Supporting Information

for manuscript entitled

Regioselective Formation of RNA Strands in the Absence of Magnesium Ions

by

Sebastian Motsch, Peter Tremmel and Clemens Richert

Table of Contents

1. General ................................................................................................................................................. 2
2. Synthesis of Reference Compounds ................................................................................................. 3
3. Oligomerization Assays ...................................................................................................................... 20
   3.1 Representative Protocols for Oligomerization Assays ............................................................... 20
   3.2 Representative $^{31}$P-NMR Spectra from Oligomerization Assays ........................................... 21
   3.4 Confirmation of Peak Assignments via Spiking with Authentic Samples ......................... 24
   3.5 Longest Detectable Oligoadenylates ......................................................................................... 25
4. Kinetics and Systems Chemistry Analysis ......................................................................................... 26
5. References for Supporting Information .......................................................................................... 36
1. General

Chemicals. Chemicals and dry solvents were purchased from commercial suppliers and used without further purification. Adenosine 5'-monophosphate (AMP, free acid) was purchased from Carbosynth Ltd. (Berkshire, UK), cytidine 5'-monophosphate (CMP, disodium salt) from Sigma Aldrich (Deisenhofen, Germany), uridine 5'-monophosphate (UMP, disodium salt) from Acros (Geel, Belgium), and guanosine 5'-monophosphate (GMP, disodium salt monohydrate) from TCI (Zwijndrecht, Belgium). Preactivated nucleotide 2-MeIm-AMP was synthesized according to the published procedures. D$_2$O (99.90% deuterated) was purchased from Euriso-Top (Saint-Aubin, France).

NMR Spectroscopy. NMR spectra were acquired on a Bruker Avance III HD spectrometer (proton resonance frequency 700.36 MHz, phosphorous frequency 283.5 MHz) or a Bruker Avance III HD-NanoBay spectrometer (proton resonance frequency 400.10 MHz, phosphorous frequency 162 MHz). $^1$H-NMR spectra were measured in the presence of either 3-(trimethylsilyl)-1-propanesulfonic acid sodium salt or 3-(trimethylsilyl)-propionic-2,2,3,3,-d$_4$ acid sodium salt as an internal standard.

MALDI-TOF MS. MALDI-TOF mass spectra were acquired on a Bruker Microflex mass spectrometer (Bruker Daltonics, Bremen, Germany) in linear negative mode. A mixture of 2:1 (v:v) of a 2,4,6-trihydroxyacetophenone solution (0.2 M in ethanol) and a diammonium citrate solution (0.1 M in H$_2$O) were used as matrix and co-matrix, respectively. ESI-HRMS. High resolution mass spectra were recorded on a micrOTOF-Q mass spectrometer (Bruker Daltonics, Bremen, Germany), either in negative or positive mode.

Chromatography. Purification of oligonucleotides via cartridge used Sep-Pak Vac C18 3cc RP-cartridges or Sep-Pak Vac QMA 3cc IEX-cartridges (Waters, Milford, MA, USA). The C18 cartridges were rinsed with MeCN and aqueous triethylammonium acetate (TEAA) buffer solution (100 mM, pH 7.5) before applying the product solution; QMA-cartridges were washed with a solution of (NH$_4$)$_2$CO$_3$ in H$_2$O (5 mM, pH 8.0). Alternatively, anion-exchange chromatography was performed on Express-Ion DEAE cellulose from GE Healthcare (Buckinghamshire, UK). Details are given in the respective sections, below.

HPLC. Reversed-phase HPLC was performed on a Rainin Dynamax Model SD-300 HPLC, using an analytical Nucleosil 120-5 C18 column (250 x 4.6 mm) from Macherey-Nagel (Düren, Germany). A gradient of MeCN in triethylammonium acetate (TEAA) buffer (0.1 M, pH 7.5) with a flow of 1 mL/min was used with detection at $\lambda =260$ nm.

Solid-Phase RNA Synthesis. Automated solid-phase synthesis was performed on an ABI Expedite 8909 DNA synthesizer (PerSeptive Biosystems, Framingham, USA) using the phosphoramidite method and the
settings recommended by the manufacturer. Reagents for oligonucleotide synthesis included controlled-pore glass (cpg) loaded with the 3'-terminal ribonucleoside and 2'-TBDMS-protected phosphoramidites from Sigma Aldrich (Taufkirchen, Germany). The 3'-TBDMS-protected phosphoramidites were from ChemGenes (Wilmington, MA). Finally, the 'chemical phosphorylation reagent' 2-[2-(4,4'-dimethoxytrityloxy)ethylsulfonyl]ethyl-(2-cyanoethyl)-(N,N-diisopropyl) phosphoramidite was synthesized in-house following a known procedure.\textsuperscript{S2}

2. Synthesis of Reference Compounds

![Scheme S1](image.png)

**Scheme S1.** Synthesis routes to reference compounds. Chain extension was by automatic RNA synthesis, subsequent coupling to dipeptido dinucleotides or pyrophosphate-capped dinucleotides was by solution phase syntheses. The route shown is for the 3',5'-regioisomer and adenine as nucleobase and is representative. See the General Protocols, below, for details.

**Synthesis of Dinucleotides (General Protocol A).** The following protocol describes the synthesis of pGG and is representative for all dinucleotides. Solid-phase synthesis started from cpg loaded with the guanosine nucleoside (80 mg, 2 \( \mu \)mol). The first coupling cycle was performed with the 3'-TBDMS protected guanosine building block and the second with 'chemical phosphorylation reagent'. Afterwards, the solid support was treated with aqueous ammonia/aqueous methylamine (1:1 v/v; "AMA") solution (750 \( \mu \)L) for 15 min at 55 °C. The supernatant was harvested, and the residue was washed three times with H\(_2\)O (3 x 500 \( \mu \)L).
The combined aqueous solutions were freed from ammonia and methylamine with a nitrogen stream directed onto the surface of the solution and then lyophilized to dryness. The resulting solid was dissolved in TEAA buffer (1 mL, 100 mM) and loaded onto a C18 SepPac cartridge (500 mg). The TBDMS-protected dinucleotide was eluted with a gradient of MeCN in TEAA buffer (0-25% MeCN, in 5% steps, 5 mL each). Product-containing fractions were collected and lyophilized. The resulting solid was dissolved in triethylamine trihydrofluoride (750 µL) and the solution was allowed to shake for 5 h at room temperature. Then, methoxytrimethylsilane (5 mL) was added, and the mixture was shaken vigorously for 5 min. The suspension was centrifuged for 5 min, and the supernatant was decanted off. The pellet was dissolved in TEAA buffer (1 mL, 100 mM) and again loaded onto a C18 SepPac cartridge (500 mg). After rinsing with TEAA buffer (4.5 mL, 100 mM) the deprotected dinucleotide was eluted with H₂O (10 mL). Fractions containing the fully deprotected dinucleotide were pooled and dried in vacuo.

