Electronic Supplementary Material (ESI) for Chemical Communications.

Prebiotic triose glycolysis promoted by co-catalytic proline and phosphate in neutral water

Álvaro F. Magalhães and Matthew W. Powner

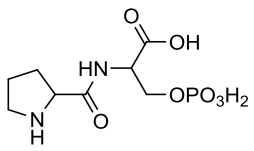
Table of Contents

General

Reagents and solvents were obtained from commercial sources and used without further purification, unless stated otherwise. Formaldehyde was purchased from Sigma-Aldrich as 37 wt. % in H_2O, containing 10-15% methanol as stabilizer. Phosphatase was purchase from Sigma-Aldrich as phosphatase, acid from wheat germ, ≥ 0.4 unit/mg solid. Deionized water was obtained from an *Elga Option 3* purification system. ^{1}H, ^{13}C NMR and ^{31}P spectra were recorded on Bruker NMR spectrometers *AVANCE Neo 700* and *AVANCE III 400* equipped with a Bruker room temperature 5 mm multinuclear gradient probe (700 MHz) and a gradient probe (400 MHz). All ^{13}C and ^{31}P spectra were proton decoupled. All chemical shifts (δ) are reported in parts per million (ppm) relative to residual solvent peaks, and ^{1}H, ^{13}C and ^{31}P chemical shifts relative to TMS were calibrated using the residual solvent peak or an internal standard [residual solvent peaks: (δ) D_2O – 4.79 ppm (^{1}H), CD_3OD – 3.31 ppm (^{1}H) and 49.00 ppm (^{13}C), $CDCl_3$ – 7.26 ppm (^{1}H) and 77.16 ppm (^{13}C)]. Nuclear assignments were made using 2D NMR homo and heteronuclear correlation spectroscopy (^{1}H – ^{13}C HSQC; ^{1}H – ^{13}C HMBC and ^{1}H – ^{31}P HMBC). Where noted, solvent suppression pulse sequence with pre-saturation and spoil gradients were used to obtain ^{1}H NMR spectra (noesygppr1d, Bruker) and ^{1}H – ^{13}C HMBC NMR spectra (hmbcgplpndprqf, Bruker). Coupling constants (*J*) are reported in Hertz (Hz). Spin multiplicities are indicated by symbols: s (singlet); d (doublet); t (triplet); q (quartet); sx (sextuplet); m (multiplet); obs (obscured); app (apparent), or a combination of these. Diastereotopic geminal (AB) spin systems coupled to one or two additional nuclei are reported as ABX and ABXY, respectively. NMR data are reported as follows: chemical shift (number of protons, multiplicity, coupling constants (*J*), nuclear assignment). Spectra were recorded at 298 K. Infrared spectra (IR) were recorded on a Shimadzu IR Tracer 100 FT-IR spectrometer as a solid or neat oil/liquid. Absorption maxima are reported in wavenumber (cm^{-1}). Mass spectra were recorded on a Waters LCT Premier QTOF connected to a Waters Autosampler Manager 2777C, Thermo Finnigan MAT900, and an Agilent LC connected to an Agilent 6510 QTOF mass spectrometer at the Department of Chemistry, University College London. Solution pH values were measured using a Mettler Toledo Seven Compact pH meter with a Mettler Toledo InLab semi-micro pH probe, or a Corning 430 pH meter with a Fisherbrand FB68801 semi-micro pH probe. The readings for H_2O and H_2O/D_2O (9:1) solutions are reported uncorrected. Flash column chromatography was performed on a Biotage Isolera One using Kinesis TELOS cartridges. Chiral HPLC analysis was performed in Agilent 1260 Infinity LC system using an YMC Chiral Amylose-SA Classical Analytical S-5 column (250 × 4.6 mm).

Synthesis of catalysts

*Synthesis of H-Pro-SOP-OH (**10**)*



10

Adapted from Raines[1], Auriel[2] and Fields[3]. Fmoc-Pro-OPfp (612 mg, 0.41 mmol) and *O*-phospho-L-serine (H-SOP-OH, 150 mg, 0.81 mmol) were dissolved in anhydrous methanol (9 mL), under a N_2 atmosphere, and diisopropylamine (458 µL, 3.24 mmol) was added. The reaction mixture was stirred at room temperature for 3 days. After that time, the solvent was removed, and the residue was purified by flash chromatography DCM/MeOH/H_2O (100:0:0 → 0:100:0 → 0:0:100), yielding the di-diisopropylamide salt as white powder, which was used without further purification. Fmoc-Pro-SOP-OH (150 mg, 0.21 mmol) was dissolved in anhydrous DCM (5.4 mL) and piperidine (5.4 mL, 0.55 mmol) was added. The reaction mixture was stirred at room temperature for 1 h and, after that time, the solvent was removed in vacuo. The residue was suspended in H_2O and the mixture centrifuged. The aqueous solution was partially evaporated and 0.5 M sodium perchlorate in acetone (5 mL) was added. The precipitate was isolated by centrifugation, dried in vacuo and then dissolved in water and eluted through Dowex™ (50W× 8 H^+-form, 2 g). The solution was lyophilised yielding an off-white powder (50 mg, 0.18 mmol, 83%). ^{1}H NMR (700 MHz, D_2O): 2.05 (2H, m), 2.14 (1H, app sx, *J* = 6.9 Hz), 2.47 (1H, dq, *J* = 7.2, 13.2 Hz), 3.35 – 3.45 (2H, ABXY, *J* = 7.2, 11.7 Hz), 4.15 – 4.31 (1H, ABX, *J* = 10.9, 5.6, 3.3 Hz), 4.45 (1H, dd, *J* = 8.5, 6.7 Hz), 4.73 (1H, obs). ^{13}C NMR (176 MHz, D_2O): 24.3, 30.2, 47.2, 54.2 (d, *J* = 7.9 Hz), 60.3, 65.1 (d, *J* = 4.3 Hz), 170.3, 172.7. ^{31}P NMR (284 MHz, D_2O): 0.25 (s, P). HRMS (ESI+) calculated for $C_8H_{15}N_2O_7P+H^+$: 283.0690; found 283.0691. IR (cm^{-1}): 2944 (br), 1722 (m), 1672 (s), 1548 (m). 1456 (m), 1383 (w), 1192 (w), 1138 (w), 959 (br). Melting point (ºC): 169 (decomp.).

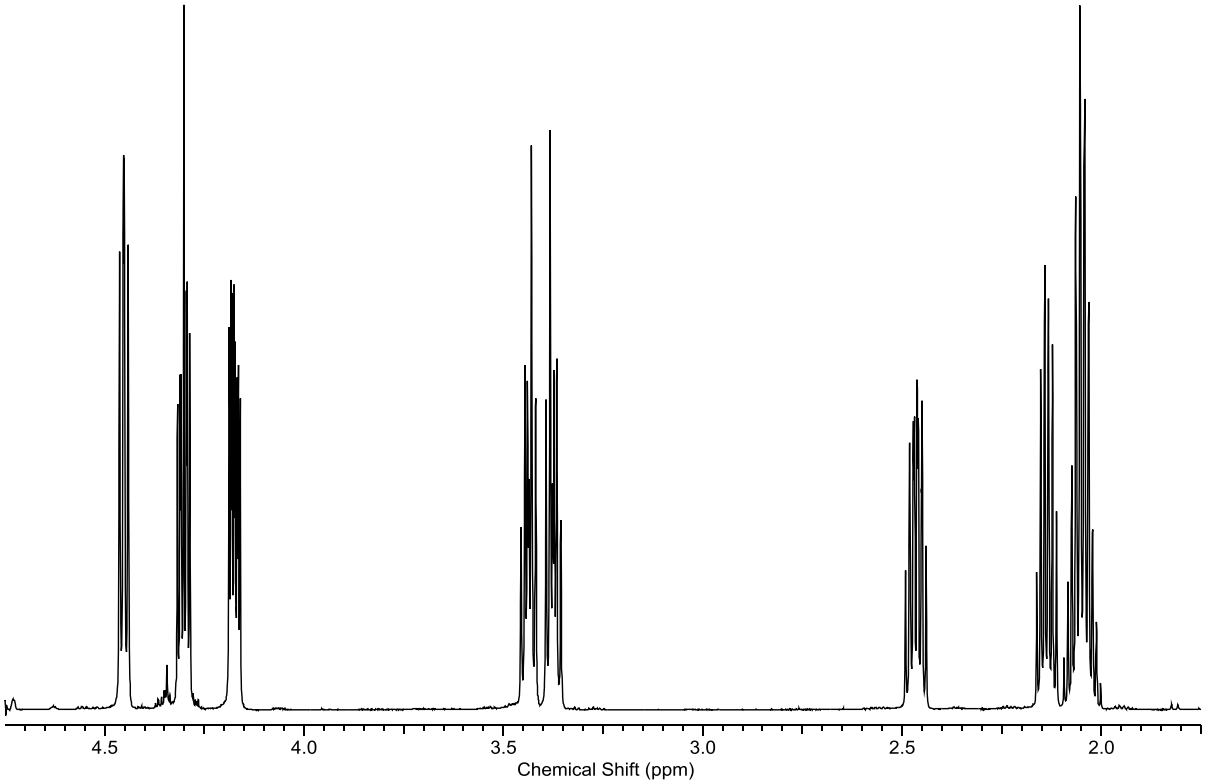


Figure S1. ^{1}H NMR (700 MHz, D_2O, noesygppr1d, 1.75 – 4.75 ppm) spectrum to show H-Pro-SOP-OH (**10**).

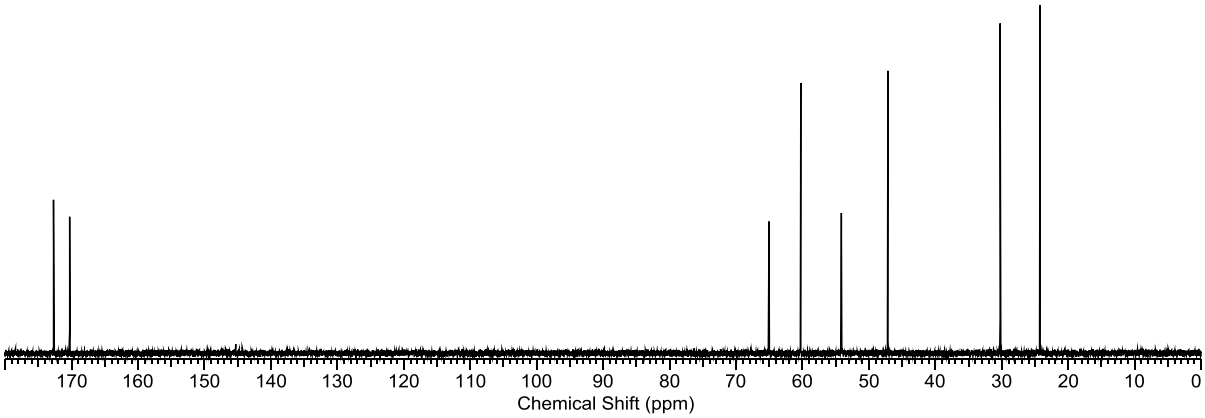


Figure S2. ^{13}C NMR (176 MHz, D_2O) spectrum to show H-Pro-SOP-OH (**10**).

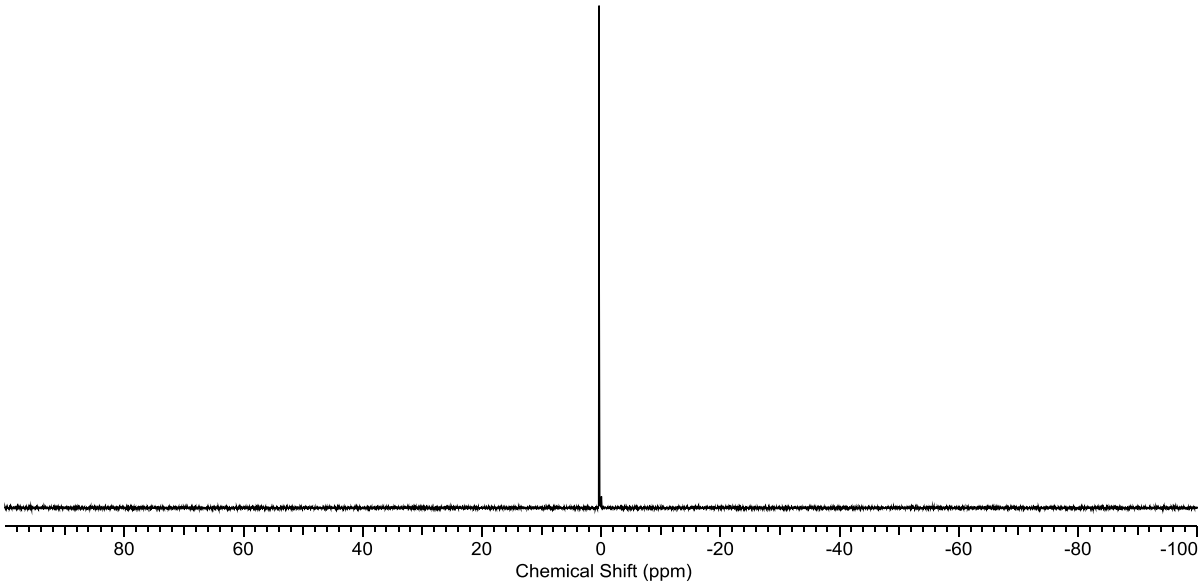
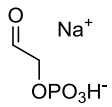


Figure S3. ^{31}P NMR (284 MHz, D_2O) spectrum to show H-Pro-SOP-OH (**10**).

Synthesis of glycolaldehyde-phosphate (6)



6

Modified from Awanda and Dedon[4] and Wong *et al*[5]. Adenosine-5'-monophosphate (10.0 g, 28.8 mmol) was dissolved in H_2O (195 mL) and Dowex™ 50W×8 (H^+-form; 95 mL) resin was added and refluxed for 30 min. The resin was removed by filtration and the volume adjusted to 475 mL with H_2O. The solution was cooled to 0ºC, sodium periodate (24.7 g, 115.2 mmol) was added and the solution was then adjusted to pH 6. The mixture was allowed to stir at room temperature for 1 h and then ethylene glycol (1.76 mL, 31.7 mmol) was added. The solution was adjusted to pH 7, cooled to 0ºC and then barium chloride (35.0 g, 144 mmol) was added. After 10 min., the mixture was filtered. Ethanol (1.9 L) was added to the filtrate. The resultant mixture was centrifuged and the white gelatinous residue was washed with ice-cold H_2O/EtOH (1:4). Then the solids were shaken with Dowex™ 50W×8 (H^+ form) resin in H_2O (1 L) until the white solids had full dissolved. The suspension was poured into a Dowex™ (50W×8; H^+ -form) column and the solution eluted. The column was washed with H_2O (100 mL) and the solution was concentrated to half its volume in vacuo. The resultant solution was adjusted to pH 2.7 and then evaporate to half its volume in vacuo. The solution was then adjusted to pH 3 and the volume reduced to approximately 5 mL. The solution was adjusted to pH 4 and the volume adjusted to 10 mL with H_2O. The aqueous solution of sodium glycolaldehyde-phosphate (**6**, 10 mL, 2.03 M, ~ 90% purity with ~ 3% of formic acid, 60% yield) was stored at -20ºC and used without further purification. ^{1}H NMR (400 MHz, D_2O): 5.17 (1H, t, *J* = 4.9 Hz, C***H***$(OH)_2$), 3.80 (2H, dd, *J* = 6.9, 4.8 Hz, C***H***$_2OPO_3$). ^{13}C NMR (101 MHz, D_2O): 89.2 (d, *J* = 8.8 Hz, ***C***$H(OH)_2$, 68.0 (d, *J* = 5.1 Hz, ***C***H_2OPO_3). ^{31}P NMR (162 MHz, D_2O): 1.39 (s, P); All data agrees with published values[6].

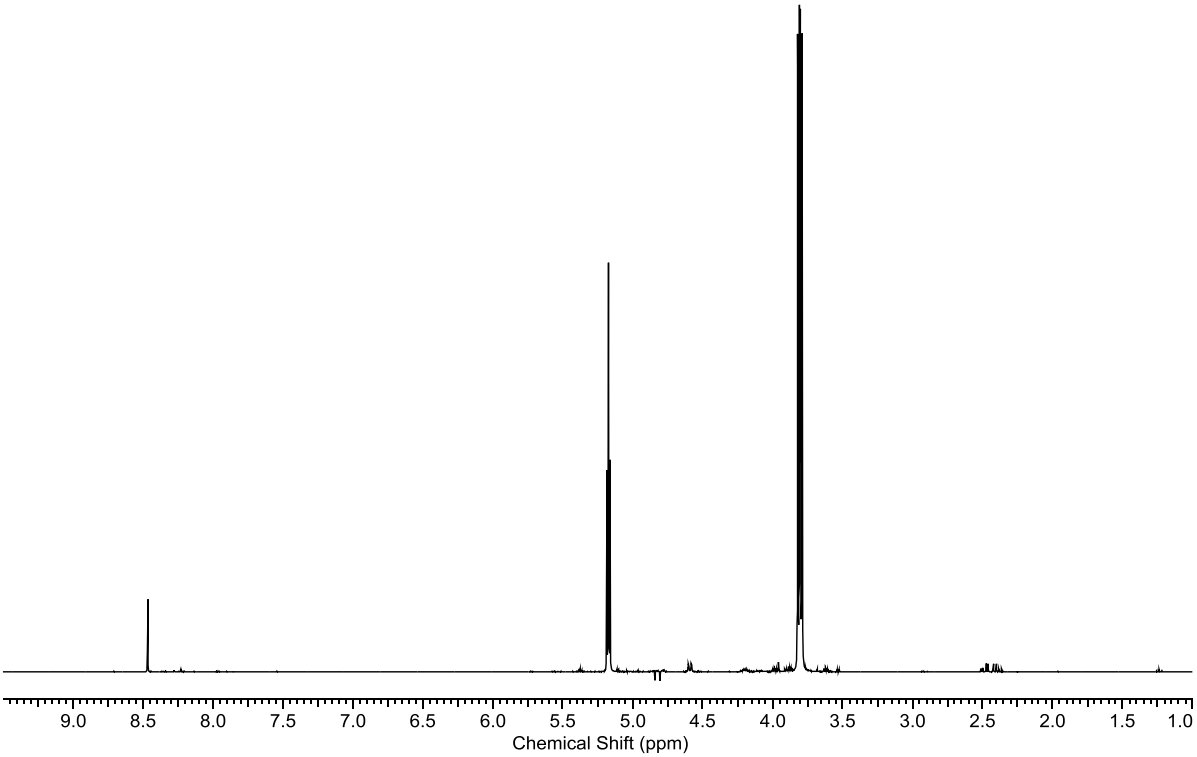


Figure S4. ^{1}H NMR (400 MHz, D_2O, noesygppr1d, 1.00 - 9.50 ppm) spectrum to show glycolaldehyde-phosphate (**6**).

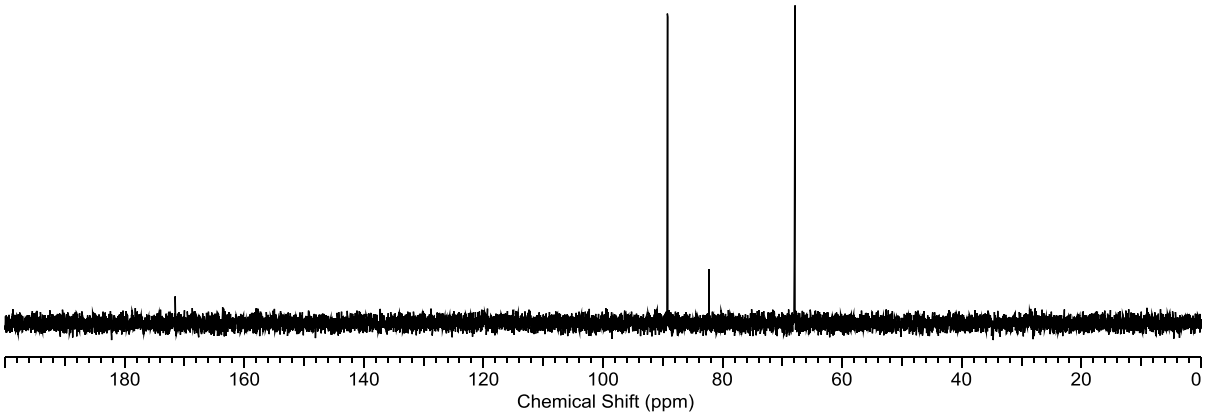


Figure S5. ^{13}C NMR (101 MHz, D_2O) spectrum to show glycolaldehyde-phosphate (**6**).

