Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water

Callum S. Foden[1,2], Saidul Islam[1,2], Christian Fernández-García[1], Leonardo Maugeri[1], Tom D. Sheppard[1] and Matthew W. Powner[1]*

[1] Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, U.K.

[2] These authors contributed equally to this work

Peptide biosynthesis is orchestrated by a complex suite of enzymes, but this must have been predated by a simple chemical synthesis at the origins of life. α-Aminonitriles—prebiotic α-amino acid precursors—are generally produced by Strecker reactions. However, cysteine's aminothiol is incompatible with nitriles. Consequently, cysteine nitrile is not stable and therefore cysteine has long been believed to be a product of evolution, not prebiotic chemistry. We report the first high-yielding, prebiotic synthesis of cysteine peptides. Our biomimetic pathway converts serine to cysteine by nitrile–activated dehydroalanine synthesis. We additionally demonstrate that *N*-acylcysteines catalyze peptide ligation, directly coupling kinetically stable—but energy-rich—α-amidonitriles to proteinogenic amines. This rare example of selective and efficient organocatalysis in water implicates cysteine at the onset of life's evolution.

Peptides and proteins are essential to all life on Earth, but their biosynthesis is achieved by a highly evolved system of enzyme catalysts[1,2]. Although the origin of protein synthesis in biology remains a mystery, life's exploitation of peptides must have predated the evolution of the complex enzymes that are now required to coordinate biosynthesis. At the origins of life, simple chemical processes must have furnished the first peptide catalysts that later evolved to become modern enzymes[3]. To elucidate these chemical mechanisms, it is essential that we reflect on the biochemical strategies of extant biology to inform the systematic evaluation of prebiotic chemistry. For example, we recently demonstrated α-peptide synthesis in water could be achieved by H_2S-mediated stoichiometric ligation[4]. Our synthesis exploited a biomimetic N-to-C terminal chain-growth mechanism. This overcame long-standing problems that had prohibited the coupling of α-aminonitriles[5-7], and avoided the irrevocable side chain modifications caused by electrophilic agents required to activate amino acids[8,9]. However, our sulfide-mediated peptide synthesis was not catalytic, and each ligation step required fluctuation between reducing and oxidizing conditions. Further reflection on the deep-seated role of thiols in non-ribosomal peptide synthesis and core metabolism[10,11] has now led us to hypothesize that cysteine may have originated as a secondary product of sulfide-mediated peptide synthesis, and we suspected that cysteine could be used to deliver (a prebiotically plausible) catalytic peptide ligation (CPL) in water.

Cysteine (**Cys**) is the primary organic source of sulfide in biology and the feedstock for essential cofactors such as glutathione (**GSH**) and co-enzyme A (**CoA**)[12]. It is also an important residue within enzyme active sites, with vital functions in catalysis, redox sensing, and electron transfer, as well as being an essential ligand in ancient iron-sulfur proteins[11,13,14]. It therefore seems almost inconceivable that cysteinyl thiols were not present during the development of nascent biological processes on the early Earth, and yet this is not the prevailing view[15-18]. Numerous, unsuccessful attempts to synthesize and isolate cysteines under prebiotically plausible conditions[7,19-21] have led to a widely-held belief that cysteine is a biological invention[17,18], as well as a late addition to the genetic code[15-18]. Here, we report a high-yielding prebiotic synthesis of cysteines and demonstrate that these cysteine peptides catalyze non-enzymatic CPL in neutral water. Our results support the hypothesis that cysteine (**Cys**) was available at the origins of life as a secondary product of serine nitrile synthesis, and that cysteines would have been a cornerstone of early catalytic activity.

Results

Biomimetic prebiotic cysteine synthesis

The Strecker synthesis of aminonitriles is widely believe to play an important role in the prebiotic origins of amino acids, and we had previously identified glycolaldehyde (**GCA**) and β-mercaptoacetaldehyde (**BMA**) as key nodes in the chemical network required for abiogenesis of RNA and peptides[3,22,23], and specifically **BMA** as a Strecker precursor of **Cys**[22]. However, although Strecker reactions are generally highly efficient[23,24], **BMA** forms intractable and insoluble mixtures in Strecker reactions (fig. S1)[7]. This is in stark contrast to **GCA**, which undergoes the Strecker reaction in excellent yield to afford a stable aminonitrile product, serine nitrile **Ser-CN** (fig. S2)[23,24]. The observed disparity between congeners **GCA** and **BMA** is likely due to the rapid reaction of β-aminothiols with nitriles in water[25], and because cysteine nitrile (**Cys-CN**) is both an β-aminothiol and an α-aminonitrile it is inherently unstable[7]. These observations suggested to us that whilst the Strecker reaction of **GCA** could yield serine (**Ser**) under prebiotic conditions, the Strecker reaction of **BMA** could not play a role in prebiotic **Cys** synthesis. To overcome this inherent stability problem, we therefore considered an alternative biomimetic pathway for **Cys** synthesis. We suspected that clues to the prebiotic synthesis of **Cys** might remain embedded within the extant biological pathway and that **GCA**, rather than **BMA**, was the key prebiotic precursor of **Cys**.

The principal mechanism by which reduced inorganic sulfur is incorporated into bioorganic compounds is through **Cys** biosynthesis[17,18]. In plants and various archaea and bacterial species, **Cys** biosynthesis begins with the enzymatic conversion of **Ser** to *O*-acetylserine (**Ser^{Ac}**) [or *O*-phosphoserine (**Sep**)] followed by a pyridoxal-5′-phosphate (**PLP**) dependent acetic acid β-elimination, and then sulfide β-addition before disassociation of the cysteine-enzyme complex (Fig. 1A). Throughout this sequence, the α-amine remains

covalently bound to **PLP** to promote elimination and to prevent the release and rapid decomposition of highly unstable dehydroalanine (**Dha**)[26]. The nitrile equivalent of **Dha**, dehydroalanine nitrile (**Dha-CN**), has been synthesized by Eschenmoser and co-workers, but was found to be extremely unstable and failed to react with H_2S to produce **Cys-CN**, instead undergoing rapid degradation, even under anhydrous conditions[27]. However, we have recently shown that *N*-acylation of α-aminonitriles is a crucial element in initiating peptide synthesis in water, and it prevents hydantoin-, diketopiperazine- and imidazole-induced peptide degradation[4,5,6]. Similarly, *N*-acylation would stabilize **Dha-CN**. *N*-Acylation would simultaneously prevent the highly favourable, but unwanted, enamine→imine tautomerization that precludes sulfur addition to **Dha**[27], and prevent the degradation of **Cys-CN** that is brought about by its free α-amine[7]. Consequently, we recognized that serine diacylation presents a simple biomimetic strategy for prebiotic dehydroalanine synthesis, and we identified *N,O*-diacetyl-serine nitrile (**Ac-Ser^{Ac}-CN**) as a key intermediate for prebiotic **Cys** synthesis (Fig. 1B). In addition to the electron-withdrawing effects of the α-nitrile, *N*,*O*-diacetylation would further enhance the acidity of the α-proton of **Ac-Ser^{Ac}-CN** as well as activating the serine hydroxyl moiety as a leaving group. We envisaged that these combined effects would promote **Ac-Dha-CN** synthesis at neutral pH, without recourse to the highly alkaline (pH >13) conditions typically required for **Dha** formation in water[28] that would also promote peptide degradation[29].

