Dehydration promotes phosphoramidate-linked amino acidyl and peptido adenosine conjugates

Supporting Information

Yaam Deckel, Joshua J. Brown, Tejaswi Senthilkumar and Albert C. Fahrenbach

School of Chemistry, University of New South Wales, Sydney, NSW 2052, Australia

Australian Centre for Astrobiology, University of New South Wales, Sydney, NSW 2052, Australia

Environment Research Unit, CSIRO, Black Mountain, Canberra, ACT, 2601

Advanced Engineering Biology Future Science Platform, Black Mountain, Canberra, ACT 2601

Revolutionary Energy Storage Systems Future Science Platform, Clayton, Melbourne, VIC, 3168

UNSW RNA Institute, University of New South Wales, Sydney, NSW 2052, Australia

Contents

General Methods

Solvents and Reagents

All reagents were used without further purification. The following reagents were purchased from Sigma-Aldrich: adenosine 5'-monophosphate disodium salt (AMP), adenosine 5'-monophosphate monohydrate, 3',5'-cyclic adenosine monophosphate sodium salt monohydrate (3',5'-cAMP), 2',3'-cyclic adenosine monophosphate sodium salt (2',3'-cAMP), glycine, glycine-$^{13}C_2$, Gly-Gly (glyclglycine), Gly-Gly-Gly (triglycine), *N*-methylglycine, ethanolamine, sodium pyrophosphate decahydrate, phosphonoacetic acid, Trizma base, triethylamine, sodium perchlorate, 1-methylimidazole, ammonium carbonate and Dowex 50W X8 cation exchange resin. Sodium phosphate monobasic, *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC), glycine methyl ester hydrochloride (GlyOMe), acetone, diethyl ether anhydrous, 10 M NaOH and 32% HCl were purchased from ChemSupply. *N,N*-dimethylglycine hydrochloride was purchased from Combi-Blocks Inc. Glycinamide hydrochloride was purchased from eMolecules, Inc. LC-MS grade water was purchased from Honeywell. LC-MS grade acetonitrile was purchased from RCL Labscan. Formic acid for LC-MS was purchased from Scharlau. Deuterium oxide was purchased from Cambridge Isotope Laboratories, Inc. All reactions were conducted in LC-MS grade water.

Samples were adjusted to the required pH ± 0.05 with the addition of 3.2% or 32% HCl solutions and 1 M or 10 M NaOH solutions. pH measurements were conducted at approximately 25 °C on a Mettler Toledo InLab Micro Pro-ISM pH probe (cat. # 51344163) equipped with a Mettler Toledo Orion Star A321 pH meter calibrated with pH 4, 7 and 10 buffers obtained from Australian Chemical Reagents.

Analysis Techniques

Nuclear Magnetic Resonance Spectroscopy

NMR spectroscopy was performed at the Mark Wainwright Analytical Centre NMR facility, located at the University of New South Wales Randwick campus. A Bruker Avance III 400 MHz spectrometer equipped with a Bruker Prodigy Cryo-Probe was used for all experiments. ^{1}H and $^{13}C\{^{1}H\}$ NMR spectroscopy was performed in 0.5 M phosphate buffer at pH 7 with 10% D_2O. $^{31}P\{^{1}H\}$ NMR spectroscopy was performed in 0.5 M Tris.HCl buffer at pH 7 with 10% D_2O. ^{1}H NMR spectra

were acquired using water suppression by excitation sculpting with 32 scans. $^{13}C\{^{1}H\}$ NMR spectra were acquired with either 1200 scans for samples without labelled glycine and 256 scans for samples with labelled glycine. $^{31}P\{^{1}H\}$ spectra were acquired with 64 scans and a D1 of 30 seconds for adequate relaxation for quantitative measurements and with a higher number of scans (128 to 256) for qualitative/semi-quantitative measurements. Data was processed with MestreNova v.15.0.0-34764.

Mass Spectrometry

QTOF-ESI mass spectrometry was performed using a Shimadzu LCMS-9030 quadrupole-time-of-flight mass spectrometer connected to a Shimadzu Nexera 40 Series UHPLC system. In most cases, samples were diluted 1 in 100 after the rehydration step (see below for DH/RH cycling experimental procedures) in LC-MS water and directly injected into the spectrometer without prior LC. Samples were run for 1.5 minutes at 0.2 mL min^{-1} with an injection volume of 5 µL using 100% water with 0.1% formic acid. Mass spectra were collected in positive mode using an interface voltage of 4.0 kV. The interface and desolvation temperatures were set to 300 °C and 526 °C, respectively, the nebulisng gas flow rate was 2 L min^{-1}, and the heating gas flow rate was 10 L min^{-1}. Data was processed with Shimadzu LabSolutions Postrun Analysis 5.110 and plotted with Excel.

General Method for DH/RH Synthesis

3',5'-cAMP sodium salt monohydrate (18 mM) and glycine (50 mM) were dissolved in LC-MS water and adjusted to pH 10 ± 0.05 with minimal volumes of 1 M and/or 5 M NaOH. 1 mL aliquots of this reaction in 1.5 mL Eppendorf tubes were then heated uncapped for 18 hours at 80 °C. For experiments with more than one DH/RH cycle, the reaction was then rehydrated with 1 mL of LC-MS grade water and left to heat capped at 60 °C for 6 hours before being uncapped to dry for another 18 hours at 80 °C. This series of steps was then repeated for up to ten cycles. For analysis by ^{1}H or ^{13}C NMR spectroscopy, the samples were rehydrated with 1 mL of 0.5 M phosphate buffer at pH 7 with 30 mM MeCN as an internal standard and 10% D_2O. For analysis by $^{31}P\{^{1}H\}$ NMR, samples were rehydrated with 1 mL of 0.5 M Tris.HCl buffer at pH 7 with 20 mM phosphonoacetic acid as an internal standard and 10% D_2O. If for analysis by MS, samples were rehydrated with LC-MS grade water. All samples for analyses were heated briefly at 60 °C after rehydration until redissolved.

Other DH/RH Reaction Methods

Alternate DH/RH reactions such as those at other pH conditions (pH 2, 2.5, 3, 4, 6 and 8), control reactions (e.g., 3',5'-cAMP alone and glycine alone) and those with alternate starting materials (2',3'-cAMP, 5'-AMP, glycine-$^{13}C_2$, glycylglycine, triglycine, ethanolamine, glycinamide, *N*-methylglycine and *N,N*-dimethylglycine*)* were prepared and reacted as described above. pH was adjusted to ± 0.05 of the desired value with the addition of 1 M and/or 5 M NaOH and 32% and/or 3.2% HCl. Reactions which were carried out without dehydration were performed in capped Eppendorf tubes using the same heating profile as described above. Monitoring the reaction during DH/RH involved sampling 10 µL aliquots from the reaction mixture at the desired time points while heating uncapped. Reactions with 5'-AMP which required a change of pH in between cycles were initially rehydrated, heated capped at 60 °C for 6 hours, and then adjusted to pH 10 ± 0.05 prior to the second cycle.

For reactions performed under vacuum, cNMPs (2',3'- and 3',5'-cAMP) (20 mM) and glycine (100 mM) were dissolved in LC-MS grade water and the solution was adjusted to pH 10 ± 0.05 with the addition of 1 M and 5 M NaOH. 100 µL aliquots of this were then dried in 4 mL glass vials as described by Verlander *et al*.[1] at 5 mbar in a desiccator over phosphorous pentoxide at room temperature for 3 days.

Synthesis and Characterisation of NMR Standards

Gly-*N*-pA

The following procedure was based on that reported by Räuchle *et al*.[2] 5'-AMP (as a monohydrate, 200 mg, 0.55 mol), EDC (525 mg, 2.74 mmol) and 2-methylimidazole (2-MeIm) (450 mg, 5.47 mmol) were dissolved in deionised water (4 mL), and the mixture was adjusted to pH 6 by addition of minimal volumes of 10 M NaOH. One equivalent of EDC (105 mg, 0.55 mmol) was added into the solution at 0.5, 1, and 1.5 hours. After 2 hours, 95% of the AMP was converted to the activated 2-MeIm-pA imidazolide as determined by $^{31}P\{^1H\}$ NMR spectroscopy. The crude 2-MeIm-pA was added to an aqueous solution of excess glycine methyl ester (GlyOMe) (1.38 g, 11.0 mmol, 4 mL). After 24 h, 95% of the 2-MeIm-pA was converted to GlyOMe-*N*-pA and Gly-*N*-pA.

Crude Gly-*N*-pA was purified on a CombiFlash EZPrep system equipped with a Biotage Sfär C18 D column (400g resin,100 Å pore width) using a binary gradient of 25 mM NH_4HCO_3 in water (pH

8, A buffer) with a gradient of 0% to 90% MeCN (B buffer) over 20 minutes. Elution of the products was monitored by UV-Vis absorption at 200 and 260 nm. The GlyOMe-*N*-pA and Gly-*N*-pA fractions were combined and then fully deprotected with NaOH at ~pH 14 for 18 hours. Then, Gly-*N*-pA was repurified using the same column using 25 mM triethylammonium bicarbonate in water (pH 7.5, A buffer) with a gradient of 0% to 90% MeCN (B buffer) over 20 mins to remove remaining (spent) EDC and excess NaOH. Gly-*N*-pA was then exchanged to its sodium salt using a DOWEX-Na^+ resin at pH 6–7, following lyophilisation. Excess sodium salts were removed by adding Gly-*N*-pA dissolved in water (100 uL) dropwise to a solution of 0.2 M sodium perchlorate in acetone/diethyl ether (1:1 v/v, 40 mL). The resulting precipitate was centrifuged, washed twice with acetone, redissolved in water, and lyophilised. The final product was characterised by NMR spectroscopy and mass spectrometry. ^{1}H NMR (400 MHz, 90% H_2O + 10% D_2O, 298 K), δ = 8.36 (s, 1H), 8.07 (s, 1H), 6.65 (s, 1H), 5.99 (d, J = 5.8 Hz, 1H), 3.96 (dd, J = 4.8, 3.1 Hz, 1H), 3.30 (d, J = 8.4 Hz, 2H);^{13}C NMR (101 MHz, D_2O, 298 K), δ = 179.1 (d, J = 9.0 Hz), 155.6, 152.8, 149.0, 139.8, 118.6, 86.9, 84.1 (d, J = 9.2 Hz), 74.3, 70.5, 63.9 (d, J = 4.8 Hz), 45.0; ^{31}P NMR (162 MHz, 90% H_2O + 10% D_2O, 298 K), δ = 8.1. ESI-MS (positive mode): $m/z_{obs.}$ = 405.0906 for $[M + H]^+$ (m/z_{calc} = 405.0910).

Gly_2-*N*-pA

Crude 2-MeIm-pA was synthesised as described above and added to a solution of excess glyclglycine (Gly_2) (1.585 g, 12 mmol, 7 mL). After 48 hours, the 2-MeIm-A was fully converted to Gly_2-*N*-pA as determined by $^{31}P\{^1H\}$ NMR spectroscopy.

Crude Gly_2-*N*-pA was purified as described above. No deprotection step was necessary. After exchange to its sodium salt using DOWEX-Na^+ at pH 6–7 and lyophilisation, the product was characterised by NMR spectroscopy and mass spectrometry. Excess salts were not removed prior to collecting characterisation data and its use as a standard. ^{1}H NMR (500 MHz, 90% H_2O + 10% D_2O, 298 K), δ = 8.30 (s, 1H), 8.09 (s, 1H), 6.65 (s, 1H), 5.96 (d, J = 5.5 Hz, 1H), 4.24 – 4.18 (m, 0H), 3.90 (dddd, J = 18.0, 16.3, 11.4, 7.4 Hz, 1H), 3.55 (d, J = 5.8 Hz, 1H), 3.34 (dd, J = 10.6, 7.2 Hz, 2H); ^{13}C NMR (101 MHz, 90% H_2O + 10% D_2O, 298 K), δ = 176.4, 174.2 (d, J = 6.9 Hz), 155.4, 152.7, 148.8, 139.6, 118.5, 87.1, 83.8 (d, J = 9.0 Hz), 74.2, 70.4, 63.9, 44.5 (d, J = 8.6 Hz), 43.1; ^{31}P NMR (162 MHz, 90% H_2O + 10% D_2O, 298 K), δ = 7.7. ESI-MS (positive mode): $m/z_{obs.}$ = 462.1116 for $[M + H]^+$ (m/z_{calc} = 462.1133).

Computational Methods

All density functional theory calculations were performed using the ORCA 5.0.3 software.[3] All structures were optimized with the M06-2X DFT functional[4] and the 6-31+G(d,p) basis set. We have opted to use diffuse functions to model the presence of negatively charged hydroxide ions and alkoxide species in the reaction mechanism.[5] The solvation model based on density (SMD)[6] was also applied during geometry optimization in order to better capture condensed phase phosphoester cleavage pathways. Optimizing the geometries in the gas phase was carried out, but it was found that the potential energy surface was too sensitive to adequately capture H^+ transfer states in this case. Transition states were captured and verified by the presence of a single imaginary frequency, and subsequent vibrational frequency calculations checks via intrinsic reaction coordinate (IRC) calculations were performed. Free energies were calculated at standard conditions (298.15 and 353.15 K) including thermal corrections obtained from vibrational frequency calculations at the same level of theory. Note, we have opted not to include a +D3(0) dispersion correction for M06-2X as recommended in the work of Goerigk et al.[7], because we are optimizing the geometry with SMD implicit solvent, and the original parameterization of SMD and Minnesota functionals both contain corresponding parameterization to account for dispersion effects on geometry. The +D3(0) term is recommended just as an addition to M06-2X by itself. Free energies of the reaction coordinate states along the reaction mechanisms are calculated relative to the reactants 3′,5′-cAMP and Gly at 298.15 K. Due to the simplicity of the way we are approximating the dehydrated solid-state (with a phosphate-adjacent Na^+ ion), we have prioritized a more expansive search of the potential energy surface instead of including additional corrections at a higher level of theory to the electronic energy.

Characterisation of the 3',5'-cAMP + Glycine Reaction by NMR Spectroscopy

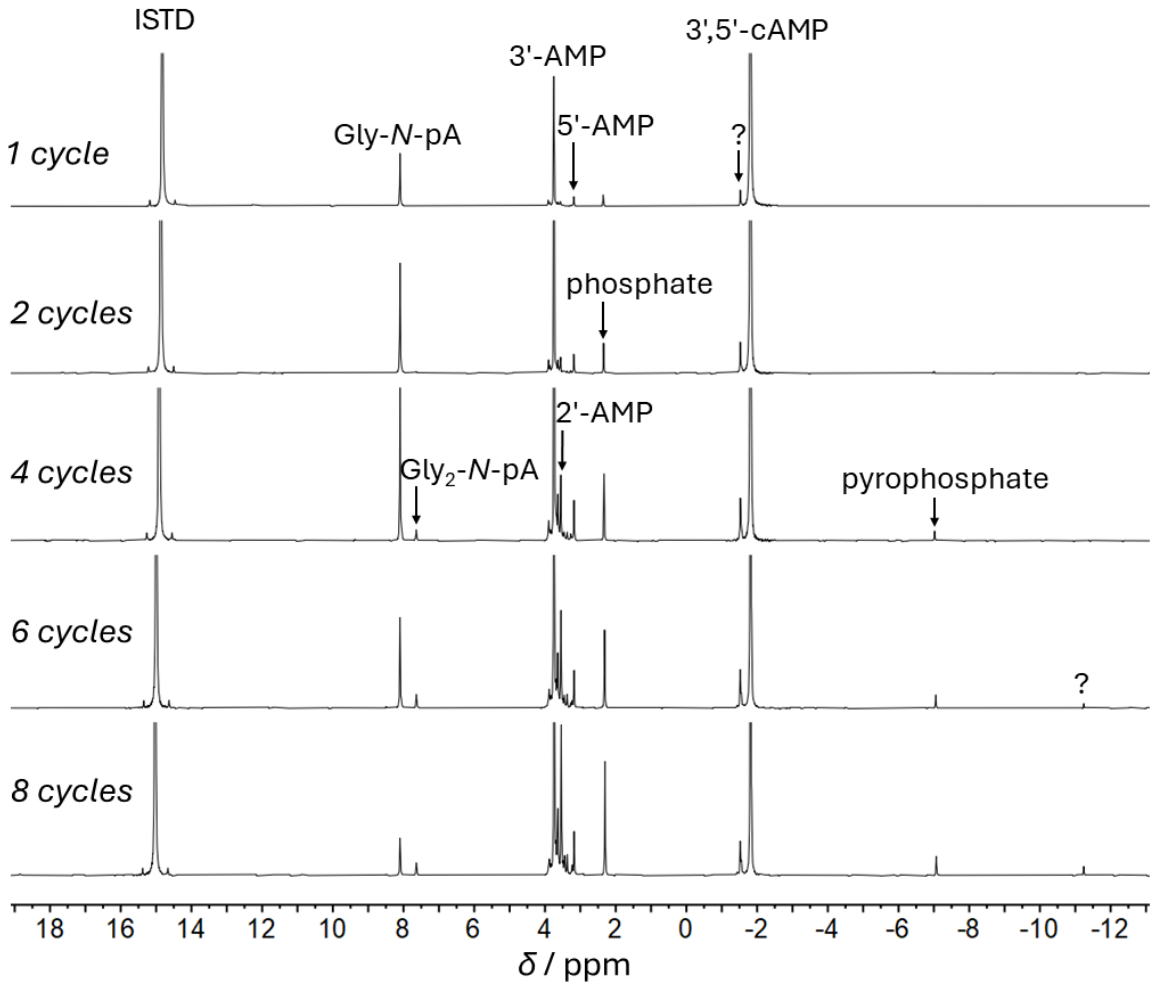


Figure S1: $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 80 °C recorded between 1 and 8 cycles. Assignment of resonances was confirmed by addition of standards (see Figs. S7 and S8). The resonances at −1.5 may represent phosphodiester linked glycine conjugated AMP dimers (see Fig. S15 for the corresponding mass spectrum recorded on this solution and Fig. S2 for more detail on this peak).

