Supporting Information

Enhanced Nonenzymatic RNA Copying with 2-Aminoimidazole Activated Nucleotides

Li Li, Noam Prywes[†], Chun Pong Tam, Derek K. O'Flaherty, Victor S. Lelyveld, Enver Cagri Izgu, Ayan Pal and Jack W. Szostak[*]

Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, United States

[*]To whom correspondence should be addressed:
CCIB 7215, Simches Research Center
185 Cambridge Street
Massachusetts General Hospital
Boston, Massachusetts 02114
E-mail: szostak@molbio.mgh.harvard.edu

[†]Present address: Department of Plant and Environmental Sciences, The Weizmann Institute of Science, P.O. Box 26, Rehovot 76100, Israel

Contents

List of Figures

List of Tables

1 Materials and Methods

1.1 Abbreviations

CV, column volume
DCM, dichloromethane
DIAD, diisopropyl azodicarboxylate
DIPEA, *N*,*N*-diisopropylethylamine
DMSO, dimethyl sulfoxide
DPDS, 2,2′-dipyridyldisulfide
EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EDTA, ethylenediaminetetraacetic acid
HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC, high-performance liquid chromatography
HRMS, high-resolution mass spectrometry
LC-MS, liquid chromatography-mass spectrometry
NMR, nuclear magnetic resonance
PAGE, polyacrylamide gel electrophoresis
Q-TOF, quadrupole time-of-flight
TEAA, triethylammonium acetate
TEAB, triethylammonium bicarbonate
THF, tetrahydrofuran
Tris, tris(hydroxymethyl)aminomethane

1.2 General information

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. All nucleoside-5′-monophosphate free-acid compounds were purchased from Santa Cruz Biotechnology (Dallas, TX). 2-bromoimidazole was purchased from AstaTech (Bristol, PA). 2-thiouridine was purchased from Carbosynth (Compton, Berkshire, UK). 2-aminoimidazole hemisulfate were purchased from Alfa Aesar (Ward Hill, MA). Reverse phase flash chromatography was performed using a prepacked RediSep Rf Gold C18Aq 50 g column from Teledyne Isco (Lincoln, NE).

^{1}H and ^{13}C NMR spectra were acquired at 400 MHz and 100 MHz, respectively, on a Varian Oxford AS-400 NMR spectrometer. Chemical shifts are reported in parts per million. ^{1}H NMR was referenced using the solvent resonance as the internal standard (HDO, 4.79 ppm at 25 °C and 4.92 ppm at 12 °C) [1]. ^{13}C NMR was referenced using the solvent resonance ($CDCl_3$, 77.2 ppm) or the CH_3 signal of acetone as the internal standard (30.89 ppm).[1] Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quadruplet, m = multiplet), integration, and coupling constants. HRMS was carried out on an Agilent 6520 QTOF LC-MS.

1.3 Procedures for nucleoside-5′-phosphoroimidazolide synthesis

1.3.1 General procedure A for nucleoside-5′-phosphoroimidazolide synthesis

To a solution of nucleoside-5′-monophosphate free acid (1 mmol, 1 equiv.), imidazole (5 mmol, 5 equiv.), triethylamine (4 mmol, 4 equiv.) in DMSO (30 mL) was added triphenylphosphine (9 mmol, 9 equiv.) and DPDS (10 mmol, 10 equiv.) under stirring at room temperature. After 30 min, the reaction mixture was added to a pre-chilled solution of acetone/diethyl ether/triethylamine (400/250/30 mL) to which 2 mL of saturated $NaClO_4$ solution in acetone had been added. After the precipitate had settled out, the majority of the supernatant was removed using pipette-suction and the remaining suspension was centrifuged at 4,000 rpm for 5 min. The pellets were washed twice with acetone (10 mL). The product was purified by reverse phase flash chromatography using gradient elution between (A) aqueous 25 mM TEAB buffer (pH 7.5) and (B) acetonitrile. The sample was eluted between 0% and 15% B over 8 CVs with a flow rate of 40 mL/min. Fractions containing products were pooled and lyophilized at -15 °C.

1.3.2 Procedures for nucleoside-5′-phosphoro-(2-aminoimidazolide) synthesis

The synthesis of nucleoside-5′-phosphoro-(2-aminoimidazolide) was carried out as described above with the exception that the pH of flash-chromatography fractions containing products were adjusted to 10 using NaOH before lyophilization.

1.3.3 General procedure B for nucleoside-5′-phosphoroimidazolide synthesis

An aqueous solution (10 mL) of nucleoside-5′-monophosphate disodium salt (1 mmol, 1 equiv.), EDC (5 mmol, 5 equiv.), and imidazole (4 mmol, 4 equiv.) was adjusted to a pH of 7 and allowed to stir for 4 h at room temperature. The product was purified by reverse phase flash chromatography using gradient elution between (A) aqueous 25 mM TEAB buffer (pH 7.5) and (B) acetonitrile. The sample was eluted between 0% and 15% B over 8 CVs with a flow rate of 40 mL/min. Fractions containing product were pooled and lyophilized at -15 °C.

1.4 Measurement of phosphoroimidazolide hydrolysis rate

Hydrolysis of phosphoroimidazolides **1a-pG** and **1j-pG** was performed at 25 °C in reaction buffers containing 200 mM HEPES (pH 8.0), 50 mM $MgCl_2$, 400 mM NaCl, 0.1% (v/v) trimethyl phosphate, and 10% (v/v) D_2O. The hydrolysis reactions were initiated by adding 5 mM **1a-pG** or **1j-pG** and monitored at 25 °C by ^{31}P NMR spectroscopy on a Varian 400 MHz NMR spectrometer (Oxford AS-400) using previously published procedures.[2] The hydrolysis rate constant of each phosphoroimidazolide was calculated from three independent experiments.

1.5 Oligonucleotide synthesis

1.5.1 Primers, templates, and complementary DNAs or RNAs

All oligonucleotides used in this study are listed in Table S2. Synthetic oligonucleotides were either purchased from Integrated DNA Technologies or prepared by solid-phase synthesis using an Expedite 8909 DNA/RNA synthesizer with reagents and phosphoramidites purchased from Glen Research (Sterling, VA) and Chemgenes (Wilmington, MA), respectively. Oligonucleotides prepared in-house were deprotected and purified by ion-exchange HPLC on a PA-100 column (Dionex, Sunnyvale, CA) using gradient elution between (A) aqueous 25 mM Tris buffer (pH 8.0) and (B) aqueous 25 mM Tris buffer (pH 8.0) with 1 M NaCl with a flow rate of 10 mL/min. The collected fractions were desalted and concentrated with Amicon Ultra-15 centrifugal filters (Millipore, Billerica, MA) to a final concentration of 100 μM.

1.5.2 Trinucleotides activated by 2-aminoimidazole

All activated trinucleotides used in this study are listed in Table S3. Trinucleotides were prepared by solid-phase synthesis using a MerMade 6 DNA/RNA synthesizer (Bioautomation, Plano, TX). The trinucleotides were chemically phosphorylated using bis(2-cyanoethyl)-*N*,*N*-diisopropyl phosphoramidite (Chemgenes). The 5′-phosphorylated trinucleotides were deprotected and purified by reverse phase flash chromatography using gradient elution between (A) aqueous 25 mM TEAB buffer (pH 7.5) and (B) acetonitrile. The sample was eluted between 0% and 10% B over 8 CVs with a flow rate of 40 mL/min. The collected fractions were dried *in vacuo*.

