LETTER

https://doi.org/10.1038/s41586-019-1371-4

Peptide ligation by chemoselective aminonitrile coupling in water

Pierre Canavelli[1], Saidul Islam[1] & Matthew W. Powner[1]*

Amide bond formation is one of the most important reactions in both chemistry and biology[1–4], but there is currently no chemical method of achieving α-peptide ligation in water that tolerates all of the 20 proteinogenic amino acids at the peptide ligation site. The universal genetic code establishes that the biological role of peptides predates life's last universal common ancestor and that peptides played an essential part in the origins of life[5–9]. The essential role of sulfur in the citric acid cycle, non-ribosomal peptide synthesis and polyketide biosynthesis point towards thioester-dependent peptide ligations preceding RNA-dependent protein synthesis during the evolution of life[5,9–13]. However, a robust mechanism for aminoacyl thioester formation has not been demonstrated[13]. Here we report a chemoselective, high-yielding α-aminonitrile ligation that exploits only prebiotically plausible molecules—hydrogen sulfide, thioacetate[12,14] and ferricyanide[12,14–17] or cyanoacetylene[8,14]—to yield α-peptides in water. The ligation is extremely selective for α-aminonitrile coupling and tolerates all of the 20 proteinogenic amino acid residues. Two essential features enable peptide ligation in water: the reactivity and pK_{aH} of α-aminonitriles makes them compatible with ligation at neutral pH and *N*-acylation stabilizes the peptide product and activates the peptide precursor to (biomimetic) N-to-C peptide ligation. Our model unites prebiotic aminonitrile synthesis and biological α-peptides, suggesting that short *N*-acyl peptide nitriles were plausible substrates during early evolution.

To improve the efficiency and selectivity of peptide ligation in water we sought to develop a mechanism for non-enzymatic peptide synthesis, which would operate via biomimetic N → C ligation in water at near-neutral pH, and we suspected that a combination of sulfur and nitrile chemistry would be required[8,9,14,18–21] (Fig. 1a). Proteinogenic α-aminonitriles (AA-CN) are readily synthesized[8,18], and their direct ligation would provide the simplest prebiotic pathway to peptides. Unfortunately, incubation of AA-CN in water results in extremely ineffective peptide synthesis[22]. α-Amino acids (AA) are widely assumed to be prebiotic precursors of peptides, but the harsh conditions (typically strongly acidic or alkaline solutions) required for AA formation from AA-CN are incompatible with the integrity of both peptides and electrophilic activating agents. Therefore, we sought a more congruent and direct pathway from AA-CN to α-peptides. Although the conversion of AA-CN to α-aminothioacids (AA-SH) has never been reported[23], harnessing the AA-CN nitrile moiety for thioacid synthesis seemed more prudent than dissipating its activation through exhaustive hydrolysis.

Maurel and Orgel have previously suggested that AA-SH[16] might offer an interesting alternative to biological thioesters[10,11]. AA-SH combine excellent aqueous stability with highly selective (electrophilic or oxidative) activation[12,14,16,24], but their prebiotic synthesis presents difficulties[25] and they undergo inefficient ligation at near-neutral pH (Supplementary Discussion)[16,26]. To overcome these problems we reconsidered the synthesis of thioacids from nitriles (Fig. 1b). Recently we reported that high-yielding nucleophilic displacement of sulfides by Gly-CN[19] is promoted by the low pK_{aH} of AA-CN in water, and we hypothesized that coupling AA-CN to the C terminus of a growing peptide would be facile at neutral pH. Importantly, we suspected that this ligation would (electronically) activate the nitrile moiety to thiolysis. Accordingly, AA-CN *N*-acylation, which appears to be essential to prevent diketopiperazine (DKP)-induced peptide degradation[27,28] (Fig. 1c), would initiate peptide synthesis by promoting thioacid synthesis.

Ferricyanide-mediated acetylation of AA-CN (50 mM) by AcSH (3 equiv.)[12,14] gave α-amidonitriles (Ac-AA-CN) in near-quantitative yield in water (Table 1). As anticipated, acylation of AA-CN activated the nitrile moiety, and quantitative conversion of Ac-AA-CN to Ac-AA-SNH_2 was observed upon incubation with H_2S (10 equiv., pH 9, room temperature, 1–4 d) (Supplementary Figs. 39–52, 64, 80). Incubation of Ac-Gly-CN (50 mM) and Gly-CN (50 mM) or acetonitrile (50 mM) with H_2S (0.25 M, pH 9, room temperature, 24 h) gave smooth conversion to Ac-Gly-SNH_2 (91%), whereas only 7% of Gly-CN was converted to Gly-SNH_2 (Supplementary Fig. 18) and acetonitrile thiolysis was not observed (Supplementary Fig. 20). This demonstrates the pronounced nitrile activation provided by acylation. Electrophilic activation is also specific to α-amidonitriles; for example, the reaction of Ac-Gly-CN and Ac-β-Ala-CN (1:1) with H_2S results in almost exclusive Ac-Gly-CN thiolysis (Fig. 1e).

Notably, we observed hydrolysis of Ac-AA-SNH_2 to Ac-AA-SH to realize effective synthesis of thioacids (Fig. 1b). This is in stark contrast to the reactivity of AA-SNH_2, for which hydrolysis to the respective AA-SH was not observed (Supplementary Discussion and Supplementary Fig. 16). Hydrolysis of Ac-AA-SNH_2 generally furnished the respective Ac-AA-SH in good yields (51%–85%; Table 1 and Supplementary Figs. 53–58, 64, 80). However, the sterically bulkier Val residue hydrolysed sluggishly to give the corresponding α-amidothioacid Ac-Val-SH in poor yield (8%; entry 9 in Table 1 and Supplementary Fig. 59). This amino acid residue is one of several notoriously problematic C-terminal ligation residues observed during the (semi)synthesis of peptides in the related process of thioester-mediated native chemical ligation[4,29,30]. Future investigation of catalytic α-amidothioacid Ac-AA-SH synthesis is warranted; however, we note that (uncatalysed) Ac-Val-CN thiolysis already delivers an Ac-Val-SH yield seven times greater than that of AA-SH analogues synthesized by electrophilic AA activation[25]. Furthermore, Ac-AA-SH are highly stable to the conditions of their formation, whereas AA-SH are destroyed by the activating agents required for their synthesis[25].

We next investigated the ligation of Ac-AA-SH. We observed that incubation of Ac-Gly-SH (50 mM) with Gly-CN (2 equiv.) and ferricyanide (3 equiv.) gave Ac-Gly_2-CN in near-quantitative yield over a broad pH range (pH 5–9, room temperature). A range of activating agents—including ferricyanide[12,14–17], cupric salts[8], cyanoacetylene[8,14] and *N*-cyanoimidazole[14]—were all found to be effective (Extended Data Table 1), showing that multiple methods of Ac-AA-SH activation towards AA-CN ligation in water are possible.