We have a long-standing interest in (prebiotic) acylation[30,31], but it was not clear if the hydroxyl moiety of **Ser-CN** could be selectively acetylated in water. Therefore, we were pleased to observe chemoselective *N*,*O*-acetylation of **Ser-CN** with thioacetic acid (**AcSH**) and ferricyanide[4,30] to produce **Ac-Ser^{Ac}-CN** in up to 91% yield within 1 h at room temperature (Supplementary Pages S13–24). Acetylation of **Ser-CN** with *N*-acetylimidazole (**NAI**)[30,31] was equally effective for **Ac-Ser^{Ac}-CN** synthesis in neutral water. We initially observed *O*-acetylserine nitrile **Ser^{Ac}-CN** as the major product (<10 min) with **NAI** (fig. S8), and attribute the nucleophilicity of the β-hydroxyl of **Ser-CN** to the pronounced electron-withdrawing effect of the α-nitrile and the remarkably low pK_{aH} of this amino-alcohol (Supplementary Pages S23–32). More importantly, we also observed the formation of a stable dehydroalanine nitrile **Ac-Dha-CN** at near-neutral pH for the first time during these acetylation reactions. Optimal formation of **Ac-Dha-CN** from **Ac-Ser^{Ac}-CN** occurred at pH 8, and after 4 d at room temperature an 85% conversion to **Ac-Dha-CN** was observed (Fig. 1C). This elimination is all the more remarkable because *N,O*-diacetylserinamide (**Ac-Ser^{Ac}-NH$_2$**) underwent near-exclusive hydrolysis at pH 8 (fig. S34–35)[32]. The switch in reactivity between α-nitrile (acetate elimination) and α-amide (acetate hydrolysis) demonstrates the benefits of α-nitrile activation for acetate elimination, and indicates that **Ac-Dha-CN** is predisposed to form in near-neutral water[28,29].

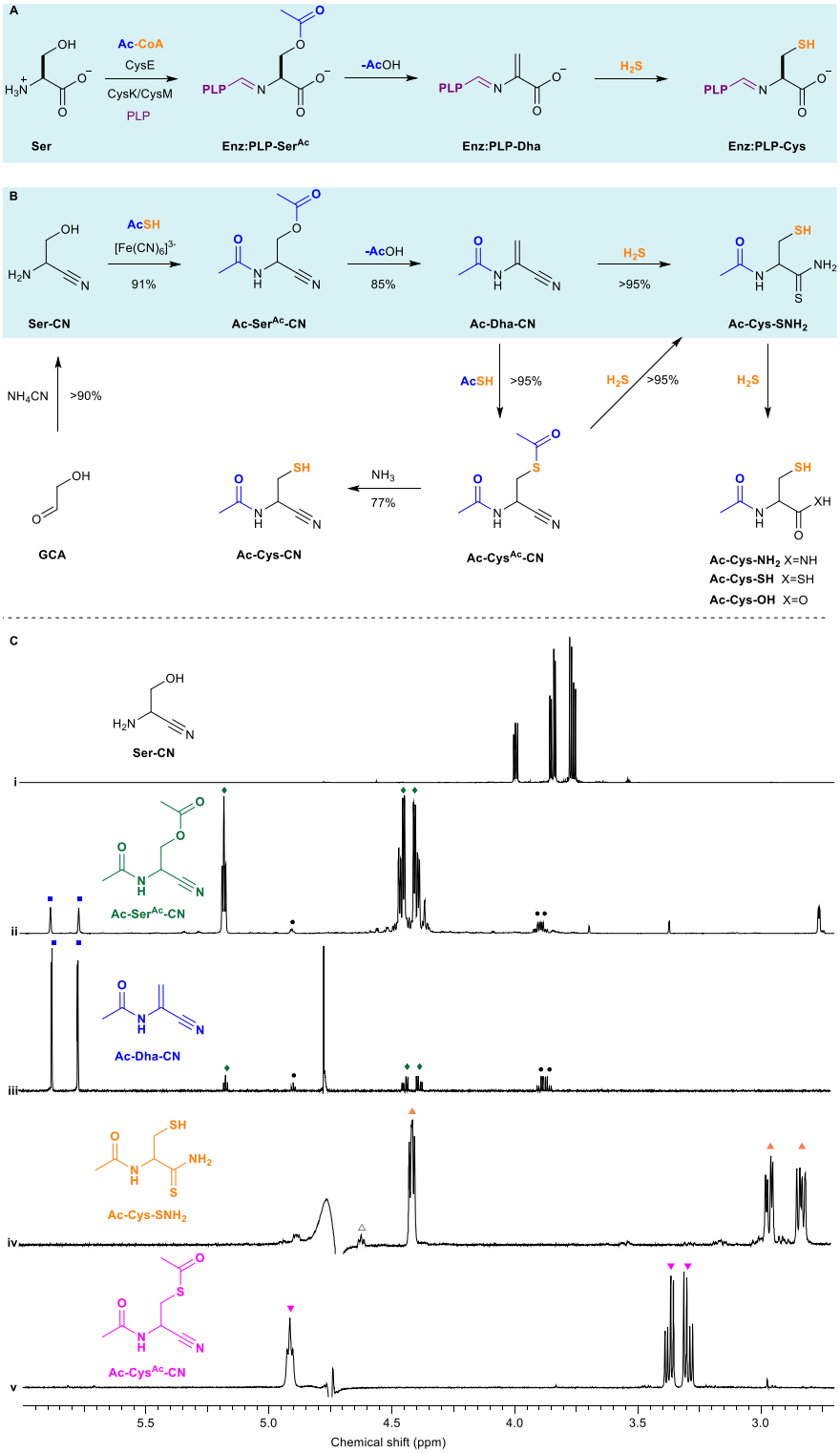

Fig. 1. Prebiotic cysteine synthesis. Biomimetic conversion of serinyl nitriles to cysteinyl nitriles. (**A**) Pyridoxal-5′-phosphate (**PLP**)-dependent enzymatic cysteine synthesis pathway. (**B**) Prebiotic synthesis of cysteines in neutral water starting from **Ser-CN**, the stable Strecker product of **GCA**. (**C**) ^{1}H NMR (700 MHz, H_2O/D_2O) spectra showing: (i) **Ser-CN**; (ii) crude NMR of the reaction of **Ser-CN** (100 mM) and **NAI** (5 equiv., 19 h, pD 7, r.t.) yielding **Ac-Ser^{Ac}-CN** (72%) and **Ac-Dha-CN** (6%); (iii) crude NMR of the elimination of **Ac-Ser^{Ac}-CN** (100 mM, 4 d, pH 8, r.t.) yielding **Ac-Dha-CN** (85%); (iv) crude NMR of the reaction of **Ac-Dha-CN** (60 mM) and H_2S (10 equiv., 4 h, pH 9, r.t.) yielding **Ac-Cys-SNH$_2$** (>95%); (v) crude NMR (600 MHz) of the reaction of **Ac-Dha-CN** (100 mM) and **AcSH** (4 equiv.) in phosphate buffer (500 mM, 12 h, pH 7, r.t.) yielding **Ac-Cys^{Ac}-CN** (>95%). • = **Ac-Ser-CN** Δ = **Ac-Cys-CN**.

We recently demonstrated a high-yielding prebiotic synthesis of phosphoserine nitrile **Sep-CN**[33], so we next investigated **Ac-Dha-CN** synthesis via phosphate elimination. Acetylation of **Sep-CN** in water with **AcSH** and ferricyanide at pH 7 yielded **Ac-Sep-CN** (80%) after 1 h. Phosphoserines typically require alkaline pH and Ba^{2+} to promote **Dha** formation[29], but the formation of **Ac-Dha-CN** (8%) was nonetheless observed after heating **Ac-Sep-CN** at 60 °C for 3 d at pH 7. The sluggish rate of phosphate elimination was enhanced by Mg^{2+}, yielding **Ac-Dha-CN** (24%) after 1 d at 60 ºC (fig. S33). This demonstrates prebiotically plausible dehydroalanine formation can be achieved by **Ser** acetylation or phosphorylation. **Ac-Dha-CN** was found to be highly stable, and we did not observe the addition of acetate, phosphate or hydroxide (even at pH 11) to this dehydroalanine, setting the stage for selective addition of inorganic sulfur to synthesize cysteine.

