N2′→P3′ phosphoramidate glycerol nucleic acid as a potential alternative genetic system

Jesse J. Chen[†], Xin Cai[†], 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

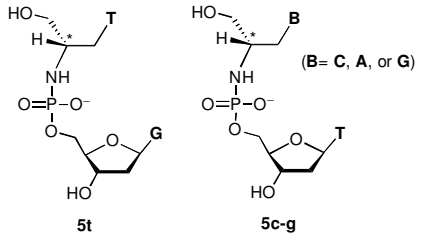
[*]E-mail: szostak@molbio.mgh.harvard.edu

Supporting Table S1. MALDI-TOF MS analysis of npGNA or GNA oligonucleotides.

Sequence*	Underlined Sequence	Calculated	Observed
3′-TTT TTT TTT T-2′-T	npGNA	2854.953	2855.66
3′-p-AAA AAA AAA -2′-T	npGNA	2754.880	2755.54
3′-CGT ACG ACA T -2′-T	npGNA	2886.985	2886.43
3′-ATG TCG TAC G-2′-T	npGNA	2917.996	2917.21
3′-CGA ATT CG -2′-T	npGNA	2370.639	2375.61
T-3′-TTT TTT TTT T-2′-T	GNA	3168.994	3169.94
T-3′-AAA AAA AAA A-2′-T	GNA	3259.128	3259.82
T-3′-CGT ACG ACA T -2′-T	GNA	3201.026	3201.33
T-3′-ATG TCG TAC G -2′-T	GNA	3232.037	3233.25
5′-GTG TAG AGT GAT TGG-3′-NH_2	n/a	4703.143	4701.59
3′-TTT TTT CCA ATC ACT CTA CAC-3′	GNA	6014.191	6014.53
3′-AAA AAA CCA ATC ACT CTA CAC-3′	GNA	6068.997	6068.84

* The non-underlined sequences are DNA.

Supporting Figures



	^{31}P NMR δ: (ppm)
5t	7.69
5c	8.00
5g	7.51
5a	7.56

Figure S1. Structures and ^{31}P-NMR chemical shifts of dinucleotide model compounds containing N2′→P3′ phosphoramidate bonds.

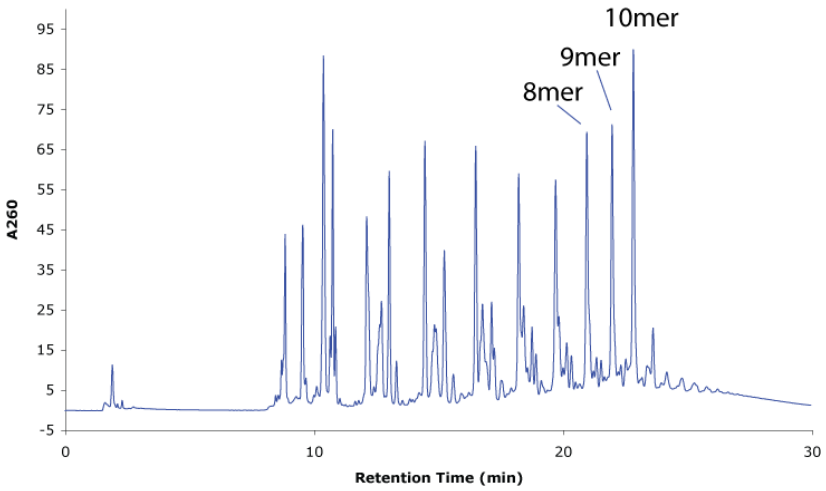


Figure S2. Anion-exchange HPLC analysis of the crude product of npGNA synthesis. Sequence: 3′-TTT TTT TTT T-2′-T, the underlined denotes the npGNA sequence. The two truncated by-products were identified by MALDI-TOF MS analysis. For the 9mer, [M+H]+ calcd: 2593.782, obsd:2594.17. For the 8mer, [M+H]+ calcd: 2332.610, obsd:2332.14.

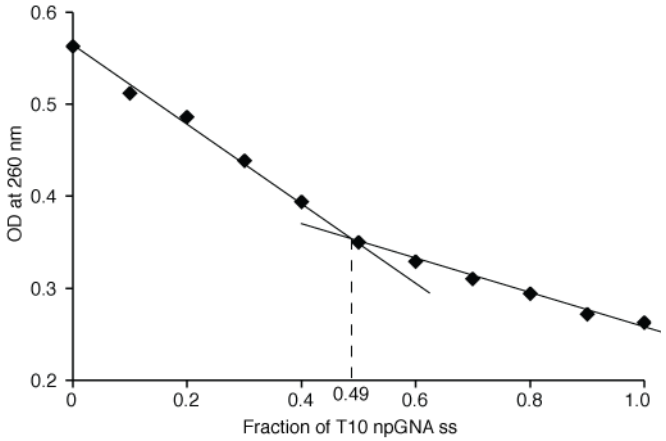


Figure S3. Mixing curve showing 1:1 stoichiometry in the 3′-AAA AAA AAA-2′-T/3′-TTT TTT TTT T-2′-T npGNA complex (the underlined portion of the sequence is npGNA, the non-underlined bases are DNA). In each sample, the sum of the concentrations of two npGNA oligomers was 4 μM.

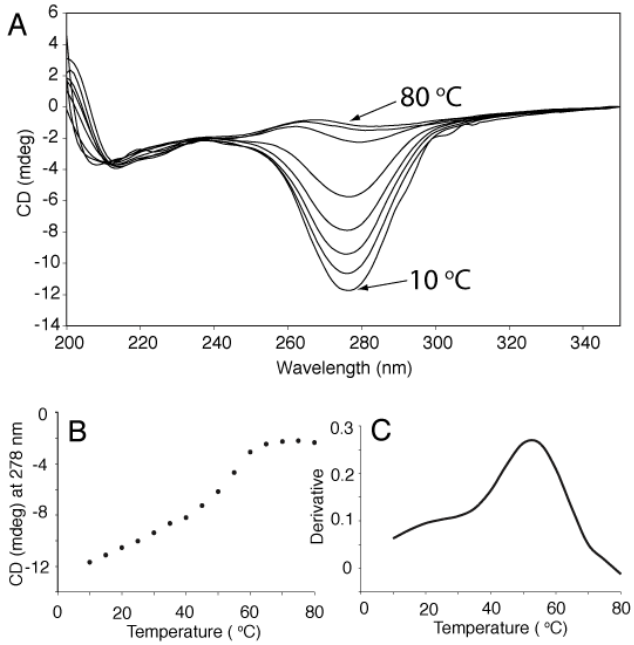


Figure S4. Circular dichroism studies on an npGNA duplex with the sequence 3′-CGT ACG ACA T-2′-T / T-2′-GCA TGC TGT A-3′ (the underlined portion of the sequence is npGNA, the non-underlined bases are DNA). The concentration of the npGNA duplex was 10 μM. **A**. CD spectra obtained in the temperature range of 10-80°C at 10°C intervals. **B**. Temperature-dependent CD signal change monitored at 278 nm. **C.** Plot of derivative of temperature-dependent CD signal change in **B**.