We then carried out an iterative one-pot AA-CN coupling without isolating the intermediate ligation products. The α-amidonitrile Ac-Gly-CN was successively homologated to afford the corresponding peptides Ac-Gly_n-CN (n = 2–5; n = 2, 71%; n = 3, 71%; n = 4, 63%; n = 5, 41%; Table 2). After four iterations of the homologation

[1]Department of Chemistry, University College London, London, UK. *e-mail: matthew.powner@ucl.ac.uk

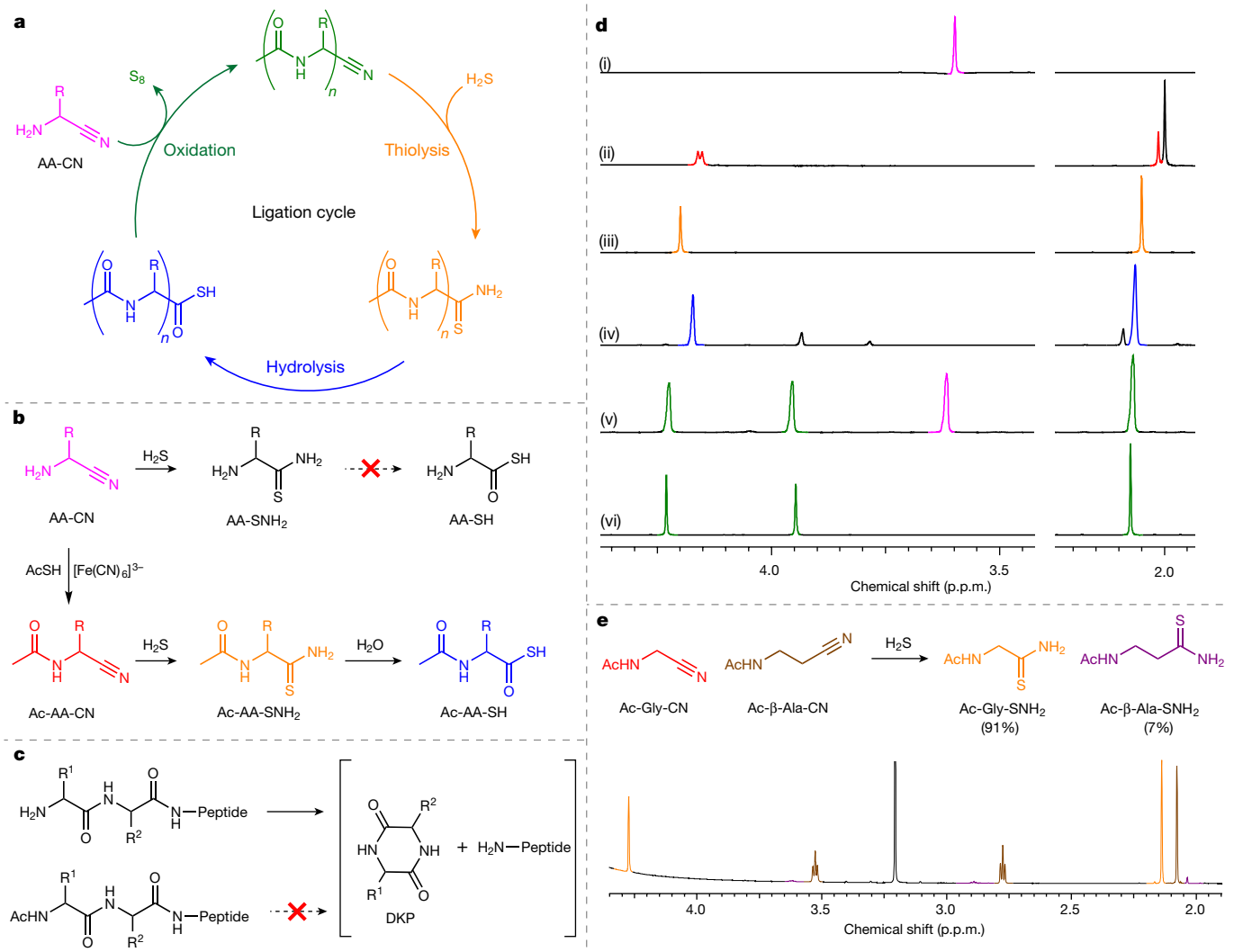


Fig. 1 | Sulfide-mediated α-aminonitrile ligation a, Iterative AA-CN ligation to give *N*-acetyl peptide nitriles (Ac-AA$_n$-CN; green) by sequential thiolysis, hydrolysis and AA-CN ligation. **b**, The thiolysis of AA-CN (magenta) to yield AA-SH (black) is not observed, whereas the thiolysis of *N*-acetyl aminonitrile (Ac-AA-CN; red) to α-amidoacyl thioacid (Ac-AA-SH; blue) is facile. **c**, Iterative truncation of peptides by DKP excision is blocked by *N*-acylation. **d**, ^{1}H nuclear magnetic resonance (NMR) spectra (600 MHz, H_2O:D_2O = 98:2, 25 °C), showing: (i) Gly-CN (magenta); (ii) Ac-Gly-CN (quantitative yield, quant.; red) synthesized by the reaction of Gly-CN (50 mM) with thioacetic acid (3 equiv.) and $K_3[Fe(CN)_6]$ (9 equiv.) in water at pH 9 at room temperature after 10 min; (iii) Ac-Gly-SNH$_2$ (quant.; orange) synthesized by the reaction of Ac-Gly-CN (50 mM) with H_2S (10 equiv.) in water at pH 9 at room temperature after 1 d; (iv) Ac-Gly-SH (81%; blue) synthesized by hydrolysis of Ac-Gly-SNH$_2$ (50 mM) at pH 9 and 60 °C after 1 d; (v) Ac-Gly$_2$-CN (quant.; green) synthesized by the reaction of Ac-Gly-SH (50 mM) with Gly-CN (2 equiv.; magenta) and $K_3[Fe(CN)_6]$ (3 equiv.) in water at pH 9 and room temperature after 20 min; (vi) pure Ac-Gly$_2$-CN. **e**, ^{1}H NMR spectrum (600 MHz, H_2O:D_2O = 98:2, 25 °C) showing the reaction of homologous amidonitriles Ac-Gly-CN (red) and Ac-β-Ala-CN (brown) with H_2S (10 equiv., pH 9, room temperature, 1 d), which strongly favours thiolysis of the proteinogenic glycyl residue to yield Ac-Gly-SNH$_2$ (orange) (Supplementary Fig. 19).

cycle, partial precipitation of Ac-Gly$_4$-CN reduced the overall coupling yield for Ac-Gly$_5$-CN synthesis (13% overall yield of Ac-Gly$_5$-CN from Ac-Gly-CN; Supplementary Figs. 209–211). This demonstrates that iterative ligation of AA-CN can be achieved in good yield in water without purification, within the limits of peptide solubility. Our ligation is highly robust and tolerates monomer-by-monomer peptide growth and fragment ligations to produce oligomers in high yield, even at low concentrations (3.1 mM; entry 17 in Table 2) and with stoichiometric (1:1) coupling partners (entries 5–17 in Table 2). To our knowledge, these are the first examples of fragment ligations with prebiotic substrates in water.

