

## SUPPORTING INFORMATION

### **The Importance of Steric Effects on the Efficiency and Fidelity of Transcription by T7 RNA Polymerase**

Sébastien Ulrich, Eric T. Kool\*

*Department of Chemistry, Stanford University  
Stanford, CA 94305-5080 (USA)*

*\*e-mail: kool@stanford.edu*

General Information .....	2
Synthesis of nonpolar nucleoside analogues B and I.....	2
Synthesis of nonpolar nucleoside analogue <sup>Me</sup> L.....	9
MALDI-TOF MS characterization of DNA templates .....	16
Enzymatic transcription reactions at low product-conversion .....	20
Enzymatic transcription reactions: comparison of templates L and <sup>Me</sup> L .....	20
Composition analysis of RNA transcripts: control experiment.....	21
Composition analysis of RNA transcripts: experiments with different DNA templates (T, H, F, L, B, and I) .....	22
Identification of the doubled bands seen in the gel electrophoresis.....	23
References.....	23

## General Information

All reagents were purchased from commercial suppliers and used without further purification unless otherwise noted. The water used in the enzymatic studies was certified RNase free.  $^1\text{H}$  NMR spectra were recorded on Inova 300, Mercury 400, or Inova 500 instruments. Chemical shifts are given in ppm. Residual solvent peaks were taken as reference (1). The coupling constants  $J$  are given in Hz. Peaks are described as singlet (s), doublet (d), triplet (t), doublet of doublet (dd), multiplet (m) and broad (br).  $^{13}\text{C}$  NMR spectra were recorded on Inova 300, Mercury 400, or Inova 500 instruments. All spectra were measured in broadband decoupled conditions. Chemical shifts are given in ppm. Residual solvent peak were taken as reference (1). High Resolution ElectroSpray Ionization Mass Spectrometry (HR-ESI-MS) analyses were performed at the Stanford University Mass Spectrometry facility. The given value represents the largest peak. The observed isotopic pattern conformed to the theoretical pattern in all cases. MALDI-TOF analyses were performed at the Stanford University Protein and Nucleic Acid facility.

Nucleoside analogues **H**, **F**, and **L** were synthesized as described previously (2, 3). Compounds **B** and **I** were prepared by a variation of previous methods, as described below:

## Synthesis of nonpolar nucleoside analogues **B** and **I**.

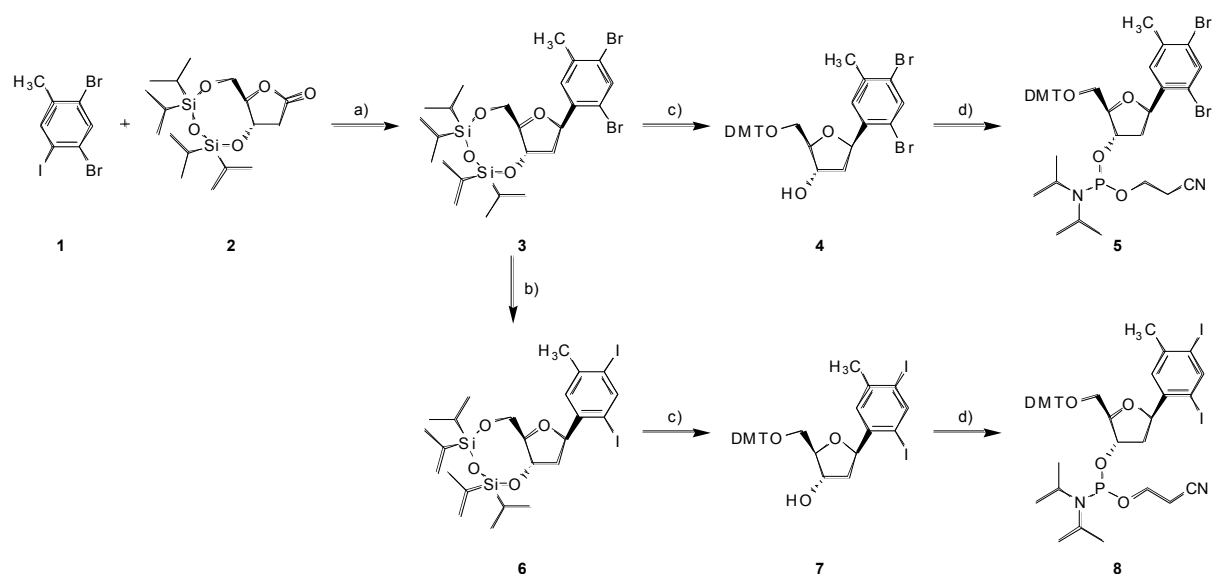


Figure S1. Synthetic route for the preparation of analogues **B** and **I**. a) i)  $i\text{PrMgCl}$ , THF,  $0^\circ\text{C}$ , ii) **2**, THF,  $0^\circ\text{C}$ , 18%; b) NaI, KI, CuI, *trans*- $N,N'$ -dimethyl-1,2-cyclohexanediamine, pentanol,  $130^\circ\text{C}$ , 4 d, 43%; c) i) TBAF, THF,  $0^\circ\text{C}$ , ii) DMT-Cl, DIPEA, pyridine, DCM; d) 2-Cyanoethyl  $N,N'$ -diisopropylchlorophosphoramidite, DIPEA, DCM.

**Compound 3.** Compound **1**(**2**) (1.5 g, 4.0 mmoles) was dissolved in anhydrous THF (4 mL) under an Ar atmosphere and cooled to  $0^\circ\text{C}$  with an ice bath.  $i\text{PrMgCl}$  (2.0 mL of a 2 M solution in THF, 4.0 mmoles) was added and the reaction mixture was stirred at  $0^\circ\text{C}$  for 1 h. A solution of compound **2**(**2**) (1 g, 2.67 mmoles) in THF anhydrous (4 mL) was added to the reaction mixture. Stirring at  $0^\circ\text{C}$  continued for 8 hours, after which a saturated solution of ammonium chloride was added. The organics were extracted with EtOAc, washed with brine, dried over  $\text{MgSO}_4$ , then under high vacuum. The crude oil was dissolved in anhydrous DCM (10 mL) under an Ar atmosphere and the solution was cooled to  $-78^\circ\text{C}$ .  $\text{Et}_3\text{SiH}$  (1.28 mL, 8.0 mmoles) and  $\text{BF}_3 \cdot \text{Et}_2\text{O}$  (1.0 mL, 8.0 mmoles) were successively added. After 4 hours the reaction was quenched at  $-78^\circ\text{C}$  by addition of a saturated solution of  $\text{NaHCO}_3$ . The organics were extracted with diethyl ether, washed with a saturated solution of  $\text{NaHCO}_3$  then brine, and dried over  $\text{Na}_2\text{SO}_4$ . Purification by column chromatography ( $\text{SiO}_2$ , hexanes $\rightarrow$ hexanes/DCM 70/30) afforded 292 mg (18%) of a colorless oil.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ): 7.68 (s, 1H), 7.47 (s, 1H), 5.24 (dd, 1H,  $J = 8.3, 5.4$ , H-1'), 4.42 (app dd, 1H,  $J = 14.6, 7.6$ , H-3'), 4.14-4.02 (m, 2H, H-5'), 3.83-3.78 (m, 1H, H-4'), 2.60-2.50 (m, 1H, H-2'  $\alpha$ ), 2.34 (s, 3H,  $\text{CH}_3$ ), 1.99-1.91 (m, 1H, H-2'  $\beta$ ), 1.12-0.92 (m, 28H, 4 x  $-\text{CH}(\text{CH}_3)_2$ );  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ): 141.6, 137.3, 135.0, 128.6, 123.6, 118.3, 85.1, 70.5, 61.8, 40.9, 22.5, 17.5, 17.4, 17.3, 17.3, 17.2, 17.1, 17.0, 13.5, 13.2, 12.8, 12.5.

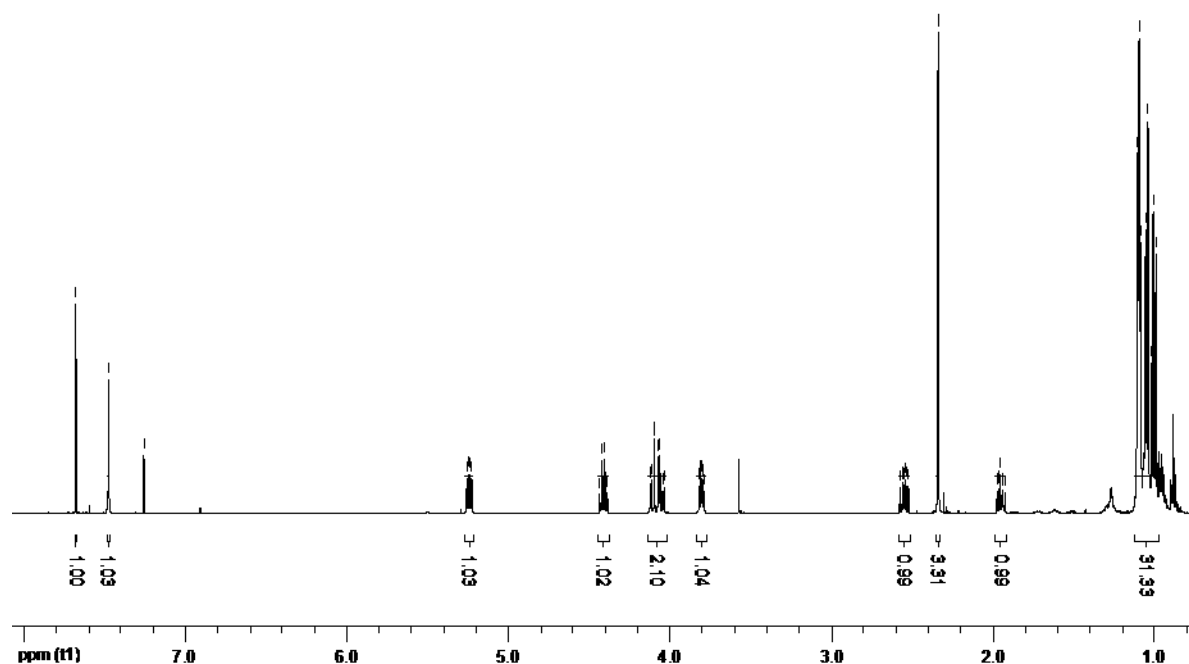


Figure S2.  $^1\text{H}$  NMR spectrum of compound **3**.

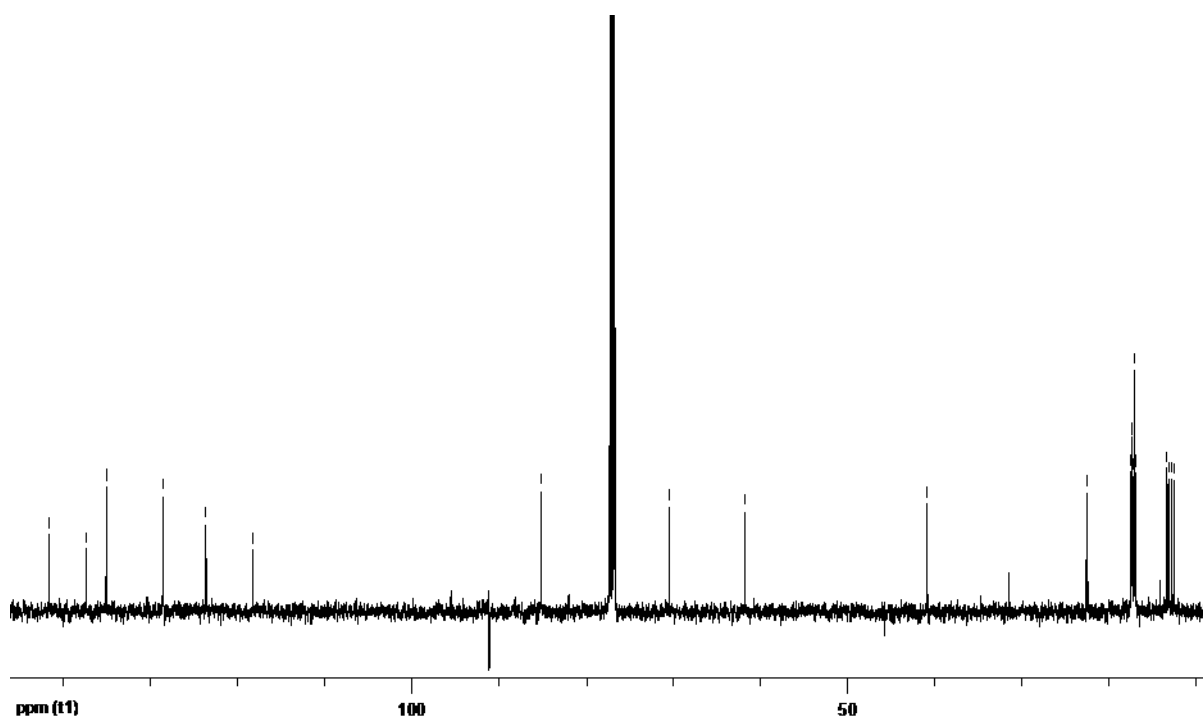


Figure S3.  $^{13}\text{C}$  NMR spectrum of compound **3**.

Compound **6**. The method was adapted from our previously reported synthesis (2). Compound **3** (315 mg, 0.52 mmoles), CuI (30 mg, 0.15 mmoles), NaI (931 mg, 6.21 mmoles), KI (1.3 g, 6.21 mmoles), and *trans*-*N,N'*-dimethyl-1,2-cyclohexanediamine (50  $\mu\text{L}$ , 0.32 mmoles) were placed in a pressure tube, under an Ar atmosphere. 1-pentanol (3 mL) was added and the reaction mixture was stirred at 130-140 $^\circ\text{C}$  for 4 d. A saturated solution of  $\text{NaHCO}_3$  was added and the organics were extracted with DCM. The organic solution was dried over  $\text{Na}_2\text{SO}_4$  and then under high vacuum. Purification by column chromatography ( $\text{SiO}_2$ , hexanes $\rightarrow$ hexanes/DCM 30/70) afforded 155 mg (43%) of product **6**.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ): 8.19 (s, 1H), 7.41 (s, 1H), 5.10 (dd, 1H,  $J = 8.5, 5.0$ , H-1'), 4.40 (app dd, 1H,  $J = 15.0, 7.5$ , H-3'), 4.11-4.04 (m, 2H, H-5'), 3.81-3.77 (m, 1H, H-4'), 2.59-2.53 (m, 1H, H-2'  $\alpha$ ), 2.37 (s, 3H,  $\text{CH}_3$ ), 1.93-1.88 (m, 1H, H-2' $\beta$ ), 1.12-

0.95 (m, 28H, 4 x -CH(CH<sub>3</sub>)<sub>2</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>): 147.6, 145.9, 142.1, 127.9, 122.5, 100.4, 91.5, 85.4, 81.1, 70.5, 61.9, 41.2, 27.9, 17.8, 17.7, 17.6, 17.5, 17.4, 17.2, 13.7, 13.5, 13.1, 12.8.

