Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

Can 5-methylcytosine analogues with extended alkyl side chain guide DNA methylation?

Delshanee Kotandeniya,[†ζ] Christopher L. Seiler,[†ζ] Jenna Fernandez,[†ζ] Suresh S. Pujari,[†ζ] Lauren Curwick,[ζ] Kristopher Murphy,[ηζ] Susith Wickramaratne,[ηζ] Shuo Yan,[ζ] Daniel Murphy,[ζ] Yuk Y. Sham,[£] and Natalia Tretyakova[†*]

[†]Department of Medicinal Chemistry, [ζ]The Masonic Cancer Center, [η]Department of Chemistry, and [£]Department of Integrative Biology and Physiology and the University of Minnesota Informatics Institute

University of Minnesota Minneapolis, Minnesota 55455

* To whom correspondence should be addressed: Masonic Cancer Center, University of Minnesota, 2231 6th St SE, Minneapolis, MN 55455, USA. phone: 612-626-3432 fax: 612-626-5135 email: trety001@umn.edu

Table of contents

Experimental procedures

Materials:

All nucleoside phosphoramidites, solvents, and solid supports required for the solid phase synthesis of DNA were acquired from Glen Research Corporation (Sterling, VA). Recombinant human DNA-methyltransferase 1 was purchased from New England Biolabs (Ipswich, MA). Phosphodiesterase I, phosphodiesterase II, and DNase I were obtained from Worthington Biochemical Corporation (Lakewood, NJ). Bovine intestinal alkaline phosphatase was procured from Sigma Aldrich Chemical Company (Milwaukee, WI) while 2'-deoxycytidine hydrochloride was from Berry & Associates Inc. (Dexter, MI). Stable isotope labeled thymidine was purchased from Cambridge Isotope Laboratories (Tewksbury, MA). The rest of the chemicals were purchased either from Fisher Scientific (Fairlawn, NJ), Sigma-Aldrich (Milwaukee, WI) or Alfa Aesar (WardHill, MA) and used without further purification unless otherwise noted. NMR was performed on a 500 MHz Bruker system. Accurate mass measurements were obtained on a LTQ Orbitrap Velos (Thermo Scientific, Pittsburgh, PA) and mass fragmentation data were collected on an Agilent 1100 LC/MSD Ion trap (Agilent Technologies, Santa Clara, CA).

Methods:

5-(Trimethylsilyl)-ethynyl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (2): A flask containing compound **1**(1.5 g, 2.6 mmol) was dried under vacuum for two days. The dried solid was transferred to a flame-dried flask and the solid was dissolved in 100 mL 3:1 dry, deoxygenated solution of DMF:Et_3N (deoxygenated by bubbling with Ar for 30 min). To the solution, 2 molar equivalents of trimethylsilylacetylene (0.73 mL, 5.2 mmol), 0.2 mol eq. of CuI (0.098 g, 0.52 mmol), and $Pd(PPh_3)_4$ (0.30 g, 0.26 mmol) were added sequentially to the flask in an Ar-glove bag. The flask was capped, and the reaction mixture was stirred for 2 days at room temperature under an atmosphere of Ar, under dark conditions. The mixture was diluted with 400 mL EtOAc and the organic layer was washed with water (100 mL x 3) and brine (twice) followed by drying the organic layer with anhydrous Na_2SO_4. The product was purified by two successive silica columns using a gradient of 100% CH_2Cl_2 to 16:1 CH_2Cl_2:2-propanol to give compound **2** in 35% yield. ESI$^+$-MS/MS: *m/z* 552.2 $[M + H]^+ \rightarrow$ *m/z* 208.1 $[M + H$ – deoxyribose$]^+$, Exact mass calc: 552.3104 $[M + H]^+$, Exact mass obs: 552.3093 $[M + H]^+$ (error 2.0 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.23 (br s, 1H), 8.07 (s, 1H), 6.24 (dd, 1H, *J* = 6.0, 6.0 Hz), 5.76 (br s, 1H), 4.35 (ddd, 1H, *J* = 6.3, 3.3, 3.3 Hz), 3.95 (d, 1H, *J* = 2.5 Hz), 3.91 (dd, 1H, *J* = 11.5, 2.0 Hz), 3.75 (dd, 1H, *J* = 11.5, 2.0 Hz), 2.43 (ddd, 1H, *J* = 13.0, 6.0, 4.0 Hz), 2.00 (m, 1H), 0.92 (s, 9H), 0.87 (s, 9H), 0.21 (s, 9H), 0.13 (d, 6H, *J* = 6.5 Hz), 0.05 (d, 6H, *J* = 3.0 Hz). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 164.9, 154.4, 144.1, 101.2, 96.2, 91.1, 88.2, 86.8, 71.7, 62.6, 42.8, 26.2, 25.9, 18.6, 18.1, 0.0, -4.5, -4.8, -5.1, -5.3.

5-Ethynyl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (3): Compound **2** (0.7 g, 1.3 mmol) was dissolved in 10 mL absolute MeOH and this solution was cooled on ice for 5 min. To this was added 3 mL of 1N KOH and the mixture was stirred at room temperature for 3 hours. The reaction was monitored by TLC using 12:1 DCM:MeOH. Once the starting material appeared to have been consumed, the pH of the solution was brought up to pH 7 by drop wise addition of 1N HCl. The solvent was then evaporated under reduced pressure and product was purified on a silica column (DCM:MeOH) to afford a light beige foam 0.55 g (90% yield). ESI^+-MS/MS: *m/z* 480.3 $[M + H]^+ \rightarrow$ *m/z* 136.0 [M + H – deoxyribose]$^+$, Exact mass calc: 480.2708 $[M + H]^+$, Exact mass obs: 480.2699 $[M + H]^+$ (error: 1.9 ppm); ^1H NMR ($CDCl_3$, 600 MHz): δ 8.21 (s, 1H), 6.26 (dd, 1H, *J* = 6.5, 6.5 Hz), 6.11 (br s, 1H), 5.63 (br s, 1H), 4.36 (m, 1H), 3.95 (m, 1H), 3.93 (dd, 1H, *J* = 9.7, 2.3 Hz), 3.77 (dd, 1H, *J* = 9.5, 1.5 Hz), 3.32 (s, 1H), 2.48 (m, 1H), 2.04 (ddd, 1H, *J* = 10.7, 5.3 Hz), 0.92 (s, 9H), 0.88 (s, 9H), 0.13 (d, 6H, *J* = 6.0 Hz), 0.06 (d, 6H, *J* = 3.0 Hz).

5-Ethyl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (4): Compound **3** (0.548 g, 1.14 mmol) was dried overnight under high vacuum and was dissolved in 20 mL of anhydrous MeOH by gently warming it in a warm water bath. The flask was capped and the solution was purged with Ar by bubbling. To this was quickly added 0.35 mol equivalents of 10% Pd-C (0.042 g, 0.4 mmol) and the reaction was stirred under hydrogen for 24 hr. at room temperature. The reaction was monitored by TLC, using 12:1 DCM:MeOH as the mobile phase. The crude reaction mixture was then passed through a Celite® column to remove the Pd-catalyst, and the column washed with MeOH to collect the product. The MeOH solvent was evaporated under reduced pressure to yield 0.547 g of white foam (99% yield). ESI^+-MS/MS: *m/z* 484.3 $[M + H]^+ \rightarrow$ *m/z* 140.0 [M + H – deoxyribose]$^+$, Exact mass calc: 484.3021 $[M + H]^+$, Exact mass obs: 484.3008 $[M + H]^+$ (error 2.7 ppm); ^1H NMR ($CDCl_3$, 500 MHz): δ 7.50 (s, 1H), 6.31 (dd, 1H, *J* = 6.5, 6.5 Hz), 4.34 (ddd, 1H, *J* = 6.0, 3.0, 3.0 Hz), 3.93 (d, 1H, *J* = 3.5 Hz), 3.83 (dd, 1H, *J* = 11.5, 3.5 Hz), 3.75 (dd, 1H, *J* = 11.5, 3.0 Hz), 2.40 (ddd, 1H, *J* = 13.0, 5.5, 3.0 Hz), 2.27 (q, 2H, *J* = 7.5 Hz), 1.93 (m, 1H), 1.15 (t, 3H, *J* = 7.5 Hz), 0.91 (s, 9H), 0.88 (s, 9H), 0.09 (s, 6H), 0.06 (d, 6H, *J* = 3.0 Hz). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 165.1, 155.9, 137.4, 107.2, 87.8, 86.1, 72.1, 63.1, 42.2, 26.1, 25.9, 21.0, 18.5, 18.1, 13.0, -4.5, -4.7, -5.3.

4-*N*-Benzoyl-5-ethyl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (5): Compound **4** (0.125 g, 0.258 mmol) which was dried overnight under high vacuum was co-evaporated three times with 5 mL each of anhydrous pyridine. This compound was then dissolved in 5 mL of anhydrous pyridine and the flask was immediately capped with a rubber septum under argon atmosphere. To the mixture was added 60 µL (0.516 mmol, 2.0 equiv) of benzoyl chloride over a period of 5 minutes at room temperature and the reaction was stirred. The reaction was monitored by TLC (30:1 DCM:MeOH) every ½ hour. Since the reaction still had a lot of starting material remaining after 2 hours, another 40 µL

(0.34 mmol, 1.3 equiv) of benzoyl chloride was added and reaction was monitored for another 2 hours, until all starting material was consumed. At the end of the reaction the solvent was removed under reduced pressure, the crude product was taken up in DCM, and the organic layer was washed three times with 50 mL water followed by drying with anhydrous Na_2SO_4 and the solvent removed under reduced pressure. The crude product was purified on a silica column, using 30:1 DCM:MeOH. After evaporation of solvent, a few drops of MeOH were added to the resulting oil to produce white needle like crystals 0.129 g (85% yield). ESI$^+$-MS/MS: Exact mass calc: 588.3284 $[M + H]^+$, Exact mass obs: 588.3267 $[M + H]^+$ (error 2.9 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 13.4 (s, 1H), 8.30 (dd, 2H, *J* = 8.5, 1.5 Hz), 7.62 (s, 1H), 7.52 (m, 1H), 7.44 (dd, 2H, *J* = 7.5, 7.5 Hz), 6.34 (dd, 1H, *J* = 7.5, 5.5 Hz), 4.41 (m, 1H), 3.98 (ddd, 1H, *J* = 5.0, 2.5 Hz), 3.98 (dd, 1H, *J* = 11.5, 3.0 Hz), 3.78 (dd, 1H, *J* = 11.5, 3.0 Hz), 2.57 (m, 2H), 2.34 (ddd, 1H, *J* = 13.5, 6.0, 2.5 Hz), 2.02 (ddd, 1H, *J* = 13.5, 7.5, 6.0 Hz), 1.25 (t, 3H, *J* = 7.5 Hz), 0.94 (s, 9H), 0.90 (s, 9H), 0.12 (d, 6H, *J* = 2.0 Hz), 0.09 (d, 6H, *J* = 4.0 Hz). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 179.7, 159.5, 147.9, 137.5, 136.4, 132.4, 130.0, 128.2, 117.5, 88.2, 85.7, 72.5, 63.2, 41.7, 26.1, 25.9, 21.6, 18.6, 18.1, 13.9, -4.5, -4.7, -5.3.

4-*N*-Benzoyl-5-ethyl-2'-deoxycytidine (6): Compound **5** (0.046 g, 0.078 mmol) was dried overnight under high vacuum and then co-evaporated three times with 5 mL each of anhydrous pyridine. The flask was immediately fitted with a rubber septum and purged with argon (1-2 min) and 1.3 mL of dry and degassed DCM was added under an atmosphere of argon. To the solution, 102 µL (0.63 mmol, 8 equiv) of triethylamine-trihydrofluoride (Et_3N•3HF) was added dropwise to the flask over a period of 30 minutes using a syringe and the reaction was stirred under an atmosphere of Ar at room temperature for 16 hours. During this time the reaction was monitored every 2 hours by TLC (30:1 DCM:MeOH, if a lot of starting material is remaining, more Et_3N.3HF is added). Upon total consumption of starting compound **5** by TLC, product (**7**) was isolated by purification on a silica column using a gradient of DCM and 30:1 DCM:MeOH resulting in a white solid 0.027 g in 96%, yield. ESI$^+$-MS/MS: *m/z* 360.2 $[M + H]^+$ → *m/z* 244.1 [M + H – deoxyribose]$^+$, Exact mass calc: 360.1554 $[M + H]^+$, Exact mass obs: 360.1553 $[M + H]^+$ (error 0.3 ppm); ^{1}H NMR (MeOD/DMSO-*d6*, 500 MHz): δ 7.47 (d, 2H, *J* = 5.5 Hz), 7.33 (s, 1H), 6.78 (t, 1H, *J* = 7.3 Hz), 6.69 (t, 2H, *J* = 7.5 Hz), 5.52 (dd, 1H, *J* = 6.8, 6.8 Hz), 3.64 (m, 1H), 3.17 (d, 1H, *J* = 3.0 Hz), 3.03 (dd, 1H, *J* = 12.0, 3.0 Hz), 2.96 (dd, 1H, *J* = 11.8, 3.3 Hz), 1.81 (q, 2H, *J* = 7.5 Hz), 1.54 (m, 1H), 1.48 (m, 1H), 0.46 (t, 3H, *J* = 7.5 Hz). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 180.6, 160.7, 149.5, 139.4, 138.5, 133.6, 130.8, 129.3, 118.0, 89.2, 87.0, 71.9, 62.5, 41.7, 21.9, 13.7.

4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-ethyl-2'-deoxycytidine (7): compound **6** (4-*N*-benzoyl-5-ethyl-2'-dC, 0.130 g, 0.361 mmol) was dried under vacuum overnight and co-evaporated three times with 6 mL each of anhydrous pyridine. The flask was immediately capped and purged with Ar and transferred to a glove bag filled with Ar.

Anhydrous pyridine (kept over activated 4Å molecular sieves) was degassed with argon for 20 min. All additions were performed in an Ar glove bag. Anhydrous, degassed pyridine (2 mL) was added to the flask containing dried nucleoside. Dimethoxytrityl chloride (DMTrCl), 0.147 g (0.434 mmol, 1.2 equiv), was added to the nucleoside-pyridine solution. Upon addition of all reagents, the flask was taken out of the glove bag, and the reaction was stirred for 16 hours at room temperature away from light. The reaction was monitored using 5:47.5:47.5 TEA:EtOAc:Hexane on TLC plates soaked in 95:5 hexane:TEA. The reaction mixture was quickly dried under high vacuum and to this was added 30 mL of DCM and the organic layer washed two times with 25 mL each of ice cold water and dried using anhydrous Na_2SO_4 and solvent removed under reduced pressure. The resulting solid was purified on a short silica-flash column packed in 95:5 Hexane:TEA using solvent gradient below: 95:5 Hexane:TEA, 5:25:70 TEA:EtOAc:Hexane, 5:47.5:47.5 TEA:EtOAc:Hexane, 95:5 EtOAc:TEA. Upon evaporation of the solvent, the resulting oily product was dissolved in 1 mL of anhydrous DCM, and dried under high vacuum overnight, to get a white foam 0.155 g of 65% yield. ESI^+-MS/MS: *m/z* 662.6 $[M + H]^+ \rightarrow$ *m/z* 303.1 [dimethoxytrityl cation]$^+$, 244.0 [M + H – deoxyribose]$^+$, Exact mass calc: 684.2680 [M + Na]$^+$, Exact mass obs: 684.2680 [M + Na]$^+$; ^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.20 (dd, 1H, *J* = 7.0, 1.0 Hz), 7.64 (s, 1H), 7.51 (m, 1H), 7.42 (m, 4H), 7.30 (m, 7H), 6.84 (m, 4H), 6.40 (dd, 1H, *J* = 6.0, 5.0 Hz), 4.57 (ddd, 1H, *J* = 5.0, 2.5, 2.5 Hz), 4.07 (ddd, 1H, *J* = 3.0, 3.0, 3.0 Hz), 3.50 (dd, 1H, *J* = 9.0, 3.0 Hz), 3.39 (dd, 1H, *J* = 8.7, 2.7 Hz), 2.60 (br s, 1H), 2.47 (dq, 1H, *J* = 11.0, 2.5 Hz), 2.32 (m, 1H), 2.23 (m, 1H), 2.09 (m, 1H), 1.00 (t, 3H, *J* = 6.3 Hz).

