 (*C*$H_3$), 48.00 (C-2′).

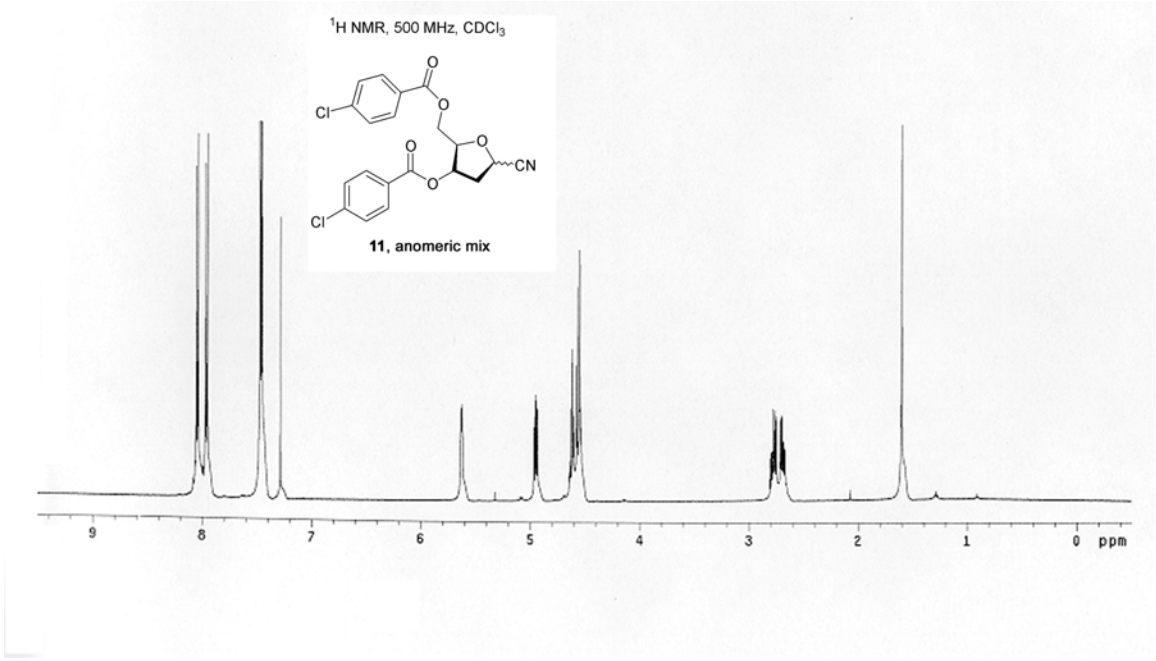
**2-Nitrobenzyl 1′-cyano-3′-*O*-(2-cyanoethyl *N*,*N*-diisopropylphosphoramidyl)-5′-*O*-(4-monomethoxytrityl)-2′-deoxy-D-ribofuranoside (7b).** Compound **15b** (0.315 g, 0.557 mmol) was coevaporated with 3x10 mL anhydrous acetonitrile, dried under vacuum for 12 h, and dissolved in 2.8 mL anhydrous $CH_2Cl_2$. Diisopropylammonium tetrazolide (0.0718 g, 0.419 mmol) was added to the stirring solution, followed by dropwise addition of 2-cyanoethyl tetraisopropylphosphorodiamidite (0.265 mL, 0.0251 g, 0.834 mmol). The reaction was stirred for 3 h, then 40 mL $CH_2Cl_2$ and 40 mL $H_2O$ were added, and the organic phase was separated. The aqueous layer was back-extracted with 2x20 mL $CH_2Cl_2$ and the combined organic layers were washed with 2x50 mL