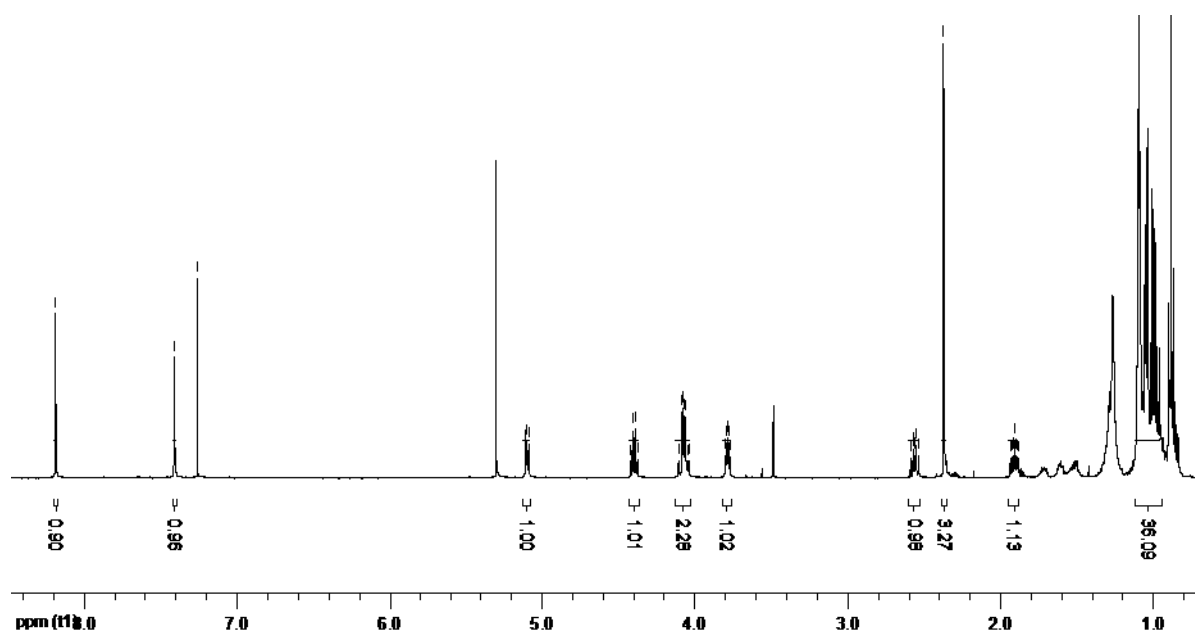


Figure S4. <sup>1</sup>H NMR spectrum of compound **6**.

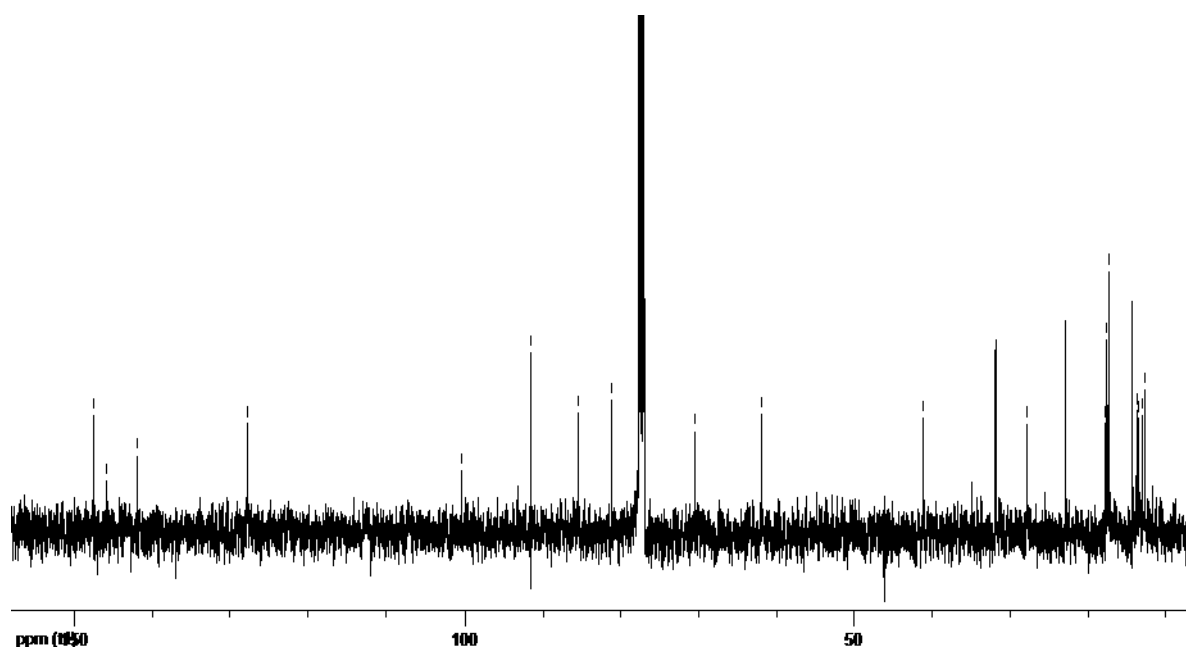


Figure S5. <sup>13</sup>C NMR spectrum of compound **6**.

Compounds **4**, **5**, **7**, and **8** were prepared according to the previously reported methods (2, 3). The identification was confirmed by NMR analysis:

Compound **4**. <sup>1</sup>H NMR (CDCl<sub>3</sub>): 7.71 (s, 1H), 7.60 (s, 1H), 7.50 (m, 2H), 7.41-7.37 (m, 4H), 7.31-7.24 (m, 3H), 6.85 (dd, 4H, J = 9.0, 0.6), 5.39 (dd, 1H, J = 9.6, 6.0), 4.43-4.40 (m, 1H), 4.12-4.08 (m, 1H), 3.81 (s, 6H), 3.45-3.35 (m, 2H), 2.55-2.48 (m, 1H), 2.25 (s, 3H), 1.95-1.86 (m, 1H); <sup>13</sup>C NMR (CDCl<sub>3</sub>): 158.4, 144.8, 140.8, 137.4, 135.9, 135.9, 135.0, 130.0, 128.8, 128.1, 127.8, 126.8, 123.6, 118.4, 113.1, 86.2, 86.0, 78.5, 74.1, 64.1, 55.1, 42.6, 22.4.

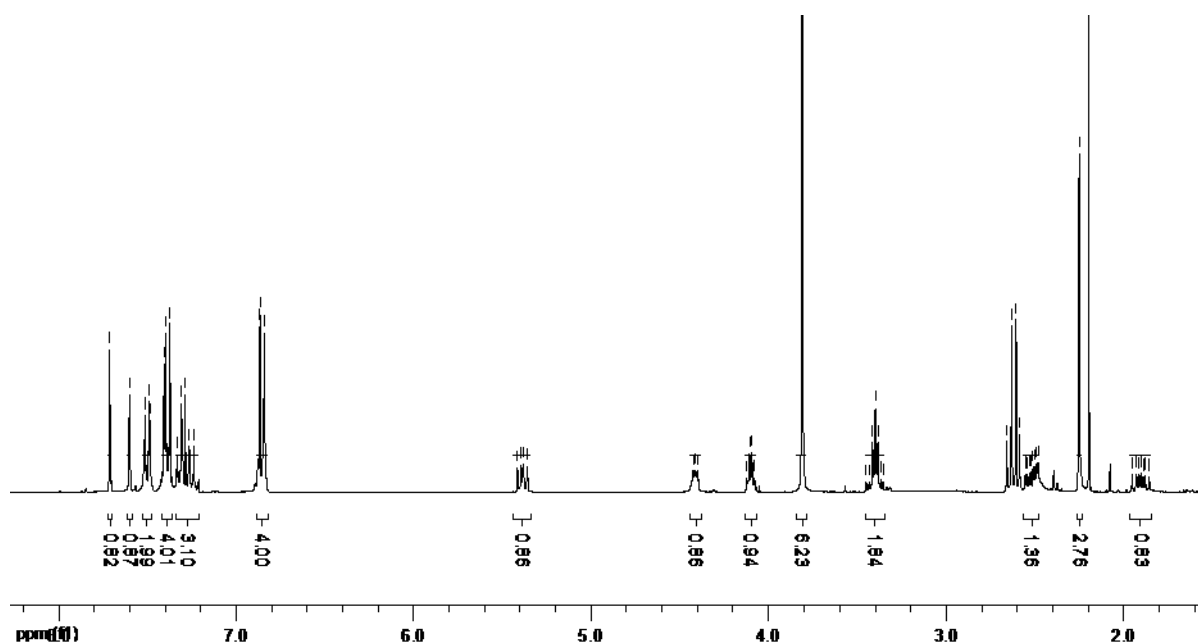


Figure S6.  $^1\text{H}$  NMR spectrum of compound 4.

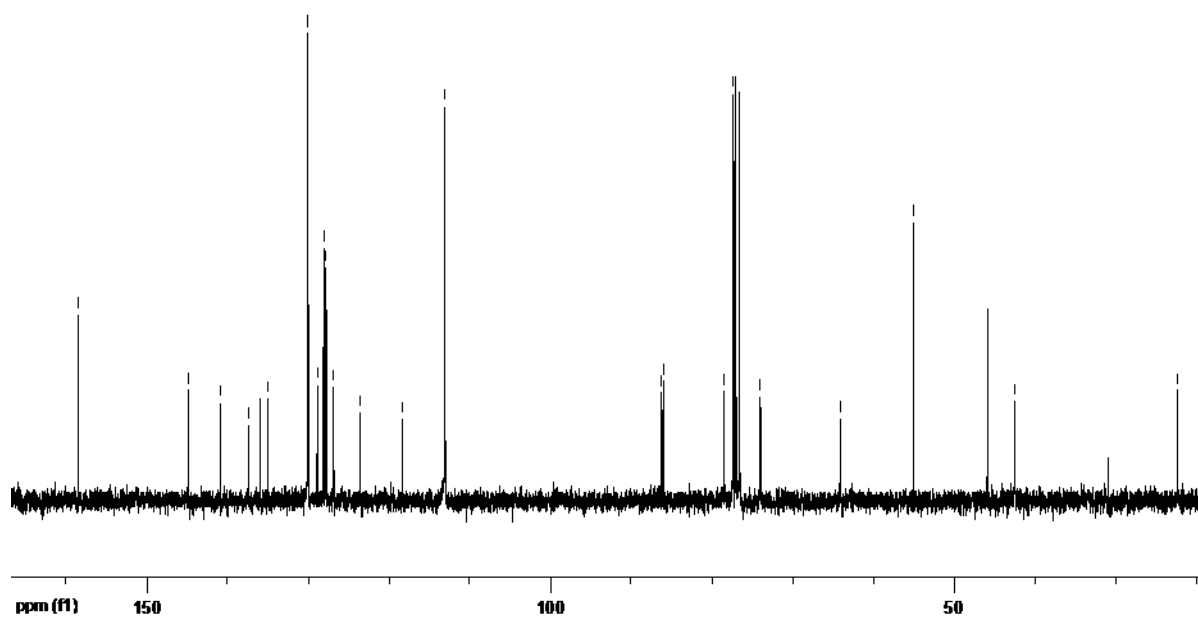


Figure S7.  $^{13}\text{C}$  NMR spectrum of compound 4.

Compound 5.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ): 7.70 (app d,  $J = 1.2$ , 1H), 7.64 (d,  $J = 8.0$ , 1H), 7.51-7.48 (m, 2H), 7.40-7.35 (m, 4H), 7.31-7.21 (m, 3H), 6.85-6.81 (m, 4H), 5.36 (m, 1H, H1'), 4.54-4.48 (m, 1H, H3'), 4.23 (m, 1H, H4'), 3.87-3.76 (m, 1H), 3.79 (app d, 6H,  $J = 3.2$ , OCH3), 3.67-3.56 (m, 3H), 3.48-3.39 (m, 1H), 3.31-3.26 (m, 1H), 2.66-2.42 (m, 3H), 2.22 (app d,  $J = 4.0$ , 3H, ArCH3), 1.90-1.86 (m, 1H, H2'), 1.21-1.17 (m, 8H), 1.08-1.02 (m, 4H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ): 158.7 (d), 145.1 (d), 141.0 (d), 137.8, 136.2 (d), 136.1 (d), 135.3 (d), 130.3 (d), 129.2 (d), 128.4 (d), 128.1, 127.1 (d), 124.0 (d), 118.8 (d), 117.8 (d), 113.3, 86.4, 85.9 (d), 85.7 (d), 79.1 (d), 75.9 (d), 75.8 (d), 64.0 (d), 58.6 (t), 55.4 (d), 43.4 (t), 42.1 (m), 24.8 (m), 22.7, 20.9, 20.6 (d), 20.4 (d);  $^{31}\text{P}$  NMR: 149.6, 148.7.