4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-ethyl-2-deoxycytidine-3'-[(2-cyanoethyl)-(*N,N*-diisopropyl)]-phosphoramidite (8): Into a flame dried round bottom flask was added compound **7** (0.075 g, 0.113 mmol), which was dried in an Ar filled glove bag under high vacuum overnight. Anhydrous DCM (kept over activated 4Å molecular sieves) was degassed with argon for 20 minutes. All additions below were performed sequentially in an Ar glove bag: 197 µL of anhydrous *N,N*-diisopropylethylamine (1.13 mmol, 5 equiv as phosphitylating reagent), 2 mL of anhydrous and degassed DCM and 51 µL of 2-Cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.23 mmol, 2 equiv). The flask was capped and the reaction stirred for 16 hours under an Ar atmosphere at room temperature away from light. After 16 hours, 250 µL of anhydrous MeOH was added to prevent the oxidation of the phosphoramidite and 200 µL of Et_3N was added to ensure that the product does not get hydrolyzed during the work-up steps. To this solution was added 30 mL of EtOAc. The organic layer was extracted three times with ice-cold brine, dried using anhydrous Na_2SO_4, and the solvent was removed under reduced pressure. The resulting solid was dissolved in a few milliliters of 5:95 TEA:EtOAc and a short silica-flash column run to isolate the product using the gradient elution described below: 95:5 Hexane:TEA, 5:25:70 TEA:EtOAc:Hexane, 5:47.5:47.5

TEA:EtOAc:hexane, 95:5 EtOAc:TEA. Evaporation of the solvent resulted in a white foam 0.059 g of 62% yield. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.28 (d, 2H, *J* = 7.6 Hz), 7.67 (s, 1H), 7.51 (t, 1H, 8.0 Hz), 7.42-7.27 (m, 11H), 6.85 (m, 4H), 6.41 (dd, 1H, *J* = 13.8, 7.6 Hz), 4.66 (m, 1H), 4.18 (m, 1H), 3.79 (s, 6H), 3.63-3.48 (m, 4H), 3.35-3.30 (m, 2H), 2.62 (t, 2H, *J* = 6.4 Hz), 2.41 (dd, 1H, *J* = 6.4, 6.4 Hz), 2.31 (m, 1H), 2.17 (m, 1H), 1.25-1.16 (m, 12H), 0.96 (t, 3H, *J* = 8.4 Hz). ^{31}P NMR ($CDCl_3$, 162 MHz): δ 147.70, 147.29.

5-Propyn-1-yl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (9): 5-Iodo-3',5'-TBS-2'-dC (**1**) (2 g, 3.44 mmol) was dried under vacuum for two days. The dried solid was transferred to a flame-dried two-necked flask inside an Ar-glove bag. To the flask was added 0.397 g of $Pd(PPh_3)_4$ (0.344 mmol, 0.1 equiv) and 0.033 g CuI (0.172 mmol, 0.05 equiv). The flask was capped and 50 mL of anhydrous, deoxygenated solution of DMF:Et_3N (4:1) was added to the flask along with constant purging of propyne gas. The pressure inside the flask was kept at a constant of 8 psi throughout the reaction. The reaction was stirred for 48 hours at room temperature, away from light. At the end of the reaction, the solvents were evaporated, and crude product was taken up in 150 mL of DCM. The organic layer was washed five times with 75 mL each of water and twice with 50 mL each of brine followed by drying organic layer with anhydrous Na_2SO_4. The product was purified by two successive silica columns with solvent gradient from 100% DCM to 9:1 DCM:MeOH to afford light a beige solid in (1.44 g) 85% yield. ESI$^+$-MS/MS: *m/z* 494.3 $[M + H]^+$ → *m/z* 150.1 [M + H – deoxyribose]$^+$, Exact mass calc: 494.2865 $[M + H]^+$, Exact mass obs: 494.2849 $[M + H]^+$ (error 3.2 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.44 (br s, 1H), 8.00 (s, 1H), 6.25 (dd, 1H, *J* = 6.0, 6.0 Hz), 5.81 (br s, 1H), 4.34 (m, 1H), 3.90 (m, 2H), 3.74 (dd, 1H, *J* = 12.0, 3.0 Hz), 2.40 (ddd, 1H, *J* = 13.5, 6.5, 4.5 Hz), 2.00 (m, 4H), 0.92 (s, 9H), 0.86 (s, 9H), 0.12 (d, 6H, *J* = 8.0 Hz), 0.04 (s, 6H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 165.2, 154.5, 142.9, 91.8, 91.7, 87.9, 86.4, 71.3, 71.1, 62.5, 42.6, 26.1, 25.8, 18.5, 18.1, 4.5, -4.5, -4.8, -5.4, -5.5.

5-Propyl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (10): To a flask containing 0.075 g of dry 5-Propyn-1-yl-3',5'-*t*-butyldimethylsilyl-2'-deoxycytidine (**9**) was added 5 mL of anhydrous MeOH. The flask was capped, and the solution was purged with Ar by bubbling for 5 min. To this was added 20 weight % of 10% Pd-C catalyst (0.015 g, 0.14 mmol) and under hydrogen gas and stirring the reaction at room temperature, away from light for 2 days. The reaction was monitored by TLC, using 12:1 DCM:MeOH as the mobile phase. The crude reaction was then passed through a Celite® column to remove the Pd-catalyst, and the column washed with MeOH to collect product **10**. The solvent was evaporated under reduced pressure to yield white crystals (0.068 g) in 90% yield. ESI$^+$-MS/MS: *m/z* 498.2 $[M + H]^+$ → *m/z* 154.1 [M + H – deoxyribose]$^+$, Exact mass calc: 498.3178 $[M + H]^+$, Exact mass obs: 498.3163 $[M + H]^+$ (error 3.0 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.46 (s, 1H), 6.30 (dd, 1H, *J* = 6.8, 6.8 Hz), 4.33 (m, 1H), 3.91 (ddd, 1H, *J* = 6.0, 3.0,

3.0 Hz), 3.82 (dd, 1H, *J* = 11.0, 3.0 Hz), 3.74 (dd, 1H, *J* = 11.0, 3.0 Hz), 2.37 (ddd, 1H, *J* = 13.0, 6.0, 3.0 Hz), 2.20 (t, 2H, *J* = 7.5 Hz), 1.91 (m, 1H), 1.51 (qt, 2H, *J* = 7.5, 7.5 Hz), 0.92 (t, 3H, *J* = 7.5 Hz), 0.90 (s, 9H), 0.86 (s, 9H), 0.08 (d, 6H, *J* = 2.0 Hz), 0.04 (d, 6H, *J* = 2.5 Hz). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 165.3, 155.8, 137.9, 105.9, 87.7, 85.9, 72.0, 63.0, 42.2, 30.0, 26.1, 25.9, 21.8, 18.5, 18.1, 13.8, -4.5, -4.8, -5.3.

4-*N*-Benzoyl-5-propyl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (11): Compound **10** (0.2 g, 0.4 mmol) was co-evaporated three times with 5 mL each of anhydrous pyridine, and the flask was flushed with Ar. To this was added 6 mL of anhydrous pyridine, followed by purging of the solution mixture with argon by bubbling for 5 min. To the flask 56 µL of benzoyl chloride (0.48 mmol, 1.2 equiv) was added by syringe over 5 minutes. This mixture was stirred at room temperature under an argon atmosphere for 4 hours. The course of the reaction was monitored by TLC (30:1 DCM:MeOH). Analysis of the crude mixture by mass spectrometry indicated the dibenzoylated product to be present. The crude product mixture was dissolved in 2.5 mL of 65:30:5 pyridine:MeOH:H_2O and this mixture was cooled on ice. To this was added 2.5 mL of cold 2N NaOH solution, in 65:30:5 pyridine:MeOH:H_2O. The reaction was stirred on ice for 10 min. Ammonium chloride (0.1 g, 1.2 equiv) was added to quench the reaction. The crude sample was dried and re-dissolved in DCM, washed three times with 30 mL water and dried with Na_2SO_4, and the organic layer was removed to result in crude solid **11**. Compound **11** was further purified by silica column using 15:1 DCM:MeOH to give colorless solid (0.217 g) in 90% yield. ESI$^+$-MS/MS: Exact mass calc: 602.3440 $[M + H]^+$, Exact mass obs: 602.3421 $[M + H]^+$ (error 3.2 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.29 (dd, 2H, *J* = 7.0, 1.5 Hz), 7.62 (s, 1H), 7.52 (t, 1H, *J* = 7.5 Hz), 7.44 (t, 2H, *J* = 7.5 Hz), 6.35 (dd, 1H, *J* = 7.5, 5.5 Hz), 4.41 (m, 1H), 3.97 (ddd, 1H, *J* = 3.0, 3.0, 3.0 Hz), 3.88 (dd, 1H, *J* = 11.3, 2.8 Hz), 3.78 (dd, 1H, *J* = 11.3, 2.8 Hz), 2.51 (m, 2H), 2.33 (ddd, 1H, *J* = 13.0, 6.0, 2.5 Hz), 2.02 (ddd, 1H, *J* = 13.5, 7.5, 6.0 Hz), 1.68 (qt, 2H, *J* = 6.5, 6.5 Hz), 1.00 (t, 3H, *J* = 7.5 Hz), 0.94 (s, 9H), 0.90 (s, 9H), 0.13 (d, 6H, *J* = 2.5 Hz), 0.09 (d, 6H, *J* = 3.5 Hz). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 179.7, 159.7, 148.0, 137.5, 136.9, 132.5, 130.0, 128.2, 116.0, 88.2, 85.6, 72.4, 63.2, 41.8, 30.2, 26.1, 25.9, 22.5, 18.6, 18.1, 14.1, -4.5, -4.7, -5.3.

4-*N*-Benzoyl-5-propyl-2'-deoxycytidine (12): Compound **11** (0.25 g, 0.502 mmol) was dried under vacuum and co-evaporated with of anhydrous pyridine (3 x 5 mL), and the flask was purged with Ar. To this was added 7 mL of anhydrous DCM to dissolve the solid, followed by purging of the solution mixture with argon by bubbling for 5 min. To the flask was injected 327 µL of Et_3N:H_3F (2.01 mmol) by syringe over 30 minutes. The reaction was stirred, away from light, for 16 hours at room temperature. The course of the reaction was monitored by TLC (30:1, DCM:MeOH). The resulting crude product was purified using column chromatography with a mobile phase of 15:1 DCM:MeOH to afford compound **12** (0.131 g) as a white solid in 92% yield. ESI$^+$-MS/MS: *m/z* 374.0 $[M + H]^+$ → *m/z* 258.0 [M + H –

deoxyribose]$^+$, Exact mass calc: 374.1710 [M + H]$^+$, Exact mass obs: 374.1704 [M + H]$^+$ (error 1.6 ppm); ^{1}H NMR (MeOD/DMSO-*d6*, 500 MHz): δ 8.49 (d, 2H, *J* = 6.0 Hz), 7.38 (s, 1H), 6.82 (t, 1H, *J* = 7.3 Hz), 6.73 (t, 2H, *J* = 7.5 Hz), 5.53 (dd, 1H, *J* = 6.5, 6.5 Hz), 3.64 (m, 1H), 3.19 (ddd, 1H, *J* = 3.0, 3.0, 3.0 Hz), 3.03 (dd, 1H, *J* = 11.7, 3.3 Hz), 2.96 (dd, 1H, *J* = 12.0, 3.5 Hz), 1.77 (m, 2H), 1.55 (m, 1H), 1.49 (m, 1H), 0.92 (m, 2H), 0.24 (t, 3H, *J* = 7.5 Hz). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 180.3, 160.8, 149.3, 140.2, 138.6, 133.7, 130.8, 129.5, 116.0, 89.3, 86.9, 71.8, 62.5, 41.7, 30.5, 22.8, 14.4.

4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-propyl-2'-deoxycytidine (13): Compound **14** (50 mg, 0.13 mmol) was dried under vacuum overnight and co-evaporated with anhydrous pyridine (3 x 5 mL) and this flask was kept under vacuum overnight. The flask was immediately capped and purged with Ar and transferred to a glove bag filled with Ar. Anhydrous pyridine (kept over activated 4Å molecular sieves) was degassed with argon for 20 min. To the solution in the glove bag, 3 mL of anhydrous, degassed pyridine solution and dimethoxytrityl chloride (DMTrCl) (0.091 g, 0.27 mmol) were added sequentially. Upon addition of all reagents, the flask was taken out of the glove bag and the reaction was stirred for 16 hours at room temperature away from light. The progress of the reaction was monitored by TLC. After completion of the reaction, the crude sample was quickly dried under high vacuum and 30 mL of DCM was added. The organic layer was washed three times with 25 mL each of ice cold H_2O, dried using anhydrous Na_2SO_4, and solvent removed under reduced pressure. The resulting solid was purified on a short silica-flash column packed in 95:5 Hexane:TEA using the solvent gradient: 95:5 Hexane:TEA, 5:25:70 TEA:EtOAc:Hexane, 5:47.5:47.5 TEA:EtOAc:Hexane, 95:5 EtOAc:TEA. Upon evaporation of solvent, the resulting oily product was dissolved in 1 mL of anhydrous DCM and allowed to dry under high vacuum overnight, to get compound **13** (0.075 g) as a white foam in 77% yield. ESI$^+$-MS/MS: *m/z* 676.1 [M + H]$^+$ → *m/z* 303.1 [dimethoxytrityl cation]$^+$, 258.1 [M + H – deoxyribose]$^+$, Exact mass calc: 698.2837 [M + Na]$^+$, Exact mass obs: 698.2833 [M + Na]$^+$ (error 0.5 ppm); ^{1}H NMR (ACN-*d*4, 400 MHz): δ 8.27 (dd, 2H, *J* = 6.8, 1.6 Hz), 7.69 (s, 1H), 7.59 (m, 1H), 7.49 (m, 4H), 7.34 (m, 7H), 6.90 (m, 4H), 6.27 (dd, 1H, *J* = 6.6, 6.6 Hz), 4.50 (ddd, 1H, *J* = 5.2, 5.2, 5.2 Hz), 4.00 (ddd, 1H, *J* = 3.6, 3.6, 3.6 Hz), 3.79 (s, 6H), 3.33 (d, 2H, *J* = 3.6 Hz), 2.35 (t, 2H, *J* = 6.4 Hz), 2.16 (m, 2H), 1.47 (m, 2H), 0.80 (t, 3H, *J* = 7.4 Hz).