**Synthesis of 5'-Pyrophosphate-Capped Dinucleotides (General Protocol B).** The following protocol is for AppAA and is representative. A sample of pAA (triethylammonium salt, 0.88 mg, 1.0 µmol, 1.0 equiv) was dissolved in a buffer solution (15 µL) containing HEPES (1.8 mg, 7.5 µmol, 500 mM) and MgCl₂ (0.11 mg, 1.2 µmol, 80 mM) at pH 7.5. Then, 2-MeIm-AMP (2.64 mg, 6.1 µmol, 6.1 equiv, 510 mM) was added, and the reaction mixture was incubated at 40 °C. After 16 h, additional 2-MeIm-AMP (1.02 mg, 2.4 µmol, 2.4 equiv) was added, and the reaction was allowed to proceed for 5 h at 40 °C. Then, the product was pre-purified by anion exchange chromatography (DEAE cellulose), using a step-gradient of ammonium bicarbonate in H₂O (0-200 mM bicarbonate, 20 mM steps, 6 mL each). Products were detected in fractions with 140-200 mM bicarbonate. The fractions were collected, freeze-dried, and the product was re-purified by HPLC (C18, MeCN gradient in 100 mM aqueous TEAA buffer, 0% to 15% in 45 min, t<sub>R</sub> = 31 min). Excess TEAA buffer was removed via C18 SepPac cartridge. The product was dissolved in TEAA buffer (1 mL, 100 mM) and loaded on the cartridge (500 mg). After washing with water (6 mL), the product was eluted with a solution of MeCN in H₂O (6 mL, 5% MeCN). The solvents were removed by lyophilization to obtain a colorless powder.

**Synthesis of Dipeptido Dinucleotides (General Protocol C).** This protocol is for GlyGlyAA and is representative. The triethylammonium salt of dinucleotide pAA (0.22 mg, 250 nmol, 1.0 equiv) was dissolved in a freshly prepared solution (D₂O, 100 µL) containing EDC hydrochloride (15.3 mg, 80 µmol, 320 equiv, 800 mM) and DMAP (1.80 mg, 15 µmol,
60 equiv, 150 mM) at pH 8.0, uncorrected for deuterium effect. After 3 h at 22 °C, the reaction mixture was added to a solution of NaOCl in acetone/Et₂O (3 mL, 0.1 M, 1:1 v/v) at 0 °C. After vigorous shaking, the suspension was kept in an ice bath for 10 min. Then, the suspension was centrifuged for 5 min, and the supernatant was decanted. The pellet thus obtained was washed with Et₂O (1 mL) twice, and then dried at 15 mbar in a desiccator. The pellet was dissolved in a solution of glycylglycine (7.92 mg, 60 µmol, 240 equiv, 300 mM) in D₂O (200 µL) at pH = 7.5. The reaction mixture was incubated for 3 h at 22 °C. Afterwards, the solution was diluted with TEAA buffer (1 mL, 100 mM) and loaded onto a C18 SepPac cartridge (500 mg). The product was eluted with a step gradient of MeCN in 100 mM TEAA buffer (0-14%, 2%-steps, 2 mL each). Product-containing fractions (6-8% MeCN) were pooled and lyophilized. Residual TEAA buffer was removed by adsorbing the product on a C18 SepPac cartridge, followed by elution with water/acetonitrile and lyophilization, as described in General Protocol B.

**Analytical Data for Dinucleotides**

The NMR data of 3’,5’-dimers pAA, pCC, pUU, and pGG were in agreement with those reported recently.\[S3]\n
![Dimer pAA](image)

**Dimer pAA.** Yield 50%. **¹H-NMR (700 MHz, D₂O):** δ [ppm] = 8.36 (s, 1 H, AH8), 8.30 (s, 1 H, AH8), 8.18 (s, 1 H, AH2), 8.02 (s, 1 H, AH2), 6.01 (d, ³J = 4.0 Hz, 1 H, A2H1’), 5.92 (d, ³J = 3.8 Hz, 1 H, A1H1’), 4.77 (t, ³J = 5.4 Hz, 1 H, A1H3’), 4.71 (dd, ³J = 4.3 Hz, 1 H, A1H2’), 4.53 (dd, ³J = 4.5 Hz, 1 H, A2H2’), 4.50-4.48 (unresolved multiplets, 2 H, A2H3’, A1H4’), 4.37-4.35 (unresolved multiplets, 2 H, A2H4’, A2H5’), 4.22-4.12 (unresolved multiplets, 3 H, A1H5’, A2H5”, A1H5”); **³¹P-NMR (162 MHz, D₂O):** δ [ppm] = 0.19 (P1), -0.83 (P2); **HRMS (ESI-TOF):** m/z calcd. for C₂₀H₂₅N₁₀O₁₃P₂ [M-H] 675.108, found 675.108.
**Dimer pCC.** Yield 72%. $^1$H-NMR (700 MHz, D$_2$O): $\delta$ [ppm] = 8.13 (d, $^3J = 7.8$ Hz, 1 H, C1H6), 8.08 (d, $^3J = 7.8$ Hz, 1 H, C2H6), 6.18 (d, $^3J = 7.8$ Hz, 1 H, C2H5), 6.15 (d, $^3J = 7.8$ Hz, 1 H, C1H5), 5.91 (d, $^3J = 3.0$ Hz, 1 H, C2H1'), 5.86 (d, $^3J = 3.3$ Hz, 1 H, C1H1'), 4.59 (ddd, $^3J = 5.4$ Hz, $^3J_{H,P} = 8.5$ Hz, 1 H, C1H3'), 4.48-4.46 (unresolved multiplets, 2 H, C1H2', C1H4'), 4.34-4.31 (unresolved multiplets, 2 H, C2H5', C2H3'), 4.29-4.25 (unresolved multiplets, 3 H, C1H5', C2H2', C2H4'), 4.14-4.11 (unresolved multiplets, 2 H, C1H5'', C2H5''); $^{31}$P-NMR (162 MHz, D$_2$O): $\delta$ [ppm] = -0.08 (P1), -1.00 (P2); HRMS (ESI-TOF): m/z calcd. for C$_{18}$H$_{25}$N$_6$O$_{15}$P$_2$ [M-H]$^-$ 627.086, found 627.088.