Pleasingly, **Ac-Dha-CN** underwent near-quantitative conversion to **Ac-Cys-SNH$_2$** upon incubation with H_2S, yielding the first prebiotic synthesis of a stable cysteinyl residue (Fig. 1C). We next investigated the prebiotic acetylating agent, **AcSH**, as a more water-soluble sulfide source at pH 7. Incubation of **Ac-Dha-CN** with **AcSH** led to quantitative thioester formation after 12 h in phosphate buffer (Fig 1C). The cysteine residue was then rapidly liberated by ammonolysis yielding **Ac-Cys-CN** (77%; fig. S44) or thiolysis to yield **Ac-Cys-SNH$_2$** in 95% yield (fig. S39). Prolonged incubation of **Ac-Cys^{Ac}-CN** with H_2S gradually yielded cysteine thioacid **Ac-Cys-SH** and amide **Ac-Cys-NH$_2$** (2:1) (fig. S47–48). High-yielding **Ser-CN→Cys^{Ac}-CN** conversion was also observed for *N*-acetylvalinylserine nitrile (**Ac-Val-Ser-CN**; Supplementary Pages S59–72) demonstrating the efficacy of cysteine synthesis within a sterically encumbered peptide substrate.

Acetyl CoA is biosynthesized from **Cys** and is a universally conserved acetylating agent for protein, carbohydrate and lipid metabolism. The relative simplicity of thioester **Ac-Cys^{Ac}-CN** suggests it may have been exploited as an acetyl CoA analogue and an activated source of acetate in (proto)metabolism[10]. Consequently, we verified the proficiency of **Ac-Cys^{Ac}-CN** as a source of activated acetate. We suspected that acetyl-transfer from **Ac-Cys^{Ac}-CN** to α-aminonitriles (**AA-CN**) would be promoted by their low pK_{aH} (e.g. **Gly-CN** pK_{aH} = 5.4), which renders **AA-CN** neutral and highly nucleophilic at neutral pH[4]. To this end, we incubated **Ac-Dha-CN**, **AcSH** and **Gly-CN** in phosphate buffer at pH 7 and room temperature. We observed *in situ* thioester formation and near-quantitative acetyl-transfer from **Ac-Cys^{Ac}-CN** to **Gly-CN**, producing **Ac-Gly-CN** (81%) and **Ac-Gly-SNH$_2$** (13%) after 3 d (fig. S67). Further roles for *N*-acylcysteine thioesters in the broader context of origins of life are currently under investigation.

Organocatalytic peptide ligations in water

Prebiotic syntheses, like biosyntheses, are necessarily multistep chemical pathways and therefore, like all multistep processes, are susceptible to diminished overall yields unless the individual reactions proceed with remarkably high efficiency[3]. In biosynthesis, enzyme catalysis achieves exquisite selectivity and high yields,

but modern enzymes are a product of billions of years of evolution. In the absence of enzymes, prebiotic chemistry had to initially exploit unevolved and directly accessible alternatives, such as organocatalysts[34]. Small molecule catalysts could have functioned as rudimentary 'enzymes' at the origins of life, and although they have been highly sought-after since the watershed rediscovery of proline catalysis, progress has been greatly hampered by the generally poor activity of organocatalysts in water[35]. However, the importance of nitriles at the origins of life[3], as well as their low background reactivity, suggested they may be ideal substrates for a highly selective and high yielding catalytic peptide ligation (CPL) in water, and warranted further investigation.

Peptide fragment ligations, such as native chemical ligation, are important reactions in chemical biology and synthetic chemistry, facilitating rapid synthesis of longer peptides from smaller subunits (Fig 2A.i)[36-38]. Fragment ligations have been proposed to play an important role in prebiotic chemistry[39], but the plausibility of these are diminished by the need for synthetically-prepared C-terminal thioesters[40]. Although we have recently demonstrated a high yielding prebiotic fragment ligation[4], this stepwise sequence requires stoichiometric H_2S and is followed by a stoichiometric activating agent (such as ferricyanide or cyanoacetylene) (Fig 2A.ii). Ideally, prebiotic fragment ligations through direct coupling of nitrile and amine fragments would side-step thioamide and thioacid intermediates[4]. Therefore, we became intrigued by the potential of thiols, such as *N*-acylcysteines, to act as organocatalysts for CPL. Reversible thiol addition to an α-amidonitrile could render peptide synthesis catalytic and redox-neutral via a highly reactive – but transient – thioimidate intermediate (Fig. 2A.iii). We anticipated that these thioimidates would be significantly more reactive than the thioesters conventionally exploited in peptide ligations (Fig 2A.i)[36,37], but would also be inherently protected from hydrolysis by reversible thiol-to-nitrile addition, and elimination back to the stable nitrile substrate. CPL would also remove the necessity for stoichiometric H_2S and ferricyanide activation during peptide fragment ligation, and therefore alleviate the geochemical constraints imposed by their mutual incompatibility, and eradicate the need to recycle spent reagents across a redox gradient[4]. Overcoming these challenges would yield a novel CPL strategy and further validate the critical role of the nitrile moiety for powering peptide ligation at the origin of life.

Eschenmoser and co-workers had previously reported **Cys**–catalyzed ammonolysis of α-amidonitriles in methanol[41]. In water, however, the addition of **Cys** to α-amidonitriles near-quantitatively yields thiazolines (Supplementary Pages S75–88), following the mechanism implicated in **Cys-CN** self-degradation[7]. Thiazoline formation is, however, completely suppressed by cysteine *N*-acylation[42], suggesting that our biomimetic **Cys** synthesis—which necessarily yields *N*-acyl-cysteines—is predisposed to furnish catalytically active cysteines. Therefore, we next tested **Ac-Cys-OH** as a catalyst for peptide ligation. We incubated **Ac-Gly-CN** and **Gly** with **Ac-Cys-OH** and, remarkably, observed CPL in water for the first time. This reaction yielded peptidyl amidine **Ac-Gly^{N}-Gly-OH** (60%) after 24 h at 60 °C and pH 7 (Fig. 3).

