

Peptide ligation by chemoselective aminonitrile coupling in water

Pierre Canavelli¹, Saidul Islam¹ & Matthew W. Powner^{1*}

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

Supplementary Information for

Peptide ligation by chemoselective aminonitrile coupling in water

Pierre Canavelli, Saidul Islam and Matthew W. Powner*

*Correspondence to: matthew.powner@ucl.ac.uk

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

CONTENTS

GENERAL AND SAFETY	3
SUPPLEMENTARY DISCUSSION.....	4
CARBONYL SULFIDE-MEDIATED N-TERMINAL HYDANTOIN FORMATION OF PEPTIDES	6
THIOLYSIS OF AMINONITRILES.....	15
THIOLYSIS OF GLYCINE NITRILE	15
ATTEMPTED HYDROLYSIS OF GLYCINE THIOAMIDE TO GLYCINE THIOACID.....	16
THIOLYSIS OF N-ACETYLGLYCINE NITRILE.....	17
CHEMOSELECTIVE THIOLYSIS OF N-ACETYLGLYCINE NITRILE IN THE PRESENCE OF GLYCINE NITRILE	18
CHEMOSELECTIVE THIOLYSIS OF N-ACETYLGLYCINE NITRILE IN THE PRESENCE OF N-ACETYL- β -ALANINE NITRILE	19
CHEMOSELECTIVE THIOLYSIS OF N-ACETYLGLYCINE NITRILE IN THE PRESENCE OF ACETONITRILE.....	20
ACETYLATION OF AMINONITRILES	21
N-ACETYLAMINOACYL THIOAMIDE SYNTHESIS.....	31
N-ACETYLAMINOACYL THIOACID SYNTHESIS	39
LIGATION OF N-ACETYLAMINOACYL THIOACIDS AND GLYCINE NITRILE WITH VARIOUS PREBIOTIC ACTIVATING AGENTS.....	47
OPTIMISATION OF FERRICYANIDE-MEDIATED LIGATION OF N-ACETYLAMINOACYL THIOACIDS	48
PREBIOTIC PEPTIDE LIGATIONS OF α -AMINOACETYL THIOACIDS	50
GENERAL PREBIOTIC LIGATION PROCEDURES	50
SUPPLEMENTARY TABLES OF LIGATION YIELDS AND HIGH RESOLUTION MASS SPECTROMETRY DATA.....	52
CHARACTERISATION OF LIGATION REACTIONS	55
ITERATIVE ELONGATION OF N-ACETYLGLYCINE NITRILE.....	162
SYNTHESIS OF AUTHENTIC AC-(GLY) ₅ -CN STANDARD	165
PREBIOTIC PEPTIDE FRAGMENT LIGATIONS.....	167
COUPLING IN THE PRESENCE OF PREBIOTIC COMPETITORS	187
PREPARATIVE SYNTHESSES OF α -ACETAMIDOTHIOACIDS	195
STEREORETENTIVE PREBIOTIC α -AMIDOTHIOACID SYNTHESIS AND LIGATION	222
SUPPLEMENTARY REFERENCES	227
SUPPLEMENTARY TABLE OF FIGURES	230

General and Safety

Reagents and solvents were obtained and used without further purification, unless specified, from the following commercial sources: *Alfa Aesar*, *Acros Organics*, *Apollo Scientific*, *BDH*, *Sigma Aldrich*, *Fluorochem*, *MerckMillipore*, *Fisher Scientific*, *VWR International*, *Carbosynth*, *Manchester Organics*, *Lancaster*, *Molekula*, *Honeywell*, *TCI* and *Santa Cruz Biotechnology*. Sodium hydrosulfide hydrate ($\text{NaSH} \cdot x\text{H}_2\text{O}$ (50%), CAS: 207683-19-0, *Acros Organics*) and sodium sulfide (Na_2S , CAS: 1313-82-2, *Sigma Aldrich*) were used without purification. Dowex® 50W×8 ion-exchange resin (200-400 mesh, Cl^- form) and Dowex® 50W×8 ion-exchange resin (200-400 mesh, H^+ form) were purchased from *Acros Organics*. Deionized water was obtained from an *Elga Option 3* purification system. ^1H , ^{13}C and ^{19}F 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), 5 mm DCH cryoprobe (600 MHz) and a gradient probe (400 and 300 MHz). All chemical shifts (δ) are reported in parts per million (ppm) relative to residual solvent peaks, and ^1H and ^{13}C chemical shifts relative to TMS were calibrated using the residual solvent peak (residual solvent peaks: (δ_{H}) D_2O – 4.79; $\text{DMSO}-d_6$ – 2.50; CDCl_3 – 7.26; CD_3OD – 3.31). Nuclear assignments were made using 2D NMR homo- and heteronuclear correlation spectroscopy (^1H – ^1H COSY; ^1H – ^{13}C HSQC; ^1H – ^{13}C HMBC). Where noted, solvent suppression pulse sequence with presaturation and spoil gradients was used to obtain ^1H NMR spectra (noesygppr1d, *Bruker*) and ^1H – ^{13}C HMBC NMR spectra (hmbcgp1pndprqf, *Bruker*). Coupling constants are reported in Hertz (Hz). Spin multiplicities are indicated by symbols: s (singlet); d (doublet); t (triplet); q (quartet); qn (quintet); spt (septet); oct (octet); m (multiplet); obs. (obscured/coincidental signals), or a combination of these. Diastereotopic geminal (AB) spin systems coupled to one or two additional nuclei are reported as ABX and ABXY, respectively. NMR data are reported as follows: chemical shift (multiplicity, coupling constants (J), number of protons, nuclear assignment). Spectra were recorded at 298 K. Melting points were determined using an *Electrothermal* standard digital apparatus for all solids and are quoted to the nearest $^\circ\text{C}$ and are uncorrected. Infrared spectra (IR) were recorded on a *Shimadzu IR Tracer 100* FT-IR spectrometer as a solid or neat oil/liquid. Absorption maxima are reported in wavenumber (cm^{-1}). Mass spectra and accurate mass measurements were recorded on a *Waters LCT Premier QTOF* connected to a *Waters Autosampler Manager 2777C*, *Thermo Finnigan MAT900*, and an *Agilent LC* connected to an *Agilent 6510 QTOF* mass spectrometer at the Department of Chemistry, University College London. 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, or a *Corning 430* pH meter with a *Fisherbrand* FB68801 semi-micro pH probe. The readings for D_2O solutions are reported as pD, and corrected according to *Covington et al.*¹ The readings for H_2O and $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1) solutions are reported uncorrected. **HCN and H_2S WARNING:** Cyanide and (hydro)sulfide are highly toxic poisons by inhalation, contact, and ingestion. They generate poisonous hydrogen cyanide ($\text{p}K_{\text{a}} = 9.2$) and hydrogen sulfide ($\text{p}K_{\text{a}} = 7.1$) gas at neutral or acidic pH. Solutions containing cyanide or (hydro)sulfide, or compounds which may generate these must be handled in a well-ventilated fumehood equipped with appropriate chemical quenchers, such as sodium hypochlorite (bleach) or iron(II) sulfate solution. Read and follow the material safety data sheet (MSDS) instructions for personnel handling, exposure, and disposal information. Also consult local safety personnel for regulations concerning proper disposal.

Supplementary Discussion

Peptide bond formation is essential in Life, and the antiquity of peptides suggests that they played a role in its early evolution. Therefore, the first peptides exploited by biology were likely to have formed non-enzymatically by an effective prebiotic mechanism that efficiently and chemoselectively ligated peptides in near-neutral water. Electrophilic amino acid **AA** activation (e.g. **NCA**-chemistry^{2,3}) has dominated prebiotic peptide chemistry, but there is an enduring need for activation to be separated from the (oligo)peptide products because the reaction of electrophilic activating agents (such as carbonyl sulfide (**COS**)²) with the free amino-terminus of peptides terminate peptide synthesis by rapid and quantitative urea and hydantoin formation (see below '*Carbonyl sulfide-mediated N-terminal hydantoin formation of peptides*'). To accommodate all proteinogenic residues, and to overcome the narrow pH-range over which α -amino acids **AA** and **NCA**-chemistry operate (pH 8.7 – 9.8), we specifically sought to move away from uncontrolled polymerisation reactions promoted by electrophilic activation of amino acids, which have failed to efficiently produce oligomers in water, lack chemoselectivity, and only operate within narrow (high) pH ranges^{2,4}.

We considered the mechanistic strategies of biological peptide synthesis to inform our analysis. There are two primary mechanisms for peptide synthesis in cells: ribosomal peptide synthesis and non-ribosomal peptide synthesis (NRPS). In contrast to chemical polymerisation strategies (such as **NCA**-chemistry), which ligate peptides from the C→N termini, biology universally synthesises peptides from the N→C termini. We were also strongly motivated to focus on prebiotic stepwise peptide ligation (rather than polymerisation reactions) by the growing number of reports that short peptides are functionally active protometabolic catalysts that can influence a wide range of important (prebiotic) processes, such as chiroselective sugar synthesis⁵ and protocell growth⁶. Furthermore, fragment ligation of peptides has previously been proposed as a strategy for peptide replication at the origins of Life⁷, but fragment ligations are incompatible with monomer activation strategies, such as **NCA**-chemistry, as they require C-terminal activation that is compatible with N-terminal NH₂. If the first biocatalysts were short peptides^{5,6}, later encoding and elongation to longer peptides would provide a simple evolutionary trajectory towards extant life. However, this transition requires there to be efficient chemistry to synthesise peptides from prebiotically plausible substrates, ideally by (biomimetic) N→C peptide ligation in water.

Thioesters have been proposed as potentially prebiotic acylating agents because they play a key role in extant biological processes, such as non-ribosomal peptide synthesis, the citric acid cycle and polyketide biosynthesis^{8,9}, but their robust prebiotic synthesis has not been demonstrated¹⁰. α -Aminothioacids (**AA-SH**) have been considered as simpler alternatives to aminoacyl thioesters¹⁰, but their prebiotic synthesis is equally, if not more, difficult. For example, α -amino acid **Phe**, COS and H₂S yields only very small amounts of the α -aminothioacid **Phe-SH** (1.4%)¹¹. Problematically, **AA-SH** are also highly unstable in the presence of COS¹¹ (and its by-product, CO₂^{3,11}), which implies that it is *unlikely that such reactions would have generated significant amounts of thioacids on the prebiotic Earth*¹¹. These problems are further compounded by their poor ligation. For example, the oxidative ligation of glutamate thioacid (**Glu-SH**) only works in the presence of CO₂ (by *in situ* **NCA** formation) and produces complex mixtures that contain only small amounts of oligomer, alongside competing cyclic products; even (Glu)₁₀ primer-promoted ligation with a 500-fold excess of **Glu-SH** yields ineffective primer extension and returns a considerable quantity of unreacted primer³. Oxidative **AA-SH** ligation, in the absence of long peptide primers (to promote peptide synthesis) and CO₂, occurs only in strongly acidic solution, and no peptide ligation is observed above pH 2¹².

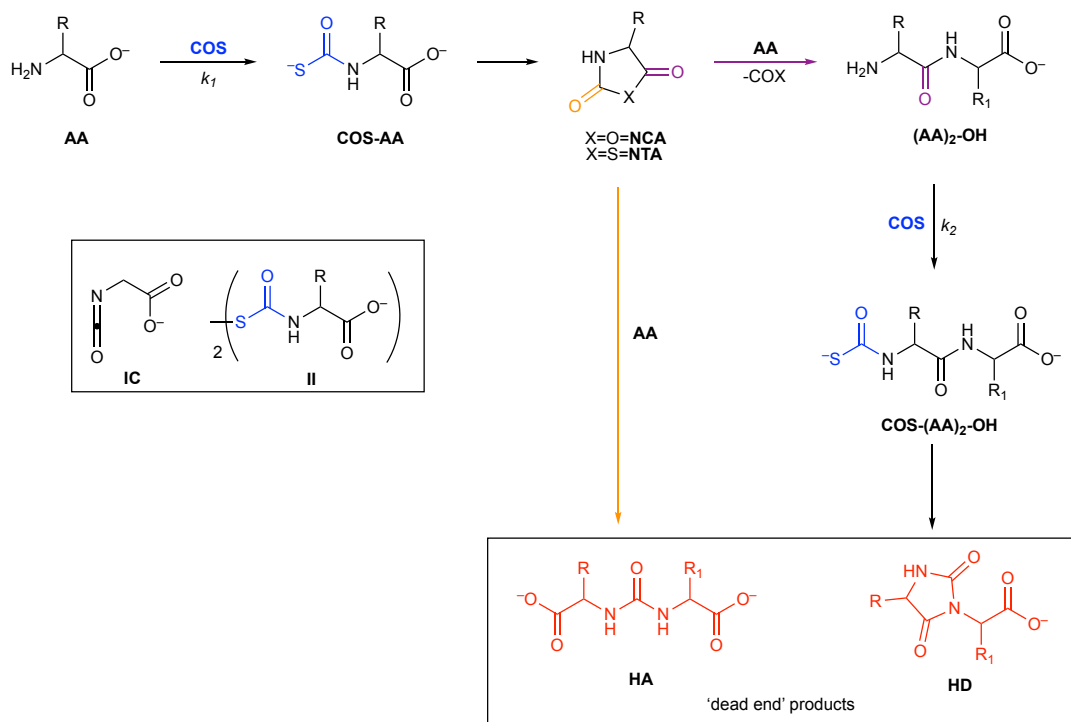
Aminonitriles as substrates for peptide thioacid synthesis in water

Conventional prebiotic peptide syntheses (starting from α -aminonitrile **AA-CN**) invoke the sequence hydrolysis→activation→ligation. However, the difficulties imposed by the harsh conditions required for **AA-CN** to **AA** hydrolysis, poor chemoselectivity of α -amino acid **AA** activation, and poor nucleophilicity of α -amino acids **AA** in neutral water all suggested that reordering the peptide synthesis sequence to ligation→activation→hydrolysis would be beneficial. Specifically, α -aminonitrile **AA-CN** ligation would permit nitrile-activated peptide ligation at neutral pH (by exploiting its low pK_{aH} and corresponding nucleophilicity at neutral pH) to an appropriately activated C-terminal peptide. Considering the spatiotemporal separation of sulfides and cyanides to be prebiotically reasonable, and recognising the emerging importance of sulfur in prebiotic chemistry¹³⁻¹⁹, we suspected that a combination of sulfur and nitrile chemistry would be required to achieve efficient non-enzymatic peptide ligation in water. The prebiotic separation and sequential reactions of sulfides has been invoked in the assembly of numerous biomolecular building blocks, including α -aminonitriles^{14,16}. In addition, we recently reported that sequential thiolysis and oxidation yields Watson-Crick base pairing nucleosides during prebiotic nucleic acid synthesis¹⁹. Therefore we set out to determine if sequential thiolysis and activation yielded chemoselective thioacid synthesis and aminonitrile coupling in water.

Orgel had previously “presumed” the reaction of organic nitriles with H_2S would readily afford thioamides, *en route* to thioacids, in water²⁰. We have not observed the addition of H_2S to unactivated nitriles, such as acetonitrile (**MeCN**), in water (even at extremely high pH and temperatures; Supplementary Figure 20). However, although incubation of **Gly-CN** (50 mM) with H_2S (10 equiv.) in water (pH 9, room temperature, 7 d) slowly furnished glycine thioamide (**Gly-SNH₂**) in quantitative yield (Supplementary Figure 14 and Supplementary Figure 15), all attempts to hydrolyse **Gly-SNH₂** to **Gly-SH** gave the deactivated by-products glycynamide (**Gly-NH₂**) and diketopiperazine (**DKP**), and their respective hydrolysis products (Supplementary Figure 16). Concerned by the propensity of unprotected *N*-terminal peptides to fragment by **DKP** cleavage (Figure 1c of manuscript)^{21,22}, and recognising that all of the significant and detrimental side reactions of **AA-CN** [as well as the destruction of **AA-SH** and peptides by COS and CO₂ (see above)]¹¹ require a free α -amine moiety, we realised that nitrile to thioacid conversion should not be pursued for α -aminonitriles (**AA-CN**), but at the C-terminus of a growing peptide, the simplest case being **Ac-AA-CN** (Figure 1b of manuscript). *N*-terminal acylation (and acetylation) would prevent the detrimental side reactions observed during thioamide hydrolysis **AA-SNH₂**, and also promote thioamide hydrolysis. Reordering the sequence of ligation and activation required for peptide coupling would then avert the problems associated with prebiotic thioacid synthesis and aminonitrile ligation (Figure 1b manuscript).

Carbonyl sulfide-mediated N-terminal hydantoin formation of peptides

The controlled syntheses of peptides through iterative coupling of *N*-carboxyanhydrides (**NCA**, X=O) and α -amino acids **AA** in water are well-documented in a seminal series of publications by an industrial team led by Ralph Hirschmann ('*The (Controlled) Synthesis of Peptides in Aqueous Medium*')²³⁻²⁷.



Supplementary Figure 1. The incompatibility of monomer-specific α -amino acid **AA** activation to form an **NCA** in the presence of growing peptides **(AA)_n-OH** ($n > 1$). **NCA** formation requires a free amino moiety and intramolecular cyclisation to achieve amino acid **AA** activation; this strategy is incompatible with the free amino terminus of peptides **(AA)_n-OH**, leading to irreversible urea (**HA**) and hydantoin (**HD**) formation. The requirement for the amino derivatisation of **AA** and a free N-terminal amino moiety of a peptide **(AA)_n-OH** for **NCA**-type C→N peptide growth is mutually incompatible. See text for further details.

Their industrial objectives warranted a comprehensive and careful investigation to elucidate the scope and limitations of **NCA** chemistry. Notably, they had encountered severe complications with **Gly-NCA** ($R=H$; $X=O$), the simplest and likely most abundant representative of prebiotic α -amino acids **AA**²⁴⁻²⁶. Substantial formation of hydantoins **HD** and hydantoic acids **HA** ("ureas") were observed with **Gly-NCA**. This reactivity is attributed to there being little discrimination of nucleophilic attack by amines between C2 and C5 of **Gly-NCA**²⁸. Further complications were observed with the histidine variant (**His-NCA**). Some of these deficiencies were remedied by the use of *N*-thiocarboxyanhydrides (**NTA**, $X=S$), the products of carbonyl sulfide (**COS**) and α -amino acids **AA**^{24,26}, but complications during **NCA** formation of **Ser**, **Thr**, **Arg**, **Cys**, **Tyr**, and **Lys** all had to be addressed by using appropriate protecting groups.

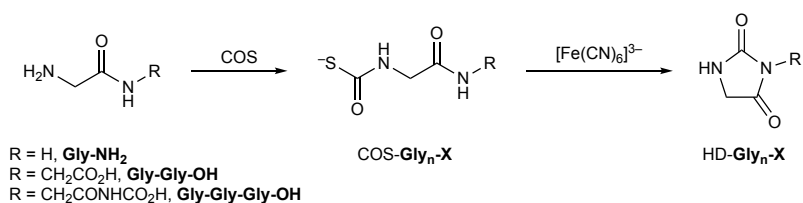
The first mention of **NCA**s as potentially prebiotic acylating agents was noted by Hirschmann when it was observed that O,S-thiocarbamate **COS-Phe** ($R=\text{CH}_2\text{C}_6\text{H}_5$) gave trace amounts of dipeptide **Phe-Phe-OH** and the corresponding hydantoic acid **HA-Phe** ($R_1=R_2=\text{CH}_2\text{C}_6\text{H}_5$; "urea")²⁶. The proposed mechanism is shown in Supplementary Figure 1. Hirschmann et al. specified that **COS** may '...offer another pathway to peptide formation under prebiotic conditions'²⁶. It is likely that aerial oxidation of **COS-Phe** promoted dipeptide **Phe-Phe-OH** formation² because it is well precedented that O,S-thiocarbamates (and the closely related S,S-dithiocarbamates) are readily activated by numerous methods that include oxidation and alkylation²⁹⁻³³. Leman, Orgel and Ghadiri investigated the rate of **NCA** formation from thiocarbamate **COS-Phe** accelerated by metal ions (Pb^{2+} , $\text{Fe}^{2+/3+}$, Zn^{2+} , Cd^{2+}), oxidants (O_2 , $[\text{Fe}(\text{CN})_6]^{3-}$) and alkylating agents, leading to the formation of predominantly dipeptides². Preactivated **COS-Phe** gave significantly higher quantities of dipeptides **(AA)₂-OH** compared to reactions carried out directly with **Phe**, where tripeptides **(AA)₃-OH** are only

observed in detectable quantities. Multiple hydantoin **HD** and urea **HA** by-products are also observed (Supplementary Figure 1)^{2,34}.

We were concerned that the **COS** needed to activate an α -amino acid **AA** to prepare **NCA**s would also destroy the peptide products. Therefore we subjected **Gly-NH₂**, **Gly-Gly-OH** and **Gly-Gly-Gly-OH** to **COS** under previously reported conditions for peptide synthesis² and observed quantitative formation of the corresponding thiocarbamate **COS-AA-X** in all cases (Supplementary Table 1). Under the conditions required for the conversion of amino acid thiocarbamates **COS-AA** into **NCA** (including metal ions, oxidation, and S-alkylation), the thiocarbamates **COS-AA-X** all underwent rapid and near-quantitative conversion to a stable hydantoin **HD** (Supplementary Table 1). These observations are in agreement with previously reported hydantoin formation of phenylthiocarbonyl dipeptide esters³⁴, and are mechanistically reminiscent of the Edman-type degradation of peptides.

Stable hydantoin **HD** formation irreversibly blocks the *N*-terminus of a peptide (growing from C→N terminus) to further elongation and terminates peptide synthesis (Supplementary Figure 1.). Thiohydantoins arising from the analogous reactions with **CS₂** have also been characterised³⁵. These observations have serious implications for the proposal that **COS** (or **CS₂**) could have mediated prebiotic peptide synthesis. Peptides formed by this mechanism are vulnerable to irrevocable *N*-terminal derivitisation under the conditions of their synthesis. This renders the peptide products incompatible with the conditions of α -amino acid **AA** activation via **COS-AA**, and necessitates prebiotic **NCA** chemistry to be a two-step process (1) **COS-AA** formation in an environment devoid of peptide; (2) **COS-AA** delivery and activation in an environment now devoid of **COS**. If **COS** is mixed with the growing peptide, **COS**-mediated termination will become manifest at the earliest stage possible (i.e. dipeptide **(AA)₂-OH**) during peptide synthesis. Furthermore, peptides **(AA)_n-OH** ($n>1$) will outcompete α -amino acids **AA** for **COS** because of the lower pK_{aH} of the *N*-terminal peptide amino groups ($k_2>k_1$)³⁵. The strict spatiotemporal separation of monomer-thiocarbamates **COS-AA** and peptides **(AA)_n-OH**, and that peptide environments needs to be devoid of **COS** has been discussed recently by Greenwald *et al.*, who report **COS**-mediated peptide synthesis requires spatiotemporal separation of thiocarbamate **COS-AA** formation and the activation event to reduce the observed amounts of hydantoin **HD** and urea **HA** by-products³⁶.

Carbonyl sulfide-mediated N-terminal hydantoin formation of peptides

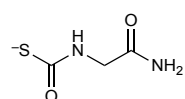


Carbonyl sulfide (**COS**) was bubbled into a stirring solution of glycineamide **Gly-NH₂**, diglycine **Gly-Gly-OH**, or triglycine **Gly-Gly-Gly-OH** (25 mM, 1.0 mmol) in 500 mM borate buffer (4 mL, pH 9.3) at room temperature until quantitative conversion to the corresponding thiocarbamate **COS-Gly_n-X** was observed by NMR spectroscopy (10 – 20 mins). An activating reagent (2 equiv.) was added and the reaction was stirred vigorously. The reaction was analysed within 20 min by NMR spectroscopy. High yielding conversion of the thiocarbamate **COS-Gly_n-X** to hydantoin **HD-Gly_n-X** was observed in all cases. Hydantoin **HD-Gly_n-X** formation was further confirmed in the case for **Gly-NH₂** by spiking with authentic imidazolidine-2,4-dione **HD-Gly-NH₂** ('hydantoin').

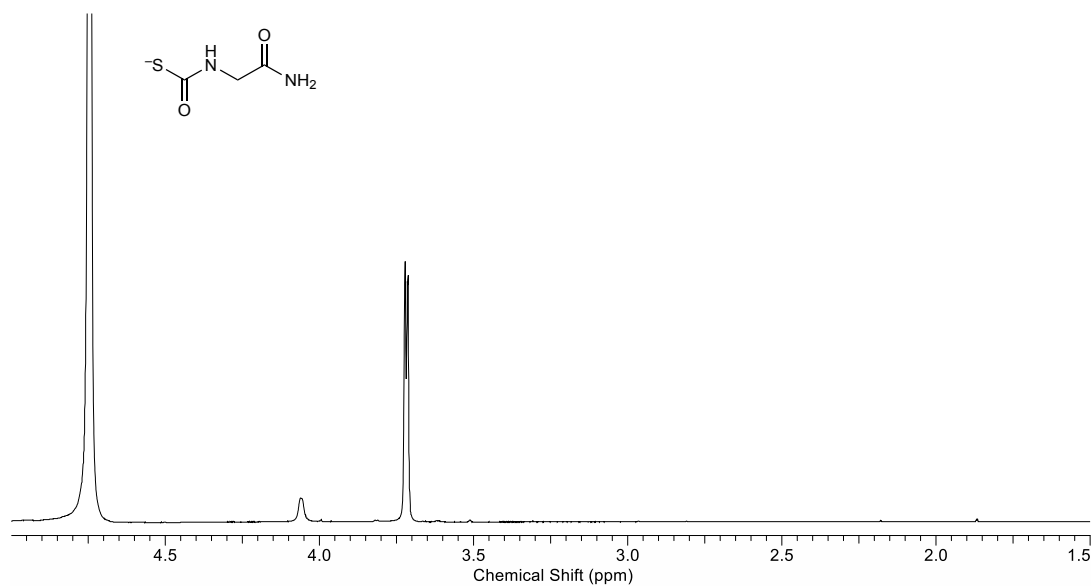
Activating agent	Hydantoin (HD-Gly_n-X) yield (%)		
	COS-Gly-NH₂ → HD-Gly-NH₂	COS-Gly₂-OH → HD-Gly₂-OH	COS-Gly₃-OH → HD-Gly₃-OH
K ₃ [Fe(CN) ₆]	>95	>95	>95
Pb(OAc) ₂	-	>95	-
HCCCN	-	>95	-
CuCl ₂	-	90	-
CuCl	-	83	-

Supplementary Table 1. Conversion of thiocarbamate **COS-Gly_n-X** (25 mM) to hydantoin **HD-Gly_n-X** upon addition of an additive (50-75 mM) in borate buffer (500 mM) at pH 9.3. (–) = reaction not carried out.

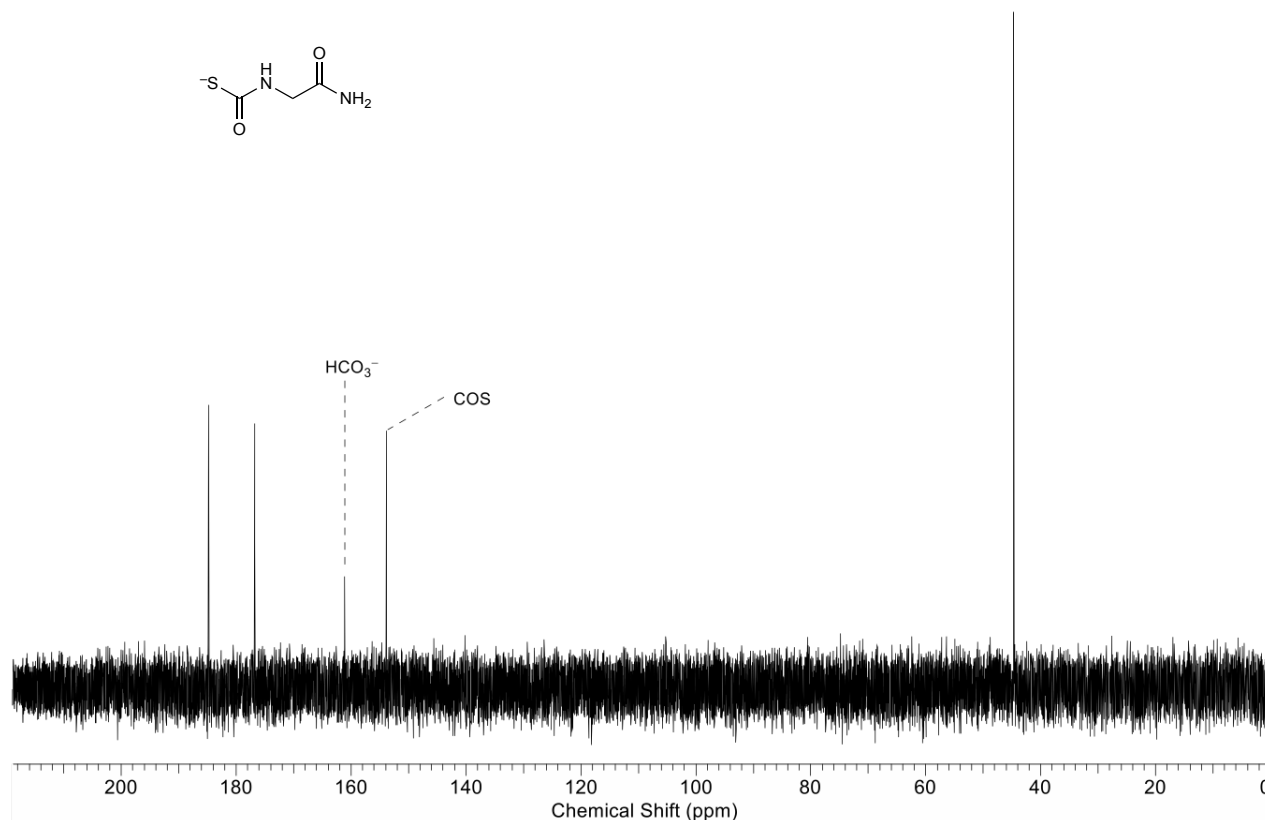
COS-Gly-NH₂



¹H NMR (700 MHz, H₂O, noesygppr1d) δ 3.72 (dd, *J* = 1.3, 5.8 Hz, 2H, (C2)–H₂). ¹³C NMR (176 MHz, H₂O) δ 184.8 (COS[−]), 176.8 (C1), 44.7 (C2).

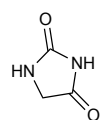


Supplementary Figure 2. ¹H NMR (700 MHz, H₂O, noesygppr1d, 1.50 – 5.00 ppm) spectrum to show **COS-Gly-NH₂**.

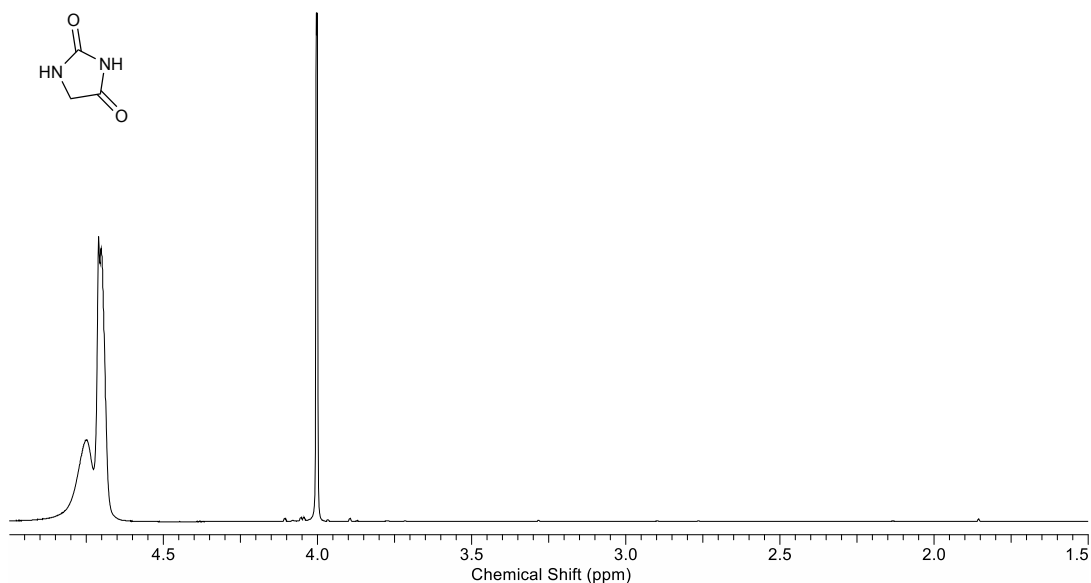


Supplementary Figure 3. ¹³C NMR (176 MHz, H₂O, 0 – 220 ppm) spectrum to show **COS-Gly-NH₂**.

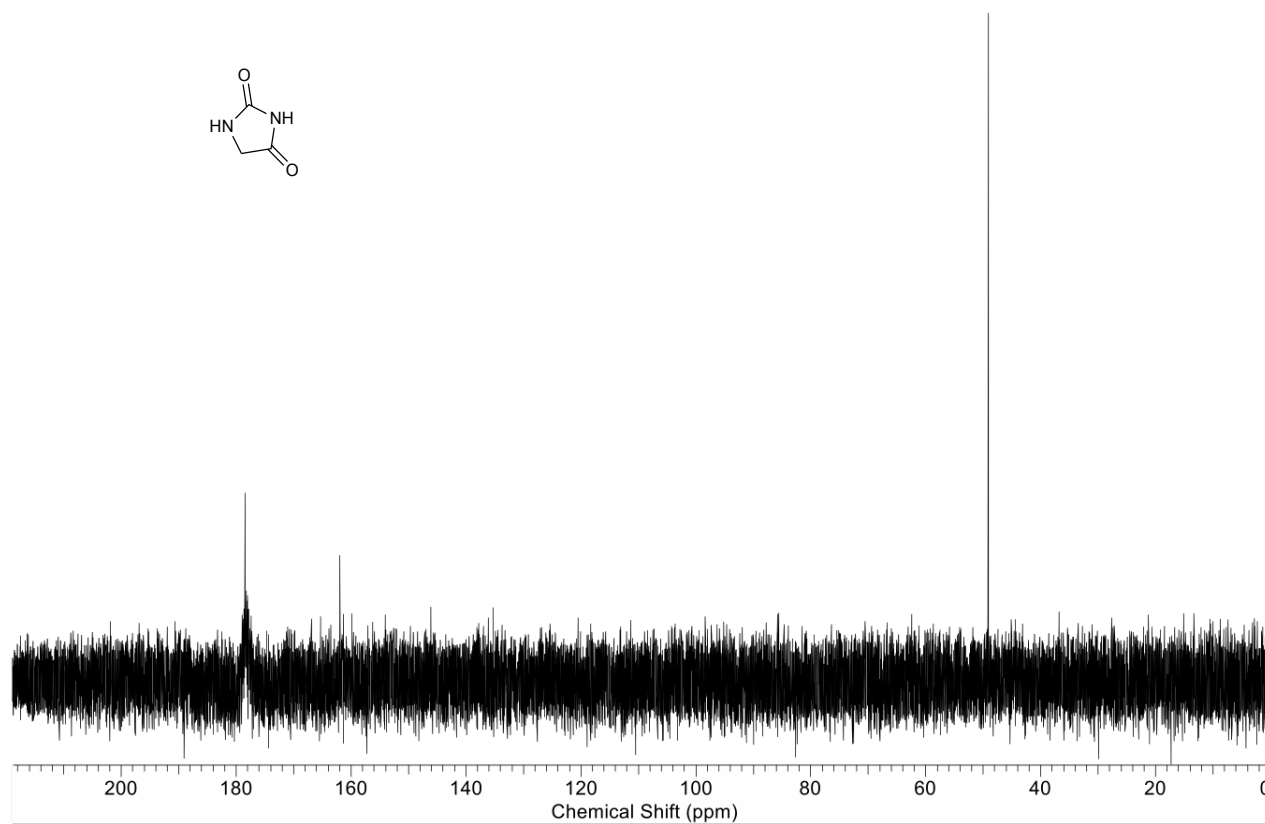
Hydantoin **HD-Gly-NH₂**



¹H NMR (700 MHz, H₂O, noesygppr1d) δ 4.00 (d, *J* = 1.8 Hz, 2H, (C5)–H₂). ¹³C NMR (176 MHz, H₂O) δ 178.5 (C4), 161.9 (C2), 49.1 (C5). HRMS-ESI [*M*+Li]⁺ calcd. for C₃H₄LiN₂O₂ 107.0433; obs. 107.0437.

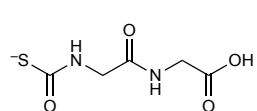


Supplementary Figure 4. ¹H NMR (700 MHz, H₂O, noesygppr1d, 1.50 – 5.00 ppm) spectrum to show **HD-Gly-NH₂**.

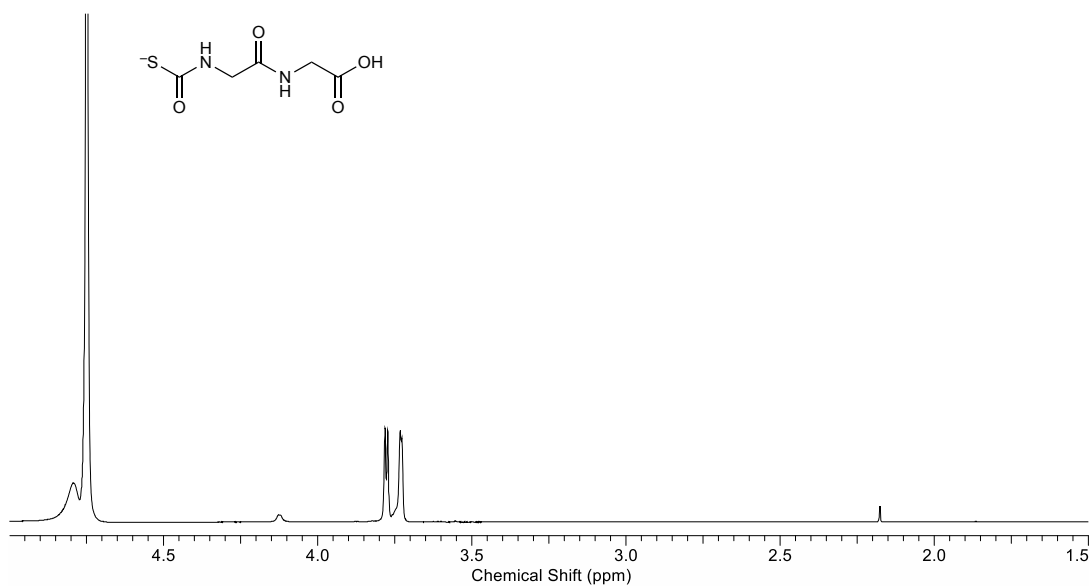


Supplementary Figure 5. ¹³C NMR (176 MHz, H₂O, 0 – 220 ppm) spectrum to show **HD-Gly-NH₂**.

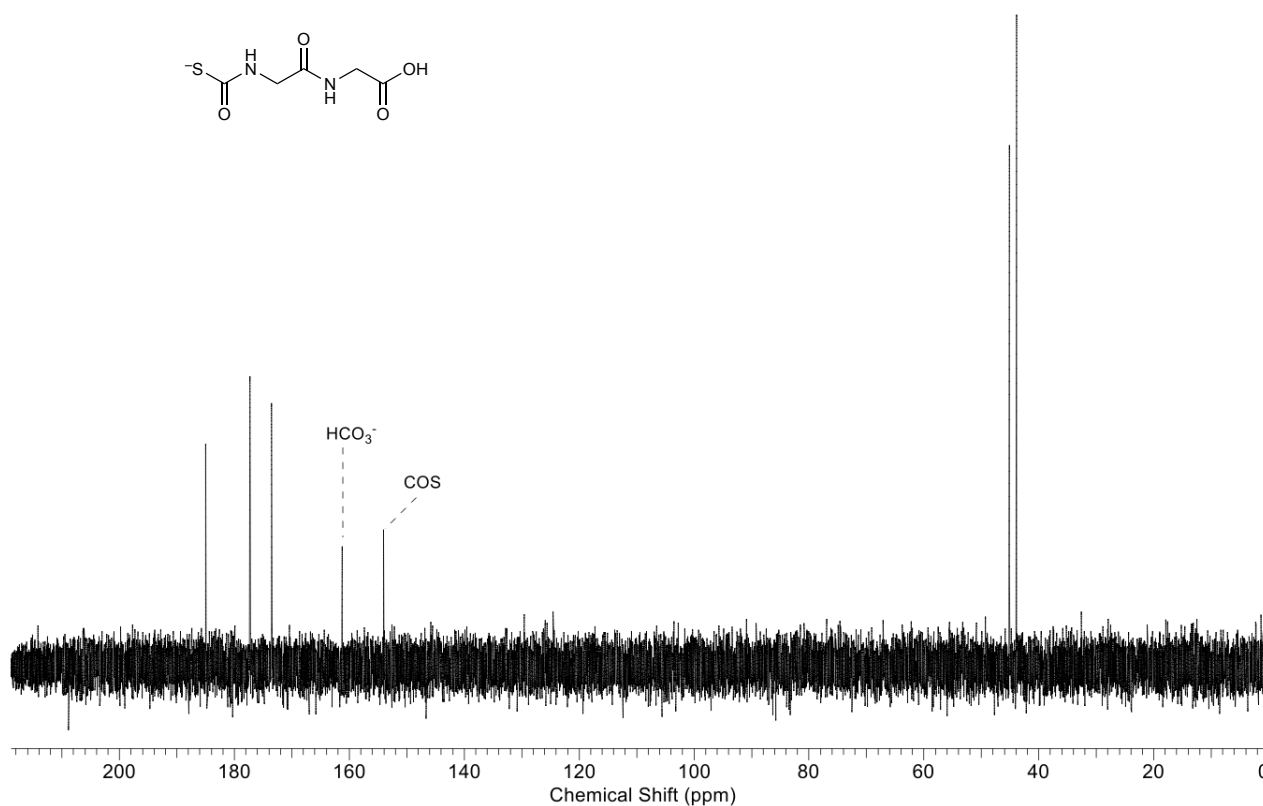
COS-Gly₂-OH



¹H NMR (700 MHz, H₂O, noesygppr1d) δ 3.78 (d, *J* = 6.1 Hz, 2H, Gly₁-(C2)-H₂), 3.73 (d, *J* = 3.8 Hz, 2H, Gly₂-(C2)-H₂). ¹³C NMR (176 MHz, H₂O) δ 185.0 (COS⁻), 177.3 (Gly₂-(C1)), 173.5 (Gly₁-(C1)), 45.1 (Gly₁-(C2)), 43.9 (Gly₂-(C2)).

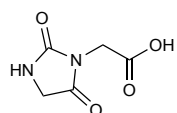


Supplementary Figure 6. ¹H NMR (700 MHz, H₂O, noesygppr1d, 1.50 – 5.00 ppm) spectrum to show **COS-Gly₂-OH**.

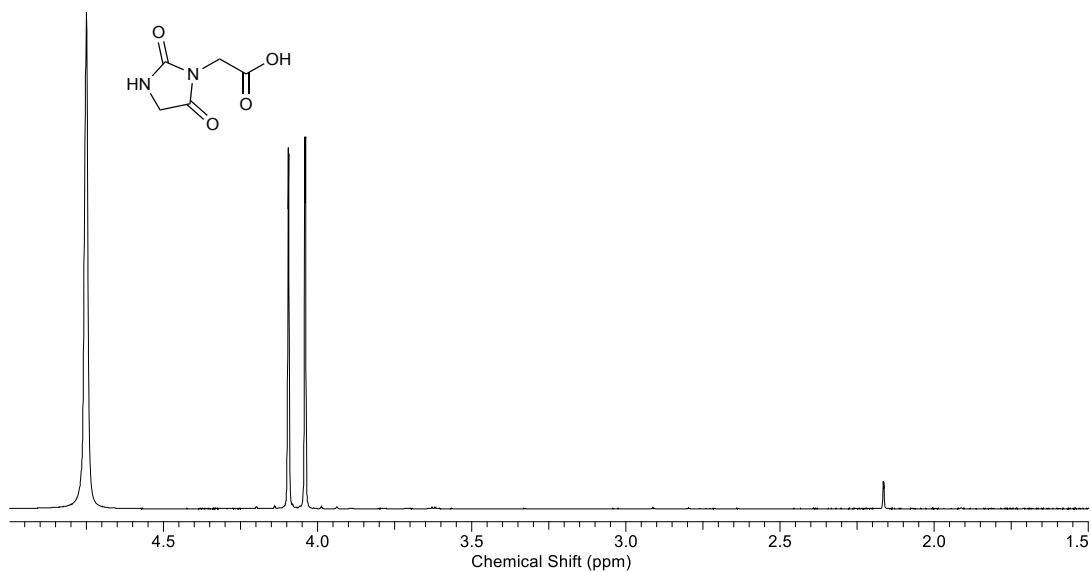


Supplementary Figure 7. ¹³C NMR (176 MHz, H₂O, 0 – 220 ppm) spectrum to show **COS-Gly₂-OH**.

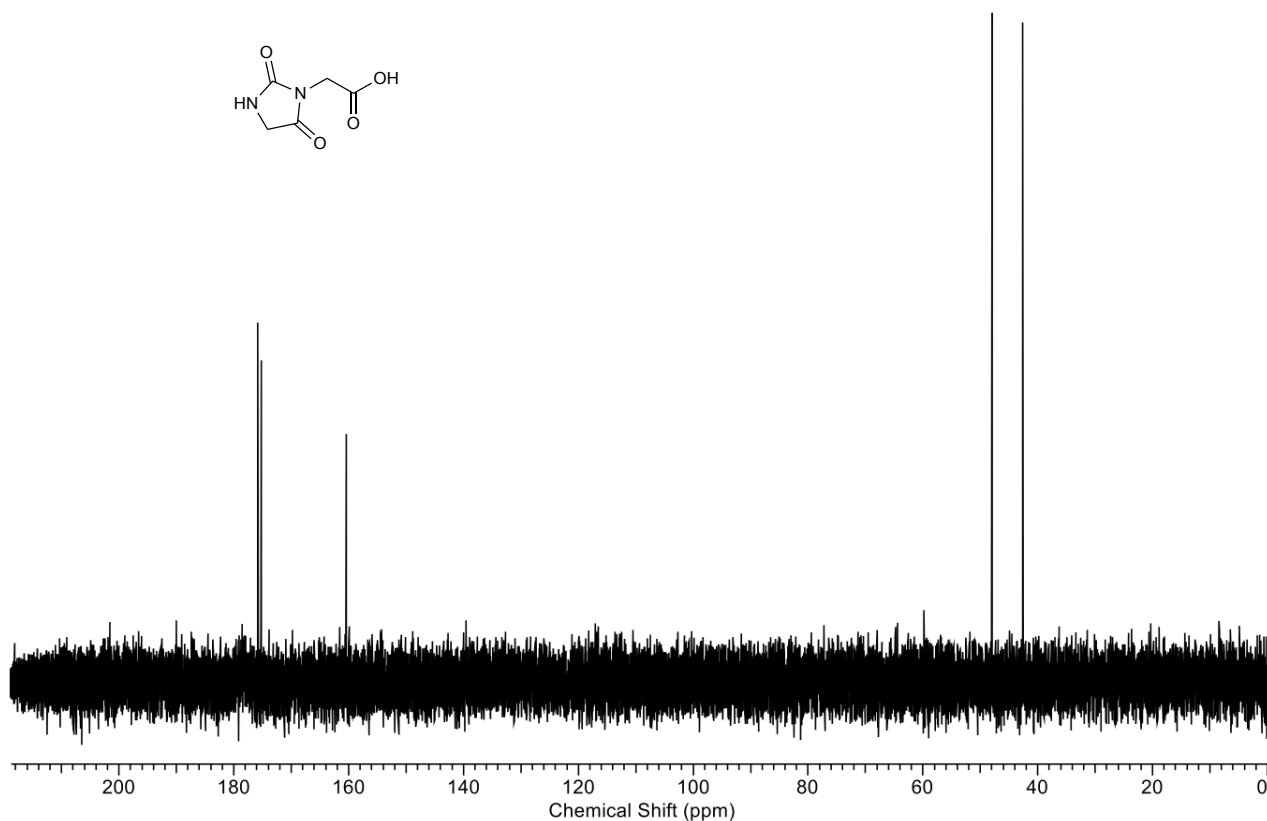
Hydantoin-3-acetic acid *HD-Gly₂-OH*



^1H NMR (700 MHz, H_2O , noesygppr1d) δ 4.09 (d, $J = 1.6$ Hz, 2H, (C5)– H_2), 4.04 (d, $J = 1.6$ Hz, 2H, $\text{CH}_2\text{CO}_2\text{H}$). ^{13}C NMR (176 MHz, H_2O) δ 175.8 (C4), 175.2 ($\text{CH}_2\text{CO}_2\text{H}$), 160.4 (C2), 48.0 (C5), 42.6 ($\text{CH}_2\text{CO}_2\text{H}$). HRMS-ESI $[\text{M}+2\text{K}-\text{H}^+]^+$ calcd. for $\text{C}_5\text{H}_5\text{K}_2\text{N}_2\text{O}_4$ 234.9523; obs. 234.9523.

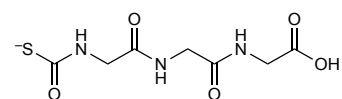


Supplementary Figure 8. ^1H NMR (700 MHz, H_2O , noesygppr1d, 1.50 – 5.00 ppm) spectrum to show **HD-Gly₂-OH**.

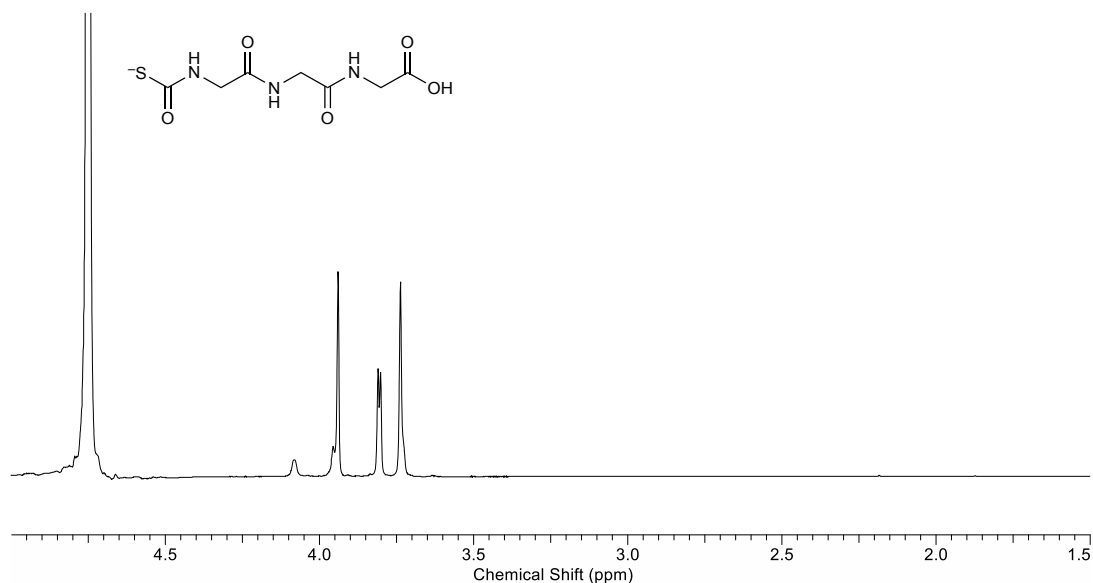


Supplementary Figure 9. ^{13}C NMR (176 MHz, H_2O , 0 – 220 ppm) spectrum to show **HD-Gly₂-OH**.

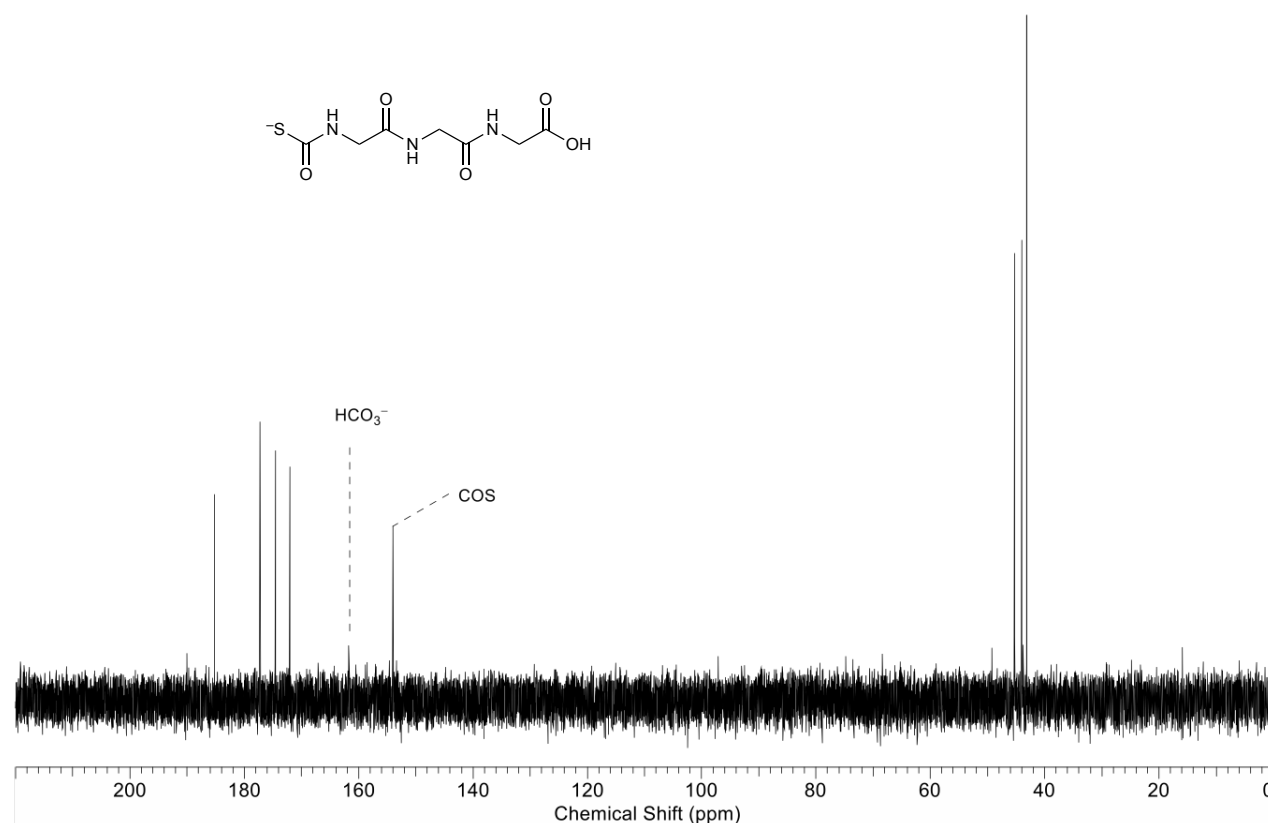
COS-Gly₃-OH



¹H NMR (700 MHz, H₂O, noesygppr1d) δ 3.94 (s, 2H, Gly₂-(C2)-H₂), 3.81 (d, *J* = 5.6 Hz, 2H, Gly₁-(C2)-H₂), 3.74 (s, 2H, Gly₃-(C2)-H₂). ¹³C NMR (176 MHz, H₂O) δ 185.2 (COS⁻), 177.3 (Gly₃-(C1)), 174.5 (Gly₁-(C1)), 172.0 (Gly₂-(C1)), 45.3 (Gly₁-(C2)), 44.0 (Gly₂-(C2)), 43.2 (Gly₃-(C2)).

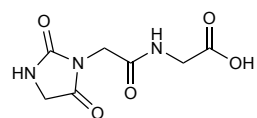


Supplementary Figure 10. ¹H NMR (700 MHz, H₂O, noesygppr1d, 1.50 – 5.00 ppm) spectrum to show **COS-Gly₃-OH**.

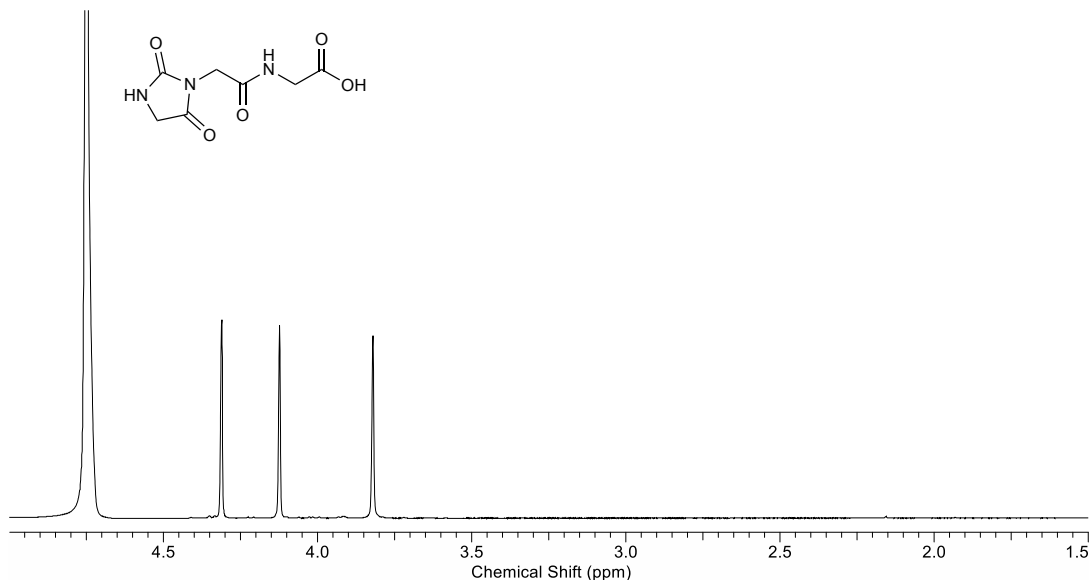


Supplementary Figure 11. ¹³C NMR (176 MHz, H₂O, 0 – 220 ppm) spectrum to show **COS-Gly₃-OH**.

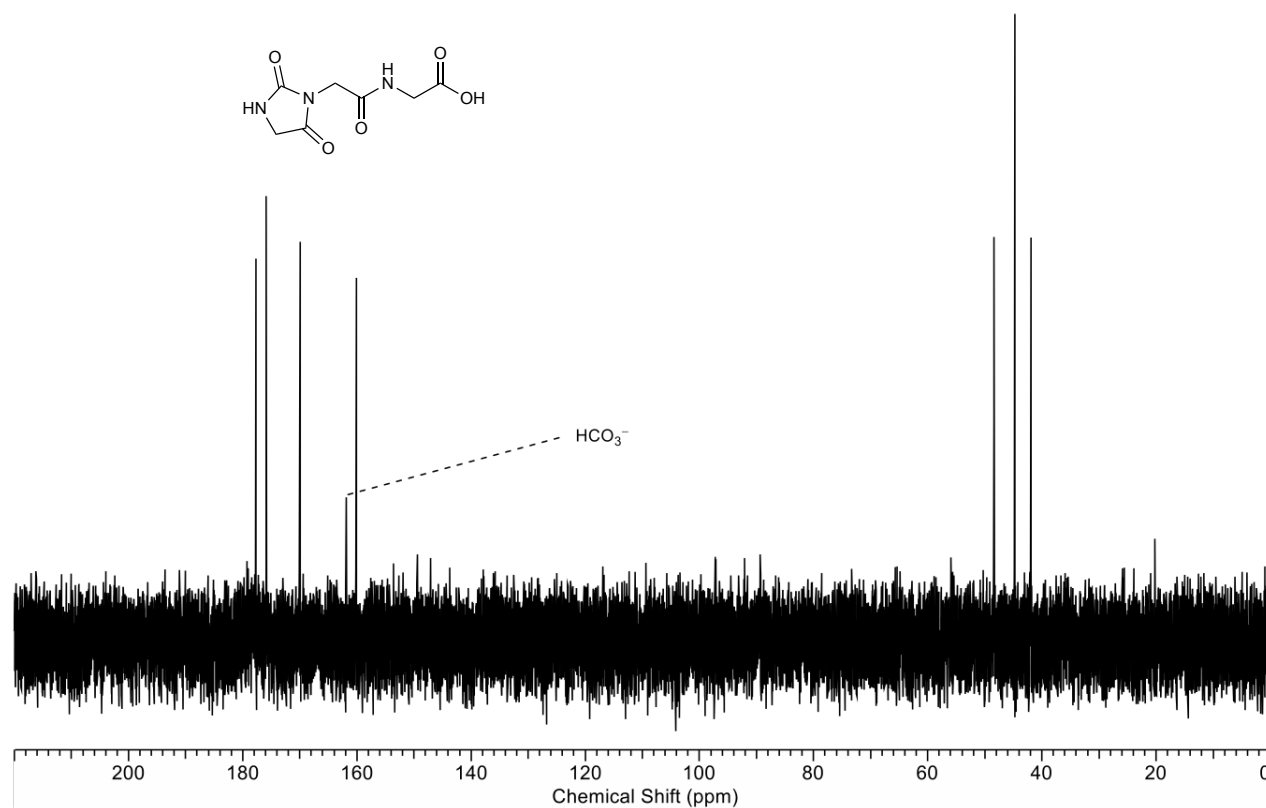
((Hydantoin-3-yl)acetyl)glycine **HD-Gly₃-OH**



^1H NMR (700 MHz, H_2O , noesygppr1d) δ 4.31 (s, 2H, Gly₁-(C2)-H₂), 4.12 (s, 2H, (C5)-H₂), 3.82 (s, 2H, Gly₂-(C2)-H₂). ^{13}C NMR (176 MHz, H_2O) δ 177.7 (Gly₂-(C1), 175.9 (C4), 170.0 (Gly₁-(C1), 160.1 (C2), 48.3 (C5), 44.8 (Gly₂-(C2), 41.9 (Gly₁-(C2)). HRMS-ESI $[\text{M}-\text{H}]^-$ calcd. for $\text{C}_7\text{H}_8\text{N}_3\text{O}_5$ 214.0464; obs. 214.0463.



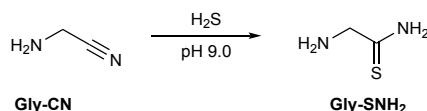
Supplementary Figure 12. ^1H NMR (700 MHz, H_2O , noesygppr1d, 1.50 – 5.00 ppm) spectrum to show **HD-Gly₃-OH**.



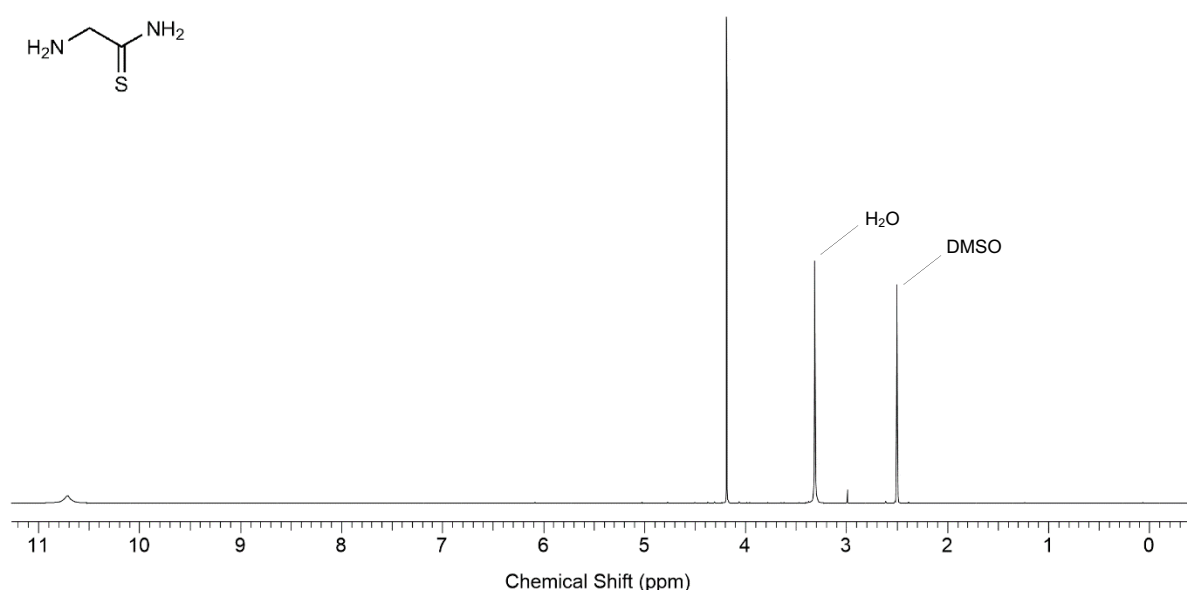
Supplementary Figure 13. ^{13}C NMR (176 MHz, H_2O , 0 – 220 ppm) spectrum to show **HD-Gly₃-OH**.

Thiolysis of Aminonitriles

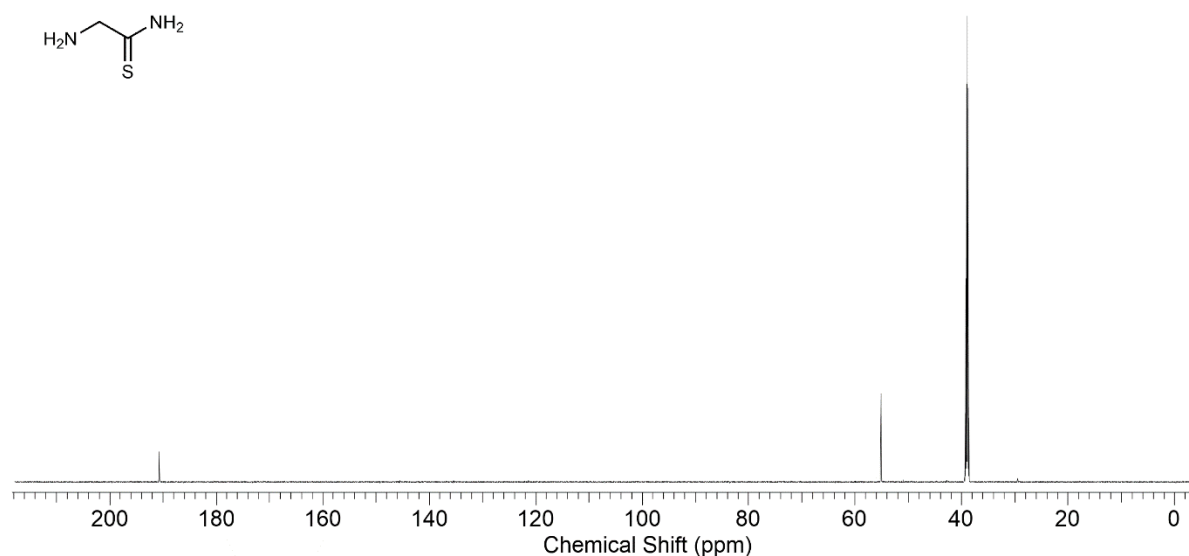
Thiolysis of glycine nitrile



Glycine nitrile (**Gly-CN**·HCl, 46.3 mg, 0.50 mmol), NaSH·xH₂O (561 mg, 5.00 mmol) and methylsulfonylmethane (0.50 mmol) were dissolved in degassed H₂O/D₂O (98:2, 10 mL). The solution was adjusted to pH 9.0 with NaOH/HCl, and NMR spectra were periodically acquired. A white precipitate was observed after 16 h, which started to crystallise after 5 d. After 7 d, NMR analysis showed complete conversion of **Gly-CN** to glycine thioamide **Gly-SNH₂**. The crystals were isolated by filtration and dissolved in DMSO-*d*₆. NMR analysis identified the solid as glycine thioamide **Gly-SNH₂** (45.1 mg, 0.50 mmol, 99%). ¹H NMR (700 MHz, DMSO-*d*₆): δ 10.71 (br. s., 1H, NH) 4.18 (s, 2H, (C2)–H₂). ¹³C NMR (176 MHz, DMSO-*d*₆): δ 190.7 (C1), 55.2 (C2). IR (solid): 3181, 3085, 2985, 2880, 2179, 1577 cm⁻¹. M.p 71.2–71.6 °C (decomp.)



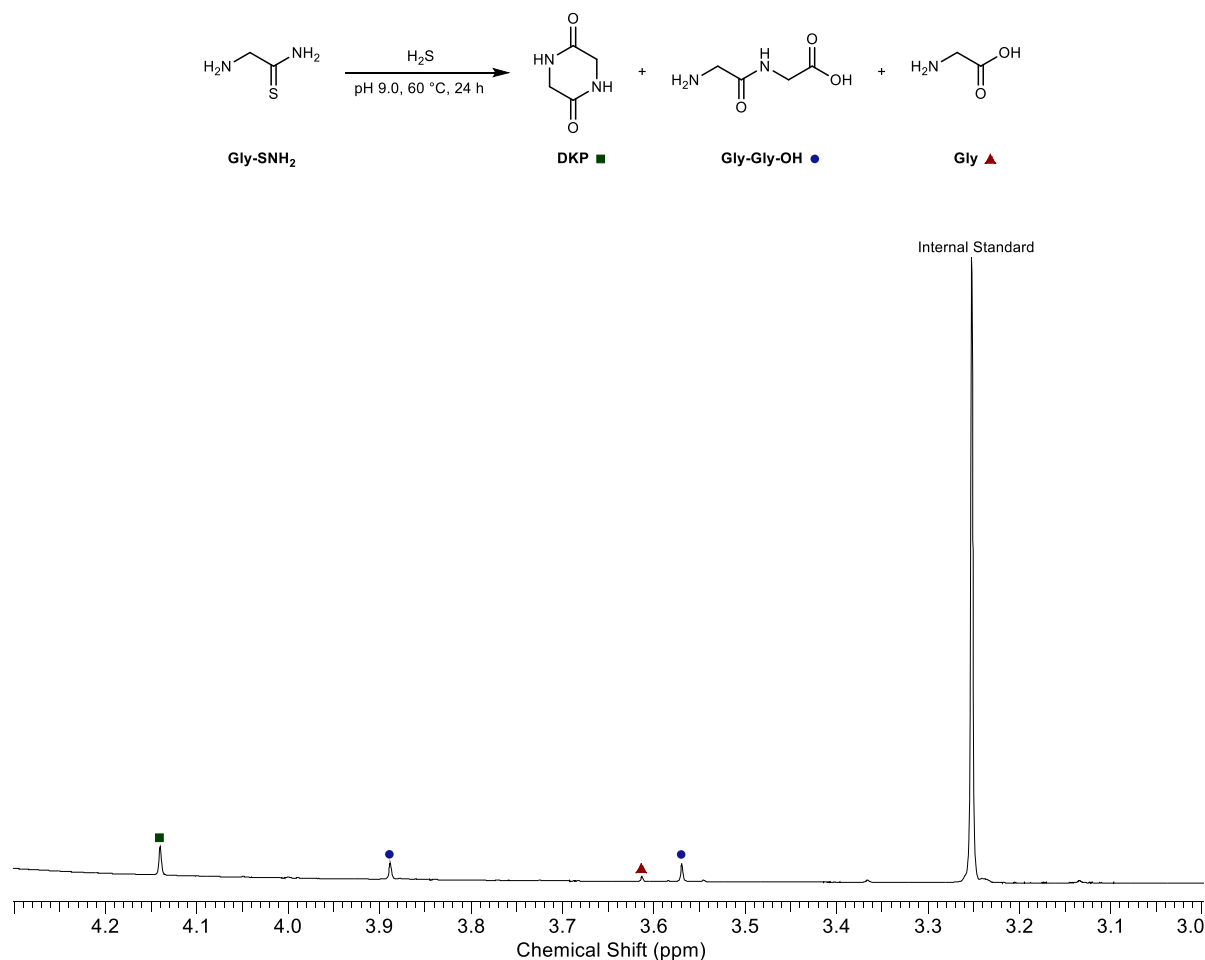
Supplementary Figure 14. ¹H NMR (700 MHz, DMSO-*d*₆, -0.40–11.20 ppm) spectrum of glycine thioamide **Gly-SNH₂**.



Supplementary Figure 15. ¹³C NMR (176 MHz, DMSO-*d*₆, 0–220 ppm) spectrum of glycine thioamide **Gly-SNH₂**.

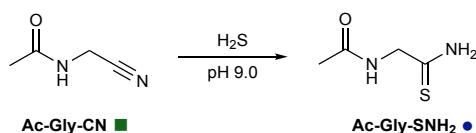
Attempted hydrolysis of glycine thioamide to glycine thioacid

Glycine thioamide **Gly-SNH₂** (9.01 mg, 0.10 mmol) was added to a solution of methylsulfonylmethane (0.10 mmol) in degassed H₂O/D₂O (98:2, 2 mL). The solution was adjusted to pH 9.0 with NaOH/HCl, and incubated at 60 °C for 24 h. After that time, NMR spectra were acquired. The presence of piperazine-2,5-dione **DKP** (25%), glycylglycine **Gly-Gly-OH** (14%) and glycine **Gly** (3%) was confirmed by ¹H–¹³C HMBC NMR analysis and spiking with pure synthetic standards. The observed products were quantified using methylsulfonylmethane (MSM) as an internal standard (Supplementary Figure 16).

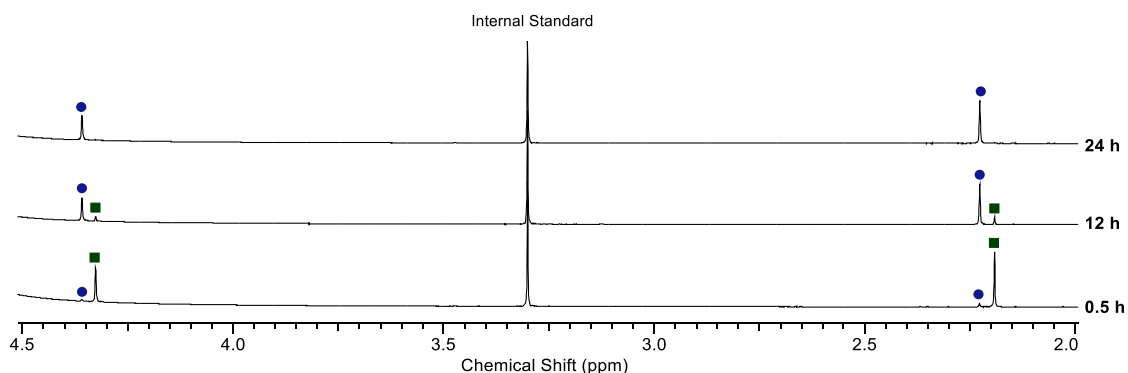


Supplementary Figure 16. ¹H NMR (600 MHz, D₂O/H₂O 2:98, 3.00–4.30 ppm) spectrum to show the incubation of glycine thioamide **Gly-SNH₂** (50 mM) in D₂O/H₂O (2:98, 2 mL) at 60 °C for 24 h.

Thiolysis of *N*-acetylglycine nitrile



N-Acetylglycine nitrile **Ac-Gly-CN** (24.5 mg, 0.25 mmol) and NaSH·xH₂O (280 mg, 2.50 mmol) were dissolved in degassed H₂O/D₂O (98:2, 5 mL). The solution was adjusted to pH 9.0 with HCl/NaOH, and stirred at room temperature. NMR spectra were periodically acquired until the quantitative conversion of *N*-acetylglycine nitrile **Ac-Gly-CN** to *N*-acetylglycine thioamide **Ac-Gly-SNH₂** was observed after 24 h (Supplementary Figure 17).

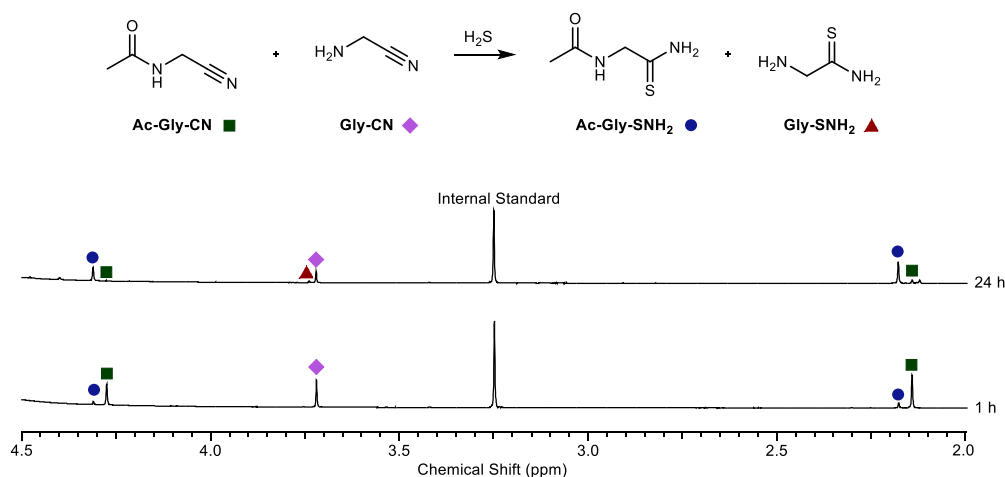


Supplementary Figure 17. ¹H NMR (600 MHz, H₂O/D₂O 98:2, pH 9.0, 2.0–4.5 ppm) spectra to show the reaction of *N*-acetylglycine nitrile (**Ac-Gly-CN**; 50 mM) with NaSH (500 mM) to yield *N*-acetylglycine thioamide (**Ac-Gly-SNH₂**) at pH 9.0 and room temperature after 0.5 h, 12 h and 24 h.

¹H NMR (600 MHz, H₂O/D₂O 98:2, pH 9.0, 2.0–4.5 ppm): *N*-Acetylglycine thioamide **Ac-Gly-SNH₂** (●): δ 4.36 (s, 2H, (C2)–H₂), 2.23 (s, 3H, (COCH₃)). *N*-Acetylglycine nitrile **Ac-Gly-CN** (■): δ 4.33 (s, 2H, (C2)–H₂), 2.19 (s, 3H, (COCH₃)).

Chemoselective thiolysis of N-acetylglycine nitrile in the presence of glycine nitrile

N-Acetylglycine nitrile **Ac-Gly-CN** (0.10 mmol), glycine nitrile hydrochloride **Gly-CN**·HCl (9.25 mg, 0.10 mmol) and NaSH·xH₂O (56.1 mg, 0.50 mmol) were dissolved in a degassed H₂O/D₂O (98:2, 2 mL), and the solution was adjusted to pH 9.0 with NaOH. The solution was stirred at room temperature for 24 h. NMR spectra were periodically acquired. The conversion of *N*-acetylglycine nitrile **Ac-Gly-CN** (■) and glycine nitrile **Gly-CN** (◆) into *N*-acetylglycine thioamide **Ac-Gly-SNH₂** (●, 91%) and glycine thioamide **Gly-SNH₂** (▲, 7%) was observed (Supplementary Figure 18).

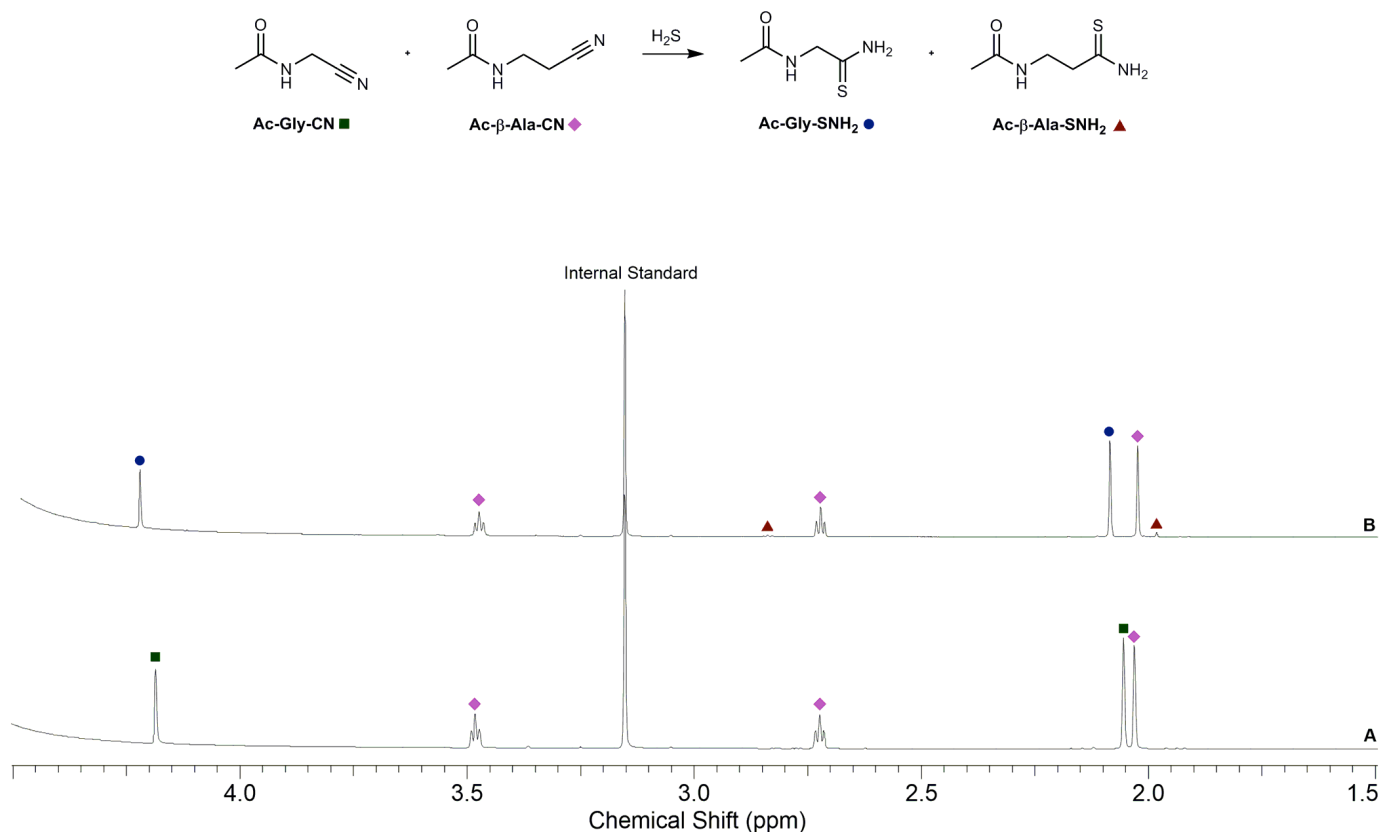


Supplementary Figure 18. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 2.00–4.50 ppm) spectra to show the reaction *N*-acetylglycine nitrile (**Ac-Gly-CN**; 50 mM), glycine nitrile (**Gly-CN**; 50 mM) and NaSH (250 mM) at pH 9.0 and room temperature after 1 h and 24 h.

¹H NMR (600 MHz H₂O/D₂O 98:2, pH 9.0, 4.50 – 2.00 ppm): *N*-Acetylglycine thioamide **Ac-Gly-SNH₂** (●): δ 4.31 (s, 2H, (C2)–H₂), 2.17 (s, 3H, (COCH₃)). Glycine thioamide **Gly-SNH₂** (▲): δ 3.73 (s, 2H, (C2)–H₂). Glycine nitrile **Gly-CN** (◆): δ 3.72 (s, 2H, (C2)–H₂). *N*-Acetylglycine nitrile **Ac-Gly-CN** (■): δ 4.27 (s, 2H, (C2)–H₂), 2.14 (s, 3H, (COCH₃)).

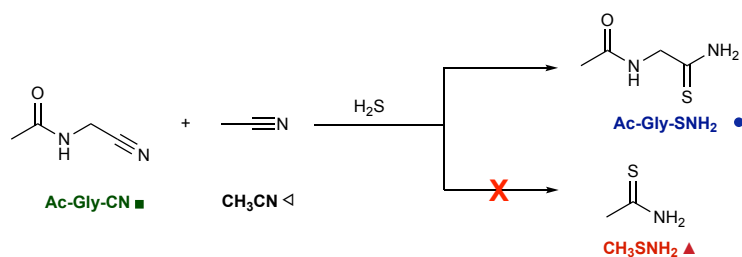
Chemoselective thiolysis of *N*-acetylglycine nitrile in the presence of *N*-acetyl- β -alanine nitrile

N-Acetylglycine nitrile **Ac-Gly-CN** (0.10 mmol), *N*-acetyl- β -alanine nitrile **Ac- β -Ala-CN** (0.10 mmol) and NaSH·xH₂O (56.1 mg, 0.50 mmol) were dissolved in a degassed H₂O/D₂O (98:2, 2 mL), and the solution was adjusted to pH 9.0 with NaOH. The solution was stirred at room temperature for 24 h. NMR spectra were periodically acquired. The conversion of *N*-acetylglycine nitrile **Ac-Gly-CN** (■) and *N*-acetyl- β -alanine nitrile **Ac- β -Ala-CN** (◆) into *N*-acetylglycine thioamide **Ac-Gly-SNH₂** (●, 99%) and *N*-acetyl- β -alanine thioamide **Ac- β -Ala-SNH₂** (▲, <1%) was observed (Supplementary Figure 19).



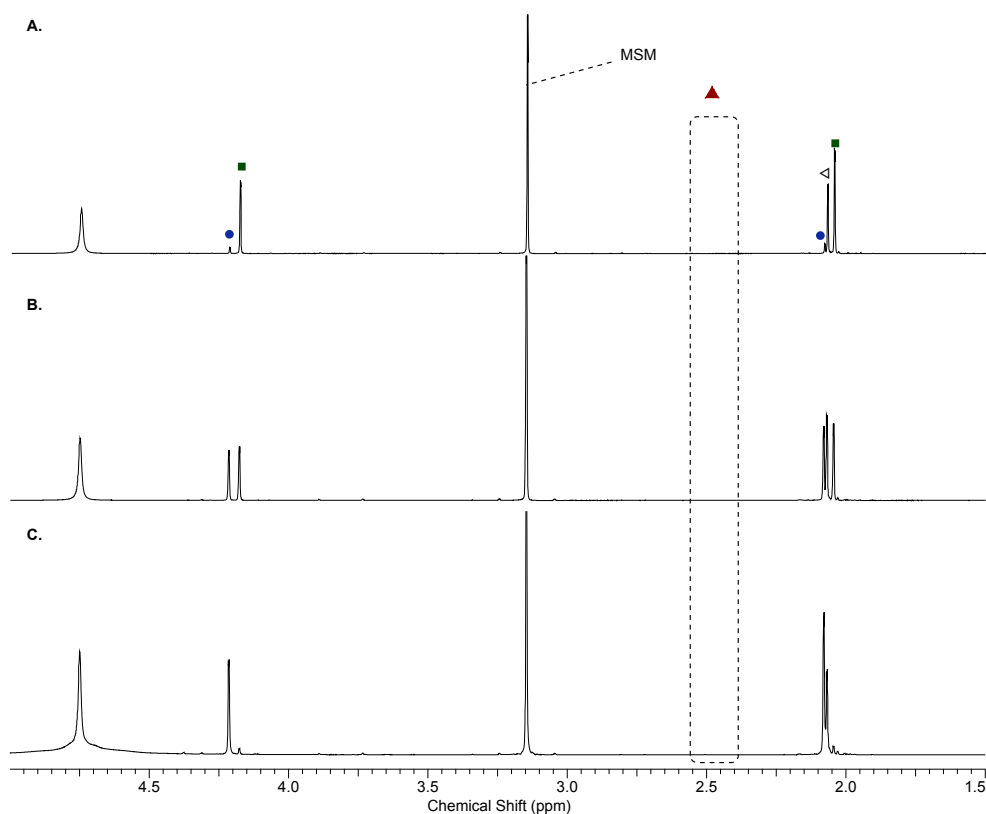
Supplementary Figure 19. ¹H NMR (700 MHz, H₂O/D₂O 98:2, 1.50–4.50 ppm) spectra to show the reaction of *N*-acetylglycine nitrile (**Ac-Gly-CN**; 50 mM), *N*-acetyl- β -alanine nitrile (**Ac- β -Ala-CN**; 50 mM) and NaSH (250 mM) at pH 9.0 and room temperature after **A**: 5 min; **B**: 24 h.

¹H NMR (700 MHz H₂O/D₂O 98:2, pH 9.0, partial assignment): *N*-Acetylglycine thioamide **Ac-Gly-SNH₂** (●): δ 4.22 (s, 2H, (C2)-H₂), 2.08 (s, 3H, (COCH₃)). *N*-Acetyl- β -alanine thioamide **Ac- β -Ala-SNH₂** (▲): δ 2.83 (obs. t, J = 6.4 Hz, 2H, (C3)-H₂), 1.98 (s, 3H, (COCH₃)). *N*-Acetyl- β -alanine nitrile **Ac- β -Ala-CN** (◆): δ 3.48 (t, J = 6.7 Hz, 2H, (C2)-H₂), 2.72 (t, J = 6.7 Hz, 2H, (C3)-H₂), 2.03 (s, 3H, (COCH₃)). *N*-Acetylglycine nitrile **Ac-Gly-CN** (■): δ 4.18 (s, 2H, (C2)-H₂), 2.06 (s, 3H, (COCH₃)).



N-Acetylglycine nitrile **Ac-Gly-CN** (0.05 mmol), acetonitrile **CH₃CN** (0.05 mmol) and NaSH·xH₂O (56.1 mg, 0.50 mmol) were dissolved in a degassed H₂O/D₂O (98:2, 1 mL), and the solution was adjusted to pH 9.0 with NaOH. The solution was stirred at room temperature for 24 h. NMR spectra were periodically acquired. The conversion of *N*-acetylglycine nitrile **Ac-Gly-CN** into *N*-acetylglycine thioamide **Ac-Gly-SNH₂** (92%) was observed, whereas **CH₃CN** remained unchanged. Reactions repeated at higher pH (pH 11 and 13) and extreme alkalinity (with 0.25 M, 2.5 M and 12.5 M NaOH), at room temperature, 60 °C and 90 °C, also gave no detectable thioacetamide **CH₃SNH₂**. Reactions carried out with 0.25 M, 2.5 M and 12.5 M NaOH led eventually to complete hydrolysis to afford glycine **Gly** and acetate **AcOH**, but thioacetamide **CH₃SNH₂** was not observed. The spectral window where the thioacetamide **CH₃SNH₂** resonance is expected in the ¹H NMR spectra (δ_H = 2.51 ppm) is boxed in Supplementary Figure 20. The absence of thioacetamide **CH₃SNH₂** formation underlines the pronounced activation of the nitrile group of **Ac-Gly-CN** through an electron-withdrawing inductive (-I) effect endowed by the acetamido-group. Our observations are in line with Cairns et. al., who reported that the addition of H₂S to MeCN requires 8500 atm and 80 °C ³⁷.

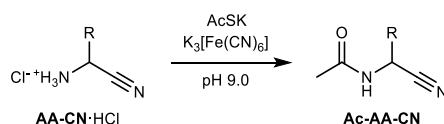
¹H NMR (700 MHz, H₂O, noesygppr1d) *N*-Acetylglycine thioamide **Ac-Gly-SNH₂** (●) δ 4.22 (s, 2H, (C2)–H₂), 2.08 (s, 3H, COCH₃). *N*-Acetylglycine nitrile **Ac-Gly-CN** (■) δ 4.18 (s, 2H, (C2)–H₂), 2.05 (s, 3H, COCH₃). Acetonitrile **CH₃CN** (▲) ¹H NMR (700 MHz, H₂O) δ 2.07 (s, 3H, CH₃)



Supplementary Figure 20. ¹H NMR (700 MHz, H₂O, noesygppr1d, 1.5–5.0 ppm) spectra to show the chemoselective thiolysis of *N*-acetylglycine nitrile (**Ac-Gly-CN**; 50 mM) in the presence of acetonitrile (**MeCN**; 50 mM) at room temperature and pH 9.0. After **A**. 5 min; **B**. 5 h; **C**. 24 h. Thiolysis of **MeCN** to thioacetamide **CH₃SNH₂** (dotted box, ▲) was not observed.

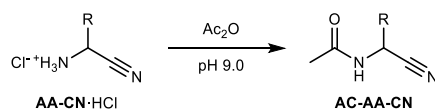
Acetylation of aminonitriles

General procedure A: Acetylation of aminonitriles with thioacetate



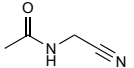
Aminonitrile hydrochloride **AA-CN·HCl** (0.1 mmol) and potassium thioacetate (0.3 mmol) were dissolved in H₂O (2 mL) and the solution was adjusted to pH 9.0 with NaOH. Potassium hexacyanoferrate(III) (296.3 mg, 0.9 mmol) was added, and the solution was stirred at room temperature for 20 min. The solution was adjusted to pH 9.0, centrifuged, and NMR spectra of the supernatant were acquired. All identified products were confirmed by signal amplification upon addition of pure synthetic standards.

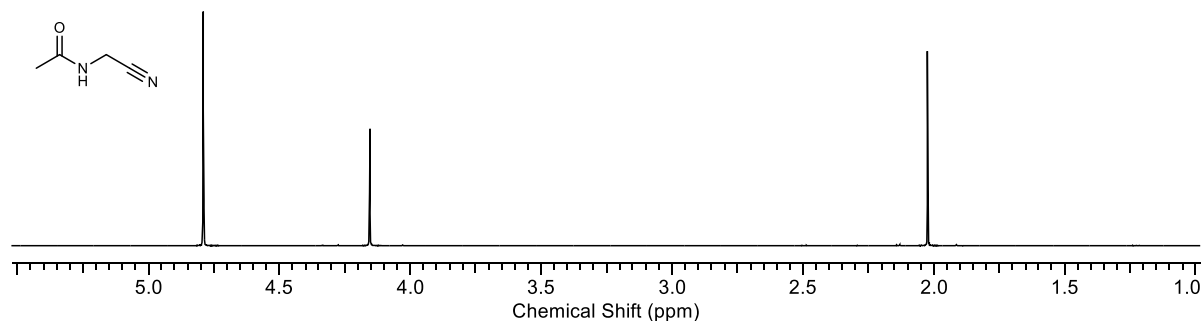
General procedure B: Acetylation of aminonitriles with acetic anhydride



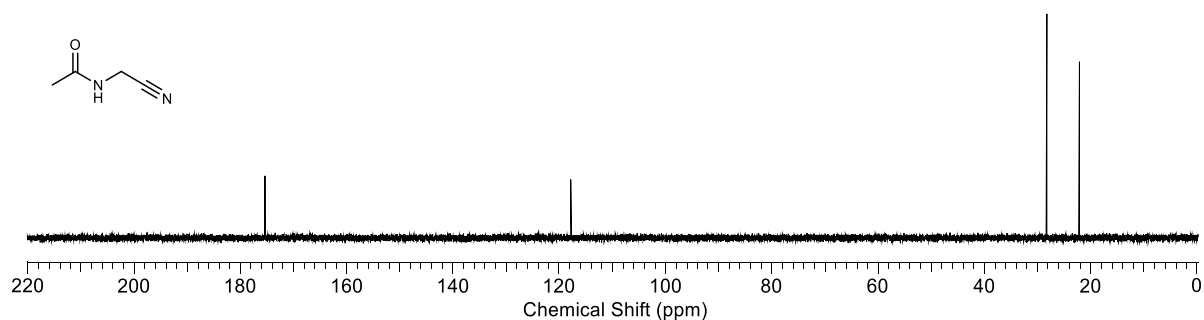
Aminonitrile hydrochloride **AA-CN·HCl** (10 mmol) was dissolved in H₂O (50 mL, 0.2 M) and the solution was adjusted to pH 9.0 with NaOH. Acetic anhydride (3.06 g, 2.86 mL, 30 mmol) was added dropwise over 10 min, whilst maintaining the solution at pH 9 with NaOH/HCl. The reaction mixture was concentrated *in vacuo* and the residue purified by flash column chromatography (SiO₂; eluting with a gradient of EtOAc/petroleum ether (40-60) 0:100 → 100:0) to afford *N*-acetylamino nitrile **Ac-AA-CN**.

N-Acetylglycine nitrile **Ac-Gly-CN**

 *N*-Acetylglycine nitrile **Ac-Gly-CN** was produced following *general procedure A* (99%). *General procedure B* using glycine nitrile hydrochloride **Gly-CN**·HCl (925 mg, 10.0 mmol) afforded *N*-acetylglycine nitrile **Ac-Gly-CN** as a white, crystalline solid (973 mg, 9.92 mmol, 99%). ¹H NMR (600 MHz, D₂O): δ 4.15 (s, 2H, (C2)–H₂), 2.03 (s, 3H, COCH₃). ¹³C NMR (151 MHz, D₂O): δ 175.4 (COCH₃), 117.8 (C1), 28.3 (C2), 22.2 (COCH₃). HRMS–CI calcd. for [C₄H₆N₂O+NH₄]⁺ 116.0818; observed: 116.0819. IR (solid): 3226, 3056, 2989, 2935, 2843, 2650, 2031, 1635, 1448, 1415 cm^{–1}. M.p. 76.3–77.0 °C ; lit. m.p. 77–79 °C.³⁸

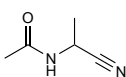


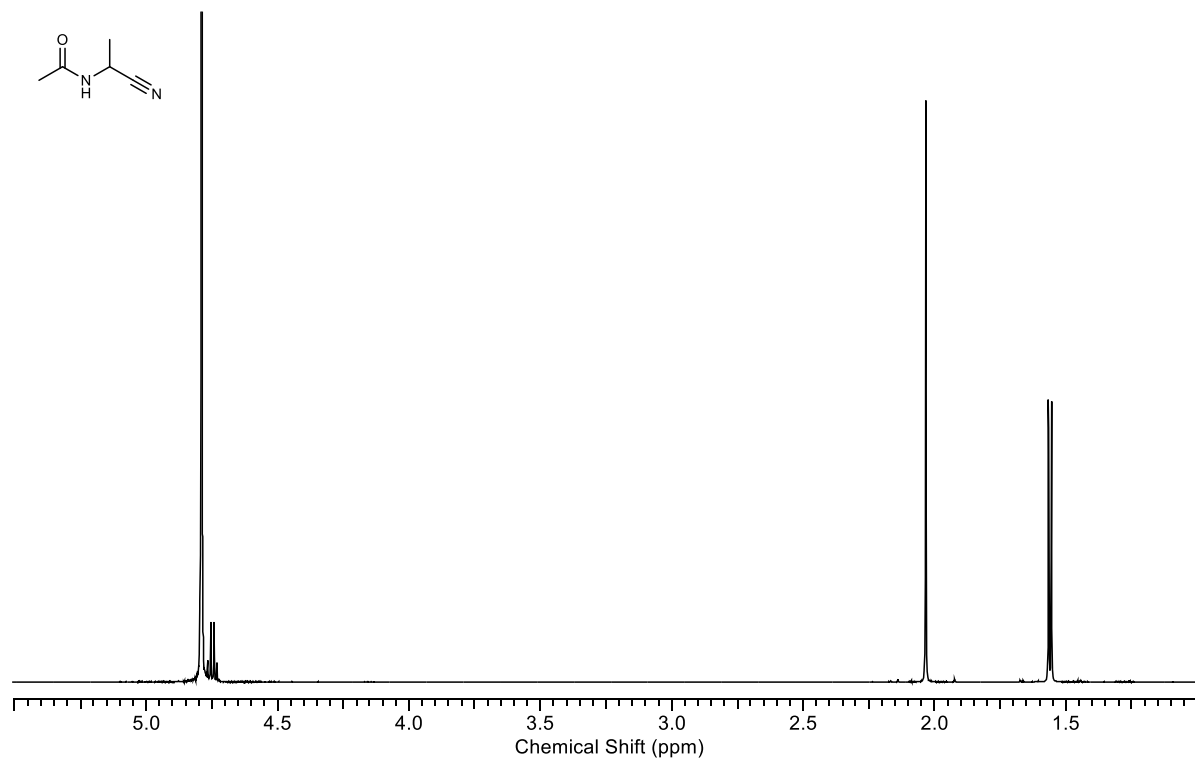
Supplementary Figure 21. ¹H NMR (600 MHz, D₂O, 1.00–5.50 ppm) spectrum to show *N*-acetylglycine nitrile **Ac-Gly-CN**



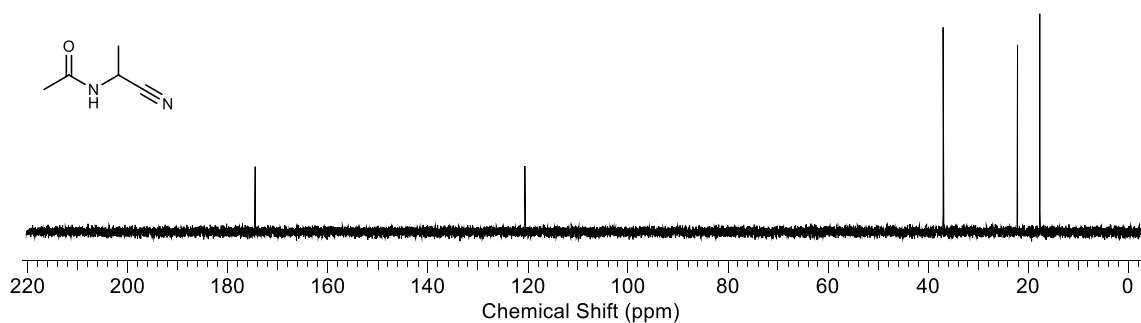
Supplementary Figure 22. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylglycine nitrile **Ac-Gly-CN**

N-Acetylalanine nitrile **Ac-Ala-CN**

 *N*-Acetylalanine nitrile **Ac-Ala-CN** was produced following *general procedure A* (99%). *General procedure B* using alanine nitrile hydrochloride **Ala-CN**·HCl³⁹ (1.07 g, 10 mmol) afforded *N*-acetylalanine nitrile **Ac-Ala-CN** as a white, crystalline solid (1.12 g, 9.99 mmol, 99%). ¹H NMR (600 MHz, D₂O): δ 4.75 (q, *J* = 7.2 Hz, 1H, (C2)–H), 2.03 (s, 3H, (COCH₃), 1.56 (d, *J* = 7.2 Hz, 3H, (C3)–H₃). ¹³C NMR (151 MHz, D₂O): δ 174.5 (COCH₃), 120.6 (C1), 37.0 (C2), 22.2 (COCH₃), 17.2 (C3). HRMS (EI⁺): calcd. for [C₅H₈N₂O]⁺ 112.0631; observed: 112.0631. IR (solid): 3238, 3048, 3011, 2954, 2838, 2807, 2243, 1734, 1633, 1357 cm⁻¹. M.p. 93.5–95.2 °C; lit. m.p. 98–101 °C.⁴⁰

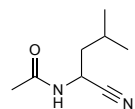


Supplementary Figure 23. ¹H NMR (600 MHz, D₂O, 1.00–5.50 ppm) spectrum to show *N*-acetylalanine nitrile **Ac-Ala-CN**

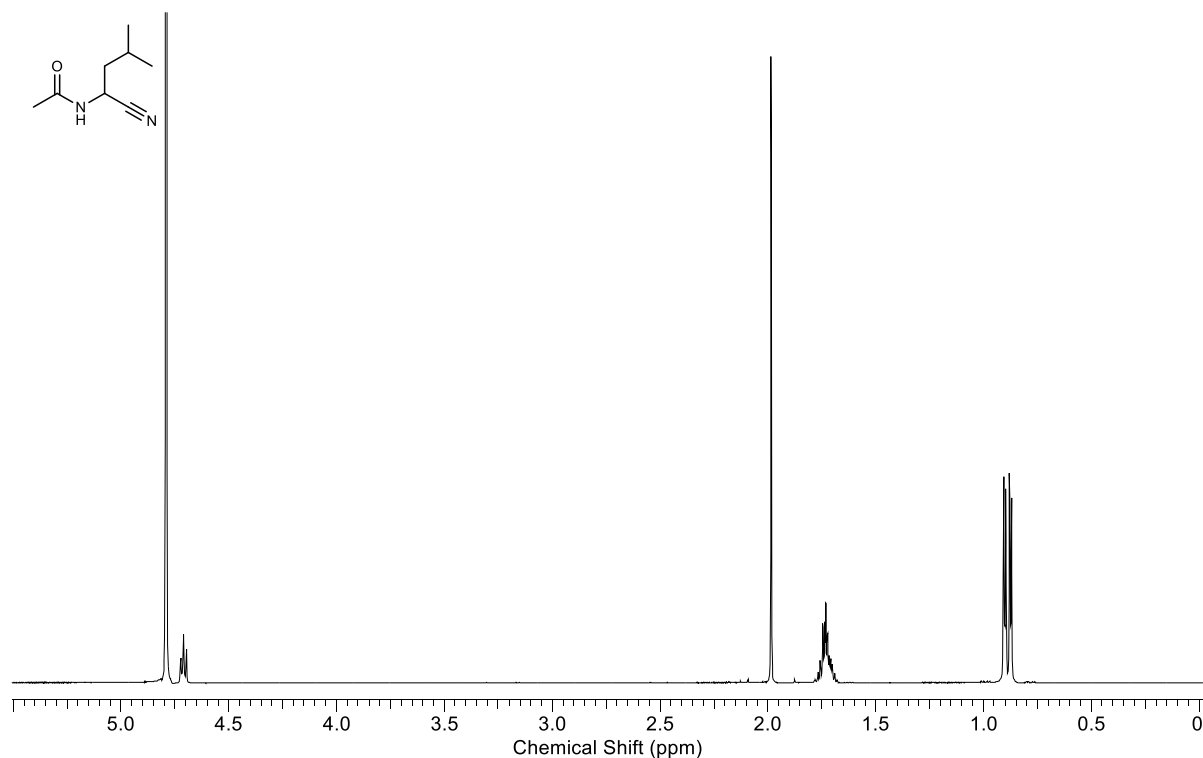


Supplementary Figure 24. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylalanine nitrile **Ac-Ala-CN**

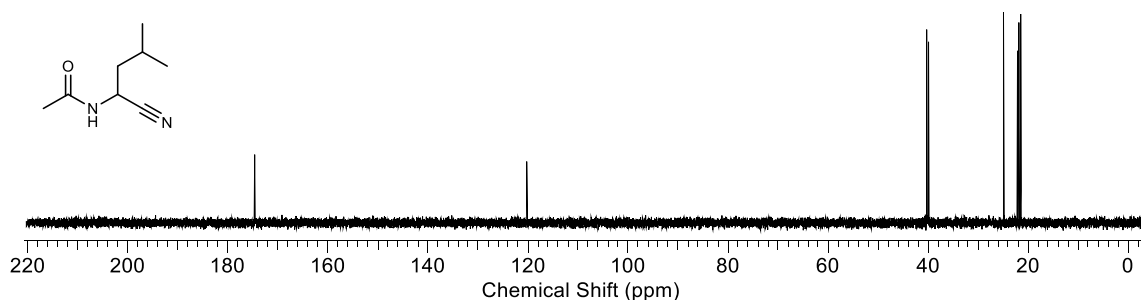
N-Acetylleucine nitrile **Ac-Leu-CN**



N-Acetylleucine nitrile **Ac-Leu-CN** was produced following *general procedure A* (99%). *General procedure B* using leucine nitrile hydrochloride **Leu-CN**·HCl³⁹ (1.49 g, 10 mmol) afforded *N*-acetylleucine nitrile **Ac-Leu-CN** as a white solid (1.51 g, 9.79 mmol, 98%). ¹H NMR (600 MHz, D₂O): δ 4.71 (t, *J* = 6.1 Hz, 1H, (C2)–H), 1.98 (s, 3H, (COCH₃)), 1.78 – 1.68 (m, 3H, (C3)–H₂, (C4)–H), 0.90 (d, *J* = 6.1, 3H, (C5)–H₃), 0.87 (d, *J* = 6.1, 3H, (C5')–H₃). ¹³C NMR (151 MHz, D₂O): δ 174.6 (COCH₃), 120.2 (C1), 40.3 (C2), 39.9 (C3), 24.9 (C4), 24.2 (COCH₃), 21.9 (C5), 21.5 (C5'). HRMS–CI [M+H]⁺: calcd. for [C₈H₁₄N₂O+H]⁺ 155.1179; observed 155.1176. IR (solid): 3285, 3046, 2965, 2938, 2876, 2156, 1657, 1531 cm^{–1}. M.p. 50.2–53.1 °C; lit. m.p. 48–49 °C.⁴¹

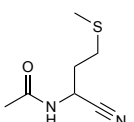


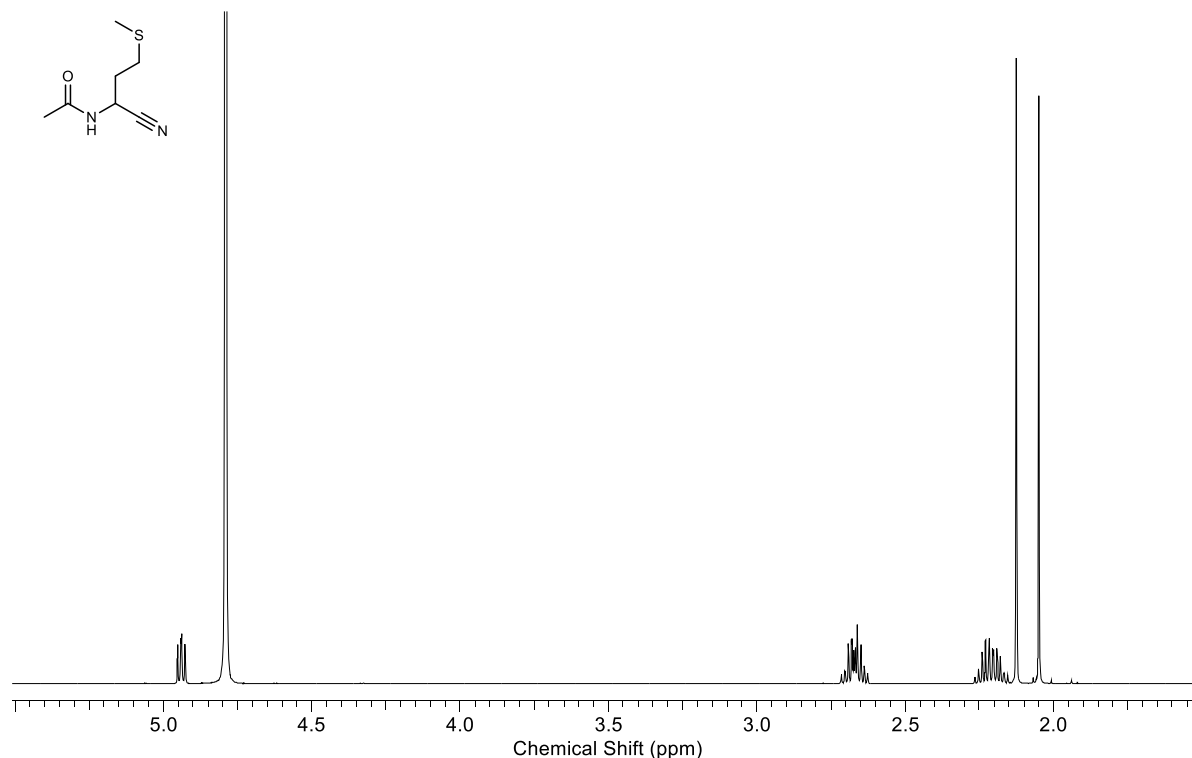
Supplementary Figure 25. ¹H NMR (600 MHz, D₂O, 0.00–5.50 ppm) spectrum to show *N*-acetylleucine nitrile **Ac-Leu-CN**



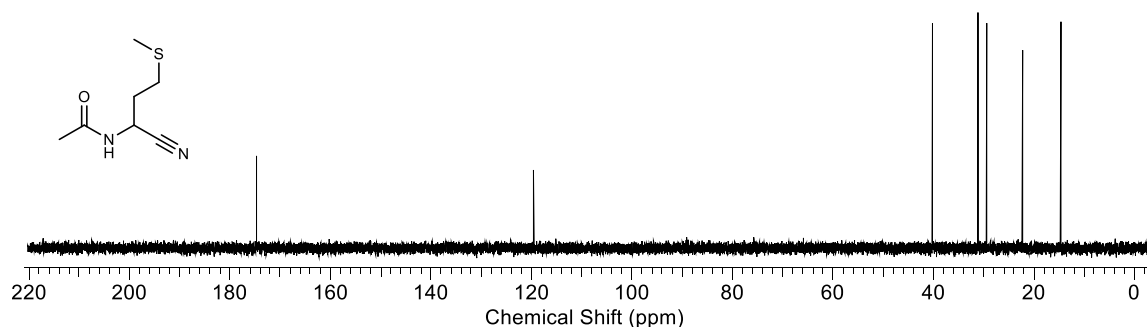
Supplementary Figure 26. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylleucine nitrile **Ac-Leu-CN**

N-Acetylmethionine nitrile **Ac-Met-CN**

 *N*-Acetylmethionine nitrile **Ac-Met-CN** was produced following *general procedure A* (99%). *General procedure B* using methionine nitrile hydrochloride **Met-CN**·HCl⁴² (1.83 g, 10 mmol) afforded *N*-acetylmethionine nitrile **Ac-Met-CN** as a white solid (1.84 g, 9.78 mmol, 98%). ¹H NMR (600 MHz, D₂O): δ 4.94 (dd, *J* = 6.8, 8.3 Hz, 1H, (C2)–H), 2.72 – 2.62 (m, 2H, (C3)–H₂), 2.27 – 2.15 (m, 2H, (C4)–H₂), 2.13 (s, 3H, SCH₃), 2.05 (s, 3H, (COCH₃)). ¹³C NMR (151 MHz, D₂O): δ 174.7 (COCH₃), 119.6 (C1), 40.2 (C2), 31.2 (C3), 29.4 (C4), 22.3 (COCH₃), 14.7 (SCH₃). HRMS–Cl [M+H]⁺: calcd. for [C₇H₁₂N₂OS+H]⁺: 173.0743; observed: 173.0745. IR (solid): 2966, 2769, 2186, 1736, 1590, 1569, 1503 cm^{–1}. M.p. 43.0–43.9 °C.