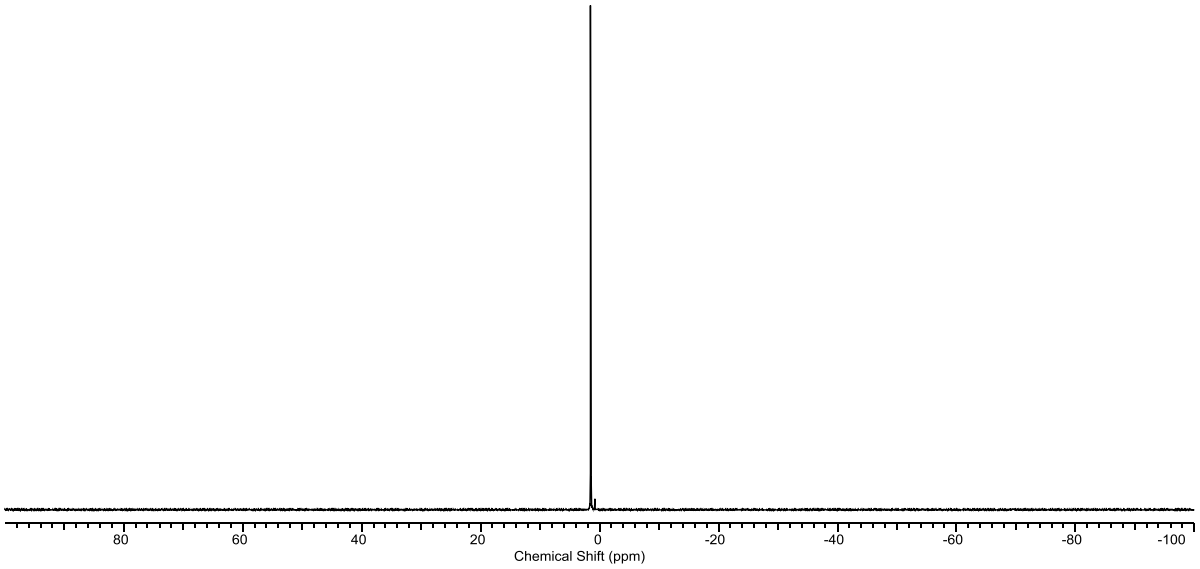
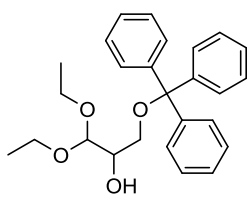


Figure S6. ^{31}P NMR (162 MHz, D_2O) spectrum to show glycolaldehyde-phosphate (**6**).

Synthesis of glyceraldehyde 2-phosphate (7)

*Synthesis of 1,1-diethoxy-3-(trityloxy)propan-2-ol (**11**)*



11

Adapted from Angrick and Rewicki[7]. *rac*-Glyceraldehyde diethyl acetal (2 g, 12.5 mmol) was dissolved in pyridine (7.5 mL) and trityl chloride (3.45 g, 12.5 mmol) was added. The mixture was stirred at 70ºC for 2 h and, after that time, poured into H_2O (20 mL) and extracted with diethyl ether (2 × 20 mL). The combined ethereal phases were washed with citric acid (0.1 M, 20 mL), brine (20 mL) and sat. $NaHCO_3$ (20 mL). The organic layer was dried over $MgSO_4$, filtered and the solvent removed in vacuo, yielding 1,1-diethoxy-3-(trityloxy)propan-2-ol as a white powder (3.5 g, 8.6 mmol, 69%). ^{1}H NMR (700 MHz, $CDCl_3$): 1.09 (3H, t, *J* = 7.0 Hz, $HC(OCH_2C\boldsymbol{H_3})_2$), 1.21 (3H, t, *J* = 7.0 Hz, $HC(OCH_2C\boldsymbol{H_3})_2$), 2.39 (1H, d, *J* = 3.9 Hz, O***H***), 3.21 – 3.34 (2H, $C\boldsymbol{H_2}OC(C_6H_5)_3$), 3.45 – 3.49 (1H, m, $HC(OCH\boldsymbol{H}CH_3)_2$), 3.54 – 3.59 (1H, m, $HC(OCH\boldsymbol{H}CH_3)_2$), 3.64 – 3.68 (1H, m, $HC(OCH\boldsymbol{H}CH_3)_2$), 3.733 – 3.77 (2H, m, $HC(OCH\boldsymbol{H}CH_3)_2$ and C***H***OH), 4.61 (1H, d, *J* = 5.8 Hz, ***H***$C(OCHHCH_3)_2$), 7.22 – 7.47 (15H, m, Ar). ^{13}C NMR (176 MHz, $CDCl_3$): 15.4 ($HC(OCH_2\boldsymbol{C}H_3)_2$), 15.5 ($HC(OCH_2\boldsymbol{C}H_3)_2$), 63.2 ($HC(O\boldsymbol{C}H_2CH_3)_2$), 63.7 ($HC(O\boldsymbol{C}H_2CH_3)_2$), 63.8 ($\boldsymbol{C}H_2OC(C_6H_5)_3$), 71.6 (***C***HOH), 86.7 ($\boldsymbol{C}(C_6H_5)_3$), 102.5 ($H\boldsymbol{C}(OCH_2CH_3)_2$), 127.1 (Ar), 127.93 (Ar), 128.8 (Ar), 144.1 (Ar). HRMS (ESI^-): m/z calcd for $C_{26}H_{30}O_4^-$: 405.2071; found 405.2070. IR (solid, cm^{-1}): 3504 (br), 2975 (m), 2927 (m), 2885 (m), 1489 (m), 1448 (m), 1104 (s), 1050 (s). 771 (s). Melting point (ºC): 79 – 81. All data agrees with published values.[7]

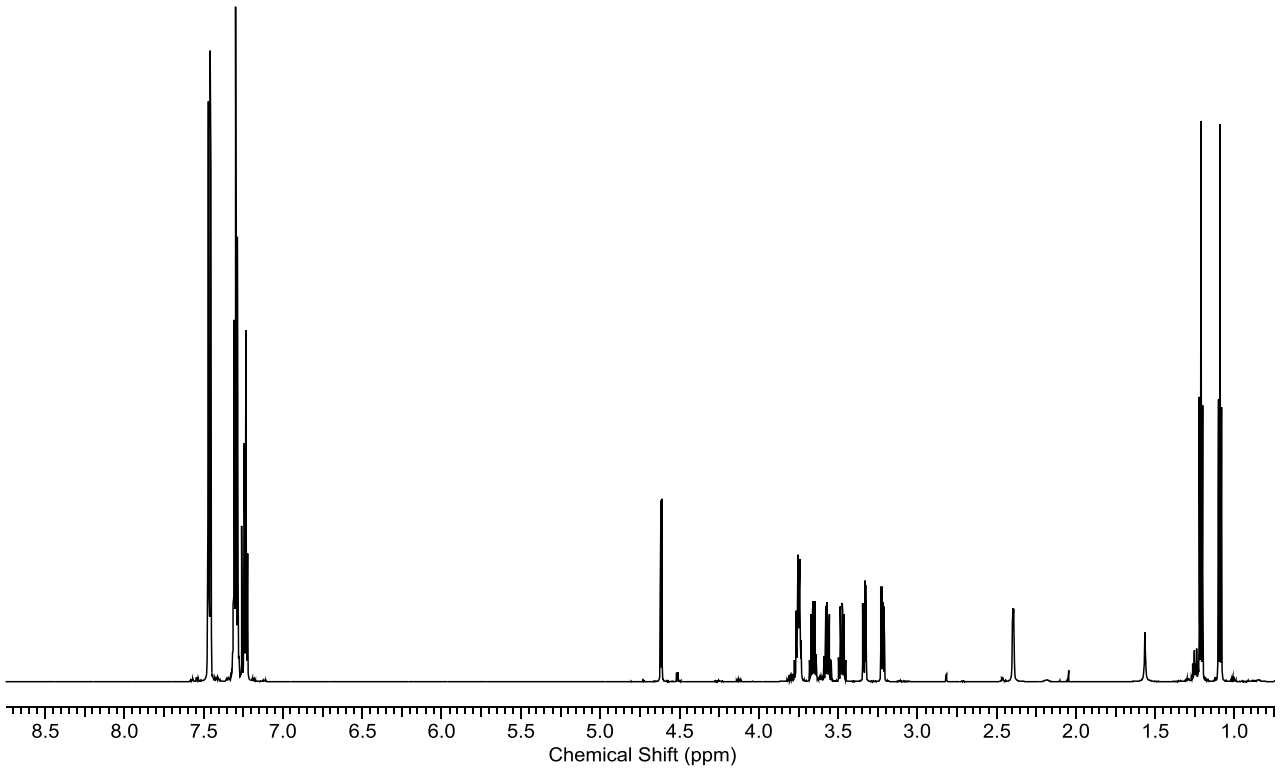


Figure S7. ^{1}H NMR (700 MHz, $CDCl_3$, 0.75 - 8.75 ppm) spectrum to show 1,1-diethoxy-3-(trityloxy)propan-2-ol (**11**).

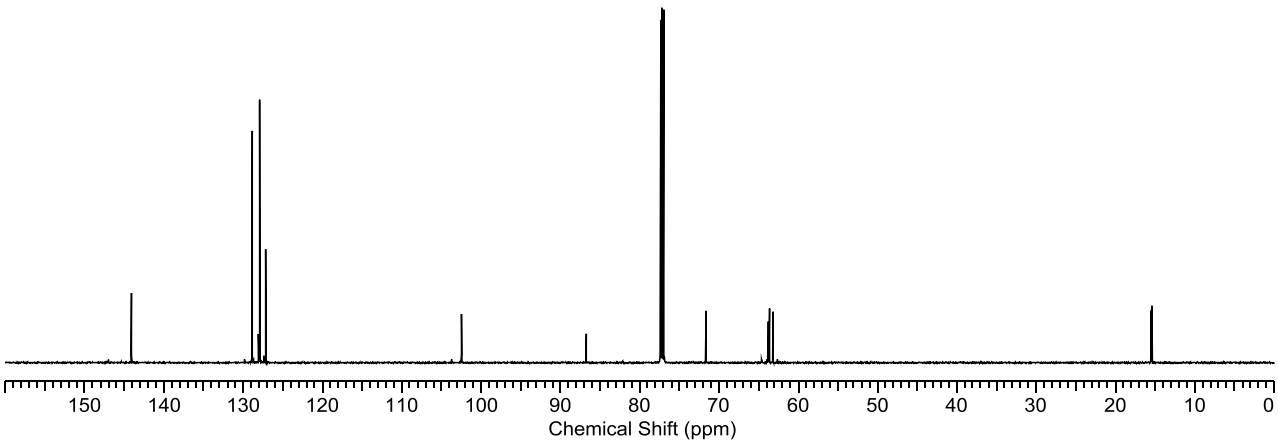
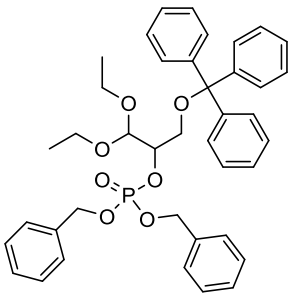


Figure S8. ^{13}C NMR (176 MHz, $CDCl_3$, 0 - 160 ppm) spectrum to show 1,1-diethoxy-3-(trityloxy)propan-2-ol (**11**).

*Synthesis of dibenzyl (1,1-diethoxy-3-(trityloxy)propan-2-yl) phosphate (**12**)*



12

Adapted from Singh *et al*[8]. 1,1-Diethoxy-3-(trityloxy)propan-2-ol (1.1 g, 2.71 mmol) and 4,5-dicyanoimidazole (390 mg, 4.06 mmol) were dissolved in anhydrous acetonitrile (2.5 mL) and the solvent removed in vacuo (on a Schlenk line). The residue was then dissolved in anhydrous acetonitrile (6 mL) and dibenzyl *N,N*-diisopropylphosphoramidite (1.4 mL, 4.06 mmol) was added. After 15 min at room temperature, the mixture was cooled to -20ºC and 4-chloro-2-methylphenoxyacetic acid (934 mg, 5.41 mmol) was added. After 10 min, diethyl ether (5 mL) and n-hexane (5 mL) was added and the precipitate removed by filtration. The filtrate was then evaporated and the residue purified by flash chromatography eluting with petroleum ether/ethyl acetate (100:0 to 0:100), yielding dibenzyl (1,1-diethoxy-3-(trityloxy)propan-2-yl) phosphate as a colourless oil (1.46 g, 2.19 mmol, 81%). ^{1}H NMR (700 MHz, $CDCl_3$): 1.06 (3H, t, *J* = 7.0 Hz; $HC(OCH_2C\boldsymbol{H_3})_2$), 1.15 (3H, t, *J* = 7.0 Hz; $HC(OCH_2C\boldsymbol{H_3})_2$), 3.32 (1H, dd, *J* = 10.5, 5.3 Hz; $CHC\boldsymbol{H}HOC(C_6H_5)_3$), 3.45 – 3.49 (1H, m, $HC(OC\boldsymbol{H}HCH_3)_2$), 3.51 – 3.57 (2H, m, $CHC\boldsymbol{H}HOC(C_6H_5)_3$) and $HC(OC\boldsymbol{H}HCH_3)_2$), 3.59 – 3.64 (1H, m, $HC(OC\boldsymbol{H}HCH_3)_2$), 3.68 – 3.72 (1H, m, $HC(OC\boldsymbol{H}HCH_3)_2$), 4.58 – 4.61 (1H, m; $C\boldsymbol{H}CH_2OC(C_6H_5)_3$), 4.77 (1 H, d, *J* = 5.7 Hz; $\boldsymbol{H}C(OCH_2CH_3)_2$), 5.02 (2H, dd, *J* = 7.2, 3.8 Hz, $POC\boldsymbol{H_2}Ph$), 5.07 (2H, dd, *J* = 7.3, 3.9 Hz, $POC\boldsymbol{H_2}Ph$), 7.19 - 7.46 (25H, m; Ar). ^{13}C NMR (176 MHz, $CDCl_3$): 15.2 ($HC(OCH_2\boldsymbol{C}H_3)_2$), 15.4 ($HC(OCH_2\boldsymbol{C}H_3)_2$), 62.9 ($HC(O\boldsymbol{C}H_2CH_3)_2$), 63.0 ($CH\boldsymbol{C}H_2OC(C_6H_5)_3$, 63.1 ($HC(O\boldsymbol{C}H_2CH_3)_2$), 69.2 (t, *J* = 5.8 Hz, $PO\boldsymbol{C}H_2Ph$), 78.0 (d, *J* = 6.5 Hz, $\boldsymbol{C}HCH_2OC(C_6H_5)_3$), 86.9 ($CHCH_2O\boldsymbol{C}(C_6H_5)_3$), 101.1 (d, *J* = 5.8 Hz, $H\boldsymbol{C}(OCH_2CH_3)_2$), 127.1 (Ar), 127.8 (Ar), 127.9 (Ar), 128.0 (Ar), 128.4 (d, *J* = 2.2 Hz, Ar)), 128.6 (Ar), 128.6 (Ar), 128.9 (Ar), 136.2 (dd, *J* = 15.1, 7.9 Hz, Ar), 143.9 (3C, Ar). ^{31}P NMR (284 MHz, $CDCl_3$): -1.48 (s, P). All data agrees with published values.[9]

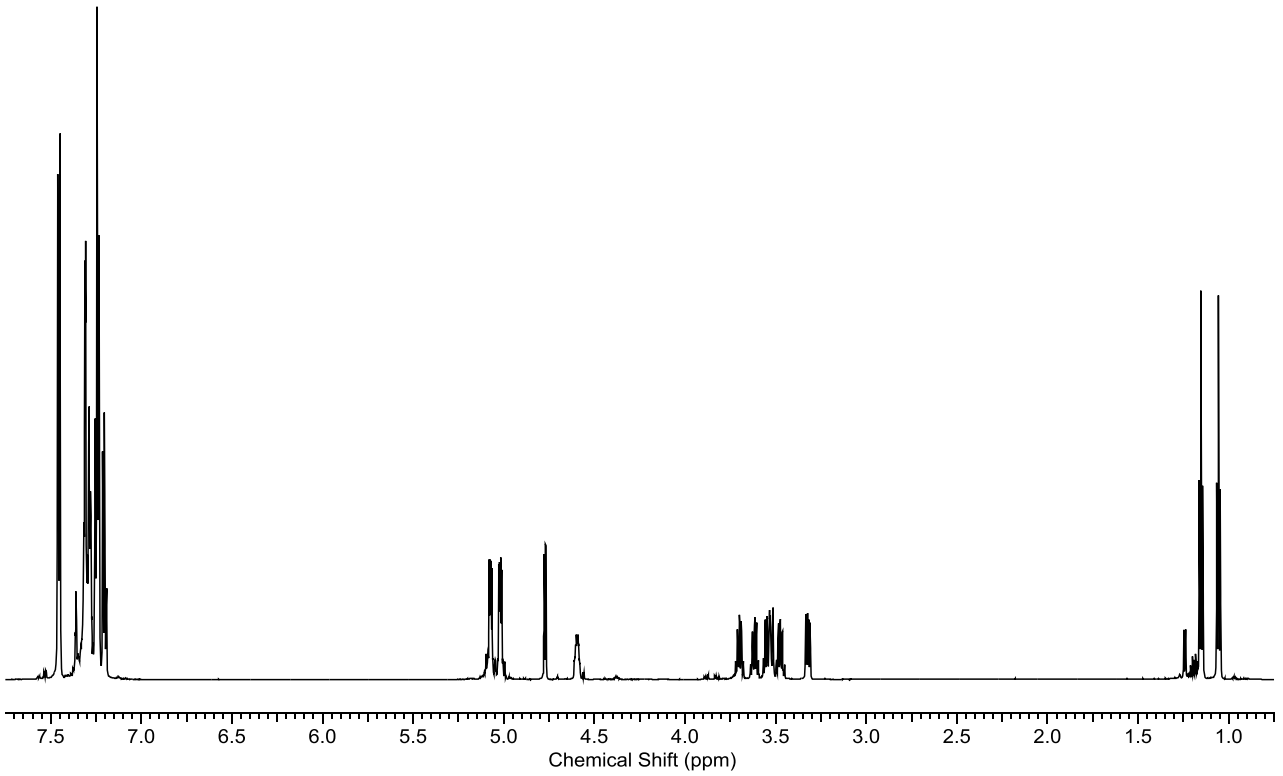


Figure S9. ^{1}H NMR (700 MHz, $CDCl_3$, 0.75 - 7.75 ppm) spectrum to show dibenzyl (1,1-diethoxy-3-(trityloxy)propan-2-yl) phosphate (**12**).

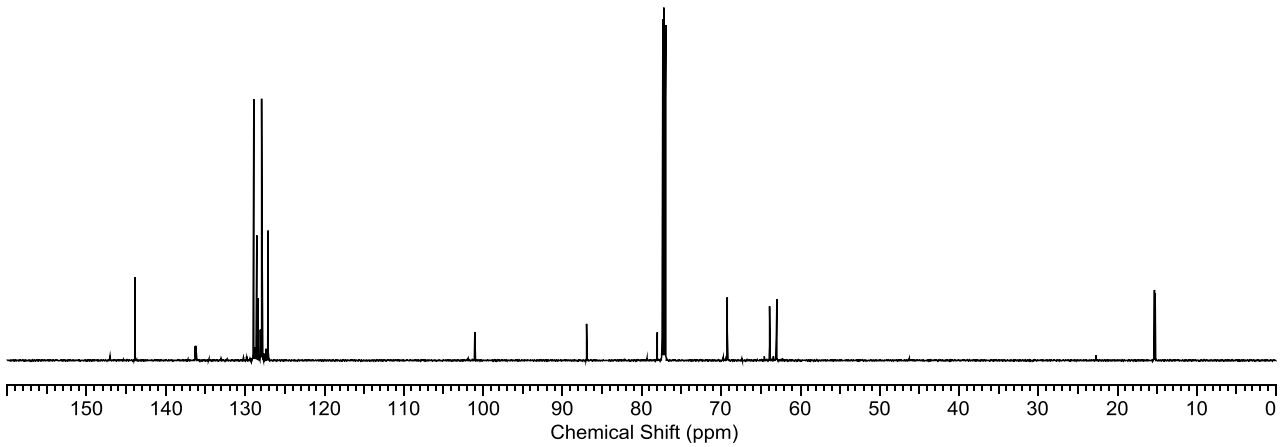


Figure S10. ^{13}C NMR (176 MHz, $CDCl_3$, 0 - 160 ppm) spectrum to show dibenzyl (1,1-diethoxy-3-(trityloxy)propan-2-yl) phosphate (**12**).