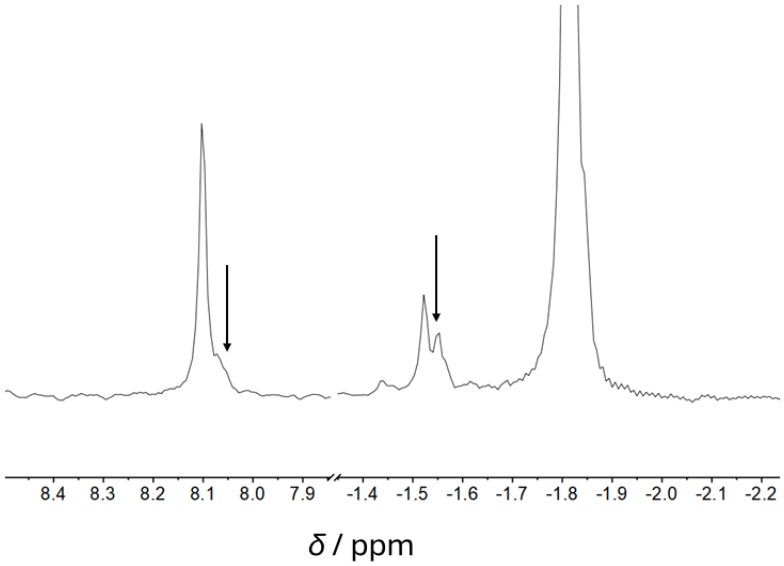


Figure S2: Partial $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 80 °C recorded at 8 cycles from Figure S1 zoomed in on the phosphoramidate and phosphodiester regions. The resonance which appears at δ = –1.55 ppm (shown by an arrow) is hypothesised to arise from the phosphodiester bond in Gly-*N*-pApA. (See Fig. S15 for MS data which supports this assignment). The shoulder at δ = 8.07 ppm may arise from the phosphoramidate bond in this molecule. Significant overlap of this resonance with Gly-*N*-pA makes quantitation difficult.

$^{31}P\{^{1}H\}$ NMR Control Reactions

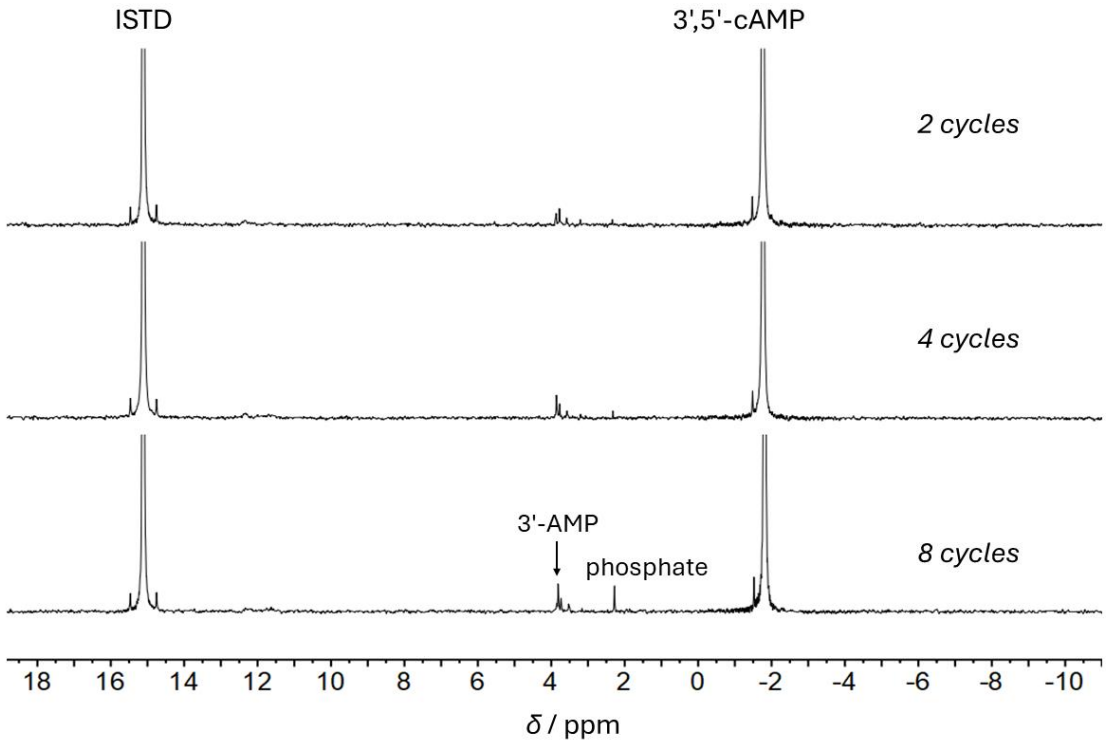


Figure S3: $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 3',5'-cAMP without glycine at pH 10 and 80 °C recorded between 2 and 8 cycles. Small amounts of phosphoester cleavage can be observed with resonances assigned as 3'-AMP and phosphate.

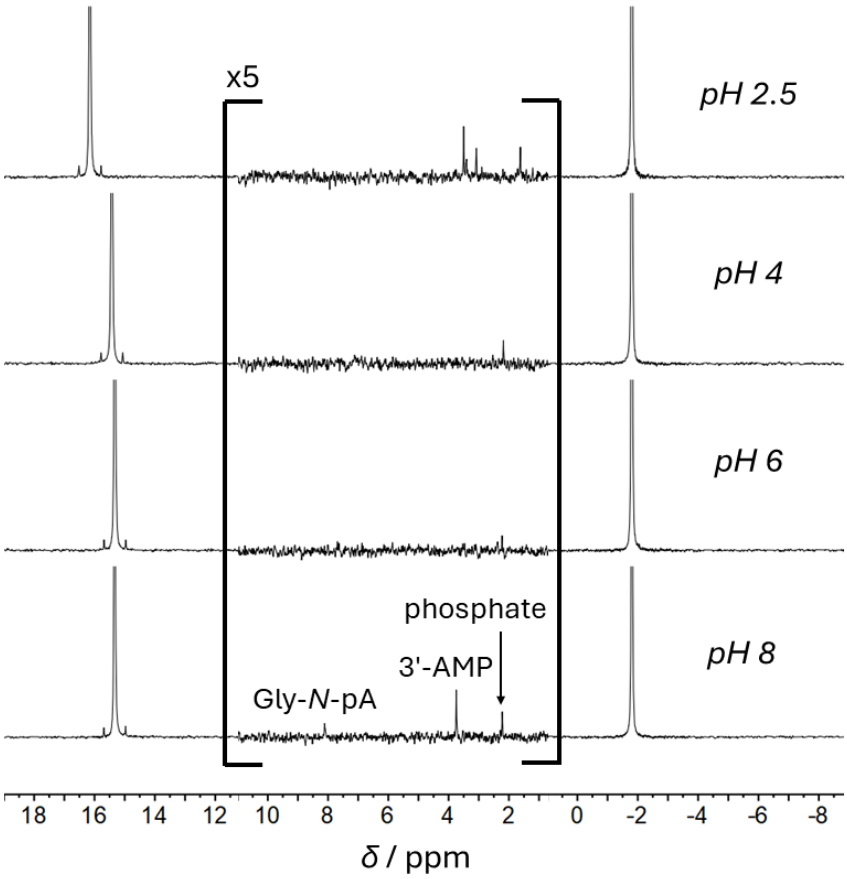


Figure S4: $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 3',5'-cAMP and glycine at 80 °C for 1 cycle at pH 2.5, 4, 6 and 8. Small resonances corresponding to Gly-*N*-pA, 3'-AMP and phosphate are observable at pH 8. The reactions at pH 4 and 6 show a resonance assigned to inorganic phosphate resulting from dephosphorylation. The reaction at pH 2.5 shows additional resonances which have not been assigned.

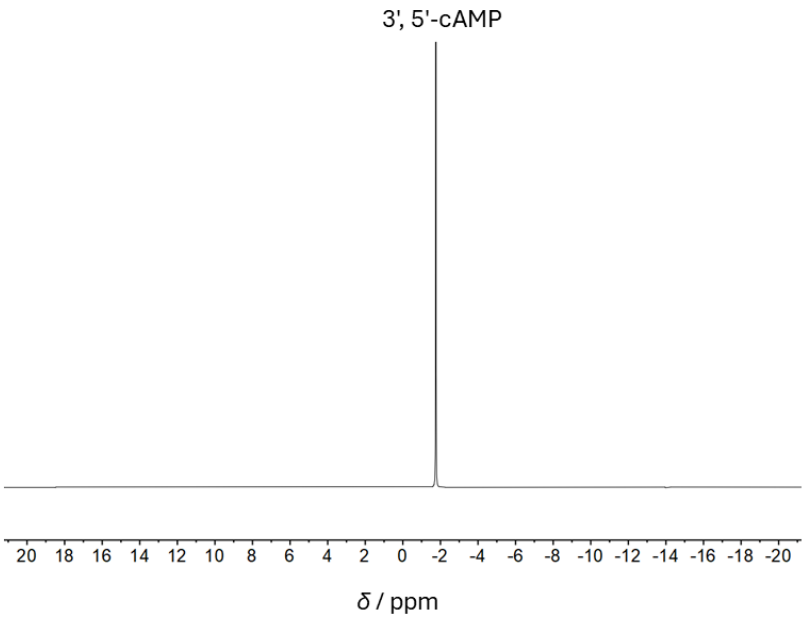


Figure S5: $^{31}P\{^{1}H\}$ NMR spectrum of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 80 °C after 6 hours of heating uncapped. 10 uL of sample was collected before full dehydration and diluted 60 times into a solution of 90% H_2O and 10% D_2O; the spectrum above was then recorded. No new resonances are observed.

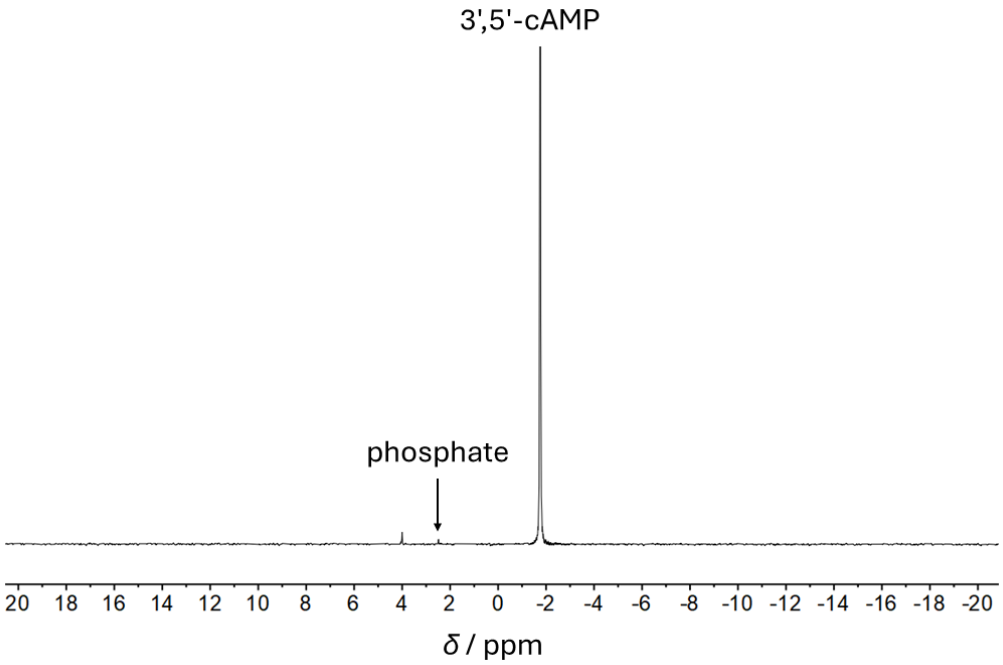


Figure S6: $^{31}P\{^{1}H\}$ NMR spectrum recorded on a reaction mixture of 3',5'-cAMP and glycine at pH 10 and 80 °C heated for18 hours in a capped vial (i.e., without dehydration). 10% D_2O was added to the solution before collecting the NMR spectrum shown above. The resonance at δ = 4.00 ppm was initially thought to be either 3'- or 5'-AMP, however, addition of standards revealed that this was not the case.

$^{31}P\{^{1}H\}$ NMR Standard Additions

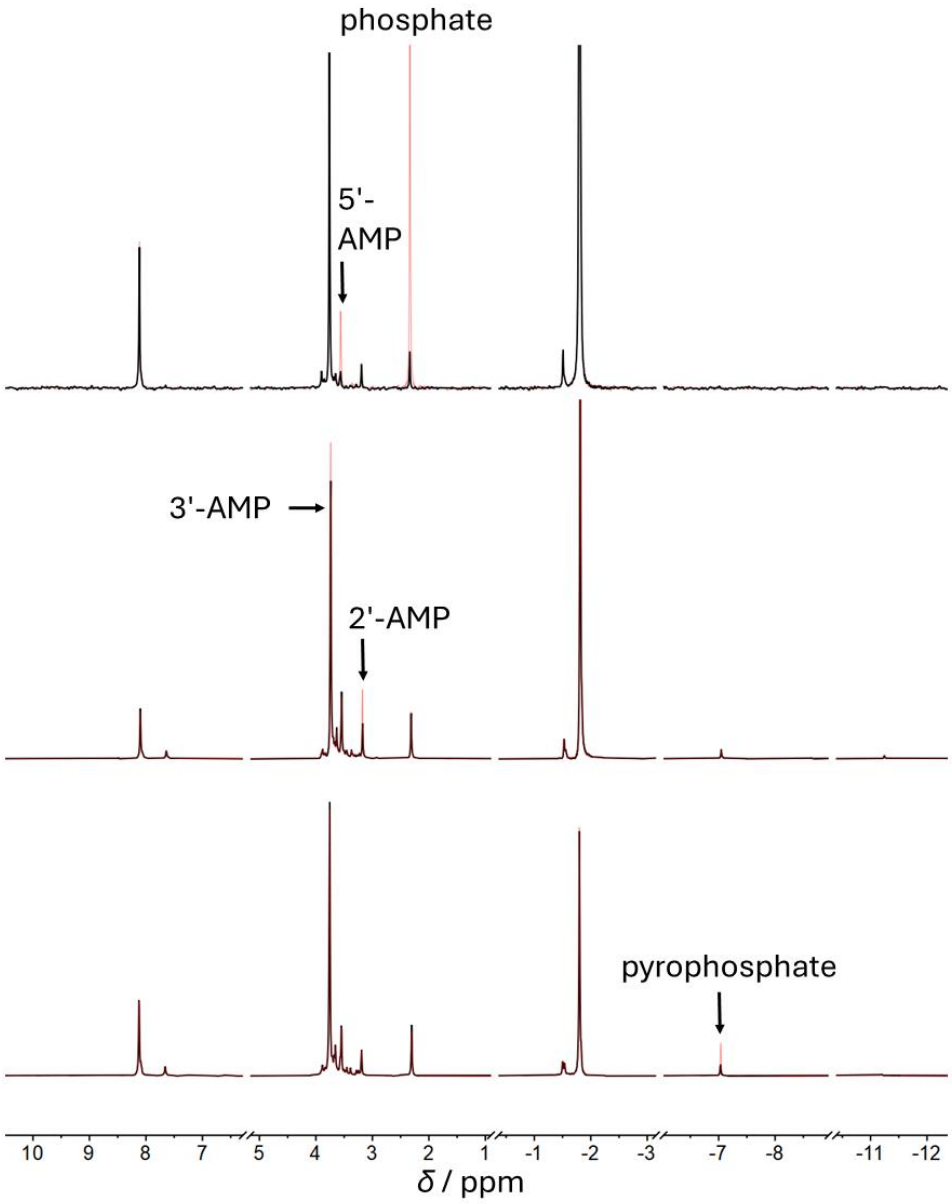


Figure S7: In black (overlaid): $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 80 °C. In red (underlaid): spectrum of the same solution after spiking with standards of 5'-AMP and inorganic phosphate (top), 2'- and 3'-AMP produced from the DH/RH reaction of 2',3'-cAMP and glycine (see Figs. S36–39) (middle) and pyrophosphate (bottom).