Activation of the C terminus of peptides and amino acids (such as Ac-Ala-OH) with electrophilic reagents (such as 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide, EDC) can result in racemization[31]. However, we observed the formation of chiral α-amidothioacid L-Ac-Ala-SH (from L-Ac-Ala-CN), and its subsequent ligation with Gly-CN proceeds with retention of stereochemistry (Supplementary Figs. 292–294). This demonstrates that enantiomeric enrichment is preserved during our peptide ligation, which is a testament to the mild ligation conditions.

We next investigated the chemoselectivity and robustness of AA-CN ligation. Stoichiometric (1:1) competition reactions between Gly-CN (50 mM) and ammonia, glycine (Gly), glycine amide (Gly-NH$_2$), β-alanine (β-Ala), β-alanine nitrile (β-Ala-CN), phosphate, propylamine, cytosine, cytidine-5′-phosphate and adenosine-5′-phosphate across a broad pH range (pH 5–9; Extended Data Table 2) were investigated. All competition reactions demonstrated outstanding selectivity for Gly-CN ligation (>80% yield) at neutral pH (Supplementary Figs. 235–244). We observed selective Gly-CN ligation in the presence of Gly-NH$_2$ (pK_{aH} = 8.4) and β-Ala-CN (pK_{aH} = 8.0) in neutral and acidic solution, but selectivity was lost at pH values above their pK_{aH}. The excellent selectivity for AA-CN ligation in neutral solution was attributed to

Table 1 | α-Amidothioacid synthesis and α-aminonitrile ligation

Entry	AA	Yield (%)			
		Ac-AA-CN[a]	Ac-AA-SNH$_2$[b]	Ac-AA-SH[c]	Ac-AA-Gly-CN[d]
1	Gly	99	99	81	99
2	Ala	99	99	85	93
3	Arg	99	99	51	64[e]
4	Leu	99	99	77	93
5	Met	99	99	70	80
6	Phe	99	99	84	78
7	Pro	99	99	72	82
8	Ser	99	99	61	87[f]
9	Val	99	99	8	92

^{1}H NMR yields are determined with an internal NMR standard. See Extended Data Table 5 for further examples of AA-CN ligations.
[a]Acetylation of AA-CN (50 mM) with AcSH (150 mM) and $K_3[Fe(CN)_6]$ (450 mM) in water (pH 9; room temperature; <20 min).
[b]Thiolysis of Ac-AA-CN (50 mM) to Ac-AA-SNH$_2$ in water by H_2S (10 equiv., pH 9, room temperature) (Supplementary Figs. 39–52, 64, 80).
[c]Hydrolysis of Ac-AA-SNH$_2$ (50 mM) to Ac-AA-SH in water with H_2S (500 mM; pH 9, 60 °C) (Supplementary Figs. 53–59, 64, 80).
[d]Ligation of Ac-AA-SH (50 mM) to Gly-CN (100 mM) in water with $K_3[Fe(CN)_6]$ (150 mM; pH 9, room temperature), unless stated otherwise.
[e]Yield for the coupling of Ac-Arg-SH (46 mM) with Gly-CN (91 mM) and $K_3[Fe(CN)_6]$ (136 mM).
[f]Yield for the coupling of Ac-Ser-SH (30 mM) with Gly-CN (61 mM) and $K_3[Fe(CN)_6]$ (92 mM).

their uniquely suppressed pK_{aH} values (for example, pK_{aH} = 5.3 for Gly-CN)[19], which renders them predominantly neutral, and consequently nucleophilic, even in weakly acidic solutions. AA-CN ligation is also observed across a broad temperature range (T = 3–60 °C), as well as at physiologically relevant concentrations (0.5 mM) (Extended Data Table 3).

Developing a universal strategy to activate and ligate peptides that accommodates all proteinogenic amino acids is problematic. Lysine and cysteine, for example, contain highly nucleophilic moieties that are incompatible with electrophilic activation[2,17,32], and aspartate and glutamate have β- and γ-carboxylate residues, respectively, in addition to the α-carboxylate that must be selectively activated and ligated[2,4,30]. α-Amidothioacid ligation is highly general and chemoselective. All investigated amino acids and their derivatives were coupled in good-to-excellent yields (Tables 1–3, Extended Data Tables 4, 5). Sterically congested and β-branched thioacid ligations were also highly effective; ligations yielding Ac-Phe-Phe-CN, Ac-Phe-Val-CN and Ac-Val-Val-CN were all rapid and high-yielding (entries 18–20 in Extended Data Table 5). We observed unprecedented protecting-group-free ligation for all 20 proteinogenic side-chain residues—including His, Asp, Lys, Cys, Ser, Thr and Tyr, which are all essential to enzyme catalysis but notoriously difficult to ligate under previously reported (prebiotic) conditions[2,4,30,32]. Although Cys is incompatible with electrophilic activating agents[2,32], it underwent highly selective ligation under our conditions to furnish Ac-Gly-Cys-OH (80%; entry 6 in Table 3) after thiol exchange (Extended Data Fig. 1a).

