Supplementary Data

for manuscript entitled

The Effect of Leaving Groups on Binding and Reactivity in Enzyme-Free Copying of DNA and RNA

by

Eric Kervio, Marilyne Sosson, and Clemens Richert

Contents

1. General Information

Reagents and Instrumentation. Anhydrous solvents were purchased and stored over molecular sieves and were used without further purification. Sodium phosphate (monobasic and dibasic form), NaCl, $MgCl_2$, AMP (free acid), CMP (disodium salt hydrate), UMP (disodium salt), and dGMP (sodium salt) were from Acros (Geel, Belgium). HEPES, 2,2'-dithiodipyridine, 2-methylimidazole, triphenylphosphine, and Dowex 50 WX8-200 cation exchange resin were from Fluka (Deisenhofen, Germany); GMP (disodium salt hydrate) and dCMP (monohydrate sodium salt) were from TCI (Zwijndrecht, Belgium). dAMP (free acid), TMP (disodium salt), HATU and HOAt were purchased from Molekula (München, Germany). Hairpin oligonucleotides (**7a-t** and **8a-u**) were from Biospring (Frankfurt, Germany) and templates (**9a-t** and **12a-u**) were from Biomers (Ulm, Germany). The 3'-aminoterminal oligonucleotide (primer **10**) was synthesized on a DNA synthesizer on controlled pored glass loaded with the aminodideoxynucleoside as previously described.[S1,S2] The triethylammonium salts of 2'-deoxynucleotides and ribonucleotides were obtained by cation exchange with the triethylammonium form of Dowex cation exchange resin. SepPak RP-C_{18} cartridges were from Waters (Eschborn, Germany). Deuterated water (99.9%) was purchased from Euriso-Top (Saclay Gif/Yvette, France).

NMR spectra were recorded on Bruker Avance 300, Avance 400 or Avance 500 spectrometers. The MALDI TOF mass spectra were measured on Bruker REFLEX IV or Microflex spectrometers, using software packages XACQ 4.0.4 or XTof 5.1. Oligonucleotides were measured in linear negative mode with a matrix/comatrix mixture of 2,4,6-trihydroxyacetophenone (0.3 M in EtOH) and diammoniumcitrate (0.1 M in H_2O) at a ratio of 2:1 (*v*/*v*).[S3] Concentrations of stock solutions were determined by UV- spectroscopy, using a NanoDrop ND-1000 spectrometer from PeqLab. Calculation of rate constants and dissociation constants were obtained by fitting data with Origin Pro 8.0. Activated nucleotides **2a-t**, **3a-t** and **5a-u** were also characterized by ESI mass spectroscopy on a Finnigan LCQ Duo spectrometer (negative mode), providing low-resolution spectra.

2. Activated Nucleotides

Triethylammonium salts of ribonucleotides and deoxynucleotides were obtained by cation exchange with the Et_3NH^+ form of Dowex resin. The syntheses of OAt-activated desoxynucleotides (**2a-t**) and ribonucleotides (**5a-u**) via activation of dNMPs (**1a-t**) and NMPs (**4a-u**) with HATU/HOAt were performed as described by Hagenbuch *et al.*[S1] for deoxynucleotides. Methylimidazolides (**3a-t**) of dNMPs (**1a-t**) and MeIm derivatives (**6a-u**) of ribonucleotides **4a-u** were prepared via Mukaiyama redox condensation, using the triethylammonium salts of NMPs and 2-methylimidazole, 2,2'-dithiodipyridine, and triphenylphosphine, as described by Lohrmann and Orgel.[S4] Purification on cartridges used Sep-Pak C18 cartridges (Waters) with a gradient of MeCN/H_2O, unless otherwise noted. In a typical protocol, the cartridge was washed with acetonitrile and then with demineralized water. An aqueous solution of the crude product of the activation reaction (50 µL for 15 mg of crude monomer) was loaded on the cartridge, washed with aqueous NaCl solution (1 M, 10 mL) and the product eluted with a gradient of acetonitrile (0% to 20%) in H_2O at ambient pressure. Fractions of 1.5 mL were collected and those fractions containing pure activated monomer were lyophilized and stored under argon at -20 °C prior to NMR titration or extension assays.

Analytical Data for Activated Nucleotides

OAt-dAMP (**2a**); yield: 15 mg (32 µmol, 28%); ESI MS (m/z) calc: 448 $[M-H]^-$, found 448.1

OAt-dCMP (**2c**); yield: 18 mg (41 µmol, 34%); ESI MS (m/z) calc: 424 $[M-H]^-$, found 424.3

OAt-dGMP (**2g**); yield: 16 mg (32 µmol, 29%); ESI MS (m/z) calc: 464 $[M-H]^-$, found 465.1

OAt-dTMP (**2t**); yield: 14 mg (30 µmol, 25%); ESI MS (m/z) calc: 439 $[M-H]^-$, found 440.5

OAt-AMP (**5a**); yield: 10 mg (22 µmol, 20%); ^{31}P-NMR (161.9 MHz, D_2O) δ = -1.15 ppm; ESI MS (m/z) calc: 464 $[M-H]^-$, found 464.2.

OAt-CMP (**5c**); yield: 9 mg (20 µmol, 17%); ^{31}P-NMR (161.9 MHz, D_2O) δ = - 0.97 ppm; ESI MS (m/z) calc: 440 $[M-H]^-$, found 441.5.

OAt-GMP (**5g**); yield: 12 mg (24 µmol, 22%); ^{31}P-NMR (121.5 MHz, D_2O) δ = - 0.92 ppm; ESI MS (m/z) calc: 480 $[M-H]^-$, found 480.3

OAt-UMP (**5u**); yield: 15 mg (33 µmol, 28%); ^{31}P-NMR (121.5 MHz, D_2O) δ = - 0.97 ppm; ESI MS (m/z) calc: 442 $[M-H]^-$, found 442.3

2-MeIm-dAMP (**3a**); yield: 21 mg (50 µmol, 35%); ^{31}P-NMR (121.5 MHz, D_2O) δ = - 8.37 ppm; ESI MS (m/z) calc: 394 $[M-H]^-$, found 395.3

2-MeIm-dCMP (**3c**); yield: 29 mg (73 µmol 48%); ^{31}P-NMR (121.5 MHz, D_2O) δ = - 8.46 ppm; ESI MS (m/z) calc: 370 $[M-H]^-$, found 371.1

2-MeIm-dGMP (**3g**); yield: 23 mg (53 µmol, 39%); ^{31}P-NMR (121.5 MHz, D_2O) δ = - 8.09 ppm; ESI MS (m/z) calc: 410 $[M-H]^-$, found 410.3

2-MeIm-dTMP (**3t**); yield: 23 mg (56 µmol, 83%); ^{31}P-NMR (121.5 MHz, D_2O) δ = - 8.18 ppm; ESI MS (m/z) calc: 385 $[M-H]^-$, found 385.2

2-MeIm-AMP (**6a**); yield: 17 mg (39 µmol, 35%); ^{31}P-NMR (161.9 MHz, D_2O) δ = - 7.82 ppm;

2-MeIm-CMP (**6c**); yield: 17 mg (41 µmol, 35%); ^{31}P-NMR (161.9 MHz, D_2O) δ= - 7.88 ppm;

2-MeIm-GMP (**6g**); yield: 15 mg (34 µmol, 32%) ^{31}P-NMR (161.9 MHz, D_2O) δ = - 7.76 ppm;

2-MeIm-UMP (**6u**); yield: 16 mg (39 µmol, 33%) ^{31}P-NMR (161.9 MHz, D_2O) δ = - 7.88 ppm;

NMR Spectra

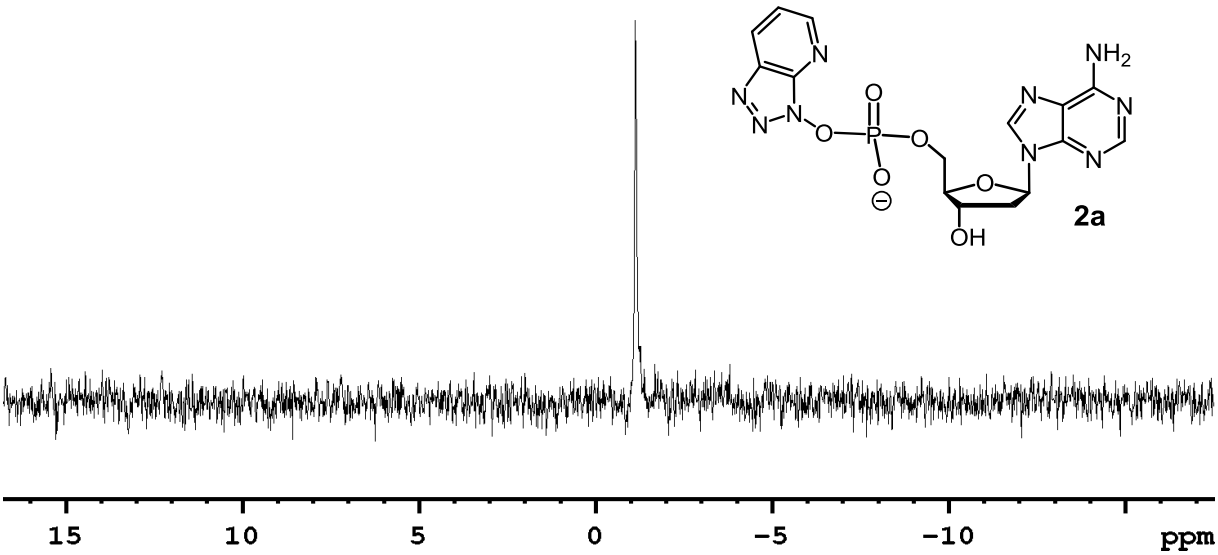


Figure S1. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of OAt-dAMP (**2a**).[S1]

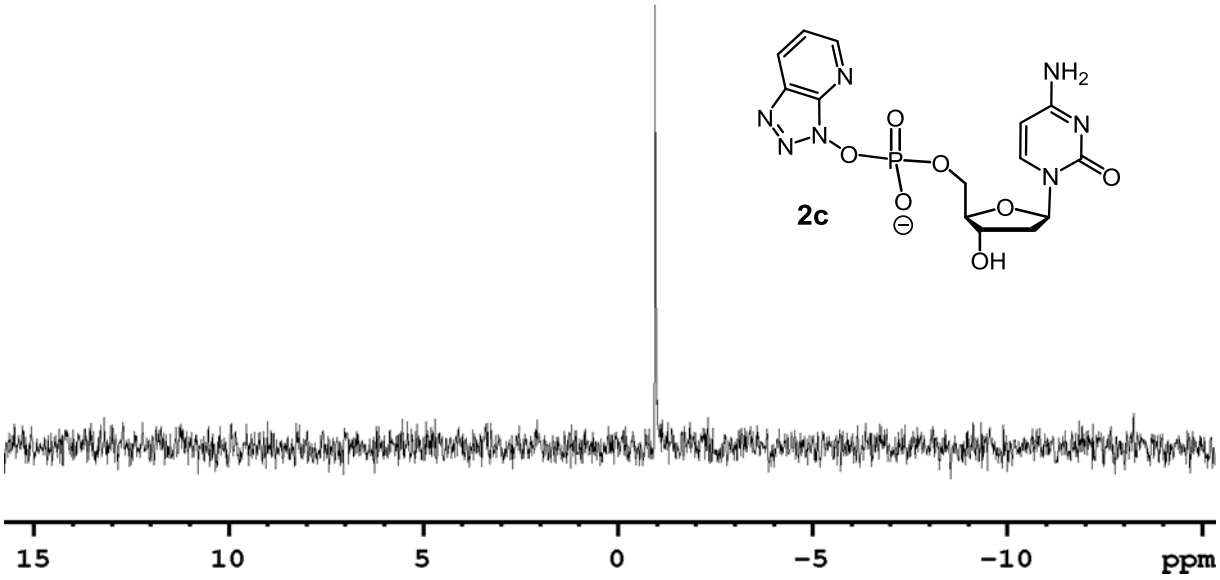


Figure S2. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of OAt-dCMP (**2c**).[S1]

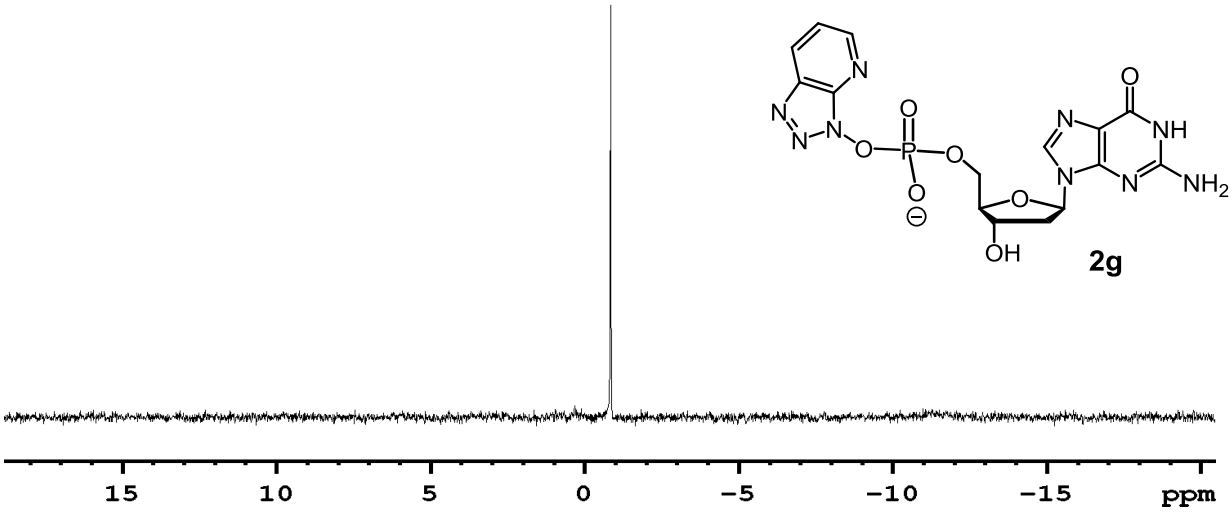


Figure S3. ^{31}P-NMR spectrum (161.9 MHz, D_2O) of OAt-dGMP (**2g**).[S1]