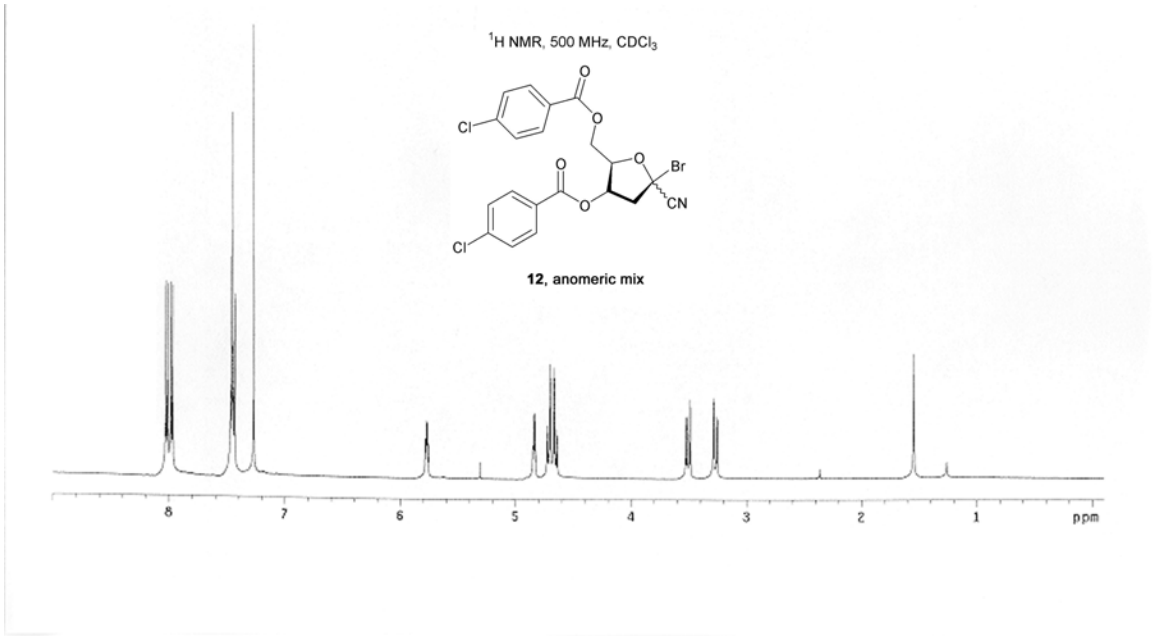
saturated aqueous $NaHCO_3$ and 2x50 mL brine. The organic phase was dried over anhydrous $Na_2SO_4$, evaporated, and the residue was purified by flash chromatography (2:3 ethyl acetate-hexanes with 1% triethylamine) to afford 0.354 g of **7b** in an 83% yield as a white foam. **7b** $^1$H NMR ($CD_3CN$): δ 8.05 (d, 2H, $J$ = 8.00 Hz, H-6), 7.69 (d, 2H, $J$ = 7.50 Hz, H-3), 7.65 (dt, 2H, $J$ = 2.50 Hz, $J$ = 7.50 Hz, H-4), 7.51 (t, 2H, $J$ = 7.50 Hz, H-5), 7.37 (m, 8H, Ar-H), 7.22 (m, 16H, Ar-H), 6.78 (s, 2H, Ar-H), 6.76 (s, 2H, Ar-H), 5.21 (d, 1H, $J$ = 14.00 Hz, Ar-C$H_2$), 5.20 (d, 1H, $J$ = 14.00 Hz, Ar-C$H_2$), 5.18 (d, 1H, $J$ = 14.00 Hz, Ar-C$H_2$), 5.17 (d, 1H, $J$ = 14.00 Hz, Ar-C$H_2$), 4.66 (m, 1H, H-3′), 4.61 (m, 1H, H-3′), 4.39 (m, 1H, H-4′), 4.35 (m, 1H, H-4′), 3.82 (m, 2H, POC$H_2$), 3.73 (s, 6H, OC$H_3$), 3.66 (m, 2H, POC$H_2$), 3.59 (m, 4H, NC$H$), 3.30 (dd, 1H, $J_{5'a,4'}$ = 4.00 Hz, $J_{5'a,5'b}$ = 10.50 Hz, H-5′$_a$), 3.25 (dd, 1H, $J_{5'a,4'}$ = 4.00 Hz, H-5′$_a$), 3.11 (m, 1H, H-5′$_b$), 3.09 (m, 1H, H-5′$_b$), 2.84 (dd, 1H, $J_{2'a,3'}$ = 6.00 Hz, $J_{2'a,2'b}$ = 13.50 Hz, H-2′$_a$), 2.83 (dd, 1H, $J_{2'a,3'}$ = 6.00 Hz, $J_{2'a,2'b}$ = 14.00 Hz, H-2′$_a$), 2.78 (dd, 1H, $J_{2'b,3'}$ = 5.00 Hz, $J_{2'b,2'a}$ = 14.00 Hz, H-2′$_b$), 2.72 (dd, 1H, $J_{2'b,3'}$ = 5.50 Hz, $J_{2'b,2'a}$ = 13.50 Hz, H-2′$_b$), 2.65 (t, 2H, $J$ = 6.00 Hz, C$H_2$CN), 2.51 (t, 2H, $J$ = 6.00 Hz, C$H_2$CN). $^{13}$C NMR ($CD_3CN$): δ 159.81, 148.25, 145.34, 145.24, 136.18, 136.09, 135.06, 133.46, 133.44, 131.25, 129.85, 129.79, 129.23, 129.21, 129.18, 128.90, 128.07, 125.85, 119.58, 119.38, 117.41, 117.37, 114.10 (*C*N), 101.40 & 101.34 (C-1′), 88.37, 88.34, 88.24, 88.19 (C-4), 87.48 (*C*-Ar$_3$), 73.88, 73.75, 73.24, 73.11 (C-3′), 65.55 & 65.50 (Ar-*C*H$_2$), 63.68 & 63.49 (C-5′), 