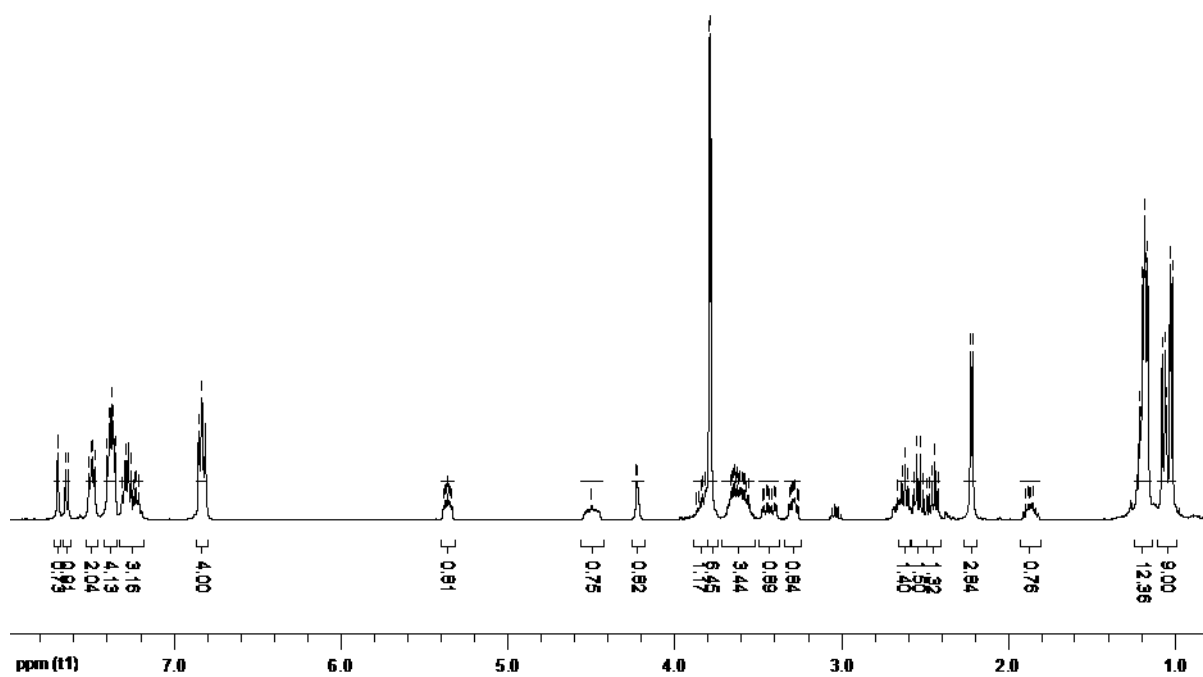


Figure S8. <sup>1</sup>H NMR spectrum of compound 5.

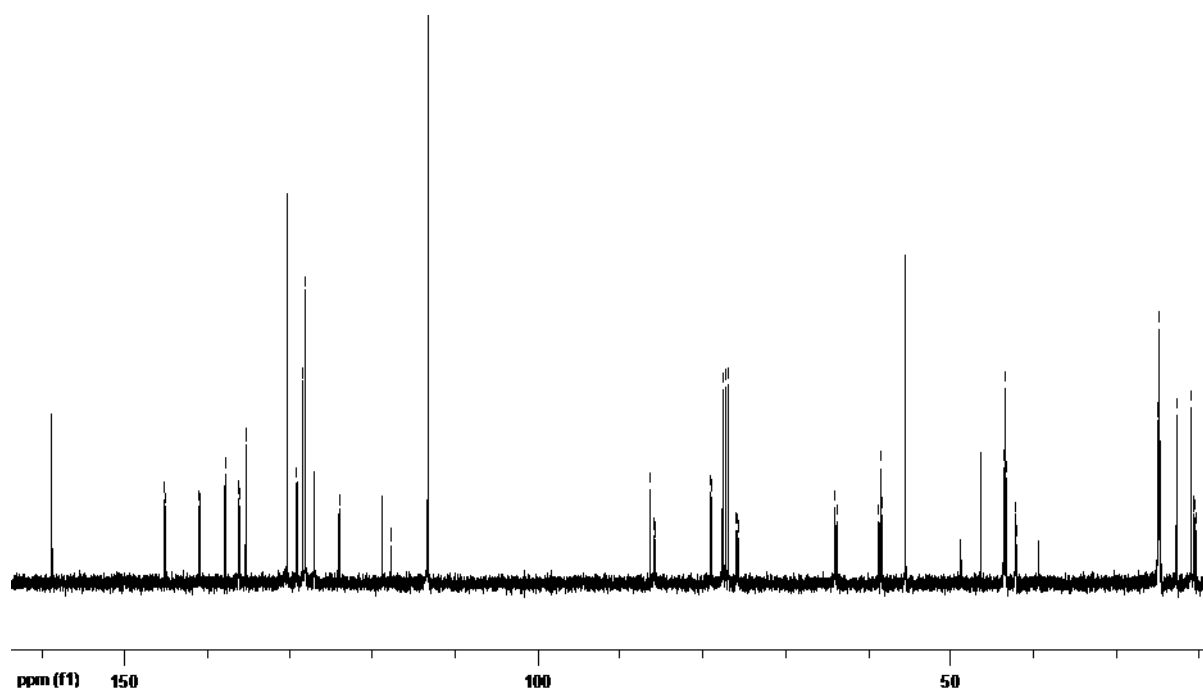


Figure S9. <sup>13</sup>C NMR spectrum of compound 5.

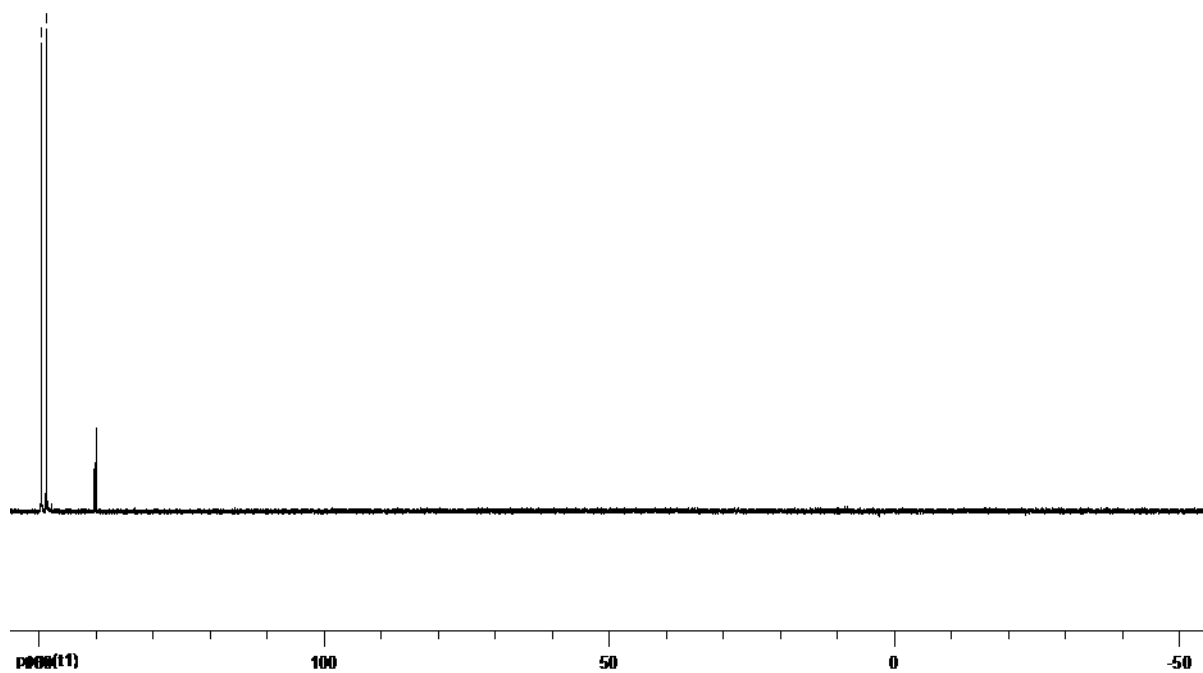


Figure S10.  $^{31}\text{P}$  NMR spectrum of compound 5.

Compound 7.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ): 8.20 (s, 1H), 7.53 (s, 1H), 7.50-7.46 (m, 2H), 7.39-7.35 (m, 4H), 7.31-7.22 (m, 3H), 6.83 (dd, 4H,  $J = 9.0, 0.9$ ), 5.22 (dd, 1H,  $J = 9.3, 6.0$ ), 4.40-4.37 (m, 1H), 4.09-4.05 (m, 1H), 3.79 (s, 6H), 3.44-3.33 (m, 2H), 2.56-2.48 (m, 1H), 2.24 (s, 3H), 1.90-1.80 (m, 1H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ): 158.4, 147.3, 144.8, 144.8, 141.8, 135.9, 135.9, 130.0, 128.1, 127.8, 127.7, 126.8, 113.1, 100.2, 93.1, 86.3, 86.2, 82.7, 74.0, 64.0, 55.2, 46.0, 42.7, 27.5.

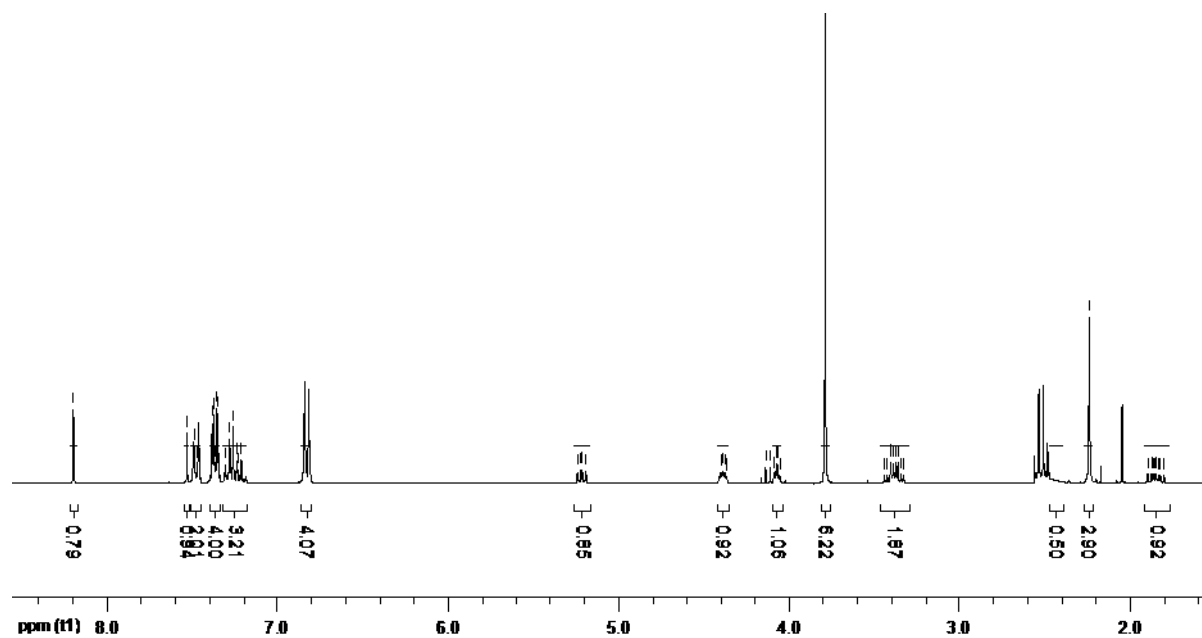


Figure S11.  $^1\text{H}$  NMR spectrum of compound 7.

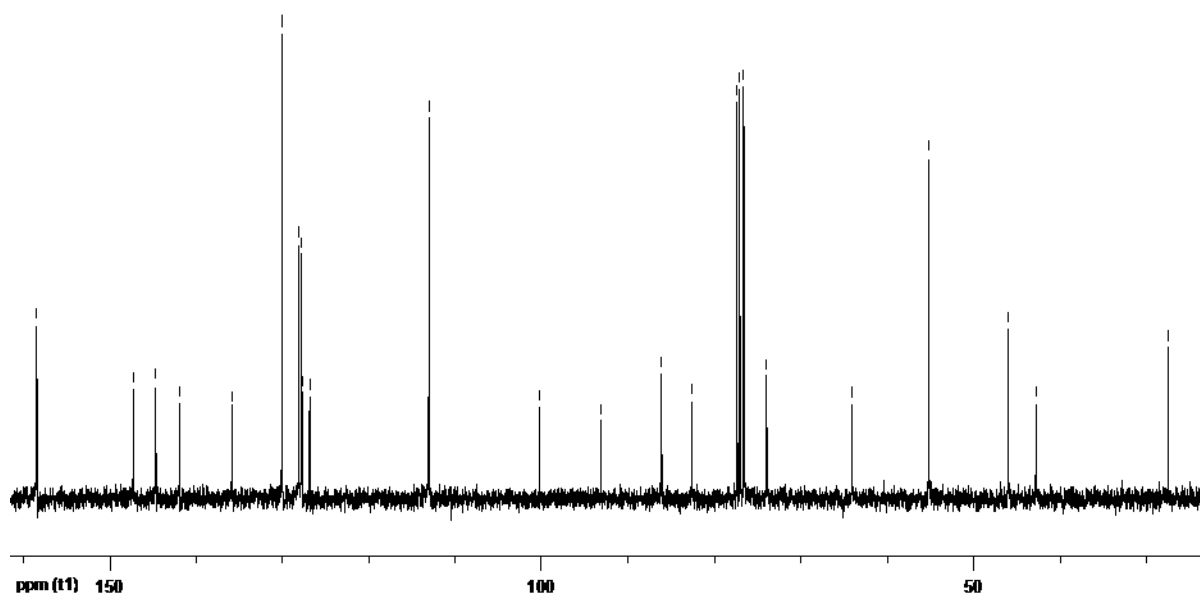


Figure S12. <sup>13</sup>C NMR spectrum of compound 7.

Compound 8. <sup>1</sup>H NMR (CDCl<sub>3</sub>): 8.20 (s, 1H), 7.58 (d, J = 7.2, 1H), 7.50-7.46 (m, 2H), 7.39-7.34 (m, 4H), 7.30-7.21 (m, 3H), 6.84-6.80 (m, 4H), 5.22-5.18 (m, 1H, H1'), 4.53-4.43 (m, 1H, H3'), 4.21 (m, 1H, H4'), 3.97-3.77 (m, 1H), 3.79 (app d, 6H, J = 3.2, OCH<sub>3</sub>), 3.67-3.54 (m, 3H), 3.48-3.39 (m, 1H), 3.29-3.24 (m, 1H), 2.69-2.41 (m, 3H), 2.23 (app d, J = 3.2, 3H, ArCH<sub>3</sub>), 1.87-1.77 (m, 1H, H2'), 1.21-1.17 (m, 8H), 1.08-1.02 (m, 4H); <sup>31</sup>P NMR: 149.7, 148.6.