To a solution of trinucleotides (10 μmol, 1 equiv.) and triethylamine (40 equiv.) in DMSO (5 mL) was added 2-aminoimidazole hemisulfate (30 equiv.), triphenylphosphine (30 equiv.), and DPDS (30 equiv.) under stirring. After 5 h the mixture was added to a pre-chilled mixture of acetone/diethyl ether/triethylamine (80/50/6 mL) to which 0.4 mL of saturated $NaClO_4$ solution in acetone had been added. The precipitate was pelleted by centrifugation at 4,000 rpm for 5 min and washed twice with acetone (10 mL). The pellet was dried under vacuum, resuspended in deionized water, and purified by reverse phase HPLC on a 250 mm × 21.2 mm (length × i.d.) Eclipse XDB-C18 PrepHT column with 7 μm particle size (Agilent, Santa Clara, CA) using gradient elution between (A) aqueous 25 mM TEAB buffer (pH 7.5) and (B) acetonitrile. The sample was eluted between 3% and 12% B over 20 min with a flow rate of 10 mL/min. The fractions containing the product were combined, and the pH was adjusted to 10. The solution was frozen in liquid nitrogen and lyophilized overnight to yield colorless powder.

1.6 Primer extension reaction

For primer extension using only monomers, the primer-template duplex was first annealed in a 20 μL solution containing 6 μM primer, 20 μM template, 50 mM HEPES (pH 7.5), 50 mM NaCl and 1 mM EDTA (pH 8.0) by heating at 94 °C for 2 min and cooling down to 25 °C at a rate of 0.1 °C/s. The final reaction mixture (20 μL) contained 5 μL of annealed primer-template duplex mixture, 50 mM $MgCl_2$, 200 mM HEPES (pH 8.0), and the reaction was initiated by addition of 20 mM activated monomers. The stock solutions (100 mM) of 2-methylimidazole activated monomers and 2-aminoimidazole activated monomers had pH around 8.3 and 9.6, respectively.

The trimer-assisted primer extension was performed as described above, with the exception that the primer-template annealing reaction contained 4.8 μM primer and 6 μM template, and the final reaction mixture contained 0.5 mM activated trimers and 10 mM activated monomers.

Aliquots (1 μL each) were removed at given time points and mixed with 29 μL of quenching buffer containing 12.5 mM EDTA (pH 8.0), 75 μM complementary DNA or RNA, and 90% (v/v) formamide. The sample was heated at 95 °C for 1 min, and a 3-μL aliquot was separated by 20% (19:1) denaturing PAGE with 7 M urea. The gel was scanned using a Typhoon 9410 scanner, and the bands were quantified using the ImageQuant TL software. The primers, templates, complementary DNAs and RNAs, and/or trimers used in the primer extension assays are listed below.

Sequence added	Reference	Primer	Template	cDNA/cRNA	Activated trimers
CGG	Figs. 1 and S1, Table 1	r002	r009	r013d	n.a.
GGG	Fig. 2, Table 1	r002	r011	or011c	n.a.
GCC	Fig. S1, Table 1	r002	r010	r014d	n.a.
CCC	Table 1	r002	r012	r015d	n.a.
AGG	Table 1	r002	r036	r039d	n.a.
UGG	Table 1	r002	r037	r040d	n.a.
s^2UGG	Table 1	r002	r037	r040d	n.a.
UGAC	Fig. 3a	r002cy3	r038	o112c	t23, t24, t15, t31
UGAC	Fig. 3b	o134P	o134T	n.a.	t23, t24, t15, t31
UU	Fig. 4a	r002cy3	o136T	o136d	t18, t02
s^2Us^2U	Fig. 4b	r002cy3	o136T	o136d	t18, t02
UAGC	Fig. 4c	r098P	o103T	o103c	t14, t05, t06, t26
AUGAGGC	Fig. 4d	r052P	r053T	r054d	t18, t02, t03, t04, t05, t06

1.7 LC-MS analysis of primer extension reaction

The primer extension reaction (30 μL) containing 10 μM primer, 12 μM template, 50 mM $MgCl_2$, 200 mM HEPES (pH 8.0), 0.5 mM of activated trimers, and 10 mM activated monomers was allowed to react for 1 day, quenched by adding 3 μL of 0.5 M EDTA (pH 8.0) and 6 μL of 5 M ammonium acetate (pH 5.0), and precipitated by 90 μL of EtOH at -80 °C for 2 hours. The precipitant was pelleted by centrifugation at 12,000 rpm for 30 minutes, washed twice by 300 μL 67% (v/v) EtOH and resuspended in 25 μL of LC-MS grade water. The solution was desalted by ion pairing reverse phase (IP-RP) purification on a C18 ZipTip pipette tip (Millipore, Billerica, MA). The tip was wetted with acetonitrile and equilibrated with 2 M TEAA prior to sample binding. Extensive washing with 10 mM TEAA was followed by elution in 50% acetonitrile. The eluate was dried under vacuum, followed by resuspension in LC-MS grade water.

The eluted sample was separated and analyzed on an Agilent 1200 HPLC coupled to an Agilent 6230 TOF equipped with a solvent degasser, column oven, autosampler, and diode array detector. The sample was separated by IP-RP-HPLC on a 100 mm × 1 mm (length × i.d.) Xbridge C18 column with 3.5 μm particle size (Waters, Milford, MA) using gradient elution between (A) aqueous 200 mM 1,1,1,3,3,3-hexafluoro-2-propanol with 1.25 mM triethylamine, pH 7.0, and (B) methanol, where the sample was eluted between 2.5% and 20% B over 28.5 min with a flow rate of 125 μL/min at 60 °C. Samples were analyzed in negative mode from 239 m/z to 3200 m/z with a scan rate of 1 spectrum/s and the following instrument settings: drying gas flow, 8 L/min; drying gas temperature, 325 °C; nebulizer pressure, 30 psig; capillary voltage, 3500 V; fragmentor, 200 V; and skimmer, 65 V. Extracted ion chromatograms were generated with the Find by Formula algorithm in Agilent's MassHunter Qualitative Analysis software using a chemical formula database of possible primer extension products. The database was generated by calculating the composition of all possible random RNA extensions beyond the primer sequence up to +8.

1.8 Preparation and characterization of phosphoroimidazolides

1.8.1 Adenosine-5′-phosphoro-(2-aminoimidazole)

The adenosine-5′-phosphoro-(2-aminoimidazole) was prepared on a 1-mmol scale using adenosine-5′-monophosphate (1 equiv.), 2-aminoimidazole hemisulfate (10 equiv.), triethylamine (13 equiv.), DPDS (10 equiv.), triphenylphosphine (9 equiv.) in DMSO (30 mL) following procedures in section 1.3.2.
^{1}H NMR (400 MHz, D_2O, 25 °C) δ: 8.19 (s, 1H), 8.06 (s, 1H), 6.62 (m, 1H), 6.42 (m, 1H), 5.97 (d, 1H, J = 5.2 Hz), 4.66 (t, 1H, J = 5.2 Hz), 4.35 (t, 1H, J = 4.8 Hz), 4.27 (m, 1H), 4.05 (m, 2H). Peaks corresponding to residual TEAB observed at 3.07 and 1.20 ppm. $^{13}C\{^1H\}$ NMR (100 MHz, D_2O, 12°C) δ: 155.9, 153.3, 152.4 (d, J = 4.0 Hz), 149.3, 140.1, 124.8 (d, J = 11.3 Hz), 119.0, 115.9 (d, J = 6.0 Hz), 87.8, 83.9 (d, J = 8.8 Hz), 74.7, 70.9, 65.8 (d, J = 5.5 Hz). Peaks corresponding to residual TEAB observed at 165.3, 47.1, and 9.0 ppm. HRMS (Q-TOF) m/z: $[M-H]^-$ Calcd for $C_{13}H_{16}N_8O_6P$ 411.0936; Found: 411.0940.