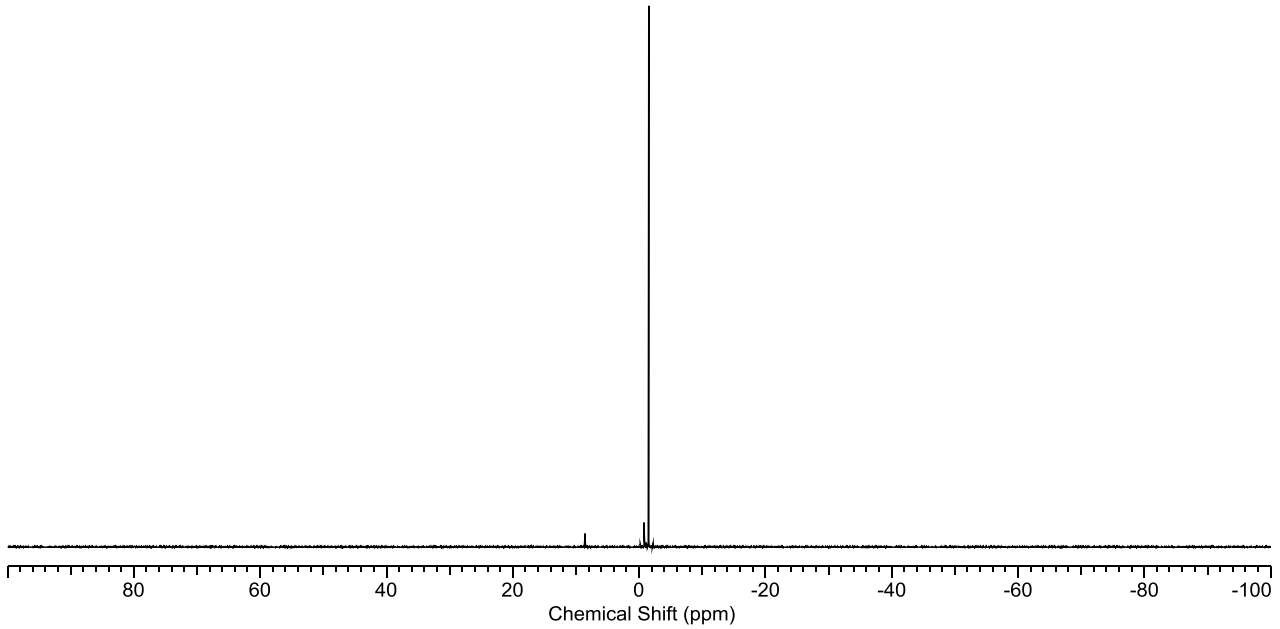
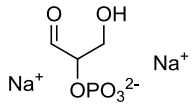


Figure S11. ^{31}P NMR (284 MHz, $CDCl_3$, -100 - 100 ppm) spectrum to show dibenzyl (1,1-diethoxy-3-(trityloxy)propan-2-yl) phosphate (**12**).

*Synthesis of glyceraldehyde 2-phosphate sodium salt (**7**)*



7

Adapted from Anastasi *et al*[9]. Dibenzyl (1,1-diethoxy-3-(trityloxy)propan-2-yl) phosphate (514 mg, 0.77 mmol) and Pd/C (10 %, 40 mg) were suspended in dioxane/H_2O (1:1, 10 mL). The mixture was then degassed by freeze/thaw and stirred under a H_2 atmosphere for 18 h at room temperature. After that time, H_2O (5 mL) was added to the mixture and then it was filtered through Celite® and the Celite® washed with dioxane/H_2O (1:1, 30 mL). The filtrate was evaporated in vacuo and the residue was dissolved in CH_2Cl_2/H_2O (2:1, 15 mL). The layers were separated, the organic layer was washed with H_2O (3 × 10 mL) and the combined aqueous phases were washed with CH_2Cl_2 (2 × 10 mL). The aqueous solution was then evaporated to dryness in vacuo, yielding a yellow oil (164 mg). The oil was dissolved in D_2O (1.6 mL) and eluted through a 0.2 µm filter. The filtrate was adjusted to pD 2.2 with NaOD (1 M) and heated at 50ºC for 2 days. After that time, the solution was adjusted to pD 6, lyophilised and the solid eluted through Dowex™ (Na^+-form, 2 g), yielding glyceraldehyde 2-phosphate (**7**) as an off-white solid (71 mg, 0.24 mmol, 31% yield, 72.5% w/w with 5% of **9** and 20% sodium phosphate). ^{1}H NMR (700 MHz, D_2O): 5.00 (1H, d, *J* = 3.8 Hz, C***H***$(OH)_2$), 4.03 (1H, m, ***H***$COPO_3$), 3.73 (1H, ABX, *J* = 4.3, 12.1 Hz, CH***H***OH), 3.69 (1H, ABX, *J* = 5.7, 12.1 Hz, CH***H***OH). ^{13}C NMR (101 MHz, D_2O): 62.1 (***C***HHOH), 77.7 (H***C***OPO_3), 90.0 (***C***H$(OH)_2$). ^{31}P NMR (284 MHz, D_2O): 2.78 (s, P). All data agrees with published values.[9]

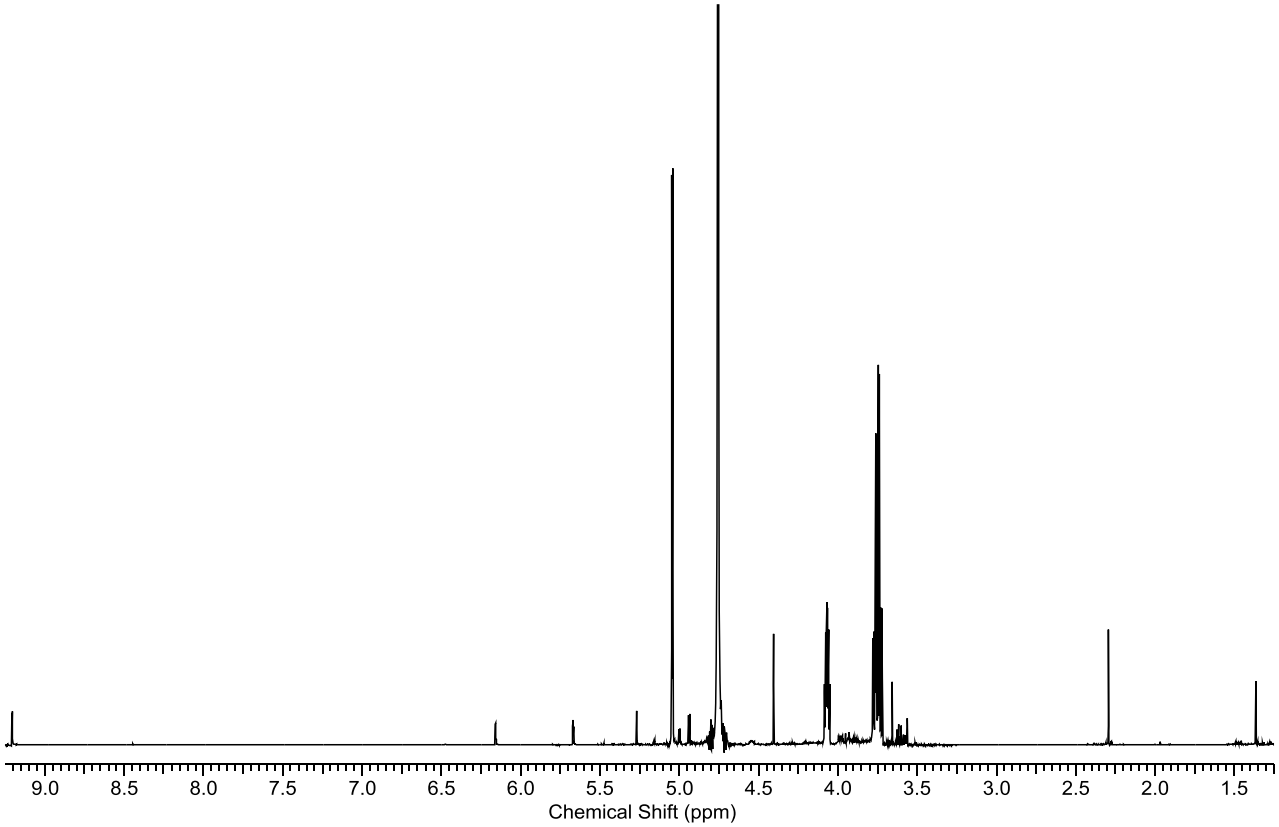


Figure S12. ^{1}H NMR (700 MHz, D_2O, noesygppr1d, 1.25 - 9.25 ppm) spectrum to show glyceraldehyde-2-phosphate (**7**).

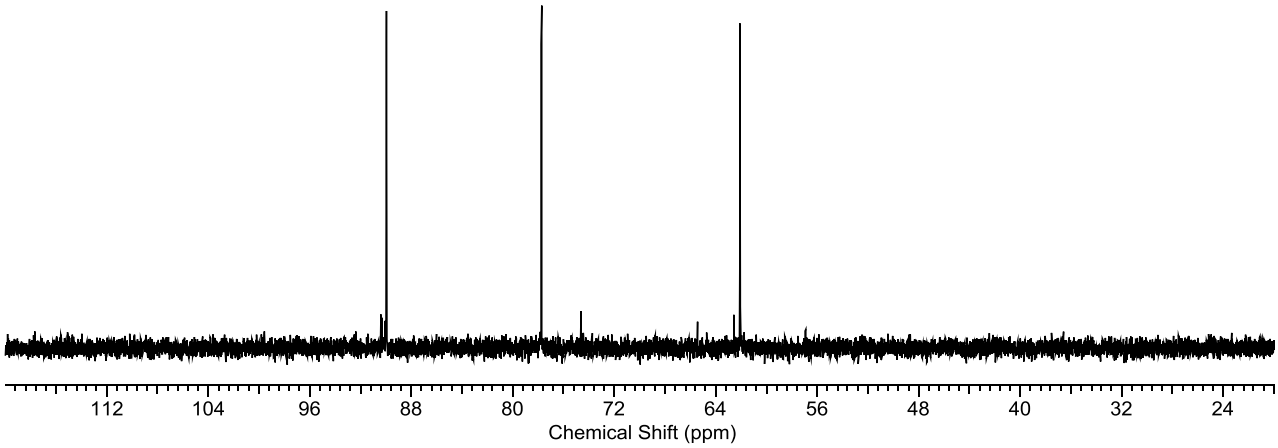


Figure S13. ^{13}C NMR (101 MHz, D_2O, 20 – 120 ppm) spectrum to show glyceraldehyde-2-phosphate (**7**).

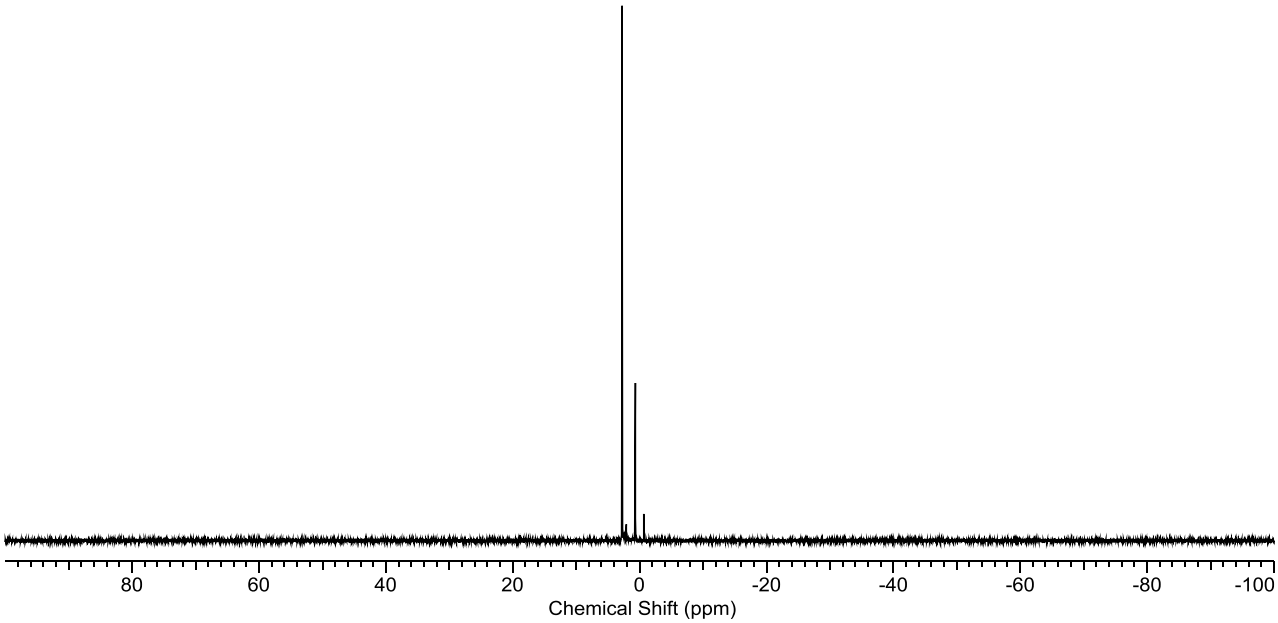
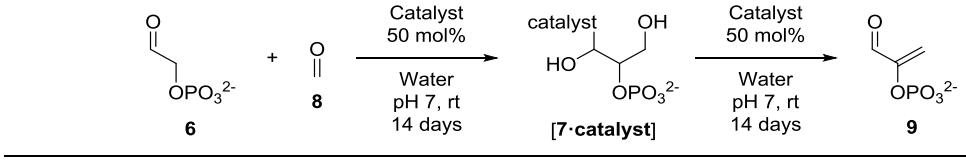


Figure S14. ^{31}P NMR (284 MHz, D_2O, -100 -100 ppm) spectrum to show glyceraldehyde-2-phosphate (**7**).

Aldol catalysts screening

Catalyst screen in water

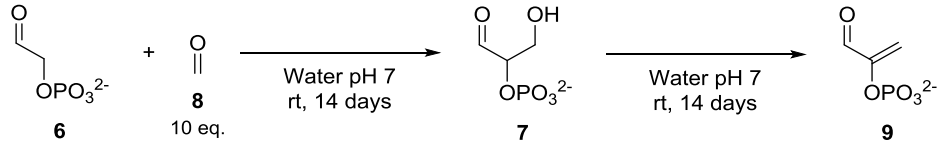
Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and catalyst (50 mol%) were dissolved in H_2O (700 µL). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and dimethyl sulfone (MSM, 18.8 µL, 18.75 µmol) were then added, and the solution pH was adjusted, as necessary, to pH 7 with NaOH/HCl (1 M). To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution. NMR spectra were then acquired regularly over 14 d. Yields for [**7·catalyst**] and **9** observed after 14 d for each catalyst are given in Figure S15.



control	Glycine	Glycine amide	Glycine methyl ester	Glycine nitrile
No catalyst				
7: nd **9**: nd	[**7·Gly**]: nd **9**: nd	[**7·Gly-NH$_2$**]: nd **9**: nd	[**7·Gly-OMe**]: nd **9**: nd	[**7·Gly-CN**]: nd **9**: nd
Sarcosine	**L-Proline**	**L-Proline amide**	L-Proline methyl ester*	L-Proline nitrile
[**7·Sar**]: 8% **9**: nd	[**7·Pro**]:10% **9**: trace	[**7·Pro-NH$_2$**]: 7% **9**: 4%	[**7·Pro-OMe**]: 7% **9**: trace	[**7·Pro-CN**]: 6% **9**: nd
Pyrrolidine	L-Lysine	Dpr-OH	L-Glutamic acid	
[**7·Pyrr**]: nd **9**: nd	[**7·Lys**]: 9% **9**: trace	[**7·Dpr**]**:** nd **9:** nd	[**7·Glu**]: nd **9**: nd	
L-Histidine	Ac-Lys-OH	Ac-Dpr-OH	H-L-Thz-OH	
[**7·His**]: nd **9**: nd	[**7·Ac-Lys**]: nd **9**: nd	[**7·Ac-Dpr**]: nd **9**: nd	[**7·Thz**]: nd **9**: nd	
Pro-*O*-Phospho-Ser-OH	Uridine 5'-monophosphate	Pro-Pro-Asp-NH$_2$	Pro-Pro-Glu-NH$_2$	
[**7·Pro-SOP**]: 4% **9**: trace	[**7·UMP**]: nd **9**: nd	[**7·Pro-Pro-Asp-NH$_2$**]: 13% **9**: trace	[**7·Pro-Pro-Glu-NH$_2$**]: 13% **9**: 3%	

Figure S15. Yields for the formation of catalyst-glyceraldehyde-2-phosphate hemiaminal [**7·catalyst**] and phosphoenol pyruvaldehyde (**9**) upon reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 1 M) with the specified catalysis (50 mol%) in H_2O at pH 7, with MSM as internal NMR standard, after 14 d. nd = not detected. *Hydrolyses to Pro-OH (> 95%) was observed.

Water (without an amine catalyst)



Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol), formaldehyde (**8**, 55.9 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were dissolved in H_2O (700 µL). The solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution. NMR spectra were then acquired regularly over 14 d. The evolution of the reaction can be observed in Figure S16.

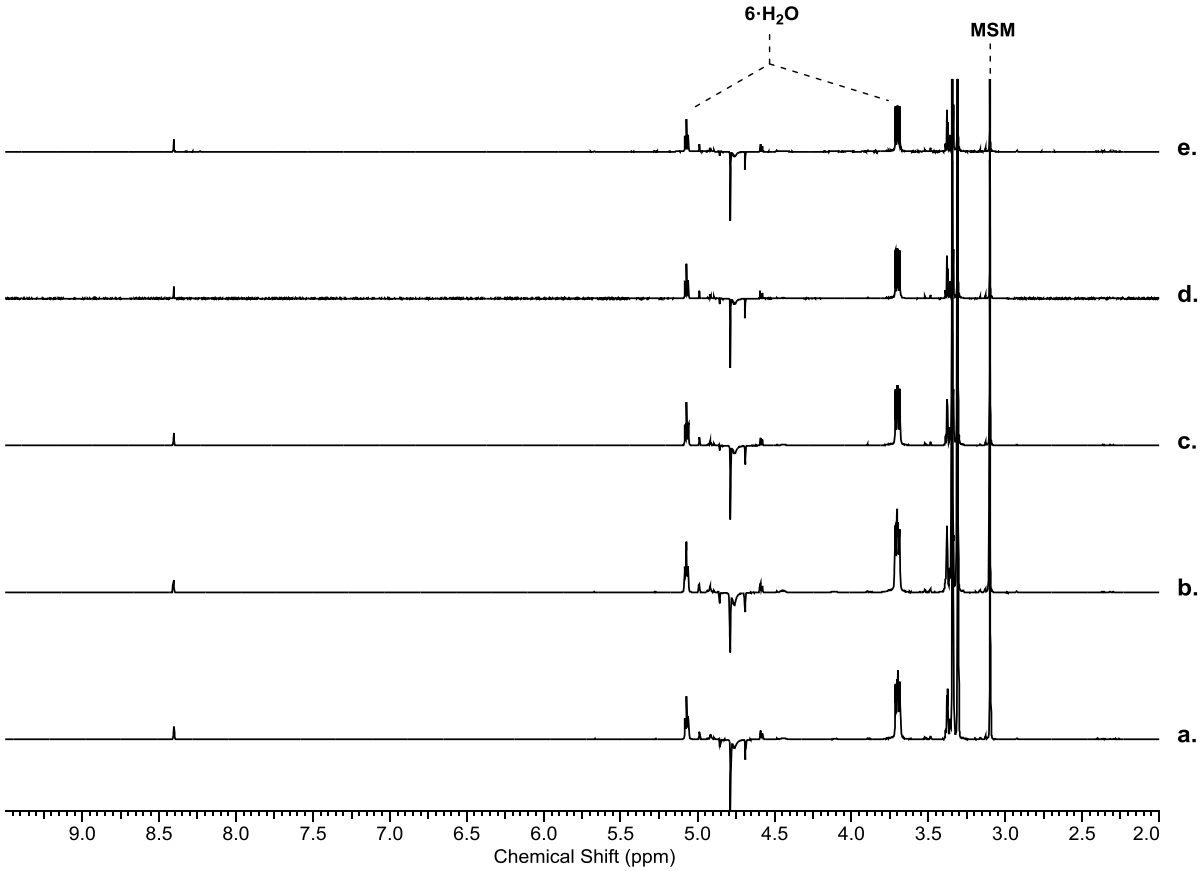
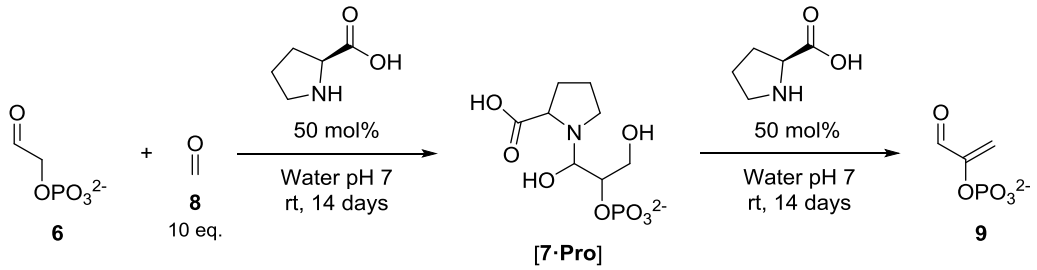


Figure S16. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 2.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 1 M) in H_2O at pH 7, with MSM (25 mM) as an internal NMR standard after: **a.** 30 min; **b.** 1 d; **c.** 3 d; **d.** 7 d; **e.** 14 d.