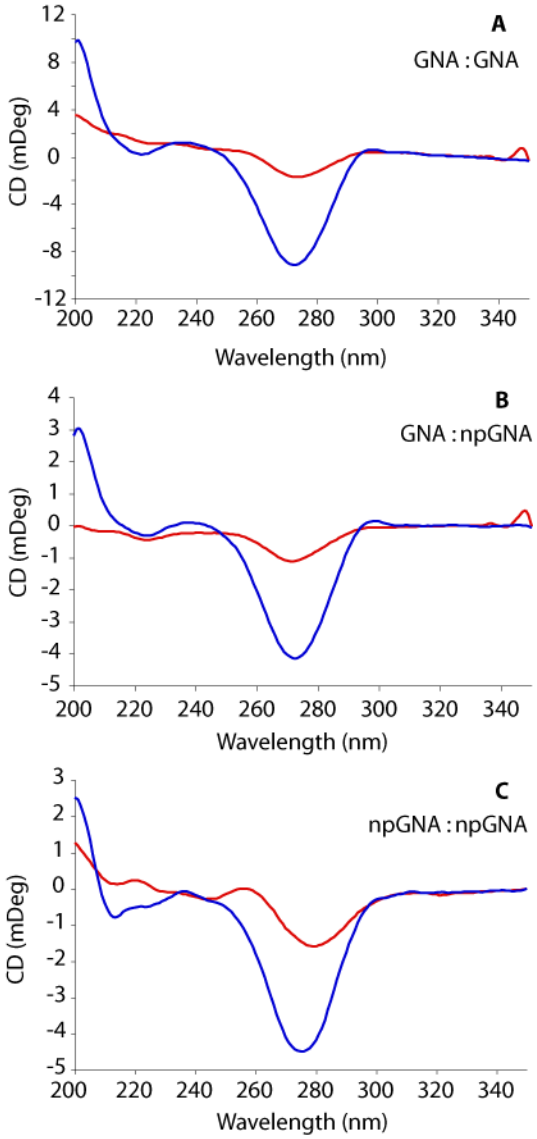


Figure S5. Circular dichroism studies showing structural similarity among various GNA and npGNA duplexes. The core sequences are the same for the three duplexes with a single thymidine at the 2′-terminus of npGNA or at 3′- and 2′-termini of GNA. **A.** GNA homoduplex with the sequence: T-3′-CGT ACG ACA T-2′-T / T-2′-GCA TGC TGT A-3′-T. **B**. GNA:npGNA heteroduplex with the sequence: T-3′-CGT ACG ACA T-2′-T (GNA) / T-2′-GCA TGC TGT A-3′ (npGNA). **C.** npGNA homoduplex with the sequence: T-3′-CGT ACG ACA T-2′-T / T-2′-GCA TGC TGT A-3′-T. (the underlined denotes the GNA or npGNA sequences). The concentration of each GNA or npGNA strand was 10 μM. The red traces are the sum of signals from single-stranded components. The blue traces are from duplexes.

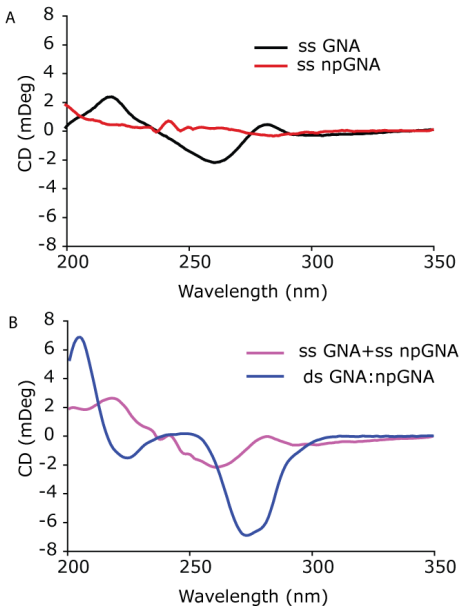


Figure S6. CD studies on T-3′-AAA AAA AAA A-2′-T(GNA): T-2′-TTT TTT TTT T-3′(npGNA) showing heteroduplex formation. (The underlined denotes the GNA or npGNA sequences). The concentration of each GNA or npGNA strand was 10 μM. **A.** CD signals of single-stranded (ss) A10 GNA (black trace) and T10 npGNA (red trace). **B.** CD signal of A10:T10 GNA:npGNA heteroduplex (blue trace) compared with the sum of CD signals of each strand (magenta trace).

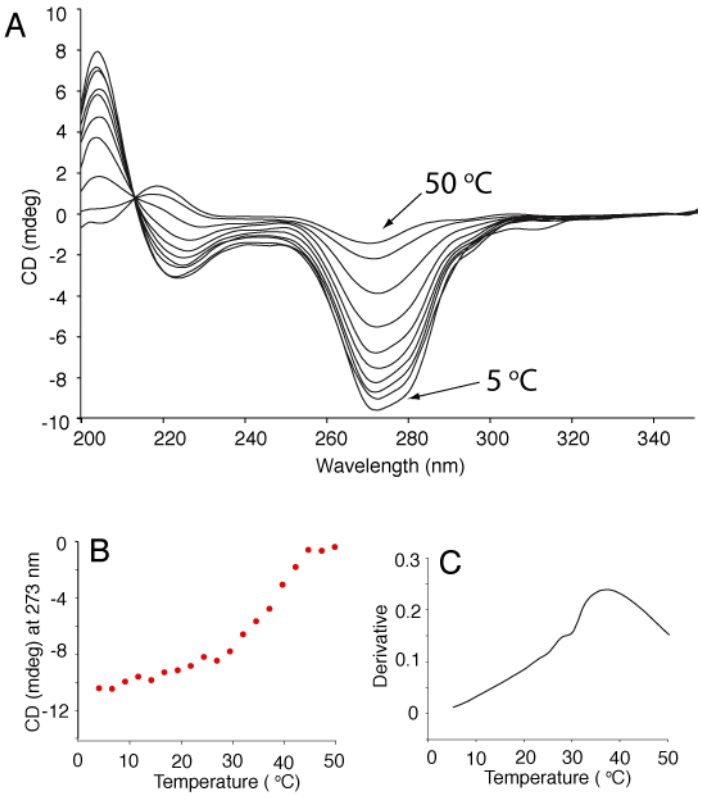


Figure S7. Temperature-dependent CD studies on the T-3′-AAA AAA AAA A-2′-T(GNA) /T-2′-TTT TTT TTT T-3′(npGNA) heteroduplex. (The underlined denotes GNA or npGNA sequences). The concentration of the heteroduplex was 10 μM. **A**. CD spectra obtained in the temperature range of 5-50°C at 5 °C intervals. **B**. Temperature-dependent CD signal change monitored at 273 nm. **C.** Plot of derivative of temperature-dependent CD signal change in **B**.

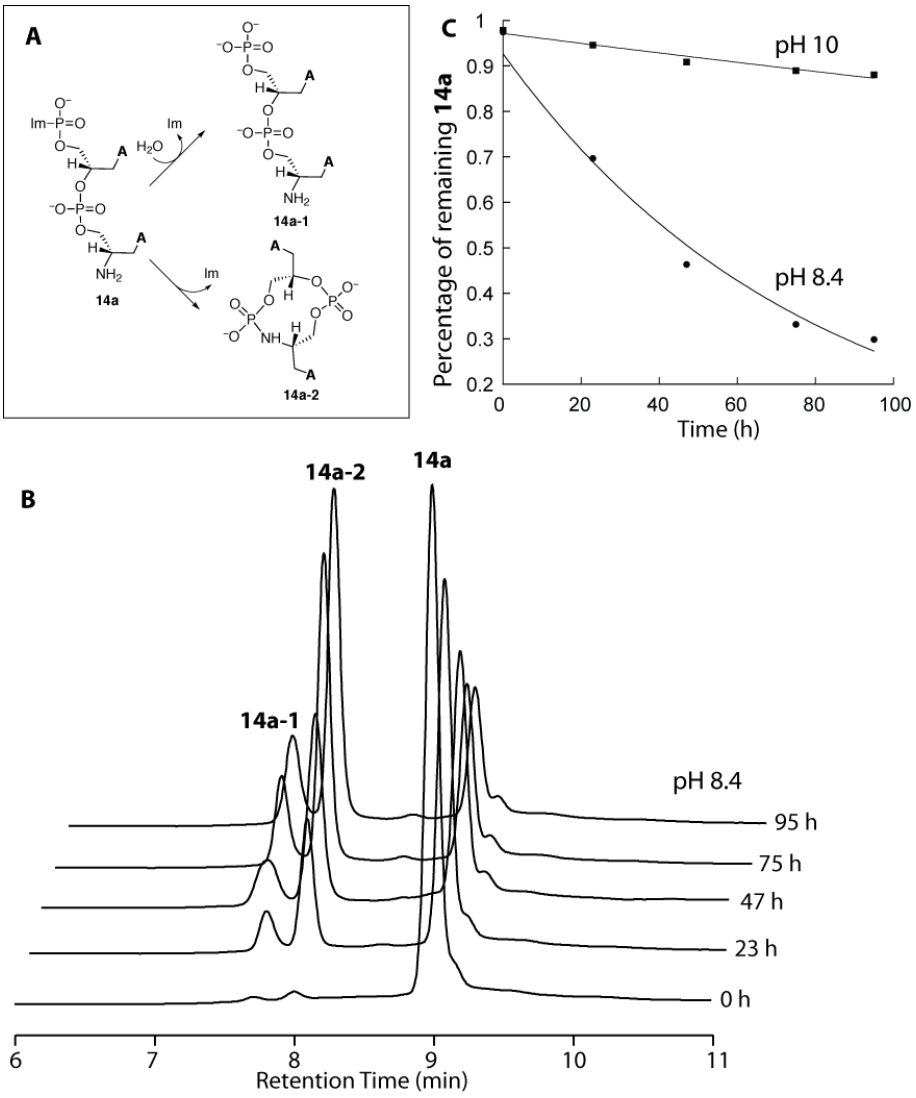


Figure S8. Stability of **14a** monitored by reverse-phase HPLC. **A**. Proposed degradation pathways of **14a**. **B**. HPLC profiles of time-dependent decomposition of **14a** at pH 8.4 and 4 °C. The breakdown products were identified by ESI-MS analysis. $[M–H]^-$ for **14a-1** calcd, 558.1128, obsd, 558.1; for **14a-2**, calcd, 540.1022, obsd, 540.1. **C**. Kinetic analysis of degradation of **14a** at pH 8.4 (solid circle, k: 1.3×10^{-2} h^{-1}, half-life: 53.7 h) and pH 10 (solid square, k: 1.1×10^{-3} h^{-1}, half-life: 615.6 h). Data were fit to a single exponential decay.