1.8.2 Cytidine-5′-phosphoro-(2-aminoimidazole)

The cytidine-5′-phosphoro-(2-aminoimidazole) was prepared on a 1-mmol scale using cytidine-5′-monophosphate (1 equiv.), 2-aminoimidazole hemisulfate (10 equiv.), triethylamine (13 equiv.), DPDS (10 equiv.), triphenylphosphine (9 equiv.) in DMSO (30 mL) following procedures in section 1.3.2.
^{1}H NMR (400 MHz, D_2O) δ: 7.79 (d, 1H, J = 7.6 Hz), 6.75 (m, 1H), 6.56 (m, 1H), 6.00 (d, 1H, J = 7.6 Hz), 5.90 (d, 1H, J = 4.4 Hz), 4.21 (t, 1H, J = 4.4 Hz), 4.15 (m, 3H), 4.04 (m, 1H). Peaks corresponding to residual TEAB observed at 3.07 and 1.21 ppm. $^{13}C\{^1H\}$ NMR (100 MHz, D_2O) δ: 166.7, 158.3, 152.5 (d, J = 3.9 Hz), 141.6, 125.0 (d, J = 11.2 Hz), 116.1 (d, J = 6.0 Hz), 97.0, 90.1, 82.9 (d, J = 8.9 Hz), 74.7, 69.9, 65.3 (d, J = 5.5 Hz). Peaks corresponding to residual TEAB observed at 166.2, 47.0, and 9.1 ppm. HRMS (Q-TOF) m/z: $[M+H]^+$ Calcd for $C_{12}H_{18}N_6O_7P$ 389.0969; Found: 389.0993.

1.8.3 Guanosine-5′-phosphoro-(2-isopropylimidazole)

The guanosine-5′-phosphoro-(2-isopropylimidazole) was prepared on a 1-mmol scale using guanosine-5′-monophosphate (1 equiv.), 2-isopropylimidazole (4 equiv.) and EDC (5 equiv.) in acetone:water (1:3, v/v, 10 mL) following procedures in section 1.3.3.

^{1}H NMR (400 MHz, D_2O, 12 °C) δ: 7.89 (s, 1H), 7.08 (s, 1H), 6.78 (s, 1H), 5.83 (d, 1H, $J = 5.2$ Hz), 4.84 (t, 1H, $J = 5.2$ Hz), 4.47 (t, 1H, $J = 5.2$ Hz), 4.25 (m, 1H), 4.01 (m, 2H), 3.34 (sep, 1H, $J = 6.8$ Hz), 1.04 (dd, 6H, $J = 6.8, 3.2$ Hz). ^{13}C{^{1}H} NMR (100 MHz, D_2O, 12°C) δ: 160.3, 158.4 (d, $J = 5.9$ Hz), 154.9, 152.2, 138.1, 126.6 (d, $J = 11.1$ Hz), 121.9 (d, $J = 6.7$ Hz), 117.1, 88.1, 83.9 (d, $J = 9.6$ Hz), 73.6, 71.0, 66.3 (d, $J = 5.3$ Hz), 27.7, 21.9 (d, $J = 15.4$ Hz). HRMS (Q-TOF) m/z: $[M-H]^-$ Calcd for $C_{16}H_{21}N_7O_7P$ 454.1246; Found: 454.1261.

1.8.4 Guanosine-5′-phosphoro-(2-phenylimidazole)

The guanosine-5′-phosphoro-(2-phenylimidazole) was prepared on a 1-mmol scale using guanosine-5′-monophosphate (1 equiv.), 2-phenylimidazole (4 equiv.) and EDC (5 equiv.) in acetone:water (1:3, v/v, 10 mL) following procedures in section 1.3.3.
^{1}H NMR (400 MHz, D_2O, 12 °C) δ: 7.76 (s, 1H), 7.58 (d, 2H, $J = 7.2$ Hz), 7.35 (m, 2H), 7.29 (m, 2H), 7.01 (s, 1H), 5.75 (d, 1H, $J = 5.6$ Hz), 4.67 (t, 1H, $J = 5.6$ Hz), 4.15 (m, 1H), 4.04 (t, 1H, $J = 4.8$ Hz), 3.96 (m, 2H). ^{13}C{^{1}H} NMR (100 MHz, D_2O, 12°C) δ: 160.4, 154.9, 152.2, 150.7 (d, $J = 4.2$ Hz), 138.0, 131.3, 129.9, 129.6, 128.5, 127.9 (d, $J = 11.1$ Hz), 124.4 (d, $J = 5.9$ Hz), 117.1, 88.1, 83.8 (d, $J = 9.9$ Hz), 73.5, 70.8, 60.1 (d, $J = 5.3$ Hz). HRMS (Q-TOF) m/z: $[M-H]^-$ Calcd for $C_{19}H_{19}N_7O_7P$ 488.1089; Found: 488.1098.

1.8.5 Guanosine-5′-phosphoro-benzimidazole

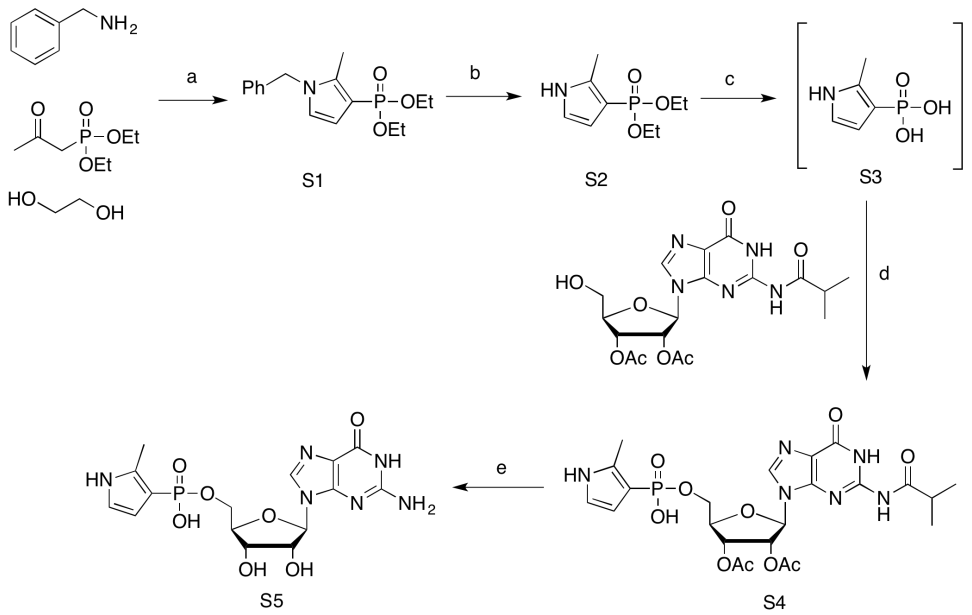
The guanosine-5′-phosphoro-benzimidazole was prepared on a 1-mmol scale using guanosine-5′-monophosphate (1 equiv.), benzimidazole (4 equiv.) and EDC (5 equiv.) in acetone:water (1:3, v/v, 10 mL) following procedures in section 1.3.3.
^{1}H NMR (400 MHz, D_2O, 12 °C) δ: 8.15 (s, 1H), 7.54 (s, 1H), 7.40 (d, 2H, $J = 8.0$ Hz), 7.03 (t, 1H, $J = 7.2$ Hz), 6.93 (t, 1H, $J = 8.0$ Hz), 5.55 (d, 1H, $J = 4.8$ Hz), 4.72 (t, 1H, $J = 5.2$ Hz), 4.38 (t, 1H, $J = 4.8$ Hz), 4.23 (m, 1H), 4.07 (m, 2H). ^{13}C{^{1}H} NMR (100 MHz, D_2O, 12°C) δ: 159.3, 153.5, 151.2, 146.0 (d, $J = 6.7$ Hz), 143.2 (d, $J = 10.6$ Hz), 138.7, 133.4 (d, $J = 4.9$ Hz), 124.0, 123.8, 118.9, 116.9, 114.1, 89.1, 84.0 (d, $J = 9.8$ Hz), 73.0, 70.9, 66.8 (d, $J = 5.5$ Hz). HRMS (Q-TOF) m/z: $[M-H]^-$ Calcd for $C_{17}H_{17}N_7O_7P^-$ 462.0933; Found: 462.0947.