L-Proline in water



Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 37.5 µmol, 50 mol%) were dissolved in H_2O (700 µL). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution. NMR spectra were then acquired regularly over 14 d. The evolution of the reaction can be observed in Figure S17.

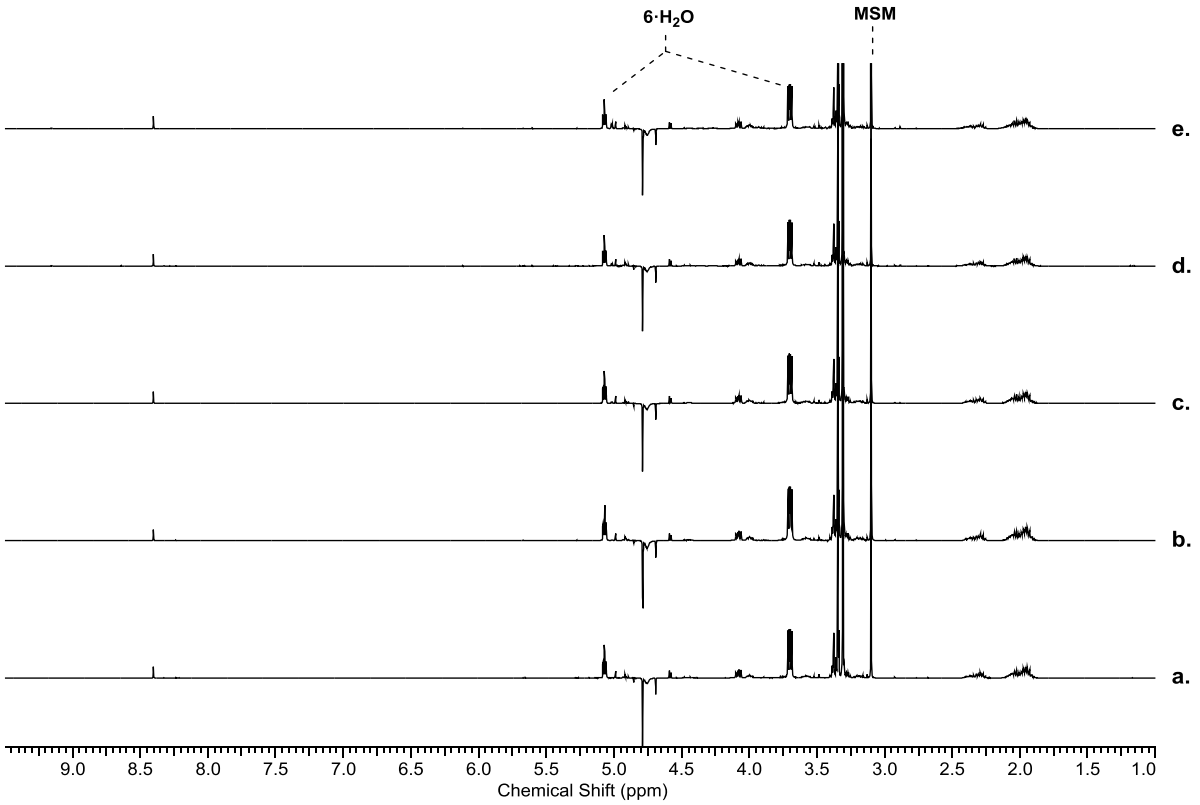
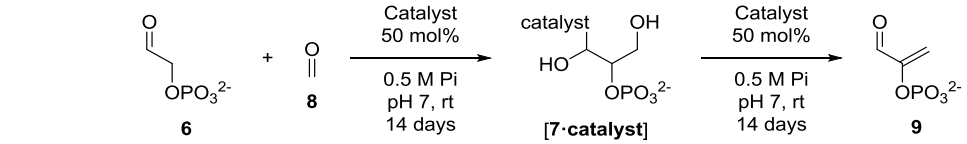


Figure S17. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 1M) and L-proline (50 mol%) in H_2O at pH 7, with MSM (25 mM) as an internal NMR standard after: **a.** 30 min; **b.** 1 d; **c.** 3 d; **d.** 6 d; **e.** 14 d.

Amine catalysts in phosphate buffer

Glycolaldehyde-phosphate (**6**, 37.00 µL, 75.0 µmol) and catalyst (50 mol%) were dissolved in phosphate buffer (700 µL, 500 mM, pH 7). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution. NMR spectra were then acquired regularly over 14 d. Yields for [**7·catalyst**] and **9** observed after 14 d for each catalyst are given in Figure S18.



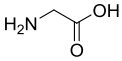
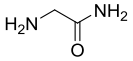
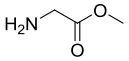
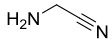
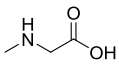
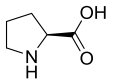
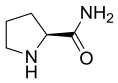
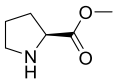
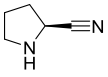
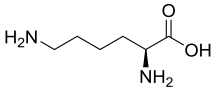
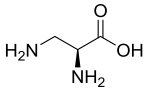
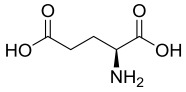
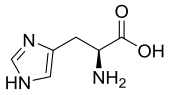
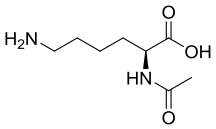
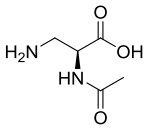
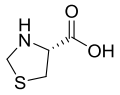
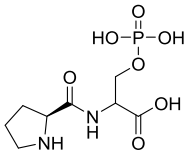
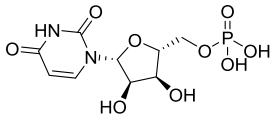
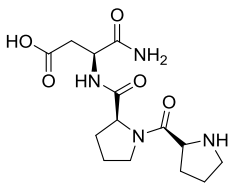
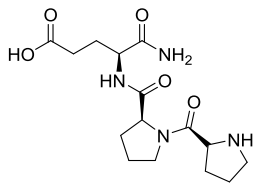
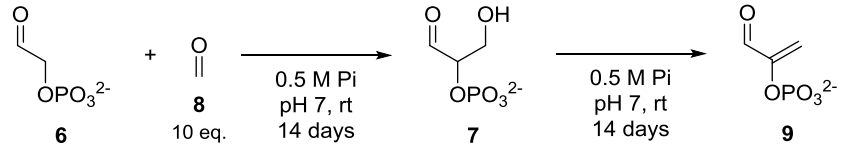
control				
No catalyst	Glycine	Glycine amide	Glycine methyl ester	Glycine nitrile
7: nd **9**: nd	[**7·Gly**]: 7% **9**: 10%	[**7·Gly-NH$_2$**]: 3% **9**: trace	[**7·Gly-OMe**]: 5% **9**: 7%	[**7·Gly-CN**]: < 2% **9**: trace
				
Sarcosine	**L-Proline**	**L-Proline amide**	**L-Proline methyl ester***	L-Proline nitrile
[**7·Sar**]: 5% **9**: 26%	[**7·Pro**]: 15% **9**: 31%	[**7·Pro-NH$_2$**]: 4% **9**: 33%	[**7·Pro-OMe**]: 15% **9**: 28%	[**7·Pro-CN**]: trace **9**: 10%
				
Pyrrolidine	L-Lysine	Dpr-OH	L-Glutamic acid	
[**7·Pyrr**]: 18% **9**: 33%	[**7·Lys**]: 2% **9**: 14%	[**7·Dpr**]**:** trace **9:** 11%	[**7·Glu**]: 16% **9**: 13%	
				
L-Histidine	Ac-Lys-OH	Ac-Dpr-OH	H-L-Thz-OH	
[**7·His**]: trace **9**: 15%	[**7·Ac-Lys**]:8% **9**: 16%	[**7·Ac-Dpr**]: 2% **9**: 5%	[**7·Thz**]: trace **9**: 5%	
				
Pro-O-Phospho-Ser-OH	Uridine 5'-monophosphate	Pro-Pro-Asp-NH$_2$	**Pro-Pro-Glu-NH$_2$**	
[**7·Pro-SOP**]: 2% **9**: 11%	[**7·UMP**]: nd **9**: nd	[**7·Pro-Pro-Asp-NH$_2$**]: trace **9**: 17%	[**7·Pro-Pro-Glu-NH$_2$**]: 18% **9**: 24%	

Figure S18. Yields for the formation of catalyst-glyceraldehyde-2-phosphate hemiaminal [**7·catalyst**] and phosphoenol pyruvaldehyde (**9**) upon reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 1 M) with the specified catalysis (50 mol%) in phosphate buffer (500 mM, pH 7), with MSM (25 mM) as internal NMR standard, after 14 d. nd = not detected. *Hydrolyses to Pro-OH (> 95%) was observed.

Phosphate buffer (without an amine catalyst)



Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) was dissolved in phosphate buffer (700 µL, 500 mM, pH 7). Formaldehyde (**8**, 55.9 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution. NMR spectra were then acquired regularly over 14 d. The evolution of the reaction can be observed in Figure S19.

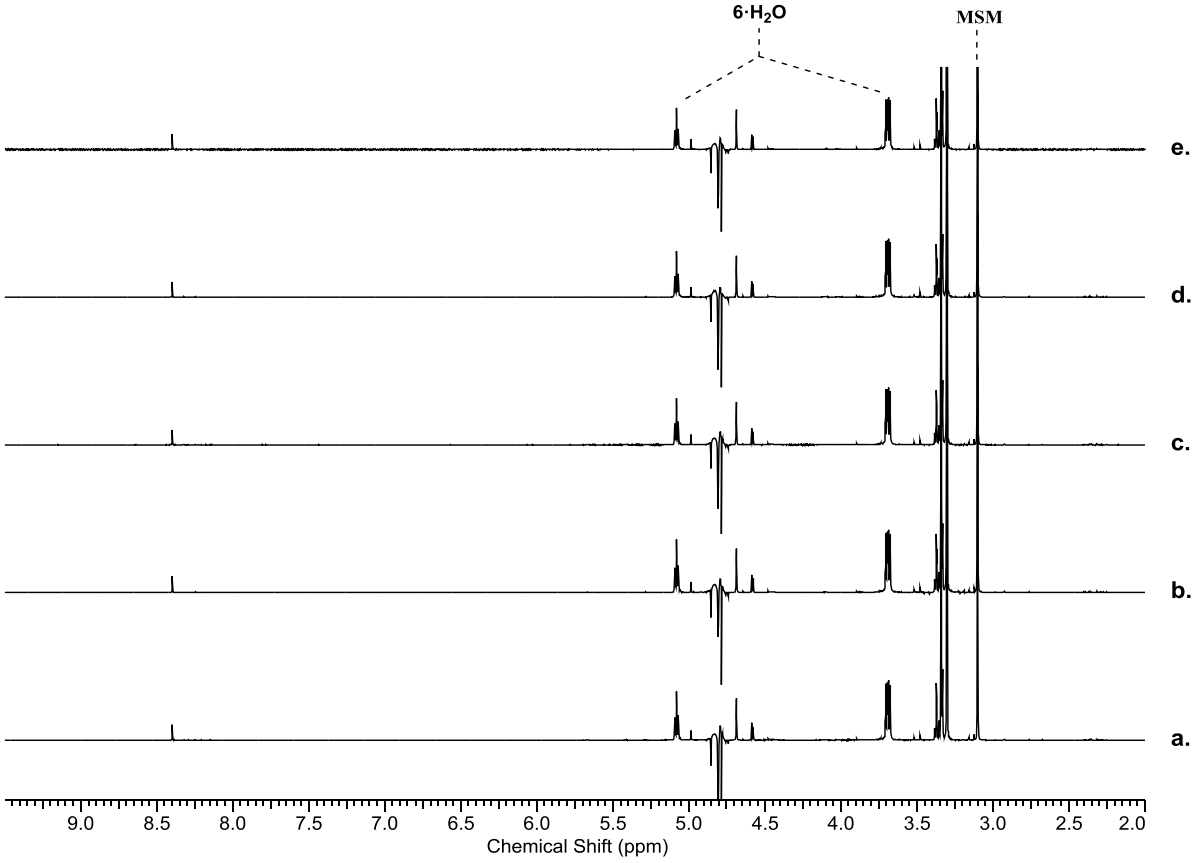
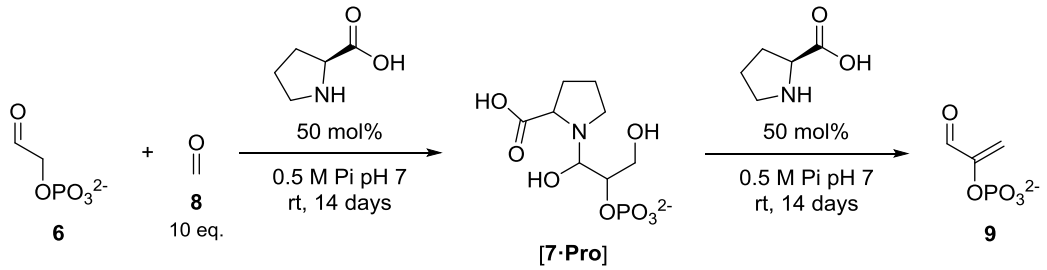


Figure S19. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 2.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 1 M) in phosphate buffer (pH 7, 500 mM) at pH 7, with MSM (25 mM) as internal NMR standard after: **a.** 30 min; **b.** 1 d; **c.** 3 d; **d.** 6 d; **e.** 14 d.

L-Proline in phosphate buffer



Glycolaldehyde-phosphate (**6**, 37.00 µL, 75.0 µmol) and L-proline (4.3 mg, 37.5 µmol, 50 mol%) were dissolved in phosphate buffer (700 µL, 500 mM, pH 7). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution. NMR spectra were then acquired regularly over 14 d. The evolution of the reaction can be observed in Figure S20.

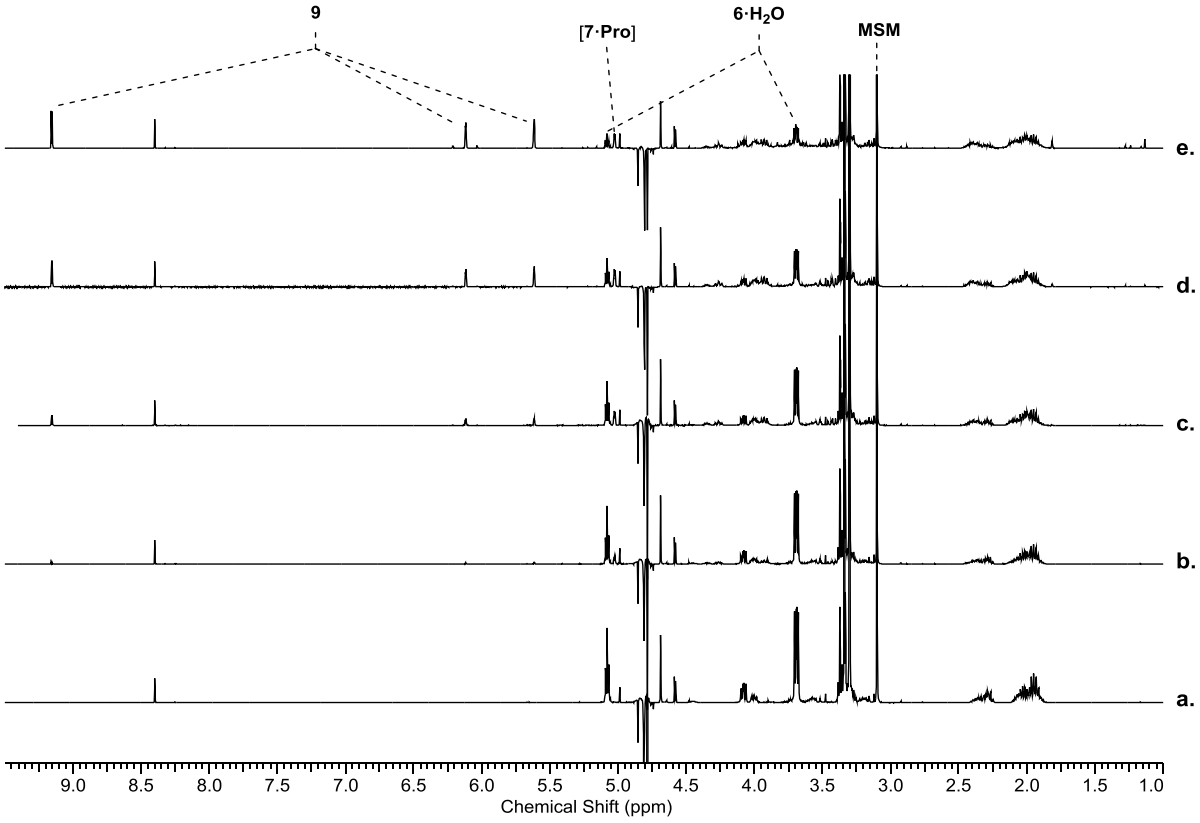


Figure S20. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**4-P**, 100 mM), formaldehyde (**8**, 1 M) and L-proline (50 mol%) in phosphate buffer (500 mM, pH 7), with MSM (25 mM) as an internal NMR standard, after: **a.** 30 min; **b.** 1 d; **c.** 3 d; **d.** 7 d; **e.** 14 d.

Reaction optimization

Supplementary discussion

As anticipated lower GAB activity was observed for imidazole than phosphate, and a relatively poor conversion to **9** (7%) was observed. Although with co-catalytic imidazole an increased yield of [**7**·**Pro**] (26%) was observed with respect to unbuffered water under comparable conditions (10% [**7**·**Pro**]) (Supplementary Table S3-S5), highlighting the uniquely effective nature of Pi as a GAB catalyst at neutral pH.

As expected, the reaction was also promoted by increased Pro loading (Supplementary Fig. 21d), and higher temperatures (Supplementary Fig. 21a). The maximum **Pro**/P_i catalyzed conversion of **6** to **9** was observed in only 48 h at 40°C, but Pi was still essential; without Pi at 40°C the conversion to **9** (2%) was (18-fold) suppressed. An even higher conversion to **9** was observed after only 7 h at 60°C, however by-products then quickly accumulated due to the facile hydrolysis of 9 to methyl glyoxal at this temperature (Supplementary Table S1-S2).[10]

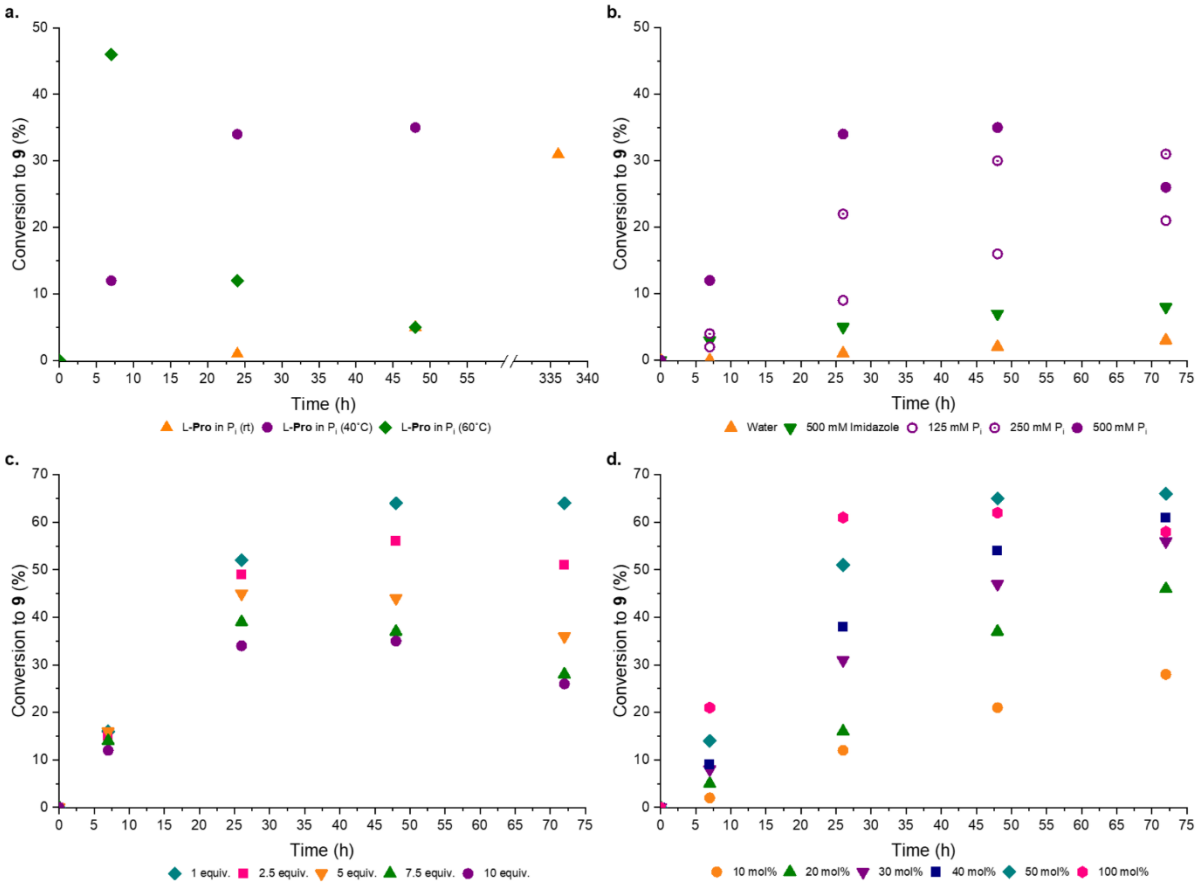
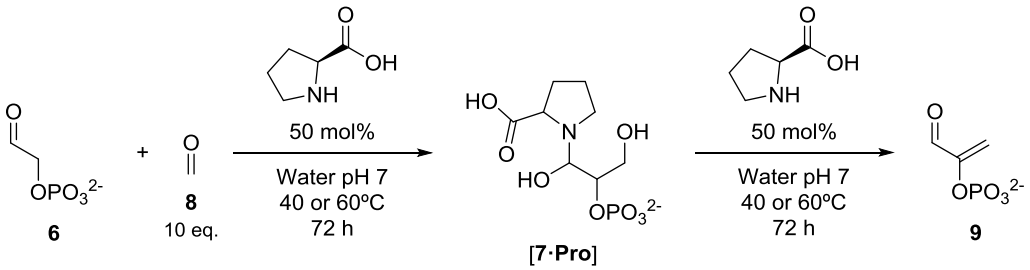


Figure S21. Proline-catalysed synthesis of phosphoenol pyruvaldehyde. Graphs to show the effect of temperature, buffer, stoichiometry, and catalyst loading. **a,** Reaction of glycolaldehyde phosphate (**6**, 100 mM) and formaldehyde (**8**, 10 equiv.) in phosphate buffer (P_i, 500 mM, pH 7) with L-**Pro** (50 mol%) at room temperature (rt), 40°C or 60°C; **b,** Reaction of **6** (100 mM) and **8** (10 equiv.) in neutral water or specified buffer (125 – 500 mM, pH 7) with L-**Pro** (50 mol%) at 40°C; **c,** Reactions of **6** (100 mM) and **8** (1 – 10 equiv.) in P_i buffer (500 mM, pH 7) with L-**Pro** (50 mol%) at 40°C; **d,** Reaction of **6** (100 mM) with **8** (100 mM) in P_i buffer (500 mM, pH 7) with L-**Pro** (10 – 100 mol%) at 40°C.

