Supporting Information

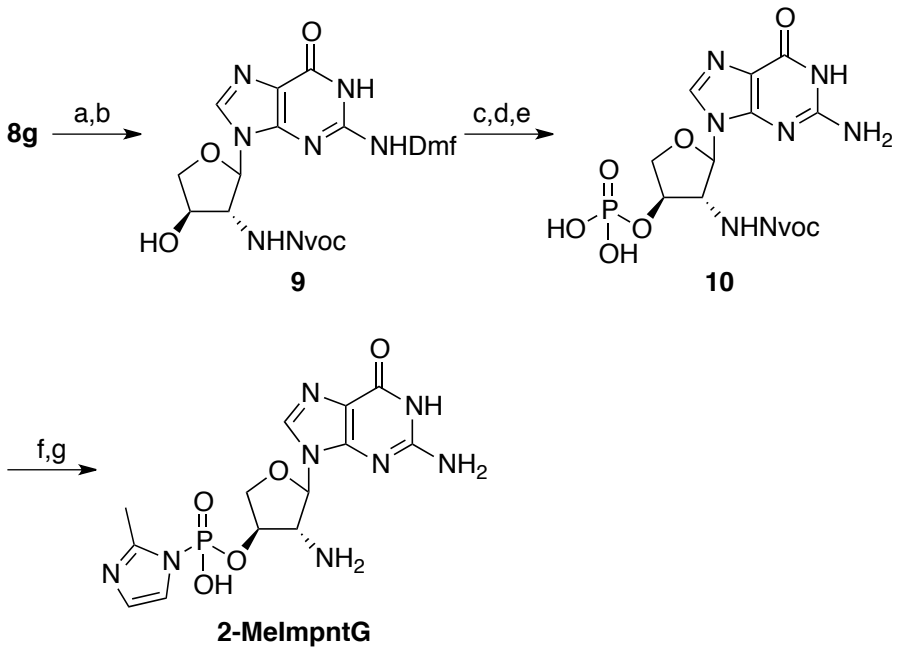
Synthesis and Non-Enzymatic Template-Directed Polymerization of 2′-Amino-2′-Deoxythreose Nucleotides

J. Craig Blain, Alonso Ricardo and Jack W. Szostak*

Howard Hughes Medical Institute and Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114, USA

*Email: szostak@molbio.mgh.harvard.edu

Scheme S1. Synthesis of the activated 2′-NH_2-TNA nucleotide 2-MeImpntG.



[a]Reagents and conditions: (a) 4,5-dimethoxy-2-nitrobenzyl chloroformate, *N*,*N*-diisopropylethylamine, DMF, 20°C, 3.5 h; (b) *N*,*N*-dimethylformamide dimethyl acetal, DMF, 20 °C, 6 h, 50%; (c) bis(2-cyanoethyl)-*N*,*N*-diisopropyl phosphoramidite, 5-(ethylthio)tetrazole, *N*,*N*-diisopropylethylamine, CH_3CN, 20 °C, 2.7 h; (d) *tert*-butyl hydroperoxide, 20 °C, 1.3 h; (e) NH_4OH, 55 °C, 18 h; (f) 2-methylimidazole, Et_3N, triphenylphosphine, 2,2′-dipyridyl disulfide, 3:2 DMSO-DMF, 20 °C, 3 h; (g) 350 nm irradiation, H_2O pH 10.5, 4 °C, 16.5 h, 6% from **9**.

