Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Role of helicity in the nonenzymatic template-directed primer extension of DNA

Supplementary Information

Sung Joon Park, Kimberley Laura Callaghan and Amanda Vera Ellis*

School of Chemical and Biomedical Engineering, The University of Melbourne, Parkville, Victoria 3010, Australia.

Table of Contents

1. Materials and Methods

1.1 Materials

Adenosine 5'-monophosphate, ammonium persulfate, boric acid, 2'-deoxycytidine-5'-monophosphate, magnesium chloride, sodium chloride, sodium hydroxide, triethylamine (TEA) and urea were purchased from Sigma Aldrich. HEPES was purchased from Formedium (England). Acetone, ethylenediaminetetraacetic acid (EDTA) and triphenylphosphine (PPh3) were purchased from ChemSupply Australia (Australia). TRIS – Ultra Pure Grade was purchased from Astral Scientific (Australia). 40% Acrylamide/bisacrylamide, 19:1 and tetramethylethylenediamine (TEMED) were purchased from Bio-rad Laboratories (USA). Gel Loading Dye, Purple (6X), no SDS, and TriDye™ Ultra Low Range DNA Ladder were purchased from New England BioLabs (USA). DEPC-Treated (nuclease-free) water, sodium perchlorate, SYBR™ Gold Nucleic Acid Gel Stain and TE buffer were purchased from Thermo Fisher Scientific (USA). 2-aminoimidazole hemisulfate, cytidine 5'-monophosphate, and 2,2'-dipyridyl-disulfide (DPS) were purchased from Tokyo Chemical Industry (Japan). Diethyl ether was purchased from RCI Labscan (Thailand). 2'-Deoxyadenosine-5'-monophosphate was purchased from Alfa Aesar (USA). *N*,*N*-Dimethylformamide (DMF) was purchased from Supelco (USA). All oligonucleotides were purchased from Integrated DNA technologies (USA). All chemicals were used as received from suppliers without further purification.

Table S1. List of oligonucleotides used.

	Length/ bases	Sequence (5'→3')
DNA Template T	30	ACTACGATGGCACTTGCAGTCAGTCTACGC
DNA Template G	30	ACTACGATGGCACGGGCAGTCAGTCTACGC
DNA Primer	15	GCGTAGACTGACTGC
FAM-labelled DNA Primer	15	5'FAM-GCGTAGACTGACTGC
DNA Helper	11	GTGCCATCGTA
RNA Template T	30	rArCrUrArCrGrArUrGrGrCrArCrUrUrGrCrArGrU rCrArGrUrCrUrArCrGrC
RNA Template G	30	rArCrUrArCrGrArUrGrGrCrArCrGrGrGrCrArGrU rCrArGrUrCrUrArCrGrC
RNA Primer	15	rGrCrGrUrArGrArCrUrGrArCrUrGrC
FAM-labelled RNA Primer	15	5'FAM-rGrCrGrUrArGrArCrUrGrArCrUrGrC
RNA Helper	11	rGrUrGrCrCrArUrCrGrUrA

1.2 General method for synthesis of 2-aminoimidazole activated nucleotides[1, 2]

TEA (132 mg, 181 μL, 1.3 mmol) was added, with stirring under a nitrogen atmosphere, to a solution of nucleotide (adenosine 5'-monophosphate, 2'-deoxyadenosine-5'-monophosphate, cytidine 5'-monophosphate or 2'-deoxycytidine-5'-monophosphate (0.1 mmol)), 2-aminoimidazole hemisulfate (0.132 g, 1.0 mmol), DPS (220 mg, 1.0 mmol) and PPh3 (236 mg, 0.9 mmol) in DMF (1.5 mL). After 30 min an additional aliquot of TEA (132 mg, 181 μL, 1.3 mmol) was added and the solution allowed to stir overnight. The resulting solid product for each reaction was removed *via* filtration and the product precipitated out of the filtrate by the addition of a pre-chilled (-4 °C) solution of acetone (20 mL), diethyl ether (12.5 mL), TEA (1.5 mL), and a saturated solution of sodium perchlorate in acetone (200 μL). Precipitation was allowed to proceed (-4 °C, 10 min). The precipitate was then removed *via*