Temperature in water

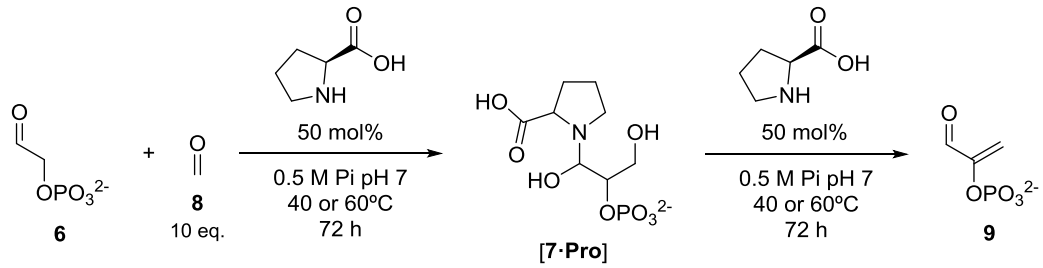


Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (0 - 50 mol%) were dissolved in H_2O (700 µL). Formaldehyde (**7**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at room temperature, 40ºC or 60ºC. NMR spectra were then acquired regularly over 14 d. The observed yields are given in Table S1

Entry	L-Pro (mM)	Temperature (ºC)	Time (h)	[**7·Pro**] (%)	**9** (%)
1	-	rt	336	nd	nd
2	-	40	72	< 2	nd
3	-	60	72	6	22
4	50	rt	336	15	trace
5	50	40	72	10	< 2
6	50	60	26	6	16

Table S1. Yields for the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 1 M) and L-proline (0 – 50 mM), with MSM (25 mM) as an internal NMR standard, in H_2O at pH 7. Time for maximum yield reported. rt = ambient room temperature. nd = not detected.

Temperature in phosphate buffer

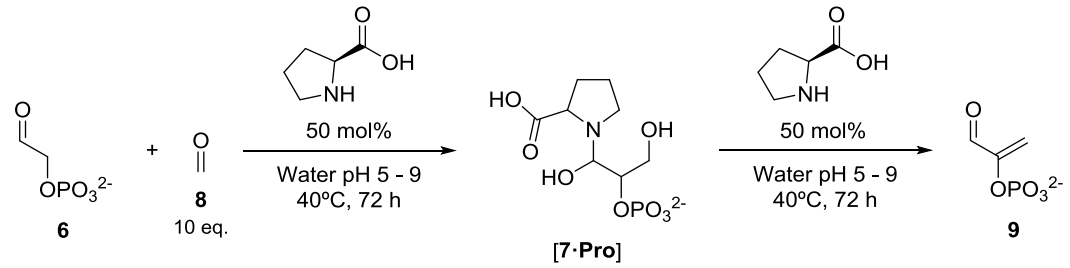


Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (0 - 50 mol%) were dissolved in phosphate buffer (700 µL, 500 mM, pH 7). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at room temperature, 40ºC or 60ºC. NMR spectra were then acquired regularly over 14 d. The observed yields are given in Table S2.

Entry	L-Pro (mM)	Temperature	Time (h)	[**7·Pro**] (%)	**9** (%)
1	-	rt	336	nd	nd
2	-	40	72	4	3
3	-	60	48	3	42
4	50	rt	336	16	31
5	50	40	48	17	35
6	50	60	7	< 2	46

Table S2. Yields for the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 1 M) and L-proline (0 – 50 mM), with MSM (25 mM) as an internal NMR standard, in phosphate buffer (500 mM, pH 7). Time for maximum yield reported. rt = ambient room temperature. nd = not detected.

pH

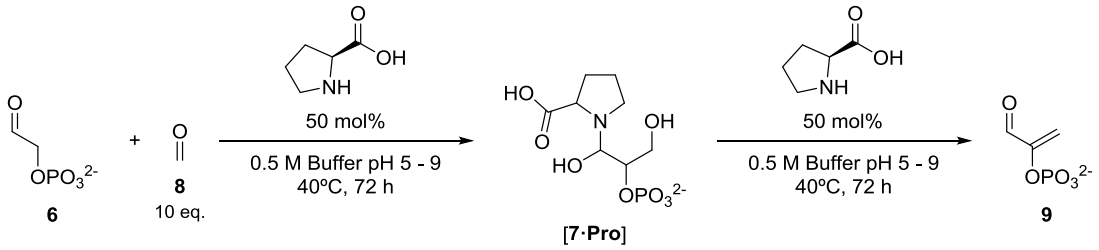


Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in H_2O (700 µL). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to the desired pH, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The observed yields are given in Table S3.

Entry	pH	Time (h)	[**7·Pro**] (%)	**7·H_2O**	**9** (%)
1	5.0	72	nd	-	4
2	6.0	72	13	-	3
3	7.0	72	12	-	3
4	8.0	72	10	-	4
5	9.0	72	28	-	19
6[a]	9.0	72	-	5	trace

Table S3. Yields for the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 1 M) and L-proline (50 mM), with MSM (25 mM) as an internal NMR standard, in H_2O at specific pH and 40ºC. nd = not detected. [a]Reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 1 M), with MSM (25 mM) as internal NMR standard, in H_2O at pH 9 and 40ºC.

pH and buffer

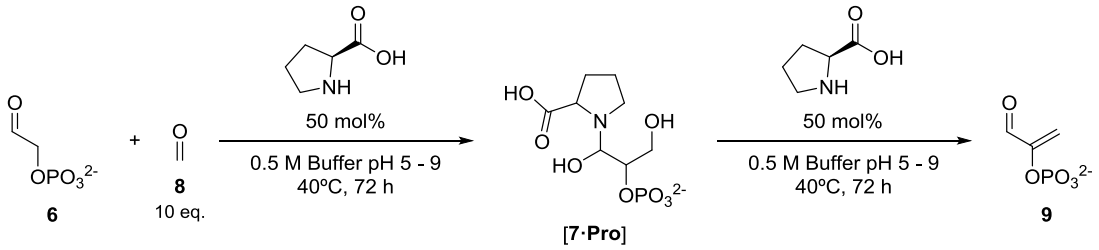


Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in specific buffer (700 µL). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to the desired pH, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The observed yields are given in Table S4.

Entry	pH	Buffer	Time (h)	[**7·Pro**] (%)	**7·H_2O** (%)	**9** (%)
1	5.0	Acetate	72	10	-	6
2	5.0	Phosphate	72	7	-	14
3	6.0	Citrate	72	15	-	10
4	6.0	Phosphate	48	11	-	31
5	7.0	Imidazole	72	28	-	7
6	7.0	Phosphate	48	17	-	35
7	8.0	Borate	72	23	-	17
8	8.0	Phosphate	26	10	-	46
9	9.0	Borate	48	8	-	30
10	9.0	Phosphate	26	10	-	44
11[a]	9.0	Phosphate	72	-	9	8

Table S4. Yields for the reaction of glycolaldehyde-phosphate (**4-P**, 100 mM), formaldehyde (**8**, 1 M) and L-proline (50 mM), with MSM (25 mM) as an internal NMR standard, at the specified pH in the specified buffer (500 mM) at 40ºC. Time for maximum yield reported. [a]Reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 1 M), with MSM (25 mM) as internal NMR standard, in phosphate buffer (500 mM) at pH 9 and 40ºC.

Phosphate concentration

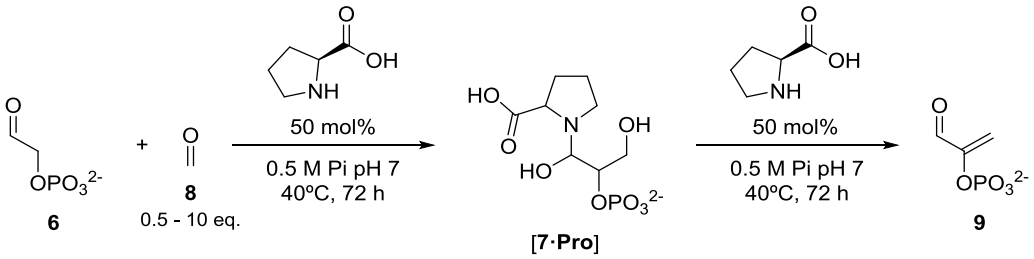


Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in phosphate buffer (700 µL, 0 – 1 M). Formaldehyde (**8**, 55.8 µL, 0.75 mmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The observed yields are given in Table S5.

Entry	[Phosphate]	Time (h)	[**7·Pro**] (%)	**9** (%)
1	0	72	12	3
2	125 mM	72	16	20
3	250 mM	72	11	30
4	500 mM	48	17	35
5	1 M	26	9	46

Table S5. Yields for the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 1 M) and L-proline (50 mM), with MSM (25 mM) as an internal NMR standard, at pH 7 in phosphate buffer (0 – 1 M) at 40ºC. Time for maximum yield reported.

Screening of formaldehyde stoichiometry in phosphate buffer[a]



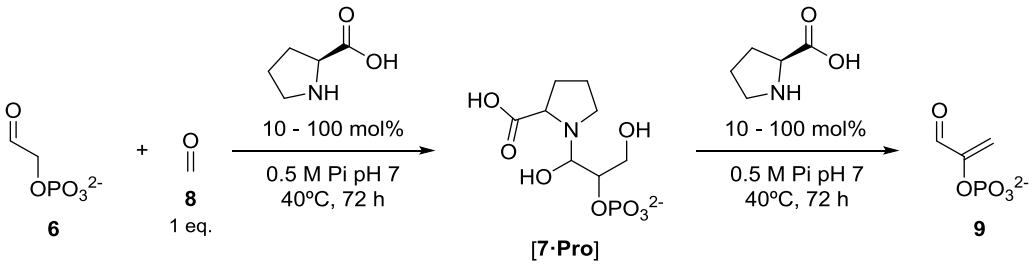
Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in phosphate buffer (700 µL). Formaldehyde (**8**, 1 -10 eq.) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The observed yields are given in Table S6.

Entry	Formaldehyde (M)	Time (h)	[**7·Pro**] (%)	**9** (%)
1	0.05	26	9 (18[a])	38 (76[a])
2	0.10	48	7	64
3	0.25	48	8	56
4	0.50	26	15	45
5	0.75	26	15	39
6	1.00	26	17	35

Table S6. Yields for the reaction of glycolaldehyde-phosphate (**4-P**, 100 mM), formaldehyde (**8**, 100 mM – 1.0 M) and L-proline (50 mM), with MSM (25 mM) as an internal NMR standard, in phosphate buffer (500 mM, pH 7) at 40ºC. [a]Yield based on **8** as limited reagent.

[a]Eschemnoser and co-workers[6] reported that the reaction glycolaldehyde-phosphate (**6**) and formaldehyde (**8**, 10 equiv.) at pH 10.7 afforded a 66% yield of **7** after 6 d at room temperature. We have incubated **6** (26.8 µL, 54.0 µmol), **8** (1 equiv.) and MSM (13.6 µmol, internal NMR standard) at pH 10.7 in H_2O at room temperature. The reaction was monitored for 6 d by NMR spectroscopy. After 6 d, we observed 40% conversation to **7**.

Catalyst loading (L-proline) in phosphate buffer



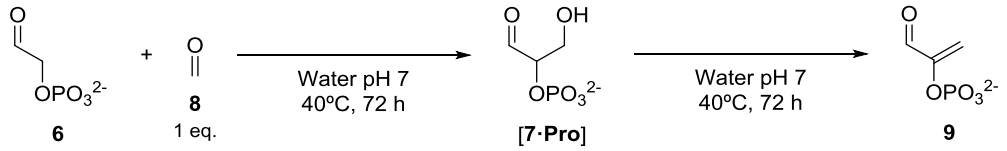
Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (10 - 100 mol%) were dissolved in phosphate buffer (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The observed yields are given in Table S7.

Entry	L-Pro (mM)	Time (h)	[**7·Pro**] (%)	**9** (%)
1	10 mM	72	< 5%	27%
2	20 mM	72	< 5%	45%
3	30 mM	72	< 5%	55%
4	40 mM	48	9%	60%
5	50 mM	48	12%	65%
6	60 mM	48	7%	63%
7	70 mM	48	7%	63%
8	80 mM	48	10%	65%
9	90 mM	48	10%	62%
10	100 mM	48	12%	62%

Table S7. Yields for the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (10 - 100 mM), with MSM (25 mM) as an internal NMR standard, in phosphate buffer (500 mM, pH 7) at 40ºC.

Control for optimized reaction conditions

Reaction in water (without L-proline)



Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) was dissolved in H_2O (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The evolution of the reaction is shown in Figure S21.

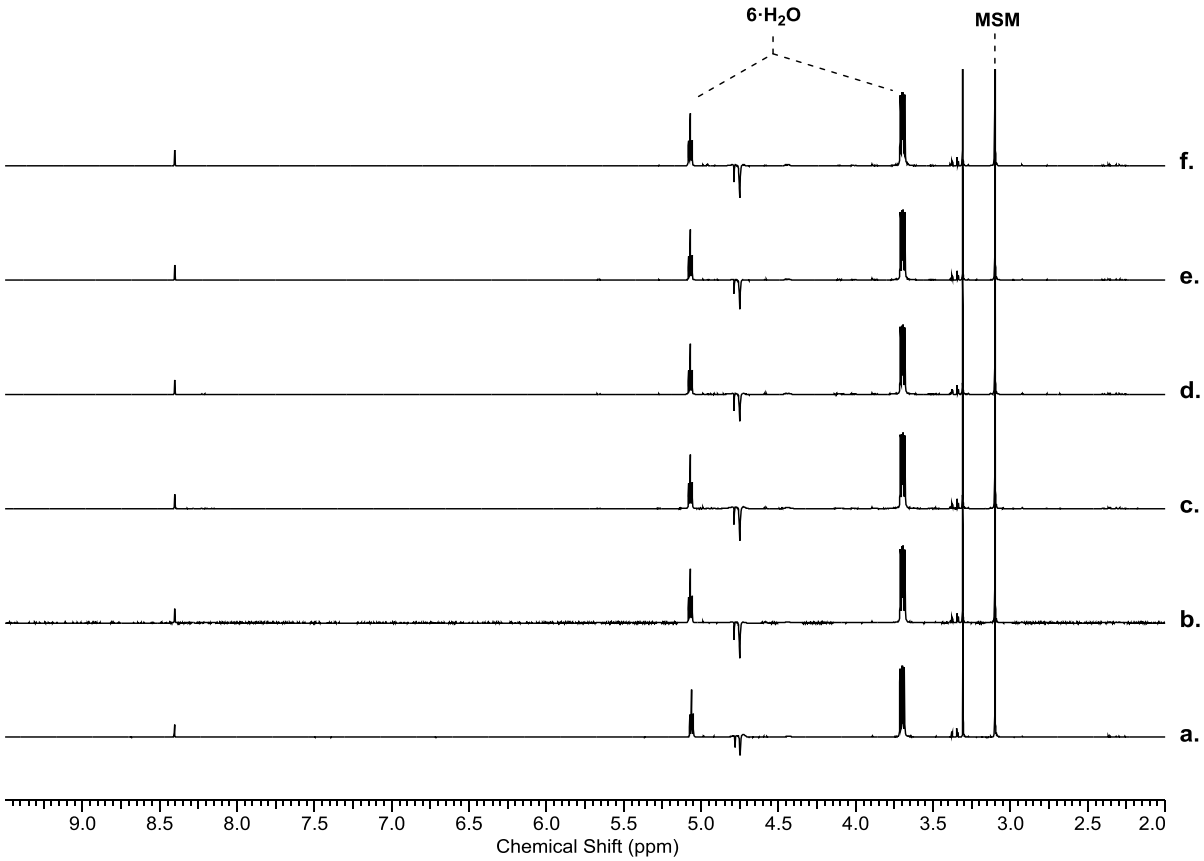
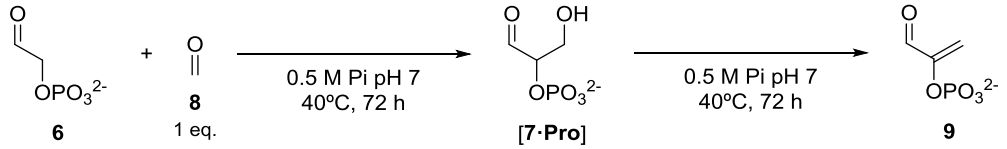


Figure S22. ^1H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 2.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 100 mM), with MSM (25 mM) as an internal NMR standard, in H_2O at pH 7 and 40ºC after: **a.** 0 h; **b.** 1 h; **c.** 7 h; **d.** 26 h; **e.** 50 h; **f.** 72 h.

Phosphate buffer (without L-proline)



Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) was dissolved in phosphate buffer (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The evolution of the reaction is shown in Figure S22.