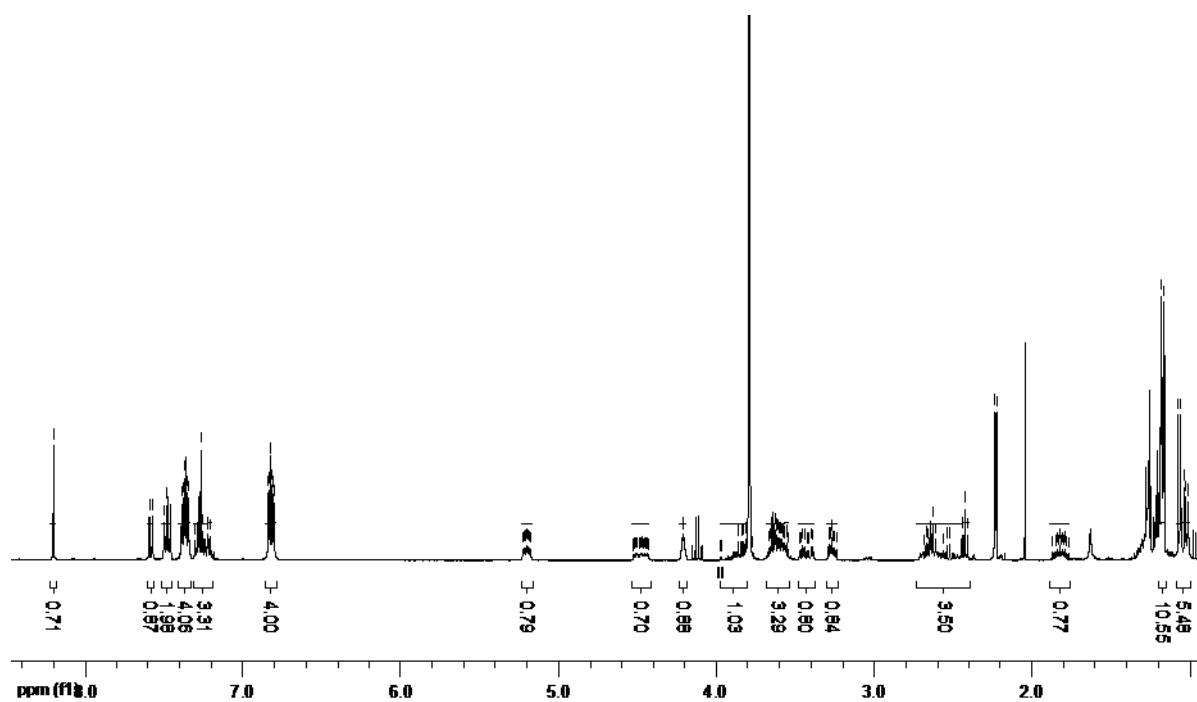


Figure S13. <sup>1</sup>H NMR spectrum of compound 8.



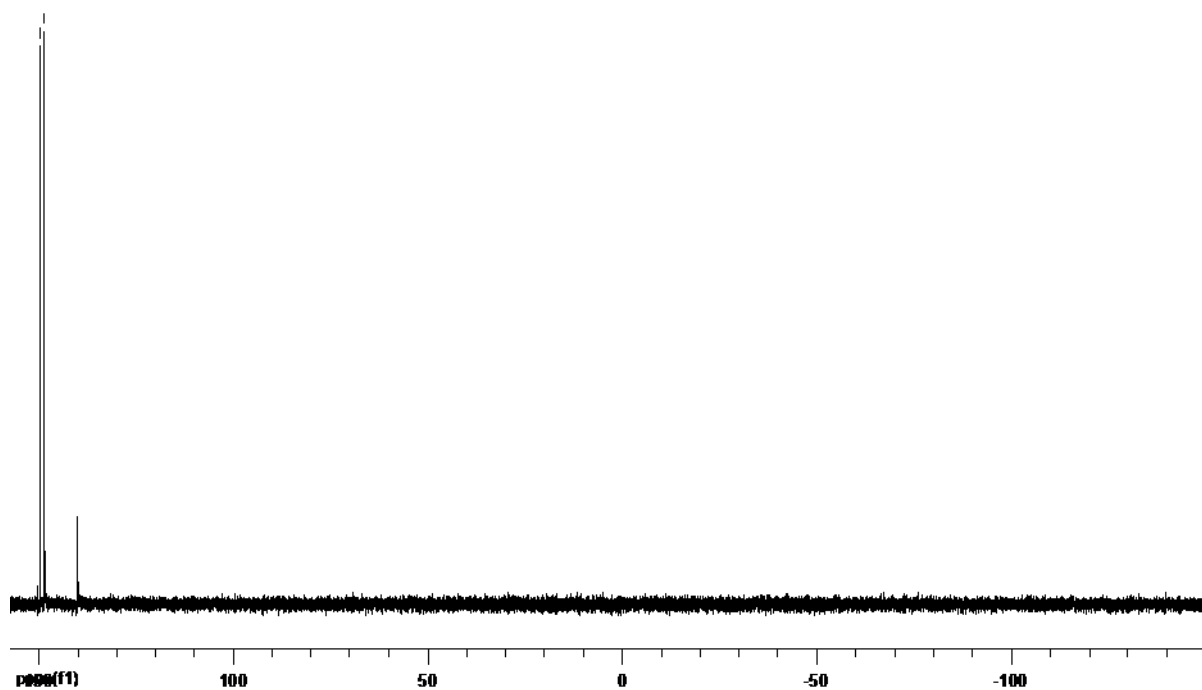


Figure S14.  $^{31}\text{P}$  NMR spectrum of compound **8**.

### Synthesis of nonpolar nucleoside analogue $^{\text{Me}}\text{L}$

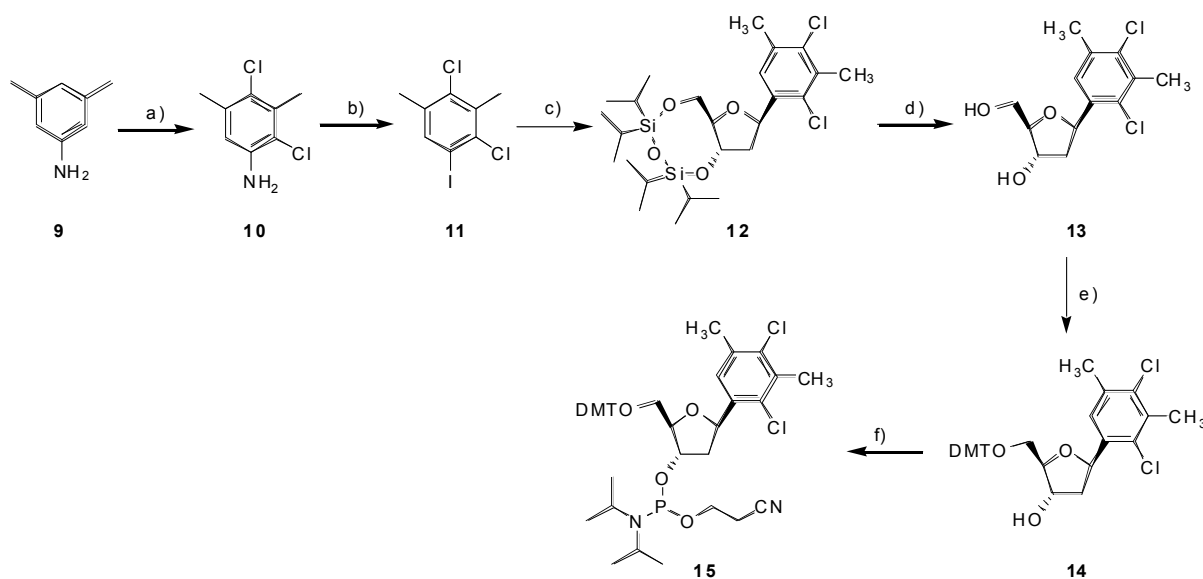


Figure S15. Synthetic route for the preparation of analogue  $^{\text{Me}}\text{L}$ . a) *N*-Chlorosuccinimide,  $\text{CCl}_4$ , 35%; b) i)  $\text{NaNO}_2$ ,  $\text{HCl}$ , then ii)  $\text{KI}$ , 80%; c) i) *n*BuLi, THF,  $-78^\circ\text{C}$ , ii) **2**, THF,  $-78^\circ\text{C}$ , 28%; d) TBAF, THF,  $0^\circ\text{C}$ , quant.; e) DMT-Cl, DIPEA, pyridine, DCM, 63%; f) 2-Cyanoethyl *N,N*-diisopropylchlorophosphoramidite, DIPEA, DCM, 54%.

Compound **10**. The method was adapted from an analogous reaction (**4**). Compound **9** (1 mL, 8.0 mmoles) was dissolved in  $\text{CCl}_4$  (4.5 mL) and the solution was cooled to  $0^\circ\text{C}$ . *N*-Chlorosuccinimide (2.14 g, 16.0 mmoles) was added portionwise and the reaction was monitored by TLC (Hexanes/EtOAc 19/1). After 4 h, water was added and the organics were extracted with DCM, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under vacuum. Purification by column chromatography ( $\text{SiO}_2$ , hexanes→hexanes/EtOAc 90/10) afforded 526 mg (35%) of product **10**.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ): 6.59 (s, 1H), 4.60-3.80 (br, 2H), 2.45 (s, 3H), 2.28 (s, 3H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ): 140.4, 135.0, 134.3, 124.7, 118.3, 115.2, 20.8, 18.2; HR-ESI-MS: calcd for  $[\text{C}_8\text{H}_9\text{NCl}_2+\text{H}]^+$  190.0185, found 190.0182.

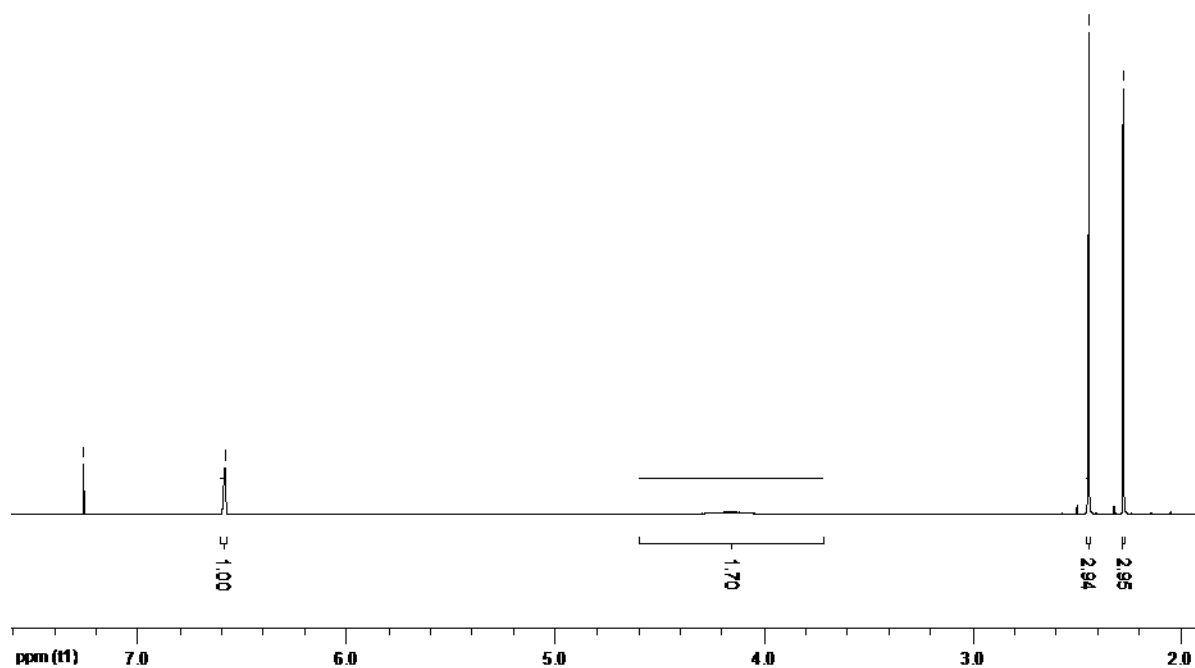


Figure S16.  $^1\text{H}$  NMR spectrum of compound **10**.

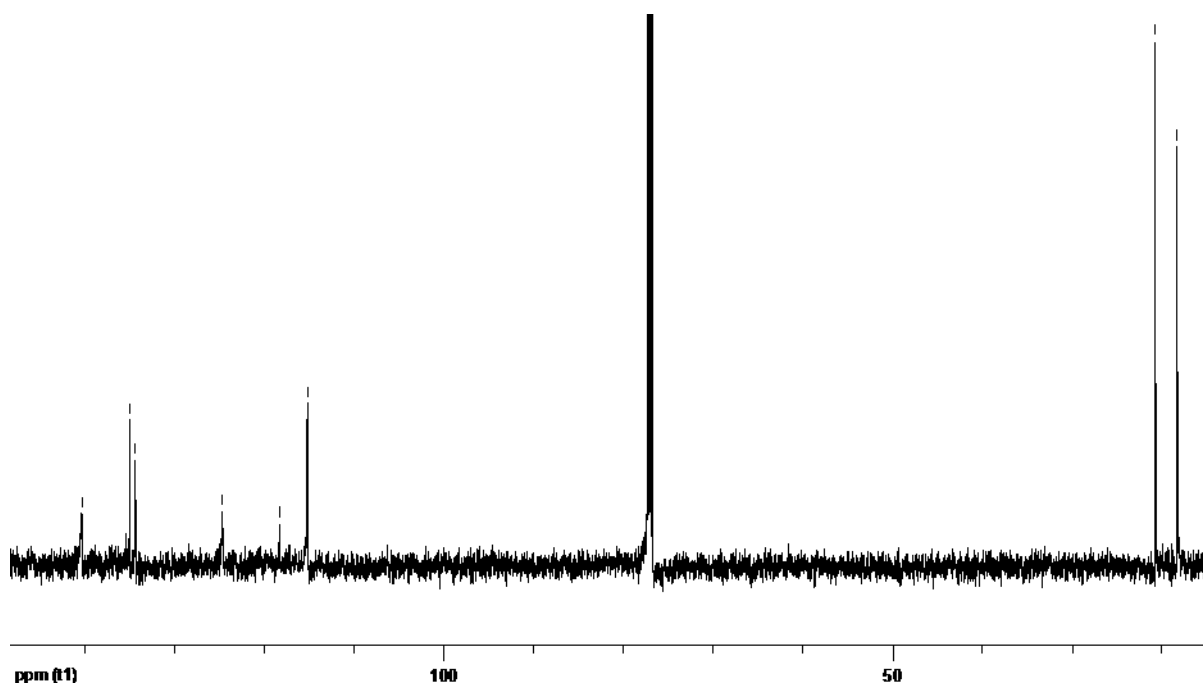


Figure S17.  $^{13}\text{C}$  NMR spectrum of compound **10**.