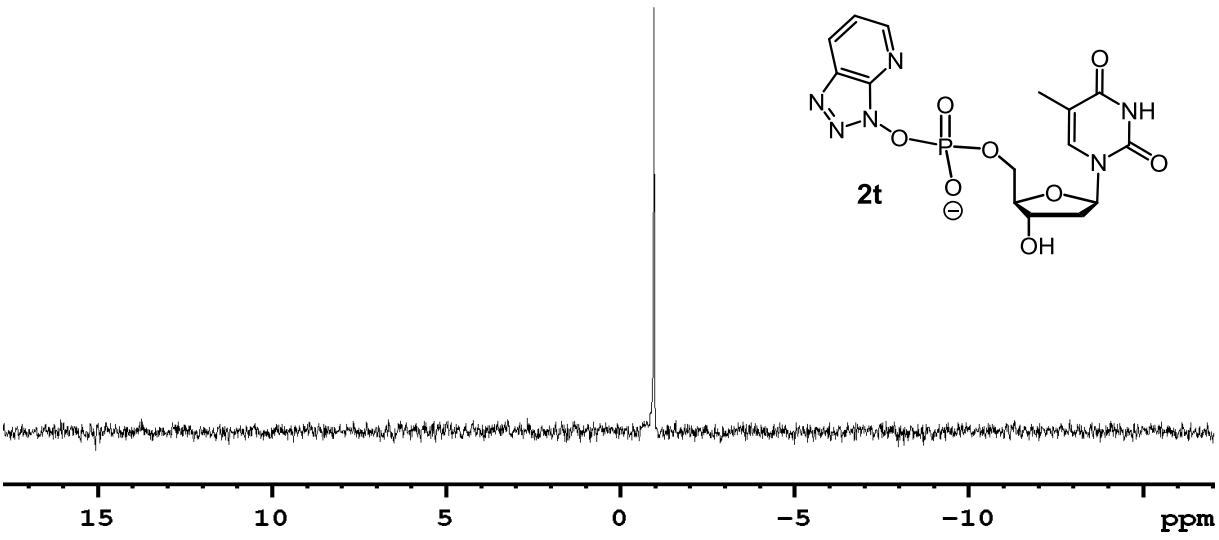


Figure S4. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of OAt-dTMP (**2t**).[S1]

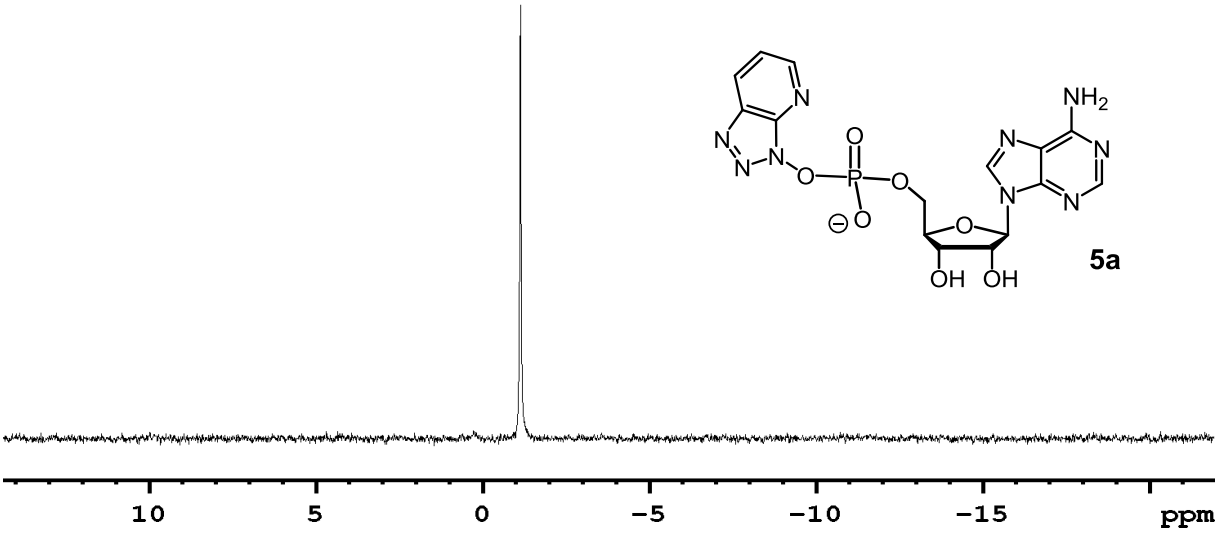


Figure S5. ^{31}P-NMR spectrum (161.9 MHz, D_2O) of OAt-AMP (**5a**); see ref. [S7] for data in a different solvent.

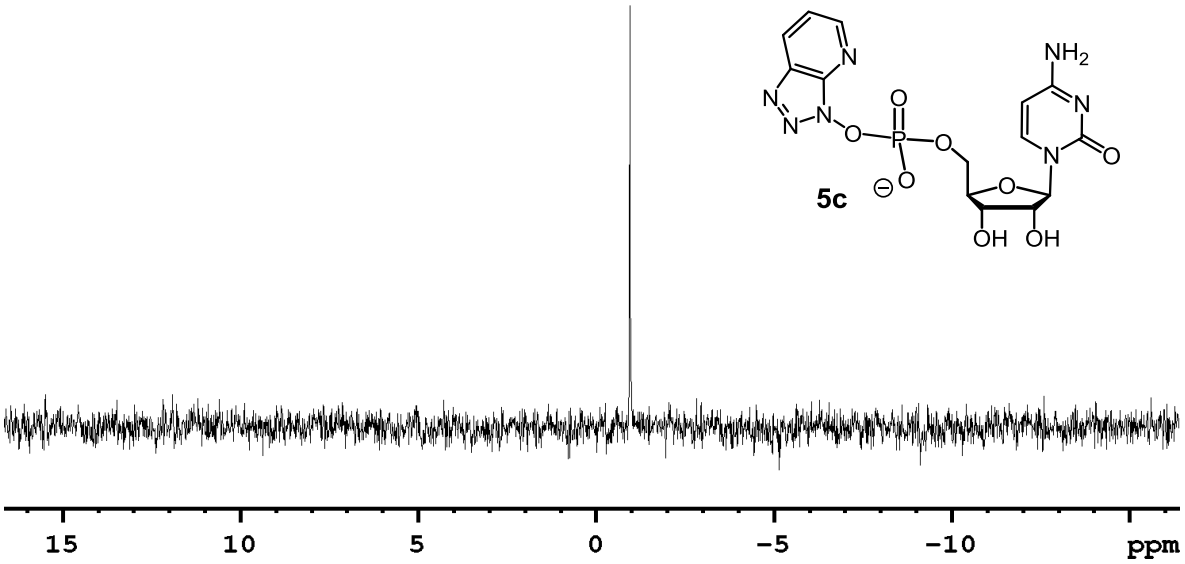


Figure S6. ^{31}P-NMR spectrum (161.9 MHz, D_2O) of OAt-CMP (**5c**); see ref. [S7] for data in a different solvent.

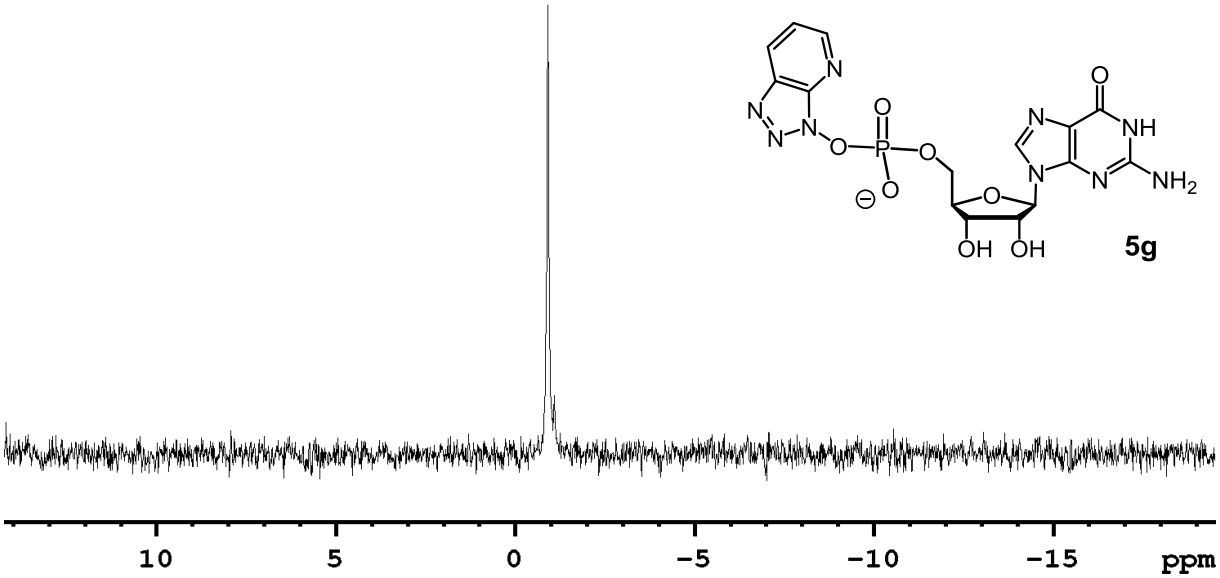


Figure S7. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of OAt-GMP (**5g**). see ref. [S7] for data in a different solvent.

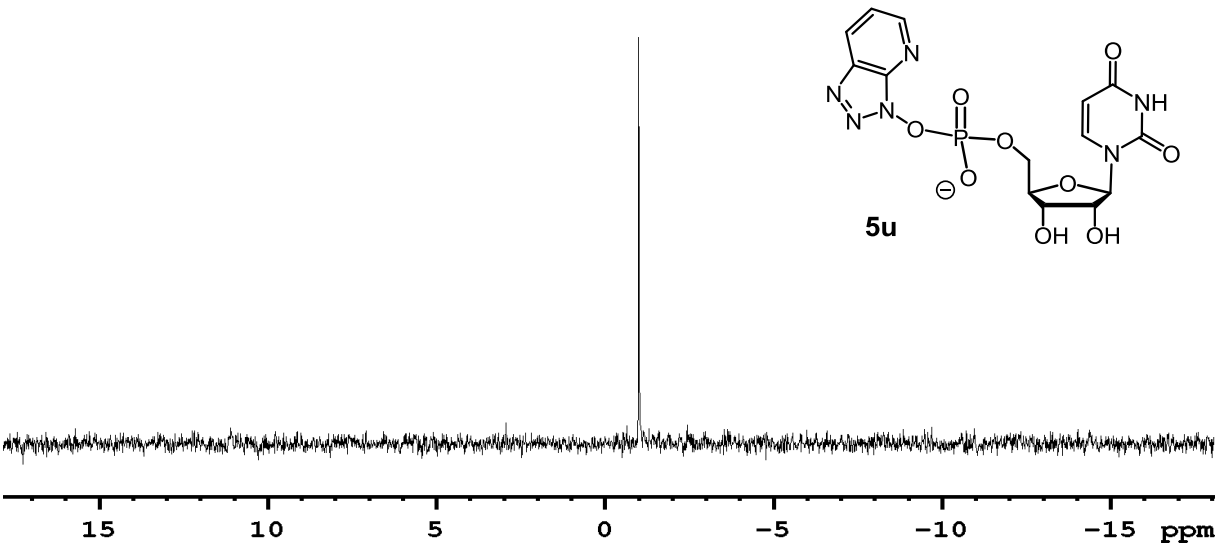


Figure S8. ^{31}P-NMR spectrum (161.9 MHz, D_2O)of OAt-UMP (**5u**).

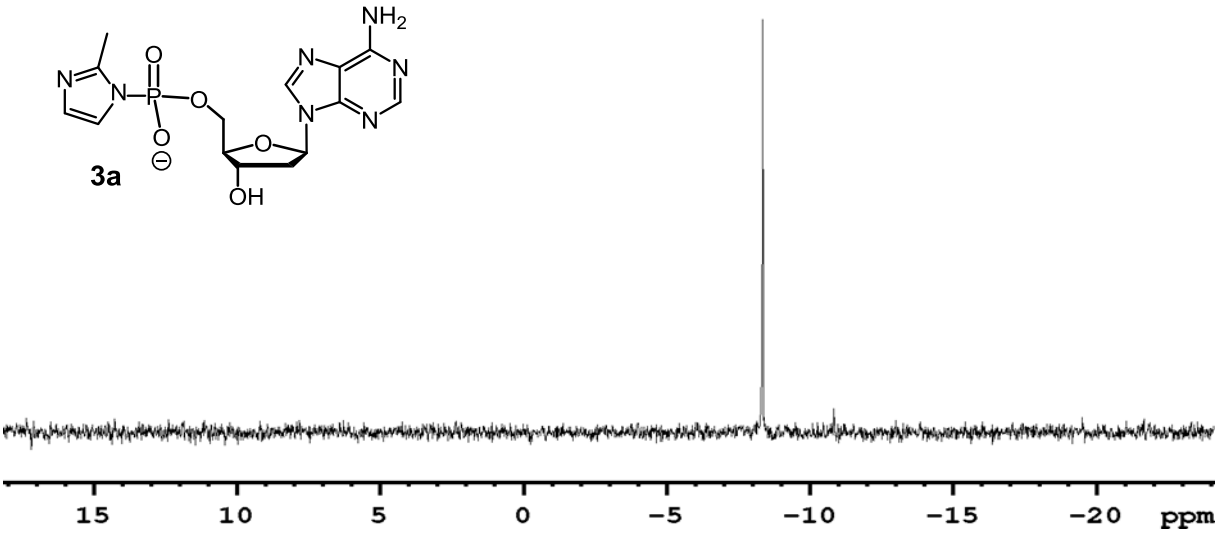


Figure S9. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of MeIm-dAMP (**3a**).

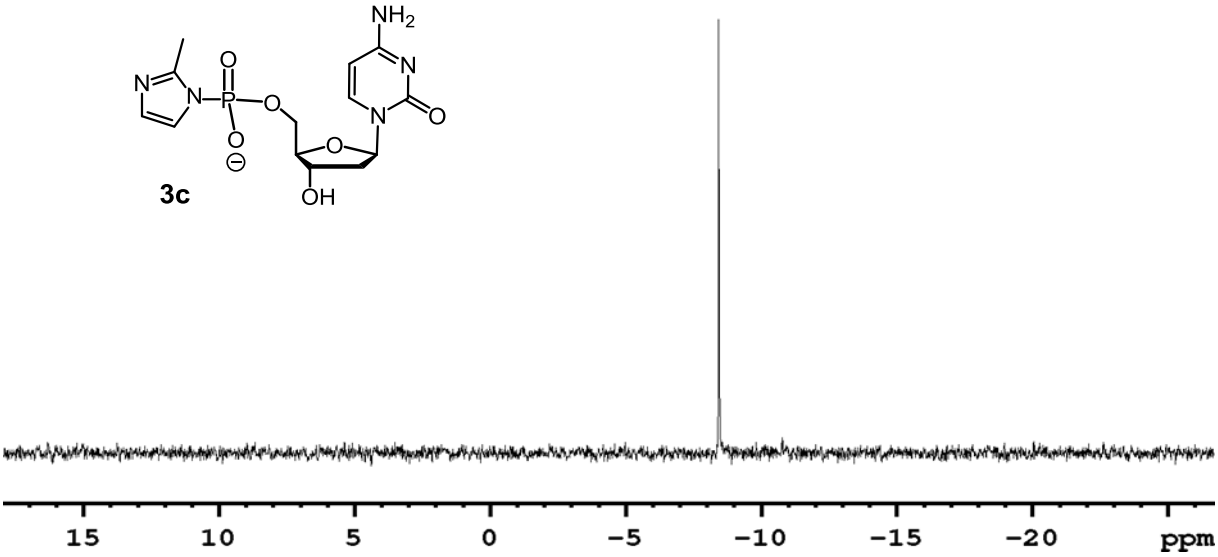


Figure S10. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of MeIm-dCMP (**3c**).