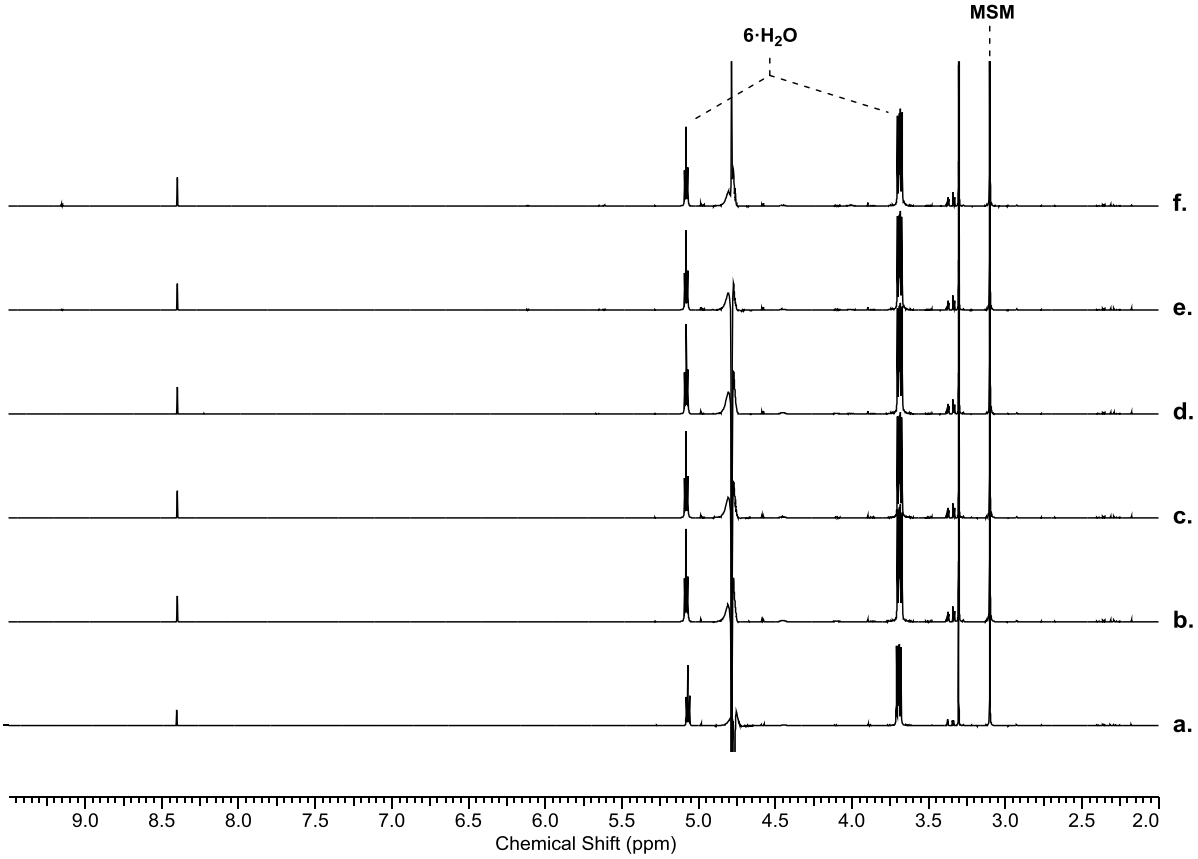
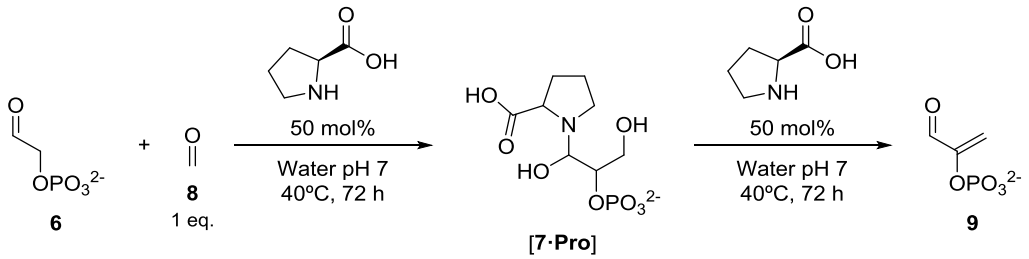


Figure S23. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 2.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM) and formaldehyde (**8**, 100 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 1 h; **c.** 7 h; **d.** 26 h; **e.** 50 h; **f.** 72 h.

L-Proline in water



Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in H_2O (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The evolution of the reaction is shown in Figure S23.

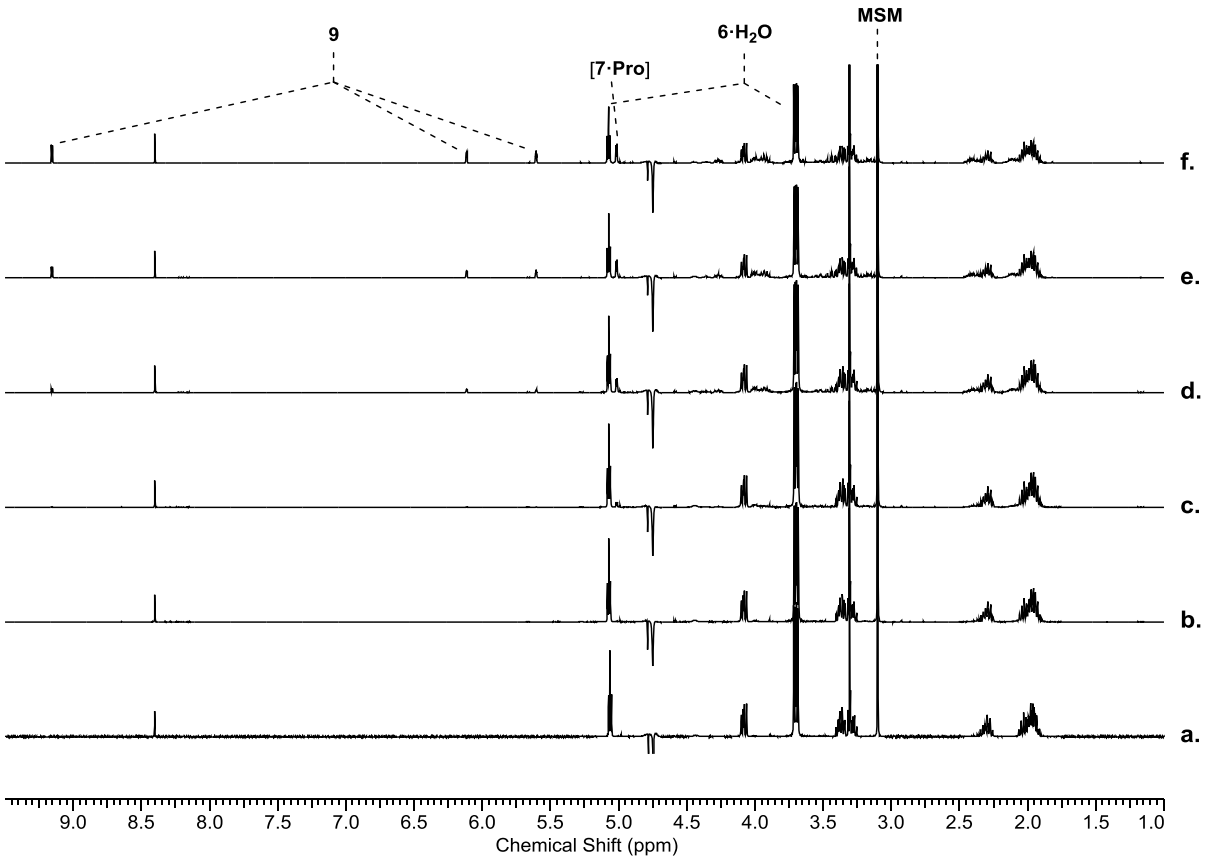
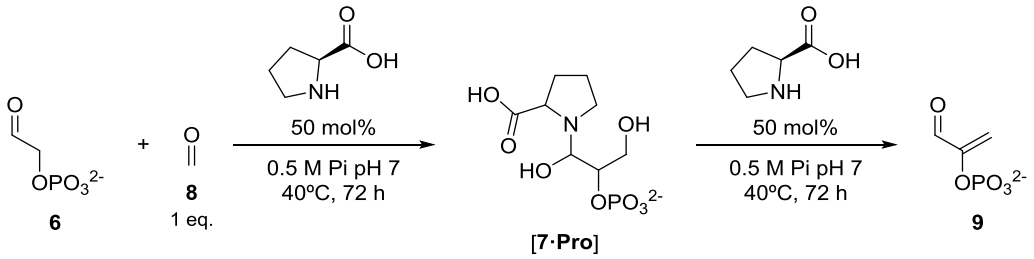


Figure S24. ^1H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM) and with MSM (25 mM) as an internal NMR standard, in H_2O at pH 7 and 40ºC after: **a.** 0 h; **b.** 1 h; **c.** 7 h; **d.** 26 h; **e.** 50 h; **f.** 72 h.

L-Proline in phosphate buffer - quadruplicates



Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in phosphate buffer (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. After 48 h, [**7·Pro**] (12%) and **9** (64%; average over 4 reactions) were observed. The evolution of the reaction can be observed in Figure S24 and Figure S25.

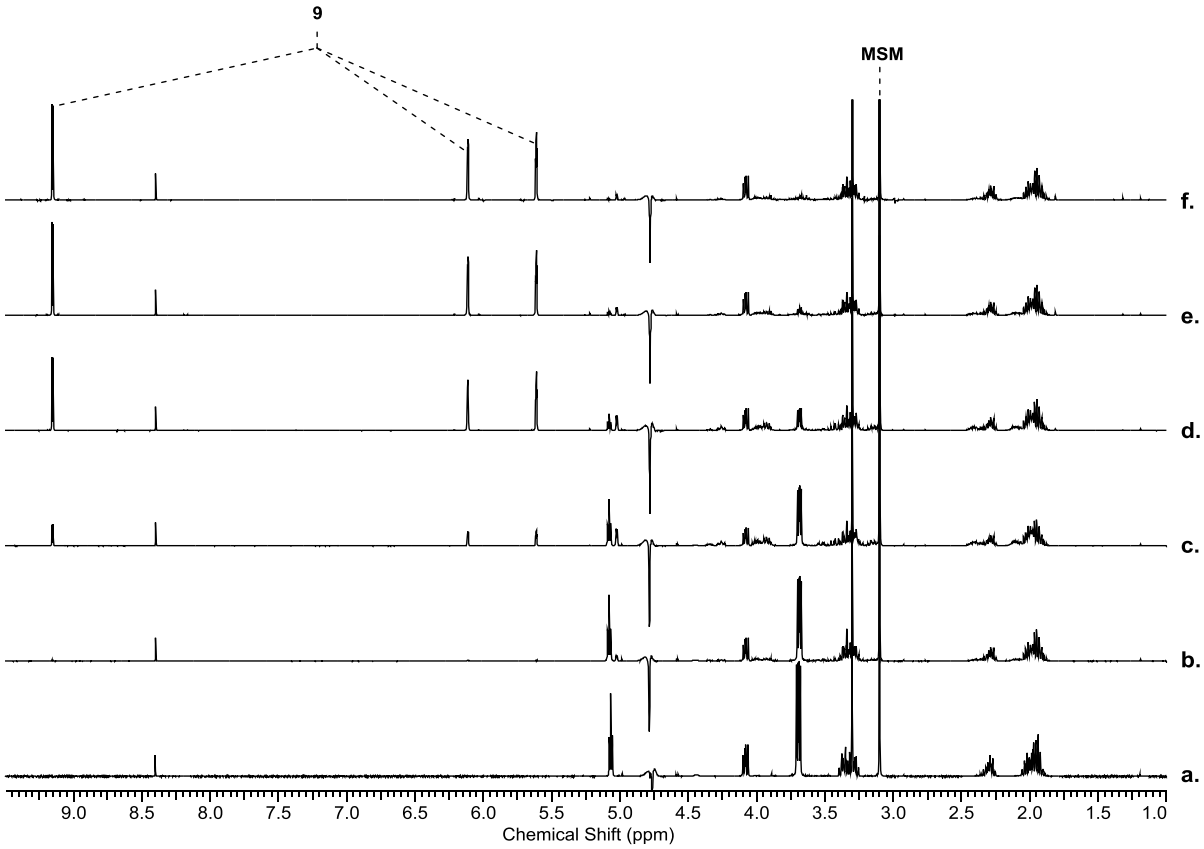


Figure S25. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 1 h; **c.** 7 h; **d.** 26 h; **e.** 50 h; **f.** 72 h.

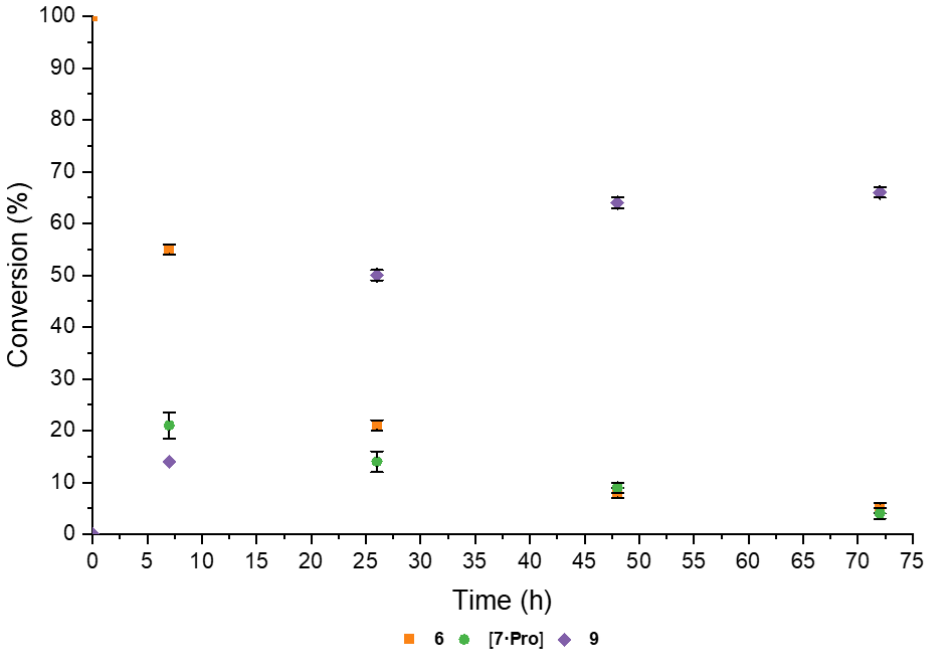


Figure S26. Evolution of the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM) with MSM (25 mM) as an internal NMR standard, in phosphate buffer (500 mM, pH 7) at 40ºC. Orange squares represents **6**; Green circles represents [**7·Pro**]; Purple diamonds represents **9**. Error bars presented as standard deviation.

Mechanistic studies

*Spiking of aldol reaction with **7***

Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in phosphate buffer (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 7 h. After that time, the reaction was monitored by NMR and then spiked with glyceraldehyde-2-phosphate (**7**, 5 mg) and NMR spectra acquired. The evolution of the reaction and the spiking can be observed in Figure S26.

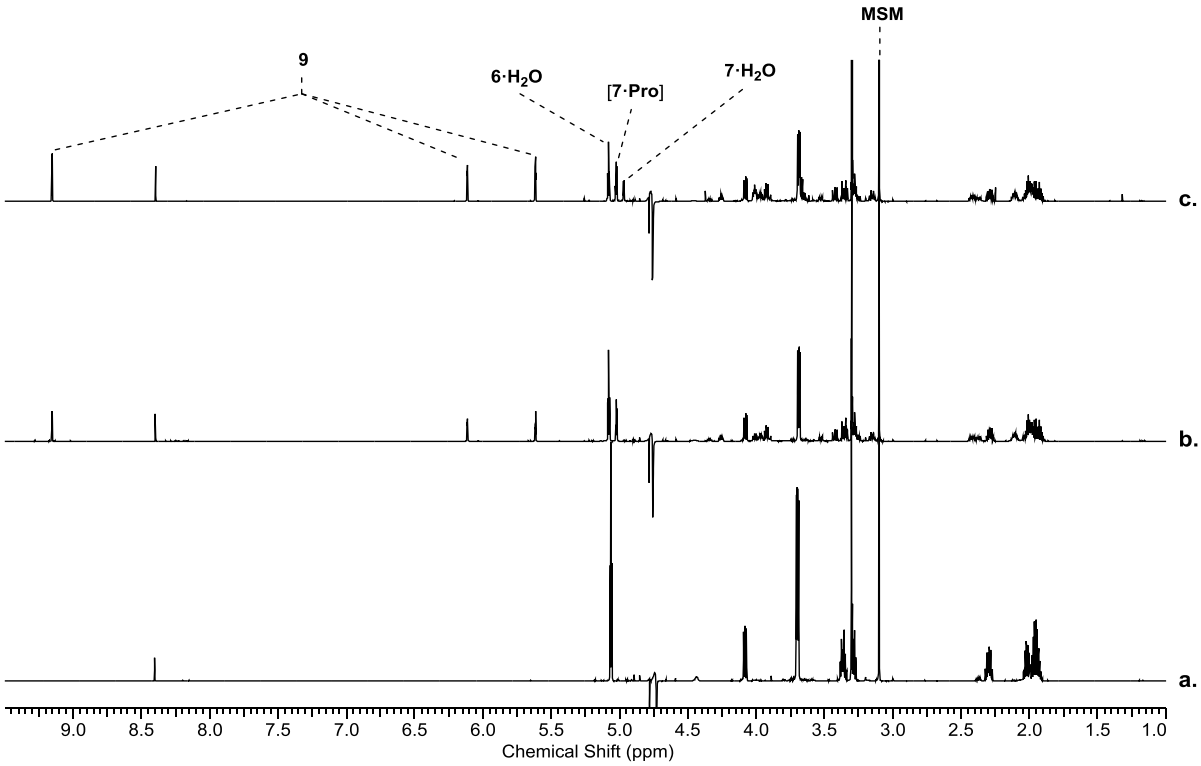


Figure S27. ^{1}H NMR (700 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 7 h. **Spectra c.** Addition of authentic glyceraldehyde-2-phosphate (**7**, 5 mg).

Studies of aldol reaction with L-proline in phosphate buffer

Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in phosphate buffer (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 7 h. The reaction was monitored by NMR over 7h and after that time analysed by ^{1}H-^{13}C HMBC and ^{1}H-^{31}P HMBC. The evolution of the reaction can be observed in Figure S27, ^{1}H-^{13}C HMBC in Figure S38-S29 and ^{1}H-^{31}P HMBC in Figure S30. A combination of mass spectrometry [HRMS (ESI): m/z calcd for $C_8H_{16}NO_8P+H^+$: 286.0686; found: 286.0683] and ^{1}H-^{13}C HMBC NMR cross-correlation indicated the observed hemiaminal [**7·Pro**].

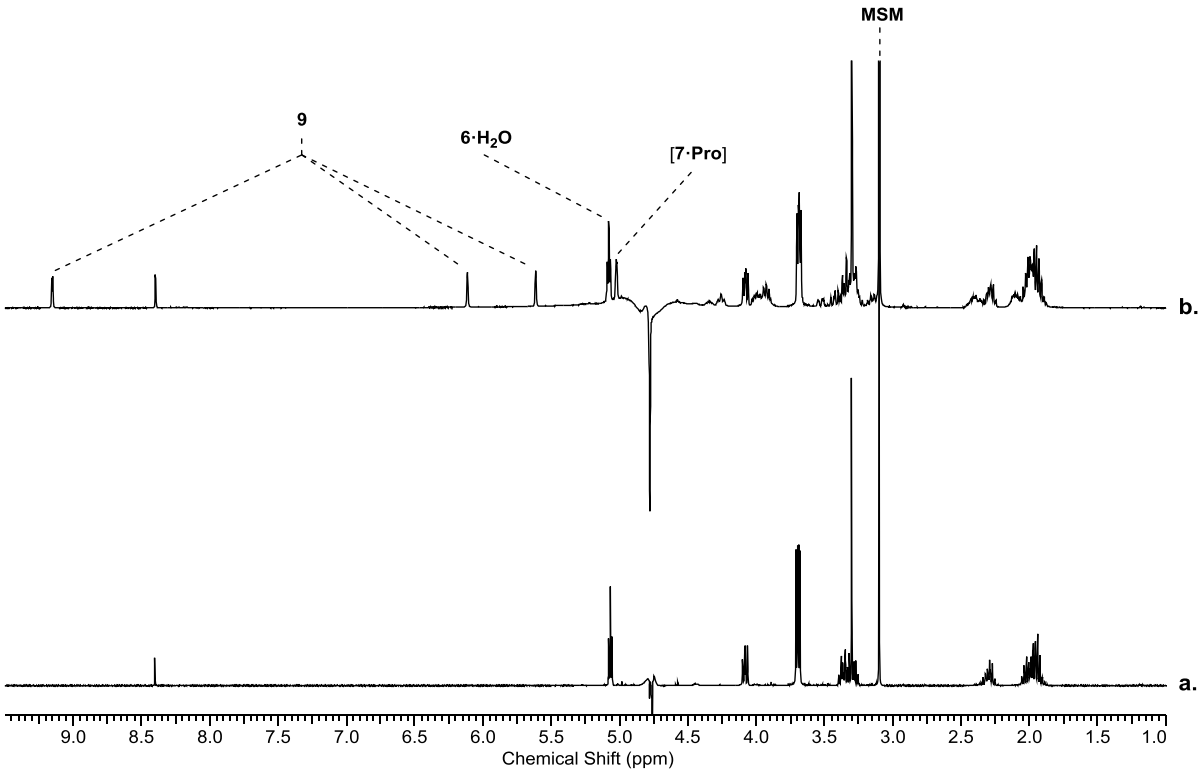


Figure S28. ^{1}H NMR (700 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 7 h.

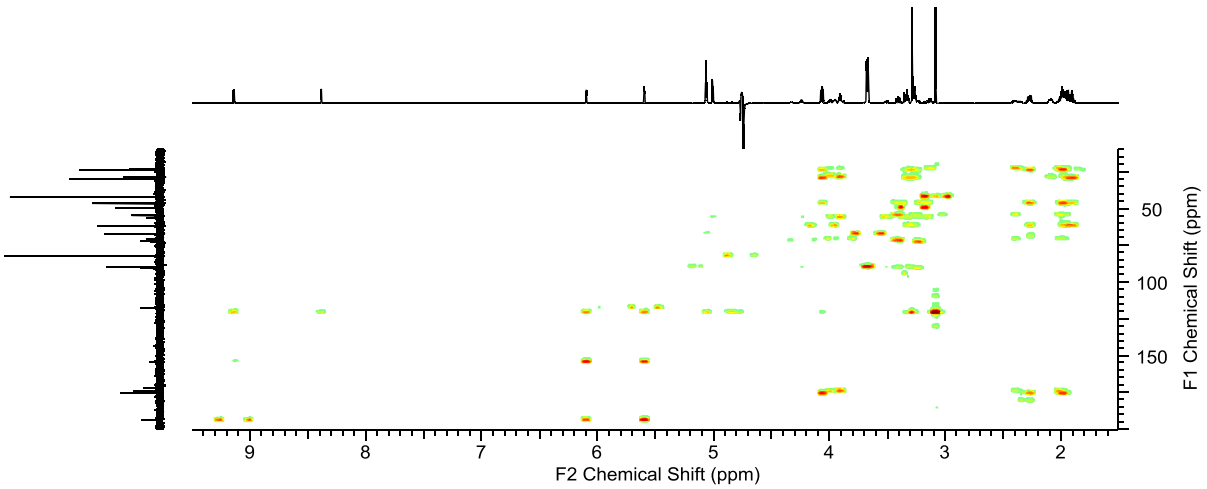


Figure S29. ^{1}H-^{13}C HMBC NMR (700, 176 MHz, H_2O/D_2O 9:1, 1.00 - 9.50 and 10 - 200 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after 7 h.