Compound **11**. Compound **10** (150 mg, 0.79 mmoles) was suspended in conc. 37% HCl (1.5 mL) and the solution was cooled to  $0^\circ\text{C}$ . A solution of  $\text{NaNO}_2$  (93 mg, 1.34 mmoles) in water (0.6 mL) was added and stirring was continued at  $0^\circ\text{C}$ . After 15 minutes a solution of KI (563 mg, 3.39 mmoles) in water (1.5 mL) was slowly added and the reaction mixture was allowed to warm up to room temperature and stirred overnight. The organics were extracted with hexanes, washed with a saturated solution of  $\text{Na}_2\text{S}_2\text{O}_3$ , dried over  $\text{MgSO}_4$ , and concentrated under vacuum. Purification by column chromatography ( $\text{SiO}_2$ , hexanes) afforded 190 mg (80%) of product **11**.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ): 7.61 (s, 1H), 2.56 (s, 3H), 2.30 (s, 3H);  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ ): 138.5, 136.4, 135.8, 135.4, 96.0, 20.4, 20.2; GC-MS: calcd for  $[\text{C}_8\text{H}_7\text{Cl}_2\text{I}]^+$  300 (100), 302 (64), 304 (10), found 300, 302, 304.

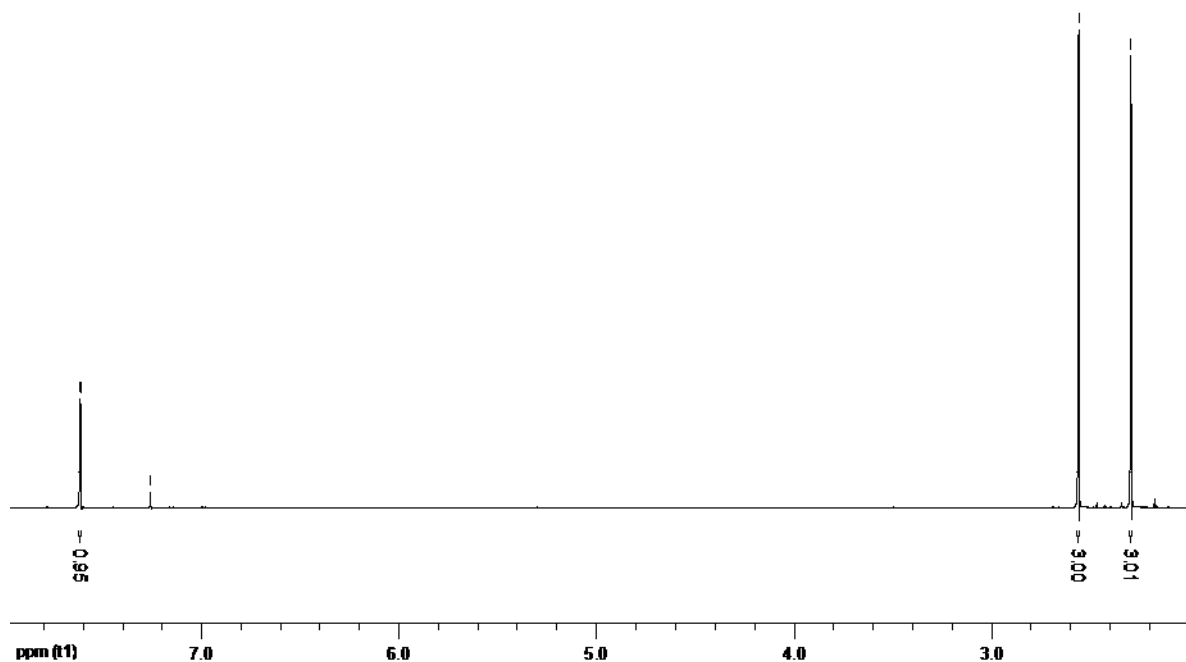


Figure S18.  $^1\text{H}$  NMR spectrum of compound **11**.

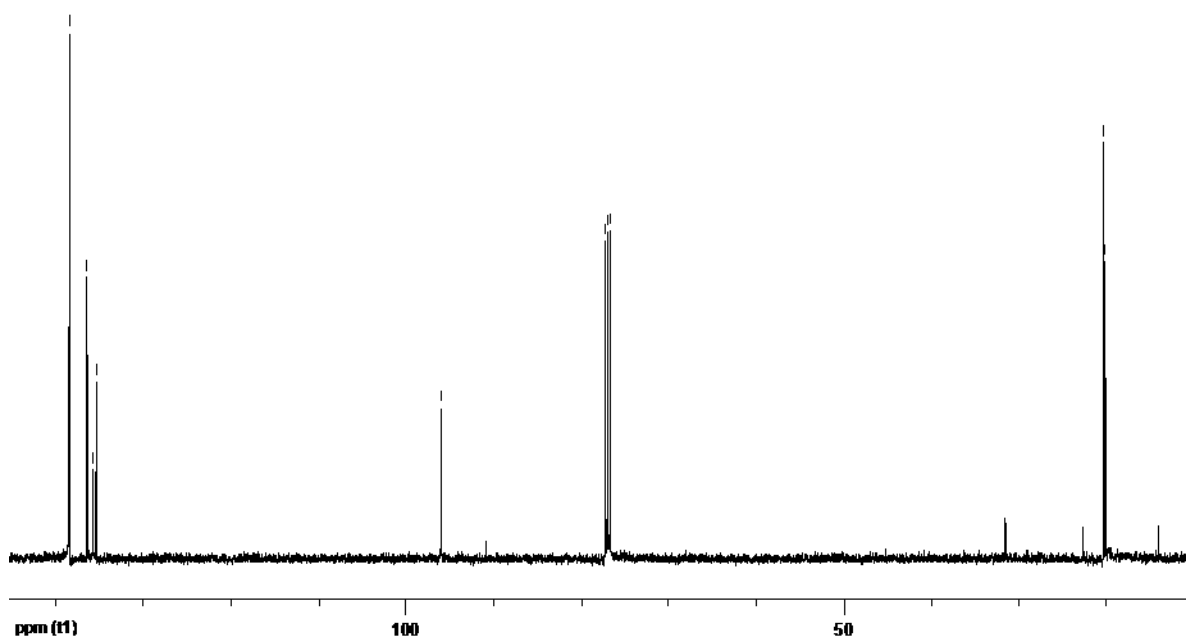


Figure S19.  $^{13}\text{C}$  NMR spectrum of compound **11**.

Compound **12**. Compound **11** (600 mg, 2.00 mmoles) was dissolved in anhydrous THF (5 mL) and cooled to  $-78^\circ\text{C}$ . Freeze-thaw degassing cycles were performed. *n*-BuLi (0.8 mL of a 2.5 M solution in hexanes, 2.00 mmoles) was slowly added and the reaction mixture was stirred for 3 minutes. A solution of compound **2** (500 mg, 1.33 mmoles) in anhydrous THF (5 mL), cooled to  $-78^\circ\text{C}$ , was then added *via* canula. After 4 hours at  $-78^\circ\text{C}$  the reaction was quenched by addition of a saturated solution of ammonium chloride. The organics were extracted with diethyl ether, dried over  $\text{MgSO}_4$ , and concentrated under vacuum. The crude oil was dissolved in anhydrous DCM (5 mL) under an Ar atmosphere and the solution was cooled to  $-78^\circ\text{C}$ . Freeze-thaw degassing cycles were performed.  $\text{Et}_3\text{SiH}$  (0.64 mL, 4.0 mmoles) and  $\text{BF}_3\cdot\text{Et}_2\text{O}$  (0.5 mL, 4.0 mmoles) were successively added. After 6 hours the reaction was quenched at  $-78^\circ\text{C}$  by addition of a saturated solution of  $\text{NaHCO}_3$ . The organics were extracted with diethyl ether, washed with a saturated solution of  $\text{NaHCO}_3$  then brine, and dried over  $\text{MgSO}_4$ . The residue was purified by column chromatography ( $\text{SiO}_2$ , hexanes/DCM 100/0 $\rightarrow$ hexanes/DCM 0/100) affording 199 mg (28%) of a colorless oil.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ ): 7.37 (s,

1H), 5.33 (dd, 1H, J = 8.5, 5.5, H-1'), 4.41 (app dd, 1H, J = 14.5, 7.0, H-3'), 4.13-4.03 (m, 2H, H-5'), 3.82-3.79 (m, 1H, H-4'), 2.57-2.52 (m, 1H, H-2'  $\alpha$ ), 2.48 (s, 3H, CH<sub>3</sub>), 2.34 (s, 3H, CH<sub>3</sub>), 1.98-1.93 (m, 1H, H-2'  $\beta$ ), 1.10-0.98 (m, 28H, 4 x -CH(CH<sub>3</sub>)<sub>2</sub>); <sup>13</sup>C NMR (CDCl<sub>3</sub>): 139.2, 134.7, 134.0, 134.0, 129.5, 125.4, 85.0, 75.4, 70.7, 61.9, 41.0, 20.9, 18.1, 17.5, 17.4, 17.4, 17.4, 17.2, 17.1, 17.0, 13.5, 13.2, 12.8, 12.5; HR-ESI-MS: calcd for [C<sub>25</sub>H<sub>42</sub>O<sub>4</sub>Cl<sub>2</sub>Si<sub>2</sub>+Na]<sup>+</sup> 555.1896, found 555.1911.

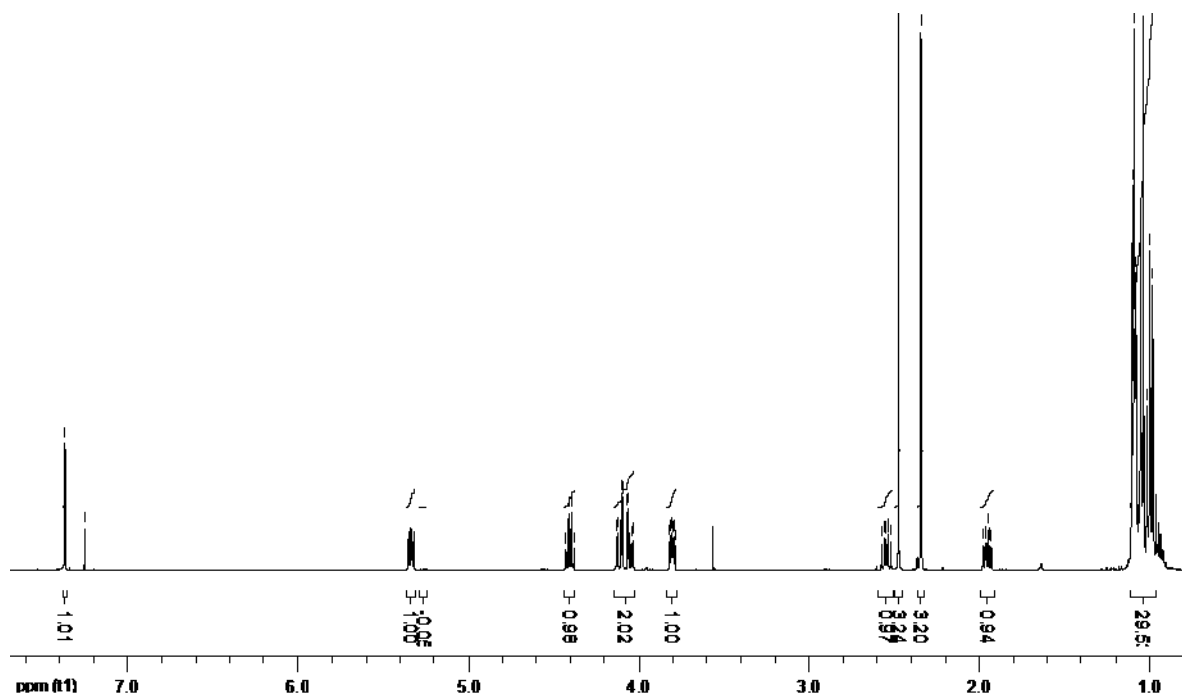


Figure S20. <sup>1</sup>H NMR spectrum of compound 12.