Supporting Information

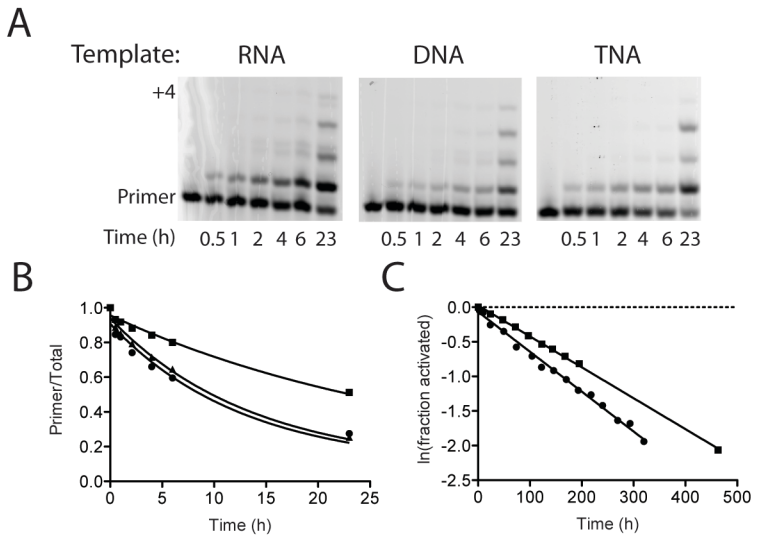


Figure S1. Non-enzymatic primer extension and hydrolysis of 2-MeImpntG in the absence of *N*-(hydroxyethyl)imidazole (HEI). (A) PAGE analysis of the products of the primer extension reaction. The conditions for the reaction are equivalent to those of Figure 4, except for the absence of HEI. (B) Plot of the amount of primer remaining as a fraction of the total lane integration against time for the RNA (•), DNA (■) and TNA (▲) templates. The plot was fit to a single exponential with a plateau fixed at 0.086, the average observed in Figure 4B, to measure pseudo-first order rate constants of 7.8×10^{-2}, 3.2×10^{-2}, and 7.4×10^{-2} h^{-1} for the RNA, DNA and TNA templates, respectively. (C) Plot of the hydrolysis of 2-MeImpntG (■) and guanosine 5′-phosphor-2-methylimidazolide (2-MeImpG) (•) measured by ^{31}P NMR with 100 mM HEPES pH 7.5, 5 mM potassium phosphate reference, and 10% D_2O at 20 °C. The natural logarithm of the fraction of activated nucleotide remaining was fit to a line; the slope yielded pseudo-first order rate constants of 4.5×10^{-3} and 5.8×10^{-3} h^{-1} for 2-MeImpntG and 2-MeImpG, respectively.

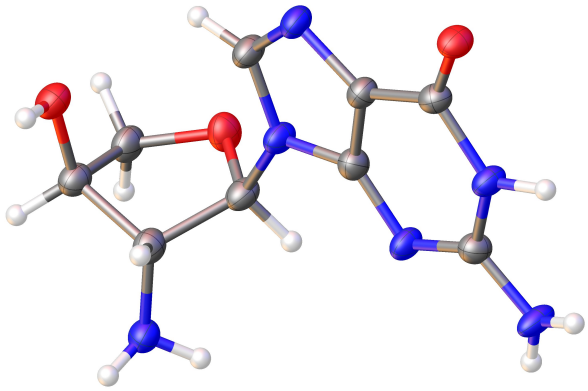


Figure S2. Thermal ellipsoid representation of the x-ray crystal structure of nucleoside **8g**.

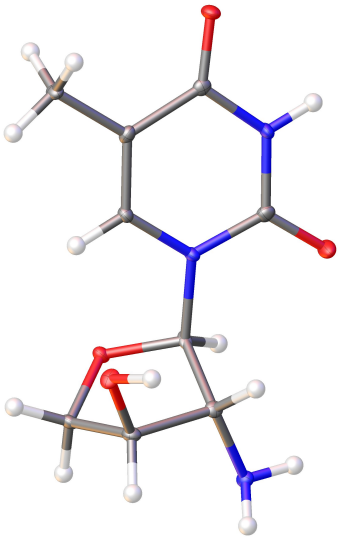


Figure S3. Thermal ellipsoid representation of the x-ray crystal structure of nucleoside **8t**.

Supporting Information

Primer extension reactions in the absence of HEI

Primer extension reactions in the absence of HEI were carried out in the same manner as in the presence of HEI except that HEI was omitted, time points were taken at 0.5, 1, 2, 4, 6 and 23 h, and the plateau for the fit of primer consumption to an exponential decay was fixed at 0.086, the average plateau observed in the presence of HEI.

Hydrolysis of 2-MeImpntG and 2-MeImpG in the absence of HEI

The 2-MeImpG monomer was synthesized as previously described[1] and purified as per 2-MeImpntG. Hydrolysis was measured under the following conditions: 3 mM 2-MeImpntG or 5 mM 2-MeImpG, 100 mM HEPES pH 7.5, 5 mM potassium phosphate reference, and 10% D_2O at 20 ^oC. Time points were taken over the course of 19 days. At each time point a ^{31}P NMR spectrum was measured and analyzed as described for hydrolysis in the presence of HEI, with the 2-MeImpG signal appearing at -10.40 ppm and the GMP signal appearing at 1.48 ppm.