1.8.6 Guanosine-5′-phosphoro-(2-methylbenzimidazole)

The guanosine-5′-phosphoro-(2-methylbenzimidazole) was prepared on a 1-mmol scale using guanosine-5′-monophosphate (1 equiv.), 2-methylbenzimidazole (4 equiv.) and EDC (5 equiv.) in acetone:water (1:3, v/v, 10 mL) following procedures in section 1.3.3.
^{1}H NMR (400 MHz, D_2O, 12 °C) δ: 7.58 (s, 1H), 7.56 (d, 1H, $J = 8$ Hz), 7.24 (d, 1H, $J = 8$ Hz), 7.00 (m, 1H), 6.91 (m, 1H), 5.59 (d, 1H, $J = 4.4$ Hz), 4.71 (t, 1H, $J = 4.8$ Hz), 4.43 (t, 1H, $J = 4.8$ Hz), 4.24 (dt, 1H, $J = 4.8, 4.4$ Hz), 4.12 (m, 2H), 2.59 (s, 3H). ^{13}C{^{1}H} NMR (100 MHz, D_2O, 12°C) δ: 159.5, 156.5 (d, $J = 6.9$ Hz), 153.5, 151.2, 141.7 (d, $J = 11$ Hz), 139.0, 136.0 (d, $J = 6.0$ Hz), 123.4, 123.1, 117.4, 117.1, 114.5, 89.6, 84.2 (d, $J = 8.7$ Hz), 73.2, 71.0, 67.0 (d, $J = 5.2$ Hz), 16.4. HRMS (Q-TOF) m/z: $[M-H]^-$ Calcd for $C_{18}H_{19}N_7O_7P$ 476.1089; Found: 476.1097.

1.8.7 Guanosine-5′-phosphoro-(2-chloroimidazole)

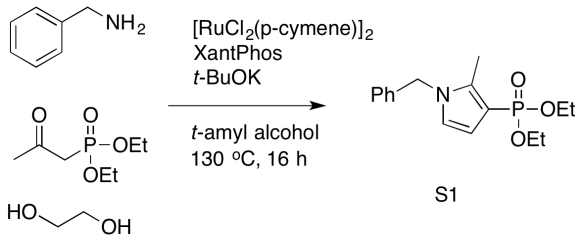
The guanosine-5′-phosphoro-(2-chloroimidazole) was prepared on a 1-mmol scale using guanosine-5′-monophosphate (1 equiv.), 2-chloroimidazole (4 equiv.) and EDC (5 equiv.) in water (10 mL) following procedures in section 1.3.3.
^{1}H NMR (400 MHz, D_2O, 12 °C) δ: 7.85 (s, 1H), 7.16 (s, 1H), 6.74 (s, 1H), 5.80 (d, 1H, $J = 5.2$ Hz), 4.81 (t, 1H, $J = 5.2$ Hz), 4.47 (t, 1H, $J = 4.8$ Hz), 4.27 (m, 1H), 4.12 (m, 2H). ^{13}C{^{1}H} NMR (100 MHz, D_2O, 12°C) δ: 160.0, 154.7, 152.1, 138.1, 133.7 (d, $J = 3.6$ Hz), 127.9 (d, $J = 10.3$ Hz), 124.5 (d, $J = 5.8$ Hz), 116.9, 88.2, 83.8 (d, $J = 9.0$ Hz), 73.8, 70.9, 66.7 (d, $J = 5.5$ Hz). HRMS (Q-TOF) m/z: $[M-H]^-$ Calcd for $C_{13}H_{14}ClN_7O_7P$ 446.0386; Found: 446.0386.



Synthetic scheme of Guanosine-5′-(2-methyl-1*H*-pyrrol-3-yl)phosphonate. Reaction conditions: (a) $[RuCl_2(p\text{-cymene})]_2$, XantPhos, *t*-BuOK, *t*-amyl alcohol, 130 °C, 16 hr; (b) Na^0, NH_3 (l), -78 °C, 30 min; (c) TMSBr, Et_3N, DCM, 4 h; then MeOH; (d) DIAD, PPh_3, DCM, 3 h; (e) 28% NH_3 in H_2O, 65 °C, 2.5 h. $[RuCl_2(p\text{-cymene})]_2$, dichloro(*p*-cymene)ruthenium(II) dimer; XantPhos, 4,5-bis(diphenylphosphino)-9,9-dimethylxanthene; TMSBr, bromotrimethylsilane; DIAD, diisopropyl azodicarboxylate.

1.8.8 Guanosine-5′-(2-methyl-1*H*-pyrrol-3-yl)phosphonate

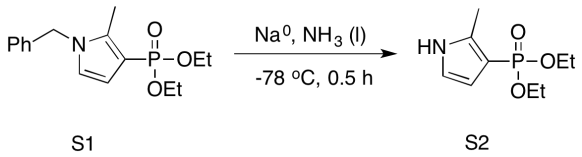
a. Diethyl (*N*-benzyl-2-methyl-3-pyrrolyl)phosphonate (compound **S1**)



This procedure was adapted from Beller and coworkers.[3] To a septa-sealed and oven-dried glass vessel, the following reagents were added: dichloro(*p*-cymene)ruthenium(II) dimer (92 mg, 0.15 mmol, 0.05 equiv.), 4,5-bis(diphenylphosphino)-9,9-dimethylxanthene (174 mg, 0.3 mmol, 0.1 equiv.), and potassium *tert*-butoxide (67 mg, 0.6 mmol, 0.2 equiv.), followed by a desiccator-dried magnetic stir bar. After three rounds of argon purging, *tert*-amyl alcohol (6 mL), benzylamine (0.48 mL, 4.4 mmol, 1.5 equiv.), ethylene glycol (0.37 mL, 6.6 mmol, 2.2 equiv.), and diethyl (2-oxopropyl)phosphonate (0.37 mL, 3 mmol, 1 equiv.) were successively delivered to the reaction vessel via needle injections. The reaction vessel was then sealed by a Teflon screw cap, and the reaction was stirred for 16 hours at 130 °C. The reaction mixture turned dark brick-red upon homogenization of solid reagents in *tert*-amyl alcohol. The solvent of the crude reaction mixture was then evaporated *in vacuo*. The crude mixture was purified by normal-phase flash chromatography on a 40 g pre-packed silica column using gradient elution between (A) DCM and (B) MeOH, where compound **S1** was eluted between 0% and 10% B over 20 CVs as a dark red oil.
^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.32 (m, 3H), 6.97 (d, 2H, $J = 6.8$ Hz), 6.61 (dd, 1H, $J = 4.4, 2.8$ Hz), 6.37 (dd, 1H, $J = 4.4, 2.8$ Hz), 5.02 (s, 2H), 4.06 (m, 4H), 2.35 (d, 3H, $J = 1.6$ Hz), 1.29 (t, 6H, $J = 6.8$ Hz)
^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ: 136.9, 136.4 (d, $J = 25.2$ Hz), 128.8, 127.7, 126.4, 121.7 (d, $J = 14.7$ Hz),