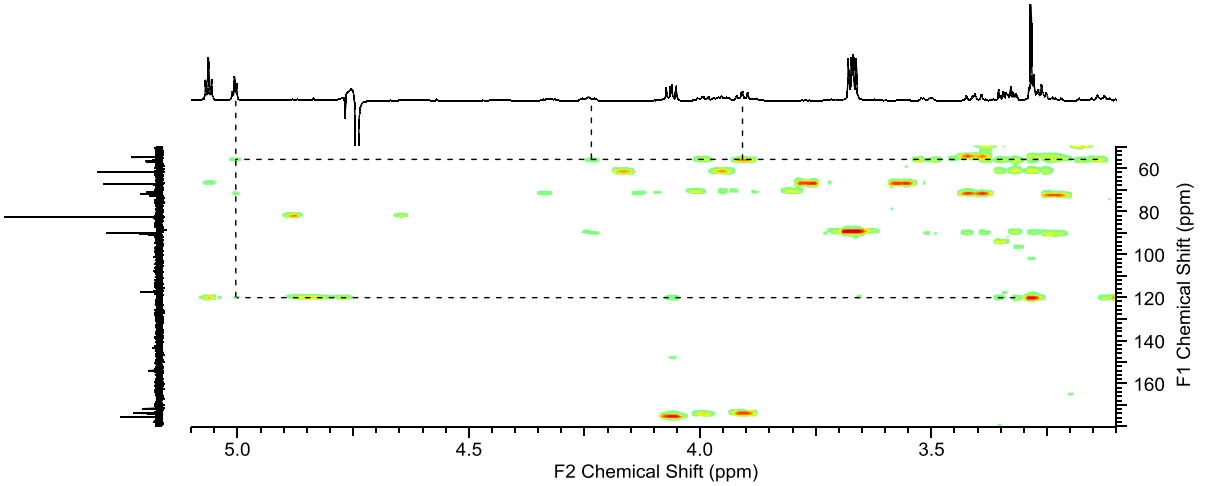


Figure S30. ^{1}H-^{13}C HMBC NMR (700, 176 MHz, H_2O/D_2O 9:1, 3.1 – 5.1 and 50 - 180 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after 7 h.

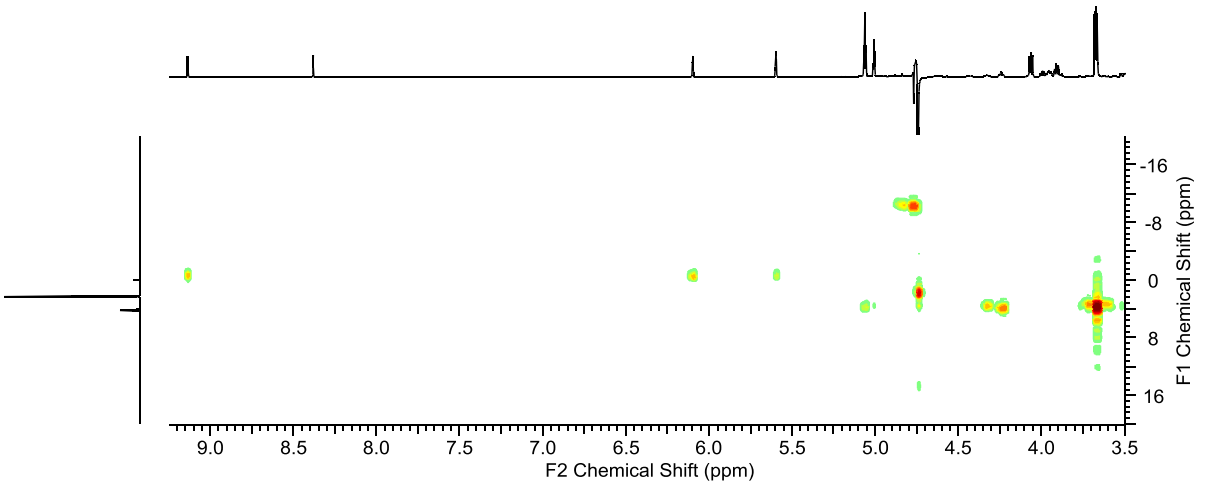


Figure S31. ^{1}H-^{31}P HMBC NMR (700, 284 MHz, H_2O/D_2O 9:1, 3.5 - 9.25 and -20 - 20 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after 7 h.

Studies of aldol products between L-proline and formaldehyde

L-proline (4.3 mg, 50 mol%, 37.5 µmol) was dissolved in phosphate buffer (700 µL) and formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 24 h. After that time, formaldehyde (**8**, 50.3 µL, 675 µmol) was added and then incubated at 40ºC for 24 h. The reaction was monitored by NMR over 48 h and after that time analysed by ^{1}H-^{13}C HMBC and ^{1}H-^{31}P HMBC. The evolution of the reaction can be observed in Figure S31.

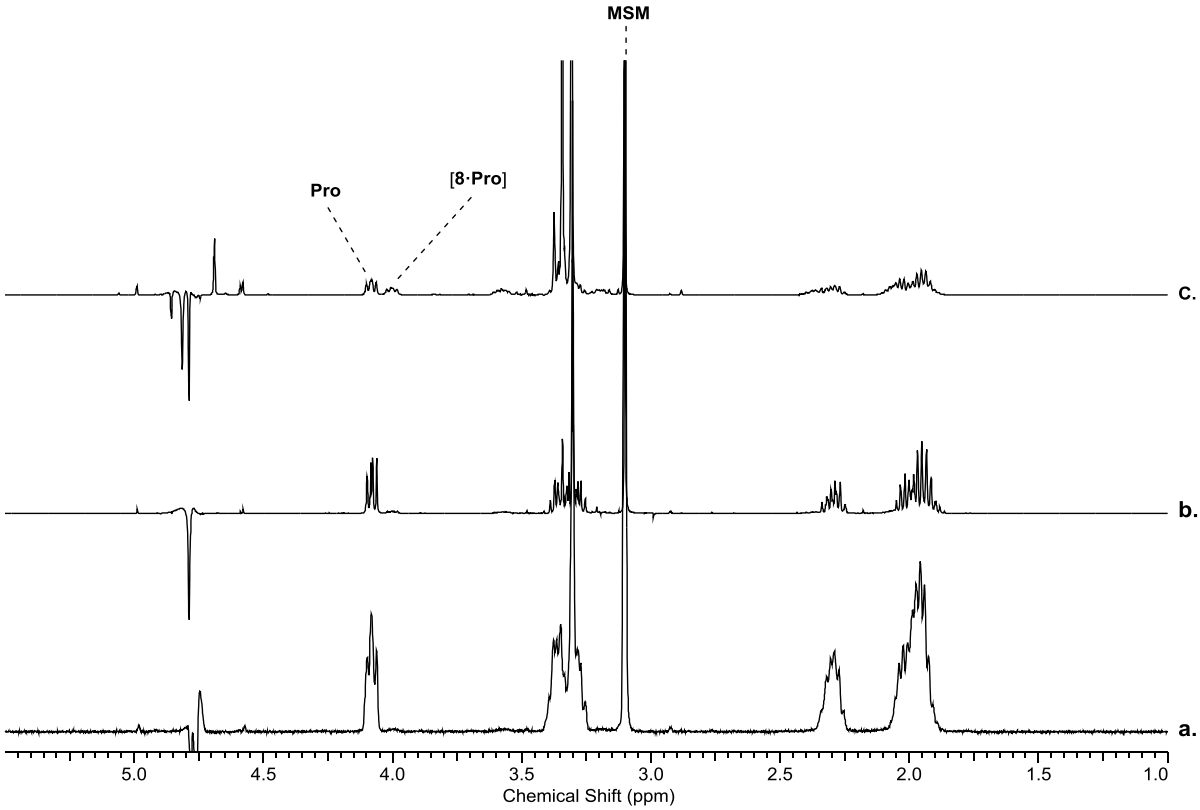


Figure S32. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 5.50 ppm) to show the reaction of formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 7 h. **c.** After addition of formaldehyde (**8**, 9 eq.) and incubation at 40ºC for 24 h.

Studies of glyceraldehyde-2-phosphate dehydration

L-Proline in water

Glyceraldehyde-2-phosphate (**7**, 5.0 mg, 16.9 µmol) and L-proline (3.9 mg, 33.9 µmol) were dissolved in H_2O/D_2O (9:1, 677 µL) and MSM (8.5 µL, 8.47 µmol) was added as an internal NMR standard. The solution pH was measured and adjusted to pH 7.0. The solution was then incubated at 40ºC. NMR spectra were periodically acquired. Trace dehydration to **9** (5%) was observed after 96h.

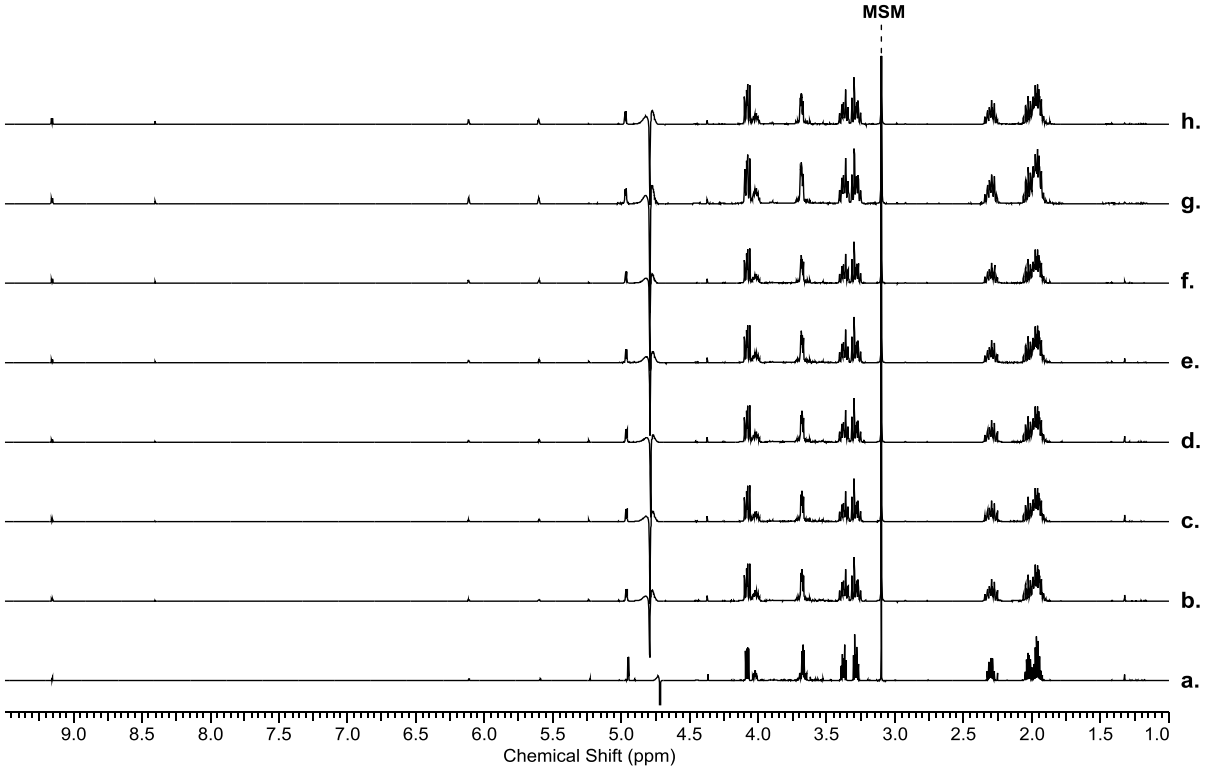


Figure S33. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reactivity of glyceraldehyde 2-phosphate (**7**, 25 mM) and L-proline (50 mM), with MSM (12.5 mM) as an internal NMR standard, in water at pH 7 and 40ºC after: **a.** 0 h (acquired at 700 MHz); **b.** 1 h; **c.** 4 h; **d.** 7.5 h; **e.** 24 h; **f.** 50 h; **g.** 72 h; **h.** 96 h.

Phosphate buffer (without L-proline)

Glyceraldehyde-2-phosphate (**7**, 14.7 mg, 49.8 µmol) was dissolved in phosphate buffer (500 µL, H_2O/D_2O 9:1, 500 mM) and MSM (12.45 µL, 12.45 µmol) was added as an internal NMR standard. The solution pH was measured and adjusted to pH 7.0. The solution was then incubated at 40ºC. NMR spectra were periodically acquired. Dehydration to **6** (40%) was observed after 108 h.

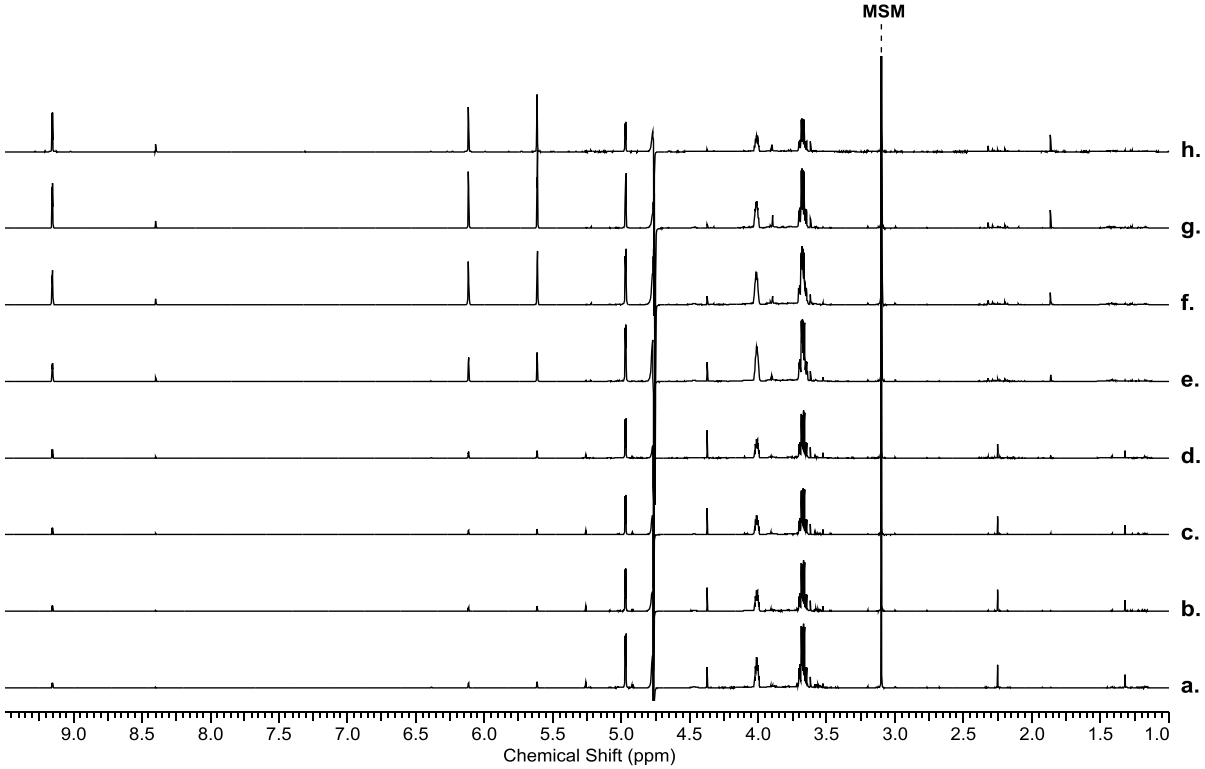


Figure S34. ^{1}H NMR (700 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the dehydration of glyceraldehyde 2-phosphate (**7**, 25 mM), with MSM (12.5 mM) as an internal NMR standard, in phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 1 h; **c.** 3 h; **d.** 6 h; **e.** 21 h; **f.** 50 h; **g.** 86 h; **h.** 108 h.

L-Proline in phosphate buffer

Glyceraldehyde-2-phosphate (**7**, 14.7 mg, 49.8 µmol) and L-proline (2.9 mg, 24.9 µmol) were dissolved in phosphate buffer (500 µL, H_2O/D_2O 9:1, 500 mM) and MSM (12.45 µL, 12.45 µmol) was added as an internal NMR standard. The solution pH was measured and adjusted to pH 7.0. The solution was then incubated at 40ºC. NMR spectra were periodically acquired. Dehydration to **6** (43%) was observed after 108 h.

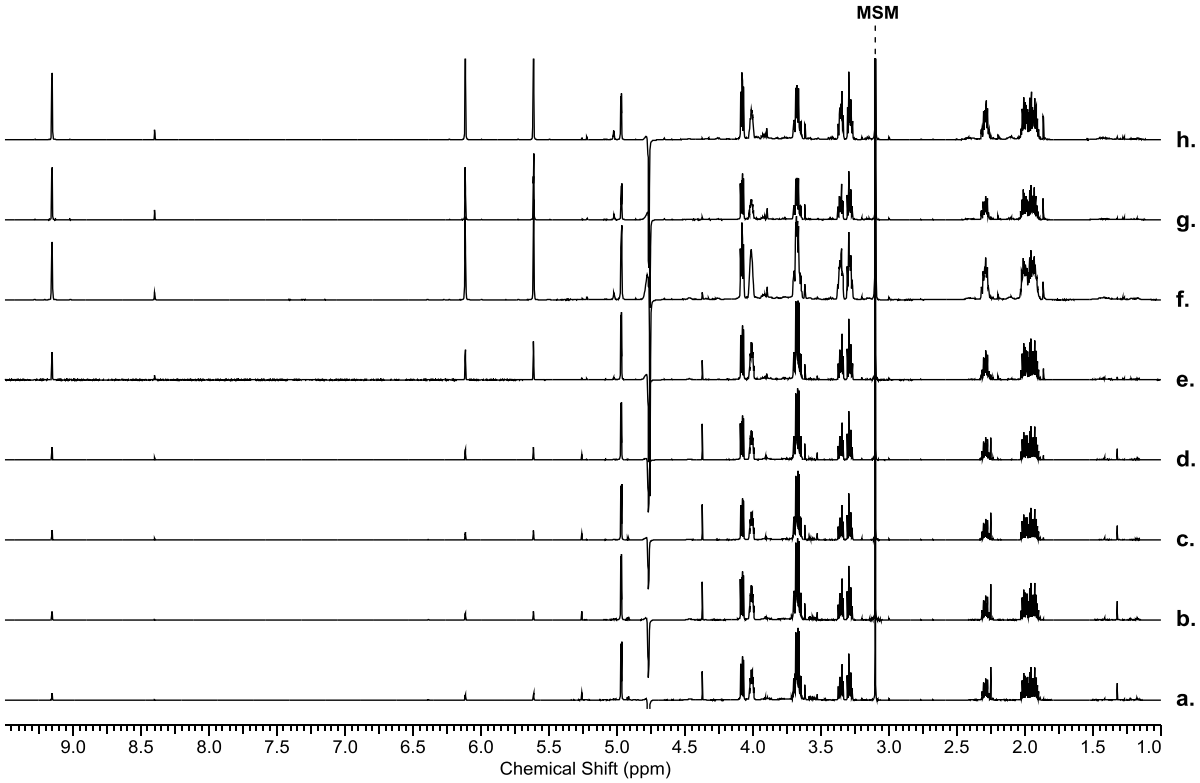


Figure S35. ^{1}H NMR (700 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the dehydration of glyceraldehyde 2-phosphate (**7**, 25 mM) with L-proline (50 mM), with MSM (12.5 mM) as an internal NMR standard, in phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 1 h; **c.** 3 h; **d.** 6 h; **e.** 21 h; **f.** 50 h; **g.** 86 h; **h.** 108 h.

