


 Cite this: *Chem. Commun.*, 2022, 58, 13519

 Received 6th October 2022,
 Accepted 14th November 2022

DOI: 10.1039/d2cc05466c

rsc.li/chemcomm

Proline and phosphate promote a near-quantitative aldol reaction between glycolaldehyde phosphate and formaldehyde at neutral pH in water. Our results demonstrate the important role of general acid-base catalysis in water and underscore the essential role that amino acid catalysis may have played in early evolution of life's core metabolic pathways.

Metabolism is orchestrated by highly evolved, genetically encoded enzymes. However, at the origins of life, before the advent of enzyme catalysis, the chemical reactions that sustained life must have been organized by the innate reactivity of simple abiotic molecules.^{1a–o} Therefore, the boundary between intrinsically and enzymatically controlled reactions lies at the heart of elucidating the origins of life. The triose glycolysis pathway (TGP) serves as a central pillar of metabolism. It is one of life's most highly conserved and densely connected metabolic pathways.^{2a,b} This underscores the TGP's deep-seated antiquity,^{3a,b} but also raises the fascinating possibility that biological triose glycolysis was foreshadowed by a prebiotic reaction pathway, and that its central role in metabolism was predisposed by the inherent reactivity of simple sugars.^{4a,b}

In the TGP, glyceraldehyde-3-phosphate (**1**) is first oxidised to glyceric acid-3-phosphate (**2**), the phosphate must then be enzymatically migrated to yield glyceric acid-2-phosphate (**3**), before enzymatic elimination yields phosphoenol pyruvate (**4**) (Fig. 1a). We have previously identified the constitutional relationship between glyceraldehyde-2-phosphate (**7**) and the carbon framework of **4**,^{4a} demonstrating phosphorylation of glycolaldehyde **5** – which yields glycolaldehyde phosphate (**6**) – controls a network of reactions that delivered not only **4**, but also the other components of the TGP. Rerouting the metabolic pathway – *i.e.*, switching the order of oxidation and elimination – yielded a chemically predisposed sequence of

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

 Álvaro F. Magalhães  and Matthew W. Powner *

reactions in which elimination occurs at the aldehyde oxidation level from **7** (Fig. 1b).

This juxtaposition between predisposed and enzymatically activated pathways, opens the intriguing (general) question of how predisposed chemical reactions would first become catalytically controlled, en route to becoming embedded within an enzyme-gated biochemical network. We therefore set out to investigate the relationship between simple, prebiotically, and biologically relevant, catalysts and the key aldol reaction required to transition between sugar **5** and phosphoenol **4**.

Aldolase enzymes greatly enhance the rate and stereoselectivity of (biochemical) aldol reactions,^{5a,b} and can be divided into two mechanistic classes. Type I aldolases have a catalytic amine in their active site and form an enamine intermediate that accelerates their cognate aldol reaction, whilst Type II aldolases coordinate a



Fig. 1 Triose glycolysis. (a) Biological enzyme (Enz.) catalysed triose glycolysis pathway (TGP); (b) proposed prebiotic triose glycolysis. (c) Proline-catalysed triose glycolysis.

Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK. E-mail: matthew.powner@ucl.ac.uk

 † Electronic supplementary information (ESI) available: Experimental procedures and spectroscopic data. See DOI: <https://doi.org/10.1039/d2cc05466c>


Lewis acid cofactor, such as Zn^{2+} , to accelerate an aldol reaction.⁶ Aldolase enzymes are clearly the product of evolution, and too structurally complex to have been available to control energy metabolism at the origins of life. However, small molecule catalysts must have been prebiotically available. If these small molecules endowed enhanced reactivity, they could then have been acted upon by evolution to provide an evolutionary trajectory to modern enzyme catalyzed metabolism.