111.7 (d, J = 11.9 Hz), 105.4 (d, J = 214.8 Hz), 61.3 (d, J = 5.1 Hz), 50.4, 16.3 (d, J = 6.7 Hz), 11.1. HRMS (Q-TOF) m/z: $[M + H]^+$ Calcd for $C_{16}H_{23}NO_3P$ 308.1416; Found: 308.1436.

b. Diethyl (2-methyl-1H-pyrrol-3-yl)phosphonate (compound **S2**)



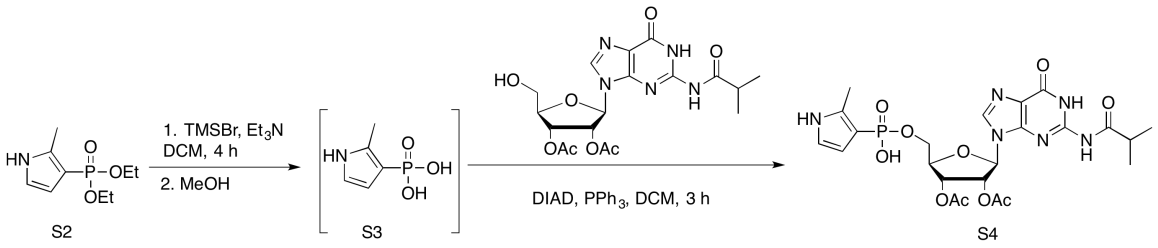
S1 S2

A Schlenk tube and stir bar were flame-dried, sealed with a rubber septum, and allowed to cool to room temperature. The atmosphere within the flask was purged with argon in three rounds. To deliver ammonia gas into the Schlenk tube, a thick-walled steel cylinder containing pressurized ammonia gas was connected to a corrosive lecture-bottle gas regulator, and the tygon tubing was connected with 12-inch, 16-gauge needle assembly. All joints of the assembly were sealed with Teflon tape. The gas-delivery assembly was securely fastened to a stainless steel frame at the back of a fume hood. The Schlenk tube was cooled in an acetone-dry ice bath, and dry ammonia gas was delivered directly into the Schlenk tube from the gas-delivery assembly. Liquid ammonia condensation was allowed to proceed until the volume reached 8 mL, after which the gas-delivery needle was removed from the Schlenk tube.

In a separate beaker, small pieces of metallic sodium (40 mg, 1.74 mmol, 5.4 equiv.) were first washed with anhydrous hexanes to wash off the paraffin oil coating and transferred into the condensed liquid ammonia to produce a deep blue mixture. A 0.25 M THF solution (1.3 mL) of phosphonate **S1** (100 mg, 0.325 mmol, 1 equiv.) was prepared in an oven-dried pear-shaped flask, added drop-wise into the sodium/ammonia mixture, and allowed to react at -78 °C for 30 min.

To quench the reaction, solid ammonium chloride (350 mg, 6.5 mmol, 20 equiv.) was added to the reaction mixture. The quenched reaction mixture was then allowed to warm up to room temperature for 4 h. The crude residue was diluted with water (30 mL) and methanol (10 mL), and the resulting solution was extracted twice with DCM (30 mL). The combined organic layer was dried with anhydrous sodium sulfate, followed by solvent evaporation *in vacuo*. The crude mixture was purified by normal phase flash chromatography on a 12 g pre-packed silica column using gradient elution between (A) DCM and (B) MeOH, where compound **S2** (60 mg, 0.28 mmol, 86%) was eluted between 0% and 20% B over 20 CVs as a pale-yellow oil.
^{1}H NMR (400 MHz, $CDCl_3$) δ: 10.17 (s, 1H), 6.55 (m, 1H), 6.25 (m, 1H), 3.98 (m, 4H), 2.35 (s, 3H), 1.27 (t, 6H, J = 7.1 Hz). $^{13}C\{^1H\}$ NMR (100 MHz, $CDCl_3$) δ: 136.5 (d, J = 25.5 Hz), 117.4 (d, J = 16.1 Hz), 112.2 (d, J = 13.4 Hz), 102.8 (d, J = 216.0 Hz), 61.6 (d, J = 5.2 Hz), 16.4 (d, J = 6.8 Hz), 12.7. HRMS (Q-TOF) m/z: $[M + H]^+$ Calcd for $C_9H_{17}NO_3P$ 218.0946; Found: 218.0943.

c. N^2-isobutyryl-2′-3′-diacetylguanosine (2-methyl-1H-pyrrol-3-yl)phosphonate (compound **S4**)



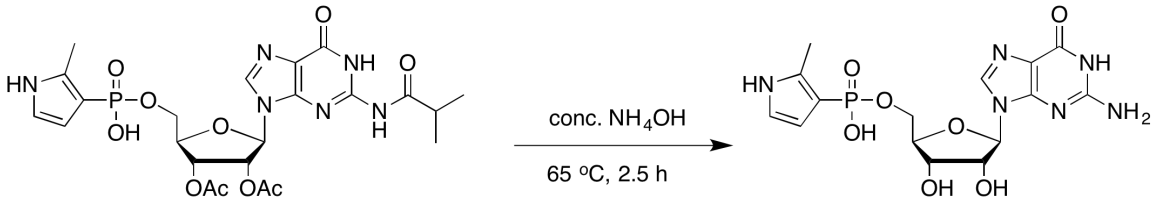
S2 S3 S4

To dried compound **S2** (113 mg, 0.52 mmol, 1 equiv.) was added anhydrous DCM (5.2 mL) and triethylamine (0.36 mL, 2.6 mmol, 5 equiv.) to yield a yellow homogeneous solution. Bromotrimethylsilane (0.68 mL, 5.2 mmol, 10 equiv.) was then added to the stirring mixture drop-wise, and the reaction mixture was allowed to stir for 4 h. An excess of methanol (10 mL) was added to the reaction mixture to quench the excess bromotrimethylsilane and to decompose the resultant bis(trimethylsilyl) phosphonate esters.[4] Solvents were evaporated *in vacuo*, and the residual volatile impurities and hydrogen bromide were removed by toluene co-evaporation in three rounds. The resultant solid residue (compound **S3**) was carried forward to the next step without further purification.

To a dried flask with crude material (compound **S3**, 1.1 equiv. assuming 100% conversion) was added N^2-isobutyryl-2′-3′-diacetylguanosine (204 mg, 0.47 mmol, 1 equiv.), triphenylphosphine (680 mg, 2.6 mmol, 5.5 equiv.) and anhydrous DCM (10.4 mL). The mixture was cooled at 0 °C for 5 min, and DIAD (0.53 mL, 2.60 mmol, 5.5 equiv.) was added to the mixture drop-wise via syringe over five minutes. After the reaction mixture clarified into a clear brown mixture, it was stirred at 0 °C for 30 min and then at room temperature for 2.5 h. The reaction was stopped by evaporating the solvent *in vacuo*, followed by normal phase flash chromatography in a pre-packed 40 g silica column. The column was first flushed with 10 CVs of 90:9:1 (v/v/v) DCM-MeOH-triethylamine to remove nonpolar impurities, followed by elution with 5 CVs of 99:1 (v/v) MeOH-triethylamine. The dried-down crude product was suspended in 9:1 (v/v) water-acetonitrile, and purified with reverse-phase flash chromatography using gradient elution between (A) 25 mM aqueous TEAB buffer (pH 7.5) and (B) acetonitrile, where compound **S4** was eluted between 0% and 25% B over 25 CVs. The relevant fractions were pooled and lyophilized to afford the title compound as a fluffy white solid.