59.70, 59.56, 59.41 (PO*C*H$_2$), 55.96 (O*C*H$_3$), 47.42 & 47.19 (N*C*H), 44.21, 44.18, 44.11, 44.08 (C-2′), 24.98, 24.92, 24.86 (*C*H$_3$), 21.14, 21.09, 21.03, 20.97 (*C*H$_2$CN). $^{31}$P NMR, 162 MHz ($CD_3CN$): δ 149.4, 149.3 (s). HRMS-FAB (*m*/*z*): $[M + Na]^+$ calcd for $C_{42}H_{47}N_4O_8P$+Na, 789.3029; found, 789.3026.

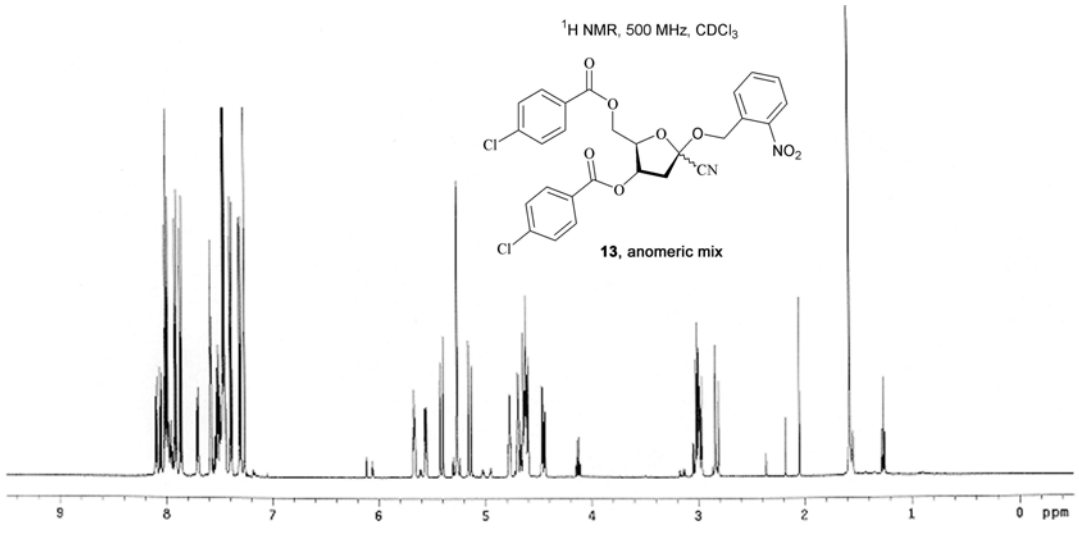
The alpha anomer **7a** was prepared by an analogous method in 86% yield starting from **15a**. Compound **7a** $^{1}$H NMR ($CD_3CN$): δ 8.10 (d, 2H, $J_{6,5}$ = 8.00 Hz, H-6), 7.87 (d, 2H, *J* = 8.00 Hz, H-3), 7.73 (dt, 2H, *J* = 3.00 Hz, *J* = 8.00 Hz, H-4), 7.55 (t, 2H, *J* = 8.00 Hz, H-5), 7.46 (m, 8H, Ar-H), 7.34 (m, 12H, Ar-H), 7.27 (m, 4H, Ar-H), 6.89 (m, 4H, Ar-H), 5.25 (d, 2H, *J* = 14.50 Hz, Ar-C$H_2$), 5.17 (d, 2H, *J* = 13.50 Hz, Ar-C$H_2$), 4.55 (m, 2H, H-3′), 4.33 (m, 1H, H-4′), 4.27 (m, 1H, H-4′), 3.77 (d, 6H, OC$H_3$), 3.73 (m, 2H, POC$H_2$), 3.59 (m, 2H, POC$H_2$), 3.53 (m, 4H, NC*H*), 3.35 (m, 2H, H-5′$_a$), 3.10 (m, 2H, H-5′$_b$), 2.98 (m, 2H, H-2′$_a$), 2.71-2.55 (m, 4H, H-2′$_b$ & C$H_2$CN), 2.42 (t, 2H, *J* = 6.00 Hz, C$H_2$CN). $^{13}$C NMR ($CD_3CN$): δ 159.81, 148.25, 145.34, 145.24, 136.18, 136.09, 135.06, 133.46, 133.44, 131.25, 129.85, 129.79, 129.23, 129.21, 129.18, 128.90, 128.07, 125.85, 119.58, 119.38, 117.41, 117.37, 114.10 (*C*N), 101.40 &101.34 (C-1′), 88.37, 88.34, 88.24, 88.19 (C-4′), 87.48 (*C*-Ar$_3$), 73.88, 73.75, 73.24, 73.11 (C-3′), 65.55 & 65.50 (Ar-*C*H$_2$), 63.68, 63.49 (C-5′), 59.70, 59.56, 59.41 (POCH$_2$), 55.96 (OCH$_3$), 47.42 & 47.19 (NCH), 44.21, 44.18, 44.11, 44.08 (C-2′), 24.98, 24.92, 24.86 (*C*H$_3$), 21.14, 21.09, 21.03, 20.97 (*C*H$_2$CN). $^{31}$P NMR, 121 MHz ($CD_3CN$): δ 150.0, 149.5 (s). HRMS-FAB (*m*/*z*): [M + Na]$^+$ calcd for $C_{42}H_{47}N_4O_8P$+Na, 789.3029; found, 789.3026.