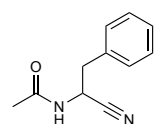


Supplementary Figure 27. ¹H NMR (600 MHz, D₂O, 1.50–5.50 ppm) spectrum to show *N*-acetylmethionine nitrile **Ac-Met-CN**

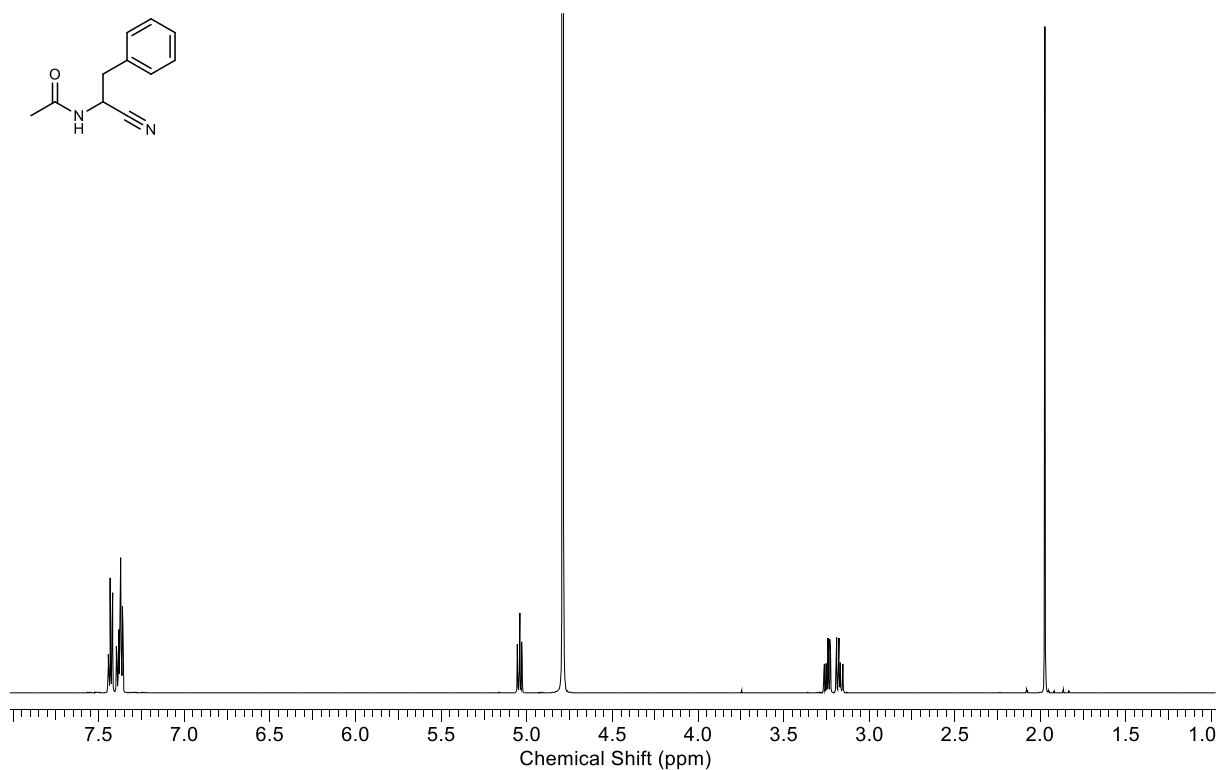


Supplementary Figure 28. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylmethionine nitrile **Ac-Met-CN**

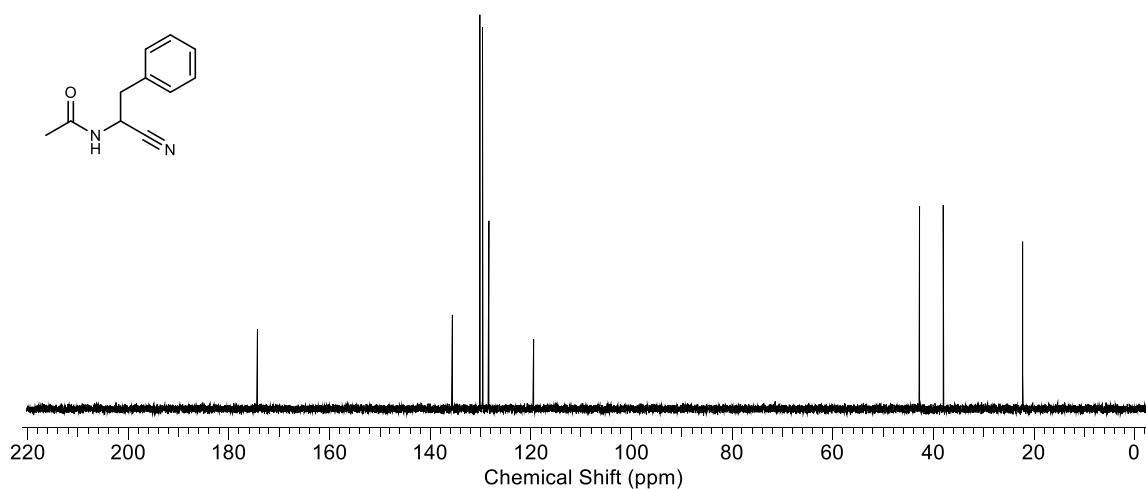
N-Acetylphenylalanine nitrile **Ac-Phe-CN**



N-Acetylphenylalanine nitrile **Ac-Phe-CN** was produced following *general procedure A* (99%). *General procedure B* using phenylalanine nitrile hydrochloride **Phe-CN**·HCl⁴¹ (1.83 g, 10 mmol) afforded *N*-acetylphenylalanine nitrile **Ac-Phe-CN** as a white solid (1.84 g, 9.78 mmol, 93%). ¹H NMR (600 MHz, D₂O): δ 7.45 – 7.35 (m, 5H, Ar), 5.04 (t, *J* = 7.4 Hz, 1H, (C2)–H), 3.25 (ABX, *J* = 7.4, 13.8 Hz, 1H, (C3)–H), 3.17 (ABX, *J* = 7.4, 13.8 Hz, 1H, (C3)–H'), 1.97 (s, 3H, (COCH₃)). ¹³C NMR (151 MHz, D₂O): δ 174.4 (COCH₃), 135.7 (Ar), 130.1 (Ar × 2), 129.6 (Ar × 2), 128.4 (Ar), 119.5 (C1), 42.8 (C2), 38.0 (C3), 22.2 (COCH₃). HRMS–ESI [M+H]⁺: calcd. for [C₁₁H₁₃N₂O+H]⁺: 189.1028; observed: 189.1032. IR (solid): 3282, 3063, 3029, 2966, 2930, 2239, 1654, 1602, 1583, 1537 cm⁻¹. M.p. 95.8–97.1 °C; lit. m.p. 94–95 °C.⁴³

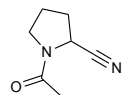


Supplementary Figure 29. ¹H NMR (600 MHz, D₂O, 1.00–8.00 ppm) spectrum to show *N*-acetylphenylalanine nitrile **Ac-Phe-CN**

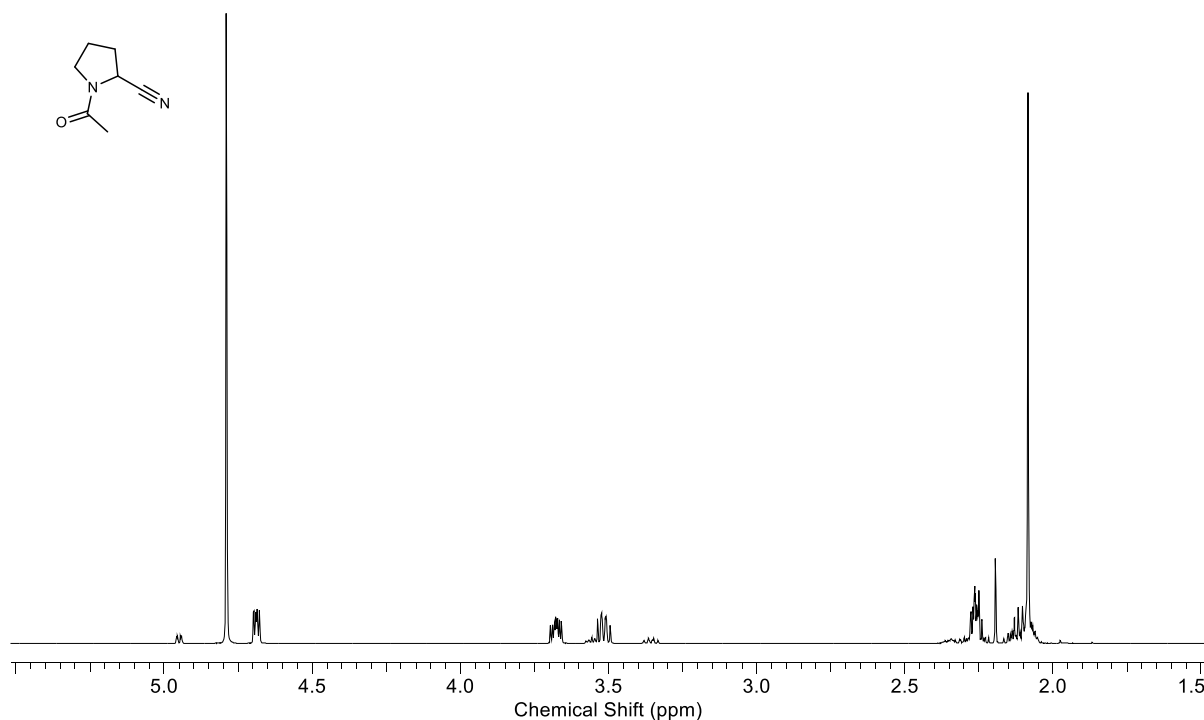


Supplementary Figure 30. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylphenylalanine nitrile **Ac-Phe-CN**

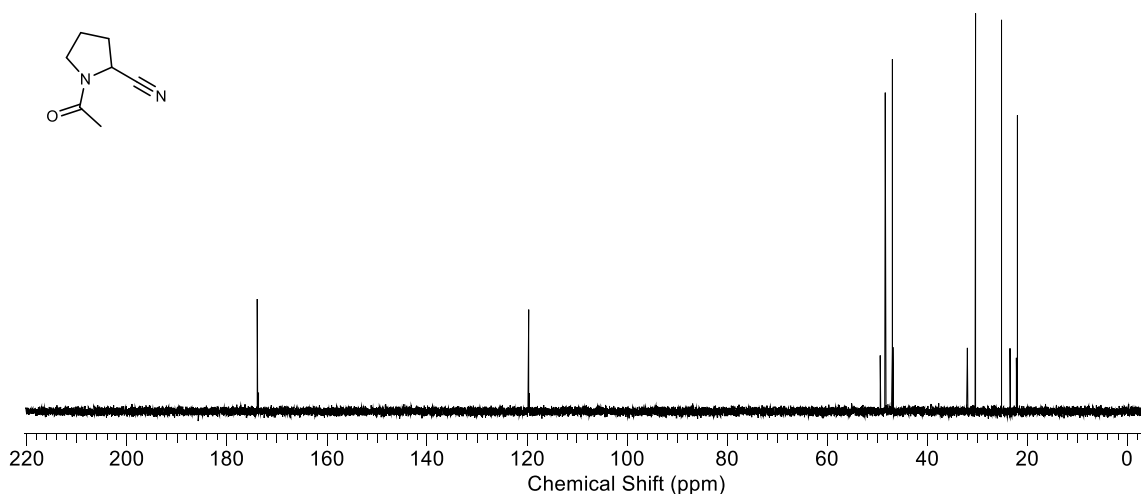
N-Acetylproline nitrile **Ac-Pro-CN**



N-Acetylproline nitrile **Ac-Pro-CN** was produced following *general procedure A* (99%). *General procedure B* using proline nitrile hydrochloride **Pro-CN**·HCl (265 mg, 2 mmol) afforded *N*-acetylproline nitrile **Ac-Pro-CN** as a clear oil (205 mg, 1.48 mmol, 74%). ^1H NMR (600 MHz, D_2O , 85:15 mixture of rotamers a/b): rotamer a: δ 4.69 (dd, $J = 4.3, 7.2$ Hz, 1H, (C2)–H), 3.70 – 3.66 (m, 1H, (C5)–H), 3.54 – 3.49 (m, 1H, (C5')–H), 2.28 – 2.23 (m, 2H, (C3)–H₂), 2.15 – 2.06 (m, 2H, (C4)–H₂), 2.08 (s, 3H, (COCH₃)); rotamer b: δ 4.94 (dd, $J = 2.1, 7.8$ Hz, 1H, (C2)–H), 3.58 – 3.54 (m, 1H, (C5)–H), 3.38 – 3.34 (m, 1H, (C5')–H), 2.36 – 2.28 (m, 2H, (C4)–H₂), 2.19 (s, 3H, (COCH₃), 2.15 – 2.06 (m, 2H, (C4)–H₂). ^{13}C NMR (151 MHz, D_2O , 85:15 mixture of rotamers a/b): rotamer a: δ 173.9 (COCH₃), 119.8 (C1), 48.4 (C2), 47.1 (C5), 30.5 (C3), 25.2 (C4), 22.0 (COCH₃); rotamer b: δ 173.8 (COCH₃), 119.6 (C1), 49.4 (C2), 46.7 (C5), 32.1 (C3), 23.6 (C4), 22.2 (COCH₃). HRMS–ESI $[\text{M}+\text{Na}]^+$: calcd. for $[\text{C}_7\text{H}_{10}\text{N}_2\text{O}+\text{Na}]^+$: 161.0685; observed: 161.0668. IR (solid): 3291, 3037, 2959, 2951, 2870, 2800, 2251 1671, 1610, 1551, 1501 cm^{-1}

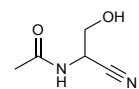


Supplementary Figure 31. ^1H NMR (600 MHz, D_2O , 1.50–5.50 ppm) spectrum to show *N*-acetylproline nitrile **Ac-Pro-CN**

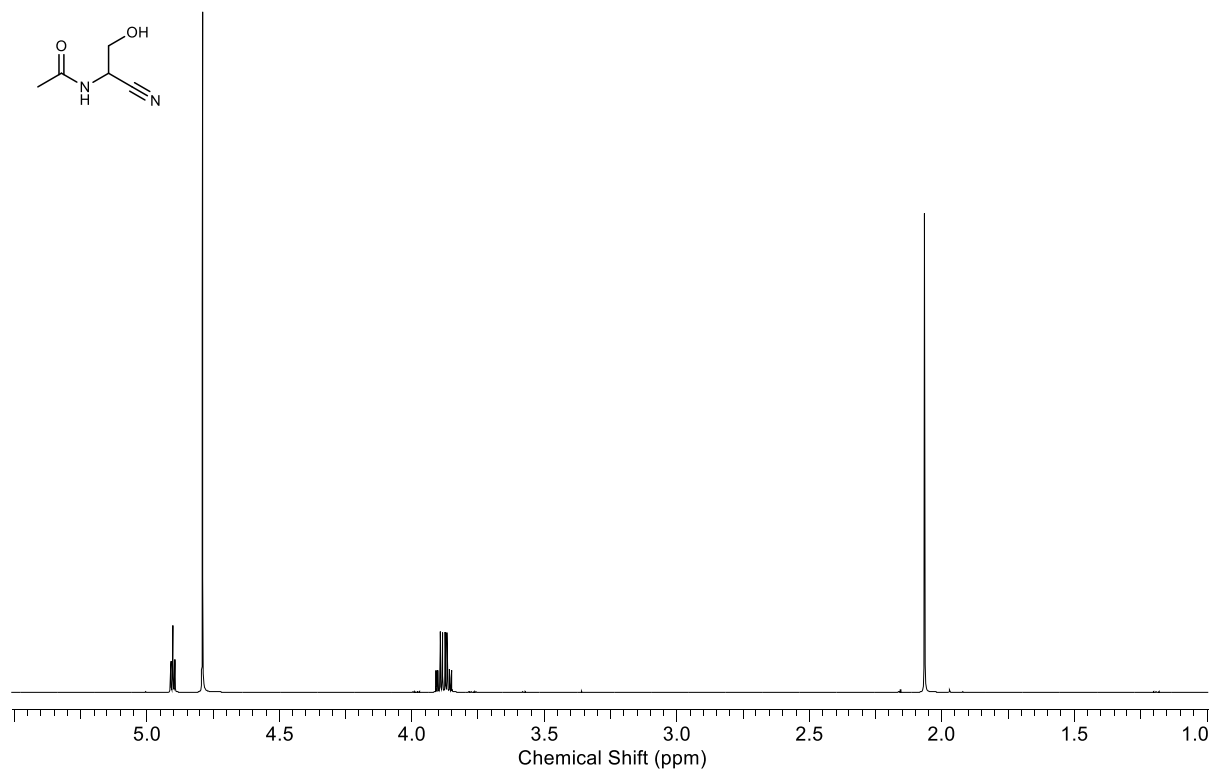


Supplementary Figure 32. ^{13}C NMR (151 MHz, D_2O , 0–220 ppm) spectrum to show *N*-acetylproline nitrile **Ac-Pro-CN**

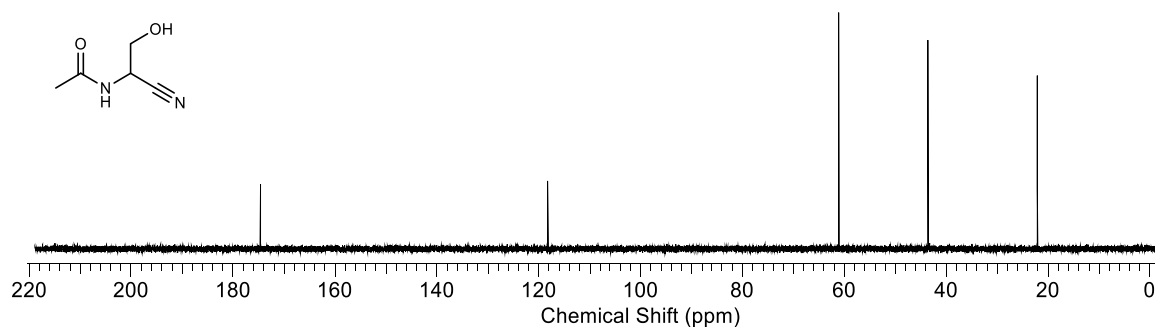
N-Acetylserine nitrile **Ac-Ser-CN**



N-Acetylserine nitrile **Ac-Ser-CN** was produced following *general procedure A* (97%). *General procedure B* using serine nitrile **Ser-CN**¹⁴ (258 mg, 3.00 mmol) afforded *N*-acetylserine nitrile **Ac-Ser-CN** as white, crystalline solid (363 mg, 2.83 mmol, 94%). ¹H NMR (700 MHz, D₂O): δ 4.90 (app. t, *J* = 5.3 Hz, 1H, (C2)–H), 3.90 (ABX, *J* = 5.3, 11.2 Hz, 1H, (C2)–H), 3.86 (ABX, *J* = 5.3, 11.2 Hz, 1H, (C2)–H'), 2.06 (s, 3H, (COCH₃)). ¹³C NMR (151 MHz, D₂O): δ 174.6 (COCH₃), 118.3 (C1), 61.1 (C3), 43.6 (C2), 22.1 (COCH₃). HRMS–Cl [M+H]⁺: calcd. for [C₅H₈N₂O₂+H]⁺: 129.0659; observed: 129.0658. IR (solid): 3456, 3262, 3016, 2970, 2970, 3016, 2970, 2939, 2241, 2125, 1738, 1620, 1527 cm^{−1}. M.p. 116.5–117.9 °C; lit. m.p. 111–112 °C.⁴⁴

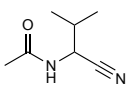


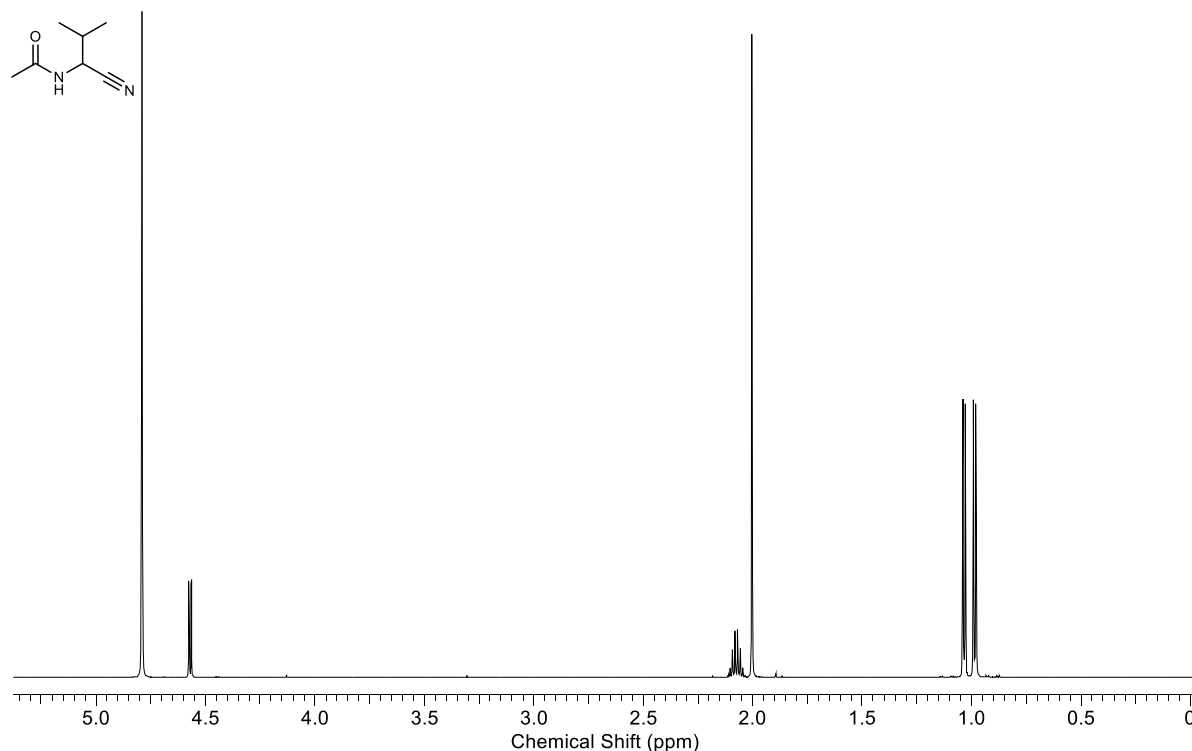
Supplementary Figure 33. ¹H NMR (700 MHz, D₂O, 1.00–5.50 ppm) spectrum to show *N*-acetylserine nitrile **Ac-Ser-CN**



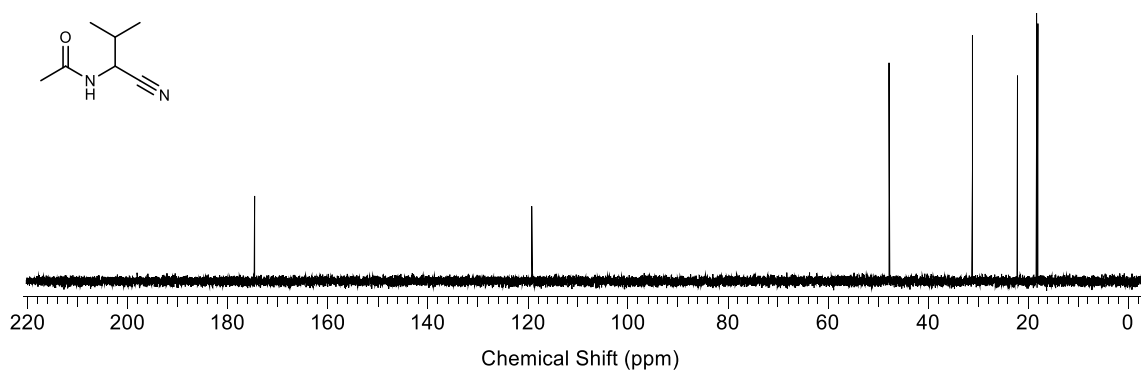
Supplementary Figure 34. ¹³C NMR (176 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylserine nitrile **Ac-Ser-CN**

N-Acetylvaline nitrile **Ac-Val-CN**

 *N*-Acetylvaline nitrile **Ac-Val-CN** was produced following *general procedure A* (99%). *General procedure B* using valine nitrile hydrochloride **Val-CN**·HCl³⁹ (1.35 g, 10 mmol) afforded *N*-acetylvaline nitrile **Ac-Val-CN** as a white solid (1.40 g, 9.97 mmol, 99%). ¹H NMR (600 MHz, D₂O): δ 4.57 (d, *J* = 7.0 Hz, 1H, (C2)–H), 2.12 – 2.03 (m, 1H, (C3)–H), 2.01 (s, 3H, (COCH₃)), 1.03 (d, *J* = 6.8 Hz, 3H, (C4)–H₃), 0.98 (d, *J* = 6.8 Hz, 3H, (C4')–H₃). ¹³C NMR (151 MHz, D₂O): δ 174.6 (COCH₃), 119.2 (C1), 47.8 (C2), 31.2 (C3), 22.2 (COCH₃), 18.4 (C4), 18.1 (C4'). HRMS–CI [M+H]⁺: calcd. for [C₇H₁₂N₂O+H]⁺: 141.1022; observed: 141.1022. IR (solid): 3454, 2970, 2849, 2636, 2575, 2133, 1736, 1583, 1555, 1514 cm^{−1}. M.p. 100.9–102.1 °C.

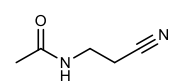


Supplementary Figure 35. ¹H NMR (600 MHz, D₂O, 0.00–5.85 ppm) spectrum to show *N*-acetylvaline nitrile **Ac-Val-CN**

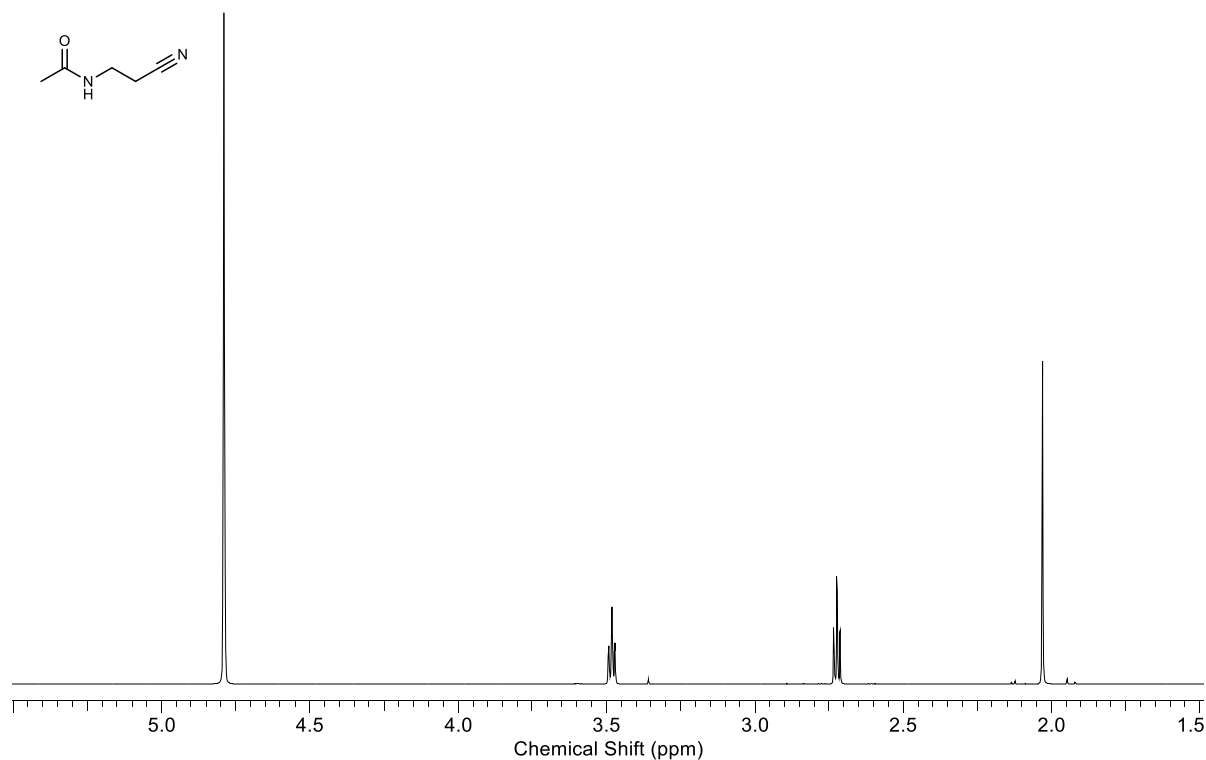


Supplementary Figure 36. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylvaline nitrile **Ac-Val-CN**

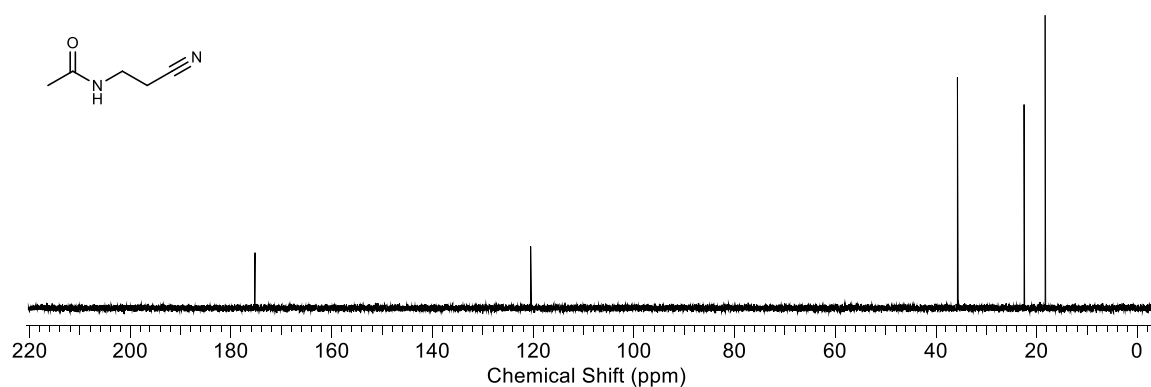
N-Acetyl- β -alanine nitrile **Ac- β -Ala-CN**



N-Acetyl- β -alanine nitrile **Ac- β -Ala-CN** was prepared following *general procedure B* using β -alanine nitrile **β -Ala-CN** (701 mg, 731 μ L, 10.0 mmol). The reaction afforded *N*-acetyl- β -alanine nitrile **Ac- β -Ala-CN** as a white solid (912 mg, 8.13 mmol, 83%). ^1H NMR (600 MHz, D_2O): δ 3.48 (t, J = 6.8 Hz, 2H, (C2)- H_2), 2.73 (t, J = 6.8 Hz, 2H, (C3)- H_2), 2.03 (s, 3H, (COCH₃)). ^{13}C NMR (151 MHz, D_2O): δ 175.3 (COCH₃), 120.5 (C1), 35.7 (C2), 22.5 (C3), 18.3 (COCH₃). HRMS-Cl $[\text{M}+\text{NH}_4]^+$: calcd. for $[\text{C}_5\text{H}_8\text{N}_2\text{O}+\text{NH}_4]^+$: 130.0975; observed: 130.0975. IR (solid): 3253, 3071, 2950, 2929, 2856, 2240, 1652, 1556 cm^{-1} . M.p. 60.7–61.5 $^\circ\text{C}$; lit. m.p. 62–63 $^\circ\text{C}$.⁴⁵



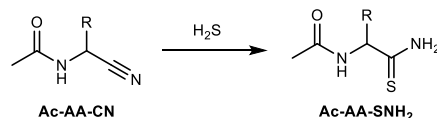
Supplementary Figure 37. ^1H NMR (600 MHz, D_2O , 1.50–5.50 ppm) spectrum to show *N*-acetyl- β -alanine nitrile **Ac- β -Ala-CN**



Supplementary Figure 38. ^{13}C NMR (151 MHz, D_2O , 0–220 ppm) spectrum to show *N*-acetyl- β -alanine nitrile **Ac- β -Ala-CN**

N-Acetylaminoacyl thioamide synthesis

General procedure C: Thiolyis of N-acetylamino nitriles



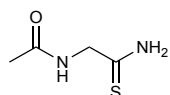
Prebiotic thiolyis of *N*-acetylamino nitriles **Ac-AA-CN** was carried out according to *general procedure C* unless stated otherwise. *N*-Acetylamino nitrile **Ac-AA-CN** (2.50 mmol) and NaSH·xH₂O (2.80 g, 25.0 mmol) was dissolved in degassed H₂O/D₂O (98:2, 50 mL, 50 mM). The solution was adjusted to pH 9.0 and stirred at room temperature for 24 h. NMR spectra were periodically acquired, until complete conversion of *N*-acetylamino nitrile **Ac-AA-CN** to *N*-acetylaminothioamide **Ac-AA-SNH₂** was observed. The solution was then sparged with argon for 15 min at pH 5.0 and concentrated *in vacuo*. The residue was purified by flash column chromatography (SiO₂; eluting with a gradient of MeOH/CH₂Cl₂ 0:100 → 10:90) to afford *N*-acetylaminothioamide **Ac-AA-SNH₂**. Yields are reported in Supplementary Table 2.

Ac-AA-CN	NMR yield (%)	Isolated yield (%)	Reaction time (h)
Gly	99	99	24
Ala	99	99	48
Arg ^[a]	99	-	36
Leu	99	32	48
Met	99	58	48
Phe	99	95	24
Pro	99	77	96
Ser ^[b]	99	94	24
Val	99	96	72

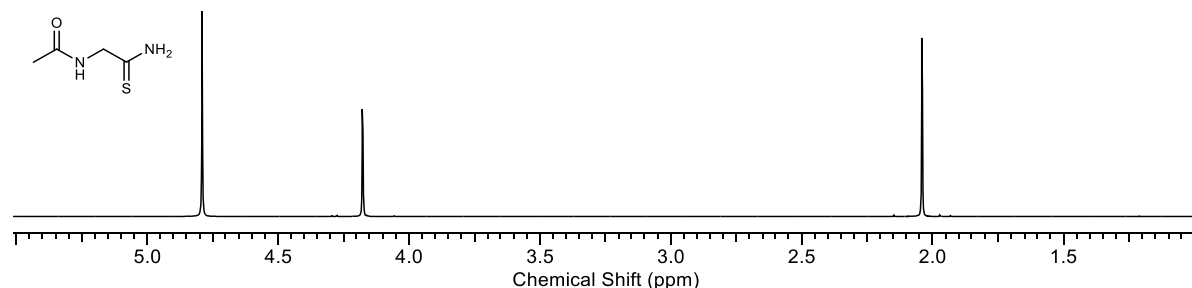
Supplementary Table 2. Formation of *N*-acetylaminothioamides **Ac-AA-SNH₂** by reaction of *N*-acetylamino nitriles **Ac-AA-CN** (50 mM) and sodium hydrosulfide (NaSH, 500 mM) at pH 9.0 and at room temperature (*general procedure C*) unless stated otherwise.

^[a] See Supplementary Page S58 for the NMR data of the thiolyis of **Ac-Arg-CN**. ^[b] See Supplementary Page S76 for the NMR data of the thiolyis of **Ac-Ser-CN**.

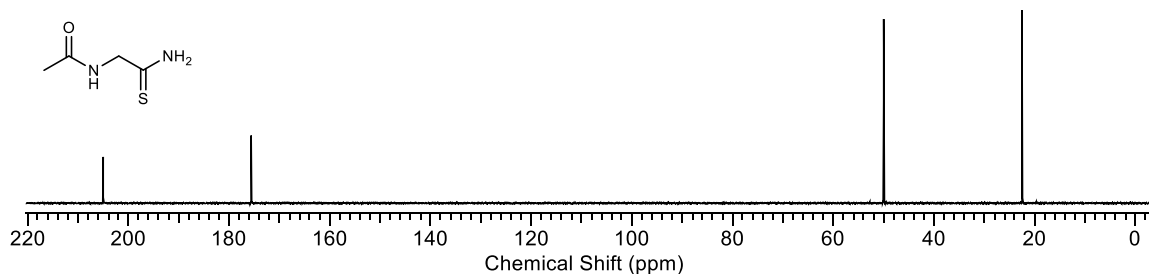
N-Acetylglycine thioamide **Ac-Gly-SNH₂**



N-Acetylglycine thioamide **Ac-Gly-SNH₂** was prepared following *general procedure C* using *N*-acetylglycine nitrile **Gly-CN** (245 mg, 2.5 mmol). The reaction afforded *N*-acetylglycine thioamide **Ac-Gly-SNH₂** as a white solid (329 mg, 2.49 mmol, 99%). ¹H NMR (600 MHz, D₂O): δ 4.18 (s, 2H, (C2)-H₂), 2.04 (s, 3H, (COCH₃)). ¹³C NMR (151 MHz, D₂O): δ 204.9 (C1), 175.6 (COCH₃), 50.0 (C2), 22.5 (COCH₃). HRMS-ESI [M-H]⁻ calcd. for [C₄H₇N₂OS-H]⁻ 131.0279; observed: 131.0278. IR (solid): 3341, 3287, 3077, 2931, 2143, 1640, 1522 cm⁻¹. M.p: 125.2–127.8 °C; lit. m.p. 124–126 °C.⁴⁶

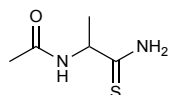


Supplementary Figure 39. ¹H NMR (600 MHz, D₂O, 1.00–5.50 ppm) spectrum to show *N*-acetylglycine thioamide **Ac-Gly-SNH₂**



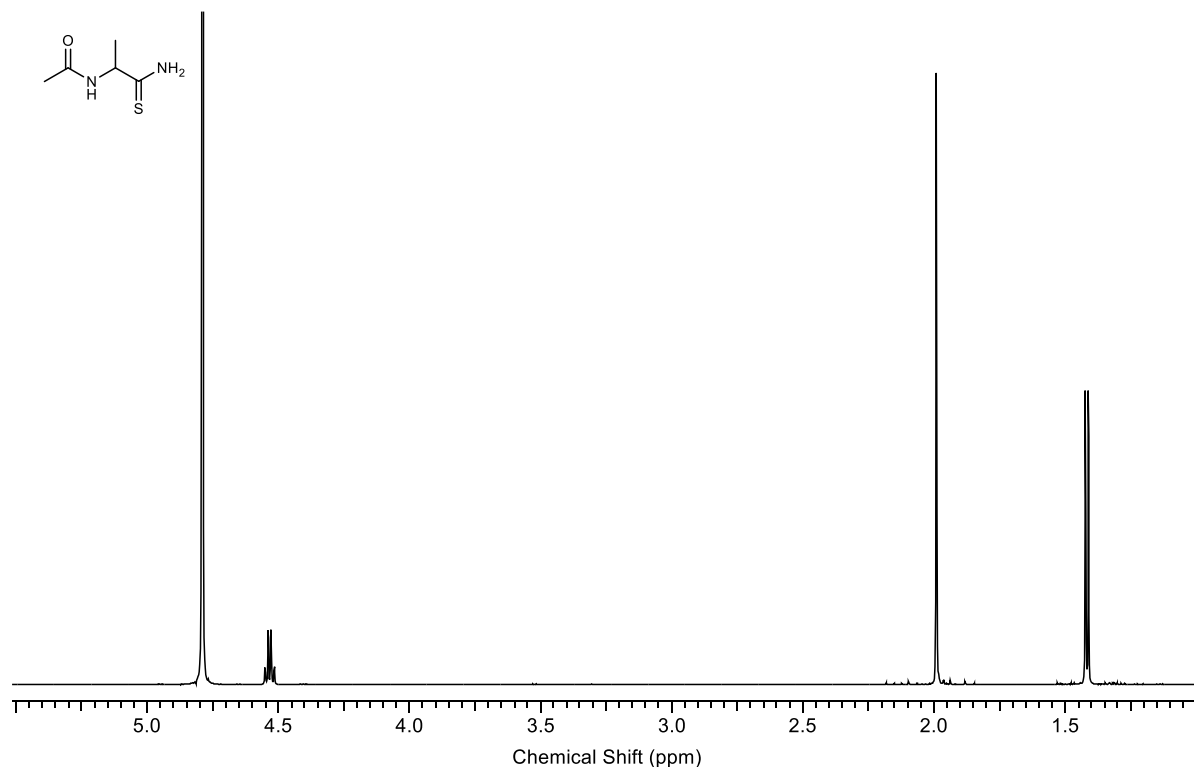
Supplementary Figure 40. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylglycine thioamide **Ac-Gly-SNH₂**

N-Acetylalanine thioamide **Ac-Ala-SNH₂**

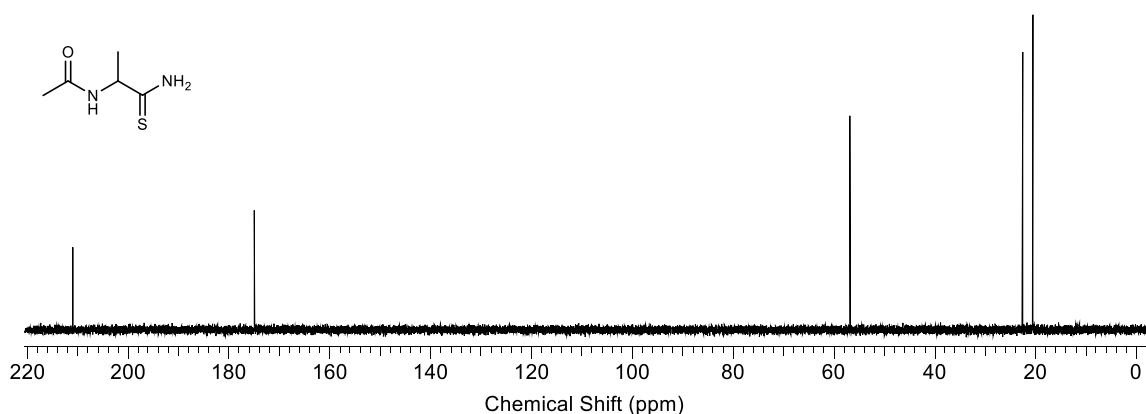


N-Acetylalanine thioamide **Ac-Ala-SNH₂** was prepared following *general procedure C* using *N*-acetylalanine nitrile **Ac-Ala-CN** (280.3 mg, 2.5 mmol). The reaction afforded *N*-acetylalanine thioamide

Ac-Ala-SNH₂ as a white solid (362.4 mg, 2.48 mmol, 99%). ¹H NMR (600 MHz, D₂O): δ 4.53 (q, *J* = 7.2 Hz, 1H, (C2)-H), 1.99 (s, 3H, (COCH₃), 1.41 (d, *J* = 7.2 Hz, 3H, (C3)). ¹³C NMR (151 MHz, D₂O): δ 210.7 (C1), 174.7 (COCH₃), 56.5 (C2), 22.3 (COCH₃), 20.3 (C3). HRMS-ESI [M+H]⁺ calcd. for [C₅H₁₀N₂OS+H]⁺: 147.0592; observed: 147.0594. IR (solid): 3292, 3233, 3097, 1623, 1541, 1524 cm⁻¹. M.p. 152.7–154.8 °C; lit. m.p. 152–154 °C.⁴⁷

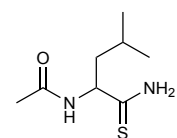


Supplementary Figure 41. ¹H NMR (600 MHz, D₂O, 1.00–5.50 ppm) spectrum to show *N*-acetylalanine thioamide **Ac-Ala-SNH₂**

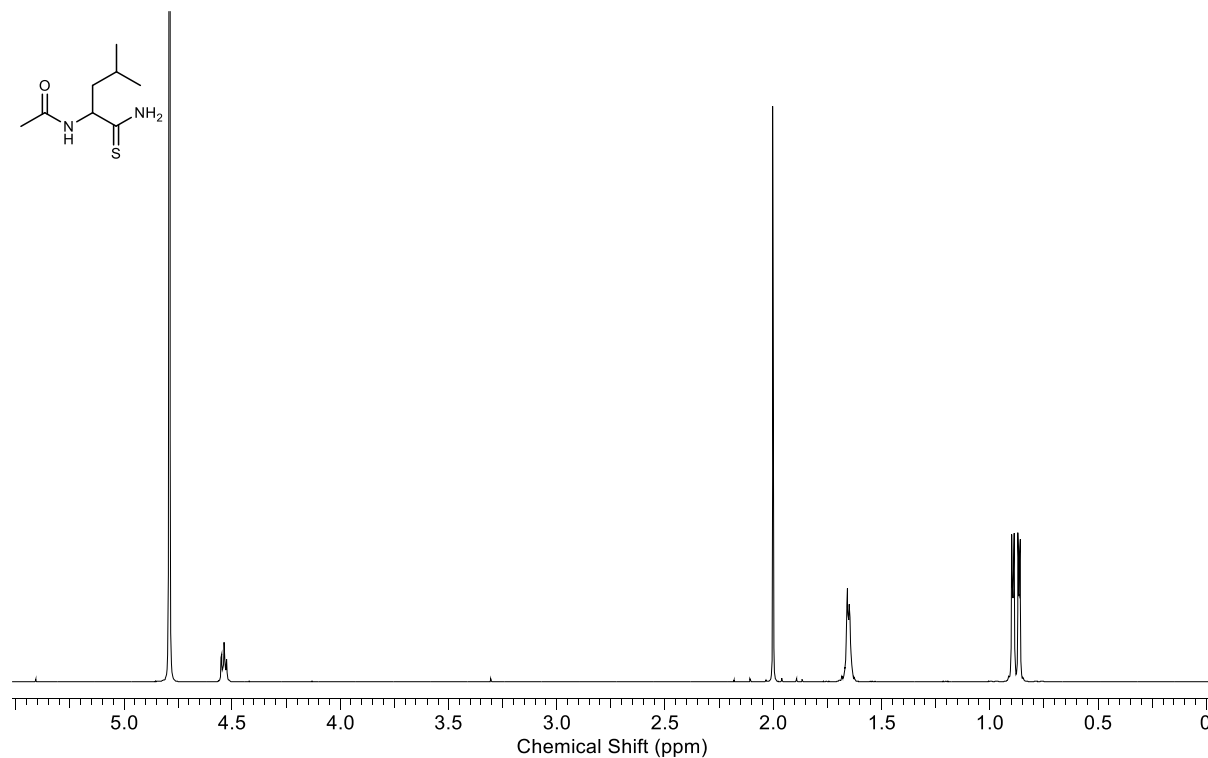


Supplementary Figure 42. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylalanine thioamide **Ac-Ala-SNH₂**

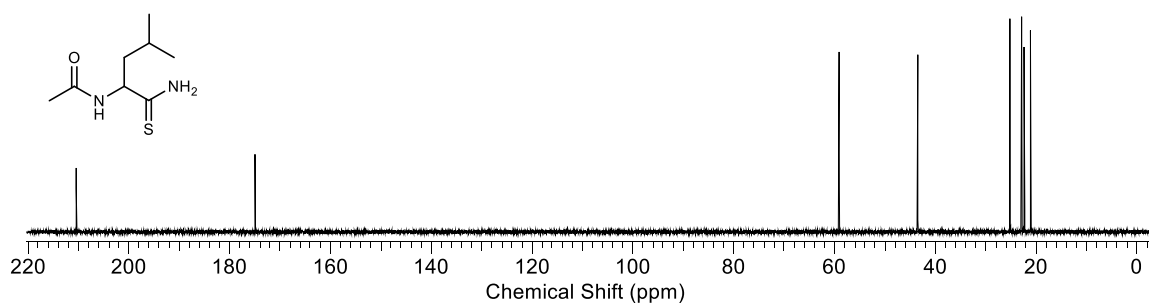
N-Acetylleucine thioamide **Ac-Leu-SNH₂**



N-Acetylleucine thioamide **Ac-Leu-SNH₂** was prepared following *general procedure C* using *N*-acetylleucine nitrile **Ac-Leu-CN** (385.5 mg, 2.5 mmol). The reaction afforded *N*-acetylleucine thioamide **Ac-Leu-SNH₂** as a white solid (149.5 mg, 0.79 mmol, 32%). ¹H NMR (600 MHz, D₂O): δ 4.54 (app. t, *J* = 7.1 Hz, 1H, (C2)–H), 2.00 (s, 3H, (COCH₃)), 1.68 – 1.63 (m, 2H, (C3)–H₂), 0.92 (d, *J* = 6.1 Hz, 3H, (C5)–H₃), 0.86 (d, *J* = 6.1 Hz, 3H, (C5')–H₃). ¹³C NMR (151 MHz, D₂O): δ 210.4 (C1), 174.9 (COCH₃), 59.1 (C2), 43.5 (C3), 25.2 (C4), 22.4 (C5), 22.4 (C5'), 21.1 (COCH₃). HRMS–EI [*M*]⁺⁺ calcd. for [C₈H₁₆N₂OS]⁺⁺: 188.0978; observed: 188.0978. IR (solid): 3255, 3086, 2956, 2930, 2915, 2869, 1652, 1532 cm^{–1}. M.p. 127.5–128.0 °C.

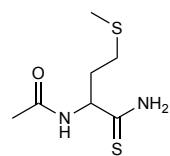


Supplementary Figure 43. ¹H NMR (600 MHz, D₂O, 0.0– 5.50 ppm) spectrum to show *N*-acetylleucine thioamide **Ac-Leu-SNH₂**

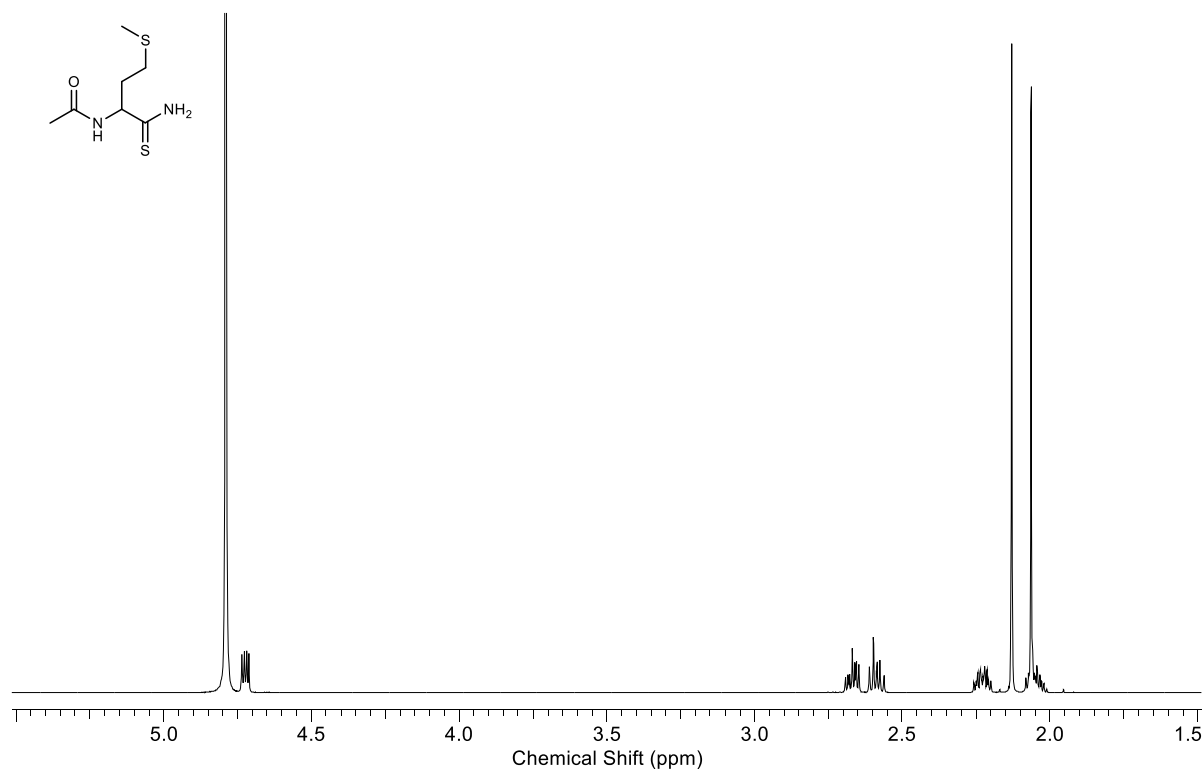


Supplementary Figure 44. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylleucine thioamide **Ac-Leu-SNH₂**

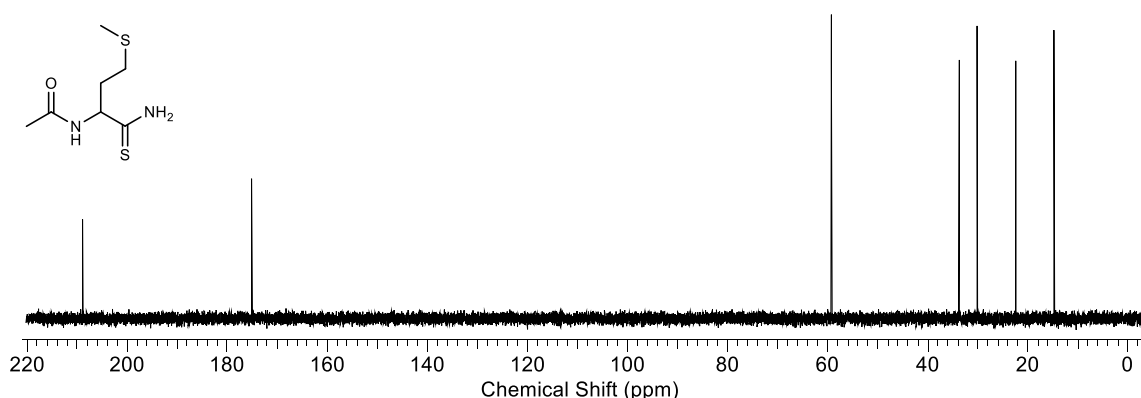
N-Acetylmethionine thioamide **Ac-Met-SNH₂**



N-Acetylmethionine thioamide **Ac-Met-SNH₂** was prepared following *general procedure C* using *N*-acetylmethionine nitrile **Ac-Met-CN** (430.6 mg, 2.5 mmol). The reaction afforded *N*-acetylmethionine thioamide **Ac-Met-SNH₂** as a white solid (298.1 mg, 1.44 mmol, 58%). ¹H NMR (600 MHz, D₂O): δ 4.72 (dd, *J* = 4.8, 9.4 Hz, 1H, (C2)–H), 2.70 – 2.64 (m, 1H, (C3)–H), 2.62 – 2.55 (m, 1H, (C3')–H), 2.26 – 2.19 (m, 1H, (C4)–H), 2.13 (s, 3H, SCH₃), 2.08 – 1.99 (obs. m, 1H, (C4')–H), 2.05 (s, 3H, (COCH₃)). ¹³C NMR (151 MHz, D₂O): δ 208.8 (C1), 175.0 (COCH₃), 52.2 (C1), 33.7 (C2), 30.1 (C4), 22.4 (COCH₃), 14.8 (SCH₃). HRMS–EI [*M*]⁺⁺ calcd. for [C₇H₁₄N₂OS₂]⁺⁺: 206.0542; observed: 206.0543. IR (solid): 3339, 3240, 3067, 2964, 2916, 1641, 1527 cm^{–1}. M.p. 103.1–105.2 °C.

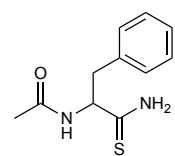


Supplementary Figure 45. ¹H NMR (600 MHz, D₂O, 1.50–5.50 ppm) spectrum to show *N*-acetylmethionine thioamide **Ac-Met-SNH₂**

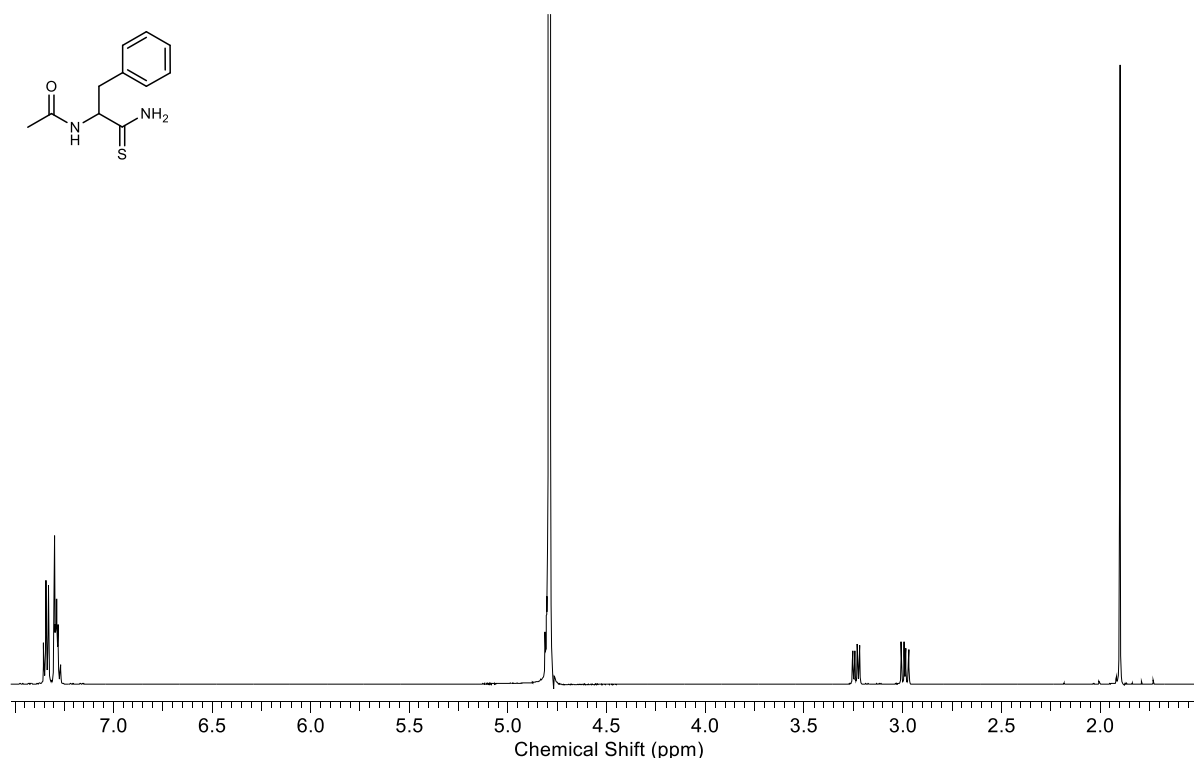


Supplementary Figure 46. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylmethionine thioamide **Ac-Met-SNH₂**

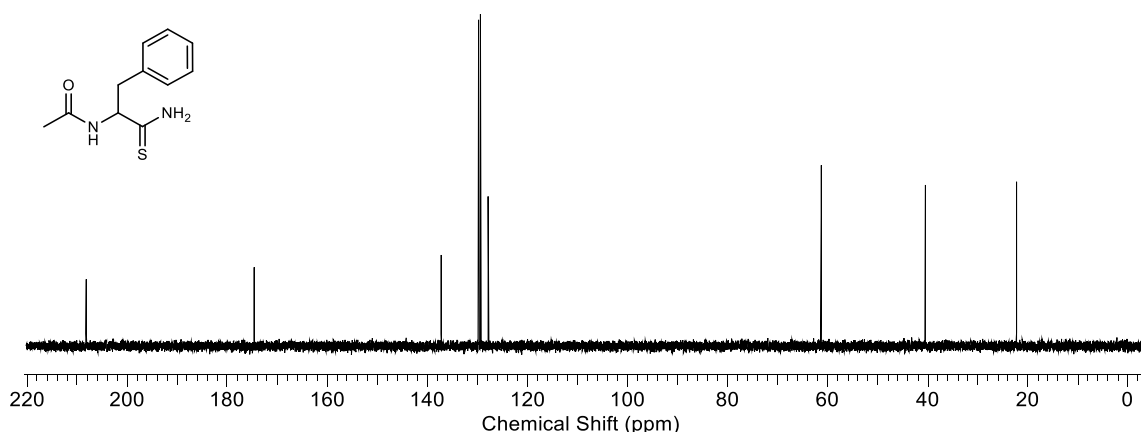
N-Acetylphenylalanine thioamide **Ac-Phe-SNH₂**



N-Acetylphenylalanine thioamide **Ac-Phe-SNH₂** was prepared following *general procedure C* using *N*-acetylphenylalanine nitrile **Ac-Phe-CN** (470.6 mg, 2.5 mmol). The reaction afforded *N*-acetylphenylalanine thioamide **Ac-Phe-SNH₂** as a white solid (529.3 mg, 2.38 mmol, 95%). ¹H NMR (600 MHz, D₂O) δ 7.36 – 7.25 (m, 5H, Ar), 4.80 (obs., 1H, (C2)–H), 3.23 (ABX, *J* = 6.1, 13.9 Hz, 1H, (C3)–H), 2.99 (ABX, *J* = 9.0, 13.9 Hz, 1H, (C3')–H), 1.90 (s, (COCH₃)). ¹³C NMR (151 MHz, D₂O): δ 208.2 (C1), 174.6 (COCH₃), 137.2 (Ar), 129.8 (Ar × 2), 129.3 (Ar × 2), 127.8 (Ar), 61.3 (C2), 40.5 (C3), 22.3 (COCH₃). HRMS–ESI [*M*+H]⁺ calcd. for [C₁₁H₁₄N₂OS]⁺: 222.0821; observed: 222.0822. IR (solid): 3362, 3293, 3136, 3025, 2954, 1653, 1628, 1519, 1500 cm⁻¹. M.p. 129.9–130.2 °C.

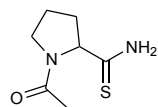


Supplementary Figure 47: ¹H NMR (600 MHz, D₂O, 1.50–7.50 ppm) spectrum to show *N*-acetylphenylalanine thioamide **Ac-Phe-SNH₂**

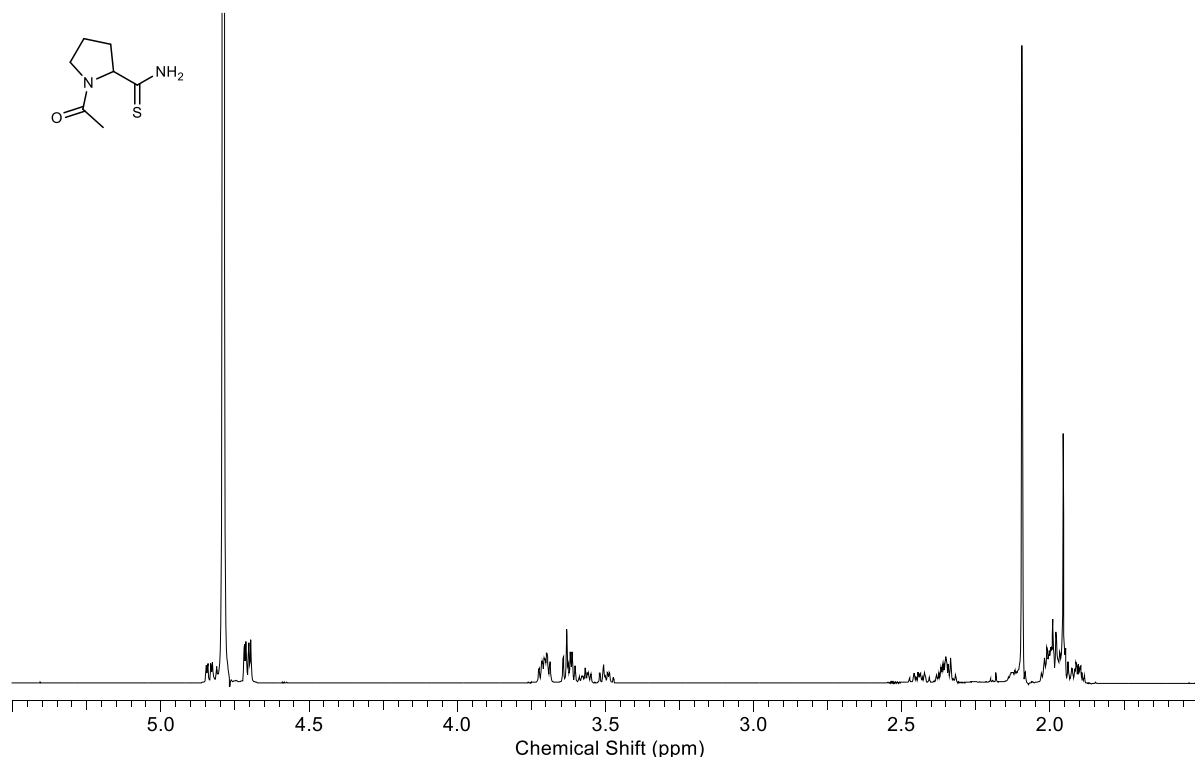


Supplementary Figure 48: ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylphenylalanine thioamide **Ac-Phe-SNH₂**

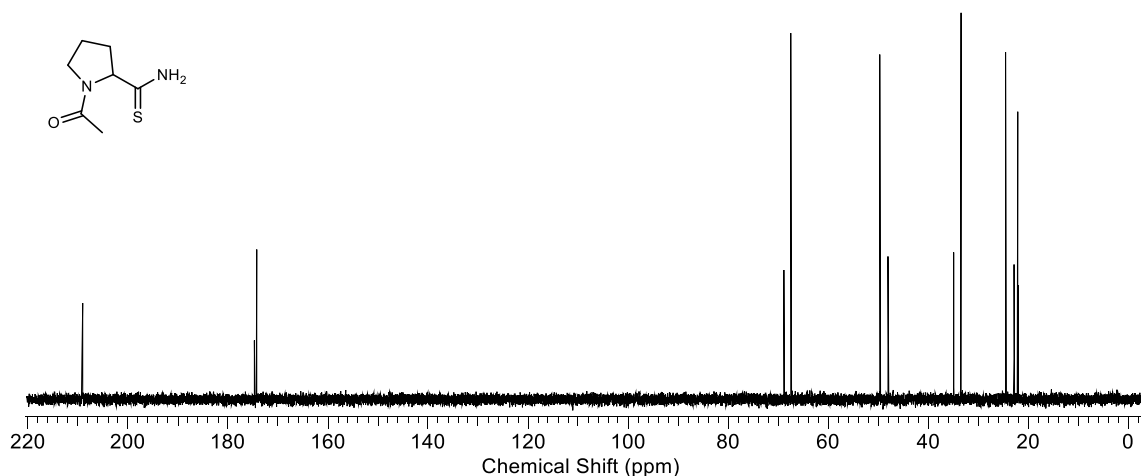
1-Acetylproline thioamide *Ac-Pro-SNH₂*



1-Acetylproline thioamide **Ac-Pro-SNH₂** was prepared following *general procedure C* using 1-acetylproline nitrile **Ac-Pro-CN** (345.4 mg, 2.5 mmol). The reaction afforded 1-acetylproline thioamide **Ac-Pro-SNH₂** as a white solid (333.7 mg, 1.93 mmol, 77%). ¹H NMR (600 MHz, D₂O, 65:35 mixture of rotamers a/b): rotamer a: δ 4.71 (dd, *J* = 3.6, 9.0 Hz, 1H, (C2)–H), 3.73 – 3.68 (m, 1H, (C5)–H), 3.65 – 3.60 (m, 1H, (C5')–H), 2.39 – 2.31 (m, 1H, (C4)–H), 2.09 (s, 3H, (COCH₃)), 2.03 – 1.97 (m, 2H, (C4')–H, (C3)–H), 1.94 – 1.88 (m, 1H, (C3')–H); rotamer b: δ 4.84 (dd, *J* = 3.1, 9.0 Hz, 1H, (C2)–H), 3.59 – 3.54 (m, 1H, (C5)–H), 3.53 – 3.46 (m, 1H, (C5')–H), 2.48 – 2.40 (m, 1H, (C4)–H), 2.15 – 2.08 (obs. m, 2H, (C4')–H, (C3)–H), 1.95 (s, 3H, (COCH₃)), 1.98 – 1.93 (obs. m, 1H, (C3')–H). ¹³C NMR (151 MHz, D₂O, mixture of rotamers a/b): rotamer a: δ 209.1 (C1), 174.3 (COCH₃), 67.5 (C2), 49.7 (C5), 33.5 (C3), 24.5 (C4), 22.2 (COCH₃); rotamer b: δ 209.2 (C1), 174.7 (COCH₃), 68.9 (C2), 48.0 (C3), 35.0 (C5), 22.9 (C4), 22.1 (COCH₃). HRMS–EI [*M*]⁺ calcd. for [C₇H₁₂N₂OS]⁺: 172.0665; observed: 172.0664. IR (solid): 3381, 3297, 3079, 2974, 2951, 2870, 1645, 1619 cm^{–1}. M.p. 159.0–161.2 °C.

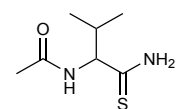


Supplementary Figure 49. ¹H NMR (600 MHz, D₂O, 1.50–5.50 ppm) spectrum to show 1-acetylproline thioamide **Ac-Pro-SNH₂**



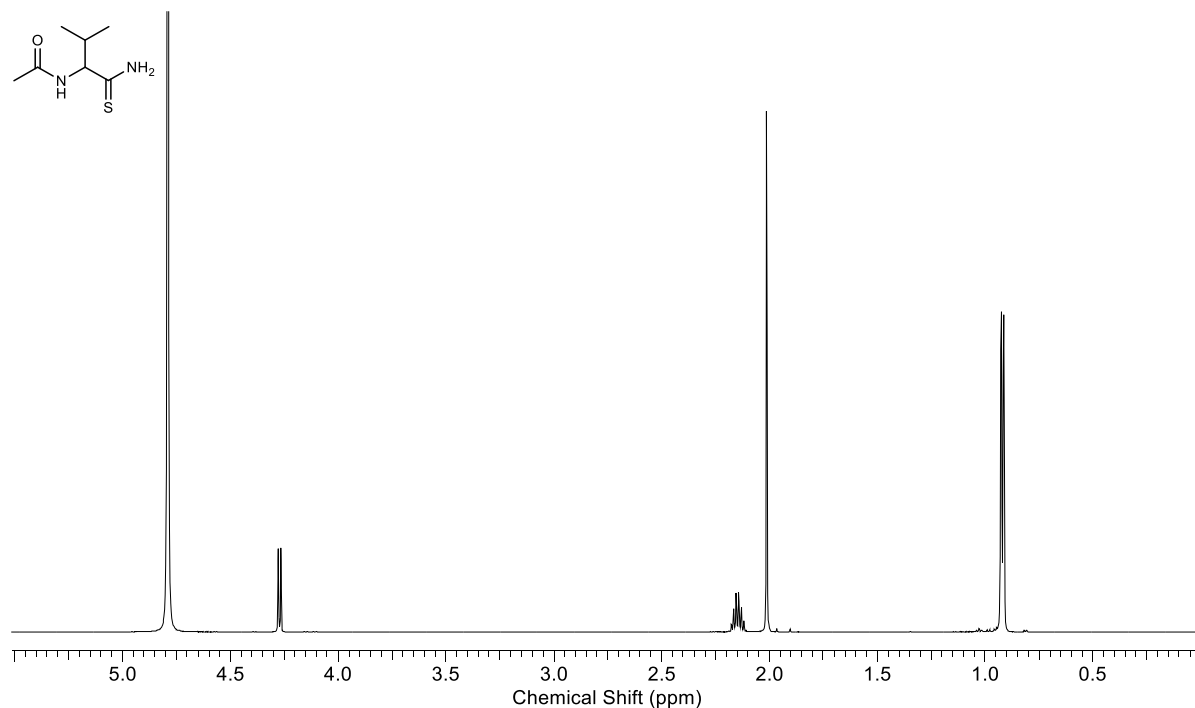
Supplementary Figure 50. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show 1-acetylproline thioamide **Ac-Pro-SNH₂**

N-Acetylvaline thioamide **Ac-Val-SNH₂**

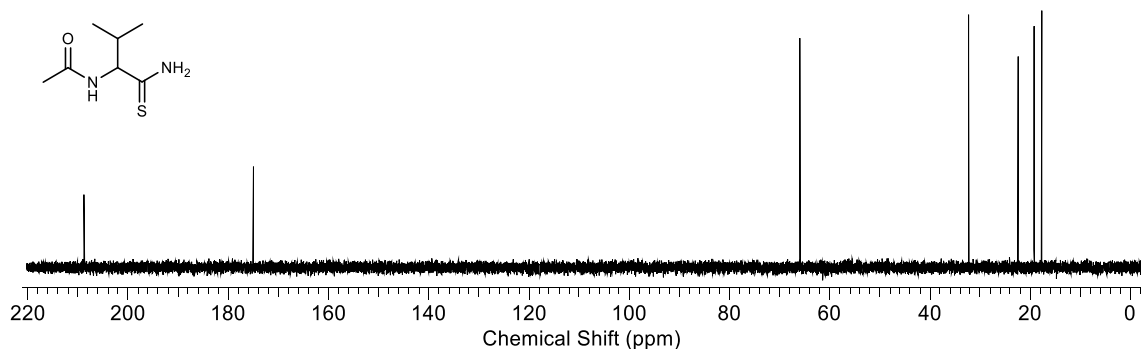


N-Acetylvaline thioamide **Ac-Val-SNH₂** was prepared following *general procedure C* using *N*-acetylvaline nitrile **Ac-Val-CN** (350.5 mg, 2.5 mmol). The reaction afforded *N*-acetylvaline thioamide

Ac-Val-SNH₂ as a white solid (419.2 mg, 2.41 mmol, 96%). ¹H NMR (600 MHz, D₂O): δ 4.27 (d, *J* = 7.5 Hz, 1H, (C2)–H), 2.19 – 2.11 (m, 1H, (C3)–H), 2.01 (s, 3H, (COCH₃), 0.92 (d, *J* = 6.8 Hz, 6H, (C4)–H₃, (C4')–H₃). ¹³C NMR (151 MHz, D₂O): δ 208.7 (C1), 175.0 (COCH₃), 65.9 (C2), 32.3 (C3), 22.4 (COCH₃), 19.2 (C4) 17.7 (C4'). HRMS-ESI [M+H]⁺ calcd. for [C₇H₁₄N₂OS]⁺: 174.0821; observed: 174.0822. IR (solid): 3269, 3091, 2966, 2869, 1656, 1535 cm^{−1}. M.p. 161.9–162.3 °C.



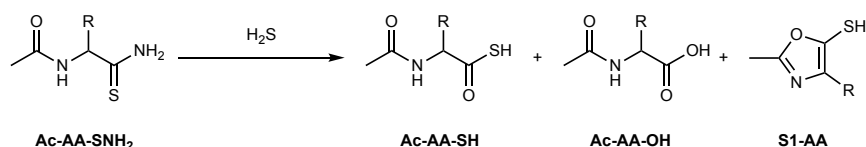
Supplementary Figure 51. ¹H NMR (600 MHz, D₂O, 0.00–5.50 ppm) spectrum to show *N*-acetylvaline thioamide **Ac-Val-SNH₂**



Supplementary Figure 52. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylvaline thioamide **Ac-Val-SNH₂**

N-Acetylaminoacyl thioacid synthesis

General procedure D: Hydrolysis of α -amidothioamides



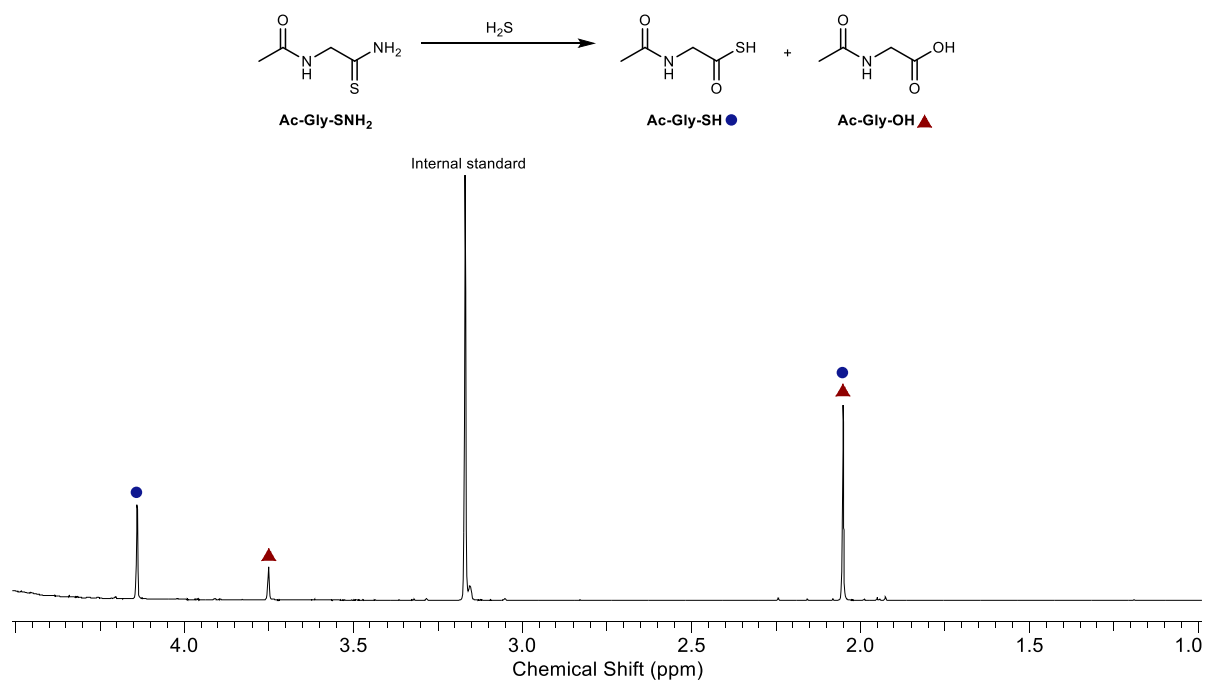
Hydrolyses of α -amidothioamides **Ac-AA-SNH₂** was carried out according to *general procedure D*, unless stated otherwise. α -Amidothioamide **Ac-AA-SNH₂** (0.05 mmol), $\text{NaSH} \cdot x\text{H}_2\text{O}$ (56.1 mg, 0.50 mmol) and methylsulfonylmethane (0.05 mmol) were dissolved in degassed $\text{H}_2\text{O}/\text{D}_2\text{O}$ (98:2, 1 mL), and the solution was adjusted to pH 9.0 with NaOH/HCl . The solution was incubated at 60 °C, whilst maintaining the solution at pH 9.0 with NaOH/HCl , and NMR spectra were periodically acquired until complete consumption of α -amidothioamide **Ac-AA-SNH₂** was observed. The presence of α -amidothioacid **Ac-AA-SH** and *N*-acetylamino acid **Ac-AA-OH** were confirmed by ^1H – ^{13}C HMBC NMR analysis, spiking or comparison of NMR data with pure synthetic standards. The reaction mixture was quantified using methylsulfonylmethane (MSM) an internal standard. A second by product of α -amidothioamide **Ac-AA-SNH₂** hydrolysis was observed in varying quantities, depending on the α -amidothioamide side-chain residue. This was tentatively assigned as the aromatic species **S1-AA** based on NMR data. Yields are reported in Supplementary Table 3.

Ac-AA-SNH₂	Time (h)	Ac-AA-SH (%)	Ac-AA-OH (%)	S1-AA (%)
Gly	24	81	19	0
Ala	72	85	8	6
Arg ^[a]	96	51	18	24
Leu	72	77	1	21
Met	72	70	6	22
Phe	48	84	0	10
Pro	168	72	26	0
Ser ^[b]	24	61	16	0
Val	72	8	9	82

Supplementary Table 3. Yields for the formation of α -amidothioacids **Ac-AA-SH** from the hydrolysis of α -amidothioamides **Ac-AA-SNH₂** (50 mM) in the presence of NaSH (500 mM) at pH 9.0 and 60 °C. ^[a] See page Supplementary Page S58 for the NMR data of the hydrolysis of **Ac-Arg-CN**. ^[b] See page Supplementary Page S76 for the NMR data of the hydrolysis of **Ac-Ser-CN**.

2-Acetamidoethanethioic S-acid **Ac-Gly-SH**

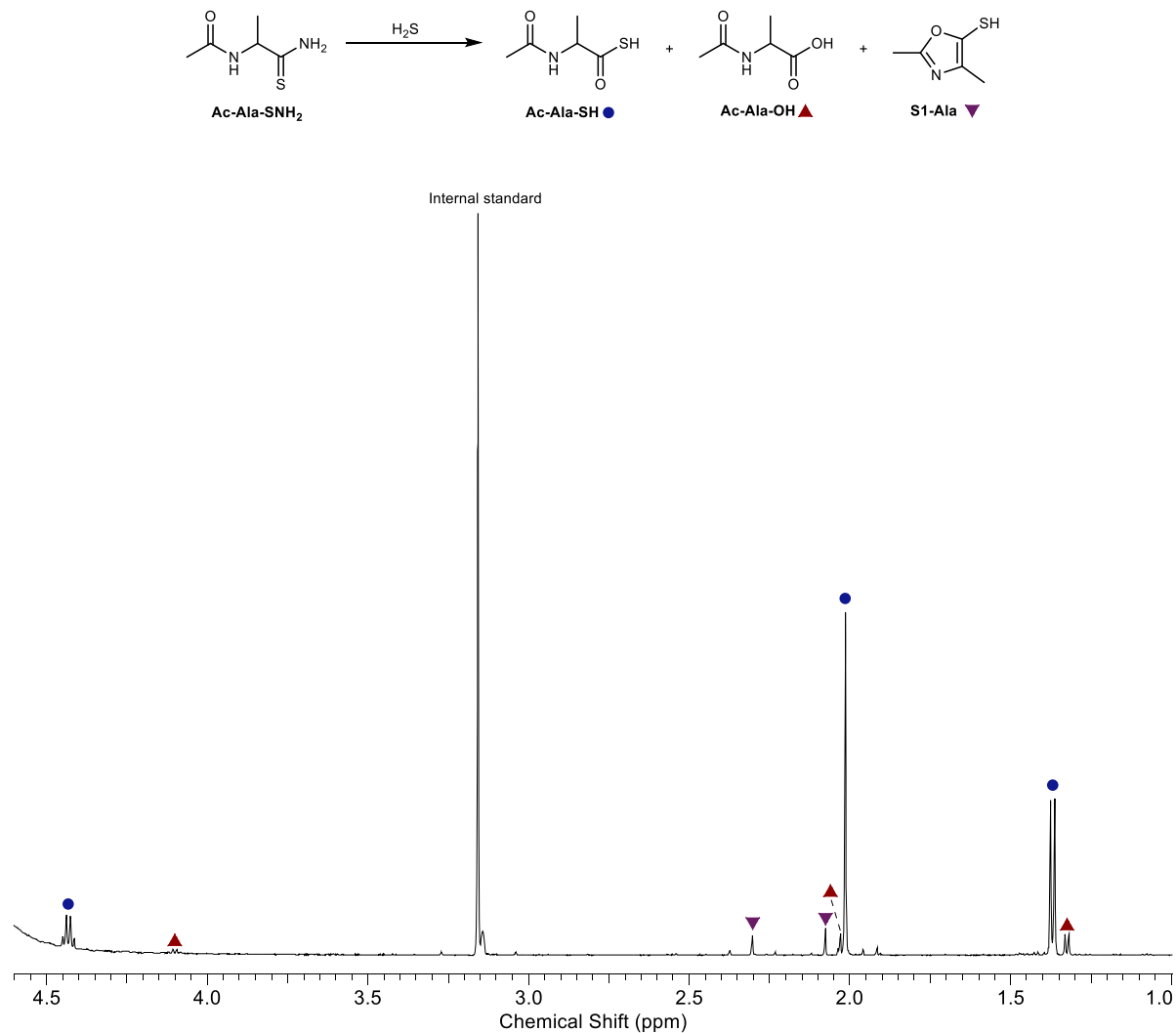
2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (81%, 24 h) was produced following *general procedure D* using *N*-acetylglycine thioamide **Ac-Gly-SNH₂** (6.61 mg, 0.05 mmol). ¹H NMR (600 MHz, H₂O/D₂O 98:2): 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (•): δ 4.14 (s, 2H, (C2)–CH₂), 2.05 (s, 3H, (COCH₃)). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.75 (s, 2H, (C2)–CH₂), 2.05 (s, 3H, (COCH₃)).



Supplementary Figure 53. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 1.0–4.5 ppm) spectrum to show the products of the reaction of *N*-acetylglycine thioamide (**Ac-Gly-SNH₂**; 50 mM) with NaSH (500 mM) at 60 °C and pH 9.0 after 24 h.

2-Acetamidopropanethioic S-acid **Ac-Ala-SH**

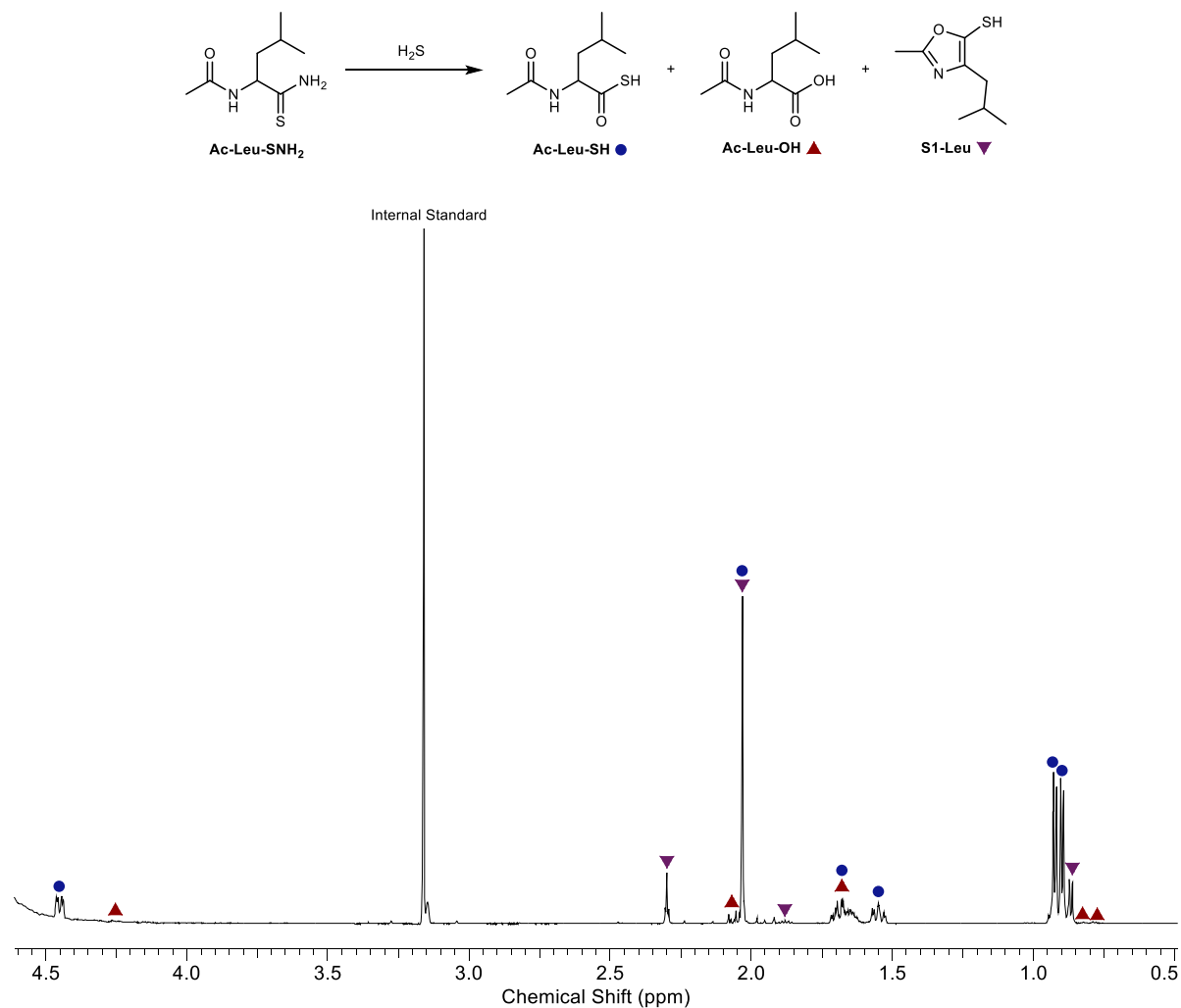
2-Acetamidopropanethioic S-acid **Ac-Ala-SH** (85%, 72 h) was produced following *general procedure D* using *N*-acetylalanine thioamide **Ac-Ala-SNH₂** (7.31 mg, 0.05 mmol). ¹H NMR (600 MHz, H₂O/D₂O 98:2): **2-Acetamidopropanethioic S-acid Ac-Ala-SH** (●): δ 4.43 (q, *J* = 7.2 Hz, 1H, (C2)–H), 2.02 (s, 3H, (COCH₃), 1.37 (d, *J* = 7.2 Hz, 3H, (C3)–H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 4.11 (q, *J* = 7.3 Hz, 1H, (C2)–H), 2.08 (s, 3H, (COCH₃), 1.32 (d, *J* = 7.3 Hz, 3H, (C3)–H₃). **S1-Ala** (▼): δ 2.31 (s, 3H, (COCH₃), 2.07 (s, (C3)–H₃).



Supplementary Figure 54. ¹H NMR (600 MHz, D₂O/H₂O 98:2, 1.0–4.6 ppm) spectrum to show the products of the reaction of *N*-acetylalanine thioamide (**Ac-Ala-SNH₂**; 50 mM) with NaSH (500 mM) at 60 °C and pH 9.0 after 72 h.

2-Acetamido-4-methylpentanethioic S-acid *Ac-Leu-SH*

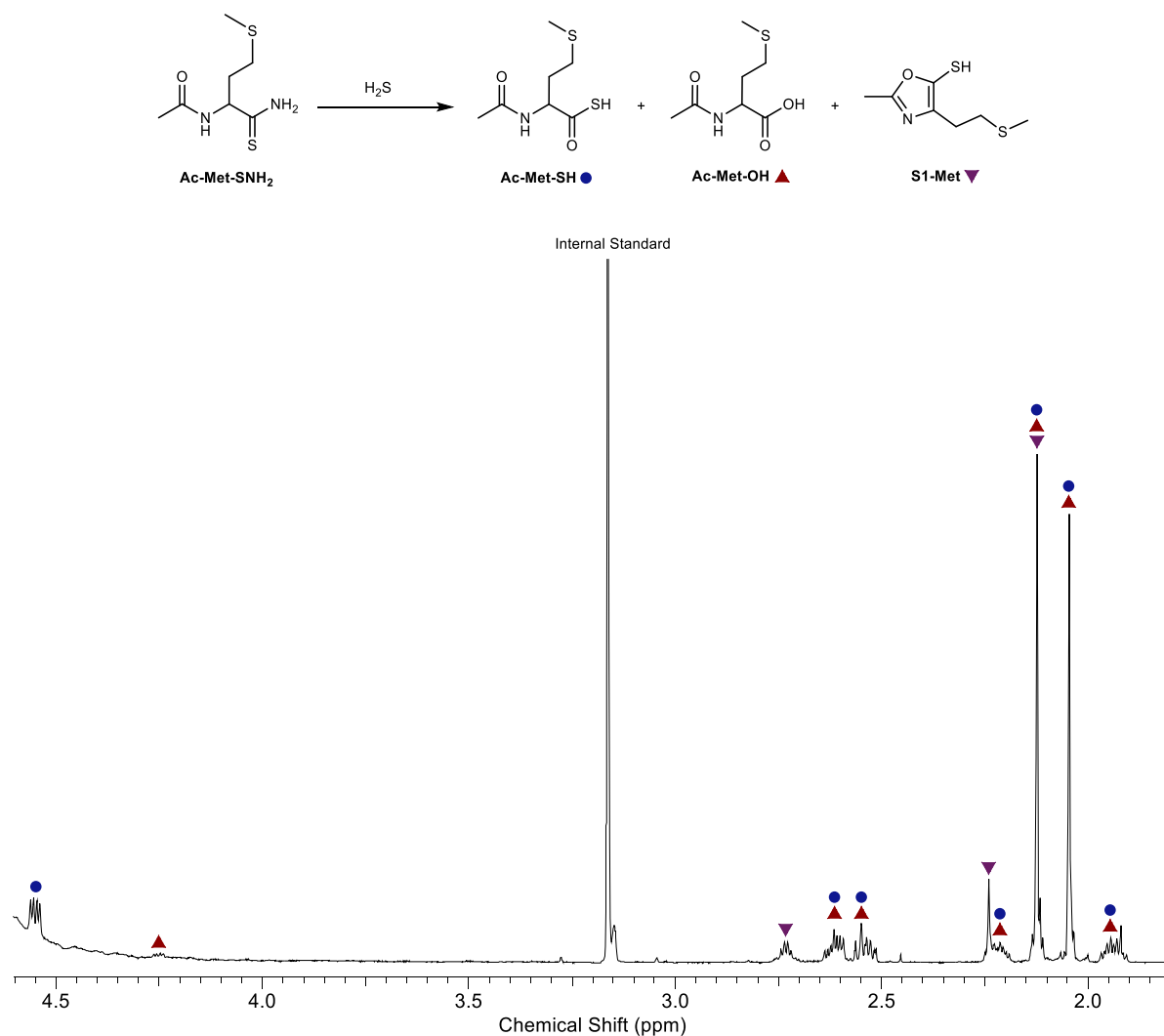
2-Acetamido-4-methylpentanethioic S-acid **Ac-Leu-SH** (77%, 72 h) was produced following *general procedure D* using *N*-acetylleucine thioamide **Ac-Leu-SNH₂** (9.41 mg, 0.05 mmol). ¹H NMR (600 MHz, H₂O/D₂O 98:2): 2-Acetamido-4-methylpentanethioic S-acid **Ac-Leu-SH** (•): δ 4.45 (dd, *J* = 3.6, 10.7 Hz, 1H, (C2)–H), 2.03 (s, 3H, (COCH₃), 1.72 - 1.61 (m, 2H, (C3)–H₂), 1.58 - 1.52 (m, 1H, (C4)–H), 0.92 (d, *J* = 6.4 Hz, 3H, (C5)–H₃), 0.90 (d, *J* = 6.4 Hz, 3H, (C5')–H₃). *N*-Acetylleucine **Ac-Leu-OH** (▲, partial assignment): δ 4.28 - 4.23 (obs. m, 1H, (C2)–H), 2.06 (s, 3H, (COCH₃), 1.72 - 1.61 (obs. m, (C3)–H₂) 1.32 (d, *J* = 7.3 Hz, 3H). **S1-Leu** (▼): δ 2.30 (t, *J* = 3.5 Hz, 2H, (C3)–H₂), 2.03 (s, 3H, (COCH₃), 1.91 - 1.84 (m, 1H, (C4)–H), 0.87 (d, *J* = 6.6 Hz, 6H, (C5)–H₃, (C5')–H₃).



Supplementary Figure 55. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 0.50–4.60 ppm) spectrum to show the products of the reaction of *N*-acetylleucine thioamide (**Ac-Leu-SNH₂**; 50 mM) with NaSH (500 mM) at 60 °C and pH 9.0 after 72 h.

2-Acetamido-4-(methylthio)butanethioic S-acid *Ac-Met-SH*

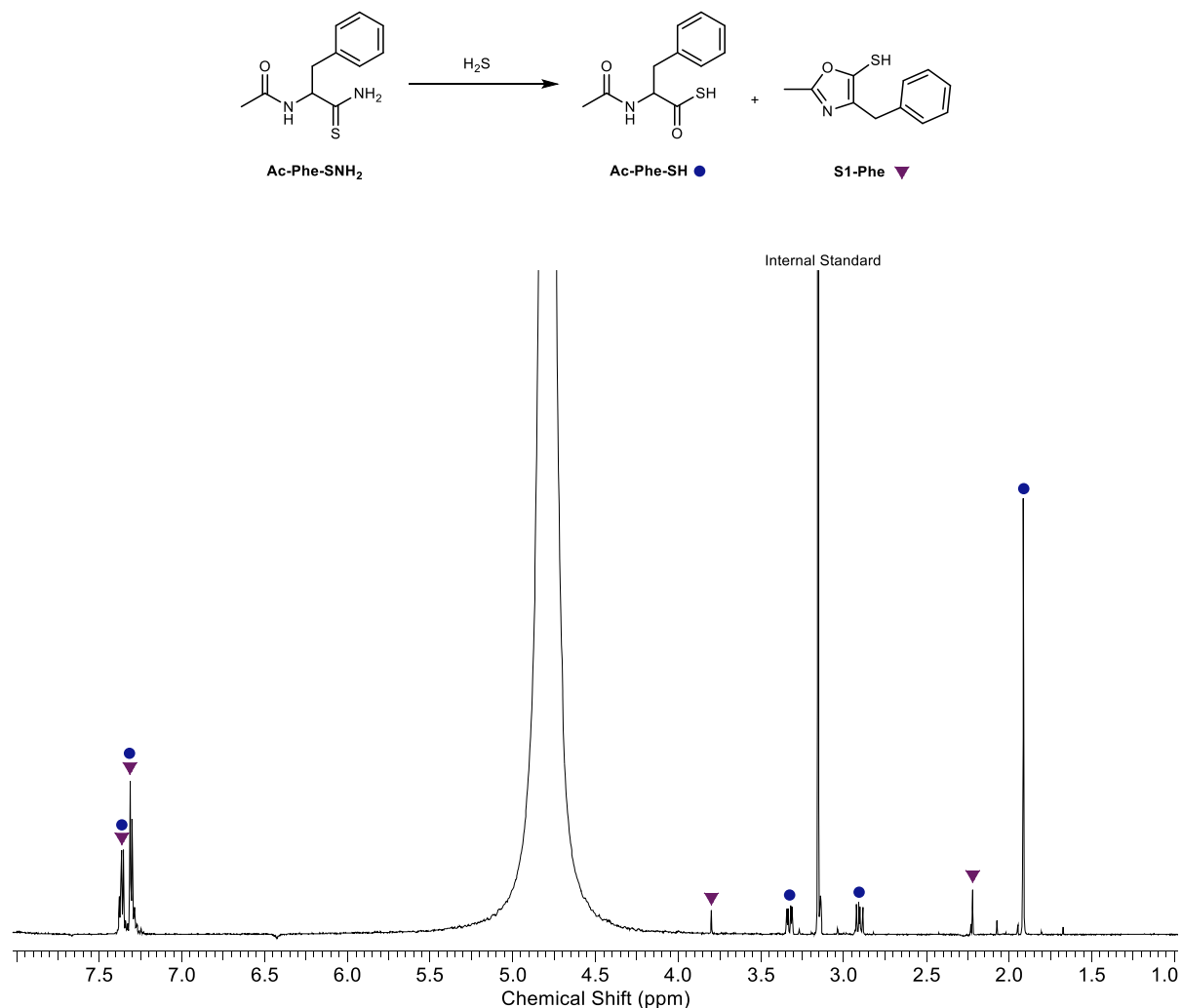
2-Acetamido-4-(methylthio)butanethioic S-acid **Ac-Met-SH** (70%, 72 h) was produced following *general procedure D* using *N*-acetylmethionine thioamide **Ac-Met-SNH₂** (10.61 mg, 0.05 mmol). ¹H NMR (600 MHz, H₂O/D₂O 98:2): 2-Acetamido-4-(methylthio)butanethioic S-acid thioacid **Ac-Met-SH** (●): δ 4.55 (dd, *J* = 4.1, 9.4 Hz, 1H, (C2)–H), 2.64 – 2.58 (m, 1H, (C3)–H), 2.57 – 2.51 (m, 1H, (C3')–H), 2.25 – 2.18 (m, 1H, (C4)–H), 2.12 (s, 3H, SCH₃), 2.05 (s, 3H, (COCH₃), 1.97 – 1.80 (m, 1H, (C4')–H). *N*-Acetylmethionine **Ac-Met-OH** (▲): δ 4.24 (dd, *J* = 4.1, 9.4 Hz, 1H, (C2)–H), 2.64 – 2.58 (m, 1H, (C3)–H), 2.57 – 2.51 (m, 1H, (C3')–H), 2.25 – 2.18 (m, 1H, (C4)–H), 2.12 (s, 3H, SCH₃), 2.05 (s, 3H, (COCH₃), 1.97 – 1.80 (m, 1H, (C4')–H). **S1-Met** (▼): δ 2.76 – 2.69 (m, 4H, (C3)–H₂, (C4)–H₂), 2.24 (s, 3H, (COCH₃), 2.12 (s, 3H, SCH₃).



Supplementary Figure 56. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 1.80–4.60 ppm) spectrum to show the products of the reaction of *N*-acetylmethionine thioamide (**Ac-Met-SNH₂**; 50 mM) with NaSH (500 mM) at 60 °C and pH 9.0 after 72 h.

2-Acetamido-3-phenylpropanethioic S-acid *Ac-Phe-SH*

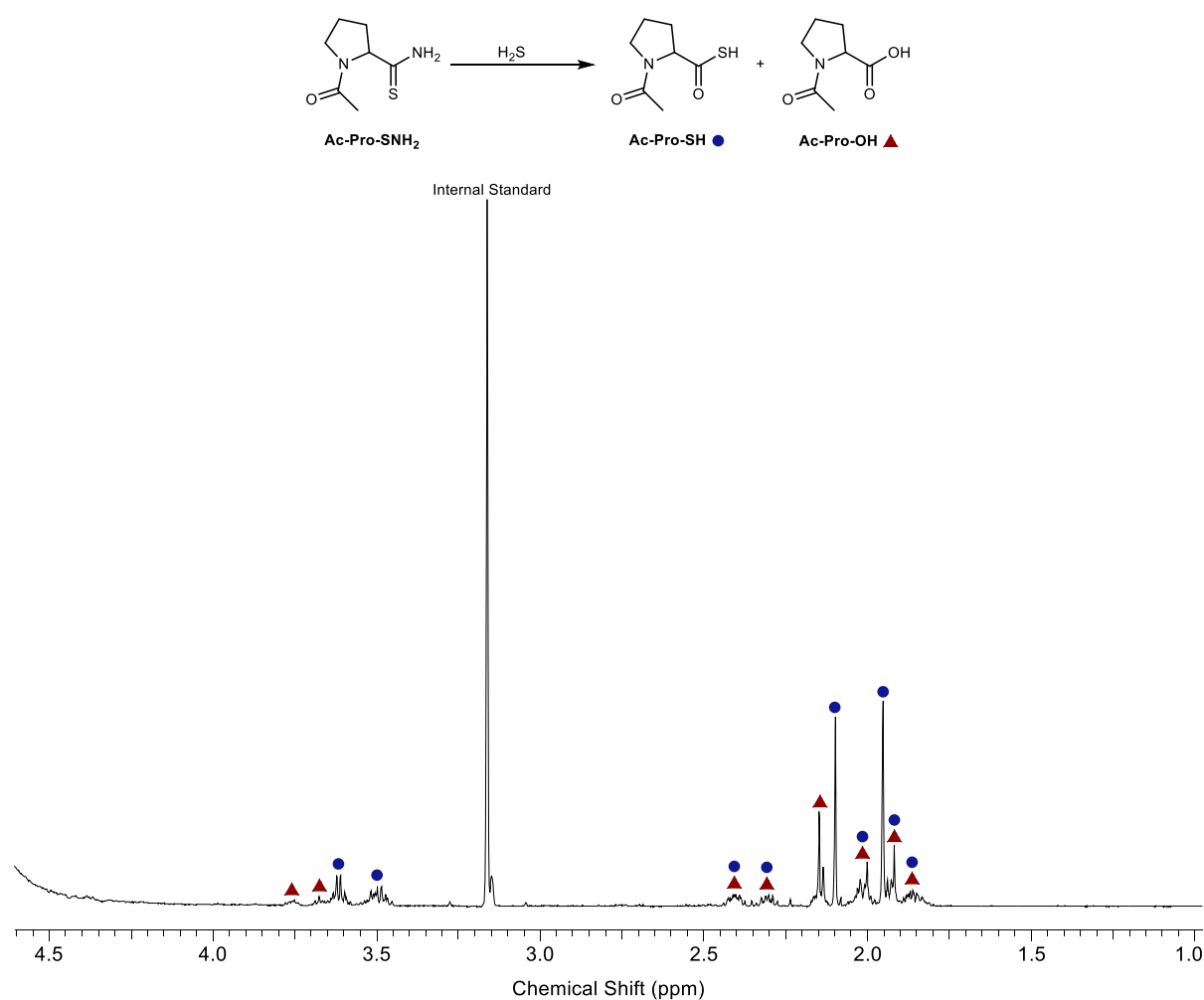
2-Acetamido-3-phenylpropanethioic S-acid **Ac-Phe-SH** (84%, 48 h) was produced following *general procedure D* using *N*-acetylphenylalanine thioamide **Ac-Phe-SNH₂** (9.41 mg, 0.05 mmol). ¹H NMR (600 MHz, H₂O/D₂O 98:2): 2-Acetamido-3-phenylpropanethioic S-acid **Ac-Phe-SH** (●, partial assignment): δ 7.40 – 7.22 (m, 5H, Ar), 3.33 (ABX, *J* = 4.8, 14.5 Hz, 1H, (C3)–H), 2.90 (ABX, *J* = 9.6, 14.5 Hz, 1H, (C3')–H), 1.87 (s, 3H, (COCH₃)). **S1-Phe** (▼): δ 7.40 – 7.22 (m, 5H, Ar), 3.80 (s, 2H, (C3)–H₂), 2.22 (s, 3H, (COCH₃)).



Supplementary Figure 57. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 0.50–4.60 ppm) spectrum to show the products of the reaction of *N*-acetylphenylalanine thioamide (**Ac-Phe-SNH₂**; 50 mM) with NaSH (500 mM) at 60 °C and pH 9.0 after 48 h.

1-Acetylpyrrolidine-2-carbothioic S-acid **Ac-Pro-SH**

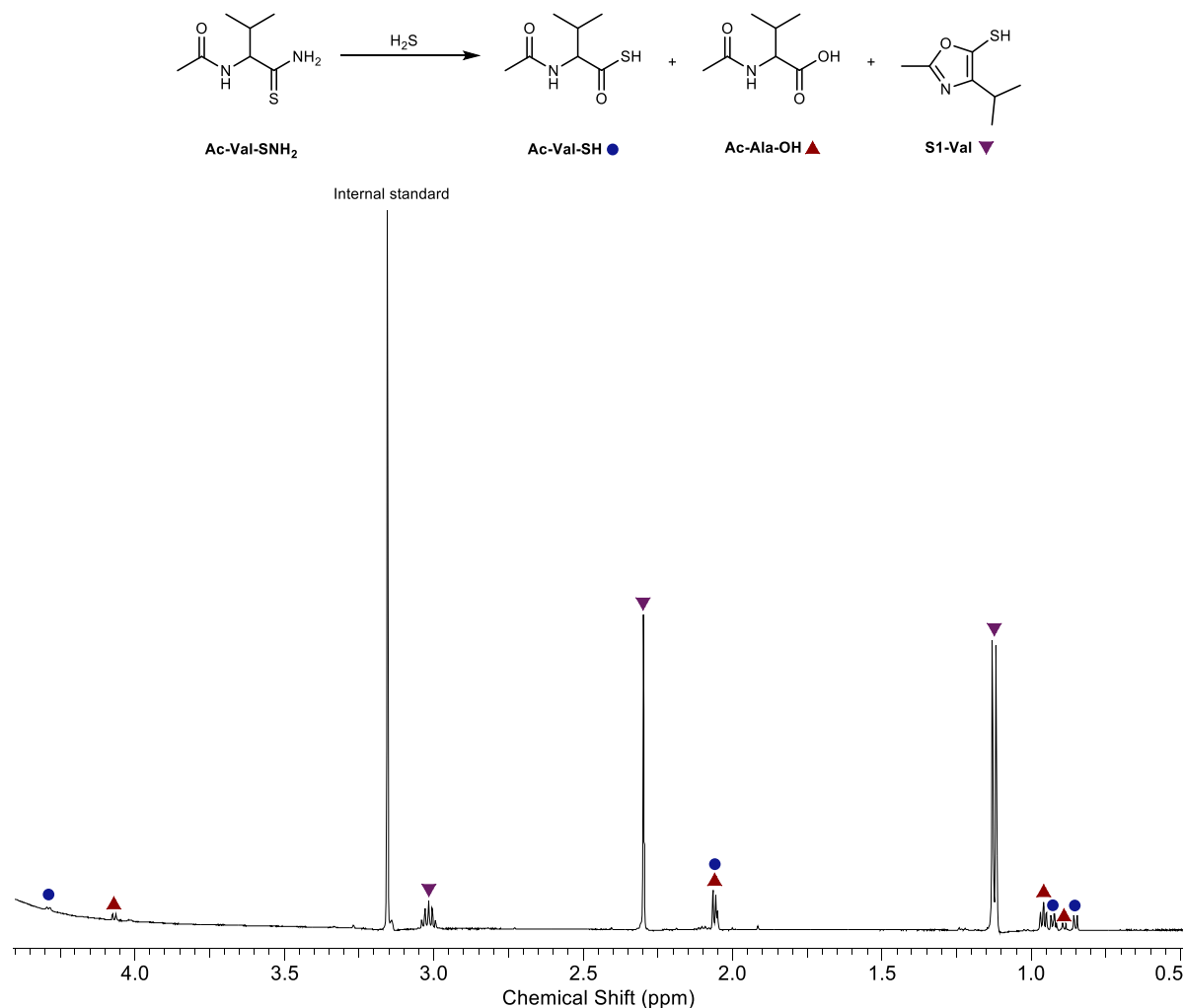
1-Acetylpyrrolidine-2-carbothioic S-acid **Ac-Pro-SH** (72%, 168 h) was produced following *general procedure D* using 1-acetylproline thioamide **Ac-Pro-SNH₂** (8.61 mg, 0.05 mmol). ¹H NMR (600 MHz, D₂O/H₂O 98:2): 1-Acetylpyrrolidine-2-carbothioic S-acid **Ac-Pro-SH** (●, 56:46 mixture of rotamers a/b, partial assignment): rotamer a: δ 3.66 - 3.57 (m, 1H, (C5)–H₂), 2.43–2.35 (m, 1H, (C3)–H), 2.34 - 2.26 (m, 1H, (C3')–H), 1.96 (s, 3H, (COCH₃), 1.89 - 1.83 (m, 2H, (C4)–H₂); rotamer b: δ 3.55 - 3.43 (m, 1H, (C5)–H₂), 2.33 - 2.27 (m, 1H, (C3)–H), 2.07 (s, 3H, (COCH₃), 2.07–1.98 (m, 1H, (C3')–H), 1.93–1.84 (m, 2H, (C4)–H₂). 1-Acetylproline **Ac-Pro-OH** (▲, 55:45 mixture of rotamers a/b, partial assignment) rotamer a: δ 3.78 - 3.74 (m, 2H, (C5)–H₂), 2.34 - 2.23 (m, 2H, (C3)–H₂), 2.15 (s, 3H, (COCH₃), 2.05 - 1.98 (m, 2H, (C4)–H₂); rotamer b: δ 3.70 - 3.65 (m, 2H, (C5)–H₂), 2.44 - 2.37 (m, 2H, (C3)–H₂), 1.93 (s, 3H, (COCH₃), 1.90 - 1.83 (m, 2H, (C4)–H₂).