^{1}H NMR (400 MHz, CD_3OD) δ: 8.18 (s, 1H), 6.53 (dd, 1H, $J = 3.5, 3.0$ Hz), 6.29 (dd, 1H, $J = 3.8, 2.9$ Hz), 6.11 (m, 2H), 5.59 (dd, 1H, $J = 5.2, 2.9$ Hz), 4.32 (m, 1H), 4.04 (m, 1H), 3.96 (m, 1H), 2.80 (m, 1H), 2.38 (d, 3H, $J = 1.6$ Hz), 2.11 (s, 3H), 1.98 (s, 3H), 1.20 (m, 6H). Peaks corresponding to residual triethylamine observed at 2.83 and 1.15 ppm. ^{13}C{^{1}H} NMR (100 MHz, CD_3OD) δ: 182.4, 171.5, 170.9, 157.9, 150.8, 150.1, 140.8, 134.5 (d, $J = 22.7$ Hz), 121.9, 116.9 (d, $J = 14.8$ Hz), 113.4 (d, $J = 12.6$ Hz), 110.6 (d, $J = 203.5$ Hz), 87.6, 84.3 (d, $J = 9.4$ Hz), 74.2, 73.5, 64.5 (d, $J = 4.0$ Hz), 36.9, 20.7, 20.4, 19.7, 19.5, 13.1. Peaks corresponding to residual triethylamine observed at 47.3 and 10.3 ppm. HRMS (Q-TOF) m/z: $[M + H]^+$ Calcd for $C_{23}H_{30}N_6O_{10}P$ 581.1761; Found: 581.1786.

d. Guanosine (2-methyl-1*H*-pyrrol-3-yl)phosphonate (compound **S5**)



S4 S5

Compound **S4** (33 mg, 0.057 mmol, 1 equiv.) was added to a high-pressure glass reaction vessel along with a stir bar and dissolved in 28% (w/v) ammonium hydroxide (1.14 mL). The vessel was sealed, and stirred for 2.5 h at 65 °C. The reaction was then allowed to equilibrate to room temperature, followed by solvent evaporation *in vacuo* to afford a white solid deposit. The crude material was resuspended in 100 mM TEAB buffer (pH 7.5), and purified by reverse phase flash chromatography using gradient elution between (A) aqueous 25 mM TEAB buffer (pH 7.5) and (B) acetonitrile, where compound **S5** was eluted between 0% and 15% B over 10 CVs with a flow rate of 40 mL/min. The yield, based on A_{260} of the pooled fractions, was 83%. The solution was then lyophilized, dissolved in 1 mL of H_2O, and precipitated in acetone:diethyl ether (12/12 mL) to which 0.1 mL of saturated $NaClO_4$ solution in acetone had been added. The precipitate was centrifuged, washed with 1:1 (v/v) acetone:diethyl ether twice, and lyophilized to afford compound **S5** as fine white granular solids.

^{1}H NMR (400 MHz, D_2O) δ: 7.88 (s, 1H), 6.65 (t, 1H, $J = 3.2$ Hz), 6.23 (t, 1H, $J = 3.2$ hz), 5.87 (d, 1H, $J = 5.6$ Hz), 4.76 (t, 1H, $J = 5.2$ Hz), 4.41 (m, 1H), 4.27 (m, 1H), 3.93 (t, 2H, $J = 4.0$ Hz), 2.22 (s, 3H). ^{13}C{^{1}H} NMR (100 MHz, D_2O) δ: 159.6, 154.5, 152.4, 138.2, 134.7 (d, $J = 23.0$ Hz), 117.2 (d, $J = 15.0$ Hz), 116.8, 112.3 (d, $J = 12.9$ Hz), 109.2, 87.6, 84.9 (d, $J = 8.9$ Hz), 74.2, 71.5, 64.1 (d, $J = 3.5$ Hz), 12.4. HRMS (Q-TOF) m/z: $[M - H]^-$ Calcd for $C_{15}H_{18}N_6O_7P$ 425.0980; Found: 425.0993.

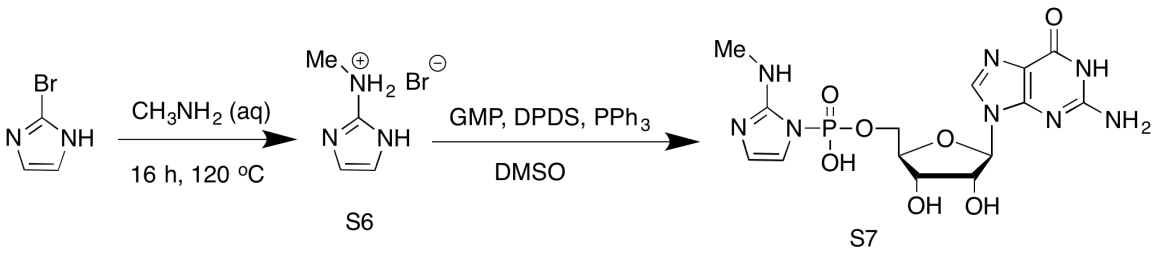
1.8.9 Guanosine-5′-phosphoro-(2-aminoimidazole)

The guanosine-5′-phosphoro-(2-aminoimidazole) was prepared on a 1-mmol scale using guanosine-5′-monophosphate (1 equiv.), 2-aminoimidazole hemisulfate (10 equiv.), triethylamine (13 equiv.), DPDS (10 equiv.), triphenylphosphine (9 equiv.) in DMSO (30 mL) following procedures in section 1.3.2.

^{1}H NMR (400 MHz, D_2O) δ: 7.86 (s, 1H), 6.63 (s, 1H), 6.43 (m, 1H), 5.84 (d, 1H, $J = 5.6$ Hz), 4.67 (t, 1H,

$J = 5.6$ Hz), 4.33 (t, 1H, $J = 4.4$ Hz), 4.23 (m, 1H), 4.03 (m, 2H). $^{13}C\{^1H\}$ NMR (100 MHz, D_2O): 166.2, 159.9, 152.4 (d, $J = 4.1$ Hz), 152.3, 136.4, 124.8 (d, $J = 11.1$ Hz), 117.9, 116.0 (d, $J = 6.0$ Hz), 87.2, 83.8 (d, $J = 8.8$ Hz), 74.1, 71.0, 66.0 (d, $J = 5.5$ Hz). HRMS (Q-TOF) m/z: $[M + H]^+$ Calcd for $C_{13}H_{18}N_8O_7P$ 429.1031; Found: 429.1050.