Materials and methods

Reagents and solvents were purchased from Sigma-Aldrich. Pyridine, triethylamine, and diisopropylethylamine were distilled from CaH_2. Compounds **2t**, **2c**, **2g**, and **2a** were prepared as described (Zhang, L.; Peritz, A. E.; Carrikk, P. J.; Meggers, E. *Synthesis* **2005**, *4*, 645-653). Flash column chromatography was performed using silica gel from Sigma-Aldrich (Grade 9385, 230-400 mesh) with solvents indicated below. ^{1}H-, ^{31}P-, and ^{13}C-NMR experiments were performed on a Varian 400MHz spectrometer. Chemical shifts are reported in ppm with reference to tetramethylsilane (TMS) or trisilyl propionic acid (TSP) (0.00 ppm) for ^{1}H, phosphoric acid (0.00 ppm) for ^{31}P, or $CDCl_3$ (77.16 ppm) for ^{13}C. Coupling constants are reported in Hz. Low-resolution mass spectrometry (LSMS) analysis was performed on a Bruker Daltonics Esquire 6000 mass spectrometer. Photolysis was performed using a Rayonet RPR-100 photo-reactor (400W, Southern New England Ultraviolet Company, Branford, Connecticut).

Matrix-Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) Mass Spectrometry (MS) A sample of ~200 pmol oligonucleotide was adsorbed on a C18 Zip Tip. Samples were eluted with 1.5 μL of a matrix solution containing a 2:1 mixture of 52.5 mg/mL 3-hydroxypicolinic acid (or 10 mg/mL 2′,4′,6′-trihydroxyacetophenone for oligonucleotides with MW < 2000) in 50% acetontrile and 0.1 M ammonium citrate in water. Eluents were directly spotted onto a stainless steel MALDI-TOF plate and were analyzed in positive mode on a MALDI-TOF mass spectrometer (PerSeptive Biosystems, Model Voyager DE).

Tm determination by thermal denaturation Thermal denaturation experiments were performed on a Varian Cary 1E spectrophotometer equipped with a programmable

temperature control. Optical absorbance was monitored at 260nm (1 nm width) with a heating rate of 1 °C/min. Samples (200 μL) contained 2.0 μM of each strand in 100mM NaCl, 10 mM sodium phosphate, pH 7.0. Melting temperatures (Tm) were calculated from first derivatives of melting curves.

Circular dichroism (CD) spectroscopy CD spectroscopic studies were carried out on an Aviv CD spectrometer (Model 202) at 25 °C. The samples (30 μL) contained 10 μM of each strand in 100mM NaCl, 10 mM sodium phosphate, pH 7.0, path length 1.0 mm. The samples were scanned from 200 to 350 nm with a 1-nm increment. The signal was recorded from the average of 10 measurements for each wavelength. A sample containing only buffer was used as the control for all measurements.

Solid-phase synthesis of npGNA oligonucleotides npGNA oligonucleotides were synthesized by oxidative amination coupling using controlled-pore glass solid support modified with T (0.5 μmol, Glen Research, Sterling, VA) (Chen, J. K.; Schultz, R. G.; Lloyd, D. H.; Gryaznov, S. M. *Nucleic Acids Research* **1995**, *23*, 2661-8). The following procedure was developed for manually delivering reagents and washing solution using syringes during each step: 1. Detritylation with 1 mL of 3% dichloroacetic acid/dichloromethane for 1 min; 2. Phosphitylation with 0.4 M 2-cyanoethyl-*N,N′*-diisopropylchlorophosphite/0.4 M diisopropylethylamine in dichloromethane (0.5 mL for 10 min); 3. Hydrolysis with 0.4 M tetrazole in 9:1 acetonitrile/water (1 mL for 5 min). 4. Coupling with 150 μL of 0.2 M **1t**, **1c**, **1a**, or **1g** and 0.2 M triethylamine in 1:1 CCl_4/acetonitrile for 20 min (C, or T) or 40 min (A or G). Between each step, the solid support was washed with acetonitrile (5 mL). The npGNA oligonucleotides were cleaved off the solid support and deprotected by concentrated ammonia at 55 °C overnight. The

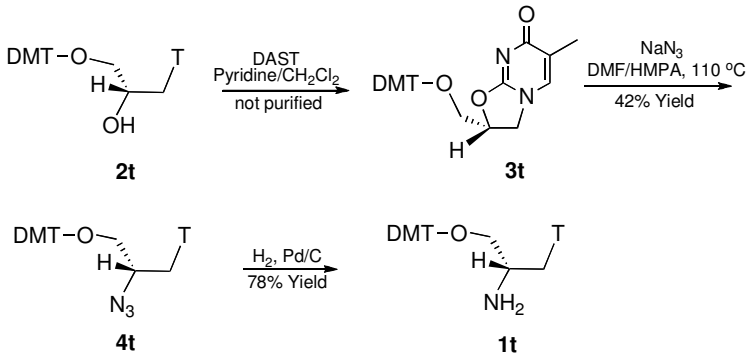
average coupling yields were 40-80% based on the trityl assay, which limits this method to synthesizing short npGNAs. All npGNA, GNA or DNA oligonucleotides used in this study were purified by polyacrylamide gel electrophoresis and characterized by MALDI-TOF MS (See Supporting Table S1).

*Stability studies of **14a*** A solution of 100 μL containing 50 μM **14a** in 0.2 M NaCl, 20 mM $MgCl_2$, and 100 mM HEPBS (*N*-(2-hydroxyethyl)piperazine-*N*′-(4-butanesulfonic acid), pH 8.4 (or 10 mM NaOH, pH 10) was incubated at 4 °C. At each time point, a 20-μL aliquot was removed and analyzed by reverse-phase HPLC using a Varian Microsorb-100 (4.6 mm × 250 mm) column. Conditions: solution A: 25 mM triethylammonium bicarbonate, 2.5 % acetonitrile, pH 7.0; solution B: 100% acetonitrile; Gradient: 0-2 min, 0% B; 2-22 min, 0-40% B. Compound, retention time: **14a**, 9.4 min, **14a-1**, 8.1 min, **14a-2**, 8.4 min.

*Polymerization of in situ generated **14a** on the T-GNA-$(T)_{10}$-T template* The reaction mixture of 10 μL contained 200 μM T-GNA-$(T)_{10}$-T template, 2 mM **13a**, 20 mM $MgCl_2$, 0.2 M NaCl, 0.1 M HEPBS, pH 8.4. The reaction was initiated by photolysis at 4 °C for 6 h and was then incubated for 3 d at 4 °C. Oligonucleotides were then precipitated from the reaction mixture by adding 90 μL of water, 50 μL of 3 M sodium acetate, pH 5.3, and 400 μL of ethanol. The pellet was washed with 80% ethanol once, air-dried, and analyzed by MALDI-TOF MS as described above. For template-independent polymerization, the reaction conditions were the same as described above except that no template was included. After ethanol precipitation of the reaction mixture, the template was co-spotted with the sample on the MALDI plate as an internal control.

*Polymerization of in situ generated **14a** in primer-extension experiments* The reaction mixture of 10 μL contained 1 μM 5′-[^{32}P]-primer/template, 2 mM **13a**, 20 mM $MgCl_2$, 0.2 M NaCl, 0.1 M HEPBS, pH 8.4. The reaction was initiated by photolysis at 4 °C for 6 h and, after another 12 h incubation at 4 °C, was analyzed by 20% denaturing polyacrylamide gel electrophoresis.

Synthesis of npGNA intermediates



Scheme S1. Synthesis of **1t**.