centrifugation (10 min, 3,500 x g) using a Hettich Rotofix 32 A and the supernatant discarded. The pellet was washed with cold acetone (2 x 5 mL), before being dissolved in an acetonitrile:water solution (1:1, 1 mL) and frozen in liquid nitrogen. Lyophilisation was then carried out using a Labconco FreeZone 4.5 Liter -105C Benchtop Freeze Dryer under reduced pressure to produce a cream-coloured solid product. These products are denoted as following: deoxyadenosine-5'-phosphoro-(2-aminoimidazole) (dA-2AI), adenosine-5'-phosphoro-(2-aminoimidazole) (rA-2AI), deoxycytidine-5'-phosphoro-(2-aminoimidazole) (dC-2AI) and cytidine-5'-phosphoro-(2-aminoimidazole) (rC-2AI).

1.3 Circular dichroism (CD) spectroscopy

For CD spectroscopy, eight different template, FAM-labelled primer, and helper complexes (systems **1-8**, Fig 2), with RNA or DNA (see Table S1 and Table S2), were prepared first by combining template solution (100 μM, 8 μL), FAM-labelled primer solution (100 μM, 8 μL), helper strand solution (100 μM, 8 μL) and Milli-Q water (46 μL) in a mixture of HEPES (pH 8.0, 500 mM, 10 μL), NaCl (500 mM, 10 μL), EDTA (pH 8.0, 10 mM, 10 μL). The resulting solution is 100 μL containing template (8 μM), FAM-labelled primer (8 μM), helper strand (8 μM), HEPES (50 mM), NaCl (50 mM) and EDTA (1 mM). This solution was then annealed by heating to 90 °C for 2 min then cooling to 25 °C at a rate of 0.1 °C/s using a Bio-rad MJ Mini Gradient Thermal Cycler (USA), resulting in annealed template-primer-helper complex solutions. Each annealed template-primer-helper complex (10 μL) was then added to a solution of HEPES (500 mM, 15 μL), $MgCl_2$ (500 mM, 7.5 μL) and Milli-Q water (17.5 μL). Each final mixture contained template (1.6 μM), FAM-labelled primer (1.6 μM), helper (1.6 μM), HEPES (160 mM), $MgCl_2$ (75 mM), NaCl (10 mM) and EDTA (0.2 mM). The CD spectra were performed on an Applied Photophysics Chirascan spectrometer. The spectra were measured under ambient conditions

in a quartz cuvette with a 0.1 mm path length, with a smoothed blanked average of three spectra measured between 220-320 nm.

Table S2. DNA/RNA variant of template, primer and helper used in each system.

	Template	Primer	Helper
1	DNA	DNA	DNA
2	DNA	DNA	RNA
3	DNA	RNA	DNA
4	DNA	RNA	RNA
5	RNA	DNA	DNA
6	RNA	DNA	RNA
7	RNA	RNA	DNA
8	RNA	RNA	RNA

1.4 General procedure for nonenzymatic primer extension

For nonenzymatic template-directed primer extension, template-primer-helper complexes were prepared by first combining a template solution (100 μM, 12.5 μL), a FAM-labelled primer solution (100 μM, 7.5 μL), a helper strand solution (100 μM, 17.5 μL) and Milli-Q water (32.5 μL) in a mixture of HEPES (pH 8.0, 500 mM, 10 μL), NaCl (500 mM, 10 μL), EDTA (pH 8.0, 10 mM, 10 μL). The resulting solution (100 μL) contains a template (12.5 μM), FAM-labelled primer (7.5 μM), helper strand (17.5 μM) and a solution of HEPES (50 mM), NaCl (50 mM) and EDTA (1 mM). The solution was then annealed by heating to 90 °C for 2 min then cooling to 25 °C at a rate of 0.1 °C/s using a Bio-rad MJ Mini Gradient Thermal Cycler (USA), resulting in annealed template-primer-helper complex solutions.