4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-propyl-2'-deoxycytidine-3'-[(2-cyanoethyl)-(*N,N*-diisopropyl)]-phosphoramidite (14): Into a flame dried round bottom flask was added 4-*N*-Benzoyl-5-*O*'-(dimethoxytrityl)-5-propyl-2'-deoxycytidine (**13**, 0.050 g, 0.074 mmol) that was dried under high vacuum-pump overnight, in an Ar filled glove bag. Anhydrous DCM (kept over activated 4 Å molecular sieves) was degassed with Ar for 20 min. All additions were performed in an Ar glove bag in the following order: anhydrous diisopropylethylamine (129 µL, 0.074 mmol), anhydrous DCM (2.5 mL) and 2-Cyanoethyl *N,N*-diisopropylchlorophosphoramidite (33 µL,

0.148 mmol). The flask was capped and the reaction mixture stirred for 15 hours under an argon atmosphere at room temperature away from light. After 15 h, 250 µL of anhydrous MeOH was added to prevent the oxidation of phosphoramidite and 200 µL of Et_3N was added to prevent hydrolysis during work-up. To this solution was added 30 mL of EtOAc and the organic layer extracted three times with ice-cold brine, dried over anhydrous Na_2SO_4 and dried under reduced pressure. The resulting solid was dissolved in a 5 mL of 5:95 TEA:EtOAc and purified on a short silica-flash column using the gradient elution described below: 95:5 hexane:TEA, 5:25:70 TEA:EtOAc:hexane, 5:47.5:47.5 TEA:EtOAc:hexane, 95:5 EtOAc:TEA. Evaporation of solvent resulted in a white foam (0.049 g) in 76% yield. ^{1}H NMR (Acetone-d_6, 400 MHz): δ 8.26 (d, 2H, *J* = 7.2 Hz), 7.82 (s, 1H), 7.57-7.44 (m, 5H), 7.39-7.23 (m, 7H), 6.91 (m, 4H), 6.35 (m, 1H), 4.77 (m, 1H), 4.21 (m, 1H), 3.90-3.80 (m, 1H), 3.77 (s, 6H), 3.75-3.58 (m, 3H), 3.47 (m, 2H), 2.76 (m, 2H), 2.62 (t, 2H, *J* = 6.0 Hz), 2.56 (m, 1H), 2.26 (m, 1H), 1.48 (m, 2H), 1.19-1.07 (m, 12H), 0.76 (t, 3H, *J* = 7.2 Hz). ^{31}P NMR (ACN-*d*4, 162 MHz): δ 148.37, 148.21.

3',5'-*O*-*t*-butyldimethylsilyl-5-vinyl-2'-deoxycytidine (15): Compound **1** (2.5 g, 4.3 mmol) was dried under high vacuum for 2 days and then coevaporated three times with 30 mL each anhydrous pyridine, and then dried under high vacuum for another two days. The flask was taken into an Ar-glove bag and to the nucleoside was added 75 mL of anhydrous deoxygenated 1-methyl-2-pyrolidinone. The resulting solution was again degassed by bubbling argon for 2 hours. The remaining additions were done inside an Ar-glove bag. To the nucleoside was added tri(2-furanyl)phosphine (0.55 g, 2.35 mmol), $Pd_2(dba)_3$ (0.504 g, 0.55 mmol) and tributylvinyltin (1.9 mL, 6.5 mmol) sequentially. The flask was sealed and the reaction mixture was stirred, away from light for 1 day at 60 °C. The reaction was monitored by TLC with 16:1 DCM:IPA and 9:1 EtOAC:hexane. Since a little starting material remained, another portion of tributylvinyltin (0.8 mL, 3.25 mmol) was added and the reaction mixture was stirred for another 12 hours. The crude reaction containing the product was divided to five portions and each portion diluted with 100 mL of EtOAC and washed four times with 25 mL each H_2O. The combined organic layers were dried with Na_2SO_4 and concentrated under reduced pressure. The resulting crude product was dissolved in 200 mL acetonitrile and washed five times with 25 mL each of hexane. The ACN layer was partitioned over hexane and the ACN was dried to afford crude compound **16**. The product was purified on a silica column using a gradient of DCM to 15:85 IPA:DCM to obtain compound **16** (1.2 g) in 58% yield. ESI$^+$-MS: Exact mass calc: 482.2865 $[M + H]^+$, Exact mass obs: 482.2850 $[M + H]^+$ (error 3.1 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.72 (s, 1H), 6.32 (dd, 1H, *J* = 17.0, 11.0 Hz), 6.26 (dd, 1H, *J* = 6.3, 6.3 Hz), 5.40 (d, 1H, *J* = 17.5 Hz), 5.24 (d, 1H, *J* = 11.0 Hz), 4.33 (d, 1H, *J* = 2.5 Hz), 3.94 (s, 1H), 3.83 (d, 1H, *J* = 11.0 Hz), 3.73 (d,

1H, *J* = 11.5 Hz), 2.42 (m, 1H), 1.94 (m, 1H), 0.86 (s, 18H), 0.04 (d, 6H, *J* = 4.5 Hz), 0.03 (s, 6H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 164.1, 155.3, 138.2, 128.4, 117.2, 105.8, 88.0, 86.4, 72.1, 63.0, 42.4 26.0, 25.8, 18.5, 18.1, -4.5, -4.8, -5.3, -5.4.

4-*N*-Benzoyl-3',5'-*O*-(*t*-butyldimethylsilyl)- 5-vinyl-2'-deoxycytidine (16): Compound 15 (0.81 g, 1.68 mmol) which was dried on high vacuum overnight then co-evaporated anhydrous pyridine (3 x 5mL) and dissolved in 20 mL of anhydrous pyridine and the flask was capped with a rubber septum under argon atmosphere. To the stirring solution was added benzoyl chloride (234 µL, 2.01 mmol) and stirred for overnight at room temperature. The reaction was checked by TLC (30:1 DCM:MeOH). Next morning, another 234 µL (2.01 mmol) of benzoyl chloride was added, and the reaction was stirred for 24 h. Solvents were evaporated under reduced pressure, and the residue was taken up in 30 mL of DCM, which was washed with water (3 x 30 mL) and dried over sodium sulfate for 30 minutes. The solution was filtered, evaporated under reduced pressure, and subjected to flash column. Purification was achieved using a gradient of DCM:MeOH to give compound **16** as a colorless solid (0.807 g, 82% yield). ESI$^+$-MS: Exact mass calc: 586.3127 $[M + H]^+$, Exact mass obs: 586.3137 $[M + H]^+$ (error 1.7 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 13.13 (s, 1H), 8.16-8.17 (d, 2H, *J* = 7.5 Hz), 7.99 (S, 1H), 7.58-7.61 (m, 1H), 7.49-7.52 (m, 2H), 6.77-6.80 (m, 1H), 6.13 (t, 1H), 5.93-5.96 (d, 1H, *J* = 6.5 Hz), 5.32-5.34 (d, 1H, *J* = 17.5 Hz), 4.39-4.40 (m, 1H), 3.80-3.90 (m, 1H), 3.81-3.84 (m, 1H), 3.74-3.77 (m, 1H), 2.36-2.42 (m, 1H), 2.21-2.27 (m, 1H), 0.87-0.88 (2s, 20H), -0.12--0.10 (2s, 12H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 141.32, 139.24, 135.35, 132.04, 131.10, 130.88, 118.07, 114.08, 90.21, 88.65, 74.57, 65.26, 28.49, 28.35, 20.75, 20.48, 20.39, -0.50, -2.07, -2.24, -2.74.

4-*N*-Benzoyl-5-vinyl-2'-deoxycytidine (17): Compound **16** (0.81 g, 1.38 mmol) was dried on high vacuum overnight then co-evaporated three times with 5 mL of anhydrous pyridine. To this was added 15 mL anhydrous DCM. The solution and flask were purged with argon for 5 minutes. To the flask 900 µL (5.5 mmol) triethylamine trihydrofluoride was added over 30 minutes. The reaction was stirred in the dark for 16 hours at room temperature. After completion of the reaction as monitored by the TLC, the reaction mixture was evaporated under reduced pressure and the residue was extracted with DCM and dried over anhydrous Na_2SO_4 and purified on a silica gel column run washed with 15:1 DCM:MeOH and eluted with a mixture of 9:1 DCM:MeOH. Yield (0.64 g, 80%). ESI$^+$-MS: Exact mass calc: 358.1397 $[M + H]^+$, Exact mass obs: 358.1398 $[M + H]^+$ (error 0.3 ppm); ^{1}H NMR ($CDCl_3$, 500 MHz): δ 8.64 (s, 1H), 8.16-8.18 (d, 2H, *J* = 7.0 Hz), 7.59 (t, 1H), 7.52 (t, 2H), 6.77-6.82 (m, 1H), 6.17 (t, 1H), 5.84-5.88 (d, 1H, *J* = 18 Hz), 5.26-5.31 (m, 3H), 4.30-4.32 (m, 1H), 3.87-3.88 (m, 1H), 3.71-3.75 (m, 1H), 3.62-3.66 (m, 1H), 2.27 (t, 2H). ^{13}C NMR ($CDCl_3$, 125 MHz): δ 141.32, 139.24, 135.35, 132.04, 131.10, 130.88, 118.07, 114.08, 90.21, 88.65, 74.57, 65.26, 28.49, 28.35.

4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-vinyl- 2'deoxycytidine (18). Compound **17** (155 mg, 0.43 mmol) was dried by repeated co-evaporation with dry pyridine (3 × 5 mL). The residue was dissolved in dry pyridine (6 mL) and 4,4'-dimethoxytrityl chloride (220 mg, 0.64 mmol) was added in portions and stirred at rt for 7h. After completion of the reaction (monitored by TLC), methanol (1 mL) was added to the reaction mixture, and stirring was continued for another half an hour. The reaction mixture was evaporated to dryness under reduced pressure and the remaining residue was dissolved in dichloromethane (25 mL) and washed with 5% aq. $NaHCO_3$ solution (50 mL) and water (50 mL). The organic layer was dried over Na_2SO_4, the solvent was evaporated under reduced pressure, and the residue was subjected to flash chromatography (silica gel, column 10 × 4 cm, eluted with CH_2Cl_2/acetone 95:5 to 85:15). Evaporation of the fractions corresponding to the product on TLC afforded **18** (0.167 g, 60%) as a pale yellow foam. TLC (silica gel, CH_2Cl_2/MeOH 9:1). R_f 0.8. Exact mass calc: 660.2704 $[M + H]^+$, Exact mass obs: 660.2688 $[M + H]^+$ (error 2.4 ppm). ^{1}H- NMR [DMSO (d_6), 500 MHz]: δ 2.26-2.29 (m, 1H), 2.37-2.40 (m, 2H), 3.25-2.52 (d, 1H), 3.72 (s, 3H), 3.73 (s, 3H), 3.98-3.99 (d, 1H), 4.28 (s, 1H), 5.08-5.10 (d, 1H), 5.37-5.38 (d, 1H), 5.70-5.74 (d, 1H), 6.19 (t, 1H), 6.44-6.50 (m, 1H), 6.87-6.90 (m, 4H), 7.21-7.27 (m, 5H), 7.31 (t, 2H), 7.38-7.40 (d, 2H), 7.51 (t, 2H), 7.60 (t, 1H), 7.96 (brs, 1H), 8.16-8.17 (d, 2H). ^{13}C- NMR [DMSO (d_6), 500 MHz]: δ 173.75, 158.61, 158. 60, 157.70, 147.25, 145.13, 139.24, 137.03, 135.86, 135.83, 133.11, 130.20, 130.17, 129.81, 128.87, 128.37, 128.16, 127.24, 115.84, 113.70, 111.90, 86.26, 70.61, 64.07, 55.49, 55.48, 31.15.

4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-vinyl-2'deoxycytidine 3'-(2-cyanoethyl)-*N, N'*-diisopropyl phosphoramidite (19). A solution of **18** (100 mg, 0.15 mmol) in dry CH_2Cl_2 (10 mL) was stirred with (i-Pr)$_2$NEt (52 µL, 39 mg, 0.30 mmol) at rt. Then, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (67 µL, 72 mg, 0.30 mmol) was added and the reaction mixture was stirred for 30 min. Upon completion of the reaction (monitored by TLC), the reaction mixture was diluted with CH_2Cl_2 (30 mL) and was poured into 5% $NaHCO_3$ solution (30 mL) and extracted with CH_2Cl_2 (3 × 20 mL). The combined organic phases were dried over Na_2SO_4, and the solvent was evaporated. The residual foam was applied to a flash chromatography column (silica gel, 8 × 3 cm, eluted with CH_2Cl_2/acetone 100:0 → 95:5). Evaporation of the fractions corresponding to the product on TLC afforded **19** (90 mg, 70%) as a colorless viscous oil. TLC (silica gel, CH_2Cl_2/acetone 9:1). R_f 0.7. ESI$^+$-MS/MS: Exact mass calc: 860.3783 $[M + H]^+$, Exact mass obs: 860.3771 $[M + H]^+$ (error 1.4 ppm). ^{31}P NMR ($CDCl_3$, 500 MHz): 149.24, 148.64.

$^{13}C_{10}{}^{15}N_2$-5-Methyl-2'-deoxycytidine (20): $^{13}C_{10}{}^{15}N_2$-Thymidine (3.5 mg, 0.014 mmol, Cambridge Isotope Labs) was dissolved in 0.7 mL dichloromethane. To the stirring solution, 360 µL triethylamine (TEA) (2.6 mmol) and 270 µL chlorotrimethyl-silane (2.2 mmol) were added. The solution was stirred for 30 min at 23 °C and then quenched with 2

mL cold 1 M sodium bicarbonate ($NaHCO_3$). Phases were separated, the organic layer dried with anhydrous magnesium sulfate ($MgSO_4$), filtered, and evaporated under reduced pressure. The resulting brown residue was co-evaporated with anhydrous toluene (3 x 1 mL). At the same time, a suspension of 1,2,4-triazole (120 mg, 1.7 mmol), TEA (370 µL, 2.7 mmol), and phosphorous oxychloride ($POCl_3$) (37 µL, 0.40 mmol) in 2 mL acetonitrile (ACN) was stirred on ice for 30 min. The residue was then dissolved in 3 mL ACN, added to the suspension of triazole, TEA, and $POCl_3$, and stirred for 2 h at room temperature. The reaction mixture was concentrated under reduced pressure and re-dissolved in 2 mL of DCM. The DCM solution was washed with 1 M cold $NaHCO_3$, separated, dried with $MgSO_4$, filtered, and the solvent removed under vacuum. The residue was dissolved in 5.5 mL of 7 N methanolic ammonia and placed with a magnetic stir bar in a Teflon tube, which had been chilled on ice for 30 min. The Teflon tube was capped and placed in a stainless steel bomb that was then sealed. The bomb was placed in an oil bath preheated to 75 °C and allowed to react for 3 days. After cooling the stainless steel bomb to below room temperature, the bomb was opened, and the remaining ammonia was allowed to evaporate overnight. The residue was dissolved in 50:50 MeOH:DCM and dried prior to reverse phase HPLC purification on an Agilent 1100 system. Purification was done using a Synergi 4u Hydro-RP 80A column [250 mm x 4.6 mm] from Phenomenex (Torrance, CA). The starting material and product were eluted isocratically using 6% MeOH in 10 mM Ammonium formate at pH 4.2 for 32 min followed by an increase to 50% ACN and subsequent equilibration to initial conditions. The product eluted at 12.5 min and starting material at 20.9 min. The desired product was desalted using Carbograph SPE cartridges (Grace Analytical, Deerfield, IL) conditioned in MeOH and then water, the product loaded onto the cartridge, followed by a water wash and 3 successive methanol elutions. Product identity was confirmed by matching the UV spectra to a commercial unlabeled standard and HPLC-ESI$^+$-MS/MS. ESI$^+$-MS/MS: *m/z* 253.9 $[M + H]^+ \rightarrow$ *m/z* 133.2 $[M + H - \text{deoxyribose}]^+$ Exact mass calc: 254.1412 $[M + H]^+$, Exact mass obs: 254.1408 $[M + H]^+$ (error 1.6 ppm).