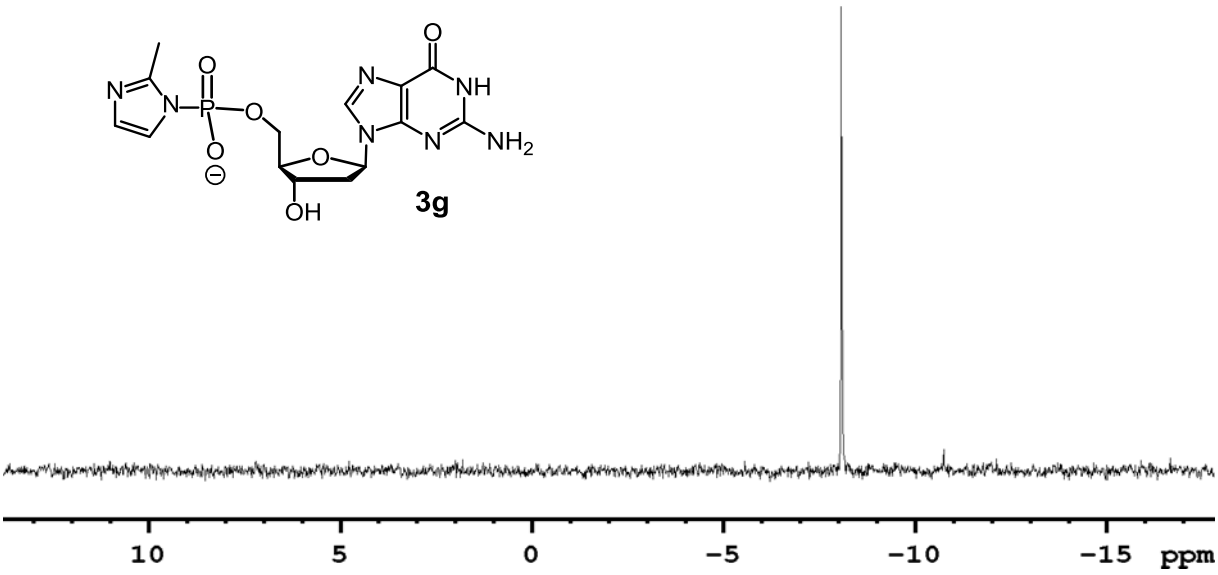


Figure S11. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of MeIm-dGMP (**3g**).

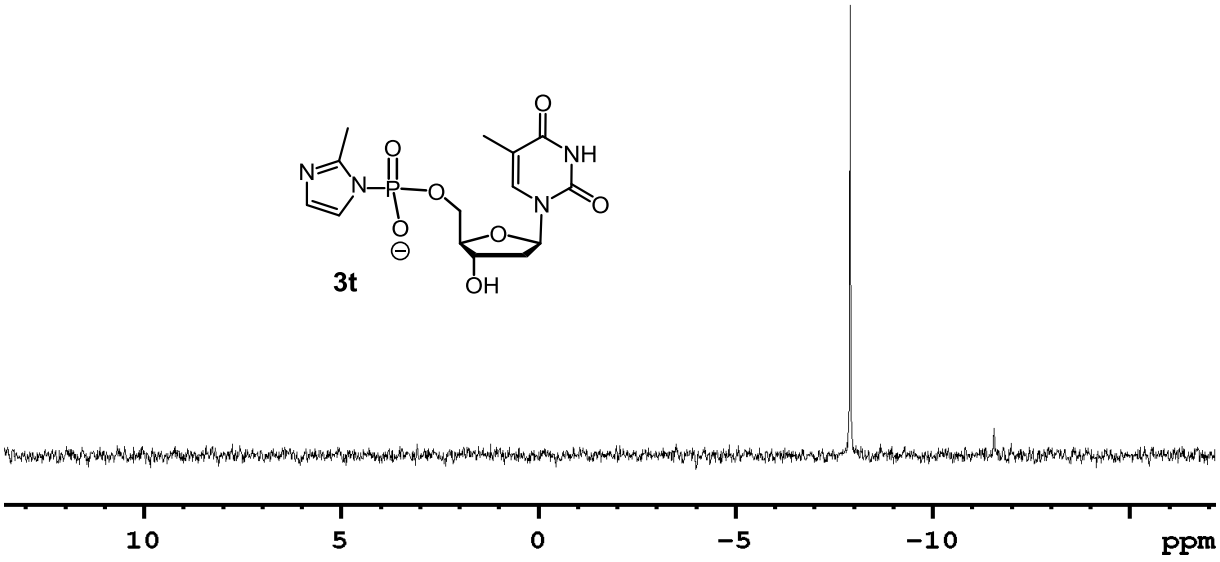


Figure S12. ^{31}P-NMR spectrum (121.5 MHz, D_2O) of MeIm-dTMP (**3t**).

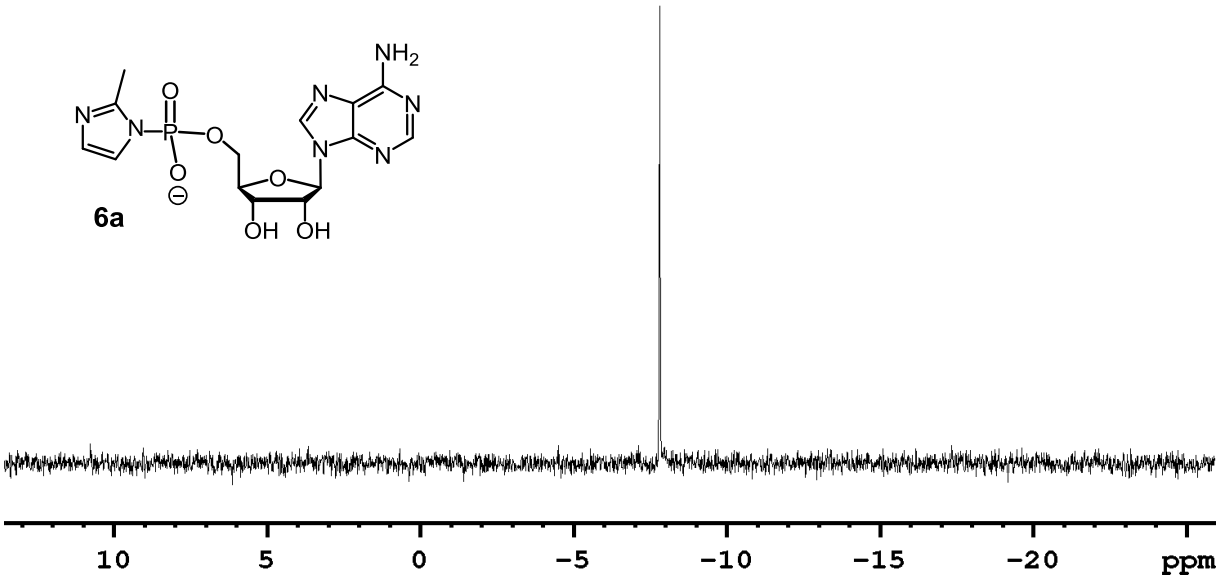


Figure S13. ^{31}P-NMR spectrum (161.9 MHz, D_2O) of MeIm-AMP (**6a**) see ref. [S6] for data in a different solvent

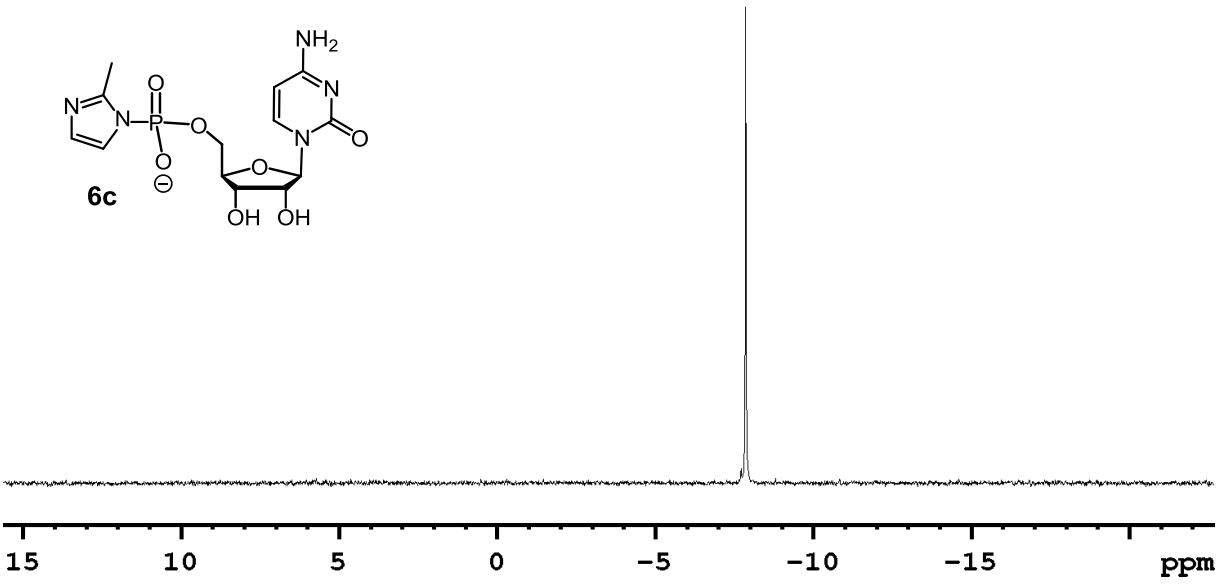


Figure S14. ^{31}P-NMR spectrum (161.9 MHz, D_2O) of MeIm-CMP (**6c**) see ref. [S6] for data in a different solvent

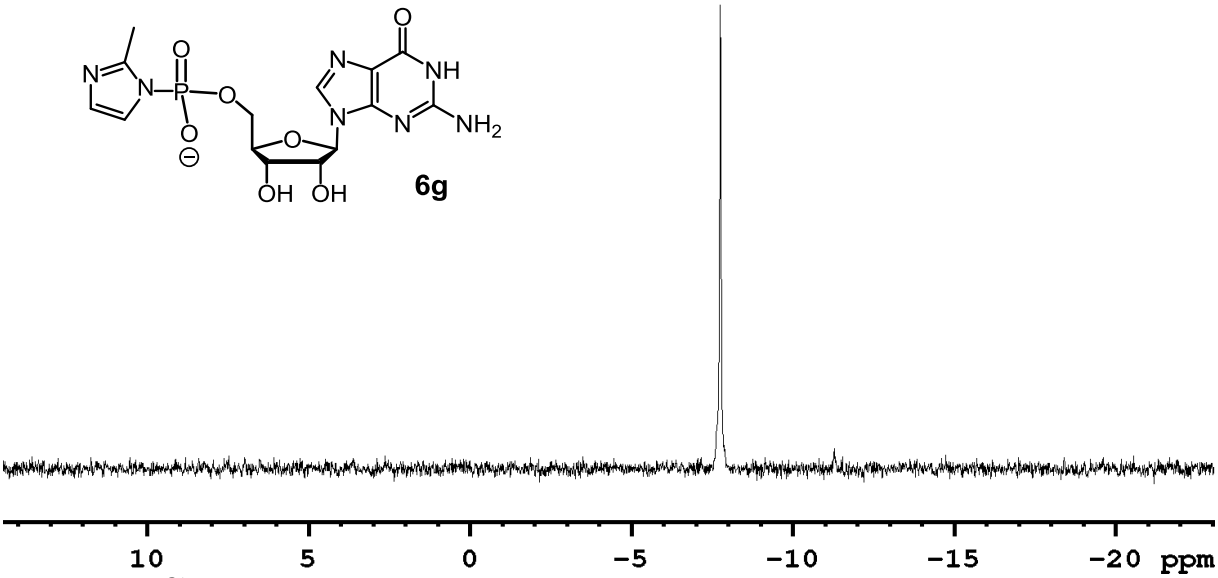


Figure S15. ^{31}P-NMR spectrum (161.9 MHz, D_2O) of MeIm-GMP (**6g**) see ref. [S6] for data in a different solvent

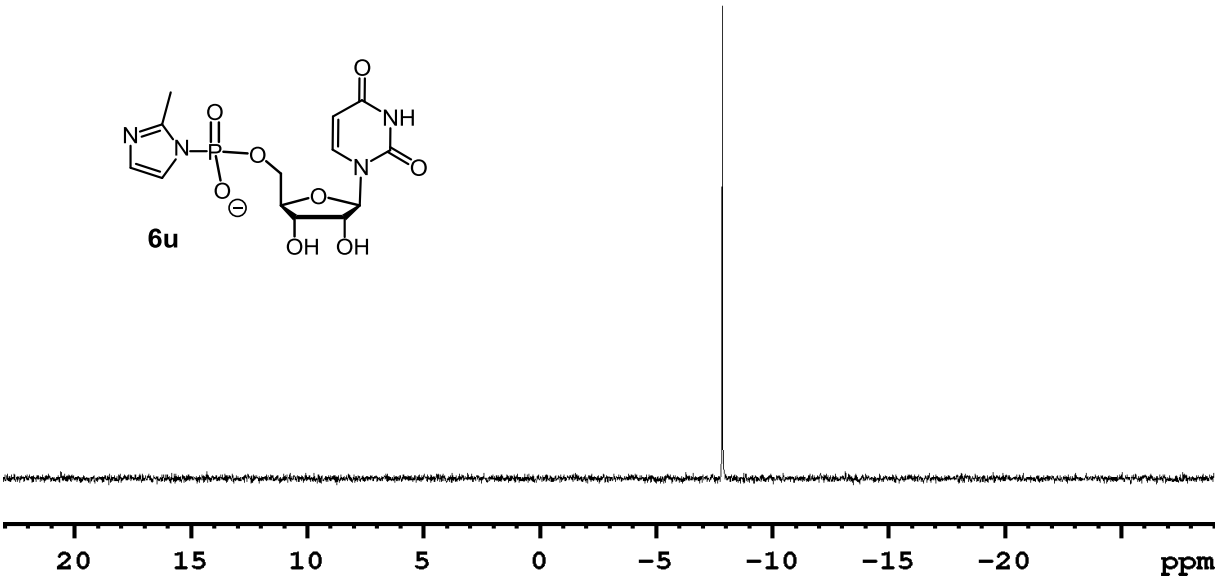


Figure S16. ^{31}P-NMR spectrum (161.9 MHz, D_2O) of MeIm-UMP (**6u**).