**NMR data for synthetic compounds *10-15* and compound *7*.**

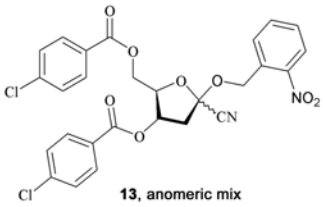



$^1$H NMR, 500 MHz, $CDCl_3$
Cl
O
O
O
O
Br
CN
Cl
O
O
**12**, anomeric mix
8
7
6
5
4
3
2
1
ppm


$^1$H NMR, 500 MHz, $CDCl_3$
Cl
O
O
O
O
O
$NO_2$
CN
Cl
O
O
**13**, anomeric mix
9
8
7
6
5
4
3
2
1
0
ppm

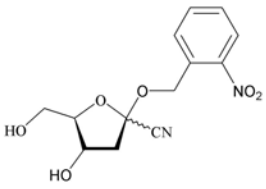
$^1$H-$^1$H COSY, 500 MHz, $CDCl_3$



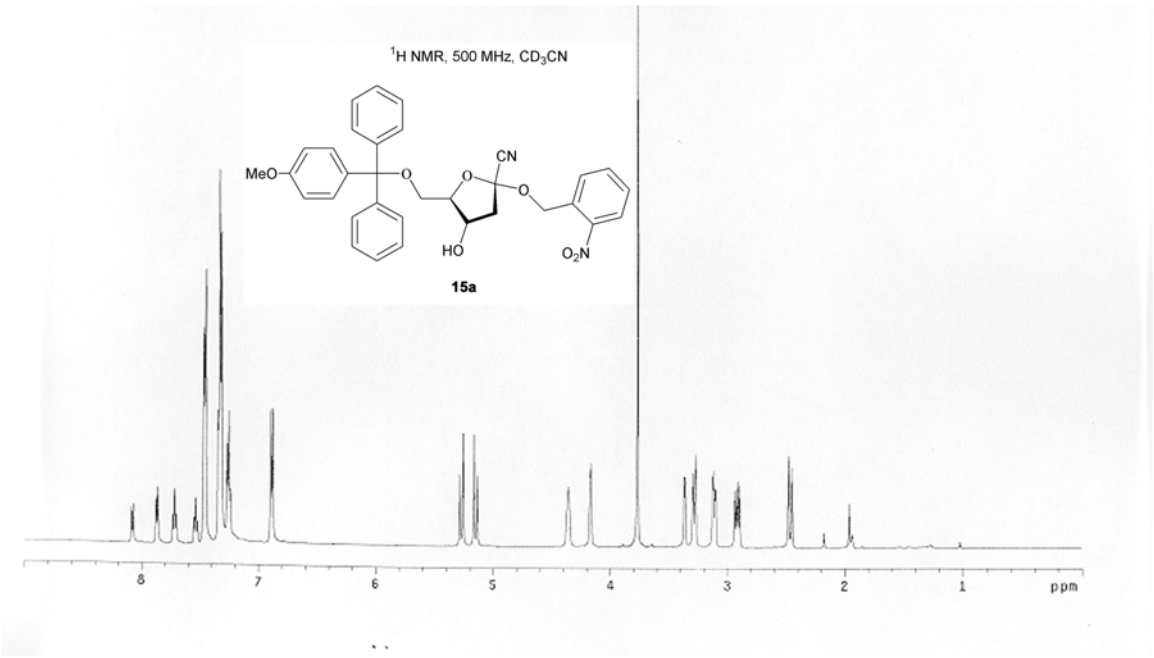
**13**, anomeric mix

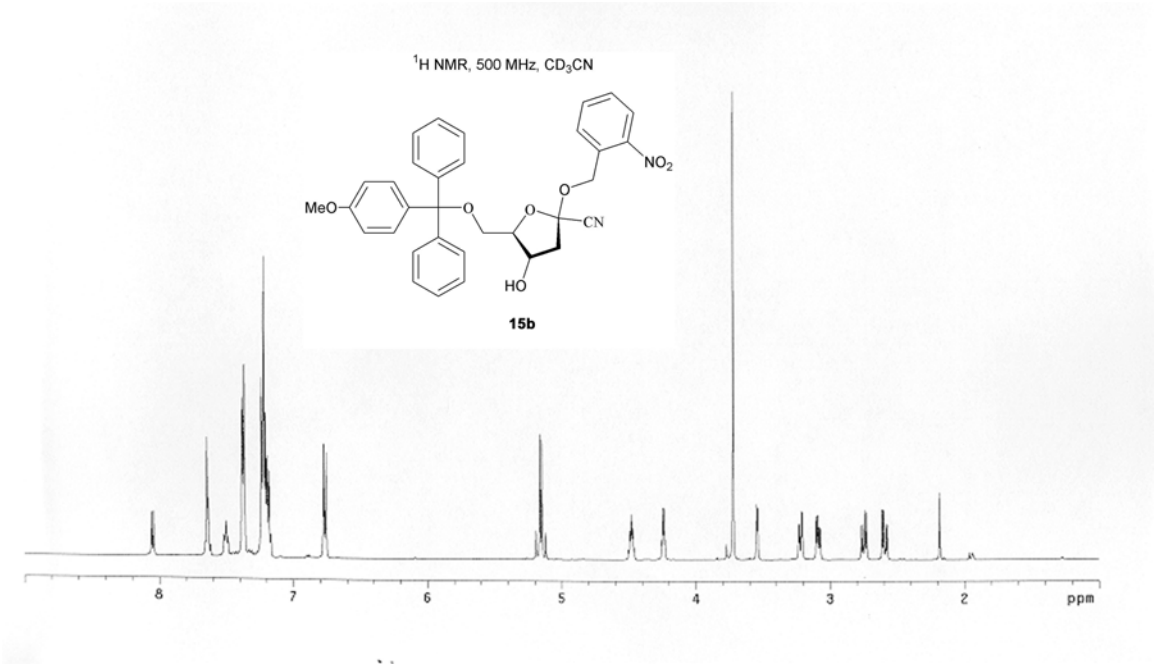


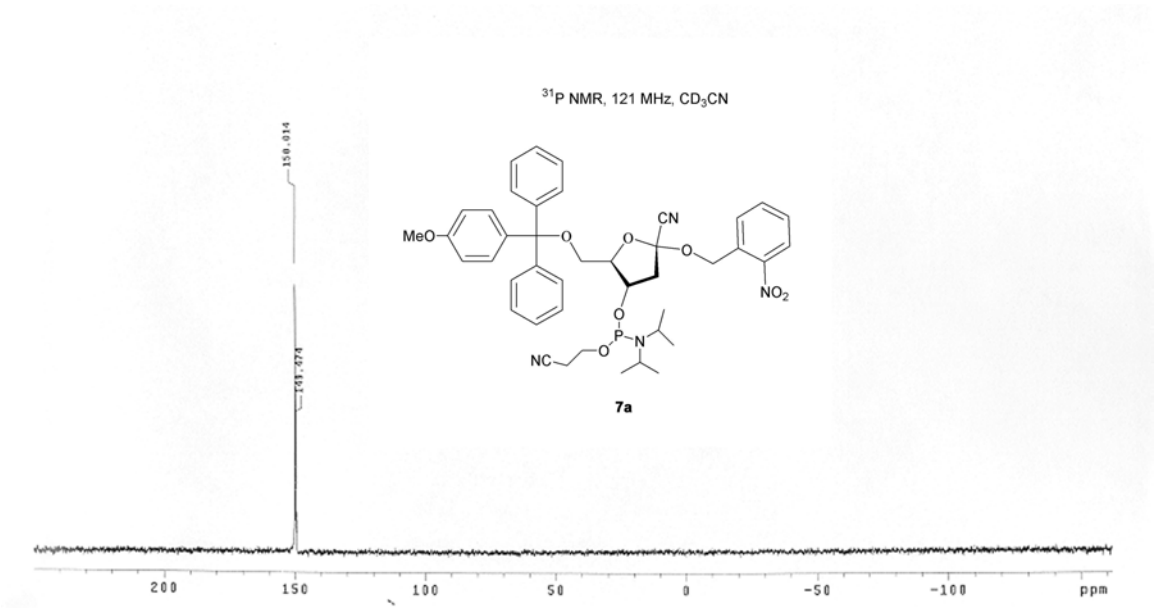

$^1$H NMR, 500 MHz, DMSO

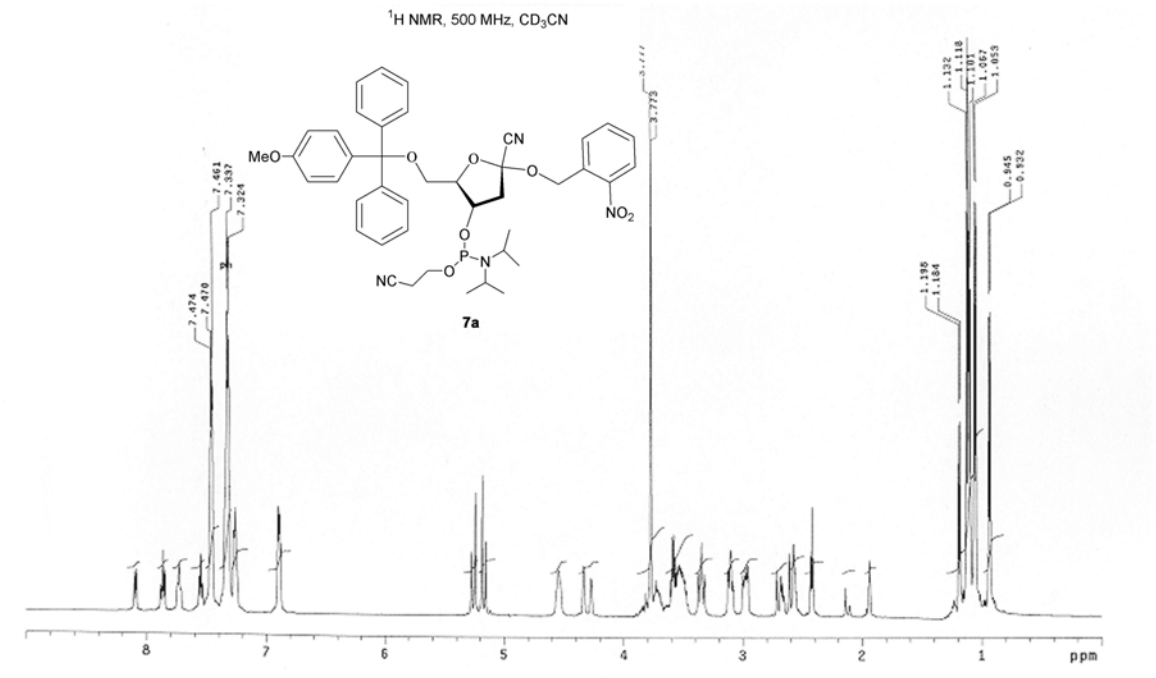


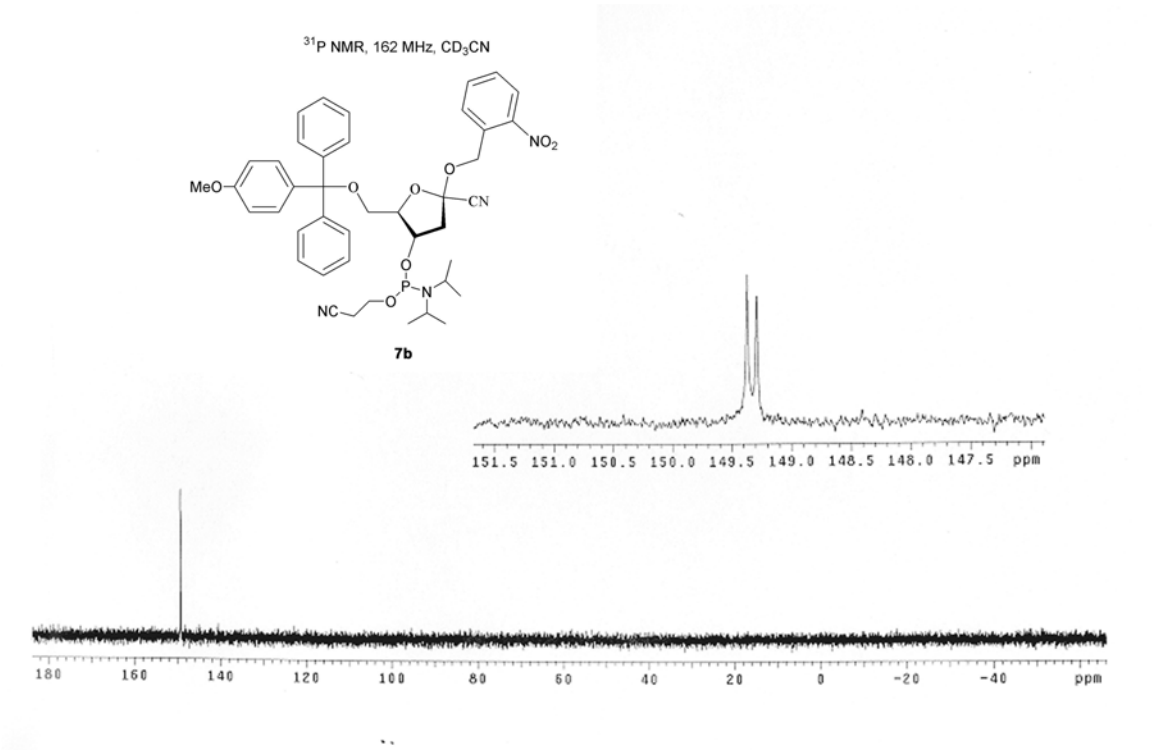
**14**, anomeric mix

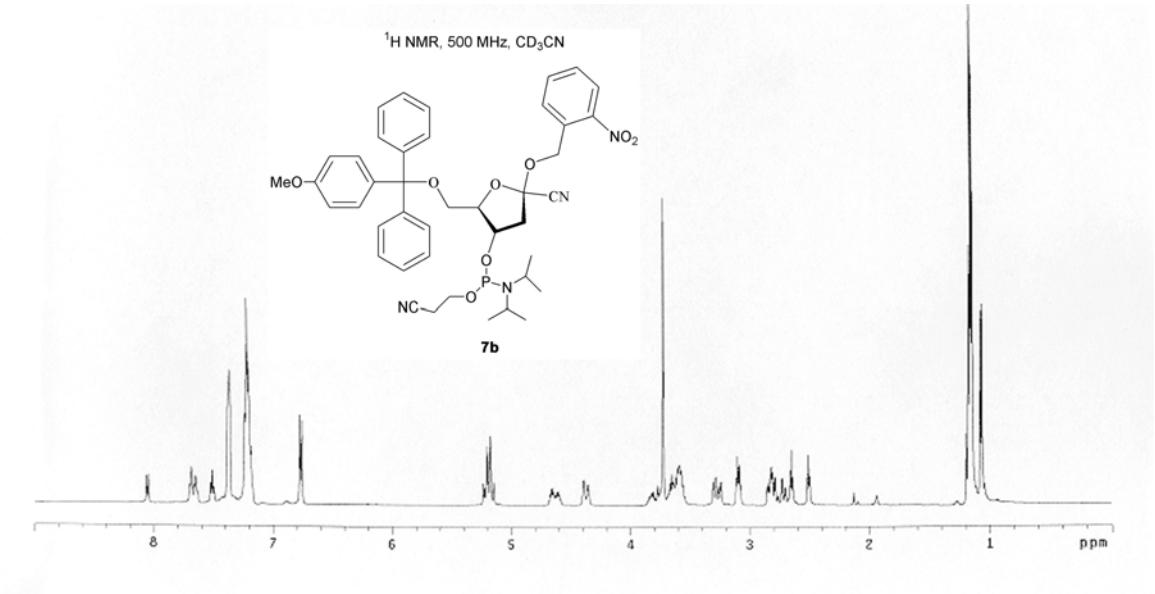

$^1$H NMR, 500 MHz, $CD_3CN$
MeO
O
O
CN
O
HO
$O_2N$
15a
8
7
6
5
4
3
2
1
ppm

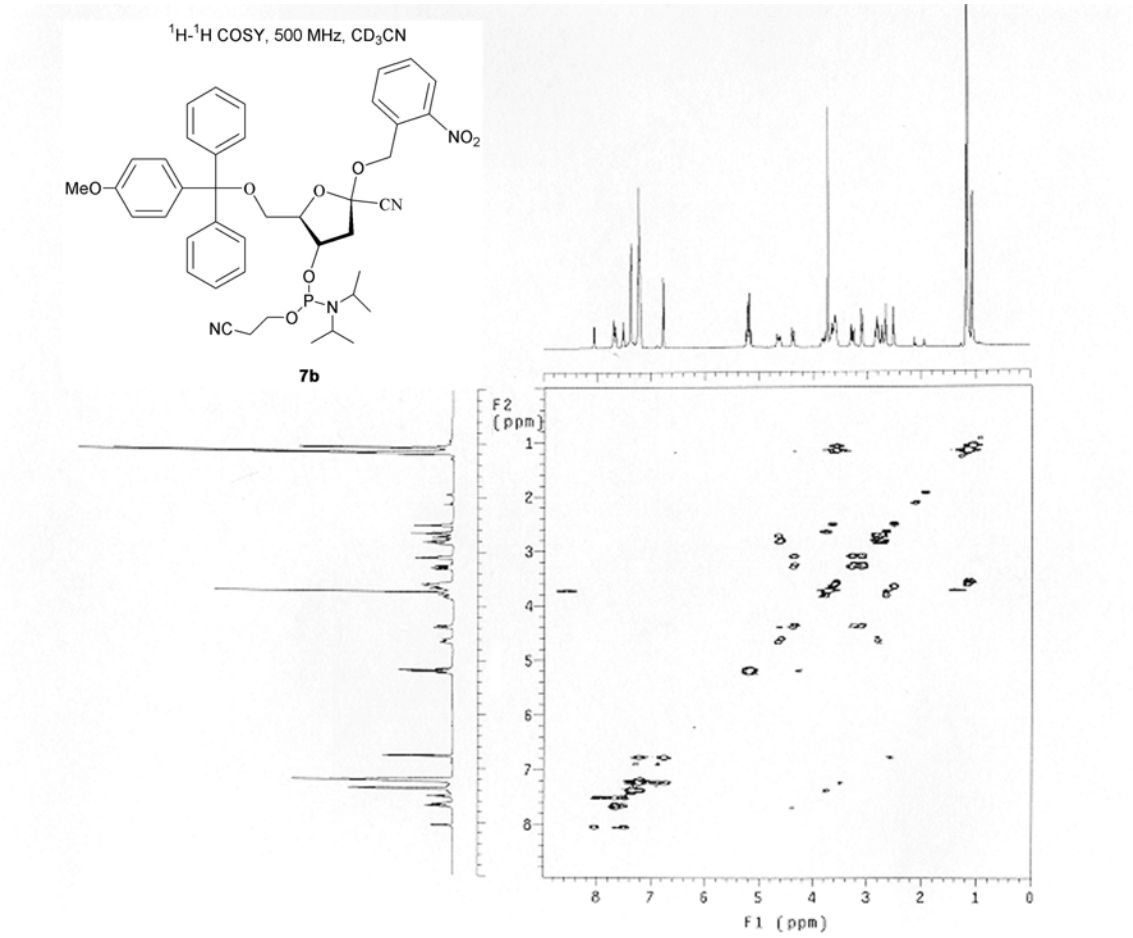

$^1$H NMR, 500 MHz, $CD_3CN$
MeO
O
O
O
CN
HO
$NO_2$
15b
8
7
6
5
4
3
2
ppm


31P NMR, 121 MHz, CD3CN
150.014
149.474
MeO
CN
NO2
NC
7a
200
150
100
50
0
-50
-100
ppm