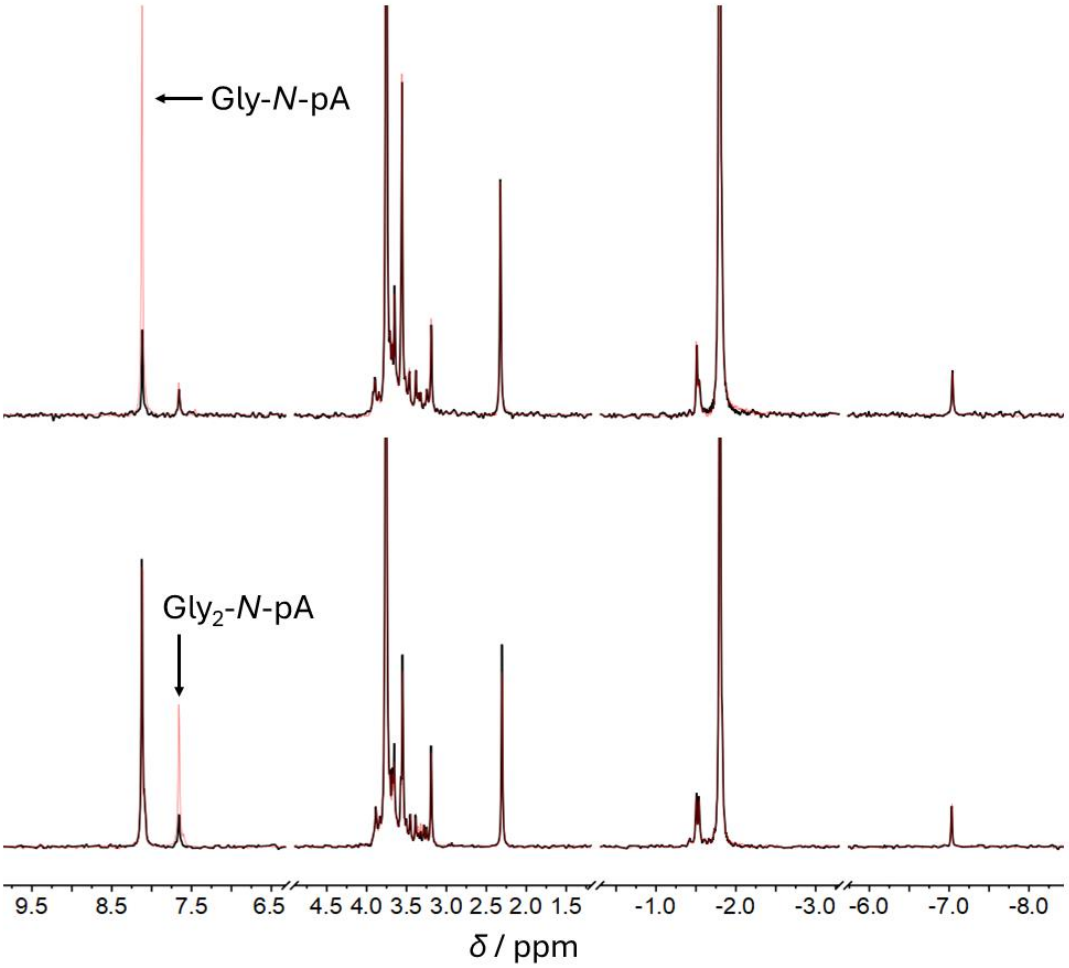


Figure S8: In black (overlaid): $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 80 °C. In red (underlaid): spectrum of the same solution after spiking with standards of Gly-*N*-pA (top) and Gly_2-*N*-pA (bottom). See Figs. S17–25 for characterisation data for these two standards.

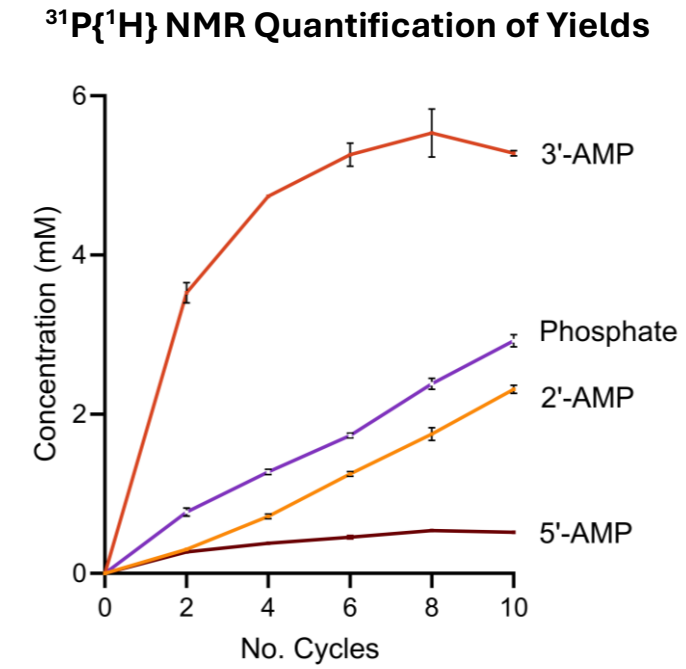


Figure S9: Concentrations of 2'-AMP (orange), 3'-AMP (red), 5'-AMP (dark red) and phosphate (purple) over 0–8 DH/RH cycles of the reaction of 18 mM 3',5'-cAMP and 50 mM glycine at pH 10 and 80 °C. Concentrations were calculated by $^{31}P\{^{1}H\}$ NMR spectroscopy through comparison to a 20 mM phosphonoacetic internal standard and averaged over triplicates. Error bars represent standard deviation.

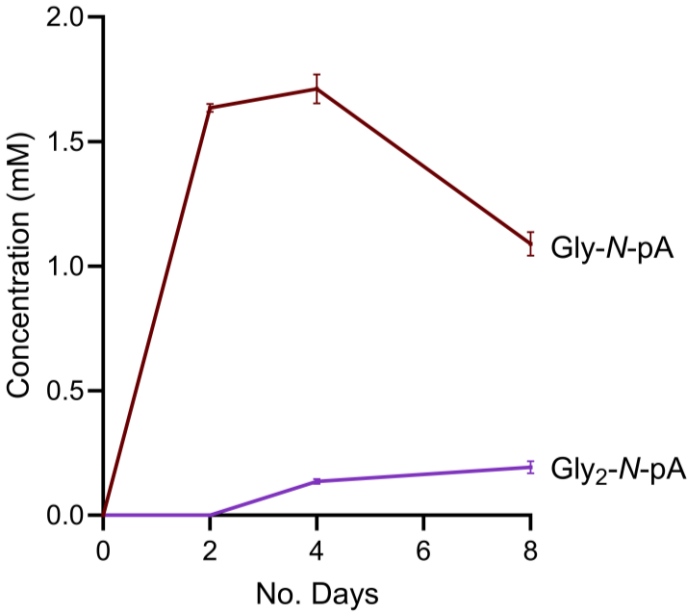


Figure S10: Concentrations of Gly-*N*-pA (dark red) and Gly_2-*N*-pA (purple) over 0–8 days after incubation of 18 mM 3',5'-cAMP and 50 mM glycine in the dry state at 80 °C, i.e., without daily rehydration. The graph shows no significant difference from the reaction with daily rehydration. Concentrations were calculated by $^{31}P\{^{1}H\}$ NMR spectroscopy through comparison to a 20 mM phosphonoacetic internal standard and averaged over triplicates. Error bars represent standard deviation.

Other NMR Experiments

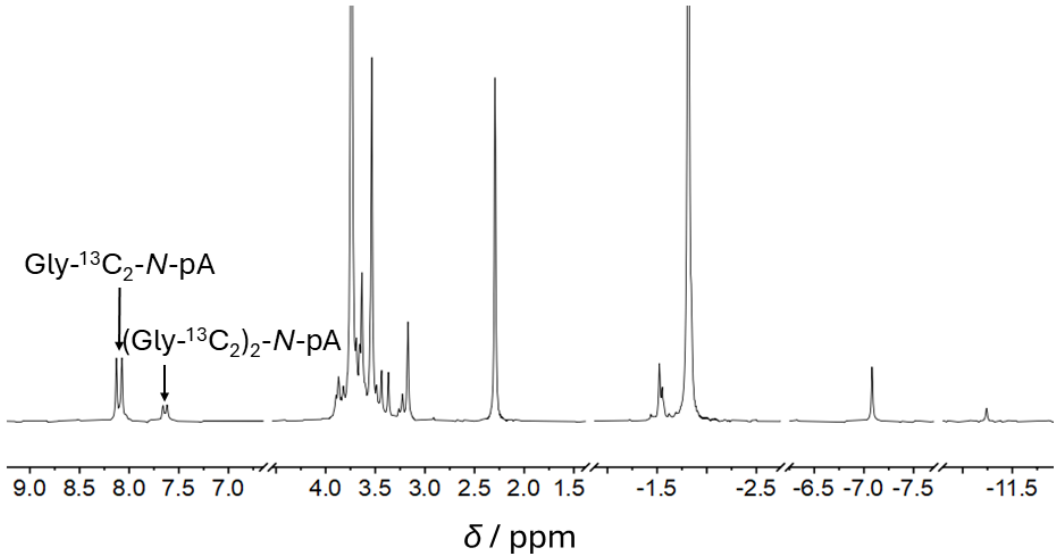


Figure S11: Partial $^{31}P\{^1H\}$ NMR spectrum of the DH/RH reaction of 3',5'-cAMP and glycine-$^{13}C_2$ at pH 10 and 80 °C for eight cycles. ^{13}C-^{31}P splitting is observed in the resonances at δ = 8.14 (J = 9.0 Hz) and 7.67 (J = 6.1 Hz) confirming their identity as phosphorous-linked glycine adducts. No other peaks show similar changes in splitting.

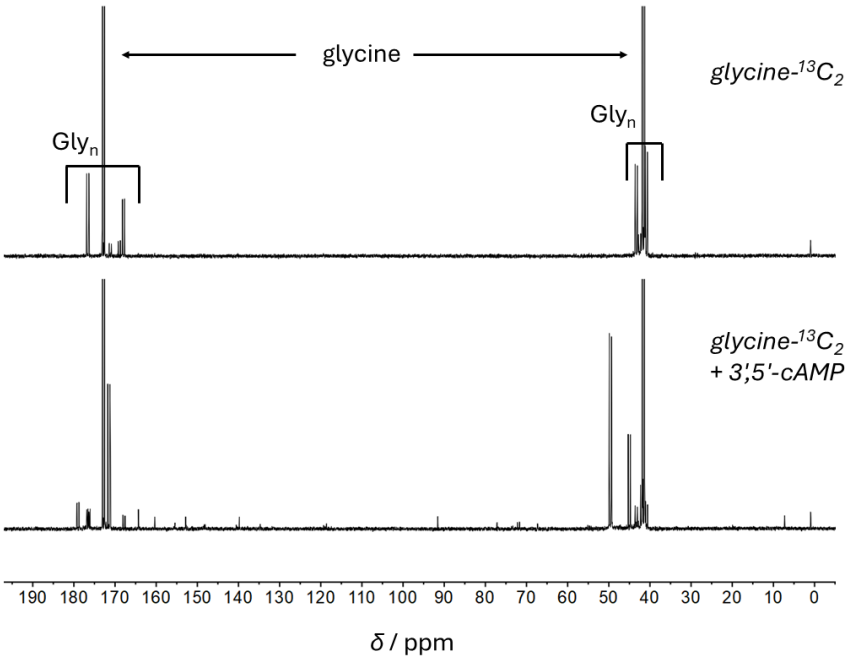


Figure S12: ^{13}C NMR spectra of the DH/RH reaction of glycine-$^{13}C_2$ alone (top) and 3',5'-cAMP and glycine-$^{13}C_2$ (bottom) at pH 10 and 80 °C for four cycles. Glycine-$^{13}C_2$ was used to increase the signal intensity of glycine-derived products. ^{13}C-^{13}C splitting is observed for all glycine adduct peaks. The resonances corresponding to glycine oligomers (labelled as Gly_n) appear with larger intensities for the reaction of glycine alone in comparison to the reaction of 3',5'-cAMP and glycine.

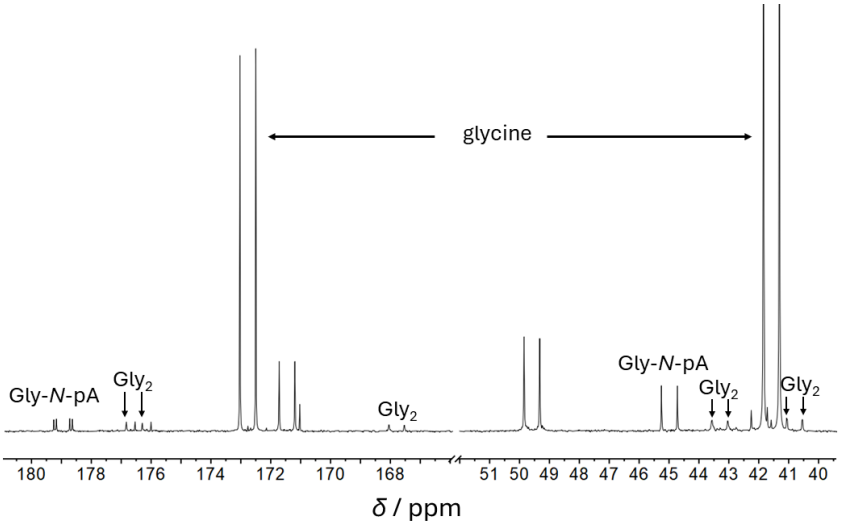


Figure S13: Partial ^{13}C NMR spectrum of the DH/RH reaction of 3',5'-cAMP and glycine-$^{13}C_2$ at pH 10 and 80 °C for four cycles shown in Fig. S12, zoomed in on the aliphatic and carbonyl regions. Glycine oligomer resonances (Gly_2) were assigned by comparison to the reaction of glycine alone (see Fig. S12). ^{13}C-^{31}P coupling is observed in the carbonyl (c) carbon but not in the alpha (α) carbon due to broadening experienced by its proximity to the amine.[8] The doublets at δ = 49.6 ppm and 171.5 ppm are unassigned, however, are unlikely to belong to a phosphate-linked product given that no ^{13}C-^{31}P coupling is observed. It is possible that these peaks belong to the adenosine-glycine conjugate observed by mass spectrometry (see Fig. S15).

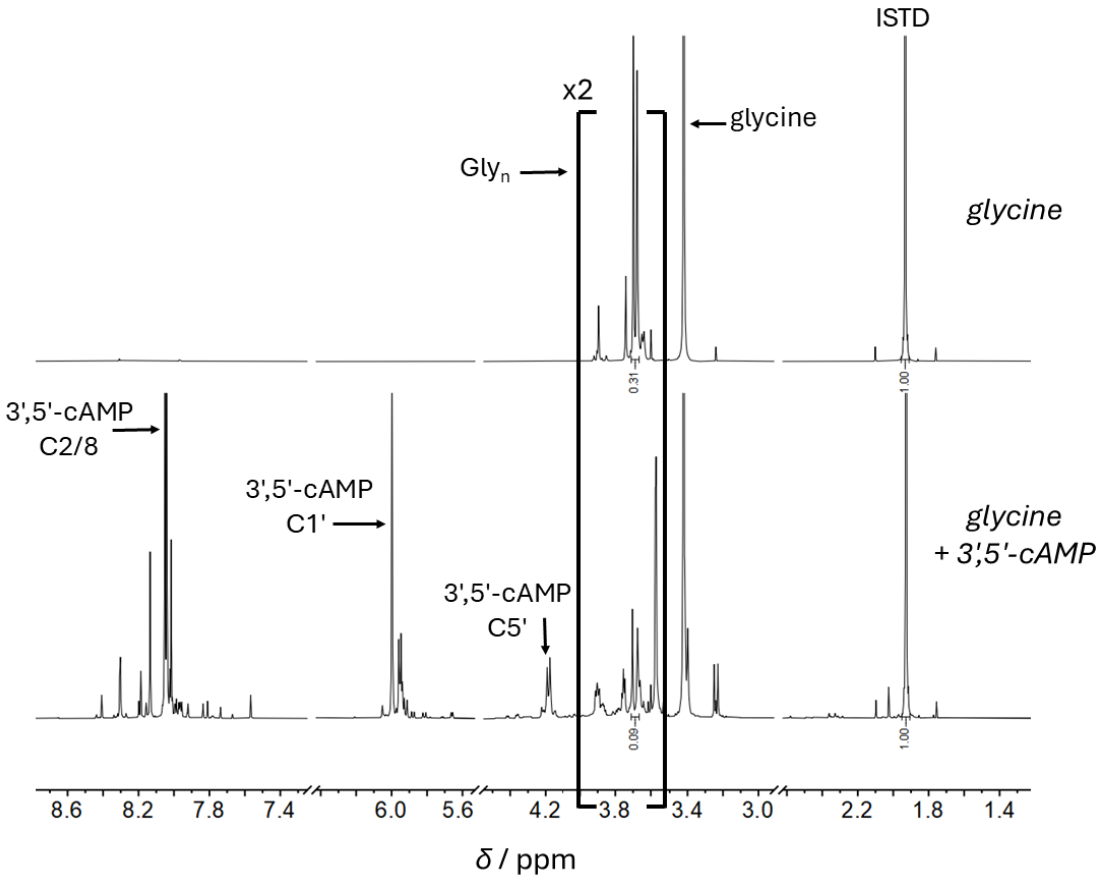


Figure S14: ^{1}H NMR spectra of the DH/RH reaction of glycine alone (top) and 3',5'-cAMP and glycine (bottom) at pH 10 and 80 °C for four cycles. The Gly_n resonances (shown in brackets) are observed are at about a third decreased intensity (Gly_2 integrations are labelled on the spectrum) for the DH/RH reaction of 3',5'-cAMP and glycine compared to the reaction of glycine alone.