Synthesis of 1t (Scheme S1)

*Synthesis of (R)-1′-(thymidine-1-yl)-3′-O-(4,4′-dimethoxytrityl)-2,2′-anhydro-2′,3′-propanediol (**3t**)* Compound **2t** (4.6 g, 9.2 mmol) was dissolved in 70 mL of anhydrous CH_2Cl_2. (Diethylamino)sulfur trifluoride (DAST) (2 mL, 15.3 mmol) and 0.8 mL anhydrous pyridine were added to the solution simultaneously at room temperature while stirring. The reaction mixture was incubated at room temperature for 15 min and was quenched by adding 10 mL of saturated $NaHCO_3$ at 0 °C. The reaction mixture was then extracted twice with saturated $NaHCO_3$ (100 mL ×2) and once with brine (100 mL), dried over Na_2SO_4, and evaporated *in vacuo* to afford **3t** as light yellow foam (4.4 g, 9.1 mmol), which was used for subsequent synthesis without further purification. ^{1}H NMR (400 MHz, $CDCl_3$) δ:7.18-7.32 (9H, m), 7.11 (1H, d, J=1.0), 6.78-6.82 (4H, m), 4.98 (1H, m), 4.21 (1H, dd, J=9.2, 9.2), 3.97 (1H, dd, 9.4, 5.2), 3.77 (6H, s), 3.64 (1H, dd, J=3.6, 10.8), 3.15 (1H, dd, J=3.2, 10.8), 1.98 (3H, d, J=1.0). ^{13}C NMR (100.5 MHz, $CDCl_3$) δ: 172.8, 160.7, 158.7, 144.1, 135.2, 135.0, 132.0, 130.0, 128.1, 127.9, 127.1, 118.4, 113.43, 113.38, 86.6, 77.4, 63.5, 55.3, 48.4, 14.1; LRMS calcd for $[M+H]^+$ ($C_{29}H_{29}N_2O_5$): 485.2076, obsd, 485.2.

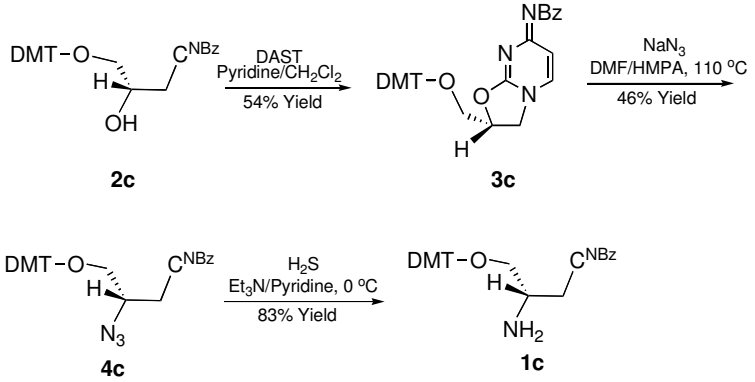
*Synthesis of (S)-1′-(thymidine-1-yl)-2′-azido-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**4t**)*

Compound **3t** (4.4 g, 9.1 mmol) and NaN_3 (3 g, 460 mmol) was dissolved in 20 mL of anhydrous dimethylformamide and 20 mL of hexamethylphosphoramide and was stirred at 100 °C for 18 h. The reaction mixture was then poured into 200 mL of ethyl acetate. The organic layer was washed twice with water (200 mL ×2), dried over Na_2SO_4, and evaporated in vacuo. The crude pellet was further purified by silica column chromatography (1% to 2% MeOH/CH_2Cl_2) to afford **4t** as white foam (2.0 g, 3.8 mmol, 42%).^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.72 (1H, br), 7.45 (1H, s), 7.43 (1H, s), 7.23-7.33 (7H, m), 6.97 (1H, s), 6.85(4H, d, J=8), 3.96 (1H, dd, J=5.6, 13.6), 3.89 (1H, m), 3.80 (6H, s), 3.43 (1H, dd, J= 8.4, 13.6), 3.37 (1H, dd, J= 2.4, 9.4), 3.19 (1H, dd, J=5.6, 9.4), 1.87 (3H, s). ^{13}C NMR (100.5 MHz, $CDCl_3$) δ:164.2, 158.8, 150.9, 144.4, 141.3, 135.44, 135.38, 130.08, 130.05, 128.14, 128.06, 127.2, 113.42, 113.40, 110.6, 87.0, 63.5, 60.4, 55.4, 49.5, 12.5; LRMS calcd for $[M+Na]^+$ ($C_{29}H_{29}N_5NaO_5$): 550.2066, obsd, 550.2.

*Synthesis of (S)-1′-(thymidine-1-yl)-2′-amino-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**1t**)*

Compound **4t** (4.6 g, 8.7 mmol) was dissolved in 100 mL of anhydrous methanol and 10% Pd on activated carbon (750 mg) was added. The reaction mixture was then shaken vigorously under H_2 (30 psi) using a hydrogenator at room temperature for 20 h. The reaction mixture was then filtered through Celite, evaporated in vacuo, and purified by silica gel chromatography (1%-4% MeOH/ CH_2Cl_2) to afford **1t** as white foam (3.4 g, 6.8 mmol, 78%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.42 (1H, s), 7.40 (1H, s), 7.22-7.31 (7H, m), 6.96 (1H, d, J=0.4), 6.83 (4H, d, J=7.2), 3.87 (1H, dd, J= 4.8, 14.0), 3.79 (6H, s), 3.64 (1H, dd, J= 7.6, 13.6), 3.28 (1H, m), 3.13 (2H, m), 1.85 (3H, s). ^{13}C NMR (100.5 MHz, $CDCl_3$) δ: 164.4, 158.7, 151.5, 144.7, 141.6, 135.8, 130.1, 128.1, 128.0, 127.1, 113.3,

110.1, 86.3, 65.6, 55.3, 52.2, 50.9, 12.4. LRMS calcd for $[M+Na]^+$ ($C_{29}H_{31}N_3NaO_5$): 524.2161, obsd, 524.2.



Scheme S2. Synthesis of **1c**.

Synthesis of 1c (Scheme S2)

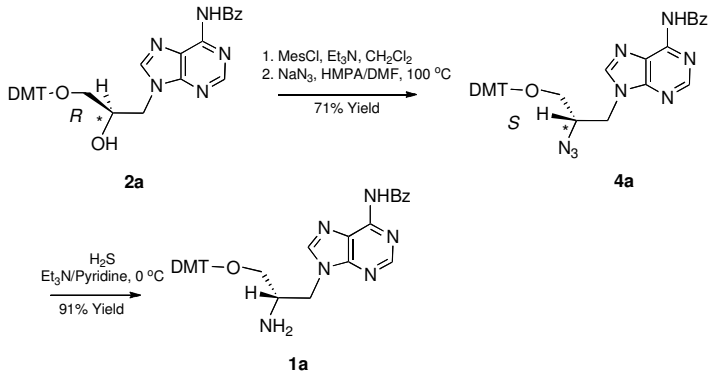
*Synthesis of (R)-1′-(4-N-benzoylcytosine-1-yl)-3′-O-(4,4′-dimethoxytrityl)-2,2′-anhydro-2′,3′-propanediol (**3c**)* Compound **2c** (2.8 g, 4.7 mmol) was dissolved in 50 mL of anhydrous CH_2Cl_2. (Diethylamino)sulfur trifluoride (DAST) (1.0 mL, 7.7 mmol) and 0.4 mL anhydrous pyridine were added to the solution simultaneously at room temperature while stirring. The reaction mixture was incubated at room temperature for 15 min and was quenched by adding 6 mL of saturated $NaHCO_3$ at 0 °C. The reaction mixture was then extracted twice with saturated $NaHCO_3$ (100 mL ×2) and once with brine (100 mL), dried over Na_2SO_4, and evaporated in vacuo. Compound **3c** was further purified by silica gel chromatography (2% MeOH/ CH_2Cl_2) as light yellow foam (1.4 g, 2.5 mmol, 54%). Samples of **3c** contained ~5% impurity due to partial decomposition during purification. However, the impurity did not seem to affect the next step of synthesis. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.17 (2H, d, J= 6.4), 7.2-7.5 (10H, m), 7.17 (1H, d, J=7.2), 6.80(4H, d, J=7.2), 5.02 (1H, m), 4.23 (1H, dd, J=9.2,9.2), 3.99 (1H, dd, J=5.2, 8.4), 3.78 (6H,s), 3.66

(1H, dd, J=3.6, 11.2), 3.18 (1H, dd, J= 3.2, 11.2); ^{13}C NMR (100.5 MHz, $CDCl_3$) δ:179.3, 163.3, 159.8, 158.8, 158.7, 144.1, 136.3, 135.7, 135.2, 134.9, 131.9, 130.0, 129.7, 128.2, 128.1, 127.9, 127.2, 113.5, 113.4, 107.2, 86.6, 77.9, 63.3, 55.3, 48.3, LRMS calcd for $[M+H]^+$ ($C_{35}H_{32}N_3O_5$): 574.2342, obsd, 574.1.