1H NMR, 500 MHz, CD3CN
7.474
7.470
7.461
7.337
7.324
3.775
1.332
1.118
1.101
1.067
1.053
0.945
0.922
1.198
1.184
MeO
CN
NO2
NC
7a
8
7
6
5
4
3
2
1
ppm


31P NMR, 162 MHz, CD3CN
MeO
NO2
CN
NC
7b
151.5 151.0 150.5 150.0 149.5 149.0 148.5 148.0 147.5 ppm
180 160 140 120 100 80 60 40 20 0 -20 -40 ppm


1H NMR, 500 MHz, CD3CN
MeO
NO2
CN
NC
7b
8 7 6 5 4 3 2 1 ppm




**Synthesis and crystallographic analysis of β-2-Nitrobenzyl 1′-cyano-3′,5′-di-(*O-p*-chlorobenzoyl)-2′-deoxy-D-ribofuranoside**. MMT-ether **15b** (0.206 g, 0.34 mmol) was stirred under nitrogen and in the dark in 20 mL detritylation solution (3% dichloroacetic acid in dichloromethane) for 45 min. The reaction was concentrated in vacuo and the free nucleoside **14b** was purified by flash chromatography (100% ethyl acetate). The purified diol **14b** was dissolved in 0.20 mL anhydrous pyridine, and 1.1 mL 4-chlorobenzoyl chloride (1.51 g, 8.65 mmol) was added dropwise to the reaction. After 90 min, the solid was broken up with 5 mL $CH_2Cl_2$ and 5 mL of $H_2O$. The $CH_2Cl_2$ layer

was separated and the $H_2O$ phase was back-extracted with 5 mL $CH_2Cl_2$. The combined organic layers were washed with 2 x 5 mL saturated $NaHCO_3$, 10 mL $H_2O$, and 10 mL brine. The solution then was dried over anhydrous $Na_2SO_4$, filtered, and evaporated to dryness. The residue was purified by flash chromatography (40% ethyl acetate in hexanes) to afford **13b** in anomerically pure form (0.089 g, 43%). Crystals of **13b** were grown from a solution of acetone:$H_2O$ (12:1) at 4 °C. Empirical Formula: $C_{27}H_{20}N_2O_8Cl_2$; Formula Weight: 571.37; Crystal Color, Habit: colorless, needle; Crystal Dimensions: 0.44 x 0.04 x 0.02 mm; Crystal System: monoclinic; Lattice Type: Primitive; Lattice Parameters: a = 12.787(3) Å, b = 5.7858(13) Å, c = 16.989(4) Å, β = 91.899(4)°, V = 1256.2(4) $Å^3$; Space Group: $P2_1$ (#4); Z = 2; GOF = 1.34; R = 0.01. See figure S1 for the crystal structure and CIF file for complete data on the crystal.