To achieve the nonenzymatic template-directed primer extension reaction, the annealed template-primer-helper complex (10 μL) was added to a solution of HEPES (500 mM, 15 μL), $MgCl_2$ (500 mM, 7.5 μL) and Milli-Q water (7.5 μL). To each solution, synthesised 2-aminoimidazole activated nucleotides (dA-2AI, rA-2AI, dC-2AI or rC-2AI) (pH 8.5, 20 mM, 10 μL) were added. Each final reaction mixture contains template (2.5 μM), FAM-labelled primer (1.5 μM), helper strand (3.5 μM), HEPES (160 mM), $MgCl_2$ (75 mM), NaCl (10 mM) and EDTA (0.2 mM). Each reaction mixture was then kept at 25 °C in a MyTemp Mini digital incubator. Aliquots (2 μL) were taken at given timepoints and mixed with formamide (6 μL) and Gel Loading Dye, Purple (6X), no SDS (2 μL) then quenched by heating at 95 °C for 2 min ready for separation by PAGE.

1.5 General procedure for gel electrophoresis

Reaction mixtures were separated by denaturing urea polyacrylamide gel electrophoresis (urea-PAGE) using a Biorad Mini-PROTEAN Tetra Cell polyacrylamide gel electrophoresis system. Urea-PAGE gels (28%, 7 M urea) were prepared by combining 40% acrylamide and bisacrylamide, 19:1 (8.4 mL), DEPC-treated water (2.4 mL), 10x TBE buffer (1.2 mL) and urea (5 g). After mixing, APS (100 mg/mL, 100 μL) and TEMED (5 μL) was added. The resulting mixture was placed between glass plates (Bio-rad Mini-PROTEAN Spacer Plates with 1.0 mm Integrated Spacers and Mini-PROTEAN Short Plates) and left to set for 1 h at room temperature.

Reaction mixtures were then run on the urea-PAGE gels at 50 °C with 1x TBE as running buffer. The gels were pre-run at 130 V for 30 min before running at 110 V for 2.5 h. The separated

gels were imaged using a Biorad Gel Doc EZ gel documentation system. The lane profiles were extracted with ImageJ2 and quantified using the peak Gaussian fitting function of OriginLab, OriginPro.

2. Characterisation

All NMR measurements were performed using a 500 MHz Bruker NEO system, measured in D_2O.

Deoxyadenosine-5'-phosphoro-(2-aminoimidazole) (dA-2AI)

Deoxyadenosine-5'-monophosphate (33.12 mg, 0.1 mmol) was combined with 2-aminoimidazole hemisulfate (132 mg, 1.0 mmol) following the protocol in Section 1.2 to give the product, deoxyadenosine-5'-phosphoro-(2-aminoimidazole) (dA-2AI), as a white powder (33.3 mg, Yield = 80%). ^{1}H-NMR (400 MHz, D_2O); δ (ppm) = 8.15 (d, 1H, J = 2.1 Hz), 8.05 (d, 1H, J = 2.1 Hz), 6.60 – 6.53 (1H, m), 6.41 (dd, 1H, J = 4.2, 2.1 Hz), 6.30 (td, 1H, J = 6.8, 1.9 Hz), 4.63 – 4.55 (1H, m), 4.19 (dd, 1H, J = 3.6, 1.8 Hz), 4.05 – 3.96 (m, 2H). ^{13}C-NMR (101 MHz, D_2O) δ 155.18, 152.43, 151.20 (d, J = 4.9 Hz), 148.27, 139.43, 122.45 (d, J = 10.8 Hz), 118.39, 115.21 (d, J = 6.0 Hz), 85.36 (d, J = 9.1 Hz), 83.80, 70.99, 65.81 – 65.47 (m), 38.84. ^{31}P-NMR (162 MHz, D_2O) δ -7.20, -9.20. LRMS (OrbiTrap) *m/z*: $[M+H]^+$ Calcd for $C_{13}H_{18}N_8O_5P^+$ 397.11, found: 397.11. $[M-H]^-$ Calcd for $C_{13}H_{16}N_8O_5P^-$ 395.10, found: 395.10.