The remarkable (type I) aldol activity of proline (**Pro**) in organic solvents,^{7a-d} and the high intracellular concentrations of amino acids, have suggested to some that *in vivo* natural products synthesis may be influenced by amino acid catalysis and, by extension, that organocatalysis may lie at the root of biological (enzyme) catalysis and even biological homochirality.^{1e,8a-i} However, despite the magnitude of these questions, and a large body of work attempting to address organocatalysis in water,^{8a-i,9a-d} this concept remains elusive. Most organocatalytic reactions employ heterogeneous media (rather than water) exploiting *e.g.* non-aqueous additives, co-solvents, extreme (>molar) substrate concentrations, emulsions or amphiphiles due to the challenges imposed by implementing organocatalysis in dilute water.^{8i,9a-d,10a-c} Although Clarke and co-workers have reported **Pro** is incapable of catalyzing the dimerization of glycolaldehyde (**5**) in water,^{8g} **Pro** does catalyze the aldol reaction of acetone and *p*-nitrobenzaldehyde in aqueous solution [with organic co-solvent, *i.e.* 10% DMSO].^{11a,b} **L-Pro** has also been reported to yield an unsubstantiated (small) quantity of enantioenriched glyceraldehyde from the reaction of glycolaldehyde (**5**) and formaldehyde (**8**).^{12a-e} Together these results highlight the ineffective nature of **Pro** catalysis in water, but hint at the potential for **Pro**, or a related small molecule catalyst, to enhance the rate of the key aldol reaction in prebiotic triose glycolysis (Fig. 1c) and provide a (small molecule) chemical link towards the evolution of enzyme-catalyzed metabolism.

The key aldol reaction of the prebiotic TGP, between glycolaldehyde phosphate (**6**, 100 mM) and formaldehyde (**8**, 10 equiv.) yields **7** (66%) in alkaline solution (pH 10.7) at room temperature (rt) after 6 days.¹³ However, with no catalyst, no reaction was observed at neutral pH (Fig. 2a and ESI,† Fig. S15, S16). We suspected neutral pH reactivity would be established with the appropriate catalyst, so we set out to study the effect of amino acids on this reaction.^{10a-c,14a,b}

We initially investigated glycine (**Gly**) and **Pro** (50 mol%, pH 7, rt; ESI,† Fig. S15). In water only **Pro** was observed to have significant reactivity, yielding **7** as a proline-hemiaminal [**7-Pro**] (10%) (Fig. 1), alongside trace amounts of the elimination product **9** (ESI,† Fig. S17), but the conversion was low with the remaining mass balance (89% **6**) unreacted after 14 days.

We suspected that [**7-Pro**] was not equilibrating with **7**. This was confirmed by studying authentically synthesized **7**.¹⁵ Upon incubation with **Pro**, **7** was observed to directly dehydrate to **9**, but only yield small amounts of [**7-Pro**] (Fig. 3 and ESI,† Fig. S33–S35). This low concentration of [**7-Pro**] (8%) stands in stark contrast to the observation that [**7-Pro**] (27%) was a more major compound (at intermediate time points) in the **Pro**-catalysed reaction of **6** and **8**, where only trace quantities of



Fig. 2 Proline/phosphate co-catalysed reaction of glycolaldehyde phosphate and formaldehyde. ¹H NMR [400 MHz, noesygpplrd, H₂O/D₂O (9 : 1)] spectra to show the reaction of glycolaldehyde phosphate (**6**, 100 mM) with formaldehyde (**8**, 100 mM) after 48 h at pH 7 and 40 °C: (a) in water; (b) with 50 mol% **L-Pro** in water; (c) with 50 mol% **L-Pro** in 500 mM phosphate.

aldehyde **7** (<1%) were observed. Furthermore, the dehydration of **7** was not notably enhanced by **Pro**; 40% and 43% conversion of **7** (25 mM) to **9** was observed with and without **Pro** (50 mol%) after 4.5 days at 40 °C and pH 7, respectively (ESI,† Fig. S33 and S35). It is likely the highly electron-withdrawing α -phosphate retards the rate of hemiaminal-to-aldehyde equilibration, and trapping **Pro** in hemiaminal [**7-Pro**] would clearly block turnover and suppress catalytic activity.¹⁶ This suggested enhancing the turnover and release of **Pro** would be highly beneficial.

Exploring other amino acids, amino amides and simple peptides did not lead to improved activity (ESI,† Fig. S15). Indeed, the activity of **Pro** was depressed by transformation into its methyl ester and nitrile derivatives (ESI,† Fig. S15), and so we continued to focus on **Pro**-catalysis.

Disappointed, but not surprised, by the poor organocatalytic activity observed in neutral water, we reflected on the distribution of charged moieties (*e.g.* in catalytic triads) in enzyme active sites.^{18a-d} We postulated a combination of enamine catalysis and general acid-base (GAB) catalysed proton transfer would be the key to effective aldol activity at neutral pH. Within enzyme active sites GAB catalysis is mediated by the structural proximity of acidic and basic moieties,^{19a,b} however, without the structured scaffold of a complex, coded polypeptide, we recognized this activity could be facily replaced by an appropriate buffer. Pleasingly, when co-catalytic phosphate (**P_i**)^{14a,20a,b} was introduced the conversion to the desired products markedly (3-fold) increased, such that co-catalytic **Pro** (50 mol%) and **P_i** (500 mM, pH 7) yielded [**7-Pro**] (17%) and phosphoenol **9** (31%) after 14 days at rt (ESI,† Fig. S20). **P_i** was