Supplementary Figure 58. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 0.50–4.60 ppm) spectrum to show the products of the reaction of *N*-acetylproline thioamide (**Ac-Pro-SNH₂**; 50 mM) with NaSH (500 mM) at 60 °C and pH 9.0 after 168 h.

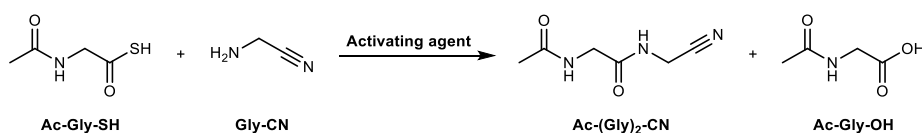
2-Acetamido-3-methylbutanethioic S-acid *Ac-Val-SH*

2-Acetamido-3-methylbutanethioic S-acid **Ac-Val-SH** (8%, 72 h) was produced following *general procedure D* using *N*-acetylvaline thioamide **Ac-Val-SNH₂** (8.61 mg, 0.05 mmol). ¹H NMR (600 MHz, D₂O/H₂O 98:2, partial assignments): 2-Acetamido-3-methylbutanethioic S-acid **Ac-Val-SH** (●): δ 4.29 (d, *J* = 5.9 Hz, 1H, (C2)–H), 2.07 (s, 3H, (COCH₃), 0.93 (d, *J* = 6.9 Hz, 3H, (C4)–H₃), 0.85 (d, *J* = 6.9 Hz, 3H, (C4')–H₃). *N*-Acetylvaline **Ac-Val-OH** (▲): δ 4.06 (d, *J* = 7.3 Hz, 1H, (C2)–H), 2.06 (s, 3H, (COCH₃), 0.96 (d, *J* = 6.9 Hz, 3H, (C4)–H₃), 0.89 (d, *J* = 6.9 Hz, 3H, (C4')–H₃). **S1-Val** (▼): δ 3.06 - 2.98 (m, 1H, (C3)–H), 2.30 (s, 3H, (COCH₃), 1.13 (d, *J* = 6.7 Hz, 6H, (C4)–H₃, (C4')–H₃).



Supplementary Figure 59. ¹H NMR (600 MHz, D₂O/H₂O 98:2, 1.00–4.60 ppm) spectrum to show the products of the reaction of *N*-acetylvaline thioamide (**Ac-Val-SNH₂**; 50 mM) with NaSH (500 mM) at 60 °C and pH 9.0 after 72 h.

Ligation of *N*-acetylaminoacyl thioacids and glycine nitrile with various prebiotic activating agents



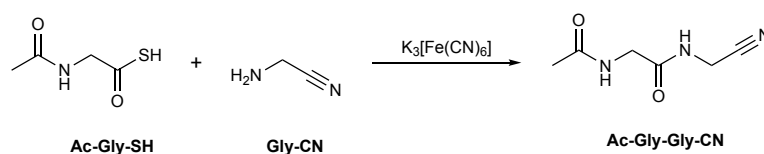
2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM, 0.10 mmol), glycine nitrile **Gly-CN** (100 mM, 0.20 mmol) and methylsulfonylmethane (50 mM, 0.10 mmol) were dissolved in degassed H₂O/D₂O (98:2, 1 mL), and the solution adjusted to pH 9.0 with NaOH/HCl. The total volume adjusted to 2 mL with H₂O/D₂O (98:2), and the specified activating agent (150 mM, 0.30 mmol) was added. The solution was stirred at room temperature for 20 min. The resulting suspension was adjusted to pH 9.0 with NaOH/HCl, centrifuged, and the supernatant was analysed by 1D and 2D NMR spectroscopy in H₂O/D₂O (98:2). The presence of **Ac-(Gly)₂-CN** and **Ac-Gly-OH** was confirmed by ¹H–¹³C HMBC NMR analysis and signal amplification or comparison with pure synthetic standards. The reaction mixture was quantified using methylsulfonylmethane as an internal standard. Yields are reported in Supplementary Table 4.

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

Supplementary Table 4. Yields for the formation of *N*-acetylglycylglycine nitrile **Ac-Gly-Gly-CN** from the coupling of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) and glycine nitrile (**Gly-CN**; 100 mM) after 20 min in the presence of an activating agent (150 mM) at room temperature

Optimisation of ferricyanide-mediated ligation of *N*-acetylaminoacyl thioacids

Effect of aminonitrile stoichiometry and pH

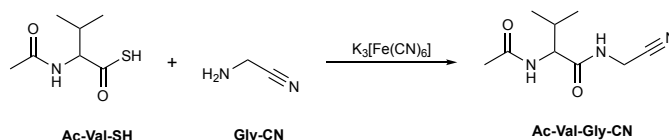


2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM, 0.10 mmol), glycine nitrile **Gly-CN** (55/100/150 mM, 0.11/0.20/0.30 mmol) and methylsulfonylmethane (50 mM, 0.10 mmol) were dissolved in H₂O/D₂O (98:2, 1 mL). The solution was adjusted to the specified pH value with HCl/NaOH and the total volume was adjusted to 2 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (150 mM, 98.8 mg, 0.30 mmol) was then added and the resultant solution stirred at room temperature for 20 min. The resulting suspension was then adjusted to pH 9.0, centrifuged and the supernatant was analysed by 1D and 2D NMR spectroscopy in H₂O/D₂O (98:2). The presence of **Ac-Gly-Gly-CN** was confirmed by ¹H–¹³C HMBC NMR analysis. The reaction mixture was quantified using methylsulfonylmethane as an internal standard. Yields are reported in Supplementary Table 5.

Gly-CN (equiv.)	Gly-CN (mM)	pH	Ac-Gly-Gly-CN (%)
1.1	55	5.0	43
1.1	55	7.0	78
1.1	55	9.0	91
2	100	5.0	95
2	100	7.0	99
2	100	9.0	99
3	150	5.0	96
3	150	7.0	99
3	150	9.0	99

Supplementary Table 5. Yields for the formation of *N*-acetylglycylglycine nitrile **Ac-Gly-Gly-CN** from the oxidative ligation of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) and glycine nitrile (**Gly-CN**; 55, 100 or 150 mM) after 20 min in the presence of K₃[Fe(CN)₆] (150 mM) at room temperature and varying pH.

Effect of reaction concentration and temperature



A solution of 2-acetamido-3-methylbutanethioic S-acid **Ac-Val-SH** (100 mM), glycine nitrile **Gly-CN** (200 mM) and methylsulfonylmethane (50 mM) at pD 9.0 was prepared in D₂O. The solution was diluted with D₂O to final reactant concentrations (given in Supplementary Table 6) and potassium hexacyanoferrate(III) (3 equiv.) was added. The reactions were monitored by ¹H NMR spectroscopy at the specified temperature (given in Supplementary Table 6) and quantified using methylsulfonylmethane as an internal standard. Yields are reported in Supplementary Table 6 (entries 1-5). The reaction carried out at 3° C with 10 mM **Ac-Val-SH** (Supplementary Table 6, entry 6) was analysed by periodic sampling (50 µL) of a batch reaction (1 mL) and quenching by 20-fold dilution with ice-cooled D₂O (1 mL, [**Ac-Val-SH**] = <0.5 mM) followed by immediate NMR analysis. The reaction carried out at 60° C with 10 mM **Ac-Val-SH** (Supplementary Table 6, entry 7) was carried out by preheating a mixture of **Ac-Val-SH** (10 mM), glycine nitrile **Gly-CN** (20 mM) and methylsulfonylmethane (50 mM) at pD 9.0 in D₂O to 60° C before the addition of solid potassium hexacyanoferrate(III) (3 equiv.) with rapid stirring for 2 min before quenching the reaction by 20-fold dilution with ice-cooled D₂O (1 mL, maximum [**Ac-Val-SH**] = <0.5 mM) followed by immediate NMR analysis. The presence of **Ac-Val-Gly-CN** was confirmed by ¹H–¹³C HMBC NMR analysis.

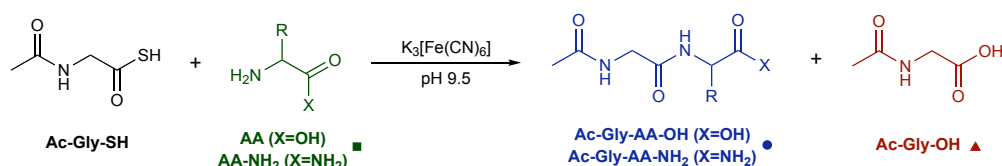
Entry	Ac-Val-SH (mM)	Gly-CN (mM)	K ₃ [Fe(CN) ₆] (mM)	Temp (° C)	Ac-Val-Gly-CN (%)									
					Time (min)									
					2	90	180	285	510	750	990	1260	1920	2700
1	0.5	1	1.5	23	–	0	2	4	17	25	31	38	41	45
2	1	2	3	23	–	4	16	30	50	57	62	62	–	–
3	2.5	5	7.5	23	–	29	57	71	80	80	80	81	–	–
4	5	10	15	23	–	75	85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

Supplementary Table 6. Yields for the formation of *N*-acetylvalylglycine nitrile **Ac-Val-Gly-CN** from the oxidative ligation of 2-acetamido-3-methylbutanethioic S-acid (**Ac-Val-SH**) and glycine nitrile (**Gly-CN**; 2 equiv) in the presence of K₃[Fe(CN)₆] (3 equiv.) at various reaction concentrations. (–) = not determined.

Prebiotic peptide ligations of α -aminoacetyl thioacids

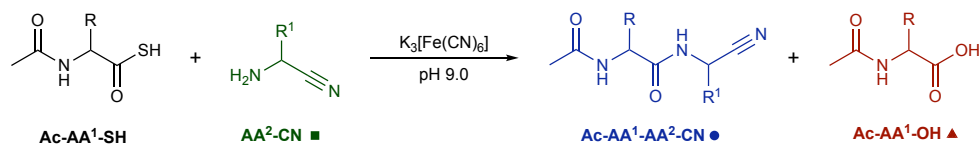
General prebiotic ligation procedures

General procedure E: Prebiotic coupling of 2-acetamidoethanethioic S-acid with α -amino acids or α -amino amides



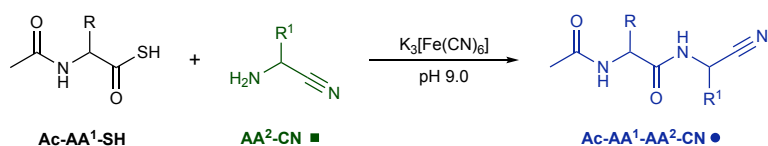
All α -amino acid **AA** and α -amino amide **AA-NH₂** couplings with 2-acetamidoethanethioic S-acid **Ac-Gly-SH** were carried out and analysed according to *general procedure E*, unless stated otherwise. **AA** or **AA-NH₂** (0.3 mmol) was dissolved in degassed H₂O/D₂O (98:2; 1 mL) and the solution was adjusted to pH 9.5 with 1 M-4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.1 mmol) was added and the total volume was adjusted to 2 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (98.8 mg, 0.3 mmol) was added and the solution was stirred at room temperature for 20 min whilst maintaining the solution at pH 9.5 with 1 M-4 M NaOH. The resulting suspension was centrifuged and the supernatant was analysed by 1D and 2D NMR spectroscopy (¹H-¹H COSY; ¹H-¹³C HSQC; ¹H-¹³C HMBC) in H₂O/D₂O (98:2). The reaction mixture was quantified using methylsulfonylmethane (MSM) as an internal standard. The presence of ligation product **Ac-Gly-AA-X** (X=OH or NH₂) was confirmed by ¹H-¹³C HMBC NMR spectral analysis and high resolution mass spectrometry. Reaction mixtures were lyophilised and dissolved in DMSO-*d*₆ or CD₃OD for further NMR spectral analysis if ¹H-¹³C HMBC cross-correlation peaks were obscured by the HOD resonance during original NMR analysis in H₂O/D₂O (98:2). Yields and mass spectrometry data are given in Supplementary Table 8 (α -amino acids **AA** (X=OH)) and Supplementary Table 9 (α -amino amides **AA-NH₂** (X=NH₂)).

General procedure F: Prebiotic coupling of α -aminoacetyl thioacids with α -aminonitriles



All α -aminonitrile **AA²-CN** couplings with α -aminoacetyl thioacids **Ac-AA¹-SH** were carried out according to *general procedure F*, unless stated otherwise. α -Aminonitrile **AA²-CN** (0.2 mmol) was dissolved in degassed H₂O/D₂O (98:2; 2 mL) and the solution was adjusted to pH 9.0 with 1 M-4 M HCl/NaOH. α -Aminoacetyl thioacid **Ac-AA¹-SH** (0.1 mmol) was added and the total volume was adjusted to 2 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (98.8 mg, 0.3 mmol) was added and the solution was stirred at room temperature for 20 min. The pH was readjusted to 9.0 using 1 M-4 M NaOH. The resulting suspension was centrifuged and the supernatant was analysed by 1D and 2D NMR spectroscopy (¹H-¹H COSY; ¹H-¹³C HSQC; ¹H-¹³C HMBC) in H₂O/D₂O (98:2). The reaction mixture was quantified using methylsulfonylmethane (MSM) as an internal standard. The presence of ligation product **Ac-AA¹-AA²-CN** was confirmed by ¹H-¹³C HMBC NMR spectral analysis and high resolution mass spectrometry. Reaction mixtures were diluted with DMSO-*d*₆ (1:49:50; D₂O:H₂O:DMSO-*d*₆), or lyophilised and dissolved in DMSO-*d*₆ or CD₃OD for further NMR spectral analysis if ¹H-¹³C HMBC cross-correlation peaks were obscured by the HOD resonance during original NMR analysis in H₂O/D₂O (98:2). Yields and mass spectrometry data are given in Supplementary Table 7.

General procedure G: Preparative prebiotic coupling of α -aminoacetyl thioacids with α -aminonitriles



α -Aminonitrile **AA²-CN** (1.0 mmol) was dissolved in H₂O (5 mL) and the solution pH was adjusted to pH 9.0 with 10M NaOH. α -Aminoacetyl thioacid **Ac-AA¹-SH** (0.5 mmol) was added and total volume was adjusted to 10 mL with H₂O. Potassium hexacyanoferrate(III) (494 mg, 1.5 mmol) was added and the solution was stirred at room temperature for 20 min. The solution was then extracted with ethyl acetate (3 x 25 mL). The combined organic layers were washed with HCl (0.1 M, 25 mL), NaHCO₃ (sat., 25 mL) and brine (sat., 25 mL), dried over MgSO₄, filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (SiO₂) to give α -aminoacetyl-aminonitrile **Ac-AA¹-AA²-CN** as a white solid. Isolated yields and mass spectrometry data are given in Supplementary Table 7.

Ac-AA ¹ -SH	AA ² -CN	pH	Ac-AA ¹ -AA ² -CN (%)	Ac-AA ¹ -OH (%)	HRMS-ESI Data for Ac-AA ¹ -AA ² -CN		
					Formula	Calculated	Found
Gly	Gly	9.0	99	1	C ₆ H ₈ N ₃ O ₂ [M-H] ⁻	154.0616	154.0613
Ala	Gly	9.0	93	7	C ₇ H ₁₁ N ₃ O ₂ K [M+K] ⁺	208.0483	208.0477
Arg	Gly	9.0	64	n.d ^e	C ₁₀ H ₁₉ N ₆ O ₂ [M+H] ⁺	255.1564	255.1544
Ile	Gly	9.0	83	17	C ₁₀ H ₁₇ N ₃ O ₂ K [M+K] ⁺	250.0952	250.0947
Leu	Gly	9.0	93	7	C ₁₀ H ₁₇ N ₃ O ₂ K [M+K] ⁺	250.0952	250.0961
Lys	Gly	12.0	50	18 ^f	C ₁₀ H ₁₉ N ₄ O ₂ [M+H] ⁺	227.1508	227.1510
		9.0	88	0 ^f			
		7.5	95	0 ^f			
		6.5	95	0 ^f			
Met	Gly	9.0	80	20	C ₉ H ₁₅ N ₃ O ₂ K [M+K] ⁺	268.0517	268.0522
Phe	Gly	9.0	78	22	C ₁₃ H ₁₅ N ₃ O ₂ Na [M+Na] ⁺	268.1056	268.1057
Pro	Gly	9.0	82	18	C ₉ H ₁₃ N ₃ O ₂ K [M+K] ⁺	234.0639	234.0646
Ser	Gly	9.0	87	13	C ₇ H ₁₁ N ₃ O ₂ Na [M+Na] ⁺	208.0693	208.0700
Val	Gly	9.0	92	8	C ₉ H ₁₅ N ₃ O ₂ K [M+K] ⁺	236.0796	236.0803
Gly	Ala	9.0	95 (90 ^a)	5 (n.d ^e)	C ₇ H ₁₁ N ₃ O ₂ [M+H] ⁺	170.0924	170.0924
Ala	Ala	9.0	85	15	C ₈ H ₁₄ N ₃ O ₂ [M+H] ⁺	184.1081	184.1080
Phe	Ala	9.0	(71 ^a)	n.d	C ₁₄ H ₁₇ N ₃ O ₂ [M] ⁺⁺	259.1315	259.1315 ^b
Gly	Arg	9.0	70	6	C ₁₀ H ₁₉ N ₆ O ₂ [M+H] ⁺	255.1564	255.1565
Gly	Asp	9.0	91	8	C ₈ H ₁₁ N ₃ O ₄ Na ₁ [M+Na] ⁺	236.0647	236.0643
Gly	Glu	9.0	74	17	C ₉ H ₁₉ N ₃ O ₄ K [M+K] ⁺	266.0538	266.0536
Gly	Ile	9.0	87 (82 ^a)	13 (n.d ^e)	C ₁₀ H ₁₈ N ₃ O ₂ [M+H] ⁺	212.1394	212.1391
Gly	Leu	9.0	89 (84 ^a)	11 (n.d ^e)	C ₁₀ H ₁₈ N ₃ O ₂ [M+H] ⁺	212.1394	212.1391
Gly	Lys	11.0	80 ^c	<5	C ₁₀ H ₁₉ N ₄ O ₂ [M+H] ⁺	227.1503	227.1508
		9.5	91 ^c	<5			
		9.0	93 ^c	<5			
		8.5	85 ^d	<5			
		7.5	78	10			
		6.5	78	20			
		5.5	53	42			
Gly	Met	9.0	93 (87 ^a)	7 (n.d ^e)	C ₉ H ₁₆ N ₃ O ₂ S [M+H] ⁺	230.0958	230.0957
Gly	Phe	9.0	88 (84 ^a)	12 (n.d ^e)	C ₁₃ H ₁₆ N ₃ O ₂ [M+H] ⁺	246.1237	246.1238
Phe	Phe	9.0	(90 ^a)	(n.d ^e)	C ₂₀ H ₂₁ N ₃ O ₂ [M] ⁺⁺	335.1628	335.1629 ^b
Gly	Pro	9.0	85 (85 ^a)	15 (n.d ^e)	C ₉ H ₁₄ N ₃ O ₂ [M+H] ⁺	196.1081	196.1081
Gly	Ser	9.0	92	8	C ₇ H ₁₁ N ₃ O ₂ Na [M+Na] ⁺	208.0693	208.0695
Gly	Thr	9.0	83	9	C ₈ H ₁₃ N ₃ O ₃ Na [M+Na] ⁺	222.0848	222.0849
Gly	Val	9.0	94 (92 ^a)	6 (n.d ^e)	C ₉ H ₁₆ N ₃ O ₂ [M+H] ⁺	198.1237	198.1237
Phe	Val	9.0	(73 ^a)	(n.d ^e)	C ₁₆ H ₂₂ N ₃ O ₂ [M+H] ⁺	288.1712	288.1719
Val	Val	9.0	(91 ^a)	(n.d ^e)	C ₁₂ H ₂₂ N ₃ O ₂ [M+H] ⁺	240.1707	240.1706

Supplementary Table 7. Yields and high resolution ESI-MS data for the products of prebiotic oxidative coupling of α -aminoacetyl thioacid **Ac-AA¹-SH** (50 mM) with α -aminonitriles **AA²-CN** (100 mM) and K₃[Fe(CN)₆] (150 mM), unless stated otherwise. ^a Isolated yield in parenthesis. ^b Data acquired using high resolution electron impact (EI) ionization mass spectrometry. ^c A mixture of mono- and di-acylated products are observed, see Supplementary Table 13 and Supplementary Page S123 for observed variation in α -, ϵ -, and α,ϵ -di-acylation with solution pH value. ^d A mixture of monoacylated products are observed, see Supplementary Table 13 and page S123 for observed variation in α - and ϵ -acylation with solution pH value. ^e Hydrolysis of **Ac-AA¹-SH** not determined ('n.d') due to product **Ac-AA¹-AA²-CN** isolation, and removal of **Ac-AA¹-OH** during purification, or due to signal overlap in the ¹H NMR spectrum in the case of **Ac-Arg-Gly-CN**. ^f *N*-Acetyl-lysine- ϵ -lactam **c-Ac- α -Lys** (pH 12 (32%), pH 9.0 (7%); pH 7.5 (<5%); pH 6.5 (<5%)) was also observed. See Supplementary Page 64 for further details.

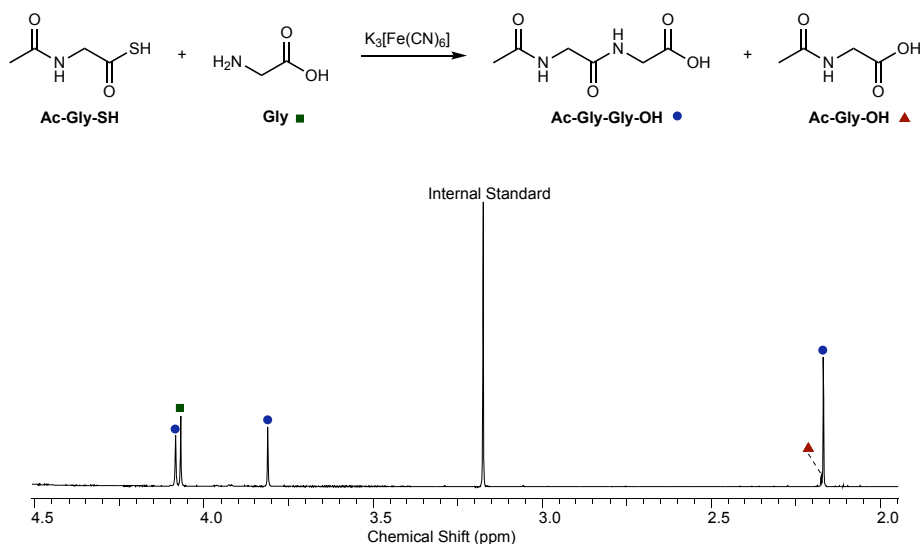
AA	pH	Ac-Gly-AA-OH (%)	Ac-Gly-OH (%)	HRMS-ESI Data for Ac-Gly-AA-OH		
				Formula	Calculated	Found
Gly	9.5	94	6	C ₆ H ₁₁ N ₂ O ₄ [M+H] ⁺	175.0713	175.0716
Ala	9.5	83	15	C ₇ H ₁₃ N ₂ O ₄ [M+H] ⁺	189.0870	189.0874
Arg	9.5	88	7	C ₁₀ H ₂₀ N ₅ O ₄ [M+H] ⁺	274.1515	274.1513
Asn	9.5	81	7	C ₈ H ₁₂ N ₃ O ₅ [M-H] ⁻	230.0777	230.0780
Asp	9.5	89	7	C ₈ H ₁₁ N ₂ O ₆ [M-H] ⁻	231.0617	231.0619
Cys	9.5	80 (84) ^a	10 (9)	C ₇ H ₁₂ N ₂ O ₄ SNa [M+Na] ⁺	243.0410	243.0412
Gln	9.5	90	10	C ₉ H ₁₆ N ₃ O ₅ [M+H] ⁺	246.1084	246.1086
Glu	9.5	92	5	C ₉ H ₁₅ N ₂ O ₆ [M+H] ⁺	247.0925	247.0926
His	9.5	95	5	C ₁₀ H ₁₅ N ₄ O ₄ [M+H] ⁺	255.1088	255.1084
Ile	9.5	84	6	C ₁₀ H ₁₉ N ₂ O ₄ [M+H] ⁺	231.1339	231.1331
Leu	9.5	86	6	C ₁₀ H ₁₇ N ₂ O ₄ [M-H] ⁻	229.1188	229.1190
Lys	9.5	94 ^b	4	C ₁₀ H ₂₀ N ₃ O ₄ [M+H] ⁺	246.1448	246.1446
	8.5	89 ^b	5			
	7.5	84 ^b	5			
	6.5	82 ^b	10			
Met	9.5	95	5	C ₉ H ₁₅ N ₂ O ₄ S [M-H] ⁻	247.0753	247.0751
Phe	9.5	90	8	C ₁₃ H ₁₇ N ₂ O ₄ [M+H] ⁺	265.1183	265.1180
Pro	9.5	89	10	C ₉ H ₁₄ N ₂ O ₄ K [M+K] ⁺	253.0591	253.0592
Ser	9.5	85	9	C ₇ H ₁₃ N ₂ O ₅ [M+H] ⁺	205.0819	205.0821
Thr	9.5	81	9	C ₈ H ₁₃ N ₂ O ₅ [M-H] ⁻	217.0824	217.0821
Trp	9.5	71	9	C ₁₅ H ₁₆ N ₃ O ₄ [M-H] ⁻	302.1141	302.1143
Tyr	9.5	23 ^c	75	C ₁₃ H ₁₇ N ₂ O ₅ [M+H] ⁺	281.1132	281.1139
Val	9.5	84	10	C ₉ H ₁₆ N ₂ O ₄ K [M+K] ⁺	255.0747	255.0748

Supplementary Table 8. Yields and high resolution ESI-MS data for the products of prebiotic oxidative coupling of 2-acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM) with α -amino acids **AA** (150 mM) and K₃[Fe(CN)₆] (150 mM). ^a Yield observed using K₃[Fe(CN)₆] (300 mM) for ligation, followed by methanethiol (600 mM, pH 10.8) reduction. Yield in parenthesis using K₃[Fe(CN)₆] (300 mM) for ligation, followed by tris(2-carboxyethyl)phosphine hydrochloride (**TCEP**) (250 mM, pH 9.5) reduction. For experimental details for **Cys** ligation see Supplementary Information page S96. ^b A mixture of mono- and di-acylated products are observed, see Supplementary Table 11 and page S118 for observed variation in α -, ϵ -, and α,ϵ -di-acylation with solution pH value. HRMS-ESI data for *N*₂,*N*₆-di(acetylglycyl)lysine **Ac-Gly- α,ϵ -Lys-OH** ($\alpha+\epsilon$): calc. for C₁₄H₂₅N₄O₆ [M+H]⁺ 345.1769, obs. 345.1777. ^c L-Tyrosine **Tyr** exhibits extremely low solubility in water (6.5 mM, pH 9.5, room temperature), therefore <5% of the maximal theoretical **Tyr** concentration (150 mM, based on initial **Tyr** stoichiometry) is expected to be dissolved at the outset of *general procedure E* in water. See also Supplementary Reference 51 for a reported solubility of **Tyr** in water (2.8 mM, no pH given). Given its low solubility (<5% of initial 150 mM **Tyr** concentration in water at pH 9.5), it is significant that 23% of the initial L-tyrosine **Tyr** was acylated following *general procedure E*. No further attempts were made to optimise **Tyr** coupling.

AA-NH ₂	pH	Ac-Gly-AA-NH ₂ (%)	Ac-Gly-OH (%)	HRMS-ESI Data for Ac-Gly-AA-NH ₂		
				Formula	Calculated	Found
Gly	9.5	93	7	C ₆ H ₁₂ N ₃ O ₃ [M+H] ⁺	174.0873	174.0875
Ala	9.5	93	6	C ₇ H ₁₄ N ₃ O ₃ [M+H] ⁺	188.1030	188.1028
Arg	9.5	74	8	C ₁₀ H ₂₁ N ₆ O ₃ [M+H] ⁺	273.1675	273.1678
Asn	9.5	87	5	C ₈ H ₁₄ N ₄ O ₄ K [M+K] ⁺	269.0652	269.0698
	8.0	76	17			
	6.5	72	17			
Asp	9.5	80	7	C ₈ H ₁₄ N ₃ O ₅ [M+H] ⁺	232.0928	232.0931
Gln	9.5	65	12	C ₉ H ₁₆ N ₄ O ₄ Na [M+Na] ⁺	267.1064	267.1067
Glu	9.5	90	10	C ₉ H ₁₆ N ₃ O ₅ [M+H] ⁺	246.1084	246.1086
His	9.5	87	4	C ₁₀ H ₁₆ N ₅ O ₃ [M+H] ⁺	254.1248	254.1249
Ile	9.5	74	7	C ₁₀ H ₁₉ N ₃ O ₃ K [M+K] ⁺	268.1064	268.1061
Leu	9.5	72	7	C ₁₀ H ₂₀ N ₃ O ₃ [M+H] ⁺	230.1499	230.1491
Lys	9.5	94 ^a	5	C ₁₀ H ₂₁ N ₄ O ₃ [M+H] ⁺	245.1608	245.1603
	8.5	86 ^a	8			
	7.5	91 ^a	6			
	6.5	84 ^a	n.d			
Met	9.5	74	10	C ₉ H ₁₇ N ₃ O ₃ SK [M+K] ⁺	286.0628	286.0626
Phe	9.5	63	11	C ₁₃ H ₁₈ N ₃ O ₃ [M+H] ⁺	264.1343	264.1347
Pro	9.5	67	9	C ₉ H ₁₅ N ₃ O ₃ K [M+K] ⁺	252.0750	252.0752
Ser	9.5	78	16	C ₇ H ₁₄ N ₃ O ₄ [M+H] ⁺	204.0979	204.0974
Thr	9.5	71	15	C ₈ H ₁₅ N ₃ O ₄ Na [M+Na] ⁺	240.0960	240.0962
Trp	9.5	71	10	C ₁₅ H ₁₈ N ₄ O ₃ K [M+K] ⁺	341.1016	341.1017
Tyr	9.5	9 ^b	91	C ₁₃ H ₁₇ N ₃ O ₄ K [M+K] ⁺	318.0856	318.0849
	8.5	22 ^b	66			
	7.5	36 ^b	54			
	6.5	56 ^b	42			
Val	9.5	72	19	C ₉ H ₁₇ N ₃ O ₃ K [M+K] ⁺	254.0907	254.0910

Supplementary Table 9. Yields and high resolution ESI-MS data for the products of prebiotic oxidative coupling of 2-acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM) with amino amides **AA-NH₂** (150 mM) and K₃[Fe(CN)₆] (150 mM). ^a A mixture of mono- and di-acylated products are observed, see Supplementary Table 12 and page S120 or observed variation in α -, ϵ -, and α,ϵ -di-acylation with solution pH value. HRMS-ESI data for *N*₂,*N*₆-di(acetylglucyl)lysineamide **Ac-Gly- α,ϵ -Lys-NH₂** ($\alpha+\epsilon$): calcd. for C₁₄H₂₆N₅O₆ [M+H]⁺ 344.1928, obs. 344.1928. ^b NMR analysis carried out at pH 9.5 after incubation at room temperature for 16 h. n.d = Not determined due to signal overlap.

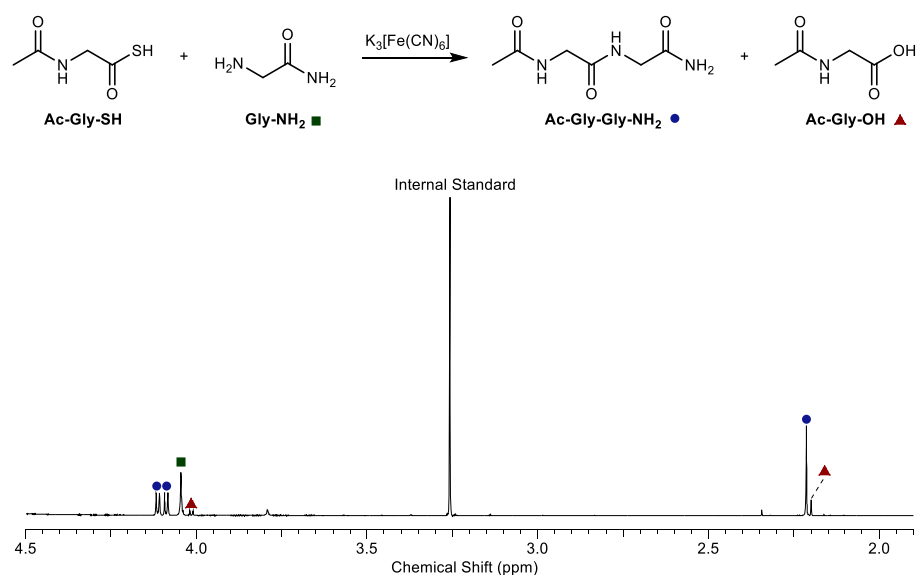
N-Acetylglycylglycine **Ac-Gly-Gly-OH**



Supplementary Figure 60. 1H NMR (600 MHz, D_2O/H_2O 2:98, 1.90–4.50 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine (**Gly**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in H_2O/D_2O (98:2, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and NMR spectra were acquired.

1H NMR (600 MHz, 2:98 $D_2O:H_2O$, pH 9.0) *N*-Acetylglycylglycine **Ac-Gly-Gly-OH** (•): δ 4.08 (s, 2H, Gly₂-(C2)-H₂), 3.81 (s, 2H, Gly₁-(C2)-H₂), 2.16 (s, 3H, (COCH₃)). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.98 (br. app. d, J = 4.8 Hz, 2H, (C2)-H₂), 2.16 (s, 3H, (COCH₃)). Glycine **Gly-OH** (■): δ 4.07 (s, 2H, (C2)-H₂).

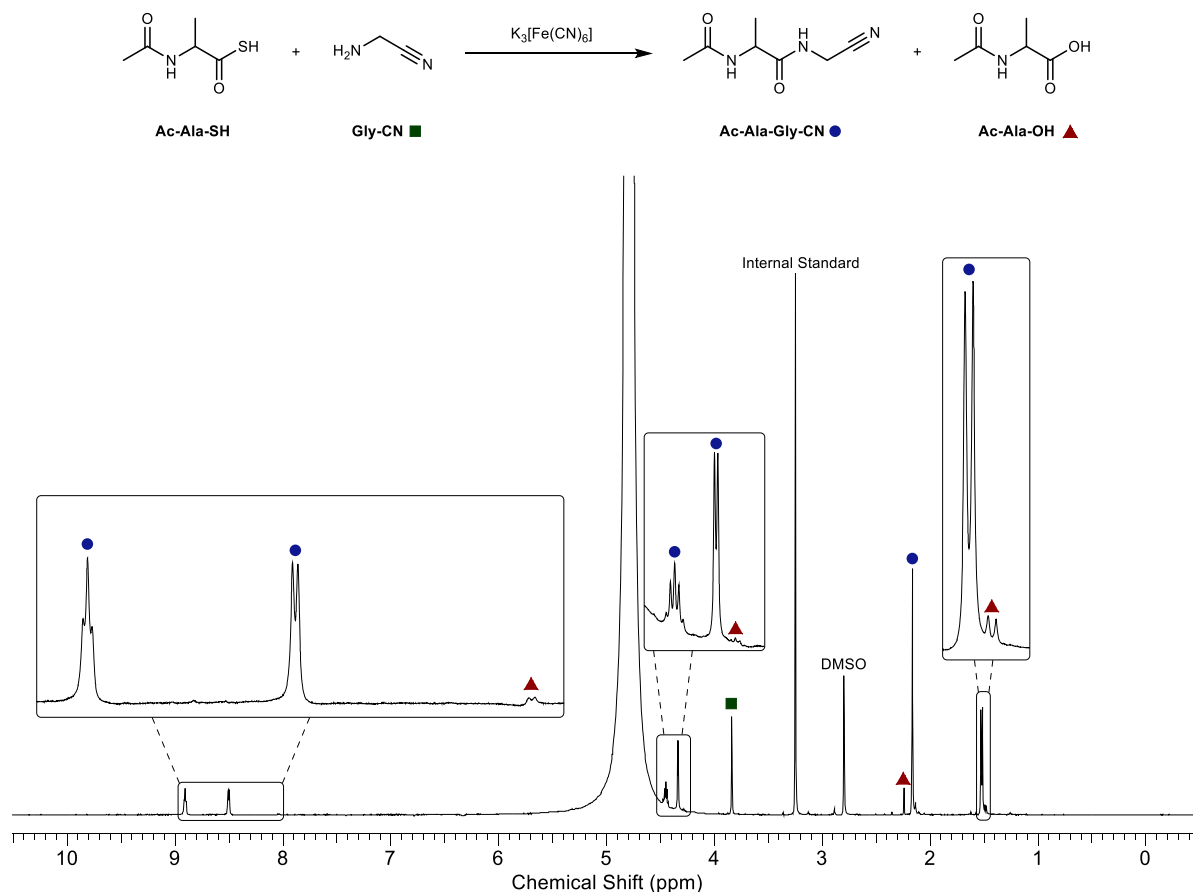
N-Acetylglycylglycinamide **Ac-Gly-Gly-NH₂**



Supplementary Figure 61. 1H NMR (600 MHz, D_2O/H_2O 2:98, 1.90–4.50 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycinamide (**Gly-NH₂**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in H_2O/D_2O (98:2, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and NMR spectra were acquired.

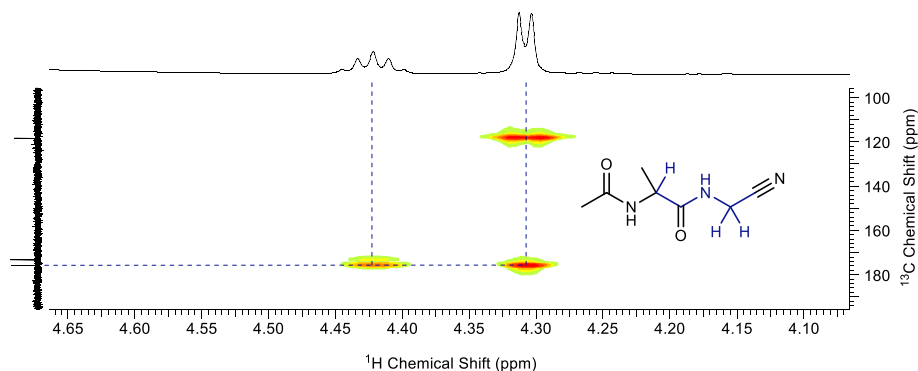
1H NMR (600 MHz, 2:98 $D_2O:H_2O$, pH 9.0): *N*-Acetylglycylglycinamide **Ac-Gly-Gly-NH₂** (•): δ 4.12 (app. d, J = 6.7 Hz, 2H, Gly₂-(C2)-H₂), 4.09 (app. d, J = 6.4 Hz, 2H, Gly₁-(C2)-H₂), 2.21 (s, 3H, (COCH₃)). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 4.02 (d, J = 6.4 Hz, 2H, (C2)-H₂), 2.20 (s, 3H, (COCH₃)). Glycinamide **Gly-NH₂** (■): δ 4.04 (s, 2H, (C2)-H₂).

N-Acetylalanylglycine nitrile **Ac-Ala-Gly-CN**



Supplementary Figure 62. 1H NMR 600 MHz, $D_2O:H_2O:DMSO-d_6$ 1:49:50, -0.50–10.50 ppm) spectrum to show the reaction of 2-acetamidopropanethioic S-acid (**Ac-Ala-SH**; 50 mM), glycine nitrile (**Gly-CN**; 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in H_2O/D_2O (98:2, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and diluted with $DMSO-d_6$ (1 mL) for analysis.

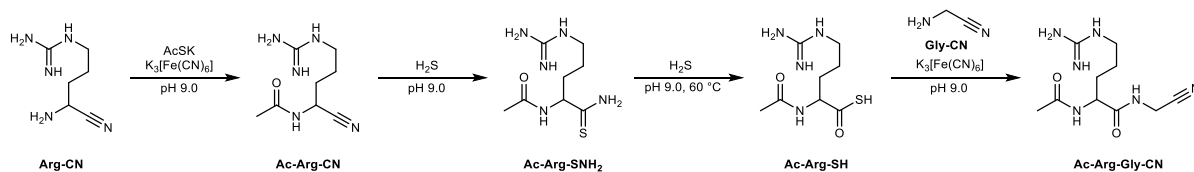
1H NMR (600 MHz, $D_2O/H_2O/DMSO-d_6$ 1:49:50): *N*-Acetylalanylglycine nitrile **Ac-Ala-Gly-CN** (●): δ 8.91 (br t, J = 5.4 Hz, 1H, Gly-NH), 8.50 (d, J = 6.2 Hz, 1H, Ala-NH), 4.48 – 4.42, (m, 1H, Ala-(C2)-H), 4.34, (d, J = 5.4 Hz, 2H, Gly-(C2)-H₂), 2.16 (s, 3H, (COCH₃), 1.52 (d, J = 7.3 Hz, 3H, Ala-(C3)-H₃). *N*-Acetylalanine **Ac-Ala-OH** (▲): δ 8.04 (br d, J = 6.8 Hz, 1H, NH), 4.30 – 4.26, (m, 1H, (C2)-H), 2.24 (s, 3H, (COCH₃), 1.49 (d, J = 7.1 Hz, 3H, (C3)-H₃). Glycine nitrile **Gly-CN** (■): δ 3.84 (s, 2H, (C2)-H₂).



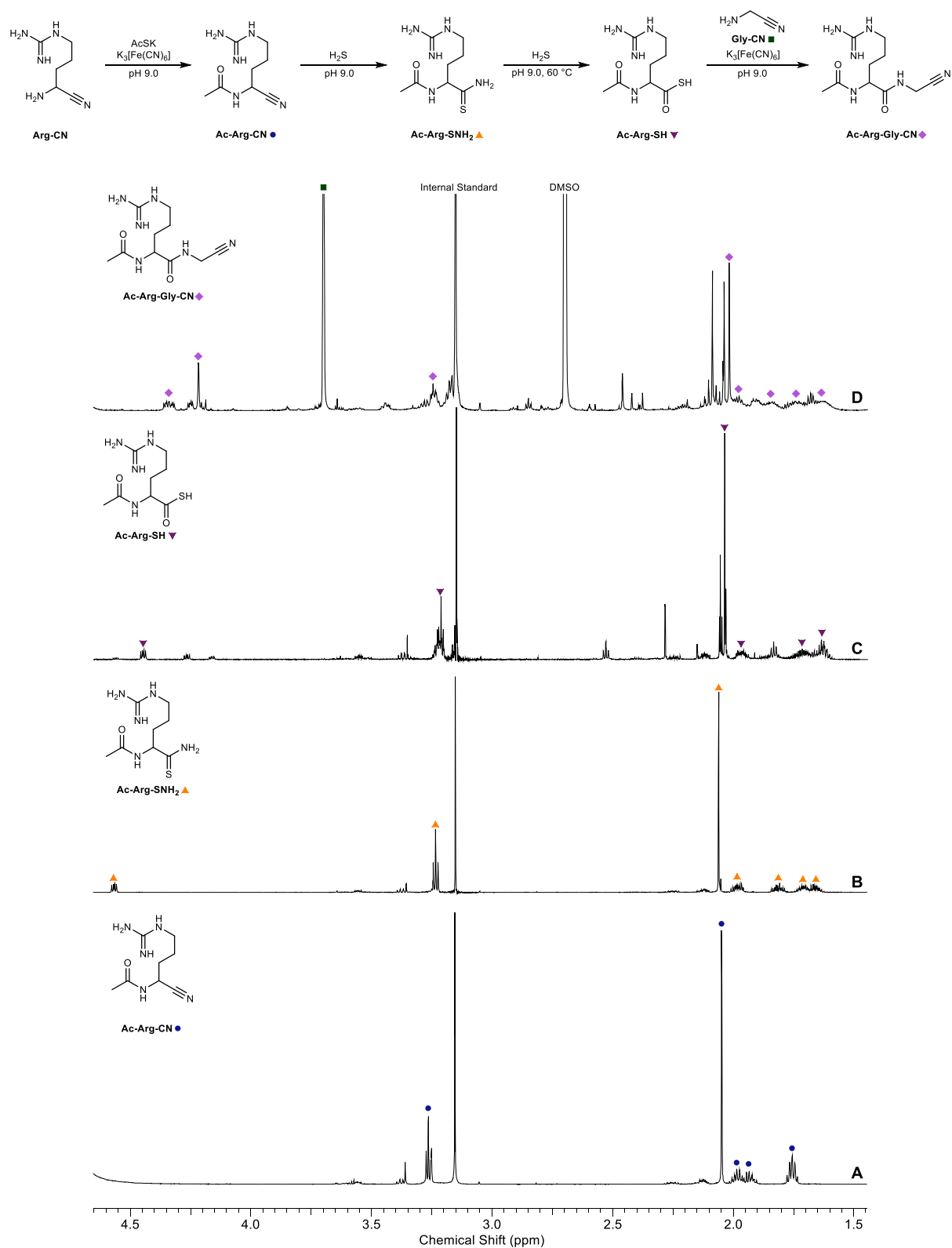
Supplementary Figure 63. 1H - ^{13}C HMBC (1H -600 MHz [4.0–4.6 ppm], ^{13}C -151 MHz [110–185 ppm], $D_2O/H_2O/DMSO-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{CH}$ and $^2J_{CH}$ coupling of Gly-(C2)-H of **Ac-Ala-Gly-CN** at 4.34 ppm with a CN resonance at 122.1 ppm and a C=O resonance at 179.6 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 62 for expanded and labelled 1H NMR spectrum.

N-Acetylarginylglycine nitrile **Ac-Arg-Gly-CN**

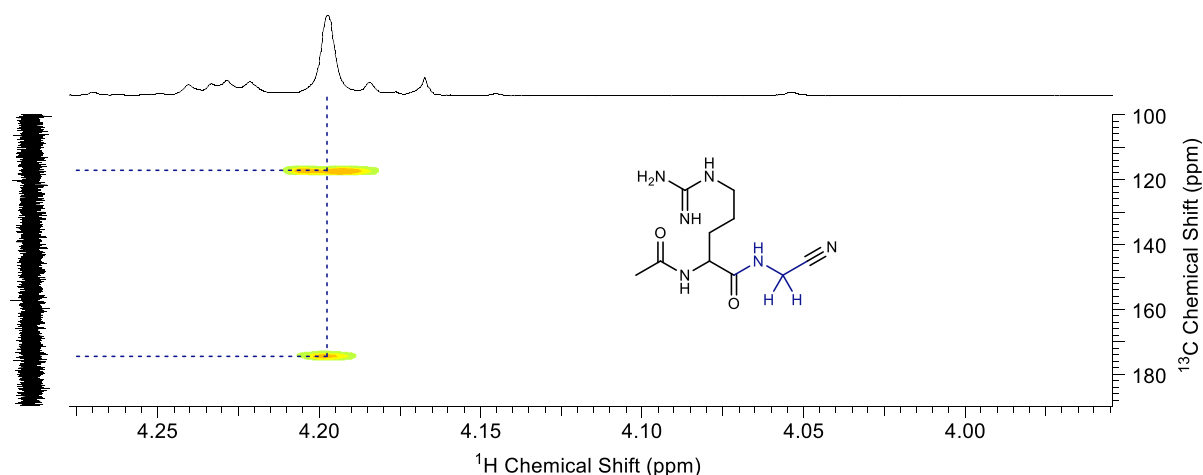
Sequential acetylation, thiolysis, hydrolysis and oxidative ligation *N*-acetylarginine nitrile with glycine nitrile



Arginine nitrile (**Arg-CN**; 0.10 mmol), potassium thioacetate (34.2 mg, 0.30 mmol) and methylsulfonylmethane (9.41 mg, 0.10 mmol) were dissolved in degassed H₂O/D₂O (98:2, 1 mL), and the solution was adjusted to pH 9.0 with NaOH/HCl. Potassium hexacyanoferrate(III) (296.3 mg, 0.90 mmol) was added, and the solution was stirred at room temperature for 20 minutes. After that time, the suspension was adjusted to pH 2.0, and the reaction was lyophilised to remove residual acetic acid. NMR analysis of the residue showed conversion of arginine nitrile **Arg-CN** to *N*-acetylarginine nitrile **Ac-Arg-CN** (89%). NaSH·xH₂O (112.1 mg, 1.00 mmol) was added, the solution was adjusted to pH 9.0 with NaOH/HCl, and incubated at room temperature whilst maintaining the solution at pH 9.0 with NaOH/HCl. After 36 h, NMR analysis showed quantitative conversion of *N*-acetylarginine nitrile **Ac-Arg-CN** to *N*-acetylarginine thioamide **Ac-Arg-SNH₂**. The resulting solution was adjusted to pH 9.0 with NaOH/HCl and incubated at 60 °C for 96 h, after which time NMR analysis showed complete hydrolysis of *N*-acetylarginine thioamide **Ac-Arg-SNH₂**. The solution was then sparged of hydrogen sulfide with argon for 30 min whilst maintaining the solution at pH 7.0 with NaOH/HCl. Glycine nitrile hydrochloride **Gly-CN-HCl** (8.41 mg, 0.091 mmol) was added, and the solution was adjusted to pH 9.0 with NaOH/HCl. Potassium hexacyanoferrate(III) (44.77 mg, 0.136 mmol) was added, and the solution was stirred at room temperature for 20 minutes. DMSO-*d*₆ (1 mL) was then added dropwise to the stirred solution at 0 °C until homogenous, and then NMR spectra were acquired. The presence of *N*-acetylarginylglycine nitrile **Ac-Arg-Gly-CN** (64%) was confirmed by ¹H–¹³C HMBC NMR analysis and high resolution mass spectrometry. The reaction mixtures were quantified by integration against methylsulfonylmethane as the internal standard. Ligation yields are given in Supplementary Table 7.



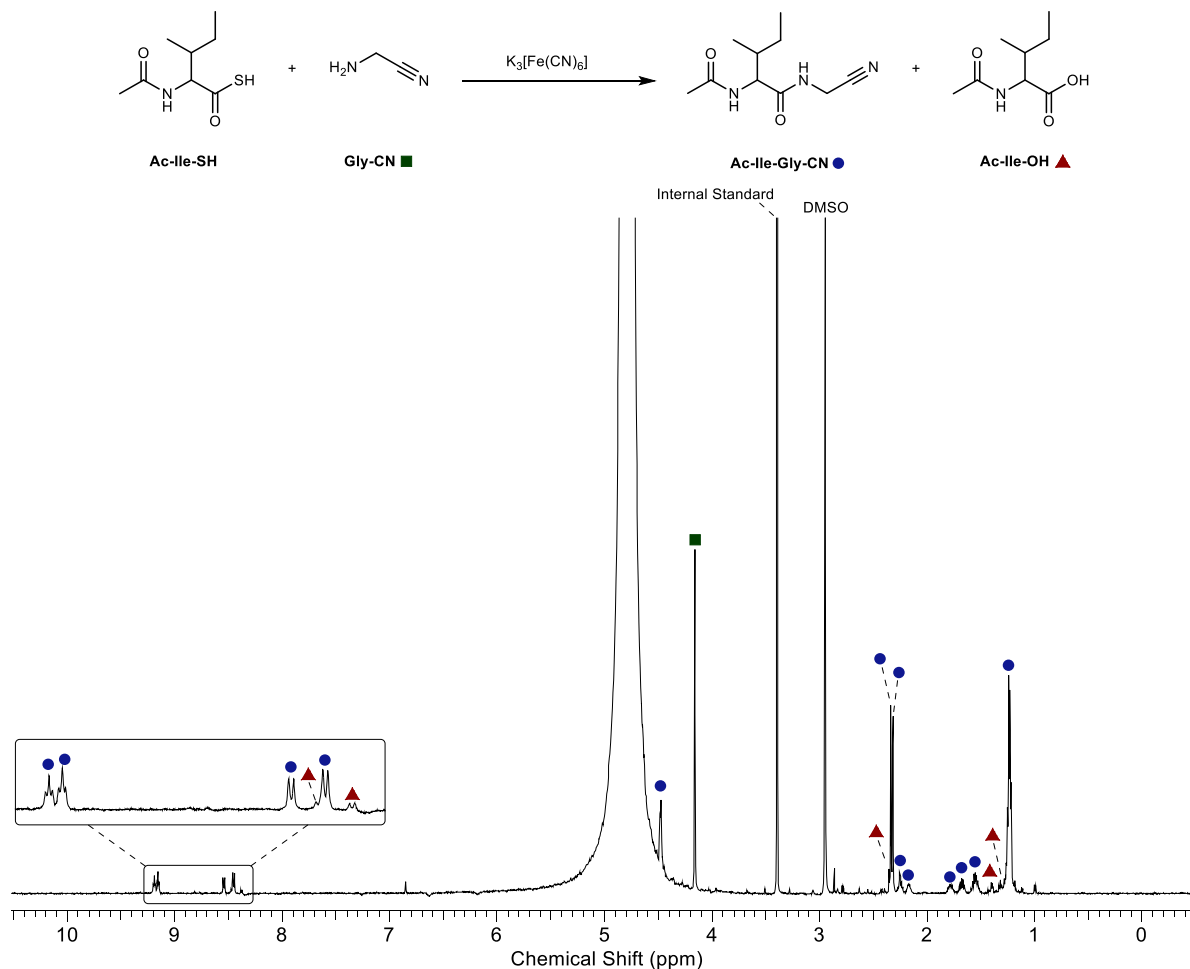
Supplementary Figure 64. ¹H NMR (700 MHz, pH 9.0, 1.45–4.65 ppm) spectra to show the sequential acetylation, thiolysis, hydrolysis and oxidative ligation of arginine nitrile **Arg-CN** with glycine nitrile **Gly-CN**. **A:** Acetylation of **Arg-CN**, H₂O/D₂O 98:2; **B:** Thiolysis of **Ac-Arg-CN** to **Ac-Arg-SNH₂**, H₂O/D₂O 98:2; **C:** Hydrolysis of **Ac-Arg-SNH₂**, H₂O/D₂O 98:2; **E:** Coupling of **Ac-Arg-SH** with **Gly-CN**, H₂O/D₂O/DMSO-*d*₆ 49:1:50.



Supplementary Figure 65. ^1H - ^{13}C HMBC (^1H -700 MHz [3.95–4.25 ppm], ^{13}C -151 MHz [100–180 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO}-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)-H₂ of **Ac-Arg-Gly-CN** at 4.22 ppm with a CN resonance at 117.5 ppm and a C=O resonance at 174.3 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 64D for expanded and labelled ^1H NMR spectrum.

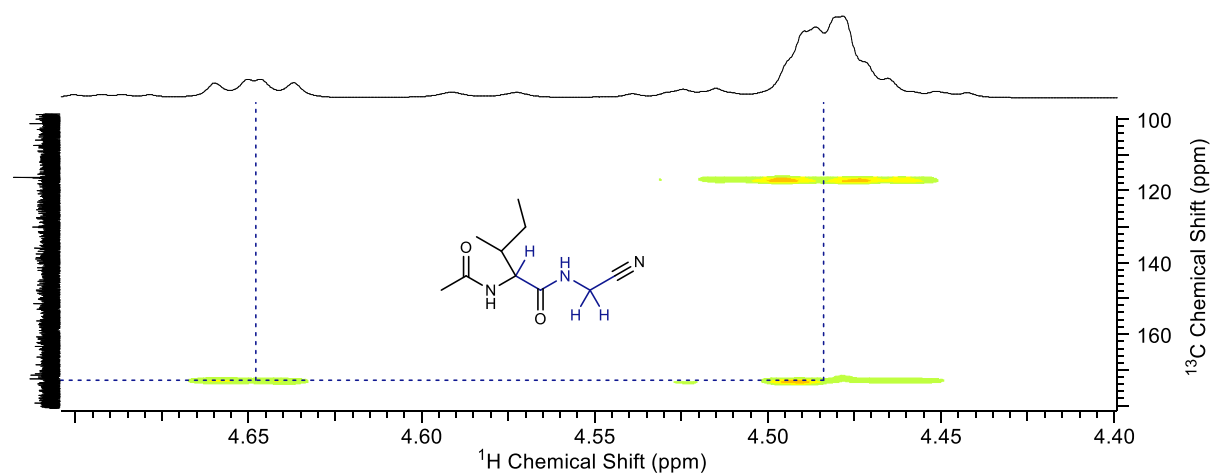
^1H NMR (700 MHz, pH 9.0) data for *Supplementary Figure 64A* (acetylation, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, pH 9.0, partial assignment): *N*-Acetylarginine nitrile **Ac-Arg-CN** (•) δ 3.26 (t, J = 7.2 Hz, 2H, (C5)-H₂), 2.05 (s, 3H, (COCH₃)), 2.02 – 1.95 (m, 1H, (C3)-H), 1.95 – 1.88 (m, 1H, (C3)-H'), 1.79 – 1.71 (m, 2H, (C4)-H₂). *Supplementary Figure 64B* (thiolysis, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, pH 9.0, partial assignment): *N*-Acetylarginine thioamide **Ac-Arg-SNH₂** (▲) δ 4.57 (dd, J = 5.3, 8.8 Hz, 1H, (C2)-H), 3.24 (t, J = 7.2 Hz, 2H, (C5)-H₂), 2.07 (s, 3H, (COCH₃)), 2.02 – 1.95 (m, 1H, (C3)-H), 1.85 – 1.77 (m, 1H, (C3)-H'), 1.76 – 1.68 (m, 1H, (C4)-H), 1.76 – 1.63 (m, 1H, (C4)-H'). *Supplementary Figure 64C* (hydrolysis, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, pH 9.0, partial assignment): 2-Acetamido-5-guanidinopentanethioic *S*-acid **Ac-Arg-SH** (▼) δ 4.47 (dd, J = 4.9, 8.6 Hz, 1H, (C2)-H), 3.24 (obs. t, 2H, (C5)-H₂), 2.05 (s, 3H, (COCH₃)), 2.00 – 1.94 (m, 1H, (C3)-H), 1.77 – 1.69 (m, 1H, (C3)-H'), 1.69 – 1.60 (m, 1H, (C4)-H₂). *Supplementary Figure 64D* (ligation, 700 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}/\text{DMSO}-d_6$ 49:1:50, pH 9.0, partial assignment): *N*-Acetylarginylglycine nitrile **Ac-Arg-Gly-CN** (◆) δ 4.37 – 4.32 (m, 1H, Arg-(C2)-H), 4.22 (s, 2H, Gly₂-(C2)-H₂), 3.25 (obs. t, 2H, Arg-(C5)-H₂), 2.02 (s, 3H, (COCH₃)), 2.01 – 1.96 (obs. m, 1H, Arg-(C3)-H), 1.88 – 1.80 (obs. m, 1H, (C3)-H'), 1.80 – 1.70 (obs. m, 1H, (C4)-H), 1.65 – 1.58 (obs. m, 1H, (C4)-H'); *Glycine nitrile Gly-CN* (■) δ 3.72 (s, 2H, (C2)-H₂).

N-Acetylisoleucylglycine nitrile **Ac-Ile-Gly-CN**



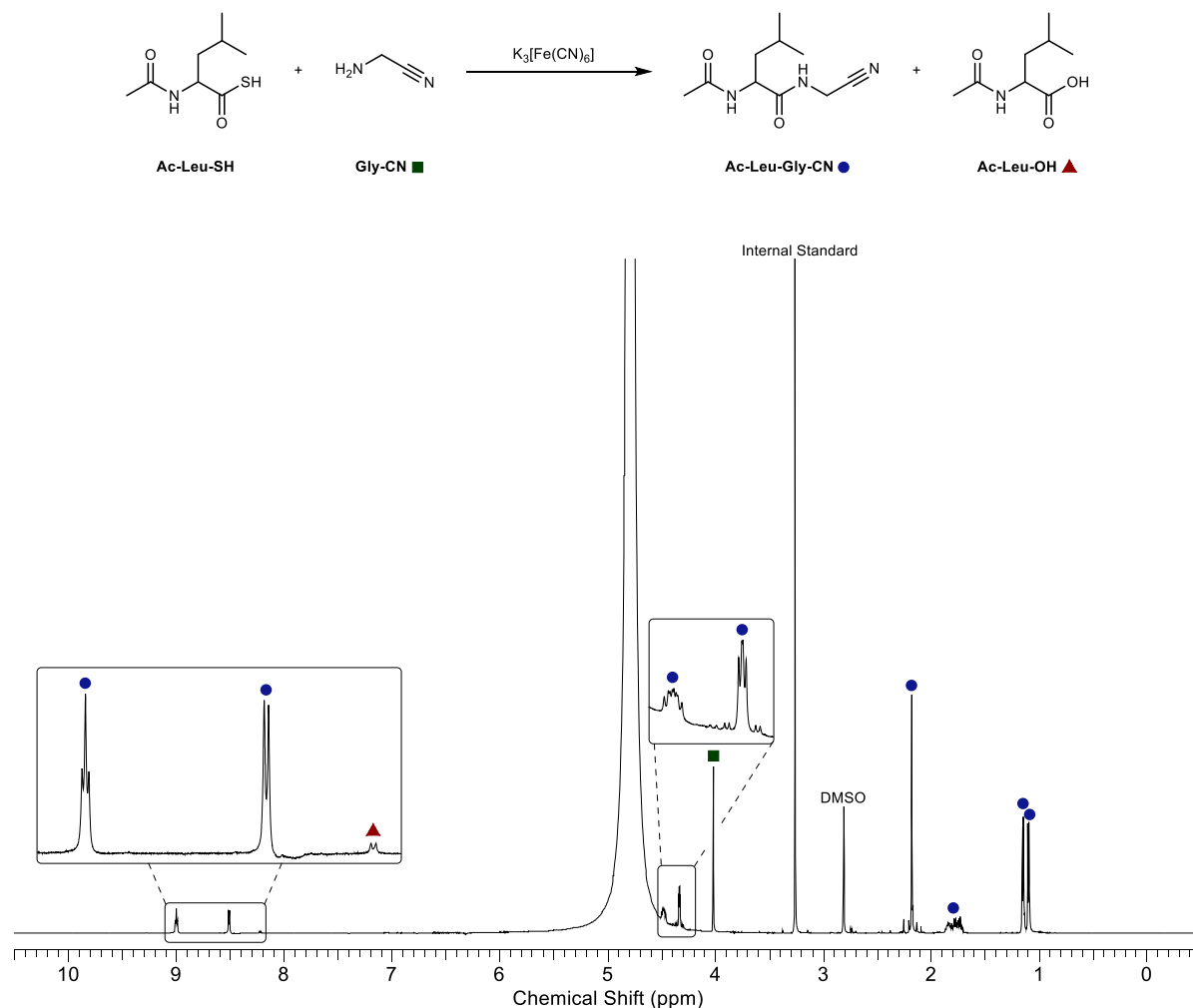
Supplementary Figure 66. ¹H NMR (600 MHz, D₂O:H₂O:DMSO-*d*₆ 1:49:50, -0.50–10.50 ppm) spectrum to show the reaction of 2-acetamido-3-methylpentanethioic S-acid (**Ac-Ile-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in D₂O/H₂O (2:98, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and diluted with DMSO-*d*₆ (1 mL) for analysis.

¹H NMR (600 MHz, 1:49:50 D₂O:H₂O:DMSO-*d*₆, -0.50–10.50 ppm): *N*-Acetylisoleucylglycine nitrile **Ac-Ile-Gly-CN** (•, 54:46 mixture of diastereoisomers a/b): diastereoisomer a (partial assignment): δ 9.16 (t, J = 5.7 Hz, 1H, Gly-NH), 8.45 (d, J = 8.1 Hz, 1H, Ile-NH), 4.50 – 4.44 (m, 2H, Gly-(C2)–H₂), 2.32 (s, 3H, (COCH₃), 2.29 – 2.22 (m, 1H, Ile-(C3)–H), 1.72 – 1.63 (m, 1H, Ile-(C4)–H), 1.59 – 1.49 (obs. m, 1H, Ile-(C4')–H), 1.27 – 1.21 (obs. m, 6H, Ile-(C3)–CH₃, Ile-(C5)–H₃); diastereoisomer b (partial assignment): δ 9.19 (t, J = 5.6 Hz, 1H, Gly-NH), 8.54 (d, J = 8.1 Hz, 1H, Ile-NH), 4.50 – 4.44 (m, 2H, Gly-(C2)–H₂), 2.34 (s, 3H, (COCH₃), 2.20 – 2.13 (m, 1H, Ile-(C3)–H), 1.83 – 1.73 (m, 1H, Ile-(C4)–H), 1.59 – 1.49 (obs. m, 1H, Ile-(C4')–H), 1.27 – 1.21 (obs. m, 6H, Ile-(C3)–CH₃, Ile-(C5)–H₃). *N*-Acetylisoleucine **Ac-Leu-OH** (▲, 51:49 mixture of diastereoisomers a/b): diastereoisomer a (partial assignment): δ 8.48 (obs. d, 1H, NH), 2.33 (s, 3H, (COCH₃), 1.34 – 1.29 (obs. m, 3H, (C5)–H₃); diastereoisomer b (partial assignment): δ 8.38 (d, J = 8.4 Hz, 1H, NH), 2.36 (s, 3H, (COCH₃), 1.44 – 1.37 (obs. m, 3H, (C5)–H₃). Glycine nitrile **Gly-CN** (■): δ 4.16 (s, 3H, (C2)–H₂).



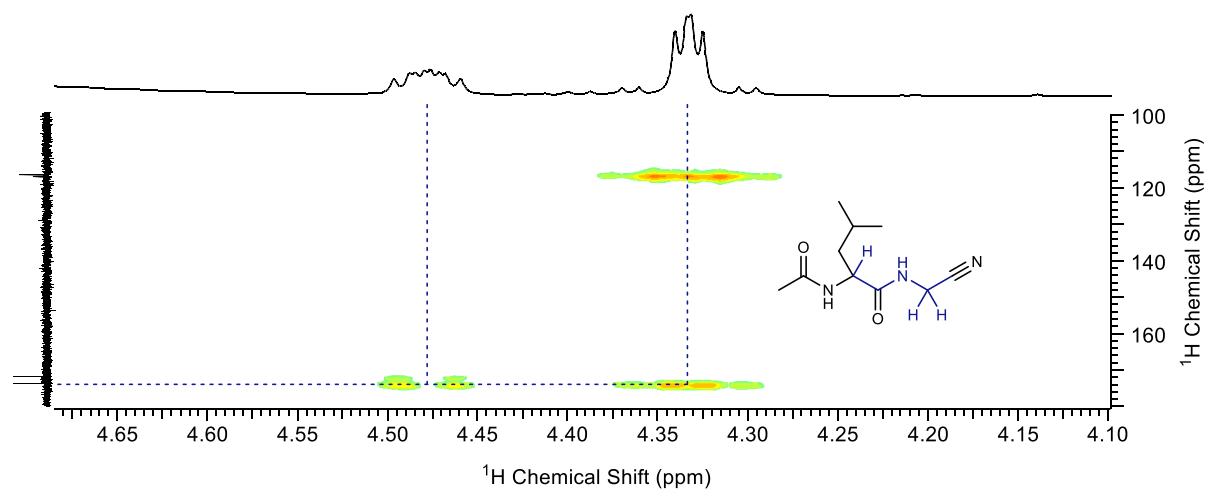
Supplementary Figure 67. ^1H - ^{13}C HMBC (^1H -600 MHz [4.40–4.70 ppm], ^{13}C -151 MHz [100–180 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO}-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)- H_2 of **Ac-Ile-Gly-CN** at 4.44–4.50 ppm with a CN resonance at 118.1 ppm and a C=O resonance at 173.4 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 66 for expanded and labelled ^1H NMR spectrum.

N-Acetylleucylglycine nitrile **Ac-Leu-Gly-CN**



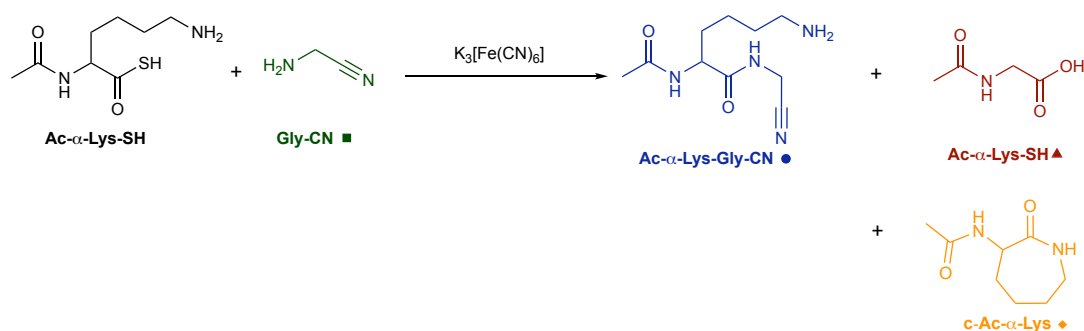
Supplementary Figure 68. ¹H NMR (600 MHz, 1:49:50 D₂O:H₂O:DMSO-*d*₆, -0.50–10.50 ppm) spectrum to show the reaction of 2-acetamido-4-methylpentanethioic S-acid (**Ac-Leu-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in D₂O/H₂O (2:98, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and diluted with DMSO-*d*₆ (1 mL) for analysis.

¹H NMR 600 MHz, 1:49:50 D₂O:H₂O:DMSO-*d*₆): *N*-Acetylleucylglycine nitrile **Ac-Leu-Gly-CN** (•): δ 9.00 (t, J = 5.6 Hz, 1H, Gly-NH), 8.51 (d, J = 7.3 Hz, 1H, Leu-NH), 4.45, (ddd, J = 5.2, 7.0, 10.0, 1H, Leu-(C2)-H), 4.33 (ABX, J = 5.6, 17.6 Hz, 1H, Gly-(C2)-H), 4.29 (ABX, J = 5.6, 17.6 Hz, 1H, Gly-(C2)-H'), 2.18 (s, 3H, (COCH₃)), 1.88 – 1.70 (m, 3H, Leu-(C3)-H₂, Leu-(C4)-H), 1.15 (d, J = 6.5 Hz, 3H, Leu-(C5)-H₃), 1.10 (d, J = 6.5 Hz, 3H, Leu-(C5')-H₃). *N*-Acetylleucine **Ac-Leu-OH** (partial assignment) (▲): δ 8.23 (d, J = 7.6 Hz, 1H, NH). Glycine nitrile **Gly-CN** (■): δ 4.03 (s, 2H, (C2)-H₂).



Supplementary Figure 69. ^1H - ^{13}C HMBC (^1H -600 MHz [4.10–4.70 ppm], ^{13}C -151 MHz [100–180 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO}-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)-H and Gly-(C2)-H' of **Ac-Leu-Gly-CN** at 4.29 and 4.33 ppm with a CN resonance at 117.2 ppm and a C=O resonance at 174.7 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 68 for expanded and labelled ^1H NMR spectrum.

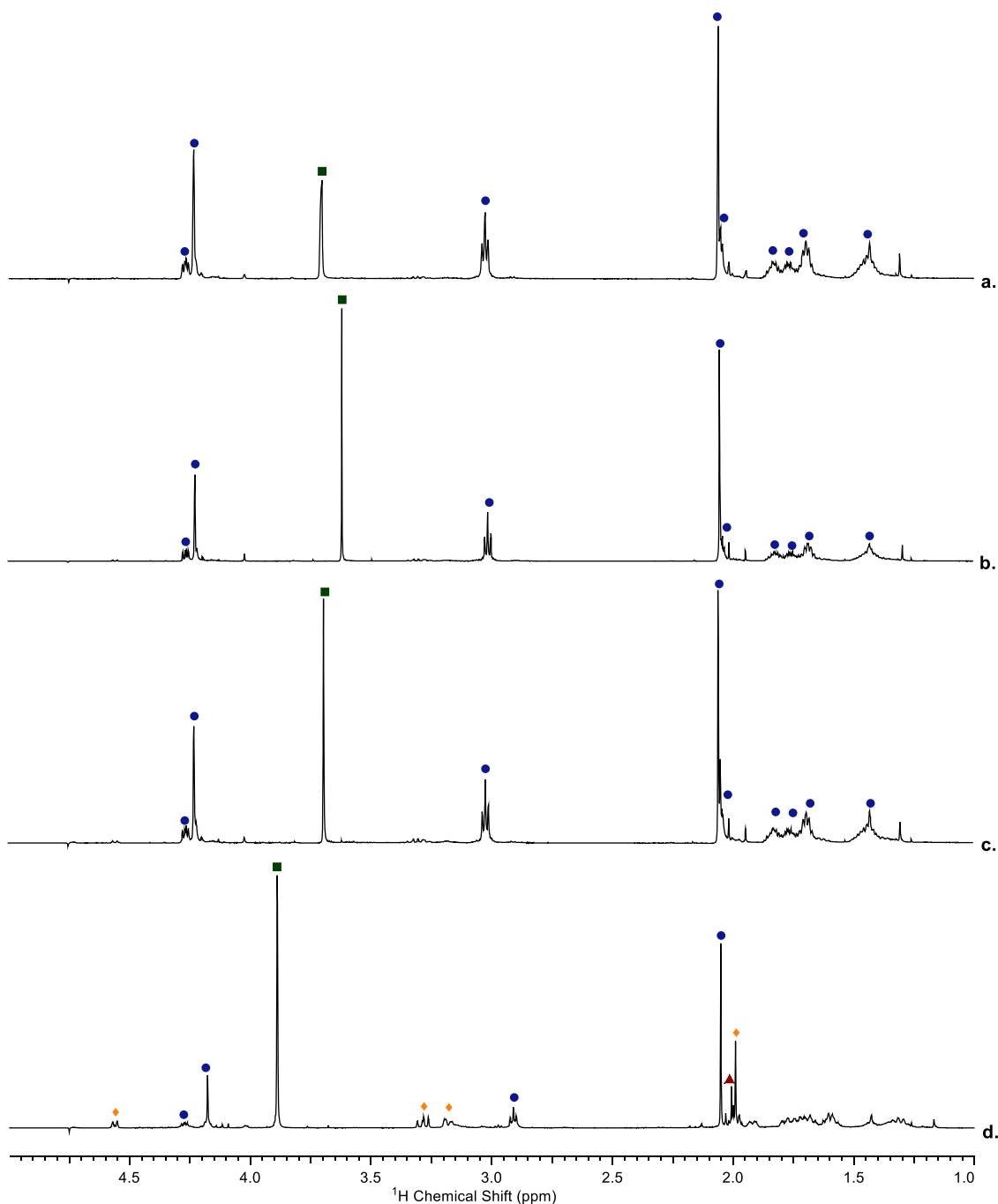
*N*₂-Acetyllysylglycine nitrile **Ac- α -Lys-Gly-CN**



To 2-acetamido-6-aminohexanethioic S-acid **Ac- α -Lys-SH** (final concentration 32 mM) in D₂O was added glycine nitrile hydrochloride **Gly-CN**·HCl (2 equiv., final concentration 64 mM). The pD was adjusted to pD 12.0, 9.0, 7.5 or 6.5, and the total volume was adjusted to 1 mL with D₂O. Potassium hexacyanoferrate(III) (32 mg, 0.10 mmol) was added and the solution was stirred at room temperature for 20 min. The solution was adjusted to pD 9.0, centrifuged, and the supernatant was analysed by 1D and 2D NMR spectroscopy at pD 9.0. The solution was then readjusted to pD 5.5 and further 1D and 2D NMR spectra were acquired. The formation of ligation products was confirmed by ¹H–¹³C HMBC NMR cross-correlation and high-resolution mass spectrometry. See Supplementary Table 7 on page S52 for yields and HRMS-ESI data, and Supplementary Table 10 for product distribution at specified pH values.

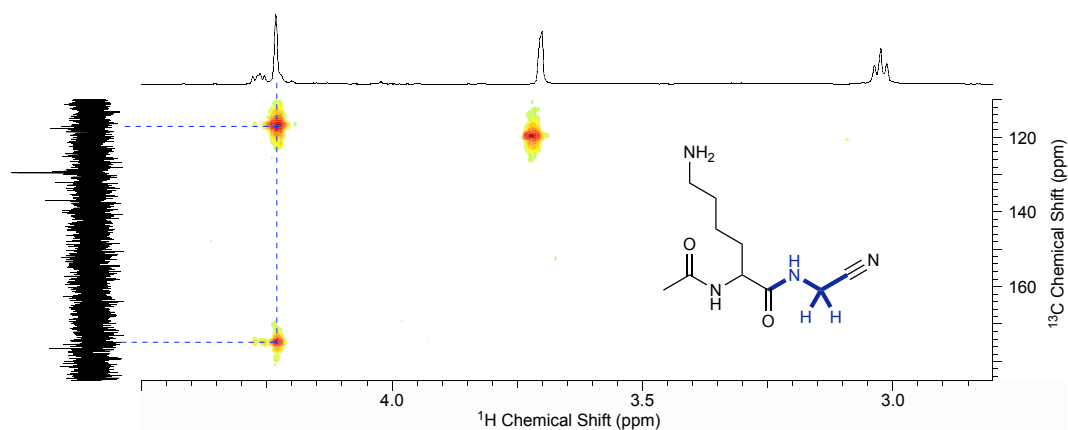
pD	Product	Yield (%)
12.0	c-Ac- α -Lys (♦)	32
	Ac- α -Lys-Gly-CN (●)	50
	Ac- α -Lys-OH (▲)	18
9.0	c-Ac- α -Lys (♦)	7
	Ac- α -Lys-Gly-CN (●)	88
	Ac- α -Lys-OH (▲)	0
7.5	c-Ac- α -Lys (♦)	<5
	Ac- α -Lys-Gly-CN (●)	95
	Ac- α -Lys-OH (▲)	0
6.5	c-Ac- α -Lys (♦)	<5
	Ac- α -Lys-Gly-CN (●)	95
	Ac- α -Lys-OH (▲)	0

Supplementary Table 10. Yields for the products of prebiotic oxidative coupling of 2-acetamido-6-aminohexanethioic S-acid **Ac- α -Lys-SH** (32 mM) with glycine nitrile **Gly-CN** (64 mM) and $K_3[Fe(CN)_6]$ (96 mM) at specified pD values.



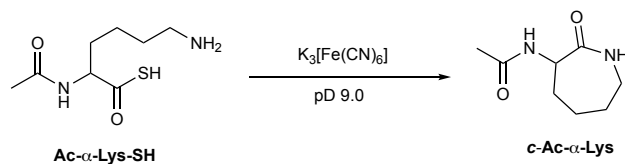
Supplementary Figure 70. ^1H NMR (600 MHz, D_2O , noesygppr1d, 1.0–5.0 ppm) spectrum to show the reaction of 2-acetamido-6-aminohexanethioic *S*-acid **Ac- α -Lys-SH** (32 mM) with glycine nitrile **Gly-CN** (64 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (96 mM) at specified pD values. **a.** pD 6.5; **b.** pD 7.5; **c.** pD 9.0; **d.** pD 12.0. **Gly-CN** (■) (C2)– H_2 resonance underwent migration during NMR data acquisition due to fluctuating pH.

*N*₂-Acetyllysylglycine nitrile **Ac- α -Lys-Gly-CN** (pD 6.5, Supplementary Figure 70a) (●) ^1H NMR (600 MHz, D_2O) δ 4.27 (dd, $J = 5.4, 8.7$ Hz, 1H, Lys-(C2)–H), 4.23 (s, 2H, Gly-(C2)– H_2), 3.02 (t, $J = 7.6$ Hz, 2H, Lys-(C6)– H_2), 2.04 (s, 3H, COCH_3), 1.86 - 1.80 (m, 1H, Lys-(C3)–H), 1.82 - 1.80 (m, 1H, Lys-(C3)– H'), 1.72 - 1.65 (m, 2H, Lys-(C5)– H_2), 1.50 - 1.40 (m, 2H, Lys-(C4)– H_2). *N*-Acetyl-lysine- ϵ -lactam **c-Ac- α -Lys** (pD 12.0, Supplementary Figure 70d, partial data) (◆) 4.57 (dd, $J = 11.6, 1.4$ Hz, 1H, (C3)–H), 3.32 (dd, $J = 15.0, 11.6$ Hz, 1H, (C7)–H), 3.21 - 3.14 (m, 1H, (C7)– H').

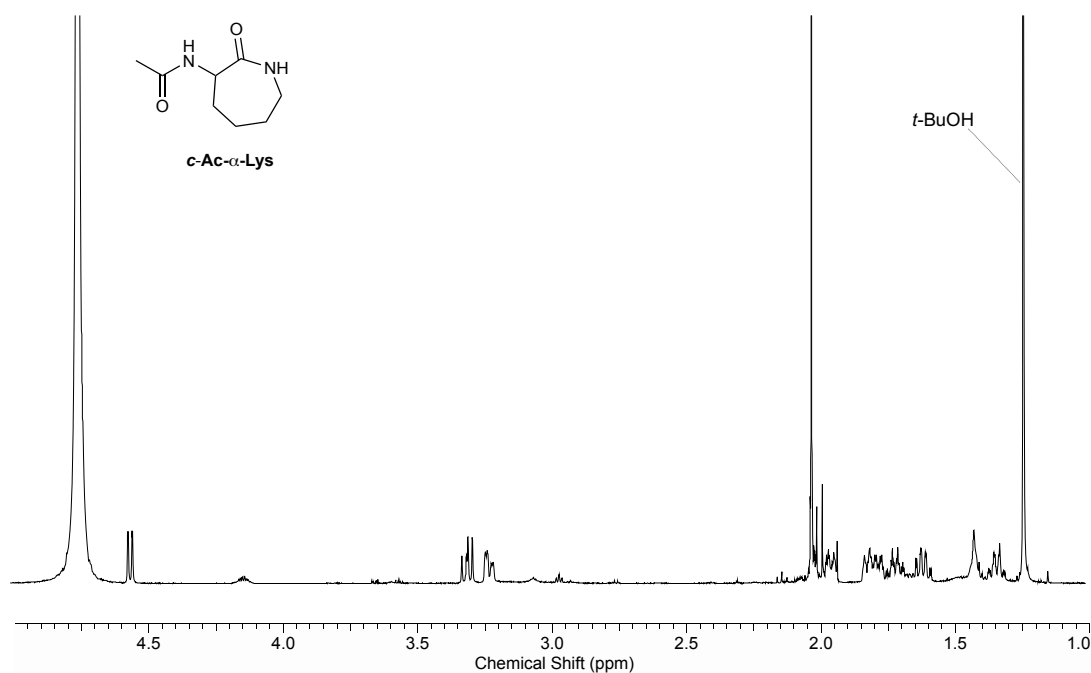


Supplementary Figure 71. ^1H - ^{13}C HMBC (^1H -600 MHz [2.80–4.50 ppm], ^{13}C -151 MHz [110–185 ppm], D_2O) spectrum to show the reaction of 2-acetamido-6-aminohexanethioic S-acid (**Ac- α -Lys-SH**; 32 mM) with glycine nitrile (**Gly-CN**; 64 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (96 mM) at pD 9.0 and room temperature. The diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)- H_2 of **Ac- α -Lys-Gly-CN** at 4.27 ppm with a CN resonance at 117.2 ppm and a C=O resonance at 174.6 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 70a for expanded and labelled ^1H NMR spectrum. **Gly-CN** (■) (C2)- H_2 resonance at 3.70 ppm resonance underwent migration during full NMR data acquisition due to pH fluctuation, especially between ^1H and ^1H - ^{13}C HMBC NMR acquisition to cause the slight off-centre appearance ^1H - ^{13}C cross-correlation peak.

N-Acetyl-lysine- ϵ -lactam **c-Ac- α -Lys**

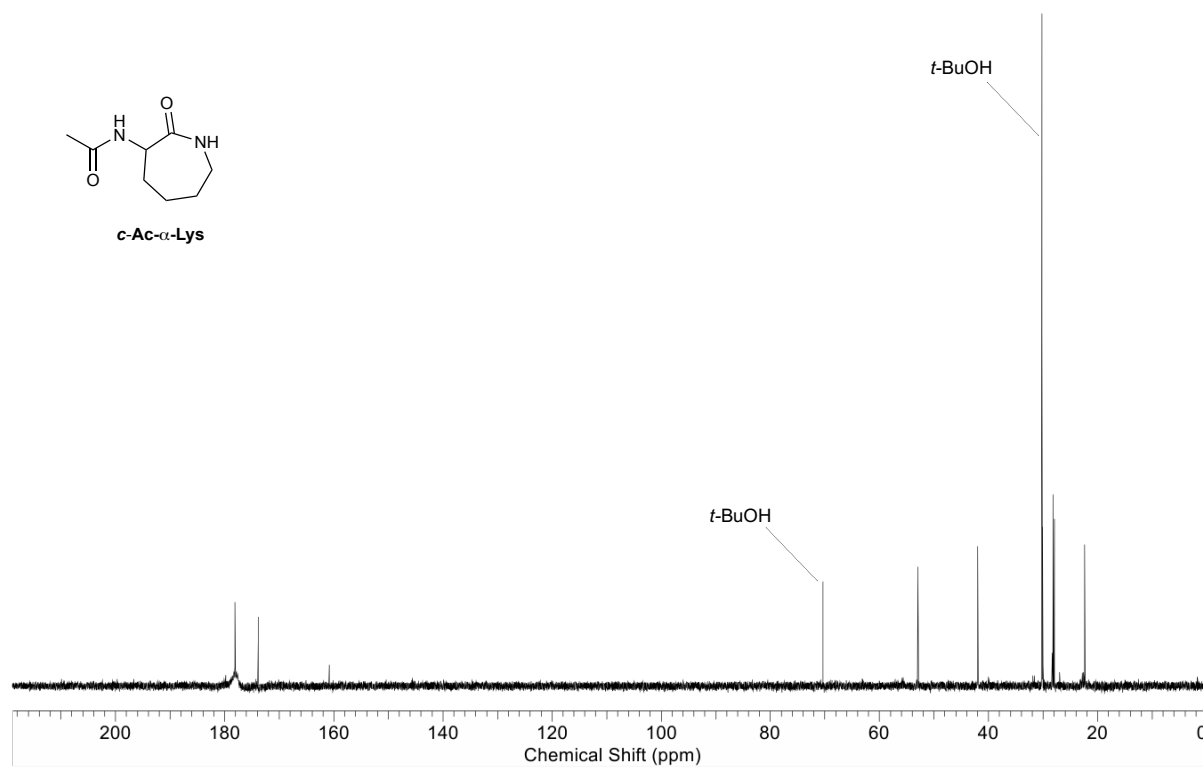


2-Acetamido-6-aminohexanethioic S-acid **Ac- α -Lys-SH** (6.1 mg, 0.03 mmol) in D₂O (1 mL) was adjusted to pD 9.0 with 4 M NaOH/HCl. Potassium hexacyanoferrate(III) (29.6 mg, 0.09 mmol) was added and the reaction was stirred for 20 min. The pD of the solution was readjusted to pD 9.0 and centrifuged. The supernatant was analysed by ¹H, ¹³C and ¹H-¹³C HMBC NMR spectroscopy and high resolution mass spectrometry to confirm the formation of *N*-acetyl-lysine- ϵ -lactam **c-Ac- α -Lys** (92%). *N*-Acetyl-lysine- ϵ -lactam **c-Ac- α -Lys** ¹H NMR (700 MHz, D₂O) δ 4.57 (dd, J = 1.6, 11.7 Hz, 1H, (C3)-H), 3.32 (dd, J = 11.7, 15.3 Hz, 2H, (C7)-H), 3.26 - 3.21 (m, 1H, (C7)-H'), 2.04 (s, 3H COCH₃), 2.00 - 1.91 (m, 1H, (C4)-H), 1.85 - 1.73 (m, 2H, (C5)-H, (C6)-H), 1.75 - 1.69 (m, 1H, (C4)-H'), 1.66 - 1.58 (m, 1H (C5)-H', 1.38 - 1.31 (m, 1H, (C6)-H'). ¹³C NMR (176 MHz, D₂O) δ 178.1 (C2), 173.9 (COCH₃), 52.9 (C3), 42.0 (C7), 30.1 (C4), 28.2 (C6), 28.0 (C5), 22.4 (COCH₃). HRMS-ESI [M+H]⁺ calc. for C₈H₁₅N₂O₂⁺ 171.1128; obs. 171.1130. *tert*-Butanol **t-BuOH** ¹H NMR (700 MHz, D₂O) δ 1.25 (s, 9H, HO-C(CH₃)₃). ¹³C NMR (176 MHz, D₂O) δ 70.4 (HO-C(CH₃)₃), 30.3 (HO-C(CH₃)₃).ⁱⁱ



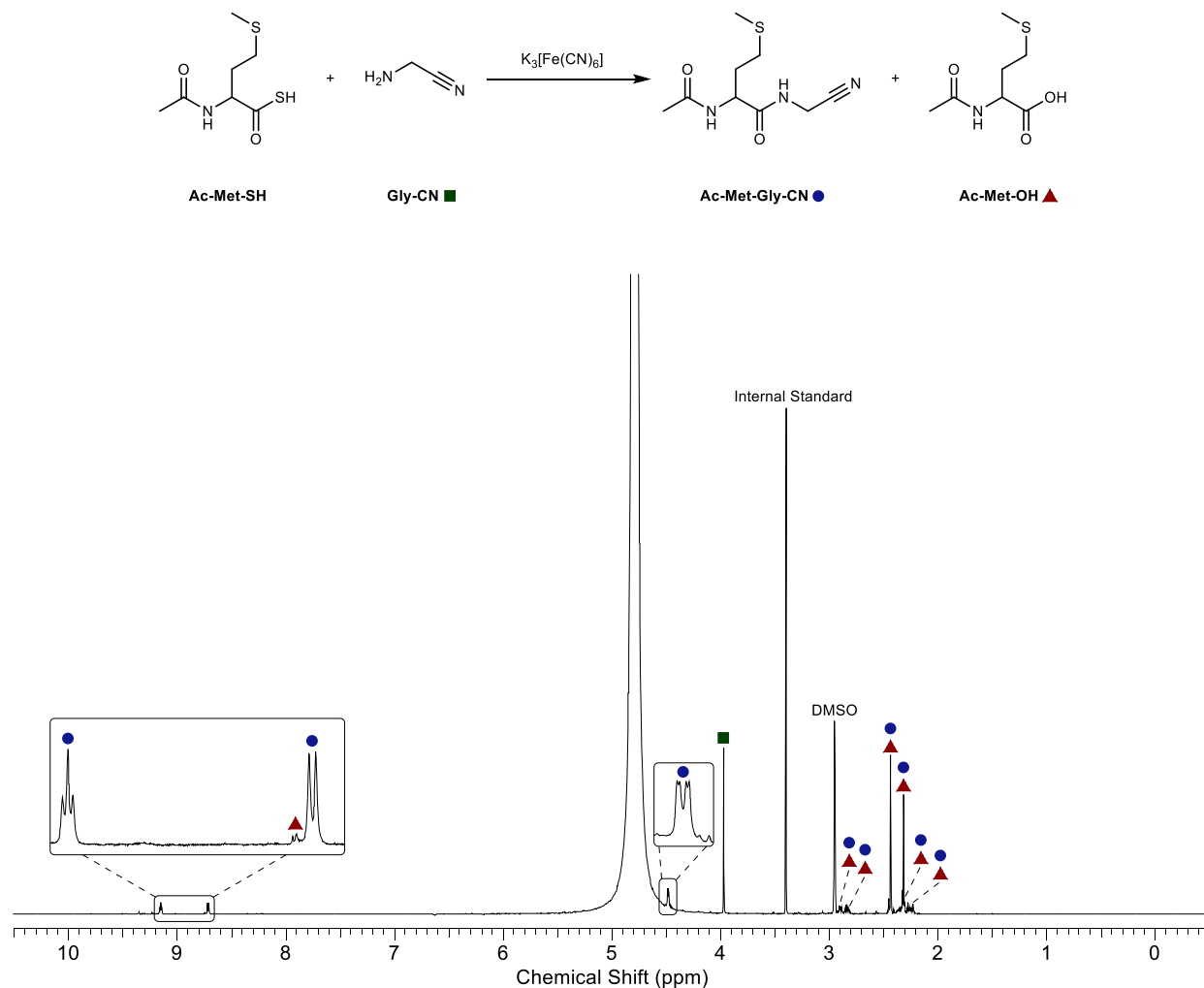
Supplementary Figure 72. ¹H NMR (700 MHz, D₂O, 1.0–5.0 ppm) spectrum of *N*-acetyl-lysine- ϵ -lactam **c-Ac- α -Lys**

ⁱⁱ Residual *t*-Butanol from and *N*- α -acetyllysine **Ac- α -Lys-OH** from the synthesis of 2-acetamido-6-ammoniohexanethioate **Ac- α -Lys-S⁻Na⁺**. See Supplementary Page S207 for further details.



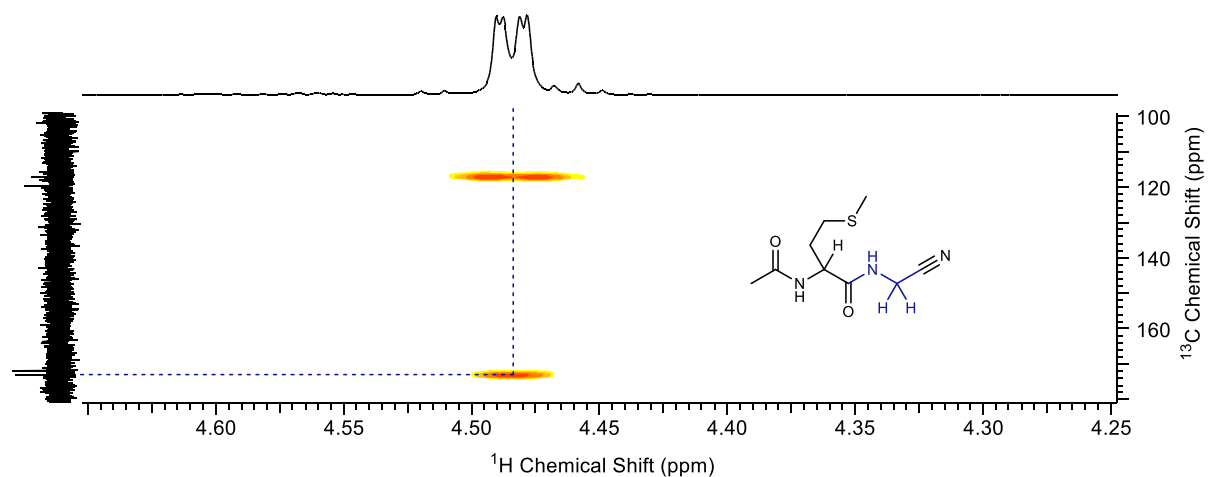
Supplementary Figure 73. ^{13}C NMR (176 MHz, D_2O , 0–220 ppm) spectrum of *N*-acetyl-lysine- ϵ -lactam **c-Ac- α -Lys**

N-Acetylmethionylglycine nitrile **Ac-Met-Gly-CN**



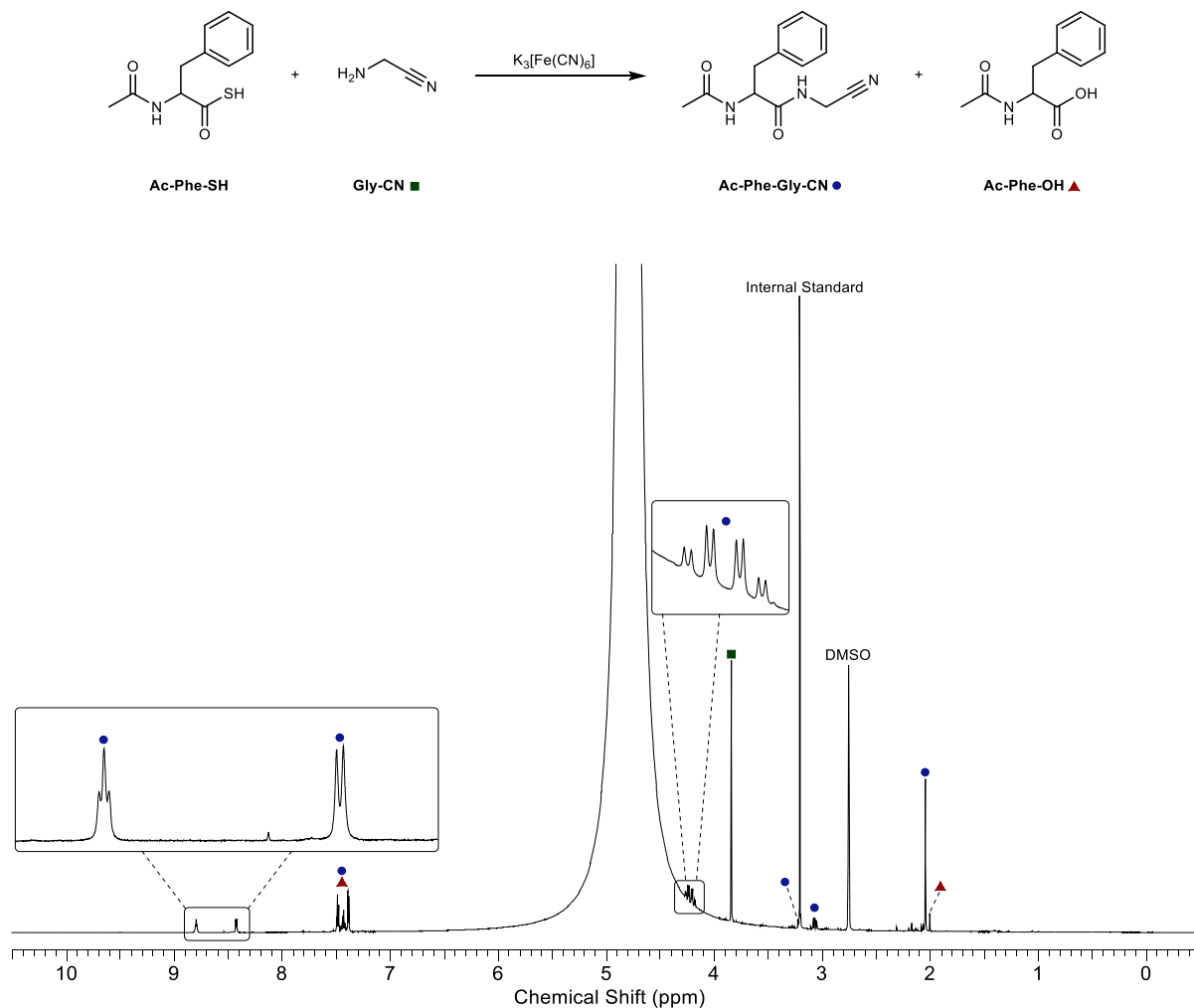
Supplementary Figure 74. ¹H NMR (600 MHz, 1:49:50 D₂O:H₂O:DMSO-*d*₆, -0.50–10.50 ppm) to show the reaction of 2-acetamido-4-(methylthio)butanethioic S-acid (**Ac-Met-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in D₂O/H₂O (2:98, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and diluted with DMSO-*d*₆ (1 mL) for analysis.

¹H NMR 600 MHz, 1:49:50 D₂O:H₂O:DMSO-*d*₆, partial assignment): *N*-Acetylmethionylglycine nitrile **Ac-Met-Gly-CN** (●): δ 9.15 (t, *J* = 5.6 Hz, 1H, Gly-NH), 8.72 (d, *J* = 7.9 Hz, 1H, Met-NH), 4.47 (ABX, *J* = 5.6, 17.7 Hz, 1H, Gly-(C2)-H), 4.43 (ABX, *J* = 5.6, 17.7 Hz, 1H, Gly-(C2)-H'), 2.96 – 2.87 (m, 1H, Met-(C3)-H), 2.86 – 2.80 (m, 1H, Met-(C3')-H), 2.44 (s, 3H, Met-(C4)-SCH₃), 2.41 – 2.37 (m, 2H, Met-(C4)-H₂), 2.32 (s, 3H, (COCH₃)). *N*-Acetylmethionine **Ac-Met-OH** (▲): δ 8.75 (br d, *J* = 7.7 Hz, 1H, NH), 2.96 – 2.87 (m, 1H, (C3)-H), 2.86 – 2.80 (m, 1H, (C3')-H), 2.43 (s, 3H, (C4)-SCH₃), 2.33 (s, 3H, (COCH₃)), 2.38 – 2.30 (m, 1H, (C4)-H), 2.28 – 2.20 (m, 1H, (C4')-H). Glycine nitrile **Gly-CN** (■): δ 3.97 (s, 2H, (C2)-H₂).



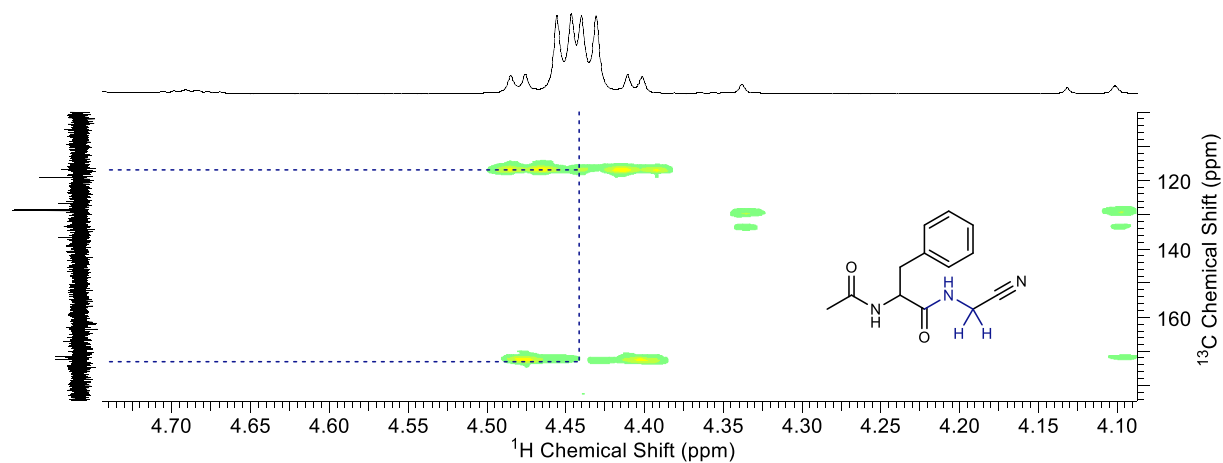
Supplementary Figure 75. ^1H - ^{13}C HMBC (^1H -600 MHz [4.25–4.65 ppm], ^{13}C -151 MHz [100–180 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO}-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)-H and Gly-(C2)-H' of **Ac-Met-Gly-CN** at 4.43 and 4.47 ppm with a CN resonance at 117.7 ppm and a C=O resonance at 173.3 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 74 for expanded and labelled ^1H NMR spectrum.

N-Acetylphenylalanylglycine nitrile **Ac-Phe-Gly-CN**



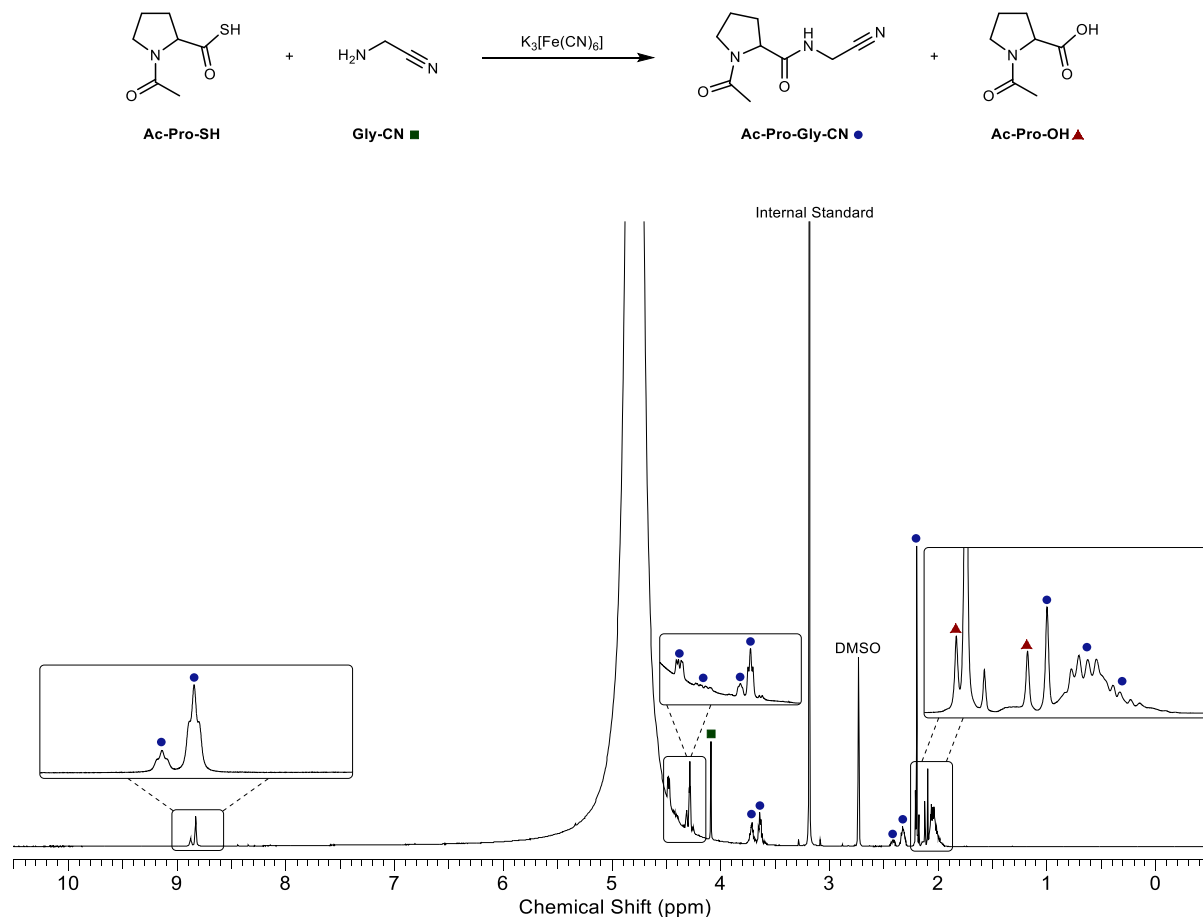
Supplementary Figure 76. ¹H NMR (600 MHz, D₂O/H₂O/DMSO-*d*₆ 1:49:50, -0.50–10.50 ppm) spectrum to show the reaction of 2-acetamido-3-phenylpropanethioic S-acid (**Ac-Phe-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in D₂O/H₂O (2:98, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and diluted with DMSO-*d*₆ (1 mL) for analysis.

¹H NMR (600 MHz, 1:49:50 D₂O:H₂O:DMSO-*d*₆, -0.50–10.50 ppm, partial assignment): *N*-Acetylphenylalanylglycine nitrile **Ac-Phe-Gly-CN** (•): δ 8.80 (t, J = 5.4 Hz, 1H, Gly-NH), 8.43 (d, J = 6.3 Hz, 1H, Phe-NH), 7.51 – 7.37 (m, 5H, Phe-Ar), 4.21 (ABX, J = 5.4, 17.6 Hz, 1H, Gly-(C2)-H), 4.15 (ABX, J = 5.4, 17.6 Hz, 1H, Gly-(C2)-H'), 3.21 (ABX, 1H, J = 6.8, 12.6 Hz, Phe-(C3)-H), 3.07 (ABX, 1H, J = 9.1, 12.6 Hz, Phe-(C3')-H), 2.04 (s, 3H, (COCH₃)). *N*-Acetylphenylalanine **Ac-Phe-OH** (▲): δ 7.51 – 7.37 (m, 5H, Ar), 2.01 (s, 3H, (COCH₃)). Glycine nitrile **Gly-CN** (■): δ 3.84 (s, 2H, (C2)-H₂).



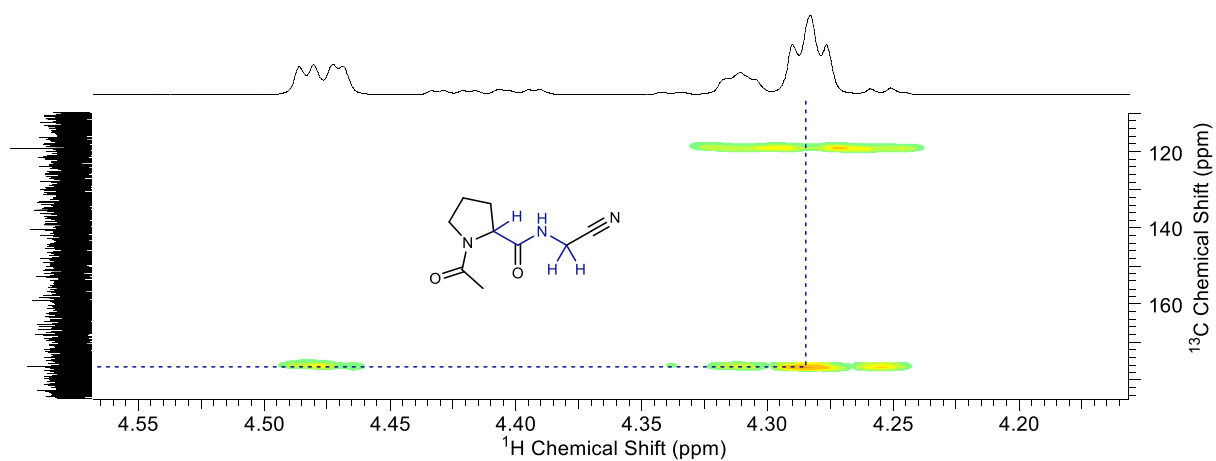
Supplementary Figure 77. ^1H - ^{13}C HMBC (^1H -600 MHz [4.10–4.75 ppm], ^{13}C -151 MHz [100–180 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO}-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)-H and Gly-(C2)-H' of **Ac-Phe-Gly-CN** at 4.15 and 4.21 ppm with a CN resonance at 116.1 ppm and a C=O resonance at 172.5 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 76 for expanded and labelled ^1H NMR spectrum.