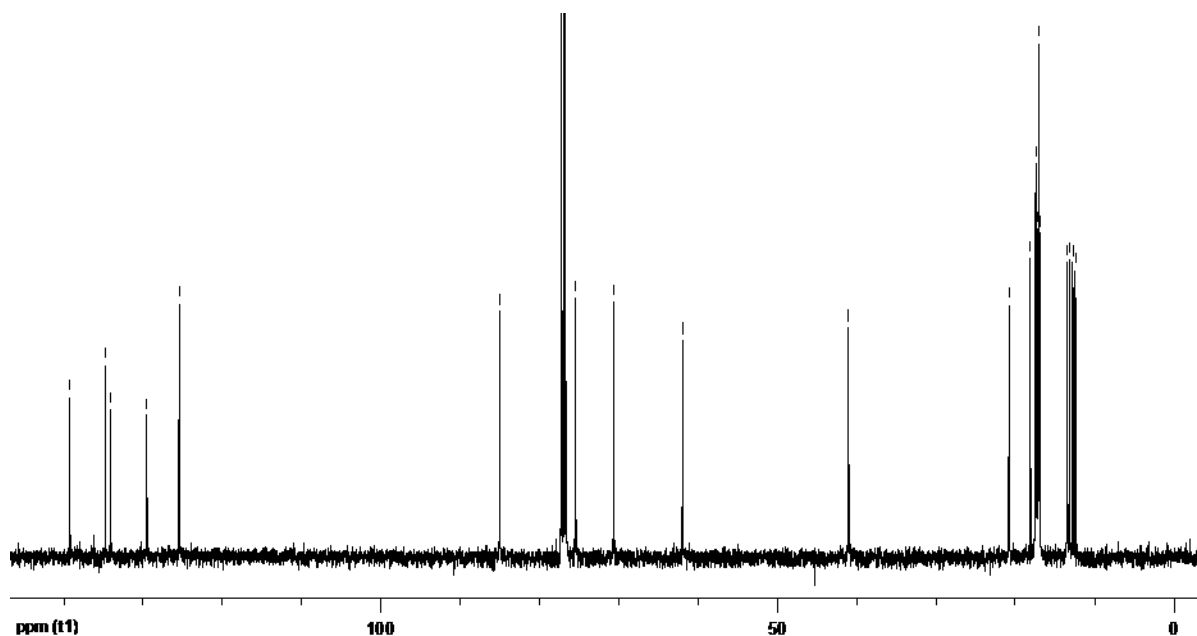


Figure S21. <sup>13</sup>C NMR spectrum of compound 12.

Compound 13. Compound 12 (199 mg, 0.37 mmoles) was dissolved in anhydrous THF (6.5 mL) under an Ar atmosphere and cooled to 0°C. TBAF (1.12 mL of a 1 M solution in THF, 1.12 mmoles) was added and the reaction mixture was stirred at room temperature. After 3 h the reaction was quenched by addition of a saturated solution of NaHCO<sub>3</sub>. The organics were extracted twice with EtOAc, dried over MgSO<sub>4</sub>, and concentrated in vacuum. The residue

was purified by column chromatography (SiO<sub>2</sub>, DCM→DCM/MeOH 95/5) which afforded 110 mg (quant.) of a white solid. <sup>1</sup>H NMR (CD<sub>3</sub>OD): 7.49 (s, 1H), 5.39 (dd, 1H, J = 10.0, 6.0, H-1'), 4.29 (m, 1H), 3.94 (m, 1H), 3.71 (m, 2H), 2.46 (s, 3H, CH<sub>3</sub>), 2.44-2.42 (m, 1H), 2.35 (s, 3H, CH<sub>3</sub>), 1.73-1.67 (m, 1H); <sup>13</sup>C NMR (CD<sub>3</sub>OD): 140.0, 136.3, 134.9, 134.9, 130.7, 127.0, 88.9, 78.3, 74.1, 63.8, 43.3, 20.9, 18.2; HR-ESI-MS: calcd for [C<sub>13</sub>H<sub>16</sub>O<sub>3</sub>Cl<sub>2</sub>+Na]<sup>+</sup> 313.0374, found 313.0381.

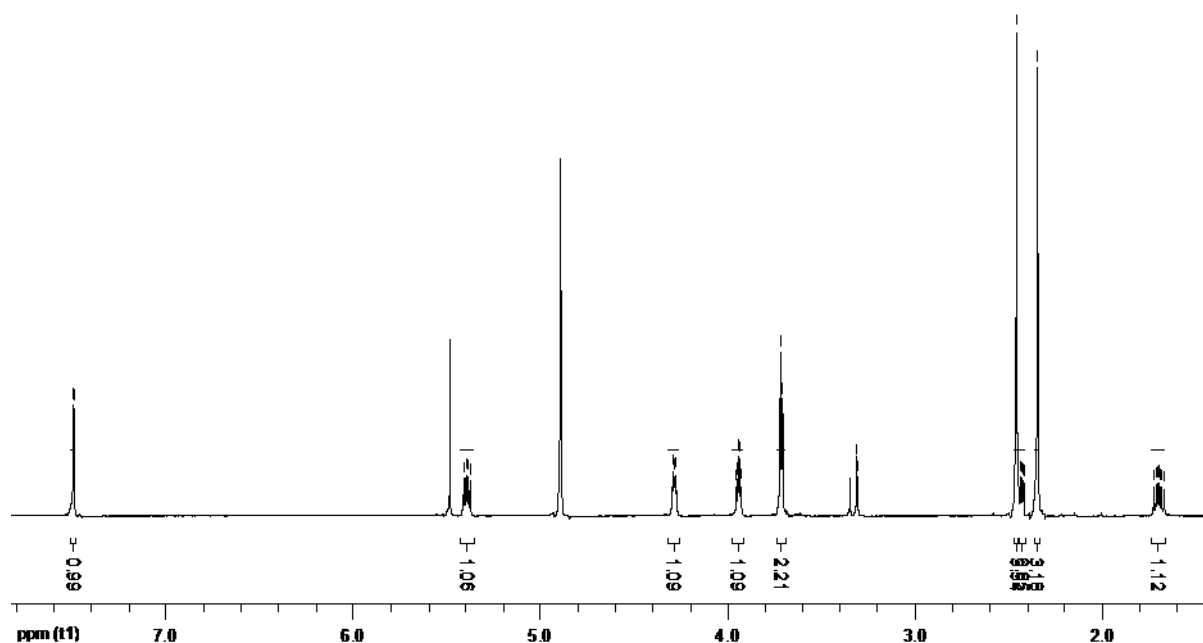


Figure S22. <sup>1</sup>H NMR spectrum of compound 13.

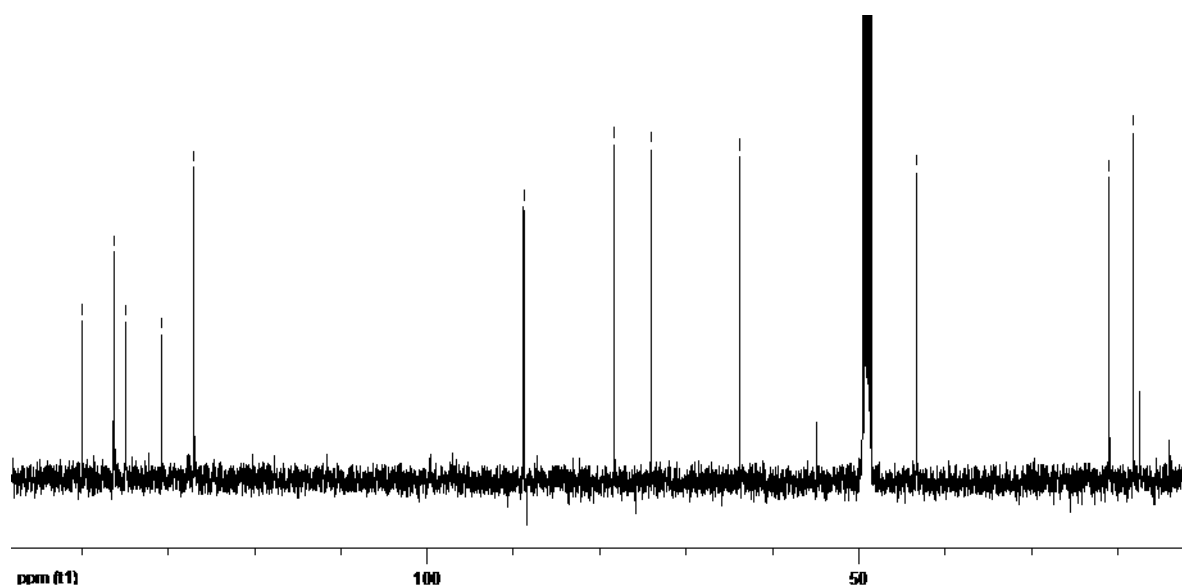


Figure S23. <sup>13</sup>C NMR spectrum of compound 13.

Compound **14**. Compound **13** (110 mg, 0.37 mmoles) was dissolved in anhydrous pyridine (2.5 mL) under an Ar atmosphere. DIPEA (0.1 mL, 0.56 mmoles) was added. A solution of DMT-Cl (158 mg, 0.47 mmoles) in anhydrous DCM (2.5 mL) was slowly added and the reaction mixture was stirred at room temperature overnight. The reaction was then quenched by addition of methanol and the solvents were removed in vacuum. The residue was purified by column chromatography (SiO<sub>2</sub> neutralized with 5% Et<sub>3</sub>N, Hexanes/EtOAc/Et<sub>3</sub>N 30/10/1) to afford 140 mg (63%) of product **14**. <sup>1</sup>H NMR (CDCl<sub>3</sub>): 7.50-7.47 (m, 3H), 7.39-7.36 (m, 4H), 7.30-7.20 (m, 3H), 6.83 (dd, 4H, J = 9.0, 1.5), 5.46 (dd, 1H, J = 9.6, 6.0), 4.38 (m, 1H), 4.08 (m, 1H), 3.79 (s, 6H), 3.38 (m, 2H), 2.49 (s, 3H), 2.49-2.45 (m, 1H), 2.26 (s, 3H), 1.91-1.85 (m, 1H); <sup>13</sup>C NMR (CDCl<sub>3</sub>): 158.4, 144.8, 138.4, 136.0, 135.9, 134.8, 134.0, 134.0, 130.1, 128.1, 127.8, 126.8, 125.5,

113.1, 86.2, 85.8, 74.1, 64.1, 55.2, 46.0, 42.4, 20.8, 18.1; HR-ESI-MS: calcd for  $[C_{34}H_{34}O_5Cl_2+Na]^+$  615.1676, found 615.1672.

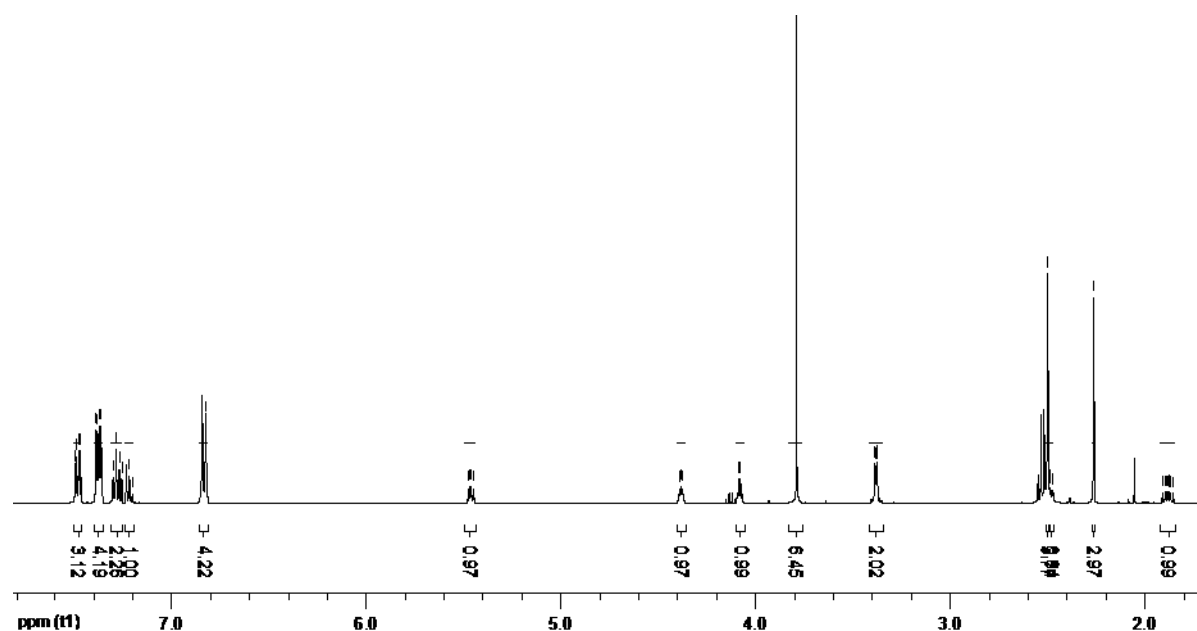


Figure S24.  $^1H$  NMR spectrum of compound **14**.

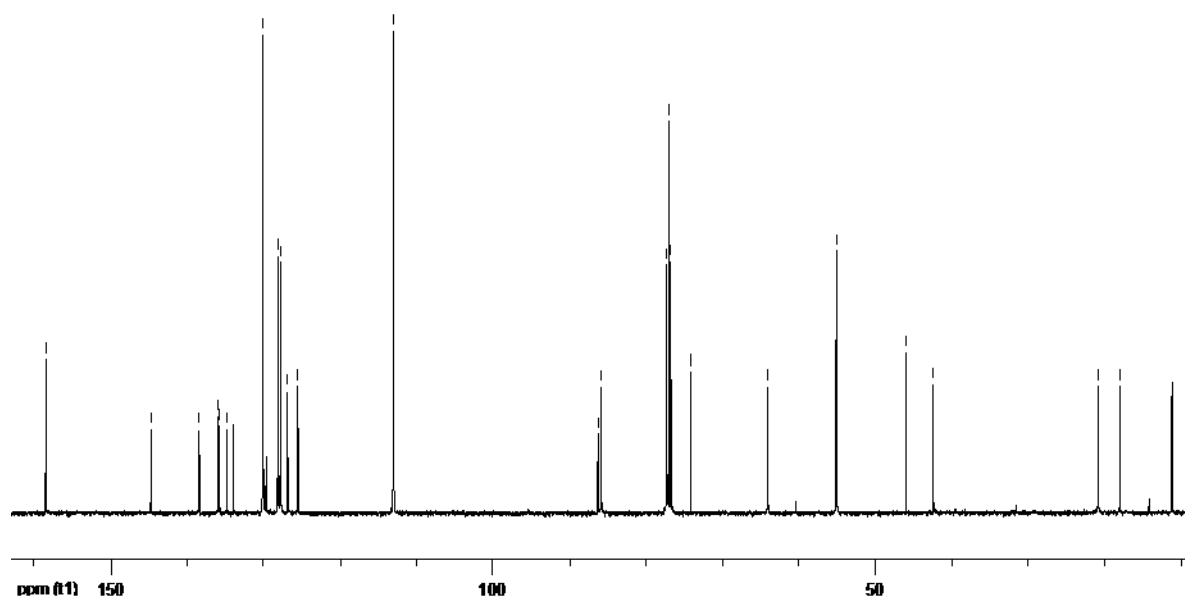


Figure S25.  $^{13}C$  NMR spectrum of compound **14**.