*Synthesis of (S)-1′-(4-N-benzoylcytosine-1-yl)-2′-azido-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**4c**)* Compound **3c** (2.0 g, 3.5 mmol) and NaN_3 (1.2 g, 17.5 mmol) was dissolved in 10 mL of dimethylformamide/hexamethylphosphoramide (1:1) and was stirred at 110 °C for 4 h. The reaction mixture was then poured into 100 mL of ethyl acetate. The organic layer was washed twice with water (200 mL ×2), dried over Na_2SO_4, and evaporated in vacuo. The crude pellet was further purified by silica column chromatography (1% to 2% MeOH/CH_2Cl_2) to afford **4c** as light yellow foam (1.0 g, 1.6 mmol, 46%).^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.80 (1H, br), 7.89 (2H, d, J=7.6), 7.21-7.63 (14H, m), 6.85(4H, d, J=8.8), 4.21 (1H, dd, J=4, 13.2), 4.12 (1H, m), 3.79 (6H, s), 3.58 (1H, dd, J= 8.8, 13.2), 3.45 (1H, dd, J= 2.8, 10), 3.18 (1H, dd, J=5.6, 10). ^{13}C NMR (100.5 MHz, $CDCl_3$) δ: 162.7, 158.8, 150.0, 144.3, 135.4, 135.3, 133.4, 133.0, 130.1, 129.2, 128.1, 128.0, 127.6, 127.1, 113.4, 96.5, 86.9, 63.4, 59.7, 55.3, 51.8. LRMS calcd for $[M+H]^+$ ($C_{35}H_{33}N_6O_5$): 617.2512, obsd, 617.3.

*Synthesis of (S)-1′-(4-N-benzoylcytosine-1-yl)-2′-amino-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**1c**)* Compound **4c** (1.4g, 2.3 mmol) and triethylamine (0.6 mL) were dissolved in 3.4 mL of anhydrous pyridine. The solution was then cooled to 0 °C on ice and hydrogen sulfide was bubbled through the reaction mixture for 15 min. The solvent was then evaporated in vacuo. The crude product was redissolved in 100 mL of CH_2Cl_2, extracted twice with saturated $NaHCO_3$ (100 mL ×2) and once with brine (100 mL), and

evaporated. Compound **1c** was further purified by silica gel chromatography (1%-3% MeOH/CH_2Cl_2) as light yellow foam (1.1 g, 1.9 mmol, 83%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.89 (2H, d, J=7.2), 7.22-7.59 (14H, m), 6.84 (4H, d, J=8.8), 4.10 (1H, dd, J=2.8, 13.2), 3.80 (1H, J= 3.6, 13.2), 3.78 (6H, s), 3.42 (1H, m), 3.22 (1H, dd, J= 4.4, 9.6), 3.10 (1H, dd, J=4, 9.6). ^{13}C NMR (100.5 MHz, $CDCl_3$) δ: 162.2, 158.7, 150.2, 144.6, 135.7, 135.6, 133.2, 130.1, 129.1, 128.11, 128.06, 127.7, 127.1, 113.3, 96.1, 86.2, 64.9, 55.3, 54.6, 50.1. LRMS calcd for $[M+H]^+$ ($C_{35}H_{35}N_4O_5$): 591.2607, obsd, 591.2.



Scheme S3. Synthesis of **1a**.

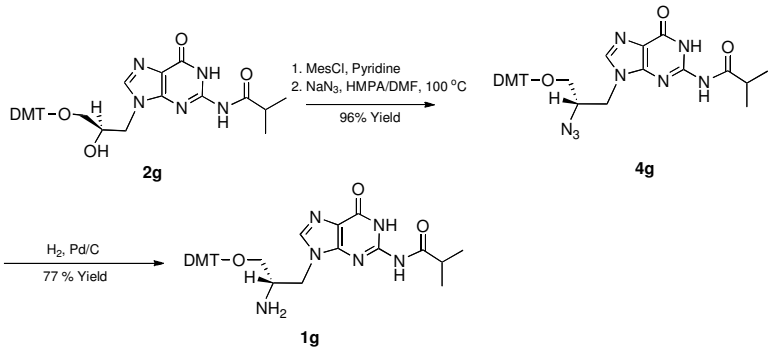
Synthesis of 1a (Scheme S3)

*Synthesis of (S)-1′-(6-N-benzoyladenine-9-yl)-2′-azido-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**4a**)* Compound **2a** (4.3 g, 7.0 mmol) was dissolved in 70 mL of anhydrous pyridine. The solution was cooled to 0 °C and methanesulfonyl chloride (1.1 mL, 14.1 mmol) was added drop-wise. The reaction mixture was then warmed up to room temperature and stirred for 4 h. The reaction was quenched by adding 5 mL of methanol at 0 °C and the solvent was then evaporated in vacuo. The crude product was dissolved in 200 mL of CH_2Cl_2, washed twice with saturated $NaHCO_3$ (100 mL ×2) and once with brine (100 mL), dried over Na_2SO_4, and evaporated in vacuo. The resulting pellet was

dissolved in 20 mL of dimethylformamide/hexamethylphosphoramide (1:1). Sodium azide (2.3 g, 35 mmol) was added to the solution and the reaction mixture was incubated at 100 °C for 16 h. The reaction mixture was then poured into 200 mL of ethyl acetate. The organic layer was washed twice with water (200 mL ×2), dried over Na_2SO_4, and evaporated in vacuo. The crude product was further purified by silica column chromatography (0.5% to 1% MeOH/CH_2Cl_2) to afford **4a** as light yellow foam (3.2 g, 5.0 mmol, 71%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.06 (1H, s), 8.77 (1H, s), 8.02 (2H, d, J= 7.2), 7.99 (1H, s), 7.23-7.61 (12H, m), 6.84 (4H, d, J= 8.8), 4.40 (1H, dd, J=3.6, 14.0), 4.18 (1H, dd, J= 8.4, 14.0), 4.00 (1H, m), 3.79 (6H, s), 3.44 (1H, dd, J= 4.0, 10.0), 3.29 (1H, dd, J=6.4, 10.0). ^{13}C NMR (100.5 MHz, $CDCl_3$) δ: 164.7, 158.8, 152.8, 152.1, 149.6, 144.3, 143.6, 135.34, 135.31, 133.7, 132.9, 130.1, 129.0, 128.2, 128.0, 127.9, 127.2, 122.9, 113.4, 87.2, 63.8, 60.9, 55.4, 44.7. LRMS calcd for $[M+H]^+$ ($C_{36}H_{33}N_8O_4$): 641.2625, obsd, 641.2.

*Synthesis of (S)-1′-(6-N-benzoyladenine-9-yl)-2′-amino-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**1a**)* Compound **4a** (1.0 g, 1.6 mmol) and triethylamine (0.6 mL) were dissolved in 3.4 mL of anhydrous pyridine at 0 °C. Hydrogen sulfide was bubbled through the solution at 0 °C for 20 min. The reaction mixture was then evaporated, re-dissolved in 50 mL of CH_2Cl_2, extracted twice with saturated $NaHCO_3$ (100 mL ×2) and once with brine (100 mL), dried over Na_2SO_4, and evaporated in vacuo. The crude product was further purified by silica gel chromatography to afford **1a** as white foam (0.9g, 1.46 mmol, 91%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.12 (1H, br), 8.75 (1H, br), 8.00-8.02(3H, m), 7.20-7.61(12H, m), 6.83 (4H, d, J=9.2), 4.43 (1H, dd, J= 4.4, 14.0), 4.22 (1H, dd, J=7.6, 14.0), 3.78 (6H, s), 3.45 (1H, m), 3.16 (2H, m); ^{13}C NMR (100.5

MHz, $CDCl_3$) δ: 164.7, 158.7, 152.6, 152.5, 149.4, 144.7, 144.1, 135.73, 135.70, 133.8, 132.8, 130.1, 128.9, 128.1, 128.0, 127.9, 127.1, 122.9, 113.3, 86.5, 65.6, 55.3, 51.4, 48.1 . LRMS calcd for $[M+H]^+$ ($C_{36}H_{35}N_6O_4$): 615.2720, obsd, 615.3.