Following the excellent selectivity of AA-CN ligation, we challenged the α-NH$_2$ selectivity with lysine, which possesses two amine nucleophiles. We observed poor selectivity for α-coupling of Lys (1.2:1 α/ε) and Lys-NH$_2$ (2.7:1 α/ε), but Lys-CN ligated with exceptional α-selectivity (>80:1 α/ε; Extended Data Fig. 1b; Supplementary Fig. 149). We then turned our attention to the coupling of AA-CN to a C-terminal lysine residue, which requires intermolecular AA-CN coupling to outcompete cyclization (Extended Data Fig. 1c). We first demonstrated that activation of Ac-α-Lys-SH (30 mM) by ferricyanide (90 mM) at pH 9.0 led to rapid lactamization (92%). This was not surprising, given the close proximity of the ε-NH$_2$ and thioacid moieties of Ac-α-Lys-SH. However, we found that Gly-CN (64 mM) successfully coupled with Ac-α-Lys-SH (32 mM). The intermolecular coupling of Gly-CN outcompeted lactamization across a broad pH range to give Ac-α-Lys-Gly-CN (88%–95%, pH 6.5–9.0; Supplementary Figs. 70–71). The chemoselective coupling of lysine residues at the C and N termini of peptides underscores that AA-CN ligation is predisposed to yield

Table 2 | Synthesis of oligomeric *N*-acetyl peptides and peptide nitriles by oxidative fragment ligation

Entry	$(AA^1)_n$	$(AA^2)_m$-X	Ac-$(AA^1)_n$–$(AA^2)_m$-X (%)
1	Gly	Gly-CN	71[a]
2	Gly_2	Gly-CN	71[b]
3	Gly_3	Gly-CN	63[c]
4	Gly_4	Gly-CN	41[d]
5	Gly_3	Ala_3-OH	65
6	Gly_3	Arg-Gly-Asp-OH	76
7	Gly_3	Gly_3-OH	90
8	Gly_3	Gly_3-CN	>95
9	Gly_3	Gly_2-His-OH	90[e]
10	Gly_3	Leu_3-OH	70
11	Gly_3	Met-Ala-Ser-OH	75
12	Gly_3	Phe-Gly_2-OH	74
13	Gly_5	Ala_3-OH	74
14	Gly_5	Gly_2-His-OH	80
15	Gly_6	Gly_3-CN	43[f]
16	Gly_5	Gly_5-OH	79 (66[g])
17	Gly_6	Gly_5-OH	>95[h] (92[g])

Ferricyanide-mediated oxidative coupling of Ac-$(AA^1)_n$-SH with $(AA^2)_m$-X (X = CN or CO_2H) to produce oligopeptides Ac-$(AA^1)_n$–$(AA^2)_m$-X. Yields of the oxidative coupling of thioacid Ac-$(AA^1)_n$-SH (25 mM) with peptide $(AA^2)_m$-X (25 mM, pD 9.5) and $K_3[Fe(CN)_6]$ (75 mM) in D_2O at room temperature, unless stated otherwise. See Supplementary Table 15 for further details.
[a–d]Ac-Gly_n-CN synthesis by iterative ligation after two, three, four and five cycles of thiolysis, hydrolysis and ligation (Supplementary Figs. 209–211).
[e]Coupling of Ac-Gly_3-SH (30 mM) with Gly_2-His-OH (25 mM, pD 9.5) with $K_3[Fe(CN)_6]$ (75 mM).
[f]Yield of Ac-Gly_9-CN is given after four sequential steps from Ac-Gly_3-SH.
[g]Yield determined by product isolation.
[h]Coupling of Ac-Gly_6-SH (3.13 mM) and Gly_5-OH (6.25 mM).

Table 3 | Chemoselective synthesis of *N*-acetyl dipeptides

Entry	AA	Ac-Gly-AA-OH (%)
1	Gly	94
2	Ala	83
3	Arg	88
4	Asn	81
5	Asp	89
6	Cys	80[a]
7	Gln	90
8	Glu	92
9	His	95
10	Ile	84
11	Leu	86
12	Lys	94[b]
13	Met	95
14	Phe	90
15	Pro	89
16	Ser	85
17	Thr	81
18	Trp	71
19	Tyr	23[c]
20	Val	84

Yields are given for the products of oxidative coupling of Ac-Gly-SH (50 mM) with AA (150 mM) with $K_3[Fe(CN)_6]$ (150 mM) in water at room temperature and pH 9.5. ^{1}H NMR yields determined with an internal NMR standard.
[a]Yield observed using $K_3[Fe(CN)_6]$ (300 mM), followed by methanethiol (600 mM, pH 10.8) reduction (see Extended Data Fig. 1a and Supplementary Figs. 112–114).
[b]The observed ratio of mono- and di-acylated products varies with solution pH (see Extended Data Fig. 1b for α-selectivity of Lys ligation at pH 7.5 and Supplementary Table 11).
[c]L-Tyrosine (Tyr) exhibits extremely low solubility in water (6.5 mM, pH 9.5, room temperature; see Supplementary Table 8).

α-peptides. To the best of our knowledge, these reactions constitute the first non-enzymatic, chemoselective and protecting-group-free intermolecular lysine ligations for native peptide bond formation at near-neutral pH[26,33].

In a clear departure from the convention that AA are essential for prebiotic peptide synthesis, we have found that their precursors, AA-CN, are predisposed to undergo selective ligation at biochemically relevant pH and concentration. *N*-Acylation initiates our peptide synthesis strategy and activates a ligated aminonitrile to thiolysis and hydrolysis to its respective α-amidothioacid. *N*-Acylation circumvents the irreversible derivatization of peptides by electrophiles (such as COS[17]; see Supplementary Discussion) and promotes (biomimetic) N → C peptide ligation. Our peptide ligation strategy requires separate sequential delivery of H_2S and an activating agent. For example, H_2S and ferricyanide are mutually reactive feedstock molecules and would need to be delivered from separate source locations. However, repeated sequential delivery of H_2S and then AA-CN and an oxidant (for example, ferricyanide), chalcophilic metal ion (for example, Cu^{2+}) or an electrophile (for example, cyanoacetylene) would yield controlled stepwise peptide ligation. Controlled synthesis, which responds to environmental or internal stimuli, is an essential element of metabolic regulation, and we speculate that coupling iterative aminonitrile ligation to metabolic (redox) cycles may lead to positive cooperative feedback during the early evolution of life.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1371-4.