Characterisation of the 3',5'-cAMP + Glycine Reaction by Mass Spectrometry

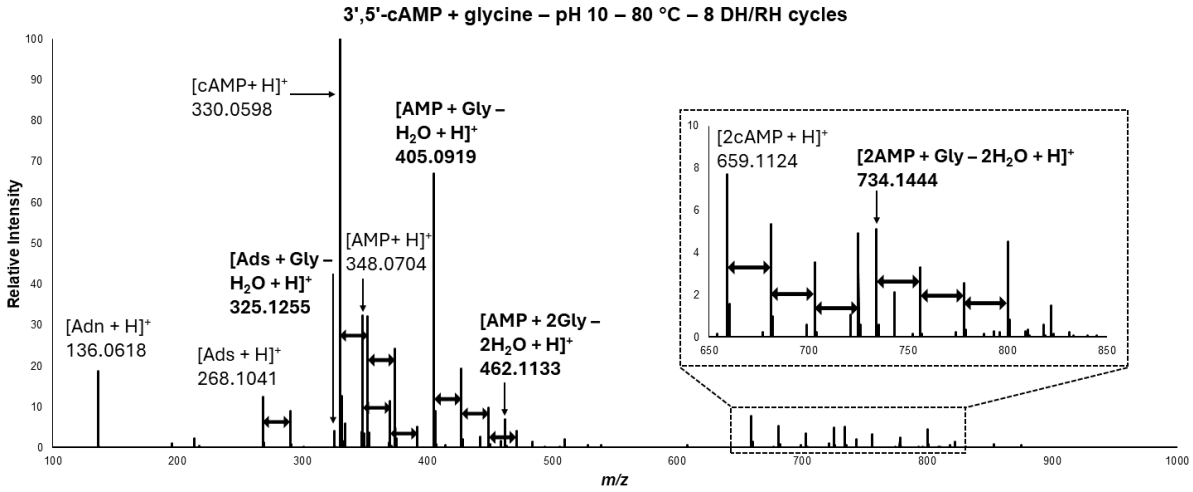


Figure S15: Fully assigned mass spectrum of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 80 °C after 8 cycles. Black arrows in between peaks represent singly charged sodium adducts. Labels in bold are assigned to AMP-glycine conjugate derivatives.

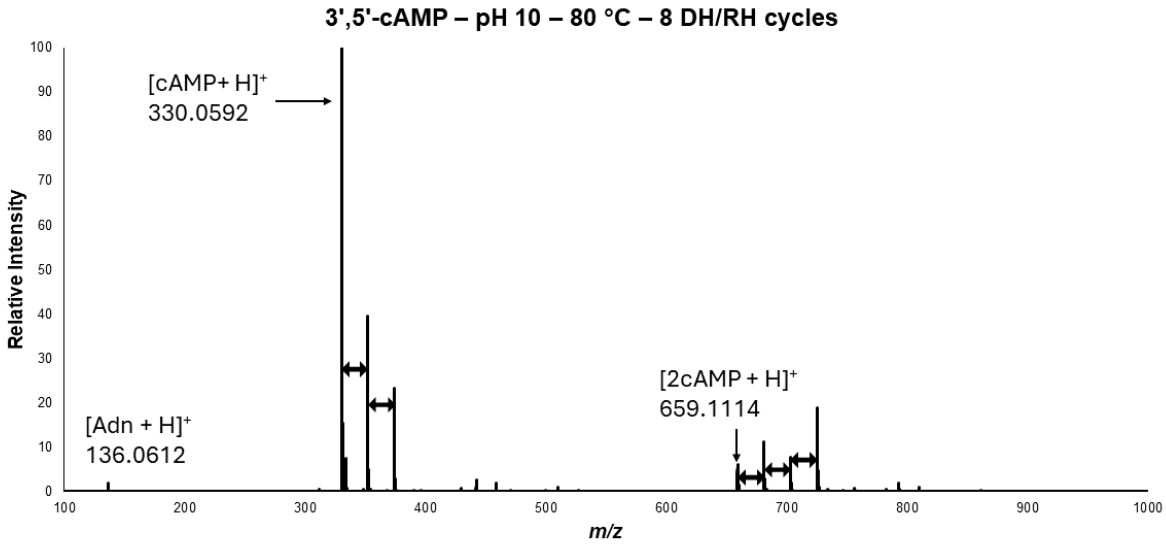


Figure S16: Mass spectrum of the DH/RH reaction of 3',5'-cAMP alone at pH 10 and 80 °C after 8 cycles. Black arrows in between peaks represent singly charged sodium adducts. Only minor decomposition is observed by the formation of adenine.

Table S1: List of mass spectrum assignments of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 80 °C for 8 cycles from Fig. S15. Products are listed alongside the calculated *m/z*, observed *m/z*, and error (ppm) between the two.

Product	$m/z_{calc.}$	$m/z_{obs.}$	Error (ppm)
$[Adn + H]^+$	136.0618	136.0610	5.88
$[Ads + H]^+$	268.1041	268.1034	2.61
$[Ads + Gly - H_2O + H]^+$	325.1255	325.1247	2.46
$[cAMP + H]^+$	330.0598	330.0591	2.12
$[AMP + H]^+$	348.0704	348.0696	2.30
$[AMP + Gly - H_2O + H]^+$	405.0919	405.0910	2.22
$[AMP + 2Gly - 2H_2O + H]^+$	462.1133	462.1125	1.73
$[2cAMP + H]^+$	659.1124	659.1113	1.67
$[2AMP + Gly - H_2O + H]^+$	734.1444	734.1430	1.91

Table S2: List of m/z_{obs} values for peaks from the mass spectra of the DH/RH reaction of 3',5'-cAMP and glycine alongside the same reaction repeated with glycine-$^{13}C_2$. Peaks which show a +2 Dalton increase per glycine subunit are highlighted in green.

Product	m/z_{obs} with glycine-^{12}C	m/z_{obs} with glycine-$^{13}C_2$
$[Adn + H]^+$	136.0610	136.0611
$[Ads + H]^+$	268.1034	268.1033
$[Ads + Gly - H_2O + H]^+$	325.1247	327.1314
$[cAMP + H]^+$	330.0591	330.0590
$[AMP + H]^+$	348.0696	348.0696
$[AMP + Gly - H_2O + H]^+$	405.0910	407.0977
$[AMP + 2Gly - 2H_2O + H]^+$	462.1125	466.1259
$[2cAMP + H]^+$	659.1113	659.1114
$[2AMP + Gly - H_2O + H]^+$	734.1430	736.1499

NMR Standards

Gly-*N*-pA

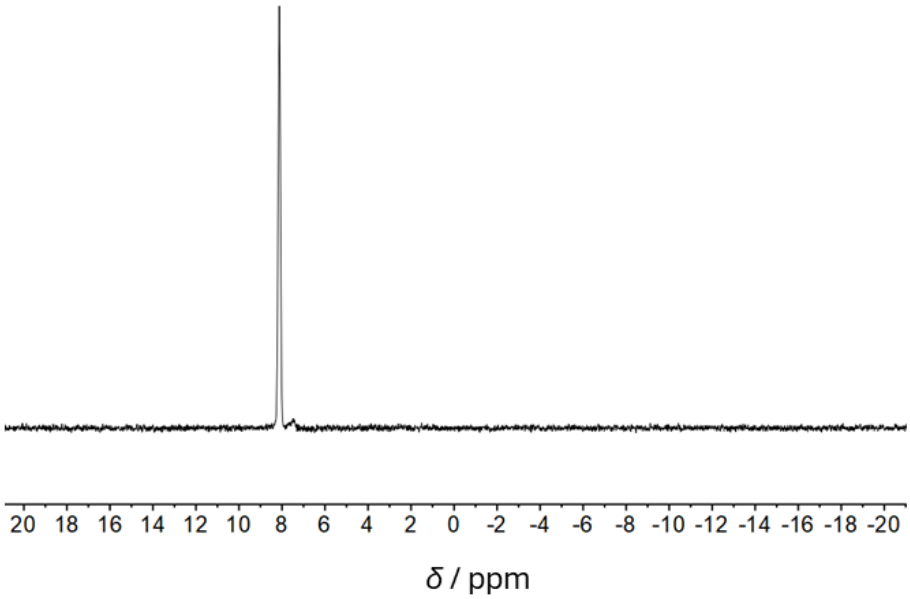


Figure S17: $^{31}P\{^{1}H\}$ NMR spectrum of the Gly-*N*-pA standard in 90% water and 10% D_2O.

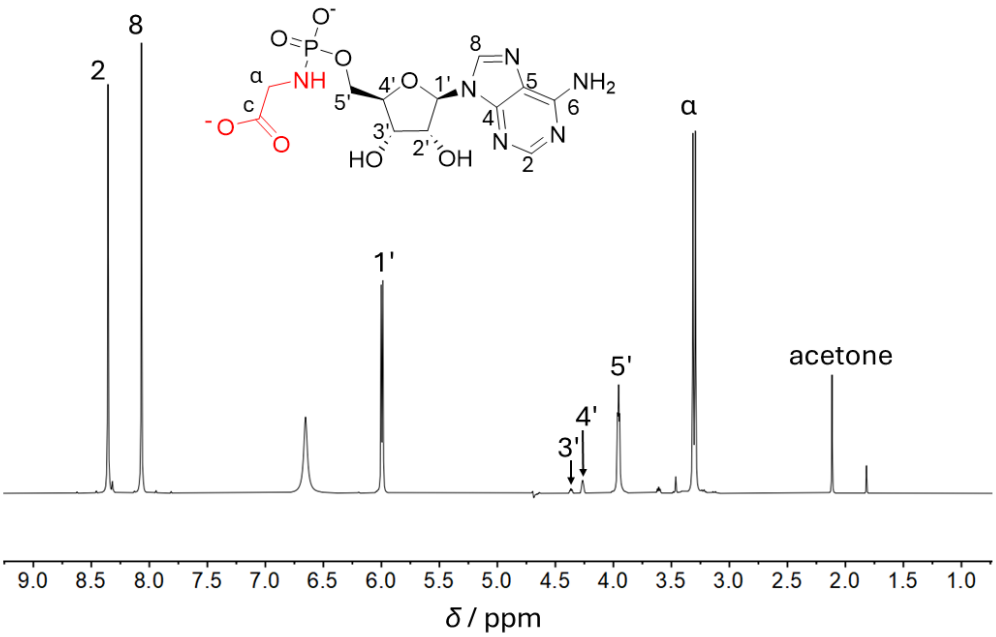


Figure S18: ^{1}H NMR spectrum of the Gly-*N*-pA standard in 90% water and 10% D_2O. The 2', 3' and 4' protons have reduced intensities due to proximity to the water peak from its excitation during water suppression.

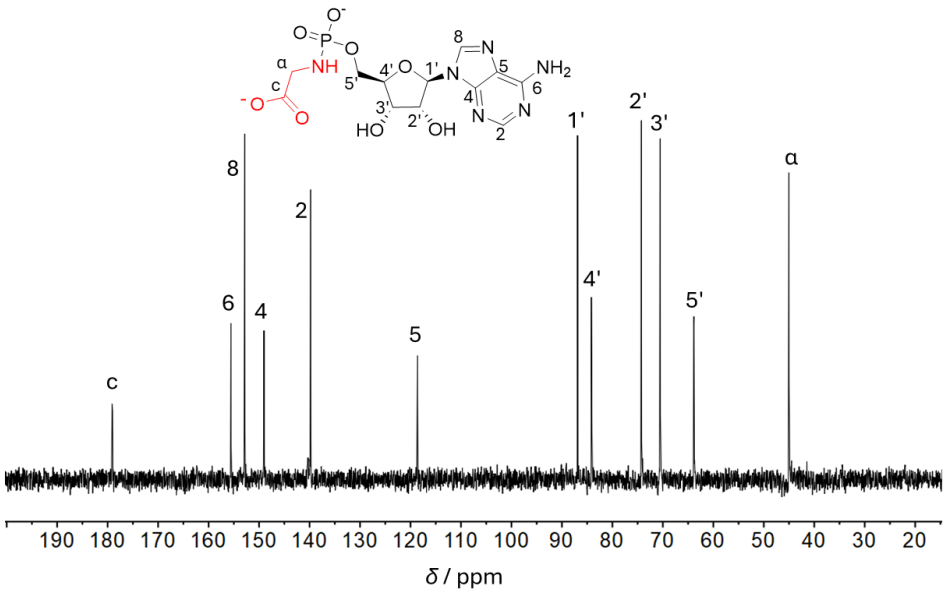


Figure S19: ^{13}C NMR spectrum of the Gly-*N*-pA standard in 100% D_2O.

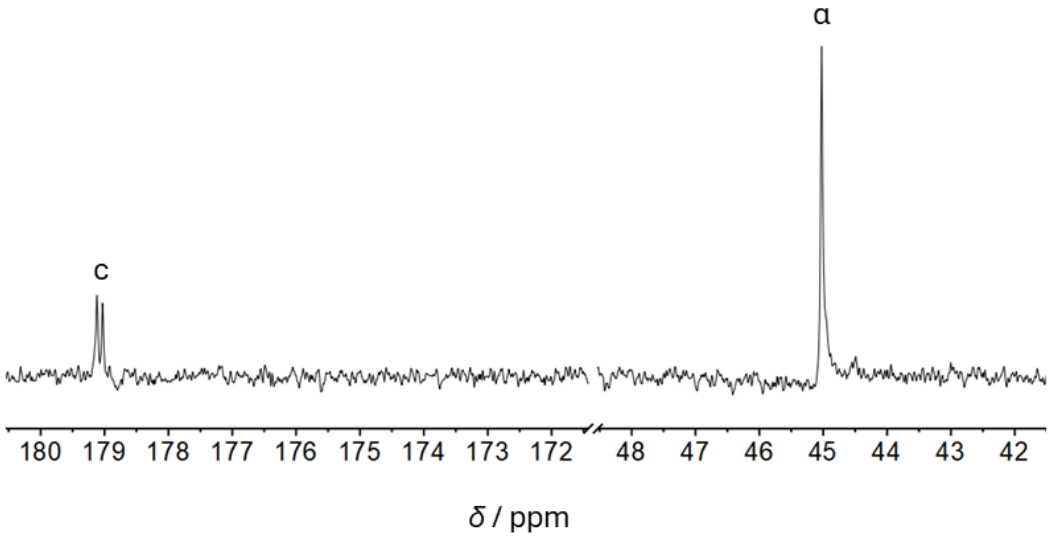


Figure S20: Partial ^{13}C NMR spectrum of the Gly-*N*-pA standard, zoomed in on the aliphatic and carbonyl regions. Shifts and splitting match with those observed after the DH/RH reaction of 3',5'-cAMP and glycine.

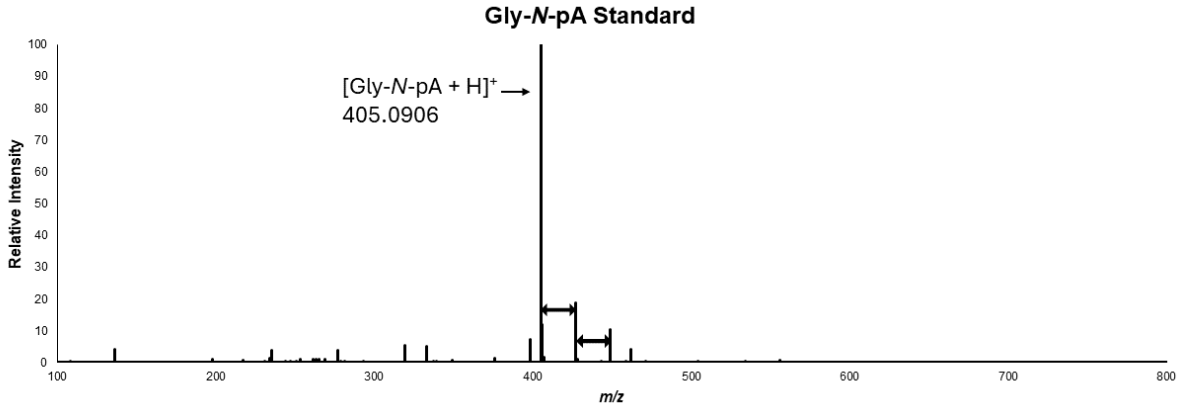


Figure S21: Mass spectrum of the Gly-*N*-pA standard. The m/z_{obs} is within a 3.21 ppm error of the m/z_{calc} for the $[M + H]^+$ molecular ion. Black arrows in between peaks represent singly charged sodium adducts.