**Dimer pUU.** Yield 45%. $^1$H-NMR (700 MHz, D$_2$O): $\delta$ [ppm] = 7.96 (d, $^3J = 8.2$ Hz, 1 H, UH6), 7.90 (d, $^3J = 8.2$ Hz, 1 H, UH6), 5.97 (d, $^3J = 5.4$ Hz, 1 H, U1H1'), 5.94 (d, $^3J = 3.8$ Hz, 1 H, U2H1'), 5.92 (d, $^3J = 8.2$ Hz, 1 H, UH5), 5.91 (d, $^3J = 8.1$ Hz, 1 H, UH5), 4.65 (ddd, $^3J_{H,P} = 7.9$ Hz, $^3J = 4.5$ Hz, 1 H, U1H3'), 4.47 (m, 1 H, U1H4'), 4.44 (dd, $^3J = 5.2$ Hz, 1 H, U1H2'), 4.34-4.33 (unresolved multiplets, 2 H, U2H2', U3H2'), 4.26-4.23 (unresolved multiplets, 2 H, U2H4', U3H5'), 4.16 (m, 1 H, U1H5'), 4.13 (m, 1 H, U2H5''), 4.09 (m, 1 H, U1H5''); $^{31}$P-NMR (162 MHz, D$_2$O): $\delta$ [ppm] = 0.00 (P1), -0.73 (P2); HRMS (ESI-TOF): m/z calcd. for C$_{18}$H$_{23}$N$_4$O$_{17}$P$_2$ [M-H]$^-$ 629.054, found 629.055.
Dimer pGG. Yield 40%. \(^{1}H\)-NMR (700 MHz, D\(_2\)O): δ [ppm] = 8.02 (s, 1 H, GH8), 7.98 (s, 1 H, G2H1'), 5.86 (d, \(^{3}J = 5.5\) Hz, 1 H, G2H2'), 4.77 (dd, \(^{3}J = 4.6\) Hz, 1 H, G1H1'), 4.49 (dd, \(^{3}J = 4.8\) Hz, 1 H, G2H3'), 4.41 (m, 1 H, G1H4'), 4.33 (m, 1 H, G2H4'), 4.26 (m, 1 H, G2H5'), 4.18 (m, 1 H, G2H5''), 4.07-4.05 (unresolved multiplets, 2 H, G1H5', G1H5''); \(^{31}P\)-NMR (162 MHz, D\(_2\)O): δ [ppm] = 0.25 (P1), -0.57 (P2); HRMS (ESI-TOF): m/z calcd. for C\(_{20}\)H\(_{25}\)N\(_{10}\)O\(_{15}\)P\(_{2}\) [M-H] 707.098, found 707.099.

Dimer pA\(^2\)A. Yield 53%. \(^{1}H\)-NMR (700 MHz, D\(_2\)O): δ [ppm] = 8.40 (s, 1 H, AH8), 8.30 (s, 1 H, AH8), 8.20 (s, 1 H, AH2), 8.18 (s, 1 H, AH2), 6.24 (d, \(^{3}J = 4.2\) Hz, 1 H, A1H1'), 5.94 (d, \(^{3}J = 4.1\) Hz, 1 H, A2H1'), 5.13 (m, 1 H, A1H2'), 4.71 (dd, \(^{3}J = 5.0\) Hz, 1 H, A1H3'), 4.45 (dd, \(^{3}J = 4.5\) Hz, 1 H, A2H2'), 4.38 (dd, \(^{3}J = 5.2\) Hz, 1 H, A2H3'), 4.36 (m, 1 H, A1H4'), 4.21 (m, 1 H, A2H4'), 4.14-4.06 (unresolved multiplets, 4 H, A1H5', A1H5'', A2H5', A2H5''); \(^{31}P\)-NMR (162 MHz, D\(_2\)O): δ [ppm] = 0.29 (P1), -1.09 (P2); HRMS (ESI-TOF): m/z calcd. for C\(_{20}\)H\(_{25}\)N\(_{10}\)O\(_{13}\)P\(_{2}\) [M-H] 765.108, found 765.109.

Dimer pC\(^2\)C. Yield 17%. \(^{1}H\)-NMR (700 MHz, D\(_2\)O): δ [ppm] = 7.94 (d, \(^{3}J = 7.2\) Hz, 1 H, CH6), 7.93 (d, \(^{3}J = 6.8\) Hz, 1 H, CH6), 6.18 (d, \(^{3}J = 2.6\) Hz, 1 H, CH5), 6.17 (m, 1 H, C1H1'),
6.03 (d, $^3J = 7.7$ Hz, 1 H, CH5), 5.86 (d, $^3J = 3.2$ Hz, 1 H, C2H1'), 4.69 (ddd, $^3J = 5.7$ Hz, $^3J_{HP} = 9.0$ Hz, 1 H, C1H2'), 4.48 (dd, 1 H, $^3J = 5.2$ Hz, $^3J = 3.2$ Hz, C1H3'), 4.31 (m, 2 H, C1H4'), 4.24 (dd, $^3J = 6.2$ Hz, $^3J = 5.2$ Hz, 1 H, C2H3'), 4.21 (dd, $^3J = 5.0$ Hz, $^3J = 3.2$ Hz, 1 H, C2H2'), 4.18-4.14 (unresolved multiplets, 2 H, C2H4', C2H5'), 4.12-4.07 (unresolved multiplets, 3 H, C1H5', C1H5'', C2H5''); $^{31}$P-NMR (162 MHz, D$_2$O): $\delta$ [ppm] = 0.16 (P1), -1.13 (P2); HRMS (ESI-TOF): $m/z$ calcd. for C$_{18}$H$_{25}$N$_6$O$_{15}$P$_2$ [M-H] 627.086, found 627.087.

**Dimer pU$_2$U.** Yield 59%. 1H-NMR (700 MHz, D$_2$O): $\delta$ [ppm] = 7.91 (d, $^3J = 8.1$ Hz, 1 H, UH6), 7.84 (d, $^3J = 8.1$ Hz, 1 H, UH6), 6.14 (d, $^3J = 6.0$ Hz, 1 H, U1H1'), 5.91 (d, $^3J = 8.1$ Hz, 1 H, UH5), 5.89 (d, $^3J = 3.8$ Hz, 1 H, U2H1'), 5.86 (d, $^3J = 8.1$ Hz, 1 H, UH5), 4.74 (ddd, $^3J = 9.3$ Hz, $^3J = 5.6$ Hz, 1 H, U1H2'), 4.46 (dd, $^3J = 5.3$ Hz, $^3J = 3.6$ Hz, 1 H, U1H3'), 4.28-4.24 (unresolved multiplets, 3 H, U2H2', U2H3', U1H4'), 4.18 (m, 1 H, U2H4'), 4.14 (m, 1 H, U2H5'), 4.09-4.06 (unresolved multiplets, 2 H, U2H5'', U1H5'), 4.02 (m, 1 H, U1H5''); $^{31}$P-NMR (162 MHz, D$_2$O): $\delta$ [ppm] = 1.30 (P1), -0.88 (P2); HRMS (ESI-TOF): $m/z$ calcd. for C$_{18}$H$_{23}$N$_4$O$_{17}$P$_2$ [M-H] 629.054, found 629.055.