Received: 16 October 2018; Accepted: 28 May 2019;
Published online: 10 July 2019

1. Constable, D. J. C. et al. Key green chemistry research areas—a perspective from pharmaceutical manufacturers. *Green Chem.* **9**, 411–420 (2007).
2. Isidro-Llobet, A., Álvarez, M. & Albericio, F. Amino acid-protecting groups. *Chem. Rev.* **109**, 2455–2504 (2009).
3. Pattabiraman, V. R. & Bode, J. W. Rethinking amide bond synthesis. *Nature* **480**, 471–479 (2011).
4. 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).
5. Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. A self-replicating peptide. *Nature* **382**, 525–528 (1996).
6. Weber, A. L. & Pizzarello, S. The peptide-catalyzed stereospecific synthesis of tetroses: a possible model for prebiotic molecular evolution. *Proc. Natl Acad. Sci. USA* **103**, 12713–12717 (2006).
7. Adamala, K. & Szostak, J. W. Competition between model protocells driven by an encapsulated catalyst. *Nat. Chem.* **5**, 495–501 (2013); corrigendum **5**, 634 (2013).
8. 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).
9. Semenov, S. N. et al. Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. *Nature* **537**, 656–660 (2016).
10. Lipmann, F. Attempts to map a process evolution of peptide biosynthesis. *Science* **173**, 875–884 (1971).
11. De Duve, C. *Blueprint for a Cell: The Nature and Origin of Life* (Neil Patterson Publishers, 1991).
12. Liu, R. & Orgel, L. E. Oxidative acylation using thioacids. *Nature* **389**, 52–54 (1997).
13. Weber, A. L. Prebiotic amino acid thioester synthesis: thiol-dependent amino acid synthesis from formose substrates (formaldehyde and glycolaldehyde) and ammonia. *Orig. Life Evol. Biosph.* **28**, 259–270 (1998).
14. Bowler, F. R. et al. Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nat. Chem.* **5**, 383–389 (2013).
15. Keefe, A. D. & Miller, S. L. Was ferrocyanide a prebiotic reagent? *Orig. Life Evol. Biosph.* **26**, 111–129 (1996).
16. Maurel, M.-C. & Orgel, L. E. Oligomerization of α-thioglutamic acid. *Orig. Life Evol. Biosph.* **30**, 423–430 (2000).
17. Leman, L., Orgel, L. & Ghadiri, M. R. Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **306**, 283–286 (2004).
18. 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).
19. Stairs, S. et al. Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides. *Nat. Commun.* **8**, 15270 (2017).
20. Islam, S. & Powner, M. W. Prebiotic systems chemistry: complexity overcoming clutter. *Chem* **2**, 470–501 (2017).
21. Roberts, S. J. et al. Selective prebiotic conversion of pyrimidine and purine anhydronucleosides into Watson–Crick base-pairing arabino-furanosyl nucleosides in water. *Nat. Commun.* **9**, 4073 (2018).
22. Chadha, M. S., Replogle, L., Flores, J. & Ponnamperuma, C. Possible role of aminoacetonitrile in chemical evolution. *Bioorg. Chem.* **1**, 269–274 (1971).
23. Paventi, M. & Edward, J. T. Preparation of α-aminothioamides from aldehydes. *Can. J. Chem.* **65**, 282–289 (1987).
24. Sheehan, J. C. & Johnson, D. A. The synthesis and reactions of *N*-acyl thiol amino acids. *J. Am. Chem. Soc.* **74**, 4726–4727 (1952).
25. Leman, L. J. & Ghadiri, M. R. Potentially prebiotic synthesis of α-amino thioacids in water. *Synlett* **28**, 68–72 (2017).
26. Okamoto, R. et al. Regioselective α-peptide bond formation through oxidation of amino thioacids. *Biochemistry* **58**, 1672–1678 (2019).
27. Steinberg, S. M. & Bada, J. L. Peptide decomposition in the neutral pH region via the formation of diketopiperazines. *J. Org. Chem.* **48**, 2295–2298 (1983).
28. Radzicka, A. & Wolfenden, R. Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *J. Am. Chem. Soc.* **118**, 6105–6109 (1996).
29. Dawson, P., Muir, T., Clark-Lewis, I. & Kent, S. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
30. Villain, M., Gaertner, H. & Botti, P. Native chemical ligation with aspartic and glutamic acids as C-terminal residues: scope and limitations. *Eur. J. Org. Chem.* **2003**, 3267–3272 (2003).
31. Danger, G. et al. 5(4*H*)-Oxazolones as intermediates in the carbodiimide- and cyanamide-promoted peptide activations in aqueous solution. *Angew. Chem. Int. Ed.* **52**, 611–614 (2013).
32. 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).
33. Zhang, L. & Tam, J. P. Lactone and lactam library synthesis by silver ion-assisted orthogonal cyclization of unprotected peptides. *J. Am. Chem. Soc.* **121**, 3311–3320 (1999).

METHODS

General and safety information. Reagents and solvents were obtained and used without further purification, unless specified. Sodium hydrosulfide hydrate ($NaSH{\cdot}xH_2O$; 50% purity) and sodium sulfide (Na_2S; >97%) were used without purification. Deionized water was obtained from an Elga Option 3 purification system. NMR spectra were recorded on Bruker NMR spectrometers (Avance Neo 700, Avance III 600, Avance III 400 and Avance 300), equipped with a Bruker room-temperature 5-mm multinuclear gradient probe (700 MHz), a 5-mm DCH cryoprobe (600 MHz) and a gradient probe (400 and 300 MHz). Where noted, a solvent suppression pulse sequence with presaturation and spoil gradients was used to obtain ^{1}H NMR spectra (noesygppr1d, Bruker) and ^{1}H–^{13}C heteronuclear multiple bond correlation (HMBC) NMR spectra (hmbcgplpndprqf, Bruker). Coupling constants are reported in hertz. Spectra were recorded at 298 K. Infrared spectra (IR) were recorded on a Shimadzu IR Tracer 100 Fourier transform (FT)-IR spectrometer as a solid or liquid. Absorption maxima are reported as wavenumbers (in cm^{-1}). Mass spectra and accurate mass measurements were recorded on a Waters LCT Premier quadrupole time-of-flight (QTOF) mass spectrometer connected to a Waters Autosampler Manager 2777C, Thermo Finnigan MAT900, and an Agilent liquid chromatography (LC) system connected to an Agilent 6510 QTOF mass spectrometer. High-performance liquid chromatography (HPLC) analysis was carried out using an Agilent Infinity 1260 LC system. Solution pH values were measured using a Mettler Toledo Seven Compact pH meter with a Mettler Toledo InLab semi-micro pH probe. The pH readings for H_2O and H_2O/D_2O (9:1) solutions are reported uncorrected. Warning: hydrogen cyanide (HCN) and hydrogen sulfide (H_2S) are highly toxic poisons by inhalation and ingestion. They generate poisonous gas at neutral or acidic pH (HCN, $pK_a = 9.2$; H_2S, $pK_a = 7.1$). Solutions containing cyanide, (hydro)sulfide or compounds that may generate these should be handled in a well ventilated fume hood equipped with appropriate chemical quenches, such as sodium hypochlorite (bleach) or iron(II) sulfate solution.

General procedures. *Acetylation of α-aminonitriles with thioacetate.* α-Aminonitrile hydrochloride (AA-CN·HCl; 50 mM) and potassium thioacetate (AcSK; 150 mM) were dissolved in H_2O (2 ml) and the solution was adjusted to pH 9.0 with NaOH. Potassium hexacyanoferrate(III) ($K_3[Fe(CN)_6]$; 450 mM) was added, and the solution was stirred at room temperature for 20 min. The solution was adjusted to pH 9.0 and centrifuged, and NMR spectra of the supernatant were acquired. Yields are reported in Table 1 and characterization data in Supplementary Information.