1.8.10 Guanosine-5′-phosphoro-(2-methylaminoimidazole)



Synthetic scheme of guanosine-5′-phosphoro-(2-methylaminoimidazole).

a. 2-(Methylamino)-1*H*-imidazole, compound **S6**

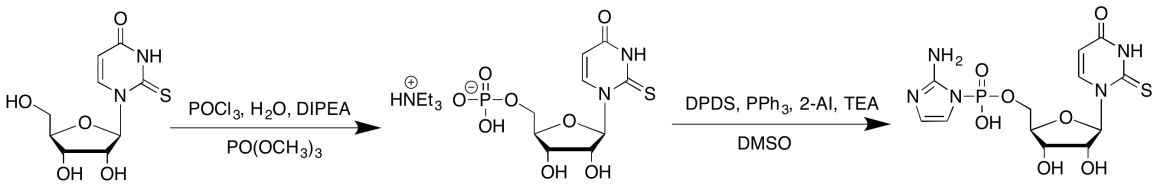
A cylindrical heavy-wall pressure vessel was charged with aqueous methylamine solution (7 mL, 40%, w/v) and 2-bromoimidazole (1 g, 6.80 mmol). After 16 h of stirring at 120 °C, the reaction was allowed to cool to room temperature. The solvent was removed *in vacuo* and the crude content was taken up in water (5 mL). The pH was adjusted to 7 using concentrated hydrobromic acid (48%, w/w). The product was purified via reverse phase flash chromatography with an acetonitrile/TEAB buffer (20 mM, pH ca. 8.0) solvent system. The product fractions were pooled into a 50 mL falcon tube and freeze-dried under high-vacuum at room temperature to afford 0.79 g (66%) of the hydrobromide salt. The product was used without further purification for the subsequent step. An analytical sample (50 mg) was additionally purified using combiflash column (silica) chromatography with a MeOH / DCM solvent system.
^1H NMR (400 MHz, D_2O) δ: 6.81 (s, 2H), 2.95 (s, 3H). $^{13}C\{^1H\}$ NMR (100 MHz, D_2O) δ: 148.45, 113.53, 29.35. Residual MeOH peak observed at 49.45 ppm. HRMS (Q-TOF) m/z: $[M + H]^+$ Calcd for $C_4H_7N_3$ 98.0713; Found: 98.0721.

b. Preparation of guanosine-5′-phosphoro-(2-methylaminoimidazole), compound **S7**

The guanosine-5′-phosphoro-(2-methylaminoimidazole), compound **S7**, was prepared on a 1-mmol scale using guanosine-5′-monophosphate (1 equiv.), imidazole **S6** (10 equiv.), triethylamine (13 equiv.), DPDS (10 equiv.), triphenylphosphine (9 equiv.) in DMSO (30 mL) following procedures in section 1.3.2.
^1H NMR (400 MHz, D_2O) δ: 7.88 (s, 1H), 6.69 (t, 1H, J = 1.5 Hz), 6.54 (t, 1H, J = 2 Hz), 5.88 (d, 1H, $J = 6$ Hz), 4.70 (t, 1H, $J = 5.5$ Hz), 4.33 (dd, 1H, $J = 5, 3.5$ Hz), 4.24 (m, 1H), 4.04-4.02 (m, 1H), 2.76 (s, 3H). $^{13}C\{^1H\}$ NMR (100 MHz, D_2O) δ: 168.0, 161.3, 154.4 (d, $J = 4.0$ Hz), 152.3, 135.9, 124.8 (d, $J = 11.5$ Hz), 118.2, 116.0 (d, $J = 6.0$ Hz), 87.0, 83.6 (d, $J = 9.0$ Hz), 73.9, 70.9, 65.8 (d, $J = 5.7$ Hz), 29.8. HRMS (ESI): m/z: $[M - H]^-$ Calcd for $C_{14}H_{19}N_8O_7P$ 441.1042; Found: 441.1034.

1.8.11 Uridine-5′-phosphoro-(2-aminoimidazole)

The uridine-5′-phosphoro-(2-aminoimidazole) was prepared on a 1-mmol scale using uridine-5′-monophosphate (1 equiv.), 2-aminoimidazole hemisulfate (10 equiv.), triethylamine (13 equiv.), DPDS (10 equiv.), triphenylphosphine (9 equiv.) in DMSO (30 mL) following procedures in section 1.3.2.
^1H NMR (400 MHz, D_2O, 12 °C) δ: 7.60 (d, 1H, $J = 7.6$ Hz), 6.76 (m, 1H), 6.57 (m, 1H), 5.96 (d, 1H, $J = 5.2$ Hz), 5.81 (d, $J = 7.6$ Hz), 4.24 (m, 1H), 4.17 (m, 2H), 4.06 (m, 2H). Peaks corresponding to residual TEAB observed at 3.12 and 1.23 ppm. $^{13}C\{^1H\}$ NMR (100 MHz, D_2O, 12°C) δ: 175.9, 159.0, 152.5 (d, $J = 4.2$ Hz), 140.7, 125.0 (d, $J = 11.2$ Hz), 116.1 (d, $J = 6.0$ Hz), 103.6, 89.4, 82.9 (d, $J = 8.8$ Hz), 74.3, 70.3, 65.7 (d, $J = 5.4$ Hz). Peaks corresponding to residual TEAB observed at 165.0, 47.1, and 9.0 ppm. HRMS (Q-TOF) m/z: $[M + H]^+$ Calcd for $C_{12}H_{17}N_5O_8P$ 390.0809; Found: 390.0829.



Synthetic scheme of 2-thiouridine-5′-phosphoro-(2-aminoimidazole).

1.8.12 2-thiouridine-5′-phosphoro-(2-aminoimidazole)

To a pre-chilled stirring solution of $POCl_3$ (0.2 mL, 2.1 mmol), H_2O (4 μL, 0.22 mmol) in trimethylphosphate (5 mL) was added 2-thiouridine (0.15 g, 0.58 mmol) and the resultant solution was allowed to stir at 0 °C. After complete solubilization of the nucleoside, DIPEA (52 μL, 0.30 mmol) was added drop-wise to the stirring reaction. Three additional portions of DIPEA were added (52 μL, 52 μL, and 20 μL, respectively) with 20 min interval. Once the starting material disappeared, the reaction was quenched using 1 M TEAB (10 mL, pH 7.6). The product was partially purified by reverse phase flash chromatography using gradient elution between (A) aqueous 25 mM TEAB (pH 7.5) and (B) acetonitrile. Fractions containing product were pooled and the solvent was removed by freeze-drying under high-vacuum at room temperature. The 2-thiouridine-5′-monophosphate was used as is for the subsequent step without further purification.

The 2-thiouridine-5′-phosphoro-(2-aminoimidazole) was prepared on a 300-μmol scale using 2-thiouridine-5′-monophosphate (1 equiv.), 2-aminoimidazole hemisulfate (10 equiv.), triethylamine (13 equiv.), DPDS (10 equiv.), triphenylphosphine (9 equiv.) in DMSO (15 mL) following procedures in section 1.3.2.
^{1}H NMR (400 MHz, D_2O) δ: 7.72 (d, 1H, $J = 8$ Hz), 6.65 (m, 1H), 6.58 (d, 1H, $J = 3$ Hz), 6.49 (m, 1H), 5.94 (d, 1H, $J = 8$ Hz), 4.17 (m, 1H), 4.12-3.94 (m, 4H). Peaks corresponding to residual TEAB observed at 3.04 ppm and 1.13 ppm. ^{13}C{^{1}H} NMR (100 MHz, D_2O) δ: 179.8, 172.3, 152.5 (d, $J = 4.0$ Hz), 141.5, 125.0 (d, $J = 11.0$ Hz), 116.1 (d, $J = 6.0$ Hz), 107.1, 94.0, 82.7 (d, $J = 9.0$ Hz), 75.4, 69.4, 65.0 (d, $J = 5.5$ Hz). Peaks corresponding to residual TEAB observed at 163.8, 47.2, and 9.0 ppm. HRMS (Q-TOF) m/z: $[M - H]^-$ Calcd for $C_{12}H_{16}N_5O_7PS$ 406.0581; Found: 406.0606.

References

[1] Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997,** *62,* 7512–7515.

[2] Li, L.; Lelyveld, V. S.; Prywes, N.; Szostak, J. W. *J. Am. Chem. Soc.* **2016,** *138,* 3986–3989.

[3] Zhang, M.; Fang, X.; Neumann, H.; Beller, M. *J. Am. Chem. Soc.* **2013,** *135,* 11384–11388.

[4] Matulic-Adamic, J.; Haeberli, P.; Usman, N. *J. Org. Chem.* **1995,** *60,* 2563–2569.