Gly_2-*N*-pA

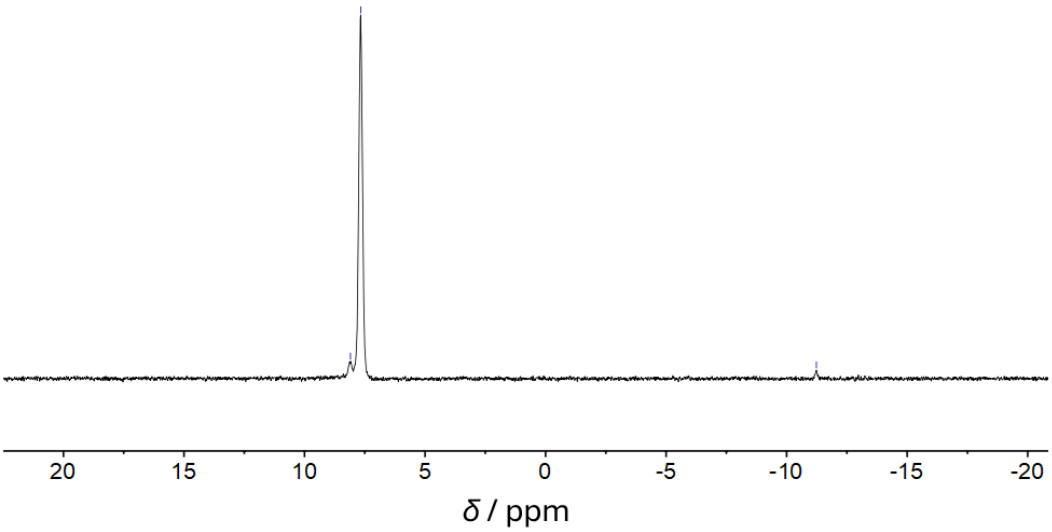


Figure S22: $^{31}P\{^1H\}$ NMR spectrum of the Gly_2-*N*-pA standard in 90% water and 10% D_2O. The resonance at δ = −11.2 ppm is unassigned yet corresponds to literature reports for a pyrophosphate-linked AMP dimer.[9]

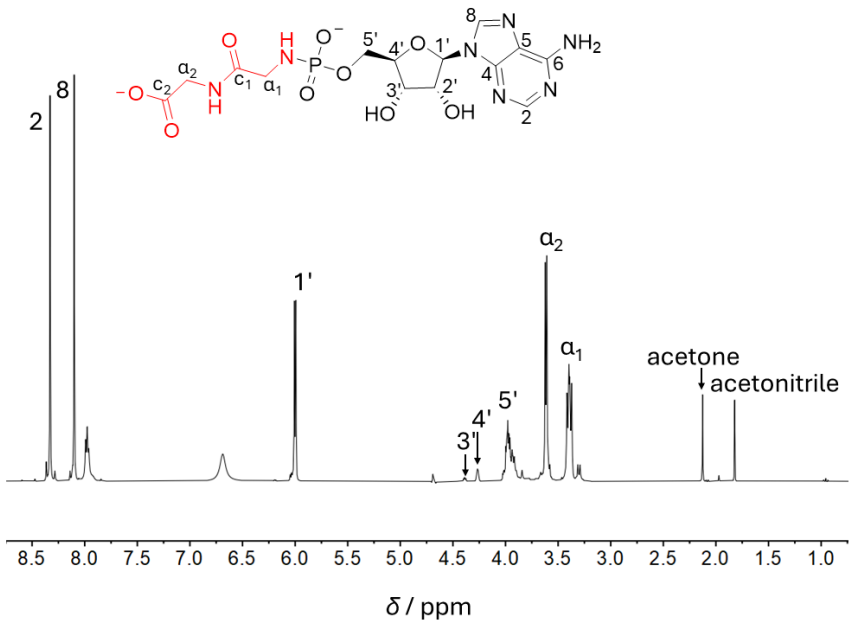


Figure S23: ^1H NMR spectrum of the Gly_2-*N*-pA standard in 90% water and 10% D_2O. The 2', 3' and 4' protons have reduced intensities due to proximity to the water peak from its excitation during water suppression. The resonances at δ = 7.97 ppm are suspected to arise from the amide N-H.

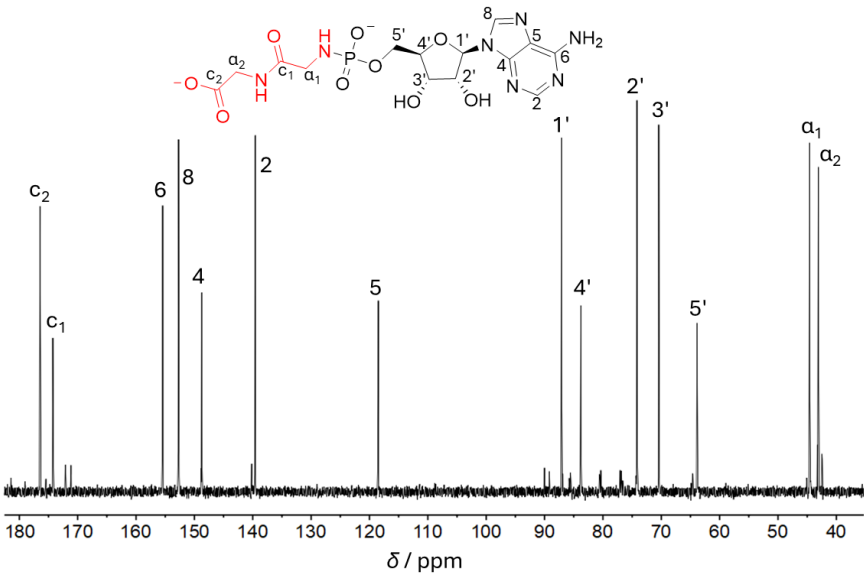


Figure S24: ^{13}C NMR spectrum of the Gly_2-*N*-pA standard in 90% water and 10% D_2O.

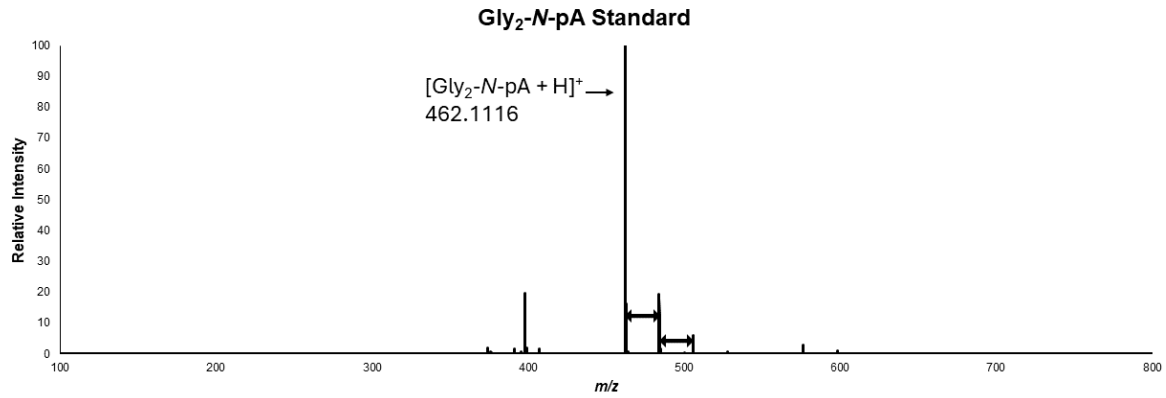
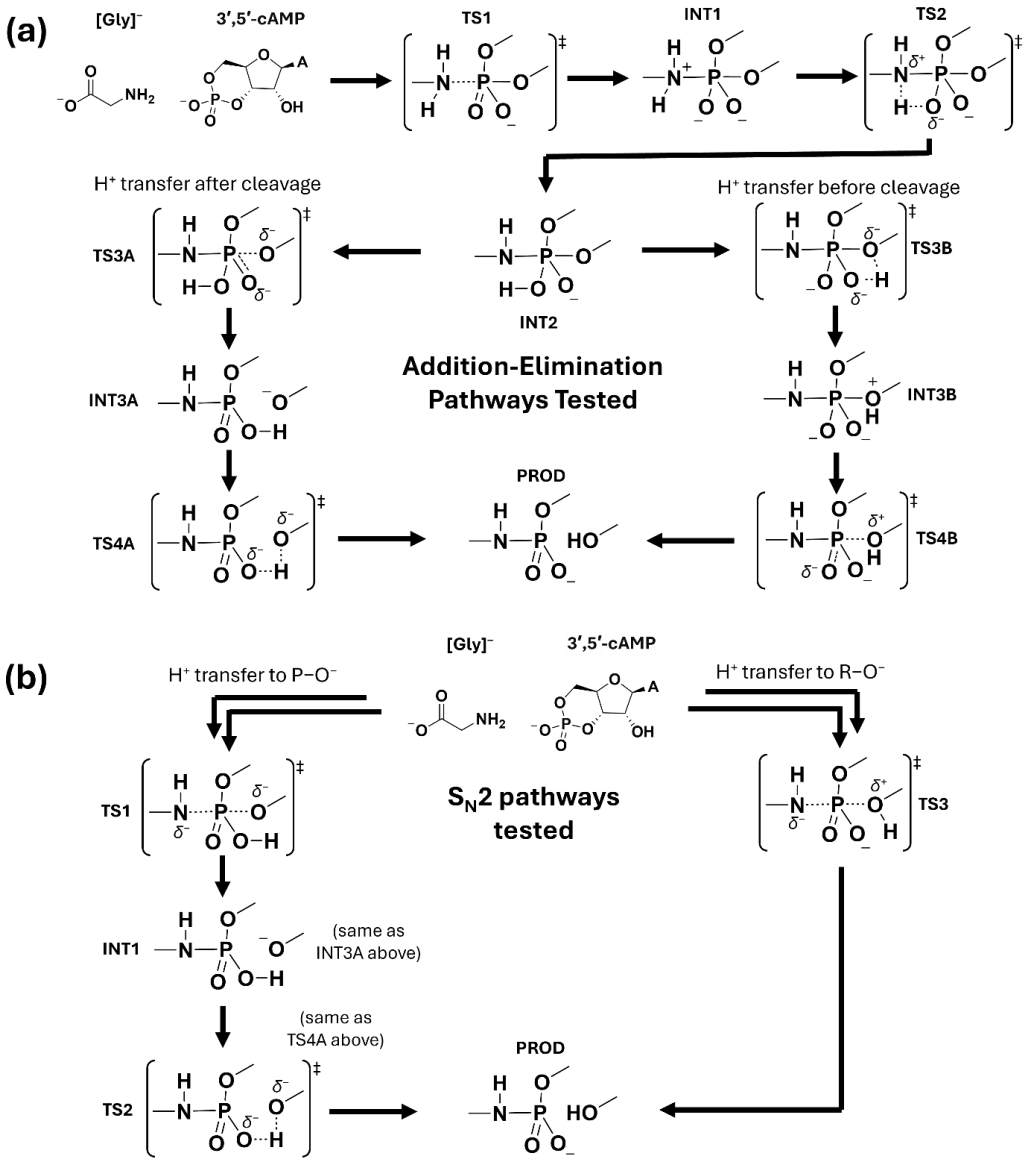


Figure S25: Mass spectrum of the Gly_2-*N*-pA standard. The m/z_{obs} is within a 3.68 ppm error of the m/z_{calc} for the $[M + H]^+$ molecular ion. Black arrows in between peaks represent singly charged sodium adducts.

Mechanism Search

A variety of initial configurations were tested to search the potential energy surface around the phosphoester cleavage for saddle points and minima corresponding to transition states and intermediate structures. The following diagrams (Scheme S1) illustrate the various addition-elimination substitution and S_N2 pathways searched before ultimately finding a consistent mechanism of converged states.



Scheme S1: Diagrams for how pathways were tested computationally for (a) addition-elimination pathways and (b) S_N2 pathways. Note, only the reaction sites and not the entire molecules are shown for clarity. For S_N2 pathways, the initial proton transfer step from the amine of Gly to one of the oxygens has been omitted, because the bimolecular S_N2 transition states were found to not exist on the potential energy surface.

For the S_N2 reactions, no stable saddle point was found corresponding to a bimolecular transition state. Therefore, these pathways were discarded. The initial glycine addition step still proved illusive for the addition-elimination pathways, and further bond constrained geometry optimizations were used to scan the N—P distance in the region of the Gly addition step, as shown in Figure S26.

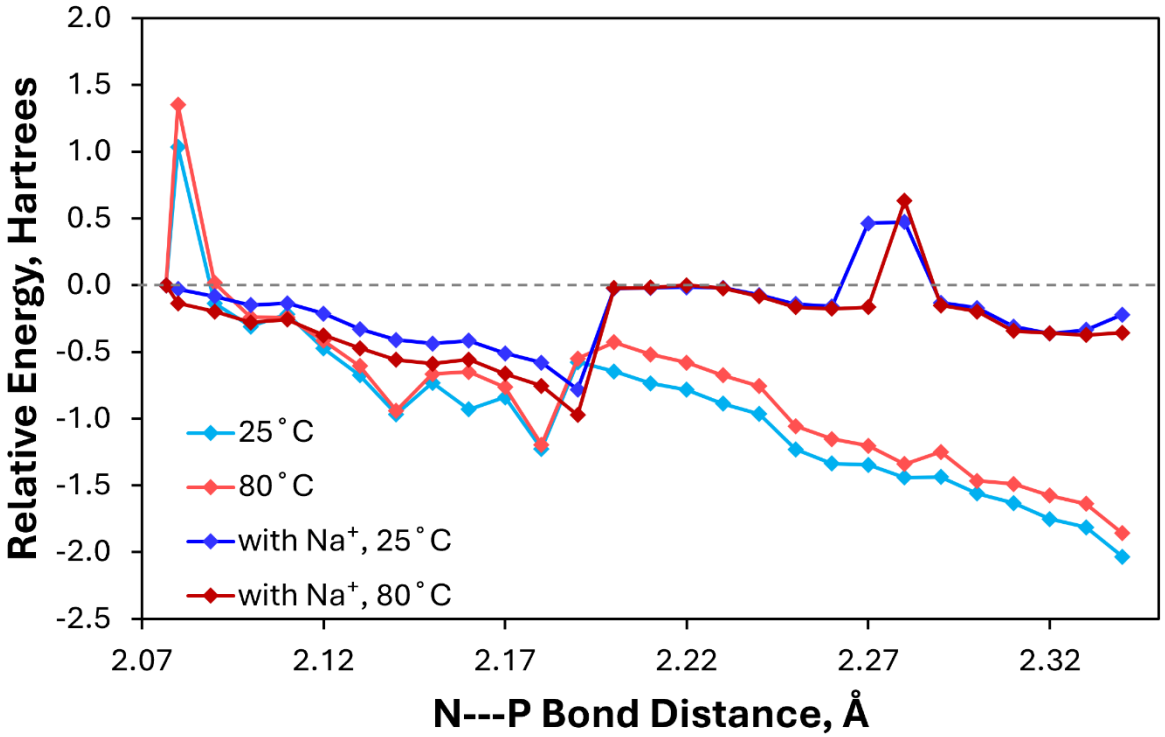


Figure S26: Bond-scan geometry optimization carried out across the N---P bond in TS1 for addition-elimination pathways. Note, energies are not plotted on an absolute scale but relative to the energy of the starting point at 2.077 Å for each plot, respectively.

This bond-scan geometry optimization revealed there were a significant number of low-lying minima and saddle points corresponding to TS1 (NH_2 attachment) and INT1 (R-NH_2-PO_4) in the region of 2.09–2.26 Å for addition-elimination pathways. The following Table S3 details the presence and number of imaginary frequencies corresponding to saddle points along this path. Note, the transition state detected at 2.08 Å was found to be a bucking/rocking saddle point associated with larger glycine reorientation.

Table S3: Imaginary frequencies for saddle points along the paths shown in Figure S26. *

N---P bond (Å)	298.15 K	353.15 K	with Na^+, 298.15 K	with Na^+, 353.15 K
2.077				
2.080	-7.1	-5.99		
2.090				
2.100				
2.110				
2.120				
2.130				
2.140				
2.150				
2.160				
2.170				
2.180				
2.190	-30.11	-38		
2.200	-50.01	-50.02	-32.66	-32.64
2.210	-62.69	-62.7	-53.59	-53.75
2.220	-73.09	-73.09	-65.97	-66.06
2.230	-81.71	-81.72	-75.38	-75.4
2.240	-93.66	-93.38	-84.03	-84.07
2.250	-103.94	-103.94	-92.32	-92.33
2.260	-109.35	-109.35	-99.33	-99.32
2.270	-114.79	-114.5	-105.57, -38.52	-105.49
2.280	-120.39	-120.42	-110.9, -38.54	-110.89, 38.53
2.290	-124.77	-124.78	-115.72	-115.72
2.300	-128.55	-128.54	-120.07	-120.02
2.310	-132.23	-132.25	-123.8	-123.86
2.320	-134.93	-134.94	-128.58	-127.29
2.330	-137.05	-137.05	-130.52	-130.52
2.340	-139.17	-139.17	-133.03	-132.99

* Blank entries indicate a minima corresponding to the INT1 structure where the negative imaginary frequency corresponds to a saddle point indicative of TS1.

From this bond scan we took the lowest energy minima state as INT1 and the highest energy saddle point as TS1 for each condition set. The following Table S4 summarizes the data points chosen.

Table S4: Bond distances determined for INT1 and TS1 structures.

	298.15 K	**353.15 K**	**with Na^+, 298.15 K**	**with Na^+, 353.15 K**
INT1	2.18 Å	2.18 Å	2.19 Å	2.19 Å

Computational Energies

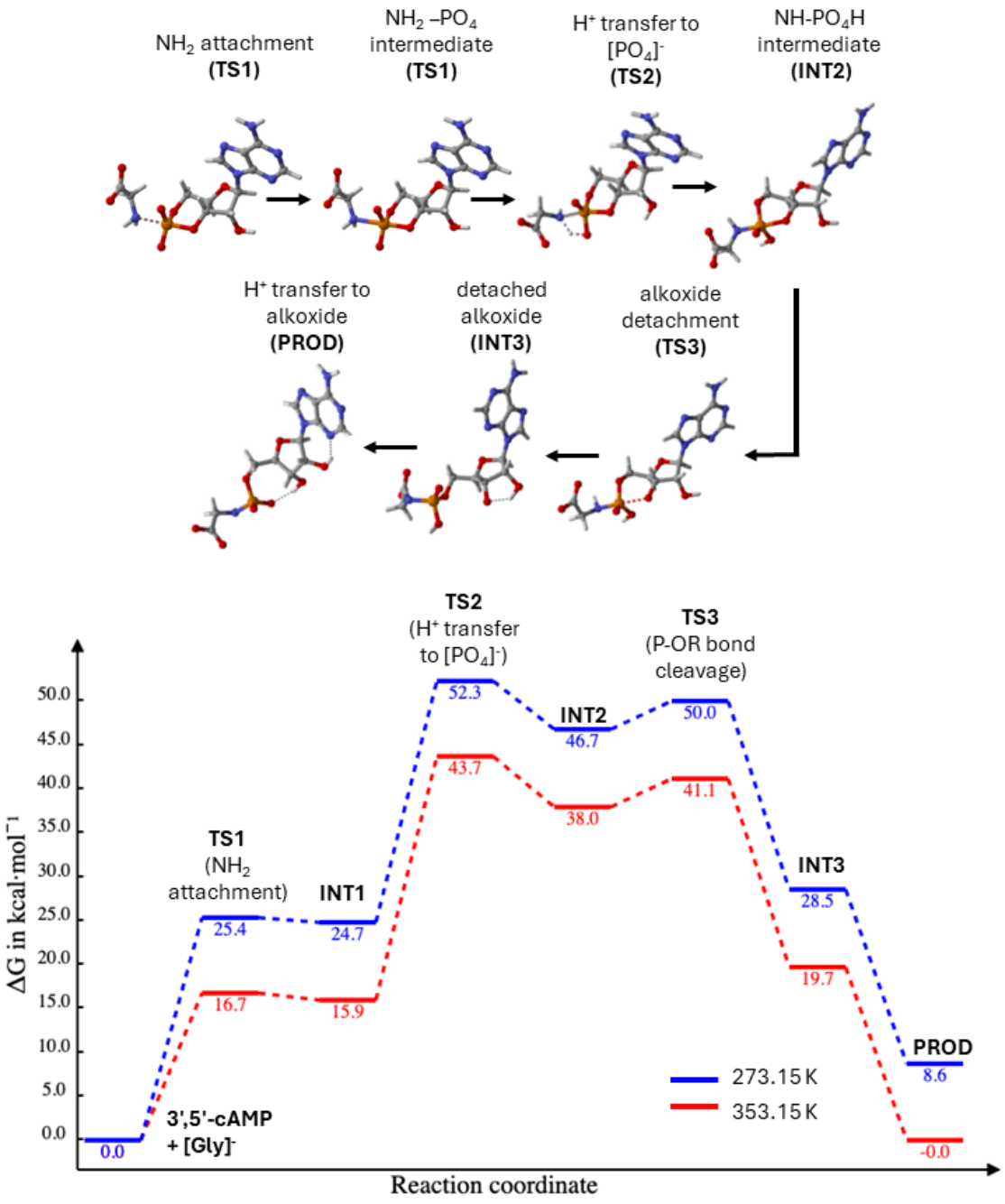


Figure S27: DFT calculated reaction free energy profile and structures of glycine addition to 3',5'-cAMP and subsequent formation of Gly-*N*-pA at 273.15 K (in blue) and 353.15 K (in red).