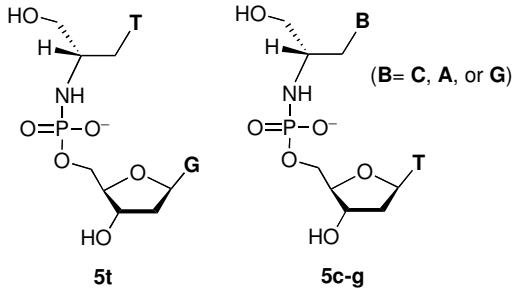
Scheme S4. Synthesis of **1g**.

Synthesis of 1g (Scheme S4)

*Synthesis of (S)-1′-(2-N-isobutyrylguanine-9-yl)-2′-azido-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**4g**)* Compound **2g** (3.6 g, 6 mmol) was dissolved in 40 mL of anhydrous pyridine. The solution was cooled to 0 °C and methanesulfonyl chloride (1.1 mL, 14.1 mmol) was added drop-wise. The reaction mixture was then warmed up to room temperature and stirred for 4 h. The reaction was quenched by adding 5 mL of methanol at 0 °C. The solvent was then evaporated and the crude product was redissolved in 100 mL of CH_2Cl_2. The solution was then washed twice with saturated $NaHCO_3$ (100 mL ×2) and once with brine (100 mL), dried over Na_2SO_4, and evaporated in vacuo. The resulting pellet was dissolved in 20 mL of dimethylformamide/hexamethylphosphoramide (1:1). Sodium azide (2 g, 30 mmol) was added to the solution and the reaction mixture was incubated at 100 °C for 2 h. The reaction mixture was then poured into 200 mL of ethyl acetate. The organic layer was washed twice with water (200 mL ×2), dried over

Na_2SO_4, and evaporated in vacuo. The crude product was further purified by silica column chromatography (0.5% to 1% MeOH/CH_2Cl_2) to afford **4g** as white foam (3.6 g, 5.8 mmol, 96%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 11.98 (1H, br), 8.63 (1H, br), 7.53 (1H, s), 7.42 (2H, d, J= 7.6), 7.21-7.31 (7H, m), 6.81 (4H, d, J= 8.8), 4.10 (1H, dd, J=4.8, 14.0), 3.98 (1H, dd, J=8.0, 14.0), 3.81(1H, m), 3.78 (6H, s), 3.28 (1H, dd, J= 4.0, 10.0), 3.20 (1H, dd, J= 6.0, 10.0), 2.68 (1H, sept, J= 7.0), 1.26 (6H, d, J= 7.0); ^{13}C NMR (100.5 MHz, $CDCl_3$) δ: 178.6, 158.8, 155.6, 148.4, 147.5, 144.3, 139.4, 135.4, 135.3, 130.07, 130.05, 128.1, 128.0, 127.2, 121.1, 113.4, 87.2, 63.5, 60.9, 55.4, 44.2, 36.6, 19.13, 19.11. LRMS calcd for $[M+Na]^+$ ($C_{33}H_{34}N_8NaO_5$): 645.2250, obsd, 645.2.

*Synthesis of (S)-1′-(2-N-isobutyrylguanine-9-yl)-2′-amino-3′-O-(4,4′-dimethoxytrityl)-3′-propanol (**1g**)* Compound **4g** (1.9 g, 3.0 mmol) was dissolved in 100 mL of anhydrous methanol and 10% Pd on activated carbon (750 mg) was added. The reaction mixture was then shaken vigorously under H_2 (30 psi) using a hydrogenator at room temperature for 20 h. The reaction mixture was then filtered through Celite and evaporated in vacuo to afford **1g** as white foam (1.4 g, 2.3 mmol, 77%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.51 (1H, s), 7.39 (1H, s), 7.38 (1H, s), 7.22-7.31 (7H, m), 6.77 (4H, d, J=7.2), 4.18 (1H, dd, J= 4.4, 14.0), 4.01 (1H, dd, J= 7.2, 14.0), 3.75 (6H, s), 3.40 (1H, m), 3.15 (1H, dd, J= 4.8, 9.2), 3.04 (1H, dd, J= 6.4, 9.0), 2.64 (1H, sept, J=6.8), 1.17 (6H, d, J= 6.8). ^{13}C NMR (100.5 MHz, $CDCl_3$) δ:179.1, 158.7, 155.7, 148.8, 147.7, 144.6, 139.7, 135.6, 135.5, 130.0, 128.0, 127.1, 120.8, 113.3, 86.5, 65.2, 55.4, 51.5, 47.9, 46.0, 36.3, 19.1. LRMS calcd for $[M+H]^+$ ($C_{33}H_{37}N_6NO_5$): 597.2820, obsd, 597.3.



Scheme S5. Structures of dinucleotide model compounds **5t**, **5c**, **5a**, and **5g**.

Synthesis and characterization of 5t-5g (Scheme S5)

Dinucleotides **5t**, **5c**, **5a**, and **5g** were synthesized via oxidative amination coupling as described above using controlled pore glass solid support modified with dG or T (2 μmol, Glen Research, Sterling, VA) as the starting material. The dinucleotides were cleaved off the solid support and deprotected by concentrated ammonia treatment at 55 °C overnight. The dinucleotides were further purified by reverse phase HPLC (Column: Varian C18 Microsorb 100, 250 × 21.4 mm; Solution A: 10 mM NH_4HCO_3, pH 6.8, Solution B: 100% acetonitrile; Gradient: 10-25% B over 15 min; Flow rate: 15 mL/min). Compound, retention time, yield: **5t**, 5.5 min, 520 nmol; **5c**, 5.1 min, 652 nmol, **5a**, 6.2 min, 390 nmol, and **5g**, 5.1 min, 190 nmol.

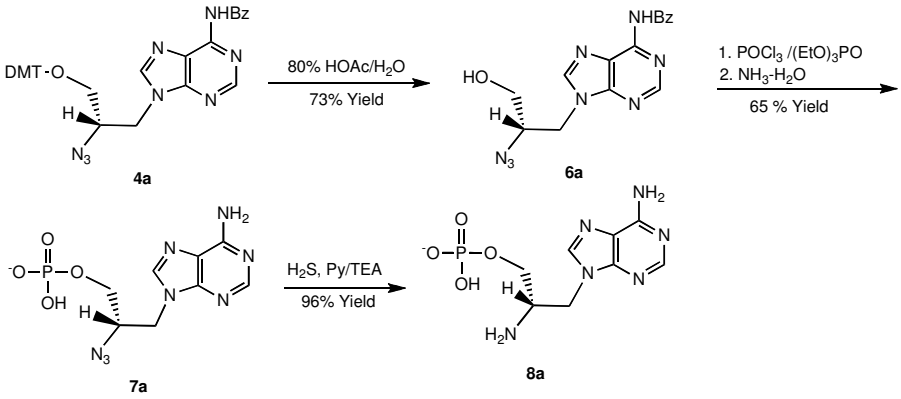
Compound **5t**: ^{1}H NMR (400 MHz, D_2O) δ: 8.39 (1H, s), 8.03 (1H, s), 7.32 (1H, s), 6.21 (1H, t, d=6.8), 4.63 (1H, m), 4.06 (1H, m), 3.87 (2H, m), 3.69 (1H, m), 3.59 (1H, dd, J= 7.6, 14), 3.48 (2H, m), 3.23 (1H, m), 2.75 (1H, m), 2.48 (1H, m), 1.67 (3H, s); ^{31}P NMR (160.84 MHz, D_2O) δ: 7.69 ppm; LRMS calcd for $[M-H]^-$:527.1404, obsd, 527.1.

Compound **5c**: ^{1}H NMR (400 MHz, D_2O) δ: 7.70 (1H, s), 7.58 (1H, d, J=7.3), 6.31 (1H, t, J=6.4), 5.87 (1H, d, J= 7.2), 4.48 (1H, m), 4.07 (1H, m), 4.01 (1H, m), 3.87 (2H, m),

3.56-3.68 (3H, m), 3.38 (1H, m), 2.20-2.36 (2H, m), 1.90 (3H, s); ^{31}P NMR (160.84 MHz, D_2O) δ: 8.00 ; LRMS calcd for $[M-H]^-$:487.1342, obsd, 487.1.

Compound **5g**: ^{1}H NMR (400 MHz, D_2O) δ: 7.84 (1H, s), 7.62 (1H, s), 6.15 (1H, t, J= 6.4), 4.39 (1H, m), 4.27 (1H, m), 3.95 (1H, dd, J=10.4, 14.4), 3.77 (2H, m), 3.63-3.69 (2H, m), 3.34 (1H, m), 3.05 (1H, m), 2.12-2.28 (2H, m), 1.90 (3H, s); ^{31}P NMR (160.84 MHz, D_2O) δ: 7.51; LRMS calcd for $[M-H]^-$:527.1404, obsd, 527.1.