Thiolysis of N-acetylaminonitriles. N-Acetylaminonitrile (Ac-AA-CN; 50 mM) and $NaSH{\cdot}xH_2O$ (10 equiv.) were dissolved in degassed H_2O/D_2O (98:2, 50 ml). The solution was adjusted to pH 9.0 and stirred at room temperature for 24 h. NMR spectra were periodically acquired, until complete conversion of Ac-AA-CN to Ac-AA-SNH_2 was observed. The solution was sparged with argon for 15 min at pH 5.0 and concentrated in vacuo. The residue was purified using flash column chromatography to afford Ac-AA-SNH_2. Yields are reported in Table 1 and characterization data in Supplementary Information.

Hydrolysis of N-acetylaminothioamide to N-acetylaminoacyl thioacids. Ac-AA-SNH_2 (50 mM), $NaSH{\cdot}xH_2O$ (10 equiv.) and methylsulfonylmethane (MSM; 50 mM) were dissolved in degassed H_2O/D_2O (98:2, 1 ml), and the solution was adjusted to pH 9.0 with NaOH/HCl. The solution was incubated at 60 °C while being maintained at pH 9.0 with NaOH/HCl, and NMR spectra were periodically acquired until complete consumption of Ac-AA-SNH_2 was observed. The Ac-AA-SH was confirmed by ^{1}H–^{13}C HMBC NMR analysis, spiking or comparison of NMR data with pure synthetic standards. The reaction mixture was quantified using MSM as an internal standard. Yields are reported in Table 1 and characterization data are given in Supplementary Information.

Oxidative coupling of Ac-Gly-SH with α-amino acids or α-amino amides. AA or α-amino amide (AA-NH_2) (150 mM) was dissolved in degassed H_2O/D_2O (98:2; 1 ml) and the solution was adjusted to pH 9.5 with HCl/NaOH. Ac-Gly-SH (50 mM) was added and the total volume was adjusted to 2 ml with H_2O/D_2O (98:2). Potassium hexacyanoferrate(III) ($K_3[Fe(CN)_6]$, 150 mM) was added and the solution was stirred at room temperature for 20 min while maintaining the solution at pH 9.5 with NaOH. The resulting suspension was centrifuged and the supernatant was analysed by one- and two-dimensional NMR spectroscopy (^{1}H–^{1}H correlated spectroscopy (COSY), ^{1}H–^{13}C heteronuclear single quantum coherence spectroscopy (HSQC) and ^{1}H–^{13}C HMBC in H_2O/D_2O (98:2). The yield was quantified using MSM as an internal standard. The ligation product (Ac-Gly–AA-X; X = OH or NH_2) was confirmed by ^{1}H–^{13}C HMBC NMR spectral analysis and high-resolution mass spectrometry (HRMS). Reaction mixtures were lyophilized and dissolved in DMSO-d_6 or CD_3OD for further NMR spectral analysis if ^{1}H–^{13}C HMBC cross-correlation peaks were obscured by the HOD resonance during the original NMR analysis in H_2O/D_2O (98:2). Yields and HRMS data are given in Table 3, Supplementary Table 8 (Ac-Gly–AA-OH), Extended Data Table 4 and Supplementary Table 9 (Ac-Gly–AA-NH_2), and characterization data are provided in Supplementary Information.

Oxidative coupling of α-aminoacetyl thioacids with α-aminonitriles. AA^2-CN (100 mM) was dissolved in degassed H_2O/D_2O (98:2; 2 ml) and the solution was adjusted to pH 9.0 with HCl/NaOH. Ac-AA^1-SH (50 mM) was added and the total volume was adjusted to 2 ml with H_2O/D_2O (98:2). Potassium hexacyanoferrate(III) ($K_3[Fe(CN)_6]$; 150 mM) was added and the solution was stirred at room temperature for 20 min. The pH was readjusted to pH 9.0 using NaOH. The resulting suspension was centrifuged and the supernatant was analysed by one- and two-dimensional NMR spectroscopy (^{1}H–^{1}H COSY, ^{1}H–^{13}C HSQC and ^{1}H–^{13}C HMBC) in H_2O/D_2O (98:2). The reaction mixture was quantified using MSM as an internal standard. The ligation product Ac-AA^1–AA^2-CN was confirmed by ^{1}H–^{13}C HMBC NMR spectral analysis and HRMS. Reaction mixtures were diluted with DMSO-d_6 (1:49:50; D_2O/H_2O/DMSO-d_6) or lyophilized and dissolved in DMSO-d_6 or CD_3OD for further NMR spectral analysis if ^{1}H–^{13}C HMBC cross-correlation peaks were obscured by the HOD resonance during the original NMR analysis in H_2O/D_2O (98:2). Yields and HRMS data are given in Table 1, Extended Data Table 5 and Supplementary Table 7, and characterization data are reported in Supplementary Information.

Preparative oxidative coupling of α-aminoacetyl thioacids with α-aminonitriles. AA^2-CN (100 mM) was dissolved in degassed H_2O (5 ml) and the solution pH was adjusted to pH 9.0 with NaOH. Ac-AA^1-SH (50 mmol) was added and the total volume was adjusted to 10 ml with H_2O. Potassium hexacyanoferrate(III) ($K_3[Fe(CN)_6]$; 150 mM) was added and the solution was stirred at room temperature for 20 min. The solution was then extracted with ethyl acetate (3 × 25 ml). The combined organic layers were washed with HCl (0.1 M, 25 ml), $NaHCO_3$ (saturated; 25 ml) and brine (saturated; 25 ml), dried over $MgSO_4$, filtered and concentrated in vacuo. The crude residue was purified by flash column chromatography to give the ligation product (Ac-AA^1–AA^2-CN) as a white solid. Isolated yields and HRMS data are given in Extended Data Table 5 and Supplementary Table 7, and characterization data are provided in Supplementary Information.

Oxidative peptide fragment ligations. Ac-$(AA^1)_n$-SH (3.1–30.0 mM) and $(AA^2)_m$-X (X = CO_2H or CN; 1–2 equiv.) were dissolved in degassed D_2O and the solution was adjusted to pD 9.5 with NaOH. Potassium hexacyanoferrate(III) ($K_3[Fe(CN)_6]$; 3 equiv.) was added and the solution was stirred at room temperature for 20 min while being maintained at pD 9.5 with NaOH. The resulting suspension was centrifuged and the supernatant was analysed by one- and two-dimensional NMR spectroscopy (^{1}H–^{1}H COSY, ^{1}H–^{13}C HSQC and ^{1}H–^{13}C HMBC). The ligation product (Ac-$(AA^1)_n$–$(AA^2)_m$-X; X = CO_2H or CN) was quantified using relative integral analysis by ^{1}H, ^{1}H–^{13}C HMBC NMR spectral analysis and HRMS. Yields and HRMS data are given in Table 2 and Supplementary Table 15, and characterization data are reported in Supplementary Information.