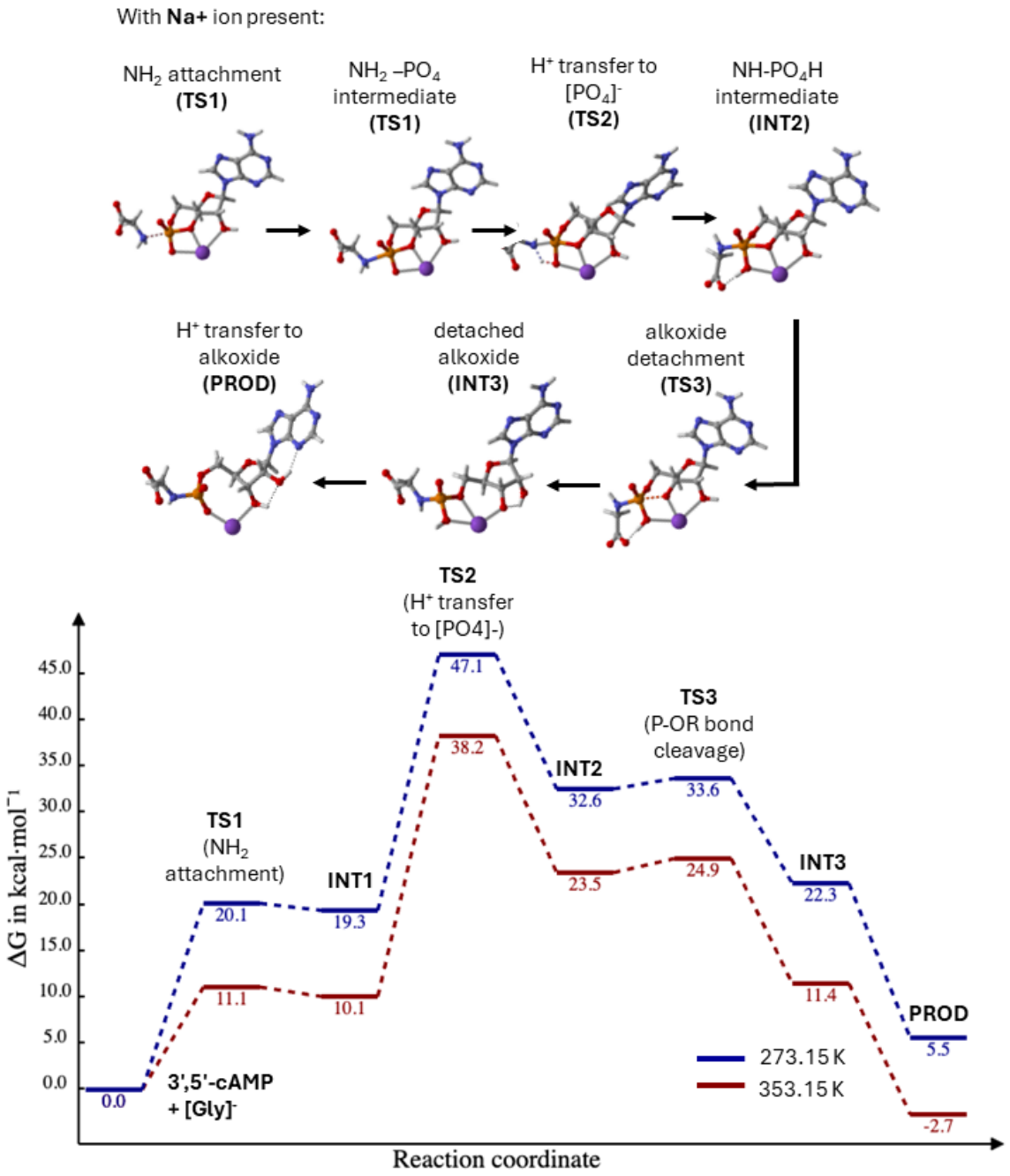


Figure S28: DFT calculated reaction free energy profile and structures of glycine addition to 3',5'-cAMP and subsequent formation of Gly-*N*-pA at 273.15 K and 353.15 K in the presence of a single desolvated Na^+ ion coordinating to the phosphate group.

Reactions with Alternative Amines

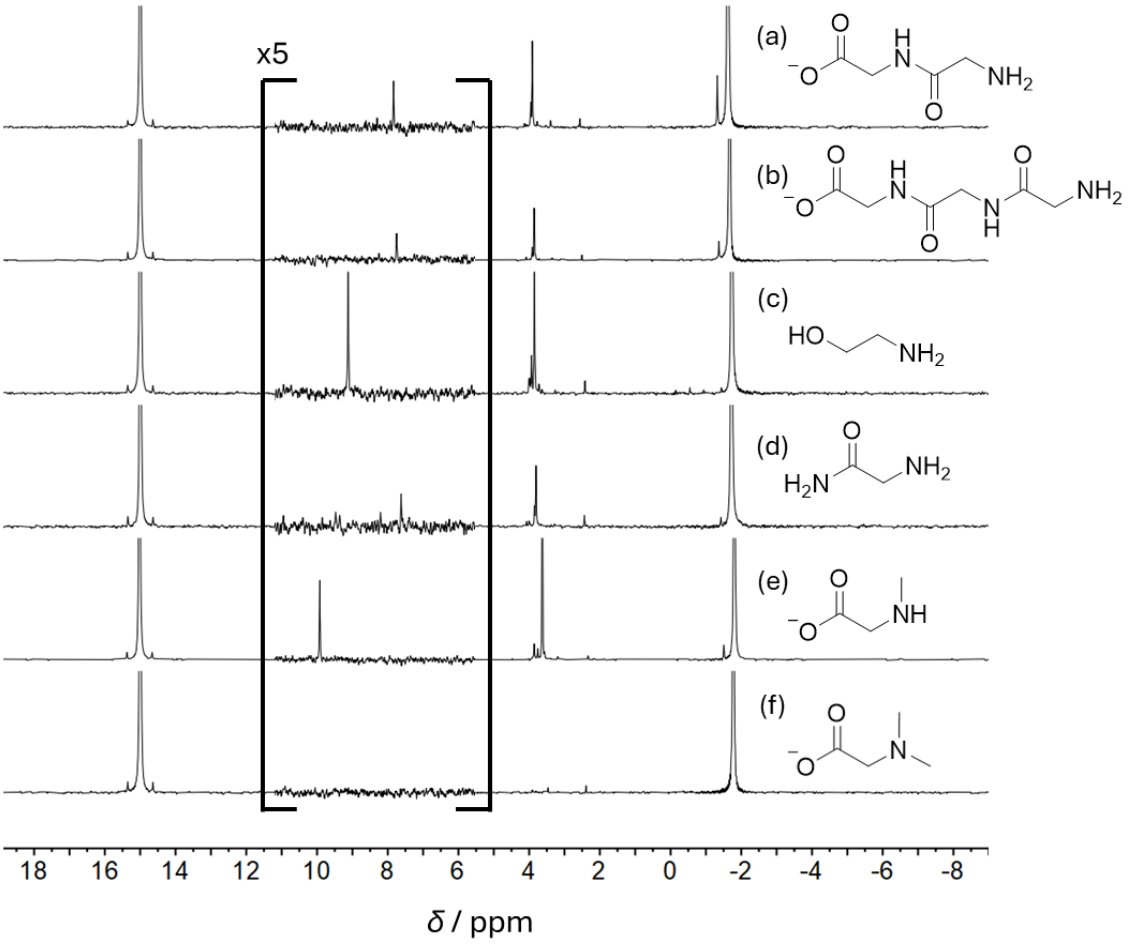


Figure S29: $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 3',5'-cAMP and a series of amines (a: glycylglycine, b: triglycine, c: ethanolamine, d: glycinamide, e: *N*-methylglycine, f: *N,N*-dimethylglycine) at pH 10 and 80 °C for 1 cycle. The reactions with amines a–e show new resonances in the phosphoramidate region (δ = ~6–10 ppm) as well as similar peaks in the phosphate monoester region (δ = ~2–5 ppm) as observed in the reaction with glycine. Amine f does not show any new resonances in the phosphoramidate region.

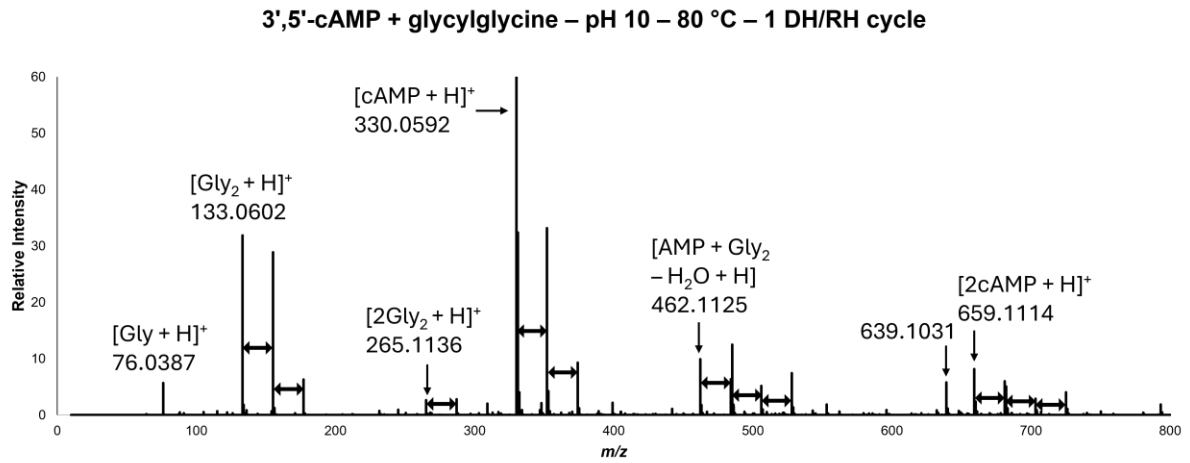


Figure S30: Mass spectrum of the DH/RH reaction of 3',5'-cAMP and glycylglycine (Gly_2) at pH 10 and 80 °C for 1 cycle. Black arrows in between peaks represent singly charged sodium adducts. The peak at *m/z* = 462.1125 corresponds to the calculated $[M + H]^+$ molecular ion of Gly_2-*N*-pA within a 1.73 ppm error.

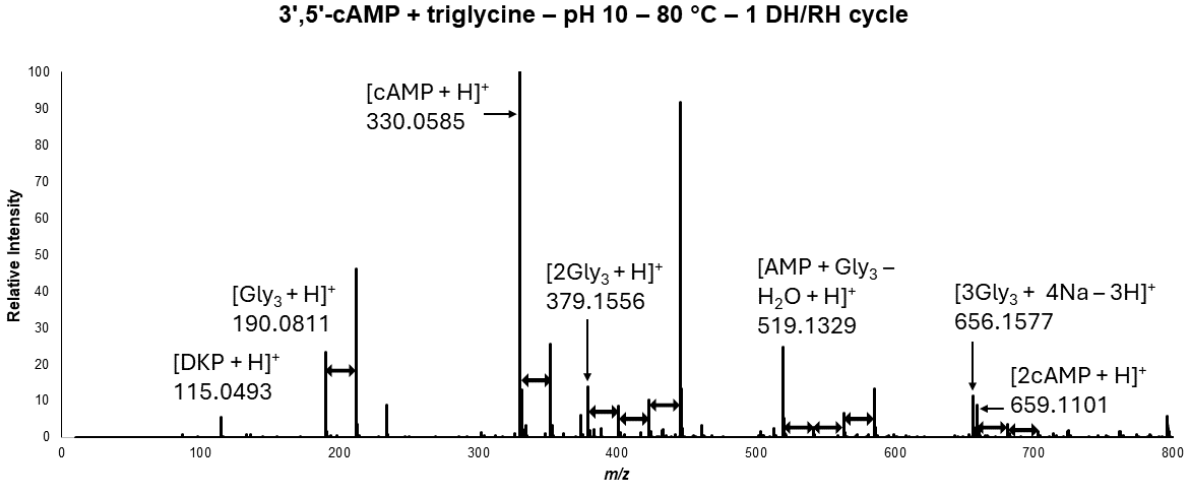


Figure S31: Mass spectrum of the DH/RH reaction of 3',5'-cAMP and triglycine (Gly_3) at pH 10 and 80 °C for 1 cycle. Black arrows in between peaks represent sodium adducts. The peak at *m/z* = 519.1329 corresponds to the calculated $[M + H]^+$ molecular ion of Gly_3-*N*-pA within a 3.66 ppm error.

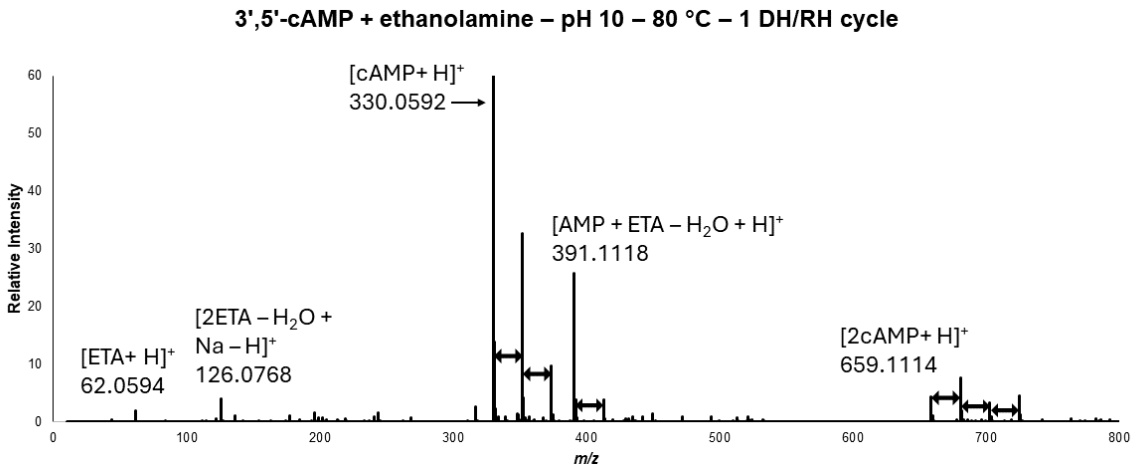


Figure S32: Mass spectrum of the DH/RH reaction of 3',5'-cAMP and ethanolamine (ETA) at pH 10 and 80 °C for 1 cycle. Black arrows in between peaks represent singly charged sodium adducts. The peak at *m/z* = 391.1118 corresponds to the calculated $[M + H]^+$ molecular ion of ETA-*N*-pA within a 2.05 ppm error.