Compound **15**. Compound **14** (60 mg, 0.10 mmoles) was dissolved in anhydrous DCM (1.0 mL) under an Ar atmosphere. DIPEA (70  $\mu$ L, 0.40 mmoles), then 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite (38  $\mu$ L, 0.15 mmoles) were added and the reaction mixture was stirred at room temperature for 3 h, then concentrated in vacuo. The residue was purified by column chromatography ( $SiO_2$  neutralized with 5%  $Et_3N$ , Hexanes/ $EtOAc/Et_3N$  60/20/1) to afford 43 mg (54%) of product **15**.  $^1H$  NMR ( $CDCl_3$ ): 7.52-7.46 (m, 3H), 7.39-7.34 (m, 4H), 7.30-7.20 (m, 3H), 6.84-6.80 (m, 4H), 5.47-5.42 (m, 1H, H1'), 4.52-4.43 (m, 1H, H3'), 4.21 (m, 1H, H4'), 3.97-3.77 (m, 1H), 3.79 (app d, 6H, J = 3.2, OCH<sub>3</sub>), 3.68-3.53 (m, 3H), 3.44-3.36 (m, 1H), 3.30-3.24 (m, 1H), 2.67-2.61 (m, 2H), 2.50 (app d, J = 3.2, 3H, ArCH<sub>3</sub>), 2.45 (m, 1H), 2.25 (app d, J = 4.4, 3H, ArCH<sub>3</sub>), 1.88-1.80 (m, 1H, H2'), 1.20-1.15 (m, 8H), 1.15-1.04 (m, 4H);  $^{31}P$  NMR: 149.5, 148.7; HR-ESI-MS: calcd for  $[C_{43}H_{51}O_6Cl_2N_2P+Na]^+$  815.2754, found 815.2748.

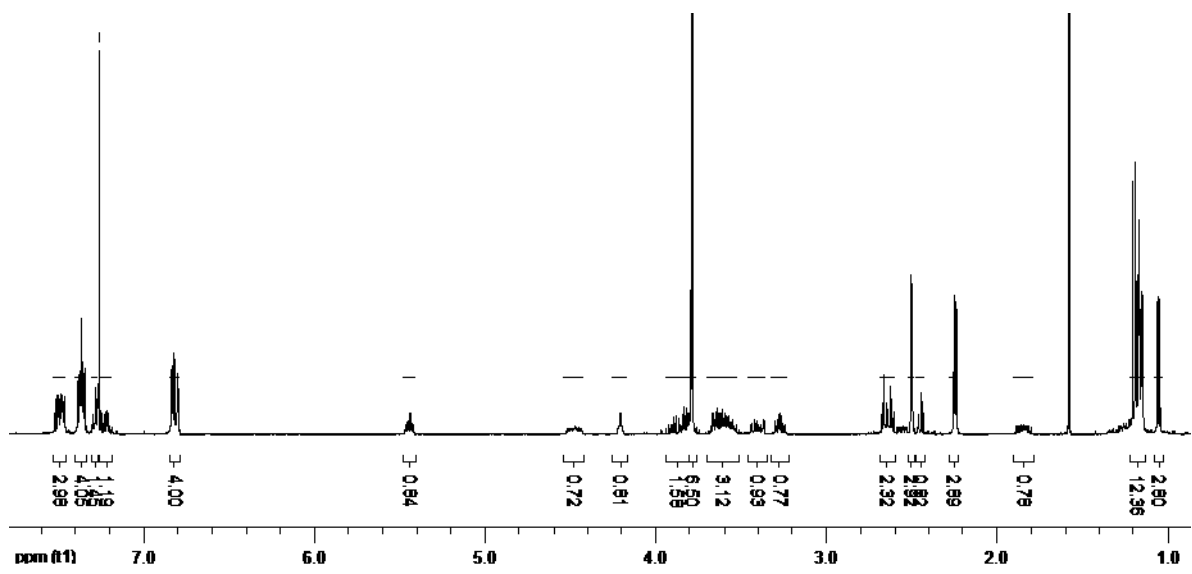


Figure S26. <sup>1</sup>H NMR spectrum of compound 15.

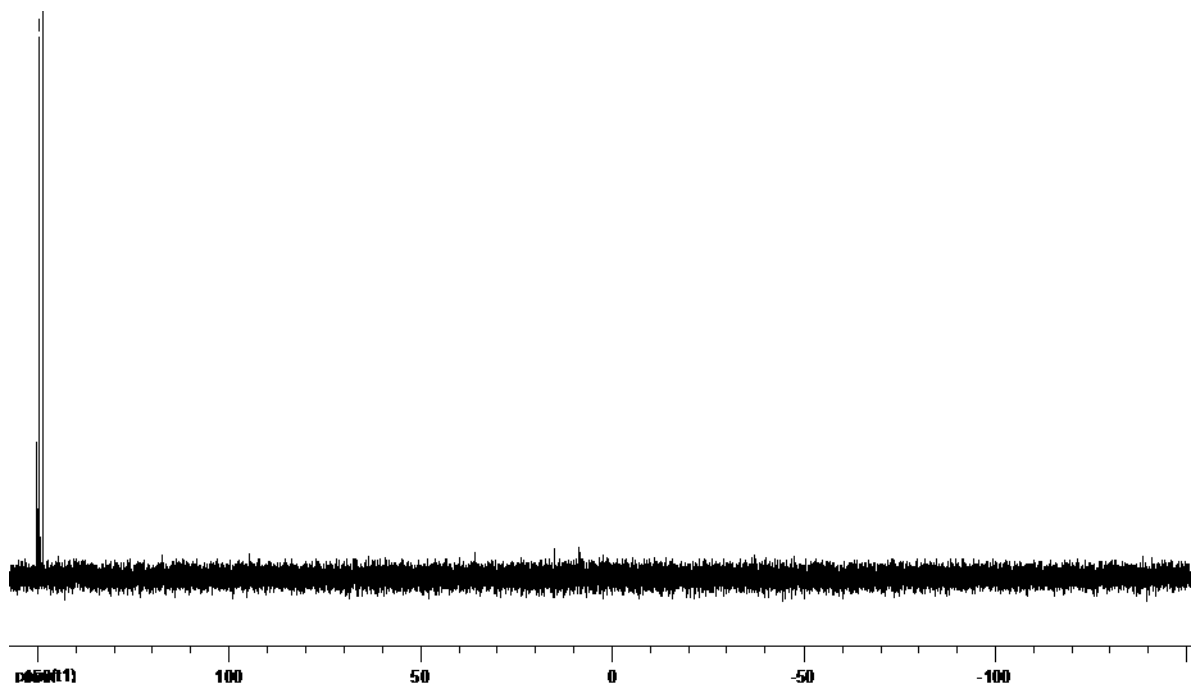


Figure S27. <sup>31</sup>P NMR spectrum of compound 15.

### MALDI-TOF MS characterization of DNA templates

X	Calcd	Found
Promoter	6438.3	6438.2
T	10703.0	10703.4
H	10669.0	10668.2
F	10705.0	10705.4
L	10737.9	10738.7
B	10826.8	10827.7
I	10920.8	10920.6
Me <sub>L</sub>	10751.9	10751.0

#### Promoter

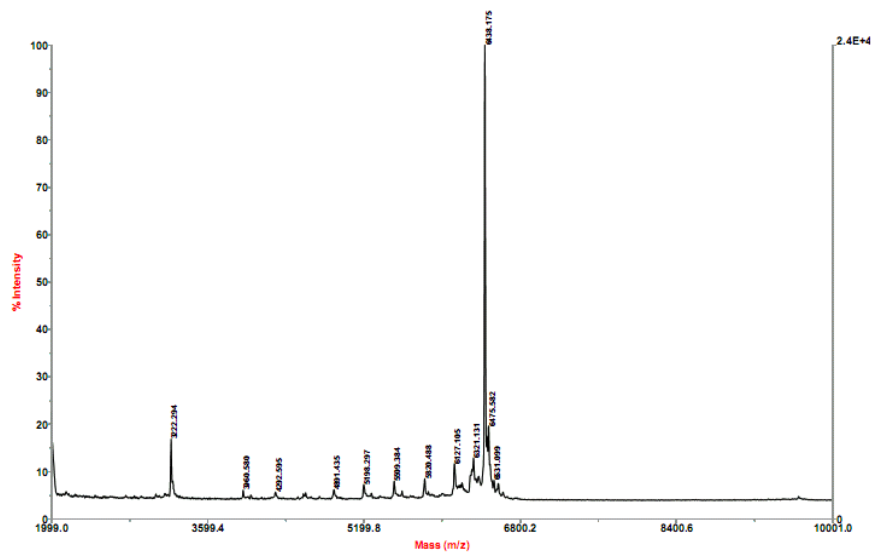


Figure S28. MALDI-TOF MS of the promoter DNA.

T



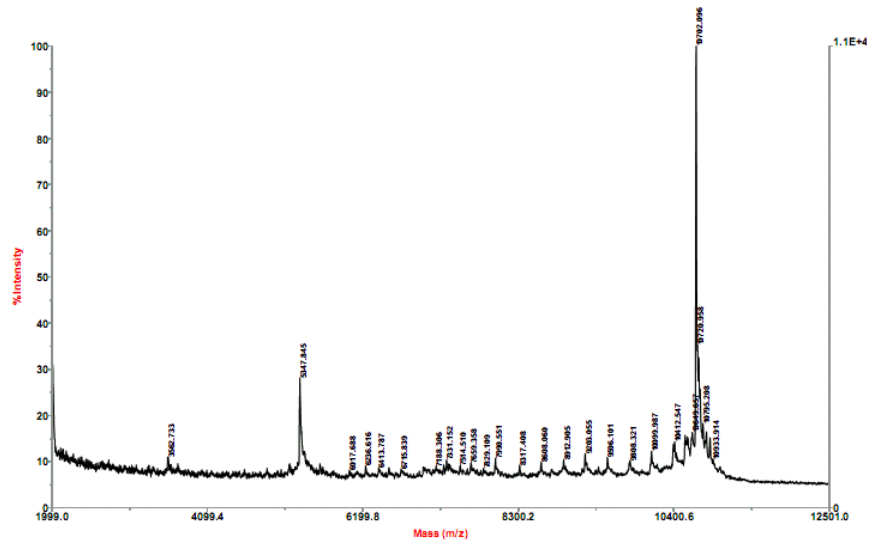


Figure S29. MALDI-TOF MS of the DNA template X=T.

H

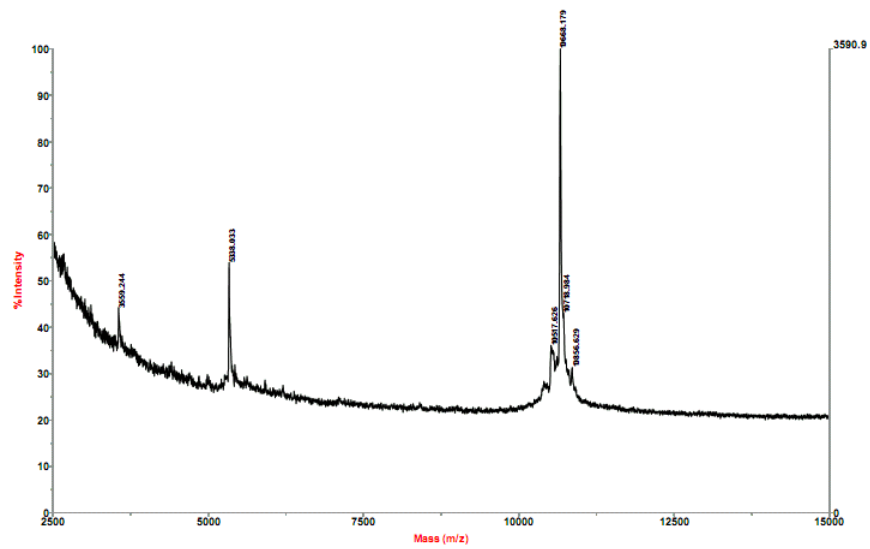


Figure S30. MALDI-TOF MS of the DNA template X=H.

F

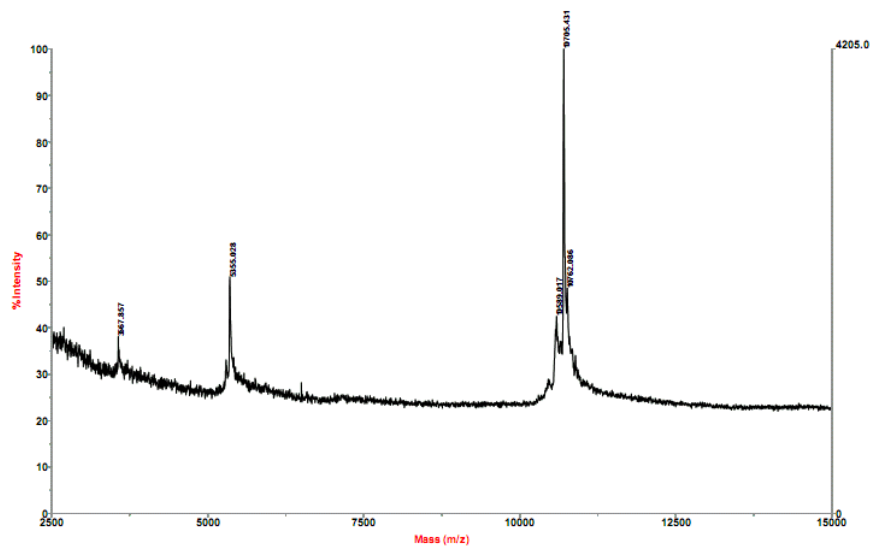


Figure S31. MALDI-TOF MS of the DNA template X=F.