3. Kinetics of the Hydrolysis of Monomers

The kinetics of the hydrolysis of MeIm-activated deoxynucleotides (**3a-t**) were measured by ^{31}P-NMR, according to the procedure described in the literature for OAt-dNMPs. [S2] Each MeIm-dNMP (**3a-t**) was dissolved in HEPES buffer (200 mM, NaCl 400 mM, $MgCl_2$ 80 mM) in D_2O (600 µL) at pH 7.0, uncorrected for deuterium effect, to reach a final concentration of 60 mM of the monomer. The solutions were then analyzed by ^{31}P-NMR at 121.5 MHz with a relaxation time of 10 s. The time course of the conversion of the activated species to the free nucleotide was fitted by using the following equation for monoexponential decay:

$$y(t) = y_o + \exp(-k_{hydr} * t)$$

(Equation S1)

where *y* is the fraction of activated species left, k_{hydr} is the rate constant of hydrolysis, and y_0 is the maximum conversion of activated monomer at infinite reaction time, as obtained by fitting.

Deoxynucleotides

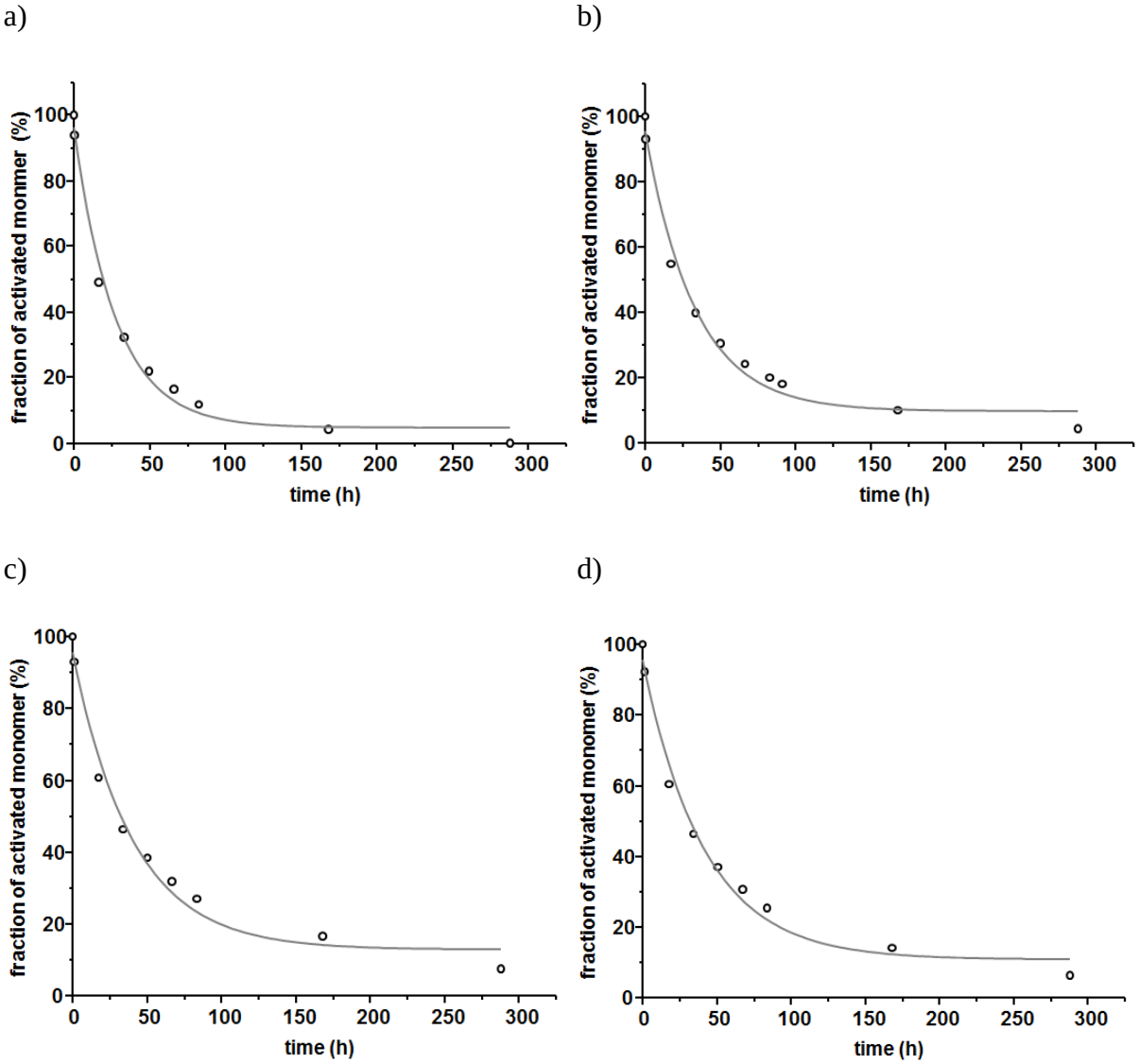


Figure S17. Kinetics of hydrolysis for 2-methylimidazolidess of 2'-deoxynucleotides. Symbols are data points and lines are fits. a) MeIm-dAMP (**3a**), b) MeIm-dCMP (**3c**), c) MeIm-dGMP (**3g**), and d) MeIm-TMP (**3t**).

Ribonucleotides

The kinetics of the hydrolysis of MeIm-activated ribonucleotides (**6a-u**) were measured at 161.9 MHz in HEPES buffer (200 mM, NaCl 400 mM, $MgCl_2$ 80 mM) in D_2O (200 µL) at pH 7.7 (uncorrected for deuterium effect), at monomer concentrations of 25 mM.

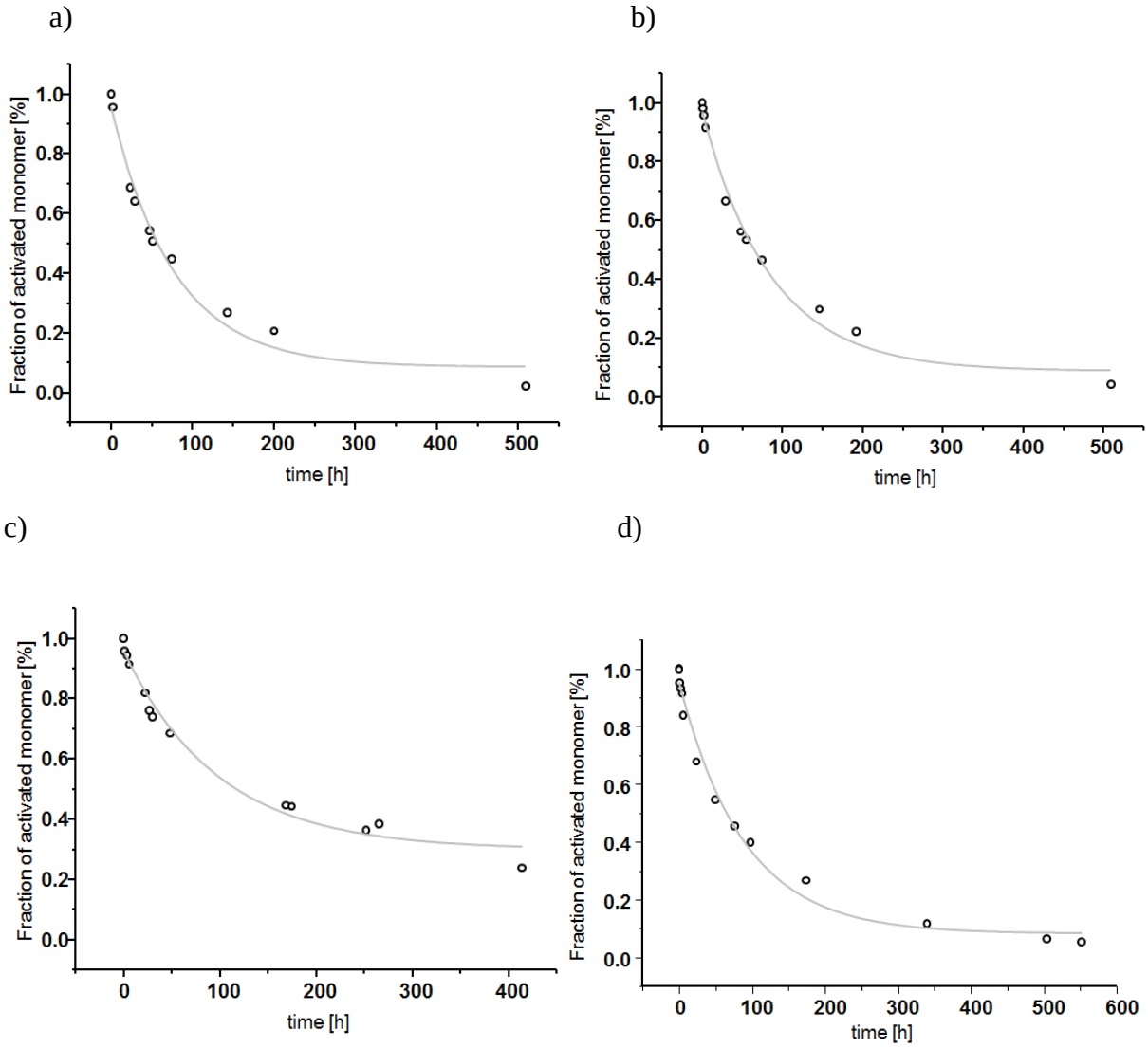


Figure S18. Kinetics of hydrolysis for 2-methylimidazolidess of ribonucleotides. Symbols are data points and lines are fits. a) MeIm-AMP (**6a**), b) MeIm-CMP (**6c**), c) MeIm-GMP (**6g**), and d) MeIm-UMP (**6u**).

4. NMR-Monitored Titration

UV-Melting Analysis of Hairpins

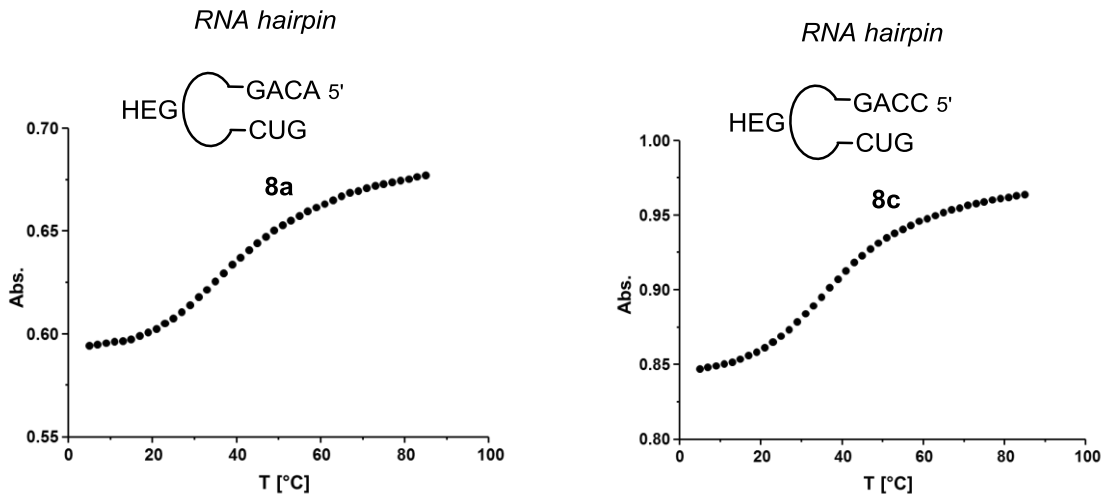


Figure S19. UV-Melting curves of RNA hairpins **8a** (left, 8 µM strand concentration) and **8c** (right, 13 µM strand concentration) in the buffer used for titrations (200 mM phosphate, 400 mM NaCl, 80 mM $MgCl_2$, pH 7.0), measured at 260 nm with a heating rate of 1 °C/min. The melting points, as obtained from the maximum of the first derivative, are 36 °C for **8a** and 36 °C for **8c**. Because folding is a unimolecular process, the melting point is independent of concentration, so that the melting curve results are also relevant for NMR-monitored titrations with the hairpins thus characterized.

Titrations. NMR titration experiments were performed using the procedure described in reference [S2], using phosphate buffer (200 mM, NaCl 400 mM, $MgCl_2$ 80 mM) at pH 7.0 or 8.9. A pH of 7.0 was used for all RNA monomers to suppress extension of the RNA hairpin. To the solution of the hairpin oligonucleotide were added aliquots of the nucleotide solution. The resulting solution was centrifuged in a set-up with a plastic tube surrounding the glass tube (2 min, 25 g) to remove any residual gas bubbles. In the samples, 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid (TPA, 0.1 mM) was used as internal reference. Chemical shifts were determined graphically from the centers of peaks. Dissociation constants K_ds were calculated using:

$$\delta shift = \frac{\delta max}{K_d} * \frac{x}{1 + \frac{x}{K_d}}$$

(Equation S2)

δmax is the maximal absolute value of the shift of the signal, *x* is the concentration of the monomer, *δshift* is the actual displacement of the chemical shift of the signal in question, and K_d is the dissociation constant.