Adenosine-5'-phosphoro-(2-aminoimidazole) (rA-2AI)

Adenosine-5'-monophosphate hydrate (34.72 mg, 0.1 mmol) was combined with 2-aminoimidazole hemisulfate (132 mg, 1.0 mmol) following the protocol in Section 1.2 to give the product, adenosine-5'-phosphoro-(2-aminoimidazole) (rA-2AI), as a white powder (30.2

mg, Yield = 77%). ^{1}H-NMR (400 MHz, D_2O); δ (ppm) = 8.18 (1H), 8.04 (dd, 1H, J = 8.4, 1.9 Hz), 6.69 – 6.63 (1H, m), 6.49 – 6.42 (1H, m), 5.97 (d, 1H, J = 5.5 Hz), 4.67 (dd, 2H, J = 6.0, 4.4 Hz), 4.43 – 4.32 (1H, m), 4.28 (1H), 4.09 (3H), 3.02 – 2.93 (m, 1H), 2.86 – 2.74 (m, 1H). ^{13}C-NMR (101 MHz, D_2O) δ 164.91 (s), 155.17, 152.59, 151.35 (d, J = 4.6 Hz), 148.59, 139.41, 122.73 (d, J = 10.4 Hz), 118.31, 115.64 (d, J = 2.1 Hz), 115.62 – 115.21 (m), 87.26, 83.19 (d, J = 8.8 Hz), 74.04, 70.20, 65.24. ^{31}P-NMR (162 MHz, D_2O) δ -7.21, -9.19. LRMS (OrbiTrap) *m/z*: $[M+H]^+$ Calcd for $C_{13}H_{18}N_8O_6P^+$ 413.11, found: 413.11. $[M-H]^-$ Calcd for $C_{13}H_{16}N_8O_6P^-$ 411.09, found: 411.09.

Deoxycytidine-5’-phosphoro-(2-aminoimidazole) (dC-2AI)

Deoxycytidine-5’-monophosphate (30.72 mg, 0.1 mmol) was combined with 2-aminoimidazole hemisulfate (132 mg, 1.0 mmol) following the protocol in Section 1.2 to give the product, deoxycytidine-5’-phosphoro-(2-aminoimidazole) (dC-2AI), as a white powder (28.5 mg, Yield = 77%). ^{1}H-NMR (400 MHz, D_2O); δ (ppm) = 7.71 (dd, 1H, J = 7.7, 1.7 Hz), 6.77 (dd, 1H, J = 3.7, 1.6 Hz), 6.62 (dd, 1H, J = 4.1, 2.0 Hz), 6.25 (t, 1H, J = 7.6 Hz), 6.00 (dd, 1H, J = 7.7, 1.8 Hz), 4.46 – 4.37 (m, 1H), 4.13 (dd, 1H, J = 4.8, 3.4 Hz), 4.11 – 4.07 (m, 0.5H), 4.07 – 3.99 (m, 1H). ^{13}C-NMR (101 MHz, D_2O) δ 166.12, 157.50, 151.45 (d, J = 4.6 Hz), 141.16 (d, J = 8.2 Hz), 122.87 (d, J = 11.5 Hz), 115.79 – 114.51 (m), 96.23 (d, J = 7.4 Hz), 86.11 (d, J = 7.7 Hz), 85.06 (d, J = 9.0 Hz), 70.76, 65.49 (d, J = 6.2 Hz), 46.70 (s), 39.47 (d, J = 15.4 Hz), 8.29 (s). ^{31}P-NMR (162 MHz, D_2O) δ -7.21, -9.18. LRMS (OrbiTrap) *m/z*: $[M+H]^+$ Calcd for $C_{12}H_{18}N_6O_6P^+$ 373.10, found: 373.10. $[M-H]^-$ Calcd for $C_{12}H_{16}N_6O_6P^-$ 371.09, found: 371.09.