Data availability

All data supporting the findings of this study are available within the main text, Extended Data Tables 1–5, Extended Data Fig. 1 and the Supplementary Information (which contains Supplementary Discussion, Supplementary Figs. 1–296, Supplementary Tables 1–16, experimental details and compound characterization data).

Acknowledgements We thank the Engineering and Physical Sciences Research Council (EP/K004980/1, EP/P020410/1), the Simons Foundation (318881, 493895) and the Volkswagen Foundation (94743) for financial support. The authors thank K. Karu (UCL Mass Spectrometry Facility), E. Samuel (Mass Spectrometry, UCL School of Pharmacy) and A. E. Aliev (NMR spectroscopy) for assistance.

Author contributions M.W.P. conceived the research. P.C., S.I. and M.W.P. designed and analysed the experiments. P.C. and S.I. contributed equally to the experiments. S.I. wrote the Supplementary Information. M.W.P and S.I. wrote the paper and Supplementary Discussion.

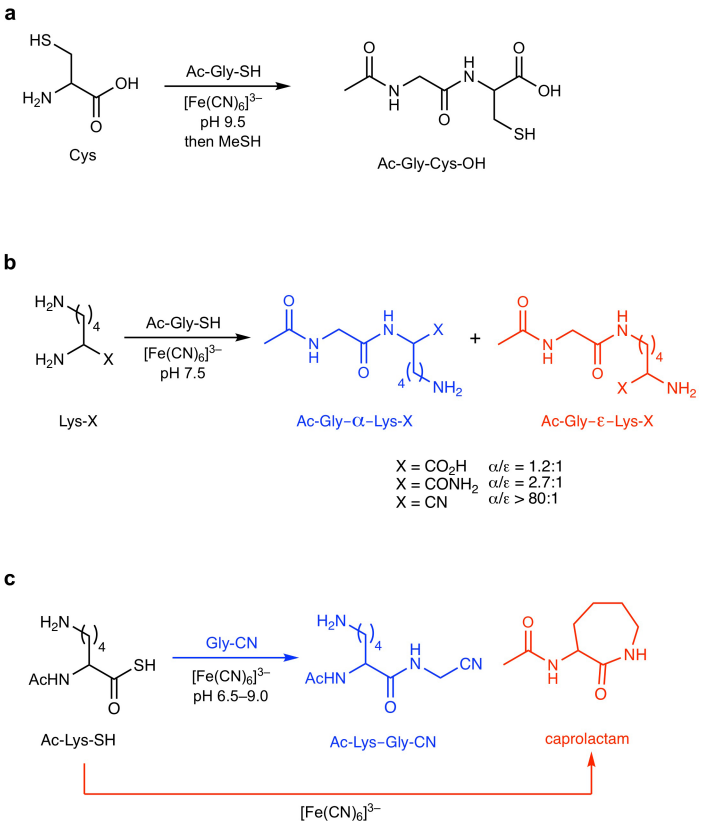
Competing interests The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1371-4.

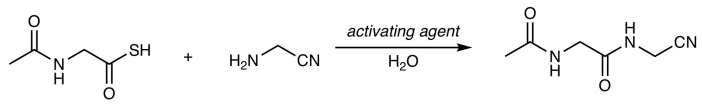
Correspondence and requests for materials should be addressed to M.W.P.

Peer review information *Nature* thanks Irene Chen and the other anonymous reviewer(s) for their contribution to the peer review of this work.

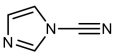


Extended Data Fig. 1 | Chemoselective native peptide bond ligations of cysteine and lysine residues. a, Ligation of Cys is notoriously challenging owing to its highly nucleophilic thiol side chain, which necessitates S-protection to prevent it outcompeting C- and/or N-terminal activation through degradation of the electrophilic activating agents. Protecting-group-free ligation of Cys (150 mM) is achieved through reaction with Ac-Gly-SH (50 mM) and $K_3[Fe(CN)_6]$ (300 mM) in water (pH 9.5, room temperature), followed by thiol reduction (MeSH, 600 mM, pH 10.8, room temperature) to give Ac-Gly-Cys-OH in high yield (80%, over two steps) (Supplementary Figs. 112–114). **b**, Lys-X coupling partners (X = CN, $CONH_2$ or CO_2H) pose greater chemoselectivity challenges because they possess two amino groups (α-NH_2 and ε-NH_2). However, pK_a-controlled native peptide ligation of Lys-CN demonstrates the pivotal role that the unusually low α-amine pK_{aH} of AA-CN[19] can play in selective ligation. Ligation of Lys-CN (100 mM) with Ac-Gly-SH (50 mM) proceeds with unprecedented selectivity in neutral water (pH 7.5, room temperature). Little or no selectivity was observed for the corresponding α-amino amide (Lys-NH_2; 150 mM) and AA (Lys; 150 mM) (Supplementary Figs. 145–151). **c**, Selective intermolecular ligation of the C-terminal lysine residue with AA-CN coupling partner Gly-CN at near-neutral pH (pH 6.5–9.0, blue; see Supplementary Fig. 70). In the absence of Gly-CN, highly efficient intramolecular caprolactam formation is observed (red).

Extended Data Table 1 | α-Amidothioacid activating agents

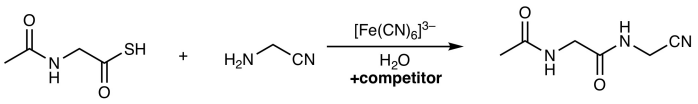


Ac-Gly-SH Gly-CN Ac-Gly–Gly-CN

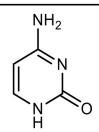
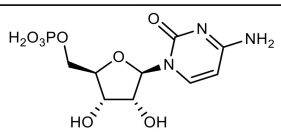
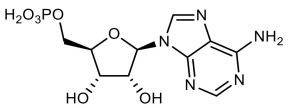
Activating agent	pH	Ac-Gly-Gly-CN (%)
	5.0	85
	7.0	74
	9.0	57
	5.0	95
	7.0	70
	9.0	61
$CuCl_2$	5.0	95
	7.0	94
	9.0	86
$K_3[Fe(CN)_6]$	5.0	91
	7.0	97
	9.0	99

Yields of the oxidative coupling of Ac-Gly-SH (50 mM) and Gly-CN (100 mM) with the specified activating agent (150 mM) after 20 min in water at room temperature. ^{1}H NMR yields were determined with an internal NMR standard.