**Dimer pG$_2$G.** Yield 50%. 1H-NMR (700 MHz, D$_2$O): $\delta$ [ppm] = 7.90 (s, 1 H, GH8), 7.80 (s, 1 H, GH8), 6.01 (d, $^3J = 3.6$ Hz, 1 H, G1H1'), 5.67 (d, $^3J = 5.3$ Hz, 1 H, G2H1'), 5.14 (ddd, $^3J = 9.0$ Hz, $^3J = 5.2$ Hz, $^3J = 3.8$ Hz, 1 H, G1H2'), 4.70 (dd, $^3J = 5.7$ Hz, 1 H, G1H3'), 4.56 (dd, $^3J = 5.2$ Hz, 1 H, G2H2'), 4.35 (dd, $^3J = 4.6$ Hz, 1 H, G2H3'), 4.25 (m, 1 H, G1H4'), 4.18 (m, 1 H, G2H4'), 4.09-4.02 (unresolved multiplets, 3 H, G1H5', G2H5', G2H5''), 3.98 (m, 1 H, G1H5''); $^{31}$P-NMR (162 MHz, D$_2$O): $\delta$ [ppm] = 1.24 (P1), -0.83 (P2); HRMS (ESI-TOF): $m/z$ calcd. for C$_{20}$H$_{25}$N$_{16}$O$_{15}$P$_2$ [M-H] 707.098, found 707.100.
Other Reference Compounds

**AppAA.** The triethylammonium salt of pAA (0.88 mg, 1.0 μmol, 1.0 equiv, 67 mM) was coupled with 2-MeIm-AMP (3.68 mg, 8.5 μmol, 8.5 equiv, 570 mM) in aqueous buffer solution (15 μL, 500 mM HEPES, 80 mM MgCl₂, pH 7.5) at 40 °C for 21 h, as described in General Protocol B. After HPLC purification (C18, MeCN gradient in 100 mM aqueous TEAA buffer, 0% to 15% in 45 min, tᵣ = 31 min) and desalting by C18 SepPac cartridge, the triethylammonium salt of the title compound was obtained as a colorless solid (0.69 mg, 530 nmol, 53%). ¹H-NMR (700 MHz, D₂O): δ [ppm] = 8.28 (s, 1 H), 8.14 (s, 1 H), 8.12 (s, 1 H), 8.08 (s, 1 H), 7.99 (s, 1 H), 7.85 (s, 1 H), 5.99 (d, ³J = 4.0 Hz, 1 H, A3H1’), 5.87 (d, ³J = 5.4 Hz, 1 H, A1H1’), 5.82 (d, ³J = 3.7 Hz, 1 H, A2H1’), 4.62 (dd, ³J = 4.2 Hz, 1 H, A2H2’), 4.55 (dd, ³J = 4.5 Hz, 1 H, A1H2’), 4.51-4.99 (unresolved multiplets, 2 H, A3H2’, A4H2’), 4.48 (dd, ³J = 5.0 Hz, 1 H, A3H3’), 4.43 (dd, ³J = 4.4 Hz, 1 H, A1H3’), 4.41 (m, 1 H, A2H5’), 4.36-4.32 (unresolved multiplets, 4 H, A3H4’, A1H4’, A1H5’, A3H5’), 4.22-4.16 (unresolved multiplets, 3 H, A2H5”, A1H5”, A3H5”); ³¹P-NMR (162 MHz, D₂O): δ [ppm] = -0.78 (P3), -11.08 (P1,P2); HRMS (ESI-TOF): m/z calcd. for C₃₀H₃₆N₁₅O₁₉P₃ [M-2H]²⁻ 501.577, found 501.578.

**AppA²⁻A.** Dinucleotide pA²⁻A in the form of the triethylammonium salt (0.70 mg, 800 nmol, 1.0 equiv, 67 mM) was coupled with 2-MeIm-AMP (2.64 mg, 6.1 μmol, 7.6 equiv, 510 mM) in aqueous buffer solution (12 μL, 500 mM HEPES, 80 mM MgCl₂, pH 7.5) for 18 h according to General Protocol B. After addition of further 2-MeIm-AMP (1.04 mg, 2.4 μmol, 3.0 equiv), the reaction mixture was incubated at 40 °C for 8 h. HPLC purification (C18, MeCN gradient in 100 mM aqueous TEAA buffer, 0% to 15% in 45 min, tᵣ = 28 min) and desalting using a C18 SepPac cartridge (500 mg) according to the General Protocol B gave AppA²⁻A (triethylammonium salt) as a colorless solid (0.62 mg, 475 nmol, 59%). ¹H-NMR
(700 MHz, D$_2$O): $\delta$ [ppm] = 8.18 (s, 1 H), 8.17 (s, 1 H), 8.15 (s, 1 H), 8.10 (s, 1 H), 8.01 (s, 1 H), 7.87 (s, 1 H), 6.04 (d, $J = 4.5$ Hz, 1 H, A2H1'), 5.89 (d, $J = 5.3$ Hz, 1 H, A1H1'), 5.82 (d, $J = 4.2$ Hz, 1 H, A3H1'), 4.95 (ddd, $J_{HP} = 9.0$ Hz, $J = 4.6$ Hz, 1 H, A2H2'), 4.63 (dd, $J = 5.2$ Hz, 1 H, A2H3'), 4.53 (dd, $J = 5.0$ Hz, 1 H, A1H2'), 4.42 (dd, $J = 4.3$ Hz, 1 H, A1H3'), 4.38 (dd, $J = 4.3$ Hz, 1 H, A3H2'), 4.35-4.32 (unresolved multiplets, 2 H, A2H4', A1H5'), 4.29-4.28 (unresolved multiplets, 2 H, A3H3', A1H4'), 4.25 (m, 1 H, A1H5"), 4.21-4.15 (unresolved multiplets, 3 H, A3H4', A2H5', A2H5''), 4.09-4.04 (unresolved multiplets, 2 H, A3H5', A3H5'').

$^{31}$P-NMR (162 MHz, D$_2$O): $\delta$ [ppm] = -1.03 (P3), -11.31 (P1+2); HRMS (ESI-TOF): $m/z$ calcd. for C$_{30}$H$_{37}$N$_{15}$O$_{19}$P$_3$ [M-H]$^-$ 1004.161, found 1004.160.

**GlyGlyAA.** The triethylammonium salt of dinucleotide pAA (0.22 mg, 250 nmol, 1.0 equiv, 1.4 mM) was dissolved in D$_2$O (100 µL, pH 8.0) and treated with EDC hydrochloride (15.3 mg, 80 µmol, 320 equiv, 800 mM) and DMAP (1.80 mg, 15 µmol, 60 equiv, 150 mM) for 4 h, as described in General Protocol C. After precipitation, the DMAP-activated dinucleotide was coupled with glycylglycine (7.92 mg, 60 µmol, 240 equiv, 300 mM) in D$_2$O (200 µL, pH 7.5) for 3 h. The crude product was purified via C18 cartridge (500 mg, gradient of MeCN in 100 mM TEAA buffer, 0-14% MeCN, 2%-steps, 3 mL each). The product eluted at 6-8% MeCN. After removing the buffer by C18 SepPac cartridge the triethylammonium salt of the title compound could be obtained as a colorless solid (0.20 mg, 180 nmol, 72%).