***Solid Phase Synthesis and Purification of Oligodeoxynucleotides*:** DNA sequences used in this study were derived from codons 153-158 of the human *p53* tumor suppressor gene (5′-CCCGGCACC CGC[$^{15}N_3$,$^{13}C_1$-G]TCCGCG-3′). Oligodeoxynucleotides containing a structurally modified C:G base pair at the first position of *p53* codon 157 were synthesized using an ABI 394 DNA Synthesizer (Applied Biosystems, California) in accordance to standard solid phase oligodeoxynucleotide synthesis protocols.[1]

The oligodeoxynucleotides were purified by reversed phase HPLC using an Agilent 1100 system interfaced with a UV-VWD detector[2, 3]. The oligodeoxynucleotides containing modified cytosine base were purified on a Supelcosil LC-18-DB column (10 mm × 250 mm, 5 µm, Supelco, Bellefonte, PA) eluted at 40°C and a flow rate of 3

mL/min. HPLC buffers were 100 mM triethylammonium acetate, pH 7.0 (A), and acetonitrile (B). A linear gradient of 16.8-26% B in 21 minutes followed by a gradient increase to 34% B in the next 19 minutes was used. Following HPLC purification, the oligodeoxynucleotides were desalted by NAP-5-illustra™ size exclusion columns (GE Healthcare, Piscataway, NJ) according to the manufacturer's protocols. The presence of C5-alkyl-dC in the oligodeoxynucleotides was confirmed by capillary HPLC-ESI$^-$-MS (Table S1),[3-5] and their concentrations were determined from the extent of 2'-deoxyguanosine in enzymatic digests using previously published protocols.[6-8]

Double stranded DNA was obtained by combining equal amounts of the complementary strands (200 pmol) in 10 mM Tris-HCl (pH 8.0), 50 mM sodium chloride buffer to achieve a concentration of 200 μM. The DNA mixtures were heated at 90°C for 5 mins and then allowed to cool slowly to room temperature. Each duplex contained a $^{15}N_3$,$^{13}C_1$-labeled guanine at the first position of codon 157, base paired to cytosine or a C5-alkylcytosine analog (Table S1).[9]

DNA Methyltransferase Experiments: Synthetic DNA duplexes (250 -1500 fmol) were incubated with human DNMT1 enzyme (0.75U, Promega Corp.), 100 mg/ml BSA, and 160 μM S-adenosylmethionine (SAM) in 1X DNMT buffer for 15 min at 37 °C. The reactions were quenched by freezing on dry ice. The enzyme was inactivated by heating at 65 ˚C (50 min). DNA was digested to 2'-deoxynucleosides in the presence of phosphodiesterase I (PDE I, 86 mU), phosphodiesterase II (PDE II, 77 mU), alkaline phosphatase (29 U), and DNase I (34 U) in 10 mM Tris-HCl/15 mM $MgCl_2$ buffer, pH 7 (35 μL total volume, 18 h at 37 °C). Samples were spiked with 2 pmol of $^{13}C_{10}{}^{15}N_2$-^{Me}dC (internal standard) and purified by SPE using Thermo Fisher Scientific Hypersep Hypercarb (50 mg/mL) cartridges. Briefly, the cartridges were equilibrated with 2 mL MeOH and 2 mL deionized water. The volume of samples was adjusted to 500 μL prior to loading onto pre-equilibrated columns. The samples were washed with 2 mL each of deionized water and eluted in 1.6 mL of MeOH.

HPLC-ESI$^+$-MS/MS of 5-methyl-dC

The amounts of ^{Me}dC present in DNA before and after incubation with DNMT1 were determined by HPLC ESI$^+$-MS/MS utilizing a Thermo TSQ Vantage mass spectrometer interfaced with a Thermo Dionex Ultimate 3000 HPLC by monitoring the transitions for 5-medC and the corresponding internal standard. The mass spectrometer settings were optimized by direct infusion of standards at the same flow rate and composition of the eluting peaks. The mass source was operated at 2700 V, sheath gas 27, declustering voltage 22 V, capillary temperature 270 ºC, S-lens 93, collision gas 1.0 mTorr, and collision energy 8 V. The following transitions were utilized: *m/z* 242.1 $[M + H]^+ \rightarrow$ *m/z*

126.1 $[M - \text{deoxyribose} + H]^+$ for ^{Me}dC and *m/z* 254.1 $[M + H]^+ \rightarrow$ *m/z* 133.2 $[M - \text{deoxyribose} + H]^+$ for the corresponding internal standard $^{13}C_{10}{}^{15}N_2$-^{Me}dC. Chromatographic separation was achieved using a Thermo Fisher Hypercarb column (0.3 mm x 100 mm, 3 μm) eluted at a flow rate of 14 μL/min at ambient temperature. HPLC solvents were 15 mM NH_4OAc (A) and ACN (B). A linear gradient of 15% to 60% B in 10 min was used, followed by an increase to 95% in 2 min and held at 95% B for 3 min. The column was returned to initial conditions over 2 minutes followed by equilibration at 15% B for 7 min. Under these conditions, both the ^{Me}dC and the $^{13}C_{10}{}^{15}N_2$-^{Me}dC internal standard eluted at 6.1 min.

^{Me}dC amounts introduced during DNMT1 reaction was calculated by comparing the HPLC-ESI-MS/MS peak areas corresponding to ^{Me}dC and $^{13}C_{10}{}^{15}N_2$-^{Me}dC internal standard. Values were adjusted by subtracting ^{Me}dC amount generated in the absence of the enzyme. The amount of DNMT1 mediated formation of ^{Me}dC was used to calculate the velocity of methyl transfer. Plots of methylation velocities vs. DNA concentrations were plotted for each DNA sequence (substrate) and fitted to a Michaelis-Menten equation below:

$$Y = V_{max} * X / (K_m + X),$$

where the V_{max} is the maximal methylation rate of DNA by DNMT1, and K_m is the concentration of substrate required to reach the half-maximal rate. Statistical differences in V_{max} were calculated in GraphPad Prism using ANOVA using Tukey correction for multiple testing.

Homology Modeling

All molecular modeling was performed using the Schrödinger modeling suite package.[10] Since both the mouse and human DNMT1 shared an 85% sequence similarity[11], homology modeling of the human DNMT1 (hDNMT1) was carried out using the published crystal structure of the mouse DNMT1 in complex with hemi-methylated DNA (PDB: 4DA4)[12] based on its reference sequence (NP_001124295.1). This method[13, 14] takes advantage of the observation that protein structure is more highly conserved than its amino acid sequence, and that small or medium changes in sequence normally result in little variation in the 3D structure.[15] The DNA sequence was modified accordingly to match the sequence used experimentally to determine DNA methylation rates. Schrödinger Maestro (Schrödinger, LLC, NY) was used to model the extended forms of 5-methylcytosine (5-ethyl-dC, 5-propyl-dC, 5-vinyl-dC) within the modeled DNA template. Each of the modeled hDNMT1 – DNA complexes was subjected to standard protein preparation protocols at physiological pH, followed by energy minimization of the hydrogen atoms using OPLS3 force field with Generalized Born implicit solvent model to optimize all hydrogen-bonding networks.[16, 17] To understand

local structural effects induced by oxidation on 5-methylcytosine to $^{ethyl}C$, $^{propyl}C$, and $^{vinyl}C$, energy minimization of the modified base was carried out with a 15Å spherical radius restraint.

Table S1. Nucleobase sequences and molecular weights of oligodeoxynucleotides employed in the present study.

Oligodeoxynucleotide ID	Sequence (5′ → 3′)	MW	
		Calc.	Obs.*
(+)*p53* exon 5-[$^{15}N_3$,$^{13}C_1$-G]	CCCGGCACCCGC[$^{15}N_3$, $^{13}C_1$-G]TCCGCG	5715.7	5714.9
(-)*p53* exon 5	CGCGGA**C**GCGGGTGCCGGG	5911.9	5910.7
(-)*p53* exon 5-^{Me}C	CGCGGA**[Methyl-C]**GCGGGTGCCGGG	5925.9	5925.4
(-)*p53* exon 5-Et-C	CGCGGA**[Ethyl-C]**GCGGGTGCCGGG	5939.9	5939.5
(-)*p53* exon 5-prop-C	CGCGGA**[Propyl-C]**GCGGGTGCCGGG	5953.9	5954.2
(-)p53 exon 5-vinyl-C	CGCGGA**[Vinyl-C]**GCGGGTGCCGGG	5937.9	5939.0

***** from HPLC-ESI-MS

Table S2. Thermal melting points of structurally modified DNA duplexes. Double-stranded DNA were obtained by dissolving equimolar amounts of the complimentary strands in 10 mM sodium phosphate buffer, pH 7.0 containing 50 mM NaCl to give a 9.7 μM final DNA concentration. DNA melting temperatures were obtained on a Varian Cary 100 Bio UV-visible spectrophotometer by linearly increasing the temperature by 0.5 °C/min from 30 and 90 °C. The melting points (T_m) was determined using Cary WinUV Thermal software (Varian, Palo Alto, CA) by averaging N = 3 runs. Data for native, methyl, ethyl, and propyl strands are reproduced from[9].

Double stranded DNA	Sequence	T_m (ºC) [a]
ds Native	5′-CCCGGCACCCGC *GTCCGCG-3′ 3′-GGGCCGTGGGCG CAGGCGC-5′	80.7 ± 0.5
ds Methyl	5′-CCCGGCACCCGC *GTCCGCG-3′ 3′-GGGCCGTGGGCG ^{Me}CAGGCGC-5′	81.5 ± 0.2
ds Ethyl	5′-CCCGGCACCCGC *GTCCGCG-3′ 3′-GGGCCGTGGGCG ^{Ethyl}CAGGCGC-5′	82.1 ± 0.2
ds Propyl	5′-CCCGGCACCCGC *GTCCGCG-3′ 3′-GGGCCGTGGGCG ^{Propyl}CAGGCGC-5′	80.2 ± 1.1
ds Vinyl	5′-CCCGGCACCCGC *G]TCCGCG-3′ 3′-GGGCCGTGGGCG ^{Vinyl}CAGGCGC-5′	83.0 ± 0.8

[a] 9.7 μM dsDNA in 50 mM NaCl, 10 mM NaH_2PO_4, pH 7.0 (pH with NaOH). *Guanine is stable-isotopically labeled $^{13}C_1{}^{15}N_3$-dG

Table S3. Kinetic parameters of DNMT1 mediated methylation of DNA duplexes containing 5-alkyl-dC.[a]

DNA Duplex	V_{max} (x10^{-2} nM/min)	K_m (nM)	V_{max}/K_m (x10^{-2} min^{-1})
ds Native	0.72 ± 0.22	0.450 ± 5.99	1.6 ± 21
ds Methyl	9.6 ± 2.9	21.8 ± 13.7	0.44 ± 0.306
ds Ethyl	2.4 ± 0.59	3.90 ± 6.42	0.62 ± 1.66

[a] The V_{max} and K_m values were determined via nonlinear regression using data from three or more individual points. The ranges in V_{max} and K_m are the standard error for regression analysis. Error was propagated for V_{max}/K_m using the equation: $\frac{Dc}{c} = \sqrt{\left(\frac{Da}{a}\right)^2 + \left(\frac{Db}{b}\right)^2}$; where a, b, and c are V_{max}, K_m, and V_{max}/K_m respectively.

Figure S1. H-NMR of 5-(Trimethylsilyl)-ethynyl-3′,5′-*O*-*t*-butyldimethylsilyl-2′-deoxycytidine (**2**)

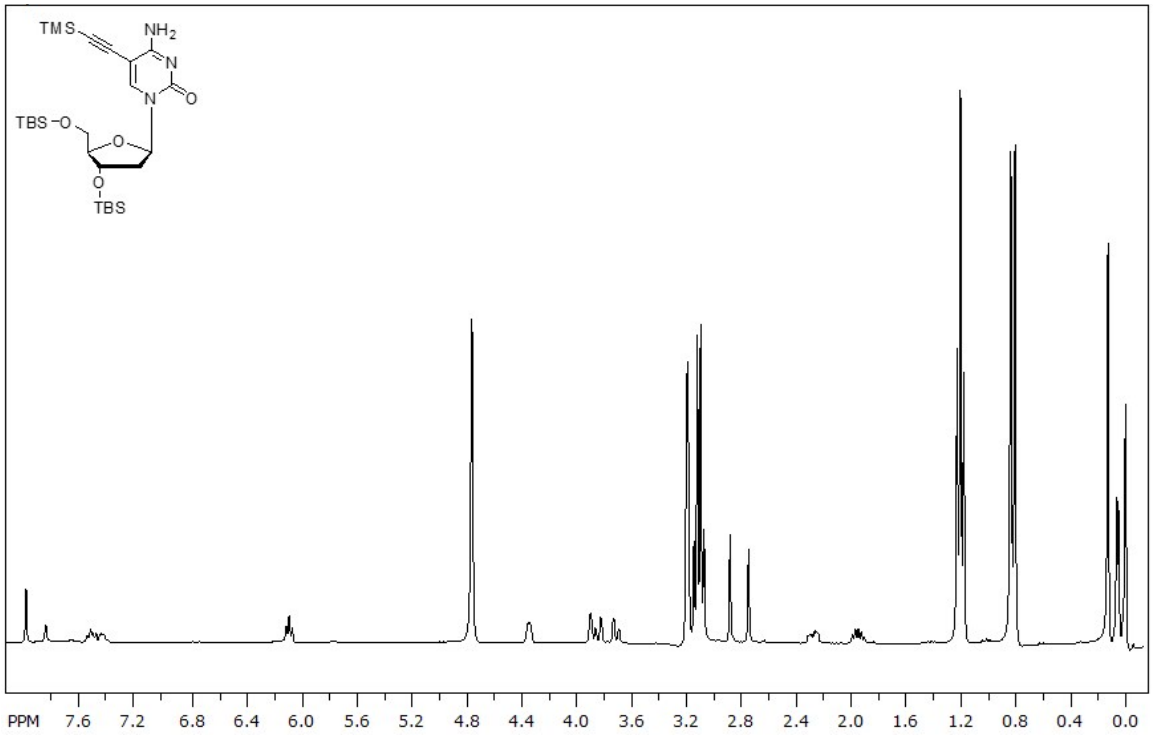


Figure S2. H-NMR of 5-Ethynyl-3′,5′-*O*-*t*-butyldimethylsilyl-2′-deoxycytidine (**3**)