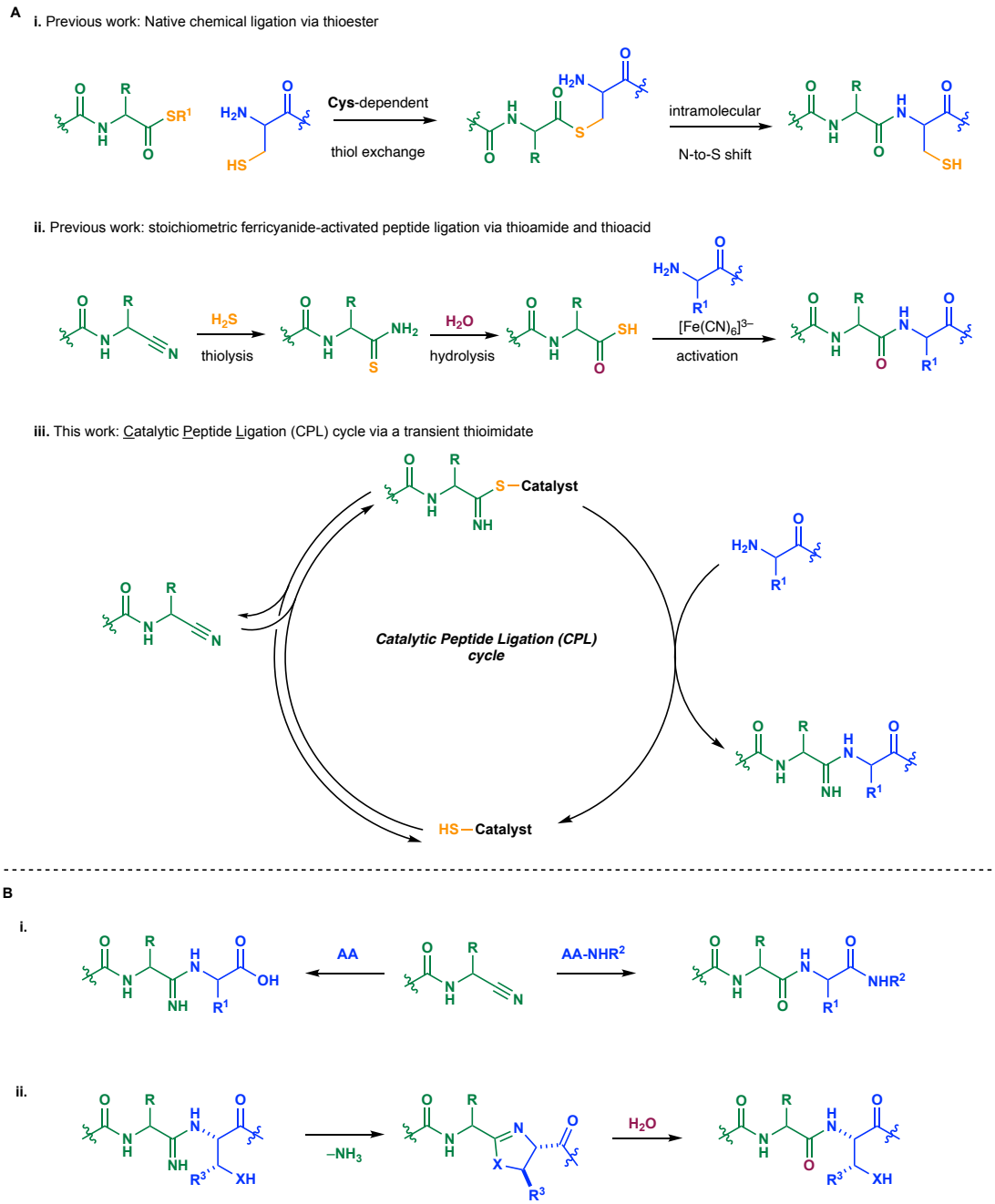

Fig. 2. Thiol-catalyzed peptide synthesis in water. (A) i. Previous work: Native chemical ligations require *C*-terminal peptide thioesters and an N-terminal cysteine to afford a peptide bond by thiol-exchange and intramolecular amidation. R^1=aryl or alkyl[36,37]. **ii. Previous work:** Stoichiometric ferricyanide-activated peptide thioacid ligation, which requires temporally separated thiolysis and activation steps[4]. **iii. This work:** Thiol-catalyzed peptide ligation via a transient thioimidate intermediate that is chemoselectively intercepted by an amine nucleophile in water to form a peptidyl amidine directly from a stable nitrile without any activating agents. (**B) i.** Peptidyl amidines persist if α-amidonitriles are coupled with an α-amino acid (**AA**), except **Ser**, **Thr** and **Cys** (see Fig. 2Bii). Intramolecular amide-assisted hydroysis of the peptidyl amidine yields the native peptide bond (e.g. **Asn** or R^2=H or peptide). **ii.** The intramolecular cyclization of serinyl (R^3=H; X=O), threoninyl (R^3=CH_3; X=O), or cysteinyl (R^3=H; X=S) residues promote stereoretentive hydrolysis of peptidyl amidines to native peptides.

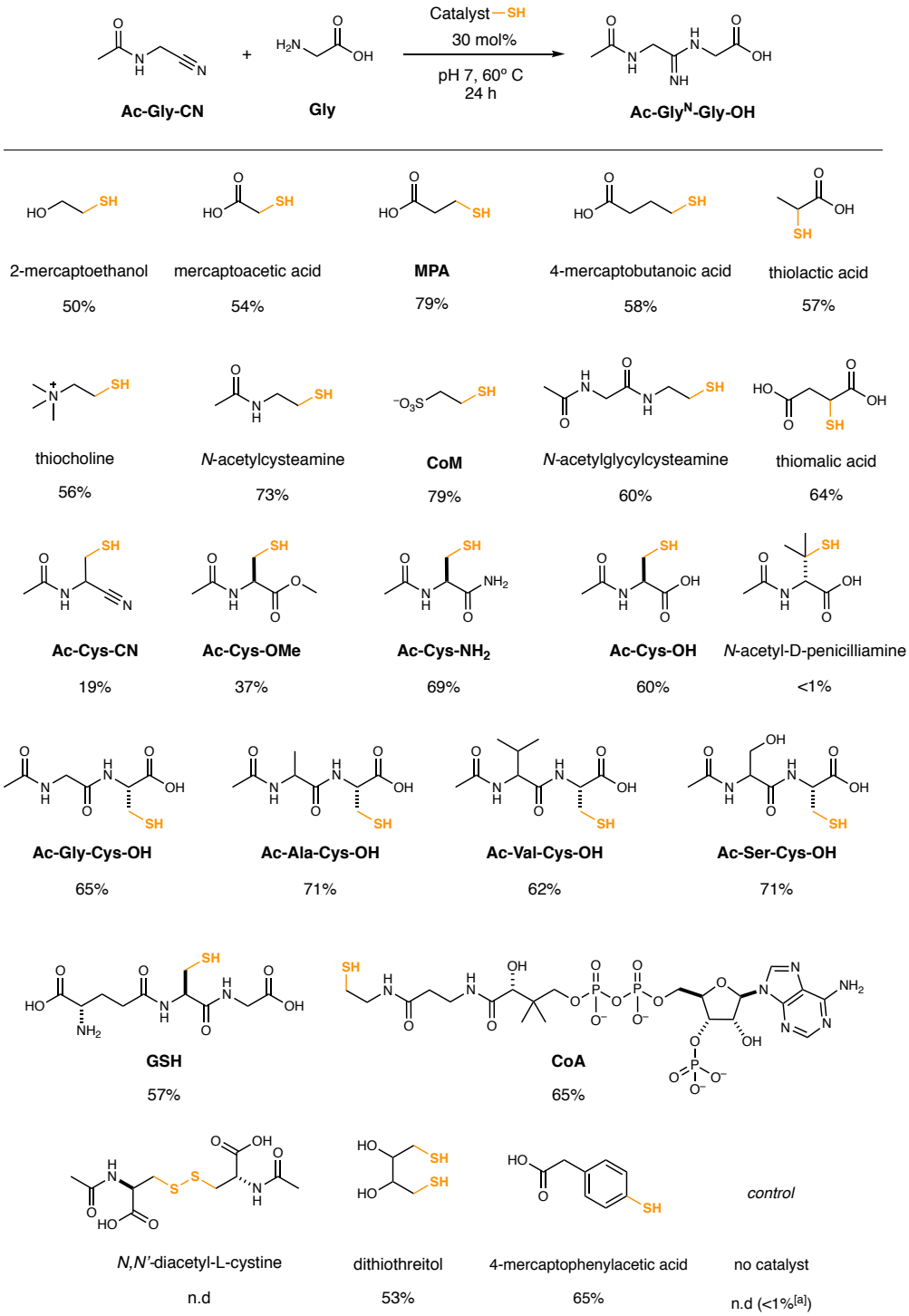

Fig. 3. Peptide ligation catalysts. Yields for the formation of **Ac-Gly^{N}-Gly-OH** by thiol-catalyzed (Catalyst–SH; 30 mol%) coupling of **Ac-Gly-CN** (200 mM) and **Gly** (200 mM) after 24 h. n.d = not detected. [a]Coupling (<1%) only observed after 7 d.