*X-ray crystallography of **8t***

An analytical sample of **8t** was crystallized from water with ethanol vapor diffusion. Data was collected from a crystal mounted on a diffractometer at 100 K. The intensities of the reflections were collected by means of a Bruker APEX II CCD along with the D8 Diffractometer (30 KeV, λ = 0.413280 Å), and equipped with an Oxford Cryosystems nitrogen open flow apparatus. The collection method involved 0.5^o scans in Phi at -5^o in 2Θ. Data integration down to 0.82 Å resolution was carried

out using SAINT V7.46 A (Bruker diffractometer, 2009) with reflection spot size optimization. Absorption corrections were made with the program SADABS (Bruker diffractometer, 2009). The structure was solved by the direct methods procedure refined by least-squares methods against F^2 using SHELXS-97 and SHELXL-97 (Sheldrick, 2008). Non-hydrogen atoms were refined anisotropically, and hydrogen atoms were allowed to ride on their respective heteroatoms.

This data was collected at the Advanced Photon Source (APS). We thank Dr. Yu-Sheng Chen at ChemMatCARS, APS, for his assistance with single-crystal data. ChemMatCARS Sector 15 is principally supported by the National Science Foundation/Department of Energy under grant number NSF/CHE-0822838. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

*X-ray crystallography of **8g***

An analytical sample of **8g** was crystallized from water with ethanol vapor diffusion. Data was collected from a crystal mounted on a diffractometer at 100 K. The intensities of the reflections were collected by means of a Bruker APEX II DUO CCD diffractometer (Cu_K radiation, λ = 1.54178 Å), and equipped with an Oxford Cryosystems nitrogen flow apparatus. The collection involved 1.0° scans in ω at 30, 55, 80 and 105° in 2θ. Data analysis was carried out as per **8t**.

Synthesis of 2-MeImpntG

4,5-Dimethoxy-2-nitrobenzyl ((3*R*,4*R*)-2-(2-(((*E*)-(dimethylamino)methylene)amino)-6-oxo-1,6-dihydro-9*H*-purin-9-yl)-4-hydroxytetrahydrofuran-3-yl)carbamate (9). Nucleoside **8g** was dissolved in anhydrous DMF (10 mL) under argon. To this solution was added *N*,*N*-diisopropylethylamine (0.11 mL, 632 µmol) and then 4,5-dimethoxy-2-nitrobenzyl chloroformate (58 mg, 210 µmol) and the reaction was stirred at room temperature, protected from light, for 3.5 h. The reaction was quenched with MeOH (2 mL) for 1 h and then concentrated *in vacuo* and coevaporated with toluene (2x10 mL). The crude was dissolved in anhydrous DMF (5 mL) under argon and to this solution was added *N*,*N*-dimethylformamide dimethyl acetal (1 mL, 7.53 mmol) and the reaction was stirred, protected from light, at room temperature for 6 h. The reaction was concentrated *in vacuo* and coevaporated with toluene (2x10 mL) and the crude was purified by flash chromatography (5 to 40% MeOH in CH_2Cl_2) to afford protected nucleoside **9** as a pale yellow solid (49 mg, 50%). R_f = 0.4, 10% MeOH in CH_2Cl_2; ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.35 (s, 1H, HN-N1), 8.62 (s, 1H, H-C=N), 8.08 (d, *J* = 8.2 Hz, 1H, H-N-Nvoc), 7.98 (s, 1H, H-C8), 7.70 (s, 1H, H-Ar), 7.15 (s, 1H, H-Ar), 5.82 (d, *J* = 3.9 Hz, 1H, H-C1′), 5.73 (d, *J* = 3.8 Hz, 1H, H-OC3′), 5.35 (d, *J* = 15.0 Hz, 1H, H-HC-Ar), 5.31 (d, *J* = 14.4 Hz, 1H, H-HC-Ar), 4.59 (m, 1H, H-C2′), 4.31 (m, 1H, H-C3′), 4.16-3.96 (m, 2H, H-C4′), 3.87 (s, 6H, H_3-COAr), 3.10 (s, 3H, H_3-C-N), 3.02 (s, 3H, H_3-C-N); ESI-MS (m/z): $[M+Na^+]^+$ calc. for $C_{22}H_{26}N_8O_9$, 569.2, obs. 569.1.

4,5-Dimethoxy-2-nitrobenzyl ((3*R*,4*R*)-2-(2-amino-6-oxo-1,6-dihydro-9*H*-purin-9-yl)-4-(phosphonooxy)tetrahydrofuran-3-yl)carbamate (10). Protected nucleoside **9** was dissolved in anhydrous acetonitrile (8 mL) with 3 Å