$^1$H-NMR (700 MHz, D$_2$O): $\delta$ [ppm] = 8.32 (s, 1 H, AH8), 8.30 (s, 1 H, AH8), 8.16 (s, 1 H, AH2), 8.04 (s, 1 H, AH2), 6.02 (d, $J = 4.3$ Hz, 1 H, A2H1'), 5.92 (d, $J = 4.1$ Hz, 1 H, A1H1'), 4.73 (dd, $J = 4.7$ Hz, 1 H, A1H2'), 4.55 (dd, $J = 4.7$ Hz, 1 H, A2H2'), 4.49 (dd, $J = 5.1$ Hz, 1 H, A2H3'), 4.45 (m, 1 H, A1H4'), 4.36 (m, 1 H, A2H4'), 4.35 (m, 1 H, A2H5'), 4.18 (m, 1 H, A2H5"), 4.09 (m, 1 H, A1H5"), 4.04 (m, 1 H, A1H5''), 3.69 (s, 2 H, Gly2Hα), 3.48 (d, $J_{HP} = 10.4$ Hz, 2 H, Gly1Hα); $^{31}$P-NMR (162 MHz, D$_2$O): $\delta$ [ppm] = 7.67 (P1), -0.75 (P2); HRMS (ESI-TOF): $m/z$ calcd. for C$_{24}$H$_{31}$N$_{12}$O$_{19}$P$_2$ [M-H]$^-$ 789.151, found 789.156.
GlyGlyA²A. According to General Procedure C, the triethylammonium salt of dinucleotide pA²A (0.22 mg, 250 nmol, 1.0 equiv, 2.5 mM) was activated with DMAP (1.80 mg, 15 µmol, 60 equiv, 150 mM) in the presence of EDC hydrochloride (15.3 mg, 80 µmol, 320 equiv, 800 mM) in D₂O (100 µL, pH 8.0) for 2 h. The resulting product was isolated via precipitation and coupled with glycine dipeptide (3.96 mg, 30 µmol, 120 equiv, 300 mM) in D₂O (100 µL, pH 8.0) for 18 h. Then, the product was purified via QMA-SepPac cartridge (500 mg, ammonium bicarbonate gradient in H₂O, 0-140 mM, 20 mM steps, 3 mL each). Fractions containing product (100 and 120 mM bicarbonate) were collected and lyophilized. The dipeptido dinucleotide (ammonium salt) was recovered as a white solid (0.20 mg, 235 nmol, 94%).

**¹H-NMR** (700 MHz, D₂O): δ [ppm] = 8.20 (s, 1 H, AH₈), 8.15 (s, 1 H, AH₈), 8.05 (s, 1 H, AH₂), 7.94 (s, 1 H, AH₂), 6.16 (d, ³J = 2.8 Hz, 1 H, A₁H¹'), 5.85 (d, ³J = 3.8 Hz, 1 H, A₂H¹'), 5.21 (m, 1 H, A₁H²'), 4.75 (dd, ³J = 6.0 Hz, 1 H, A₁H³'), 4.39 (dd, ³J = 5.3 Hz, 1 H, A₂H³'), 4.34 (dd, ³J = 4.3 Hz, 1 H, A₂H²'), 4.28-4.26 (unresolved multiplets, 2 H, A₁H⁴', A₂H⁴'), 4.25 (m, 1 H, A₂H⁵'), 4.13 (m, 1 H, A₂H⁵''), 4.00 (m, 1 H, A₁H⁵'), 3.92 (m, 1 H, A₁H⁵''), 3.61 (d, ³J = 2.4 Hz, 2 H, Gly₂Hα), 3.17 (m, 2 H, Gly₁Hα);

**³¹P-NMR** (162 MHz, D₂O): δ [ppm] = 7.73 (P₁), -0.93 (P₂);

**MALDI-TOF-MS:** m/z calcd. for C₂₄H₃₁N₁₂O₁₅P₂ [M-H]⁺ 789, found 789.
NMR Spectra of New Compounds

Figure S1. $^1$H-NMR spectrum of dinucleotide pA^{2'}.A (700 MHz, D$_2$O).

Figure S2. $^{31}$P-NMR spectrum of dinucleotide pA^{2'}.A (162 MHz, D$_2$O).
Figure S3. $^1$H-NMR spectrum of dinucleotide pC$^{2'}$C (700 MHz, D$_2$O).

Figure S4. $^{31}$P-NMR spectrum of dinucleotide pC$^{2'}$C (162 MHz, D$_2$O).
Figure S5. $^1$H-NMR spectrum of dinucleotide pU$^{3'}$U (700 MHz, D$_2$O).

Figure S6. $^{31}$P-NMR spectrum of dinucleotide pU$^{3'}$U (162 MHz, D$_2$O).
Figure S7. $^1$H-NMR spectrum of dinucleotide pG$^{2'}$G (700 MHz, D$_2$O).

Figure S8. $^{31}$P-NMR spectrum of dinucleotide pG$^{2'}$G (162 MHz, D$_2$O).
Figure S9. $^1$H-NMR spectrum of pyrophosphate AppAA (700 MHz, D$_2$O).

Figure S10. $^{31}$P-NMR spectrum of pyrophosphate AppAA (162 MHz, D$_2$O).
Figure S11. $^1$H-NMR spectrum of pyrophosphate AppA$^2$A (700 MHz, D$_2$O).

Figure S12. $^{31}$P-NMR spectrum of pyrophosphate AppA$^2$A (162 MHz, D$_2$O).
Figure S13. $^1$H-NMR spectrum of dipeptido dinucleotide GlyGlyAA (700 MHz, D$_2$O).

Figure S14. $^{31}$P-NMR spectrum of dipeptido dinucleotide GlyGlyAA (162 MHz, D$_2$O).
Figure S15. $^1$H-NMR spectrum of dipeptido dinucleotide GlyGlyA$^2$A (700 MHz, D$_2$O).

Figure S16. $^{31}$P-NMR spectrum of dipeptido dinucleotide GlyGlyA$^2$A (162 MHz, D$_2$O).
3. Oligomerization Assays

3.1 Representative Protocols for Oligomerization Assays

Oligomerization of Ribonucleotides. The following protocol is for the oligomerization of AMP and is representative for the NMP oligomerization assays. To AMP (free acid, 34.7 mg, 100 µmol), 1-ethylimidazole (7.0 µL, 75 µmol), and (where necessary) MgCl$_2$ (0-3.8 mg, 0-40 µmol) was added D$_2$O (450 µL), followed by addition of NaOH solution (up to 10 µL, 10 M) to aid dissolution, vortexing, and adjusting the pH to 7.5, again using 10 M NaOH solution in D$_2$O. The resulting solution was cooled to 0 °C. Then, EDC hydrochloride (76.7 mg, 400 µmol) was added to obtain a volume of 500 µL of a solution containing AMP (0.2 M), EDC (0.8 M), 1-EtIm (0.15 M) with or without MgCl$_2$ (0-80 mM). The reaction was allowed to proceed at 0 °C for 24 h. The mixture was then transferred to an NMR tube and analyzed by $^{31}$P-NMR (162 MHz) at room temperature. Reaction mixtures containing 80 mM MgCl$_2$ were diluted by adding a solution of EDTA (15.0 mg, 40 µmol) in D$_2$O (250 µL) at pH 8.0, before NMR measurements were performed.