Compound **5a**: ^{1}H NMR (400 MHz, D_2O) δ: 8.20 (1H, s), 7.98 (1H, s), 7.29 (1H, s), 6.09 (1H, t, J= 6.8), 4.43 (1H, m), 4.33 (1H, m), 4.10 (1H, dd, J= 10.4, 14.4), 3.80 (2H, m), 3.66-3.73 (2H, m), 3.39 (2H, m), 2.18 (1H, m), 2.04 (1H, m); 1.81 (3H, s); ^{31}P NMR (160.84 MHz, D_2O) δ: 7.56; LRMS calcd for $[M-H]^-$:511.1455, obsd, 511.1.



Scheme S6. Synthesis of **8a**.

Synthesis of 8a (Scheme S6)

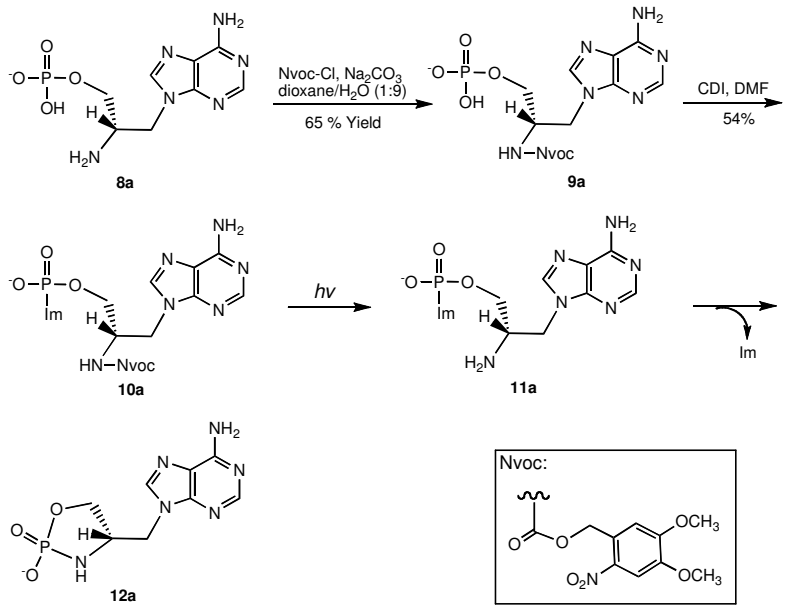
*Synthesis of (S)-1′-(6-N-benzoyladenine-9-yl)-2′-azido-3′-propanol (**6a**)* Compound **4a** (0.49g, 0.77 mmol) was dissolved in 10 mL 3% dichloroacetic acid in dichloromethane and the reaction mixture was stirred at room temperature for 15 min. The reaction was quenched by adding 1 mL of methanol and solvent was evaporated in vacuo. The crude

product was then purified by silica gel chromatography (2.5-4% methanol in dichloromethane) to afford **6a** as white solid (Yield: 190 mg, 0.56 mmol, 73%). ^{1}H NMR (400 MHz, $CDCl_3$) δ: 9.34 (1H, s), 8.76 (1H, s), 8.06 (1H, s), 8.03 (2H, d, J=7.2), 7.61 (1H, dd, J= 7.6, 7.6), 7.52 (2H, dd, J= 8.0, 8.0), 4.67 (1H, br), 4.46 (2H, m), 4.00 (1H, m), 3.74 (1H, dd, J=4.8, 12), 3.55 (1H, dd, J=6.8, 12); ^{13}C NMR (100.5 MHz, $CDCl_3$) δ: 164.9, 152.7, 152.3, 149.9, 144.0, 133.5, 133.1, 129.0, 128.0, 122.7, 61.3, 60.8, 43.7; LRMS calcd for $[M+H]^+$ ($C_{15}H_{15}N_8O_2$) : 339.1318, obsd, 339.0.

*Synthesis of (S)-1′-(adenine-9-yl)-2′-azido-3′-phosphopropanol (**7a**)* Compound **6a** (91 mg, 0.268 mmol) was rendered anhydrous by co-evaporating with dry dimethylformamide (1 mL ×3) and was dissolved in 2 mL of freshly distilled triethylphosphate at 0 °C. Freshly distilled $POCl_3$ (123 μL, 1.34 mmol, 5 eq.) was added to the reaction mixture. The solution was incubated at 0 °C for 3 h and was quenched by adding 50 mL H_2O at 0 °C. The pH of the solution was adjusted to ~ 7 by adding 1 M NaOH solution. The solvent was then evaporated in vacuo and the resulting pellet was dissolved in concentrated ammonia to remove the benzoyl protection group on the nucleobase. Crude compound **7a** (161 μmol) was purified by anion-exchange liquid chromatography using DEAE A-25 resin with a linear gradient of 0-0.5 M triethylammonium bicarbonate (final yield: 105 μmol, 65%). ^{1}H NMR (400 MHz, D_2O) δ: 8.01 (1H, s), 7.99 (1H, s), 4.27 (1H, dd, J=3.2, 14.4), 4.12 (1H, dd, J=8.8, 14.4), 4.02 (1H, m), 3.87 (1H, m), 3.75 (1H, m); ^{31}P NMR (160.84 MHz, D_2O) δ: 1.39; LRMS calcd for $[M-H]^-$ ($C_8H_{10}N_8O_4P$): 313.0563, obsd, 312.9.

*Synthesis of (S)-1′-(adenine-9-yl)-2′-amino-3′-phosphopropanol (**8a**)* Compound **7a** (105 μmol) was dissolved in a mixture of triethylamine (1.5 mL) and pyridine (8.5 mL).

Hydrogen sulfide was then bubbled through the mixture at 0 °C for 1 h. The solvent was then evaporated in vacuo and the crude product **8a** was purified by reverse-phase preparative HPLC (Column: Varian C18 Microsorb 100, 250 × 21.4 mm; Solution A: 10 mM triethylammonium acetate, pH 7.0, Solution B: 100% acetonitrile; Gradient: 0-50% B over 20 min; Flow rate: 15 mL/min). Compound, retention time, yield: **8a**, 6.4 min, 101 μmol (96%). ^{1}H NMR (400 MHz, D_2O) δ: 8.23 (1H, s), 8.19 (1H, s), 4.57 (2H, m), 4.03 (1H, m), 3.97 (1H, m), 3.87 (1H, m); ^{31}P NMR (160.84 MHz, D_2O) δ: 4.73; LRMS calcd for $[M-H]^-$ ($C_8H_{12}N_6O_4P$): 287.0658, obsd, 286.8.



Scheme S7. Synthesis of **10a** and its photolytic conversion to **11a** and **12a**

Synthesis of 10a and its photolytic conversion (Scheme S7)

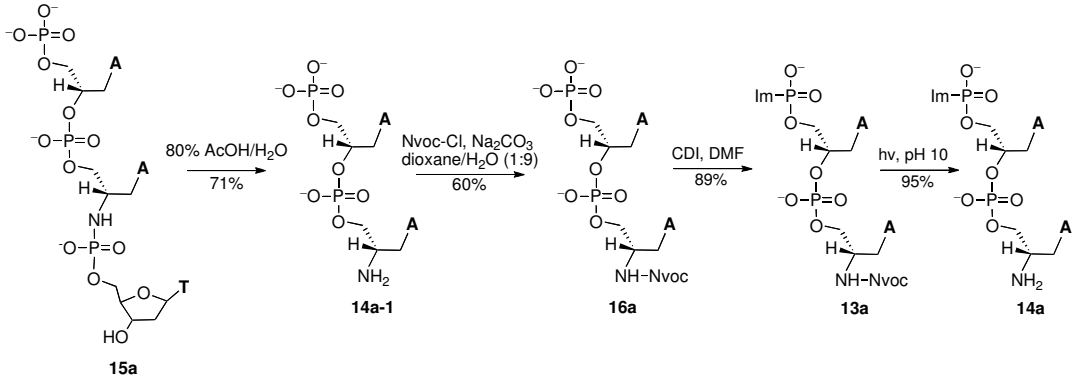
*Synthesis of **9a*** Compound **8a** (20 μmol) was dissolved in 2 mL of 0.5 M Na_2CO_3 solution. Nvoc-Cl (100 μmol, 5 eq.) in 1 mL of dioxane was then added. The reaction mixture was incubated at room temperature overnight and was then subjected to reverse-phase HPLC purification as described above for **8a**. Compound, retention time, yield: **9a**,

15.2 min, 13 μmol (65%) ^{1}H NMR (400 MHz, D_2O) δ: 8.01 (1H, s), 7.75 (1H, s), 7.64 (1H, s), 6.57 (1H, s), 4.89 (1H, d, J=14.0), 4.59 (1H, d, J=14.0), 4.41 (1H, dd, J= < 1, 14.4), 4.24 (1H, dd, J=13.6, 14.2), 4.12 (1H, m), 4.02 (2H, m), 3.96 (3H, s), 3.89 (3H, s); ^{31}P NMR (160.84 MHz, D_2O) δ: 4.86; LRMS calcd for $[M-H]^-$ ($C_{18}H_{21}N_7O_{10}P$) : 526.1088, obsd, 525.9.