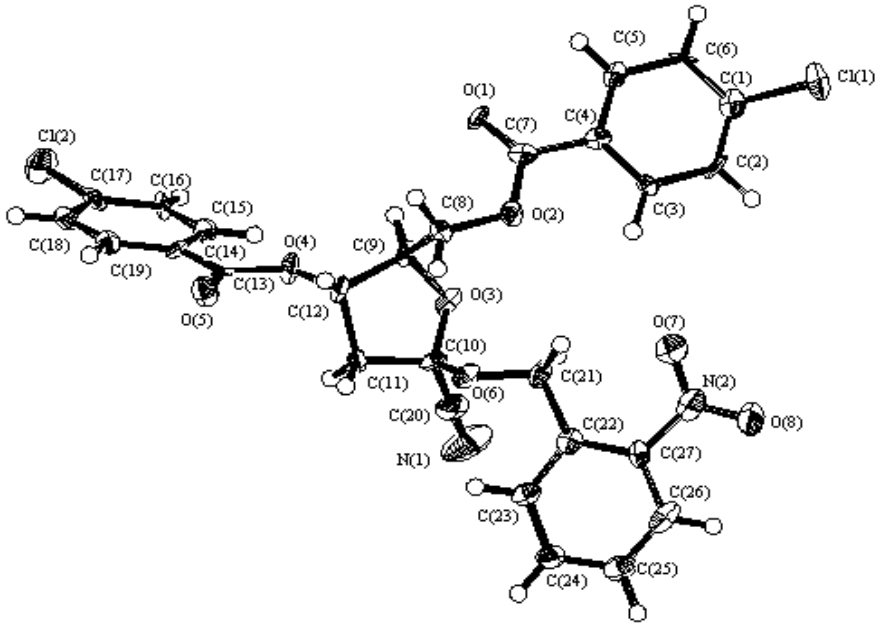



**Figure S1**. Crystal structure of compound **13b**, with the C1′ 2-nitrobenzyl substituent in the β-configuration.

**DNA Synthesis, Deprotection, and Purification.** Standard unmodified DNA oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA) and were purified by denaturing polyacrylamide gel electrophoresis (PAGE) and desalted using G-25 Sephadex columns (NAP, Amersham). Modified oligonucleotides were synthesized on a Pharmacia Gene Assembler Plus at the 1.3 μmole scale using phosphoramidites with labile base protecting groups for the natural nucleotides (Glen Research, "Ultramild": Ac for dC, *i*-Pr-PAC for dG, and PAC for dA). Three caged-analog oligonucleotides, each containing a different DNA sequence flanking the caged analog (**8**), were prepared: **16**, **18**, and **19**. One of these oligonucleotides, **16**, of sequence 5′–TGTGCC–***8***–AACTTACCGT–3′, was used as a model system for subsequent proof of the decaging methodology. For syntheses of caged-analog containing oligonucleotides, phosphoramidite **7a** or **7b** was dissolved to a concentration of 0.15 M in anhydrous $CH_3CN$. A coupling time of 6 min was used for insertion of the analog during DNA synthesis. After coupling, analog-containing DNA was detritylated using 3% dichloroacetic acid in $CH_2Cl_2$ for 6 min (at 2.5 mL/min) to remove the 5′-MMT group prior to addition of the next nucleotide. Coupling yields for each nucleotide addition were calculated by UV/visible spectrometric analysis of the DMT and MMT cation effluents produced at the beginning of each cycle (Damha, M. J., *et al. Nucleic Acids Res.* **1990**, *18*, 3813-3815). Average coupling yields for the phosphoramidite beta anomer **7b** were ~70%, with an average coupling yield of >90% for addition of the next nucleotide. The alpha anomer **7a** was introduced with an average coupling yield of ~64% and was extended in >95% yield. After the synthesis, the oligonucleotide was deprotected by treatment with 1.5 mL of 1.0 M methanolic $NH_3$ at 24 °C for 16 h. Following

deprotection, the $NH_3$/MeOH was removed using a Speedvac concentrator, and the resulting material was redissolved in TE buffer (10 mM Tris-HCl, 1.0 mM EDTA, pH 7.0) and purified by either PAGE or RP-HPLC.

For routine gel purifications, oligonucleotides were loaded on 20% PAGE (29:1 acrylamide:bisacrylamide, 8 M urea) gels in 1x gel loading buffer (1x GLB; 25% (v/v) glycerol, 0.05 M EDTA, 0.5% (w/v) SDS, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol). After electrophoresis, using 1x TBE (0.18 M Tris, 0.18 M borate, 2 mM $Na_2EDTA$) as a running buffer, the oligonucleotide was imaged by UV shadowing, excised with a sterilized razor blade and eluted from the gel slice (10 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, pH 7.5) at room temperature (4 °C for lactone-containing DNA). The eluent and the gel pieces were combined and filtered through a 0.2 micron filter, and DNA was recovered by precipitation with one volume of isopropanol from the filtrate solution which had been supplemented to 0.3 M NaOAc (pH 5.2). The oligonucleotide was redissolved in TE buffer (pH 7.0) and was desalted by G-25 Sephadex chromatography, eluting into a storage buffer of 0.5x TE (pH = 7.0).

Crude synthetic, gel-purified, or lesion-containing oligonucleotides were purified by RP-HPLC using a Waters series 600 HPLC and a Beckman Ultrasphere reverse-phase C-18 column (10 mm x 25 cm). DNA samples were loaded in TE buffer and were eluted with an acetonitrile gradient (0-20%) in 0.1 M triethylammonium acetate buffer at a flow rate of 3 mL/min over 40 min. Peaks were monitored at 254 nm, or across the spectral width by diode array detection. Oligonucleotide-containing fractions were collected, frozen, and lyophilized to dryness using a Speedvac concentrator. The lyophilized material was redissolved in TE buffer (pH = 7.0) and stored at –20 °C. All purified

oligonucleotide solutions were quantitated by measuring the UV absorption at 260 nm and using extinction coefficients calculated by the nearest-neighbors method (Richards, E. G., in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids, Vol. 1*; 3rd ed.; Fasman, G. D., Ed.; CRC Press: Cleveland, 1975). The nitrobenzyl nucleoside analog was assigned an extinction coefficient equivalent to a thymidine residue for these calculations.

**Electrospray mass spectrometric characterization of DNA oligonucleotides**. Electrospray mass spectral (ES-MS) data for oligonucleotides was obtained on a Micromass Quattro I spectrometer in negative ion electrospray ionization mode. Samples for ES-MS were prepared by RP-HPLC. A large scale (~10 nmol) purification of oligonucleotide **16** was performed by RP-HPLC using the conditions described above. The caged-analog DNA (**16**) eluted at ~31.5 min under these conditions. The peak was collected, frozen, and evaporated to dryness using a Speedvac concentrator. The sample of **16** was redissolved in ~20 μL of MeOH:$H_2O$ (1:1), and was analyzed by ES-MS. Oligonucleotide **16**, calcd: 5188.3, found *m/z* = 5188.0.