Peptido RNA Assay with AMP and glycine. A D$_2$O solution containing AMP (34.7 mg, 100 µmol), glycine (7.54 mg, 100 µmol), and 1-ethylimidazole (7.0 µL, 75 µmol) was prepared, the volume was adjusted to 450 µL, and the pH was adjusted to a value of 7.5. Then, EDC hydrochloride (76.7 mg, 400 µmol) was added, and the solution was incubated at 0 °C for 2 d. Subsequently, the mixture was transferred to an NMR tube and analyzed by $^{31}$P-NMR.
3.2 Representative $^{31}$P-NMR Spectra from Oligomerization Assays

![Figure S17](image1.png)

**Figure S17.** Phosphodiester region of $^{31}$P-NMR spectra (162 MHz, D$_2$O) from AMP oligomerization assays (pH 7.5, 0 °C, 24 h). A) In the presence of 80 mM MgCl$_2$, and B) in the absence of MgCl$_2$.

![Figure S18](image2.png)

**Figure S18.** Phosphodiester region of $^{31}$P-NMR spectra (162 MHz, D$_2$O) from CMP oligomerization assays (pH 7.5, 0 °C, 24 h). A) In the presence of 80 mM MgCl$_2$, and B) in the absence of MgCl$_2$. 
Figure S19. Phosphodiester region of $^{31}$P-NMR spectra (162 MHz, D$_2$O) from UMP oligomerization assays (pH 7.5, 0 °C, 24 h). A) In the presence of 80 mM MgCl$_2$, and B) in the absence of MgCl$_2$.

Figure S20. Phosphodiester region of $^{31}$P-NMR spectra (162 MHz, D$_2$O) from GMP oligomerization assays (pH 7.5, 0 °C). A) In the presence of 80 mM MgCl$_2$ (20 mM GMP, 8 d reaction time), and B) in the absence of MgCl$_2$ (200 mM GMP, 24 h reaction time).
Figure S21. $^{31}$P-NMR spectra (162 MHz, D$_2$O) from peptido RNA assays with AMP and glycine (pH 7.5, 0 °C, 2 d); A) in the presence of 80 mM MgCl$_2$ and B) in the absence of MgCl$_2$. The insets show expansions of the phosphodiester regions. The dominant peaks labeled "peptido RNA" are caused by phosphoramidates.
3.4 Confirmation of Peak Assignments via Spiking with Authentic Samples

The proposed assignment of peaks in $^{31}$P-NMR was confirmed by spiking with authentic samples. For this, a reaction mixture (1.5 mL) initially containing AMP (200 mM), EDC hydrochloride (800 mM) and 1-EtIm (150 mM) that had reacted for 24 h at 0 °C was split into three aliquots (500 μL each). Then, pAA (1 μmol) was added to the first solution as a solid, pA$^2$A (1 μmol) was added to the second solution, and the third solution was left unchanged. The $^{31}$P-NMR spectra (162 MHz, 22 °C) of all three samples were recorded immediately afterwards. This protocol is representative.

Figure S22. Confirmation of peak assignment in $^{31}$P-NMR spectra of AMP oligomerization assays by spiking with authentic samples. A) Spiking with pAA or pA$^2$A (assay with starting pH 7.5, no MgCl$_2$); B) spiking with AppAA or AppA$^2$A (assay with starting pH 7.5, 80 mM MgCl$_2$); C) Spiking with cyc-pAA (assay at initial pH 7.5, no MgCl$_2$).
Figure S23. Confirmation of peak assignment in $^{31}$P-NMR spectra of AMP and glycine oligomerization assay by spiking with authentic samples. A) Spiking with pAA or pA$^2$A (assay with starting pH 7.5, no MgCl$_2$); B) spiking with GlyGlyAA or GlyGlyA$^2$A (assay with starting pH 7.5, no MgCl$_2$).

3.5 Longest Detectable Oligoadenylates

This protocol describes the determination of long oligoadenylates (up to A$_{11}$) from AMP oligomerisation assay (pH 7.0) and is representative. An aliquot (50 $\mu$L) from a reaction mixture initially containing AMP (0.2 M), EDC hydrochloride (0.8 M) and 1-EtIm (0.15 M), that was kept for 7 d at 0 °C in an ice box, was diluted with aqueous Tris buffer solution (2 mL, 20 mM Tris hydrochloride, pH 7.5) and loaded onto a QMA-SepPac cartridge (500 mg, prewashed with Tris buffer solution). Oligomerization products were eluted with a gradient of (NH$_4$)$_2$CO$_3$ in Tris buffer solution (3 mL fractions, 50 mM steps). Aliquots (75 $\mu$L) of fractions with high (NH$_4$)$_2$CO$_3$ concentrations (300 - 400 mM) were taken and diluted with H$_2$O (675 $\mu$L). The samples were lyophilized and the solids obtained were dissolved in H$_2$O (10 $\mu$L), followed by MALDI-TOF MS analysis.
4. Kinetics and Systems Chemistry Analysis

The approach used to measure kinetics by NMR and obtaining rate constants for a systems chemistry analysis of the reaction system was performed analogously to that reported recently for more complex reaction mixtures containing ribonucleotides, organocatalysts and an amino acid.[54]

Representative Protocols

Assay with EDC and EI in MOPS buffer. A buffer solution (400 μL, H₂O) containing MOPS (105 mg, 0.5 mmol, 1.25 M) and 1-ethylimidazole (7.2 μL, 75 μmol, 0.187 M) was adjusted to pH 7.5 and D₂O (50 μL) was added. Then, the solution was cooled to ice temperature, and EDC hydrochloride (77 mg, 0.4 mmol) was added to obtain a solution with the desired concentrations of EDC (800 mM), MOPS (1 M) and 1-ethylimidazole (0.15 M). The solution was transferred to an ice-cold NMR tube. Then, ¹H-NMR-spectra were recorded with presaturation (700 MHz, zgpr, d1 = 20s) at different time intervals and 274 K. The concentrations of EDC (linear/cyclic) and EDU were obtained by integration of signals from the CH₃-group of the ethyl chain.

Assay with AMP and EDC in MOPS buffer. An aqueous buffer solution (400 μL) containing AMP (35 mg, 0.1 mmol, 200 mM) and MOPS (105 mg, 0.5 mmol, 1.25 M) was adjusted to pH 7.5 with NaOH solution (10 M). Then, D₂O (50 μL) was added, and the resulting solution was cooled to 0 °C, followed by addition of EDC hydrochloride (77 mg, 0.4 mmol). The volume of the solution was adjusted to 0.5 mL with H₂O to give the desired concentrations of AMP (200 mM), EDC (800 mM) and Mops (1 M). The ice-cold solution was immediately placed into an NMR tube and subsequently transferred to the NMR spectrometer set to 1 °C. The first spectrum was acquired 30 min after EDC addition. The following spectra were measured hourly, starting 1 h after the start of the reaction. ¹H-NMR spectra were recorded with presaturation during the recycle delay (d1 = 20s), ³¹P-NMR spectra with Bruker pulse program zgig30 (recycle delay d1 = 10 s).