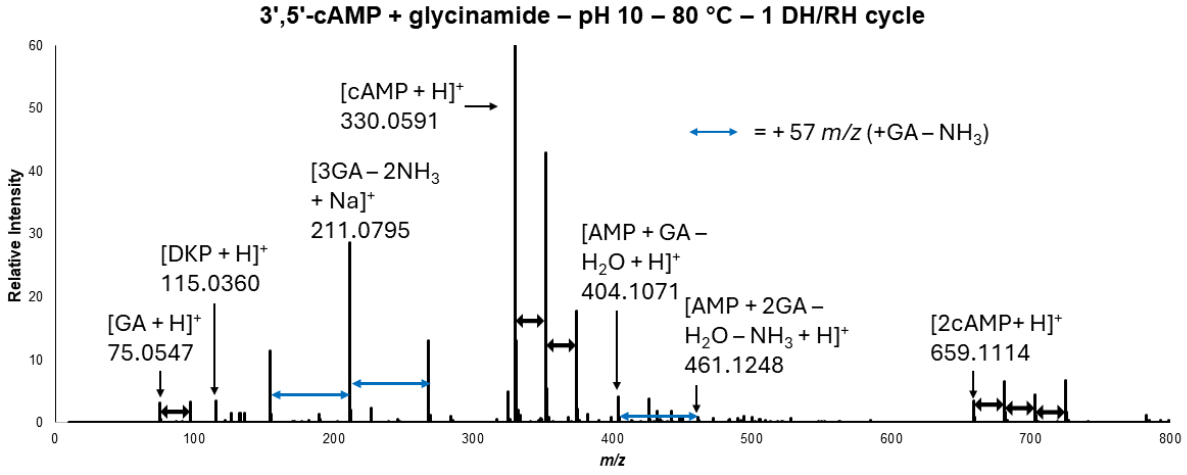


Figure S33: Mass spectrum of the DH/RH reaction of 3',5'-cAMP and glycinamide (GA) at pH 10 and 80 °C for 1 cycle. Black arrows in between peaks represent singly charged sodium adducts. Blue arrows between peaks show glycinamide conjugation. Glycinamide oligomers form regardless of the presence of 3',5'-cAMP, i.e., when glycinamide is dried down alone (data not shown). The peak at *m/z* = 404.1071 corresponds to the calculated $[M + H]^+$ molecular ion of GA-*N*-pA within a 1.98 ppm error.

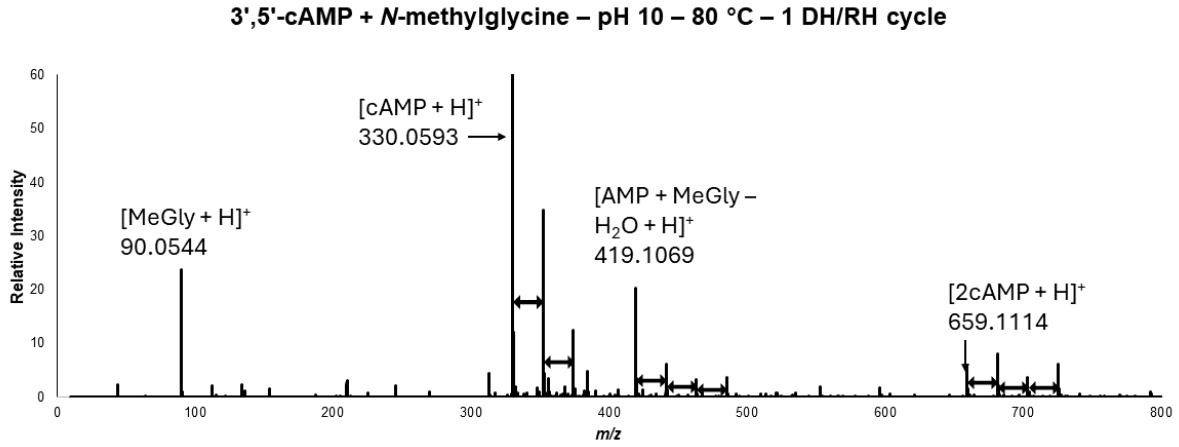


Figure S34: Mass spectrum of the DH/RH reaction of 3',5'-cAMP and *N*-methylglycine (MeGly) at pH 10 and 80 °C for 1 cycle. Black arrows in between peaks represent singly charged sodium adducts. The peak at *m/z* = 419.1069 corresponds to the calculated $[M + H]^+$ molecular ion of the MeGly-*N*-pA conjugate within a 1.43 ppm error.

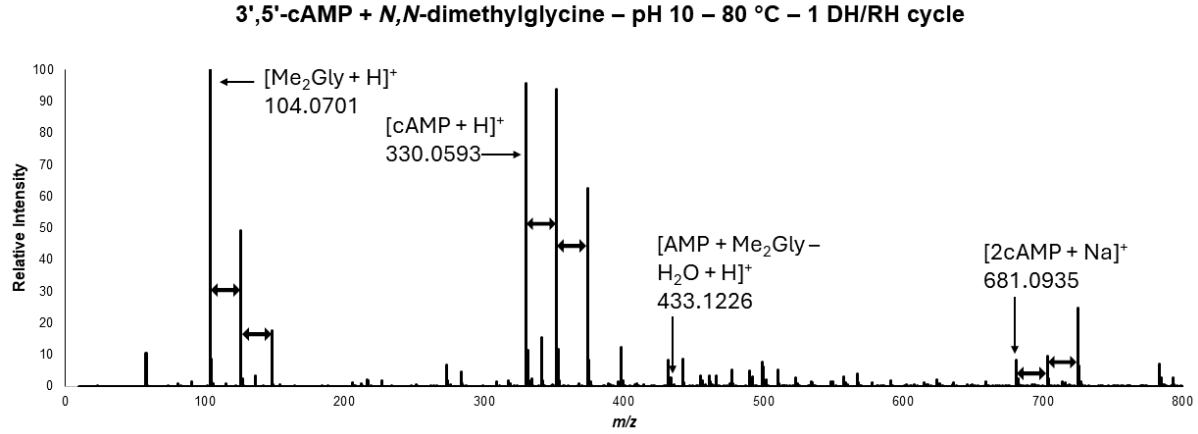


Figure S35: Mass spectrum of the DH/RH reaction of 3',5'-cAMP and *N,N*-dimethylglycine (Me_2Gly) at pH 10 and 80 °C for 1 cycle. Black arrows in between peaks represent singly charged sodium adducts. A small peak at *m/z* = 433.1226 corresponds to the calculated $[M + H]^+$ molecular ion of the Me_2Gly-*N*-pA conjugate within a 1.39 ppm error. This product is only formed in trace amounts as corroborated by the lack of an observable phosphoramidate resonance in the $^{31}P\{^1H\}$ NMR spectrum (see Fig. S29, amine f).

2',3'-cAMP + Glycine Reaction

2',3'-cAMP + Glycine Under Standard DH/RH Conditions

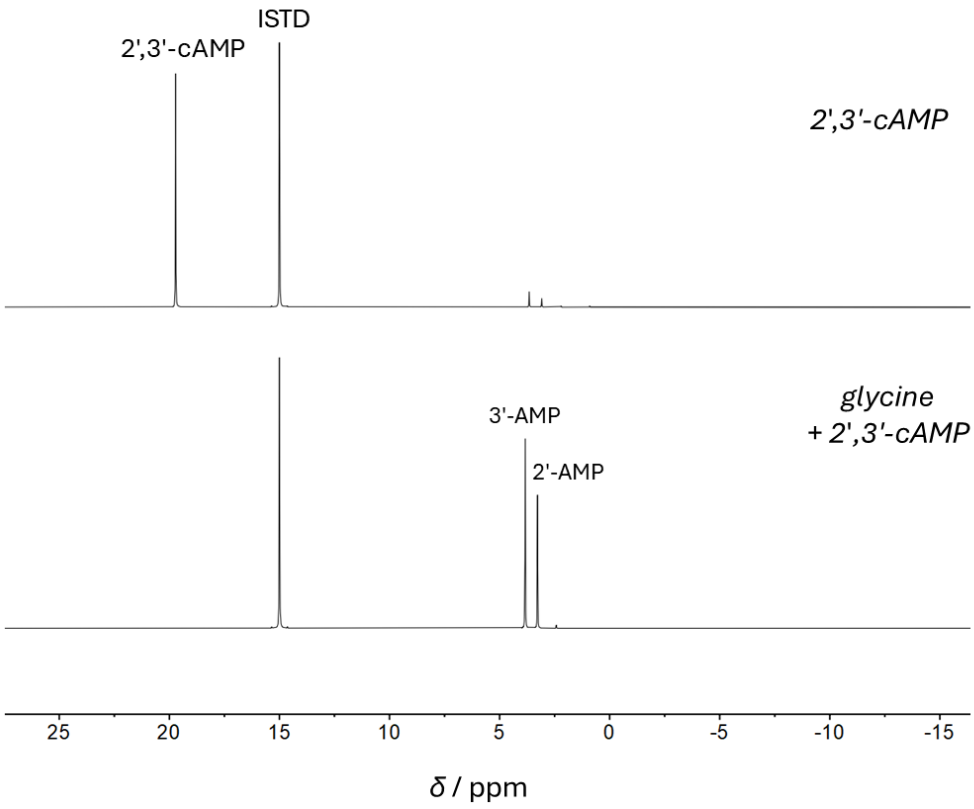


Figure S36: $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 2',3'-cAMP (top spectrum) and 2',3'-cAMP and glycine (bottom spectrum) at pH 10 and 80 °C for 1 cycle. Glycine increases the rate of conversion of 2',3'-cAMP to 2'- and 3'-AMP, and all of the 2',3'-cAMP is consumed after 1 cycle.

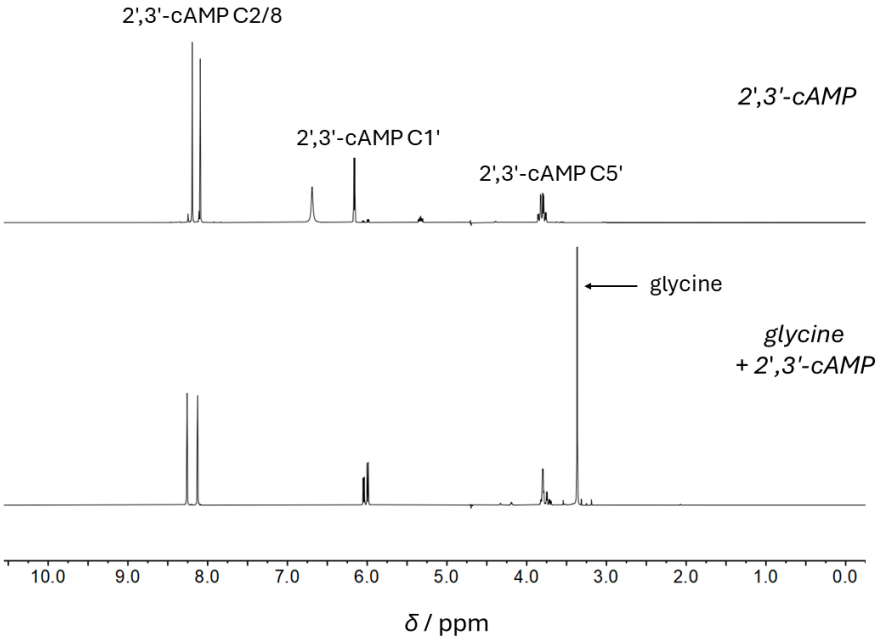


Figure S37: ^{1}H NMR spectra of the DH/RH reaction of 2',3'-cAMP alone (top spectrum) and 2',3'-cAMP and glycine (bottom spectrum) at pH 10 and 80 °C for 1 cycle.

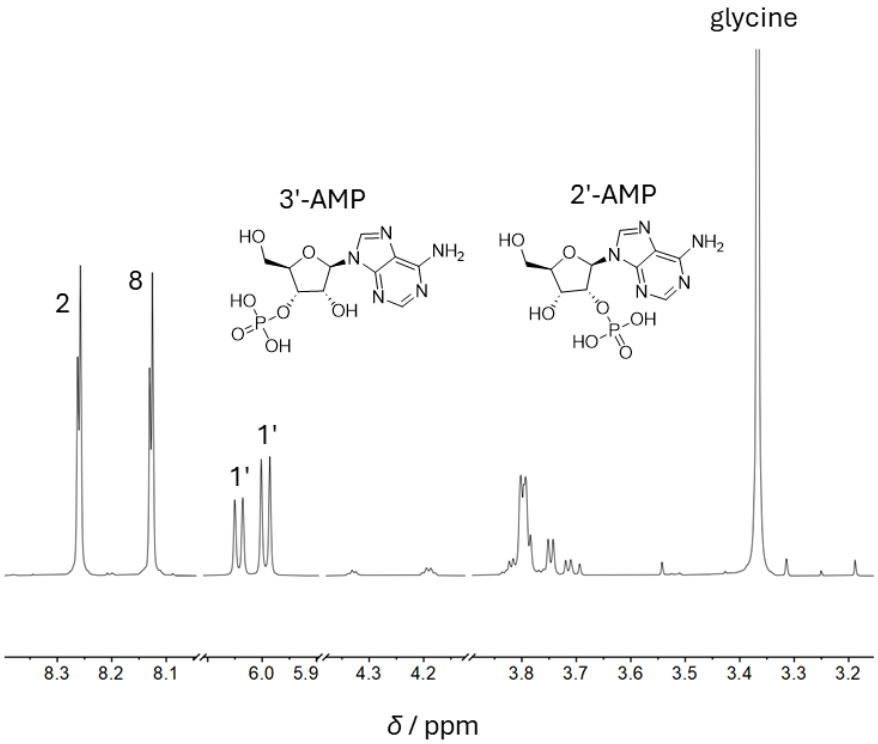


Figure S38: Partial ^{1}H NMR spectrum of the DH/RH reaction of 2',3'-cAMP and glycine at pH 10 and 80 °C for 1 cycle. Two separate C2, C8, and C1' peaks are visible demonstrating that two isomers are formed.

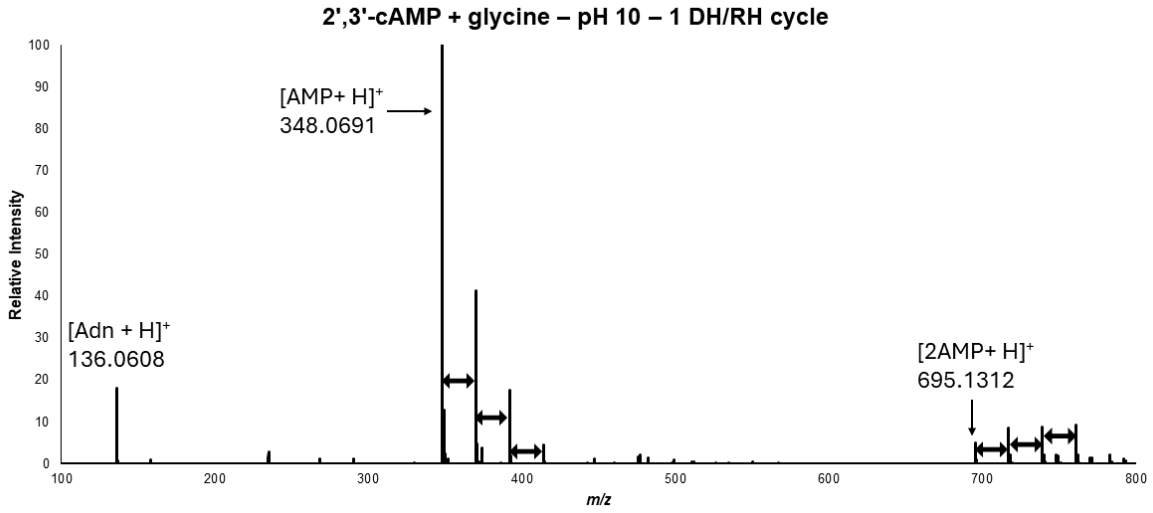


Figure S39: Mass spectrum of the DH/RH reaction of 2',3'-cAMP and glycine at pH 10 and 80 °C for 1 cycle. Black arrows in between peaks represent singly charged sodium adducts.

2',3'-cAMP + Glycine Under Vacuum DH/RH Conditions

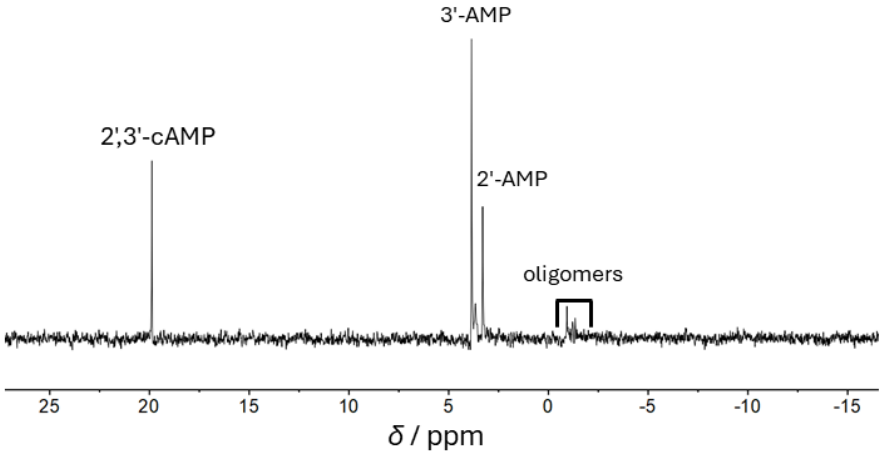


Figure S40: $^{31}P\{^{1}H\}$ NMR spectrum of the DH/RH reaction of 2',3'-cAMP and glycine at pH 10 and 25 °C under vacuum for three days. Small resonances are visible around 0 ppm indicating oligomer formation. The results confirm Verlander *et al.*'s[1] observations of oligomer formation but reveal that 2'- and 3'-AMP are the major products of this reaction.