Extended Data Table 2 | α-Aminonitrile ligation in the presence of nucleophilic competitors

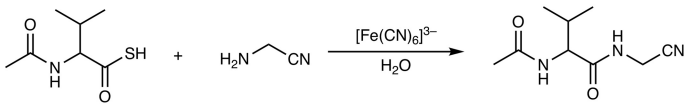


Ac-Gly-SH Gly-CN Ac-Gly–Gly-CN

Competitor	pH	Ac-Gly-Gly-CN (%)	By-product (%)
$Gly\text{-}NH_2$	5.0	66	27
	7.0	59	39
	9.0	14	86
Gly	5.0	82	9
	7.0	81	17
	9.0	79	19
NH_3	5.0	75	3
	7.0	95	3
	9.0	77	22
β-Ala	5.0	93	5
	7.0	89	8
	9.0	90	9
$CH_3CH_2CH_2NH_2$	5.0	90	n.d
	7.0	98	n.d
	9.0	91	5
H_3PO_4	5.0	77	n.d
	7.0	85	<1
	9.0	69	19
β-Ala-CN	5.0	52	21
	7.0	59	29
	9.0	51	41
	5.0	64	n.d
	7.0	89	n.d
	9.0	92	n.d
	5.0	73	n.d
	7.0	83	n.d
	9.0	90	n.d
	5.0	72	n.d
	7.0	91	n.d
	9.0	84	n.d

Yields of the oxidative coupling of Ac-Gly-SH (50 mM) and Gly-CN (100 mM) with $K_3[Fe(CN)_6]$ (150 mM) in the presence of the specified stoichiometric competitor (100 mM) after 20 min in water at room temperature. ^1H NMR yields were determined with an internal NMR standard. See Supplementary Figs. 235–244 for further details. n.d, not detected.

Extended Data Table 3 | α-Aminonitrile ligation at various concentrations and temperatures

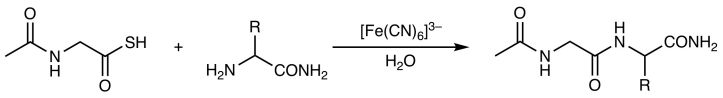


Ac-Val-SH Gly-CN Ac-Val–Gly-CN

Entry	[Ac-Val-SH] (mM)	[Gly-CN] (mM)	[$K_3[Fe(CN)_6]$] (mM)	Temp (° C)	Ac-Val-Gly-CN (%)									
					Time (min)	2	90	180	285	510	750	990	1260	1920	2700
1	0.5	1	1.5	23		–	0	2	4	17	25	31	38	41	45
2	1	2	3	23		–	4	16	30	50	57	62	62	–	–
3	2.5	5	7.5	23		–	29	57	71	80	80	80	81	–	–
4	5	10	15	23		–	75	85	85	86	87	87	87	–	–
5	10	20	30	23		–	83	84	86	86	87	87	87	–	–
6	10	20	30	3		–	50	–	75	–	–	–	78	–	–
7	10	20	30	60		85	–	–	–	–	–	–	–	–	–

Yields of the oxidative coupling of Ac-Val-SH (1 equiv.) and Gly-CN (2 equiv.) with $K_3[Fe(CN)_6]$ (3 equiv.) at specified concentration and temperature. ^1H NMR yields determined with an internal NMR standard. (–) = not determined.

Extended Data Table 4 | Chemoselective synthesis of *N*-acetyl dipeptidyl amides

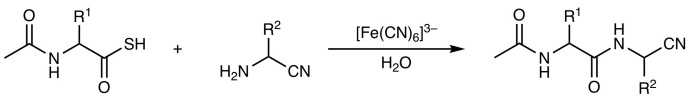


Ac-Gly-SH AA-NH_2 Ac-Gly–AA-NH_2

Entry	AA-NH_2	Ac-Gly-AA-NH_2 (%)
1	Gly	93
2	Ala	93
3	Arg	74
4	Asn	87
5	Asp	80
6	Gln	65
7	Glu	90
8	His	87
9	Ile	74
10	Leu	72
11	Lys	94[a]
12	Met	74
13	Phe	63
14	Pro	67
15	Ser	78
16	Thr	71
17	Trp	71
18	Tyr	56[b]
19	Val	72

Yields of the oxidative coupling of Ac-Gly-SH (50 mM) and AA-NH_2 (150 mM) with $K_3[Fe(CN)_6]$ (150 mM) in water at room temperature and pH 9.5. ^{1}H NMR yields were determined with an internal NMR standard.
[a]The observed ratio of mono- and di-acylated products varies with solution pH (see Extended Data Fig. 1b for α-selectivity of Lys-NH_2 ligation at pH 7.5 and Supplementary Table 12).
[b]Reaction carried out at pH 6.5 (see Supplementary Table 9).

Extended Data Table 5 | Chemoselective synthesis of *N*-acetyl dipeptidyl nitriles



Ac-AA^1-SH AA^2-CN Ac-AA^1–AA^2-CN

Entry	Ac-AA^1-SH	AA^2-CN	Ac-AA^1-AA^2-CN (%)
1	Ala	Ala	85
2	Gly	Ala	95
3	Gly	Arg	70
4	Gly	Asp	91
5	Gly	Glu	74
6	Gly	Ile	87
7	Gly	Leu	89
8	Gly	Lys	93[a]
9	Gly	Met	93
10	Gly	Phe	88
11	Gly	Pro	85
12	Gly	Ser	92
13	Gly	Thr	83
14	Gly	Val	94
15	Ile	Gly	83
16	Lys	Gly	88[b]
17	Phe	Ala	71[c]
18	Phe	Phe	90[c]
19	Phe	Val	73[c]
20	Val	Val	91[c]

Yields of the oxidative coupling of Ac-AA^1-SH (50 mM) and AA^2-CN (100 mM) with $K_3[Fe(CN)_6]$ (150 mM) in water at room temperature and pH 9.0. ^{1}H NMR yields were determined with an internal NMR standard, unless stated otherwise.

[a]The observed ratio of mono- and di-acylated products varies with solution pH (see Extended Data Fig. 1b for α-selectivity of Lys-CN ligation at pH 7.5 and Supplementary Table 13).

[b]Yield for the coupling of Ac-Lys-SH (32 mM) with Gly-CN (64 mM) and $K_3[Fe(CN)_6]$ (96 mM).

[c]Isolated yield.