To obtain the concentrations of the species of interest from the ³¹P-NMR spectra (AppA, AMP, EDC-AMP), the sum of all integrals was set to 200. The value of the integral of symmetrical pyrophosphate AppA was divided by two to compensate for the symmetry of the analyte. To determine EDC consumption, the sum of all methyl-signals of the ethyl moiety of EDC (linear and cyclic) as well as EDU in the ¹H-NMR spectra were integrated.
Representative NMR spectra

Figure S24. Rate of EDC hydrolysis. (A) $^1$H-NMR spectrum (400 MHz, H$_2$O/D$_2$O 9:1) from an assay in MOPS buffer (0.8 M EDC hydrochloride, 0.15 M 1-EtIm, 1 M MOPS, H$_2$O/D$_2$O 9:1, pH 7.5, 0 °C) after 24 h. (B) $^1$H-NMR spectrum (700 MHz, H$_2$O/D$_2$O 9:1) from an assay in MOPS buffer in the presence of AMP (0.8 M EDC hydrochloride, 0.2 M AMP, 0.15 M 1-EtIm, 1 M MOPS, H$_2$O/D$_2$O 9:1, pH 7.5, 1 °C) after 24 h. Note the difference in the intensity of the signal for the urea (EDU).

Figure S25. Monitoring the carbodiimide adduct (isourea) EDC-AMP in $^{31}$P-NMR spectra (283.5 MHz, H$_2$O/D$_2$O 9:1). AMP oligomerization in the absence of ethylimidazole, after (A) 0.5 h and (B) 24 h. Conditions: 0.2 M AMP, 0.8 M EDC hydrochloride, 1 M MOPS, H$_2$O/D$_2$O 9:1, pH 7.5, 0 °C.
Figure S26. Monitoring the formation of imidazolium nucleotide EI-AMP. $^1$H-NMR spectra (700 MHz, H$_2$O/D$_2$O, 9:1) from AMP oligomerization assays in the presence of ethylimidazole (A) after 0.5 h and (B) after 24 h. Conditions: 0.2 M AMP, 0.15 M EtIm 0.8 M EDC hydrochloride, 1 M MOPS, H$_2$O/D$_2$O 9:1, pH 7.5, 0 °C.
Mathematical Model

Figure S29, below, lists the reactions for which rate constants were determined. The time dependent concentrations $X[t]$ of the species involved were calculated based on a set of rate equations that were combined in a set of coupled differential equations. The differential equations were integrated using the numerical differential equation solver NDsolve of the Mathematica Software package 11.3. Rate constants were set so as to yield 'best fits' of the experimental data in a series of assay systems of increasing complexity, starting with the hydrolysis of EDC in MOPS buffer (the simplest) and ending with the most complex system containing AMP, EDC, and 1-ethylimidazole. The parameter values determined in the simpler assays were left unchanged in the following, more complex assays, and only those rate constants newly appearing in a given assay were adjusted.

![Reaction Diagram]

**Figure S27.** Reactions and corresponding rate constants considered in the systems chemistry model of the oligomerization of AMP in the presence of EDC and 1-ethylimidazole.
Listed below is the set of rate equations employed. With this set, the time-dependent concentrations \( X[t] \) of the species involved were calculated. As mentioned above, this is a set of coupled equations, so that individual fits to just one plot of a concentration change of one molecular species over time does not suffice to obtain satisfactory solutions. Starting from the hydrolysis of EDC, rate constants were determined via a series of assays of increasing complexity. Representative data is shown below, starting with Figure S28.

\[
\begin{align*}
    \frac{d[A]}{dt} &= -k_{AA} AEI[t] A[t] - k_{AppA} AEI[t] A[t] + kh_2 AEI[t] - k_E EDC[t] A[t] + km_E AEDC[t] \\
                      &\quad + kh_1 AEDC[t] - k_{EAA} AEDC[t] A[t] - k_{EPP} AEDC[t] A[t] \\
    \frac{d[AEI]}{dt} &= -k_{AA} AEI[t] A[t] - k_{AppA} AEI[t] A[t] - kh_2 AEI[t] - k_{AA} AEI[t] AEI[t] \\
                      &\quad + k_{Etm} AEDC[t] EI[t] \\
    \frac{d[AEDC]}{dt} &= k_E EDC[t] A[t] - km_E AEDC[t] - kh_1 AEDC[t] - k_{Etm} AEDC[t] EI[t] \\
                      &\quad - k_{EAA} AEDC[t] A[t] - k_{EPP} AEDC[t] A[t] \\
    \frac{d[AA]}{dt} &= k_{AA} AEI[t] A[t] + k_{EAA} AEDC[t] A[t] \\
    \frac{d[AppA]}{dt} &= k_{AppA} AEI[t] A[t] + k_{EPP} AEDC[t] A[t] \\
    \frac{d[EDU]}{dt} &= kh_0 EDC[t] + kh_1 AEDC[t] + kh_3 EIEDC[t] + k_{Etm} AEDC[t] EI[t] + k_{EAA} AEDC[t] A[t] \\
                      &\quad + k_{EPP} AEDC[t] A[t] \\
    \frac{d[EDC]}{dt} &= -kh_0 EDC[t] - k_{Etm} EDC[t] EI[t] - k_E EDC[t] A[t] + km_E AEDC[t] + km_{EtQ} EIEDC[t] \\
    \frac{d[EIEDC]}{dt} &= k_{Etm} EDC[t] EI[t] - km_{EtQ} EIEDC[t] - kh_3 EIEDC[t] \\
    \frac{d[EI]}{dt} &= k_{AA} AEI[t] A[t] + k_{AppA} AEI[t] A[t] + kh_2 AEI[t] - k_{EtQ} EDC[t] EI[t] \\
                      &\quad + km_{EtQ} EIEDC[t] + kh_3 EIEDC[t] - k_{Etm} AEDC[t] EI[t] + k_{AA} AEI[t] AEI[t] \\
    \frac{d[AAEI]}{dt} &= k_{AA} AEI[t] AEI[t]
\end{align*}
\]
**Assay with EDC and 1-Ethylimidazole (without AMP)**

The starting point of the determination of rate constants was the assay with 1-ethylimidazole and EDC in buffer. Based on the earlier finding that EI-EDC only dissociates but does not hydrolyze, the constant $k_{h0}$ for the hydrolysis of EDC in the buffer system with 1-ethylimidazole was found to be identical to that reported previously. Figure S28a and Figure S28b shows the agreement of the experimental data and the theoretical curve, calculated with $k_{h0}$. Next, the rate of formation of EI-EDC was set by comparing the experimental data in the current magnesium-free buffer with those obtained earlier, in magnesium-containing medium. Again, the data agreed with the rate constant found earlier (Figure S28 c), indicating that magnesium ions do not affect the reaction of the two net-neutral or cationic organic molecules.