Cytidine-5’-phosphoro-(2-aminoimidazole) (rC-2AI)

The cytidine-5’-monophosphate disodium salt (36.72 mg, 0.1 mmol) was combined with 2-aminoimidazole hemisulfate (132 mg, 1.0 mmol) following the protocol in Section 1.2 to give

the product, cytidine-5'-phosphoro-(2-aminoimidazole) (rC-1AI), as a white powder (13.3 mg, Yield = 43%). ^{1}H-NMR (400 MHz, D_2O); δ (ppm) = 7.88 – 7.61 (m, 1H), 6.85 – 6.77 (m, 1H), 6.62 (dd, 1H, J = 4.2, 2.0 Hz), 6.07 – 6.02 (m, 1H), 5.94 (dd, 1H, J = 4.5, 2.2 Hz), 4.32 – 4.23 (m, 1H), 4.23 – 4.11 (m, 3H), 4.11 – 3.98 (m, 2H), 1.96 – 1.88 (m). ^{13}C-NMR (101 MHz, D_2O) δ 166.22, 157.72, 151.63 (d, J = 4.4 Hz), 141.13, 123.30 (d, J = 12.4 Hz), 115.50 (d, J = 6.5 Hz), 96.40, 89.62, 82.24 (d, J = 9.1 Hz), 74.11, 69.31, 64.83 (d, J = 5.9 Hz), 23.32. ^{31}P-NMR (162 MHz, D_2O) δ -7.12, -9.15. LRMS (OrbiTrap) *m/z*: $[M+H]^+$ Calcd for $C_{12}H_{18}N_6O_7P^+$ 389.10, found: 389.10. $[M-H]^-$ Calcd for $C_{12}H_{16}N_6O_7P^-$ 387.08, found: 387.08.

3. Reaction kinetics of nonenzymatic template-directed primer extension

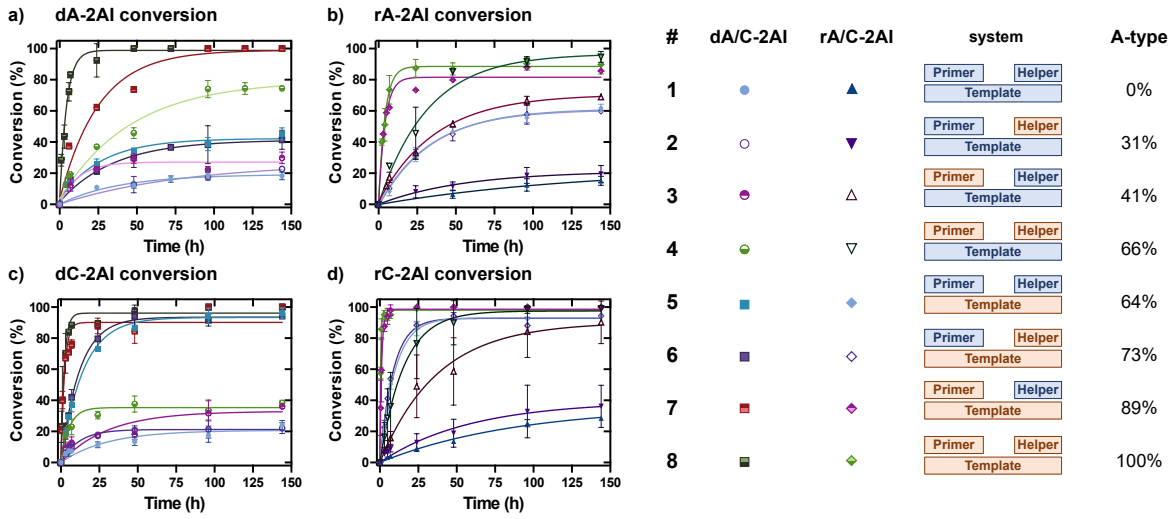

Figure S1. Percentage conversion of nonenzymatic template-directed primer extension reaction of the 8 template-primer-helper complex systems versus time (h) with an incoming 2-aminoimidazole activated nucleotide based on (a) deoxyadenosine (dA-2AI), (b) adenosine (rA-2AI), (c) deoxycytidine (dC-2AI), and (d) cytidine (rC-2AI). Error bars show the standard deviation of 3 replicates.

1. L. Li, N. Prywes, C. P. Tam, D. K. O'Flaherty, V. S. Lelyveld, E. C. Izgu, A. Pal and J. W. Szostak, *J Am Chem Soc*, 2017, **139**, 1810-1813.
2. D. Jovanovic, P. Tremmel, P. S. Pallan, M. Egli and C. Richert, *Angew Chem Int Ed Engl*, 2020, **59**, 20154-20160.