Importantly, no activating agents were required to induce ligation. Moreover, peptide ligation no longer requires a *C*-terminal thioacid[4] or thioester and an *N*-terminal cysteine residue as the nucleophilic ligation partner[36,37,40]. A broad spectrum of cysteine derivatives were excellent catalysts for peptidyl amidine synthesis in water, including **Ac-Cys-NH$_2$** (69%), *N*-acylcysteine peptides (57–71%), as well as **CoA** (65%) and simple thiols such as co-enzyme M (**CoM**; 79%). Catalysis is essential to promoting this ligation, and the catalytic

potency of these simple thiols is highlighted by the lack of background reactivity. In control experiments (with no thiol catalyst) no coupling was observed, and **Ac-Gly^{N}-Gly-OH** (<1%) could only be detected after 7 d (fig. S85). Furthermore, very little cysteine-catalyzed hydrolysis was observed in the absence of an amine nucleophile; incubating **Ac-Gly-CN** with **Ac-Cys-OH** (30 mol%, pH 7, 60 ºC) resulted in only 6% hydrolysis after 24 h (fig. S86). These results underscore the outstanding kinetic stability of α-amidonitriles and the generality of thiol-catalysis, where only a severely hindered tertiary thiol failed to catalyze CPL.

CPL is remarkably specific and selective for proteinogenic α-peptide synthesis (Fig. 4). For example, the reaction of **Ac-Gly-CN** and **Ac-β-Ala-CN** (1:1; 200 mM) with **Gly** (200 mM) and **Ac-Cys-OH** (30 mol%) results in exclusive α-amidonitrile coupling to furnish **Ac-Gly^{N}-Gly-OH** (65%) with no detectable β-amidonitrile coupling of **Ac-β-Ala-CN**. We also observed only α-ligation upon challenging CPL with *N*-acetylglutamine dinitrile **Ac-Glx-CN**. Finally, we only observed coupling of proteinogenic **Ala** in competition with α,α-disubstituted (non-proteinogenic amino acid) α-aminoisobutyric acid (**Aib**). The observed selectivity may have been an essential element in the emergence of proteinogenic α-peptides in extant biology[3,4,23].

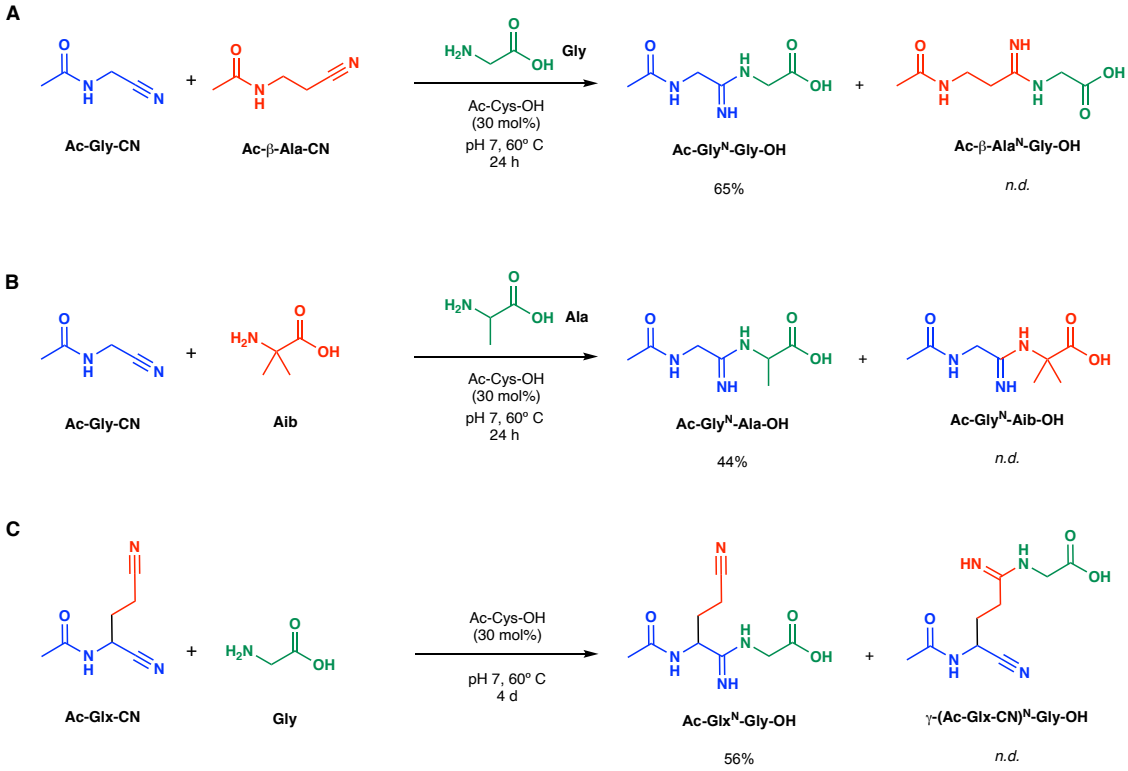

Fig. 4. Chemo- and regioselective organocatalytic proteinogenic α-peptide ligation. CPL selectively ligates native α-peptidyl bonds. Selective ligation of: (**A**) **Ac-Gly-CN** (200 mM) with **Gly** (200 mM) affords **Ac-Gly^{N}-Gly-OH** in a stoichiometric competition with **Ac-β-Ala-CN** (200 mM). Non-proteinogenic **Ac-β-Ala^{N}-Gly-OH** not observed. (**B**) **Ac-Gly-CN** (200 mM) with **Ala** (200 mM) to afford **Ac-Gly^{N}-Ala-OH** in a stoichiometric competition with **Aib** (200 mM). Non-proteinogenic **Ac-Gly^{N}-Aib-OH** not observed. (**C**) **Ac-Glx-CN** (200 mM) with **Gly** (200 mM) to afford **Ac-Gly^{N}-Gly-OH**. Non-proteinogenic **γ-(Ac-Glx-CN)^{N}-Gly-OH** not observed. n.d. = not detected.

Thiol-catalyzed coupling of α-amidonitriles with α-amino acids is highly general; all proteinogenic α-amino acids coupled with **Ac-Gly-CN** to give peptidyl amidines **Ac-Gly^{N}-AA-OH** in good yields at pH 7 and 60 °C (Table 1). However, dipeptides derived from **Ser**, **Thr** and **Asn** underwent pronounced amidine hydrolysis to the corresponding peptides (**Ac-Gly^{N}-AA-OH→Ac-Gly-AA-OH**). In the reactions of **Ser** and **Thr** we observed oxazoline intermediates, suggesting that intramolecular catalysis by the amino acid side chain was responsible for rapid amidine hydrolysis (Fig 2B.). Having also observed that the peptidyl amidine derived from **Asn** hydrolyzed, we envisaged that amino amides (and therefore peptides) would behave similarly. We next examined whether amides intramolecularly catalyze amidine hydrolysis by coupling proteinogenic α-amino amides (**AA-NH$_2$**). All **AA-NH$_2$** resulted in selective dipeptide (**Ac-Gly-AA-NH$_2$**) synthesis, irrespective of their side chain (Table 1). This observed intramolecular hydrolysis is significant because, whilst peptidyl amidines undergo racemization as expected[41], *N*-terminal **Ser**, **Thr**, and **Cys** undergo stereoretentive coupling (Fig 2B.). The origins of biological homochirality remain a formidable challenge, requiring an as-yet unknown symmetry-breaking event, and until that event prebiotic syntheses must produce racemic mixtures[43]. The observed racemization of peptidyl amidines may have limited impact in a racemic (prebiotic) environment, however, the stereoretentive couplings of peptides with nucleophilic side chains and the intramolecular amide-catalyzed hydrolysis of peptidyl amidines offers a route to investigate dynamic kinetic resolution of peptide stereochemistry that may underpin peptide chiral resolution in oligomers, rather than α-amino acid (or α-aminonitrile) monomers[44].