N-Acetylprolylglycine nitrile **Ac-Pro-Gly-CN**



Supplementary Figure 78. 1H NMR spectrum (700 MHz, 1:49:50 $D_2O:H_2O:DMSO-d_6$, -0.50–10.50 ppm) for the reaction of 1-acetylpyrrolidine-2-carboxylic S-acid (**Ac-Pro-SH**; 50 mM), glycine nitrile (**Gly-CN**; 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in D_2O/H_2O (2:98, 1 mL) at pH 9.0 and room temperature. After 20 min, the pH of the solution was adjusted to 9.0 and diluted with $DMSO-d_6$ (1 mL) for analysis.

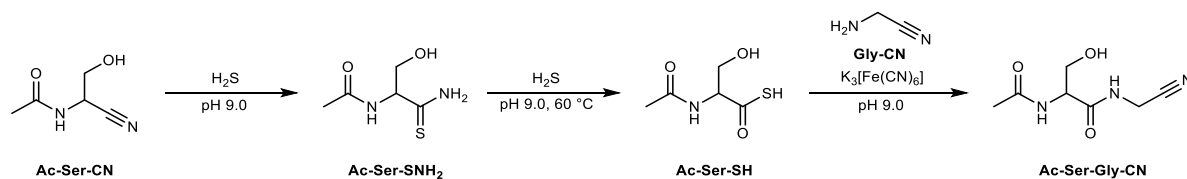
N-Acetylprolylglycine nitrile **Ac-Pro-Gly-CN** 1H NMR (600 MHz, 1:49:50 $D_2O:H_2O:DMSO-d_6$, -0.50–10.50 ppm): *N*-Acetylprolylglycine nitrile **Ac-Pro-Gly-CN** (•, 78:22 mixture of rotamers): major rotamer δ 8.83 (br t, J = 5.6 Hz, 1H, Gly-NH), 4.48 (br dd, J = 5.6, 8.4 Hz, 1H, Pro-(C2)-H), 4.29 (app. br. t, J = 5.0 Hz, 2H, Gly-(C2)-H₂), 3.76 – 3.67 (m, 1H, Pro-(C5)-H), 3.66 – 3.61 (m, 1H, Pro-(C5)-H'), 2.36 – 2.27 (m, 1H, Pro-(C3)-H), 2.20 (s, 3H, (COCH₃), 2.09 – 1.97 (m, 3H, Pro-(C3)-H'); Pro-(C4)-H₂); minor rotamer δ 8.88 (br t, J = 5.1 Hz, 1H, Gly-NH), 4.44 – 4.39 (m, 1H, Pro-(C2)-H), 4.31 (app. br. t, 2H, Gly-(C2)-H₂), 3.76 – 3.67 (m, 1H, Pro-(C5)-H), 3.66 – 3.61 (m, 1H, Pro-(C5)-H'), 2.45 – 2.38 (m, 1H, Pro1-(C3)-H), 2.10 (s, 3H, (COCH₃), 2.09 – 1.97 (m, 3H, Pro-(C3)-H'); (Pro)-(C4)-H₂). *N*-Acetylproline **Ac-Pro-OH** (▲, 55:45 mixture of rotamers): major rotamer (partial assignment) δ 2.21 (s, 3H, (COCH₃); minor rotamer (partial assignment) δ 2.11 (s, 3H, (COCH₃). Glycine nitrile **Gly-CN** (■): δ 4.09 (s, 3H, (C2)-H₂).



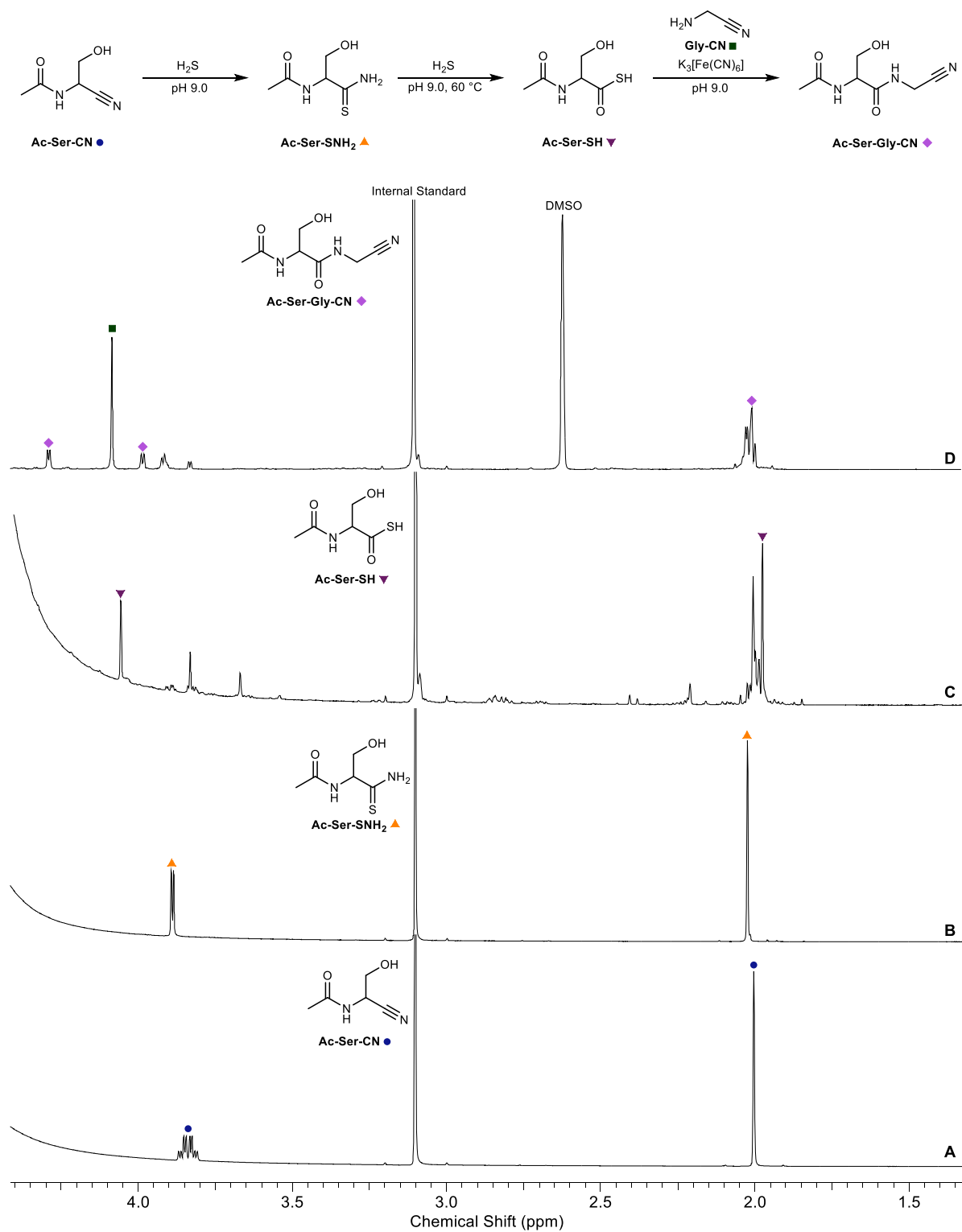
Supplementary Figure 79. ^1H - ^{13}C HMBC (^1H -700 MHz [4.10–4.55 ppm], ^{13}C -151 MHz [100–180 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO}-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)- H_2 of **Ac-Pro-Gly-CN** (major rotamer) at 4.29 ppm with a CN resonance at 118.8 ppm and a C=O resonance at 176.2 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 78 expanded and labelled ^1H NMR spectrum.

N-Acetylserylglycine nitrile **Ac-Ser-Gly-CN**

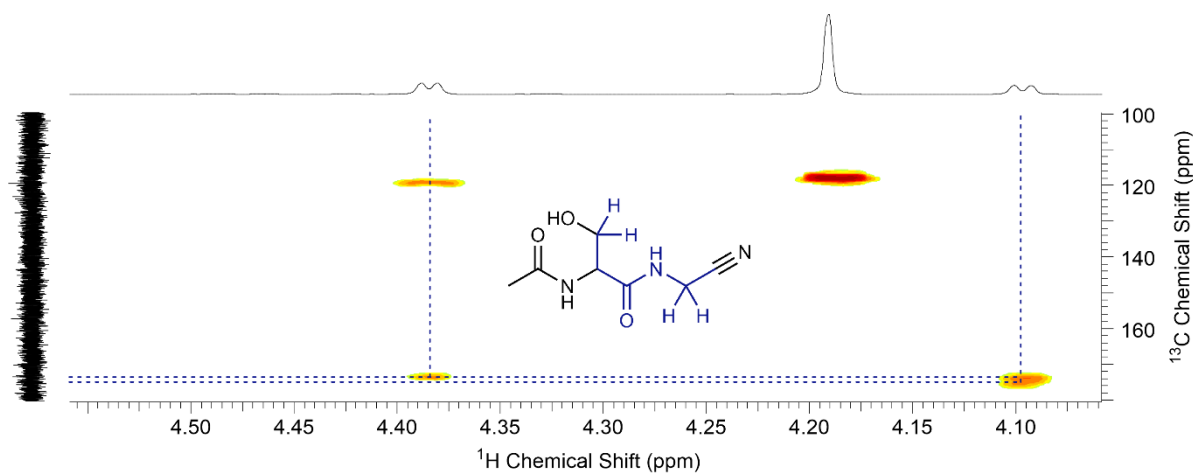
Sequential thiolysis, hydrolysis and oxidative ligation of *N*-acetylserine nitrile with glycine nitrile



N-Acetylserine nitrile **Ac-Ser-CN** (6.42 mg, 0.05 mmol), $\text{NaSH} \cdot x\text{H}_2\text{O}$ (56.1 mg, 0.50 mmol) and methylsulfonylmethane (0.05 mmol) were dissolved in degassed $\text{H}_2\text{O}/\text{D}_2\text{O}$ (98:2, 1 mL), and the solution was adjusted to pH 9.0 with NaOH/HCl . The solution was incubated at room temperature whilst maintaining the solution at pH 9.0 with NaOH/HCl . After 24 h, NMR analysis showed quantitative conversion of **Ac-Ser-CN** to *N*-acetylserine thioamide **Ac-Ser-SNH₂**, which was confirmed by ^1H - ^{13}C HMBC NMR analysis. The yield was quantified using the methylsulfonylmethane signal as an internal standard (99%). The resulting solution was adjusted to pH 9.0 with NaOH/HCl and incubated at 60 °C for a further 24 h. NMR analysis showed complete consumption of *N*-acetylserine thioamide **Ac-Ser-SNH₂**. The presence of 2-acetamido-3-hydroxypropanethioic *S*-acid **Ac-Ser-SH** was confirmed by ^1H - ^{13}C HMBC NMR analysis. The yield was quantified using the methylsulfonylmethane signal as an internal standard (61%). The solution was sparged with argon for 30 min whilst maintaining the solution at pH 7.0 with NaOH/HCl . Glycine nitrile **Gly-CN** (0.061 mmol) was added, and the solution was adjusted to pH 9.0 with NaOH/HCl , followed by the addition of potassium hexacyanoferrate(III) (30.13 mg, 0.092 mmol). The solution was stirred at room temperature for 20 minutes. The resulting suspension was then adjusted to pH 9.0 and cooled to 0 °C before dilution with $\text{DMSO}-d_6$ (1 mL). NMR spectra were acquired and the presence of *N*-acetylserylglycine nitrile **Ac-Ser-Gly-CN** was confirmed by ^1H - ^{13}C HMBC NMR analysis and high resolution mass spectrometry. The yield of **Ac-Ser-Gly-CN** was determined by integration against the methylsulfonylmethane signal.



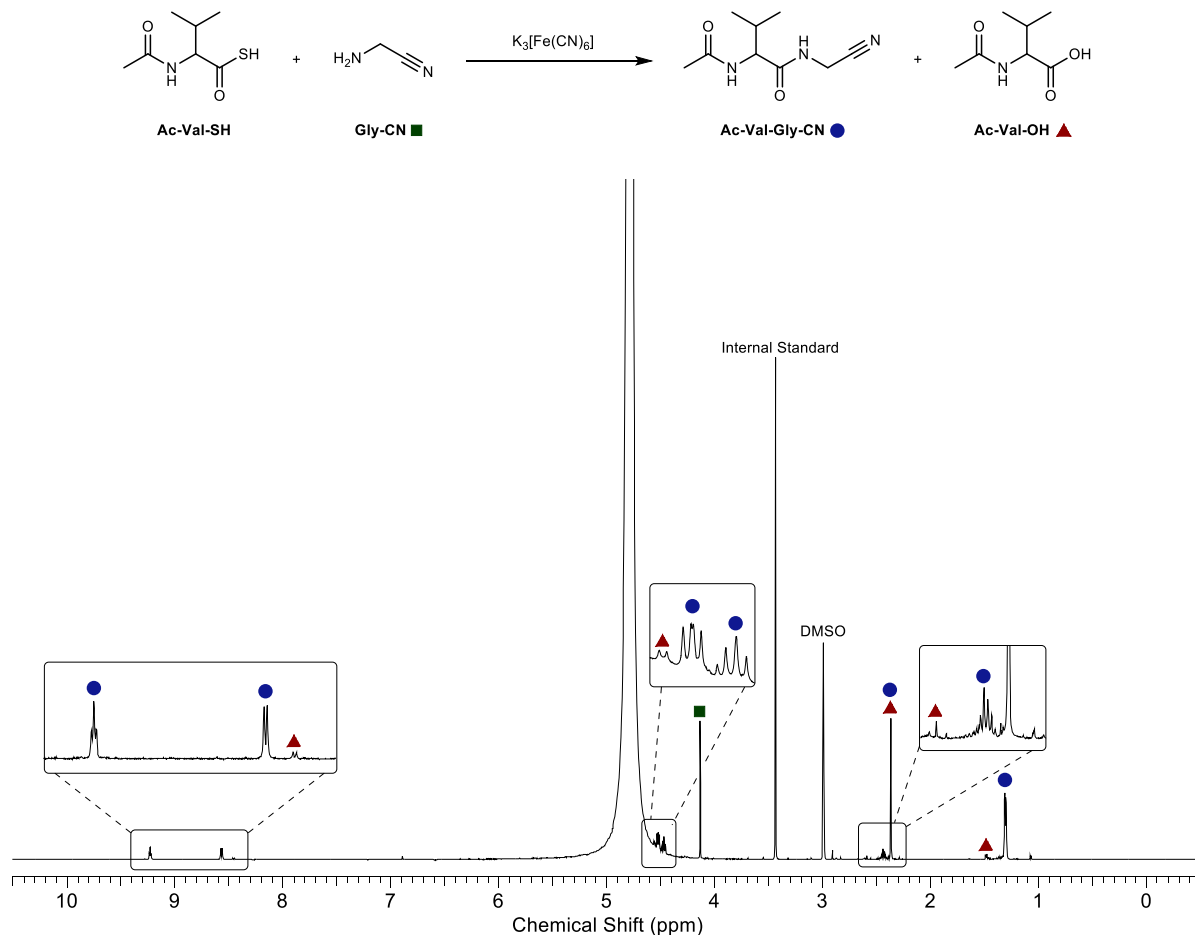
Supplementary Figure 80. ¹H NMR (700 MHz, pH 9.0, 1.35–4.40 ppm) spectra to show the sequential thiolysis, hydrolysis and oxidative ligation with glycine nitrile **Gly-CN** of *N*-acetylserine nitrile **Ac-Ser-CN**. **A:** pure **Ac-Ser-CN**, H₂O/D₂O 98:2; **B:** Thiolysis of **Ac-Ser-CN** to **Ac-Ser-SNH₂**, H₂O/D₂O 98:2; **C:** Hydrolysis of **Ac-Ser-SNH₂** to **Ac-Ser-SH**, H₂O/D₂O 98:2; **D:** Coupling of **Ac-Ser-SH** with **Gly-CN**, D₂O/H₂O/DMSO-*d*₆, noesygppr1d, 1:49:50.



Supplementary Figure 81. ^1H - ^{13}C HMBC (^1H -700 MHz [4.05–4.55 ppm], ^{13}C -151 MHz [100–180 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO-}d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of Gly-(C2)-H₂ of **Ac-Ser-Gly-CN** at 4.38 ppm with a CN resonance at 119.1 ppm and a C=O resonance at 174.6 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 80D expanded and labelled ^1H NMR spectrum.

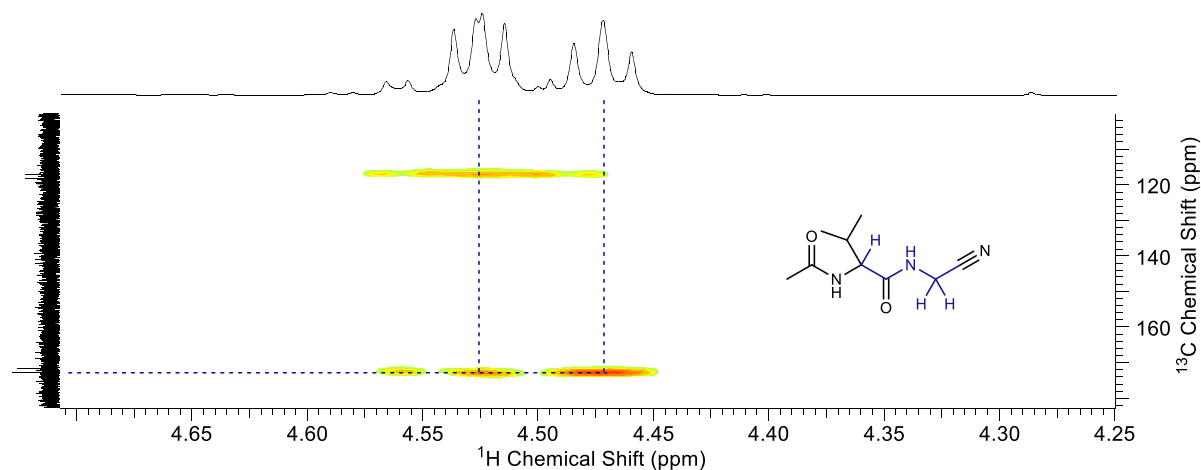
^1H NMR (700 MHz, pH 9.0) data for: *Supplementary Figure 80A* (pure starting material, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, partial assignment): *N*-Acetylserine nitrile **Ac-Ser-CN** (●) δ 3.85 (ABX, J = 5.3, 11.2 Hz, 1H, (C2)-H), 3.83 (ABX, J = 5.3, 11.2 Hz, 1H, (C2)-H'), 2.02 (s, 3H, (COCH₃)). *Supplementary Figure 80B* (thiolysis, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, pH 9.0, partial assignment): *N*-Acetylserine thioamide **Ac-Ser-SNH₂** (▲) δ 3.86 (d, J = 5.4 Hz, 2H, (C2)-H₂), 1.97 (s, 3H, (COCH₃)). *Supplementary Figure 80C* (hydrolysis, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, partial assignment): 2-Acetamido-3-hydroxypropanethioic S-acid **Ac-Ser-SH** (▼) δ 4.08 (br s, 2H, (C2)-H₂), 1.99 (s, 3H, (COCH₃)). *Supplementary Figure 80D* (ligation, $\text{H}_2\text{O}/\text{D}_2\text{O}/\text{DMSO-}d_6$ 49:1:50, partial assignment): *N*-Acetylseryl glycine nitrile **Ac-Ser-Gly-CN** (◆) δ 4.23 (br d, 5.0 Hz, 2H, Gly-(C2)-H₂), 3.96 (d, J = 5.8 Hz, 2H, Ser-(C2)-H₂), 1.98 (s, 3H, (COCH₃)); glycine nitrile (■) δ 3.95 (s, 2H, (C2)-H₂).

N-Acetylvalylglycine nitrile **Ac-Val-Gly-CN**



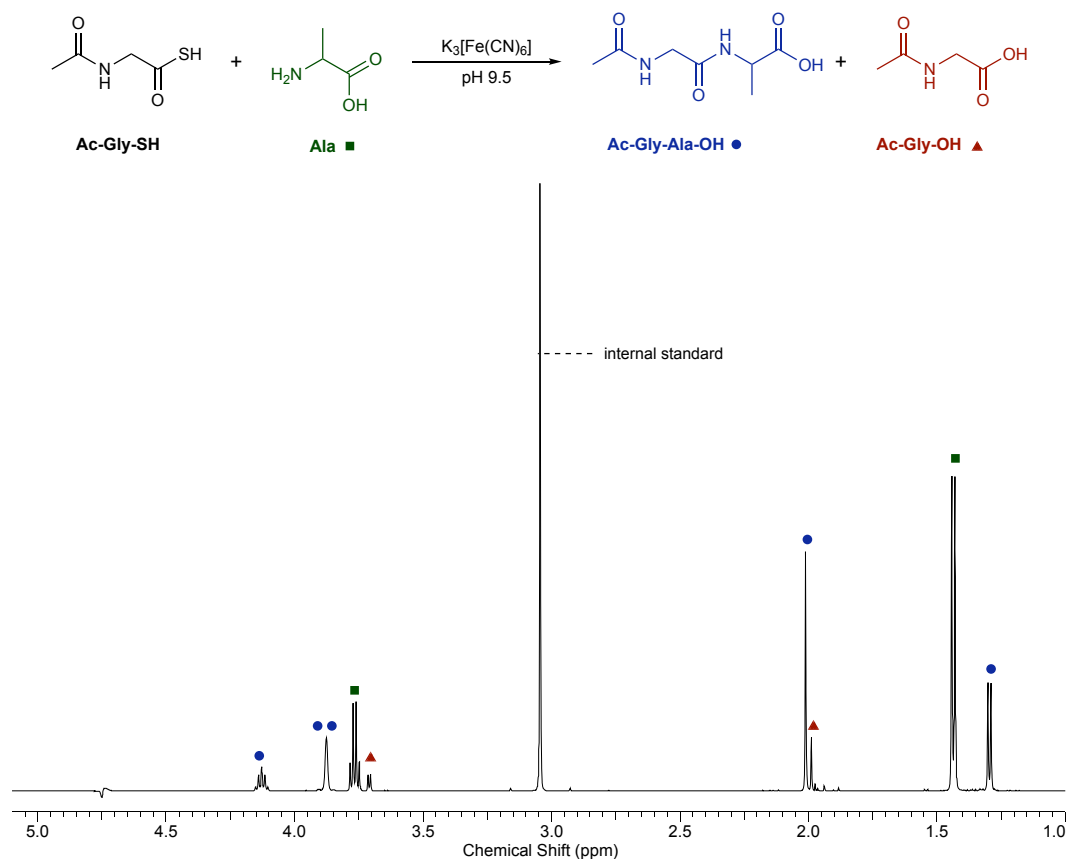
Supplementary Figure 82. 1H NMR (600 MHz, $D_2O/H_2O/DMSO-d_6$ 1:49:50, -0.50–10.50 ppm) spectrum showing the reaction of 2-acetamido-3-methylbutanethioic S-acid (**Ac-Val-SH**; 50 mM), glycine nitrile (**Gly-CN**; 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) in D_2O/H_2O (2:98, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and diluted with $DMSO-d_6$ (1 mL) for analysis.

1H NMR (600 MHz, $D_2O/H_2O/DMSO-d_6$ 1:49:50): *N*-Acetylvalylglycine nitrile **Ac-Val-Gly-CN** (•): δ 9.23 (br t, J = 5.6 Hz, 1H, Gly-NH), 8.57 (d, J = 7.9 Hz, 1H, Val-NH), 4.52 (dd, J = 5.6, 7.6 Hz, 1H, Val-(C2)-H), 4.47, (app. t, J = 7.6 Hz, 2H, Gly-(C2)-H₂), 2.48 – 2.41 (m, 1H, Val-(C3)-H), 2.37 (s, 3H, (COCH₃)), 1.33 – 1.28 (m, 6H, Val1-(C4)-H₃, Val1-(C4')-H₃). *N*-Acetylvaline **Ac-Val-OH** (▲): δ 8.46 (br d, J = 8.0 Hz, 1H, NH), 4.56, (app. d, J = 6.2 Hz, 1H, (C2)-H), 2.63 – 2.57 (m, 1H, (C3)-H), 2.37 (s, 3H, (COCH₃)), 1.50 – 1.47 (m, 6H, (C4)-H₃, (C4')-H₃). Glycine nitrile **Gly-CN** (■): δ 4.14 (s, 2H, (C2)-H₂).



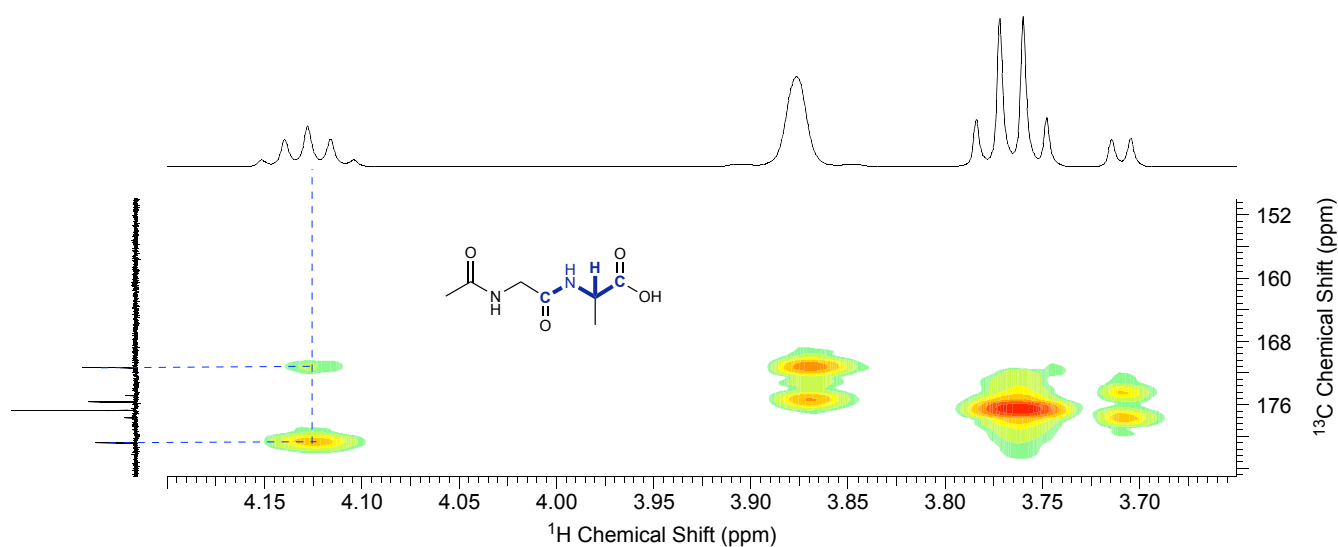
Supplementary Figure 83. 1H - ^{13}C HMBC (1H -600 MHz [4.25–4.70 ppm], ^{13}C -151 MHz [110–180 ppm], $D_2O/H_2O/DMSO-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{CH}$ and $^2J_{CH}$ coupling of Gly-(C2)-H of **Ac-Val-Gly-CN** at 4.52 ppm with a CN resonance at 118.3 ppm and a C=O resonance at 174.8 ppm, which is characteristic of peptide bond formation of **Gly-CN**. See Supplementary Figure 82 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylalanine **Ac-Gly-Ala-OH**



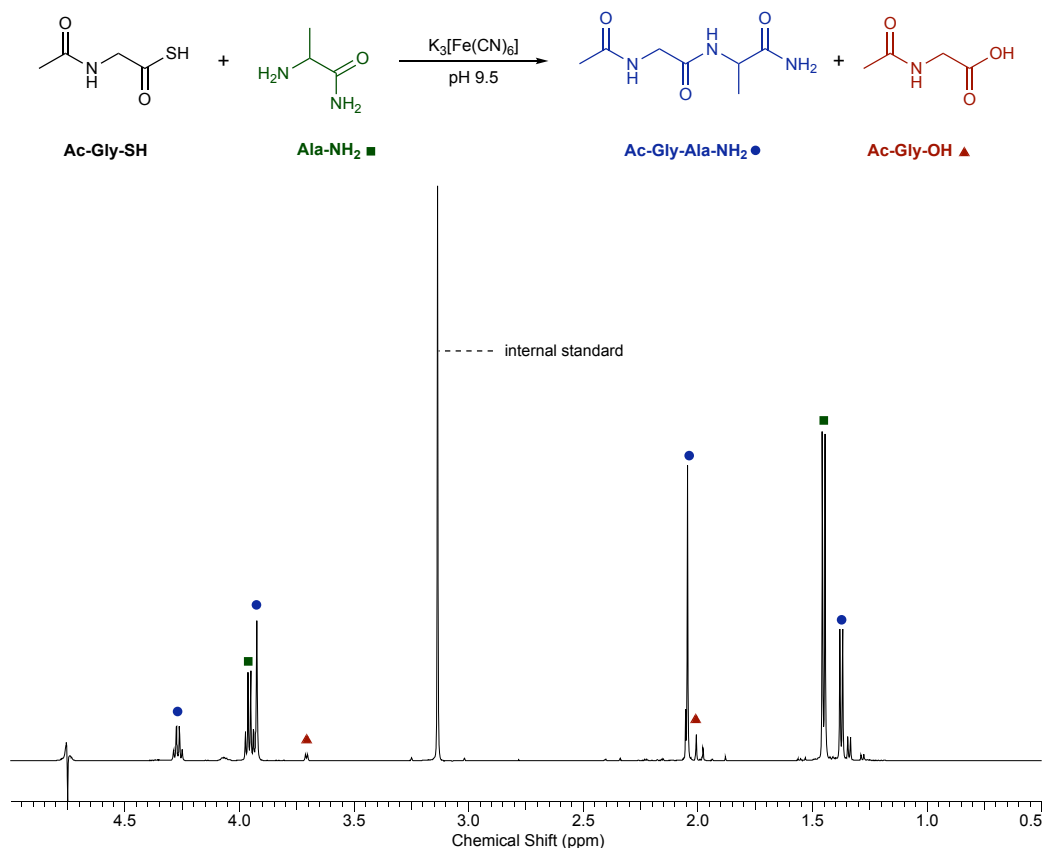
Supplementary Figure 84. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.0–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-alanine (**Ala**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylalanine **Ac-Gly-Ala-OH** (•): δ 4.13 (app. qn, $J = 7.2$ Hz, 1H, Ala-(C2)–H), 3.89 (AB, $J = 16.8$ Hz, 1H, Gly-(C2)–H), 3.85 (AB, $J = 16.8$ Hz, 1H, Gly-(C2')–H), 2.01 (s, 3H, (COCH₃)), 1.30 (d, $J = 7.3$ Hz, 3H, Ala-(C3)–H₃). L-Alanine **Ala** (■): δ 3.77 (q, $J = 7.1$ Hz, 1H, (C2)–H), 1.35 (d, $J = 7.1$ Hz, 3H, (C3)–H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.70 (app. d, $J = 5.8$ Hz, 2H, (C2)–H₂), 1.99 (s, 3H, (COCH₃)).



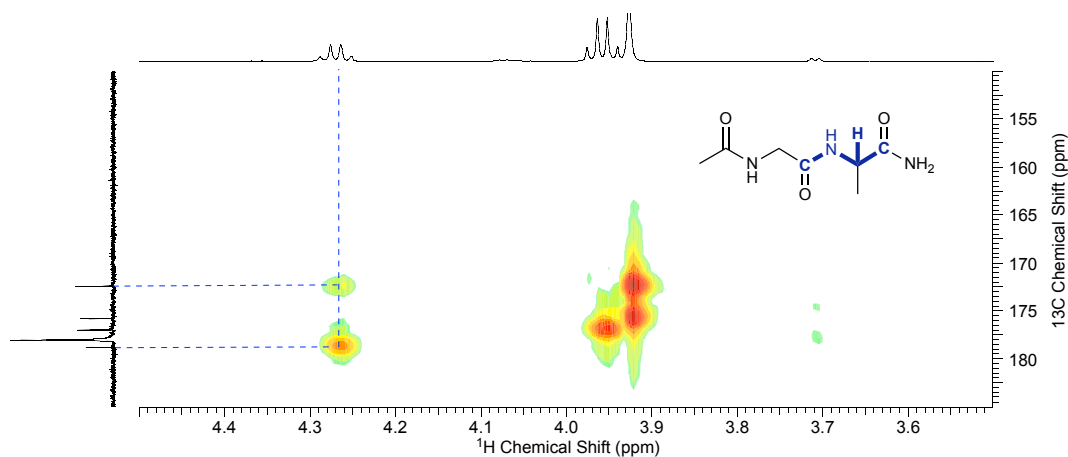
Supplementary Figure 85. ¹H–¹³C HMBC (¹H–600 MHz [3.65–5.00 ppm], ¹³C–151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ²J_{CH} and ³J_{CH} coupling of Ala-(C2)–H of **Ac-Gly-Ala-OH** at 4.13 ppm with two C=O resonances at 171.3 and 180.9 ppm, which is characteristic of peptide bond formation of **Ala**. See Supplementary Figure 84 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylalaninamide **Ac-Gly-Ala-NH₂**



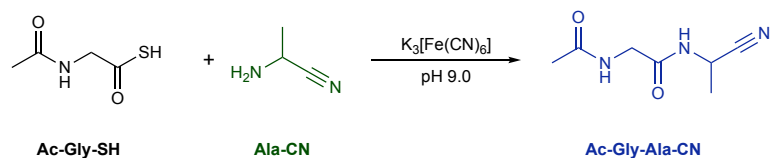
Supplementary Figure 86. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-alaninamide (**Ala-NH₂**; 150 mM) and K₃[Fe(CN)₆] (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylalaninamide **Ac-Gly-Ala-NH₂** (●): δ 4.27 (q, *J* = 7.3 Hz, 1H, Ala-(C2)–H), 3.93 (s, 2H, Gly-(C2)–H₂), 2.05 (s, 3H, (COCH₃)), 1.37 (d, *J* = 7.3 Hz, 3H, Ala-(C3)–H₃). *L*-Alaninamide **Ala-NH₂** (■): δ 3.96 (q, *J* = 7.1 Hz, 1H, (C2)–H), 1.45 (d, *J* = 7.1 Hz, 3H, (C3)–H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.71 (app. d, *J* = 5.5 Hz, 2H, (C2)–H₂), 2.01 (s, 3H, (COCH₃)).

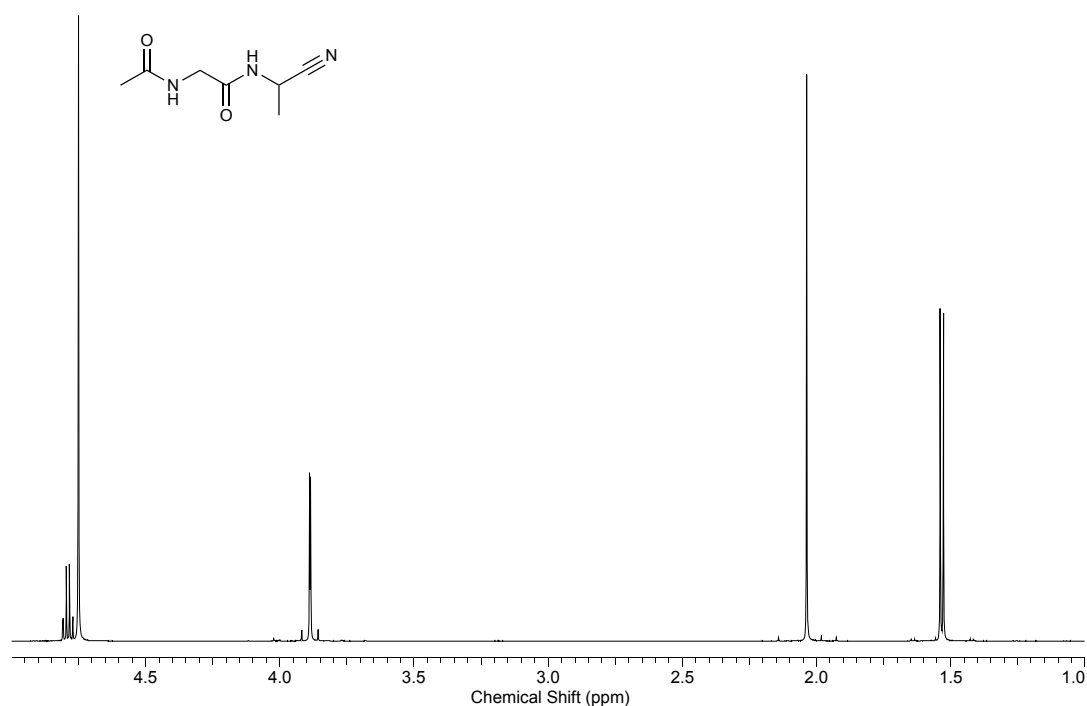


Supplementary Figure 87. ¹H–¹³C HMBC (¹H–600 MHz [3.5–4.5 ppm], ¹³C–151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ²*J*_{CH} and ³*J*_{CH} coupling of Ala-(C2)–H of **Ac-Gly-Ala-NH₂** at 4.27 ppm with two C=O resonances at 172.4 and 178.8 ppm, which is characteristic of peptide bond formation of **Ala-NH₂**. See Supplementary Figure 86 for expanded and labelled ¹H NMR spectrum.

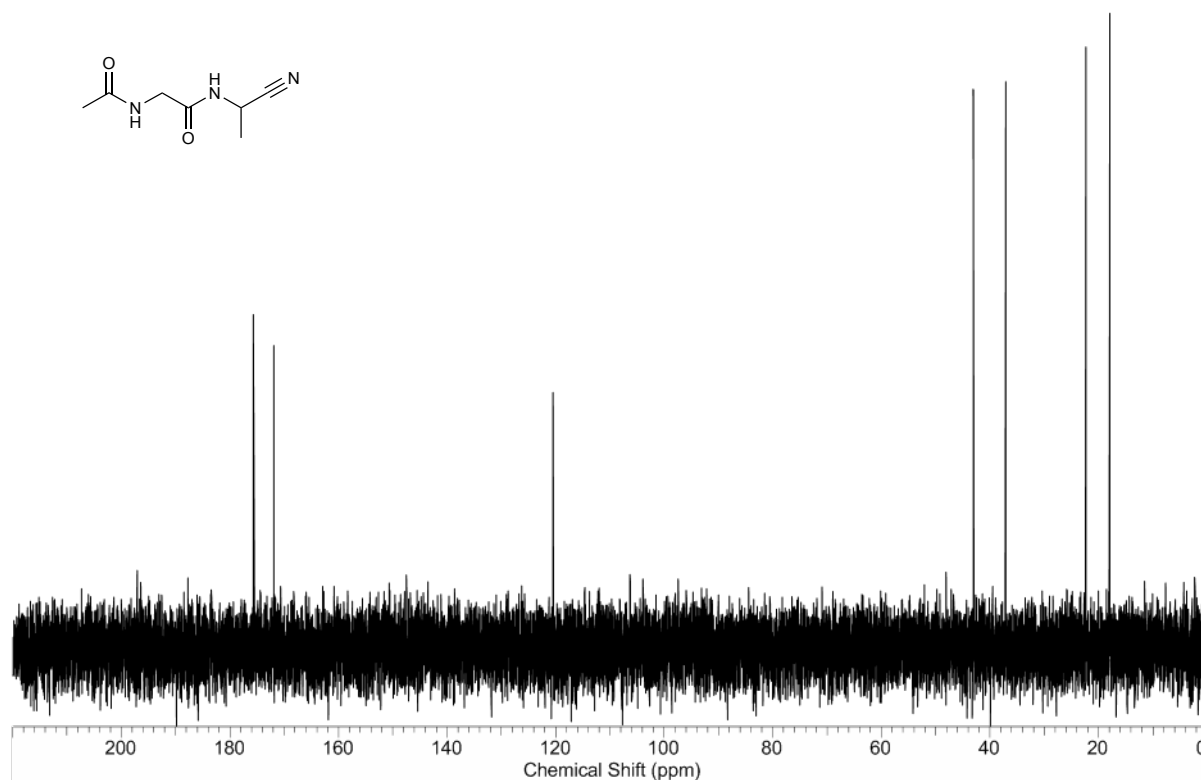
N-Acetylglycylalanine nitrile **Ac-Gly-Ala-CN**



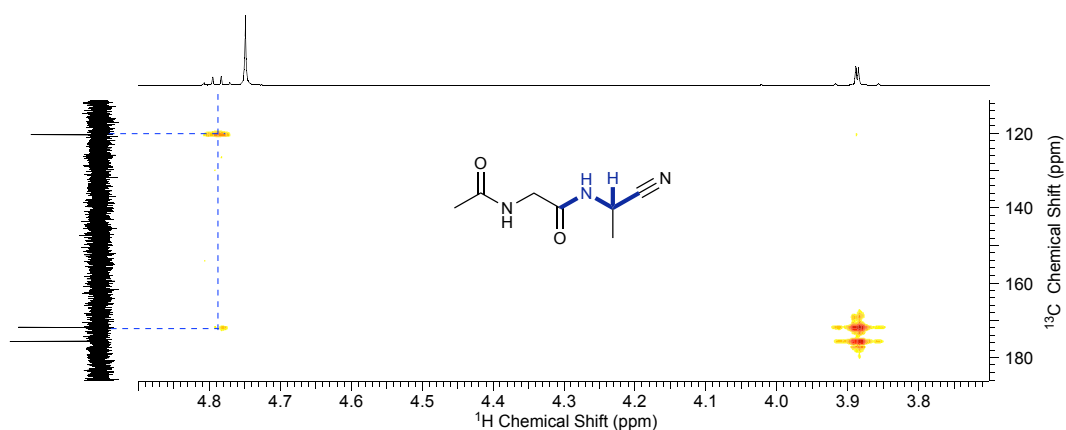
Prepared following *general procedure G* on a 0.1 mmol scale, from sodium-2-acetamido-ethanethioate **Ac-Gly-S⁻Na⁺** (15.5 mg, 0.1 mmol) and *rac*-alanine nitrile hydrochloride **Ala-CN**·HCl³⁹ (21.3 mg, 0.2 mmol). **Ac-Gly-Ala-CN** was isolated as a white solid (15.2 mg, 90%) after flash column chromatography (SiO₂, EtOAc (100%)). *R_f* 0.16 (100% EtOAc). ¹H NMR (600 MHz, D₂O) δ 4.79 (q, *J* = 7.2 Hz, 1H, Ala-(C2)-H), 3.90 (AB, *J* = 17.0 Hz, 1H, Gly-(C2)-H), 3.87 (d, *J* = 17.0 Hz, 1H, Gly-(C2)-H'), 2.04 (s, 3H, (COCH₃), 1.53 (d, *J* = 7.2 Hz, 3H, Ala-(C3)-H₃). ¹³C NMR (151 MHz, D₂O) δ 175.6 (COCH₃), 171.9 (Gly-(C1)), 120.5 (Ala-(C1)), 43.1 (Gly-(C2)), 37.1 (Ala-(C2)), 22.4 (COCH₃), 18.0 (Ala-(C3)). HRMS-ESI-MS [*M*+H]⁺ calcd. for C₇H₁₂N₃O₂ 170.0924; obs. 170.0924. IR (solid, cm⁻¹) 2463, 2360, 2331, 1672, 1629, 1556.



Supplementary Figure 88. ¹H NMR (600 MHz, D₂O, 1.0–5.0 ppm) spectrum of purified *N*-acetylglycylalanine nitrile **Ac-Gly-Ala-CN**



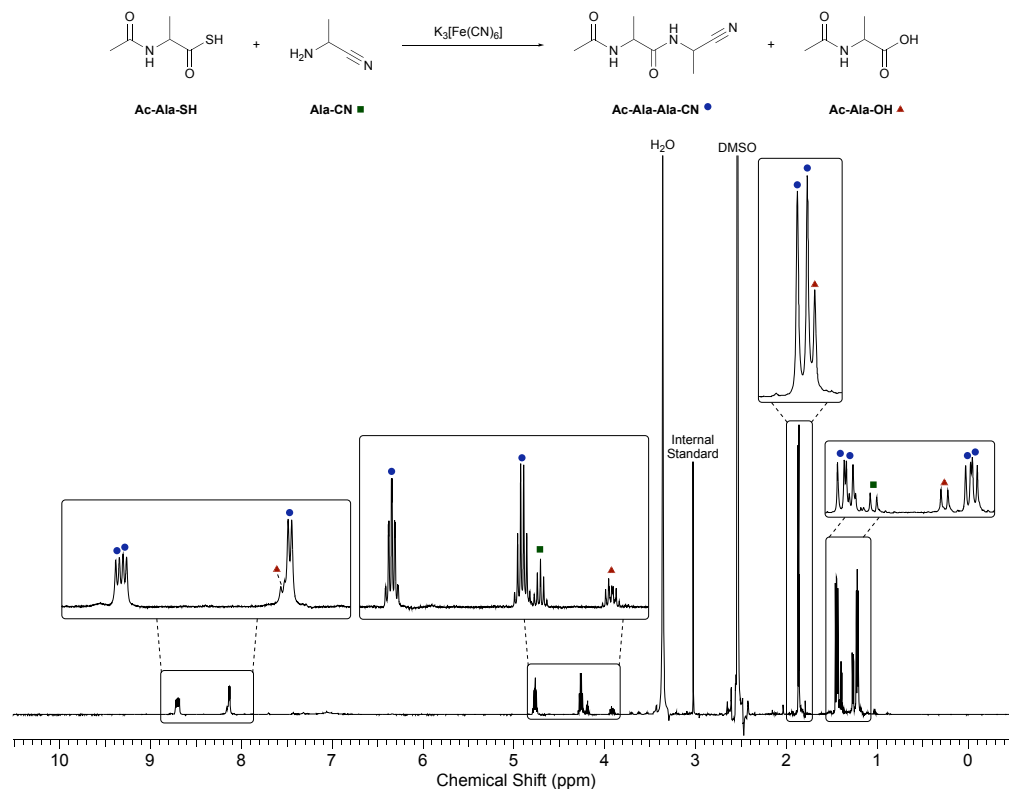
Supplementary Figure 89. ^{13}C NMR (151 MHz, D_2O , 0–220 ppm) spectrum of purified *N*-acetylglycylalanine nitrile **Ac-Gly-Ala-CN**



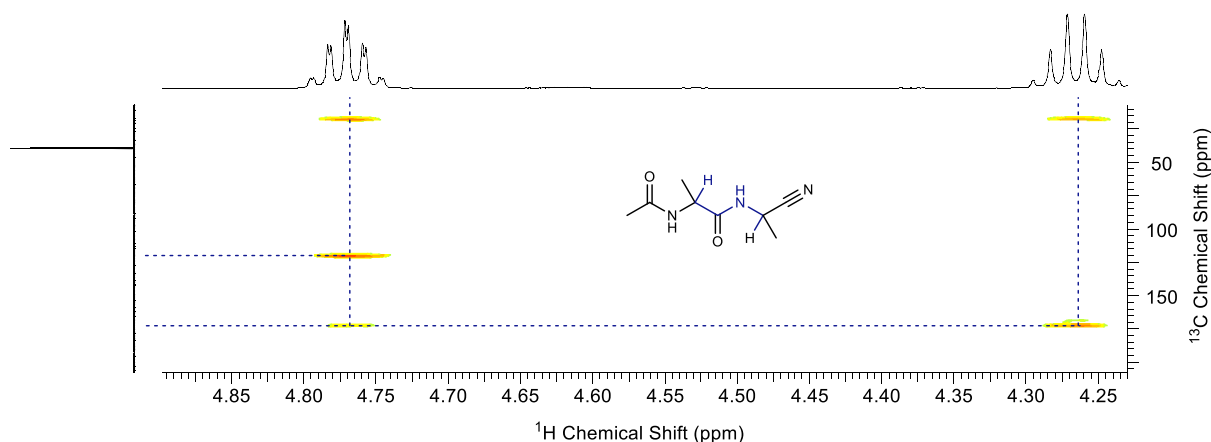
Supplementary Figure 90. ^1H – ^{13}C HMBC (^1H –600 MHz [3.7–4.9 ppm], ^{13}C –151 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Ala-(C2)–H of **Ac-Gly-Ala-CN** at 4.79 ppm with CN resonance at 120.5 ppm and a C=O resonance at 171.9 ppm, which is characteristic of peptide bond formation of **Ala-CN**. See Supplementary Figure 88 for expanded ^1H NMR spectrum.

N-Acetylalanylalanine nitrile **Ac-Ala-Ala-CN**

2-Acetamidopropanethioic S-acid **Ac-Ala-SH** (50 mM, 0.05 mmol), alanine nitrile hydrochloride **Ala-CN**·HCl³⁹ (100 mM, 0.1 mmol) and methylsulfonylmethane (50 mM, 0.05 mmol) were dissolved in H₂O/D₂O (98:2, 1 mL), and the solution was adjusted to pH 9.0 with NaOH/HCl. Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added, and the solution was stirred at room temperature for 20 minutes. The resulting suspension was then adjusted to pH 9.0 and lyophilised. The resulting lyophilisate was dissolved in DMSO-*d*₆ (1 mL) and NMR spectra were acquired. The presence of **Ac-Ala-Ala-CN** was confirmed by ¹H–¹³C HMBC NMR spectral analysis and high resolution mass spectrometry. The reaction mixture was quantified using methylsulfonylmethane (MSM) as an internal standard. Yields are given in Supplementary Table 7.



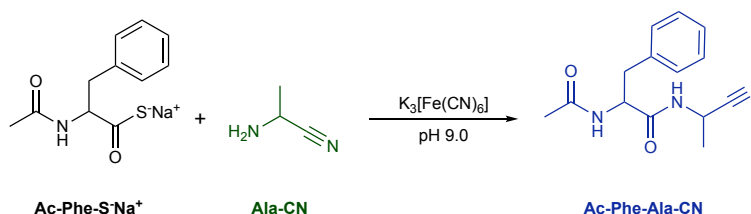
Supplementary Figure 91. ¹H NMR (600 MHz, DMSO-*d*₆, -0.50–10.50 ppm) spectrum to show the reaction of 2-Acetamidopropanethioic S-acid (**Ac-Ala-SH**; 50 mM), alanine nitrile (**Ala-CN**, 100 mM) and K₃[Fe(CN)₆] (150 mM) in D₂O/H₂O (2:98, 1 mL) at pH 9.0 and room temperature. After 20 min, the solution was adjusted to pH 9.0 and diluted with DMSO-*d*₆ (1 mL) for analysis.



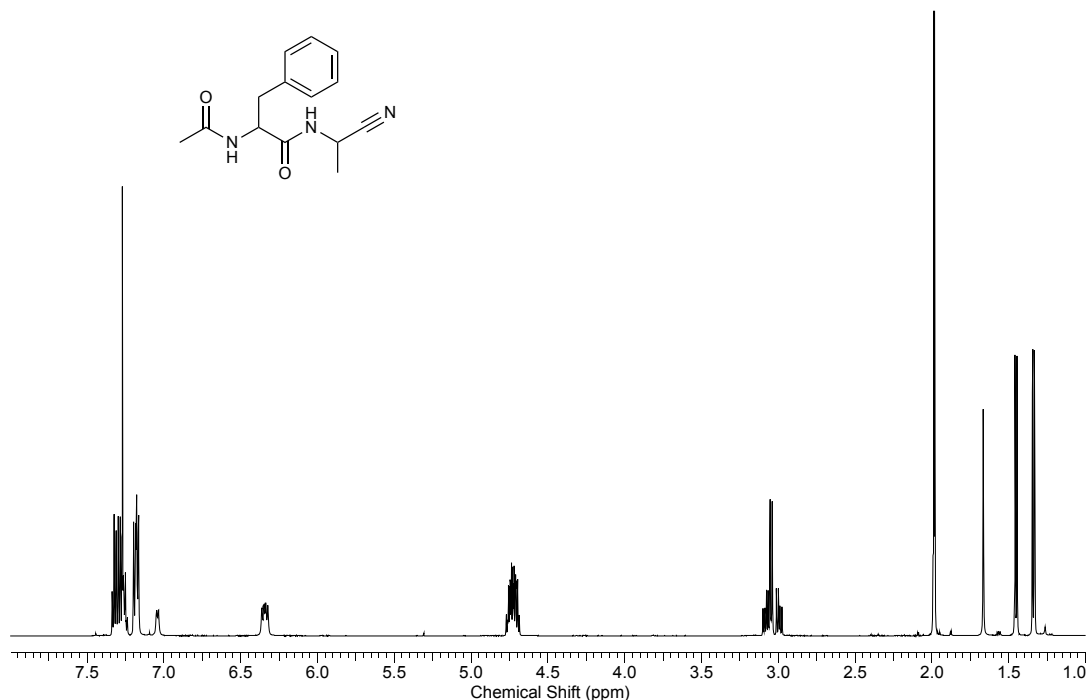
Supplementary Figure 92. ¹H–¹³C HMBC (¹H–600 MHz [4.25–4.90 ppm], ¹³C–151 MHz [20–200 ppm], DMSO-*d*₆) spectrum showing the diagnostic ²J_{CH} and ²J_{CH} coupling of Ala₂–(C2)–H of **Ac-Ala-Ala-CN** at 4.77 ppm with a CN resonance at 118.6 ppm and a C=O resonance at 174.4 ppm, which is characteristic of peptide bond formation of **Ala-CN**. See Supplementary Figure 91 for expanded and labelled ¹H NMR spectrum.

^1H NMR (600 MHz, DMSO- d_6): *N*-Acetylalanylalanine nitrile **Ac-Ala-Ala-CN** (•, 50:50 mixture of diastereoisomers a/b): diastereoisomer a: δ 8.72 (d, J = 7.2 Hz, 1H, Ala₂-NH), 8.14 (d, J = 7.2, 1H, Ala₁-NH), 4.77 (app. quint, J = 7.2 Hz, 1H, Ala₂-(C2)-H), 4.27 (app. quint., J = 7.2 Hz, 1H, Ala₁-(C2)-H), 1.88 (s, 3H, (COCH₃), 1.45 (d, J = 7.2 Hz, 3H, Ala₂-(C3)-H₃), 1.23 (d, J = 7.2 Hz, 3H, Ala₁-(C3)-H₃); diastereoisomer b: δ 8.69 (d, J = 7.2 Hz, 1H, Ala₂-NH), 8.14 (d, J = 7.2, 1H, Ala₁-NH), 4.77 (app. quint. J = 7.2, 1.3 Hz, 1H, Ala₂-(C2)-H), 4.27 (app. quint. J = 7.2 Hz, 1H, Ala₁-(C2)-H), 1.87 (s, 3H, (COCH₃), 1.46 (d, J = 7.2 Hz, 3H, Ala₂-(C3)-H₃), 1.22 (d, J = 7.2 Hz, 3H, Ala₁-(C3)-H₃). *N*-Acetylalanine **Ac-Ala-OH** (▲): δ 8.16 (obs. d, J = 7.8 Hz, 1H, NH), 3.95 – 3.89 (m, 1H, (C2)-H), 1.86 (s, 3H, (COCH₃), 1.27 (d, J = 7.8 Hz, 3H, (C3)-H₃). Alanine nitrile **Ala-CN** (■): δ 4.20 (app. quint., J = 7.3 Hz, 1H, (C2)-H), 1.37 (d, J = 7.3 Hz, 3H, (C3)-H₃).

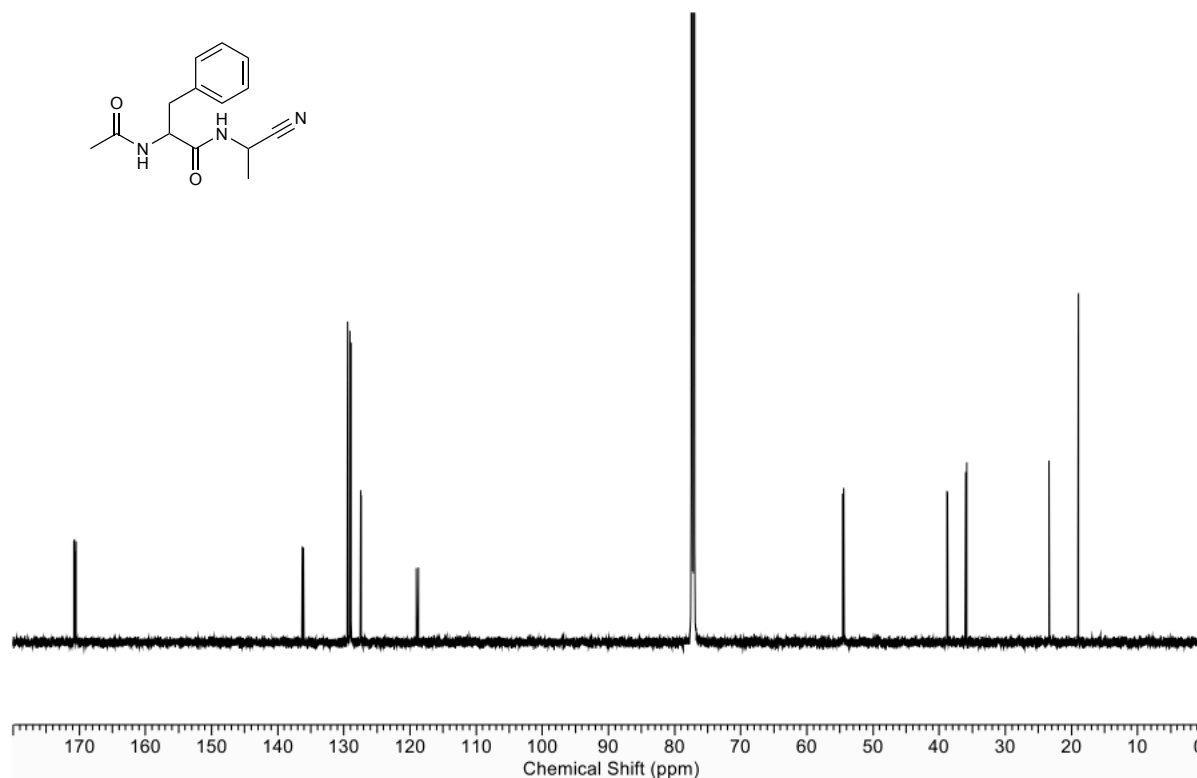
N-Acetylphenylalanylalanine nitrile **Ac-Phe-Ala-CN**



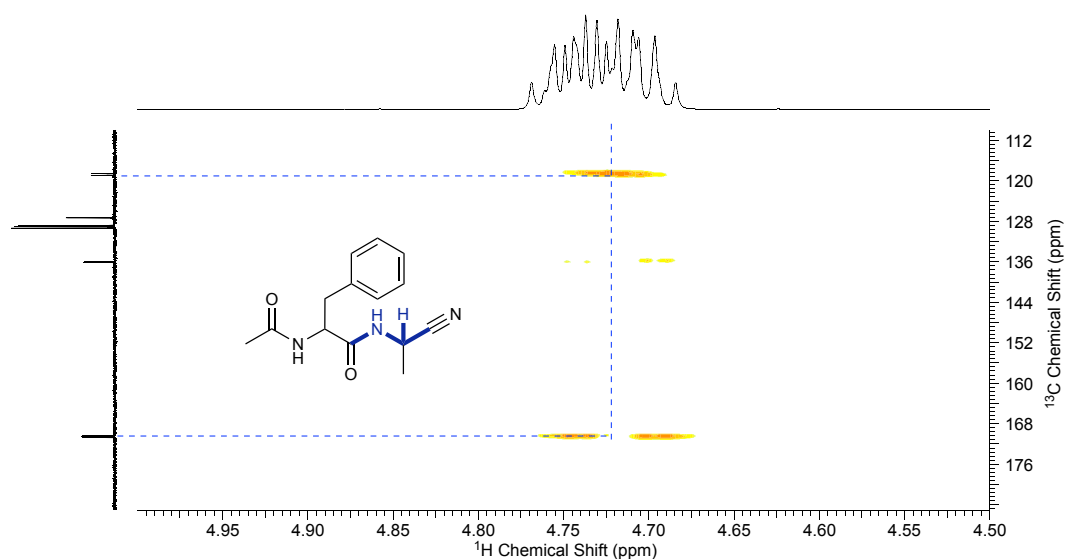
Prepared following *general procedure G* from sodium-2-acetamido-3-phenylpropanethioate **Ac-Phe-S⁻Na⁺** (123 mg, 0.5 mmol) and *rac*-alanine nitrile hydrochloride **Ala-CN**·HCl³⁹ (106 mg, 1.0 mmol). **Ac-Phe-Ala-CN** was isolated as a white solid (91 mg, 71%) after flash column chromatography (SiO₂; CH₂Cl₂/MeOH (100:0 to 95:5)). *R_f* 0.31 (CH₂Cl₂/MeOH, 95:5). ¹H NMR (600 MHz, CDCl₃, 1:1 mixture of diastereoisomers) δ 7.36 - 7.23 (m, 6H, Ar), 7.19 - 7.17 (obs, 1H, Phe-NH), 7.22 - 7.14 (m, 4H, Ar), 7.04 (d, *J* = 7.1 Hz, 1H, Phe-NH), 6.35 (d, *J* = 8.0 Hz, 1H, Ala-NH), 6.33 (d, *J* = 8.1 Hz, 1H, Ala-NH), 4.78 - 4.67 (m, 4H, Phe-(C2)-H x 2, Ala-(C2)-H x 2), 3.08 (ABX, *J* = 6.7, 13.7 Hz, 1H, Phe-(C3)-H), 3.05 (d, *J* = 7.2 Hz, 2H, Phe-(C3)-H₂), 2.99 (ABX, *J* = 8.2, 13.7 Hz, 1H, Phe-(C3')-H), 1.99 (s, 3H, (COCH₃)), 1.98 (s, 3H, (COCH₃)), 1.45 (d, *J* = 7.2 Hz, 3H, Ala-(C3)-H₃), 1.34 (d, *J* = 7.2 Hz, 3H, Ala-(C3)-H₃). ¹³C NMR (151 MHz, CDCl₃, 1:1 mixture of diastereoisomers) δ 170.6 (CO), 170.5 (CO), 170.4 (CO), 170.3 (CO), 136.1 (Ar), 135.9 (Ar), 129.3 (Ar x 2), 129.2 (Ar x 2), 128.8 (Ar x 2), 128.7 (Ar x 2), 127.3 (Ar), 127.2 (Ar), 118.8 (CN), 118.5 (CN), 54.4 (Phe-(C2)), 54.2 (Phe-(C2)), 38.6 (Phe-(C3)), 38.5 (Phe-(C3)), 35.8 (Ala-(C2)), 35.7 (Ala-(C2)), 23.2 (Ala-(C3)), 23.1 (Ala-(C3)), 18.8 (COCH₃ x 2). HRMS-EI [*M*]⁺⁺ calcd. for C₁₄H₁₇N₃O₂ 259.131528; obs. 259.131518. IR (solid, cm⁻¹) 3226, 3059, 3032, 2931, 1649, 1541.



Supplementary Figure 93. ¹H NMR (600 MHz, CDCl₃, 1.0–8.0 ppm) spectrum of isolated *N*-acetylphenylalanylalanine nitrile **Ac-Phe-Ala-CN**.

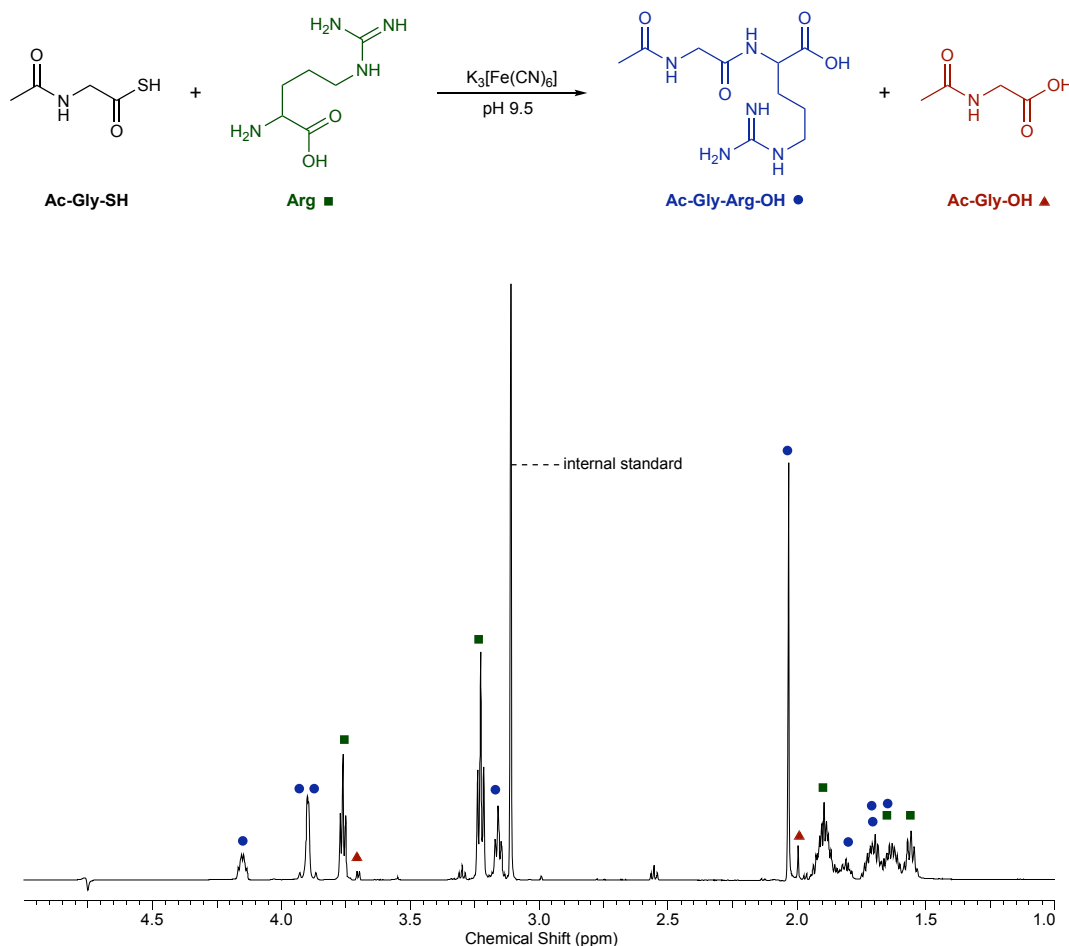


Supplementary Figure 94. ¹³C NMR (151 MHz, CDCl₃, 0–180 ppm) spectrum of isolated *N*-phenylalanylalanine nitrile **Ac-Phe-Ala-CN**.



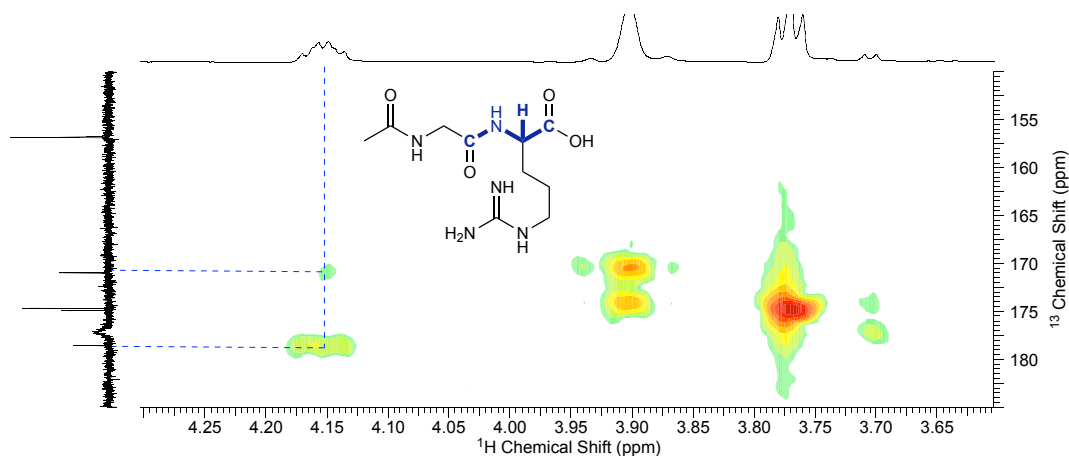
Supplementary Figure 95. ¹H–¹³C HMBC (¹H–600 MHz [4.50–5.00 ppm], ¹³C–151 MHz [110–185 ppm], CDCl₃) spectrum showing the diagnostic ²J_{CH} and ²J_{CH} coupling of Ala-(C2)–H of **Ac-Phe-Ala-CN** at 4.67–4.78 ppm with CN resonances at 118.5 and 118 ppm and C=O resonances between 170.3–170.6 ppm, which is characteristic of peptide bond formation of **Ala-CN**. See Supplementary Figure 93 for expanded ¹H NMR spectrum.

N-Acetylglycylarginine **Ac-Gly-Arg-OH**



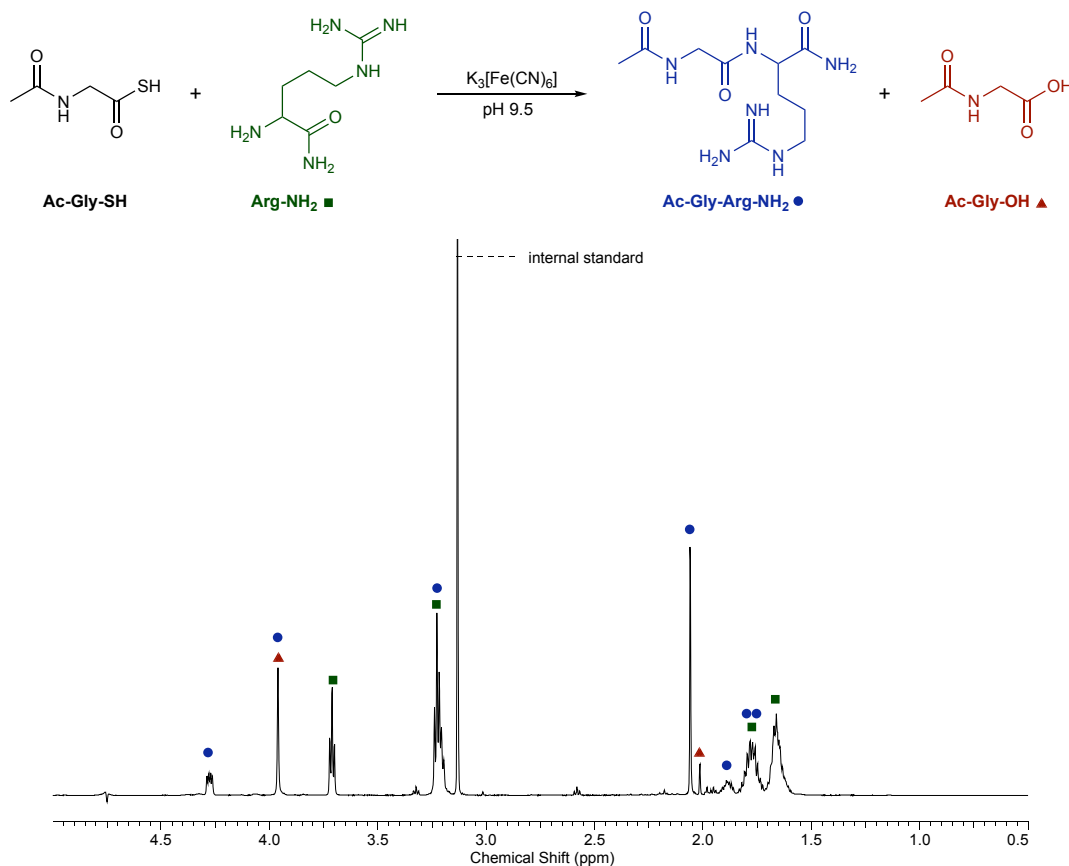
Supplementary Figure 96. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 1.0–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-arginine (**Arg**, 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycylarginine **Ac-Gly-Arg-OH** (●): δ 4.18 - 4.12 (m, 1H, Arg-(C2)-H), 3.91 (AB, $J = 17.5$ Hz, 1H, Gly-(C2)-H), 3.88 (AB, $J = 17.5$ Hz, 1H, Gly-(C2')-H), 3.16 (t, $J = 6.9$ Hz, 2H, Arg-(C5)-H₂), 2.03 (s, 3H, (COCH₃), 1.85 - 1.77 (m, 1H, Arg-(C3)-H), 1.75 - 1.67 (m, 2H, Arg-(C3')-H, Arg-(C4)-H), 1.67 - 1.60 (m, 1H, Arg-(C4')-H). *L*-Arginine **Arg** (■): δ 3.76 (t, $J = 6.1$ Hz, 1H, (C2)-H), 3.23 (t, $J = 6.9$ Hz, 2H, (C5)-H₂), 1.96 - 1.84 (m, 2H, (C3)-H₂), 1.68 - 1.60 (m, 1H, (C4)-H), 1.59 - 1.51 (m, 1H, (C4')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.70 (d, $J = 5.8$ Hz, 2H, (C2)-H₂), 2.00 (s, 3H, (COCH₃).



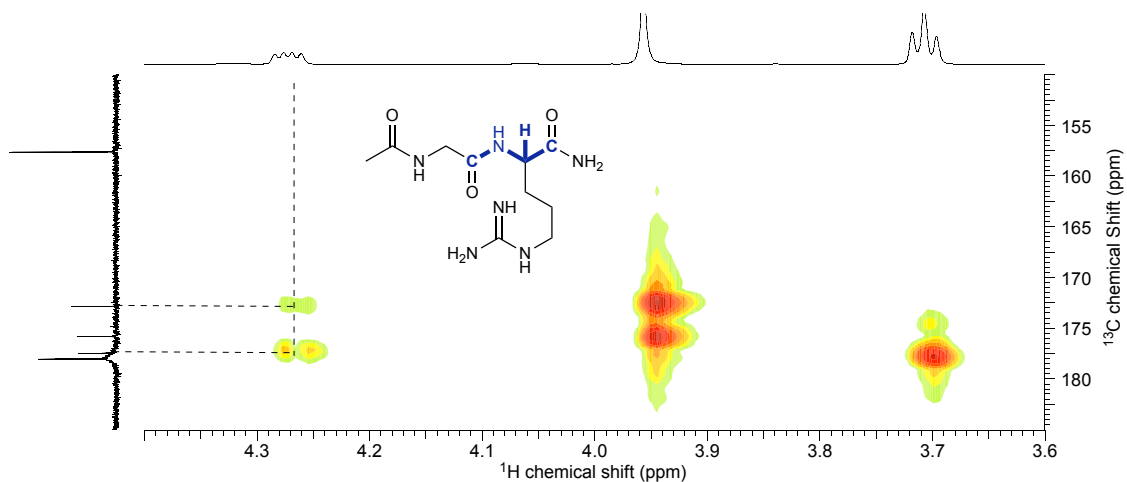
Supplementary Figure 97. 1H - ^{13}C HMBC (1H -600 MHz [3.6–4.3 ppm], ^{13}C -151 MHz [150–185 ppm], H_2O/D_2O 98:2) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Arg-(C2)-H of **Ac-Gly-Arg-OH** centred at 4.15 ppm with two C=O resonances at 170.9 and 178.5 ppm, which is characteristic of peptide bond formation of **Arg**. See Supplementary Figure 96 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylargininamide **Ac-Gly-Arg-NH₂**



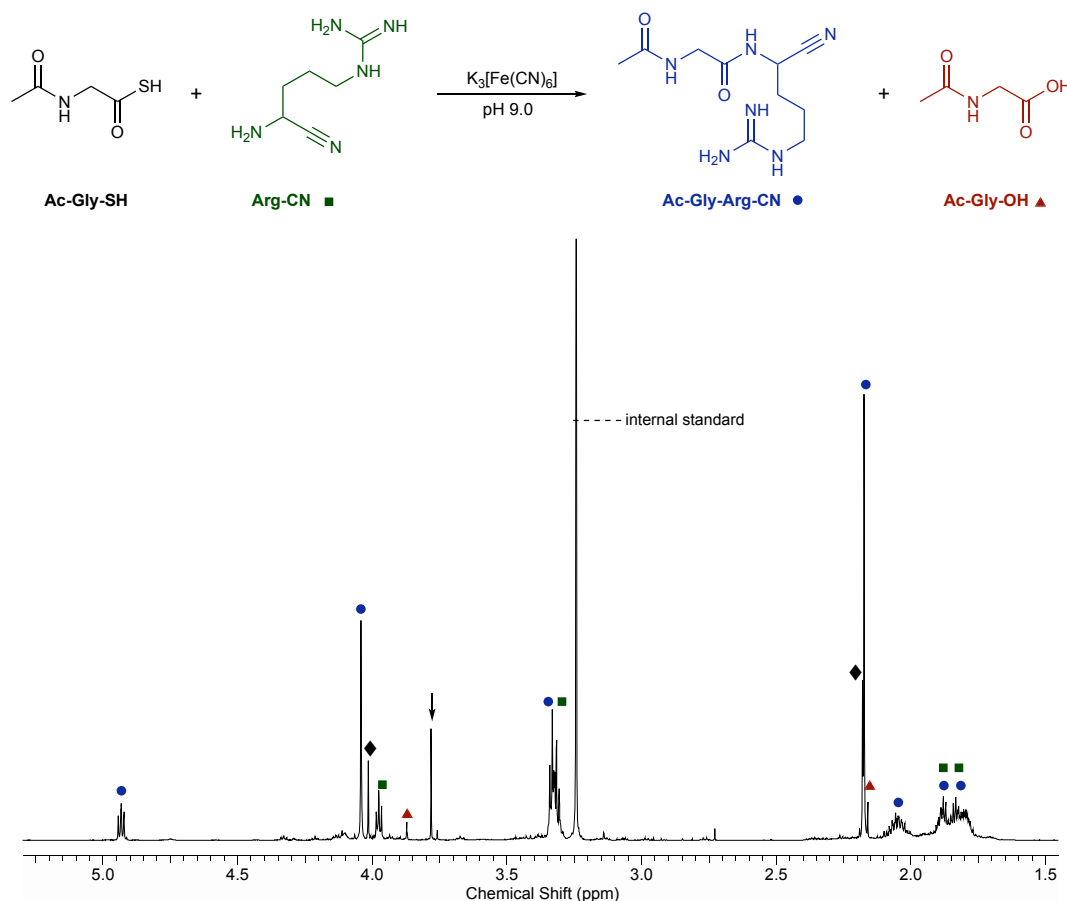
Supplementary Figure 98. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-argininamide (**Arg-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-acetylglycylargininamide **Ac-Gly-Arg-NH₂** (●): δ 4.28 (dd, J = 4.7, 9.1 Hz, 1H, Arg-(C2)–H), 3.96 (s, 2H, Gly-(C2)–H₂), 3.27 - 3.17 (m, 2H, Arg-(C5)–H₂), 2.07 (s, 3H, (COCH₃), 1.93 - 1.83 (m, 1H, Arg-(C3)–H), 1.79 - 1.72 (m, 3H, Arg-(C3')–H, Arg-(C4)–H₂). *L*-Argininamide **Arg-NH₂** (■): δ 3.71 (t, J = 6.5 Hz, 1H, (C2)–H), 3.27 - 3.17 (m, 2H, (C5)–H₂), 1.84 - 1.72 (m, 2H, (C3)–H₂), 1.71 - 1.59 (m, 2H, (C4)–H₂). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.96 (obs., 2H, (C2)–H₂), 2.01 (s, 3H, (COCH₃)).



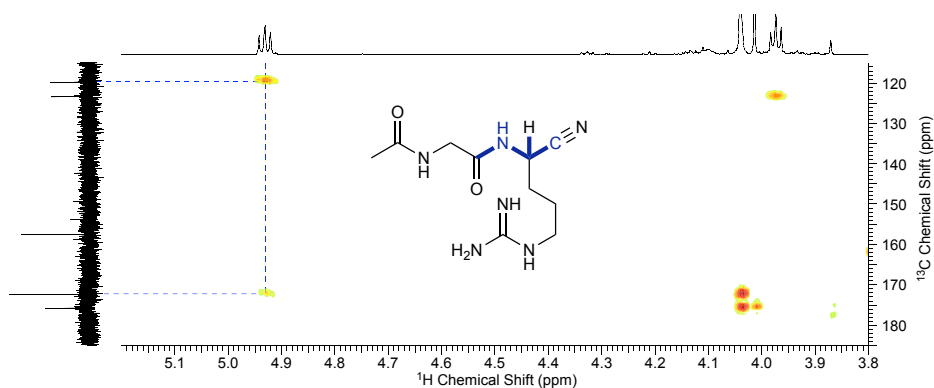
Supplementary Figure 99. ¹H–¹³C HMBC (¹H–600 MHz [3.6–4.4 ppm], ¹³C–151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Arg-(C2)–H of **Ac-Gly-Arg-NH₂** at 4.28 ppm with two C=O resonances at 172.8 and 177.4 ppm, which is characteristic of peptide bond formation of **Arg-NH₂**. See Supplementary Figure 98 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylarginine nitrile **Ac-Gly-Arg-CN**



Supplementary Figure 100. 1H NMR (700 MHz, D_2O , noesygppr1d, 1.5–5.2 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with arginine nitrile (**Arg-CN**; 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pD 9.0 and room temperature.

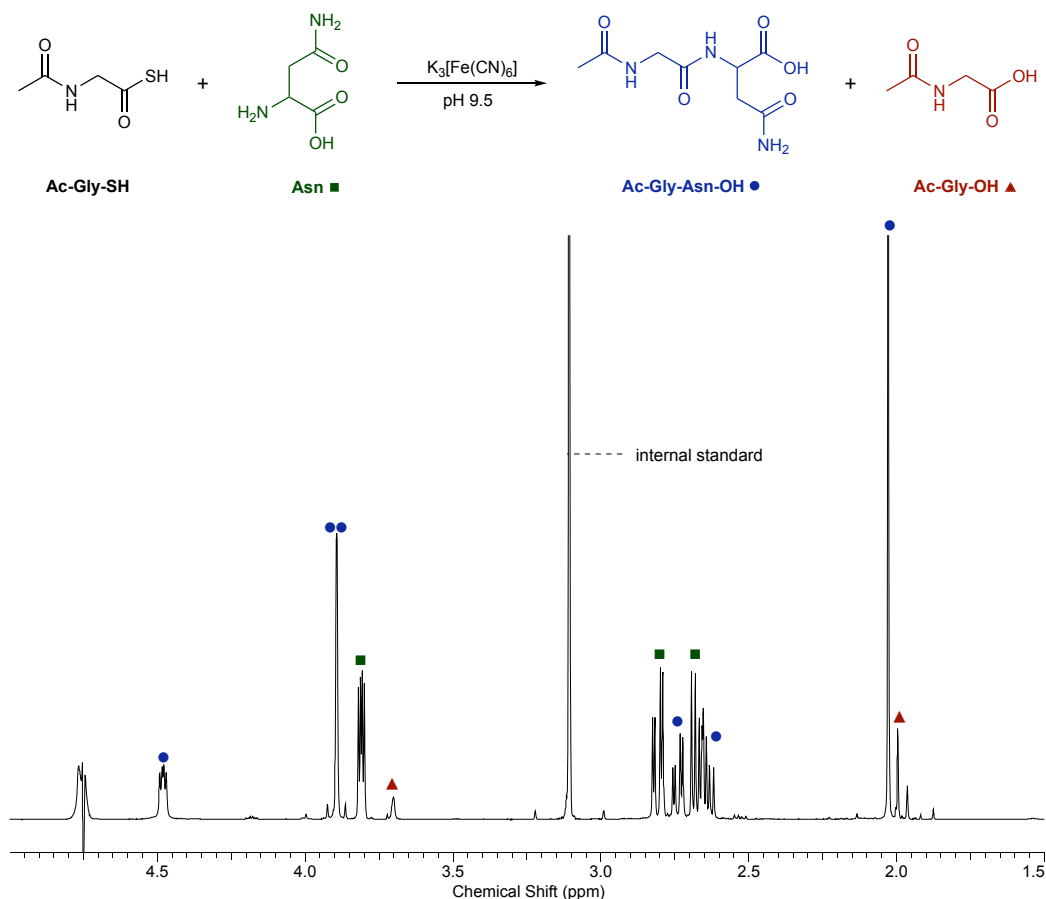
1H NMR (700 MHz, D_2O , noesygppr1d) *N*-Acetylglycylarginine nitrile **Ac-Gly-Arg-CN** (●): δ 4.94 (dd, $J = 7.0, 8.1$ Hz, 1H, Arg-(C2)–H), 4.05 (s, 2H, Gly-(C2)–H₂), 3.32 (t, $J = 6.8$ Hz, 2H, Arg-(C5)–H₂), 2.17 (s, 3H, (COCH₃)), 2.09 - 2.02 (m, 2H, Arg-(C3)–H₂), 1.86 - 1.81 (m, 2H, Arg-(C4)–H₂). Arginine nitrile **Arg-CN**¹⁴ (■): δ 3.98 (t, $J = 6.8$ Hz, 1H, (C2)–H), 3.33 (t, $J = 6.9$ Hz, 2H, (C5)–H₂), 1.93 - 1.85 (m, 2H, (C3)–H₂), 1.84 - 1.77 (m, 2H, (C3)–H₂). *N*-Acetylglycinamide **Ac-Gly-NH₂** (◆): δ 4.02 (s, 2H, (C2)–H₂), 2.18 (s, 3H, (COCH₃)). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.88 (s, 2H, (C2)–H₂), 2.16 (s, 3H, (COCH₃)). *O*-Methylisoureaⁱⁱⁱ (↓): δ 3.79 (s, 1H, OCH₃).



Supplementary Figure 101. 1H - ^{13}C HMBC (1H -700 MHz [3.8–5.2 ppm], ^{13}C -176 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Arg-(C2)–H of **Ac-Gly-Arg-CN** at 4.94 ppm with CN resonance at 119.0 ppm and a C=O resonance at 171.9 ppm, which is characteristic of peptide bond formation of **Arg-CN**. See Supplementary Figure 100 for expanded and labelled 1H NMR spectrum.

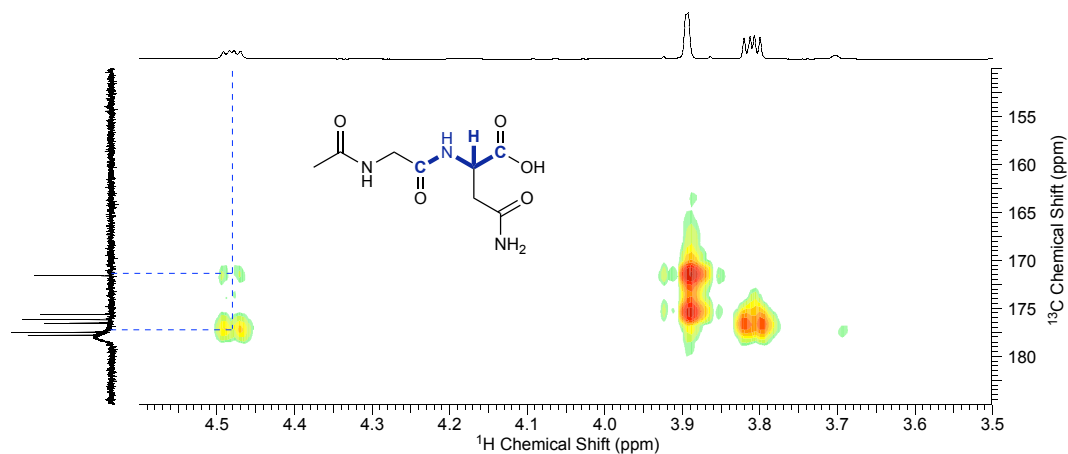
ⁱⁱⁱ *O*-Methylisourea was present as a low level and non-interfering impurity from its use during the preparative synthesis of a batch of arginine nitrile **Arg-CN** used in this ligation reaction. See Supplementary Reference 14 for the synthesis of **Arg-CN**.

N-Acetylglycylasparagine **Ac-Gly-Asn-OH**



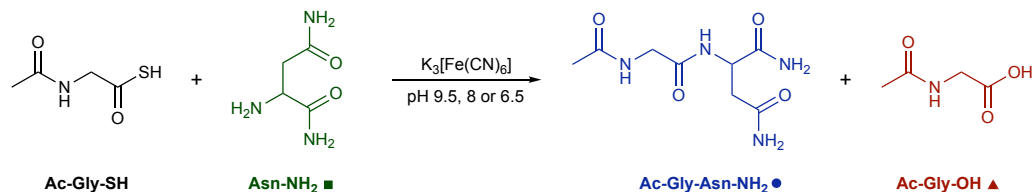
Supplementary Figure 102. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-asparagine (**Asn**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycylasparagine **Ac-Gly-Asn-OH** (●): δ 4.48 (dd, $J = 4.8, 8.3$ Hz, 1H, Asn-(C2)–H), 3.91 (AB, $J = 18.0$ Hz, 1H, Gly-(C2)–H), 3.88 (AB, $J = 18.0$ Hz, 1H, Gly-(C2')–H), 2.74 (ABX, $J = 4.8, 15.3$ Hz, 1H, Asn-(C3)–H), 2.64 (ABX, $J = 8.3, 15.3$ Hz, 1H, Asn-(C3')–H), 2.03 (s, 3H, (COCH₃)). *L*-Asparagine **Asn** (■): δ 3.81 (dd, $J = 4.5, 7.9$ Hz, 1H, (C2)–H), 2.81 (ABX, $J = 4.5, 16.0$ Hz, 1H, (C3)–H), 2.67 (ABX, $J = 7.9, 16.0$ Hz, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.70 (br. s., 2H, (C2)–H₂), 2.00 (s, 3H, (COCH₃)).

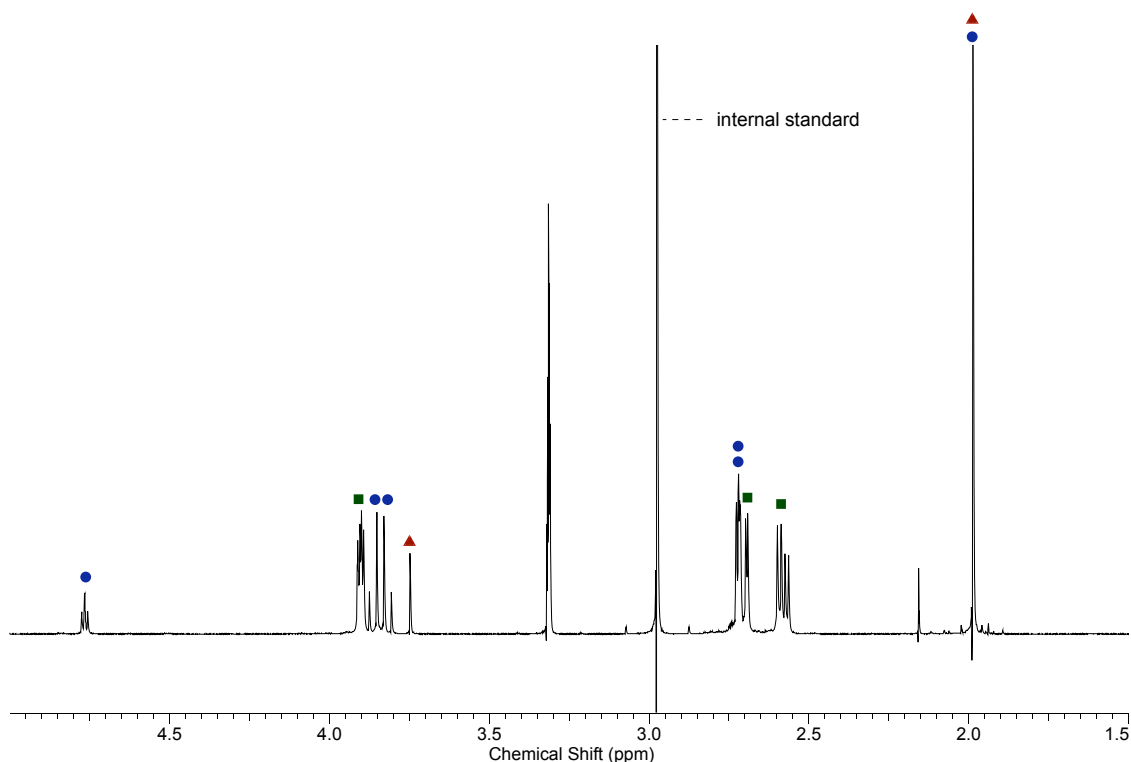


Supplementary Figure 103. 1H - ^{13}C HMBC (1H –600 MHz [3.5–4.6 ppm], ^{13}C –151 MHz [150–185 ppm], H_2O/D_2O 98:2) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Asn-(C2)–H of **Ac-Gly-Asn-OH** at 4.48 ppm with two C=O resonances at 171.6 and 177.5 ppm, which is characteristic of peptide bond formation of **Asn**. See Supplementary Figure 102 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylasparaginamide **Ac-Gly-Asn-NH₂**

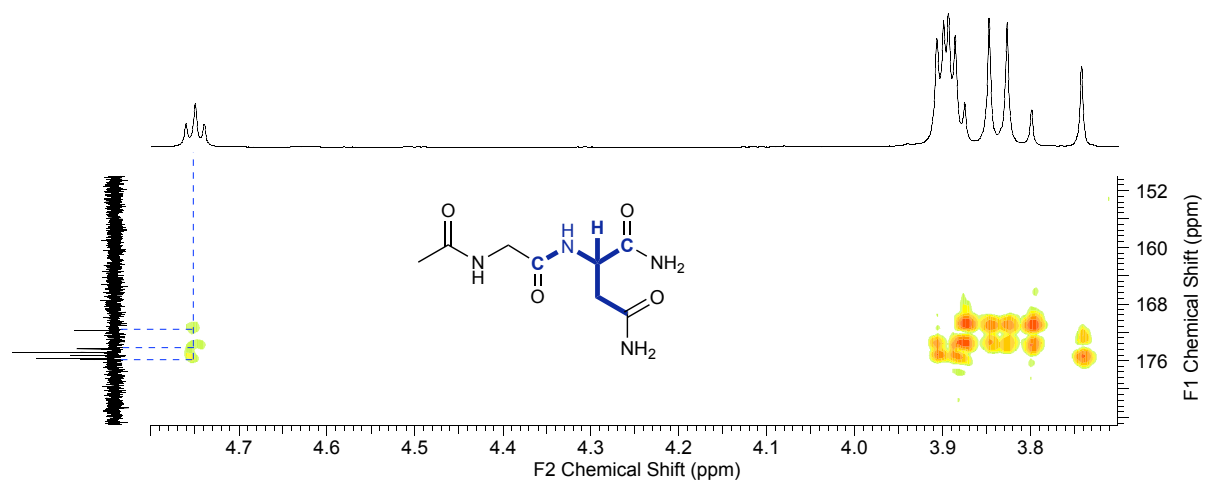


L-Asparaginamide **Asn-NH₂** (16.8 mg, 0.1 mmol) was dissolved in degassed H₂O/D₂O (98:2; 0.5 mL) and the solution was then adjusted to pH 9.5 with 4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and the total volume was adjusted to 1 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added and the solution was stirred at room temperature for 20 min, whilst maintaining the solution at pH 9.5, 8, or 6.5, as specified, with 4 M NaOH. The resulting suspension was centrifuged and the supernatant was then analysed immediately by NMR spectroscopy in H₂O/D₂O (98:2). The formation of **Ac-Gly-Asn-NH₂** (●) was quantified using methylsulfonylmethane (MSM) as an internal standard. The reaction mixture was then lyophilised and dissolved in CD₃OD for further ¹H–¹³C HMBC NMR spectral analysis, due to the coincidence of the Asn-(C2)–H resonance of **Ac-Gly-Asn-NH₂** with the HOD peak in H₂O/D₂O (98:2) NMR spectra. Yields are given in Supplementary Table 8.



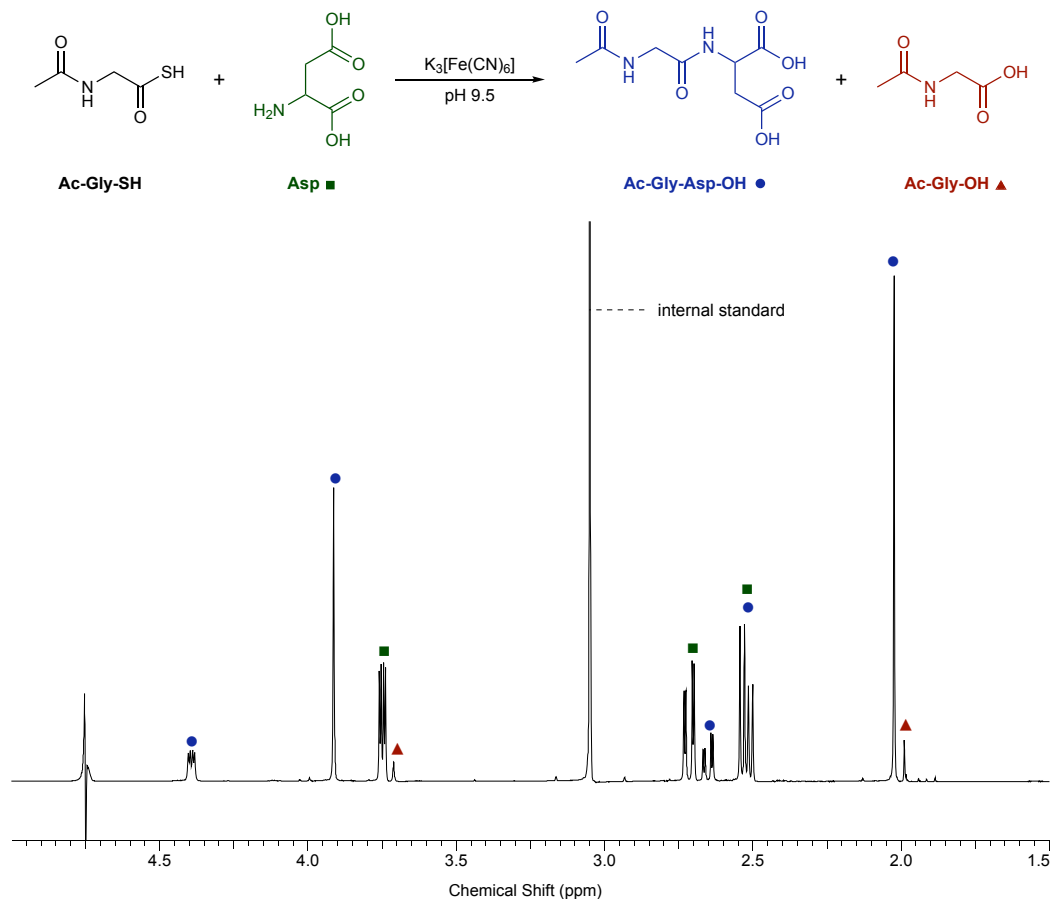
Supplementary Figure 104. ¹H NMR (700 MHz, CD₃OD, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-asparaginamide (**Asn-NH₂**; 150 mM) and K₃[Fe(CN)₆] (150 mM) at pH 8.0 and room temperature, after the solution was lyophilised and the resultant residue was dissolved in CD₃OD. Yields were obtained prior to lyophilisation.

¹H NMR (700 MHz, CD₃OD, noesygppr1d) *N*-Acetylglycylasparaginamide **Ac-Gly-Asn-NH₂** (●): δ 4.77 (t, *J* = 6.2 Hz, 1H, Asn-(C2)–H), 3.86 (AB, *J* = 16.7 Hz, 1H, Gly-(C2)–H), 3.82 (AB, *J* = 16.7 Hz, 1H, Gly-(C2')–H), 2.72 (d, *J* = 6.8 Hz, 2H, Asn-(C3)–H₂), 1.99 (s, 3H, (COCH₃)). L-Asparaginamide **Asn-NH₂** (■): δ 3.90 (dd, *J* = 4.6, 7.9 Hz, 1H, (C2)–H), 2.71 (ABX, *J* = 4.6, 16.1 Hz, 1H, (C3)–H), 2.58 (dd, *J* = 7.9, 16.1 Hz, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.75 (s, 2H, (C2)–H₂), 1.99 (s, 3H, (COCH₃)).



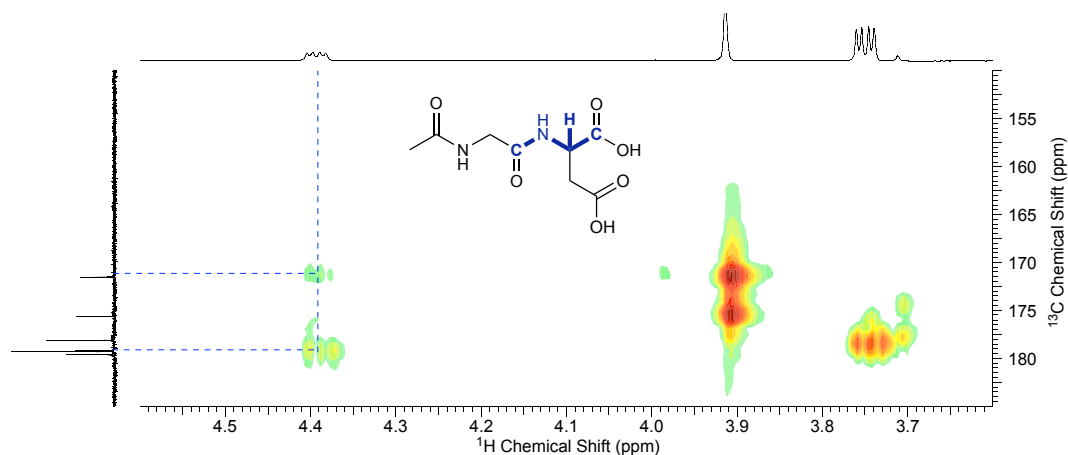
Supplementary Figure 105. ^1H - ^{13}C HMBC (^1H -700 MHz [3.7–4.8 ppm], ^{13}C -176 MHz [150–185 ppm], CD_3OD) spectrum showing the diagnostic $^2J_{\text{CH}}$ and two $^3J_{\text{CH}}$ couplings of Asn-(C2)-H of **Ac-Gly-Asn-NH₂** at 4.77 ppm with three C=O resonances at 172.4, 175.0 and 176.3 ppm, which is characteristic of peptide bond formation of **Asn-NH₂**. See Supplementary Figure 104 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycylaspartic acid **Ac-Gly-Asp-OH**



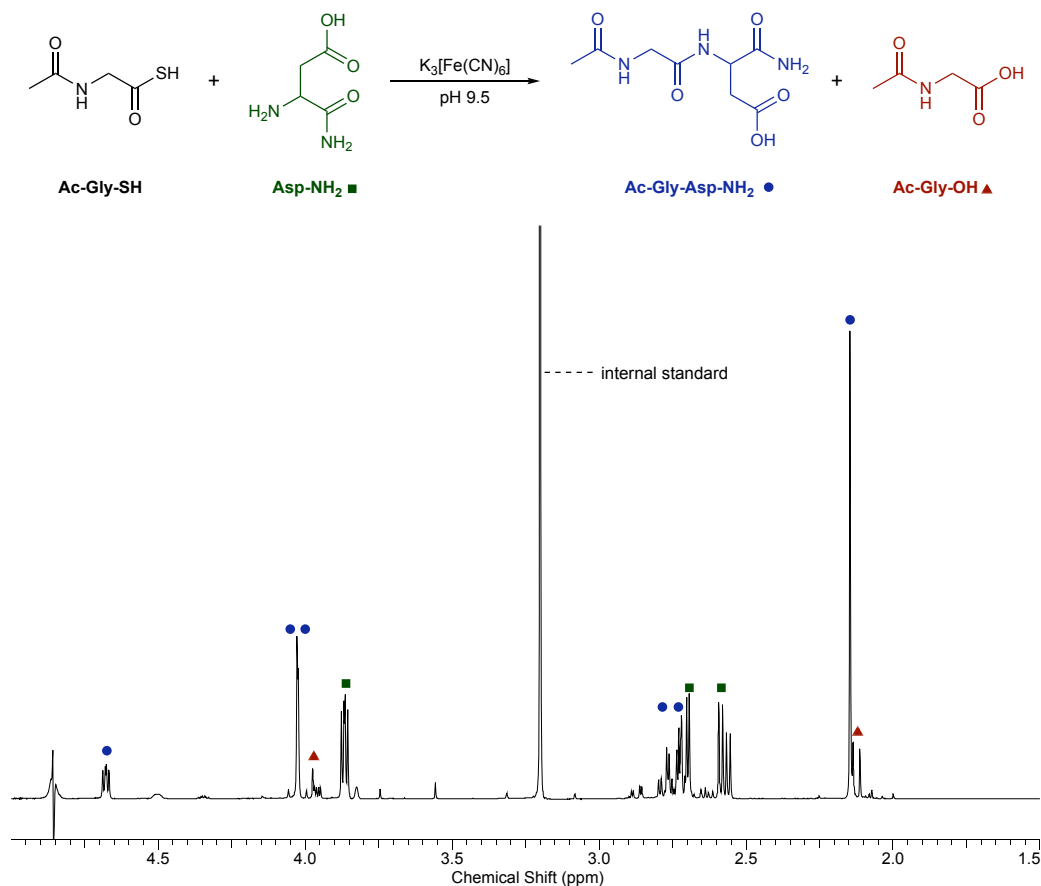
Supplementary Figure 106. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-aspartic acid (**Asp**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycylaspartic acid **Ac-Gly-Asp-OH** (●): δ 4.39 (dd, $J = 4.1$, 9.0 Hz, 1H, Asp-(C2)–H), 3.91 (s, 2H, Gly-(C2)–H₂), 2.65 (ABX, $J = 4.1$, 15.7 Hz, 1H, Asp-(C3)–H), 2.52 (obs., 1H, Asp-(C3')–H), 2.02 (s, 3H, (COCH₃)). L-Aspartic acid **Asp** (■): δ 3.75 (dd, $J = 3.9$, 8.8 Hz, 1H, (C2)–H), 2.72 (ABX, $J = 3.9$, 16.6 Hz, 1H, (C3)–H), 2.52 (ABX, $J = 8.8$, 16.6 Hz, 1H, (C')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.71 (s, 2H, (C2)–H₂), 1.99 (s, 3H, (COCH₃)).



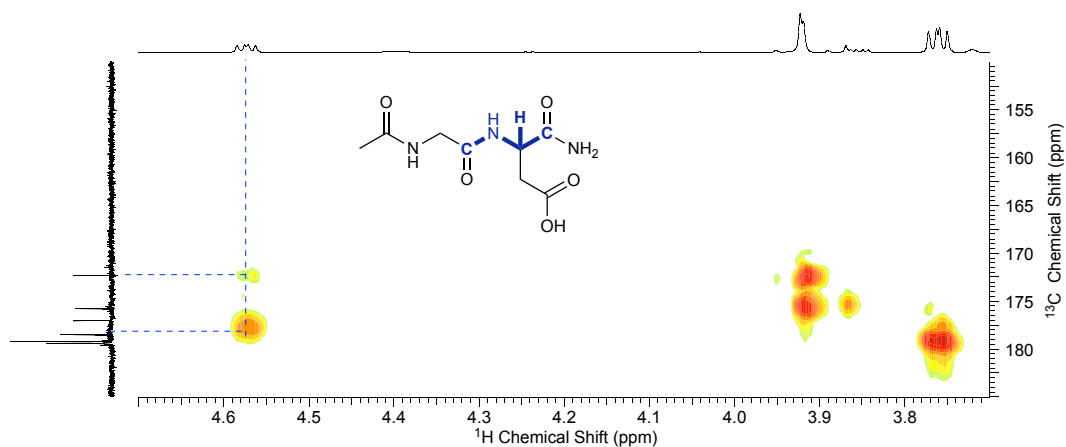
Supplementary Figure 107. 1H - ^{13}C HMBC (1H -600 MHz [3.6–4.6 ppm], ^{13}C -151 MHz [150–185 ppm], H_2O/D_2O 98:2) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Asp-(C2)–H of **Ac-Gly-Asp-OH** at 4.39 ppm with two C=O resonances at 171.5 and 179.1 ppm, which is characteristic of peptide bond formation of **Asp**. See Supplementary Figure 106 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylaspartamide **Ac-Gly-Asp-NH₂**



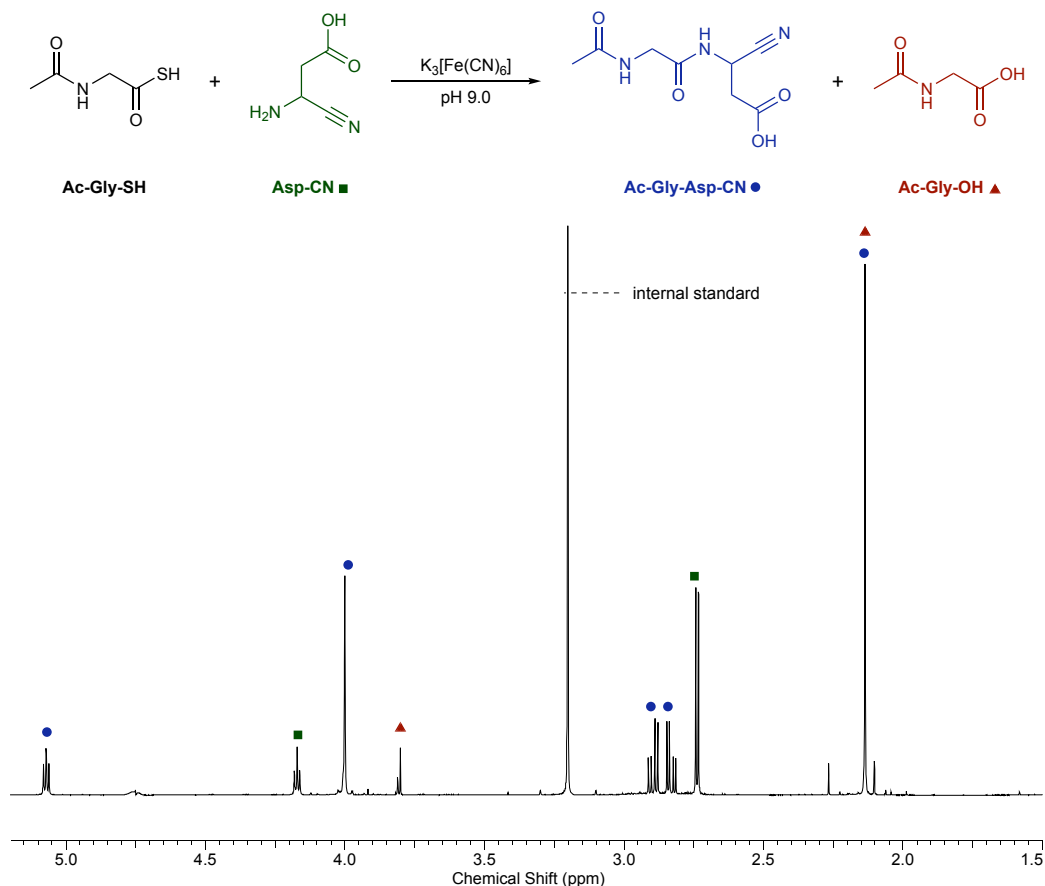
Supplementary Figure 108. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-aspartamide (**Asp-NH₂**; 150 mM) and K₃[Fe(CN)₆] (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylaspartamide **Ac-Gly-Asp-NH₂** (●): δ 4.57 (dd, *J* = 5.2, 7.8 Hz, 1H, Asp-(C2)-H), 3.94 (AB, *J* = 17.2 Hz, 1H, Gly-(C2)-H), 3.90 (AB, *J* = 17.2 Hz, 1H, Gly-(C2')-H), 2.67 (ABX, *J* = 5.2, 15.7 Hz, 1H, Asp-(C3)-H), 2.63 - 2.59 (obs., 1H, Asp-(C3')-H), 2.04 (s, 3H, (COCH₃)). *L*-Aspartamide **Asp-NH₂** (■): δ 3.76 (dd, *J* = 5.5, 7.8 Hz, 1H, (C2)-H), 2.61 (ABX, *J* = 5.5, 15.8 Hz, 1H, (C3)-H), 2.47 (dd, *J* = 7.8, 15.8 Hz, 1H, (C3')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.87 (s, 2H, (C2)-H₂), 2.01 (s, 3H, (COCH₃)).