*Synthesis of **10a*** Compound **9a** (25 μmol) was rendered anhydrous by evaporating with dry dimethylformamide (1 mL ×3) and was then re-dissolved in 1 mL of dimethylformamide. Carbonyl diimidazole (25 mg, 125 μmol, 5 eq.) in 1 mL of dimethylformamide was added to the mixture. The solution was then incubated at room temperature overnight. The product, **10a**, was precipitated from the reaction by adding acetone (20 mL), ether (15 mL), and 0.2 g $NaClO_4$ followed by incubation at -20 °C for 1 h. The pellet was washed once was 20 mL of 1:1 acetone/ether and was air-dried to afford **10a** as light yellow solid (13.5 μmol, 54%). ^{31}P NMR (160.84 MHz, D_2O) δ: -7.0; LRMS calcd for $[M-H]^-$ ($C_{21}H_{23}N_9O_9P$) : 576.1356, obsd, 575.9.

*Conversion of **10a** to **11a** or **12a** by photolysis* Compound **10a** (1 μmol) was dissolved in 500 μL of 10 % D_2O/H2O and the pH of the solution was adjusted either to 13 by 1M NaOH solution or to 7 by adding 50 μL 0.5 M sodium phosphate (pH 7.0). The solutions were UV radiated for 5 h at 4 °C. The resulting mixtures were analyzed by ^{31}P NMR and by ESI-MS. At pH 13, **10a** was quantitatively converted to the desired product, **11a**. ^{31}P NMR (160.84 MHz, D_2O) δ: -8.2; LRMS calcd for $[M-H]^-$ ($C_{11}H_{14}N_8O_3P$) : 337.0926, obsd, 336.9. At pH 7, **11a** was not detected in the reaction mixture. Instead, **10a** was quantitatively converted to the cyclic nucleotide **12a**. ^{31}P NMR (160.84 MHz, D_2O) δ: 31.6. The unusually large ^{31}P chemical shift was consistent with the previously reported

value for a phosphoramidate group in a cyclic, 5-membered ring (28 ppm for 2′-amino-uridine-2′,3′-cyclic phosphoramidate) (Thomson, J. B.; Patel, B. K.; Jimenez, V.; Eckart, K.; Eckstein, F. *J. Org. Chem.* **1996**, *61*, 6273-6281). LRMS calcd for [M-H]$^-$ ($C_8H_{10}N_6O_3P$) : 269.0552, obsd, 269.0.



Scheme S8. Synthesis of **13a** and its photolytic conversion to **14a**

Synthesis of 13a and its conversion to 14a (Scheme S8)

*Solid-phase synthesis of trinucleotide **15a*** Compound **15a** was prepared using a combination of solid-phase oxidative amination and phosphoramidite chemistry. Controlled-pore glass beads (110μmol/g, Glen Research) pre-charged with thymidine (total loading, 20 μmol, 180 mg) were first coupled with **1a** (100 μmol, 5 eq.) using oxidative coupling as described above. The dinucleotide with a phosphoramidate linkage was then further extended with *S*-adenosine-glyceronucleotide by standard phosphoramidite chemistry as described before {Zhang, 2005 #55}. The resulting trinucleotide was then phosphorylated using the chemical phosphorylation reagent from Glen Research (3-(4,4'-dimethoxytrityloxy)-2,2-dicarboxyethyl-propyl-(2-cyanoethyl)-(*N*,*N*′-diisopropyl)-phosphoramidite, 100 μmol, 5 eq.). The resin was then deprotected by incubating with concentrated ammonia at 55 °C overnight. The crude product (6.5 μmol

based on calculated ε_{260}: 38.5 $mM^{-1}cm^{-1}$) was further purified by reverse-phase preparative HPLC (Column: Varian C18 Microsorb 100, 250 × 21.4 mm; Solution A: 10 mM NH_4HCO_3, pH 6.8, Solution B: 100% acetonitrile; Gradient: 0-40% B over 20 min; Flow rate: 15 mL/min). Compound, retention time, yield: **15a**, 6.6 min, 2.8 μmol (43%). ^{31}P NMR (160.84 MHz, D_2O) δ: 6.81, 4.76, 0.49; LRMS calcd for [M-H]$^-$ ($C_{26}H_{35}N_{13}O_{15}P_3$) : 862.1588, obsd, 862.2.

*Conversion of **15a** to dinucleotide **14a-1** by acid hydrolysis* Compound **15a** (2.8 μmol) was treated with 5 mL of 80% acetic acid in water at room temperature. The progress of hydrolysis was monitored by ^{31}P NMR and ~80% of the starting material was hydrolyzed in 3 d. The reaction mixture was then evaporated to remove solvent. The crude product was purified by reverse-phase HPLC as described for **15a**. Compound, retention time, yield: **14a-1**, 5.5 min, 2.0 μmol (71%, based on calculated ε_{260}: 30 $mM^{-1}cm^{-1}$). ^{31}P NMR (160.84 MHz, D_2O) δ: 4.61, 0.39; LRMS calcd for [M-H]$^-$ ($C_{16}H_{22}N_{11}O_8P_2$) : 558.1128, obsd, 558.1.

*Synthesis of 2′-NH_2 protected dinucleotide **16a** from **14a-1*** Compound **14a-1** (2 μmol) was dissolved in 1.5 mL of 0.5 M Na_2CO_3 solution. Nvoc-Cl (10 μmol, 5 eq.) in 0.5 mL of dioxane was then added. The reaction mixture was incubated at room temperature overnight and was then subjected to reverse-phase HPLC purification (Column: Varian C18 Microsorb 100, 250 × 21.4 mm; Solution A: 25 mM triethylammonium bicarbonate, 2.5 % acetonitrile, pH 7.0, Solution B: 100% acetonitrile; Gradient: 0-40% B over 20 min; Flow rate: 15 mL/min). Compound, retention time, yield: **16a**, 12 min, 1.2 μmol (60 %, based on calculated ε_{260}: 32.2 $mM^{-1}cm^{-1}$). ^{31}P NMR (160.84 MHz, D_2O) δ: 4.41, 0.45; LRMS calcd for [M-H]$^-$ ($C_{26}H_{31}N_{12}O_{14}P_2$) : 797.1558, obsd, 797.1.

*Synthesis of 3′-imidazole activated dinucleotide **13a*** Compound **16a** (0.76 μmol) was rendered anhydrous by evaporating with dry dimethylformamide (1mL ×3) and was then re-dissolved in 0.5 mL of dimethylformamide. A solution of 1M carbonyl diimidazole (10 μL, 13 eq.) in dimethylformamide was added to the mixture. The solution was then incubated at room temperature overnight. The solvent was evaporated and the crude product was purified by reverse-phase HPLC as described for **16a**. Compound, retention time, yield: **13a**, 14 min, 0.68 μmol (89%, based on calculated ε_{260}: 32.2 $mM^{-1}cm^{-1}$). LRMS calcd for $[M-H]^-$ ($C_{29}H_{33}N_{14}O_{13}P_2$) : 847.1827, obsd, 847.2.

*Photolytic conversion of **13a** to **14a*** Compound **13a** (2 nmol) was dissolved in 20 μL of 10 mM NaOH, pH 10. The mixture was then transferred to a microcapillary pipette (100 μL, Drummond Scientific) and was UV radiated for 5 h at 4 °C. The reaction mixture was then analyzed by analytical reverse-phase HPLC (Column: Varian C18 Microsorb 100, 250 × 4.6 mm; Solution A: 25 mM triethylammonium bicarbonate, 2.5 % acetonitrile, pH 7.0, Solution B: 100% acetonitrile; Gradient: 0-40% B over 20 min; Flow rate: 15 mL/min). Compound, retention time, yield: **13a**, 9.4 min, 1.9 nmol (95%). LRMS calcd for $[M-H]^-$ ($C_{19}H_{24}N_{13}O_7P_2$) : 608.1397, obsd, 608.1.