DNA Hairpin 7a

a) b)

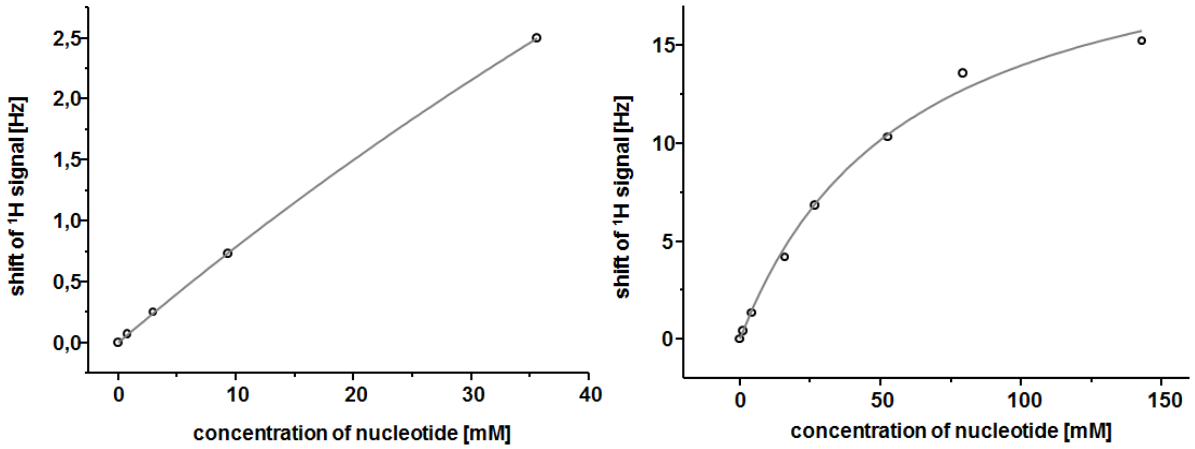


Figure S20. Plots of NMR titration data for:

a) Chemical shift of A1H2 signal of hairpin 5'-**A**CAG(HEC)CTG-3' versus concentration of MeIm-TMP **3t** (pH 7.0)

K_d : 236 mM

b) Chemical shift of A1H2 signal of hairpin 5'-**A**CAG(HEC)CTG-3' versus concentration of OAt-TMP **2t** (pH 8.9)

K_d : 71 mM

DNA Hairpin 7c

a) b)

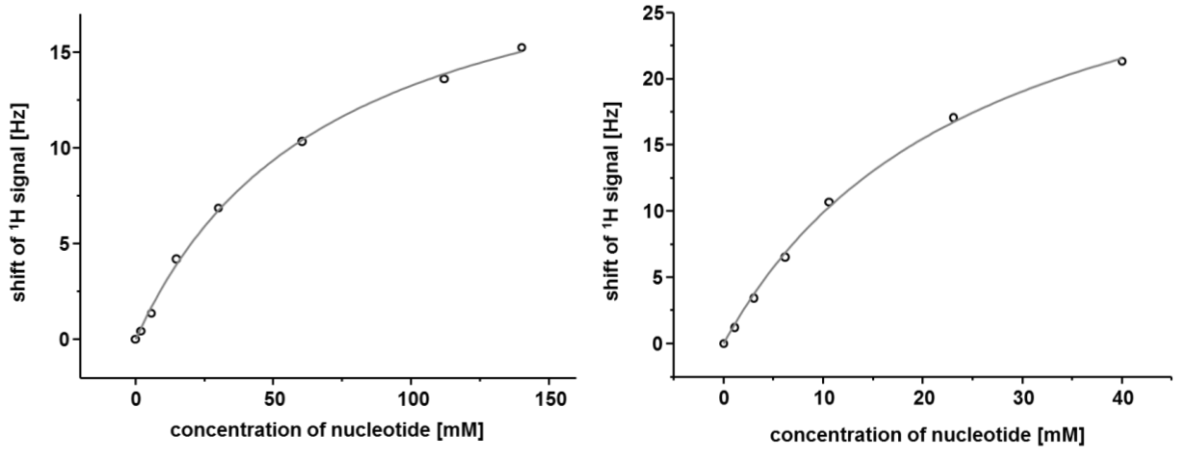


Figure S21. Plots of NMR titration data for:

a) Chemical shift of C1H6 signal of hairpin 5'-**C**CAG(HEG)CTG-3' versus concentration of MeIm-dGMP **3g** (pH 7.0),

K_d : 25 mM

b) Chemical shift of C1H6 signal of hairpin 5'-**C**CAG(HEG)CTG-3' versus concentration of OAt-dGMP **2g** (pH 8.9)

K_d : 26 mM

DNA Hairpin 7g

a) b)

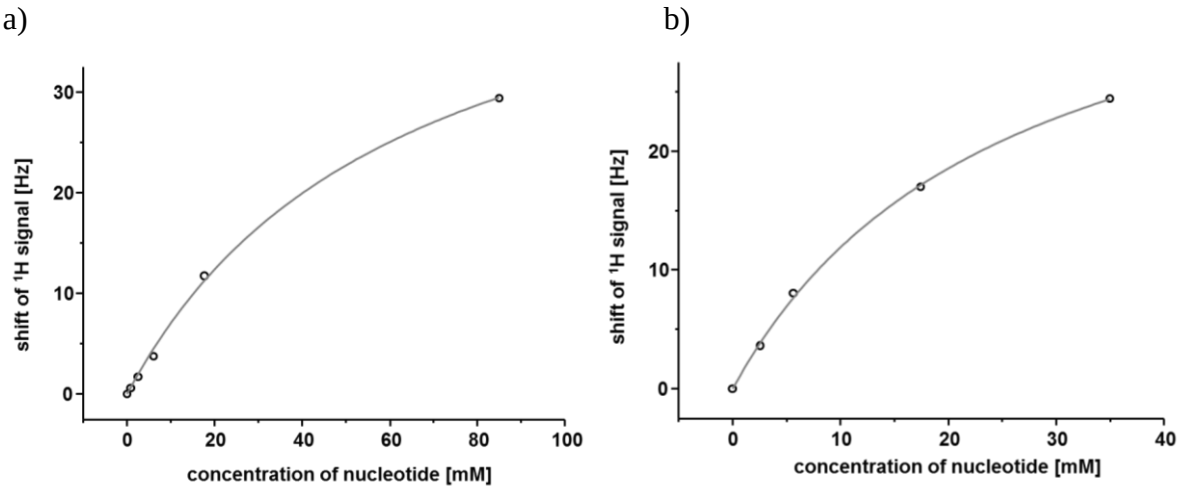


Figure S22. Plots of NMR titration data for:

a) Chemical shift of G1H8 signal of hairpin 5'-**G**CAG(HEG)CTG-3' versus concentration of MeIm-dCMP **3c** (pH 7.0).

K_d : 26 mM

b) Chemical shift of G1H8 signal of hairpin 5'-**G**CAG(HEG)CTG-3' versus concentration of OAt-dCMP **2c** (pH 8.9).

K_d : 25 mM

DNA Hairpin 7t

a) b)

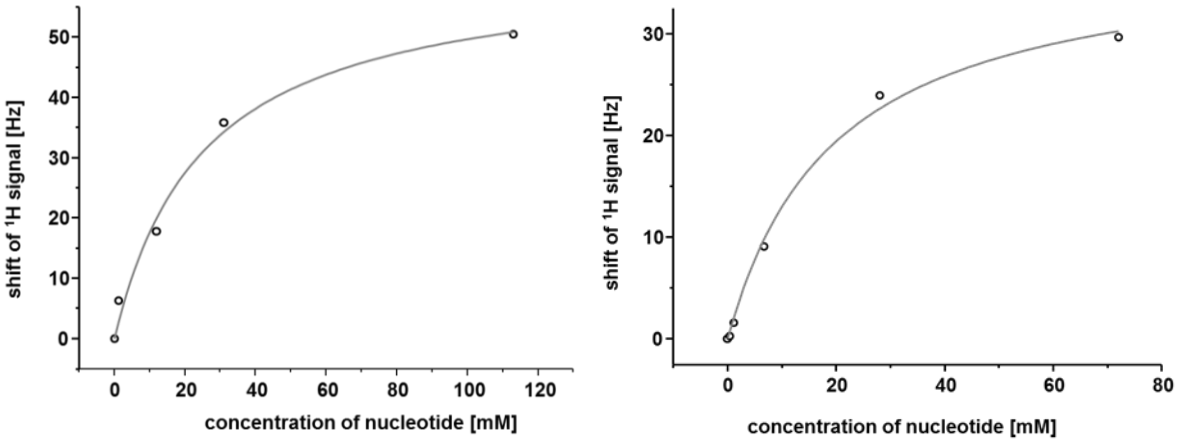


Figure S23. Plots of NMR titration data for:

a) Chemical shift of T1H6 signal of hairpin 5'-**T**CAG(HEG)CTG-3' versus concentration of MeIm-dAMP **3a** (pH 7.0).

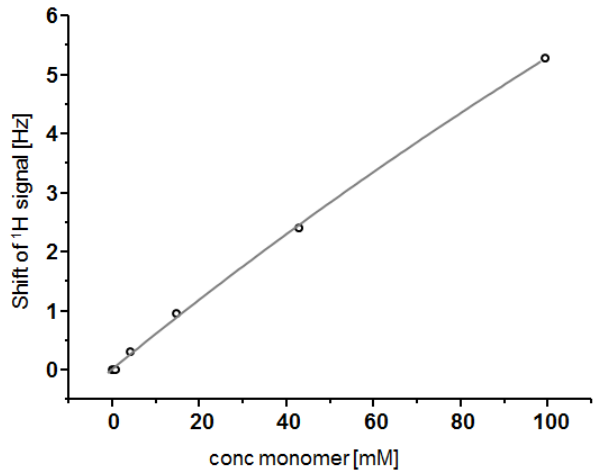
K_d : 37 mM

b) Chemical shift of T1H6 signal of hairpin 5'-**T**CAG(HEG)CTG-3' versus concentration of OAt-dAMP **2a** (pH 8.9).

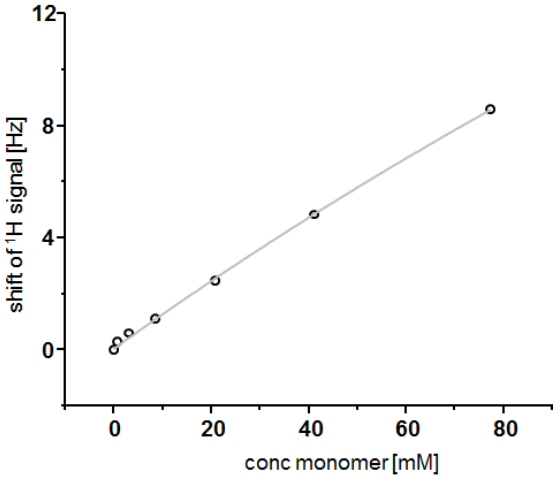
K_d : 20 mM

RNA Hairpin 8a

a)



b)



c)

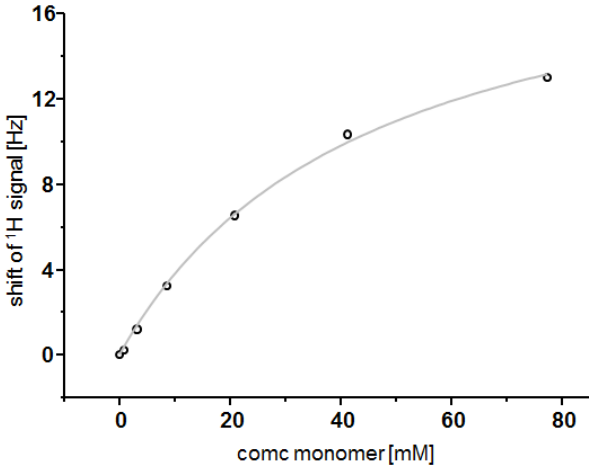


Figure S24. Plots of NMR titration data for:

a) Chemical shift of A1H2 signal of hairpin 5'-**A**CAG(HEC)CUG-3' versus concentration of UMP (**3u**) **at** pH 7.0

K_d : > 500 mM

b) Chemical shift of A1H2 signal of haiprin 5'-**A**CAG(HEC)CUG-3' versus concentration of MeIm-UMP (**6u**) at pH 7.0

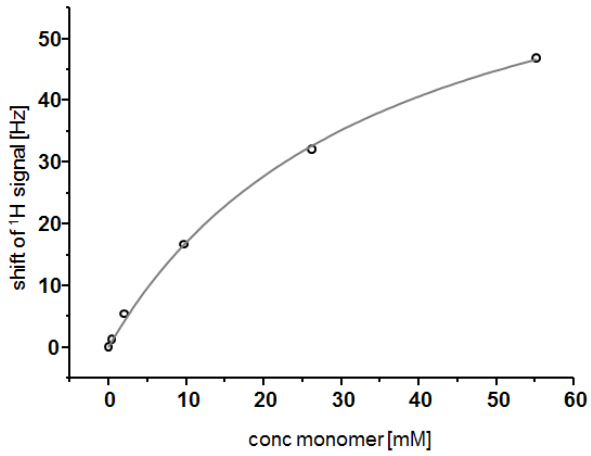
K_d : > 500 mM

c) Chemical shift of A1H2 signal of hairpin 5'-**A**CAG(HEC)CUG-3' versus concentration of OAt-UMP (**5u**) at pH 7.0

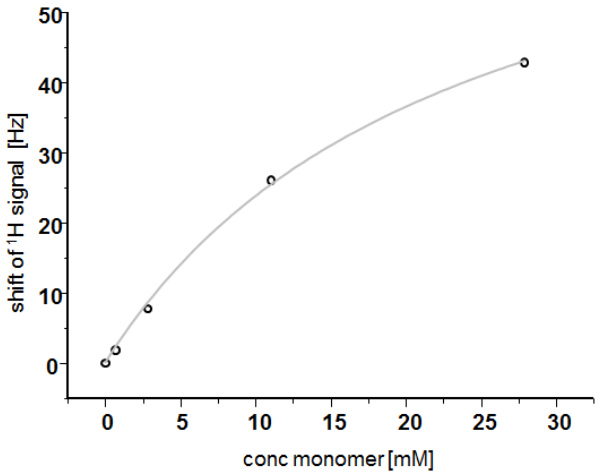
K_d : 43 mM

RNA Hairpin 8c

a)



b)



c)

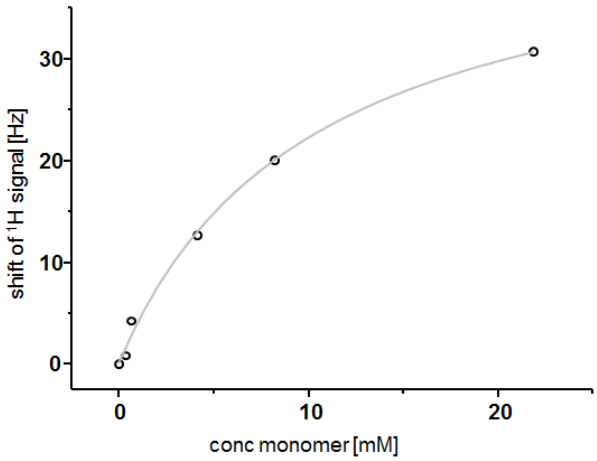


Figure S25. Plots of NMR titration data for:

a) Chemical shift of C1H6 signal of hairpin 5'-**C**CAG(HEG)CUG-3' versus concentration of GMP (**4g**) at pH 7.0

K_d : 35 mM

b) Chemical shift of C1H6 signal of hairpin 5'-**C**CAG(HEG)CUG-3' versus concentration of MeIm-GMP (**6g**) at pH 7.0

K_d : 23 mM

c) Chemical shift of C1H6 signal of hairpin 5'-**C**CAG(HEG)CUG-3' versus concentration of OAt-GMP (**5g**) at pH 7.0

K_d : 10 mM

RNA Hairpin 8g in pure D_2O

a) b)

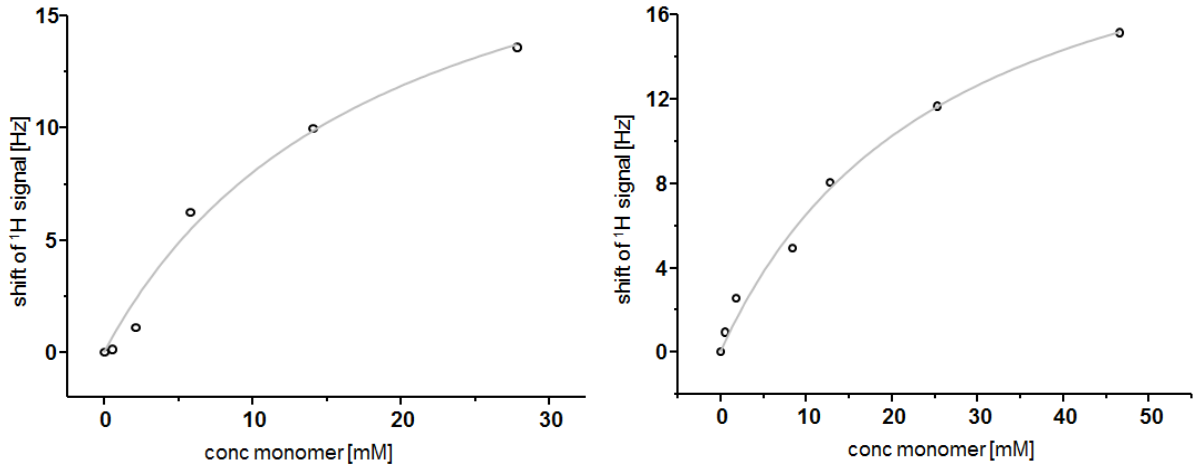


Figure S26. Plots of NMR titration data for:

a) Chemical shift of G1H8 signal of hairpin 5'-**G**CAG(HEG)CUG-3' versus concentration of CMP **4c** (pH 7.0 D_2O).