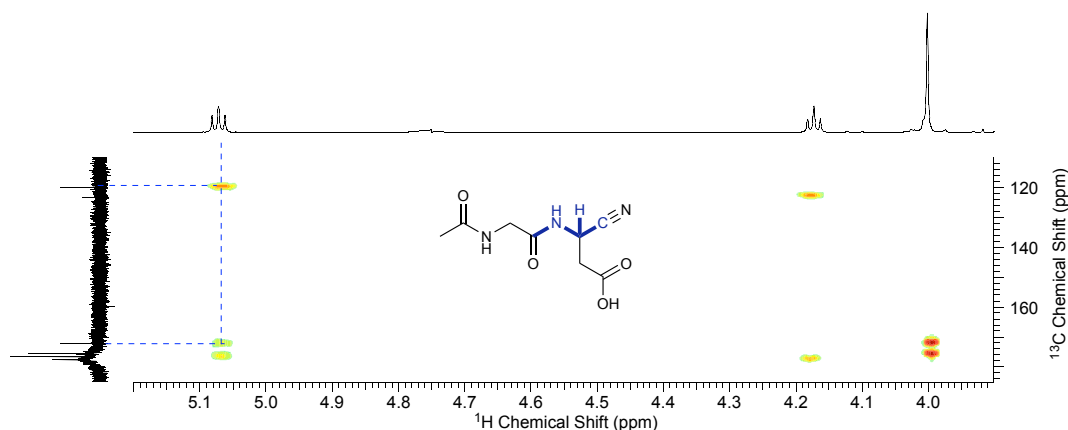
Supplementary Figure 109. ¹H-¹³C HMBC (¹H-600 MHz [3.7–4.7 ppm], ¹³C-151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ²*J*_{CH} and ³*J*_{CH} coupling of Asp-(C2)-H of **Ac-Gly-Asp-NH₂** at 4.57 ppm with two C=O resonances at 172.4 and 178.5 ppm, which is characteristic of peptide bond formation of **Asp-NH₂**. See Supplementary Figure 108 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylaspartic acid nitrile **Ac-Gly-Asp-CN**



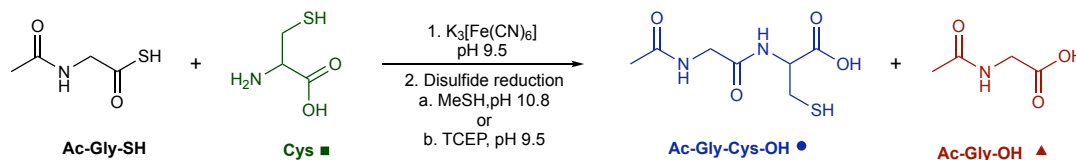
Supplementary Figure 110. 1H NMR (700 MHz, D_2O , noesygppr1d, 1.5–5.2 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with aspartic acid nitrile (**Asp-CN**; 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pD 9.0 and room temperature.

1H NMR (700 MHz, D_2O , noesygppr1d) *N*-Acetylglycyl-aspartic acid nitrile **Ac-Gly-Asp-CN** (●): δ 5.07 (dd, $J = 6.4, 7.3$ Hz, 1H, Asp-(C2)–H), 4.00 (s, 2H, Gly-(C2)–H₂), 2.90 (ABX, $J = 7.3, 16.4$ Hz, 1H, Asp-(C3)–H), 2.83 (ABX, $J = 6.4, 16.4$ Hz, 1H, Asp-(C3')–H), 2.14 (s, 3H, (COCH₃)). Aspartic acid nitrile **Asp-CN**⁴⁸ (■): δ 4.17 (t, $J = 6.5$ Hz, 1H, (C2)–H), 2.74 (d, $J = 6.5$ Hz, 2H, (C3)–H₂). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.80 (s, 2H, (C2)–H₂), 2.14 (s, 3H, (COCH₃)).



Supplementary Figure 111. 1H - ^{13}C HMBC (1H –700 MHz [3.7–5.2 ppm], ^{13}C –176 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Asp-(C2)–H of **Ac-Gly-Asp-CN** at 5.07 ppm with a CN resonance at 120.0 ppm and a C=O resonance at 172.1 ppm, which is characteristic of peptide bond formation of aspartic acid nitrile **Asp-CN**. See Supplementary Figure 110 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylcysteine **Ac-Gly-Cys-OH**

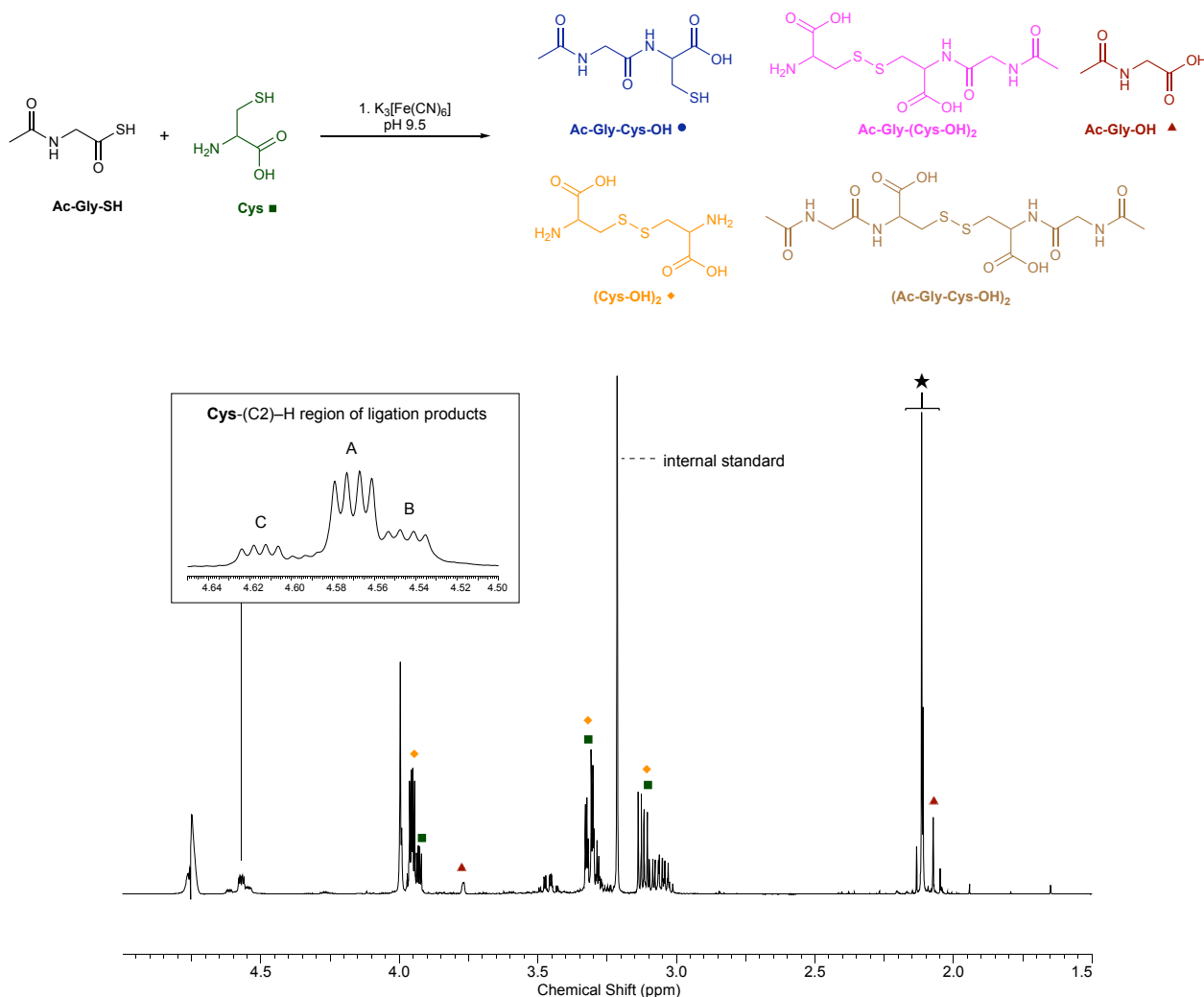


L-Cysteine **Cys** (18.2 mg, 0.15 mmol) was dissolved in degassed H_2O/D_2O (98:2; 0.5 mL) and the solution was adjusted to pH 9.5 with 4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and the total volume was adjusted to 1 mL with H_2O/D_2O (98:2). Potassium hexacyanoferrate(III) (98.8 mg, 0.3 mmol) was added and the solution was stirred at room temperature for 20 min whilst maintaining the solution at pH 9.5 with 4 M NaOH. The resulting suspension was centrifuged and the supernatant was then analysed by 1D and 2D NMR spectroscopy in H_2O/D_2O (98:2). The formation of the expected mixture of *N*-acylated cysteine and *N*-acylated cystine peptides **Ac-Gly-Cys-OH**, **Ac-Gly-(Cys-OH)₂** and **(Ac-Gly-Cys-OH)₂** was observed alongside cystine (**Cys-OH)₂** and residual **Cys** by 1H and 1H - ^{13}C HMBC analysis (Supplementary Figure 112 and Supplementary Figure 113, see below) and high resolution mass spectrometry^{iv} before disulfide reduction. Disulfide reduction was carried out by incubation with methanethiol–sodium salt^v (**MeS-Na⁺**, 42 mg, 0.6 mmol) for 5 min at pH 10.8 (Supplementary Figure 114, see below), or with tris(2-carboxyethyl)phosphine hydrochloride (**TCEP**; 62.5 mg, 0.25 mmol) for 5 min at pH 9.5 (Supplementary Figure 115 and Supplementary Figure 116, respectively- see below). Further NMR analysis then revealed complete disulfide reduction to yield *N*-acetylglycylcysteine **Ac-Gly-Cys-OH** and L-cysteine **Cys**, which were quantified using methylsulfonylmethane (MSM) as an internal standard and confirmed by 1H - ^{13}C HMBC NMR spectral analysis and high resolution mass spectrometry.

^{iv} HRMS-ESI data for: **Ac-Gly-(Cys-OH)₂** $[M+H]^+$ calcd. for $C_{10}H_{18}N_3O_6S_2$ 340.0644; obs. 340.0632; **(Ac-Gly-Cys-OH)₂** $[M+H]^+$ calcd. for $C_{14}H_{23}N_4O_8S_2$ 439.0952; obs. 439.0964.

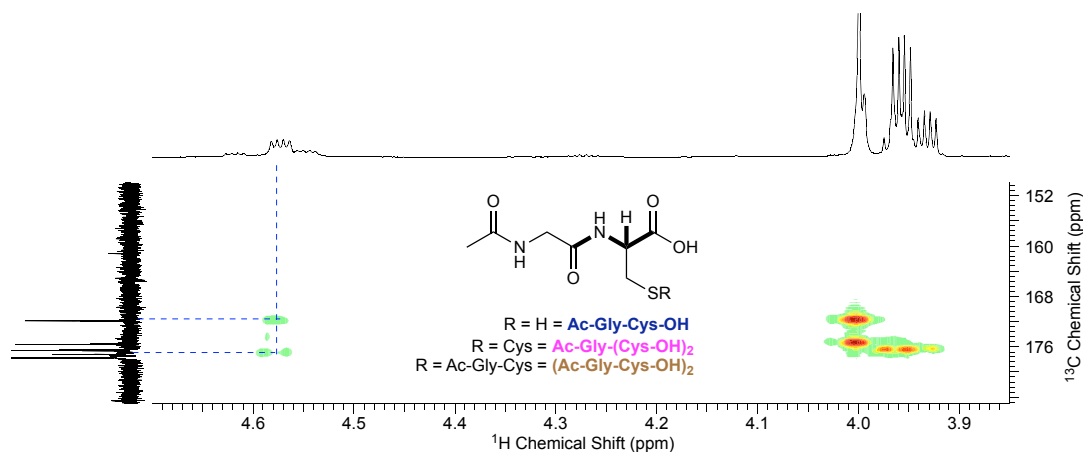
^v Disulfide reduction with methanethiol (**MeSH**; pKa 10.4) was carried out at pH 10.8 to minimise reaction hazards. **MeSH** is a highly toxic and flammable gas. **MeSH** has an extremely repulsive and putrid stench, with very low odour threshold (1 ppb). The reaction with **MeSH** was handled with care and in a well-ventilated fumehood, and rapidly analysed. All materials that came into contact with **MeSH** were washed with dilute sodium hypochlorite solution.

Before disulfide reduction



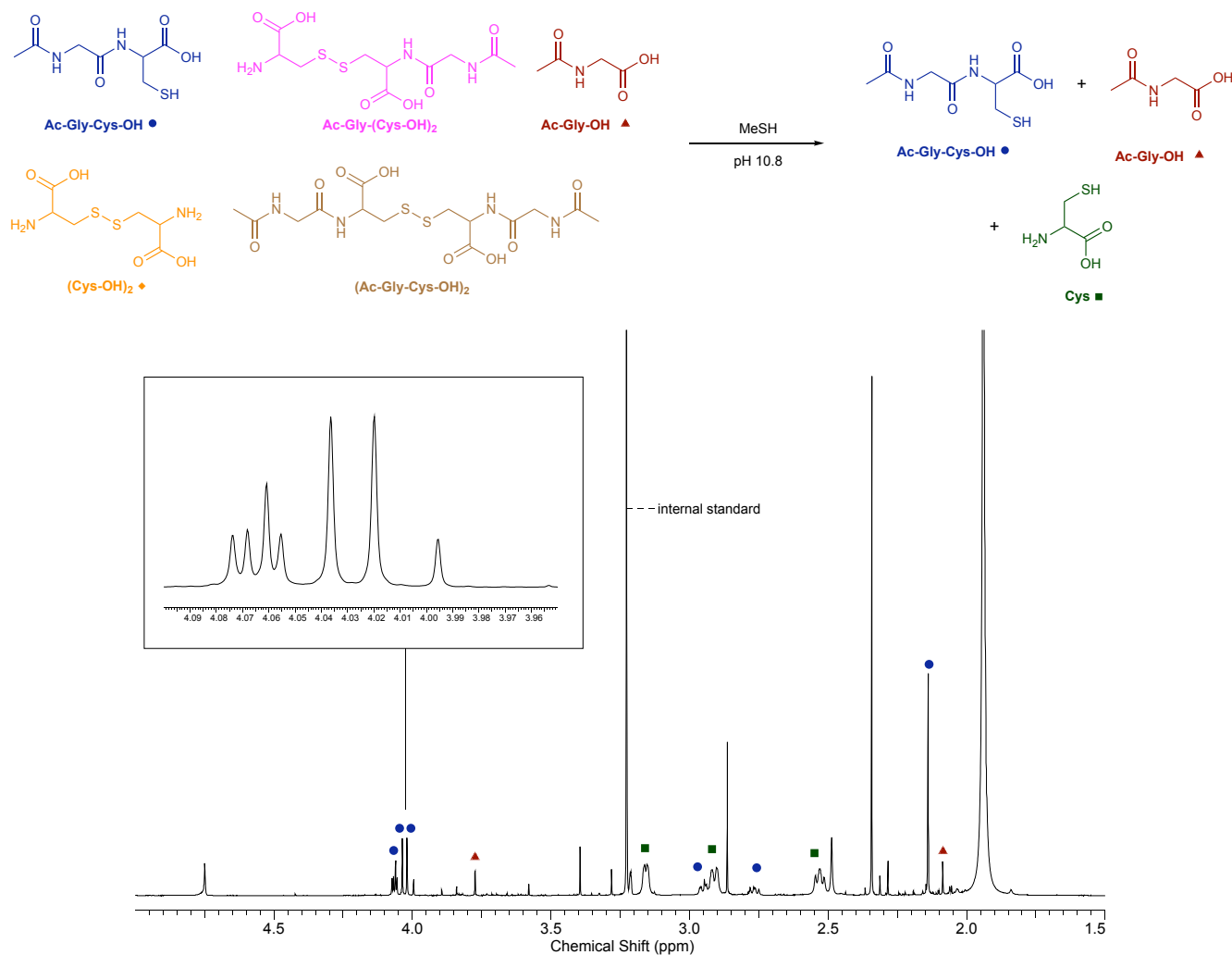
Supplementary Figure 112. ¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-cysteine (**Cys**; 150 mM) and $K_3[Fe(CN)_6]$ (300 mM) at pH 9.5 and room temperature (before reduction with **MeSH** – see Supplementary Figure 114, or **TCEP** – see Supplementary Figure 115). Inset: ¹H NMR (700 MHz, H₂O/D₂O 98:2, 4.50–4.65 ppm) showing new downfield shifted Cys-(C2)-H resonances (**A**:**B**:**C**, 60:27:13) of *N*-acylated cysteine and *N*-acylated cysteine peptides **Ac-Gly-Cys-OH**, **Ac-Gly-(Cys-OH)₂** and **(Ac-Gly-Cys-OH)₂**. NMR resonance peaks for L-cysteine **Cys** and L-cystine **(Cys-OH)₂** were specifically assigned, but those for **Ac-Gly-Cys-OH**, **Ac-Gly-(Cys-OH)₂** and **(Ac-Gly-Cys-OH)₂** could not be unambiguously identified, and are labelled **A**, **B**, and **C** in the NMR spectrum. However, upon reduction (with either **MeSH**, pH 10.8 or **TCEP**, pH 9.5) the three resonances for peptides **Ac-Gly-Cys-OH**, **Ac-Gly-(Cys-OH)₂** and **(Ac-Gly-Cys-OH)₂** and the amino acids **(Cys-OH)₂** and **Cys** coalesced into a single *N*-acylated cysteine residue **Ac-Gly-Cys-OH** and L-cysteine **Cys**, respectively. See Supplementary Figure 114 (**MeSH**) and Supplementary Figure 115 (**TCEP**) for ¹H NMR spectrum after disulfide reduction, resulting in coalescence of ¹H NMR resonances from a mixture of five cysteine-sulfide/disulfides to only *N*-acetylglycylcysteine **Ac-Gly-Cys-OH** and L-cysteine **Cys**. COCH₃ resonances for **Ac-Gly-Cys-OH**, **Ac-Gly-(Cys-OH)₂** and **(Ac-Gly-Cys-OH)₂** (denoted with a ★ between 2.1 - 2.2 ppm) could not be unambiguously assigned.

¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d, Supplementary Figure 112) *Cys-peptide A* (partial data): δ 4.57 (dd, *J* = 4.0, 8.3 Hz, 1H, Cys-(C2)-H). *Cys-peptide B* (partial data): δ 4.54 (dd, *J* = 4.0, 8.5 Hz, 1H, Cys-(C2)-H). *Cys-peptide C* (partial data): δ 4.62 (dd, *J* = 4.0, 8.1 Hz, 1H, Cys-(C2)-H). *L*-cysteine **Cys** (■): δ 3.93 (dd, *J* = 4.2, 8.4 Hz, 1H, (C2)-H), 3.32 - 3.29 (obs, 1H, (C3)-H), 3.08 (ABX, *J* = 8.4, 14.1 Hz, 1H, (C3')-H). *L*-cystine **(Cys-OH)₂** (◆): δ 3.95 (dd, *J* = 4.3, 7.9 Hz, 2H, (C2)-H), 3.32 (ABX, *J* = 4.3, 14.4 Hz, 2H, (C3)-H), 3.12 (ABX, *J* = 7.9, 14.4 Hz, 2H, (C3')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.77 (app. d, *J* = 2.7 Hz, 2H, (C2)-H₂), 2.07 (s, 3H, (COCH₃)).



Supplementary Figure 113. ^1H – ^{13}C HMBC (^1H –700 MHz [3.85–4.70 ppm], ^{13}C –176 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Cys-(C2)–H residues, with the major peptide centred at 4.57 ppm with two C=O resonances at 171.9 and 177.1 ppm. See Supplementary Figure 114 (**MeSH**) and Supplementary Figure 115 (**TCEP**) for ^1H NMR spectrum after disulfide reduction, resulting in coalescence of ^1H NMR resonances from a mixture of five cysteine-sulfide/disulfides (**Ac-Gly-Cys-OH**, **Ac-Gly-(Cys-OH) $_2$** , **(Ac-Gly-Cys-OH) $_2$** , **(Cys-OH) $_2$** and **Cys**) to only *N*-acetylglycylcysteine **Ac-Gly-Cys-OH** and L-cysteine **Cys**.

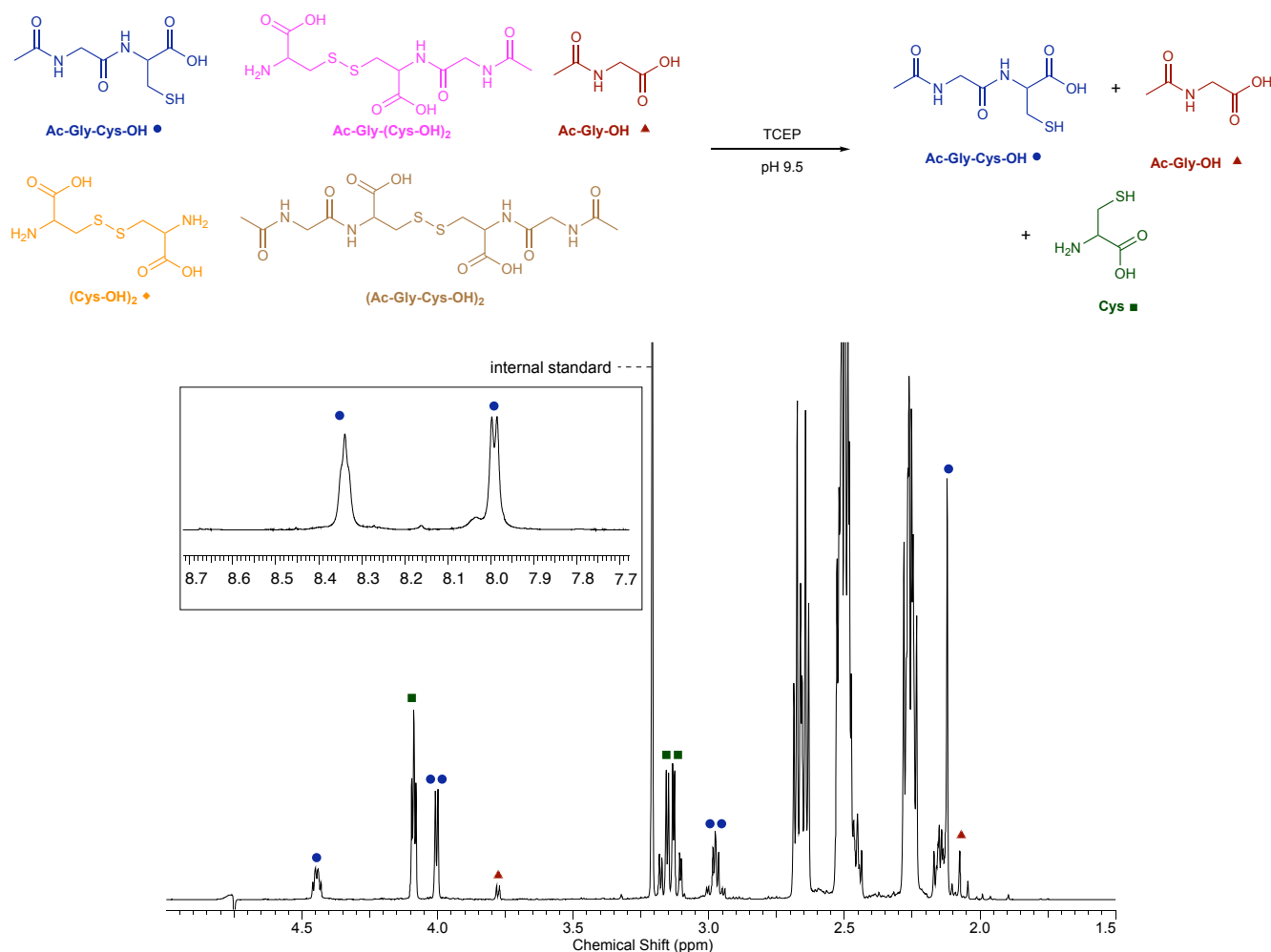
Disulfide reduction- methanethiol



Supplementary Figure 114. ^1H NMR (700 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the products after disulfide reduction (**MeSH**; 600 mM) at pH 10.8, of the mixture arising from the reaction 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-cysteine (**Cys**, 150 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (300 mM) at pH 9.5 and room temperature. See Supplementary Figure 112 for ^1H NMR spectrum before disulfide reduction. Inset: ^1H NMR (700 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygppr1d, 3.95–4.10 ppm) showing the Cys-(C2)–H and the highly diastereotopic Gly-(C2)– H_2 resonances for *N*-acetylglycylcysteine **Ac-Gly-Cys-OH**. See Supplementary Figure 115 for comparison to **TCEP** reduction, showing ^1H NMR spectrum after disulfide reduction with **TCEP**, also to only *N*-acetylglycylcysteine **Ac-Gly-Cys-OH** and L-cysteine **Cys**.

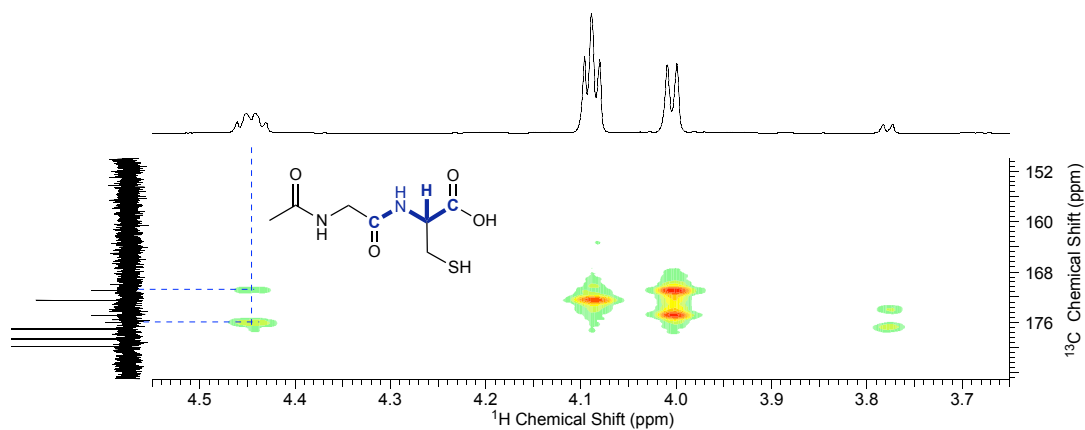
^1H NMR (700 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygppr1d, Supplementary Figure 114, pH 10.8) *N*-Acetylglycylcysteine **Ac-Gly-Cys-OH** (●): δ 4.06 (dd, $J = 4.4, 9.2$ Hz, 1H, Cys-(C2)-H), 4.04 (AB, $J = 17.0$ Hz, 1H, Gly-(C2)-H), 4.01 (d, $J = 17.0$ Hz, 1H, Gly-(C2')-H), 2.96 - 2.92 (obs., 1H, Cys-(C2)-H), 2.77 (ABX, $J = 9.2, 13.0$ Hz, 1H, Cys-(C2')-H), 2.14 (s, 3H, (COCH₃)). *L*-cysteine **Cys** (■): δ 3.16 (dd, $J = 3.3, 9.8$ Hz, 1H, (C2)-H), 2.91 (ABX, $J = 3.3, 11.8$ Hz, 1H, (C3)-H), 2.53 (ABX, $J = 9.8, 11.8$ Hz, 1H, (C3')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.77 (s, 2H, (C2)-H₂), 2.09 (s, 3H, (COCH₃)).

Disulfide reduction- tris(2-carboxyethyl)phosphine hydrochloride (TCEP)



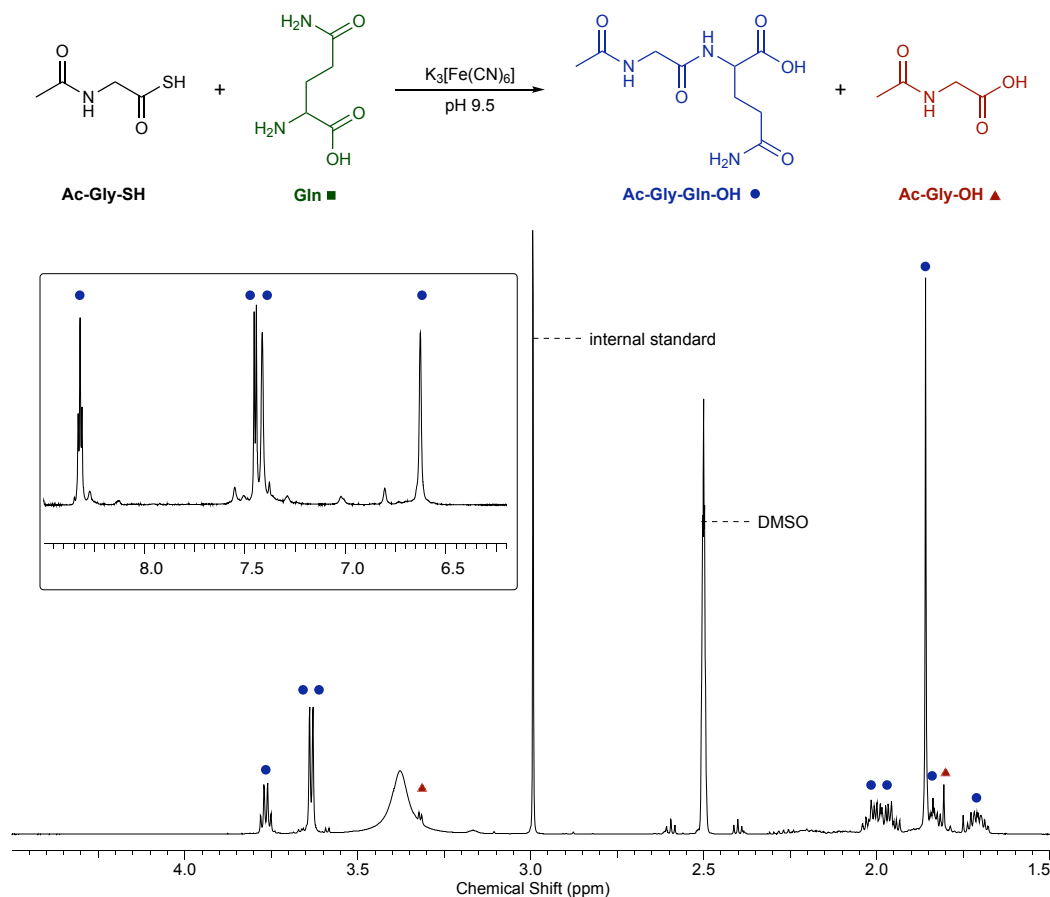
Supplementary Figure 115. ^1H NMR (600 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the products after disulfide reduction (TCEP; 250 mM) at pH 9.5, of the mixture arising from the reaction 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-cysteine (**Cys**, 150 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (300 mM) at pH 9.5 and room temperature. See Supplementary Figure 112 for ^1H NMR spectrum before disulfide reduction. Inset: ^1H NMR (600 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, 7.7–8.7 ppm) showing the Gly-NH and Cys-NH of *N*-acetylglycylcysteine **Ac-Gly-Cys-OH**, which underwent partial deuteration. See Supplementary Figure 114 for comparison to **MeSH** reduction, showing ^1H NMR spectrum after disulfide reduction with **MeSH**, also to only *N*-acetylglycylcysteine **Ac-Gly-Cys-OH** and L-cysteine **Cys**.

^1H NMR (600 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygppr1d, Supplementary Figure 115, pH 9.5) *N*-Acetylglycylcysteine **Ac-Gly-Cys-OH** (●): δ 8.36 (br. t, $J = 5.5$ Hz, 1H, Gly-NH), 8.01 (d, $J = 7.3$ Hz, 1H, Cys-NH), 4.45 (ddd, $J = 4.4, 6.1, 7.3$ Hz, 1H, Cys-(C2)-H), 4.00 (app. d, $J = 5.8$ Hz, 2H, Gly-(C2)-H), 2.99 (ABX, $J = 4.4, 13.9$ Hz, 1H, Cys-(C3)-H), 2.95 (ABX, $J = 6.1, 13.9$ Hz, 1H, Cys-(C3)-H'), 2.12 (s, 3H, (COCH₃)). *L*-cysteine **Cys** (■): δ 4.09 (dd, $J = 4.2, 5.3$ Hz, 1H, (C2)-H), 3.17 (ABX, $J = 5.3, 14.9$ Hz, 1H, (C3)-H), 3.12 (dd, $J = 4.2, 14.9$ Hz, 1H, (C3)-H'). *N*-Acetylglycine **Ac-Gly-OH** (▲) (partial assignment): δ 3.78 (d, $J = 5.9$ Hz, 2H, (C2)-H₂), 2.07 (s, 2H, (COCH₃)).



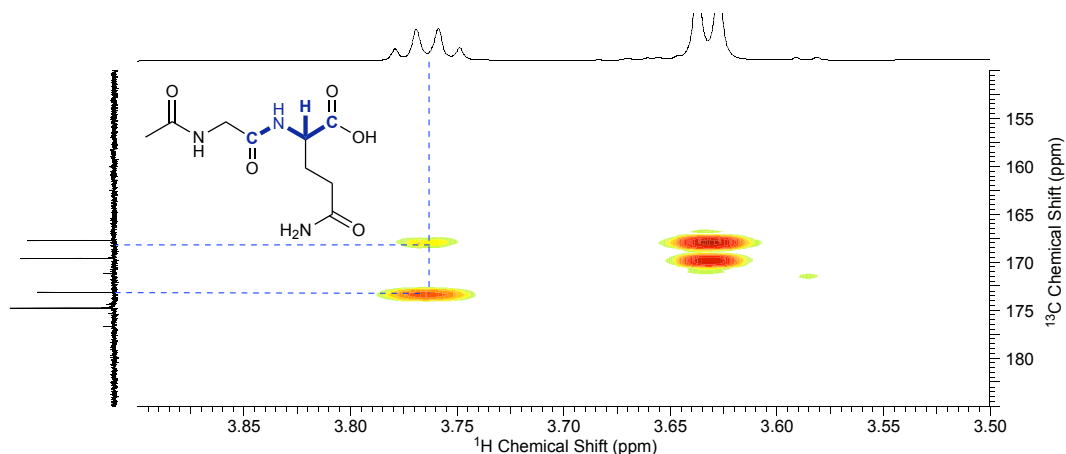
Supplementary Figure 116. ^1H - ^{13}C HMBC (^1H -600 MHz [3.65–4.55 ppm], ^{13}C -151 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) expansion showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Cys-(C2)-H of **Ac-Gly-Cys-OH** at 4.45 ppm with two C=O resonances at 171.6 and 176.9 ppm, which is characteristic of peptide bond formation of **Cys**. See Supplementary Figure 115 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycylglutamine **Ac-Gly-Gln-OH**



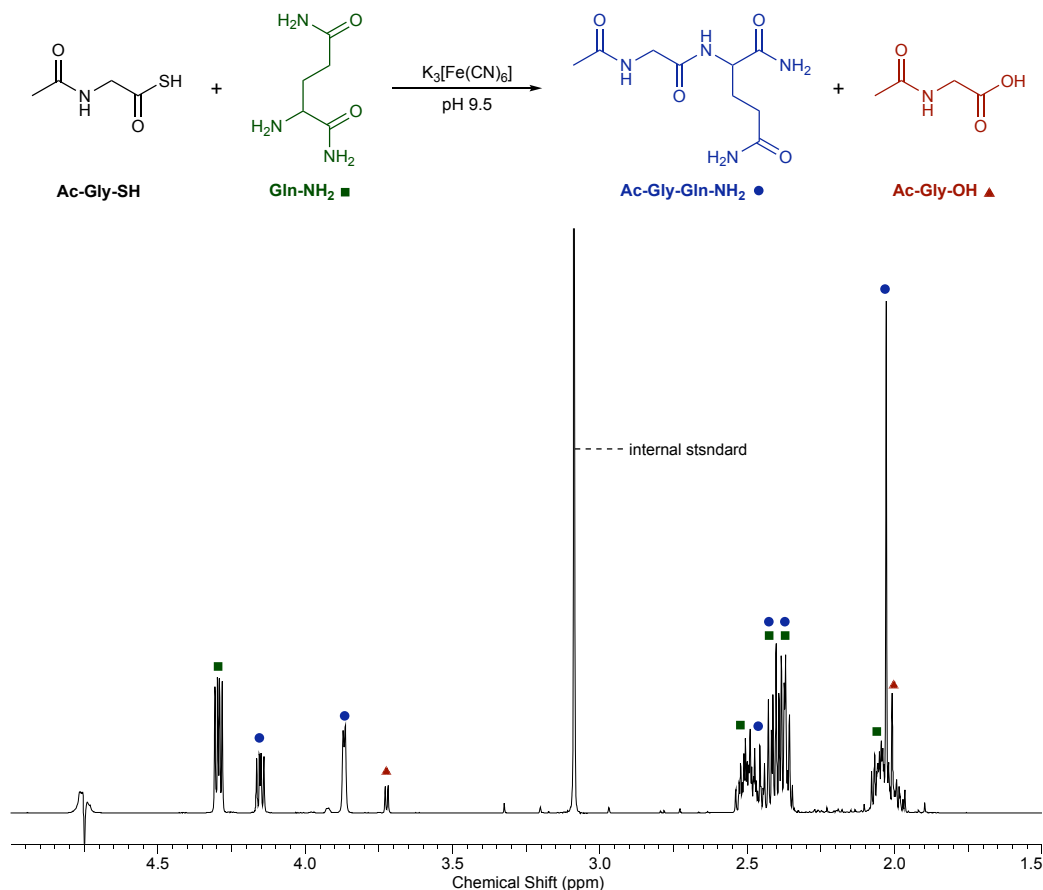
Supplementary Figure 117. ¹H NMR (600 MHz, DMSO-*d*₆, 1.5–4.5 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-glutamine (**Gln**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature after the solution was lyophilised and the resultant residue was dissolved in DMSO-*d*₆. Yields were obtained prior to lyophilisation. Residual L-glutamine **Gln** was not observed after lyophilisation (in DMSO-*d*₆), but was observed prior to lyophilisation in NMR spectra acquired in H₂O/D₂O (not shown). Inset: ¹H NMR (600 MHz, DMSO-*d*₆, 6.4–8.4 ppm) showing the amide NH resonances for *N*-acetylglycylglutamine **Ac-Gly-Gln-OH**.

¹H NMR (600 MHz, DMSO-*d*₆) *N*-Acetylglycylglutamine **Ac-Gly-Gln-OH** (●): δ 8.32 (t, *J* = 5.9 Hz, 1H, Gly-(C2)–NH), 7.45 (d, *J* = 6.7 Hz, 1H, Gln-(C2)–NH), 7.41 (br. s., 1H, Gln-(C5)–NH), 6.63 (br. s., 1H, Gln-(C5')–NH), 3.77 (app. q, *J* = 6.0 Hz, 1H, Gln-(C2)–H), 3.63 (app. d, *J* = 5.9 Hz, 2H, Gly-(C2)–H₂), 2.05 – 1.91 (m, 2H, (C4)–H₂), 1.86 (s, 3H, (COCH₃)), 1.85 – 1.81 (m, 1H, (C3)–H), 1.76 – 1.66 (m, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.87 (s, 2H, (C2)–H₂), 2.01 (s, 3H, (COCH₃)).



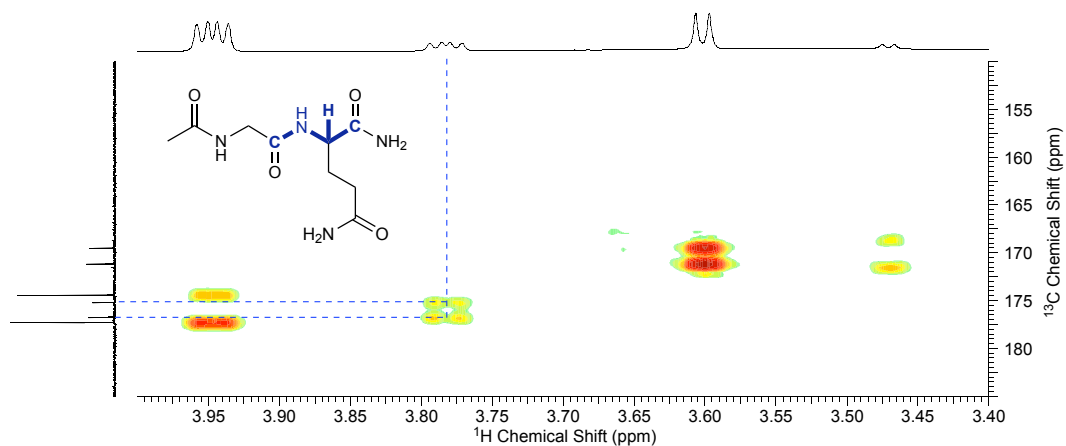
Supplementary Figure 118. ¹H–¹³C HMBC (¹H–600 MHz [3.5–3.9 ppm], ¹³C–151 MHz [150–185 ppm], DMSO-*d*₆) spectrum showing the diagnostic ²*J*_{CH} and ³*J*_{CH} coupling of Gln-(C2)–H of **Ac-Gly-Gln-OH** at 3.77 ppm with two C=O resonances at 167.7 and 173.1 ppm, which is characteristic of peptide bond formation of **Gln**. See Supplementary Figure 117 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylglutamine amide **Ac-Gly-Gln-NH₂**



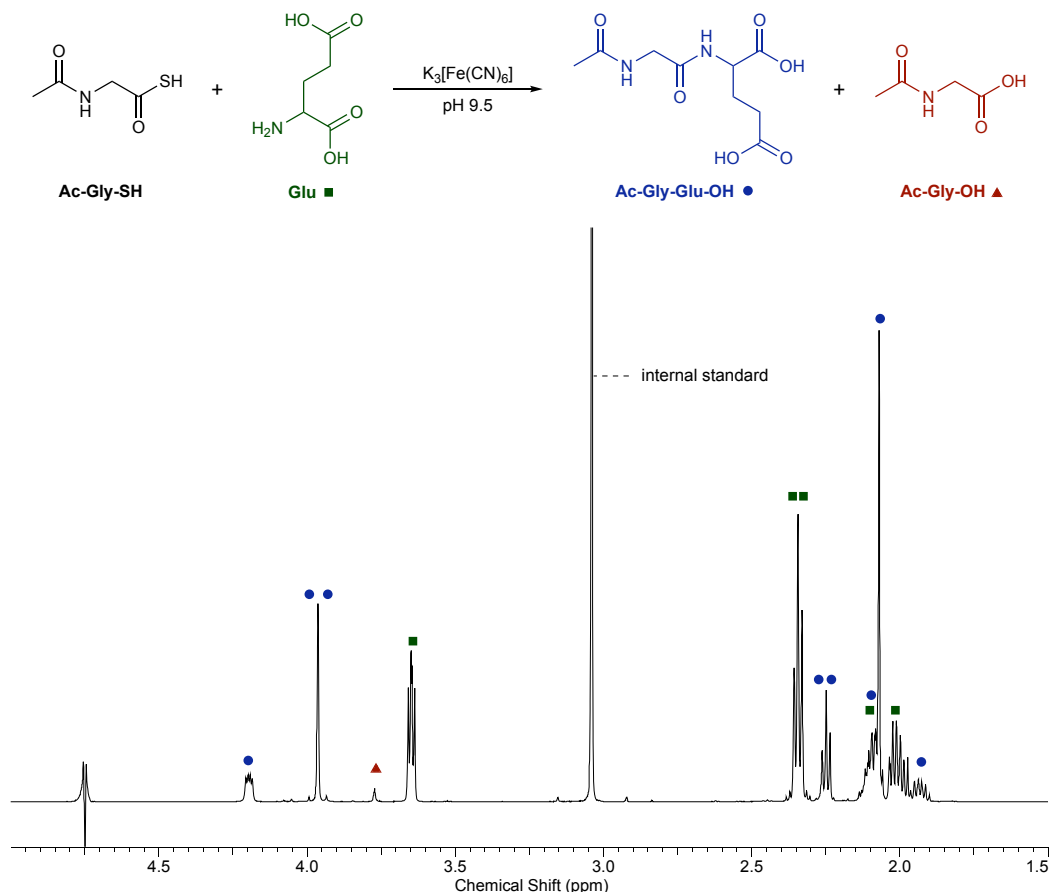
Supplementary Figure 119. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-glutamine amide (**Gln-NH₂**, 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylglutamine amide **Ac-Gly-Gln-NH₂** (●): δ 4.15 (dd, $J = 5.8, 9.0$ Hz, 1H, Gln-(C2)-H), 3.87 (app. d, $J = 4.5$ Hz, 2H, Gly-(C2)-H₂), 2.48 - 2.43 (m, 1H, Gln-(C3)-H), 2.44 - 2.34 (m, 2H, Gln-(C4)-H₂), 2.03 (s, 3H, (COCH₃)), 2.02 - 1.94 (m, 1H, Gln-(C3')-H). *L*-Glutamine amide **Gln-NH₂** (■): δ 4.29 (dd, $J = 5.2, 9.1$ Hz, 1H, (C2)-H), 2.55 - 2.45 (m, 1H, (C3)-H), 2.44 - 2.32 (m, 2H, (C4)-H₂), 2.09 - 2.03 (m, 1H, (C3')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.72 (app. d, $J = 5.9$ Hz, 2H, (C2)-H₂), 2.03 (s, 3H, (COCH₃)).



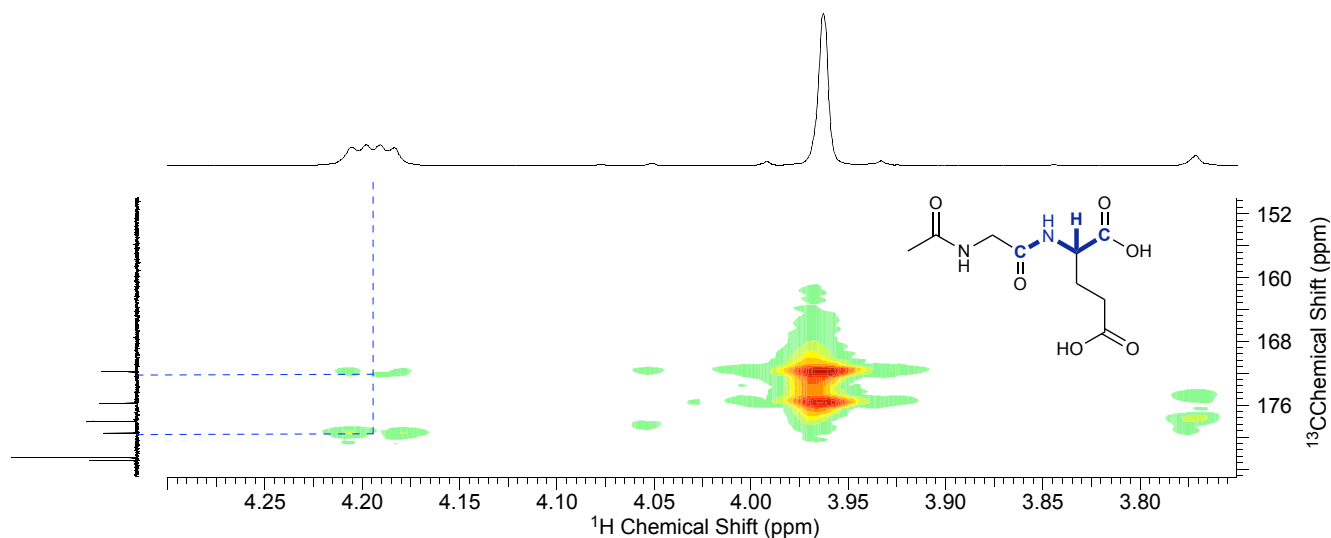
Supplementary Figure 120. ¹H-¹³C HMBC (¹H-600 MHz [3.4–4.0 ppm], ¹³C-151 MHz [150–185 ppm], DMSO-*d*₆) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Gln-(C2)-H of **Ac-Gly-Gln-NH₂** at 3.78 ppm with two C=O resonances at 175.2 and 176.9 ppm, which is characteristic of peptide bond formation of **Gln-NH₂**. See Supplementary Figure 119 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylglutamic acid **Ac-Gly-Glu-OH**



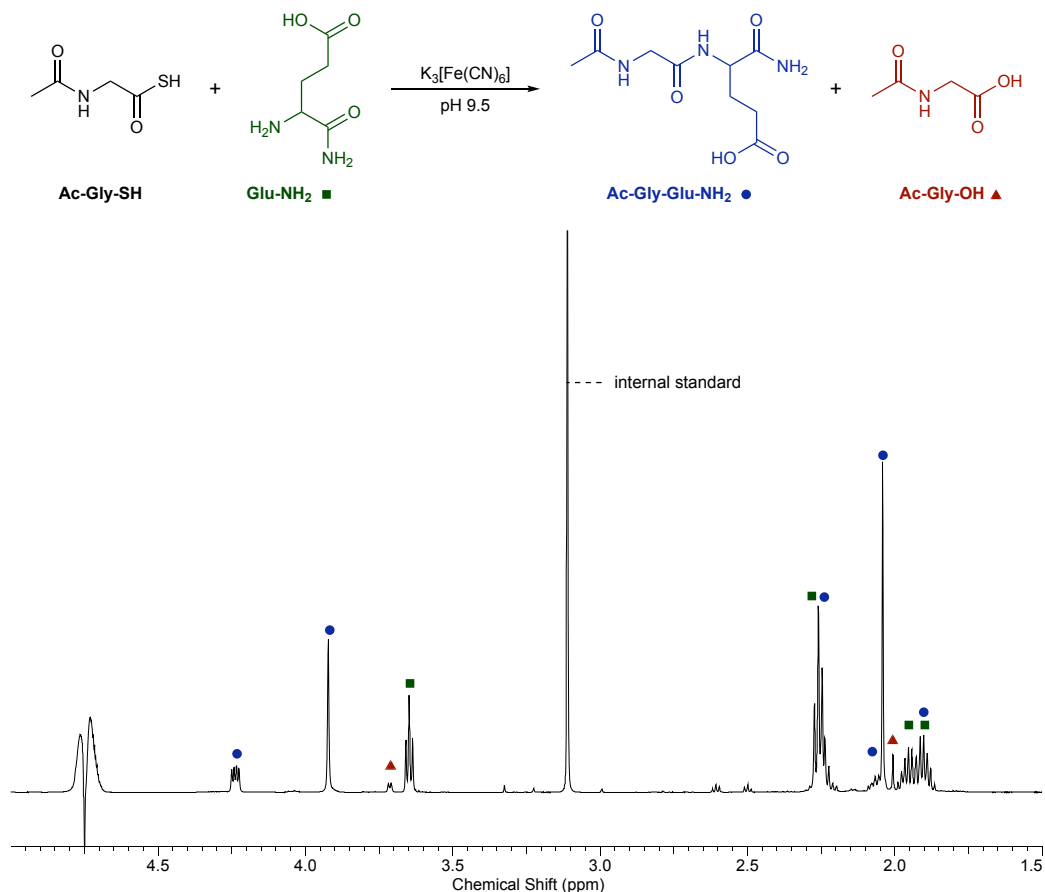
Supplementary Figure 121. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-glutamic acid (**Glu**, 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycylglutamic acid **Ac-Gly-Glu-OH** (●): δ 4.20 (dd, $J = 4.4, 8.7$ Hz, 1H, Glu-(C2)-H), 3.98 (AB, $J = 17.3$ Hz, 1H, Gly-(C2)-H), 3.95 (AB, $J = 17.3$ Hz, 1H, Gly-(C2')-H), 2.25 (t, $J = 8.0$ Hz, 2H, Glu-(C4)-H₂), 2.07 (s, 3H, (COCH₃), 2.14 - 2.05 (m, 1H, Glu-(C3)-H), 1.97 - 1.89 (m, 1H, Glu-(C3')-H). L-Glutamic acid **Glu** (■): δ 3.65 (dd, $J = 5.1, 7.0$ Hz, 1H, (C2)-H), 2.37 - 2.31 (m, 2H, (C3)-H₂), 2.14 - 2.05 (m, 1H, (C4)-H), 2.04 - 1.96 (m, 1H, (C4')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ (partial assignment): δ 3.77 (s, 2H, (C2)-H₂).



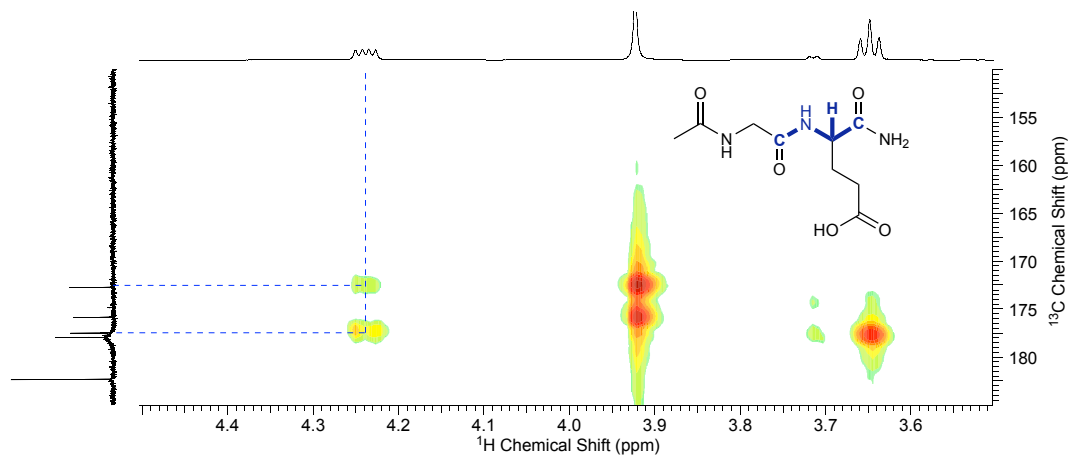
Supplementary Figure 122. 1H - ^{13}C HMBC (1H -600 MHz [3.6–4.3 ppm], ^{13}C -151 MHz [150–185 ppm], H_2O/D_2O 98:2) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Glu-(C2)-H of **Ac-Gly-Glu-OH** at 4.20 ppm with two C=O resonances at 171.8 and 179.5 ppm, which is characteristic of peptide bond formation of **Glu**. See Supplementary Figure 121 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylglutamic acid amide **Ac-Gly-Glu-NH₂**



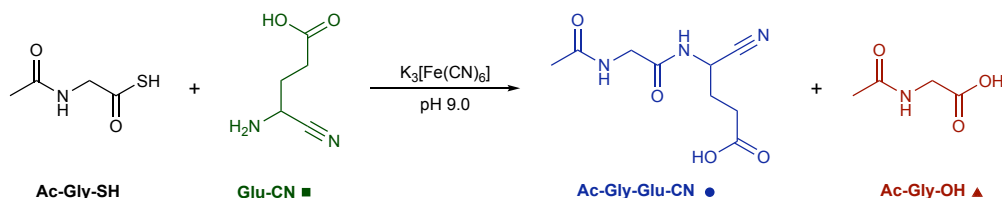
Supplementary Figure 123. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-glutamic acid amide (**Glu-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylglutamic acid amide **Ac-Gly-Glu-NH₂** (●): δ 4.24 (dd, J = 4.8, 9.4 Hz, 1H, Glu-(C2)–H), 3.92 (s, 2H, Gly-(C2)–H₂), 2.29 - 2.20 (m, 2H, Glu-(C4)–H₂), 2.09 - 2.04 (m, 1H, Glu-(C3)–H), 2.04 (s, 3H, (COCH₃)), 1.93 - 1.86 (m, 1H, Glu-(C3')–H). *L*-Glutamic acid amide **Glu-NH₂** (■): δ 3.65 (t, J = 6.6 Hz, 1H, (C2)–H), 2.28 - 2.24 (m, 2H, (C4)–H₂), 1.99 - 1.92 (m, 1H, (C3)–H), 1.92 - 1.86 (m, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.71 (app. d, J = 5.4 Hz, 2H, (C2)–H₂), 2.01 (s, 3H, (COCH₃)).

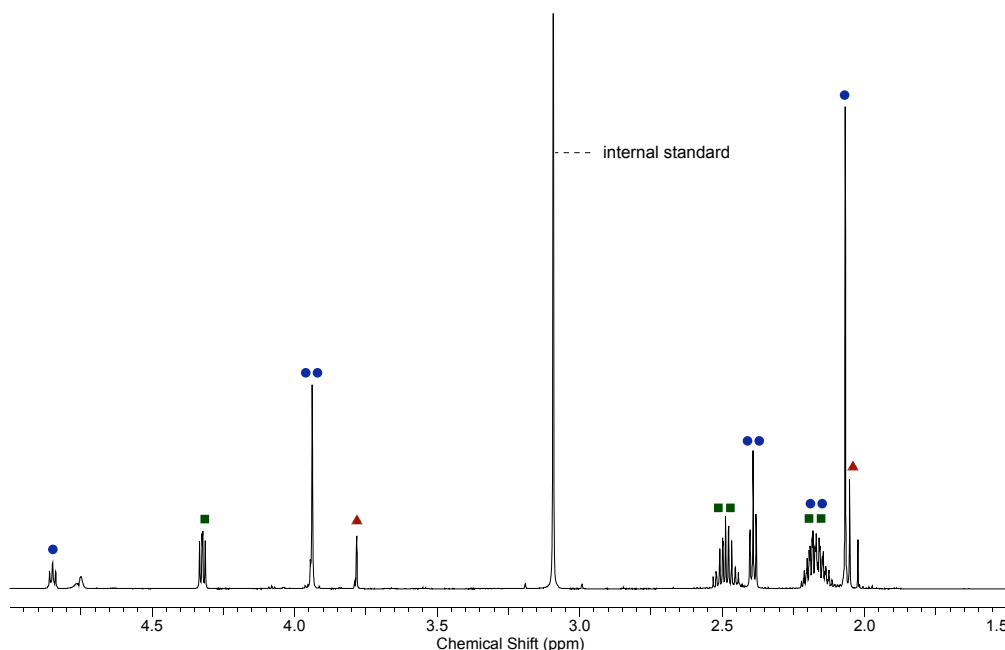


Supplementary Figure 124. ¹H-¹³C HMBC (¹H-600 MHz [3.5–4.5 ppm], ¹³C-151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ²J_{CH} and ³J_{CH} coupling of Glu-(C2)–H of **Ac-Gly-Glu-NH₂** at 4.24 ppm with two C=O resonances at 172.7 and 177.5 ppm, which is characteristic of peptide bond formation of **Glu-NH₂**. See Supplementary Figure 123 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylglutamic acid nitrile **Ac-Gly-Glu-CN**

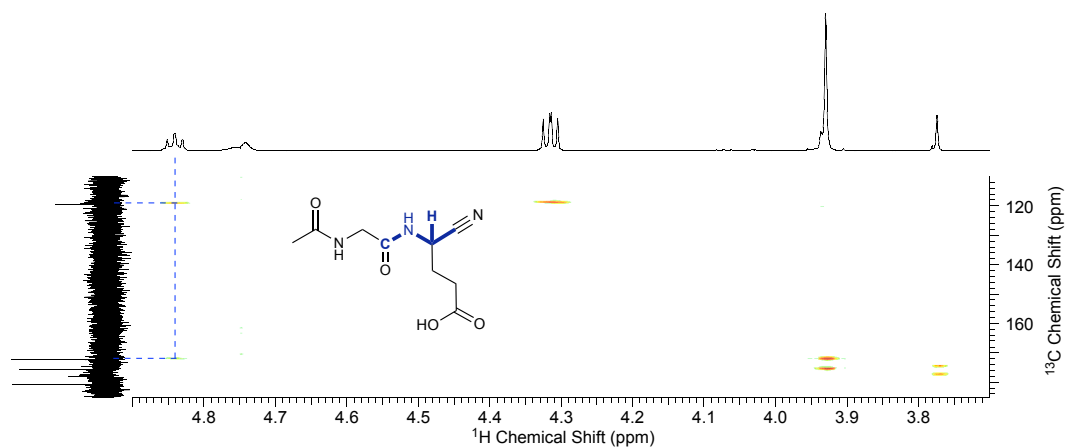


rac-Glutamic acid nitrile **Glu-CN**⁴⁹ (12.8 mg, 0.1 mmol) was dissolved in D₂O (0.5 mL) and the solution was adjusted to pH 9.0 with 4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and total volume was adjusted to 1 mL with D₂O. Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added and the solution was stirred at room temperature for 20 min. The solution was adjusted to pD 6 and centrifuged. The supernatant was analysed by 1D and 2D NMR spectroscopy and the reaction mixture was quantified using methylsulfonylmethane (MSM) as an internal standard. The presence of **Ac-Gly-Glu-CN** (●) was confirmed by ¹H–¹³C HMBC NMR spectral analysis and high resolution mass spectrometry.



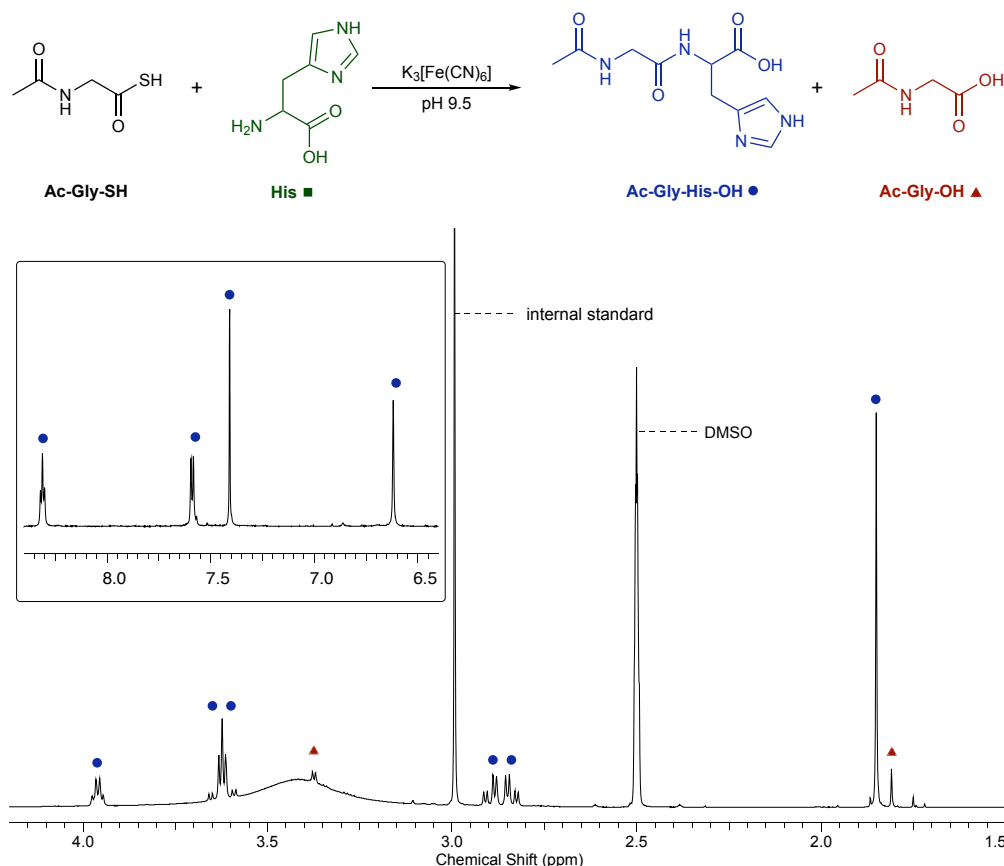
Supplementary Figure 125. ¹H NMR (700 MHz, D₂O, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with *rac*-glutamic acid nitrile (**Glu-CN**; 100 mM) and K₃[Fe(CN)₆] (150 mM) at pD 9.0 and room temperature. NMR analysis was carried out at pD 6 to reduce signal overlaps.

¹H NMR (700 MHz, D₂O, noesygppr1d) *N*-Acetylglycylglutamic acid nitrile **Ac-Gly-Glu-CN** (●): δ 4.85 (dd, *J* = 7.2, 7.9 Hz, 1H, Glu-(C2)–H), 3.95 (AB, *J* = 17.4 Hz, 1H, Gly-(C2)–H), 3.91 (d, *J* = 17.4 Hz, 1H, Gly-(C2')–H), 2.39 (t, *J* = 7.4 Hz, 2H, Glu-(C4)–H₂), 2.18 - 2.11 (m, 2H, Glu-(C3)–H₂), 2.07 (s, 3H, (COCH₃)). *rac*-Glutamic acid nitrile **Glu-CN** (■): δ 4.32 (dd, *J* = 6.3, 8.1 Hz, 1H, (C2)–H), 2.53 - 2.43 (m, 2H, (C4)–H₂), 2.22 - 2.13 (m, 2H, (C3)–H₂). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.78 (s, 2H, (C2)–H₂), 2.05 (s, 3H, (COCH₃)).



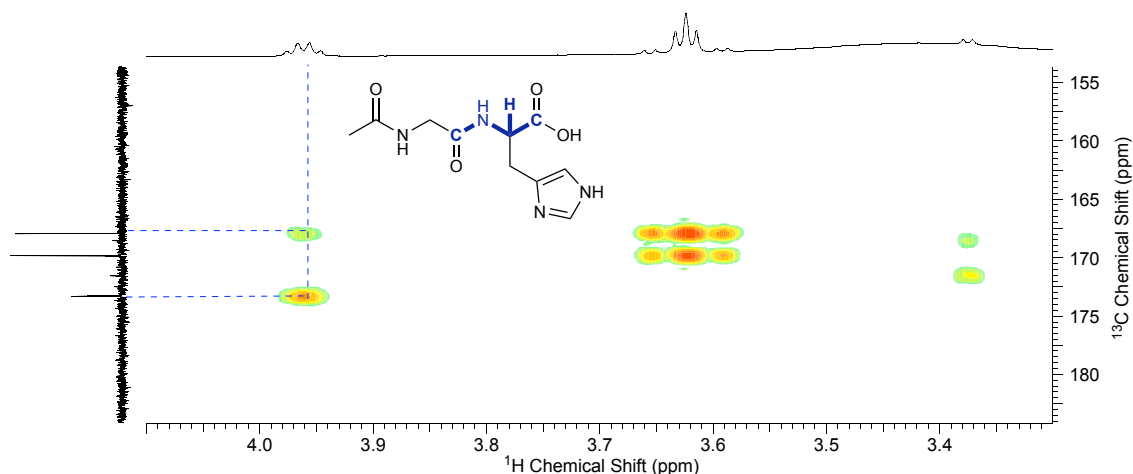
Supplementary Figure 126. ^1H - ^{13}C HMBC (^1H -700 MHz [3.7–4.9 ppm], ^{13}C -176 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Glu-(C2)-H of **Ac-Gly-Glu-CN** at 4.85 ppm with CN resonance at 119.2 ppm and a C=O resonance at 172.2 ppm, which is characteristic of peptide bond formation of *rac*-glutamic acid nitrile **Glu-CN**. See Supplementary Figure 125 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycylhistidine **Ac-Gly-His-OH**



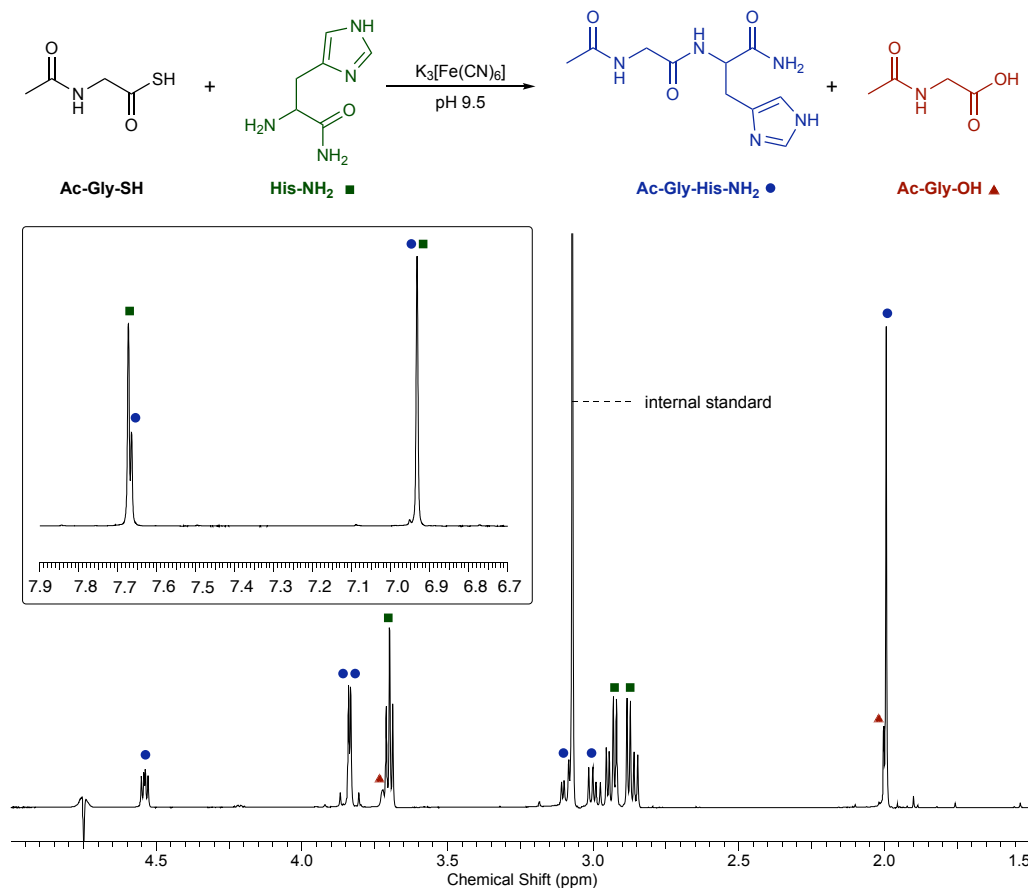
Supplementary Figure 127. 1H NMR (600 MHz, DMSO- d_6 , 1.5–4.2 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-histidine (**His**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature, after the solution was lyophilised and the resultant residue was dissolved in DMSO- d_6 . Yields were obtained prior to lyophilisation. Residual **His** was not observed following lyophilisation (in DMSO- d_6), but was observed prior to lyophilisation in the NMR spectra acquired in H_2O/D_2O (not shown). Inset: 1H NMR (600 MHz, DMSO- d_6 , 6.4–8.4 ppm) showing the aromatic CH and amide NH resonances for *N*-acetylglycylhistidine **Ac-Gly-His-OH**.

1H NMR (600 MHz, DMSO- d_6) *N*-Acetylglycylhistidine **Ac-Gly-His-OH** (●): 8.31 (t, $J = 5.8$ Hz, 1H, Gly-(C2)–NH), 7.59 (d, $J = 6.7$ Hz, 1H, His-(C2)–NH), 7.41 (s, 1H, Ar), 6.62 (s, 1H, Ar), 3.96 (app. q, $J = 5.8$ Hz, 1H, His-(C2)–H), 3.64 (ABX, $J = 5.8, 16.5$ Hz, 1H, Gly-(C2)–H), 3.61 (ABX, $J = 5.8, 16.5$ Hz, 1H, Gly-(C2')–H), 2.90 (ABX, $J = 5.7, 14.7$ Hz, 1H, His-(C3)–H), 2.84 (ABX, $J = 5.6, 14.7$ Hz, 1H, His-(C3')–H), 1.85 (s, 3H, (COCH₃)). *N*-Acetylglycine **Ac-Gly-OH** (▲) δ 3.38 (app. d, $J = 5.0$ Hz, 2H, (C2)–H), 1.81 (s, 3H, (COCH₃)).



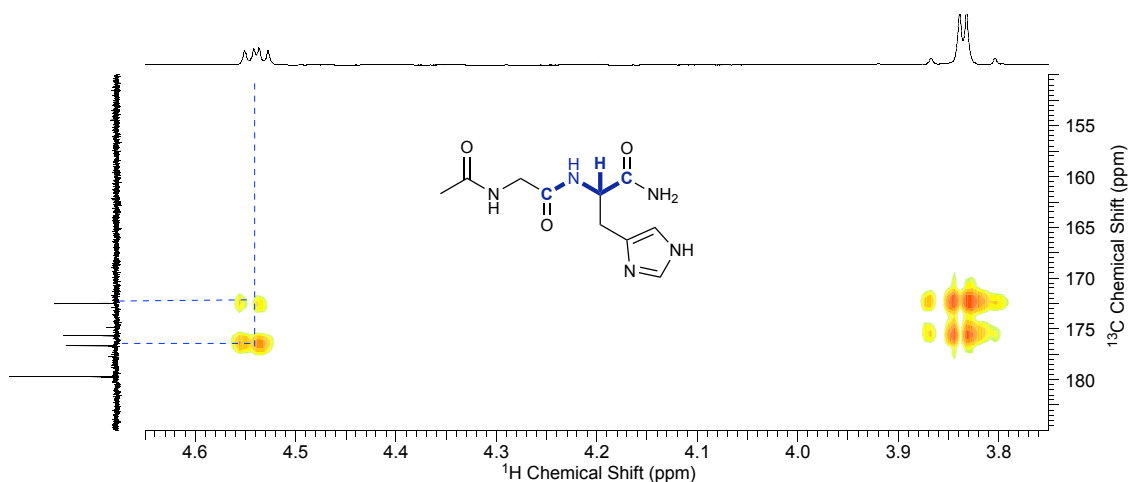
Supplementary Figure 128. 1H - ^{13}C HMBC (1H -600 MHz [3.3–4.1 ppm], ^{13}C -151 MHz [150–185 ppm], DMSO- d_6) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of His-(C2)–H of **Ac-Gly-His-OH** at 3.96 ppm with two C=O resonances at 167.9 and 173.27 ppm, which is characteristic of peptide bond formation of **His**. See Supplementary Figure 127 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylhistidinamide **Ac-Gly-His-NH₂**



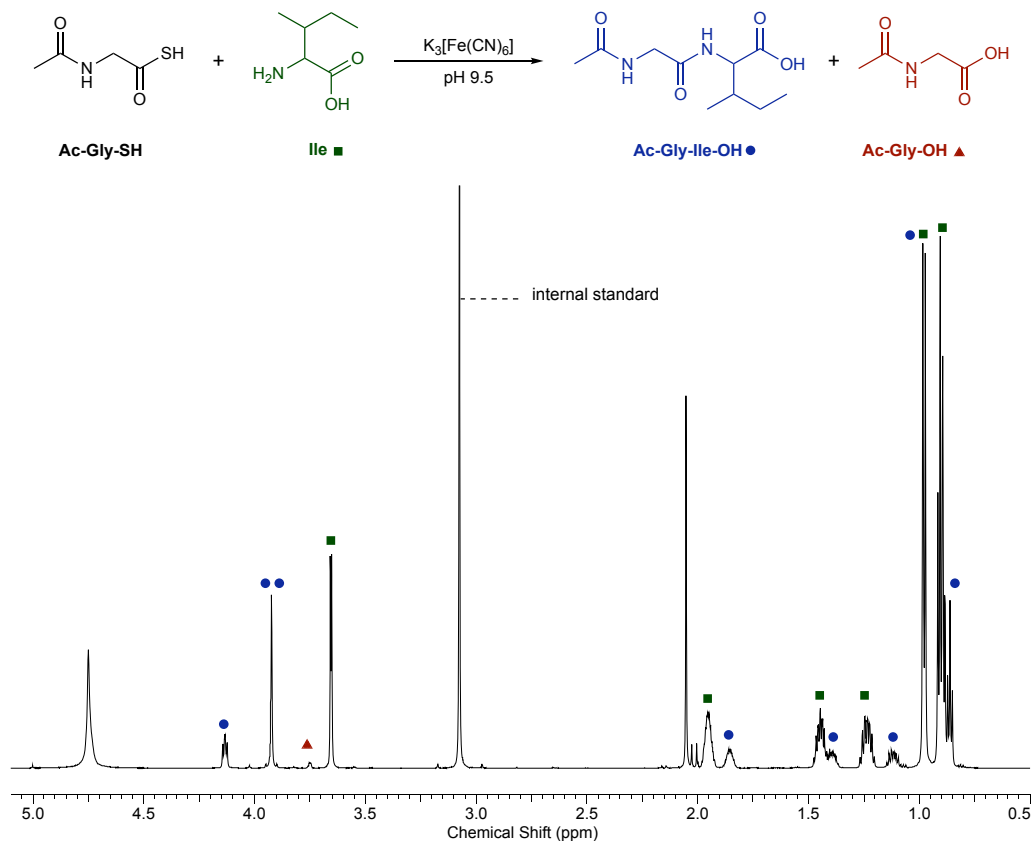
Supplementary Figure 129. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-histidinamide (**His-NH₂**, 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature. Inset: ¹H NMR (600 MHz, H₂O/D₂O 98:2, 6.7–7.9 ppm) showing the aromatic resonances for *N*-acetylglycylhistidinamide **Ac-Gly-His-NH₂** and residual **His-NH₂**.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylhistidinamide **Ac-Gly-His-NH₂** (●): δ 7.66 (s, 1H, Ar), 6.93 (s, 1H, Ar), 4.54 (dd, $J = 5.4, 8.6$ Hz, 1H, His-(C2)–H), 3.85 (AB, $J = 17.0$ Hz, 1H, Gly-(C2)–H), 3.82 (AB, $J = 17.0$ Hz, 1H, Gly-(C2')–H), 3.09 (ABX, $J = 5.4, 14.8$ Hz, 1H, His-(C3)–H), 3.00 (ABX, $J = 8.6, 14.8$ Hz, 1H, His-(C3')–H), 1.99 (s, 3H, (COCH₃)). *L*-Histidinamide **His-NH₂** (■): δ 7.67 (s, 1H, Ar), 6.93 (s, 1H, Ar), 3.70 (app. t, $J = 6.6$ Hz, 1H, (C2)–H), 2.94 (ABX, $J = 6.1, 14.6$ Hz, 1H, (C3)–H), 2.87 (ABX, $J = 7.0, 14.6$ Hz, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲) δ 3.72 (br. s, 2H, (C2)–H₂), 2.00 (s., 3H, (COCH₃)).



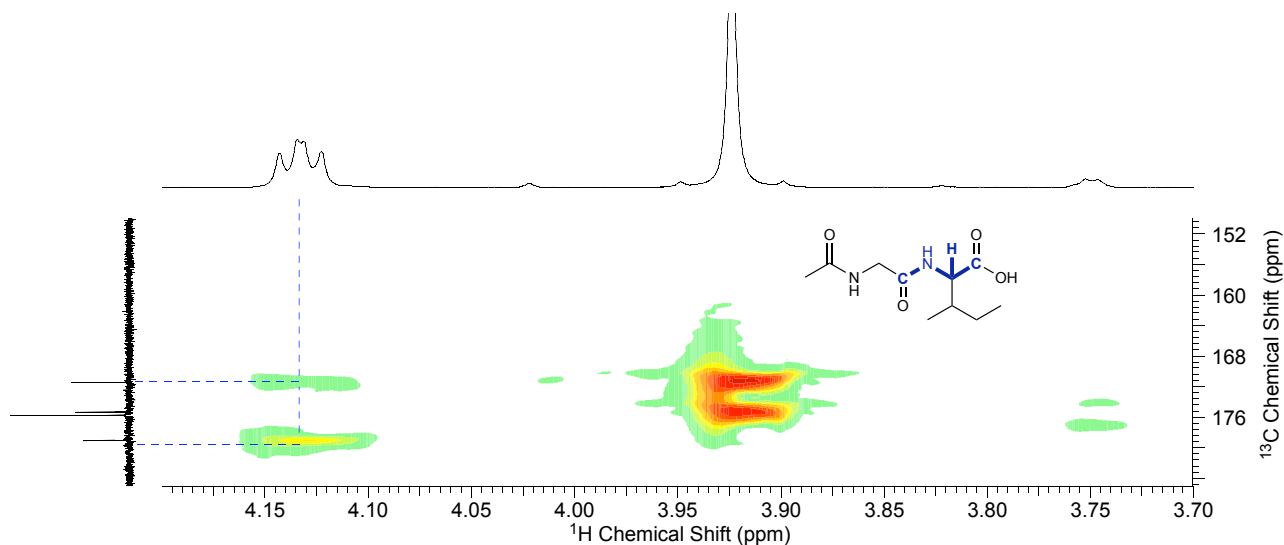
Supplementary Figure 130. ¹H–¹³C HMBC (¹H–600 MHz [3.75–4.65 ppm], ¹³C–151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of His-(C2)–H of **Ac-Gly-His-NH₂** at 4.54 ppm with two C=O resonances at 172.5 and 176.3 ppm, which is characteristic of peptide bond formation of L-histidinamide **His-NH₂**. See Supplementary Figure 129 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylisoleucine **Ac-Gly-Ile-OH**



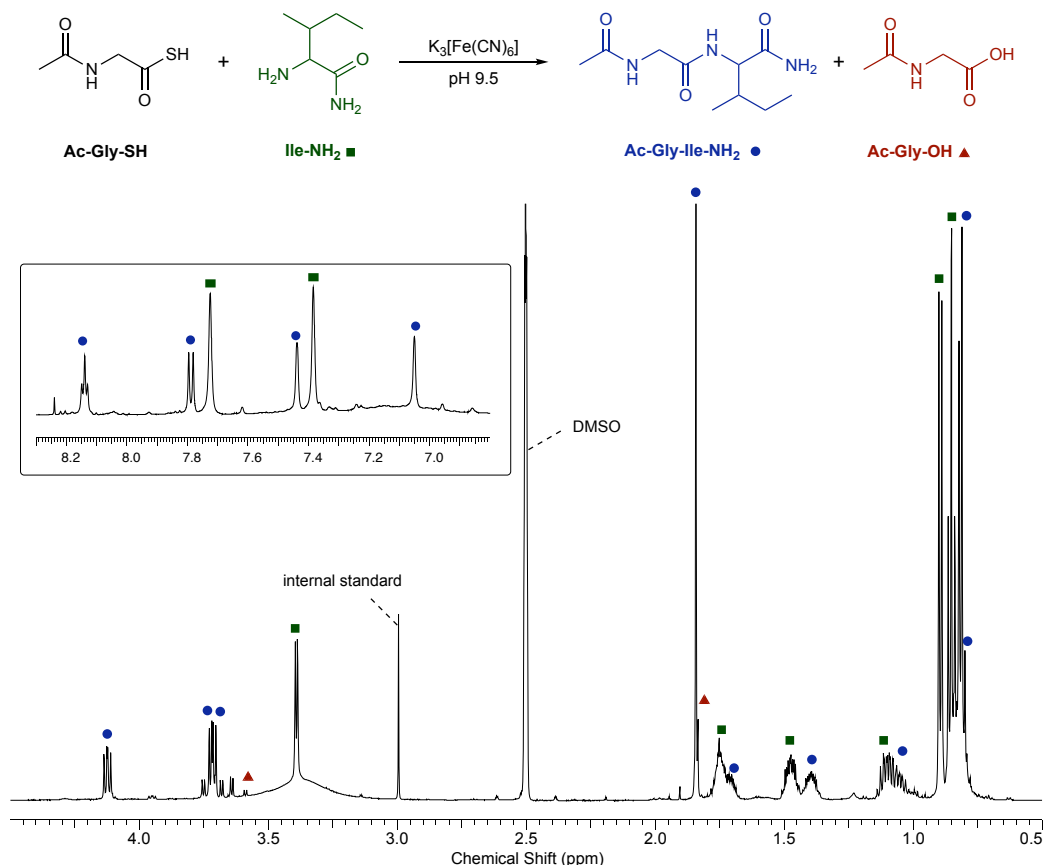
Supplementary Figure 131. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-isoleucine (**Ile**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycylisoleucine **Ac-Gly-Ile-OH** (●): δ 4.13 (dd, $J = 6.1, 8.1$ Hz, 1H, Ile-(C2)-H), 3.94 (AB, $J = 17.3$ Hz, 1H, Gly-(C2)-H), 3.91 (AB, $J = 17.3$ Hz, 1H, Gly-(C2')-H), 2.05 (s, 3H, (COCH₃), 1.90 - 1.82 (m, 1H, Ile-(C3)-H), 1.42 - 1.35 (m, 1H, Ile-(C4)-H), 1.16 - 1.08 (m, 1H, Ile-(C4')-H), 0.98 (d, $J = 7.0$ Hz, 3H, Ile-(C3)-CH₃), 0.86 (t, $J = 7.4$ Hz, 3H, Ile-(C5)-H₃). *L*-Isoleucine **Ile** (■): δ 3.66 (d, $J = 4.0$ Hz, 1H, (C2)-H), 1.99 - 1.92 (m, 1H, (C3)-H), 1.45 (ddd, $J = 5.1, 7.4, 13.4$ Hz, 1H, (C4)-H), 1.23 (ddd, $J = 7.4, 9.0, 13.4$ Hz, 1H, (C4')-H), 0.98 (d, $J = 7.0$ Hz, 3H, (C3)-H₃), 0.91 (t, $J = 7.5$ Hz, 3H, (C5)-H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲) (partial assignment): δ 3.75 (app. d, $J = 3.9$ Hz, 2H, (C2)-H₂).



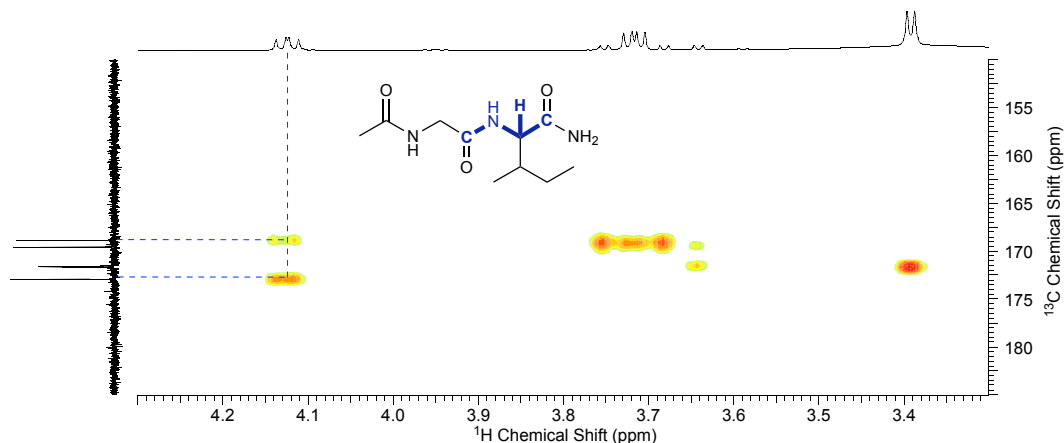
Supplementary Figure 132. 1H - ^{13}C HMBC (1H -600 MHz [3.7–4.2 ppm], ^{13}C -151 MHz [150–185 ppm], H_2O/D_2O 98:2) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Ile-(C2)-H of **Ac-Gly-Ile-OH** at 4.13 ppm with two C=O resonances at 172.4 and 180.0 ppm, which is characteristic of peptide bond formation of **Ile**. See Supplementary Figure 131 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylisoleucinamide **Ac-Gly-Ile-NH₂**



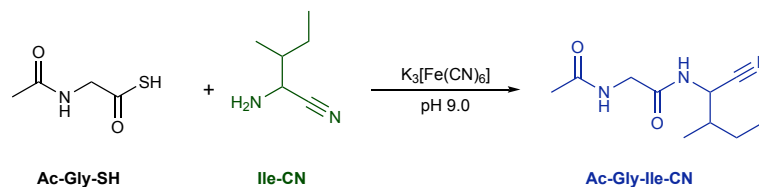
Supplementary Figure 133. ¹H NMR (600 MHz, DMSO-*d*₆, 0.5–4.5 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-isoleucinamide (**Ile-NH₂**, 150 mM), and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature, after the solution was lyophilised and the resultant residue was dissolved in DMSO-*d*₆. Yields were obtained prior to lyophilisation. Inset: ¹H NMR (600 MHz, DMSO-*d*₆, 6.9–8.3 ppm) showing the NH resonances for *N*-acetylglycylisoleucinamide **Ac-Gly-Ile-NH₂**.

¹H NMR (600 MHz, DMSO-*d*₆) *N*-Acetylglycylisoleucinamide **Ac-Gly-Ile-NH₂** (●): δ 8.14 (t, *J* = 5.8 Hz, 1H, Gly-(C2)–NH), 7.79 (d, *J* = 9.0 Hz, 1H, Ile-(C2)–NH), 7.44 (s, 1H, Ile-(C1)–NH), 7.05 (br. s., 1H, Ile-(C1')–H), 4.12 (dd, *J* = 6.9, 9.0 Hz, 1H, Ile-(C2)–H), 3.74 (ABX, *J* = 5.8, 16.5 Hz, 1H, Gly-(C2)–H), 3.70 (ABX, *J* = 5.8, 16.5 Hz, 1H, Gly-(C2')–H), 1.84 (s, 3H, (COCH₃)), 1.73–1.68 (m, 1H, Ile-(C3)–H), 1.43–1.36 (m, 1H, Ile-(C4)–H), 1.08–1.02 (m, 1H, Ile-(C4')–H), 0.82 (d, *J* = 7.0 Hz, 3H, Ile-(C3)–CH₃), 0.83–0.78 (m, 3H, Ile-(C5)–H₃). *L*-Isoleucinamide **Ile-NH₂** (■): δ 7.73 (br. s., 1H, (C1)–NH), 7.39 (br. s., 1H, (C1')–NH), 3.39 (d, *J* = 5.3 Hz, 1H, (C2)–H), 1.78–1.72 (m, 1H, (C3)–H), 1.52–1.43 (m, 1H, (C4)–H), 1.15–1.06 (m, 1H, (C4')–H), 0.89 (d, *J* = 6.9 Hz, 3H, (C3)–CH₃), 0.85 (t, *J* = 7.4 Hz, 3H, (C5)–H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.59 (app. d, *J* = 5.7 Hz, 2H, (C2)–H₂), 1.83 (s, 3H, (COCH₃)).

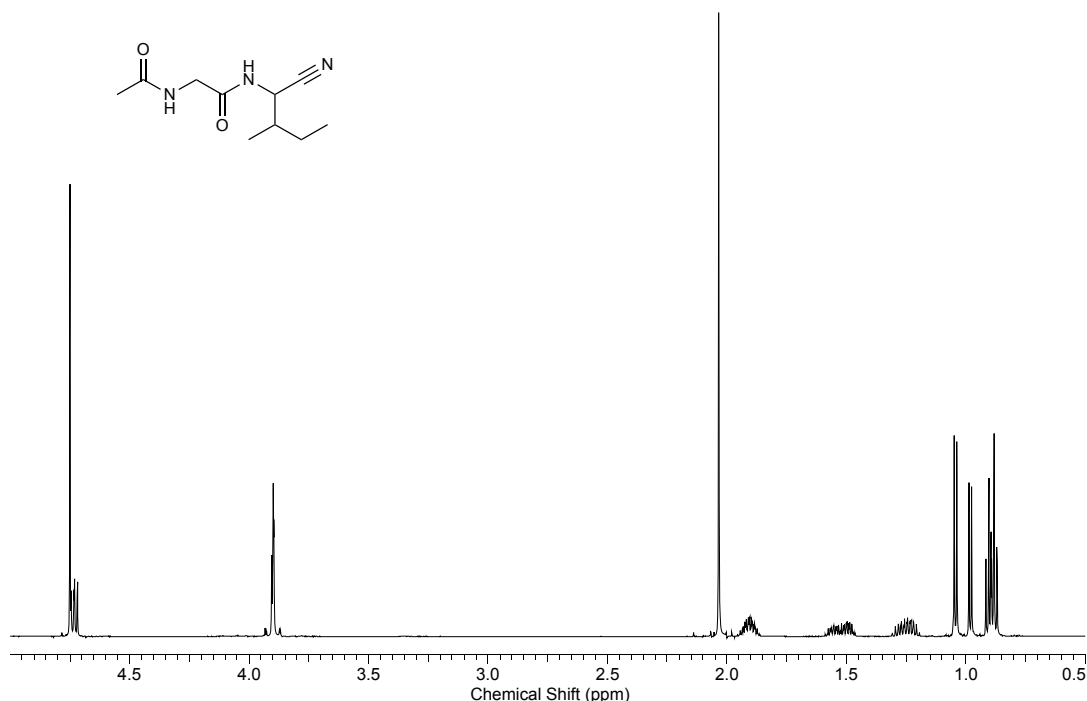


Supplementary Figure 134. ¹H–¹³C HMBC (¹H–600 MHz [3.3–4.3 ppm], ¹³C–151 MHz [150–185 ppm], DMSO-*d*₆) spectrum showing the diagnostic ²*J*_{CH} and ³*J*_{CH} coupling of Ile-(C2)–H of **Ac-Gly-Ile-NH₂** at 4.12 ppm with two C=O resonances at 168.8 and 172.9 ppm, which is characteristic of peptide bond formation of **Ile-NH₂**. See Supplementary Figure 133 for expanded and labelled ¹H NMR spectrum.

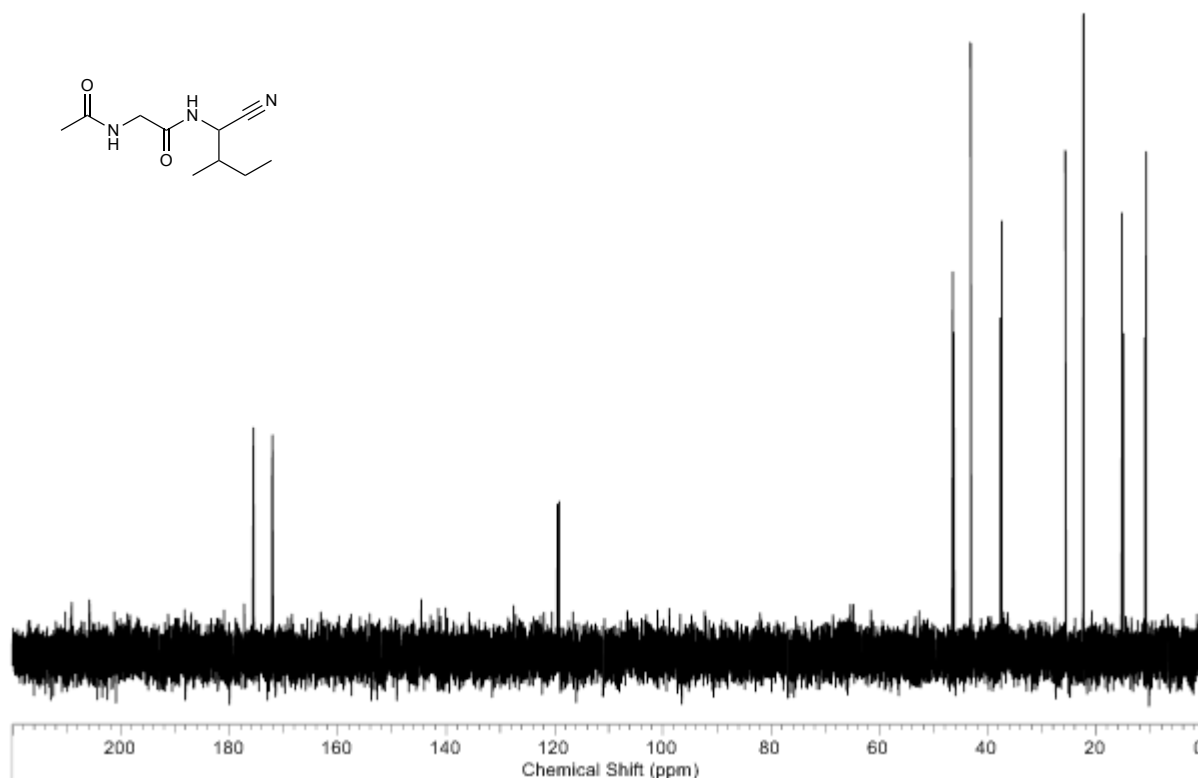
N-Acetylglycylisoleucine nitrile **Ac-Gly-Ile-CN**



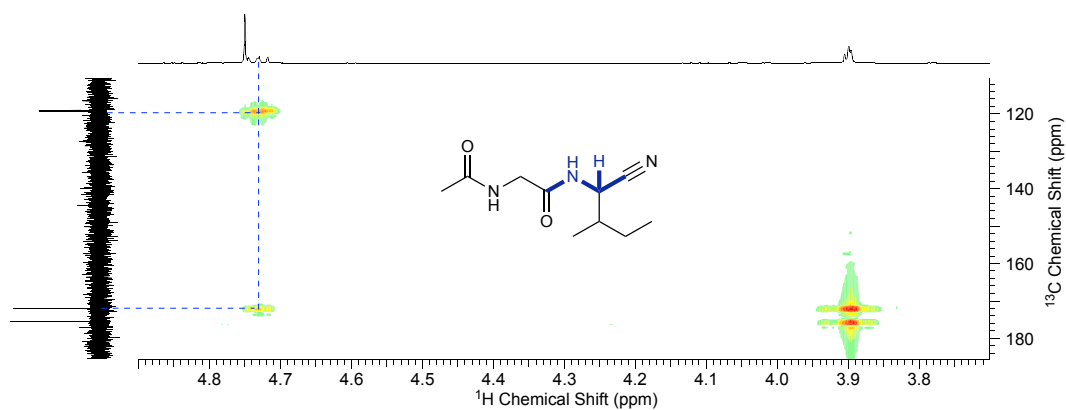
Prepared following *general procedure G* on a 0.1 mmol scale, from sodium-2-acetamido-ethanethioate **Ac-Gly-S-Na⁺** (15.5 mg, 0.1 mmol) and *rac*-isoleucine nitrile hydrochloride **Ile-CN·HCl**³⁹ (29.7 mg, 0.2 mmol). **Ac-Gly-Ile-CN** was isolated as a white solid (17.3 mg, 82%, as a mixture of two diastereoisomers a:b, 57:43) after flash column chromatography (SiO₂, EtOAc (100%)). *R_f* 0.32 (100% EtOAc). ¹H NMR (600 MHz, D₂O, diastereoisomer a (major)) δ 4.72 (d, *J* = 7.2 Hz, 1H, Ile-(C2)-H), 3.92 (AB, *J* = 16.9 Hz, 1H, Gly-(C2)-H), 3.88 (d, *J* = 16.9 Hz, 1H, Gly-(C2)-H'), 2.03 (s, 3H, (COCH₃)), 1.95 - 1.86 (m, 1H, Ile-(C3)-H), 1.59 - 1.46 (m, 1H, Ile-(C4)-H), 1.32 - 1.19 (m, 1H, Ile-(C4')-H), 1.04 (d, *J* = 6.8 Hz, 3H, Ile-(C3)-CH₃), 0.88 (t, *J* = 7.5 Hz, 3H, Ile-(C5)-H₃). ¹³C NMR (151 MHz, D₂O, diastereoisomer a (major)) δ 175.6 (COCH₃), 172.0 (Gly-(C1)), 119.1 (Ile-(C1)), 46.5 (Ile-(C2)), 43.2 (Gly-(C2)), 37.4 (Ile-(C3)), 25.6 (Ile-(C4)), 22.3 (COCH₃), 15.3 (Ile-(C3)-CH₃), 10.8 (Ile-(C5)). ¹H NMR (600 MHz, D₂O, diastereoisomer b (minor)) δ 4.74 (d, *J* = 7.1 Hz, 1H, Ile-(C2)-H), 3.91 (AB, *J* = 17.0 Hz, 1H, Gly-(C2)-H), 3.87 (d, *J* = 17.0 Hz, 1H, Gly-(C2)-H'), 2.03 (s, 3H, (COCH₃)), 1.95 - 1.86 (m, 1H, Ile-(C3)-H), 1.59 - 1.46 (m, 1H, Ile-(C4)-H), 1.32 - 1.19 (m, 1H, Ile-(C4')-H), 0.98 (d, *J* = 6.8 Hz, 3H, Ile-(C3)-CH₃), 0.90 (t, *J* = 7.4 Hz, 3H, Ile-(C5)-H₃). ¹³C NMR (151 MHz, D₂O, diastereoisomer b (minor)) δ 175.6 (COCH₃), 172.1 (Gly-(C1)), 119.4 (Ile-(C1)), 46.3 (Ile-(C2)), 43.2 (Gly-(C2)), 37.7 (Ile-(C3)), 25.6 (Ile-(C4)), 22.3 (COCH₃), 15.0, (Ile-(C3)-CH₃), 11.0 (Ile-(C5)). HRMS-ESI-MS [*M*+H]⁺ calcd. for C₁₀H₁₈N₃O₂ 212.1394; obs. 212.1391. IR (solid, cm⁻¹) 3315, 3207, 3026, 2463, 2358, 2331, 1670, 1629, 1556.



Supplementary Figure 135. ¹H NMR (600 MHz, D₂O, 0.5–5.0 ppm) spectrum of purified *N*-acetylglycylisoleucine nitrile **Ac-Gly-Ile-CN**

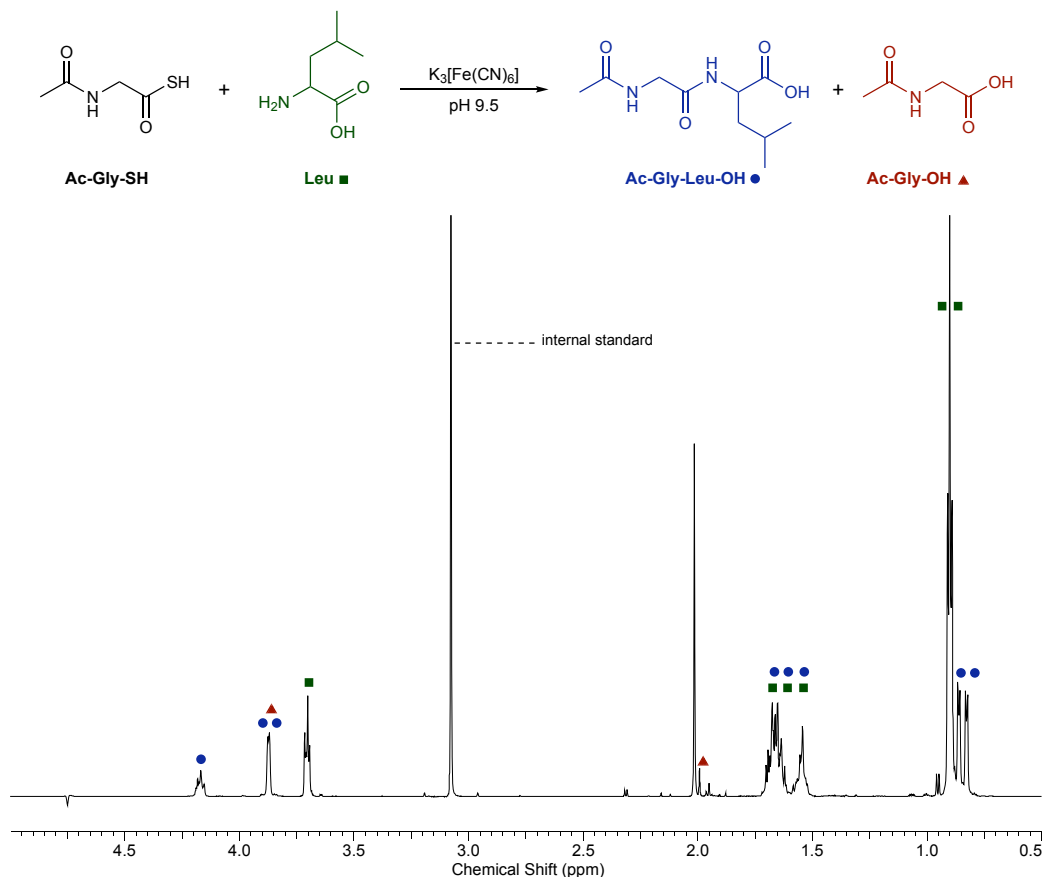


Supplementary Figure 136. ¹H NMR (600 MHz, D₂O, 0–220 ppm) spectrum of purified *N*-acetylglycylisoleucine nitrile **Ac-Gly-Ile-CN**



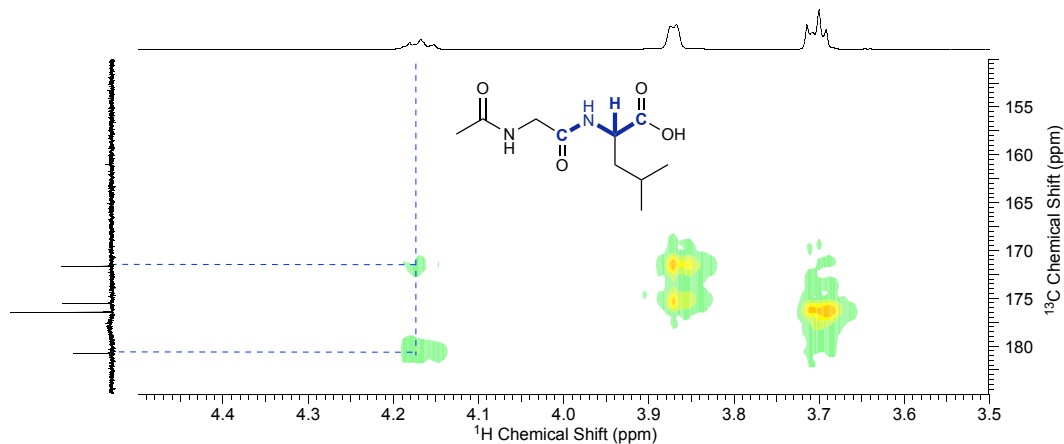
Supplementary Figure 137. ¹H–¹³C HMBC (¹H–600 MHz [3.7–4.9 ppm], ¹³C–151 MHz [110–185 ppm], D₂O) spectrum showing the diagnostic ²J_{CH} and ³J_{CH} coupling of Ile-(C2)-H of **Ac-Gly-Ile-CN** at 4.72–4.74 ppm with CN resonances at 119.1 and 119.4 ppm and C=O resonances at 172.0 and 172.1 ppm, which is characteristic of peptide bond formation of **Ile-CN**. See Supplementary Figure 135 for expanded ¹H NMR spectrum.

N-Acetylglycylleucine **Ac-Gly-Leu-OH**



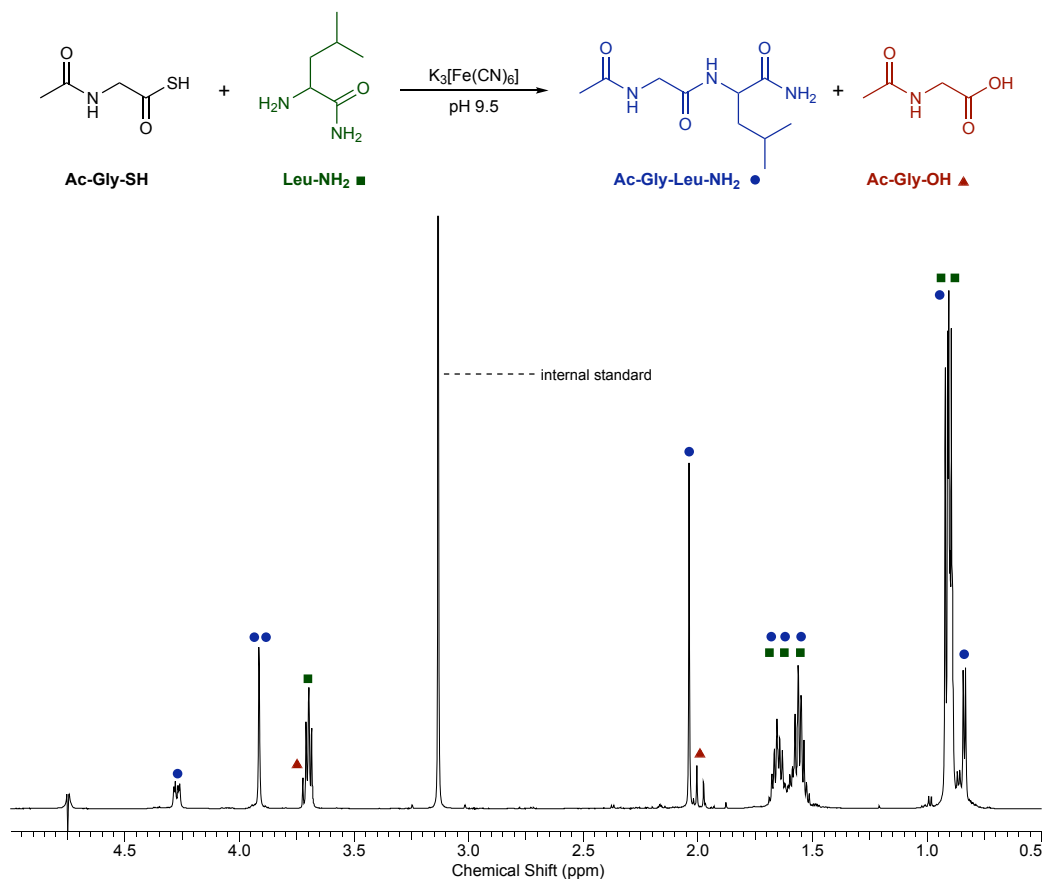
Supplementary Figure 138. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-leucine (**Leu**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycylleucine **Ac-Gly-Leu-OH** (●): δ 4.20 - 4.14 (m, 1H, Leu-(C2)-H), 3.89 (AB, $J = 17.9$ Hz, 1H, Gly-(C2)-H), 3.86 (AB, $J = 17.9$ Hz, 1H, Gly-(C2')-H), 2.01 (s, 3H, (COCH₃)), 1.72 - 1.50 (m, 3H, Leu-(C3)-H₂, Leu-(C4)-H), 0.86 (d, $J = 5.8$ Hz, 3H, Leu-(C5)-H₃), 0.83 (d, $J = 5.6$ Hz, 3H, Leu-(C5)-H₃). L-Leucine **Leu** (■): δ 3.73 - 3.67 (m, 1H, (C2)-H), 1.72 - 1.50 (m, 3H, (C3)-H, (C4)-H₂), 0.91 (d, $J = 5.8$ Hz, 3H, (C5)-H₃), 0.90 (d, $J = 5.8$ Hz, 3H, (C5')-H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.73 - 3.67 (obs., 2H), 1.99 (s, 3H, (COCH₃)).