Measurement of enantiomeric ratio of glyceraldehyde 2-phosphate

In triplicate, glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in H_2O (700 µL). Formaldehyde (**8**, 1.9 µL, 25 µmol) was added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution. The solution was then incubated at 40ºC and monitor by NMR spectroscopy. After 72 h 20% conversion to [**7·Pro**] was observed. The reaction pH was adjusted to pH 4.8 with HCl (1 M) and acid phosphatase (9.3 mg) was then added and the mixture was heated at 37ºC for 10 min. The three reactions were then combined, and the solvent removed in vacuo. The residue was dissolved in H_2O (480 µL) and TFA/H_2O (1:9, 960 µL) was added. Acetonitrile (1.5 mL) was then added, followed by a saturated solution of 2,4-dinitrophenylhydrazine in acetonitrile (1.5 mL).[11] The reaction mixture was incubated at room temperature for 30 min and filter (0.2 µm filter). 500 µL of the crude filtrate was then analysed by HPLC (YMC Chiral Amylose-SA Classical Analytical S-5 column (250 mm × 4.6 mm) with a flow rate of 0.5 mL min^{-1} and a gradient from n-hexane/ethanol (100:0 to 0:100) over 120 min; see Figure S35) and the remaining solution was purified by preparative TLC (eluting first with 100% ethyl acetate, Rf = 0.7; and second with 5% methanol/dichloromethane (5:95), Rf = 0.4)[12]. The purified glyceraldehyde dinitrophenylhydrazone where then analysed by HPLC (YMC Chiral Amylose-SA Classical Analytical S-5 column (250 mm × 4.6 mm) with a flow rate of 0.5 mL min^{-1} and a gradient from n-hexane/ethanol (100:0 to 0:100) over 120 min; see Figure S36)[12]. The retention times for L-glyceraldehyde dinitrophenylhydrazone and D-glyceraldehyde dinitrophenylhydrazone were 69 min and 91 min, respectively. L-glyceraldehyde/D-glyceraldehyde (1:3.1 e.r.) was observed.

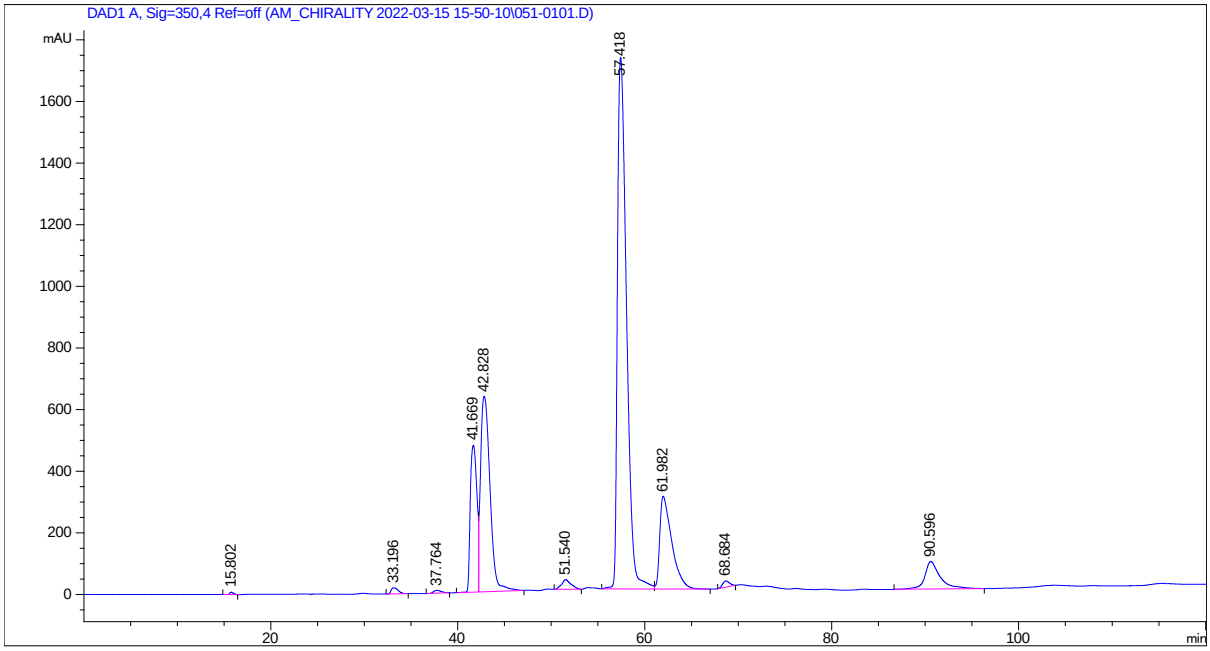


Figure S36. Chiral HPLC chromatogram of crude glyceraldehyde dinitrophenylhydrazone (15 µL injection).

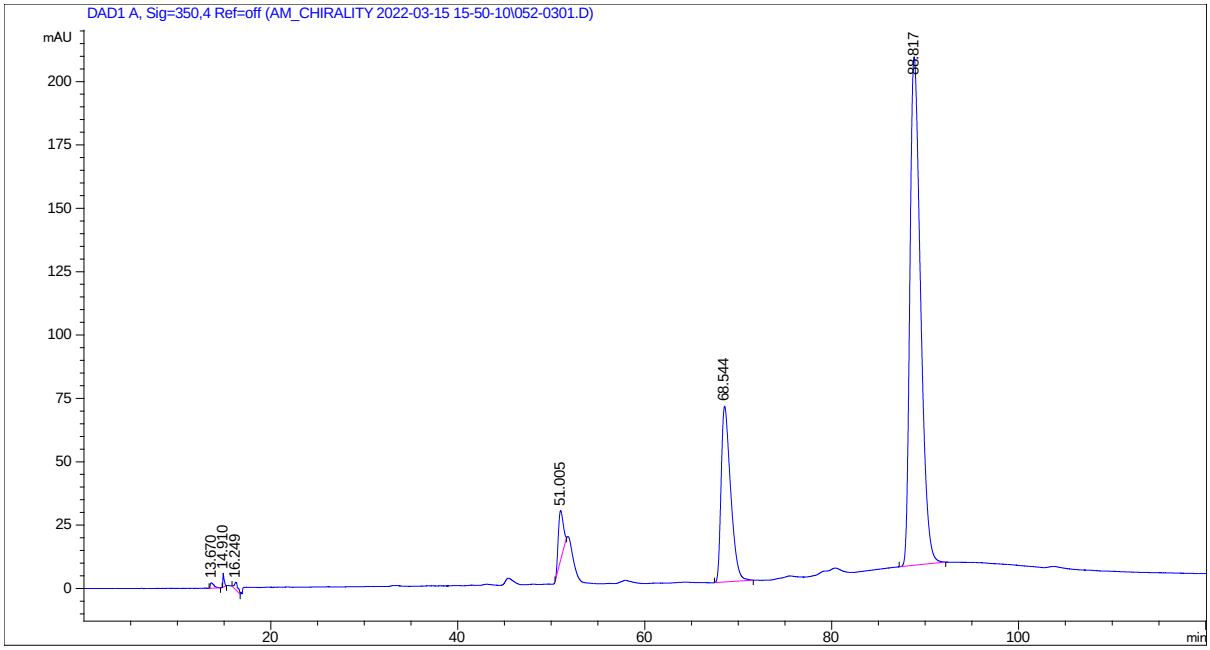


Figure S37. Chiral HPLC chromatogram of preparative TLC purified glyceraldehyde dinitrophenylhydrazone (40 µL injection).

Phosphatase control experiment

rac-Glyceraldehyde 2-phosphate (**7**, 10.7 mg, 50 µmol) in H_2O (1 mL) at pH 4.8 was incubated with acid phosphatase (9 mg) at 37ºC for 10 min. After that time, 160 µL of the mixture was quenched with TFA in H_2O (1:9, 320 µL). The reaction we diluted with acetonitrile (500 µL) and a saturated solution of 2,4-dinitrophenylhydrazine in acetonitrile (500 µL) was added.[11] The reaction was incubated at room temperature for 30 minutes and then filtered (0.2 µm filter). The solution was analysed by HPLC (YMC Chiral Amylose-SA Classical Analytical S-5 column (250 x 4.6 mm)) with a flow rate of 0.5 mL/min and a gradient from n-hexane/ethanol (100:0 to 0:100) over 120 minutes.[12] The retention times for L-glyceraldehyde dinitrophenylhydrazone and D-glyceraldehyde dinitrophenylhydrazone were 69 and 89 min, respectively. L-glyceraldehyde/D-glyceraldehyde (1:1) was observed.

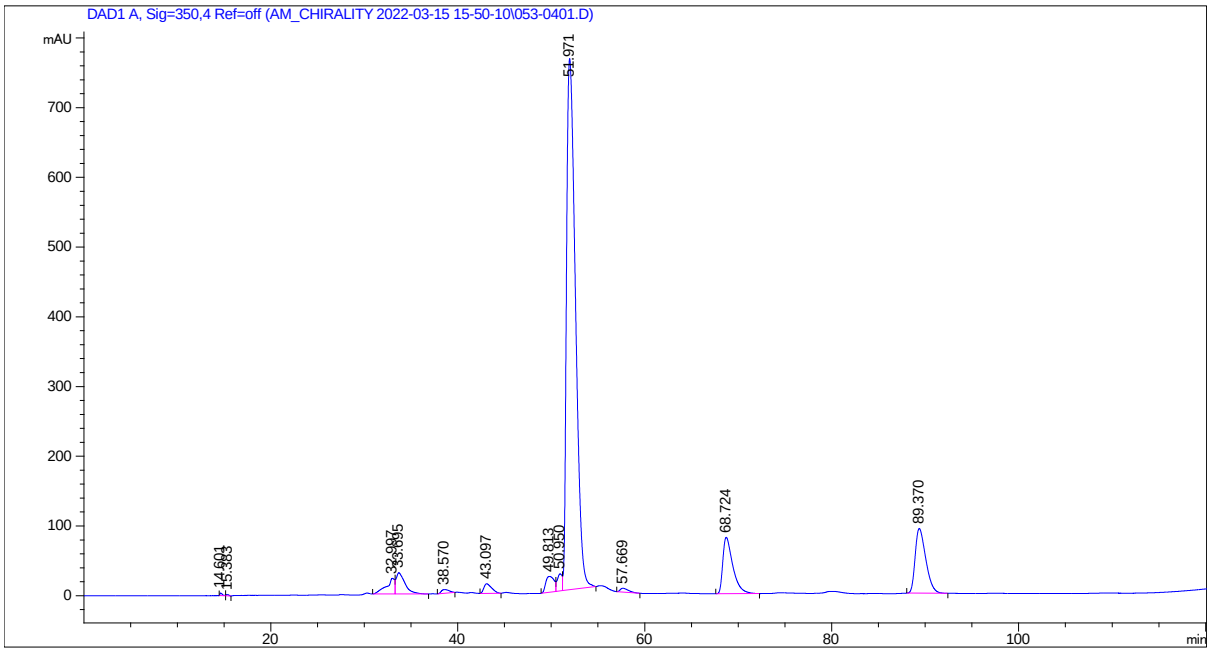


Figure S38. Chiral HPLC chromatogram of glyceraldehyde (5 µL injection).

GAB catalysis of P_i – high salt concentration control

Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in 0.5 M NaCl (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were then acquired regularly over 3 d. The evolution of the reaction can be observed in Figure S39.

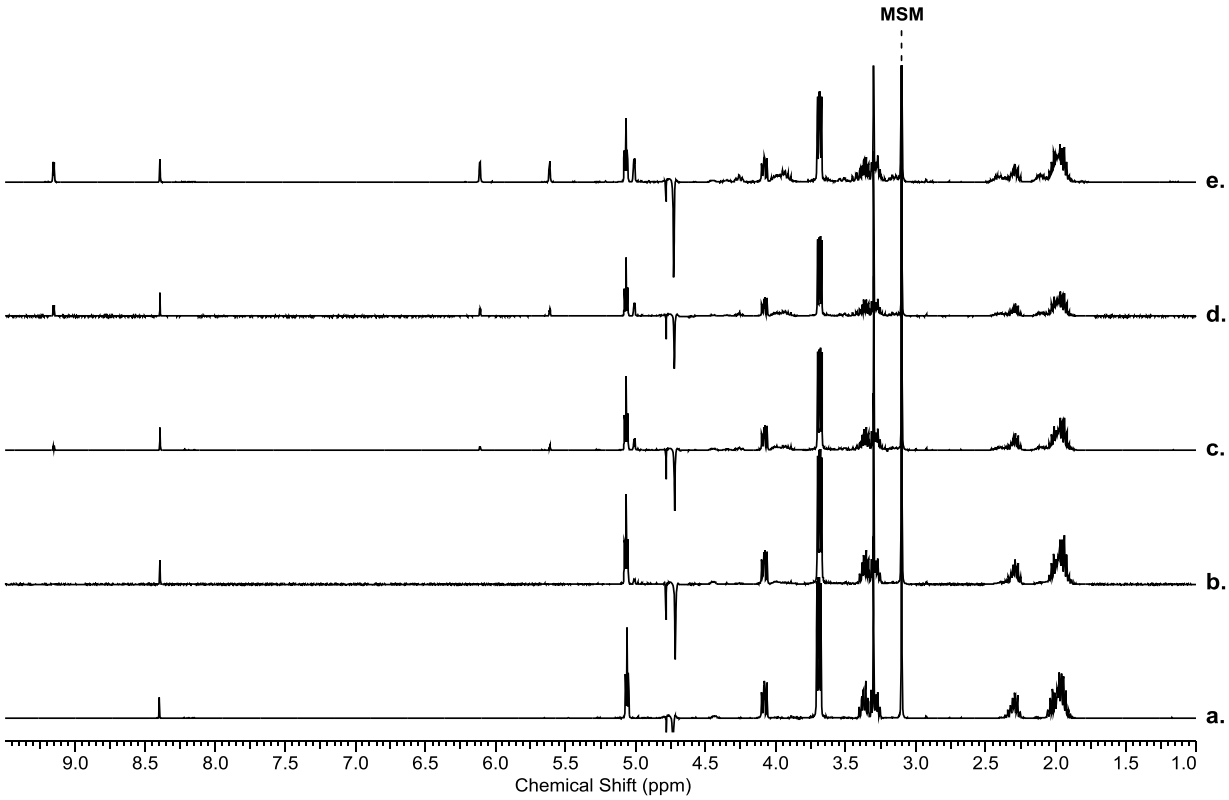


Figure S39. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in sodium chloride aqueous solution (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 7 h; **c.** 26 h; **d.** 48 h; **e.** 72 h.

Compatibility with carbonate-rich lakes[b]

Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in carbonate saturated phosphate buffer[c] (700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. After 48 h, [**7·Pro**] (9%) and **9** (60%) were observed. The evolution of the reaction can be observed in Figure S40.

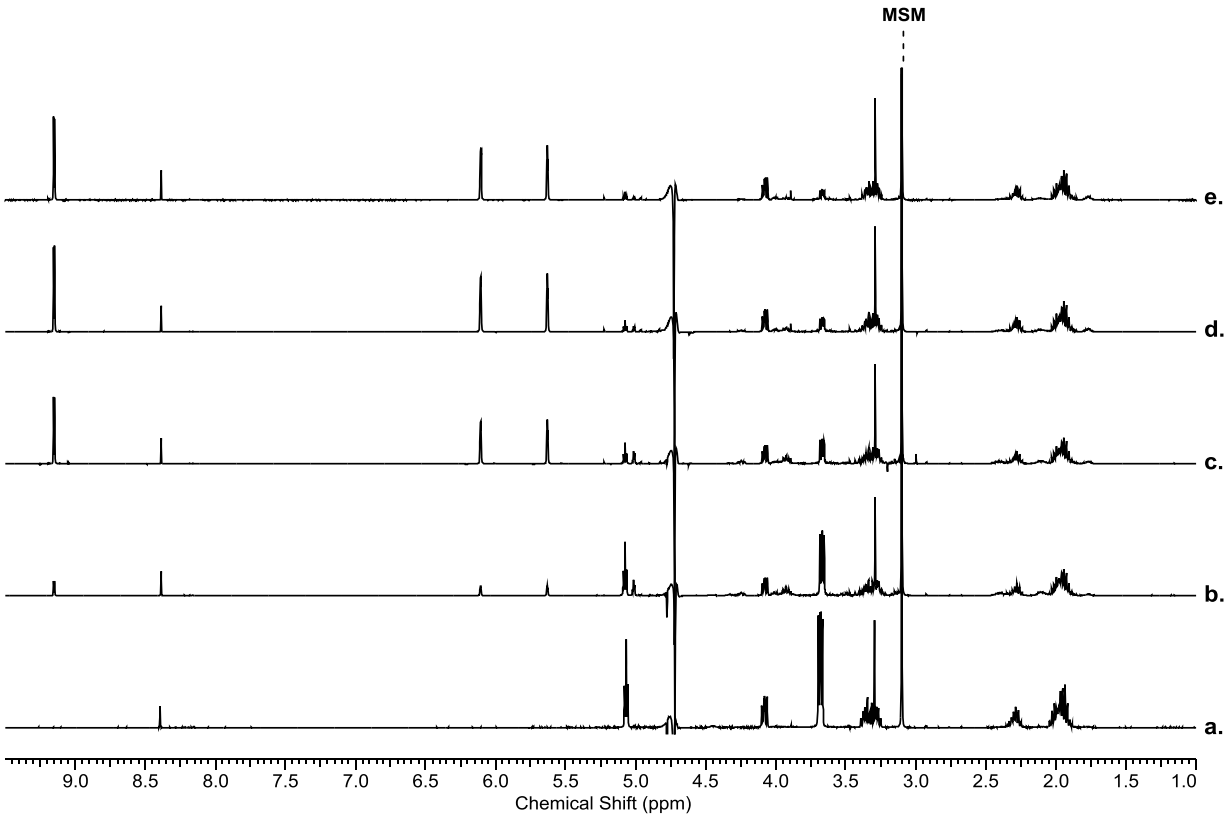


Figure S40. ^{1}H NMR (400 MHz, H_2O/D_2O 9:1, noesygppr1d, 1.00 - 9.50 ppm) to show the reaction of glycolaldehyde-phosphate (**6**, 100 mM), formaldehyde (**8**, 100 mM) and L-proline (50 mM), with MSM (25 mM) as internal standard, in carbonate saturated phosphate buffer (500 mM, pH 7) at 40ºC after: **a.** 0 h; **b.** 7 h; **c.** 26 h; **d.** 48 h; **e.** 72 h.

[b] Control reaction with carbonate buffer was performed and showed equal result to reaction in water: Glycolaldehyde-phosphate (**6**, 37.0 µL, 75.0 µmol) and L-proline (4.3 mg, 50 mol%, 37.5 µmol) were dissolved in carbonate buffer (500 mM, 700 µL). Formaldehyde (**8**, 5.6 µL, 75 µmol) and MSM (1 M, 18.8 µL, 18.75 µmol) were added and the solution pH was measure and adjusted to pH 7, if necessary. To the reaction (495 µL) was added D_2O (55 µL) to afford a H_2O/D_2O (9:1) solution and then incubated at 40ºC for 72 h. NMR spectra were acquired at the beginning and end of reaction.

[c] Carbonate saturated phosphate buffer was produced by addition of inorganic carbonate to a solution of 0.5 M phosphate buffer until saturation. The solution was then filtered and used in the reaction.

References

1 R. T. Raines, M. D. Shoulders and J. A. Hodges, US Pat., 2007275897A1, 2007.
2 D. Auriel, H.-G. Lerchen, D. Immler, J. Wesener and A. Schumacher, EU Pat., 1477493A1, 2004.
3 G. B. Fields, in *Peptide Synthesis Protocols*, M. W. Pennington and B. M. Dunn, Humana Press, Totowa, NJ, 1995, Methods for Removing the Fmoc Group, pp 17–27.
4 M. Awada and P. C. Dedon, *Chem. Res. Toxicol.*, 2001, **14**, 1247–1253.
5 C.-H. Wong, A. Pollak, S. D. McCurry, J. M. Sue, J. R. Knowles and G. M. Whitesides, in *Methods in Enzymology*, Academic Press, 1982, Vol. 89, [18] Synthesis of ribulose 1,5-bisphosphate: Routes from glucose 6-phosphate (via 6-phosphogluconate) and from adenosine monophosphate (via ribose 5-phosphate), pp 108–121.
6 D. Müller, S. Pitsch, A. Kittaka, E. Wagner, C. E. Wintner and A. Eschenmoser, *Helv. Chim. Acta*, 1990, **73**, 1410–1468.
7 M. Angrick and D. Rewicki, *Liebigs Ann. Chem.*, 1982, **1982**, 366–371.
8 J. Singh, N. Steck, D. De, A. Hofer, A. Ripp, I. Captain, M. Keller, P. A. Wender, R. Bhandari and H. J. Jessen, *Angew. Chem. Int. Ed.*, 2019, **58**, 3928–3933.
9 C. Anastasi, F. F. Buchet, M. A. Crowe, M. Helliwell, J. Raftery and J. D. Sutherland, *Chem. Eur. J.*, 2008, **14**, 2375–2388.
10 A. J. Coggins, PhD thesis, University College London (UK), 2017.
11 J. Yu, A. X. Jones, L. Legnani and D. G. Blackmond, *Chem. Sci.*, 2021, **12**, 6350–6354.
12 R. Breslow, V. Ramalingam and C. Appayee, *Origins Life Evol. Biospheres*, 2013, **43**, 323–329.