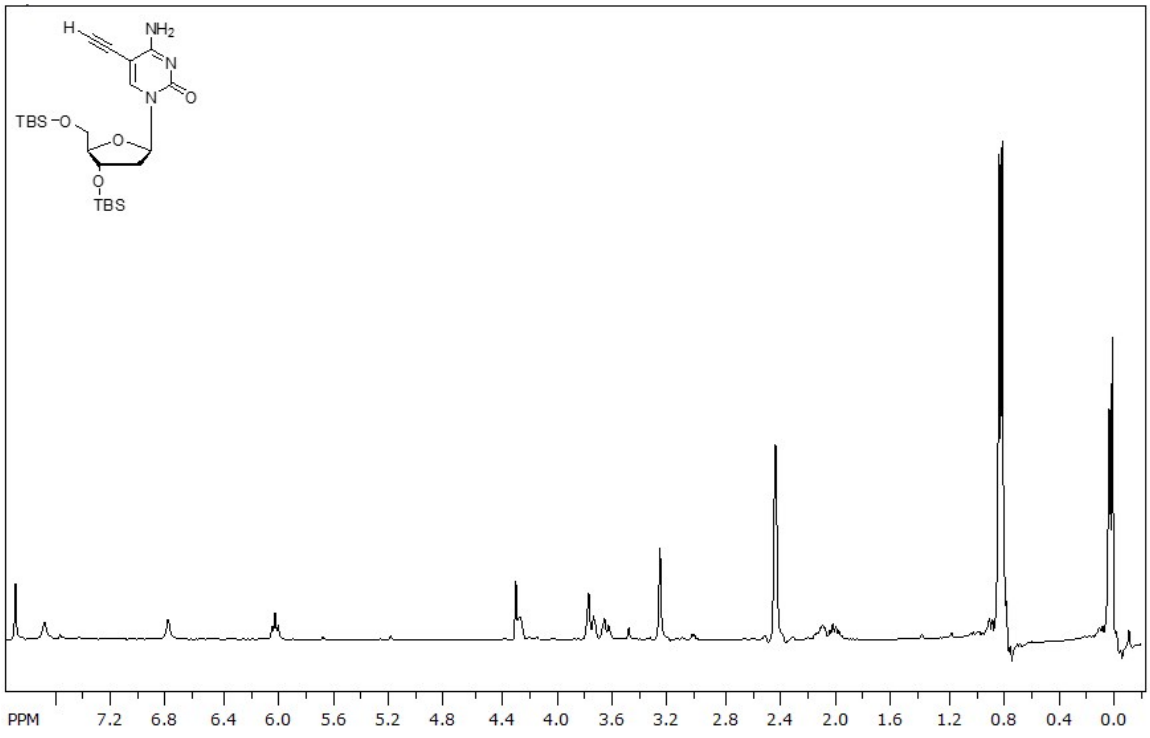


Figure S3. H-NMR of 5-Ethyl-3′,5′-*O*-*t*-butyldimethylsilyl-2′-deoxycytidine (**4**)

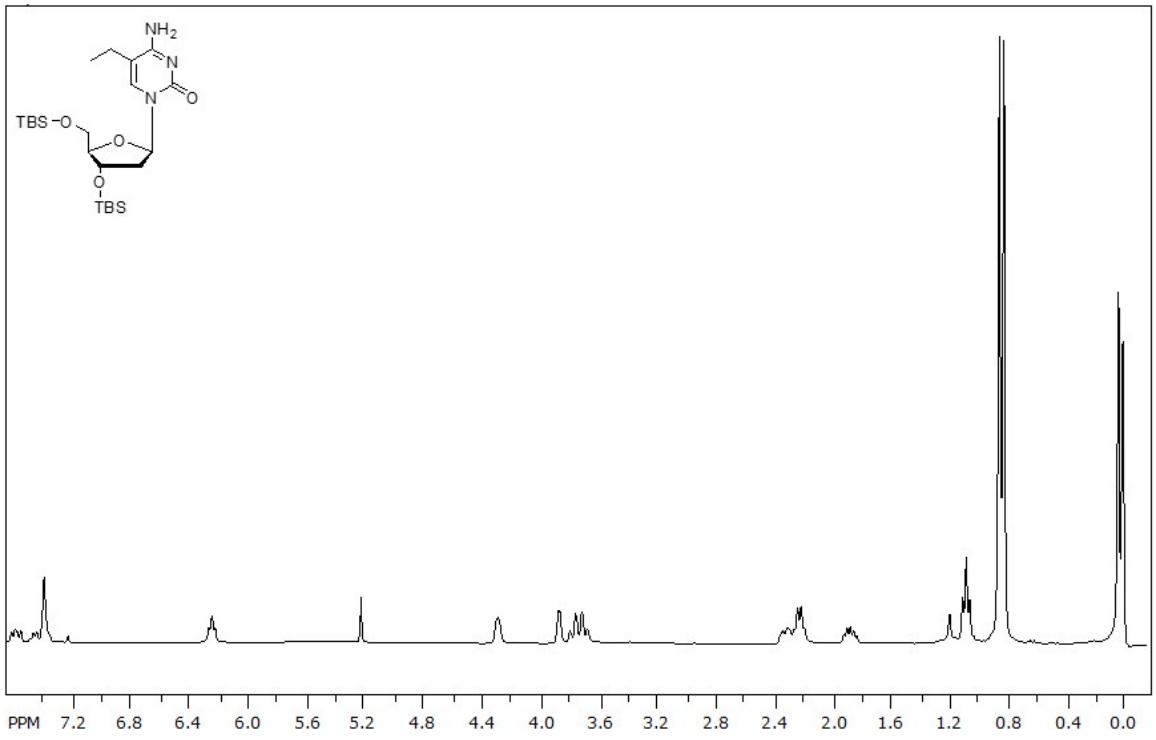


Figure S4. H-NMR of 4-*N*-Benzoyl-5-ethyl-3′,5′-*O*-*t*-butyldimethylsilyl-2′-deoxycytidine (**5**)

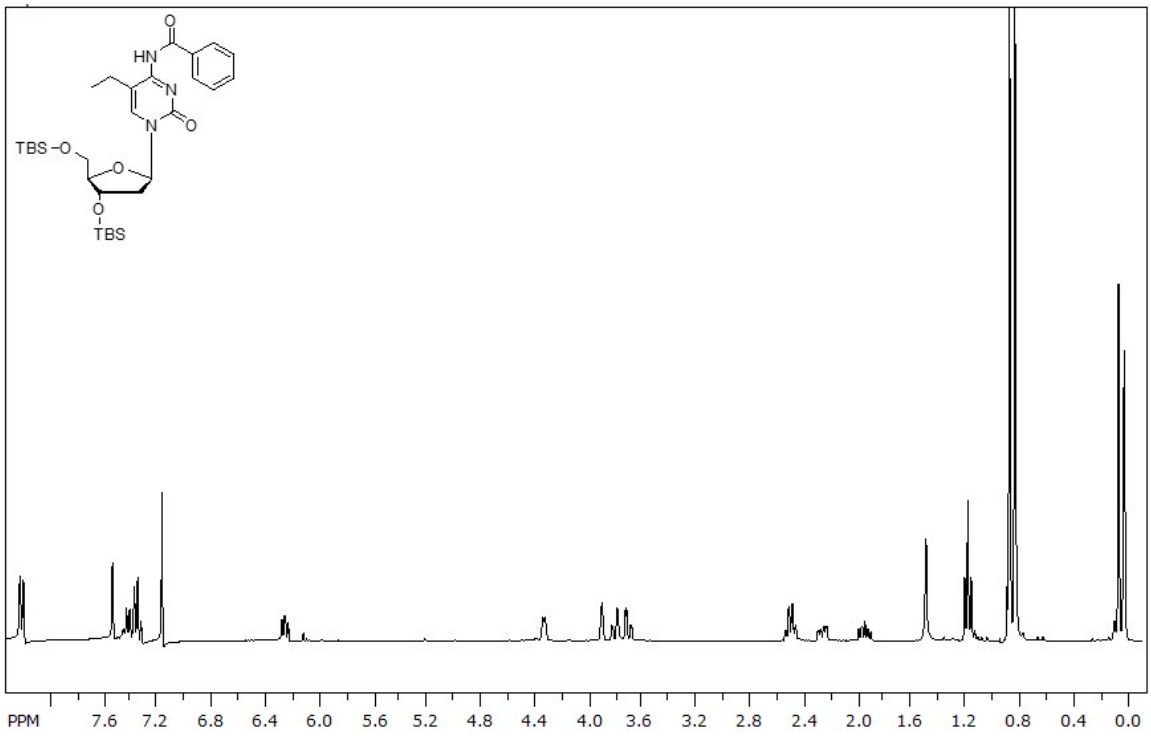


Figure S5. H-NMR of 4-*N*-Benzoyl-5-ethyl-2'-deoxycytidine (**6**)

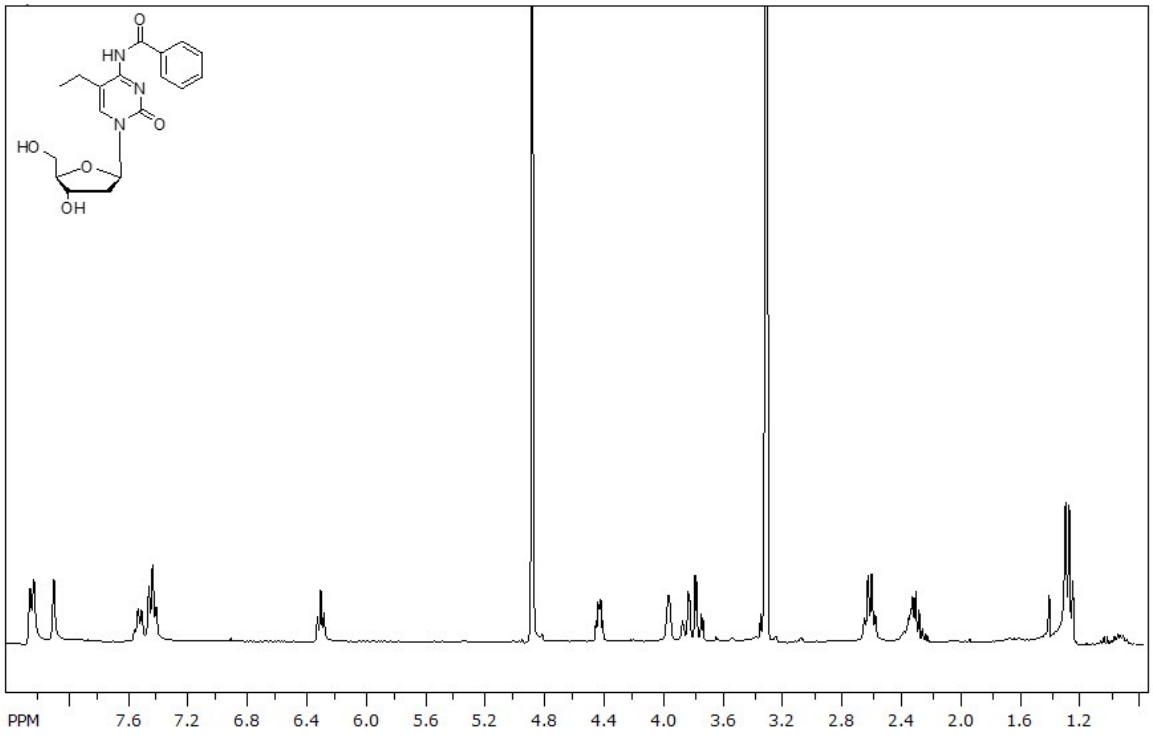


Figure S6. H-NMR of 4-*N*-Benzoyl-5′-*O*-(dimethoxytrityl)-5-ethyl-2′-deoxycytidine (**7**)

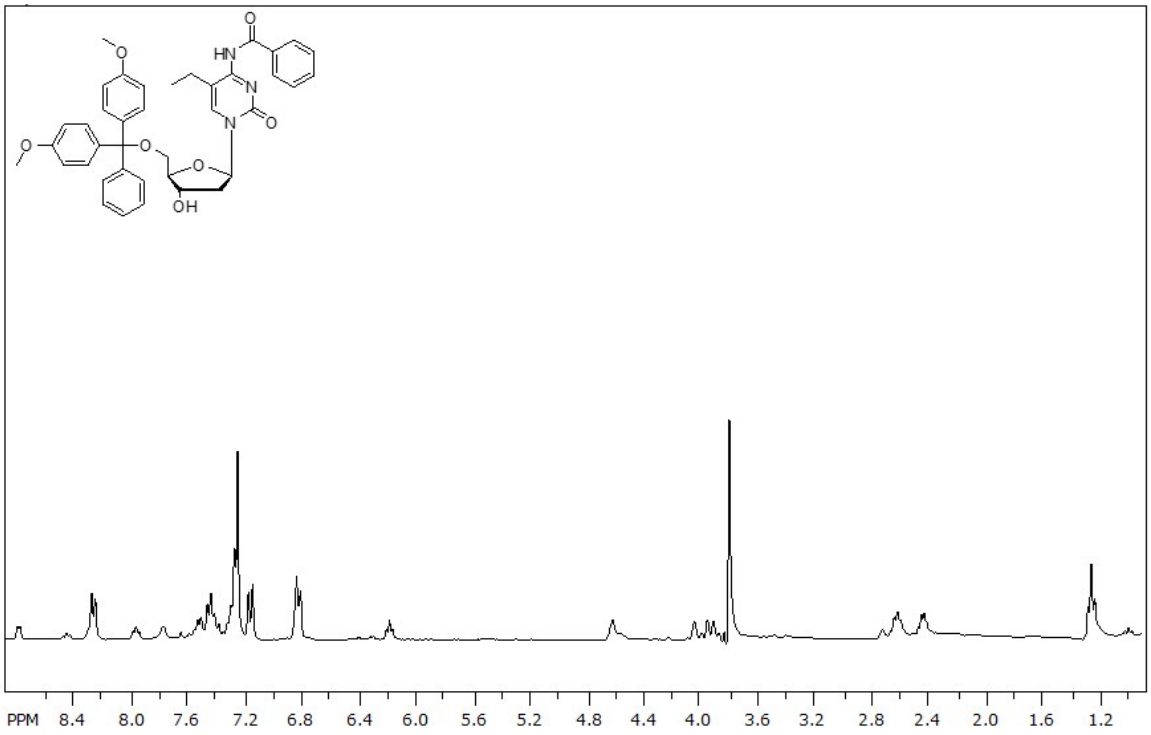


Figure S7. P-NMR of 4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-ethyl-2-deoxycytidine-3'-[(2-cyanoethyl)-(*N,N*-diisopropyl)]-phosphoramidite (**8**)

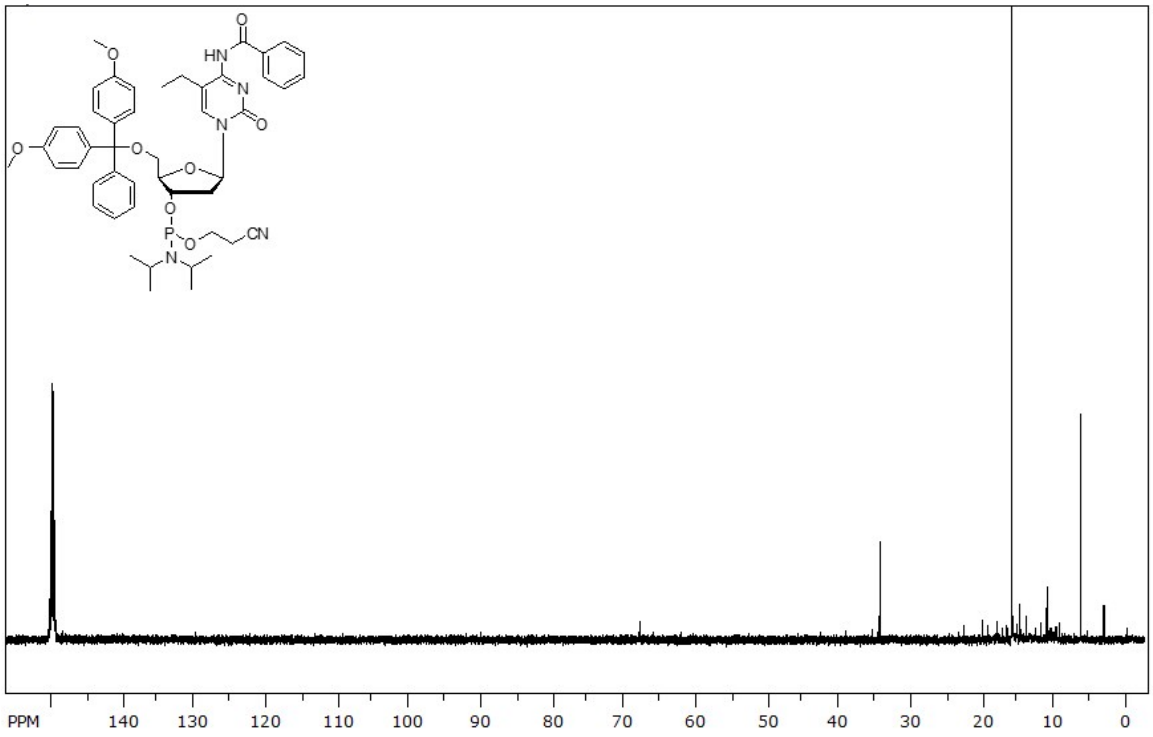


Figure S8. H-NMR of 5-Propyn-1-yl-3′,5′-*O*-*t*-butyldimethylsilyl-2′-deoxycytidine (**9**)