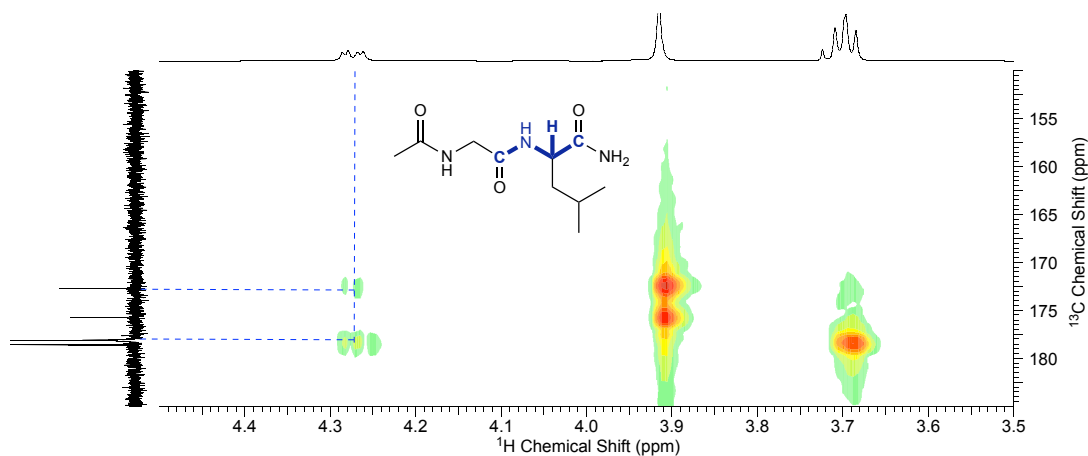
Supplementary Figure 139. 1H - ^{13}C HMBC (1H -600 MHz [3.5–4.5 ppm], ^{13}C -151 MHz [150–185 ppm], H_2O/D_2O 98:2) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Leu-(C2)-H of **Ac-Gly-Leu-OH** centred at 4.17 ppm with two C=O resonances at 171.6 and 180.7 ppm, which is characteristic of peptide bond formation of **Leu**. See Supplementary Figure 138 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylleucinamide **Ac-Gly-Leu-NH₂**



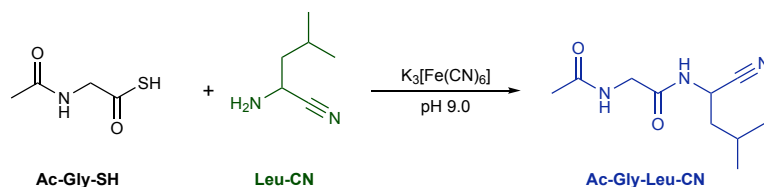
Supplementary Figure 140. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-leucinamide (**Leu-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylleucinamide **Ac-Gly-Leu-NH₂** (●): δ 4.27 (dd, $J = 4.0, 10.6$ Hz, 1H, Leu-(C2)–H), 3.92 (s, 2H, Gly-(C2)–H₂), 2.04 (s, 3H, (COCH₃), 1.70 - 1.47 (m, 3H, Leu-(C3)–H, Leu-(C4)–H₂), 0.91 - 0.89 (obs., 3H, Leu-(C5)–H₃), 0.84 (d, $J = 6.1$ Hz, 3H, Leu-(C5')–H₃). *L*-Leucinamide **Leu-NH₂** (■): δ 3.70 (dd, $J = 6.7, 7.9$ Hz, 1H, (C2)–H), 1.70 - 1.50 (m, 3H, Leu-(C3)–H, Leu-(C4)–H₂), 0.92 (d, $J = 6.6$ Hz, 3H, Leu-(C5)–H₃), 0.90 (d, $J = 6.6$ Hz, 3H, Leu-(C5')–H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.73 - 3.71 (s, 2H, (C2)–H₂), 2.00 (s, 3H, (COCH₃).

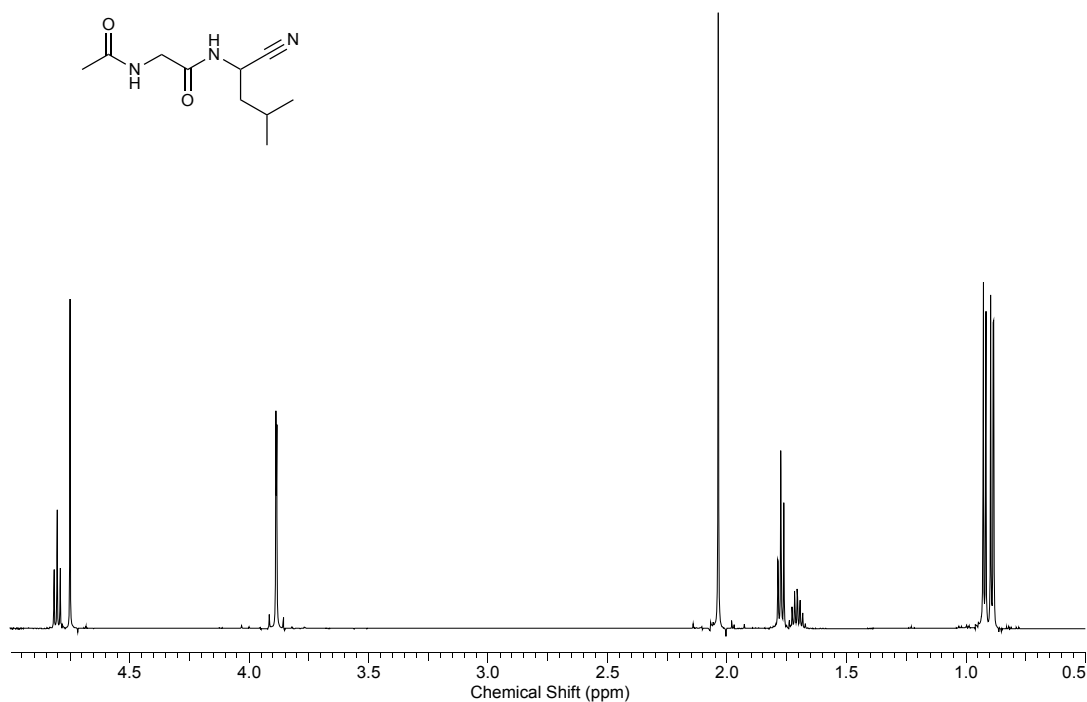


Supplementary Figure 141. ¹H–¹³C HMBC (¹H–600 MHz [3.5–4.5 ppm], ¹³C–151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Leu-(C2)–H of **Ac-Gly-Leu-NH₂** at 4.27 ppm with two C=O resonances at 172.7 and 178.1 ppm, which is characteristic of peptide bond formation of **Leu-NH₂**. See Supplementary Figure 140 for expanded and labelled ¹H NMR spectrum.

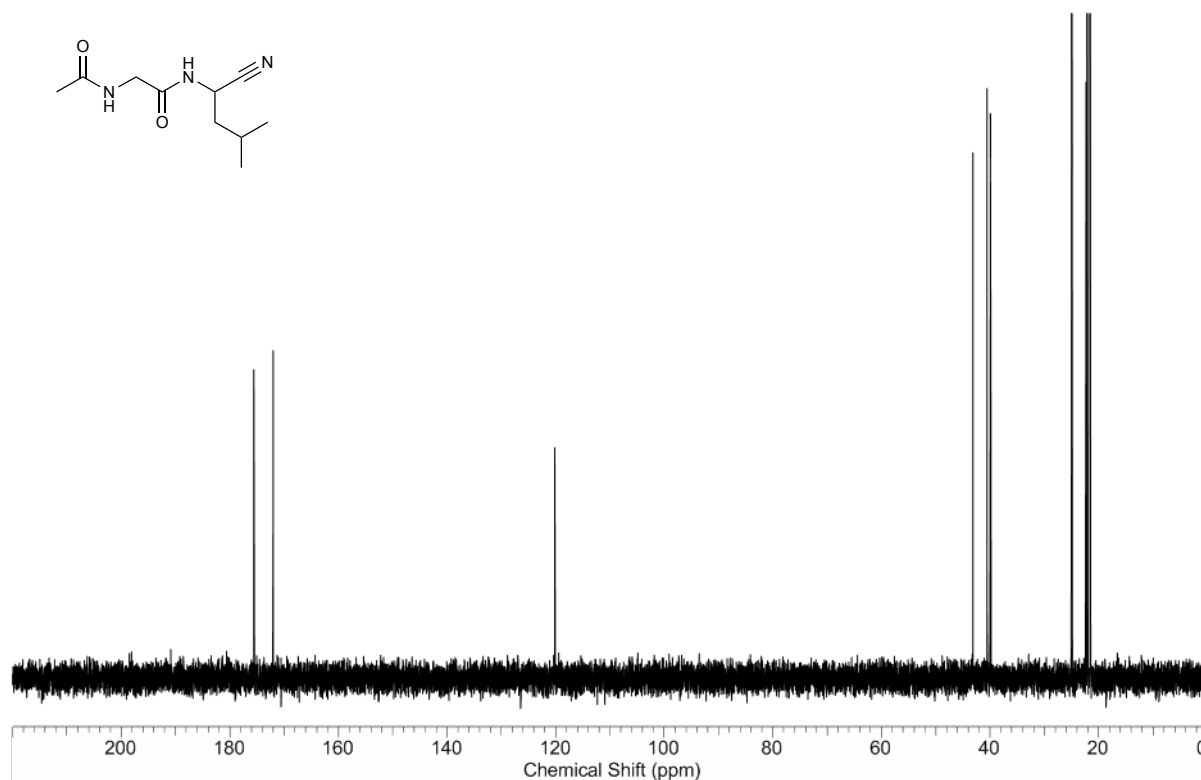
N-Acetylglycylleucine nitrile **Ac-Gly-Leu-CN**



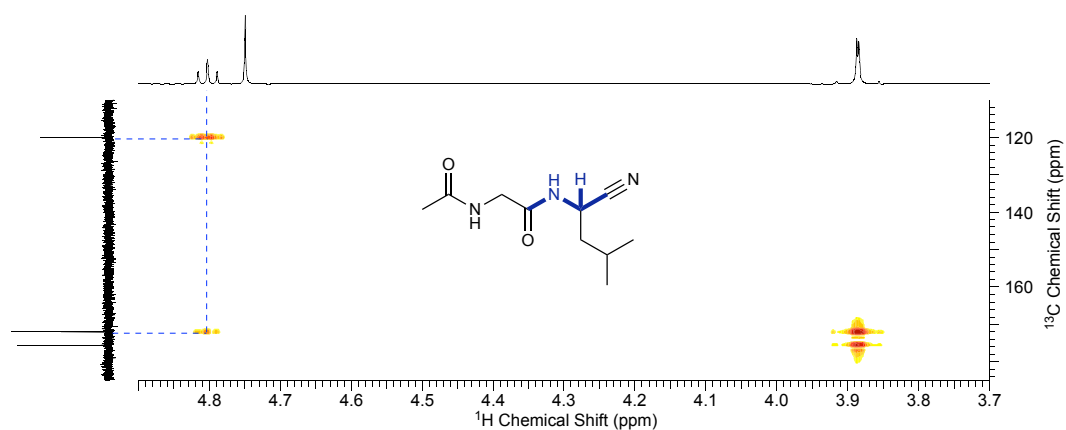
Prepared following *general procedure G* on a 0.1 mmol scale, from sodium-2-acetamido-ethanethioate **Ac-Gly-S-Na⁺** (15.5 mg, 0.1 mmol) and *rac*-leucine nitrile hydrochloride **Leu-CN·HCl**³⁹ (29.7 mg, 0.2 mmol). **Ac-Gly-Leu-CN** was isolated as a white solid (21.1 mg, 90%) after flash column chromatography (SiO₂, EtOAc (100%)). *R_f* 0.20 (100% EtOAc). ¹H NMR (600 MHz, D₂O) δ 4.80 (t, *J* = 7.9 Hz, 1H, Leu-(C2)-H), 3.90 (AB, *J* = 17.1 Hz, 1H, Gly-(C2)-H), 3.87 (AB, *J* = 17.1 Hz, 1H, Gly-(C2)-H'), 1.80 - 1.75 (m, 2H, Leu-(C3)-H₂), 1.74 - 1.67 (m, 1H, Leu-(C4)-H), 0.92 (d, *J* = 6.6 Hz, 3H, Leu-(C5)-H₃), 0.89 (d, *J* = 6.5 Hz, 3H, Leu-(C5')-H₃). ¹³C NMR (151 MHz, D₂O) δ 175.5 (COCH₃), 172.1 (Gly-(C1)), 120.1 (Leu-(C1)), 43.2 (Gly-(C2)), 40.5 (Leu-(C3)), 39.9 (Leu-(C2)), 25.0 (Leu-(C4)), 22.4 (COCH₃), 22.1 (Leu-(C5)), 21.6 (Leu-(C5')). HRMS-ESI-MS [*M*+H]⁺ calcd. for C₁₀H₁₈N₃O₂ 212.1394; obs. 212.1391. IR (solid, cm⁻¹) 3309, 3207, 3045, 2970, 2463, 2360, 2331, 1672, 1629, 1556.



Supplementary Figure 142. ¹H NMR (600 MHz, D₂O, 1.0–5.0 ppm) spectrum of purified *N*-acetylglycylleucine nitrile **Ac-Gly-Leu-CN**

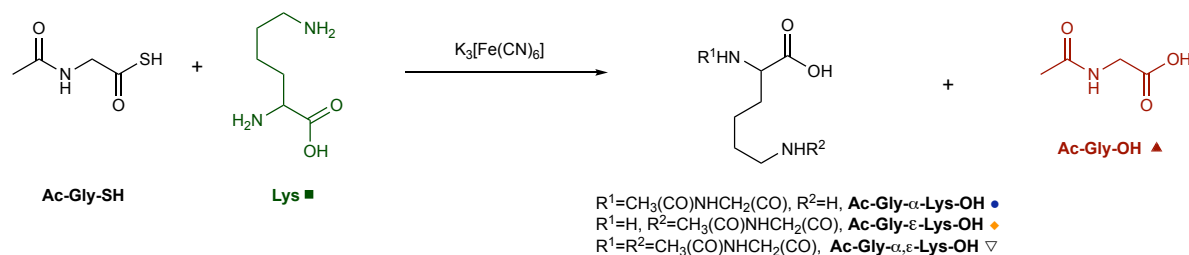


Supplementary Figure 143. ^1H NMR (600 MHz, D_2O , 0–220 ppm) spectrum of purified *N*-acetylglycylleucine nitrile **Ac-Gly-Leu-CN**

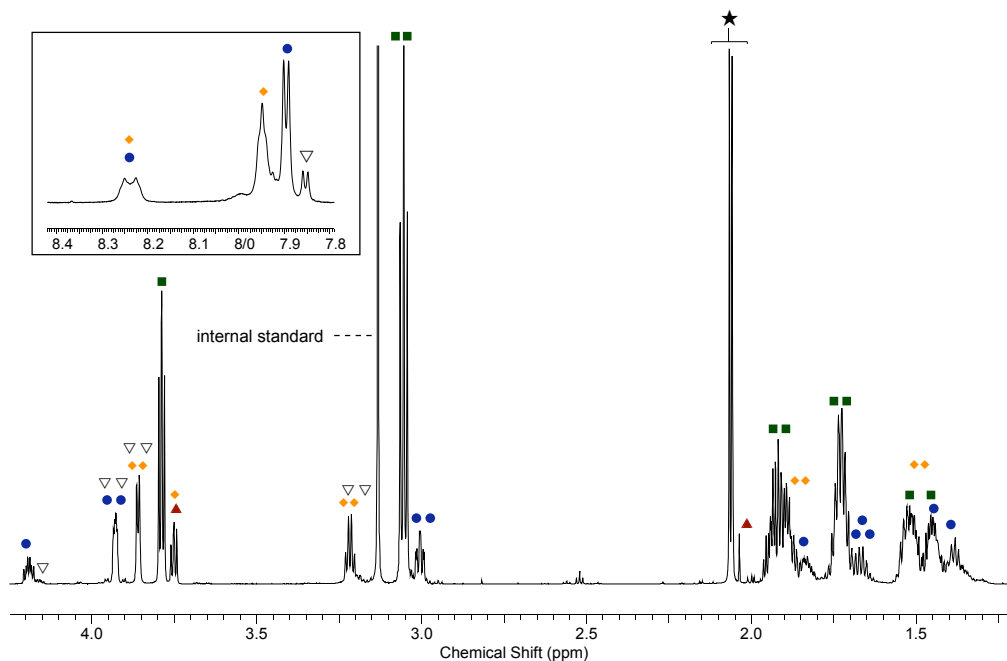


Supplementary Figure 144. ^1H – ^{13}C HMBC (^1H –600 MHz [3.7–4.9 ppm], ^{13}C –151 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Leu-(C2)–H of **Ac-Gly-Leu-CN** at 4.80 ppm with CN resonance at 120.1 ppm and a C=O resonance at 172.1 ppm, which is characteristic of peptide bond formation of **Leu-CN**. See Supplementary Figure 142 for expanded ^1H NMR spectrum.

N-Acetylglycyllysine **Ac-Gly-Lys-OH**



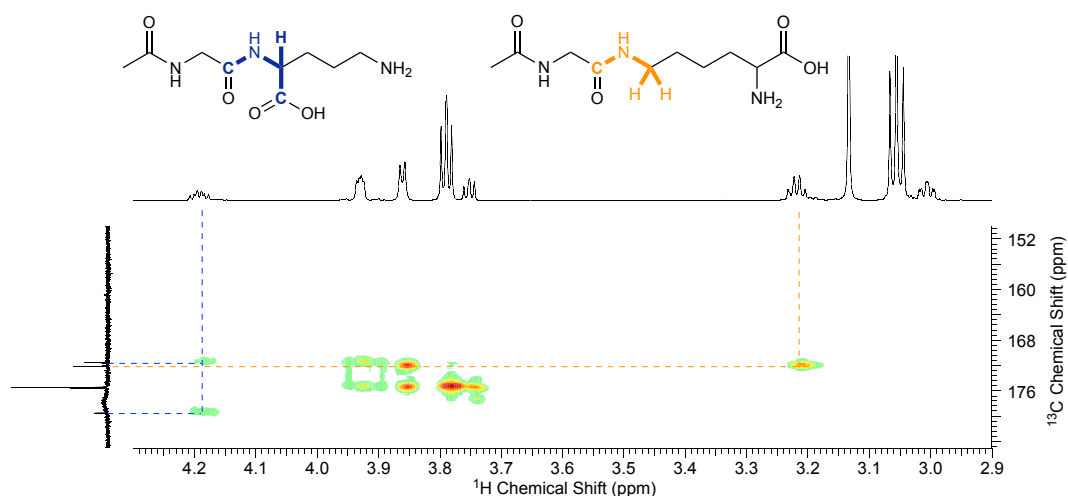
L-Lysine hydrochloride **Lys**-HCl (27.4 mg, 0.15 mmol) was dissolved in H₂O/D₂O (98:2; 0.5 mL) and the solution was adjusted to pH 9.5, 8.5, 7.5 or 6.5, as specified, with 4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and total volume was adjusted to 1 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added and the solution was stirred at room temperature for 20 min whilst maintaining the solution at the specified initial pH value with 4 M HCl/NaOH. The solution was then adjusted to pH 9.5 for analysis, and the reaction mixture was centrifuged. The supernatant was analysed by 1D and 2D NMR spectroscopy, and quantified using methylsulfonylmethane (MSM) as an internal standard. The formation of ligation products was confirmed by ¹H–¹³C HMBC NMR spectral analysis and high resolution mass spectrometry. See Supplementary Table 8 on page S53 for HRMS-ESI data, and Supplementary Table 11 (page S118) for reaction yields and product distribution at specified pH values.



Supplementary Figure 145. ¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.25–4.20 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-lysine (**Lys**; 150 mM) and K₃[Fe(CN)₆] (150 mM) at pH 9.5 and room temperature. Inset: ¹H NMR (700 MHz, H₂O/D₂O 98:2, 7.8–8.4 ppm) showing the NH amide resonances, which exhibit partial deuteration. COCH₃ resonances for **Ac-Gly-α-Lys-OH**, **Ac-Gly-ε-Lys-OH** and **Ac-Gly-α,ε-Lys-OH** (denoted with a ★ between 2.1–2.2 ppm) could not be unambiguously assigned.

¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*₂-(Acetylglycyl)lysine **Ac-Gly-α-Lys-OH** (•) (partial assignment): δ 8.27–8.18 (m, 1H, Gly-NH), 7.89 (d, *J* = 7.4 Hz, 1H, Lys-NH_α), 4.19 (dt, *J* = 4.8, 8.1 Hz, 1H, Lys-(C2)–H), 3.94 (ABX, *J* = 5.8, 17.3 Hz, 1H, Gly-(C2)–H), 3.91 (ABX, *J* = 5.4, 17.3 Hz, 1H, Gly-(C2)–H'), 3.02–2.99 (m, 2H, Lys-(C6)–H₂), 1.70–1.67 (m, 1H, Lys-(C3)–H), 1.70–1.63 (m, 3H, Lys-(C3')–H, Lys-(C5)–H₂), 1.55–1.42 (m, 2H, Lys-(C4)–H₂). *N*₆-(Acetylglycyl)lysine **Ac-Gly-ε-Lys-OH** (♦) (partial assignment): δ 8.27–8.18 (m, 1H, Gly-NH), 7.95 (br. t, *J* = 6.3 Hz, 1H, Lys-NH_ε), 3.86 (app. d, *J* = 5.4 Hz, 2H, Gly-(C2)–H₂), 3.75 (t, *J* = 6.3 Hz, 1H, Lys-(C2)–H), 3.22 (app. q, *J* = 6.4 Hz, 2H, Lys-(C6)–H₂), 1.76–1.66 (m, 2H, Lys-(C3)–H₂), 1.56–1.48 (m, 2H, Lys-(C4)–H₂). *N*₂,*N*₆-Di(acetylglycyl)lysine **Ac-**

Gly- α,ϵ -Lys-OH (∇) (partial assignment): δ 7.85 (d, J = 7.6 Hz, 1H, Lys-NH α), 4.16 (dt, J = 4.1, 8.1 Hz, 1H, Lys-(C2)-H), 3.96 - 3.90 (obs., 2H, Gly α -(C2)-H $_2$), 3.87 - 3.82 (obs., 2H, Gly ϵ -(C2)-H $_2$), 3.20 - 3.16 (obs., 2H, Lys-(C6)-H $_2$). **Lysine Lys** ($\blacksquare$): 3.79 (dd, J = 5.7, 6.4 Hz, 1H, (C2)-H), 3.05 (t, J = 7.6 Hz, 2H, (C6)-H $_2$), 1.97 - 1.86 (m, 2H, (C3)-H $_2$), 1.69 - 1.78 (m, 2H, (C5)-H $_2$), 1.56 - 1.49 (m, 1H, (C4)-H), 1.42 - 1.49 (d, 1H, (C4')-H). **N-Acetylglycine Ac-Gly-OH** ($\blacktriangle$) δ 3.75 - 3.73 (obs., 2H, (C2)-H $_2$), 2.04 - 2.02 (m, 3H, (COCH $_3$)).

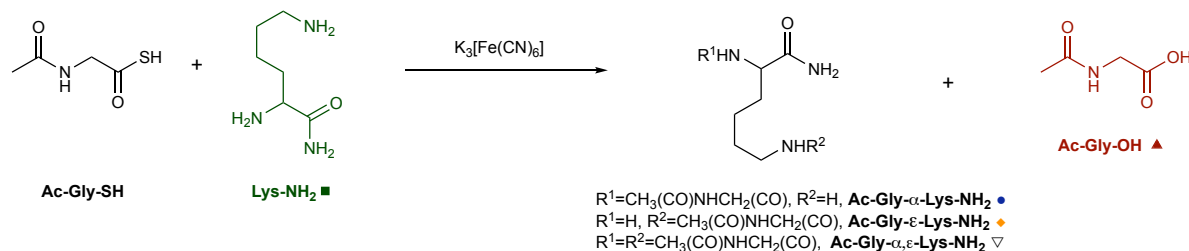


Supplementary Figure 146. ^1H - ^{13}C HMBC (^1H -700 MHz [2.90–4.30 ppm], ^{13}C -176 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Lys-(C2)-H of **Ac-Gly- α -Lys-OH** ($\bullet$) at 4.19 ppm with two C=O resonances at 171.6 and 179.4 ppm (shown in blue), which is characteristic of peptide bond formation of **Lys** by reaction of the α -NH $_2$. The side chain acylation—by reaction of the **Lys** ϵ -NH $_2$ to yield **Ac-Gly- ϵ -Lys-OH** ($\blacklozenge$)—was confirmed by the $^3J_{\text{CH}}$ coupling of Lys-(C6)-H at 3.22 ppm with a C=O resonance at 172.2 ppm (shown in orange). See Supplementary Figure 145 for expanded and labelled ^1H NMR spectrum.

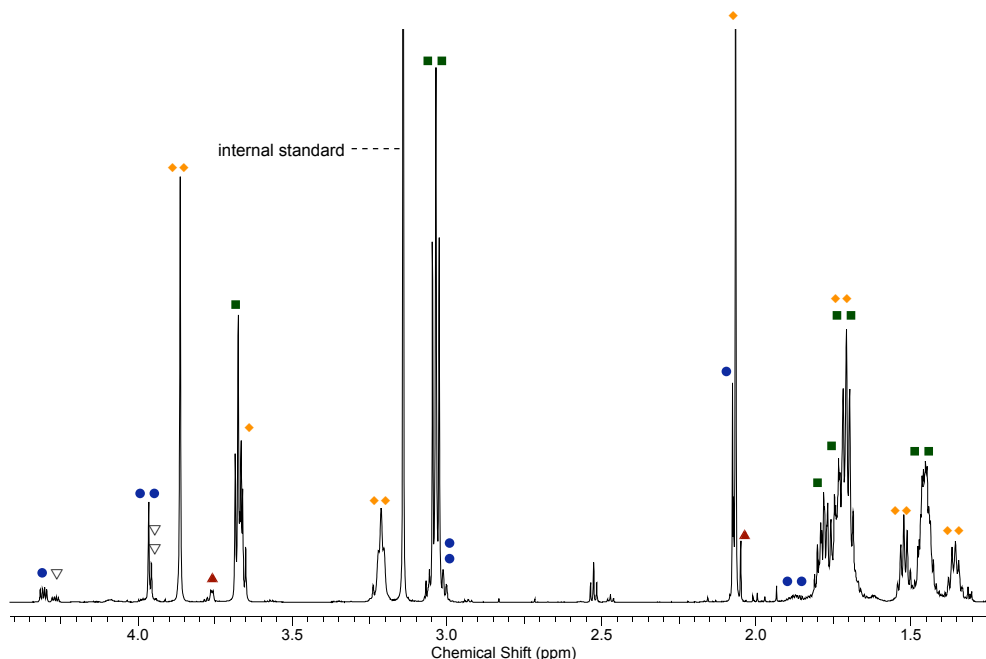
pH	Product	Yield (%)	R ¹ (α)	R ² (ϵ)	α : ϵ acylation
9.5	Ac-Gly- ϵ -Lys-OH ($\blacklozenge$)	46	H	Ac-Gly	1.0 : 1.2
	Ac-Gly- α -Lys-OH ($\bullet$)	41	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-OH (∇)	7	Ac-Gly	Ac-Gly	
	Ac-Gly-OH ($\blacktriangle$)	4	-	-	
8.5	Ac-Gly- ϵ -Lys-OH ($\blacklozenge$)	40	H	Ac-Gly	1.1 : 1.0
	Ac-Gly- α -Lys-OH ($\bullet$)	43	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-OH (∇)	6	Ac-Gly	Ac-Gly	
	Ac-Gly-OH ($\blacktriangle$)	5	-	-	
7.5	Ac-Gly- ϵ -Lys-OH ($\blacklozenge$)	36	H	Ac-Gly	1.2 : 1.0
	Ac-Gly- α -Lys-OH ($\bullet$)	42	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-OH (∇)	6	Ac-Gly	Ac-Gly	
	Ac-Gly-OH ($\blacktriangle$)	5	-	-	
6.5	Ac-Gly- ϵ -Lys-OH ($\blacklozenge$)	34	H	Ac-Gly	1.2 : 1.0
	Ac-Gly- α -Lys-OH ($\bullet$)	40	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-OH (∇)	8	Ac-Gly	Ac-Gly	
	Ac-Gly-OH ($\blacktriangle$)	10	-	-	

Supplementary Table 11. Yields for the products of prebiotic oxidative coupling of 2-acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM) with L-lysine **Lys** (150 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (150 mM) at specified pH value.

N-Acetylglycyllysineamide **Ac-Gly-Lys-NH₂**

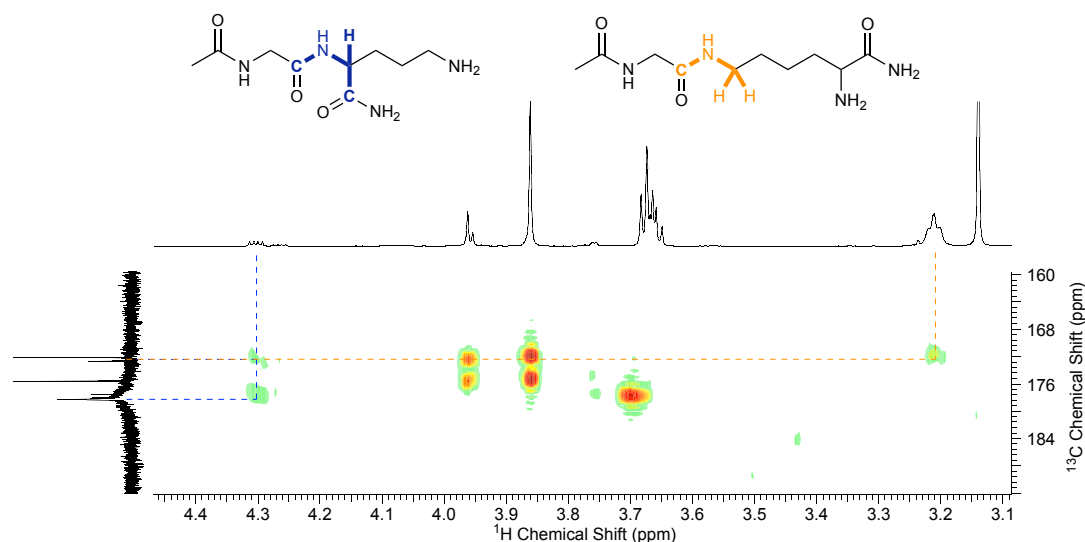


L-Lysinamide hydrochloride **Lys-NH₂·HCl** (27.2 mg, 0.15 mmol) was dissolved in H₂O/D₂O (98:2; 0.5 mL) and the solution was adjusted to pH 9.5, 8.5, 7.5 or 6.5 as specified with 4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and total volume was adjusted to 1 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added and the solution was stirred at room temperature for 20 min, whilst maintaining the solution at the specified initial pH with 4 M HCl/NaOH. The solution was then adjusted to pH 9.5, and the reaction mixture was centrifuged. The supernatant was analysed by 1D and 2D NMR spectroscopy. The reaction composition was quantified using methylsulfonylmethane (MSM) as an internal standard, and the formation of ligation products was confirmed by ¹H-¹³C HMBC NMR spectral analysis and high resolution mass spectrometry. See Supplementary Table 8 on Supplementary Page S53 for HRMS-ESI data, and Supplementary Table 12 (Supplementary Page S120) for reaction yields and product distribution at specified pH values.



Supplementary Figure 147. ¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.23–4.42 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-lysine (**Lys-NH₂**; 150 mM) and K₃[Fe(CN)₆] (150 mM) at pH 9.5 and room temperature.

¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*₂-(Acetylglycyl)lysineamide **Ac-Gly-α-Lys-NH₂** (●) (partial assignment): δ 4.31 (dd, *J* = 4.8, 9.5 Hz, 1H, Lys-(C2)-H), 3.96 (s, 2H, Gly-(C2)-H₂), 3.02 - 2.99 (m, 2H, Lys-(C6)-H₂), 2.08 (s, 3H, (COCH₃)), 1.90 - 1.84 (m, 2H, Lys-(C3)-H₂). *N*₆-(Acetylglycyl)lysineamide **Ac-Gly-ε-Lys-NH₂** (◆): δ 3.86 (s, 2H, Gly-(C2)-H₂), 3.66 (t, *J* = 6.7 Hz, 1H, Lys-(C2)-H), 3.21 (t, *J* = 6.3 Hz, 2H, Lys-(C6)-H₂), 2.07 (s, 3H), 1.75 - 1.69 (m, 2H, (C3)-H), 1.55 - 1.49 (m, 3H, (C3')-H, (C5)-H₂), 1.39 - 1.32 (m, 2H, (C4)-H₂). *N*₂,*N*₆-Di(acetylglycyl)lysineamide **Ac-Gly-α,ε-Lys-NH₂** (◻) (partial assignment): δ 4.27 (dd, *J* = 4.9, 9.4 Hz, 1H, Lys-(C2)-H), 3.95 (s, 2H, Gly-(C2)-H₂). *L*-Lysineamide **Lys-NH₂** (■): δ 3.68 (t, *J* = 6.6 Hz, 1H, (C2)-H), 3.04 (t, *J* = 7.6 Hz, 2H, (C6)-H₂), 1.82 - 1.72 (m, 2H, (C3)-H₂), 1.74 - 1.67 (m, 2H, (C5)-H₂), 1.50 - 1.40 (m, 2H, (C4)-H₂). *N*-Acetylglycine **Ac-Gly-OH** (▲) δ 3.76 (app. d, *J* = 4.9 Hz, 2H), 2.05 (s, 3H, (COCH₃)).

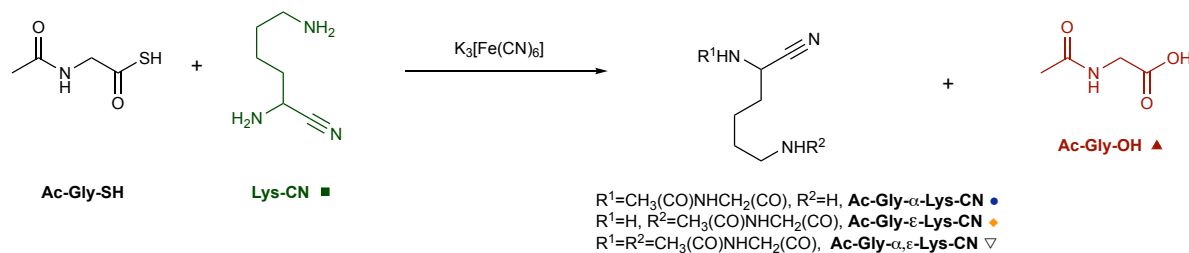


Supplementary Figure 148. ^1H – ^{13}C HMBC (^1H –700 MHz [3.10–4.45 ppm], ^{13}C –176 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Lys-(C2)–H of N_2 -(acetylglycyl)lysineamide **Ac-Gly- α -Lys-NH $_2$** (●) at 4.31 ppm with two C=O resonances at 172.2 and 177.6 ppm (shown in blue), which is characteristic of peptide bond formation of **Lys-NH $_2$** by reaction of the α -NH $_2$. The side chain acylation–by reaction of the **Lys-NH $_2$** ϵ -NH $_2$ to yield N_6 -(acetylglycyl)lysineamide **Ac-Gly- ϵ -Lys-NH $_2$** (◆)–was confirmed by the $^3J_{\text{CH}}$ coupling of Lys-(C6)–H at 3.21 ppm with a C=O resonance at 172.2 ppm (shown in orange). See Supplementary Figure 147 for expanded and labelled ^1H NMR spectrum.

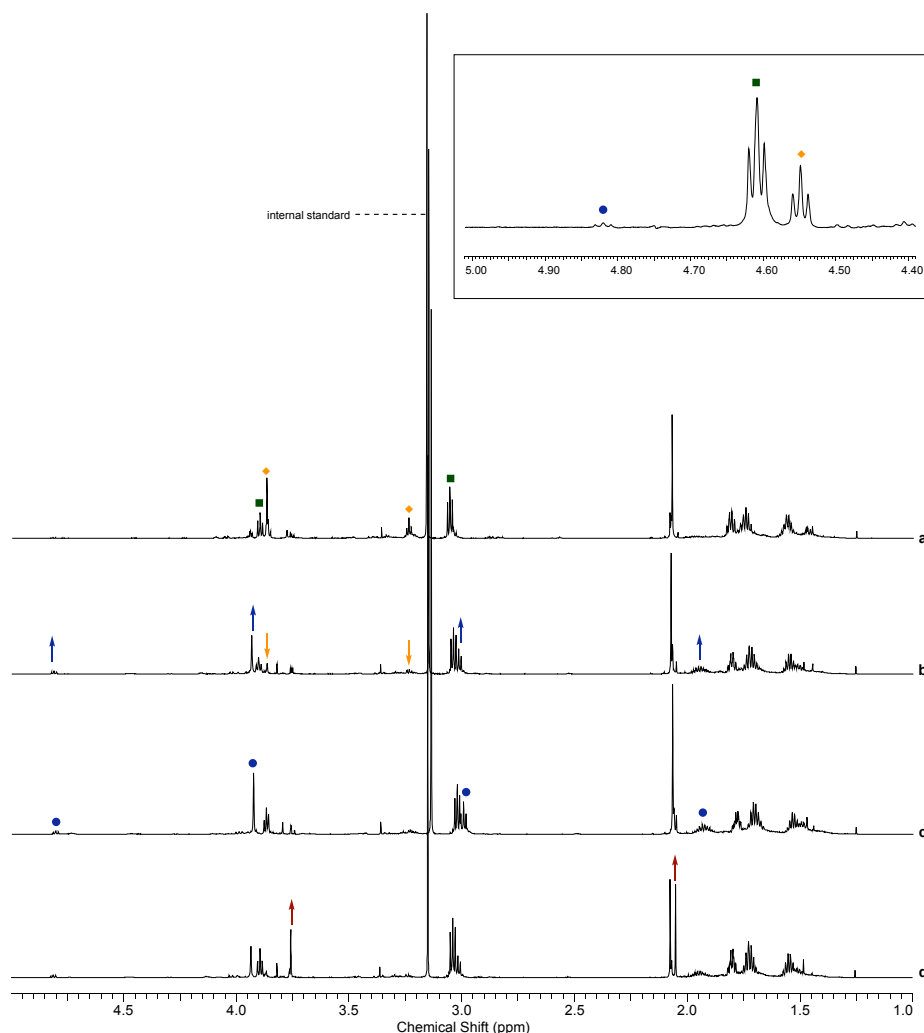
pH	Product	Yield (%)	R ¹ (α)	R ² (ϵ)	α : ϵ acylation
9.5	Ac-Gly- ϵ -Lys-NH $_2$ (◆)	73	H	Ac-Gly	1.0 : 4.9
	Ac-Gly- α -Lys-NH $_2$ (●)	15	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-NH $_2$ (∇)	6	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	5	-	-	
8.5	Ac-Gly- ϵ -Lys-NH $_2$ (◆)	42	H	Ac-Gly	1.0 : 1.2
	Ac-Gly- α -Lys-NH $_2$ (●)	36	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-NH $_2$ (∇)	8	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	8	-	-	
7.5	Ac-Gly- ϵ -Lys-NH $_2$ (◆)	23	H	Ac-Gly	2.7 : 1.0
	Ac-Gly- α -Lys-NH $_2$ (●)	61	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-NH $_2$ (∇)	7	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	6	-	-	
6.5	Ac-Gly- ϵ -Lys-NH $_2$ (◆)	21	H	Ac-Gly	2.8 : 1.0
	Ac-Gly- α -Lys-NH $_2$ (●)	58	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-NH $_2$ (∇)	5	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	n.d	-	-	

Supplementary Table 12. Yields for the products of prebiotic oxidative coupling of 2-acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM) with L-lysineamide **Lys-NH $_2$** (150 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (150 mM) at specified pH values. n.d = Not determined due to signal overlap in ^1H NMR spectra.

N-Acetylglycyllysine nitrile **Ac-Gly-Lys-CN**



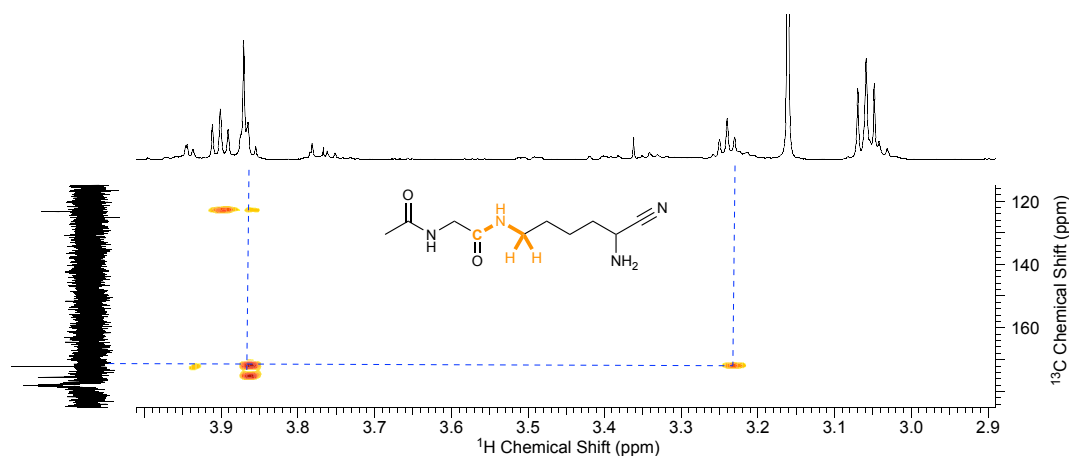
Lysine nitrile **Lys-CN**⁴⁹ (12.7 mg, 0.10 mmol) in D₂O (0.9 mL) was adjusted to pD 11, 9.5, 9.0, 8.5, 7.5, 6.5 or 5.5 with 4 M DCl/NaOD. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and the total volume was adjusted to 1 mL with D₂O. Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added and the solution was stirred at room temperature for 20 min. The solution was adjusted to pD 9.0, centrifuged, and the supernatant was analysed by 1D and 2D NMR spectroscopy at pD 9.0. The solution was then readjusted to pD 5.5 and further 1D and 2D NMR spectra were acquired. The reaction composition was quantified using methylsulfonylmethane (MSM) as an internal standard, and the formation of ligation products was confirmed by ¹H–¹³C HMBC NMR cross-correlation and high-resolution mass spectrometry. See Supplementary Table 7 on page S52 for HRMS-ESI data, and Supplementary Table 13 (page S123) for reaction yields and product distribution at specified pH values.



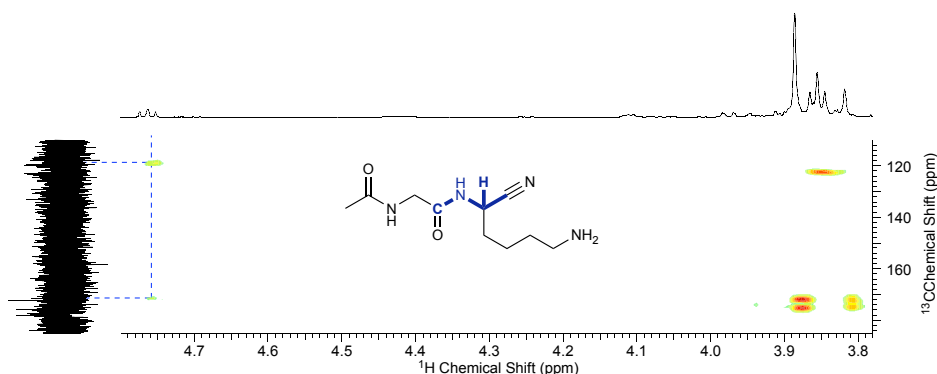
Supplementary Figure 149. ¹H NMR (700 MHz, D₂O, noesygppr1d, 1.0–5.0 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with lysine nitrile (**Lys-CN**, 100 mM) and K₃[Fe(CN)₆] (150 mM) at room temperature at specified pD (a. pD 11.0; b. pD 9.5; c. pD 7.5; d. pD 5.5). *N*₆-(Acetylglycyl)lysine nitrile **Ac-Gly-ε-Lys-CN** (♦) is the predominant species at pD 11.0 (spectrum a.) but depletes on lowering to pD 9.5 (spectrum b.; depleting of **Ac-Gly-ε-Lys-CN** shown with ↓). This coincides with an increase in *N*₂-(acetylglycyl)lysine nitrile **Ac-Gly-α-Lys-CN** (•) (spectrum b.; increase of **Ac-Gly-α-Lys-CN** shown with ↑), which is now the major ligation product throughout all pD from 5.5 to 9.5. Despite a gradual increase in *N*-acetylglycine **Ac-Gly-OH** (▲) below pD 7.5 (spectrum d., increase in **Ac-Gly-OH** shown with ↑), *N*₂-(acetylglycyl)lysine nitrile **Ac-Gly-α-Lys-CN** (•) persists as the major species. Inset: ¹H NMR (700 MHz, 4.40–5.00 ppm, pD 5.5) spectrum of the reaction of 2-

acetamidoethanethioic S-acid **Ac-Gly-SH** and lysine nitrile **Lys-CN** at pD 11, followed by analysis at pD 5.5, which shows a discernable downfield shift for the (C2)–H resonance of *N*₆-(acetylglycyl)lysine nitrile **Ac-Gly- ϵ -Lys-CN** (◆) and lysine nitrile **Lys-CN** (■) as a consequence of α -NH₂ protonation. The ¹H NMR chemical shift of the (C2)–H resonance of *N*₂-(acetylglycyl)lysine nitrile **Ac-Gly- ϵ -Lys-CN** (●) remains relatively unchanged, which is further evidence of α -acylation. It should be noted that the intensity of the (C2)–H resonance of *N*₂-(acetylglycyl)lysine nitrile **Ac-Gly- ϵ -Lys-CN** (●) is low due to it coinciding with the HOD signal and the resultant solvent suppression by the noesygppr1d NMR pulse sequence.

¹H NMR (700 MHz, D₂O, noesygppr1d) *N*₂-(Acetylglycyl)lysine nitrile **Ac-Gly- ϵ -Lys-CN** (●, assigned from the reaction at pD 7.5 (Supplementary Figure 149c)): δ 4.80 (t, *J* = 7.5 Hz, 1H, Lys-(C2)–H), 3.93 (s, 2H, Gly-(C2)–H₂), 3.00 (t, *J* = 7.6 Hz, 2H, Lys-(C6)–H₂), 2.07 (s, 3H, (COCH₃), 1.97 - 1.88 (m, 2H, Lys-(C3)–H₂), 1.82 - 1.75 (m, 2H, Lys-(C5)–H₂), 1.54 - 1.45 (m, 2H, Lys-(C4)–H₂). *N*₆-(Acetylglycyl)lysine nitrile **Ac-Gly- ϵ -Lys-CN** (◆, assigned from the reaction at pD 11 (Supplementary Figure 149a)): δ 3.87 (s, 2H, Gly-(C2)–H₂), 3.86 (t, *J* = 7.4 Hz, 1H, Lys-(C2)–H), 3.24 (t, *J* = 6.8 Hz, 2H, Lys-(C6)–H₂), 2.07 (s, 3H, (COCH₃), 1.78 - 1.70 (m, 2H, Lys-(C3)–H₂), 1.59 - 1.51 (m, 2H, Lys-(C5)–H₂), 1.50 - 1.42 (m, 2H, Lys-(C4)–H₂). *N*₂,*N*₆-Di(acetylglycyl)lysine nitrile **Ac-Gly- α,ϵ -Lys-CN** (∇, assigned from the reaction at pD 11 (Supplementary Figure 149a)): (partial assignment): δ 3.89 (s, 2H, Gly-(C2)–H₂), 3.18 - 3.14 (m, 2H, Lys-(C6)–H₂), 1.97 - 1.85 (m, 2H, Lys-(C3)–H₂). Lysine nitrile **Lys-CN** (■, assigned from the reaction at pD 11.0 (Supplementary Figure 149a)): δ 3.90 (t, *J* = 7.1 Hz, 1H, (C2)–H), 3.06 (t, *J* = 7.6 Hz, 2H, (C6)–H₂), 1.84 - 1.79 (m, 2H, (C3)–H₂), 1.78 - 1.71 (m, 2H, (C5)–H₂), 1.60 - 1.51 (m, 2H, (C4)–H₂). *N*-Acetylglycine **Ac-Gly-OH** (▲, assigned from the reaction at pD 5.5 (Supplementary Figure 149d)) δ 3.76 (s, 2H, (C2)–H₂), 2.05 (s, 3H, (COCH₃)).



Supplementary Figure 150. ¹H–¹³C HMBC (¹H–700 MHz [2.90–4.00 ppm], ¹³C–176 MHz [110–185 ppm], D₂O) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with lysine nitrile (**Lys-CN**; 100 mM) and K₃[Fe(CN)₆] (150 mM) at pD 11.0 and room temperature. A diagnostic ³*J*_{CH} coupling of Lys-(C6)–H₂ of *N*₆-(acetylglycyl)lysine nitrile **Ac-Gly- ϵ -Lys-CN** (◆) at 3.24 ppm with a C=O resonance at 172.0 ppm is highlighted. This is characteristic of peptide bond formation of lysine nitrile **Lys-CN** by reaction of the ϵ -NH₂ (side chain acylation), the predominant ligation species at pD 11.0. See Supplementary Figure 149a for an expansion of the corresponding ¹H NMR spectrum.



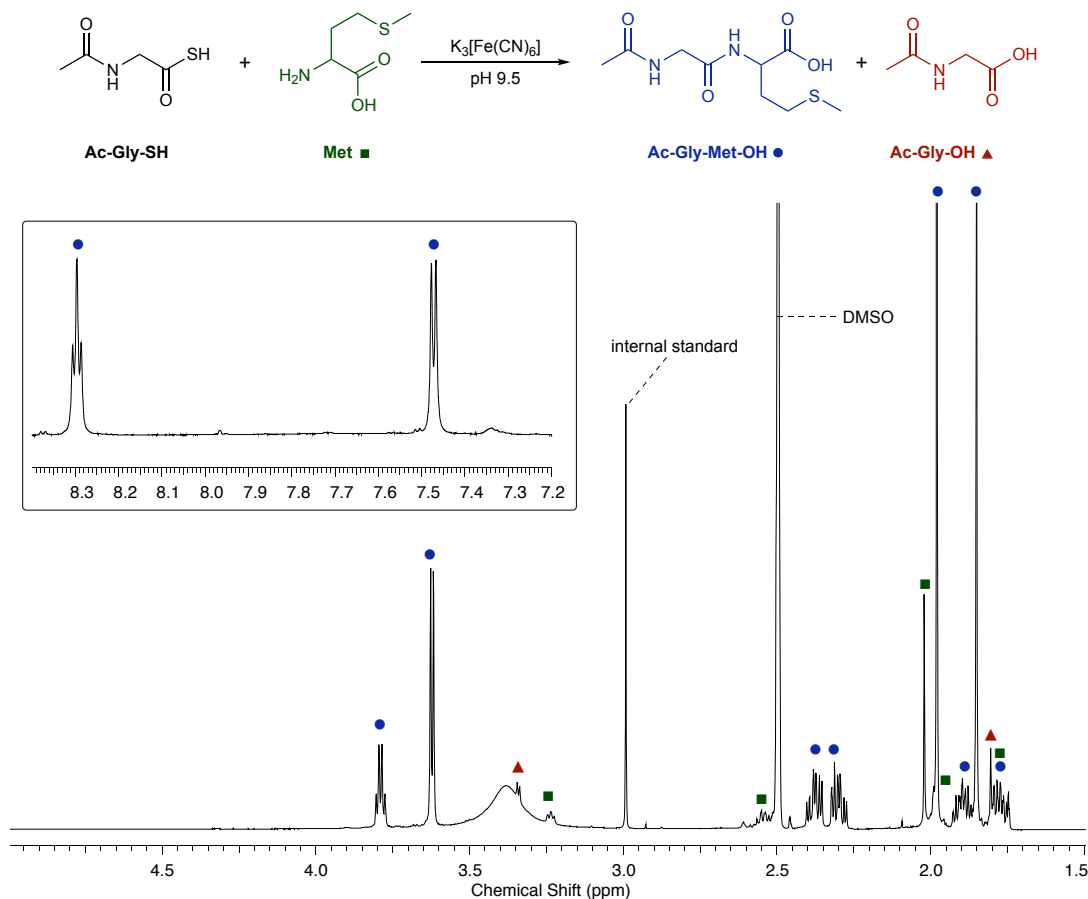
Supplementary Figure 151. ¹H–¹³C HMBC (¹H–700 MHz [3.78–4.80 ppm], ¹³C–176 MHz [110–185 ppm], D₂O) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with lysine nitrile (**Lys-CN**; 100 mM) and K₃[Fe(CN)₆] (150 mM) at pD 9.5 and room temperature. A diagnostic ²*J*_{CH} and ³*J*_{CH} coupling of Lys-(C2)–H of *N*₂-(acetylglycyl)lysine nitrile **Ac-Gly- ϵ -Lys-CN** (●) at 4.80 ppm with a CN resonance at 119.1 ppm and a C=O resonance at 171.8 ppm is highlighted. This is characteristic of peptide

bond formation of lysine nitrile **Lys-CN** by reaction of the α -NH₂, the predominant ligation species observed in the pD 5.5-9.5 range. See Supplementary Figure 149b for the expanded corresponding ¹H NMR spectrum.

pH	Product	Yield (%)	R ¹ (α)	R ² (ϵ)	$\alpha : \epsilon$ acylation
11.0	Ac-Gly- ϵ -Lys-CN (◆)	45	H	Ac-Gly	1.0 : 2.3
	Ac-Gly- α -Lys-CN (●)	20	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-CN (▽)	15	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	<5	-	-	
9.5	Ac-Gly- ϵ -Lys-CN (◆)	22	H	Ac-Gly	2.8 : 1.0
	Ac-Gly- α -Lys-CN (●)	62	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-CN (▽)	7	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	<5	-	-	
9.0	Ac-Gly- ϵ -Lys-CN (◆)	25	H	Ac-Gly	2.5 : 1.0
	Ac-Gly- α -Lys-CN (●)	62	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-CN (▽)	6	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	<5	-	-	
8.5	Ac-Gly- ϵ -Lys-CN (◆)	10	H	Ac-Gly	7.5 : 1.0
	Ac-Gly- α -Lys-CN (●)	75	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-CN (▽)	ND	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	<5	-	-	
7.5	Ac-Gly- ϵ -Lys-CN (◆)	ND	H	Ac-Gly	>80 : 1.0
	Ac-Gly- α -Lys-CN (●)	78	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-CN (▽)	ND	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	10	-	-	
6.5	Ac-Gly- ϵ -Lys-CN (◆)	0	H	Ac-Gly	>80 : 1.0
	Ac-Gly- α -Lys-CN (●)	78	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-CN (▽)	0	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	20	-	-	
5.5	Ac-Gly- ϵ -Lys-CN (◆)	0	H	Ac-Gly	>53 : 1.0
	Ac-Gly- α -Lys-CN (●)	53	Ac-Gly	H	
	Ac-Gly- α,ϵ -Lys-CN (▽)	0	Ac-Gly	Ac-Gly	
	Ac-Gly-OH (▲)	42	-	-	

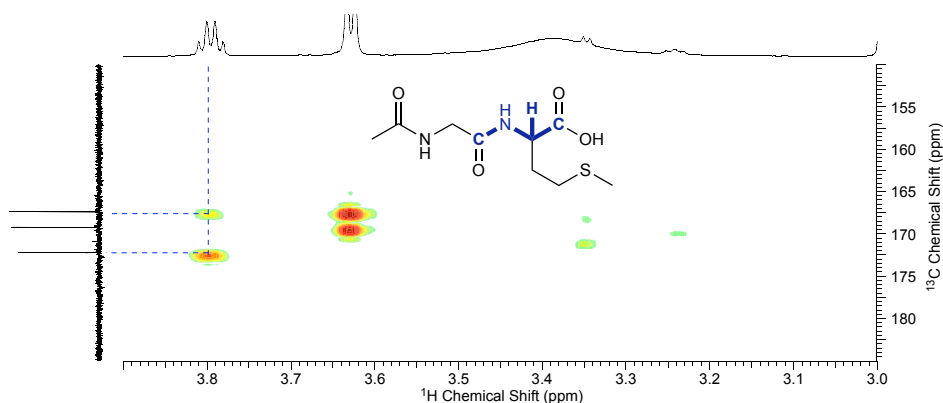
Supplementary Table 13. Yields for the products of prebiotic oxidative coupling of 2-acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM) with lysine nitrile **Lys-CN** (100 mM) and K₃[Fe(CN)₆] (150 mM) at specified pD values. ND = Not detected by ¹H NMR.

N-Acetylglycylmethionine **Ac-Gly-Met-OH**



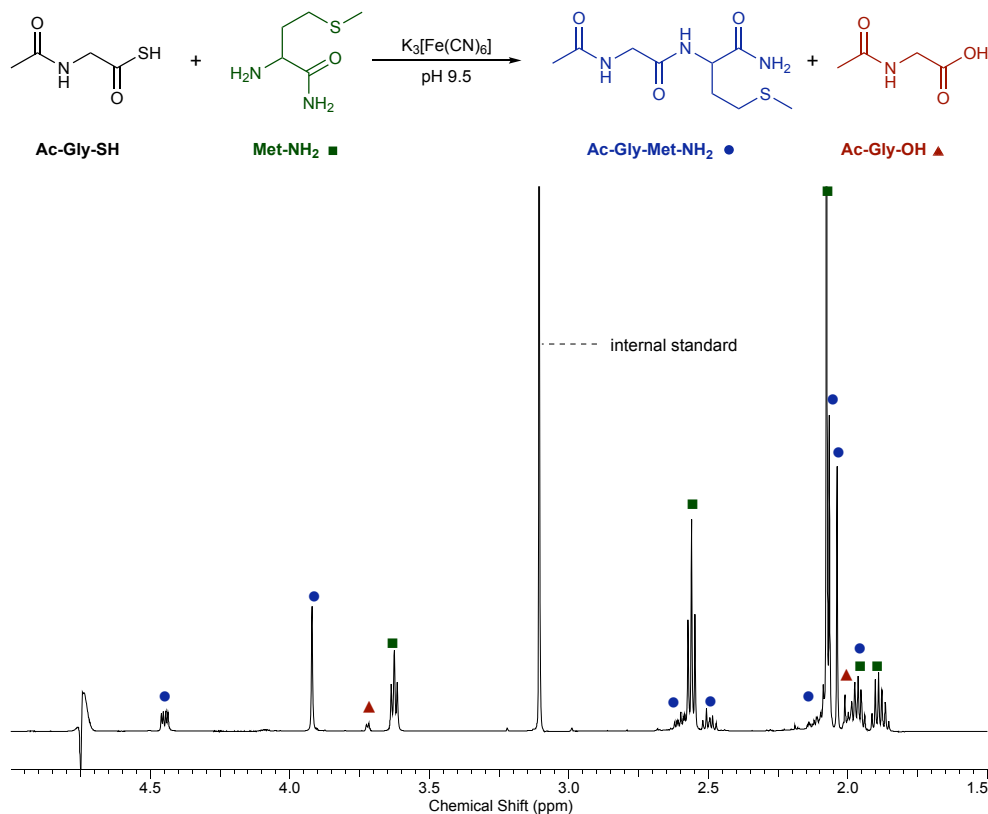
Supplementary Figure 152. 1H NMR (600 MHz, DMSO- d_6 , 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with DL-methionine (**Met**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature, after the solution was lyophilised and the resultant residue was dissolved in DMSO- d_6 . Yields were obtained prior to lyophilisation. DL-Methionine **Met** observed as a minor species due to its low solubility in DMSO- d_6 , but was fully accounted for in NMR spectra in H_2O/D_2O (not shown). Inset: 1H NMR (600 MHz, DMSO- d_6 , 7.2–8.4 ppm) showing the NH resonances for *N*-acetylglycylmethionine **Ac-Gly-Met-OH**.

1H NMR (600 MHz, DMSO- d_6) *N*-Acetylglycylmethionine **Ac-Gly-Met-OH** (●): δ 8.30 (t, J = 5.9 Hz, 1H, Gly-(C2)-NH), 7.47 (d, J = 6.5 Hz, 1H, Met-(C2)-NH), 3.79 (app. q, J = 5.8 Hz, 1H, Met-(C2)-H), 3.62 (d, J = 5.9 Hz, 2H, Gly-(C2)-H₂), 2.38 (ddd, J = 12.6, 11.6, 5.2, 1H, Met-(C4)-H), 2.30 (ddd, J = 12.6, 11.6, 5.2, 1H, Met-(C4')-H), 1.98 (s, 3H, SCH₃), 1.93 - 1.86 (m, 1H, Met-(C3)-H), 1.85 (s, 3H, (COCH₃), 1.80 - 1.73 (m, 1H, Met-(C3')-H). DL-Methionine **Met** (■): δ 3.24 (dd, J = 5.5, 6.7 Hz, 1H, (C2)-H), 2.58 - 2.51 (m, 2H, (C4)-H₂), 2.02 (s, 3H, SCH₃), 2.00 - 1.96 (obs., 1H, (C3)-H), 1.82 - 1.77 (obs., 1H, (C3')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.34 (app. d, J = 4.7 Hz, 2H, (C2)-H₂), 1.80 (s, 3H, (COCH₃)).



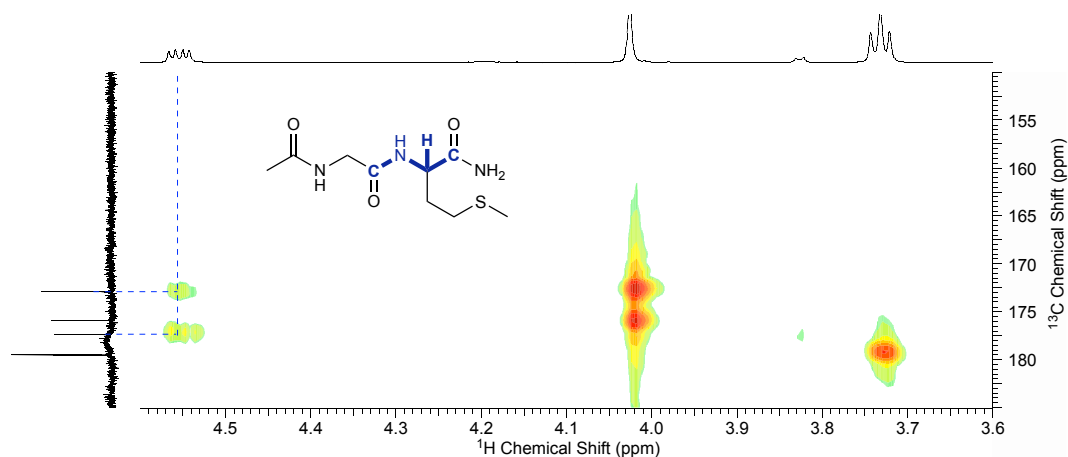
Supplementary Figure 153. 1H - ^{13}C HMBC (1H -600 MHz [3.0–3.9 ppm], ^{13}C -151 MHz [150–185 ppm], DMSO- d_6) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Met-(C2)-H of **Ac-Gly-Met-OH** at 3.79 ppm with two C=O resonances at 167.4 and 172.2 ppm, which is characteristic of peptide bond formation of **Met**. See Supplementary Figure 152 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylmethioninamide **Ac-Gly-Met-NH₂**



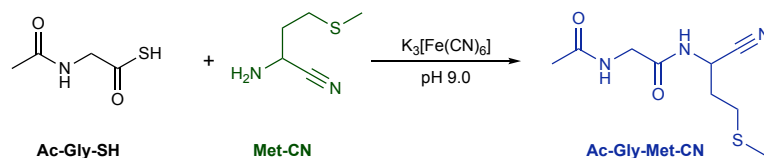
Supplementary Figure 154. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygprr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-methioninamide (**Met-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygprr1d) *N*-Acetylglycylmethioninamide **Ac-Gly-Met-NH₂** (●): δ 4.45 (dd, J = 4.4, 9.8 Hz, 1H, Met-(C2)–H), 3.92 (s, 2H, Gly-(C2)–H₂), 2.62 - 2.58 (m, 1H, Met-(C4)–H), 2.53 - 2.45 (m, 1H, Met-(C4')–H), 2.15 - 2.09 (m, 1H, Met-(C3)–H), 2.07 (s, 3H, SCH₃), 2.04 (s, 3H, (COCH₃)), 2.02 - 1.96 (m, 1H, Met-(C3')–H). *L*-Methioninamide **Met-NH₂** (■): δ 3.63 (t, J = 6.7 Hz, 1H, (C2)–H), 2.56 (t, J = 7.6 Hz, 2H, (C4)–H₂), 2.08 (s, 3H, (COCH₃)), 2.01 - 1.93 (m, 1H, (C3)–H), 1.92 - 1.84 (m, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.75 (AB, J = 16.1 Hz, 1H, (C2)–H), 3.70 (AB, J = 16.1 Hz, 1H, (C2')–H), 2.01 (s, 3H, (COCH₃)).

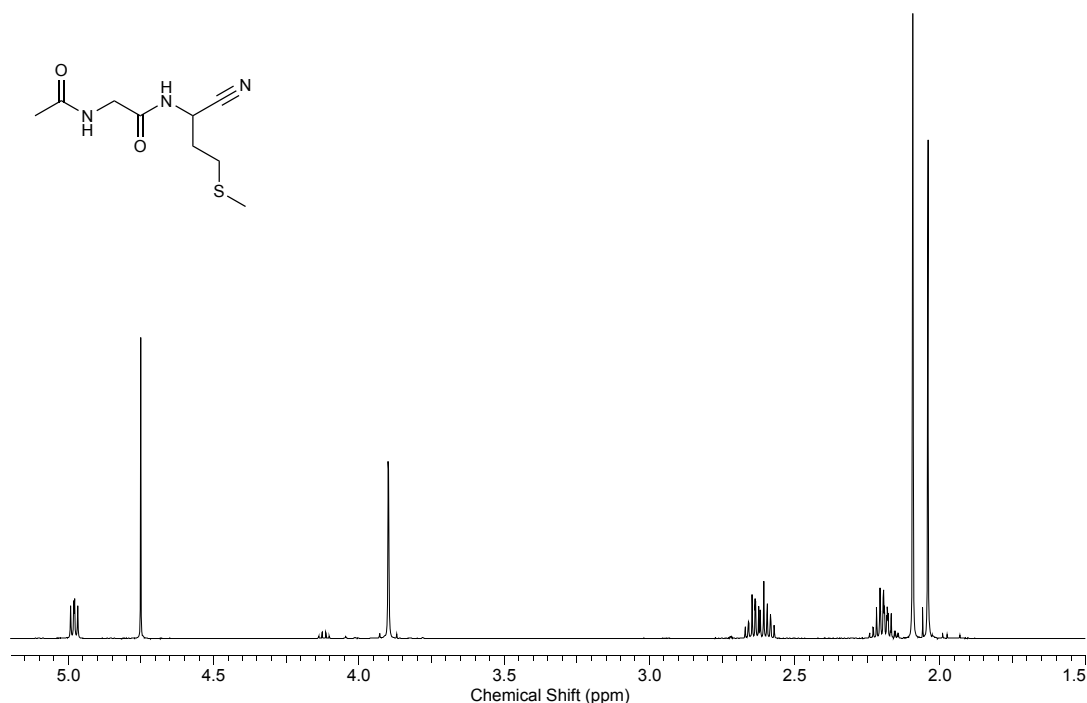


Supplementary Figure 155. ¹H-¹³C HMBC (¹H-600 MHz [3.6–4.6 ppm], ¹³C-151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Met-(C2)–H of **Ac-Gly-Met-NH₂** at 4.45 ppm with two C=O resonances at 172.9 and 177.4 ppm, which is characteristic of peptide bond formation of **Met-NH₂**. See Supplementary Figure 154 for expanded and labelled ¹H NMR spectrum.

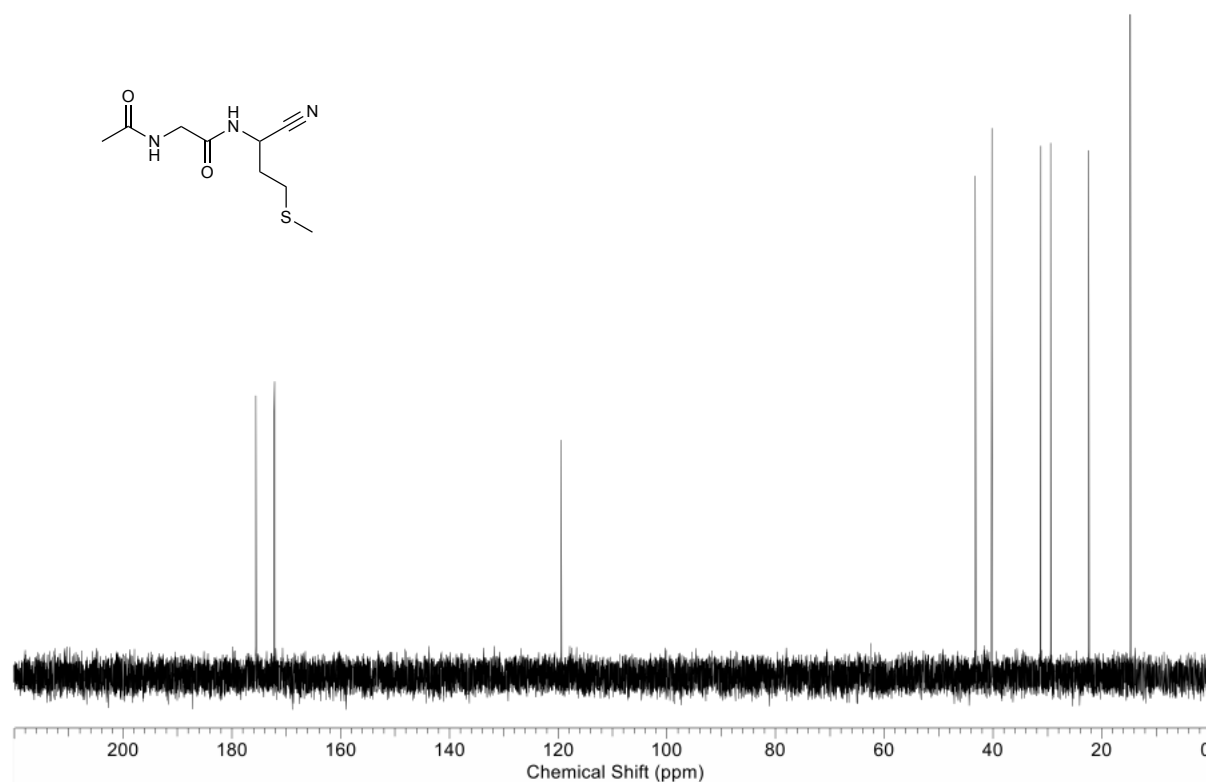
N-Acetylglycylmethionine nitrile **Ac-Gly-Met-CN**



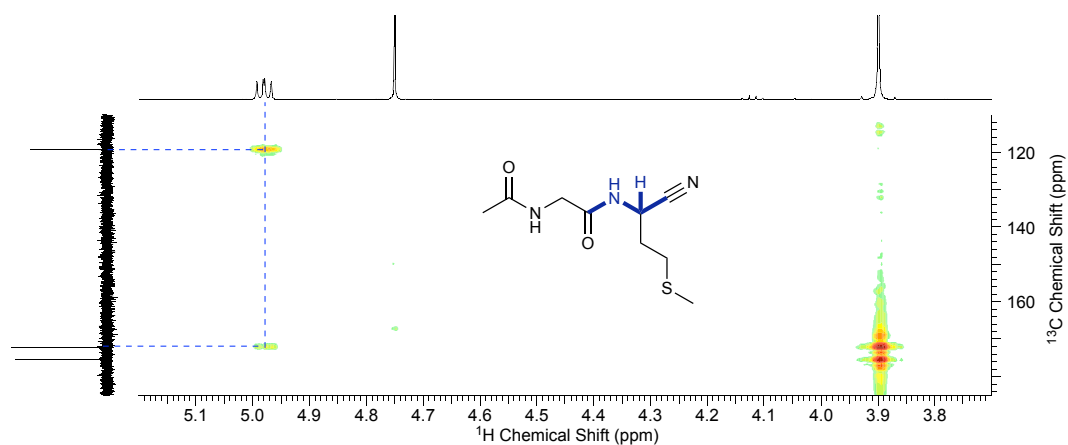
Prepared following *general procedure G* on a 0.1 mmol scale, from sodium-2-acetamido-ethanethioate **Ac-Gly-S⁻Na⁺** (15.5 mg, 0.1 mmol) and *rac*-methionine nitrile hydrochloride **Met-CN·HCl**⁴² (33.3 mg, 0.2 mmol). **Ac-Gly-Met-CN** was isolated as a pale-yellow oil, which solidified on standing (19.9 mg, 87%) after flash column chromatography (SiO₂, EtOAc (100%)). *R_f* 0.40 (100% EtOAc). ¹H NMR (600 MHz, D₂O) δ 4.98 (dd, *J* = 6.6, 8.4 Hz, 1H, Met-(C2)-H), 3.91 (AB, *J* = 17.2 Hz, 1H, Gly-(C2)-H), 3.88 (d, *J* = 17.2 Hz, 1H, Gly-(C2)-H'), 2.68 - 2.56 (m, 2H, Met-(C4)-H₂), 2.25 - 2.13 (m, 2H, Met-(C3)-H₂), 2.09 (s, 3H, Met-SCH₃), 2.04 (s, 3H, COCH₃). ¹³C NMR (151 MHz, D₂O) δ 175.6 (COCH₃), 172.2 (Gly-(C1)), 119.4 (Met-(C1)), 43.2 (Gly-(C2)), 40.2 (Met-(C2)), 31.3 (Met-(C3)), 29.4 (Met-(C4)), 22.4 (COCH₃), 14.8 (Met-SCH₃). HRMS-ESI-MS [*M*+H]⁺ calcd. for C₉H₁₆N₃O₂S 230.0958; obs. 230.0957. IR (solid, cm⁻¹) 3255, 2045, 2463, 2360, 2331, 1674, 1631, 1556, 1535.



Supplementary Figure 156. ¹H NMR (600 MHz, D₂O, 1.5–5.2 ppm) spectrum of purified *N*-acetylglycylmethionine nitrile **Ac-Gly-Met-CN**

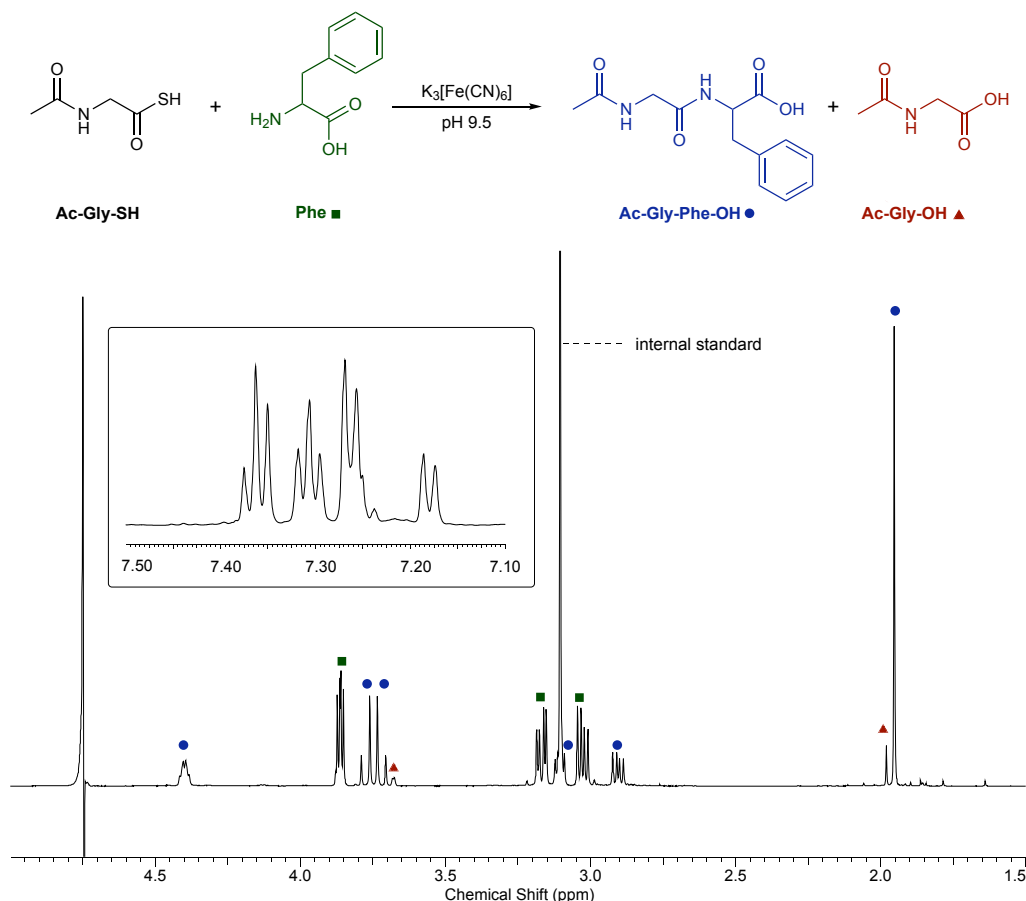


Supplementary Figure 157. ^{13}C NMR (151 MHz, D_2O , 0–220 ppm) spectrum of purified *N*-acetylglycylmethionine nitrile **Ac-Gly-Met-CN**



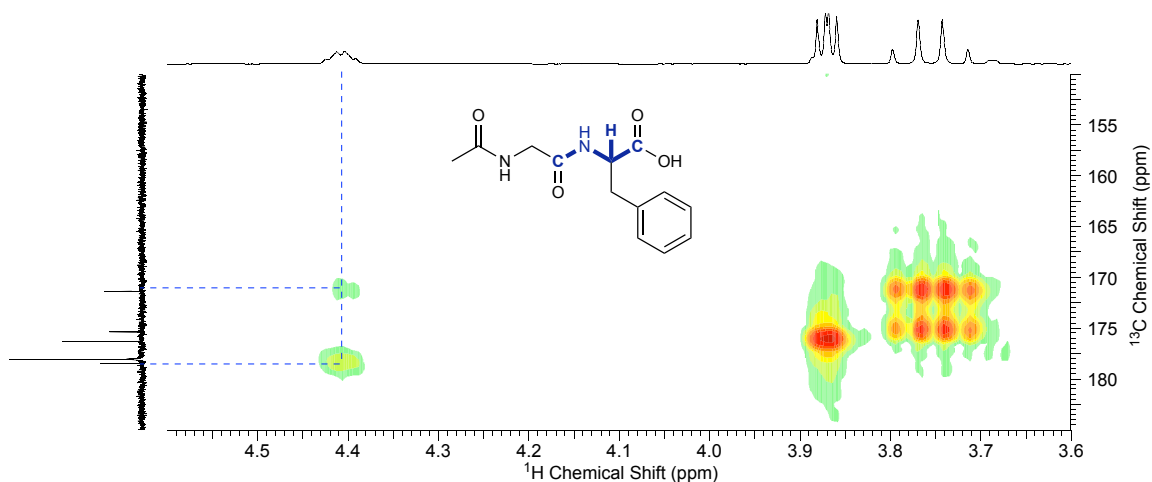
Supplementary Figure 158. ^1H – ^{13}C HMBC (^1H –600 MHz [3.7–5.2 ppm], ^{13}C –151 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Met-(C2)–H of **Ac-Gly-Met-CN** at 4.98 ppm with a CN resonances at 119.4 ppm and a C=O resonances at 172.2 ppm, which is characteristic of peptide bond formation of **Met-CN**. See Supplementary Figure 160 for expanded ^1H NMR spectrum.

N-Acetylglycylphenylalanine **Ac-Gly-Phe-OH**



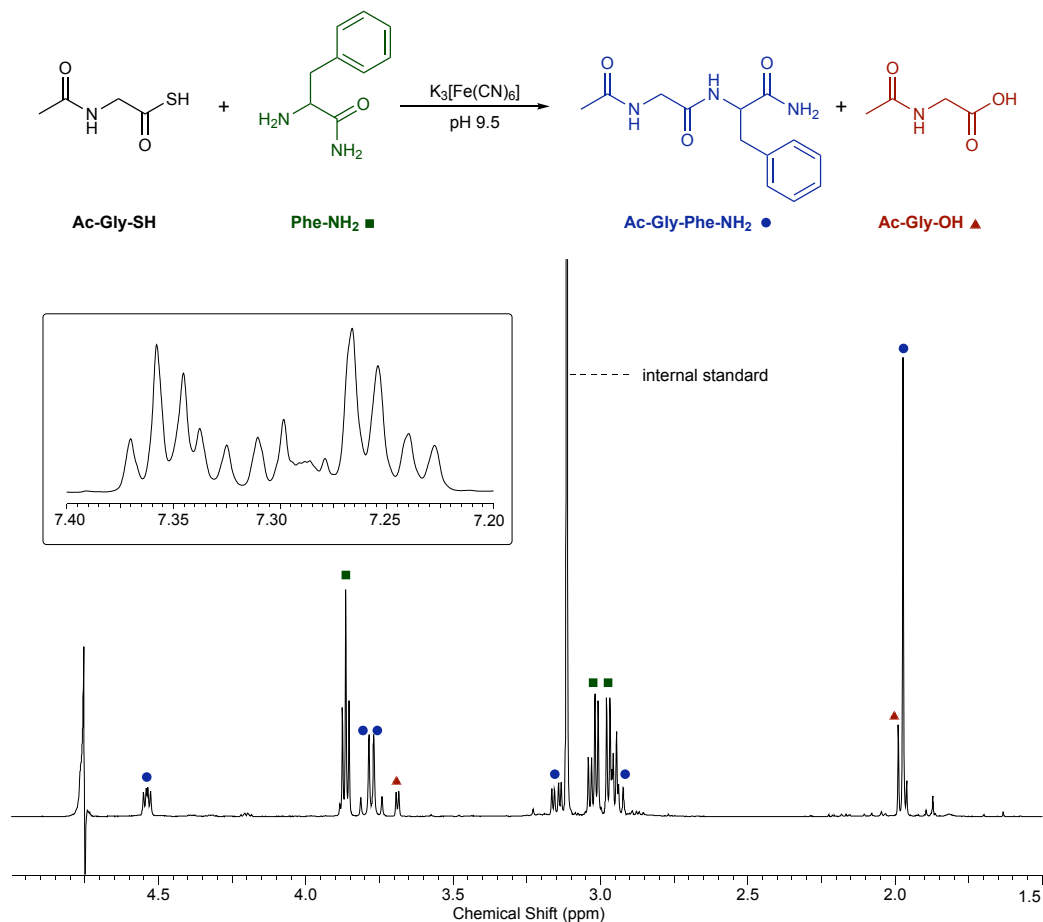
Supplementary Figure 159. ^1H NMR (600 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygprr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-phenylalanine (**Phe**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature. Inset: ^1H NMR (600 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, 7.10–7.50 ppm) showing the aromatic resonances for *N*-acetylglycylphenylalanine **Ac-Gly-Phe-OH** and residual L-phenylalanine **Phe**.

^1H NMR (600 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygprr1d) *N*-Acetylglycylphenylalanine **Ac-Gly-Phe-OH** (●): δ 7.40 - 7.13 (m, 5H, Ar), 4.42 - 4.37 (m, 1H, Phe-(C2)-H), 3.78 (AB, J = 17.0 Hz, 1H, Gly-(C2)-H), 3.72 (AB, J = 17.0 Hz, 1H, Gly-(C2')-H), 3.11 (ABX, J = 5.2, 13.9 Hz, 1H, Phe-(C3)-H), 2.91 (dd, J = 8.1, 13.9 Hz, 1H, Phe-(C3')-H), 1.95 (s, 3H, (COCH₃)). L-Phenylalanine **Phe** (■): δ 7.40 - 7.13 (m, 5H, Ar), 3.86 (dd, J = 5.4, 7.6 Hz, 1H, (C2)-H), 3.17 (ABX, J = 5.4, 14.4 Hz, 1H, (C3)-H), 3.03 (ABX, J = 7.6, 14.4 Hz, 1H, (C3')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.68 (app. d, J = 3.9 Hz, 2H, (C2)-H₂), 1.98 (s, 3H, (COCH₃)).



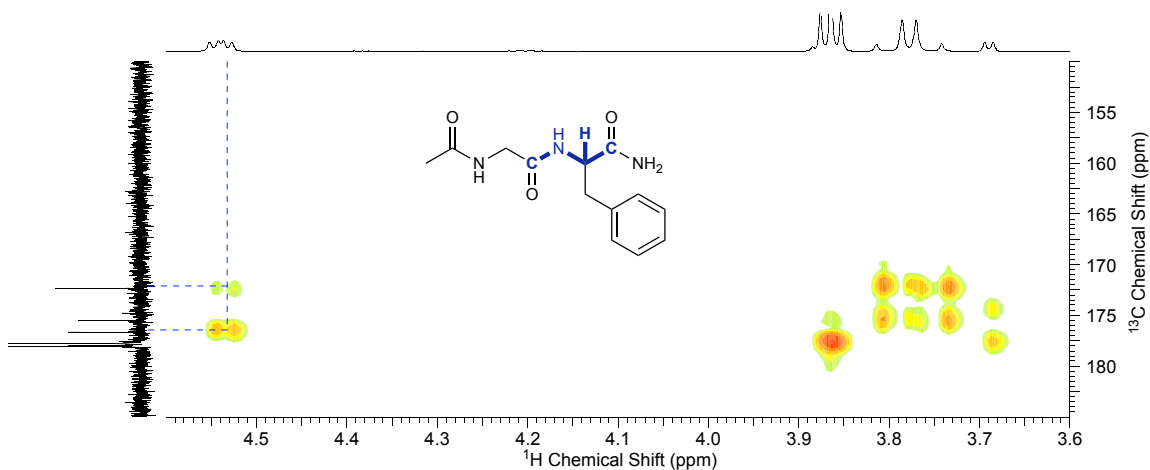
Supplementary Figure 160. ^1H - ^{13}C HMBC (^1H -600 MHz [3.6–4.6 ppm], ^{13}C -151 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Phe-(C2)-H of **Ac-Gly-Phe-OH** centred at 4.40 ppm with two C=O resonances at 171.4 and 178.5 ppm, which is characteristic of peptide bond formation of **Phe**. See Supplementary Figure 159 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycylphenylalaninamide **Ac-Gly-Phe-NH₂**



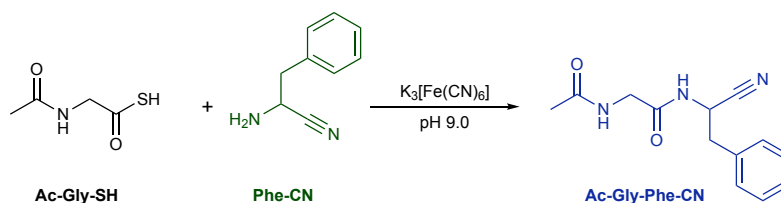
Supplementary Figure 161. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-phenylalaninamide (**Phe-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature. Inset: ¹H NMR (600 MHz, H₂O/D₂O 98:2, 7.2–7.4 ppm) showing the aromatic resonances for *N*-acetylglycylphenylalaninamide **Ac-Gly-Phe-NH₂** and residual **Phe-NH₂**.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylphenylalaninamide **Ac-Gly-Phe-NH₂** (●): δ 7.38 – 7.21 (m, 5H, Ar), 4.54 (dd, $J = 5.7, 9.2$ Hz, 1H, Phe-(C2)–H), 3.80 (AB, $J = 17.0$ Hz, 1H, Gly-(C2)–H), 3.76 (AB, $J = 17.0$ Hz, 1H, Gly-(C2')–H), 3.15 (ABX, $J = 5.7, 14.0$ Hz, 1H, Phe-(C3)–H), 2.94 (ABX, $J = 9.2, 14.0$ Hz, 1H, Phe-(C3')–H), 1.97 (s, 3H, (COCH₃)). *L*-Phenylalaninamide **Phe-NH₂** (■): δ 7.38 – 7.21 (m, 5H, Ar), 3.86 (t, $J = 6.9$ Hz, 1H, (C2)–H), 3.02 (ABX, $J = 6.9, 13.0$ Hz, 1H, (C3)–H), 2.96 (ABX, $J = 6.9, 13.0$ Hz, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.69 (app. d, $J = 5.8$ Hz, 2H, (C2)–H₂), 1.99 (s, 3H, (COCH₃)).

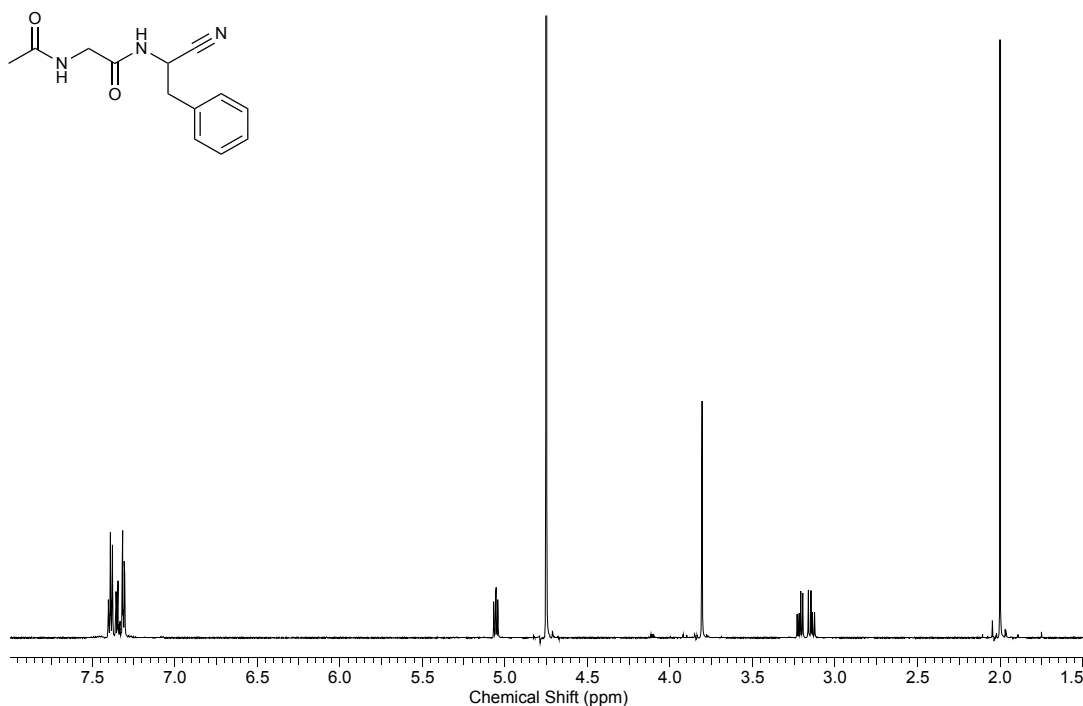


Supplementary Figure 162. ¹H–¹³C HMBC (¹H–600 MHz [3.6–4.6 ppm], ¹³C–151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Phe-(C2)–H of **Ac-Gly-Phe-NH₂** at 4.54 ppm with two C=O resonances at 172.4 and 176.7 ppm, which is characteristic of peptide bond formation of **Phe**. See Supplementary Figure 161 for expanded and labelled ¹H NMR spectrum.

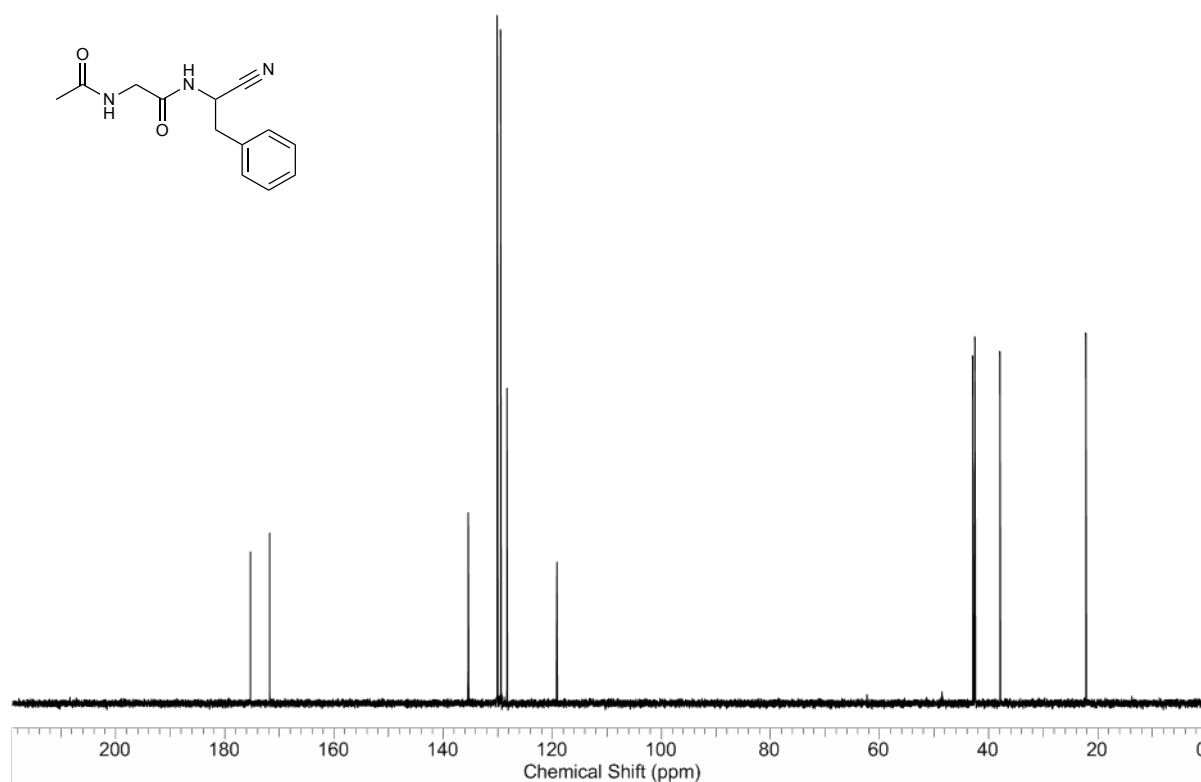
N-Acetylglycylphenylalanine nitrile **Ac-Gly-Phe-CN**



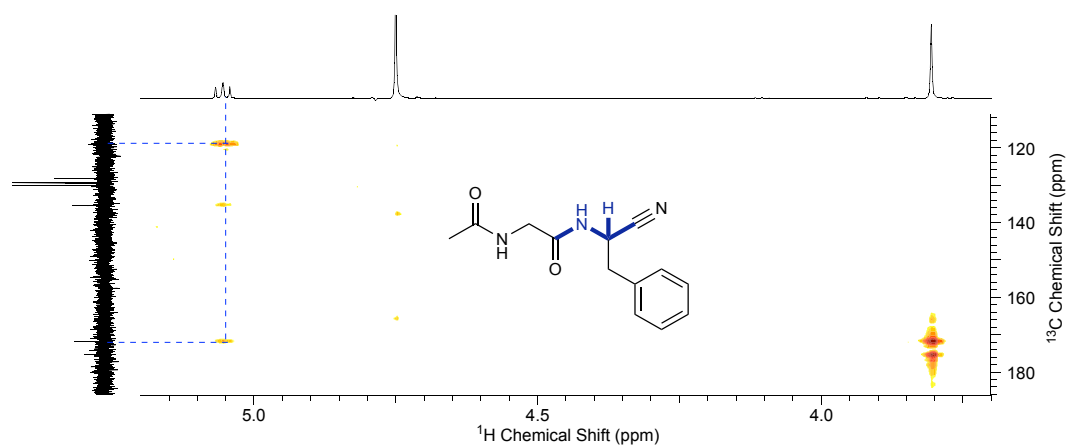
Prepared following *general procedure G* on a 0.1 mmol scale, from sodium-2-acetamido-ethanethioate **Ac-Gly-S⁻Na⁺** (15.5 mg, 0.1 mmol) and *rac*-phenylalanine nitrile hydrochloride **Phe-CN·HCl**⁴¹ (36.5 mg, 0.2 mmol). **Ac-Gly-Phe-CN** was isolated as a white solid (20.6 mg, 84%) after flash column chromatography (SiO₂, EtOAc (100%)). *R_f* 0.40 (100% EtOAc). ¹H NMR (600 MHz, D₂O) δ 7.41 - 7.37 (m, 2H, Phe-Ar), 7.36 - 7.33 (m, 1H, Phe-Ar), 7.32 - 7.29 (m, 2H, Phe-Ar), 5.06 (dd, *J* = 7.0, 8.0 Hz, 1H, Phe-(C1)-H), 3.81 (s, 2H, Gly-(C2)-H₂), 3.21 (ABX, *J* = 7.0, 13.8 Hz, 1H, Phe-(C3)-H), 3.14 (ABX, *J* = 8.0, 13.8 Hz, 1H, Phe-(C3)-H'), 2.00 (s, 3H, COCH₃). ¹³C NMR (176 MHz, D₂O) δ 175.3 (COCH₃), 171.7 (Gly-(C2)), 135.4 (Ar), 130.0 (Ar × 2), 129.4 (Ar × 2), 128.3 (Ar), 119.1 (Phe-(C1)), 42.9 (Gly-(C2)), 42.5 (Phe-(C2)), 37.9 (Phe-(C3)), 22.2 (COCH₃). HRMS-ESI-MS [*M*+H]⁺ calcd. for C₁₃H₁₆N₃O₂ 246.1237; obs. 246.1238. IR (solid, cm⁻¹) 3250, 3062, 2929, 2374, 2337, 1672, 1620, 1552.



Supplementary Figure 163. ¹H NMR (600 MHz, D₂O, 1.5–8.0 ppm) spectrum of purified *N*-acetylglycylphenylalanine nitrile **Ac-Gly-Phe-CN**

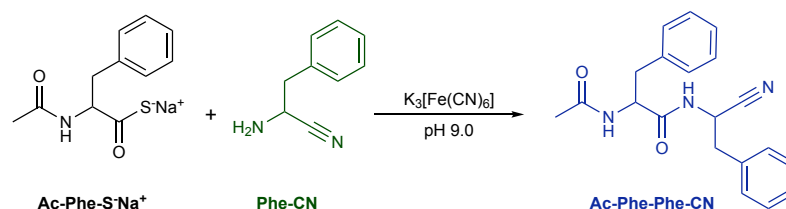


Supplementary Figure 164. ^{13}C NMR (700 MHz, D_2O , 0–220 ppm) spectrum of purified *N*-acetylglycylphenylalanine nitrile **Ac-Gly-Phe-CN**

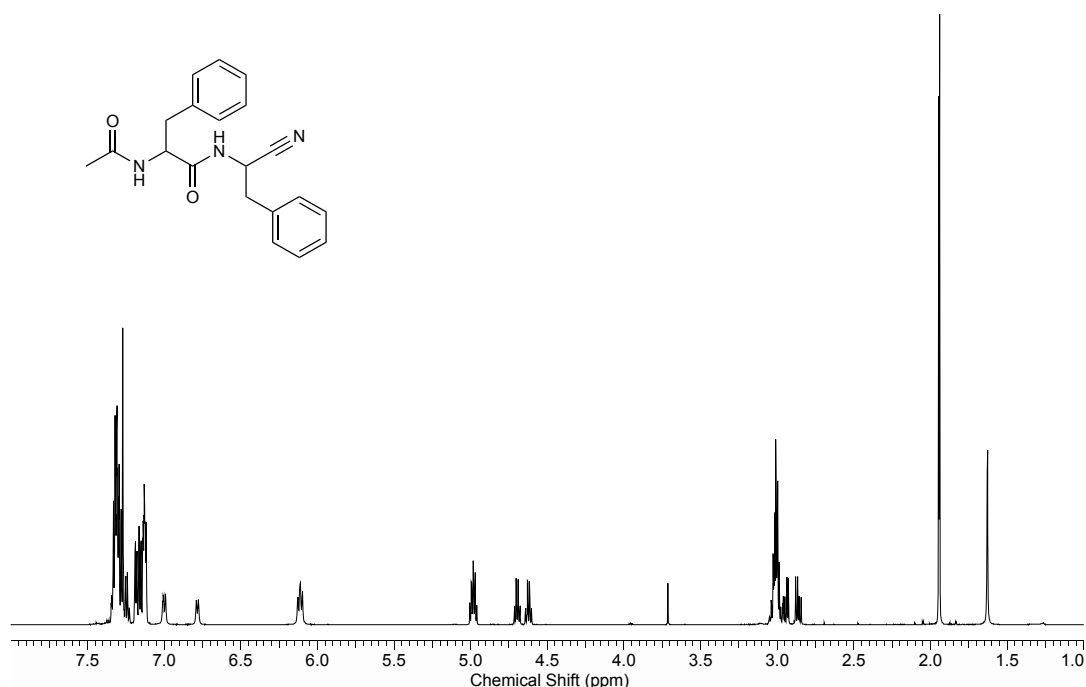


Supplementary Figure 165. ^1H – ^{13}C HMBC (^1H –600 MHz [3.7–5.2 ppm], ^{13}C –151 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Phe-(C2)–H of **Ac-Gly-Phe-CN** at 5.06 ppm with a CN resonances at 119.1 ppm and a C=O resonances at 171.7 ppm, which is characteristic of peptide bond formation of **Phe-CN**. See Supplementary Figure 163 for expanded ^1H NMR spectrum.

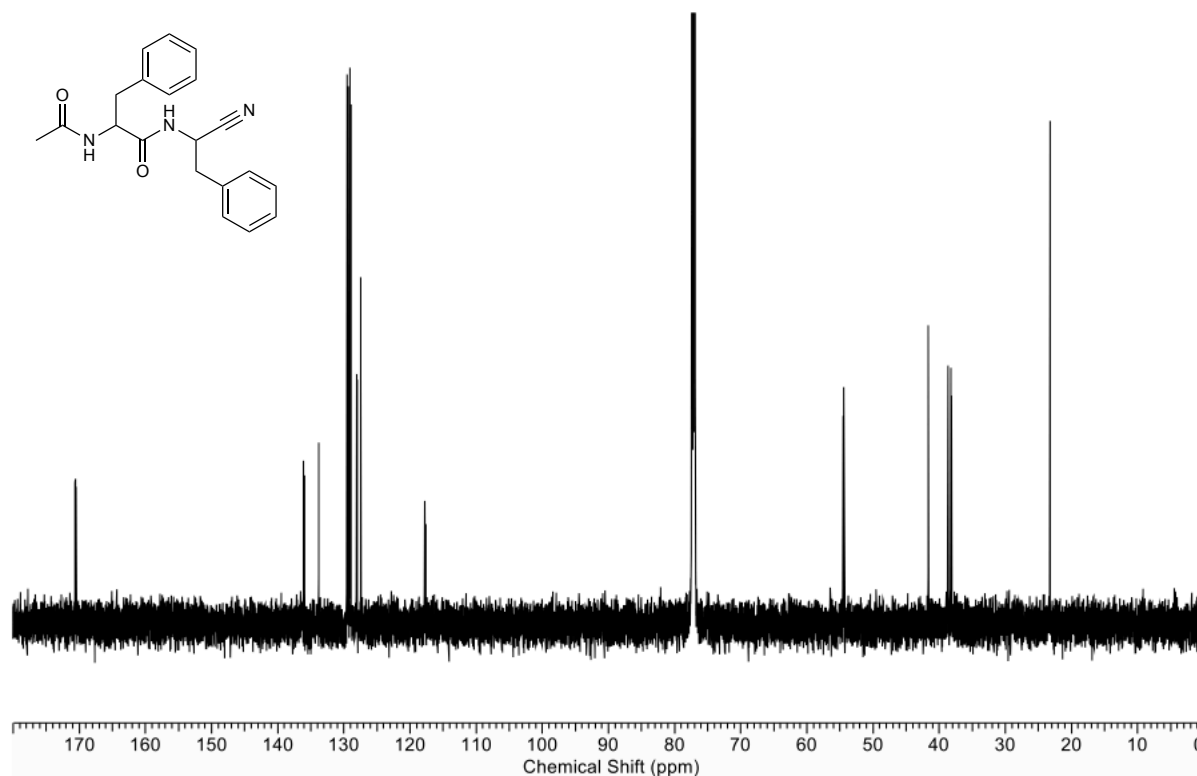
N-Acetylphenylalanylphenylalanine nitrile **Ac-Phe-Phe-CN**



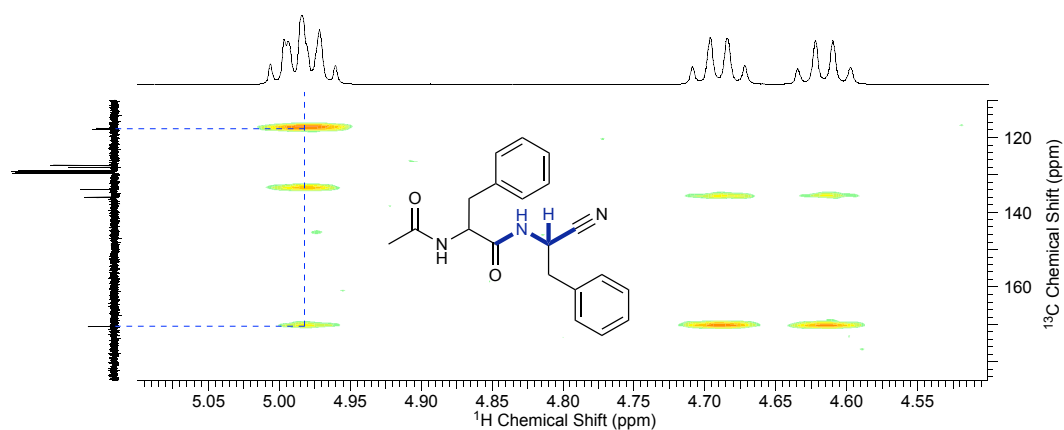
Prepared following *general procedure G* from sodium-2-acetamido-3-phenylpropanethioate **Ac-Phe-S-Na⁺** (123 mg, 0.5 mmol) and *rac*-phenylalanine nitrile hydrochloride **Phe-CN**·HCl⁴¹ (182 mg, 1.0 mmol). **Ac-Phe-Phe-CN** was isolated as a white solid (151 mg, 90%) after flash column chromatography (SiO₂, CH₂Cl₂/MeOH (100:0 to 95:5)). *R_f* 0.50 (CH₂Cl₂/MeOH, 95:5). ¹H NMR (600 MHz, CDCl₃, 53:46 mixture of diastereoisomers a/b; partial assignment of diastereoisomers) δ 7.35 - 7.22 (m, 12H, Ar), 7.20 - 7.10 (m, 8H, Ar), 7.00 (d, *J* = 8.2 Hz, 1H, Phe₂^a-NH), 6.78 (d, *J* = 8.0 Hz, 1H, Phe₂^b-NH), 6.12 (d, *J* = 8.1 Hz, 1H, Phe₁-NH), 6.10 (d, *J* = 8.4 Hz, 1H, Phe₁-NH), 5.01 - 4.95 (m, 2H, Phe₂^a-(C2)-H, Phe₂^b-(C2)-H), 4.70 (app. q, *J* = 7.6 Hz, 1H, Phe₁^a-(C2)-H), 4.62 (app. q, *J* = 7.5 Hz, 1H, Phe₁^b-(C2)-H), 3.05 - 2.97 (m, 6H, Phe₁^a-(C3)-H₂, Phe₂^b-(C3)-H₂ x 2), 2.94 (ABX, *J* = 6.0, 13.8 Hz, 1H, Phe₂^a-(C3)-H), 2.86 (ABX, *J* = 7.5, 13.8 Hz, 1H, Phe₂^a-(C3')-H), 1.95 (s, 3H, (COCH₃)^b), 1.94 (s, 3H, (COCH₃)^a). ¹³C NMR (151 MHz, CDCl₃, 53:46 mixture of diastereoisomers a/b; partial assignment of diastereoisomers) δ 170.6 (CO), 170.6 (CO), 170.5 (CO), 170.4 (CO), 136.1 (Ar), 136.0 (Ar), 133.8 (Ar), 133.7 (Ar), 129.6 (Ar x 2), 129.5 (Ar x 2), 129.4 (Ar x 2), 129.3 (Ar x 2), 129.1 (Ar x 2), 129.0 (Ar x 2), 129.0 (Ar x 2), 128.9 (Ar x 2), 128.1 (Ar), 128.0 (Ar), 127.43 (Ar), 127.43 (Ar), 117.8 (CN), 117.6 (CN), 54.5 (Phe₂-(C2)), 54.4 (Phe₂-(C2)), 41.7 (Phe₁-(C2)), 41.6 (Phe₁-(C2)), 38.7 (Phe₂-(C3)), 38.6 (Phe₂-(C3)), 38.3 (Phe₁-(C3)), 38.2 (Phe₁-(C3)), 23.2 (COCH₃ x 2). HRMS-EI-MS [*M*]⁺⁺ calcd. for C₂₀H₂₁N₃O₂ 335.1628; obs. 335.1629. IR (solid, cm⁻¹) 3307, 3278, 3064, 3028, 2966, 2922, 1666, 1645, 1535.



Supplementary Figure 166. ¹H NMR (600 MHz, CDCl₃, 1.0–8.0 ppm) spectrum of isolated *N*-acetylphenylalanylphenylalanine nitrile **Ac-Phe-Phe-CN**.

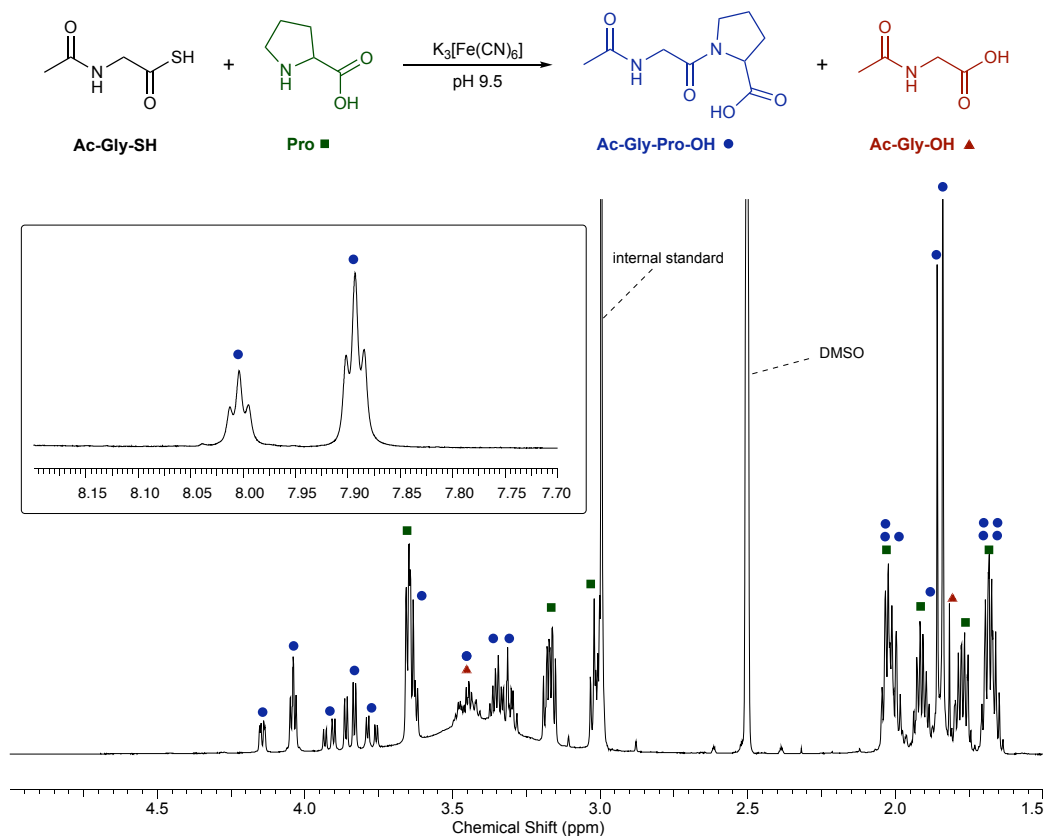


Supplementary Figure 167. ^{13}C NMR (151 MHz, CDCl_3 , 0–180 ppm) spectrum of isolated *N*-acetylphenylalanylphenylalanine nitrile **Ac-Phe-Phe-CN**.