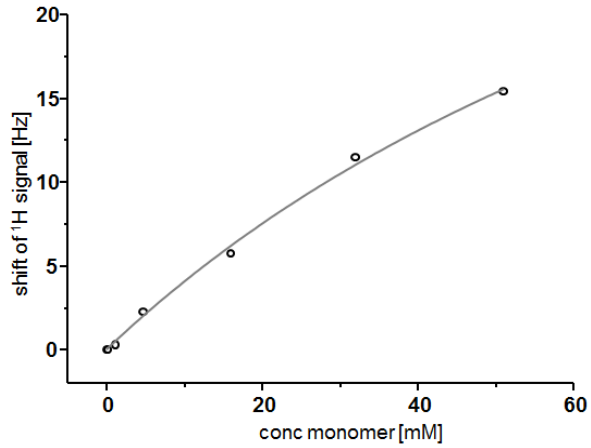
K_d : 19 mM

b) Chemical shift of G1H8 signal of hairpin 5'-**G**CAG(HEG)CUG-3' versus concentration of MeIm-CMP **6c** at pH 7.0 in D_2O.

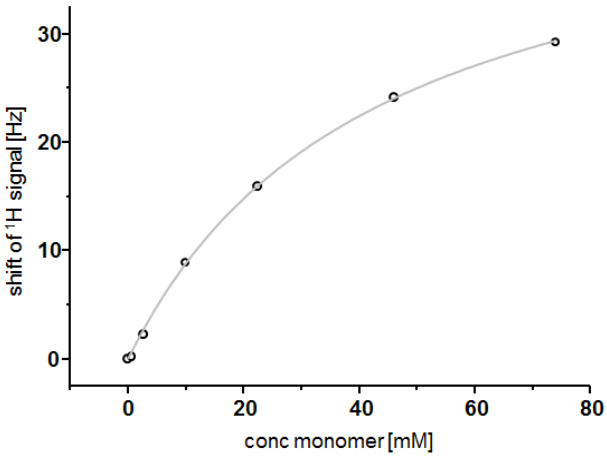
K_d : 27 mM

RNA Hairpin 8u

a)



b)



c)

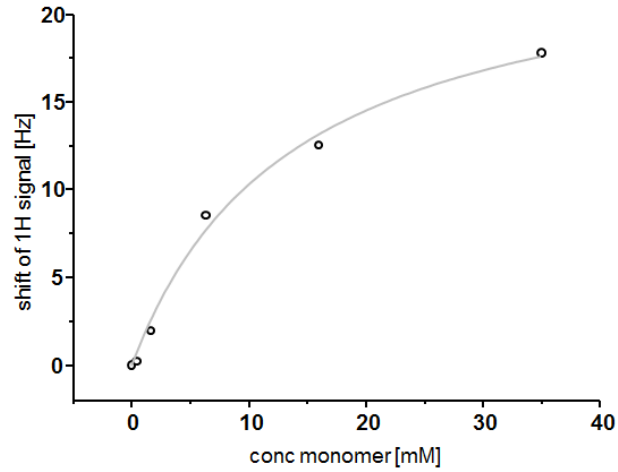


Figure S27. Plots of NMR titration data for:

a) Chemical shift of U1H6 signal of hairpin 5'-**U**CAG(HEG)CUG-3' versus concentration of AMP (**4a)** at pH 7.0

K_d : 90 mM

b) Chemical shift of U1H6 signal of hairpin 5'-**U**CAG(HEG)CTG-3' versus concentration of MeIm-AMP (**6a**) at pH 7.0

K_d : 42 mM

c) U1H6 signal of sequence 5'-**U**CAG(HEG)CTG-3' versus concentration of OAt-AMP (**5a**) at pH 7.0

K_d : 14 mM

5. Primer Extension Assays

The protocol for primer extensions was adapted from that given in reference [S5]. Assays were performed in a final volume of 10 µL. For DNA, a mixture of 3'-aminoprimer **10** (36 µM) and DNA template (**9a-t**) (54 µM, 1.5 eq) was annealed by heating to 60 °C for 5 min in HEPES buffer (200 mM, NaCl 400 mM, $MgCl_2$ 80mM) and then slowly cooling to 20°C. Assays were performed at pH 7.0 for MeIm-dNMP (**3a-t**) and at pH 8.9 for OAt-dNMP (**2a-t**). For RNA, a duplex of primer **13** (36 µM) and template (**12a-u**) (36 µM) was used. Assays with MeIm-NMP (**6a-u**) were performed at pH 7.7 and assays with OAt-NMP (**5a-u**) were run at pH 8.9. Assays were at different concentrations of an activated monomer were run in parallel.

Assays were run over a period 24-72 h (5 days for RNA), depending on the monomer concentration and the rate of the reaction. Kinetic rates were obtained by fitting with a monoexponential model[S5] to obtain k' values using the following equation :

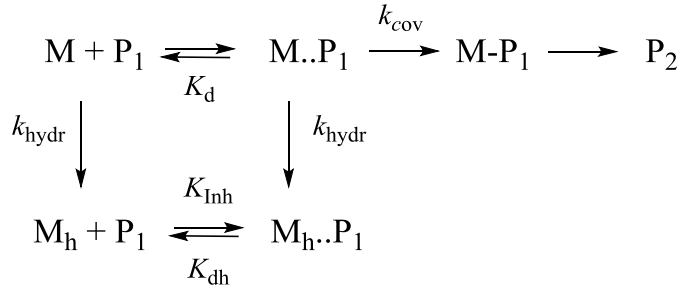
$$P_2(t) = y_0 * (1 - e^{-k.t})$$

(Equation S3)

where P_2 is the fraction of extended primer at time t, k is the pseudo-first order rate constant and y_0 is the maximum conversion at infinite time obtained by fitting.

6. Mathematical Model for Primer Extension

The model used was adapted from previous work.[S2] The following scheme shows the equations used.



Here, M is the activated monomer, M_h the hydrolyzed monomer (free nucleotide), P_1 the primer:template complex, $M_{(h)}..P_1$ the non-covalent complex of free nucleotide and primer-template duplex, M-P_1 is the covalent intermediate of the extension reaction, and P_2 is the product of the reaction (extended primer, bound to the template). The dissociation constants K_d and K_{dh} are for the complexes of activated monomer and free monomer to the primer:template complex, respectively. The kinetic constant for the hydrolysis of activated monomers (k_{hydr}) was assumed to be independent of the presence or absence of the primer:template complex. The rate constant k_{cov} is as the rate of the formation of the covalent step(s) of the extension reaction. The initial phase of a reaction (reaction time < 1/5 of $t_{1/2}$ of hydrolysis of monomer), during which the concentration of hydrolyzed monomer is negligible, was used to determine the k_{cov} values, using a monoexponential model (Equation S3).

k_{cov} was determined as following: $$k_{cov} = \frac{k}{\alpha}$$

with the fraction of binding site occupied by the monomer given as $$\alpha = \frac{[M]_0}{K_d + [M]_0}$$

The kinetic and thermodynamic constants were then entered into the following equations :

$$Y_{P_2}(t) = 1 - \exp\left[\frac{-k_{cov}[M]_0}{K_d k_{hydr} B}\left(k_{hydr}t - ln\left(\frac{B + Ae^{k_{hydr}t}}{B + A}\right)\right)\right]$$

(Equation 1)

For a full description of the model equation, see reference [S2].

7. Additional Simulated Kinetics

Aminoterminal DNA primer and OAt esters

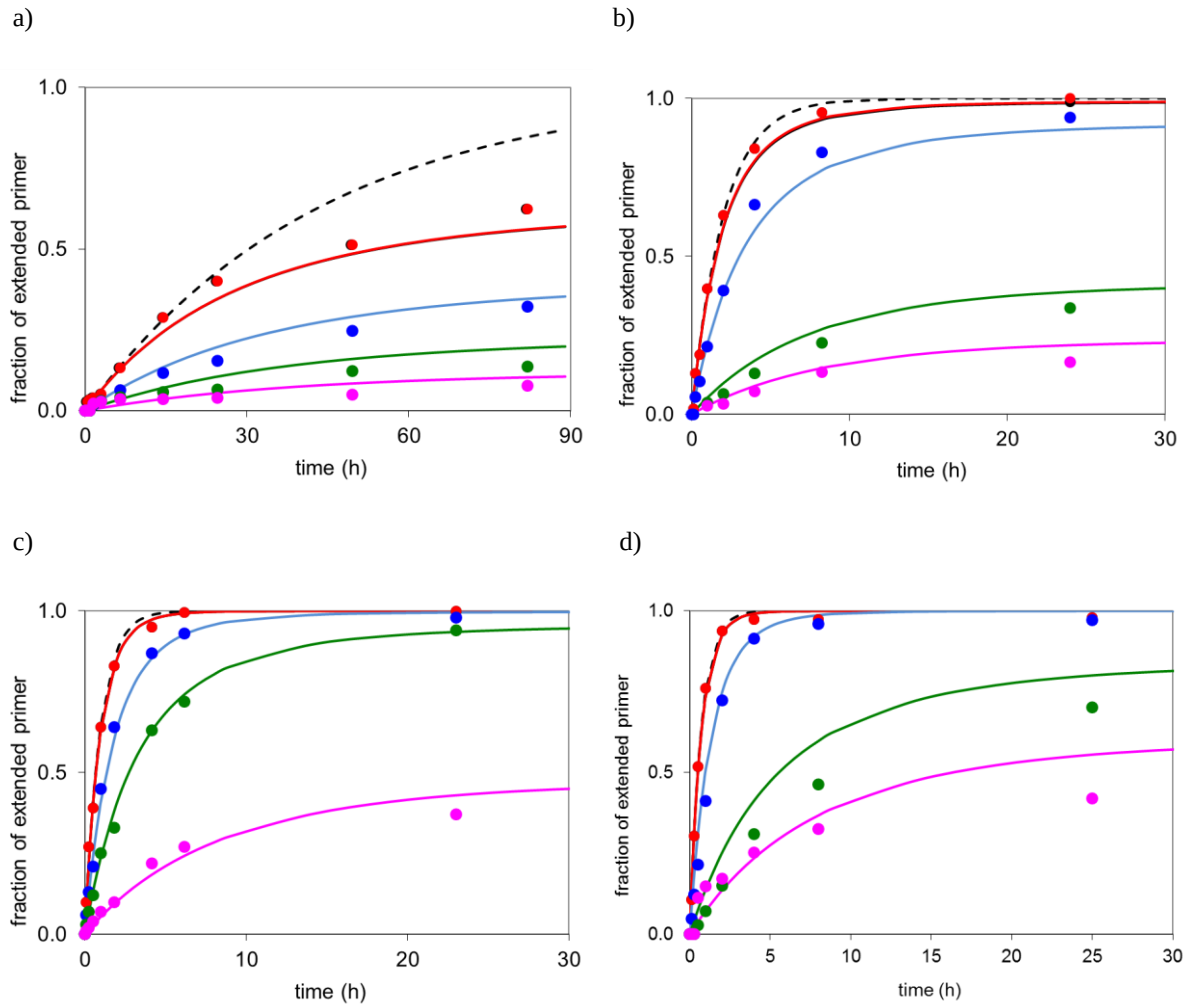


Figure S28. Simulated time-dependent yield of extended DNA primer with and without hydrolysis and inhibition. Initial concentrations of 36 µM primer were used. Filled circles represent experimental values with different concentrations of OAt-activated monomer **2a-t**. Solid lines are values calculated by equation (S2).
a) Template **9a** and the following concentrations of (OAt-dTMP) **2t**; red 7.2 mM, blue 3.6 mM, green 0.72 mM and pink 0.18 mM. Values used for the simulation: K_d = 71 mM; K_{dinh} = 240 mM; k_{hydr} = 0.044 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition for (k = 0.159 h^{-1}) at 7.2 mM OAt-dTMP, with k_{cov} = 1.73 h^{-1}.
b) Template **9t** and the following concentrations of (OAt-dAMP) **2a**; red 3.6 mM, blue 1.8 mM, green 0.36 mM and pink 0.18 mM. Values used for the simulation: K_d = 20 mM; K_{dinh} = 38 mM; k_{hydr} = 0.109 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.436 h^{-1}) at 3.6 mM OAt-dAMP, with k_{cov} =3.24 h^{-1}.
c) Template **9c** and the following concentrations of (OAt-dGMP) **2g**; red 3.6 mM, blue 1.8 mM, green 0.9 mM and pink 0.18 mM. Values used for the simulation: K_d = 26 mM; K_{dinh} = 27 mM; k_{hydr} = 0.093 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.84 h^{-1}) at 3.6 mM OAt-dGMP, with k_{cov} = 8.61 h^{-1}.
d) Template **9g** and the following concentrations of (OAt-dCMP) **2c**; red 3.6 mM, blue 1.8 mM, green 0.36 mM and pink 0.18 mM. Values used for the simulation: K_d = 25 mM; K_{dinh} = 34 mM; k_{hydr} = 0.087 h^{-1}; broken black line for a hypothetical reaction without inhibition (k = 1.25 h^{-1}) at 3.6 mM OAt-dCMP, with k_{cov} = 9.9 h^{-1}.