L

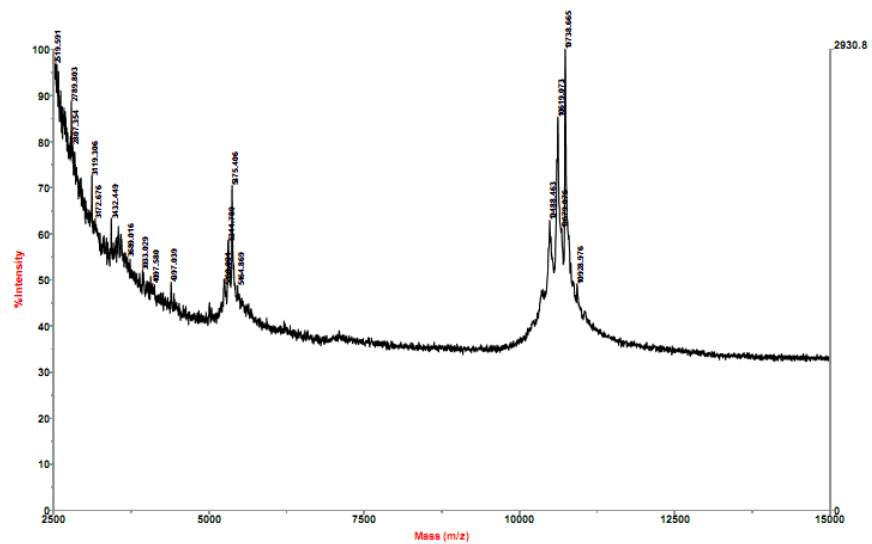


Figure S32. MALDI-TOF MS of the DNA template X=L.

B



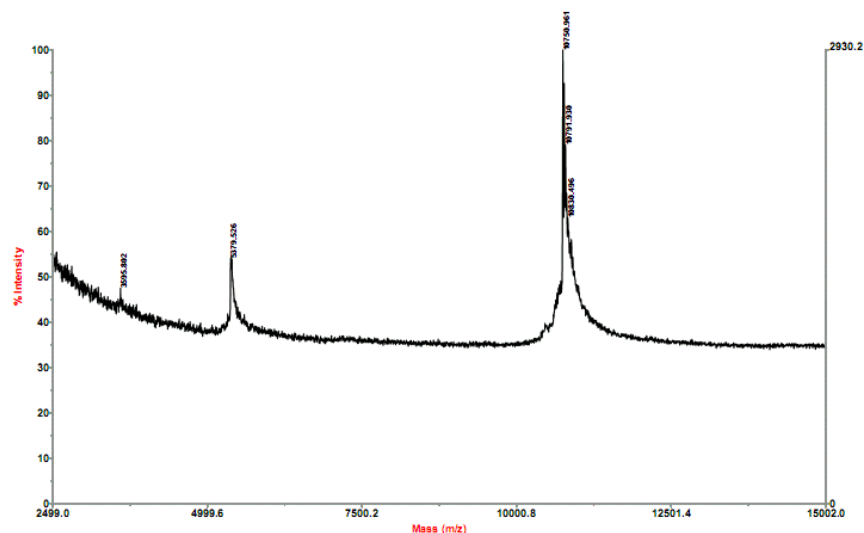


Figure S35. MALDI-TOF MS of the DNA template X=<sup>Me</sup>L.

### Enzymatic transcription reactions at low product-conversion

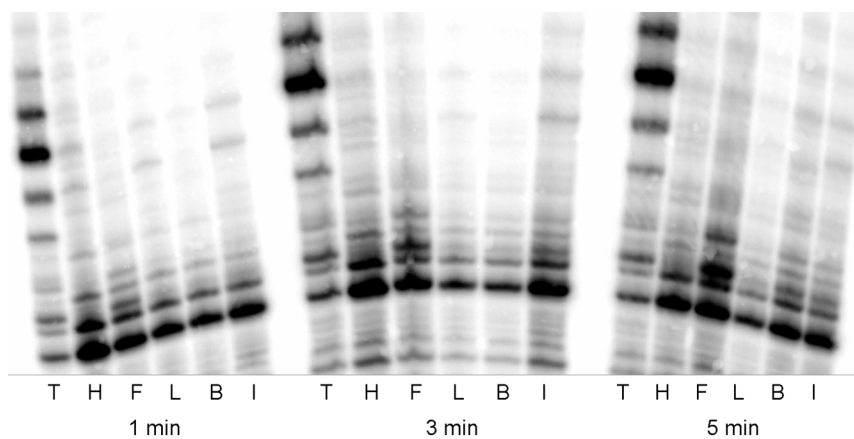


Figure S36. 20% PAGE of enzymatic transcription reaction with different DNA templates (T, H, F, L, B, and I) with different reaction times (1, 3, and 5 mins)

### Enzymatic transcription reactions: comparison of templates L and <sup>Me</sup>L

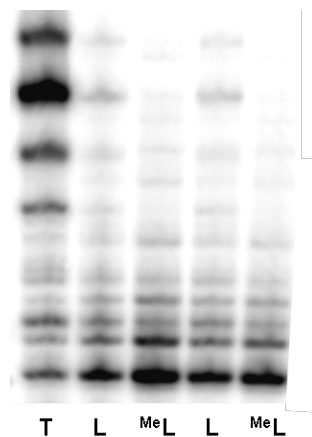


Figure S37. 20% PAGE of enzymatic transcription reaction with different DNA templates (T, L, and <sup>Me</sup>L). Reaction time: 30 mins.

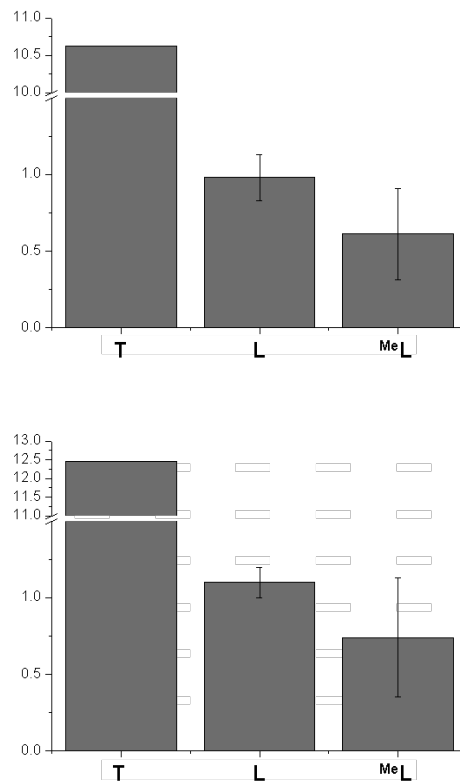


Figure S38. Comparison of the transcription efficiencies of insertion (bottom) and extension (top) between different DNA templates (X= T, L, and MeL). Data obtained from equation (1) where  $m=N'-1$  (insertion) or  $N'$  (extension). Reaction time: 30 mins (2 replicates). Error bars represent one standard deviation.

### Composition analysis of RNA transcripts: control experiment

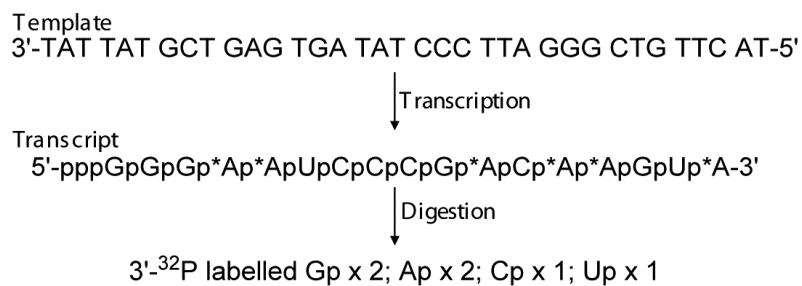
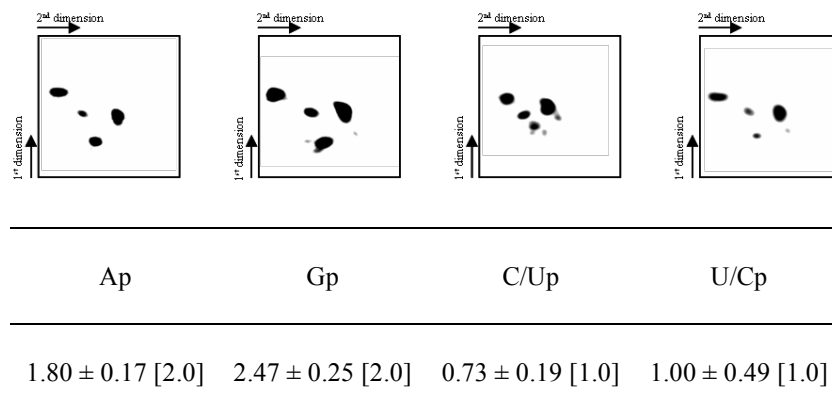


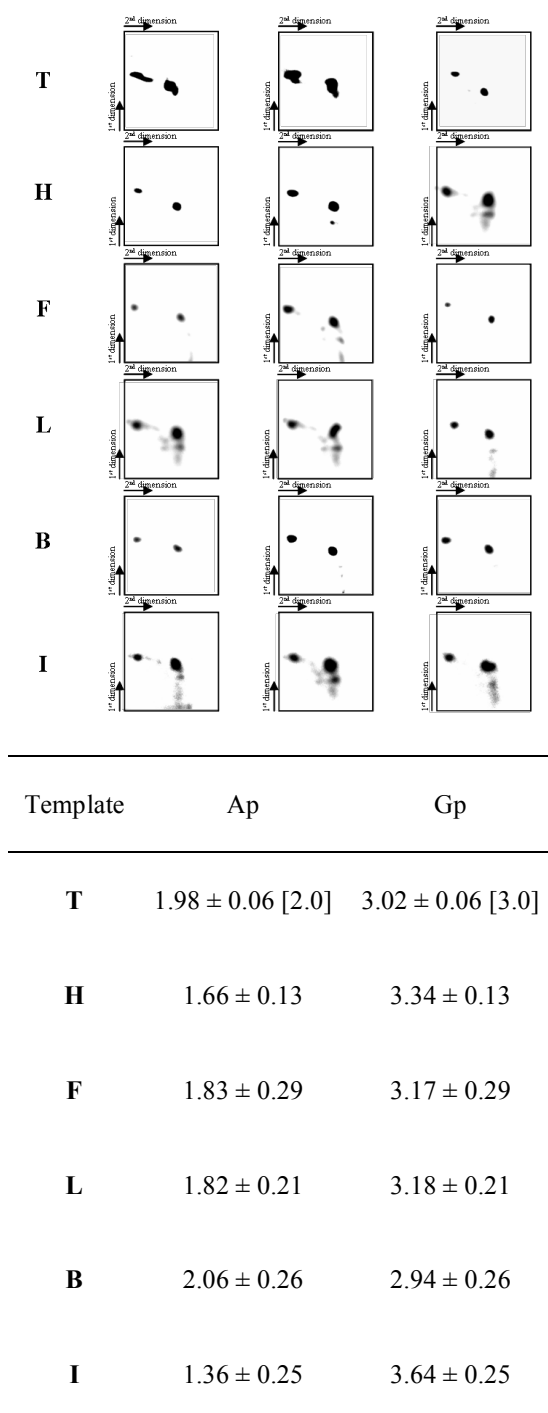
Figure S39. Sequence of a DNA template used as a control since the digestion of its corresponding full transcript yields a set of four different 3'-radiolabelled nucleoside monophosphate.



The numbers in bracket represent the theoretical value

Figure S40. 2D TLC and corresponding quantification data.

Composition analysis of RNA transcripts: experiments with different DNA templates (T, H, F, L, B, and I)



The numbers in bracket represent the theoretical value

Figure S41. 2D TLC and corresponding quantification data.

### Identification of the doubled bands seen in the gel electrophoresis

To test whether the presence of two bands is the result of a misincorporation triggered by the nonpolar nucleoside that was introduced within the DNA template we carried out a composition analysis of the two bands separately. The template **L** was chosen for this experiment since equal amounts of both bands were seen in the gel electrophoresis (see Figure 3 in the main text). The two N<sup>+</sup>+1 bands were excised and digested separately as described in the experimental section. The results of the 2D-TLC analyses show no significant differences in the composition of the two bands (Figure S42).

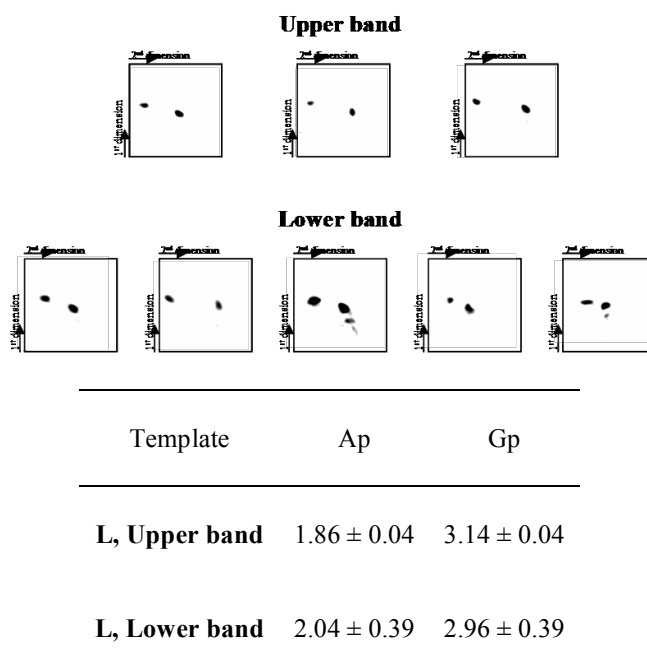


Figure S42. 2D TLC and corresponding quantification data.

### References

1. Gottlieb, H. E., Kotlyar, V., and Nudelman, A. (1997) NMR chemical shifts of common laboratory solvents as trace impurities, *J. Org. Chem.* 62, 7512-7515.
2. Kim, T. W., and Kool, E. T. (2004) A set of nonpolar thymidine nucleoside analogues with gradually increasing size, *Org. Lett.* 6, 3949-3952.
3. Kim, T. W., and Kool, E. T. (2005) A series of nonpolar thymidine analogues of increasing size: DNA base pairing and stacking properties, *J. Org. Chem.* 70, 2048-2053.
4. Hirano, K., Biju, A. T., and Glorius, F. (2009) Copper-Catalyzed Synthesis of 2-Unsubstituted, N-Substituted Benzimidazoles, *J. Org. Chem.* 74, 9570-9572.