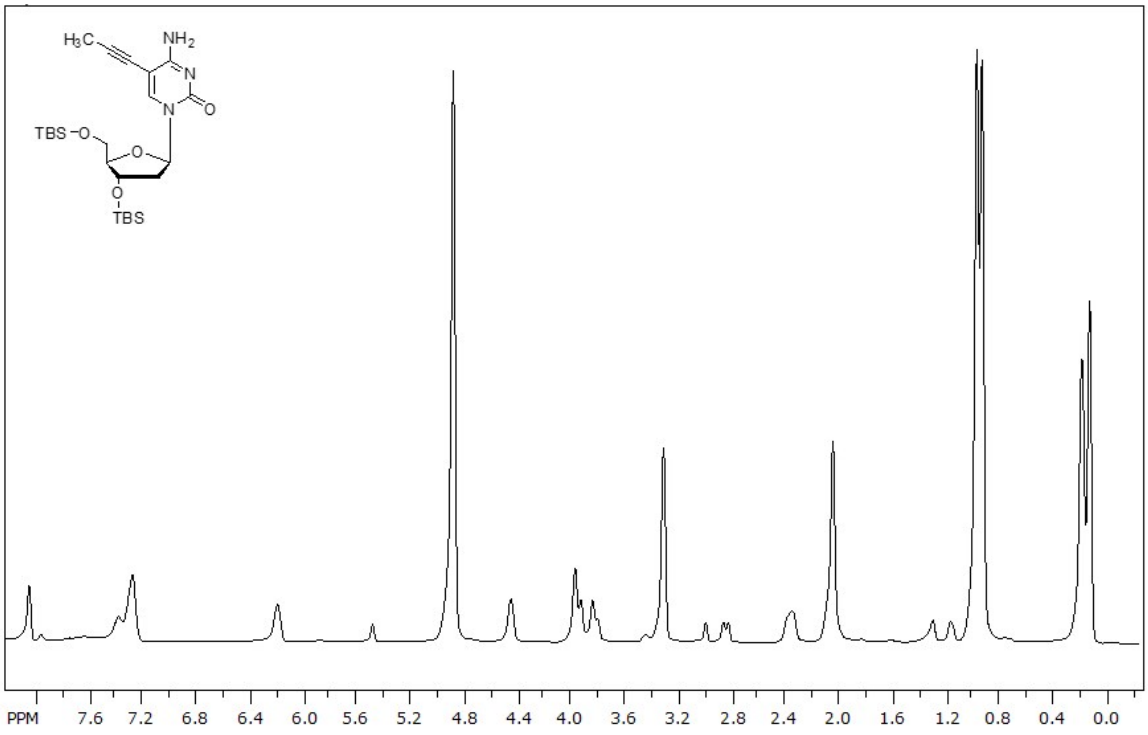


Figure S9. H-NMR and COSY of 5-Propyl-3',5'-*O*-*t*-butyldimethylsilyl-2'-deoxycytidine (**10**)

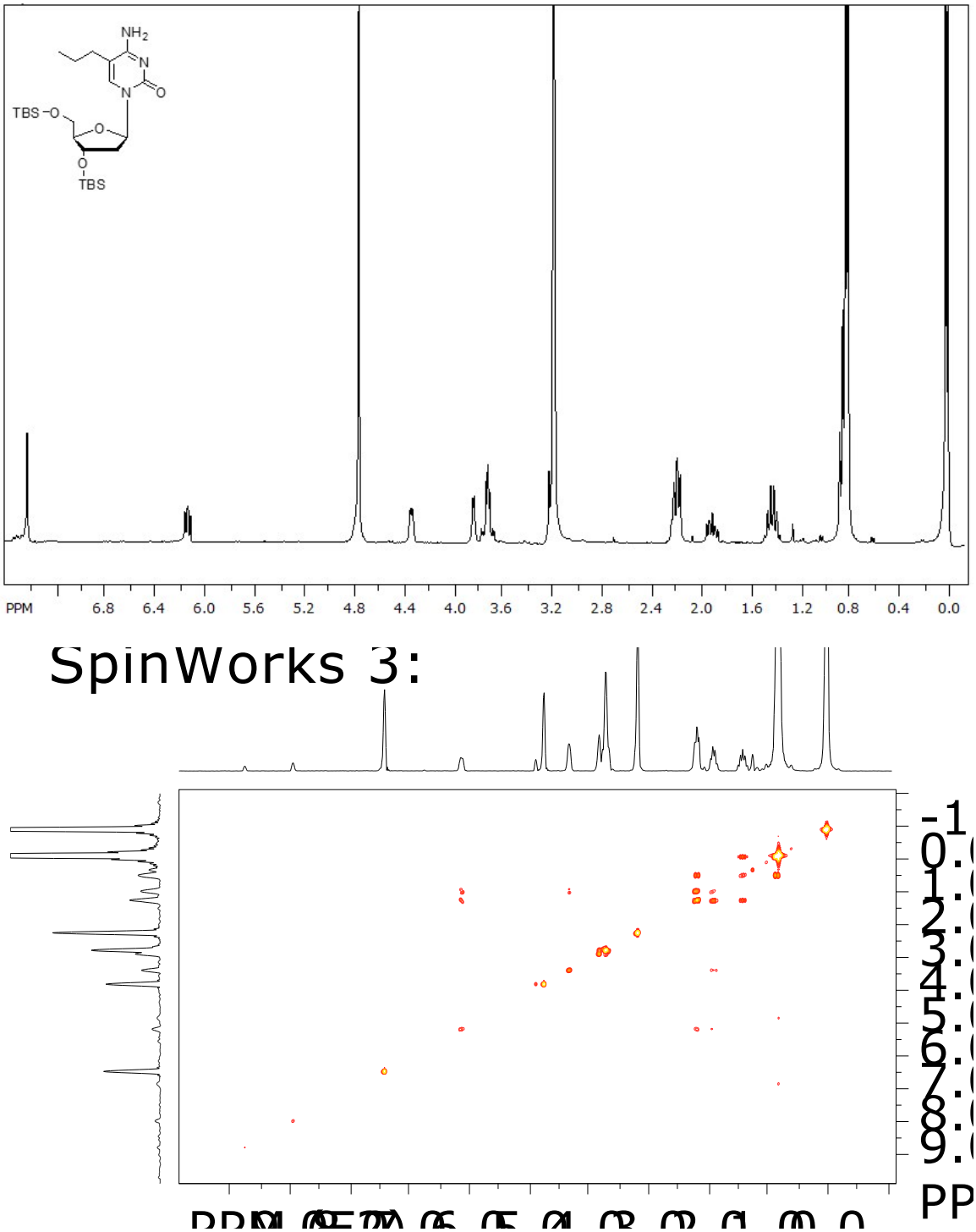


Figure S10. H-NMR of 4-*N*-Benzoyl-5-propyl-3′,5′-*O*-*t*-butyldimethylsilyl-2′-deoxycytidine (**11**)

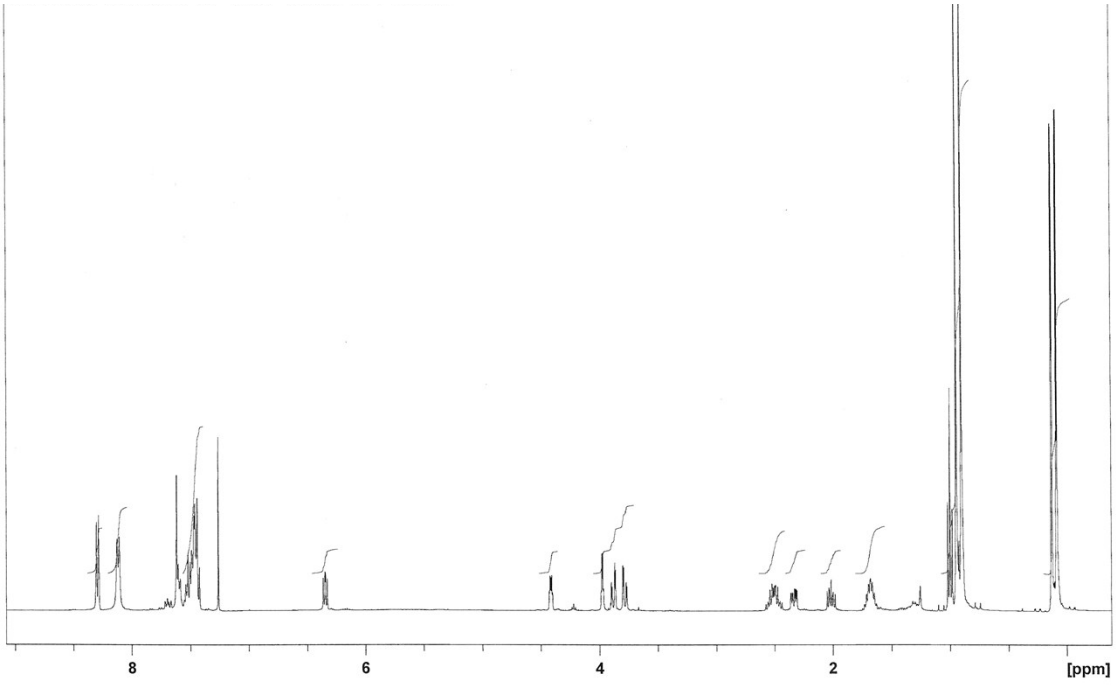


Figure S11. H-NMR of 4-*N*-Benzoyl-5-propyl-2'-deoxycytidine (**12**)

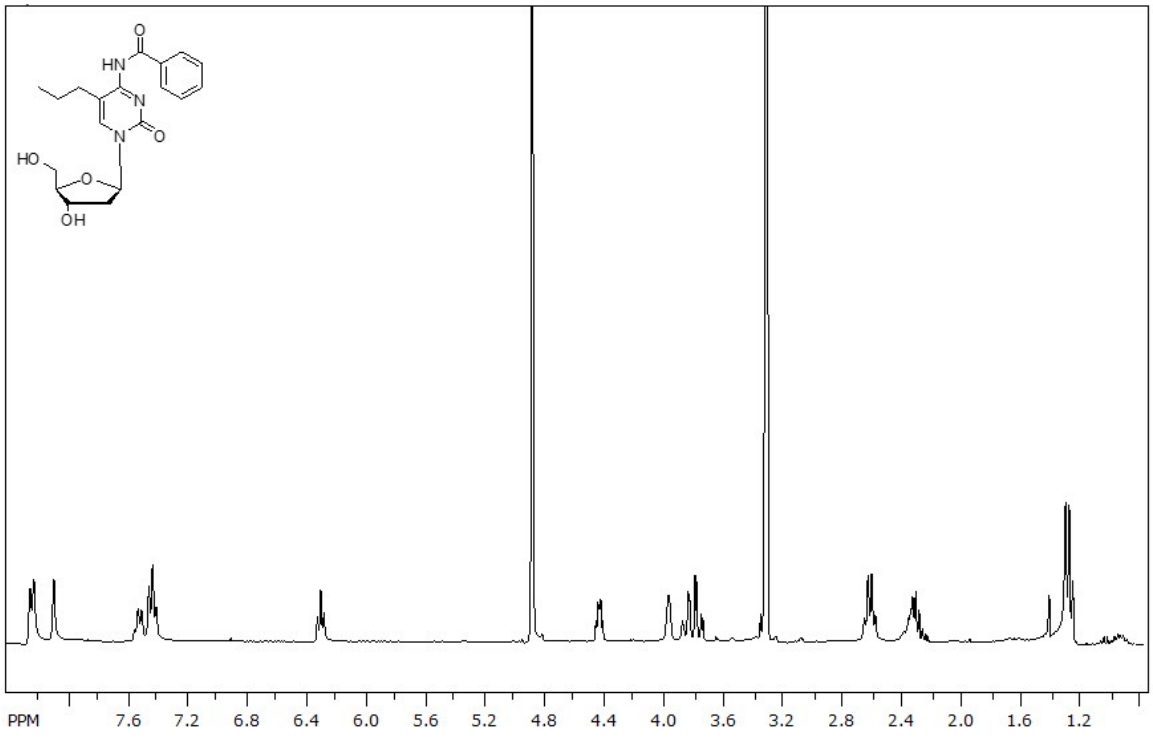


Figure S12. H-NMR of 4-*N*-Benzoyl-5′-*O*-(dimethoxytrityl)-5-propyl-2′-deoxycytidine (**13**)

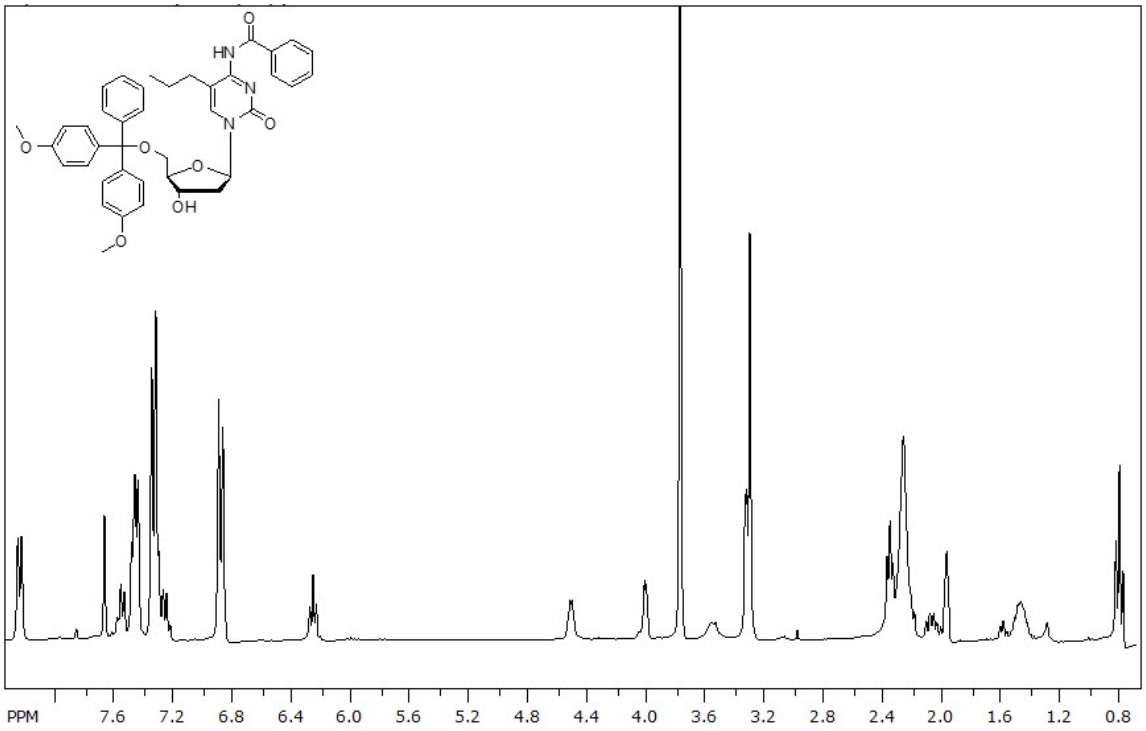


Figure S13. H-NMR of 3′,5′-*O*-*t*-butyldimethylsilyl-5-vinyl-2′-deoxycytidine (**15**)