Aminoterminal primer and methylimidazolides

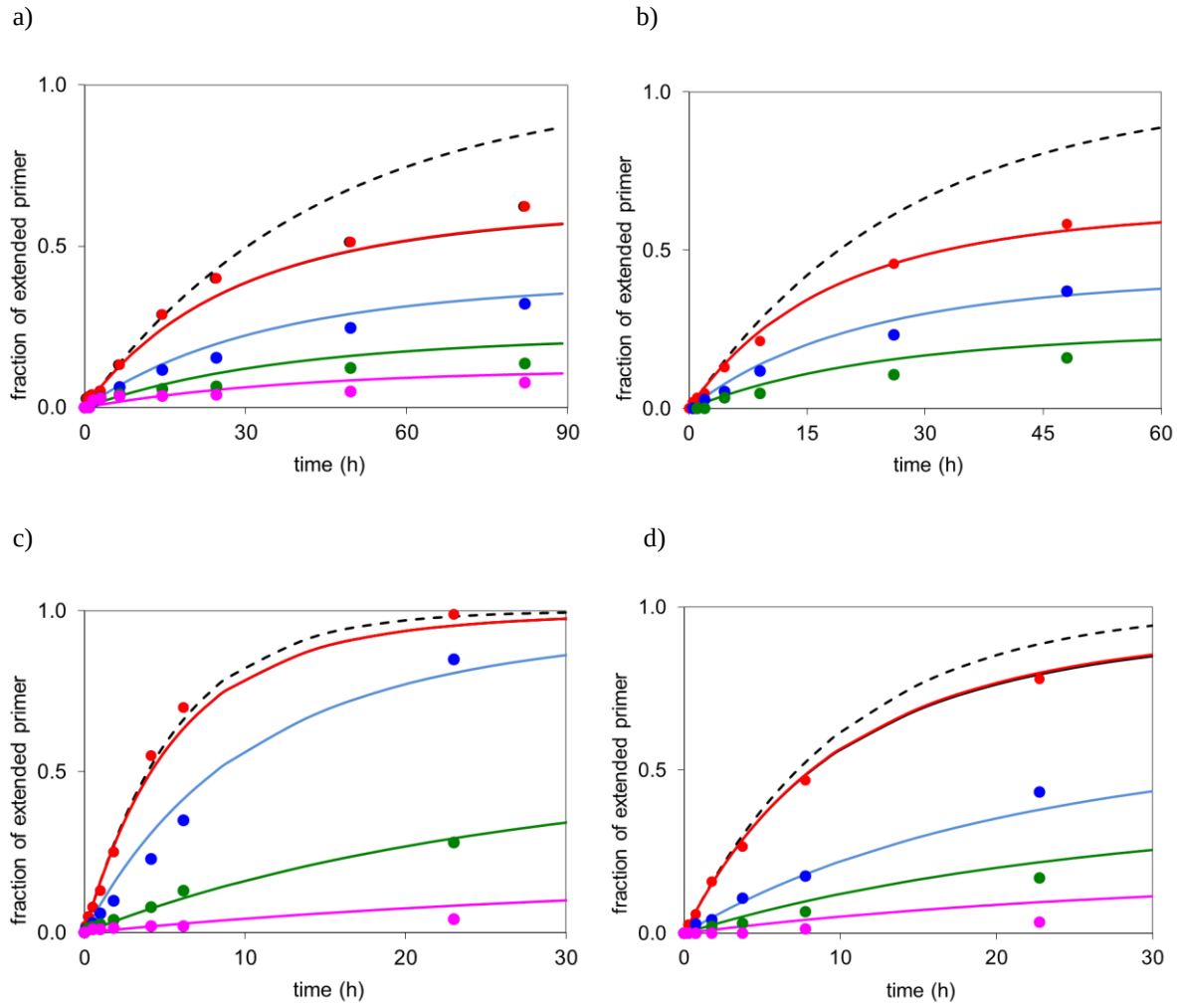


Figure S29. Simulated time-dependent yield of extended DNA primer with and without hydrolysis and inhibition. Initial concentrations of 36 µM primer were used. Filled circles represent experimental values with different concentrations of 2-Methylimidazole activated monomer **3a-t**. Solid lines are values calculated by equation (S2). The value of kcov is obtained by dividing the exponential k-value by the occupation number α given by $[M]_0/(Kd+[M]_0)$. Note that the same value of k_{cov} was used here for the simulation, using initial rate for the highest monomer concentration.

a) Template **9a** and the following concentrations of (MeIm-dTMP) **3t**; red 14.4 mM, blue 7.2 mM, green 3.6 mM and pink 1.8 mM. Values used for the simulation: K_d = 236 mM; K_{dinh} = 240 mM; k_{hydr} = 0.024 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.023 h^{-1}) at 14.4 mM MeIm-dTMP, with k_{cov} =0.35 h^{-1}.

b) Template **9t** and the following concentrations of (MeIm-dAMP) **3a**; red 5 mM, blue 2.52 mM and green 1.26 mM. Values used for the simulation: K_d = 37 mM; K_{dinh} = 38 mM; k_{hydr} = 0.037 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition for (k = 0.037 h^{-1}) at 5 mM MeIm-dAMP, with k_{cov} =0.31 h^{-1}.

c) Template **9c** and the following concentrations of (MeIm-dGMP) **3g**; red 3.6 mM, blue 1.8 mM, green 0.36 mM and pink 0.09 mM. Values used for the simulation: K_d = 25 mM; K_{dinh} = 27 mM; k_{hydr} = 0.025 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.17 h^{-1}) at 3.6 mM MeIm-dGMP, with k_{cov} =1.40 h^{-1}.

d) Template **9g** and the following concentrations of (MeIm-dCMP) **3c**; red 7.2 mM, blue 1.8 mM, green 0.9 mM and pink 0.36 mM. Values used for the simulation: K_d = 26 mM; K_{dinh} = 34 mM; k_{hydr} = 0.03 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.09 h^{-1}) at 3.6 mM MeIm-dCMP, with k_{cov} =0.42 h^{-1}.

RNA primer and OAt ester

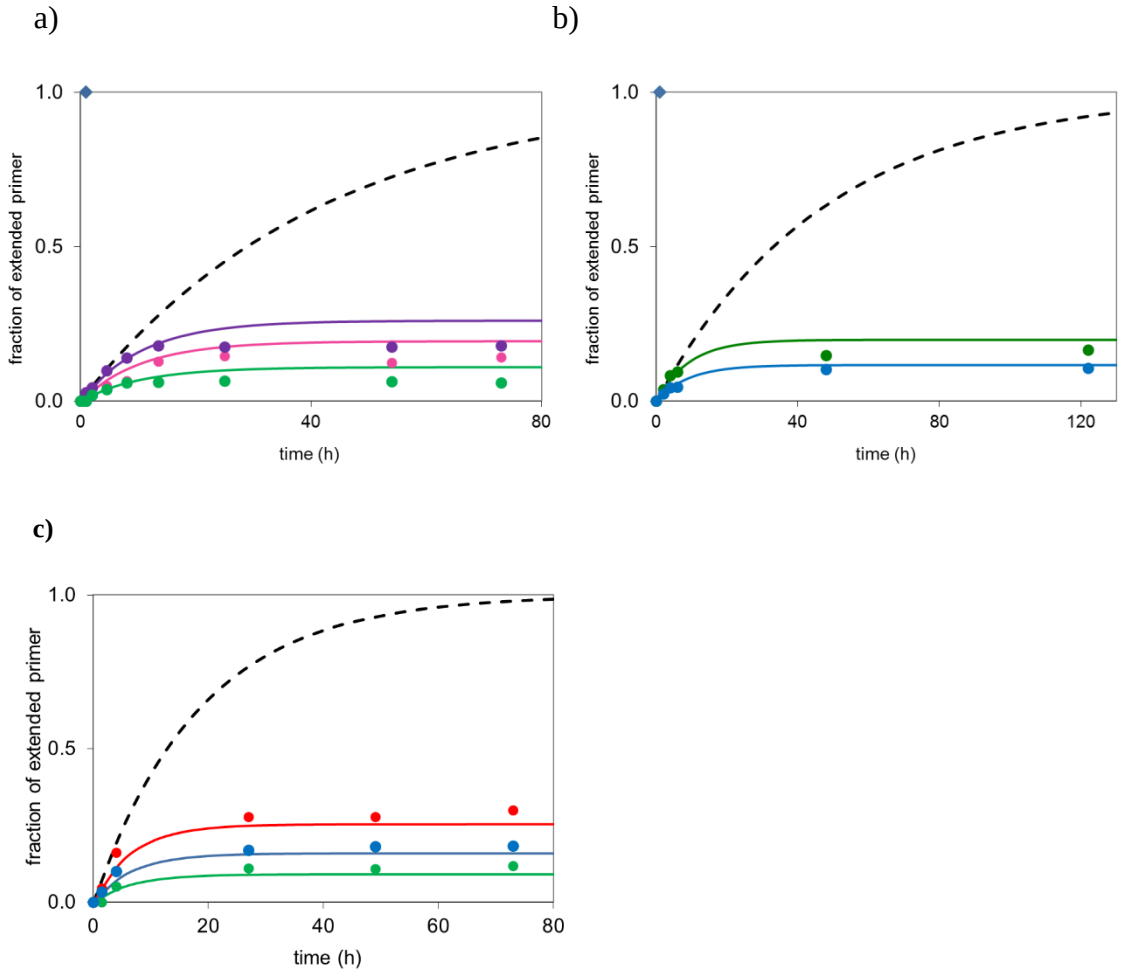


Figure S30. Simulated time-dependent yield of extended RNA primer with and without hydrolysis and inhibition. Initial concentrations of 36 µM primer were used. Filled circles represent experimental values with different concentrations of OAt-activated monomer **5a-u**. Solid lines are values calculated by equation (S2). The value of k_{cov} is obtained by dividing the exponential k-value by the occupation number α given by $[M]_0/(Kd+[M]_0)$. Note that the same value of k_{cov} was used here for the simulation, using initial rate for the highest monomer concentration.
a) Template **12a** and the following concentrations of (OAt-UMP) **5u**; violet 21.6 mM, pink 14.4 mM, green 7.2 mM. Values used for the simulation: K_d = 43 mM; K_{dinh} = 500 mM; k_h = 0.095 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.024 h^{-1}) at 21.6 mM OAt-UMP, with k_{cov} = 0.074 h^{-1}.
b) Template **12u** and the following concentrations of (OAt-AMP) **5a**; green 7.2 mM, blue 3.6 mM. Values used for the simulation: K_d = 14 mM; K_{dinh} = 90 mM; k_h = 0.112 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.021 h^{-1}) at 7.2 mM OAt-AMP, with k_{cov} = 0.062 h^{-1}.
c) Template **12c** and the following concentrations of (OAt-GMP) **5g**; green 7.2 mM, blue 3.6 mM, red 1.8 mM. Values used for the simulation: K_d = 11 mM; K_{dinh} = 35 mM; k_{hydr} = 0.147 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.062 h^{-1}) at 7.2 mM OAt-GMP, with k_{cov} = 0.095 h^{-1}.

RNA primer and methylimidazolides

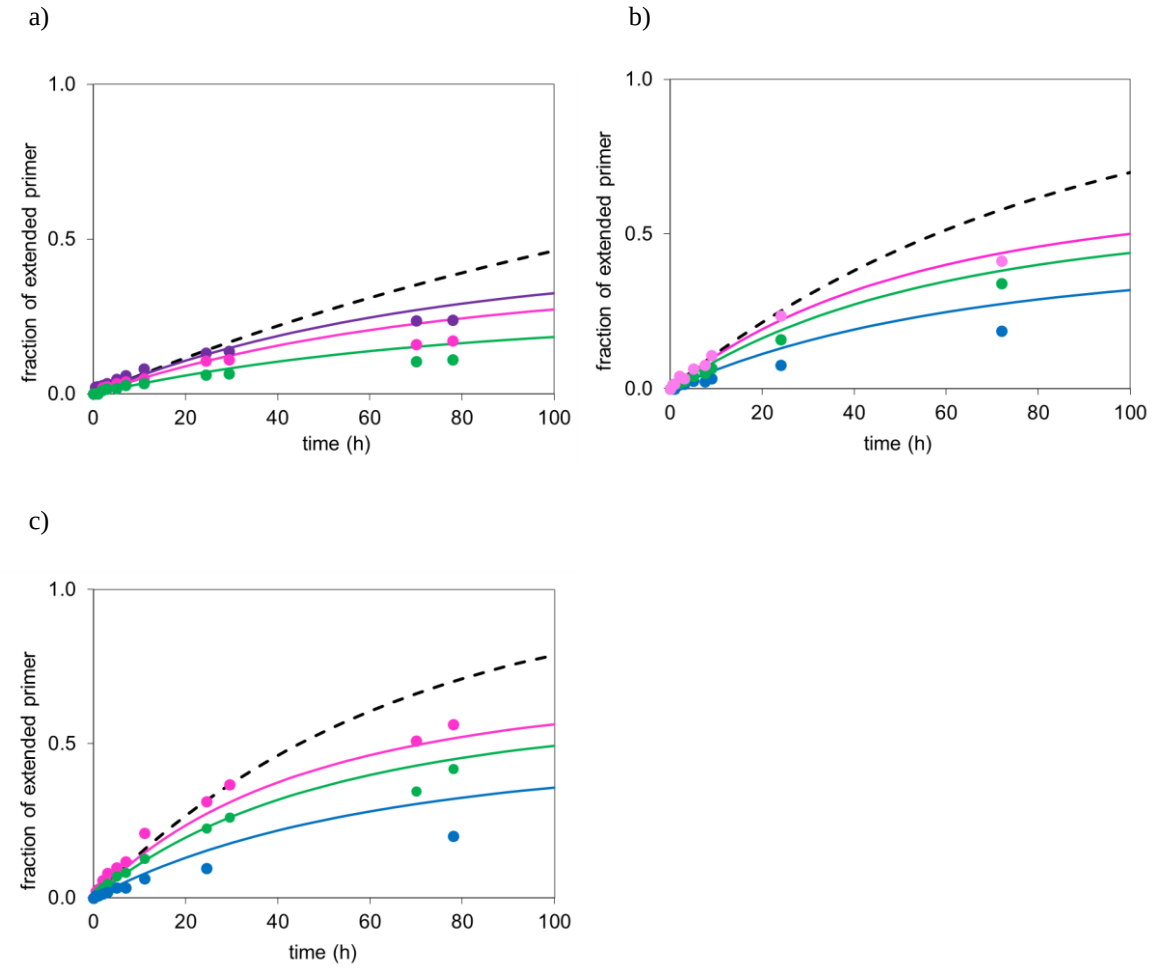


Figure S31. Simulated time-dependent yield of extended RNA primer with and without hydrolysis and inhibition. Initial concentrations of 36 µM primer were used. Filled circles represent experimental values with different concentrations of MeIm-activated monomer **6a-u**. Solid lines are values calculated by equation (S2). The value of k_{cov} is obtained by dividing the exponential k-value by the occupation number α given by $[M]_0/(Kd+[M]_0)$. Note that the same value of k_{cov} was used here for the simulation, using initial rate for the highest monomer concentration.
a) Template **12u** and the following concentrations of (MeIm-AMP) **6a**; green violet 60 mM, pink 40 mM, green 20 mM. Values used for the simulation: K_d = 42 mM; K_{dinh} = 90 mM; k_h = 0.013 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.006 h^{-1}) at 60 mM MeIm-rAMP, with k_{cov} = 0.011 h^{-1}.
b) Template **12c** and concentrations of (MeIm-GMP) **6g** used; pink 36 mM, green 24 mM, blue 12 mM. Values used for the simulation: K_d = 23 mM; K_{dinh} = 27 mM; k_{hydr} = 0.012 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.012 h^{-1}at 36 mM MeIm-rGMP, with k_{cov} = 0.020 h^{-1}.
c) Template **12c** and concentrations of (MeIm-CMP) **6c**; pink 30 mM, green 20 mM, blue 10 mM. Values used for the simulation: K_d = 27 mM; K_{dinh} = 18 mM; k_{hydr} = 0.012 h^{-1}. The dashed black line is for a hypothetical reaction without inhibition (k = 0.016 h^{-1}) at 30 mM MeIm-rCMP, with k_{cov} = 0.029 h^{-1}.