Having demonstrated the tolerance of all proteinogenic aminoacyl residues at the ligation junction in thiol-catalyzed dipeptide synthesis, we explored the feasibility of organocatalytic peptide fragment ligations. The success of this ligation hinged upon the *N*-terminal peptide fragment not undergoing diketopiperazine-induced decomposition. We recently demonstrated the prebiotic synthesis of *N*-acylpeptide nitriles (**Ac-AA$_n$-CN**) in water by sulfide-mediated iterative ligation of α-aminonitriles (Fig. 2A.ii)[4], and therefore we chose to investigate **Ac-Gly$_3$-CN** as a prebiotically plausible peptide nitrile in fragment ligation. Peptide nitrile **Ac-Gly$_3$-CN** (100 mM) readily coupled with various peptides with 3-mercaptopropionic acid (**MPA**) as a catalyst at neutral pH to give excellent yields of peptides, even with stoichiometric (1:1) coupling partners (Fig. 5). The simplicity and generality of CPL, and the lack of any detectable diketopiperazine formation underscores the privileged nature of thiol-catalyzed peptide ligations in water.

Table 1. Organocatalytic ligation of *N*-acetylaminonitriles with α-amino acids and α-amino amides. Yields for **Ac-Cys-OH** (30 mol%) catalyzed formation of peptidyl amidines (**Ac-AA^N-AA^1-X**) and peptides (**Ac-AA-AA^1-X**) by coupling of **Ac-AA-CN** (200 mM) with **AA^1-X** (1 equiv., pH 7, 60 °C, 24 h).

Entry	Ac-AA-CN	AA^1-X	Ac-AA^N-AA^1-X (%)	Ac-AA-AA^1-X (%)	Entry	Ac-AA-CN	AA^1-X	Ac-AA^N-AA^1-X (%)	Ac-AA-AA^1-X (%)
1	Ala	Gly-OH	60 (>95[a])	-	23	Gly	Lys-OH	70[e]	-
2	Ala	Ala-OH	25 (83[a])	-	24	Gly	Lys-NH_2	25[f]	52[g]
3	Gly	Gly-OH	60 (>95[a])	-	25	Gly	DL-Met-OH	72	-
4	Gly	Gly-NH_2	21	52	26	Gly	Met-NH_2	5	62
5	Gly	DL-Ala-OH	43 (79[b])	-	27	Gly	Phe-OH	21 (52[c])	-
6	Gly	D-Ala-NH_2	3	63	28	Gly	Phe-NH_2	8	52
7	Gly	Arg-OH	37 (78[c])	-	29	Gly	Pro-OH	58	-
8	Gly	Arg-NH_2	14	56	30	Gly	Pro-NH_2	-	21 (67[h])
9	Gly	Asn-OH	9	45	31	Gly	Ser-OH	-	61 (74[i])
10	Gly	Asn-NH_2	-	72[d]	32	Gly	Ser-NH_2	-	68[j] (75[i])
11	Gly	Asp-OH	58	-	33	Gly	Thr-OH	-	51 (80[i])
12	Gly	Asp-NH_2	6	58	34	Gly	Thr-NH_2	-	69[k] (85[i])
13	Gly	Gln-OH	56	-	35	Gly	Trp-OH	32	5
14	Gly	Gln-NH_2	-	43	36	Gly	Trp-NH_2	4	45
15	Gly	Glu-OH	58	-	37	Gly	Tyr-OH	20	-
16	Gly	Glu-NH_2	-	64	38	Gly	Tyr-NH_2	3	62
17	Gly	His-OH	73	-	39	Gly	Val-OH	42 (79[l])	6
18	Gly	His-NH_2	-	67	40	Gly	D-Val-NH_2	7	50
19	Gly	Ile-OH	55	-	41	Glx	Gly-OH	33 (56[m])	-
20	Gly	Ile-NH_2	12	47	42	Ser	Gly-OH	61 (90[a])	-
21	Gly	Leu-OH	53	-	43	Ser	Ala-OH	25 (71[a])	-
22	Gly	D-Leu-NH_2	5	54	44	Val	Gly-OH	3 (79[n])	-

[a]**AA^1-X** (2 equiv.) and **Ac-Cys-OH** (2 equiv.). [b]**DL-Ala-OH** (3 equiv.). [c]**AA^1-OH** (2 equiv.). [d]**Ac-Cys-OH** (5 equiv.). [e]Combined yield: ***N*^{2}-(Ac-Gly^N)-Lys-OH** (43%), ***N*^{6}-(Ac-Gly^N)-Lys-OH** (24%), ***N*^{2},*N*^{6}-bis(Ac-Gly^N)-Lys-OH** (<5%). [f]Amidine is N^6 coupling product only (***N*^{6}-(Ac-Gly^N)-Lys-NH_2**). [g]Amide is N^2 coupling product only (***N*^{2}-(Ac-Gly)-Lys-NH_2**). [h]**Pro-NH_2** (2 equiv.) and **Ac-Cys-OH** (5 equiv.). [i]**MPA** catalyst (1 equiv.), 48 h. [j]2-(Acetamidomethyl)-4,5-dihydrooxazole-4-carboxamide (6%) observed. [k]2-(Acetamidomethyl)-5-methyl-4,5-dihydrooxazole-4-carboxamide (<5%) observed. [l]**Val-OH** (2 equiv.), 96 h. [m]4 d. [n]**Gly-OH** (2 equiv.), **MPA** catalyst (2 equiv.), 96 h.

A

Entry	AA_n-OH	Ac-Gly_3-AA_n-OH (%)	HRMS-ESI		
			Formula	Calculated	Found
1	Met-Gly-OH	80	$C_{15}H_{26}N_5O_7S$ $[M+H]^+$	420.1547	420.1552
2	Ala-Ala-Ala-OH	90	$C_{17}H_{28}N_6O_8$ $[M+H]^+$	445.2041	445.2044
3	Ala-Gly-Ala-OH	84	$C_{16}H_{27}N_6O_8$ $[M+H]^+$	431.1885	431.1893
4	Gly-Ala-Gly-OH	87	$C_{15}H_{25}N_6O_8$ $[M+H]^+$	417.1728	417.1739
5	Gly-Gly-Gly-OH	89	$C_{14}H_{23}N_6O_8$ $[M+H]^+$	403.1572	403.1570
6	Gly-Gly-His-OH	89	$C_{18}H_{27}N_8O_8$ $[M+H]^+$	483.1946	483.1968
7	Leu-Leu-Leu-OH	76[i]	$C_{26}H_{47}N_6O_8$ $[M+H]^+$	571.3450	571.3460
8	Met-Ala-Ser-OH	77	$C_{19}H_{33}N_6O_9S$ $[M+H]^+$	521.2024	521.2022
9	Phe-Gly-Gly-OH	77[ii]	$C_{21}H_{27}N_6O_8$ $[M-H]^-$	491.1889	491.1890