molecular sieves (10) under argon. To this solution was added 5-(ethylthio)tetrazole (20 mg, 154 μmol) and then bis(2-cyanoethyl)-*N*,*N*-diisopropyl phosphoramidite (28 mg, 103 μmol) and the reaction was stirred at room temperature for 2.7 h protected from light. To the reaction was added a 70% aqueous solution of *tert*-butyl hydroperoxide (200 μLμ, 1.45 mmol) and the reaction was stirred for another 1.3 h. Finally, the reaction was diluted with ammonium hydroxide (15 mL) and stirred at 55 °C in a pressure vessel for 18 h. The reaction was concentration *in vacuo* and purified by preparative HPLC with 2 to 15% acetonitrile in 10 mM triethylammonium bicarbonate pH 7.5 over 15 min to afford 90 mg of the protected nucleotide **10** as a triethylammonium salt after lyophilization. Due to excess triethylammonium the actual yield of the reaction was not determined and the product was used directly in the subsequent activation reaction. ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.63 (s, 1H, H-N1), 8.35 (d, *J* = 6.9 Hz, 1H, H-N-Nvoc), 7.99 (s, 1H, H-C8), 7.69 (s, 1H, H-Ar), 7.15 (s, 1H, H-Ar), 5.73 (d, *J* = 4.8 Hz, 1H, H-C1′), 5.33 (d, *J* = 14.7 Hz, 1H, H-HC-Ar), 5.28 (d, *J* = 14.3 Hz, 1H, H-HC-Ar), 4.69 (m, 1H, H-C2′), 4.54 (m, 1H, H-C3′), 4.12 (dd, *J* = 9.1, 4.4 Hz, H-C4′), 4.07 (dd, *J* = 9.1, 5.2 Hz, H-C4′), 3.88 (s, 3H, H_3-COAr), 3.86 (s, 3H, H_3-COAr); ^{13}P NMR (160.8 MHz, DMSO-d_6) δ 0 (phosphoric acid), -1.12; ESI-MS (m/z): $[M\text{-}H^+]^-$ calc. for $C_{19}H_{22}N_7O_{12}P$, 570.1, obs. 570.1.

9-(2′-Amino-2′-deoxy-α-L-threofuranosyl)guanine 3′-phosphor-2-methylimidazolide (2-MeImpntG). Protected nucleotide **10** was dissolved in a mixture of DMSO (3 mL) and DMF (2 mL) under argon and protected from light. To this solution was added 2-methylimidazole (181 mg, 2.21 mmol), triethylamine (0.1 mL,

717 μmol), triphenylphosphine (129 mg, 492 μmol), and finally 2,2′-dipyridyl disulfide (150 mg, 681 μmol) and the reaction was stirred at room temperature for 3 h. The reaction was precipitated in a mixture of acetone (300 mL), diethyl ether (190 mL), triethylamine (20 mL) and an acetone solution saturated with sodium perchlorate (1.5 mL). The precipitate was collected by centrifugation and washed with a 1:1 acetone-ether mixture (3x40 mL). The crude pellet was dissolved in water brought to pH 10.5 with 10 N NaOH. This solution was stirred under constant 350 nm irradiation for 16.5 h. The product was purified directly from the reaction solution by preparative HPLC with 2 to 9% acetonitrile in 10 mM triethylammonium bicarbonate pH 7.5 over 12 min to afford the product as a triethylammonium salt after lyophilization. The yield of activated nucleotide 2-MeImpntG was measured by the UV spectrophotometry, assuming an extinction coefficient of 14.09 mM^{-1} cm^{-1}, and found to be 4 μmol (6% from **9**). ^{1}H NMR (400 MHz, D_2O) δ 7.90 (s, 1H, H-C8), 7.06 (app s, 1H, H-Im), 6.82 (app s, 1H, H-Im), 5.77 (app s, 1H, H-C1′), 4.59 (under DHO, H-C3′), 4.47 (d, *J* = 11.1 Hz, 1H, H-C4′), 4.36 (d, *J* = 10.9 Hz, 1H, H-C4′), 3.70 (app s, 1H, H-C2′), 3.21 (q, *J* = 7.3 Hz, 0.6x6H, $NH(CH_2CH_3)_3^+$), 2.20 (s, 3H, H_3-C), 3.21 (t, *J* = 7.3 Hz, 0.6x9H, $NH(CH_2CH_3)_3^+$); ^{13}C NMR (100 MHz, DMSO-d_6) δ 160.1, 158.8, 153.6, 150.3, 137.0, 121.2, 90.9, 80.2, 74.6, 61.4, 13.7; ^{31}P NMR (160.8 MHz, 10% D_2O) δ 1.15 (hydrolyzed), 0 (orthophosphate), -12.02 (activated); HRMS (m/z): $[M-H^+]^-$ calc. for $C_{13}H_{17}N_8O_5P$, 395.0987, obs. 395.0972.

References

(1) Joyce, G. F.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1984**, *176*, 279–306.