8. Data for Unpurified Methylimidazolide 3g

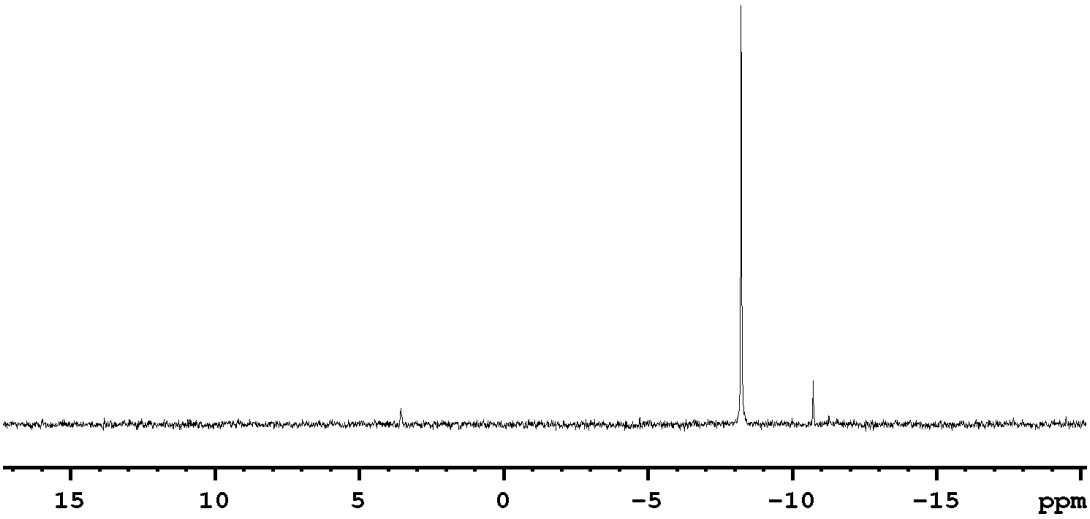
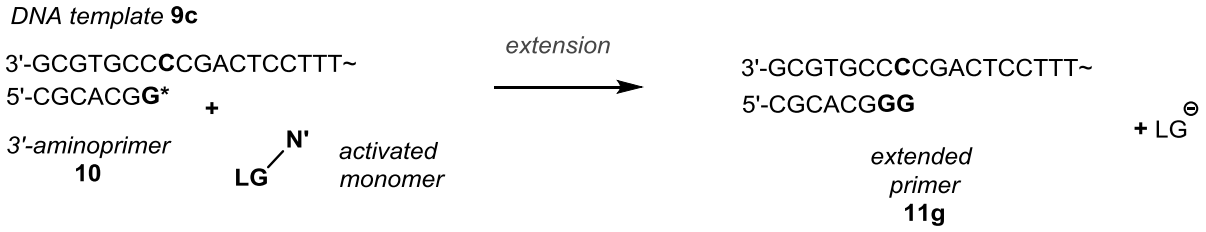
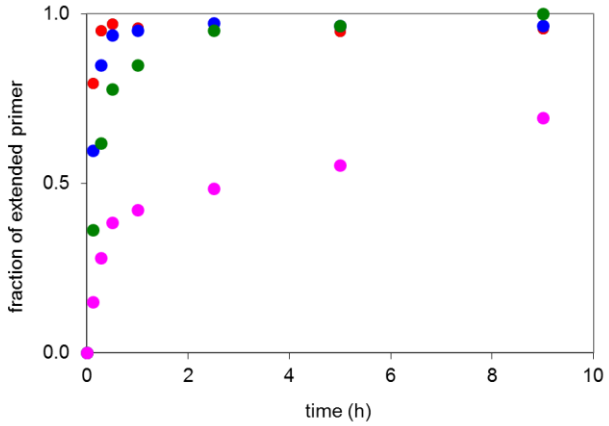


Figure S32. ^{31}P NMR spectrum (121.5 MHz, D_2O) of MeIm-dGMP (**3g**) prior to RP18 cartridge purification. The peak at -8.1 ppm corresponds to MeIm-dGMP and that at +3.5 ppm to the hydrolyzed, free nucleotide. The peak at -10.8 ppm is caused by the impurity, which appears to cause kinetics with a 'burst phase' at the beginning of the extension reaction (*vide infra*). The signal for this impurity gives an integration value of 8% of that of the main product. Assuming that the impurity has the structure shown below, which contains two phosphate groups, this was translated into 4% side product and 96% true methylimidazolide of the reactive species.



a)



b)

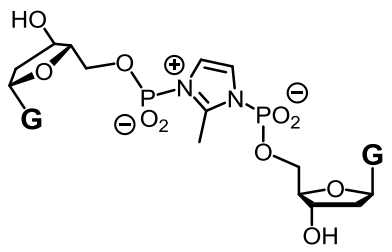


Figure S33. a) Extension of aminoterminal primer **10** (36 µM) on template **9c** when unpurified **3g** was employed. Filled circles represent experimental values at different concentrations of **3g** without chromatographic purification, as obtained by precipitation only. Concentrations of (MeIm-dGMP) **3g** used: red 7.2 mM, blue 3.6 mM, green 1.8 mM and pink 0.9 mM. Note biphasic kinetics.
b) Hypothetical structure of impurity responsible for the burst phase: an imidazoliumbisphosphate.

Monoexponential Fit of the Initial Phase of the Kinetics

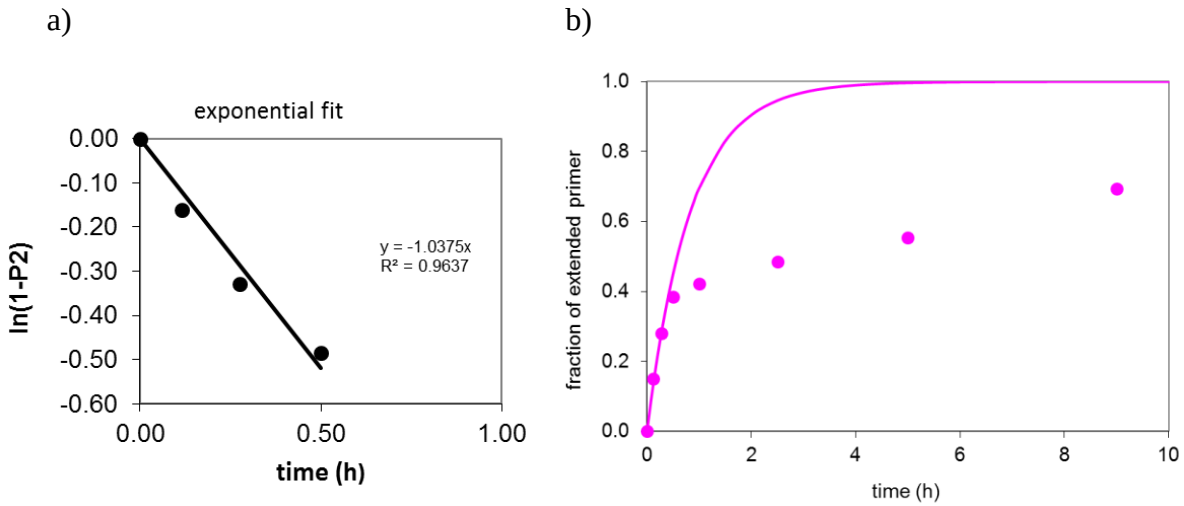


Figure S34. Monoexponential fit to the first four data points of the assay at 0.9 mM initial concentration of unpurified MeIm-dGMP **3g**. a) Fit to the initial phase of the kinetics, b) full set of experimental data points with fit function as continuous line. Assuming that the side product makes up 4% of the reactive species (*vide supra*), a rate constant k of 2.9 $\times 10^4$ $M^{-1}h^{-1}$ is obtained, which makes the initial phase approx. 600 times faster than the reaction of the purified MeIm-dGMP (**3g**). We note that this analysis is based on very few data points and that it ignores the contribution of the reaction of the main reactive species that is also present in the solution.

9. Conformational Search

Calculations were performed using MAESTRO software, version 7.5.106, (Schrödinger Inc.), using the following set-up. For the minimization of the structure without solvent, MMFS was used as force field and the PRGC option for the gradient. Calculations were run without constraints. The fixed convergence threshold was set to 0.02. A fixed energy window of 50 kJ/mol was used, and 1000 steps were applied as limit. To superimpose the structures obatined, an alignment of the carbon atoms 1' and 4' of the ribose rings was used.

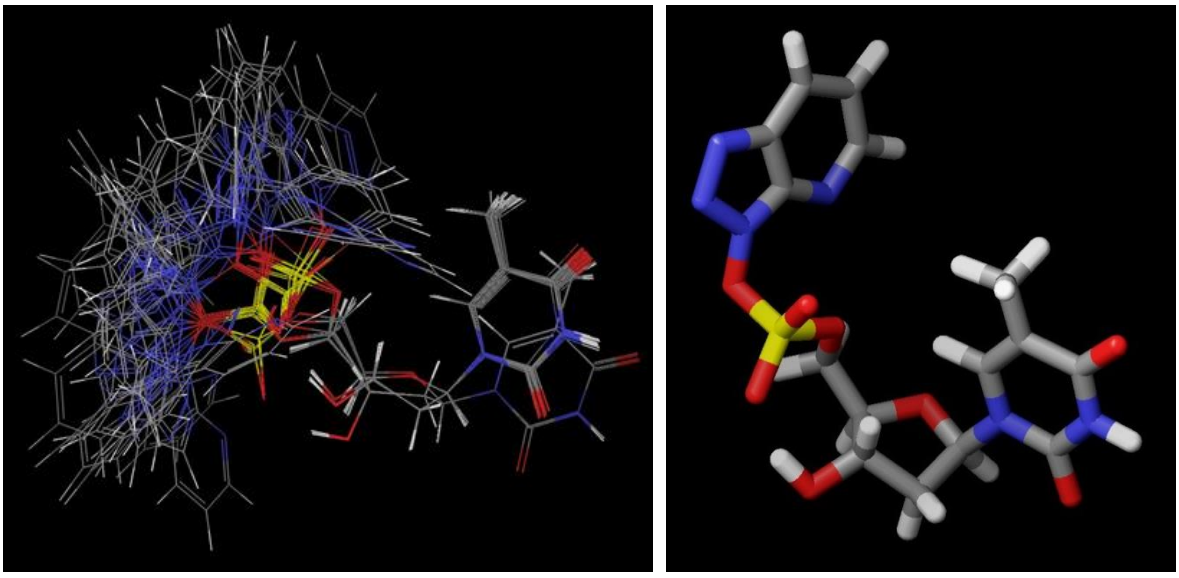


Figure S35. Structure of **2t**. Left: Overlay of representative structures of OAt-TMP **2t**, obtained by an *in silico* conformational search. Note that for a number of conformations showing, the attack of the nucleophile on the phosphorus center is hindered by the methyl group of the nucleobase, as shown in the individual conformation shown on the right-hand side.

10. References for Supporting Information

S1. Hagenbuch, P., Kervio, E., Hochgesand A., Plutowski, U., and Richert C. (2005) Chemical primer extension: efficiently determining single nucleotides in DNA. *Angew. Chem. Int. Ed.*, 44, 6588-6592.

S2. Kervio, E., Claasen, B., Steiner, U. E., and Richert C. (2014) The strength of the template effect attracting nucleotides to naked DNA. *Nucleic Acids Res*., 42, 7409-7420.

S3. Sarracino, D., and Richert, C. (1996) Quantitative MALDI-TOF MS of oligonucleotides and a nuclease assay. *Bioorganic & Medicinal Chemistry Letters*, 6, 2543-2548.

S4. Lohrmann, R., and Orgel, L.E. (1978) Preferential formation of (2'-5')-linked internucleotide bonds in non-enzymatic reactions, *Tetrahedron*, 34, 853-854. Kanavarioti, A., Stronach, M.W., Ketner, R. J., and Hurley, T. B. (1995) Large steric effect in the substitution reaction of amines with phosphoimidazolide-activated nucleosides. *J. Org. Chem*., 60, 632-637.

S5. Kervio, E. Hochgesand, A. Steiner, U. E., and Richert, C. (2010) Templating efficiency of naked DNA. *Proc. Natl. Acad. Sci*.,107, 12074-12079.

S6. Kurz, M., Göbel, K., Hartel, C., Göbel, M. W., (1998) Acridine-labeled primers as tools for the study of nonenzymatic RNA oligomerization. *Helv. Chim. Acta*, 81, 1156-1180 Hartel, C., Göbel M. W., (2000) Substitution of adenine by purine-2,6-diamine improves the nonenzymatic oligomerization of ribonucleotides on templates containing thymidine. *Helv. Chim. Acta*, 83, 2541-2549

S7. Vogel, S. R., Deck, C., Richert, C., (2005) Accelerating chemical replication steps of RNA involving activated ribonucleotides and downstream-binding elements. *Chem. Commun.*, 4922-4924