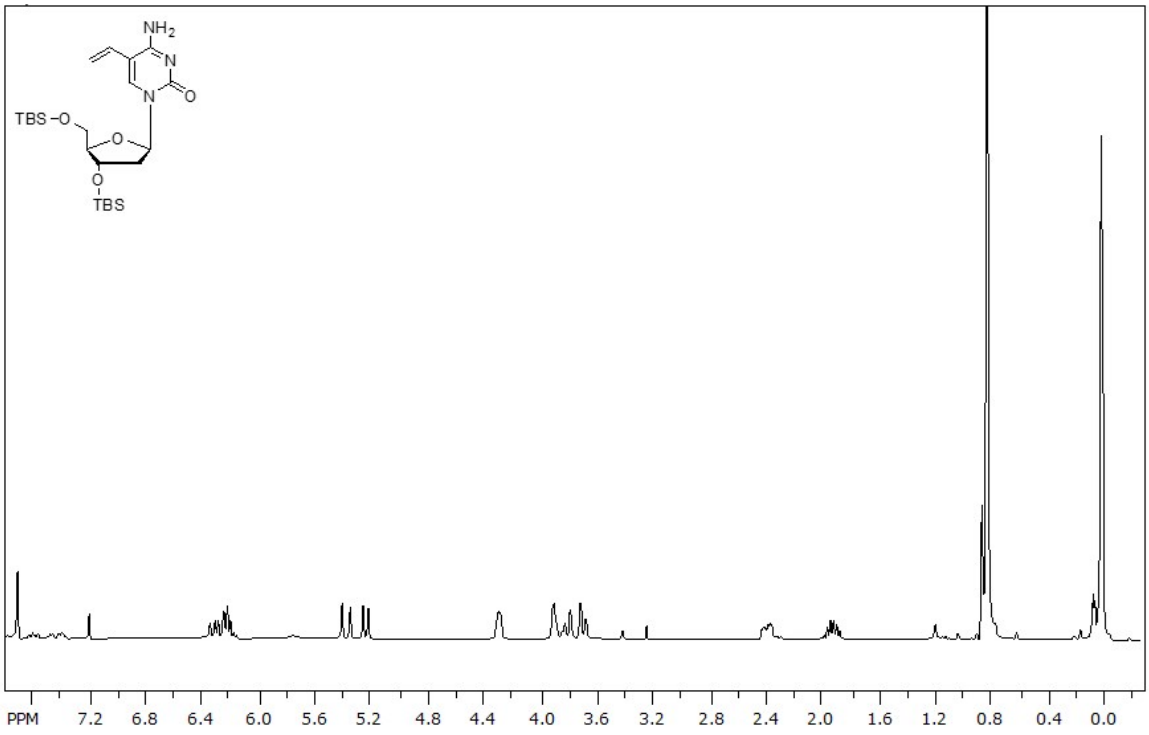


Figure S14. H-NMR of 4-*N*-Benzoyl-3′,5′-*O*-(*t*-butyldimethylsilyl)- 5-vinyl-2′-deoxycytidine (**16**)

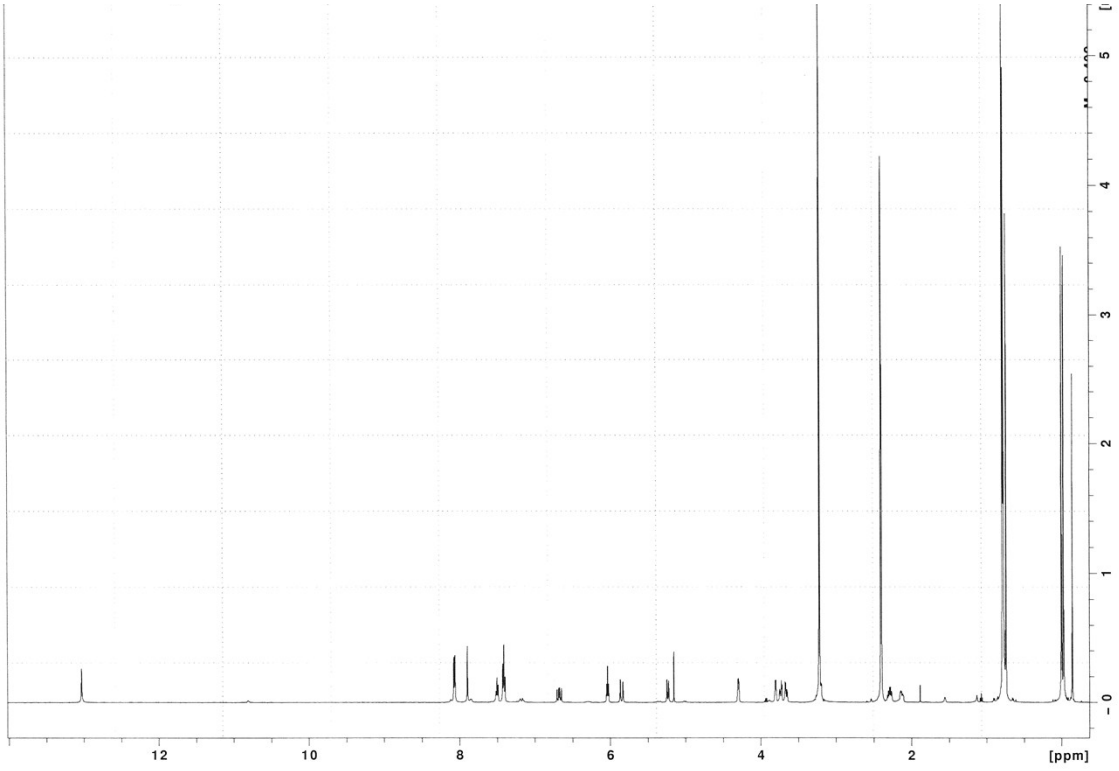


Figure S15. H-NMR and COSY of 4-*N*-Benzoyl-5-vinyl-2′-deoxycytidine (**17**)

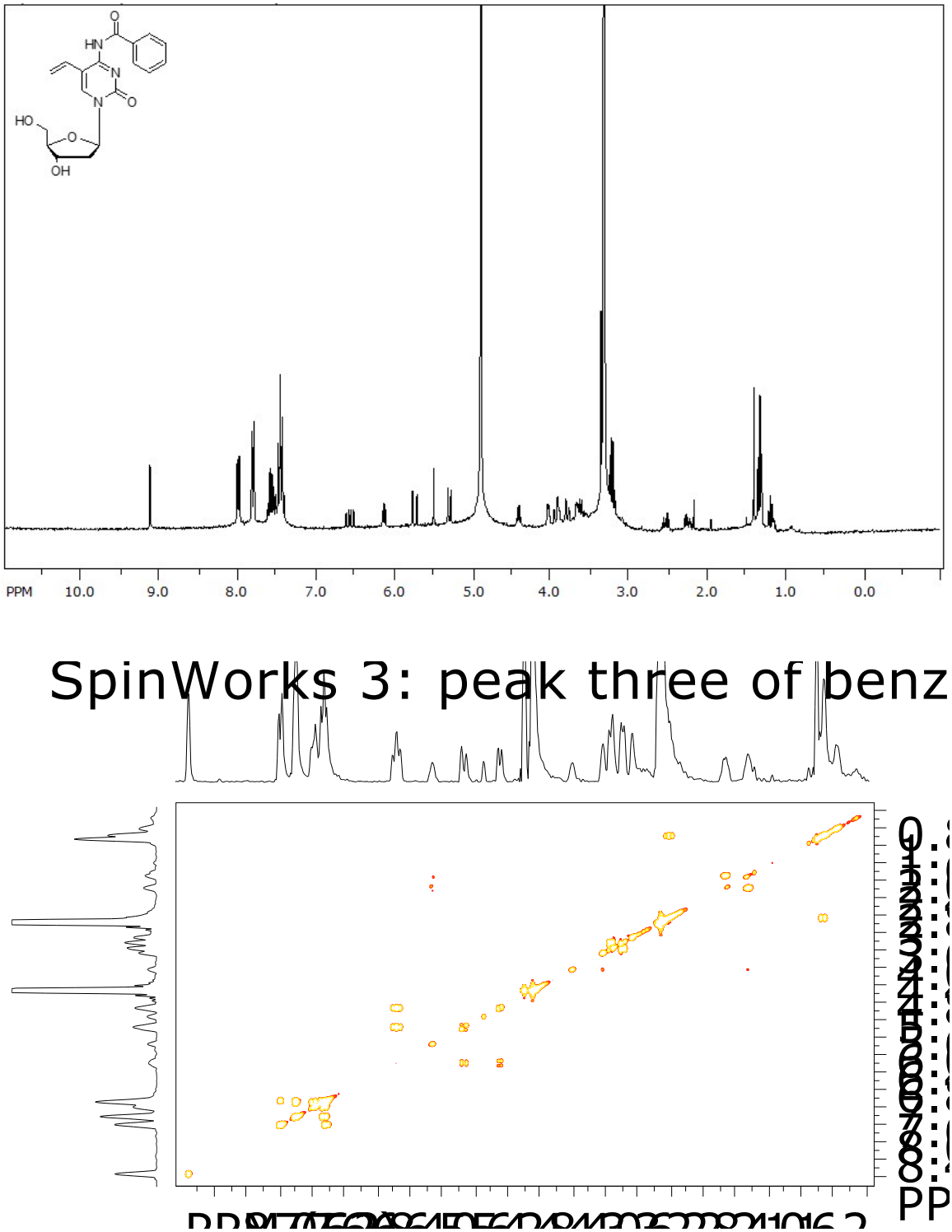


Figure S16. H-NMR of 4-*N*-Benzoyl-5′-*O*-(dimethoxytrityl)-5-vinyl- 2′deoxycytidine (**18**)

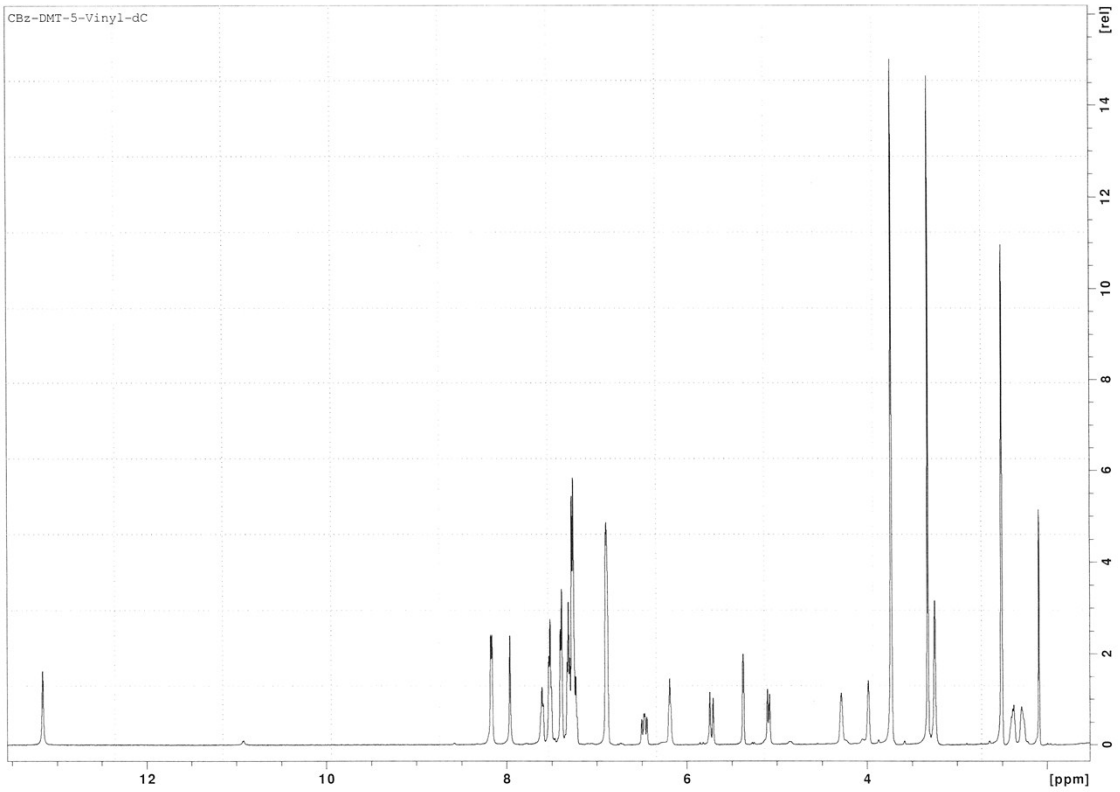


Figure S17. P-NMR of 4-*N*-Benzoyl-5'-*O*-(dimethoxytrityl)-5-vinyl-2'deoxycytidine 3'-(2-cyanoethyl)-*N*, *N*'-diisopropyl phosphoramidite (**19**)

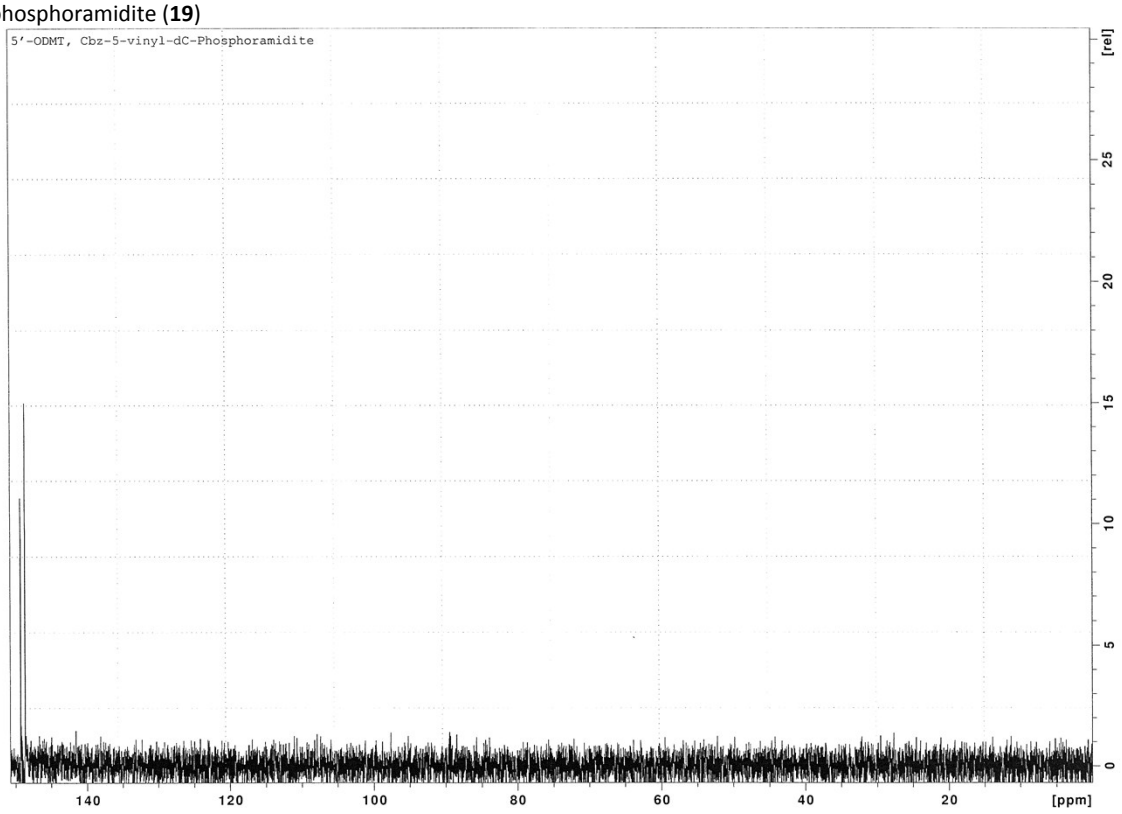


Figure S18. Deconvoluted ESI mass spectrum of Oligo: 5′-CGCGGA**[Et-C]**GCGGGTGCCGGG-3′