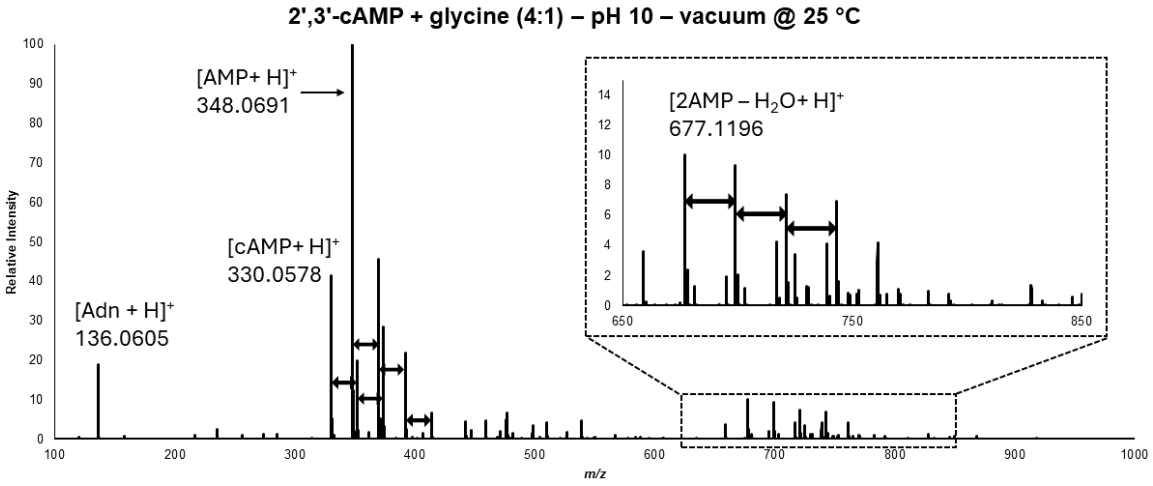


Figure S41: Mass spectrum of the DH/RH reaction of 2',3'-cAMP and glycine at pH 10 and 25 °C under vacuum for three days. Black arrows in between peaks represent singly charged sodium adducts. The peak at 677.1196 corresponds to the $[M+H]^+$ molecular ion for an AMP dinucleotide within a 4.87 ppm error. The reaction of 2',3'-cAMP alone under the same conditions does not reveal any evidence for oligomerisation (data not shown).

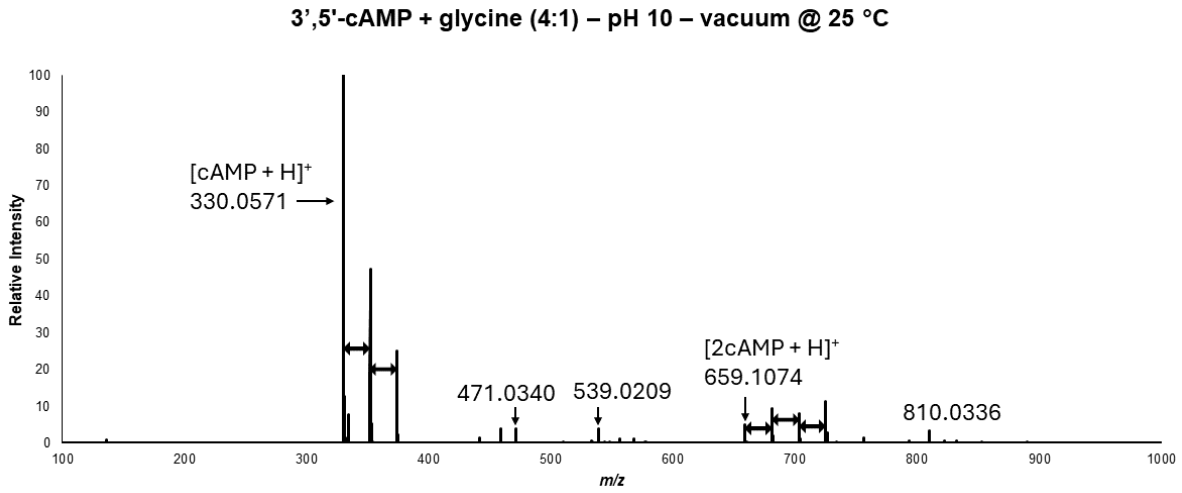


Figure S42: Mass spectrum of the DH/RH reaction of 3',5'-cAMP and glycine at pH 10 and 25 °C under vacuum for three days. Black arrows in between peaks represent singly charged sodium adducts. No peak corresponding to an AMP dinucleotide is observed. The $^{31}P\{^{1}H\}$ NMR spectrum of this reaction additionally shows no new resonances (data not shown).

5'-AMP + Glycine Reaction

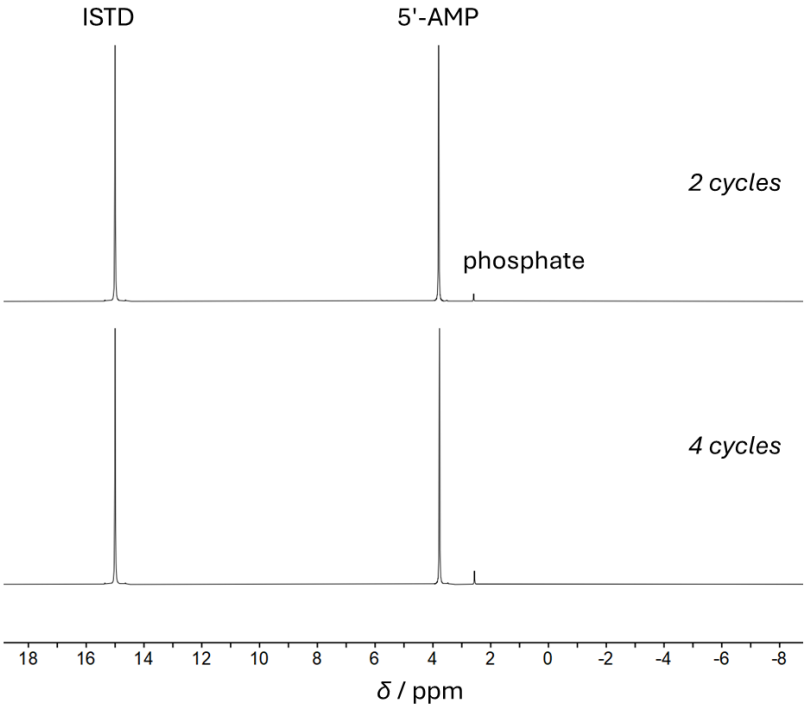


Figure S43: $^{31}P\{^{1}H\}$ NMR spectrum of the DH/RH reaction of 5'-AMP and glycine at pH 10 and 80 °C for two and four cycles. The reaction shows that only minor dephosphorylation occurs.

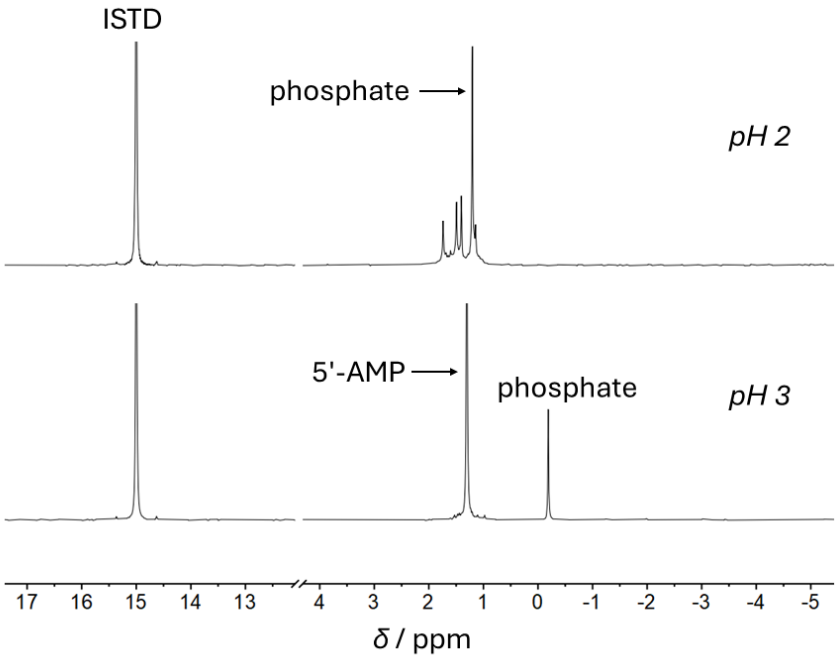


Figure S44: $^{31}P\{^{1}H\}$ NMR spectra of the DH/RH reaction of 5'-AMP and glycine at pH 2 (top spectra) and pH 3 (bottom spectra) and 80 °C for one cycle. The phosphate peak was confirmed by spiking with a standard (data not shown). pH induced shift of this peak is observed between the two spectra.

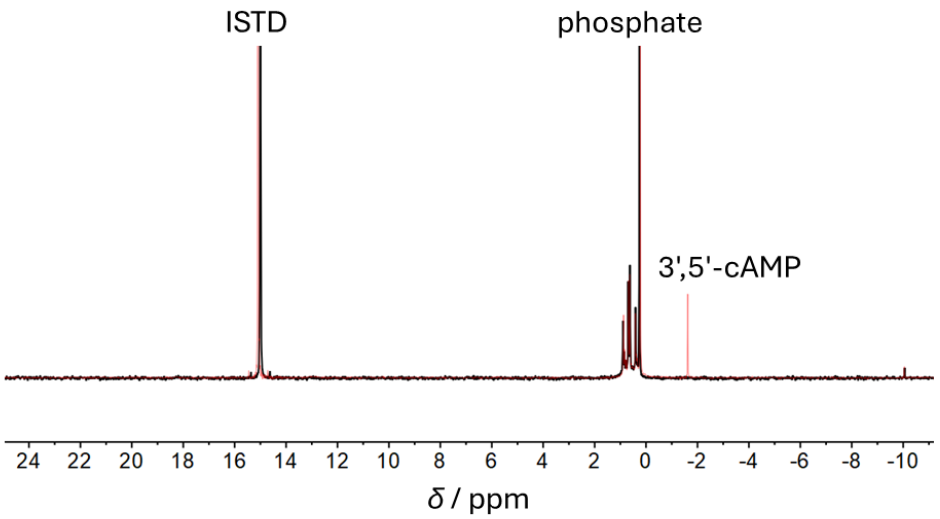


Figure S45: $^{31}P\{^{1}H\}$ NMR spectrum of the DH/RH reaction of 5'-AMP and glycine at pH 2 and 80 °C (in black, overlaid). The mass spectrum of the reaction shows the formation of a dehydrated AMP product (see Figs. S46 and S47). Spiking with 3',5'-cAMP (in red, underlaid) shows that this is not the isomer formed. Additionally, no peaks appear around ~20 ppm (see Fig. S36) indicating that it is also not the 2',3'- cyclised isomer.

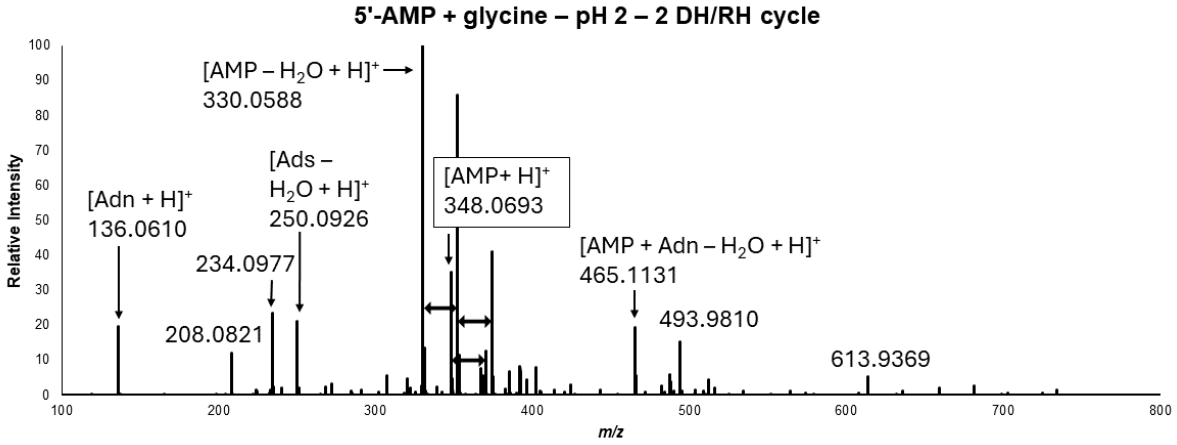


Figure S46: Mass spectrum of the DH/RH reaction of 5'-AMP and glycine at pH 2 and 80 °C for two cycles. Black arrows in between peaks represent singly charged sodium adducts. The peak at 330.0588 corresponds the calculated $[M + H]^+$ molecular ion for a dehydrated AMP isomer within a 3.03 ppm error.

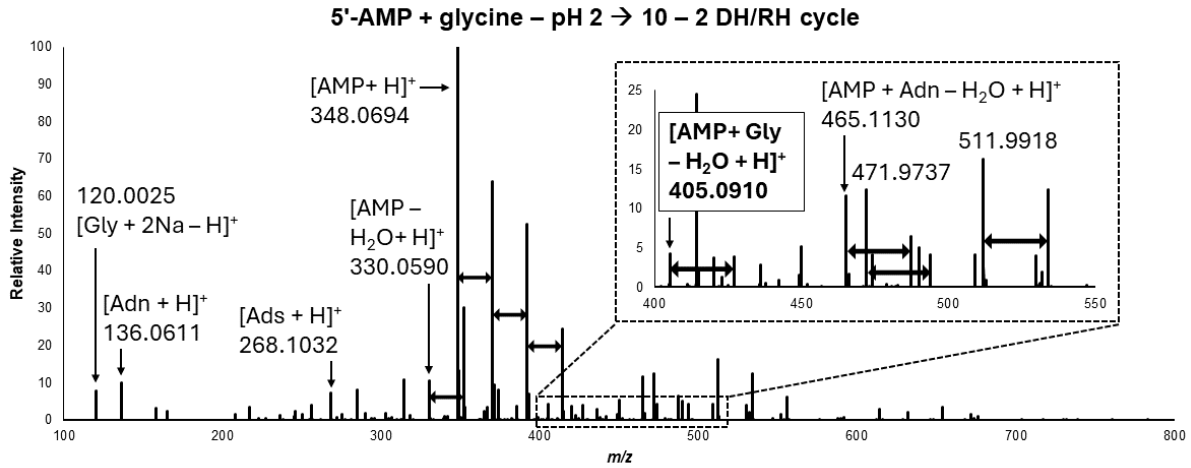


Figure S47: Mass spectrum of the DH/RH reaction of 5'-AMP and glycine at pH 2 for one cycle and then pH 10 for a second cycle. Both cycles were conducted at 80 °C. Black arrows in between peaks represent singly charged sodium adducts. A peak at 405.0910 appears which corresponds to the calculated $[M + H]^+$ molecular ion for an AMP-glycine adduct within a 2.22 ppm error.

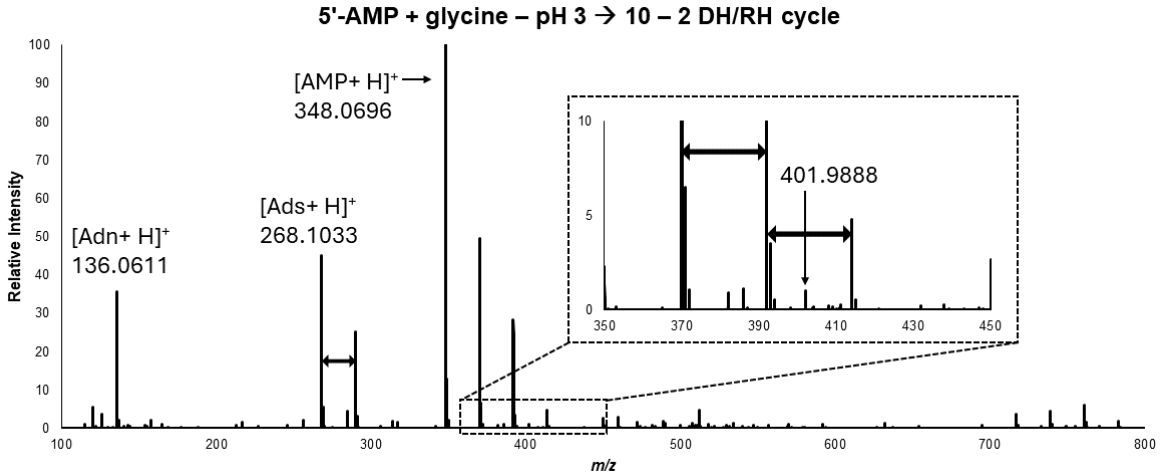


Figure S48: Mass spectrum of the DH/RH reaction of 5'-AMP and glycine at pH 3 for one cycle and pH 10 for a second cycle. Both cycles were conducted at 80 °C. Black arrows in between peaks represent sodium adducts. The dehydrated AMP isomer which forms at pH 2 does not form at pH 3. The callout shows that there is no peak around 405.0919 indicating that a more acidic pH is necessary for an initial cycle to activate AMP (most likely through intramolecular dehydration) or reaction with glycine.

References

1 M. S. Verlander, R. Lohrmann and L. E. Orgel, *J. Mol. Evol.*, 1973, **2**, 303–316.
2 M. Räuchle, G. Leveau and C. Richert, *Eur. J. Org. Chem.*, 2020, **2020**, 6966–6975.
3 F. Neese, *WIREs Comput. Mol. Sci.*, 2022, **12**, e1606.
4 M. Walker, A. J. A. Harvey, A. Sen and C. E. H. Dessent, *J. Phys. Chem. A*, 2013, **117**, 12590–12600.
5 T. Clark, J. Chandrasekhar, G. W. Spitznagel and P. V. R. Schleyer, *J. Comput. Chem.*, 1983, **4**, 294–301.
6 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.
7 Goerigk et al., *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.
8 H. Griesser, P. Tremmel, E. Kervio, C. Pfeffer, U. E. Steiner and C. Richert, *Angew. Chem. Int. Ed.*, 2017, **56**, 1219–1223.
9 M. Jauker, H. Griesser and C. Richert, *Angew. Chem. Int. Ed.*, 2015, **54**, 14564–14569.