Fig. 3 Proline-catalysed aldol-elimination cascade. (a) ^1H NMR [400 MHz, noesygpprld, $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9 : 1)] spectrum to show the reaction of glycolaldehyde phosphate (**6**, 100 mM) with formaldehyde (**8**, 100 mM) in phosphate buffer (500 mM, pH 7) at 40 °C after 48 h; (b and c) ^1H NMR [700 MHz, noesygpprld, $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9 : 1)] spectra to show the reaction of glyceraldehyde-2-phosphate (**7**, 100 mM) in phosphate buffer (500 mM, pH 7) with 50 mol% *L*-Pro at 40 °C after: (b) 0.5 h; (c) 48 h; (d) ^1H NMR [400 MHz, noesygpprld, $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9 : 1)] spectrum to show the reaction of **6** (100 mM) with **8** (100 mM) in phosphate buffer (500 mM, pH 7) with 50 mol% *L*-Pro at 40 °C after 7 h.

observed to enhance the performance of all amine catalysts tested (ESI,† Fig. S18). To confirm *P_i* was acting as a GAB catalyst we investigated the effect of *P_i* concentration and, as expected, increased *P_i* concentration enhanced the reaction (Fig. 4a).

Previously a high concentration and large excess of **8** (10 equiv.) has been employed in these aldol reactions,^{4a,13} however, in the *Pro* catalysed reaction we observed a dramatic improvement in the conversion of **6** to **9** when the initial concentration of **8** was decreased. Indeed, the maximum yield of **9** (Fig. 4b) was observed with 1 equiv. of formaldehyde **8**. This stoichiometric (1 : 1) reaction, catalyzed by *Pro*/*P_i*, thus afforded the highest reported yield of **9**. It seemed likely **8** could suppress the aldol reaction through competitive sequestration of *Pro* (into *e.g.* [**8-Pro**], ESI,† Fig. S32), but it was perhaps more noteworthy that higher initial concentrations of **8** resulted in more rapid degradation of **9**, suggesting **8** accelerates the hydrolysis of **9**. In line with these observations, we next tested setting **8** as the limiting reagent, and pleasingly we found that substoichiometric **8** (0.5 equiv.) led to a near-quantitative combined yield of [**7-Pro**] (18%) and **9** (76%) (ESI,† Table S6).

Finally, *L-Pro* was observed to induce enantioenrichment of **7**, furnishing more D-7 (ee = 50%; ESI,† Fig. S36–S38). This reaction pairs *L*-amino acid (*i.e.* *L-Pro*) catalysis with *D*-sugar-phosphates, mirroring life's symmetry relationship between



Fig. 4 Proline-catalysed synthesis of phosphoenol pyruvaldehyde. Graphs to show the effect of buffer and stoichiometry. (a) 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; (b) 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.

amino acids and sugar-phosphates. It is therefore of note that **7** is the first (simplest) chiral sugar-phosphate that can be synthesized by an aldol reaction from two achiral precursors.

The co-catalytic action of phosphate and amino acids, but especially *Pro*, markedly accelerate phosphoenol synthesis in the reported prebiotic TGP.^{4a} These results highlight, the often overlooked, importance of GAB catalysis for promoting proton transfer in neutral water, and how GAB catalysts can be readily paired with a complementary co-catalyst to enhance the co-catalysts activity and promote otherwise ineffective reaction. GAB catalysis is an essential element of enzyme activity because enzymes have evolved to operate in neutral water, and so GAB catalysis must also be essential to understanding the transition from chemical to biochemical reactivity.

It is also of note, with respect to aldol catalysis, that even in water a simple small molecule catalyst, like *Pro*, can remain substrate-bound following aldol addition. The GAB activity of *P_i* likely facilitates proton transfer during enamine catalysis, but also, importantly, promotes the elimination of [**7-Pro**] to complete the catalytic cycle (Fig. 1c). Geochemical models suggest high concentrations of *P_i* (> 1 M) would accumulate coincidentally with high concentrations of carbonate in natural (prebiotic) environments.²¹ Interestingly, we observe carbonate has no detrimental effect on the observed *Pro*/*P_i* catalysed aldol (ESI,† Fig. S39 and S40). Further investigation of the co-catalytic action of chiral amines and phosphates, as well as the potential to enhance chiroselectivity through (hemi-)aminal dynamic