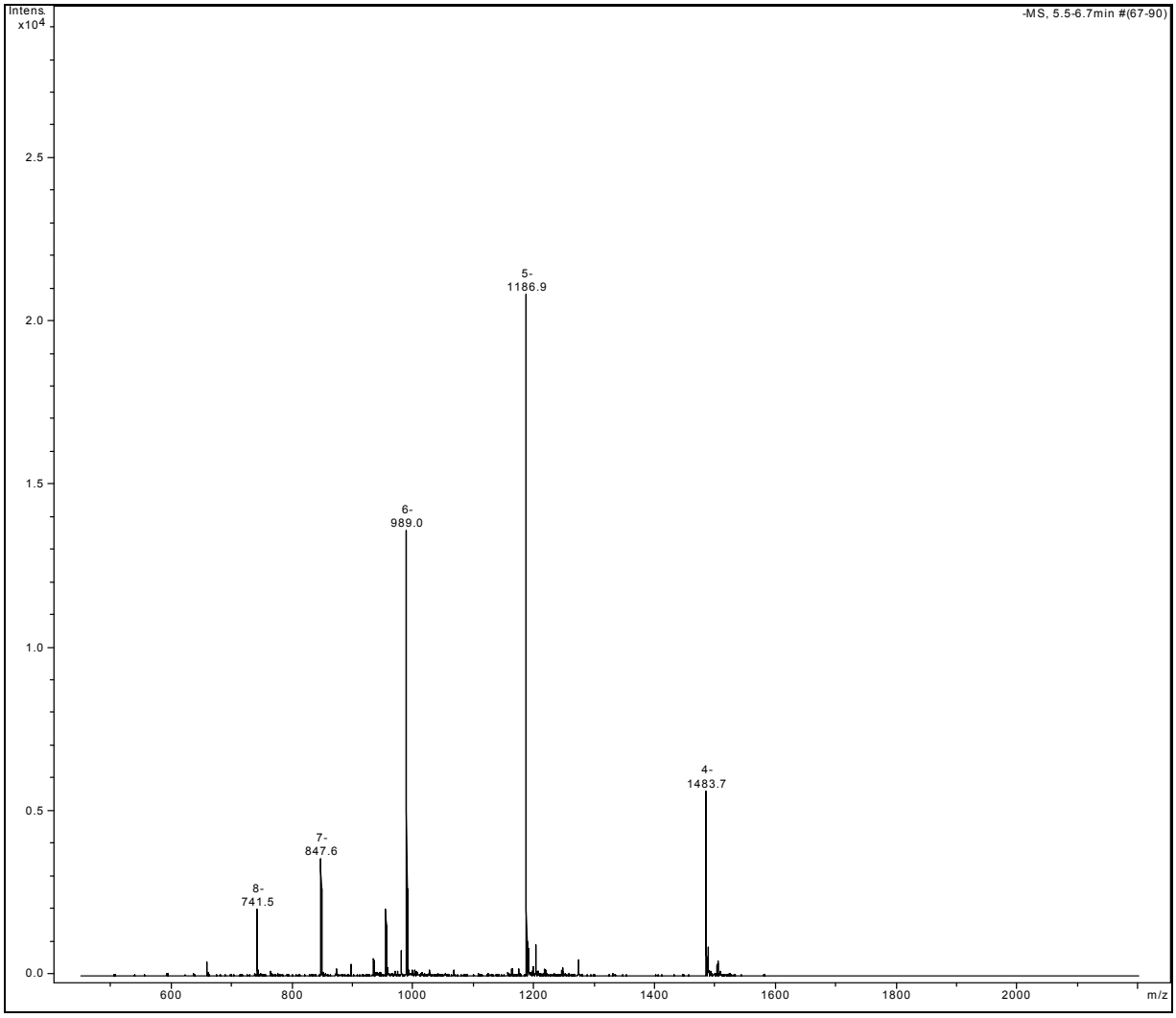


Figure S19. Deconvoluted spectrum of Oligo: 5′-CGCGGA**[Pr-C]**GCGGGTGCCGGG-3′

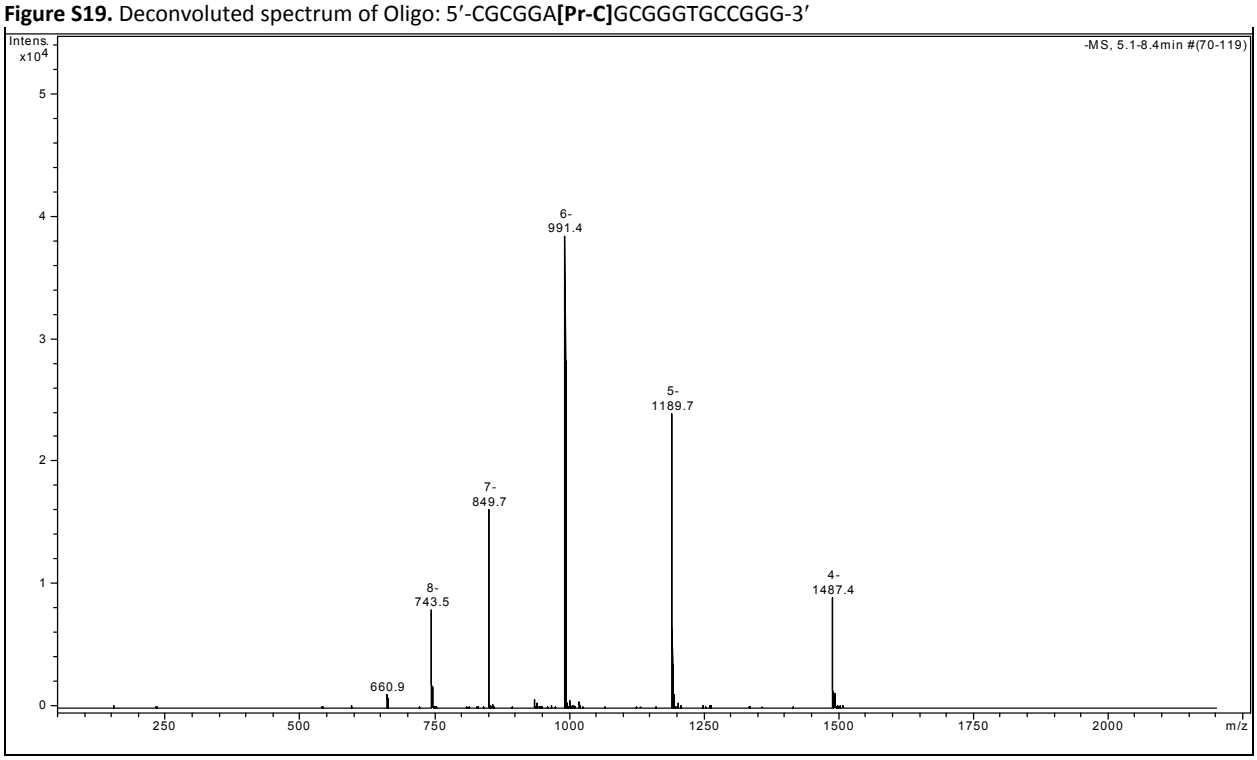


Figure S20. Deconvoluted spectrum of Oligo 5’-CGCGGA**[Pr-C]**GCGGGTGCCGGG-3’

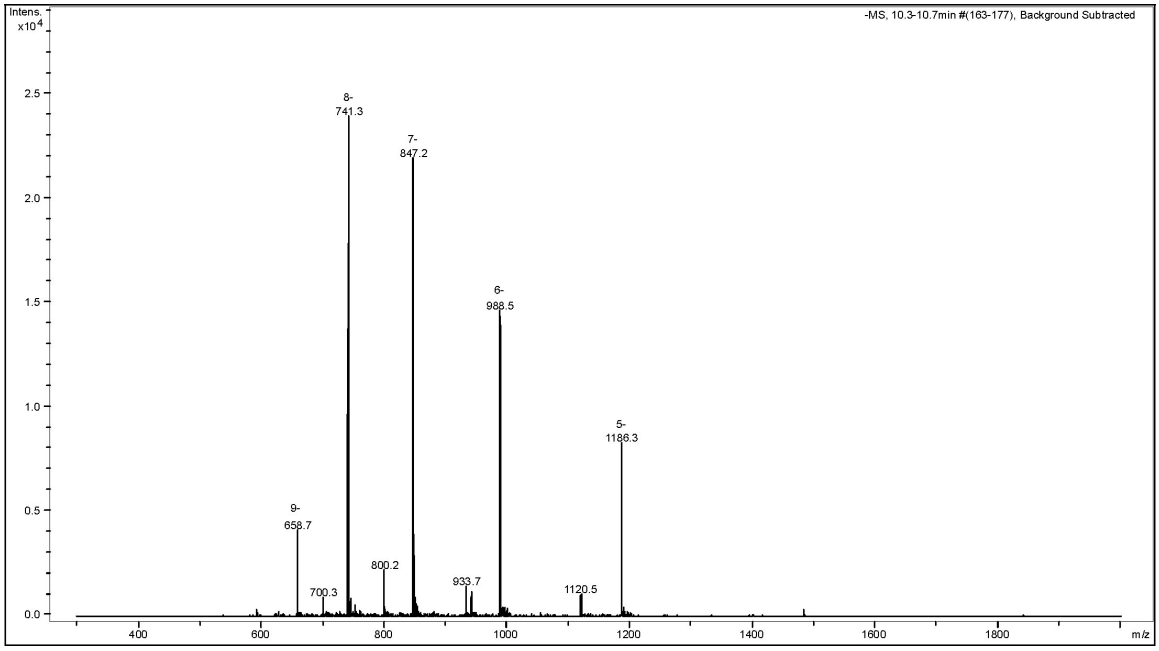
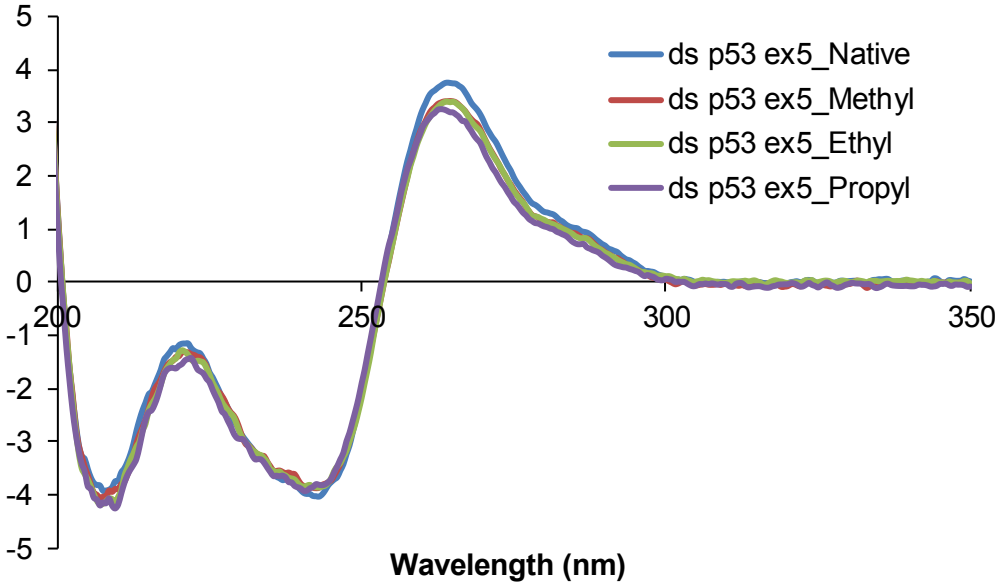


Figure S21. Circular Dichroism (CD) Spectroscopy of oligonucleotides used in this work.[9] Double-stranded DNA were obtained by dissolving equimolar amounts of the complimentary strands in 10 mM sodium phosphate buffer, pH 7.0 containing 50 mM NaCl to give a 9.7 µM final DNA concentration. The CD spectra were obtained with on a Jasco J-815 Spectropolarimeter scanning the wavelengths from 350 to 200 nm in a 1mm quartz cuvette, using a wavelength gradient of 0.5 nm. Spectra below are the average of N = 3 scans.



References

1. M. J. Gait, *Oligonucleotide synthesis: a practical approach*, IRL Press, Washington, DC, 1984.
2. B. Matter, G. Wang, R. Jones and N. Tretyakova, *Chem. Res. Toxicol*, 2004, **17**, 731-741.
3. R. Ziegel, A. Shallop, P. Upadhyaya, R. Jones and N. Tretyakova, *Biochemistry*, 2004, **43**, 540-549.
4. N. Tretyakova, B. Matter, R. Jones and A. Shallop, *Biochemistry*, 2002, **41**, 9535-9544.
5. R. Ziegel, A. Shallop, R. Jones and N. Tretyakova, *Chem. Res. Toxicol*, 2003, **16**, 541-550.
6. R. Guza, M. Rajesh, Q. Fang, A. E. Pegg and N. Tretyakova, *Chem. Res. Toxicol*, 2006, **19**, 531-538.
7. R. Guza, L. Ma, Q. Fang, A. E. Pegg and N. Y. Tretyakova, *J. Biol. Chem*, 2009, **284**, 22601-22610.
8. M. Rajesh, G. Wang, R. Jones and N. Tretyakova, *Biochemistry*, 2005, **44**, 2197-2207.
9. R. Guza, D. Kotandeniya, K. Murphy, T. Dissanayake, C. Lin, G. M. Giambasu, R. R. Lad, F. Wojciechowski, S. Amin, S. J. Sturla, R. H. Hudson, D. M. York, R. Jankowiak, R. Jones and N. Y. Tretyakova, *Nucleic Acids Res*, 2011, **39**, 3988-4006.
10. *Journal*, 2015.
11. J. Song, O. Rechkoblit, T. H. Bestor and D. J. Patel, *Science*, 2011, **331**, 1036-1040.
12. J. Song, M. Teplova, S. Ishibe-Murakami and D. J. Patel, *Science*, 2012, **335**, 709-712.
13. S. Wickramaratne, S. Ji, S. Mukherjee, Y. Su, M. G. Pence, L. Lior-Hoffmann, I. Fu, S. Broyde, F. P. Guengerich, M. Distefano, O. D. Scharer, Y. Y. Sham and N. Tretyakova, *J Biol Chem*, 2016, **291**, 23589-23603.
14. J. Tang, K. Maddali, Y. Pommier, Y. Y. Sham and Z. Wang, *Bioorg Med Chem Lett*, 2010, **20**, 3275-3279.
15. V. K. Vyas, R. D. Ukawala, M. Ghate and C. Chintha, *Indian J Pharm Sci*, 2012, **74**, 1-17.
16. W. L. Jorgensen, D. S. Maxwell and J. TiradoRives, *J Am Chem Soc*, 1996, **118**, 11225-11236.
17. W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J Am Chem Soc*, 1990, **112**, 6127-6129.