**Figure S28.** Calculated curves and experimental data points, as obtained by NMR, for assay with EDC (0.8 M) and 1-ethylimidazole at 1 °C, in 1 M MOPS buffer, pH 7.5. A) Consumption of EDC. B) Formation of EDU. C) Fits and experimental data for EI-EDC. The agreement between experimental data and model is $R^2 = 0.99$ (EDC), $R^2 = 0.92$ (EDU), and $R^2 = 0.99$ (EI-EDC).
Assay with EDC and AMP (without 1-Ethylimidazole)

The formation of EDU can occur via different pathways. From the assay above, the rate constant $k_{d0}$ for the direct hydrolysis of EDC, was known. In the assay with AMP, pyrophosphates and phosphodiesters are formed. Each endergonic step forming them consumes EDC and produces EDU. Both product formation and hydrolysis are responsible for only a minor part of the EDU formation, and the only plausible pathway for further EDU production is the hydrolysis of EDC-AMP. Accordingly, the constants $k_E$ for the formation and $k_{h1}$ for the hydrolysis of EDC-AMP were adjusted to fit the experimental data for EDC and EDU, in accordance with the data for the active species EDC-AMP. This led to the results shown in Figure S29, below.

Figure S29. Experimental data and mathematical fits for assay with AMP (0.2 M) and EDC (0.8 M) at 1 °C, pH 7.5 in 1 M MOPS buffer. A) Concentration data points and fits for the consumption of EDC (black) and the formation of EDU (purple). B) Fits and experimental data for EDC-AMP (green), phosphodiesters AA (red) and pyrophosphates AppA (blue).
**Assay with EDC, AMP and 1-Ethylimidazole**

The starting point for fitting data from the last assay, containing all components for oligomerization, was the set of constants determined from the earlier assays. The first new rate constants to be determined were those for the formation and hydrolysis of EI-AMP, in agreement with the kinetics of EDC and EDU. The formation occurs via EDC-AMP, a species that also feeds the formation of pyrophosphates and phosphodiesters. Accordingly, the rate constants were determined by fitting to the kinetics of all of species involved. This gave the initial values for $k_{AA}$ and $k_{AppA}$. All rate constants determined in this, the most complex assay, were iteratively optimized until the best fit was obtained, using the mathematical model with its coupled differential equations. (We cannot rigorously rule out low-level hydrolysis of phosphodiesters under our experimental conditions. However, control experiments with short RNA hairpins under our experimental conditions did not show detectable hydrolysis products, even after more than one week.) The agreement between experimental and theoretical values, as calculated with the model, is shown in Figure 30, below.

![Figure S30](image)

**Figure S30.** Fits and experimental data for assay with AMP (0.2 M), 1-ethylimidazole (0.15 M) and EDC (0.8 M) in 1 M MOPS buffer pH 7.5 at 1°C. A) Kinetics of EDC (black) and EDU (purple). B) EI-AMP (green), EDU (purple, plotted again for comparison), pyrophosphates (blue) and phosphodiesters (red). At the high AMP concentration chosen, a slight shift of the pH sets in over time, most probably causing the imperfect fit of the data for EI-AMP.
Table S1. Rate constants for reactions of the reaction network underlying the oligomerization of AMP with EDC and 1-ethylimidazole in the absence of MgCl$_2$, as determined by NMR monitoring.

| Assay Components$^{[a]}$ | Name of rate constant$^{[b]}$ | Order$^{[c]}$ of reaction | Value of rate constant |
|--------------------------|-------------------------------|---------------------------|------------------------|
| EDC                      | $k_{h0}$                      | 1$^{{\text{st}}}$        | $4 \times 10^4$ h$^{-1}$ |
| AMP, EDC                 | $k_{EAA}$                     | 2$^{{\text{nd}}}$        | $2.0 \times 10^{-6}$ mM$^{-1}$ h$^{-1}$ |
|                          | $k_{EPP}$                     | 2$^{{\text{nd}}}$        | $4 \times 10^{-4}$ mM$^{-1}$ h$^{-1}$ |
|                          | $k_E$                         | 2$^{{\text{nd}}}$        | $1.15 \times 10^{-4}$ mM$^{-1}$ h$^{-1}$ |
|                          | $k_{h1}$                      | 1$^{{\text{st}}}$        | $4$ h$^{-1}$            |
|                          | $k_{mE}$                      | 1$^{{\text{st}}}$        | $0.1$ h$^{-1}$          |
| AMP, EDC, EI             | $k_{AA}$                      | 2$^{{\text{nd}}}$        | $1.3 \times 10^{-5}$ mM$^{-1}$ h$^{-1}$ |
|                          | $k_{AppA}$                    | 2$^{{\text{nd}}}$        | $9 \times 10^{-5}$ mM$^{-1}$ h$^{-1}$ |
|                          | $k_{EtIm}$                    | 2$^{{\text{nd}}}$        | $0.5$ mM$^{-1}$ h$^{-1}$ |
|                          | $k_{h2}$                      | 1$^{{\text{st}}}$        | $0.03$ h$^{-1}$         |
|                          | $k_{EtQ}$                     | 2$^{{\text{nd}}}$        | $2.2 \times 10^{-5}$ mM$^{-1}$ h$^{-1}$ |
|                          | $k_{mEtQ}$                    | 1$^{{\text{st}}}$        | $1.3 \times 10^{-2}$ h$^{-1}$ |
|                          | $k_{h3}$                      | 1$^{{\text{st}}}$        | $< 10^{-3}$ h$^{-1}$    |

$^{[a]}$ Components of assay used to determine rate constant.
$^{[b]}$ See Figure S27 for a list of the reactions.
$^{[c]}$ Pseudo-first order reaction (hydrolysis) or second order reaction. Significant digits reflect estimated uncertainty.
**Figure S31.** Graphical representation of the fluxes through the two steps of activation, as well as the coupling step, of the reaction network underlying the oligomerization of AMP for the first day of the assay, as obtained by integration using the mathematical model. It can be discerned that the flux through the final coupling step is near-identical in the absence and the presence and the absence of magnesium chloride. The more successful activation steps compensate for the lower reactivity in the phosphodiester-forming step in the absence of MgCl₂. In the activation process, the rates of formation and hydrolysis of the zwitterionic EI-AMP are both unaffected by the presence of magnesium ions. The formation of EDC-AMP from AMP and EDC is 2.6-fold faster in the absence of MgCl₂, whereas the hydrolysis of EDC-AMP is only 2.2-fold faster in the absence of the divalent cations (compare Table S1 and Table 1 in ref. S4).
5. References for Supporting Information

S1. Kanavarioti, A., Bernasconi, C.F. and Baird, E. (1998) Effects of monomer and template concentration on the kinetics of nonenzymatic template-directed oligoguanylate synthesis. *J. Am. Chem. Soc.*, 120, 8575-8580.

S2. Durand, M., Chevrie, K., Chassignol, M., Thuong, N.T. and Maurizot, J.C. (1990) Circular dichroism studies of an oligodeoxyribonucleotide containing a hairpin loop made of a hexaethylene glycol chain: conformation and stability. *Nucleic Acids Res.*, 18, 6353-6359.

S3. Sosson, M., Pfeffer, D., Richert, C. (2019) Enzyme-free ligation of dimers and trimers to RNA primers. *Nucleic Acids Res.*, 47, 3836-3845

S4. Tremmel, P., Griesser, H., Steiner, U.E. and Richert, C. (2019) How small heterocycles make a reaction network of amino acids and nucleotides efficient in water. *Angew. Chem. Int. Ed.*, 58, 13087-13092.