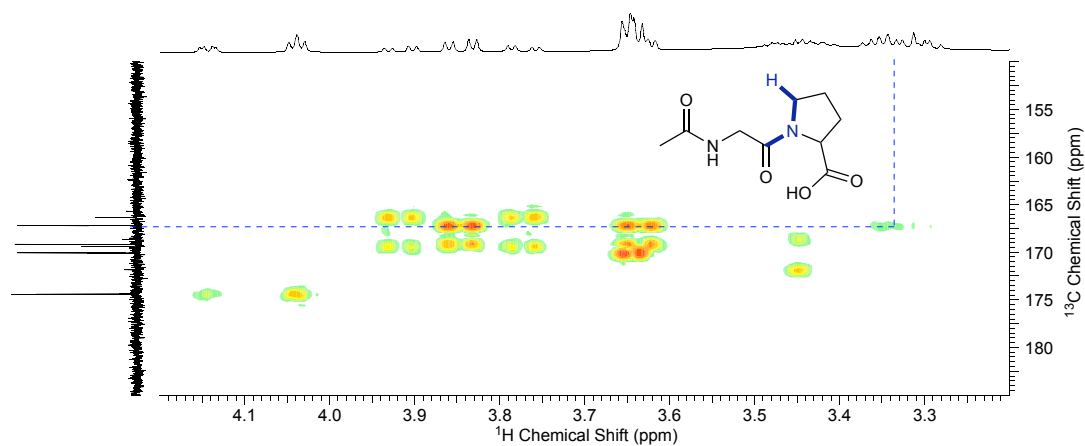
Supplementary Figure 168. ^1H – ^{13}C HMBC (^1H –600 MHz [4.50–5.10 ppm], ^{13}C –151 MHz [110–185 ppm], CDCl_3) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^2J_{\text{CH}}$ coupling of $\text{Phe}_2(\text{C}2)\text{--H}$ of **Ac-Phe-Phe-CN** at 4.95–5.01 ppm with CN resonances at 117.6 and 117.8 ppm and C=O resonances between 170.4–170.6 ppm, which is characteristic of peptide bond formation of **Phe-CN**. See Supplementary Figure 166 for expanded ^1H NMR spectrum.

N-Acetylglycylproline **Ac-Gly-Pro-OH**



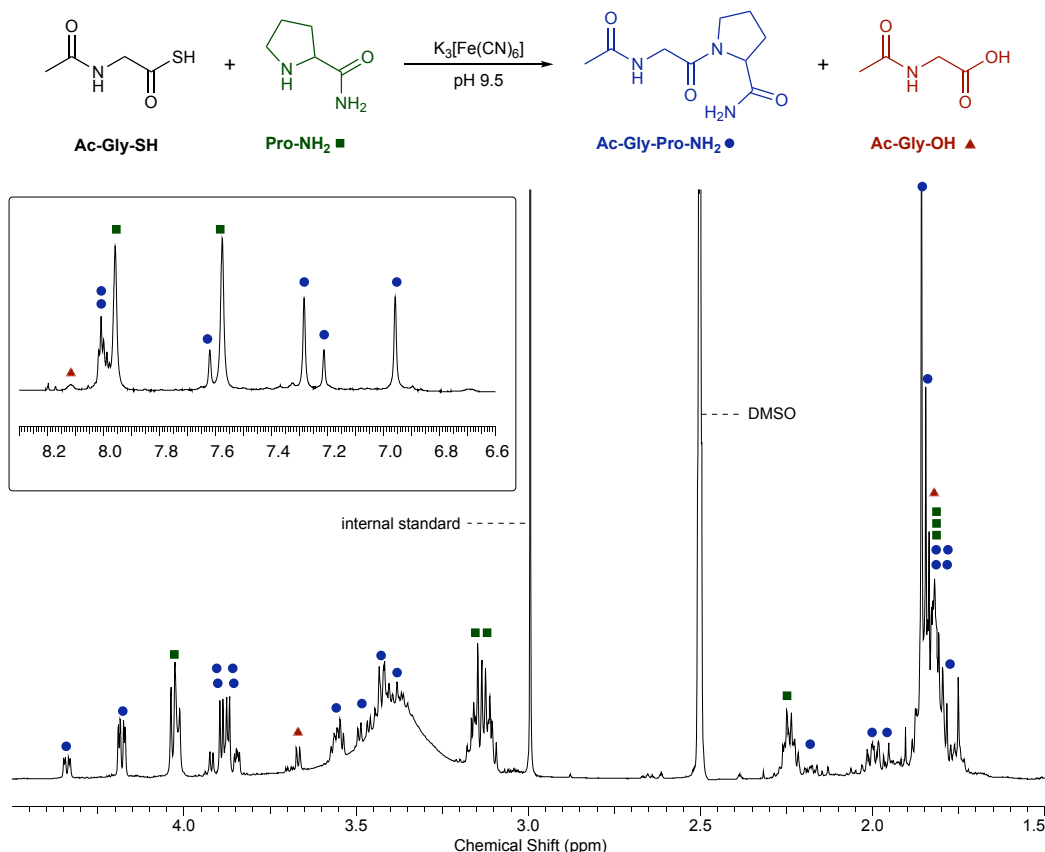
Supplementary Figure 169. ^1H NMR (600 MHz, $\text{DMSO-}d_6$, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-proline (**Pro**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature, after the solution was lyophilised and the resultant residue was dissolved in $\text{DMSO-}d_6$. Yields were obtained prior to lyophilisation. Inset: ^1H NMR (600 MHz, $\text{DMSO-}d_6$, 7.7–8.2 ppm) showing the NH resonances for the two rotamers of *N*-acetylglycylproline **Ac-Gly-Pro-OH**.

^1H NMR (600 MHz, $\text{DMSO-}d_6$, mixture of rotamers *a/b*, 67:33) *N*-Acetylglycylproline **Ac-Gly-Pro-OH_a** (major rotamer (*a*)) (●): δ 7.90 (t, J = 5.7 Hz, 1H, Gly-(C2)–NH), 4.04 (app. t, J = 5.6 Hz, 1H, Pro-(C2)–H), 3.85 (ABX, J = 5.7, 16.8 Hz, 1H, Gly-(C2)–H), 3.64 – 3.62 (obs., 1H, Gly-(C2')–H), 3.38 – 3.27 (m, 2H, Pro-(C5)–H₂), 2.06 – 1.99 (m, 2H, Pro-(C3)–H₂), 1.84 (s, 3H, (COCH₃)), 1.71 – 1.64 (m, 2H, Pro-(C4)–H₂). *N*-Acetylglycylproline **Ac-Gly-Pro-OH_b** (minor rotamer (*b*)) (●): δ 8.01 (t, J = 5.3 Hz, 1H, Gly-(C2)–NH), 4.14 (dd, J = 2.7, 8.6 Hz, 1H, Pro-(C2)–H), 3.92 (ABX, J = 5.3, 17.0 Hz, 1H, Gly-(C2)–H), 3.77 (ABX, J = 5.3, 17.0 Hz, 1H, Gly-(C2')–H), 3.51 – 3.40 (m, 2H, (C5)–H₂), 2.03 – 1.97 (m, 1H, (C3)–H), 1.86 (s, 3H, (COCH₃)), 1.89 – 1.83 (m, 1H, (C3')–H), 1.70 – 1.64 (m, 2H, (C4)–H₂). *L*-Proline **Pro** (■): δ 3.65 (dd, J = 5.8, 8.4 Hz, 1H, (C2)–H), 3.17 (ddd, J = 5.8, 7.1, 11.2 Hz, 1H, (C5)–H), 3.04 – 3.00 (m, 1H, (C5')–H), 2.02 – 1.98 (m, 1H, (C3)–H), 1.94 – 1.88 (m, 1H, (C3')–H), 1.81 – 1.74 (m, 1H, (C4)–H), 1.71 – 1.66 (m, 1H, (C4')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.46 – 3.44 (obs., 2H, (C2)–H₂), 1.82 (s, 3H, (COCH₃)).



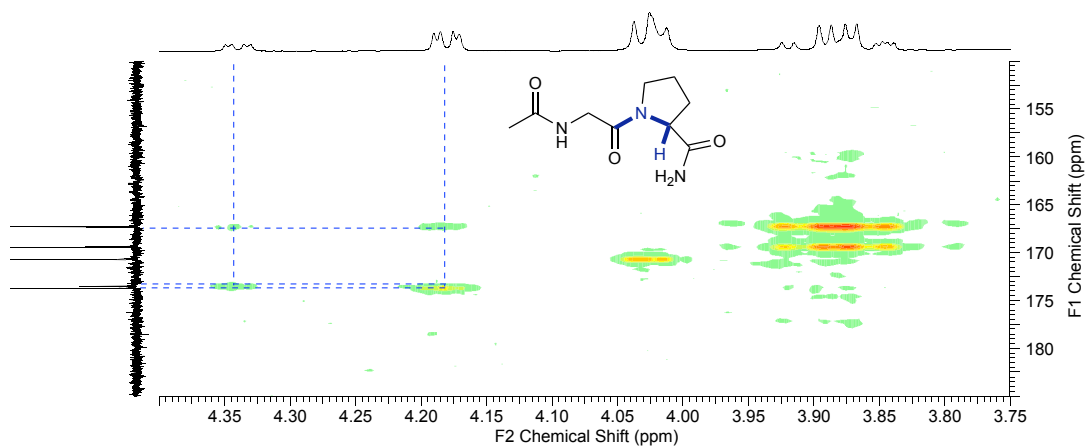
Supplementary Figure 170. ^1H – ^{13}C HMBC (^1H –600 MHz [3.4–4.2 ppm], ^{13}C –151 MHz [150–185 ppm], $\text{DMSO-}d_6$) spectrum showing a $^3J_{\text{CH}}$ coupling of Pro-(C5)–H of the major rotamer (rotamer a) of **Ac-Gly-Pro-OH_a** centred at 3.33 ppm with the Gly-(C1) C=O resonance at 167.2, which is characteristic of peptide bond formation of **Pro**. See Supplementary Figure 169 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycylprolinamide **Ac-Gly-Pro-NH₂**



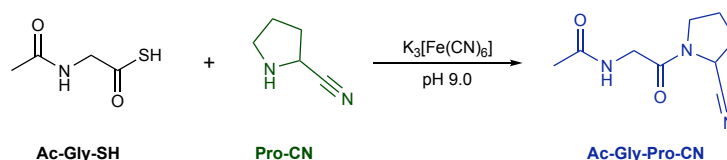
Supplementary Figure 171. ¹H NMR (600 MHz, DMSO-*d*₆, 1.5–4.5 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-prolinamide (**Pro-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature, after the solution was lyophilised and the resultant residue was dissolved in DMSO-*d*₆. Yields were obtained prior to lyophilisation. Inset: ¹H NMR (600 MHz, DMSO-*d*₆, 6.6–8.3 ppm) showing the NH resonances for the two rotamers (rotamer a/b) of *N*-acetylglycylprolinamide **Ac-Gly-Pro-NH₂**.

¹H NMR (600 MHz, DMSO-*d*₆, mixture of rotamers a/b, 72:28) *N*-Acetylglycylprolinamide **Ac-Gly-Pro-NH_{2a}** (major rotamer (a)) (●): δ 8.01 (t, *J* = 5.6 Hz, 1H, Gly-(C2)–NH), 7.28 (br. s, 1H, Pro-(C1)–NH), 6.96 (s, 1H, Pro-(C1')–NH), 4.18 (dd, *J* = 2.9, 8.8 Hz, 1H, Pro-(C2)–H), 3.91 (ABX, *J* = 5.6, 17.1 Hz, 1H, Gly-(C2)–H), 3.86 (ABX, *J* = 5.6, 17.1 Hz, 1H, Gly-(C2')–H), 3.59–3.53 (m, 1H, Pro-(C5)–H), 3.40–3.35 (m, 1H, Pro-(C5')–H), 2.04–1.95 (m, 1H, Pro-(C3)–H), 1.86 (s, 3H, (COCH₃)), 1.84–1.80 (m, 2H, Pro-(C4)–H₂), 1.79–1.71 (m, 1H, Pro-(C3')–H). *N*-Acetylglycylprolinamide **Ac-Gly-Pro-NH_{2b}** (minor rotamer (b)) (●): δ 8.00–7.97 (m, 1H, Gly-(C2)–NH), 7.62 (br. s., 1H, Pro-(C1)–NH), 7.21 (br. s., 1H, Pro-(C1')–NH), 4.34 (dd, *J* = 2.9, 8.5 Hz, 1H, Pro-(C2)–H), 3.88–3.83 (obs., 2H, Gly-(C2)–H₂), 3.51–3.44 (m, 1H, Pro-(C5)–H), 3.40–3.35 (m, 1H, Pro-(C5')–H), 2.21–2.13 (m, 1H, Pro-(C3)–H), 2.04–1.96 (m, 1H, Pro-(C3')–H), 1.85 (s, 3H, (COCH₃)), 1.84–1.77 (m, 2H, Pro-(C4)–H₂). *L*-Prolinamide **Pro-NH₂** (■): δ 7.96 (br. s., 1H, NH), 7.58 (br. s., 1H, NH), 4.04–3.99 (m, 1H, (C2)–H), 3.19–3.07 (m, 2H, (C5)–H₂), 2.29–2.20 (m, 1H, (C3)–H), 1.89–1.77 (m, 3H, (C3')–H, (C4)–H₂). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 8.15–8.09 (br. s, 1H, NH), 3.67 (app. d, *J* = 5.8 Hz, 2H, (C2)–H), 1.84 (br. s., 3H, (COCH₃)).

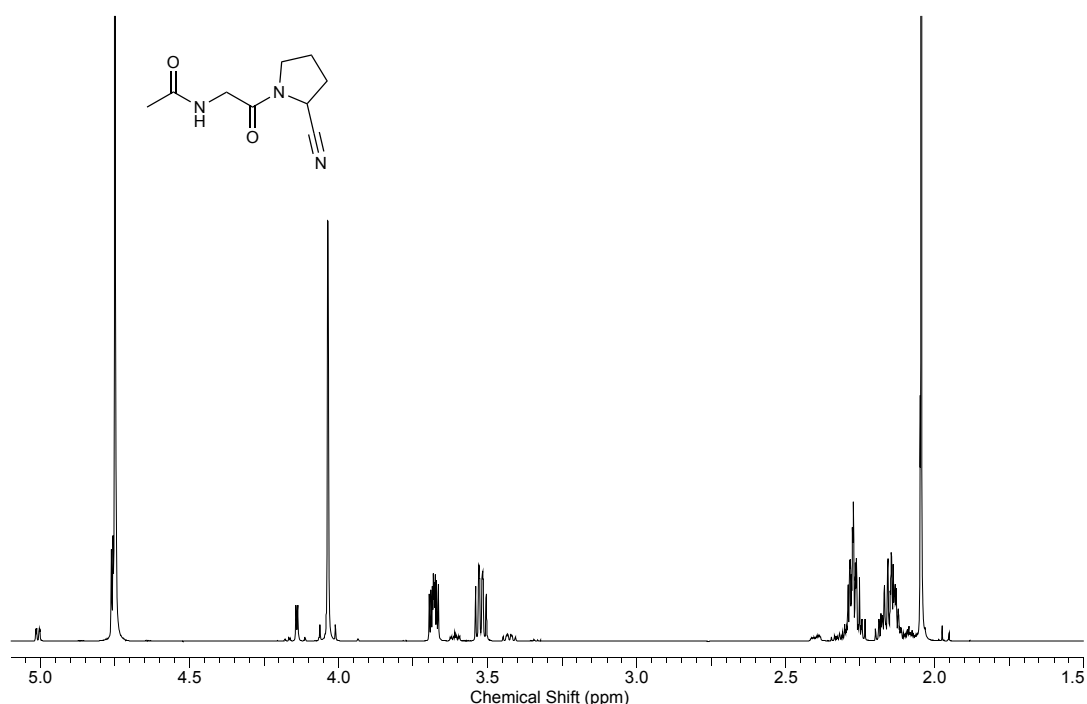


Supplementary Figure 172. ^1H - ^{13}C HMBC (^1H -600 MHz [3.75–4.40 ppm], ^{13}C -151 MHz [150–185 ppm], $\text{DMSO-}d_6$) spectrum showing the $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ couplings of Pro-(C2)-H of both rotamers (a/b) of **Ac-Gly-Pro-NH₂** at 4.18 and 4.34 ppm with C=O resonances at 167.32 and 173.5 (minor rotamer b **Ac-Gly-Pro-NH_{2b}**) and 167.30 and 173.7 ppm (major rotamer a **Ac-Gly-Pro-NH_{2a}**), which is characteristic of peptide bond formation of L-prolinamide **Pro-NH₂**. See Supplementary Figure 171 for expanded and labelled ^1H NMR spectrum.

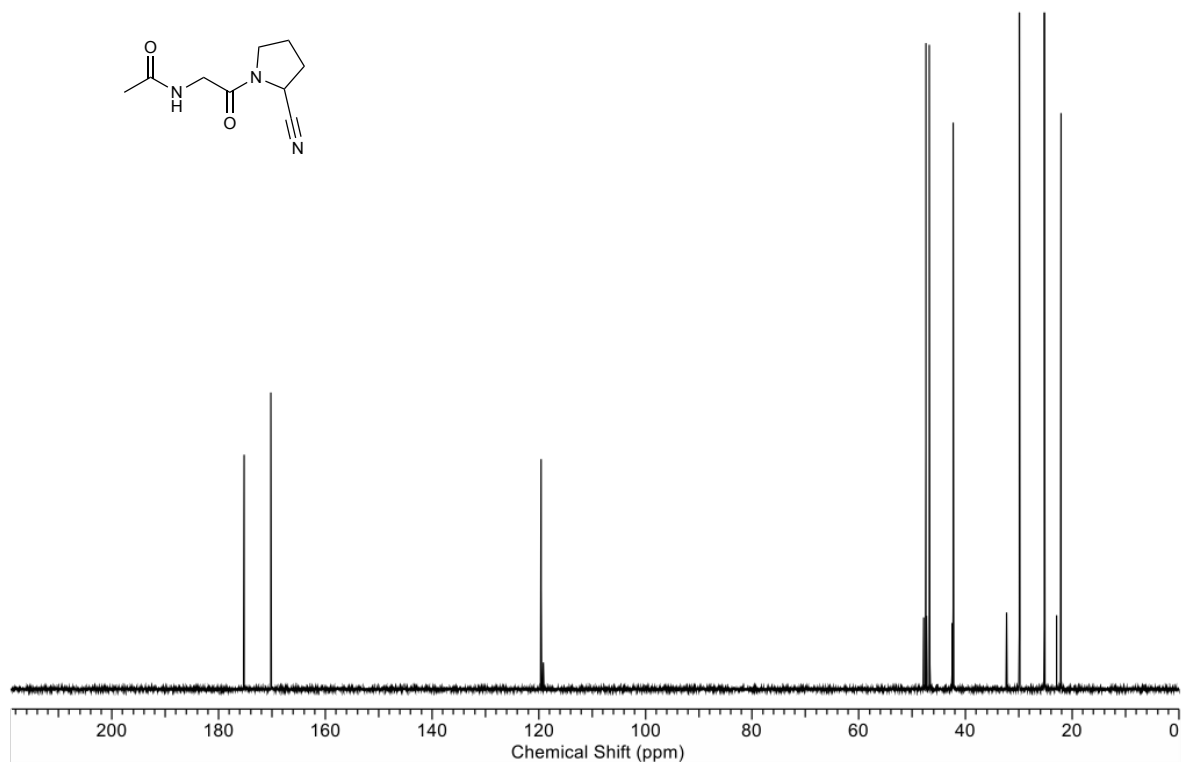
N-Acetylglycylproline nitrile **Ac-Gly-Pro-CN**



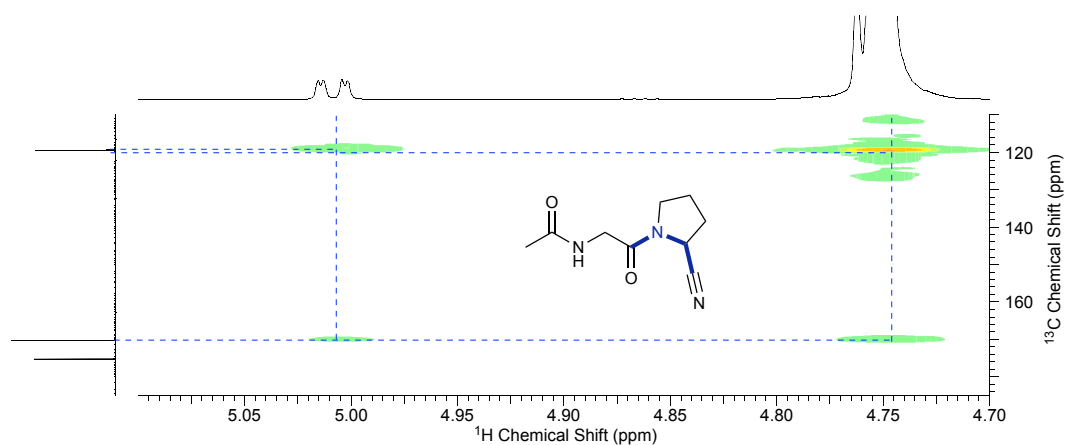
Prepared following *general procedure G* on a 0.1 mmol scale, from sodium-2-acetamido-ethanethioate **Ac-Gly-S⁻Na⁺** (15.5 mg, 0.1 mmol) and S-proline nitrile hydrochloride **Pro-CN-HCl** (26.5 mg, 0.2 mmol). **Ac-Gly-Pro-CN** was isolated as a colourless oil (16.6 mg, 85%) after flash column chromatography (SiO₂, EtOAc:MeOH (80:20)). *R_f* 0.33 (EtOAc:MeOH (80:20)), and as a mixture of two rotamers (a:b; 89:11). ¹H NMR (700 MHz, D₂O, major rotamer a) δ 4.76 (obs., 1H, Pro-(C2)-H), 4.05 (AB, *J* = 17.7 Hz, 1H, Gly-(C2)-H), 4.02 (AB, *J* = 17.7 Hz, 1H, Gly-(C2)-H), 3.68 (ddd, *J* = 4.1, 7.4, 9.8 Hz, 1H, Pro-(C5)-H), 3.52 (ddd, *J* = 7.4, 8.3, 9.7 Hz, 1H, Pro-(C5)-H'), 2.30 - 2.23 (m, 2H, Pro-(C3)-H₂), 2.20 - 2.12 (m, 2H, Pro-(C4)-H₂), 2.05 (s, 3H, (COCH₃)). ¹³C NMR (176 MHz, D₂O) δ 175.3 (COCH₃), 170.2 (Gly-(C1)), 119.6 (Pro-(C1)), 47.5 (Pro-(C2)), 46.8 (Pro-(C5)), 42.3 (Gly-(C2)), 29.9 (Pro-(C3)), 25.2 (Pro-(C4)), 22.1 (COCH₃). ¹H NMR (700 MHz, D₂O, minor rotamer b) δ 5.01 (dd, *J* = 1.8, 7.8 Hz, 1H, Pro-(C2)-H), 4.15 (AB, *J* = 16.9 Hz, 1H, Gly-(C2)-H), 4.13 (AB, *J* = 16.9 Hz, 1H, Gly-(C2)-H'), 3.61 (ddd, *J* = 3.2, 7.7, 11.2 Hz, 1H, Pro-(C5)-H), 3.43 (ddd, *J* = 7.7, 9.9, 10.5 Hz, 1H, Pro-(C5)-H'), 2.42 - 2.37 (m, 1H, Pro-(C3)-H), 2.35 - 2.30 (m, 1H, Pro-(C3)-H'), 2.12 - 2.07 (m, 1H, Pro-(C4)-H), 2.07 - 2.03 (m, 1H, Pro-(C4)-H'), 2.05 (s, 3H, (COCH₃)). ¹³C NMR (176 MHz, D₂O) δ 175.2 (COCH₃), 170.2 (Gly-(C1)), 119.2 (Pro-(C1)), 47.9 (Pro-(C2)), 47.3 (Pro-(C5)), 42.5 (Gly-(C2)), 32.3 (Pro-(C3)), 22.9 (Pro-(C4)), 22.2 (COCH₃). HRMS-ESI-MS [M+H]⁺ calcd. for C₉H₁₄N₃O₂ 196.1081; obs. 196.1081. IR (oil, cm⁻¹) 3315, 3224, 3045, 2463, 2360, 2333, 1664, 1631, 1552.



Supplementary Figure 173. ¹H NMR (700 MHz, D₂O, 1.5–5.1 ppm) spectrum of purified *N*-acetylglycylproline nitrile **Ac-Gly-Pro-CN**

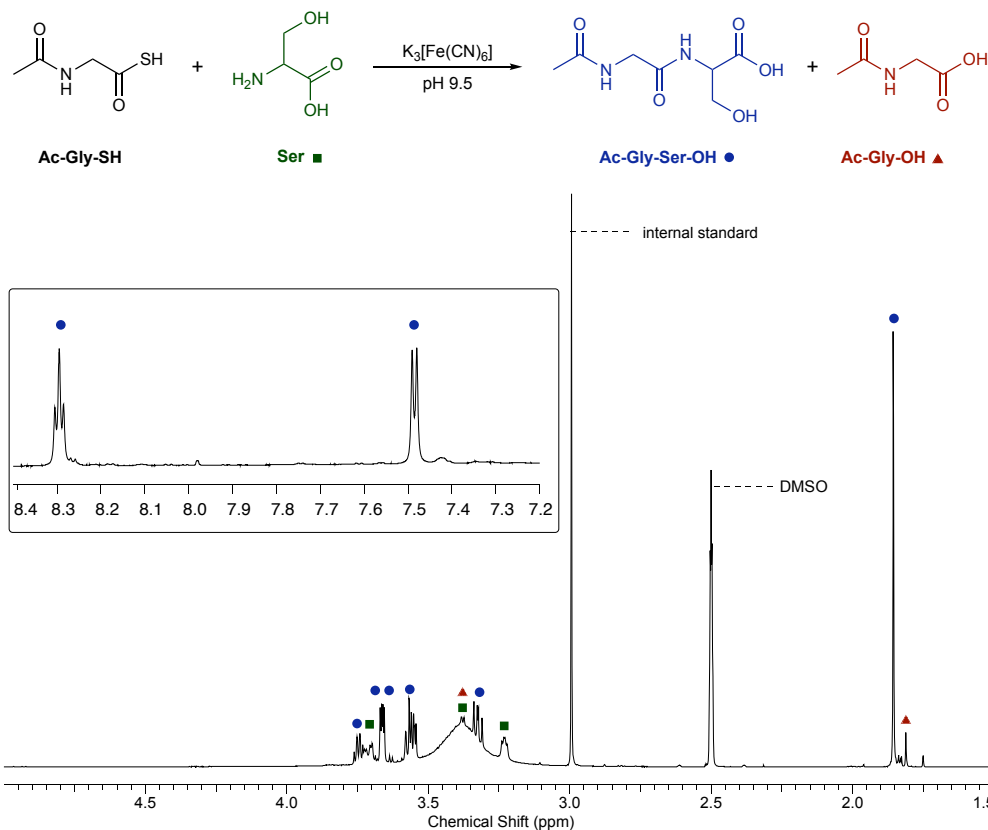


Supplementary Figure 174. ^{13}C NMR (176 MHz, D_2O , 0–220 ppm) spectrum of purified *N*-acetylglycylproline nitrile **Ac-Gly-Pro-CN**



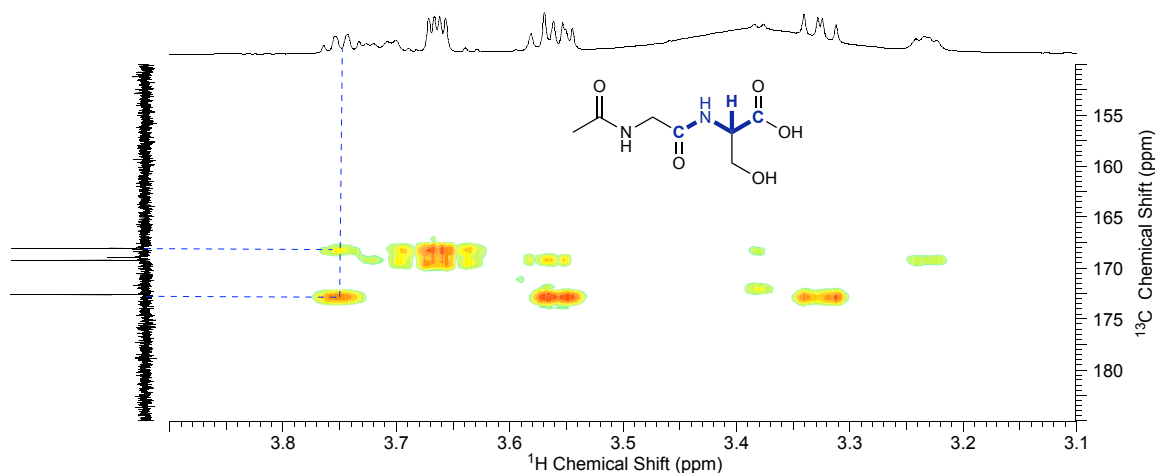
Supplementary Figure 175. ^1H – ^{13}C HMBC (^1H –700 MHz [4.7–5.1 ppm], ^{13}C –176 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Pro-(C2)–H of **Ac-Gly-Pro-CN** (major rotamer a) at 4.76 ppm with a CN resonance at 119.6 ppm and a C=O resonance at 170.2 ppm, and Pro-(C2)–H of **Ac-Gly-Pro-CN** (minor rotamer b) at 5.01 ppm with a CN resonance at 119.2 ppm and a C=O resonance at 170.2 ppm, which is characteristic of peptide bond formation of **Pro-CN**. See Supplementary Figure 173 for expanded ^1H NMR spectrum.

N-Acetylglycylserine **Ac-Gly-Ser-OH**



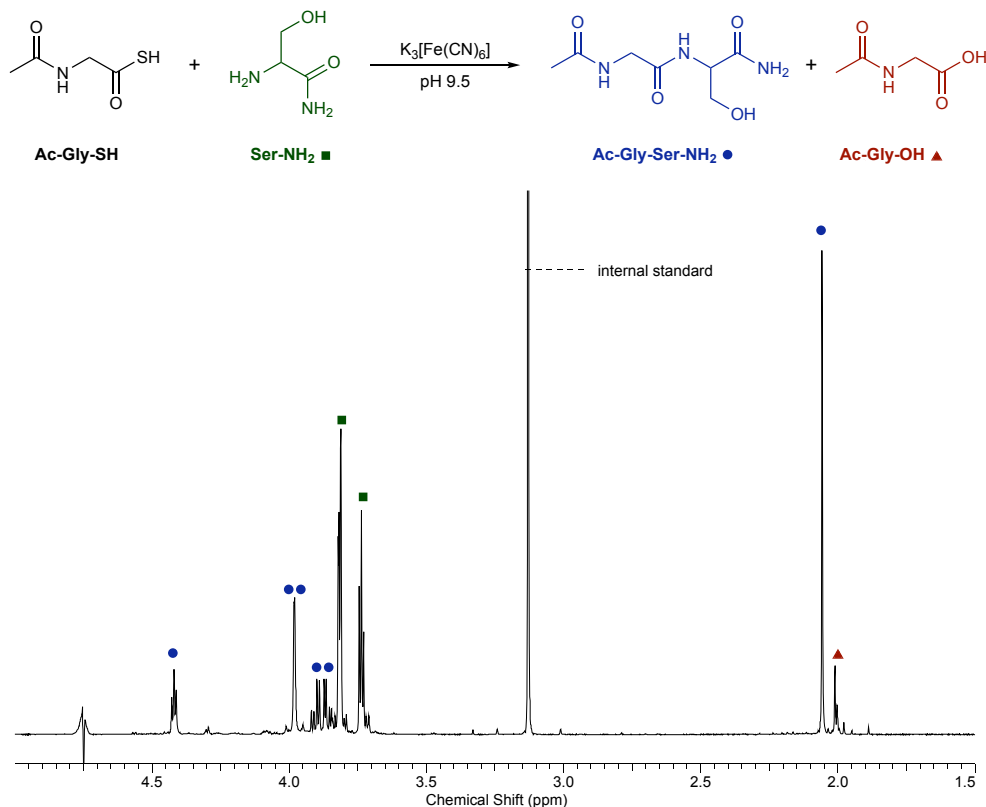
Supplementary Figure 176. 1H NMR (600 MHz, $DMSO-d_6$, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-serine (**Ser**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature, after the solution was lyophilised and the resultant residue was dissolved in $DMSO-d_6$. Yields were obtained prior to lyophilisation. Inset: 1H NMR (600 MHz, $DMSO-d_6$, 7.2–8.4 ppm) showing the NH resonances for *N*-acetylglycylserine **Ac-Gly-Ser-OH**.

1H NMR (600 MHz, $DMSO-d_6$) *N*-Acetylglycylserine **Ac-Gly-Ser-OH** (●): δ 8.30 (t, J = 5.9 Hz, 1H, Gly-(C2)–NH), 7.49 (d, J = 6.1 Hz, 1H, Ser-(C2)–NH), 3.77 – 3.73 (m, 1H, Ser-(C2)–H), 3.68 (ABX, J = 5.9, 16.2 Hz, 1H, Gly-(C2)–H), 3.65 (ABX, J = 5.9, 16.2 Hz, 1H, Gly-(C2')–H), 3.58 – 3.55 (m, 1H, Ser-(C3)–H), 3.32 (ABX, J = 7.3, 9.7 Hz, 1H, Ser-(C3')–H), 1.86 (s, 3H, (COCH₃)). L-Serine **Ser** (■): δ 3.71 (ABX, J = 4.4, 11.5 Hz, 1H, (C3)–H), 3.35 – 3.31 (obs., 1H, (C3')–H), 3.23 (dd, J = 4.4, 6.6 Hz, 1H, (C2)–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.38 (app. d, J = 4.8 Hz, 2H, (C2)–H₂), 1.81 (s, 3H, (COCH₃)).



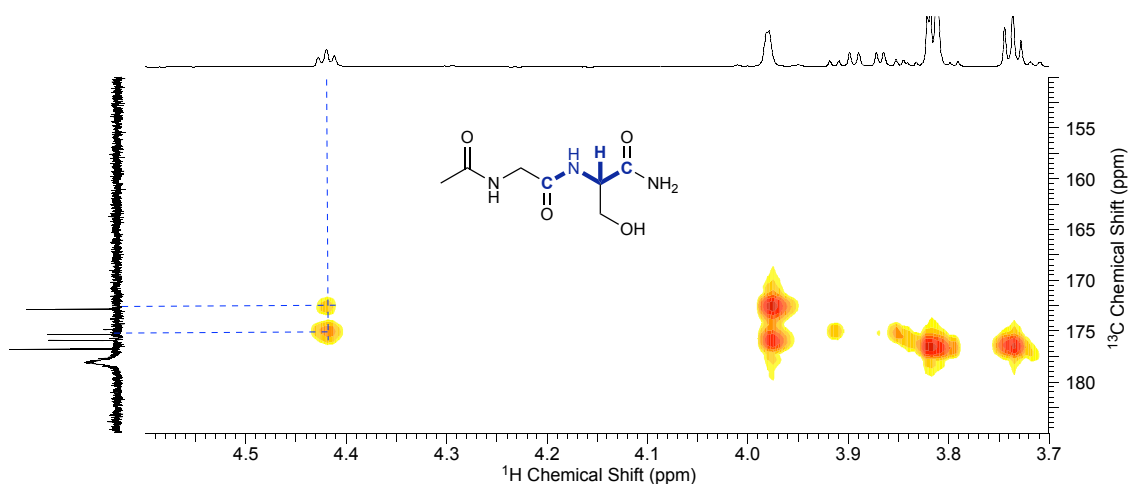
Supplementary Figure 177. 1H – ^{13}C HMBC (1H –600 MHz [3.6–3.9 ppm], ^{13}C –151 MHz [150–185 ppm], $DMSO-d_6$) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Ser-(C2)–H of **Ac-Gly-Ser-OH** centred at 3.75 ppm with two C=O resonances at 172.6 and 168.0 ppm, which is characteristic of peptide bond formation of **Ser**. See Supplementary Figure 176 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylserinamide **Ac-Gly-Ser-NH₂**



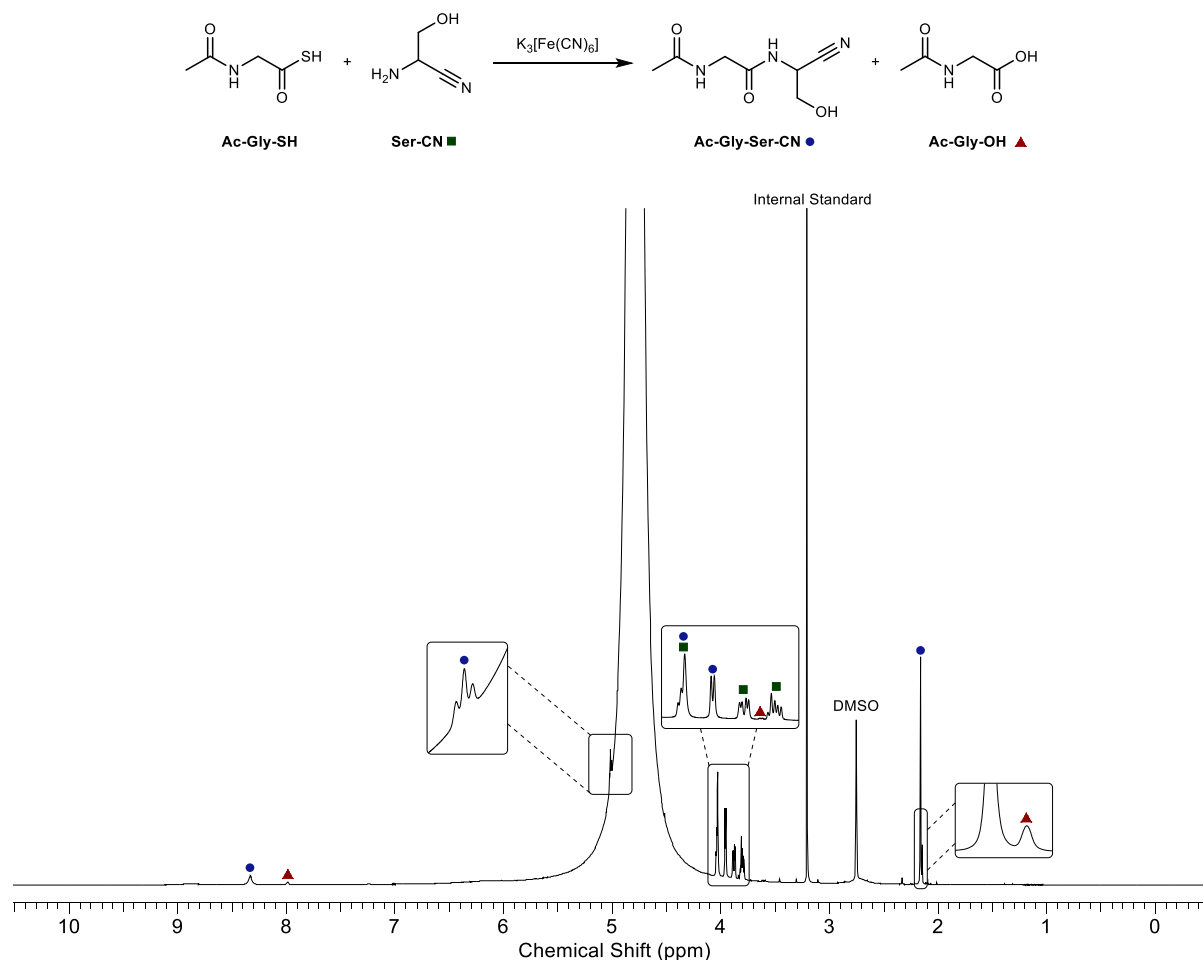
Supplementary Figure 178. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-serinamide (**Ser-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylserinamide **Ac-Gly-Ser-NH₂** (●): δ 4.42 (app. t, $J = 4.8$ Hz, 1H, Ser-(C2)-H), 4.00 (AB, $J = 17.4$ Hz, 1H, Gly-(C2)-H), 3.96 (AB, $J = 17.4$ Hz, 1H, Gly-(C2')-H), 3.90 (ABX, $J = 5.5, 11.8$ Hz, 1H, Ser-(C3)-H), 3.86 (ABX, $J = 4.2, 11.8$ Hz, 1H, Ser-(C3')-H), 2.06 (s, 3H, (COCH₃)). *L*-Serinamide **Ser-NH₂** (■): δ 3.83 - 3.80 (m, 2H, (C3)-H₂), 3.76 - 3.72 (m, 1H, (C2)-H). *N*-Acetylglycine **Ac-Gly-OH** (▲) (partial assignment): δ 2.01 (s, 3H, (COCH₃)).



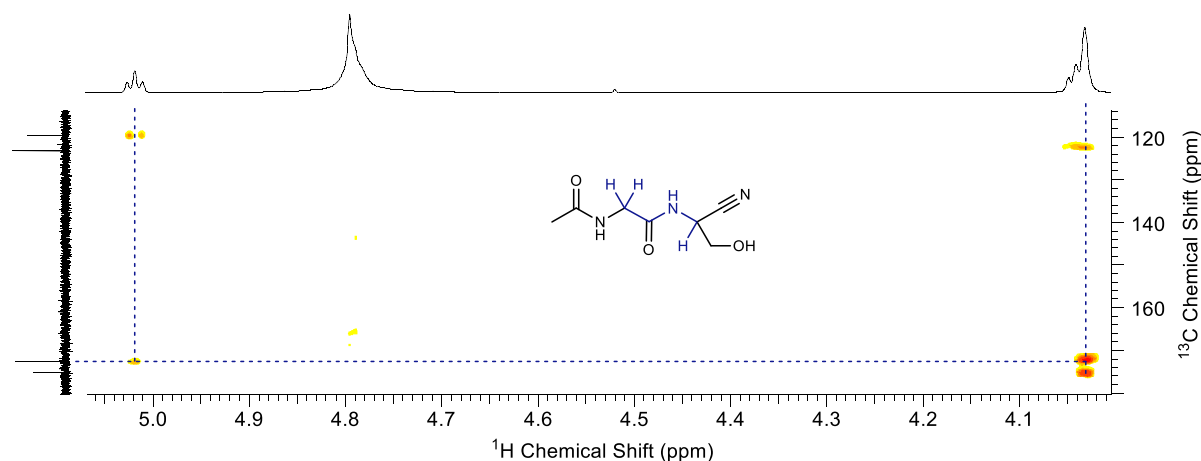
Supplementary Figure 179. ¹H-¹³C HMBC (¹H-600 MHz [3.7–4.6 ppm], ¹³C-151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Ser-(C2)-H of **Ac-Gly-Ser-NH₂** at 4.42 ppm with two C=O resonances at 172.8 and 175.3 ppm, which is characteristic of peptide bond formation of **Ser-NH₂**. See Supplementary Figure 178 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylserine nitrile **Ac-Gly-Ser-CN**



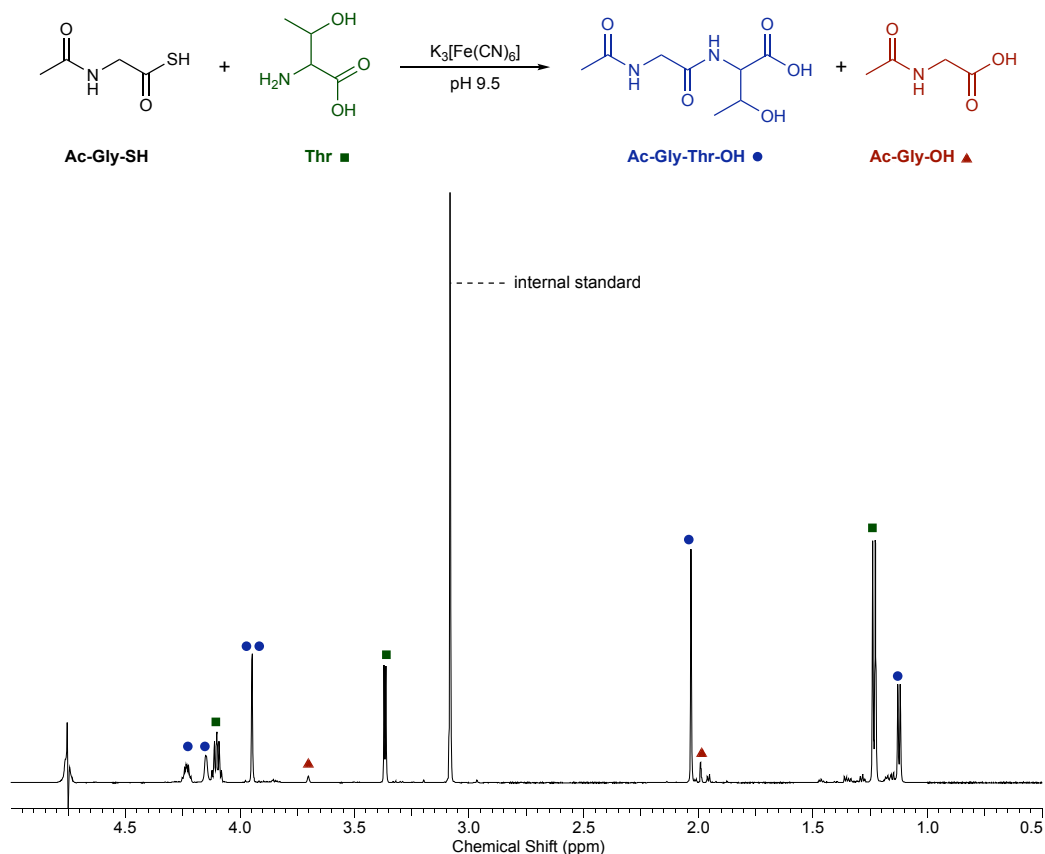
Supplementary Figure 180. 1H NMR (700 MHz, $D_2O/H_2O/DMSO-d_6$ 1:49:50, -0.50–10.50 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), serine nitrile (**Ser-CN**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at room temperature and pH 9.0 after 20 min.

1H NMR (700 MHz, $D_2O/H_2O/DMSO-d_6$ 1:49:50): *N*-Acetylglycylserine nitrile **Ac-Gly-Ser-CN** (•): δ 8.33 (br s, 1H, Ser-NH), 5.02, (obs. t, $J = 5.6$ Hz, 1H, Ser-(C2)-H), 4.03 (s, 2H, Gly-(C2)-H₂), 3.96 (d, $J = 5.6$ Hz, 2H, Ser-(C3)-H₂), 2.16 (s, 3H, (COCH₃)). Serine nitrile **Ser-CN**¹⁴ (■): δ 4.05 – 4.01 (obs. m, 1H, (C2)-H), 3.88 (ABX, $J = 4.3, 11.3$ Hz, 1H, (C3)-H), 3.80 (ABX, $J = 6.2, 11.3$ Hz, 1H, (C3)-H'). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 7.99 (br s, 1H, NH), 3.81 (app. d, $J = 5.8$ Hz, 2H, (C2)-H₂), 2.14 (s, 3H, (COCH₃)).



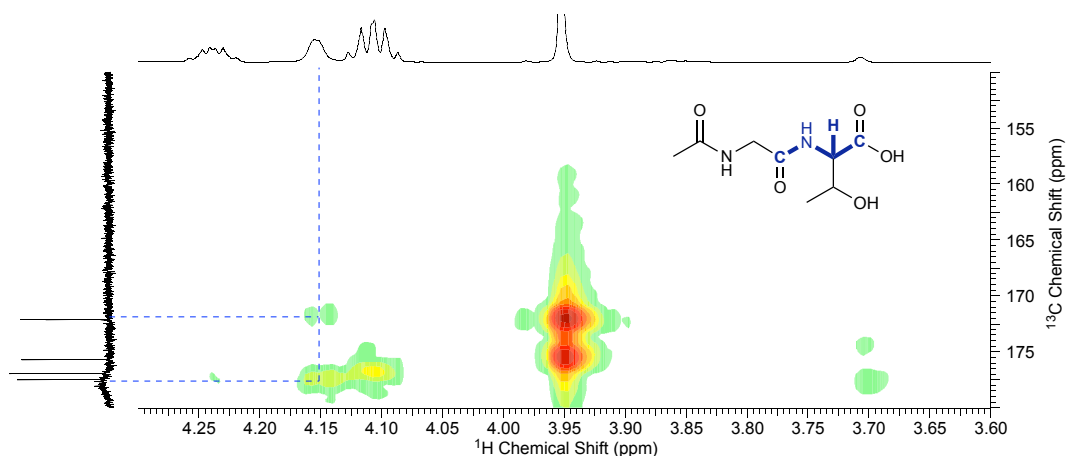
Supplementary Figure 181. 1H - ^{13}C HMBC (1H -700 MHz [4.05–5.05 ppm], ^{13}C -151 MHz [115–180 ppm], $D_2O/H_2O/DMSO-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Ser-(C2)-H of **Ac-Gly-Ser-CN** at 5.02 ppm with a CN resonance at 119.5 ppm and a C=O resonance at 172.2 ppm, which is characteristic of peptide bond formation of **Ser-CN**. See Supplementary Figure 180 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylthreonine **Ac-Gly-Thr-OH**



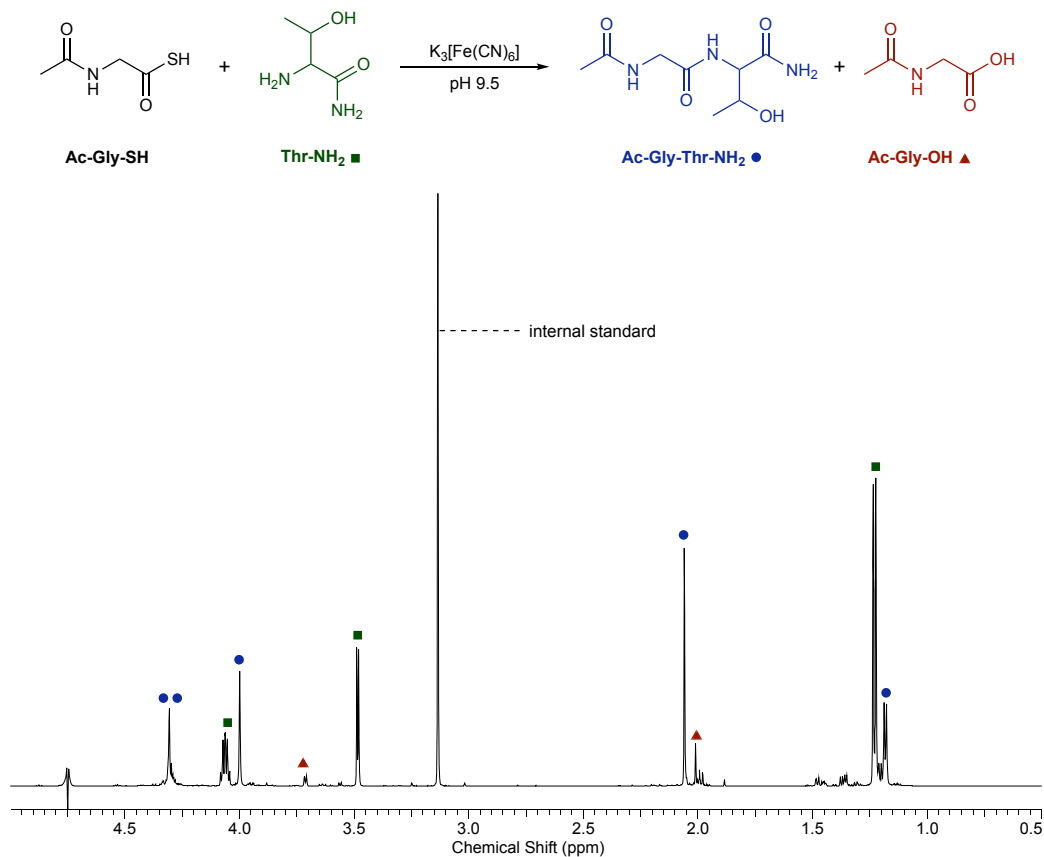
Supplementary Figure 182. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-threonine (**Thr**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycylthreonine **Ac-Gly-Thr-OH** (●): δ 4.26 – 4.20 (m, 1H, (C3)–H), 4.15 (br. s., 1H, (C2)–H), 3.98 (AB, $J = 17.4$ Hz, 1H, Gly-(C2)–H), 3.93 (AB, $J = 17.4$ Hz, 1H, Gly-(C2')–H), 2.03 (s, 3H, (COCH₃)), 1.12 (d, $J = 6.5$ Hz, 3H, (C4)–H₃). L-Threonine **Thr** (■): δ 4.10 (dq, $J = 4.8, 6.5$ Hz, 1H, (C3)–H), 3.37 (d, $J = 4.8$ Hz, 1H, (C2)–H), 1.23 (d, $J = 6.5$ Hz, 3H, (C4)–H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.70 (s, 2H, (C2)–H₂), 1.99 (s, 3H, (COCH₃)).



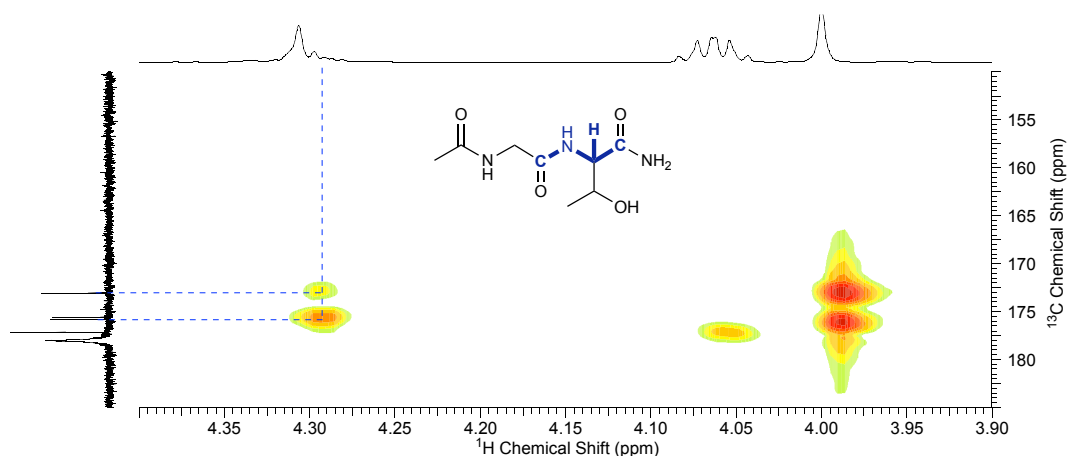
Supplementary Figure 183. 1H - ^{13}C HMBC (1H -600 MHz [3.6–4.3 ppm], ^{13}C -151 MHz [150–185 ppm], H_2O/D_2O 98:2) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Thr-(C2)–H of **Ac-Gly-Thr-OH** at 4.15 ppm with two C=O resonances at 172.1 and 177.5 ppm, which is characteristic of peptide bond formation of **Thr**. See Supplementary Figure 182 for expanded and labelled 1H NMR spectrum.

N-Acetylglycylthreoninamide **Ac-Gly-Thr-NH₂**



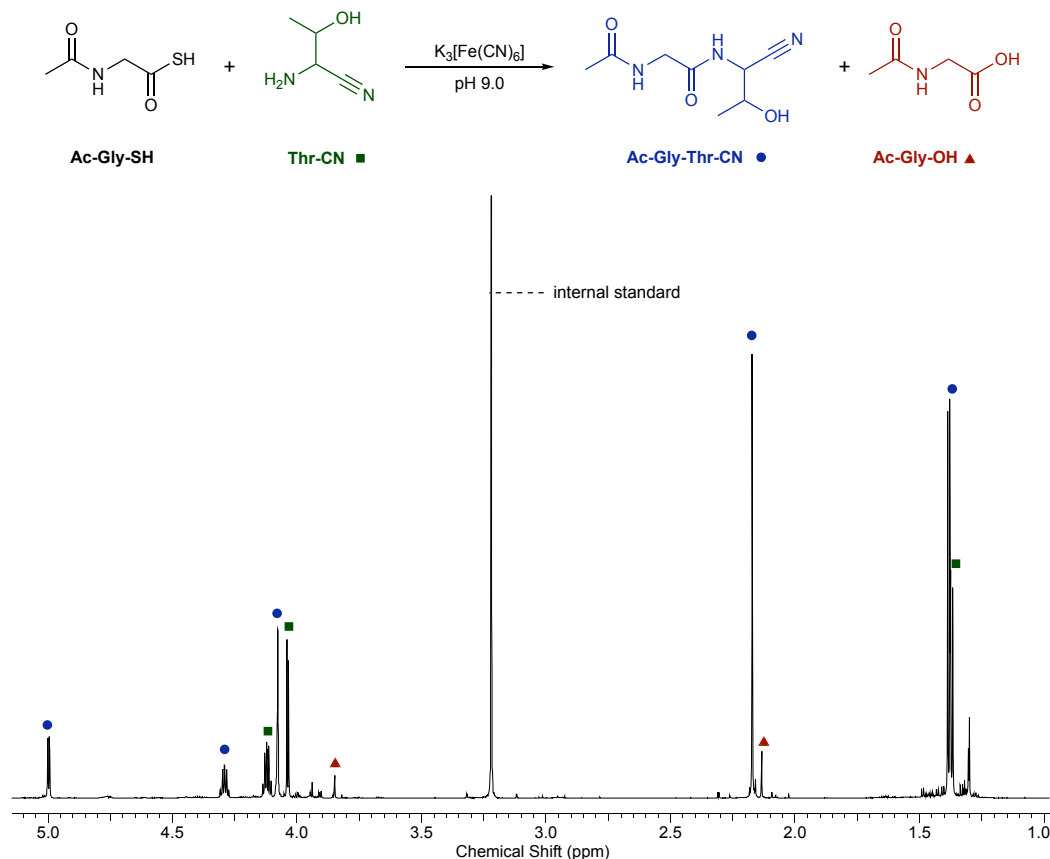
Supplementary Figure 184. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-threoninamide (**Thr-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylthreoninamide **Ac-Gly-Thr-NH₂** (●): δ 4.35 - 4.26 (m, 2H, Thr-(C3)–H, Thr-(C2)–H), 4.00 (s, 2H, Gly-(C2)–H₂), 2.06 (s, 3H, (COCH₃)), 1.23 (d, $J = 6.5$ Hz, 3H, Thr-(C4)–H₃). *L*-Threoninamide **Thr-NH₂** (■): δ 4.06 (dq, $J = 4.7, 6.5$ Hz, 1H, (C3)–H), 3.48 (d, $J = 4.7$ Hz, 1H, (C2)–H), 1.23 (d, $J = 6.5$ Hz, 3H, (C4)–H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.71 (app. d, $J = 5.6$ Hz, 2H, (C2)–H₂), 2.01 (s, 3H, (COCH₃)).



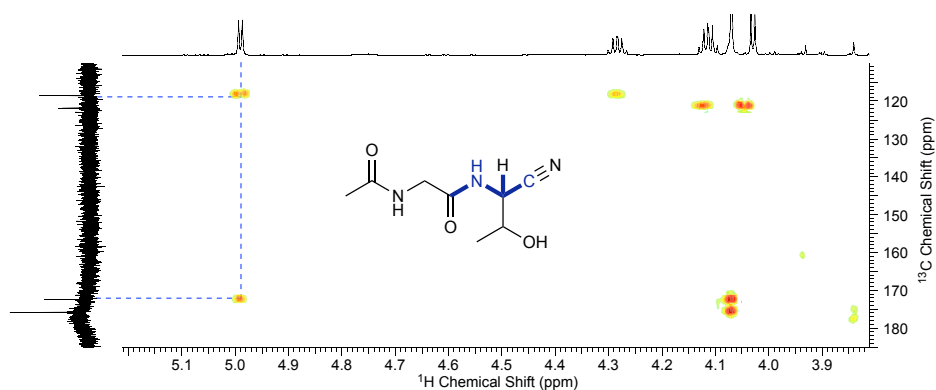
Supplementary Figure 185. ¹H-¹³C HMBC (¹H-600 MHz [3.9–4.4 ppm], ¹³C-151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ²J_{CH} and ³J_{CH} coupling of Thr-(C2)–H of **Ac-Gly-Thr-NH₂** centred at 4.29 ppm with two C=O resonances at 173.0 and 175.6 ppm, which is characteristic of peptide bond formation of **Thr-NH₂**. See Supplementary Figure 184 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylthreonine nitrile **Ac-Gly-Thr-CN**



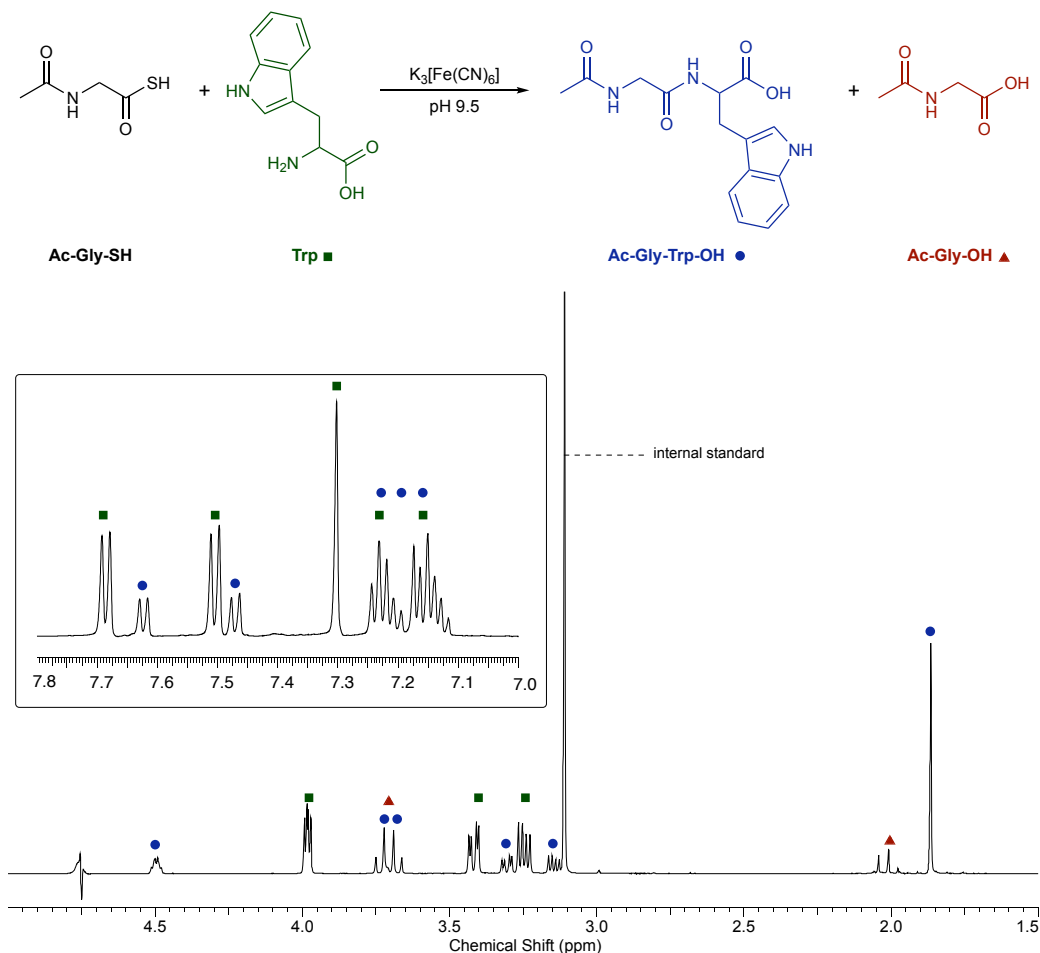
Supplementary Figure 186. 1H NMR (700 MHz, D_2O , noesygppr1d, 1.0–5.2 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with threonine nitrile (**Thr-CN**; 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pD 9.0 and room temperature.

1H NMR (700 MHz, D_2O , noesygppr1d) *N*-Acetylglycylthreonine nitrile **Ac-Gly-Thr-CN** (•): δ 5.00 (d, $J = 5.3$ Hz, 1H, Thr-(C2)-H), 4.29 (dq, $J = 5.3, 6.4$ Hz, 1H, Thr-(C3)-H), 4.08 (s, 2H, Gly-(C2)-H₂), 2.17 (s, 3H, (COCH₃)), 1.37 (d, $J = 6.4$ Hz, 3H, Thr-(C4)-H₃). Threonine nitrile **Thr-CN**⁵⁰ (■): δ 4.12 (dq, $J = 5.1, 6.3$ Hz, 1H, (C3)-H), 4.04 (d, $J = 5.1$ Hz, 1H, (C2)-H), 1.38 (d, $J = 6.3$ Hz, 3H, (C4)-H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.85 (s, 2H, (C2)-H₂), 2.13 (s, 3H, (COCH₃)).



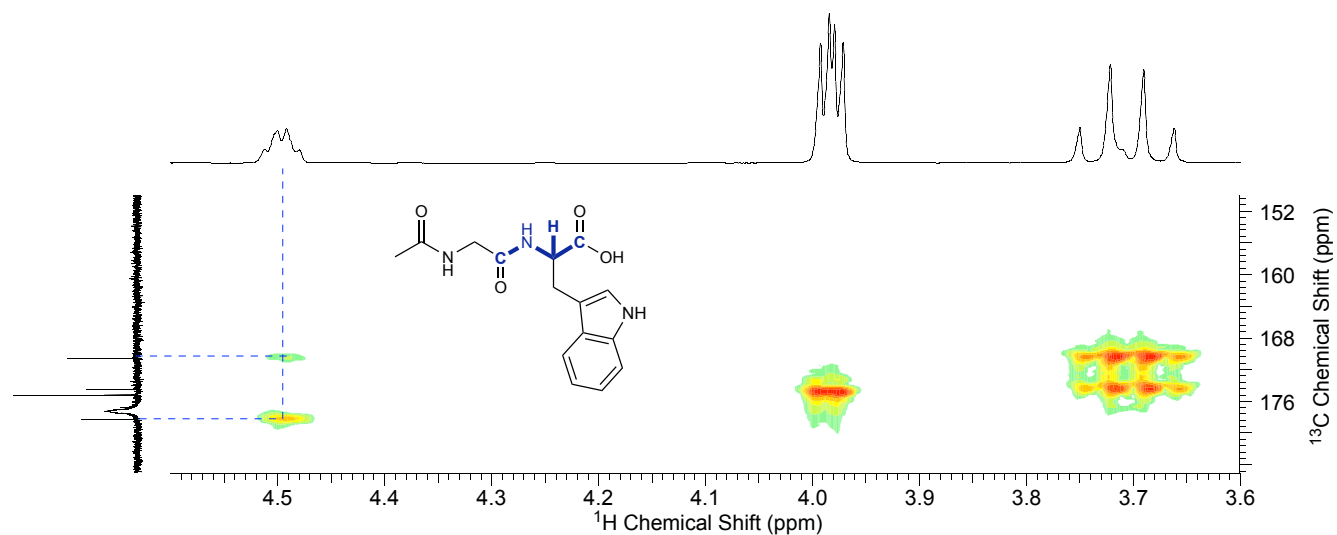
Supplementary Figure 187. 1H - ^{13}C HMBC (1H -700 MHz [3.8–5.2 ppm], ^{13}C -176 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{CH}$ and $^3J_{CH}$ coupling of Thr-(C2)-H of **Ac-Gly-Thr-CN** at 5.00 ppm with CN resonance at 118.3 ppm and a C=O resonance at 172.3 ppm, which is characteristic of peptide bond formation of **Thr-CN**. See Supplementary Figure 186 for expanded and labelled 1H NMR spectrum.

N-Acetylglycyltryptophan **Ac-Gly-Trp-OH**



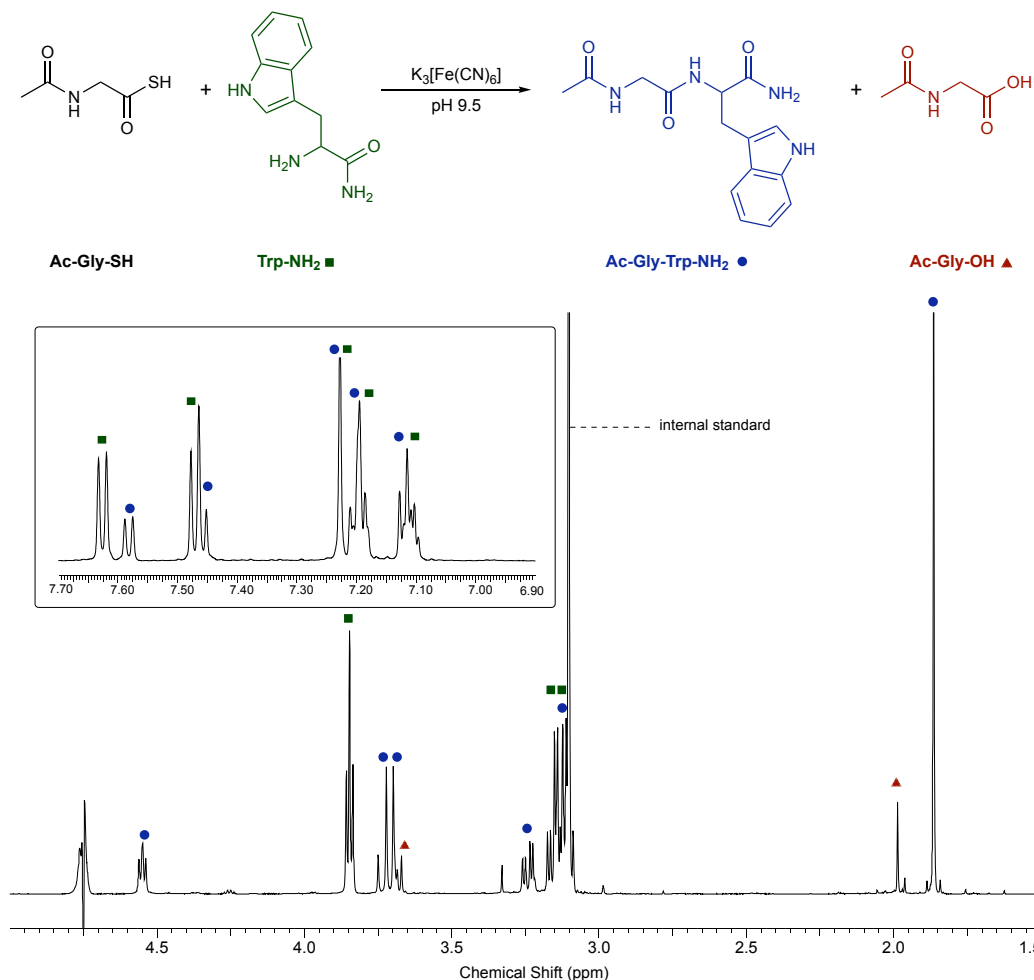
Supplementary Figure 188. 1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-tryptophan (**Trp**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature. Inset: 1H NMR (600 MHz, H_2O/D_2O 98:2, 7.0–7.8 ppm) showing the aromatic resonances for *N*-Acetylglycyltryptophan **Ac-Gly-Trp-OH** and residual **Trp**.

1H NMR (600 MHz, H_2O/D_2O 98:2, noesygppr1d) *N*-Acetylglycyltryptophan **Ac-Gly-Trp-OH** (●): δ 7.77 (d, J = 7.8 Hz, 1H, Ar), 7.62 (d, J = 8.3 Hz, 1H, Ar), 7.42 – 7.22 (m, 3H, Ar), 4.50 (dd, J = 5.0, 7.3 Hz, 1H, Trp-(C2)–H), 3.74 (AB, J = 17.1 Hz, 1H, Gly-(C2)–H), 3.68 (AB, J = 17.1 Hz, 1H, Gly-(C2')–H), 3.30 (ABX, J = 5.0, 14.9 Hz, 1H, Trp-(C3)–H), 3.14 (ABX, J = 7.3, 14.9 Hz, 1H, Trp-(C3')–H), 1.86 (s, 3H, (COCH₃)). *L*-Tryptophan **Trp** (■): δ 7.69 (d, J = 8.0 Hz, 1H, Ar), 7.50 (d, J = 8.3 Hz, 1H, Ar), 7.30 (s, 1H, Ar), 7.25 – 7.20 (m, 1H, Ar), 7.19 – 7.12 (m, 1H, Ar), 3.98 (dd, J = 4.9, 7.9 Hz, 1H, (C2)–H), 3.42 (ABX, J = 4.9, 15.3 Hz, 1H, (C3)–H), 3.25 (ABX, J = 7.9, 15.3 Hz, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.72 – 3.70 (obs., 2H, (C2)–H₂), 2.01 (s, 3H, (COCH₃)).



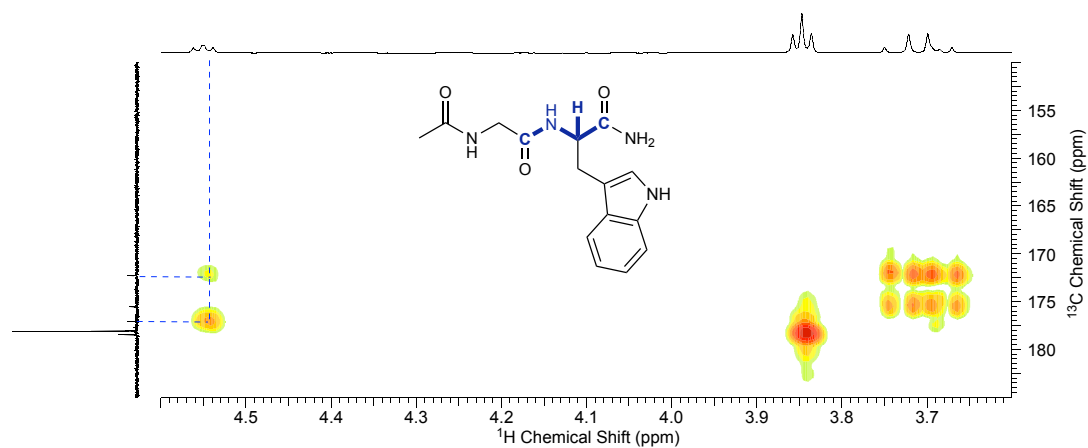
Supplementary Figure 189. ^1H - ^{13}C HMBC (^1H -600 MHz [3.6–4.6 ppm], ^{13}C -151 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Trp-(C2)-H of **Ac-Gly-Trp-OH** at 4.50 ppm with two C=O resonances at 171.3 and 179.2 ppm, which is characteristic of peptide bond formation of **Trp**. See Supplementary Figure 188 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycyltryptophanamide **Ac-Gly-Trp-NH₂**



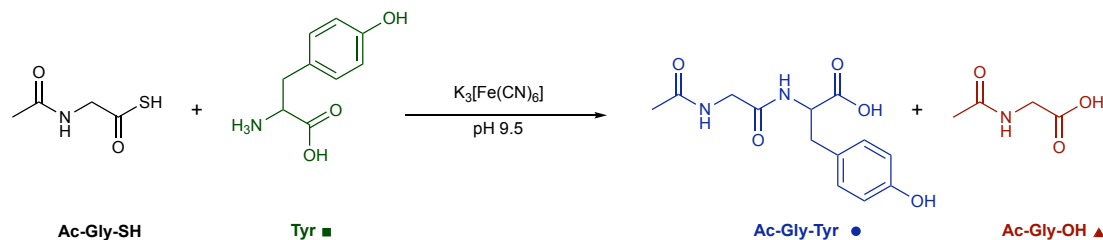
Supplementary Figure 190. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-tryptophanamide (**Trp-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature. Inset: ¹H NMR (600 MHz, H₂O/D₂O 98:2, 6.9–7.7 ppm) showing the aromatic resonances for *N*-acetylglycyltryptophanamide **Ac-Gly-Trp-NH₂** and residual **Trp-NH₂**.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycyltryptophanamide **Ac-Gly-Trp-NH₂** (●): δ 7.58 (d, *J* = 7.8 Hz, 1H, Ar), 7.46 (d, *J* = 8.0 Hz, 1H, Ar), 7.23 (s, 1H, Ar), 7.22 - 7.17 (m, 1H, Ar), 7.14 - 7.09 (m, 1H, Ar), 4.55 (dd, *J* = 6.1, 7.5 Hz, 1H, Trp-(C2)-H), 3.74 (AB, *J* = 17.1 Hz, 1H, Gly-(C2)-H), 3.68 (AB, *J* = 17.1 Hz, 1H, Gly-(C2')-H), 3.24 (ABX, *J* = 6.1, 14.9 Hz, 1H, Trp-(C3)-H), 3.14 - 3.10 (obs., 1H, Trp-(C3')-H), 1.86 (s, 3H, (COCH₃)). *L*-Tryptophanamide **Trp-NH₂** (■): δ 7.63 (d, *J* = 8.0 Hz, 1H, Ar), 7.47 (d, *J* = 8.0 Hz, 1H, Ar), 7.23 (s, 1H, Ar), 7.22 - 7.17 (m, 1H, Ar), 7.14 - 7.08 (m, 1H, Ar), 3.85 (t, *J* = 6.5 Hz, 1H, (C2)-H), 3.16 (ABX, *J* = 6.5, 14.7 Hz, 1H, (C3)-H), 3.10 (ABX, *J* = 6.5, 14.7 Hz, 1H, (C3')-H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.69 - 3.68 (obs., 2H, (C2)-H₂), 1.99 (s, 3H, (COCH₃)).

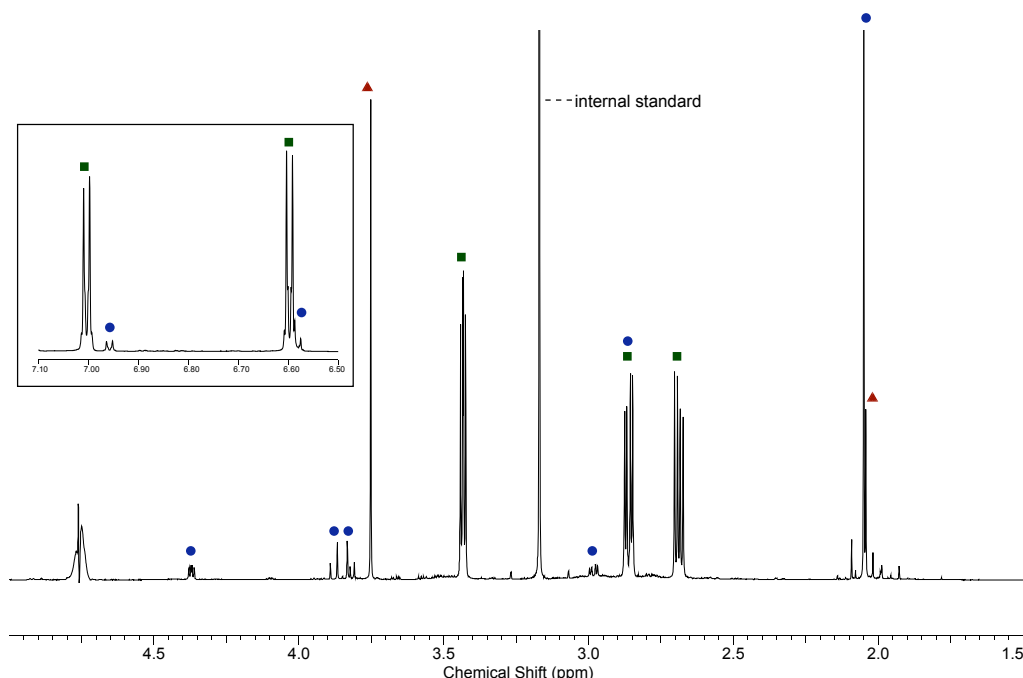


Supplementary Figure 191. ^1H - ^{13}C HMBC (^1H -600 MHz [3.6–4.6 ppm], ^{13}C -151 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Trp-(C2)-H of **Ac-Gly-Trp-NH₂** at 4.55 ppm with two C=O resonances at 172.3 and 177.1 ppm, which is characteristic of peptide bond formation of **Trp-NH₂**. See Supplementary Figure 190 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycyltyrosine **Ac-Gly-Tyr-OH**

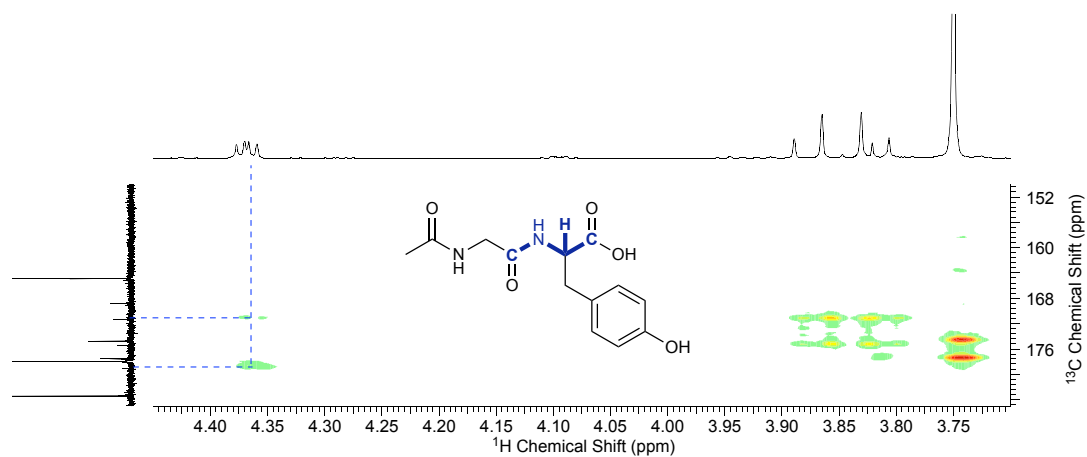


L-Tyrosine **Tyr** (27.2 mg, 0.15 mmol) was dissolved in H₂O/D₂O (98:2; 0.5 mL) and the solution was adjusted to pH 9.5 with 4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and total volume was adjusted to 1 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added and the solution was stirred at room temperature for 20 min. Due to the low solubility of L-tyrosine **Tyr**, the solution was then adjusted to pH 12 with 4 M NaOH (to promote product solubility)⁵¹, and then centrifuged. The supernatant was analysed by 1D and 2D NMR spectroscopy and the reaction mixture was quantified using methylsulfonylmethane (MSM) as an internal standard. The presence of **Ac-Gly-Tyr-OH** (●) was confirmed by ¹H–¹³C HMBC NMR spectral analysis and high resolution mass spectrometry.



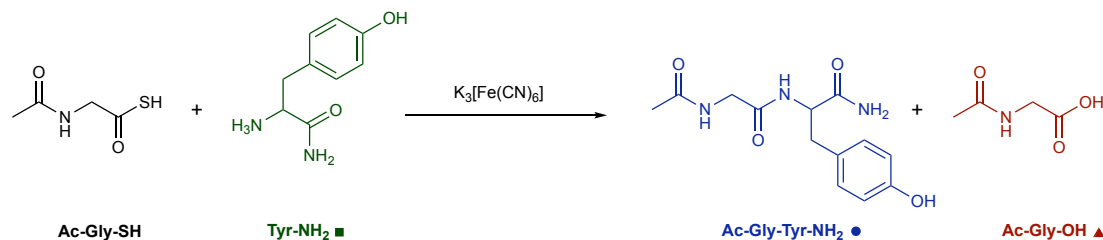
Supplementary Figure 192. ¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-tyrosine (**Tyr**; 150 mM) and K₃[Fe(CN)₆] (150 mM) at pH 9.5 and room temperature. NMR analysis carried out at pH 12 to completely solubilize **Tyr**. Inset: ¹H NMR (700 MHz, H₂O/D₂O 98:2, 6.50–7.10 ppm) showing the aromatic resonances for *N*-acetylglycyltyrosine **Ac-Gly-Tyr-OH** and residual **Tyr**.

¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycyltyrosine **Ac-Gly-Tyr-OH** (●): δ 6.96 (d, *J* = 8.5 Hz, 2H, Ar-(C2)–H), 6.58 (d, *J* = 8.5 Hz, 2H, Ar-(C3)–H), 4.37 (dd, *J* = 5.0, 7.5 Hz, 1H, Tyr-(C2)–H), 3.88 (AB, *J* = 17.1 Hz, 1H, Gly-(C2)–H), 3.82 (AB, *J* = 17.1 Hz, 1H, Gly-(C2')–H), 2.98 (ABX, *J* = 5.0, 14.0 Hz, 1H, Tyr-(C3)–H), 2.88 - 2.85 (obs., 1H, Tyr-(C3')–H), 2.04 (s, 3H, (COCH₃)). *L*-Tyrosine **Tyr** (■): 7.00 (d, *J* = 8.5 Hz, 2H, Ar-(C2)–H), 6.60 (d, *J* = 8.5 Hz, 2H, Ar-(C3)–H), 3.43 (dd, *J* = 5.3, 7.3 Hz, 1H, (C2)–H), 2.86 (ABX, *J* = 5.3, 13.8 Hz, 1H, (C3)–H), 2.69 (ABX, *J* = 7.3, 13.8 Hz, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲) δ 3.75 (s, 3H, (C2)–H₂), 2.05 (s, 3H, (COCH₃)).

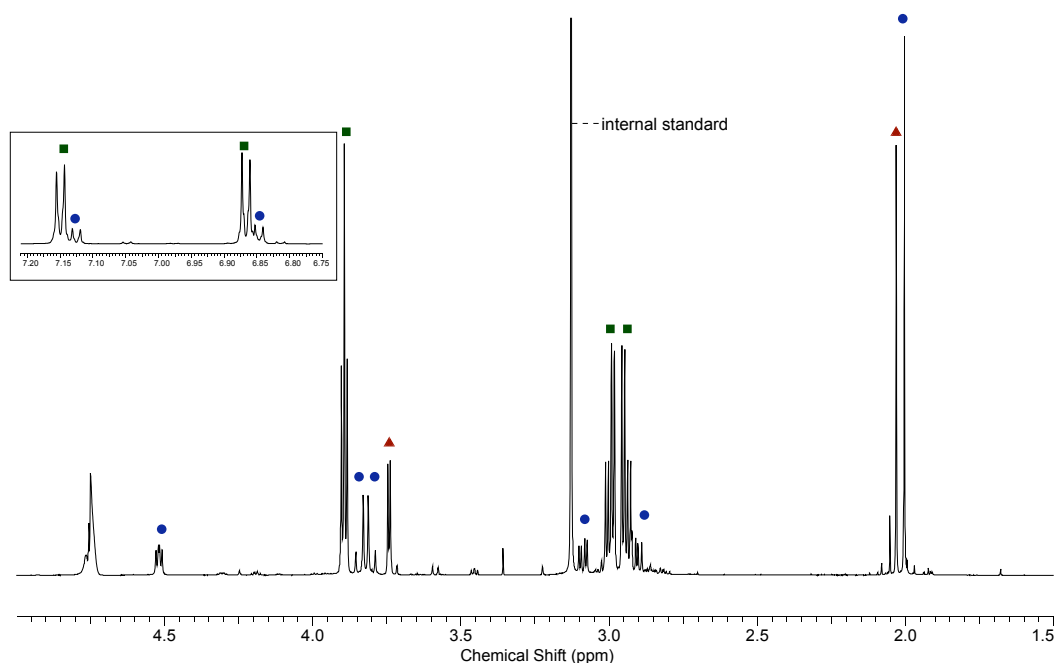


Supplementary Figure 193. ^1H - ^{13}C HMBC (^1H -600 MHz [3.70–4.45 ppm], ^{13}C -151 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Tyr-(C2)-H of **Ac-Gly-Tyr-OH** centred at 4.37 ppm with two C=O resonances at 171.3 and 178.5 ppm, which is characteristic of peptide bond formation of **Tyr**. See Supplementary Figure 192 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycyltyrosinamide **Ac-Gly-Tyr-NH₂**

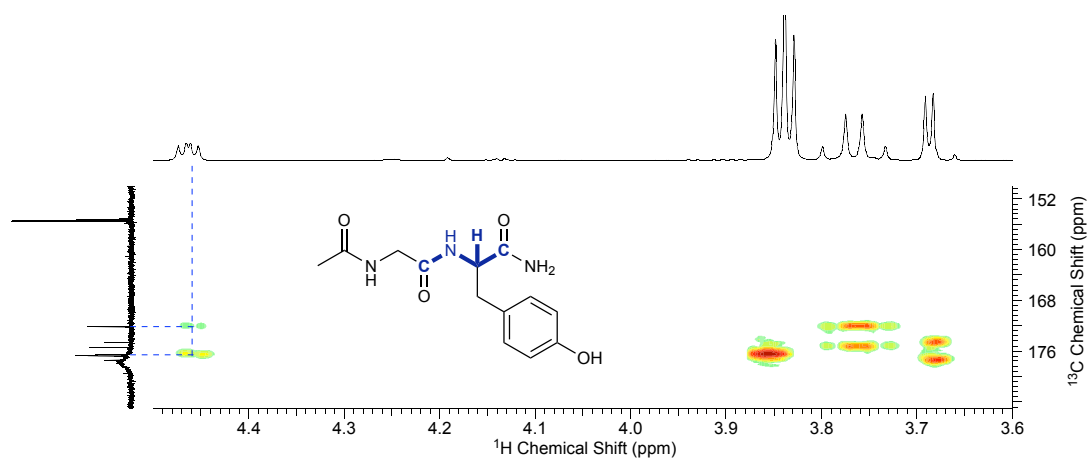


L-Tyrosinamide hydrochloride **Tyr-NH₂·HCl** (32.5 mg, 0.15 mmol) was dissolved in H₂O/D₂O (98:2; 0.5 mL) and the solution was adjusted to pH 9.5, 8.5, 7.5 or 6.5 as specified with 4 M HCl/NaOH. 2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (0.05 mmol) was added and total volume was adjusted to 1 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (49.4 mg, 0.15 mmol) was added and the solution was stirred at room temperature for 20 min whilst maintaining the starting pH with 4 M HCl/NaOH. The solution was then adjusted to pH 9.5 and incubated for 16 h at room temperature. The reaction mixture was centrifuged, and the supernatant was analysed by 1D and 2D NMR spectroscopy. The reaction mixture was quantified using methylsulfonylmethane (MSM) as an internal standard, and the presence of **Ac-Gly-Tyr-NH₂** (●) was confirmed by ¹H–¹³C HMBC NMR spectral analysis and high resolution mass spectrometry.



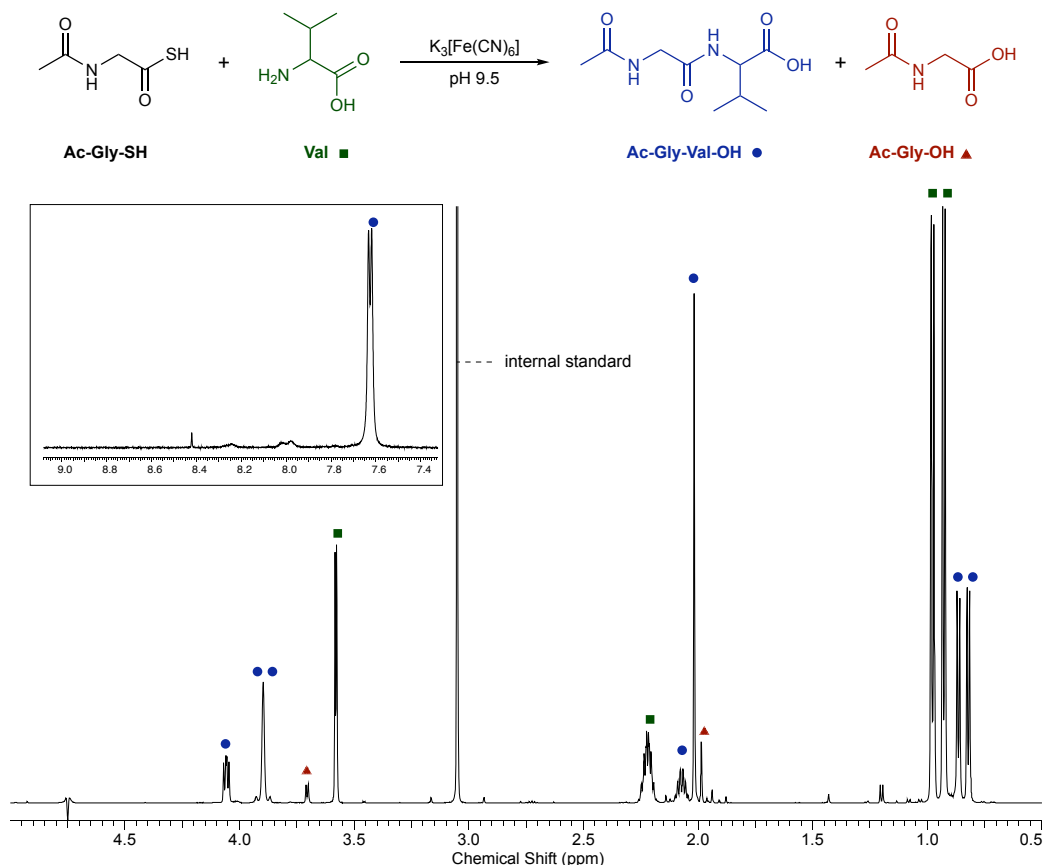
Supplementary Figure 194. ¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-tyrosinamide (**Tyr-NH₂**; 150 mM) and K₃[Fe(CN)₆] (150 mM) at pH 6.5 and room temperature. NMR analysis was carried out after incubating the reaction at pH 9.5 for 16 h. Inset: ¹H NMR (700 MHz, H₂O/D₂O 98:2, 6.75–7.20 ppm) showing the aromatic resonances for *N*-acetylglycyltyrosinamide **Ac-Gly-Tyr-NH₂** and residual L-tyrosinamide **Tyr-NH₂**.

¹H NMR (700 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycyltyrosinamide **Ac-Gly-Tyr-NH₂** (●): δ 7.13 (d, *J* = 8.5 Hz, 2H, Ar-(C2)–H), 6.85 (d *J* = 8.5 Hz, 2H, Ar-(C3)–H), 4.52 (dd, *J* = 5.8, 8.9 Hz, 1H, Tyr-(C2)–H), 3.84 (AB, *J* = 16.8 Hz, 1H, Gly-(C2)–H), 3.80 (AB, *J* = 16.8 Hz, 1H, Gly-(C2')–H), 3.09 (ABX, *J* = 5.8, 14.3 Hz, 1H, Tyr-(C3)–H), 2.91 (ABX, *J* = 8.9, 14.3 Hz, 1H, Tyr-(C3')–H), 2.00 (s, 3H, (COCH₃)). *L*-Tyrosinamide **Tyr-NH₂** (■): 7.15 (d, *J* = 8.5 Hz, 2H, Ar-(C2)–H), 6.86 (d, *J* = 8.5 Hz, 2H, Ar-(C3)–H), 3.89 (app. t, *J* = 6.8 Hz, 1H, (C2)–H), 3.00 (ABX, *J* = 6.8, 14.1 Hz, 1H, (C3)–H), 2.94 (dd, *J* = 6.8, 14.1 Hz, 1H, (C3')–H). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.74 (d, *J* = 5.6 Hz, 2H, (C2)–H₂), 2.03 (s, 3H, (COCH₃)).



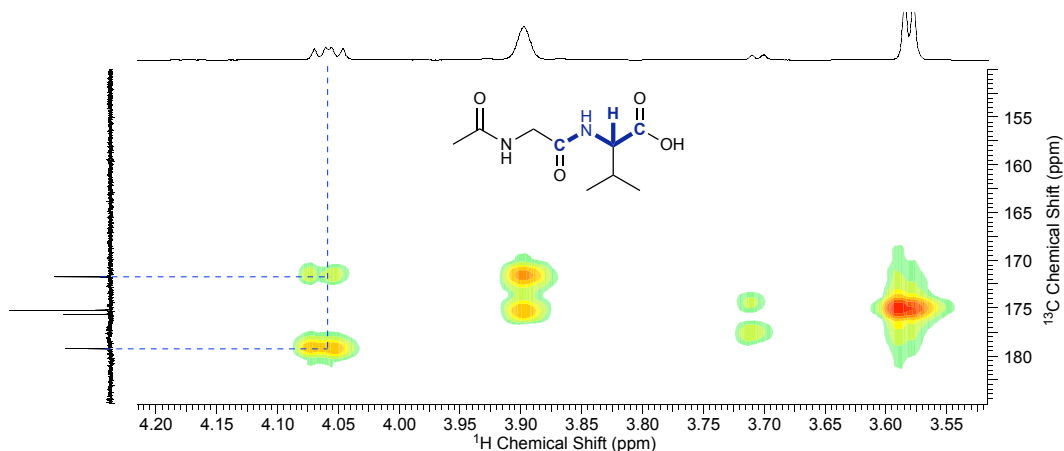
Supplementary Figure 195. ^1H - ^{13}C HMBC (^1H -700 MHz [3.60–4.50 ppm], ^{13}C -176 MHz [150–185 ppm], $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Tyr-(C2)-H of **Ac-Gly-Tyr-NH₂** at 4.52 ppm with two C=O resonances at 172.2 and 176.5 ppm, which is characteristic of peptide bond formation of **Tyr-NH₂**. See Supplementary Figure 194 for expanded and labelled ^1H NMR spectrum.

N-Acetylglycylvaline **Ac-Gly-Val-OH**



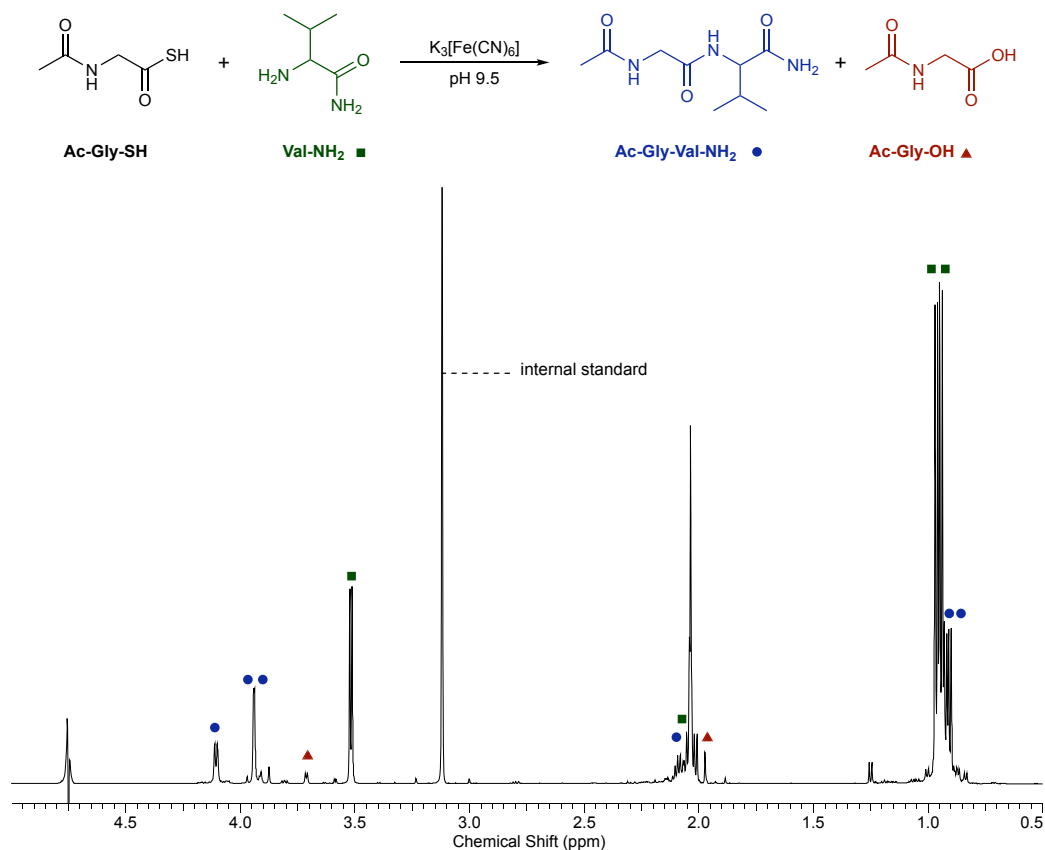
Supplementary Figure 196. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-valine (**Val**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature. Inset: ¹H NMR (600 MHz, 7.40–9.10 ppm) showing the remnants of a Val-(C2)-NH resonance, which is undergoing deuteration.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylvaline **Ac-Gly-Val-OH** (●): δ 7.63 (d, J = 8.4 Hz, 1H, Val-(C2)-NH), 4.06 (dd, J = 5.7, 8.4 Hz, 1H, Val-(C2)-H), 3.91 (AB, J = 18.2 Hz, 1H, Gly-(C2)-H), 3.87 (AB, J = 18.2 Hz, 1H, Gly-(C2')-H), 2.11–2.04 (m, 1H, Val-(C3)-H), 2.02 (s, 3H, (COCH₃)), 0.86 (d, J = 6.9 Hz, 3H, Val-(C4)-H₃), 0.82 (d, J = 6.9 Hz, 3H, Val-(C4')-H₃). *L*-Valine **Val** (■): δ 3.58 (d, J = 4.3 Hz, 1H, (C2)-H), 2.26–2.16 (m, 1H, (C3)-H), 0.98 (d, J = 7.0 Hz, 3H, (C4)-H₃), 0.93 (d, J = 7.0 Hz, 3H, (C4')-H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.70 (app. d, J = 5.8 Hz, 2H, (C2)-H₂), 1.99 (s, 3H, (COCH₃)).



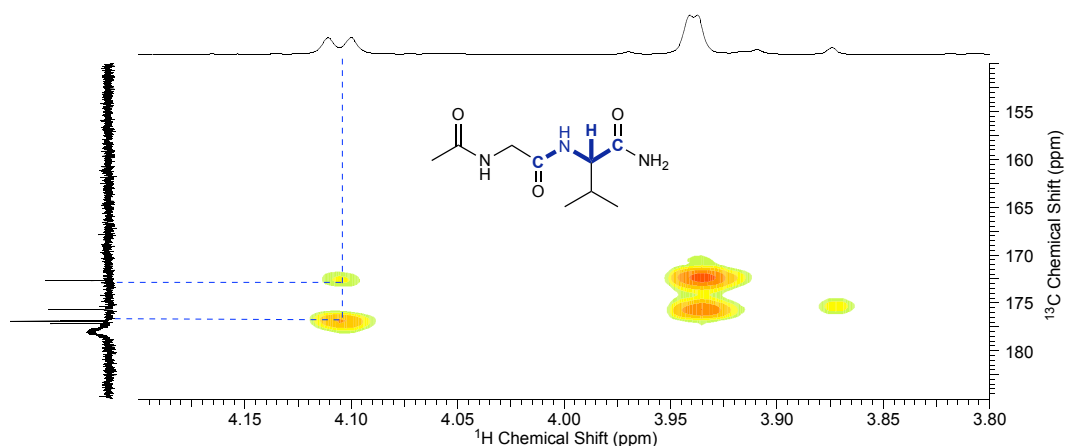
Supplementary Figure 197. ¹H-¹³C HMBC (¹H-600 MHz [3.5–4.2 ppm], ¹³C-151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Val-(C2)-H of **Ac-Gly-Val-OH** at 4.06 ppm with two C=O resonances at 171.7 and 179.2 ppm, which is characteristic of peptide bond formation of **Val**. See Supplementary Figure 196 for expanded and labelled ¹H NMR spectrum.

N-Acetylglycylvalinamide **Ac-Gly-Val-NH₂**



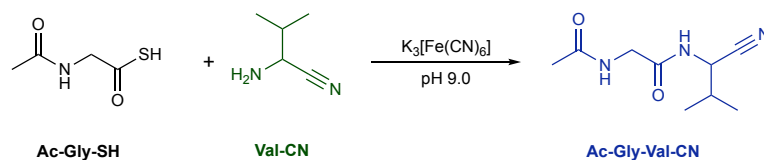
Supplementary Figure 198. ¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d, 1.5–5.0 ppm) spectrum to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM) with L-valinamide (**Val-NH₂**; 150 mM) and $K_3[Fe(CN)_6]$ (150 mM) at pH 9.5 and room temperature.

¹H NMR (600 MHz, H₂O/D₂O 98:2, noesygppr1d) *N*-Acetylglycylvalinamide **Ac-Gly-Val-NH₂** (●): δ 4.11 (d, $J = 6.5$ Hz, 1H, Val-(C2)-H), 3.96 (AB, $J = 17.0$ Hz, 1H, Gly-(C2)-H), 3.92 (AB, $J = 17.0$ Hz, 1H, Gly-(C2')-H), 2.12 - 2.06 (m, 1H, Val-(C3)-H), 2.04 (s, 3H, (COCH₃)), 0.92 (d, $J = 6.9$ Hz, 3H, Val-(C4)-H₃), 0.90 (d, $J = 6.9$ Hz, 3H, Val-(C4')-H₃). *L*-Valinamide **Val-NH₂** (■): δ 3.52 (d, $J = 5.9$ Hz, 1H, (C2)-H), 2.08 - 1.98 (m, 1H, (C3)-H), 0.96 (d, $J = 6.9$ Hz, 3H, (C4)-H₃), 0.94 (d, $J = 6.9$ Hz, 3H, (C4')-H₃). *N*-Acetylglycine **Ac-Gly-OH** (▲): δ 3.71 (app. d, $J = 5.4$ Hz, 2H, (C2)-H₂), 1.97 (s, 3H, (COCH₃)).

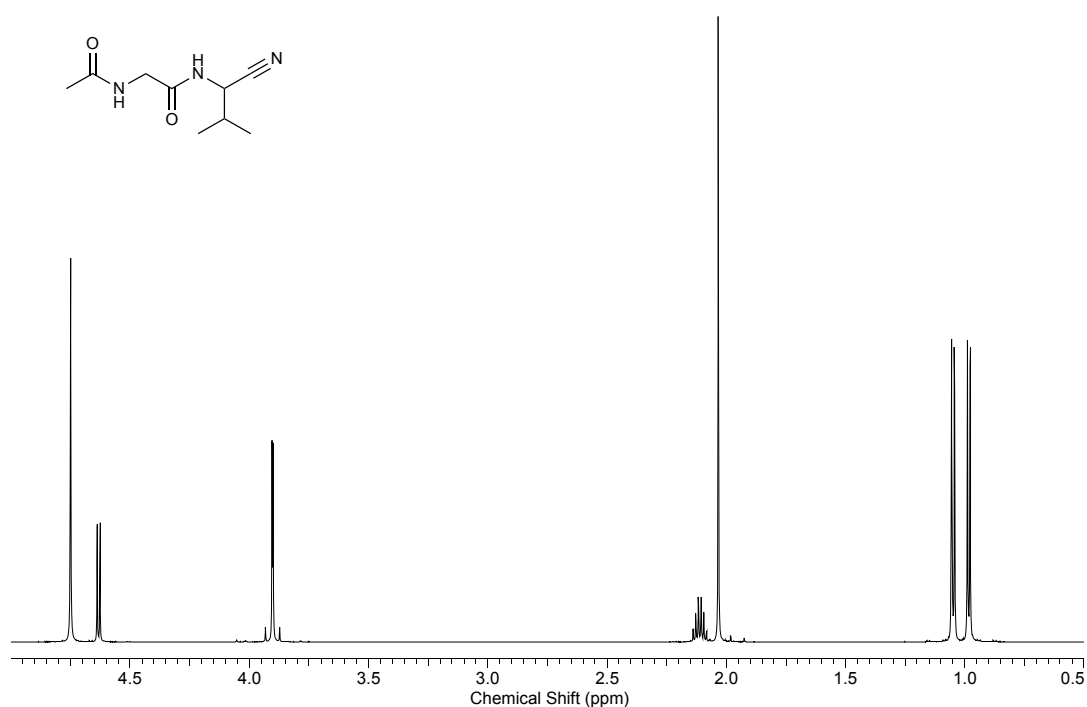


Supplementary Figure 199. ¹H-¹³C HMBC (¹H-600 MHz [3.8–4.2 ppm], ¹³C-151 MHz [150–185 ppm], H₂O/D₂O 98:2) spectrum showing the diagnostic ² J_{CH} and ³ J_{CH} coupling of Val-(C2)-H of **Ac-Gly-Val-NH₂** at 4.11 ppm with two C=O resonances at 172.7 and 176.9 ppm, which is characteristic of peptide bond formation of **Val-NH₂**. For Supplementary Figure 198 for expanded and labelled ¹H NMR spectrum.