2 Supplementary Figures

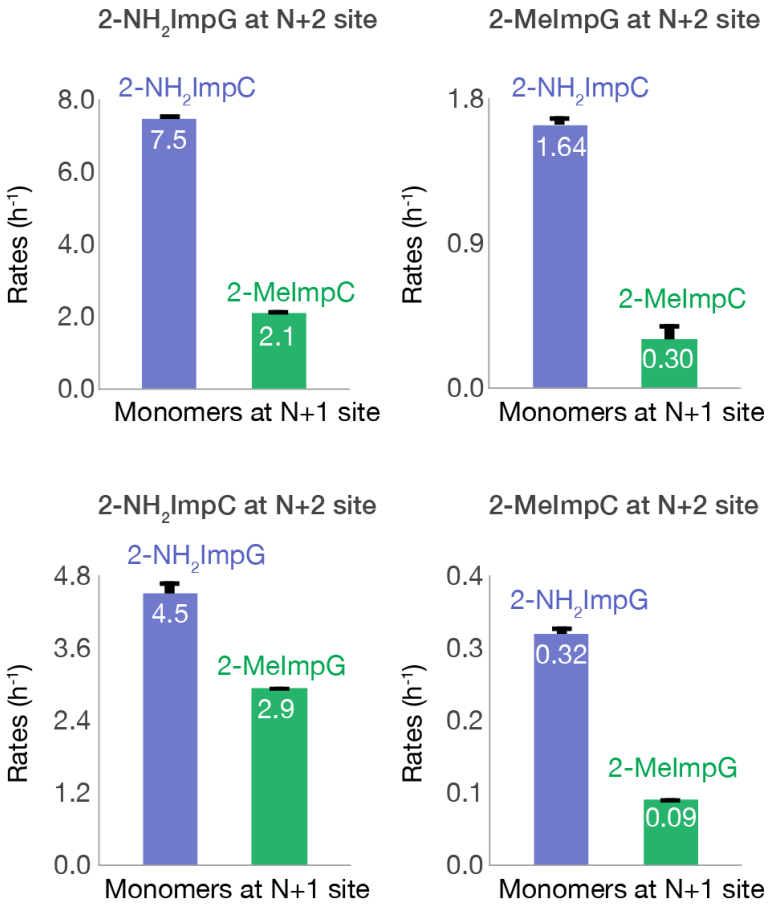


Figure S1: **Monomers activated by 2-AI leaving group react faster at N+1 sites.** Reaction rates of 2-AI activated monomers (blue) versus 2-MI activated monomers (green) at N+1 sites with different monomers at N+2 sites. Reaction conditions are the same as reported in Figure 1. Error bars represent s.e.m. derived from triplicated experiments.

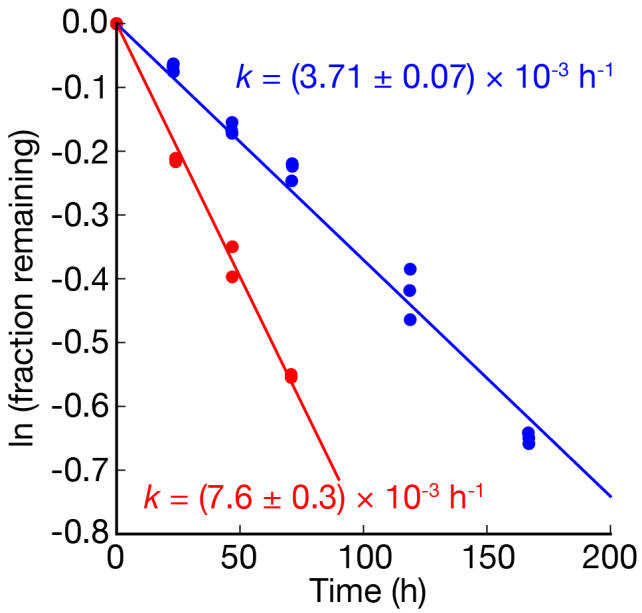


Figure S2: **Kinetic analysis of monomer hydrolysis at primer-extension condition.** Plot of the remaining monomer fraction as a function of time for 2-aminoimidazole activated monomers (red) and 2-methylimidazole activated monomers (blue). Reaction conditions: 5 mM monomers, 50 mM HEPES buffer (pH 8.0), 50 mM $MgCl_2$, 400 mM NaCl, and 10% D_2O. Hydrolysis rates were determined from linear least-squares fits of the data from three independent experiments.

3 Supplementary tables

Table S1: **Observed vs. calculated neutral masses for possible primer extension products (+1 to +7) observed by high resolution LC-MS.** The +3* product corresponds compositionally to the addition of s^2U rather than A in the third position. The +7 extension product cannot be distinguished within the mass resolution of the instrument from an alternative double-mutant (+7†) that replaces G and U with A and s^2U, although we see no intermediate products larger than +3* consistent with this double-mutant.

	Mass (obs., Da)	Mass (calc., Da)	Error (ppm)	Retention time (min)
Primer	3856.5718	3856.5772	-1.42	15.5
+1	4178.5767	4178.5797	-0.72	16.4
+2	4523.6175	4523.6271	-2.12	17.1
+3*	4845.6166	4845.6296	-2.67	17.7
+3	4852.6661	4852.6796	-2.78	17.8
+4	5157.7180	5157.7209	-0.57	18.4
+5	5502.7495	5502.7684	-3.42	18.9
+6	5847.7957	5847.8158	-3.44	19.4
+7	6153.8165	6153.8411	-4.00	19.8
+7†	6153.8186	6153.8233	-0.77	19.8

Table S2: Sequences of oligonucleotides used in this study[†]

Name	R or DNA	Sequence ($5'\rightarrow 3'$)
r002	RNA	FAM-AGUGAGUAACGC
r002cy3	RNA	Cy3-AGUGAGUAACGC
r009	RNA	AACCCGGCGUUACUCACU
r010	RNA	AAGGGCGCGUUACUCACU
r011	RNA	AACCCCGCGUUACUCACU
r012	RNA	AAGGGGGCGUUACUCACU
r013d	DNA	AGTGAGTAACGCCGGGTT
r014d	DNA	AGTGAGTAACGCGCCCTT
r015d	DNA	AGTGAGTAACGCCCCCTT
or011c	RNA	AGUGAGUAACGCGGGGUU
r036	RNA	AACCCUGCGUUACUCACU
r037	RNA	AACCCAGCGUUACUCACU
r038	RNA	AACCGUCAGCGUUACUCACU
r039d	DNA	AGTGAGTAACGCAGGGTT
r040d	DNA	AGTGAGTAACGCTGGGTT
o112c	RNA	AGUGAGUAACGCUGACGGUU
o134P	RNA	AGUGAGUAACGC
o134T	RNA	AACCGUCAGCGUUACUCACUAAAAA
o136T	RNA	CUCAAGCGUUACUCACU
o136d	DNA	AGTGAGTAACGCTTGAG
r098P	RNA	Cy3-AAGGUCACCG
o103T	RNA	UGCGCUACGGUGACCU
o103C	RNA	AGGUCACCGUAGCGCA
r052P	RNA	FAM-CGCUCGACUG
r053T	RNA	GCGCCUCAUCAGUCGAGCG
r054d	DNA	CGCTCGACTGATGAGGCGC

[†]FAM, fluorescein; Cy3, cyanine 3.

Table S3: Sequences of 2-aminoimidazole activated trinucleotides used in this study

Name	Sequence ($5'\rightarrow 3'$)
t02	GAG
t03	AGG
t04	GGC
t05	GCG
t06	CGC
t14	AGC
t15	CGG
t18	UGA
t23	GAC
t24	ACG
t26	GCA
t31	GGU