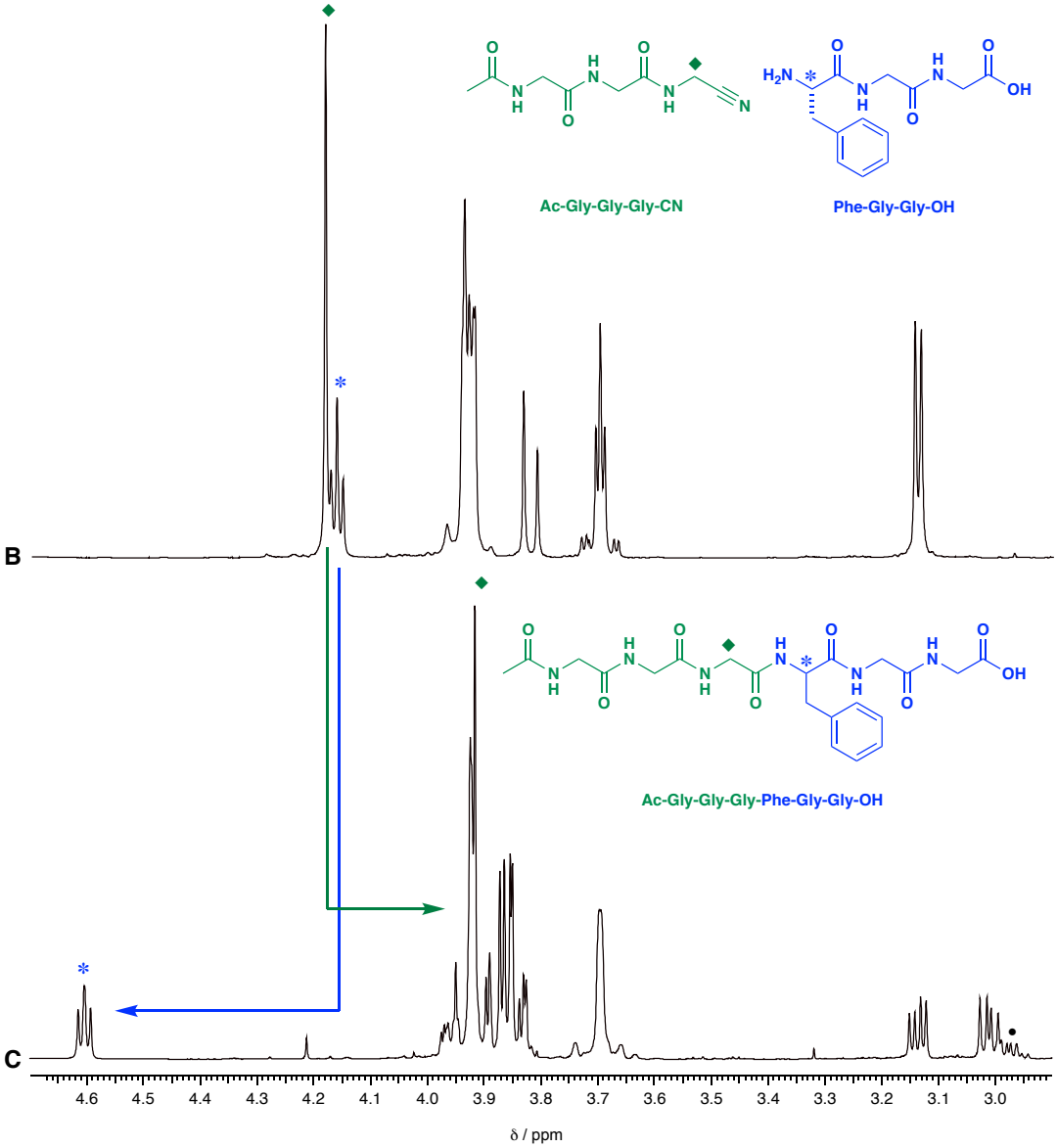

Fig. 5. Organocatalytic peptidyl-nitrile to peptide fragment ligations. A. Yields and mass spectrometry data for 3-mercaptopropionic acid (**MPA**; 160 mM) catalyzed ligation of **Ac-Gly_3-CN** (100 mM) and peptides **AA_n-OH** (100 mM, 24 h, 60 °C, pH 7.0). [i]48 h. [ii]**Ac-Gly_2-Gly^N-Phe-Gly_2-OH** (•, 12%) also observed. Total ligation yield was 89%. ^{1}H NMR (700 MHz, H_2O) spectrum to show: (**B**) **Ac-Gly_3-CN** and **Phe-Gly_2-OH**. (**C**) crude products of **Ac-Gly_3-CN**, **Phe-Gly_2-OH** and **MPA** after 24 h at 60 °C.

Discussion

We have combined a systems chemistry[3,45] approach (where reactants from different stages of a pathway are allowed to interact) with analysis of the strategy of modern biosynthesis to resolve two long-standing conundrums at the origins of life: the chemical origins of cysteine and catalytic peptide ligation. We note that extant cysteine biosyntheses[17,18] bear a striking resemblance to the prebiotic synthesis we have outlined. However, the non-enzymatic reactions described here are contingent on an α-nitrile, rather than complex enzymes and cofactors. The α-amidonitrile moiety is not only prebiotically plausible, but also ideally poised to activate dehydroalanine formation at near-neutral pH, and provides the in-built energy required to drive peptide synthesis. It is important to note that the α-amidonitrile has a unique balance of kinetic stability and thermodynamic reactivity to deliver aqueous ligation without requiring any electrophilic activation, and the highly selective and reversible reaction that thiols undergo with α-amidonitriles makes thiol-catalyzed nitrile ligations a powerful and selective system for peptide ligation in water. The amidine intermediates observed during CPL suggest that ligation occurs by intermolecular addition of the nucleophilic coupling partner to a catalyst-bound thioimidate (Fig. 2). Importantly, our ligation (in contrast to thioester acylations[36,37]) tolerates all amino acid side chain residues without protection to deliver high yields in neutral water, and is not limited to ligation of *N*-terminal cysteine residues via intramolecular amidation.

It is of note that thiol catalysis alleviates geochemical restrictions enforced by the mutual incompatibility of reagents, and reduces the number of intermediate steps to streamline our previous peptide fragment ligation (Fig. 2A.ii)[4] in a manner akin to enzyme catalysis. It is remarkable that a single amino acid residue, cysteine, provides robust catalysis for peptide ligation in water. This is strong evidence for the catalytic role of simple cysteines (and thiols) foreshadowing enzymes at the origin of life; it is now easier to envisage the (reciprocal) evolution of (coded) catalytic peptides that catalyze peptide ligation[39]. The inherent catalytic activity of simple cysteinyl peptides makes them an excellent starting point from which to evolve more complex (folded) enzymes[14,39] and catalysts for protometabolic reactions in an abiotic environment. Finally, our data support a scenario in which nitriles served as an early energy currency on the primordial Earth, perhaps acting as a forerunner to ATP and thioesters that drive reactions in extant biology[10,11].

Data availability All data supporting the findings of this study are available within the main text and supplementary materials.

Acknowledgment We thank EPSRC (EP/K004980/1, EP/M507970/1, EP/P020410/1), Simons Foundation (318881FY19, 493895) and Volkswagen Foundation (94743) for financial support, K. Karu, M. Puchnarewicz

and E. Samuel (Mass Spectrometry), A. E. Aliev (NMR spectroscopy), P. Canavelli, B. Thoma, and J. Fairchild for preliminary work and experimental assistance.

Author Contributions M.W.P. conceived the research. C.S.F., S.I., C.F-G., T.D.S and M.W.P. designed and analyzed the experiments. C.S.F., S.I., C.F-G. and L.M. conducted the experiments. S.I. assembled and wrote the Supplementary Materials with contributions from C.S.F. and C.F-G. C.S.F. and S.I. contributed equally to the experiments and are listed alphabetically. S.I. and M.W.P. wrote the manuscript with assistance from T.D.S.

Competing interests

The authors declare no competing financial interests.