Lactone-containing oligonucleotide **17** was prepared by performing a photolytic decaging of ~10 nmole of **16** under standard conditions (see below), followed by direct injection of the photolysis reaction on the RP-HPLC column. Lactone-containing DNA **17** eluted with a retention time of ~29.5 min under the standard RP-HPLC conditions. After purification, the sample was isolated and analyzed by ES-MS as described for **16**. Oligonucleotide **17**, calcd: 5026.2, found *m/z* = 5026.4.

**Oligonucleotide 5′-radiolabeling.** Oligonucleotide (20 pmol) was incubated at 37 °C for 40 min with 24.5 units of polynucleotide kinase (USB) in a 20 μL reaction

containing 10 mM Tris-acetate, 10 mM magnesium acetate, and 50 mM potassium acetate and 24 pmol (8.4 Ci/μL) [γ-$^{32}$P]ATP (7000 Ci/mmol). The reaction was quenched with an equal reaction volume of 2x GLB, and the sample was purified by 20% PAGE. Radiolabeled DNA bands were imaged on film, excised with a sterilized razor blade, and the oligonucleotide was retrieved by electroelution using a Schleicher & Schuell Elutrap®. Electroelution was performed at 4 °C for 1 h at 250 volts using TAE buffer (40 mM Tris-acetate, 1 mM EDTA) at pH 7.0. The oligonucleotide was quantitated by liquid scintillation counting.

**Photolytic generation of DNA 2-deoxyribonolactone lesions**. (Figure S2) Oligonucleotide **16** was purified by RP-HPLC and a portion was radiolabeled. A solution of caged-DNA **16** was prepared (0.5 μM unlabeled and 2.5 nM labeled) in 10 mM HEPES and 0.1 mM EDTA at pH 7.5. This sample was irradiated for 45 min in a 1.7 mL microfuge tube by an RPR-200 Rayonet Chamber Reactor (Southern New England Ultraviolet) using sixteen 350 nm bulbs (14 W each, reactor temperature ~35 °C). Reaction aliquots were removed at 30 s, 1, 5, 20, and 40 min, and stored on ice. After all timed aliquots were collected, an equal volume of freshly prepared 20 mM NaOH was added to each (final pH 11.8), and the irradiated samples were incubated at 50 °C for 35 min, combined with an equal volume of 2x GLB and analyzed by 20% PAGE. A sample of **17** was prepared by removal of a second aliquot at 40 min, but this material was not heated under alkaline conditions. A third sample of **17** was prepared by removal of a third aliquot at 40 min, and was treated with an equal volume of 200 mM piperidine, then incubated at 95 °C for 30 min. Radiolabeled **16** was used as a marker for the full-length material.

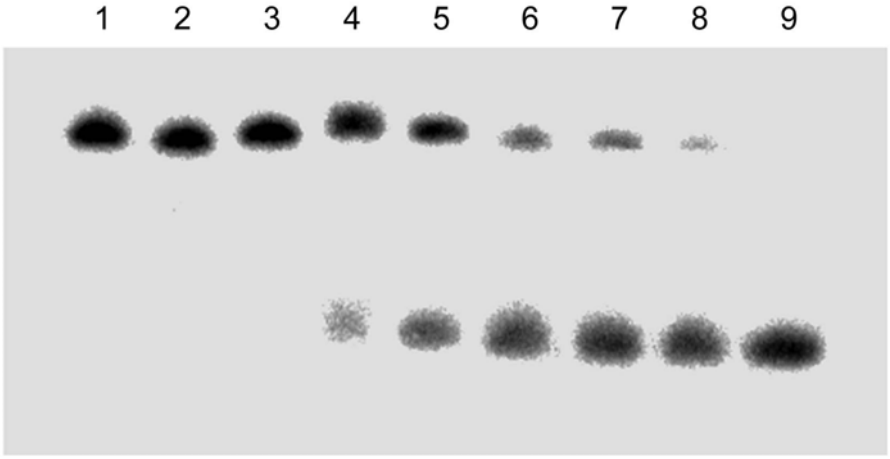



**Figure S2**. Gel electrophoretic demonstration of the photolytic production and cleavage of deoxyribonolactone lesions within DNA. Oligonucleotide **16** was irradiated at 350 nm in 10 mM HEPES and 0.1 mM EDTA (pH 7.5). Aliquots of the irradiation reaction were removed (lanes 2-9) and treated with 20 mM NaOH for 30 min at 50 °C (lanes 3-8) or 200 mM piperidine for 30 min at 95 °C (lane 9) . Lanes: (1) **16** marker; (2) t = 40 min, no heat or base (3% cleavage); (3) t = 0 (6%); (4) t = 5 min (24%); (5) t = 10 min (48%); (6) t = 20 min (71%); (7) t = 30 min (83%); (8) t = 40 min (92%); (9) t = 40 min (99%).

**Characterization data for other oligonucleotides containing the caged analog and the lactone lesion**. Two other oligonucleotide sequences were designed to demonstrate the sequence generality of our method for the photolytic introduction of lactone abasic sites within DNA. Oligonucleotides **18a**, **18b**, and **19b** and their corresponding lactone-containing DNA strands were prepared using the same protocols

described for **16** and **17**. Mass spectral data for these oligonucleotides was obtained by ES-MS. For each sequence, ES-MS is given for both the sequence containing the caged lactone analog **8**, and the lactone **2**. In the sequence, **8a** indicates the use of the α-phosphoramidite (**7a**) in the synthesis of the oligo, and **8b** indicates use of the β-phosphoramidite (**7b**).

**5′–ACCTGC–*8a*–GCGATGGACA–3′ (18a).** ES-MS: *m*/*z* with caged analog **8a,** calcd. 5247.4; found 5247.6. With lactone **2**, calcd. 5085.3; found 5085.1.

**5′–ACCTGC–*8b*–GCGATGGACA–3′ (18b).** ES-MS: *m*/*z* with caged analog **8b,** calcd. 5247.4; found 5247.3. With lactone **2**, calcd. 5085.3; found 5084.6.

**5′–GACGGA–*8b*–TCCCTACGCA –3′ (19b).** ES-MS: *m*/*z* with caged analog **8b,** calcd. 5167.3; found 5166.8. With lactone **2**, calcd. 5005.2; found 5004.8.