References

1. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
2. Fischbach, M. A. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496 (2006).
3. Islam, S. & Powner, M. W. Prebiotic systems chemistry: complexity overcoming clutter. *Chem* **2**, 470–501 (2017).
4. Canavelli, P., Islam, S. & Powner, M. W. Peptide ligation by chemoselective aminonitrile coupling in water. *Nature* **571**, 54–549 (2019).
5. Chadha, M. S., Replogle, L., Flores, J. & Ponnamperuma, C. Possible role of aminoacetonitrile in chemical evolution. *Bioorg. Chem.* **1**, 269–274 (1971).
6. Shalayel, I., Coulibaly, S., Ly, K. D., Milet, A. & Vallée, Y. The reaction of aminonitriles with aminothiols: A way to thiol-containing peptides and nitrogen heterocycles in the primitive earth ocean. *Life* **8**, doi:10.3390/life8040047 (2018).
7. Shalayel, I. *et al.* Cysteine chemistry in connection with abiogenesis. *Eur. J. Org. Chem.* **2020***,* 3019–3023, doi. 10.1002/ejoc.202000089 (2020).
8. Leman, L., Orgel, L. & Ghadiri, M. R. Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **306**, 283–286 (2004).
9. Griesser, H., Bechthold, M., Tremmel, P., Kervio, E. & Richert, C. Amino acid-specific, ribonucleotide-promoted peptide formation in the absence of enzymes. *Angew. Chem. Int. Ed.* **56**, 1224–1228 (2017).
10. De Duve, C. *Blueprint for a Cell: The Nature and Origin of Life* (Neil Patterson Publishers, 1991).
11. Goldford, J. E., Hartman, H., Smith, T. F. & Segrè, D. Remnants of an ancient metabolism without phosphate. *Cell* **168**, 1126–11134 (2017).

12. Poole, L. B. The basics of thiols and cysteines in redox biology and chemistry. *Free Radic. Biol. Med.* **80**, 148–157 (2015).
13. Eck, R. V. & Dayhoff, M. O. Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. *Science* **152**, 363–366 (1966).
14. Bonfio, C. *et al.* UV-light-driven prebiotic synthesis of iron–sulfur clusters. *Nat. Chem.* **9**, 1229–1234 (2017).
15. Wong, J. T.-F. A co-evolution theory of the genetic code. *Proc. Natl Acad. Sci. USA* **72**, 1909–1912 (1975).
16. Trifonov, E. N. Consensus temporal order of amino acids and evolution of the triplet code. *Gene* **261**, 13–151 (2000).
17. Sauerwald, A. *et al.* RNA-dependent cysteine biosynthesis in Archaea. *Science* **307**, 1969–1972 (2005).
18. Fujishima, K. *et al.* Reconstruction of cysteine biosynthesis using engineered cysteine-free enzymes. *Sci. Rep.* **8**, 1776, (2018).
19. Khare, B. N. & Sagan, C. Synthesis of cystine in simulated primitive conditions. *Nature* **232**, 577–579 (1971).
20. Weber, A. L. & Miller, S. L. Reasons for the occurrence of the twenty coded protein amino acids. *J. Mol. Evol.* **17**, 27–284, (1981).
21. Parker, E. T. *et al.* Primordial synthesis of amines and amino acids in a 1958 Miller H_2S-rich spark discharge experiment. *Proc. Natl Acad. Sci. USA* **108**, 5526–5531 (2006).
22. Powner, M. W., Zheng, S.-L. & Szostak, J. W. Multicomponent assembly of proposed DNA precursors in water. *J. Am. Chem. Soc.* **134**, 13889–13895 (2012).
23. Islam, S., Bučar, D.-K. & Powner, M. W. Prebiotic selection and assembly of proteinogenic amino acids and natural nucleotides from complex mixtures. *Nat. Chem.* **9**, 584–589 (2017).
24. Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D. & Sutherland, J. D. Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat. Chem.* **7**, 301–307 (2015).
25. Chen, Z. *et al.* Exploring the condensation reaction between aromatic nitriles and amino thiols to optimize in situ nanoparticle formation for the imaging of proteases and glycosidases in cells. *Angew. Chem. Int. Ed.* **59**, 327–3279 (2020).
26. Borchert, A. J., Ernst, D. C. & Downs, D. M. Reactive enamines and imines in vivo: Lessons from the RidA paradigm. *Trends Biochem. Sci.* **44**, 849–860 (2019).
27. Ksander, G. *et. al.* Chemie der α-aminonitrile 1. Mitteilung einleitung und wege zu uroporphyrinogen-octanitrilen. *Helv. Chim. Acta* **70**, 1115–1172 (1987).
28. Photaki, I. Transformation of serine to cysteine. β-Elimination reactions in serine derivatives. *J. Am. Chem. Soc.* **85**, 1123–1126 (1963).

29. Byford, M. F. Rapid and selective modification of phosphoserine residues catalysed by Ba2+ ions for their detection during peptide microsequencing. *Biochem. J.* **280**, 261–265 (1991).
30. Bowler, F. R. *et al.* Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nat. Chem.* **5**, 383–389 (2013).
31. Fernandez-Garcia, C. & Powner, M. W. Selective acylation of nucleosides, nucleotides, and glycerol-3-phosphocholine in water. *Synlett* **28**, 78–83 (2017).
32. Anderson, B. M., Cordes, E. H. & Jencks, W. P. Reactivity and catalysis in reactions of the serine hydroxyl group and of O-acyl serines. *J. Biol. Chem.* **236**, 455–463 (1961).
33. Coggins, A. J. & Powner, M. W. Prebiotic synthesis of phosphoenol pyruvate by α-phosphorylation-controlled triose glycolysis. *Nat. Chem.* **9**, 310–317 (2017).
34. Barbas III, C. F. Organocatalysis lost: modern chemistry, ancient chemistry, and an unseen biosynthetic apparatus. *Angew. Chem. Int. Ed.* **47**, 42–47 (2008).
35. van der Helm, M. P., Klemm, B. & Eelkema, R. Organocatalysis in aqueous media. *Nat. Rev. Chem.* **3**, 491–508 (2019).
36. Dawson, P., Muir, T., Clark-Lewis, I. & Kent, S. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
37. Kulkarni, S. S., Sayers, J., Premdjee, B. & Payne, R. J. Rapid and efficient protein synthesis through expansion of the native chemical ligation concept. *Nat. Rev. Chem.* **2**, 0122 (2018).
38. Pattabiraman, V. R. & Bode, J. W. Rethinking amide bond synthesis. *Nature* **480**, 471 (2011).
39. Söding, J. & Lupas, A. N. More than the sum of their parts: On the evolution of proteins from peptides. *BioEssays* **25**, 837–846 (2003)
40. Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. A self-replicating peptide. *Nature* **382**, 525 (1996).
41. Moser, H. *et al.* Poly(dipeptamidinium)-salze: Definition und methoden zur präparativen herstellung. *Helv. Chim. Acta* **69**, 1224–1262 (1986).
42. Lange, U. E. W., Schäfer, B., Baucke, D., Buschmann, E. & Mack, H. A new mild method for the synthesis of amidines. *Tetrahedron Lett.* **40**, 7067–7070 (1999).
43. Hein, J. E. & Blackmond, D. G. On the origin of single chirality of amino acids and sugars in biogenesis. *Acc. Chem. Res.* **45**, 2045–2054 (2012).
44. Beaufils, D., Danger, G., Boiteau, L., Rossi, J.-C. & Pascal, R. Diastereoselectivity in prebiotically relevant 5(4H)-oxazolone-mediated peptide couplings. *Chem. Commun.* **50**, 3100–3102 (2014).
45. Szostak, J. W. Systems chemistry on early Earth. *Nature* **459**, 171–172 (2009).