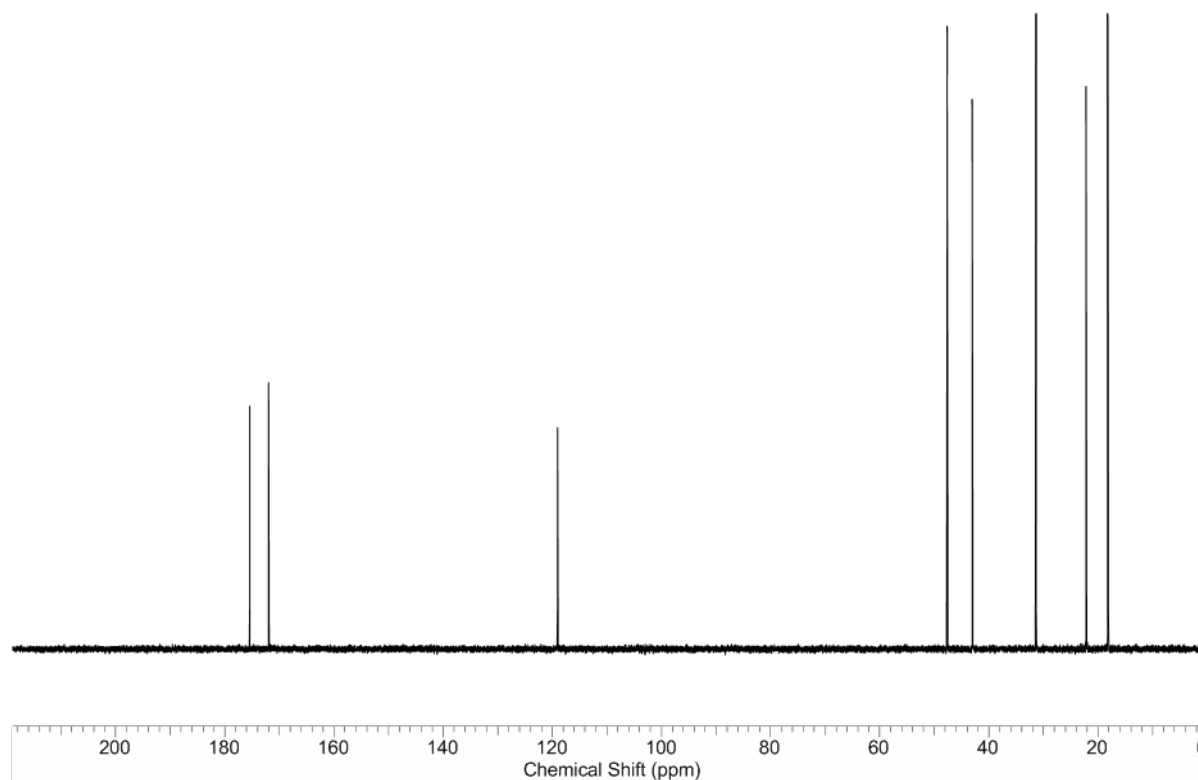
N-Acetylglycylvaline nitrile **Ac-Gly-Val-CN**



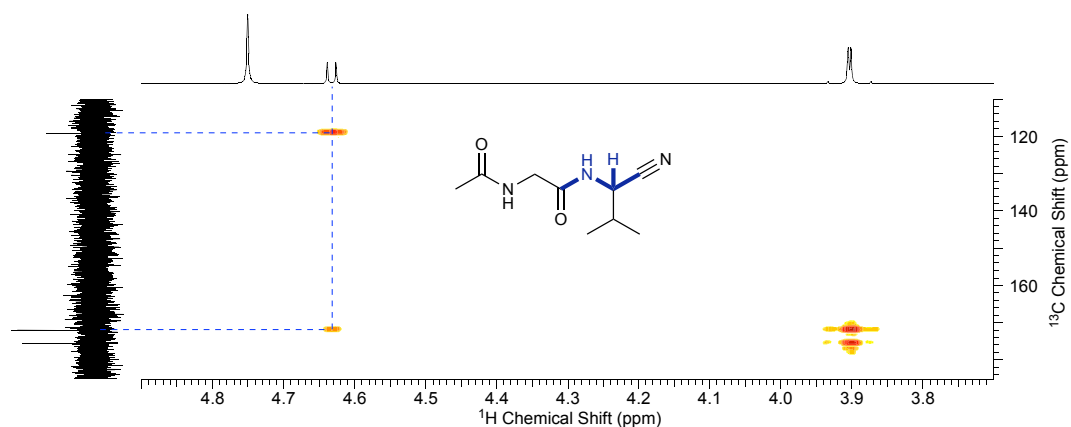
Prepared following *general procedure G* on a 0.1 mmol scale, from sodium-2-acetamido-ethanethioate **Ac-Gly-S⁻Na⁺** (15.5 mg, 0.1 mmol) and *rac*-valine nitrile hydrochloride **Val-CN·HCl**³⁹ (26.9 mg, 0.2 mmol). **Ac-Gly-Val-CN** was isolated as a white solid (18.1 mg, 92%) after flash column chromatography (SiO₂, EtOAc (100%)). *R_f* 0.60 (100% EtOAc). ¹H NMR (600 MHz, D₂O) δ 4.63 (d, *J* = 7.4 Hz, 1H, Val-(C2)-H), 3.92 (AB, *J* = 17.0 Hz, 1H, Gly-(C2)-H), 3.89 (d, *J* = 17.0 Hz, 1H, Gly-(C2)-H'), 2.16 - 2.07 (m, 1H, Val-(C3)-H), 2.03 (s, 3H, (COCH₃)), 1.05 (d, *J* = 6.8 Hz, 3H, Val-(C4)-H₃), 0.98 (d, *J* = 6.8 Hz, 3H, Val-(C4)-H₃'). ¹³C NMR (151 MHz, D₂O) δ 175.5 (COCH₃), 171.9 (Gly-(C1)), 119.0 (Val-(C1)), 47.6 (Val-(C2)), 43.0 (Gly-(C2)), 31.3 (Val-(C3)), 22.2 (COCH₃), 18.2 (Val-(C4)), 18.2 (Val-(C4')). HRMS-ESI-MS [*M*+H]⁺ calcd. for C₉H₁₆N₃O₂ 198.1237; obs. 198.1237. IR (solid, cm⁻¹) 3323, 3199, 2972, 2466, 2358, 2337, 1672, 1629, 1560.



Supplementary Figure 200. ¹H NMR (600 MHz, D₂O, 0.5–5.0 ppm) spectrum of purified *N*-acetylglycylvaline nitrile **Ac-Gly-Val-CN**

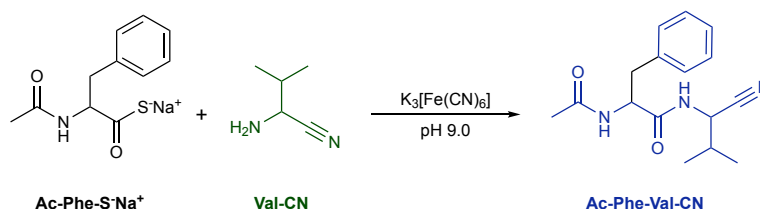


Supplementary Figure 201. ^{13}C NMR (176 MHz, D_2O , 0-220 ppm) spectrum of purified *N*-acetylglycylvaline nitrile **Ac-Gly-Val-CN**

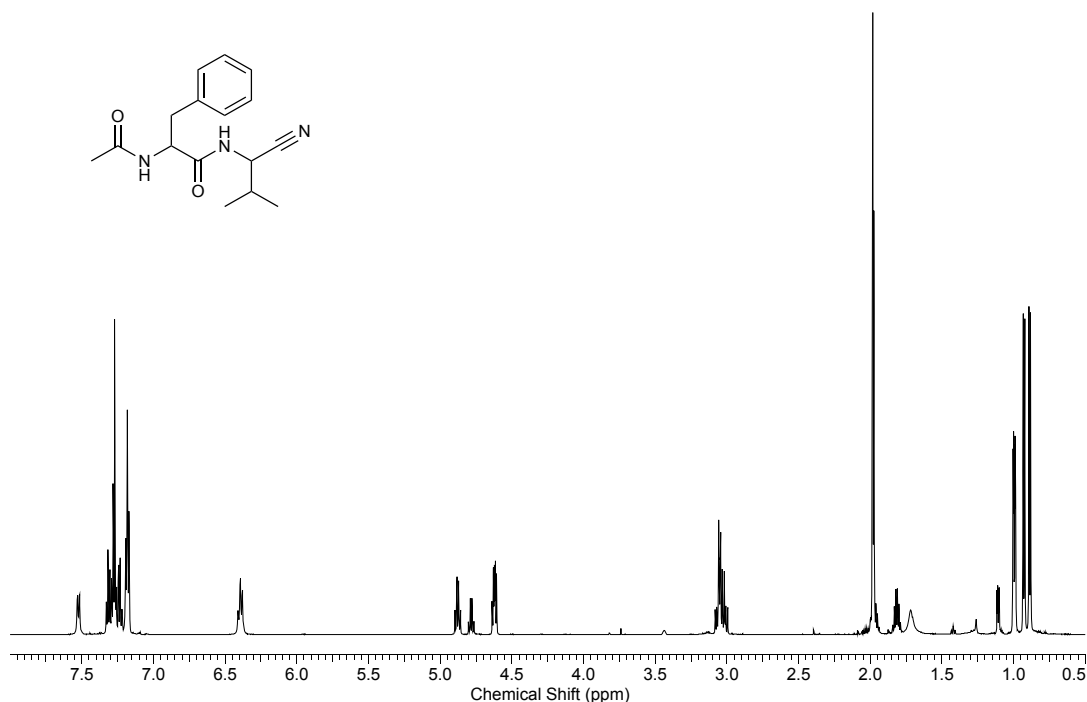


Supplementary Figure 202. ^1H - ^{13}C HMBC (^1H -600 MHz [3.7–5.0 ppm], ^{13}C -151 MHz [110–185 ppm], D_2O) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Val-(C2)-H of **Ac-Gly-Val-CN** at 4.63 ppm with a CN resonances at 119.0 ppm and a C=O resonances at 171.9 ppm, which is characteristic of peptide bond formation of **Val-CN**. See Supplementary Figure 160 for expanded ^1H NMR spectrum.

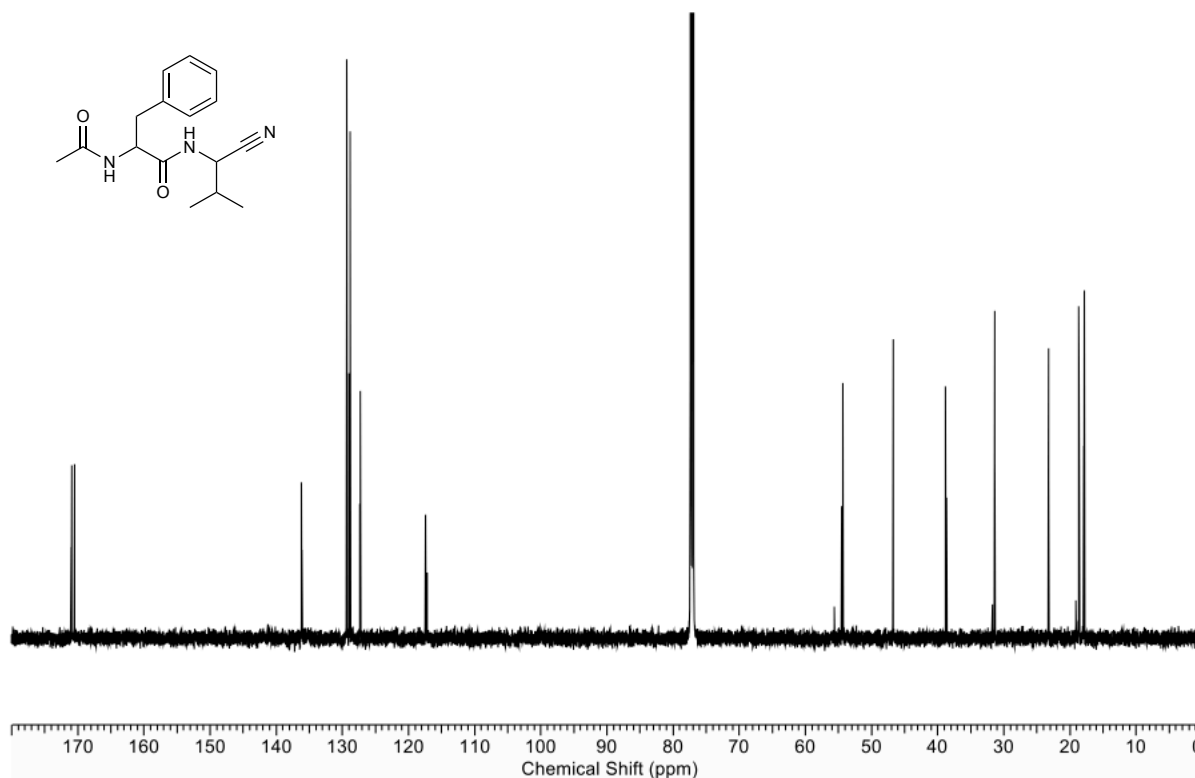
N-Acetylphenylalanylvaline nitrile **Ac-Phe-Val-CN**



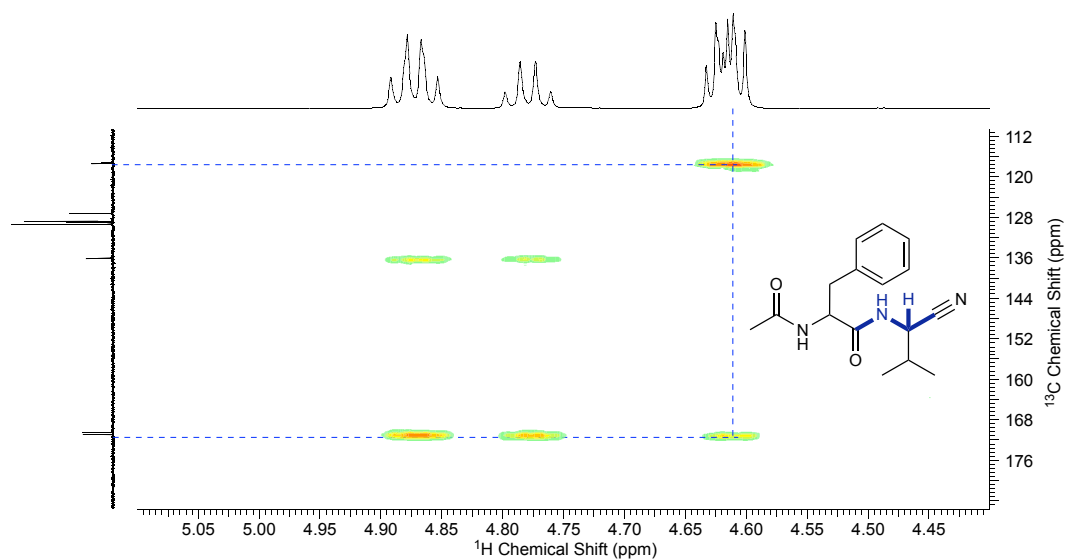
Prepared following *general procedure G* from sodium-2-acetamido-3-phenylpropanethioate **Ac-Phe-S-Na⁺** (123 mg, 0.5 mmol) and *rac*-valine nitrile hydrochloride **Val-CN**·HCl³⁹ (135 mg, 1.0 mmol). **Ac-Phe-Val-CN** was isolated as a white solid (105 mg, 73%) after flash column chromatography (CH₂Cl₂/MeOH, 100:0 to 95:5). *R_f* 0.31 (CH₂Cl₂/MeOH, 95:5). ¹H NMR (600 MHz, CDCl₃, 65:35 mixture of diastereoisomers a/b; partial assignment of diastereoisomers). δ 7.52 (d, *J* = 8.4 Hz, 1H, Val^a-NH), 7.34 - 7.25 (m, 6H, Ar), 7.24 (obs., 1H, Val^b-NH), 7.18 (m, 4H, Ar), 6.41 (d, *J* = 7.4 Hz, 1H, Phe^b-NH), 6.35 (d, *J* = 8.3, 1H, Phe^a-NH), 4.88 (dt, *J* = 6.7, 8.3 Hz, 1H, Phe^a-(C2)-H), 4.78 (app. q, *J* = 7.4 Hz, 1H, Phe^b-(C2)-H), 4.64 - 4.59 (m, 2H, Val^a-(C2)-H, Val^b-(C2)-H), 3.06 (ABX, *J* = 6.7, 13.6 Hz, 1H, Phe^a-(C3)-H), 3.05 (d, *J* = 7.4 Hz, 2H, Phe^b-(C3)-H₂), 3.01 (ABX, *J* = 8.3, 13.6 Hz, 1H, Phe^a-(C3')-H), 1.98 (s, 3H, COCH₃), 1.98 (s, 3H, COCH₃), 2.01 - 1.93 (m, 1H, Val^b-(C3)-H), 1.85 - 1.78 (m, 1H, Val^a-(C3)-H), 1.00 (d, *J* = 6.8 Hz, 3H, Val^b-(C4)-H₃), 0.99 (d, *J* = 6.7 Hz, 3H, Val^b-(C4')-H₃), 0.93 (d, *J* = 6.8 Hz, 3H, Val^a-(C4)-H₃), 0.89 (d, *J* = 6.7 Hz, 3H, Val^a-(C4')-H₃). ¹³C NMR (151 MHz, CDCl₃, 65:35 mixture of diastereoisomers a/b; partial assignment of diastereoisomers) δ 171.0 (COCH₃), 170.9 (COCH₃), 170.5 (Phe-(C1)^a), 170.5 (Phe^b-(C1)), 136.2 (Ar), 136.1 (Ar), 129.4 (Ar x 2), 129.3 (Ar x 2), 129.0 (Ar x 2), 128.8 (Ar x 2), 127.4 (Ar), 127.3 (Ar), 117.4 (Val^a-(C1)), 117.2 (Val^b-(C1)), 54.5 (Phe^b-(C2)), 54.3 (Phe^a-(C2)), 46.75 (Val^b-(C2)), 46.73 (Val^a-(C2)), 38.8 (Phe^a-(C3)), 38.6 (Phe^b-(C3)), 31.4 (Val-(C3)), 31.4 (Val-(C3)), 23.25 (COCH₃^a), 23.20 (COCH₃^b), 18.7 (Val-(C4)), 18.6 (Val-(C4)), 17.9 (Val-(C4)), 17.8 (Val-(C4)). HRMS-ESI [M+H]⁺ calcd. for C₁₆H₂₂N₃O₂ 288.1712; obs. 288.1719. IR (solid, cm⁻¹) 3275, 3064, 2966, 2933, 2247, 1689, 1647, 1637, 1533.



Supplementary Figure 203. ¹H NMR (600 MHz, CDCl₃, 1.0–8.0 ppm) spectrum of purified *N*-acetylphenylalanylvaline nitrile **Ac-Phe-Val-CN**.

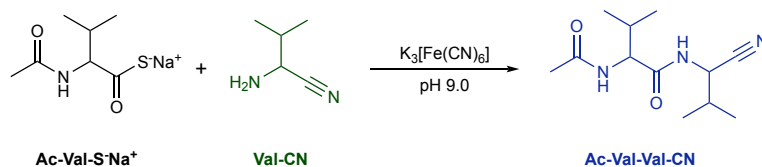


Supplementary Figure 204. ^{13}C NMR (151 MHz, CDCl_3 , 0–180 ppm) spectrum of purified *N*-acetylphenylalanylvaline nitrile **Ac-Phe-Val-CN**.

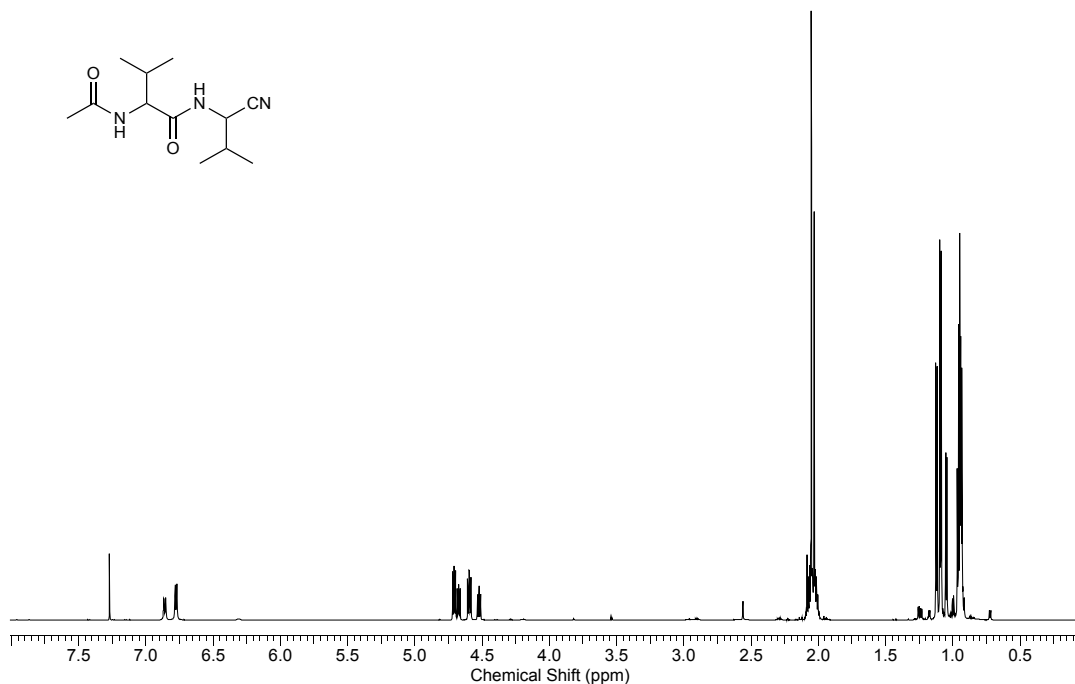


Supplementary Figure 205. ^1H – ^{13}C HMBC (^1H –600 MHz [4.40–5.10 ppm], ^{13}C –151 MHz [110–185 ppm], CDCl_3) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Val-(C2)–H of **Ac-Phe-Val-CN** at 4.59–4.64 ppm with CN resonances at 117.2 and 117.4 ppm and C=O resonances at 170.5 ppm, which is characteristic of peptide bond formation of **Val-CN**. See Supplementary Figure 203 for expanded ^1H NMR spectrum.

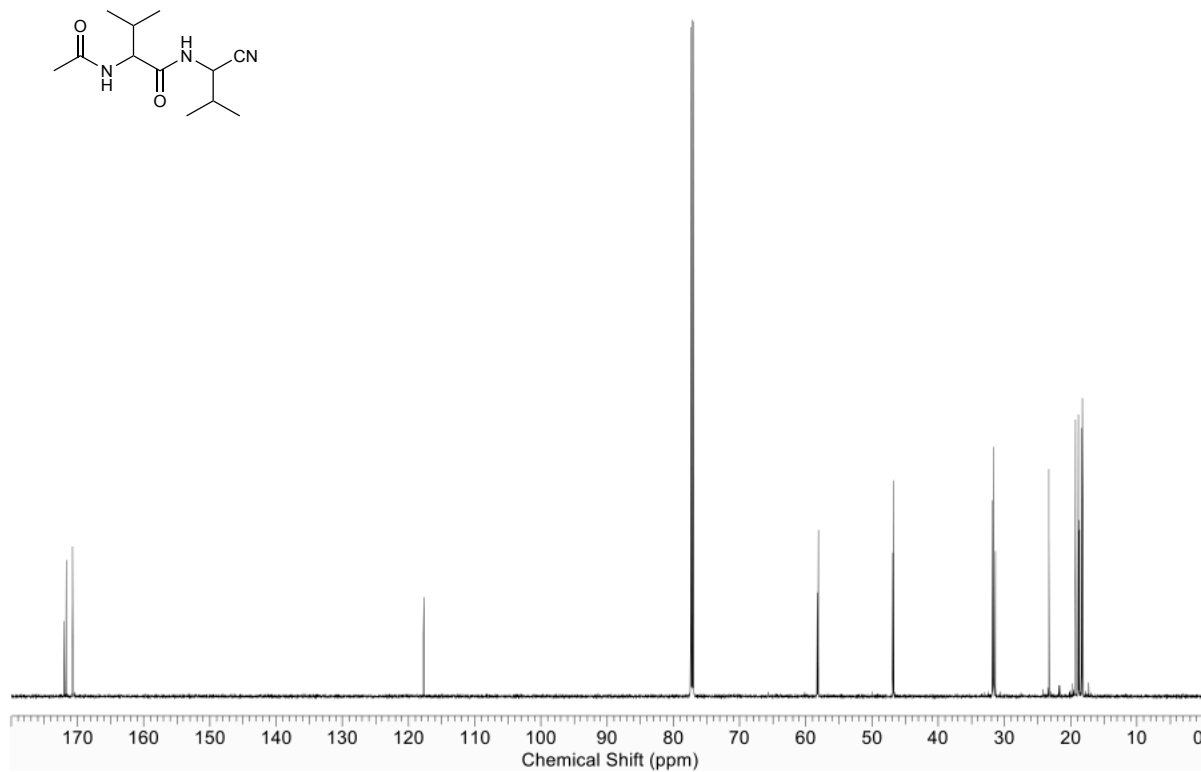
N-Acetylvalylvaline nitrile **Ac-Val-Val-CN**



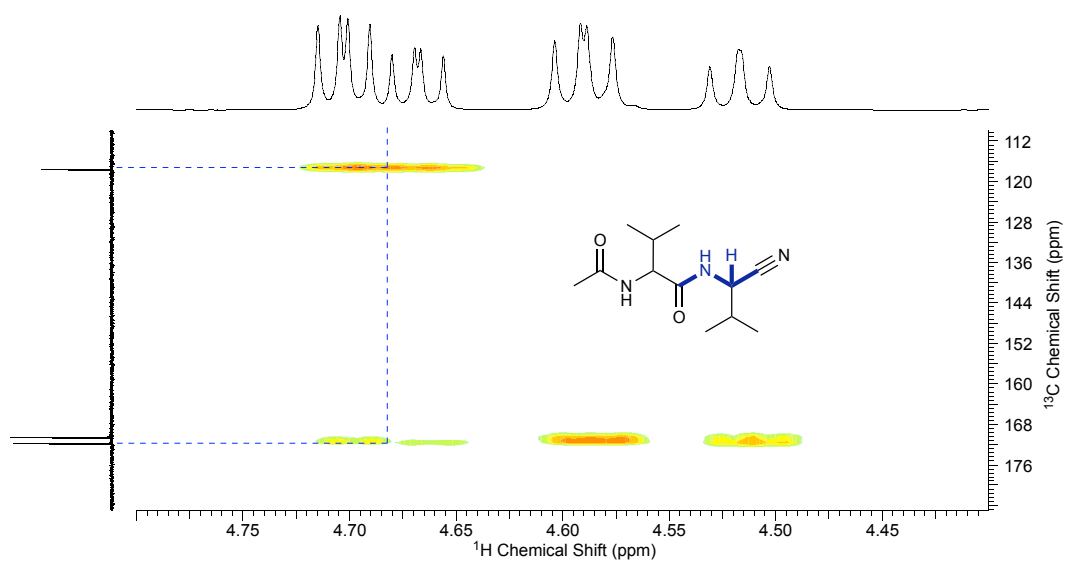
Prepared following *general procedure G* from sodium-2-acetamido-3-methylbutanethioate **Ac-Val-S-Na⁺** (99 mg, 0.5 mmol) and *rac*-valine nitrile hydrochloride **Val-CN·HCl**³⁹ (135 mg, 1.0 mmol). **Ac-Val-Val-CN** was isolated as a white solid (109 mg, 91%) after flash column chromatography (pet. ether (40-60)/Et₂O, 1:1). *R_f* 0.25 (pet. ether (40-60)/Et₂O, 1:1). ¹H NMR (700 MHz, CDCl₃, 61:39 mixture of diastereoisomers a/b) δ 8.43 (d, *J* = 8.4 Hz, 1H, Val₂^a-NH), 8.24 (d, *J* = 8.2 Hz, 1H, Val₂^b-NH), 6.86 (d, *J* = 9.0 Hz, 1H, Val₁^b-NH), 6.78 (d, *J* = 9.1 Hz, 1H, Val₁^a-NH), 4.71 (dd, *J* = 6.2, 8.4 Hz, 1H, Val₂^a-(C2)-H), 4.67 (dd, *J* = 6.4, 8.2 Hz, 1H, Val₂^b-(C2)-H), 4.60 (dd, *J* = 7.3, 9.1 Hz, 1H, Val₁^a-(C2)-H), 4.52 (dd, *J* = 7.9, 9.0 Hz, 1H, Val₁^b-(C2)-H), 2.05 (s, 3H, COCH₃^a), 2.03 (s, 3H, COCH₃^b), 2.09 - 1.99 (m, 4H, Val₁^a-(C3)-H, Val₁^b-(C3)-H, Val₂^a-(C3)-H, Val₂^b-(C3)-H), 1.12 - 0.92 (m, 24H, Val₁^a-(C4)-H₆, Val₁^b-(C4)-H₆, Val₂^a-(C4)-H₆, Val₂^b-(C4)-H₆). ¹³C NMR (176 MHz, CDCl₃, 61:39 mixture of diastereoisomers a/b) δ 172.0 (Val₁^b-(C1)), 171.7 (Val₁^a-(C1)), 170.8 (Val^b-(CO)), 170.7 (Val^a-(CO)), 117.8 (Val₂^b-(C1)), 117.7 (Val₂^a-(C1)), 58.3 (Val₂^b-(C2)), 58.1 (Val₂^a-(C2)), 46.9 (Val₁^b-(C2)), 46.8 (Val₁^a-(C2)), 31.8 (Val-(C3)), 31.70 (Val-(C3)), 31.64 (Val-(C3)), 31.4 (Val-(C3)), 23.3 (COCH₃^a), 23.2 (COCH₃^b), 19.3 (Val-(C4)), 19.2 (Val-(C4)), 18.9 (Val-(C4)), 18.8 (Val-(C4)), 18.7 (Val-(C4)), 18.4 (Val-(C4)), 18.2 (Val-(C4)), 18.2 (Val-(C4)). HRMS-ESI [M+H]⁺ calcd. for C₁₂H₂₂N₃O₂ 240.1707; obs. 240.1706. IR (solid, cm⁻¹) 3291, 3248, 3036, 2963, 2930, 2873, 2349, 1682, 1636, 1546.



Supplementary Figure 206. ¹H NMR (700 MHz, CDCl₃, 1.0–8.0 ppm) spectrum of purified *N*-acetylvalylvaline nitrile **Ac-Val-Val-CN**



Supplementary Figure 207. ^{13}C NMR (176 MHz, CDCl_3 , 0–180 ppm) spectrum of purified *N*-acetylvalylvaline nitrile **Ac-Val-Val-CN**.



Supplementary Figure 208. ^1H – ^{13}C HMBC (^1H –700 MHz [4.40–4.80 ppm], ^{13}C –151 MHz [110–185 ppm], CDCl_3) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Val₂-(C2)–H of **Ac-Val-Val-CN** at 4.67–4.71 ppm with CN resonances at 117.7 and 117.8 ppm and C=O resonances at 171.7 and 172.0 ppm, which is characteristic of peptide bond formation of **Val-CN**. See Supplementary Figure 206 for expanded ^1H NMR spectrum.

Iterative elongation of *N*-acetylglycine nitrile

The homologation of *N*-acetylglycine nitrile **Ac-Gly-CN** was performed by iteration of the following protocol: Hydrogen sulfide (**H₂S caution - see General and Safety on page S3**) was bubbled for 2 h through a stirred suspension of **Ac-(Gly)_n-CN** in H₂O/D₂O (98:2, 100 mM) and Dowex® (50w×8 SH⁻ form,^{vi} 0.1 mg/mmol **Ac-(Gly)_n-CN**) at pH 9.0. The reaction was then stirred at room temperature for 6 h, filtered, and adjusted to pH 9.0 with KOH/HCl. Quantitative conversion of **Ac-(Gly)_n-CN** to **Ac-(Gly)_n-SNH₂** was observed by NMR spectroscopy. The solution was then heated at 60 °C for 24 h and near-quantitative conversion of **Ac-(Gly)_n-SNH₂** to **Ac-(Gly)_n-SH** was observed. Residual H₂S was removed by sparging with argon for 30 min at pH 5.0. Glycine nitrile hydrochloride (**Gly-CN·HCl**; 1.2 equiv.) was added and the solution adjusted to pH 9.0 with KOH. Potassium hexacyanoferrate(III) (3.0 equiv.) was added and the solution was stirred at room temperature for 30 mins. The reaction mixture was then centrifuged and the supernatant analysed by NMR spectroscopy. ¹H-¹³C HMBC NMR spectra were acquired after each iteration. A characteristic high frequency ¹H NMR resonance for the C-terminal Gly-(C2)-H₂ was observed upon coupling of glycine nitrile **Gly-CN** with **Ac-(Gly)_n-SH** to produce the ligation product **Ac-(Gly)_{n+1}-CN**. The Gly-(C2)-H₂ ¹H NMR resonance of **Gly-CN** (δ_H = 3.70 ppm, at pH 9.0) has a discernibly lower frequency chemical shift compared to Gly-(C2)-H₂ ¹H NMR resonance of the ligation product **Ac-(Gly)_{n+1}-CN** (δ_H = 4.20 ppm, at pH 9.0). The ligation yields were determined by integration against the internal standard added at the beginning of the first iteration cycle (methylsulfonylmethane, (MSM)). The first iteration was performed starting with **Ac-Gly-CN** (0.98 g, 10.0 mmol) and methylsulfonylmethane (MSM) (0.94 mg, 10.0 mmol) in H₂O/D₂O (98:2, 100 mL, 100 mM). The next iteration was carried out on an aliquot of the reaction mixture (50% of the total volume)^{vii}. After the final iteration (iteration 4), the solution was cooled to 0 °C and diluted two-fold by slow addition of DMSO-*d*₆. The resulting solution was filtered sequentially through Dowex® (OH⁻) and Dowex® (H⁺), and spiked with authentic **Ac-(Gly)₅-CN**.^{viii} Yields of each ligation step, and overall yield of the iteration sequence starting with **Ac-Gly-CN**, are given in Supplementary Table 14. ¹H NMR spectra of each ligation step to produce **Ac-(Gly)_n-CN** are shown in Supplementary Figure 209. A stacked ¹H NMR spectra for iterative cycle 1 (**Gly-CN**→**Ac-Gly-CN**→**Ac-Gly-SNH₂**→**Ac-Gly-SH**→ **Ac-(Gly)₂-CN**) is shown in Figure 1d of the article associated with this Supporting Information document. A representative collection of ¹H-¹³C HMBC NMR for iteration 2 are shown in Supplementary Figure 210.

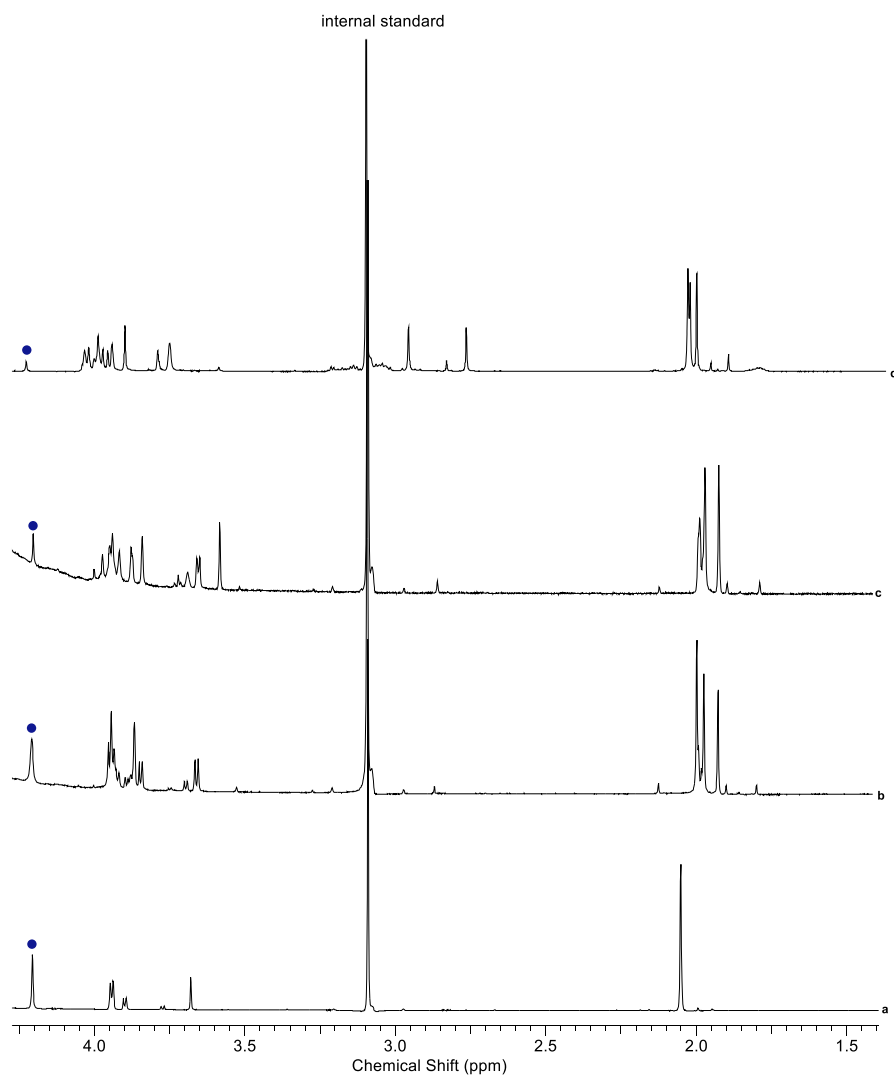
Iteration	Starting material	Product	Yield (step) %	Yield (from Ac-Gly-CN) %
1	Ac-Gly-CN	Ac-(Gly)₂-CN	71	71
2	Ac-(Gly)₂-CN	Ac-(Gly)₃-CN	71	51
3	Ac-(Gly)₃-CN	Ac-(Gly)₄-CN	63	32
4	Ac-(Gly)₄-CN	Ac-(Gly)₅-CN	41	13

Supplementary Table 14. Elongation of *N*-acetylglycine nitrile **Ac-Gly-CN** (100 mM) by iterative thiolysis, hydrolysis, and ligation with glycine nitrile **Gly-CN** (1.2 equiv.) in the presence of K₃[Fe(CN)₆] (3.0 equiv.)

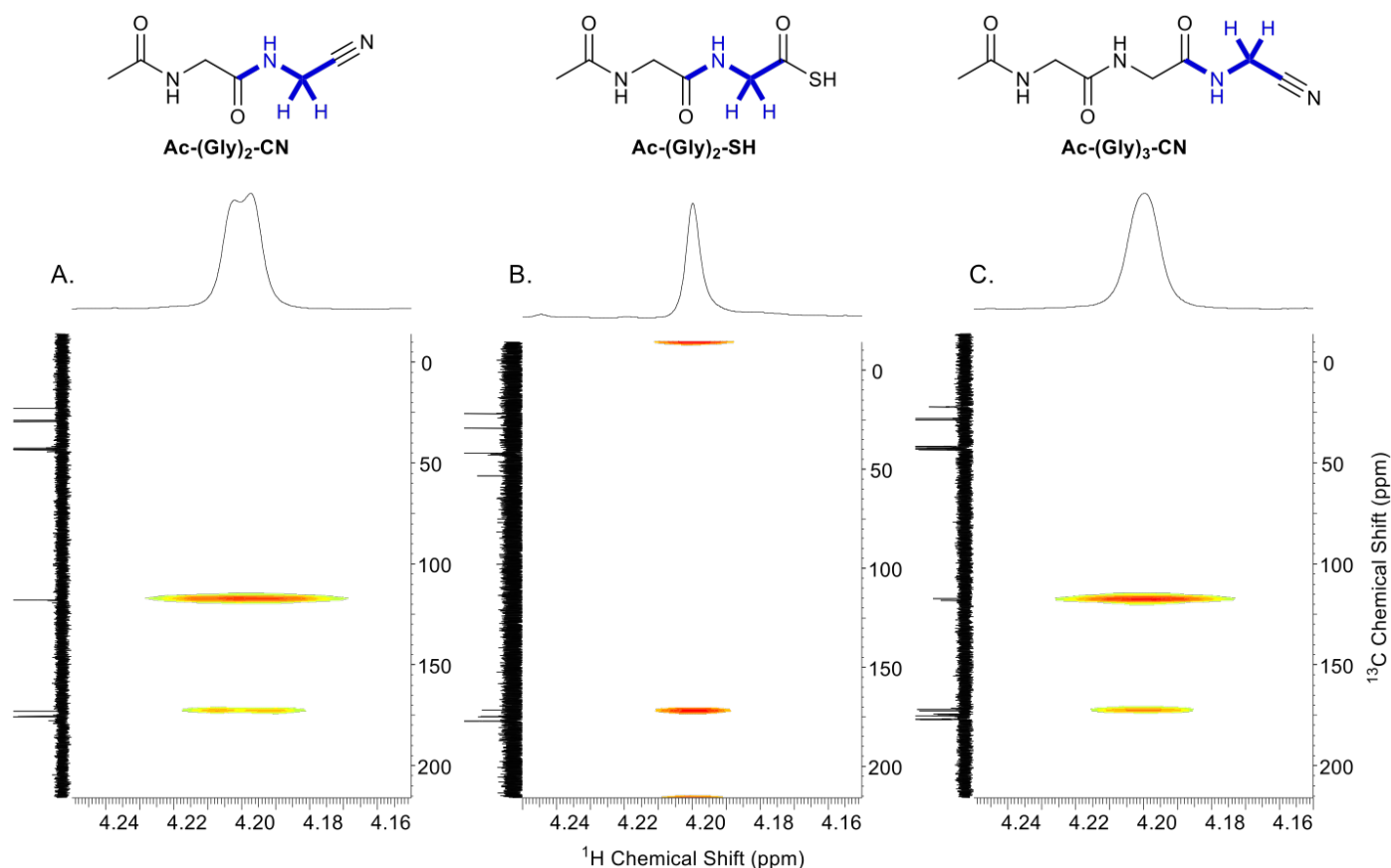
^{vi} Dowex® (SH⁻ form) was used as the counter-cation in place of a metal cation (Na⁺ or K⁺). This prevented the accumulation of metal cations that resulted in marked reduction in line-broadening in ¹H NMR spectra following multiple iterations of peptide elongation.

^{vii} The reaction batch was split after the first iteration cycle and retained as a contingency measure in the event of an experimental error resulting in restarting the iterative elongation sequence from **Ac-Gly-CN**.

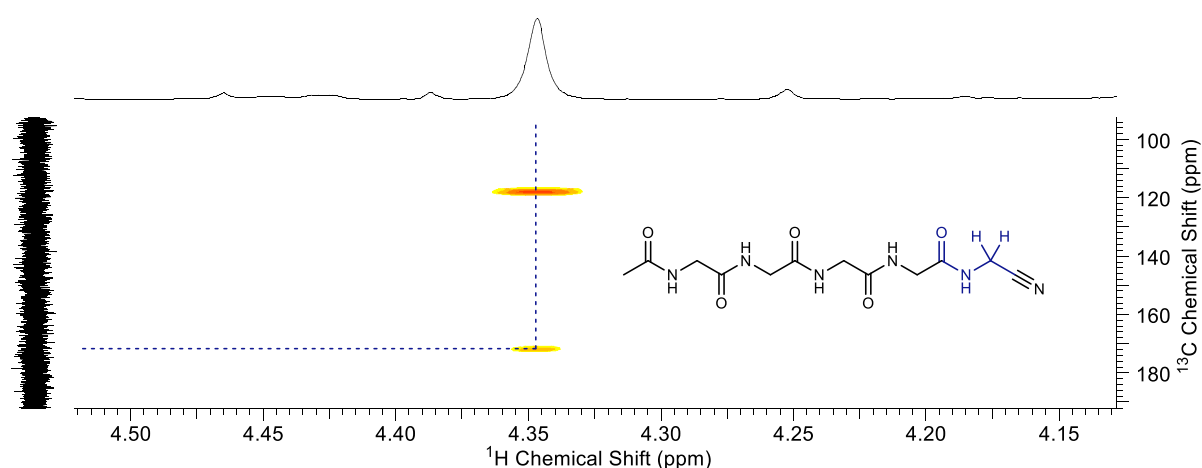
^{viii} The synthesis of an authentic sample of **Ac-(Gly)₅-CN** is described on page Supplementary Page S166.



Supplementary Figure 209. ^1H NMR (600 MHz, pH 9.0, 1.40–4.50 ppm) spectra to show the iterative elongation of *N*-acetylglycine nitrile **Ac-Gly-CN** through iterative thiolysis, hydrolysis, and ligation. **a**: 1st iteration, **Ac-(Gly)₂-CN**, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2; **b**: 2nd iteration, **Ac-(Gly)₃-CN**, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2; **c**: 3rd iteration, **Ac-(Gly)₄-CN**, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2; **d**: 4th iteration, **Ac-(Gly)₅-CN**, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, noesygppr1d.



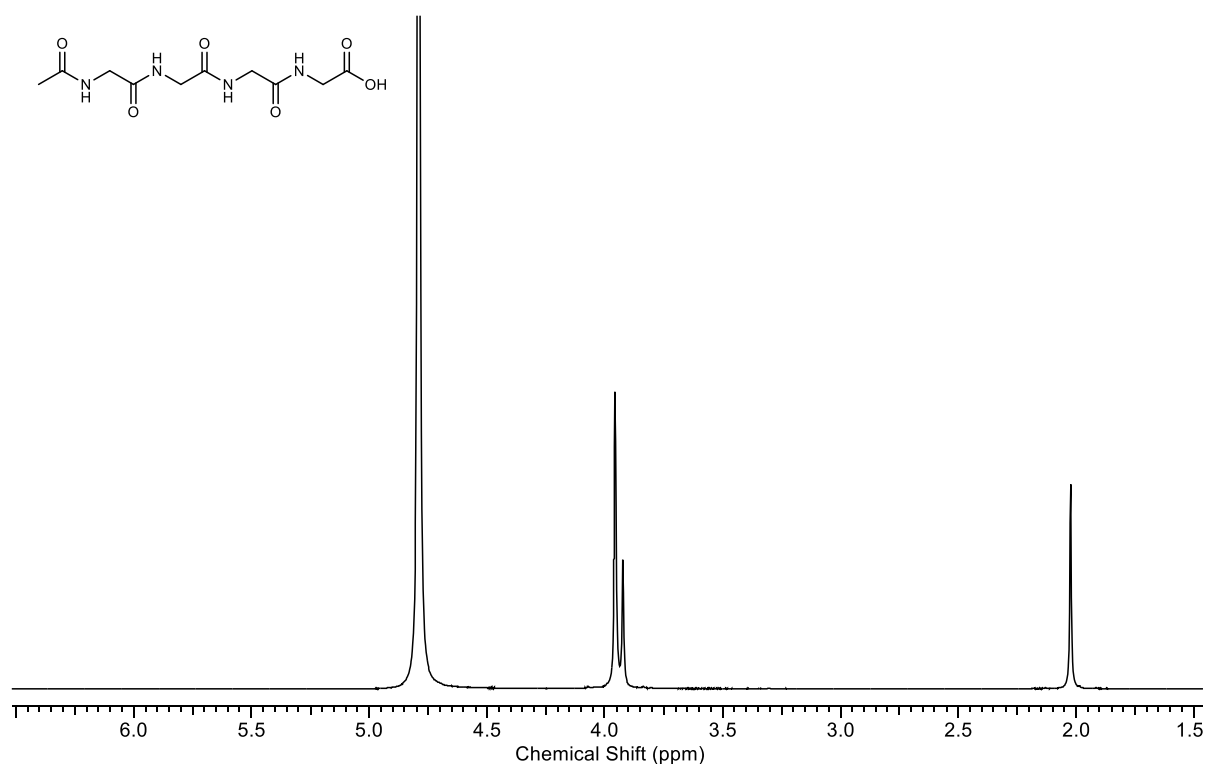
Supplementary Figure 210. ^1H - ^{13}C HMBC (^1H -700 MHz [4.15–4.25 ppm], ^{13}C -176 MHz [−14.0–216.0 ppm], D_2O) spectra showing the $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Gly₂-(C2)-H₂ of **A. Ac-(Gly)₂-CN**, **B. Ac-(Gly)₂-SH**, and the $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Gly₃-(C2)-H₂ of **C. Ac-(Gly)₃-CN**. The characteristic Gly₂-(C2)-H₂ $^2J_{\text{CH}}$ ($\delta_{\text{CN}} = 117.54$ ppm, Supplementary Figure 210A) of **Ac-(Gly)₂-CN** (the product of iteration cycle 1, Supplementary Figure 209a) clearly disappears upon thiolysis and hydrolysis to its corresponding thioacid **Ac-(Gly)₂-SH**, giving rise to the appearance of a higher frequency Gly₂-(C2)-H₂ $^2J_{\text{CH}}$ ($\delta_{\text{CO}_2\text{S}} = 215.7$ ppm, Supplementary Figure 210B), which underwent a distinct fold-over (or *aliasing*)⁵² resulting in part of the Gly₂-(C2)-H₂ $^2J_{\text{CH}}$ HMBC cross-peak appearing at the edge of the ^{13}C spectral window at 215.7 ppm and −14.2 ppm. Ligation of **Ac-(Gly)₂-SH** with **Gly-CN** gives the final product of iteration cycle 2, **Ac-(Gly)₃-CN**, which is characterised by the appearance of a new Gly₃-(C2)-H₂ $^2J_{\text{CH}}$ ($\delta_{\text{CN}} = 117.53$ ppm, Supplementary Figure 210A). The $^3J_{\text{CH}}$ of the terminal glycyl residue remains intact throughout the sequence (Supplementary Figure 210A, $\delta_{\text{C}} = 173.2$; Supplementary Figure 210B, $\delta_{\text{C}} = 171.6$; Supplementary Figure 210C, $\delta_{\text{C}} = 172.6$), demonstrating that the terminal glycyl residue is peptide-bound. A ^1H - ^{13}C HMBC NMR spectrum of the final ligation product **Ac-(Gly)₅-CN** is shown in Supplementary Figure 211 showing the characteristic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Gly₅-(C2)-H₂ of **Ac-(Gly)₅-CN**.



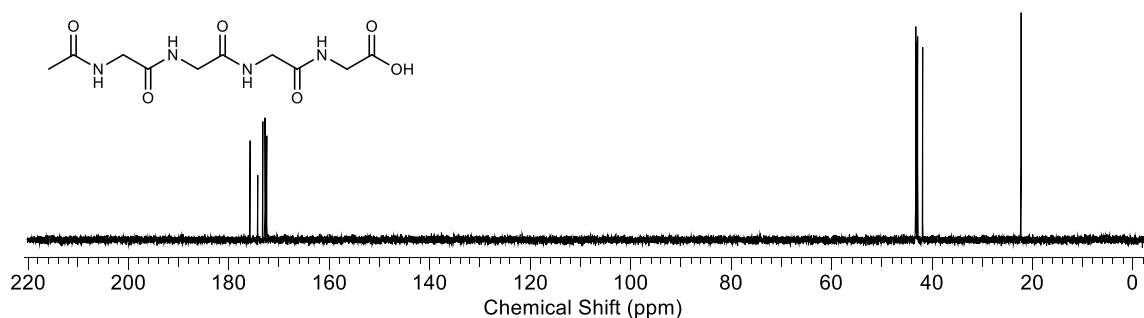
Supplementary Figure 211. ^1H - ^{13}C HMBC (^1H -600 MHz [4.15–4.50 ppm], ^{13}C -151 MHz [100–190 ppm], $\text{D}_2\text{O}/\text{H}_2\text{O}/\text{DMSO}-d_6$ 1:49:50) spectrum showing the diagnostic $^2J_{\text{CH}}$ and $^3J_{\text{CH}}$ coupling of Gly₅-(C2)-H₂ of **Ac-(Gly)₅-CN** at 4.35 ppm with a CN resonance at 118.0 ppm and a C=O resonance at 173.9 ppm, which is characteristic of peptide bond formation of glycine nitrile **Gly-CN**.

N-Acetylglycylglycylglycylglycylglycine **Ac-Gly-Gly-Gly-Gly-OH**

CC(=O)NCC(=O)NCC(=O)NCC(=O)NCC(=O)O Glycylglycylglycine (1.0 g, 5.27 mmol) and triethylamine (642 mg, 0.88 mL, 6.34 mmol) were dissolved in DMF (50 mL). *N*-Acetylglycine pentafluorophenyl ester (1.49 g, 5.27 mmol) was added, and the solution was stirred at room temperature for 2 h. After that time, the solution was concentrated *in vacuo*. The residue was dissolved in H₂O (10 mL) at pH 1.0 and extracted with DCM (3 x 50 mL). The aqueous layer was kept at 4 °C for 24 h. After that time, a white precipitate formed, which was isolated by filtration and washed with H₂O (1 x 20 mL), methanol (1 x 20 mL) and diethyl ether (1 x 50 mL) to afford *N*-acetylglycylglycylglycylglycylglycine **Ac-Gly-Gly-Gly-Gly-OH** as a white powder (1.09 g, 3.78 mmol, 71%). ¹H NMR (600 MHz, D₂O): δ 3.96 (br s, 6H, Gly₂-(C2)-H₂), (Gly₃-(C2)-H₂), (Gly₄-(C2)-H₂), 3.92 (s, 2H, Gly₁-(C2)-H₂), 2.03 (s, 3H, (CO)CH₃). ¹³C NMR (151 MHz, D₂O): δ 175.6 (CO)CH₃, 174.1 (Gly₁-(C1)), 173.1 (Gly₂-(C1)), 172.7 (Gly₃-(C1)), 172.4 (Gly₁-(C2)), 43.1 (Gly₁-(C2)), 43.0 (Gly₃-(C2)), (Gly₂-(C2)), 41.7 (Gly₄-(C2)), 22.6 (CO)CH₃. HRMS (ES⁺): calcd. for [C₁₀H₁₆N₄O₆Na]⁺: 311.0962; observed: 311.0967. IR (solid): 3330, 3260, 3083, 2978, 2937, 1707, 1637, 1551 cm⁻¹. M.p. 245.0–246.6 °C (decomp.).

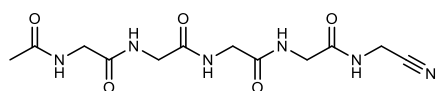


Supplementary Figure 212. ¹H NMR (600 MHz, D₂O, 1.50–6.50 ppm) spectrum to show *N*-acetylglycylglycylglycylglycylglycine **Ac-Gly-Gly-Gly-Gly-OH**



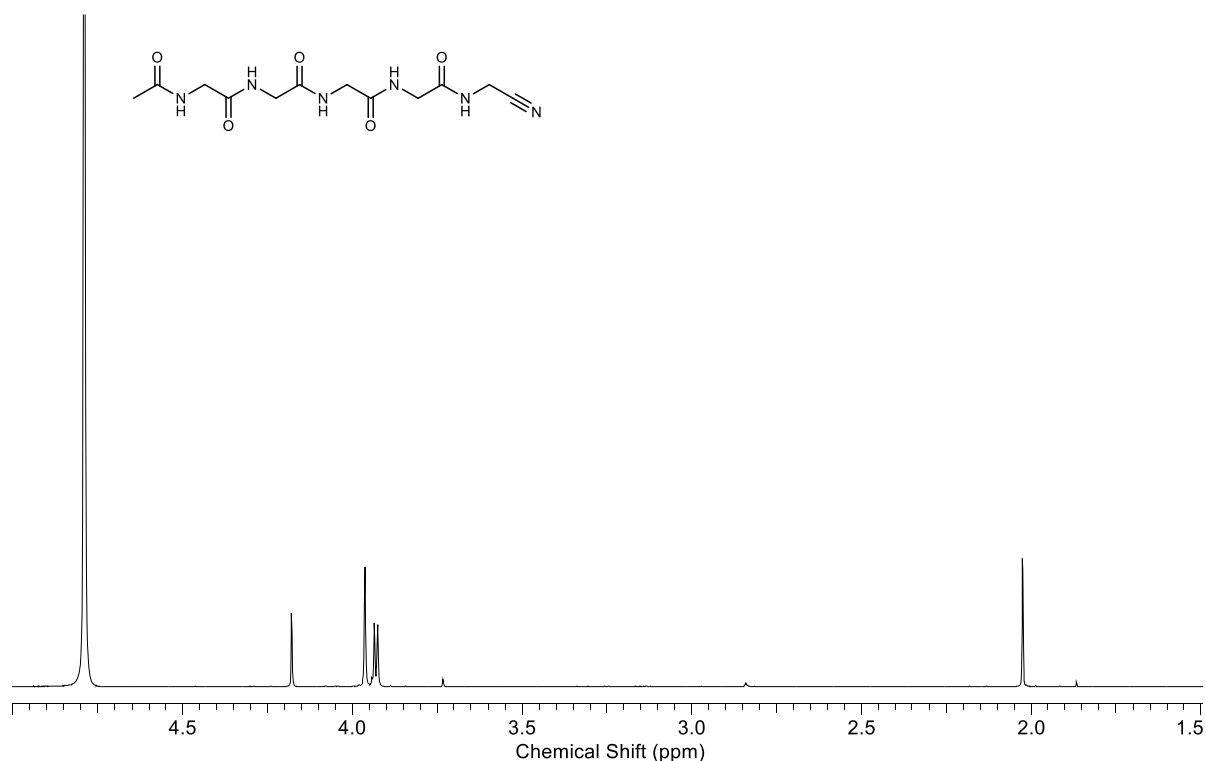
Supplementary Figure 213. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylglycylglycylglycylglycylglycine **Ac-Gly-Gly-Gly-Gly-OH**

N-Acetylglycylglycylglycylglycylglycine nitrile **Ac-Gly-Gly-Gly-Gly-Gly-CN**

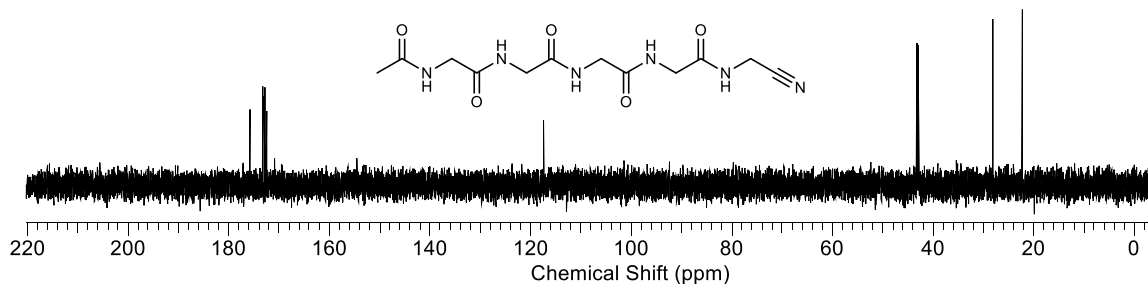


N-Acetylglycylglycylglycylglycylglycine **Ac-Gly-Gly-Gly-Gly-OH** (500 mg, 1.73 mmol), glycine nitrile hydrochloride **Gly-CN·HCl** (480 mg, 5.19 mmol) and triethylamine (875 mg, 1.21 mL, 8.65 mmol) were dissolved in DMF (25 mL).

EDC·HCl (497 mg, 2.60 mmol) was added, and the solution was stirred at room temperature for 4 h. After that time, the solution was concentrated *in vacuo*. The crude residue was dissolved in H₂O (20 mL) and refrigerated at 4 °C for 24 h. After that time, a white precipitate formed, which was isolated by filtration and washed with H₂O (5 x 25 mL), methanol (5 x 25 mL) and ether (5 x 25 mL) to afford *N*-acetylglycylglycylglycylglycylglycine nitrile **Ac-Gly-Gly-Gly-Gly-Gly-CN** as a white powder (203 mg, 0.62 mmol, 36%). ¹H NMR (600 MHz, D₂O): δ 4.18 (s, 2H, Gly₅-(C2)-H₂), 3.96 (s, 4H, Gly₃-(C2)-H₂), Gly₄-(C2)-H₂), 3.94 (s, 2H, Gly₂-(C2)-H₂), 3.96 (s, 2H, Gly₁-(C2)-H₂), 2.02 (s, 3H, (CO)CH₃). ¹³C NMR (151 MHz, D₂O): δ 175.8 (COCH₃), 173.3 (Gly₁-(C1)), 173.0 (Gly₃-(C1)), 172.8 (Gly₂-(C1)), 172.5 (Gly₄-(C1)), 117.4 (Gly₅-(C1)), 42.2 (Gly₁-(C2)), 43.1 (Gly₃-(C2)), 42.9 (Gly₂-(C2)), 42.8 (Gly₄-(C2)), 29.1 (Gly₅-(C2)), 22.3 (COCH₃). HRMS (ES⁺): calcd. for [C₁₂H₁₈N₆O₅+H]⁺ 327.1411; observed: 327.1419. IR (solid): 3368, 3329, 3312, 3229, 3134, 2198, 1742, 1687, 1565, 1501 cm⁻¹. M.p. 270.5–271.9 °C (decomp.).



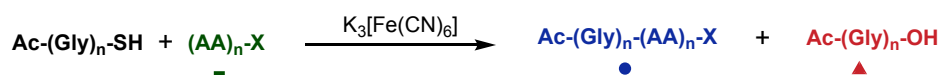
Supplementary Figure 214. ¹H NMR (600 MHz, D₂O, 1.50–5.00 ppm) spectrum to show *N*-acetylglycylglycylglycylglycylglycine nitrile **Ac-Gly-Gly-Gly-Gly-Gly-CN**



Supplementary Figure 215. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show *N*-acetylglycylglycylglycylglycylglycine nitrile **Ac-Gly-Gly-Gly-Gly-Gly-CN**

Prebiotic peptide fragment ligations

General procedure J: Prebiotic peptide fragment ligations

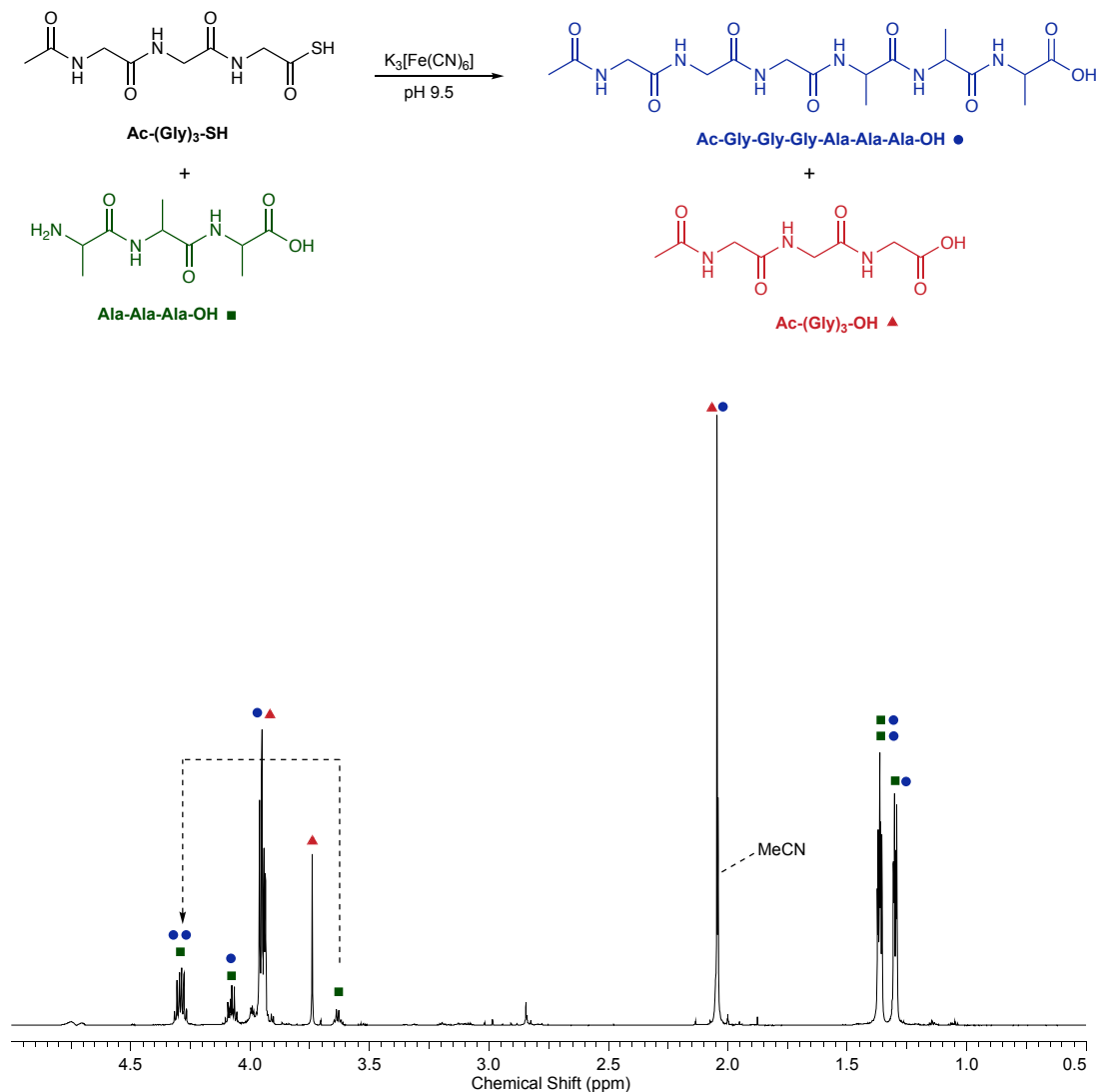


All prebiotic peptide fragment ligations were carried out according to *general procedure K*, unless stated otherwise. Amidothioacid **Ac-Gly_n-S-Na⁺** (*n* = 3, 5 or 6; 3.1 – 30.0 mM) and peptide **AA_n-X** (*X* = CO₂H or CN; *n* = 3 or 5; 1 – 2 equiv.) were dissolved in degassed D₂O and the solution was adjusted to pD 9.5 with 4 M NaOH. Potassium hexacyanoferrate(III) (3 equiv.) was added and the solution was stirred at room temperature for 20 min whilst maintaining the solution at pD 9.5 with 1 M – 4 M NaOH. The resulting suspension was centrifuged and the supernatant was analysed by 1D and 2D NMR spectroscopy (¹H–¹H COSY; ¹H–¹³C HSQC; ¹H–¹³C HMBC). The presence of ligation product **Ac-Gly_n-AA_n-X** (*X*=CO₂H or CN) was quantified using relative integral analysis by ¹H, ¹H–¹³C HMBC NMR spectral analysis, and high resolution mass spectrometry. Yields and HRMS-ESI data are given in Supplementary Table 15.

Ac-(Gly) _n -SH	(AA) _n -X	Ac-(Gly) _n -(AA) _n -X (%)	HRMS-ESI Data for Ac-(Gly) _n -(AA) _n -X		
			Formula	Calculated	Found
Ac-Gly-Gly-Gly-SH	Ala-Ala-Ala-OH	65	C ₁₇ H ₂₉ N ₆ O ₈ [M+H] ⁺	445.2041	445.2044
Ac-Gly-Gly-Gly-SH	Arg-Gly-Asp-OH	76	C ₂₀ H ₃₄ N ₉ O ₁₀ [M+H] ⁺	560.2429	560.2416
Ac-Gly-Gly-Gly-SH	Gly-Gly-Gly-OH	90	C ₁₄ H ₂₃ N ₆ O ₈ [M+H] ⁺	403.1572	403.1571
Ac-Gly-Gly-Gly-SH	Gly-Gly-His-OH	90 ^[a]	C ₁₈ H ₂₇ N ₈ O ₈ [M+H] ⁺	483.1946	483.1947
Ac-Gly-Gly-Gly-SH	Gly-Gly-Gly-CN	>95 ^[b]	C ₁₄ H ₂₁ N ₇ O ₆ Li [M+Li] ⁺	389.1704	389.1703
Ac-Gly-Gly-Gly-SH	Leu-Leu-Leu-OH	70	C ₂₆ H ₄₇ N ₆ O ₈ [M+H] ⁺	571.3450	571.3453
Ac-Gly-Gly-Gly-SH	Met-Ala-Ser-OH	75	C ₁₉ H ₃₃ N ₆ O ₉ S [M+H] ⁺	521.2024	521.2028
Ac-Gly-Gly-Gly-SH	Phe-Gly-Gly-OH	74	C ₂₁ H ₂₉ N ₆ O ₈ [M+H] ⁺	493.2041	493.2039
Ac-Gly-Gly-Gly-Gly-Gly-SH	Ala-Ala-Ala-OH	74	C ₂₁ H ₃₅ N ₈ O ₁₀ [M+H] ⁺	559.2476	559.2472
Ac-Gly-Gly-Gly-Gly-Gly-SH	Gly-Gly-His-OH	80	C ₂₂ H ₃₃ N ₁₀ O ₁₀ [M+H] ⁺	597.2381	597.2361
Ac-Gly-Gly-Gly-Gly-Gly-Gly-SH	Gly-Gly-Gly-CN	43 ^[c]	C ₂₀ H ₃₀ N ₁₀ O ₉ Li [M+Li] ⁺	561.2358	561.2352
Ac-Gly-Gly-Gly-Gly-Gly-SH	Gly-Gly-Gly-Gly-Gly-OH	79 ^[d] (66 ^[e])	C ₂₂ H ₃₃ N ₁₀ O ₁₂ [M-H] [–]	629.2280	629.2259
Ac-Gly-Gly-Gly-Gly-Gly-Gly-SH	Gly-Gly-Gly-Gly-Gly-OH	>95 ^[f] (92 ^[g])	C ₂₄ H ₃₆ N ₁₁ O ₁₃ [M-H] [–]	686.2494	686.2478

Supplementary Table 15. Yields and the products of oxidative coupling of amidothioacid **Ac-Gly_n-SH** (*n* = 3, 5 or 6; 3.1–30 mM) with peptide **AA_n-X** (*n* = 3 or 5; 1–2 equiv.; *X* = OH, CN) and K₃[Fe(CN)₆] (3 equiv.). Reactions typically carried out with **Ac-Gly_n-SH** (25 mM) and peptide **AA_n-X** (25 mM, *X* = OH, CN) and K₃[Fe(CN)₆] (3 equiv.) unless stated otherwise. ^[a] Ligation carried out with **Ac-Gly₃-SH** (30 mM) and **Gly-Gly-His-OH** (25 mM). ^[b] Isolated yield determined by quantification of residual **Ac-Gly₃-OH** and **Gly₃-CN** in the supernatant by ¹H NMR spectroscopy using formate as the internal standard (see Supplementary Information Page S174), and further confirmed upon subsequent conversion of **Ac-Gly₆-CN** after one iterative ligation cycle into **Ac-Gly₁₁-OH** (entry 13; see Supplementary Information Page S184). ^[c] Yield of **Ac-Gly₉-CN** quantified using MSM as an internal standard after reaction of **Ac-Gly₆-SH** (9.0 mM) and **Gly₃-CN** (12.5 mM). The yield of **Ac-Gly₉-CN** (50%) was determined over four steps from **Ac-Gly₃-SH** to **Ac-Gly₉-CN**. ^[d] Yield determined by quantification of residual **Gly₅-OH** in the supernatant by ¹H NMR spectroscopy using MSM as an internal standard. ^[e] Isolated yield determined after hydrolysis of the **Ac-Gly₁₀-OH** precipitate to glycine **Gly** and acetic acid **AcOH** in 6 M HCl. See Supplementary Information Page S182 for further details. ^[f] Yield of ligation after reaction of **Ac-Gly₆-SH** (3.13 mM, produced from **Ac-Gly₆-CN** by thiolysis to **Ac-Gly₆-SNH₂** and hydrolysis to **Ac-Gly₆-SH**) and **Gly₅-OH** (6.25 mM). Yield determined by quantification of residual **Gly₅-OH** in the supernatant by ¹H NMR spectroscopy using MSM as an internal standard. See Supplementary Information Page S184 for further details. ^[g] Isolated yield determined by quantitative hydrolysis of isolated **Ac-Gly₁₁-OH** precipitate to glycine **Gly** and acetic acid **AcOH** in 6 M HCl. See Supplementary Information Page S186 for further details.

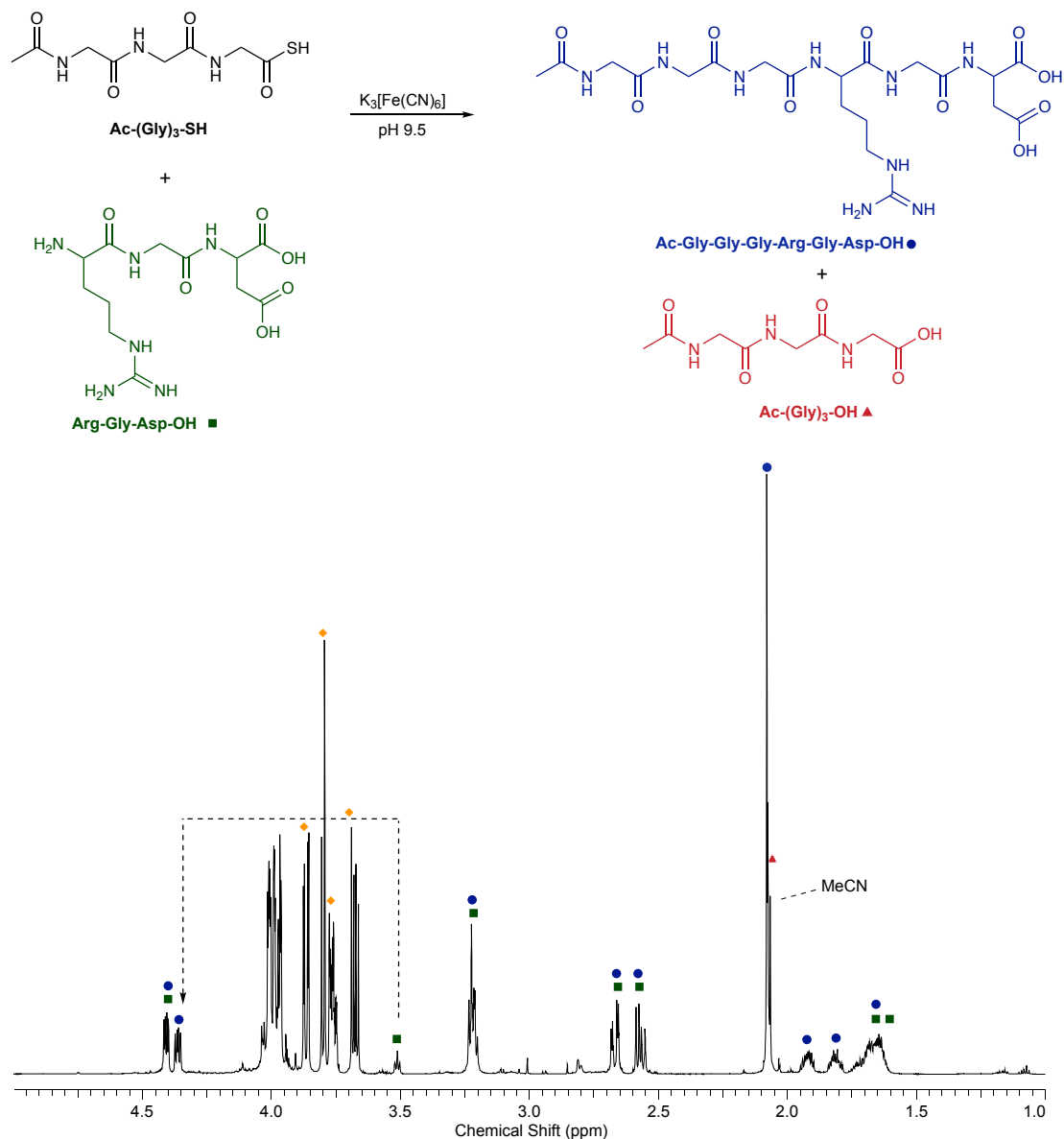
N-Acetylglycylglycylglycylalanylalanylalanine **Ac-Gly-Gly-Gly-Ala-Ala-Ala-OH**



Supplementary Figure 216. 1H NMR (700 MHz, D_2O , noesygppr1d, 0.5–5.0 ppm) spectrum to show the reaction of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid (**Ac-Gly₃-SH**; 25 mM), alanylalanylalanine (**Ala-Ala-Ala-OH**; 25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D_2O (1 mL) at pD 9.5 and room temperature. The dashed lines and arrows highlight the characteristic change in chemical shift of $Ala_1(C2)-H$ of **Ala-Ala-Ala-OH** upon acylation by **Ac-Gly₃-SH** to give **Ac-Gly-Gly-Gly-Ala-Ala-Ala-OH** (+0.63 ppm).

1H NMR (700 MHz, D_2O , noesygppr1d): *N*-Acetylglycylglycylglycylalanylalanylalanine **Ac-Gly-Gly-Gly-Ala-Ala-Ala-OH** (•): δ 4.32 - 4.26 (m, 2H, $Ala_4(C2)-H$, $Ala_6(C2)-H$), 4.07 (q, $J = 7.2$ Hz, 1H, $Ala_5(C2)-H$), 3.96 - 3.93 (m, 6H, $Gly_1(C2)-H_2$, $Gly_2(C2)-H_2$, $Gly_3(C2)-H_2$), 2.05 (s, 3H, $COCH_3$), 1.36 (obs., 6H, $Ala_4(C3)-H_3$, $Ala_6(C3)-H_3$), 1.30 (d, $J = 7.2$ Hz, 1H, $Ala_5(C3)-H_3$). *N*-Acetylglycylglycylglycine **Ac-Gly₃-OH** (▲): δ 3.97 - 3.93 (obs., 4H, $2 \times Gly(C2)-H_2$), 3.74 (s, 2H, $Gly(C2)-H_2$), 2.05, (s, 3H, $COCH_3$). *Alanylalanylalanine* **Ala-Ala-Ala-OH** (■): δ 4.29 (q, $J = 7.0$ Hz, 1H, $Ala_3(C2)-H$), 4.09 (q, $J = 7.2$ Hz, 1H, $Ala_2(C2)-H$), 3.63 (q, $J = 7.0$ Hz, 1H, $Ala_1(C2)-H$), 1.36 (d, $J = 7.0$ Hz, 3H, $Ala_3(C3)-H_3$), 1.30 (d, $J = 7.2$ Hz, 3H, $Ala_2(C3)-H_3$), 1.30 (d, $J = 7.0$ Hz, 3H, $Ala_1(C3)-H_3$).

N-Acetylglycylglycylglycylarginylglycylaspartic acid **Ac-Gly-Gly-Gly-Arg-Gly-Asp-OH**

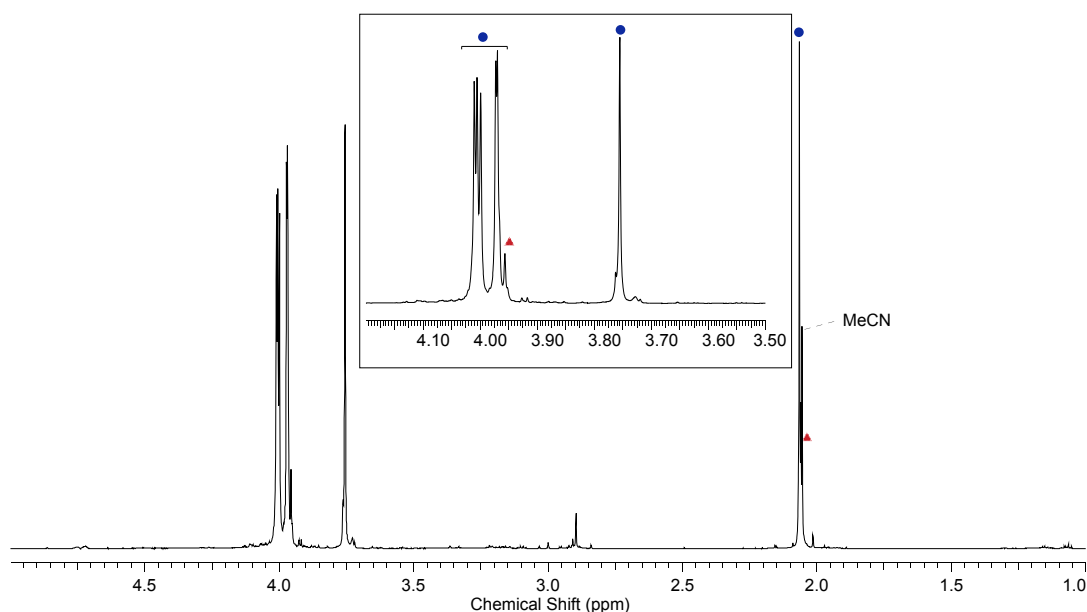
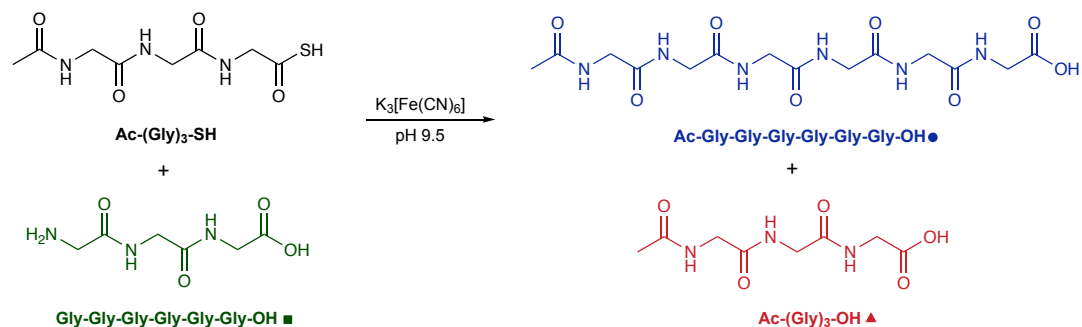


Supplementary Figure 217. 1H NMR (700 MHz, D_2O , noesygppr1d, 1.0–5.0 ppm) spectrum to show the reaction of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid (**Ac-Gly₃-SH**; 25 mM), arginylglycylaspartic acid (**Arg-Gly-Asp-OH**; 25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D_2O (1 mL) at pD 9.5 and room temperature. The dashed lines and arrows highlight the characteristic change in chemical shift of Arg-(C2)-H of **Arg-Gly-Asp-OH** upon acylation by **Ac-Gly₃-SH** to give **Ac-Gly-Gly-Gly-Arg-Gly-Asp-OH** (+0.85 ppm).^{ix}

1H NMR (700 MHz, D_2O , noesygppr1d): *N*-Acetylglycylglycylglycylarginylglycylaspartic acid **Ac-Gly-Gly-Gly-Arg-Gly-Asp-OH** (•) (partial assignment): δ 4.41 (ABX, $J = 3.9, 8.9$ Hz, 1H, Asp-(C2)-H), 4.36 (dd, $J = 5.4, 9.0$ Hz, 1H, Arg-(C2)-H), 3.22 (t, $J = 7.0$ Hz, 2H, Arg-(C5)-H₂), 2.67 (dd, $J = 3.9, 15.9$ Hz, 1H, Asp-(C3)-H), 2.57 (dd, $J = 8.9, 15.9$ Hz, 1H, Asp-(C3)-H'), 2.08 (s, 3H, COCH₃), 1.95 - 1.89 (m, 1H, Arg-(C3)-H), 1.85 - 1.78 (m, 1H, Arg-(C3)-H'), 1.71 - 1.60 (m, 2H, Arg-(C4)-H₂). *N*-Acetylglycylglycylglycine **Ac-Gly₃-OH** (▲) (partial assignment): δ 2.08 (s, 3H, COCH₃). Arginylglycylaspartic acid **Arg-Gly-Asp-OH** (■): δ 4.41 (ABX, $J = 3.9, 8.9$ Hz, 1H, Asp-(C2)-H), 3.51 (t, $J = 6.4$ Hz, 1H, Arg-(C2)-H), 3.21 (t, $J = 6.7$ Hz, 2H, Arg-(C5)-H₂), 2.67 (dd, $J = 3.9, 15.8$ Hz, 1H, Asp-(C3)-H), 2.57 (dd, $J = 8.9, 15.8$ Hz, 1H, Asp-(C3)-H'), 1.77 - 1.56 (m, 4H, Arg-(C3)-H₂, Arg-(C4)-H₂).

^{ix} **Arg-Gly-Asp-OH** (batch number FA350751601, CAS 99896-85-2, 33% purity, determined by 1H NMR spectroscopy using MSM as an internal reference standard) was supplied by Carbosynth as a mixture and used without further purification. The major contaminant (•) was purported by Carbosynth to be 2-(3,6-dioxopiperazin-2-yl)acetic acid (or 'cyclo-(Asp-Gly)'). The impurity did not effect the outcome of the ligation reaction and was returned unchanged, alongside the coupling product.

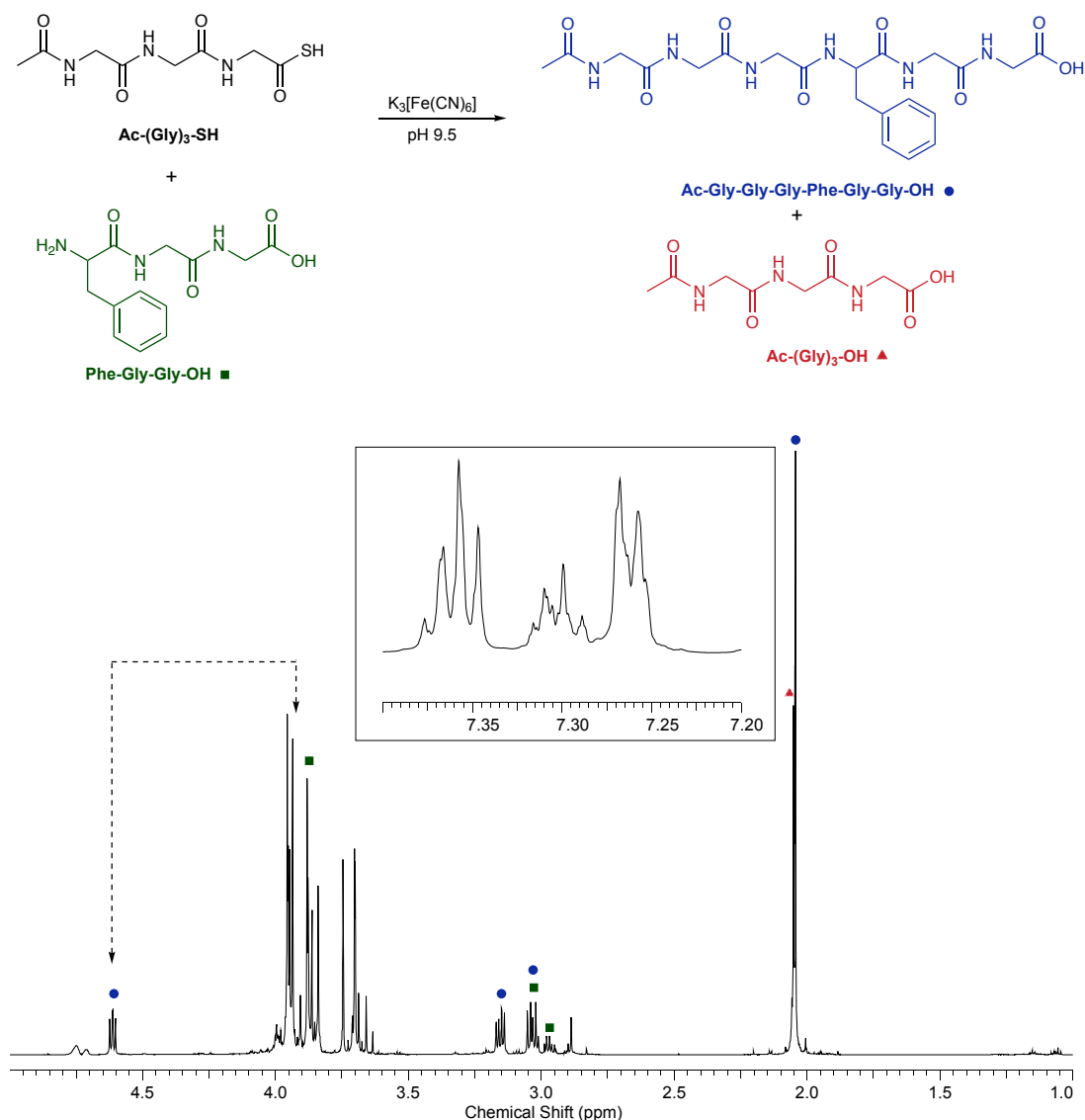
N-Acetylglycylglycylglycylglycylglycylglycine **Ac-Gly-Gly-Gly-Gly-Gly-Gly-OH**



Supplementary Figure 218. ¹H NMR (700 MHz, D₂O, noesygppr1d, 1.0–5.0 ppm) spectrum to show the reaction of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid (**Ac-Gly₃-SH**; 25 mM), glycylglycylglycine (**Gly-Gly-Gly-OH**; 25 mM) and K₃[Fe(CN)₆] (75 mM) in D₂O (1 mL) at pD 9.5 and room temperature. Inset: ¹H NMR (700 MHz, D₂O, 7.2–7.4 ppm) spectral region showing six distinct resonances for each of the Gly-(C2)-H₂ residues of **Ac-Gly-Gly-Gly-Gly-Gly-Gly-OH**. Residual **Gly-Gly-Gly-OH**, where the resonance for Gly₁-(C2)-H₂ is observed at approximately within 3.40 – 3.50 ppm at pD 9.5, was not detected.

¹H NMR (700 MHz, D₂O, noesygppr1d): *N*-Acetylglycylglycylglycylglycylglycylglycine **Ac-Gly-Gly-Gly-Gly-Gly-Gly-OH** (•): δ 4.01 (s., 2H, Gly-(C2)-H₂), 4.01 (s, 2H, Gly-(C2)-H₂), 4.00 (s, 2H, Gly-(C2)-H₂), 3.97 (s, 2H, Gly-(C2)-H₂), 3.97 (s, 2H, Gly-(C2)-H₂), 3.76 (s, 2H, Gly-(C2)-H₂), 2.07 (s, 3H, COCH₃). *N*-Acetylglycylglycylglycine **Ac-Gly₃-OH** (▲) (partial assignment): δ 3.95 (s, 2H, Gly-(C2)-H₂), 2.06 (s., 3H, COCH₃).

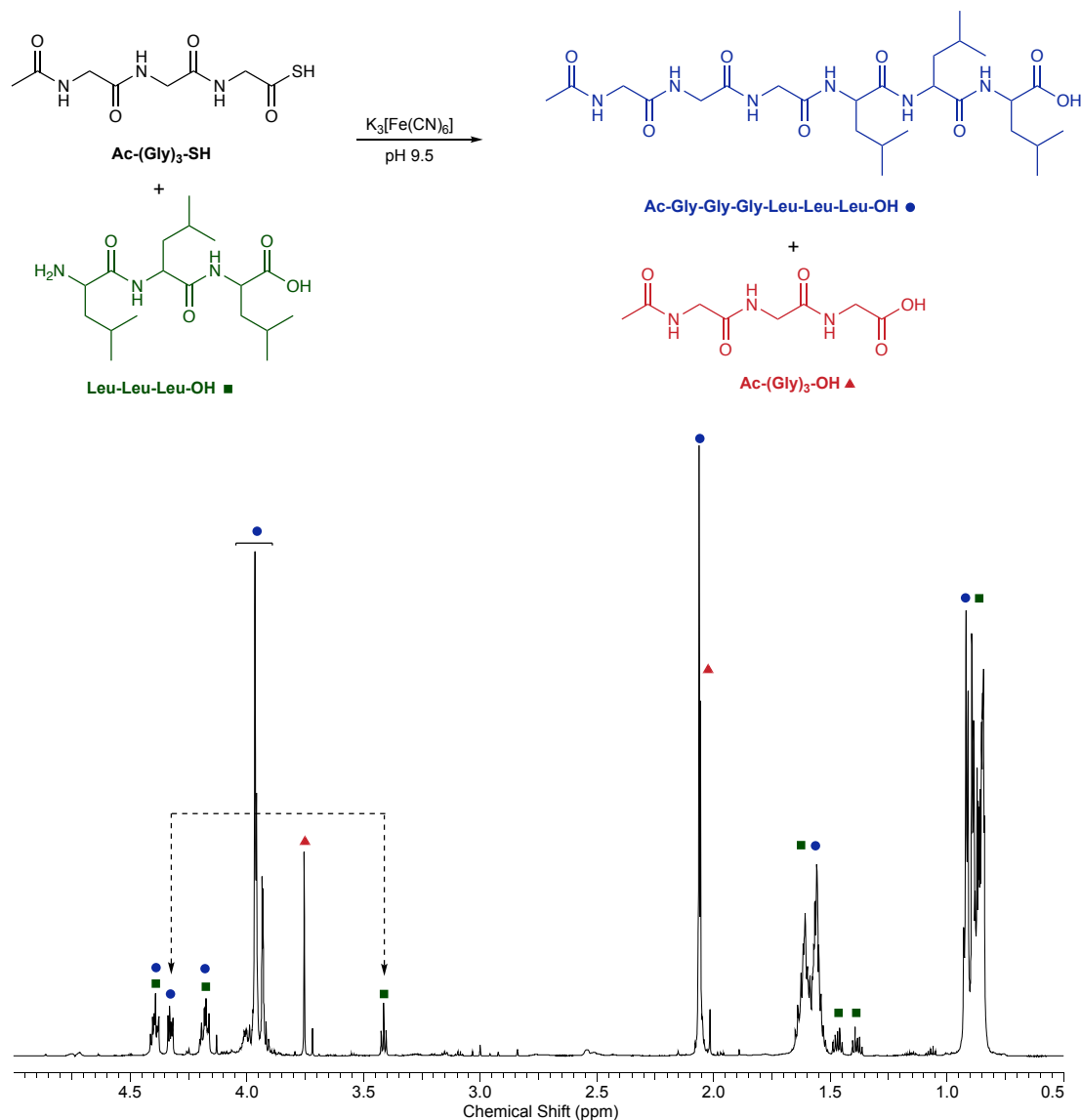
N-Acetylglycylglycylglycylphenylalanylglycylglycine **Ac-Gly-Gly-Gly-Phe-Gly-Gly-OH**



Supplementary Figure 219. 1H NMR (700 MHz, D_2O , noesygppr1d, 1.0–5.0 ppm) spectrum to show the reaction of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid (**Ac-Gly₃-SH**; 25 mM), phenylalanylglycylglycine (**Phe-Gly-Gly-OH**; 25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D_2O (1 mL) at pD 9.5 and room temperature. The dashed lines and arrows highlight the characteristic change in chemical shift of Phe₁-(C2)-H of **Phe-Gly-Gly-OH** upon acylation by **Ac-Gly₃-SH** to give **Ac-Gly-Gly-Gly-Phe-Gly-Gly-OH** (+0.74 ppm). Inset: 1H NMR (700 MHz, D_2O , 7.2–7.4 ppm) spectral region showing the aromatic protons.

1H NMR (700 MHz, D_2O , noesygppr1d): *N*-Acetylglycylglycylglycylphenylalanylglycylglycine **Ac-Gly-Gly-Gly-Phe-Gly-Gly-OH** (•): δ 7.38 - 7.34 (m, 2H, Ar), 7.32 - 7.28 (m, 1H, Ar), 7.28 - 7.24 (m, 2H, Ar), 4.61 (ABX, J = 7.0, 8.5 Hz, 2H, Phe-(C2)-H), 3.96 - 3.70 (5 $\times$ s, 10H, Gly-(C2)-H₂), 3.15 (dd, J = 7.0, 13.9 Hz, 1H, Phe-(C3)-H'), 3.04 (dd, J = 8.5, 13.9 Hz, 1H, Phe-(C3)-H''), 2.04 (s, 3H, COCH₃). *N*-Acetylglycylglycylglycine **Ac-Gly₃-OH** (▲) (partial assignment): δ 2.05 (s, 3H, COCH₃). Phenylalanylglycylglycine **Phe-Gly-Gly-OH** (■): δ 7.38 - 7.33 (m, 2H, Ar), 7.32 - 7.29 (m, 1H, Ar), 7.28 - 7.25 (m, 2H, Ar), 3.96 - 3.70 (2 $\times$ s, 4H, Gly-(C2)-H₂), 3.89 - 3.86 (obs, 1H, Phe-(C2)-H), 3.04 - 3.01 (obs, 1H, Phe-(C3)-H'), 2.97 (dd, J = 7.7, 13.5 Hz, 1H, Phe-(C3)-H'').

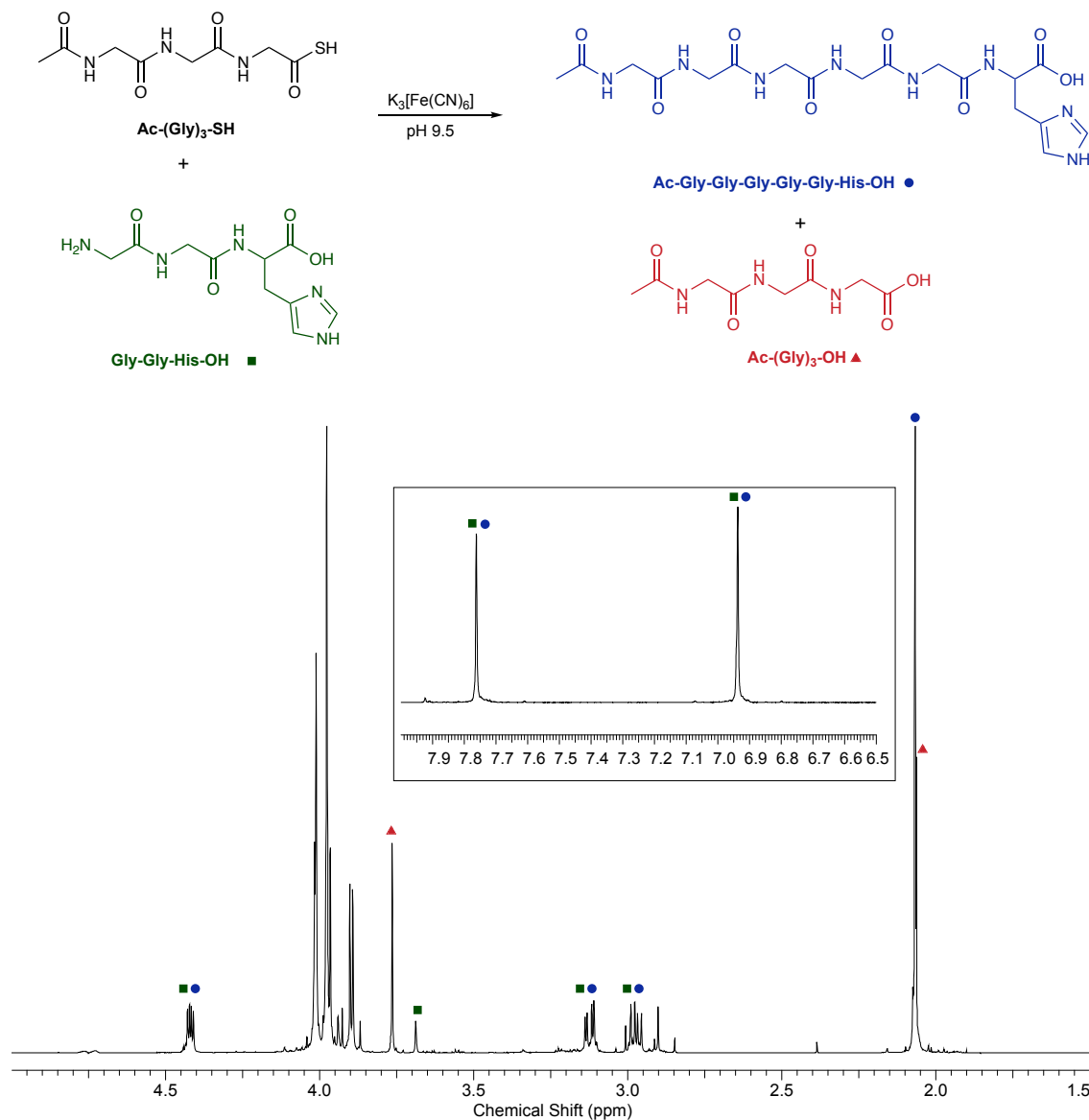
N-Acetylglycylglycylglycylleucylleucylleucine **Ac-Gly-Gly-Gly-Leu-Leu-Leu-OH**



Supplementary Figure 220. ¹H NMR (700 MHz, D₂O, noesygppr1d, 0.50–5.0 ppm) spectrum to show the reaction of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid (**Ac-Gly₃-SH**; 25 mM), leucylleucylleucine (**Leu-Leu-Leu-OH**; 25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D₂O (1 mL) at pD 9.5 and room temperature. The dashed lines and arrows highlight the characteristic change in chemical shift of Leu₁-(C2)-H of **Leu-Leu-Leu-OH** upon acylation by **Ac-Gly₃-SH** to give **Ac-Gly-Gly-Gly-Leu-Leu-Leu-OH** (+0.92 ppm).

¹H NMR (700 MHz, D₂O, noesygppr1d): *N*-Acetylglycylglycylglycylleucylleucylleucine **Ac-Gly-Gly-Gly-Leu-Leu-Leu-OH** (•): δ 4.43 - 4.37 (m, 1H, Leu₆-(C2)-H), 4.33 (dd, $J = 5.2, 9.6$ Hz, 1H, Leu₄-(C2)-H), 4.21 - 4.15 (m, 1H, Leu₅-(C2)-H), 3.99 - 3.90 (m, 6H, Gly₁-(C2)-H₂, Gly₂-(C2)-H₂, Gly₃-(C2)-H₂), 2.06 (s, 3H, COCH₃), 1.66 - 1.51 (m, 9H, Leu₄-(C3)-H₂, Leu₅-(C3)-H₂, Leu₆-(C3)-H₂, Leu₄-(C4)-H, Leu₅-(C4)-H, Leu₆-(C4)-H), 0.95 - 0.81 (m, 18H, Leu₄-(C5)-H₃', Leu₅-(C5)-H₃', Leu₆-(C5)-H₃', Leu₄-(C5)-H₃'', Leu₅-(C5)-H₃'', Leu₆-(C5)-H₃''). *N*-Acetylglycylglycylglycine **Ac-Gly₃-OH** (partial assignment) (▲): δ 3.75 (s, 2H, Gly-(C2)-H₂), 2.06 (s, 3H, COCH₃). *Leucylleucylleucine* **Leu-Leu-Leu-OH** (■): δ 4.43 - 4.37 (m, 1H, Leu₃-(C2)-H), 4.21 - 4.15 (m, 1H, Leu₂-(C2)-H), 3.41 (t, $J = 7.3$ Hz, 1H, Leu₁-(C2)-H), 1.66 - 1.51 (m, 7H, Leu₂-(C3)-H₂, Leu₃-(C3)-H₂, Leu₁-(C4)-H, Leu₂-(C4)-H, Leu₃-(C4)-H), 1.47 (td, 1H, $J = 7.0, 13.6$ Hz, 1H, Leu₁-(C3)-H'), 1.38 (dt, 1H, $J = 7.4, 14.8$ Hz, 1H, Leu₁-(C3)-H''), 0.95 - 0.81 (m, 18H, Leu₁-(C5)-H₃, Leu₂-(C5)-H₃, Leu₃-(C5)-H₃, Leu₁-(C5)-H₃', Leu₂-(C5)-H₃', Leu₃-(C5)-H₃').

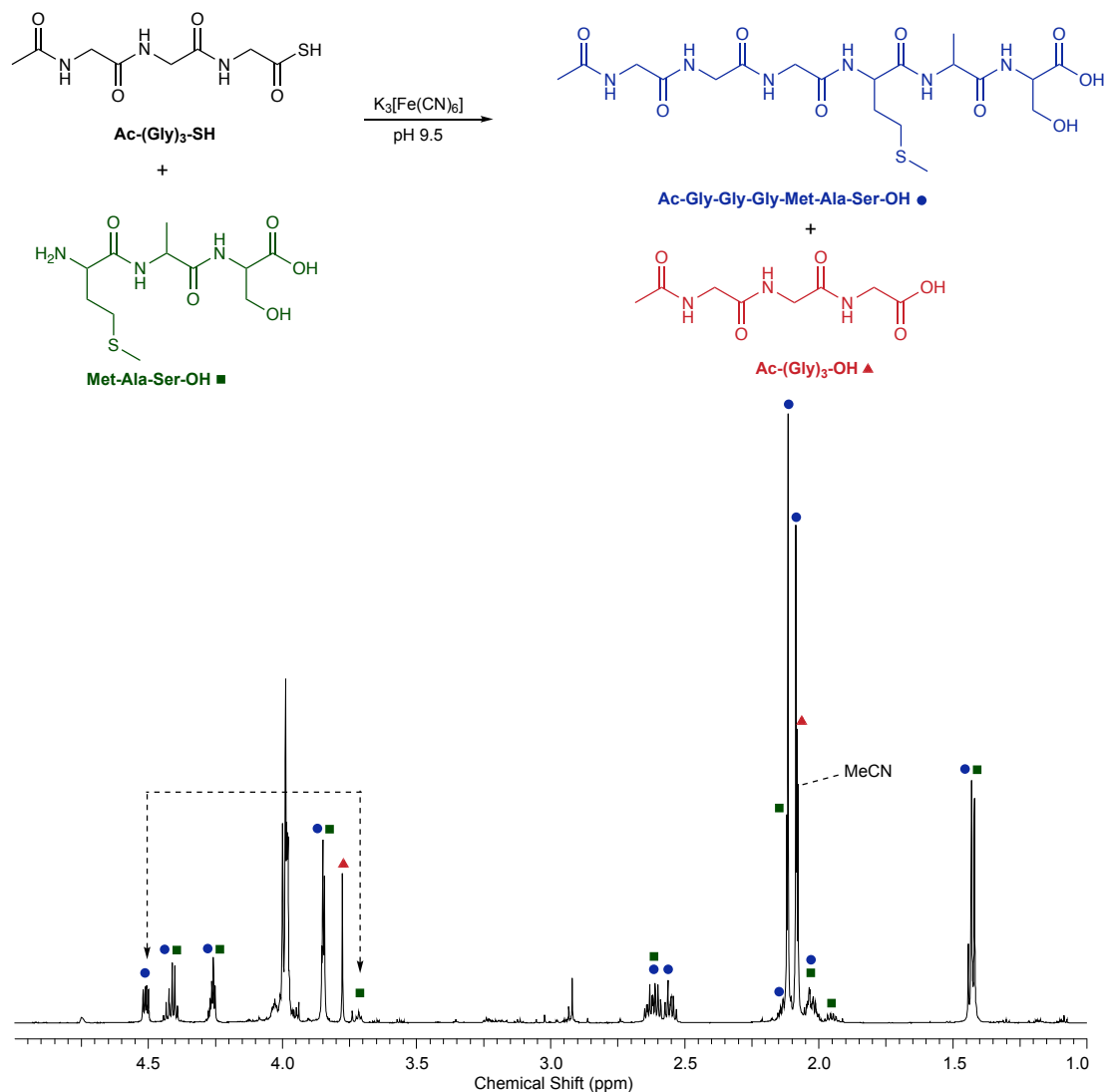
N-Acetylglycylglycylglycylglycylglycylhistidine **Ac-Gly-Gly-Gly-Gly-His-OH**



Supplementary Figure 221. ¹H NMR (700 MHz, D₂O, noesygppr1d, 1.50–5.00 ppm) spectrum to show the reaction of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid (**Ac-Gly₃-SH**; 30 mM), glycylglycylhistidine (**Gly-Gly-His-OH**; 25 mM) and K₃[Fe(CN)₆] (75 mM) in D₂O (1 mL) at pD 9.5 and room temperature. Inset: ¹H NMR (700 MHz, D₂O, 6.5–8.0 ppm) spectral region showing the aromatic protons.

¹H NMR (700 MHz, D₂O, noesygppr1d): *N*-Acetylglycylglycylglycylglycylglycylhistidine **Ac-Gly-Gly-Gly-Gly-His-OH** (•) (partial assignment): δ 7.76 (s, 1H, Ar), 6.94 (s, 1H, Ar), 4.42 (ABX, J = 4.7, 8.8 Hz, 1H, His-(C2)-H), 3.12 (dd, J = 4.7, 15.1 Hz, 1H, His-(C3)-H), 2.97 (dd, J = 8.8, 15.1 Hz, 1H, His-(C3)-H'), 2.07 (s, 3H, COCH₃). *N*-Acetylglycylglycylglycine **Ac-Gly₃-OH** (▲) (partial assignment): δ 3.76 (s, 2H, Gly-(C2)-H₂), 2.07 (s, 3H, COCH₃). Glycylglycylhistidine **Gly-Gly-His-OH** (■) (partial assignment): δ 7.76 (s, 1H, Ar), 6.94 (s, 1H, Ar), 4.45 - 4.42 (obs, 1H, His-(C2)-H), 3.69 (s, 2H, Gly₁-(C2)-H₂), 3.14 - 3.09 (obs, 1H, His-(C3)-H), 3.00 - 2.96 (obs, 1H, His-(C3)-H').

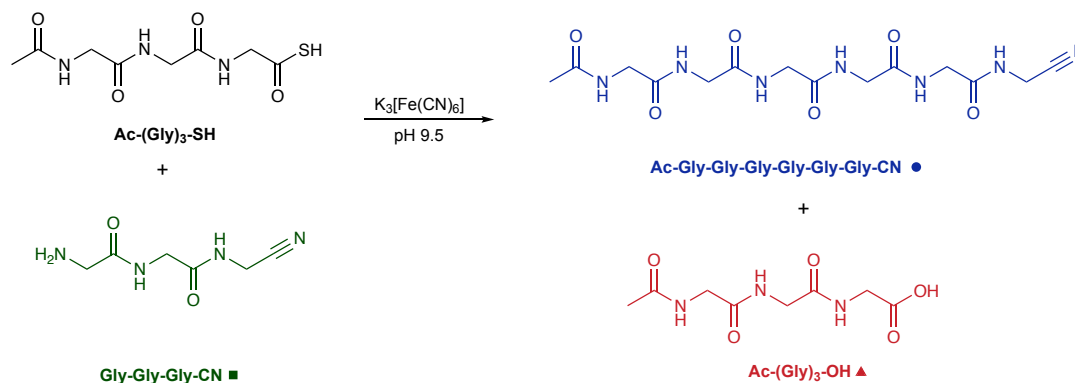
N-Acetylglycylglycylglycylmethionylalanylserine **Ac-Gly-Gly-Gly-Met-Ala-Ser-OH**



Supplementary Figure 222. 1H NMR (700 MHz, D_2O , noesygppr1d, 1.0–5.0 ppm) spectrum to show the reaction of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid (**Ac-Gly₃-SH**; 25 mM), methionylalanylserine (**Met-Ala-Ser-OH**; 25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D_2O (1 mL) at pD 9.5 and room temperature. The dashed lines and arrows highlight the characteristic change in chemical shift of Met-(C2)-H of **Met-Ala-Ser-OH** upon acylation by **Ac-Gly₃-SH** to give **Ac-Gly-Gly-Gly-Met-Ala-Ser-OH** (+0.82 ppm).

1H NMR (700 MHz, D_2O , noesygppr1d): *N*-Acetylglycylglycylglycylmethionylalanylserine **Ac-Gly-Gly-Gly-Met-Ala-Ser-OH** (•) (partial assignment): δ 4.51 (dd, $J = 5.2, 9.2$ Hz, 1H, Met-(C2)-H), 4.41 (q, $J = 7.2$ Hz, 1H, Ala-(C2)-H), 4.26 (t, $J = 4.7$ Hz, 1H, Ser-(C2)-H), 3.85 (d, $J = 4.7$ Hz, 2H, Ser-(C3)-H₂), 2.66 - 2.60 (m, 1H, Met-(C3)-H), 2.60 - 2.52 (m, 1H, Met-(C3)-H'), 2.16 - 2.11 (m, 1H, Met-(C4)-H), 2.12 (s, 3H, SCH₃), 2.09 (s, 3H, COCH₃), 1.98 - 1.90 (m, 1H, Met-(C4)-H'), 1.43 (d, $J = 7.2$ Hz, 3H, Ala-(C3)-H₃). *N*-Acetylglycylglycylglycine **Ac-Gly₃-OH** (▲) (partial assignment): δ 3.78 (s, 2H, Gly-(C2)-H₂), 2.09 (s, 3H, COCH₃). *Methionylalanylserine* **Met-Ala-Ser-OH** (■): δ 4.43 (q, $J = 7.2$ Hz, 1H, Ala-(C2)-H), 4.27 (t, $J = 4.7$ Hz, 1H, Ser-(C2)-H), 3.86 (d, $J = 4.7$ Hz, 2H, Ser-(C3)-H₂), 3.69 (t, $J = 6.6$ Hz, 1H, Met-(C2)-H), 2.64 - 2.60 (m, 1H, Met-(C3)-H₂), 2.17 - 2.12 (m, 1H, Met-(C4)-H), 2.12 (s, 3H, COCH₃), 1.97 - 1.91 (m, 1H, Met-(C4)-H'), 1.44 (d, $J = 7.2$ Hz, 3H, Ala-(C3)-H₃).

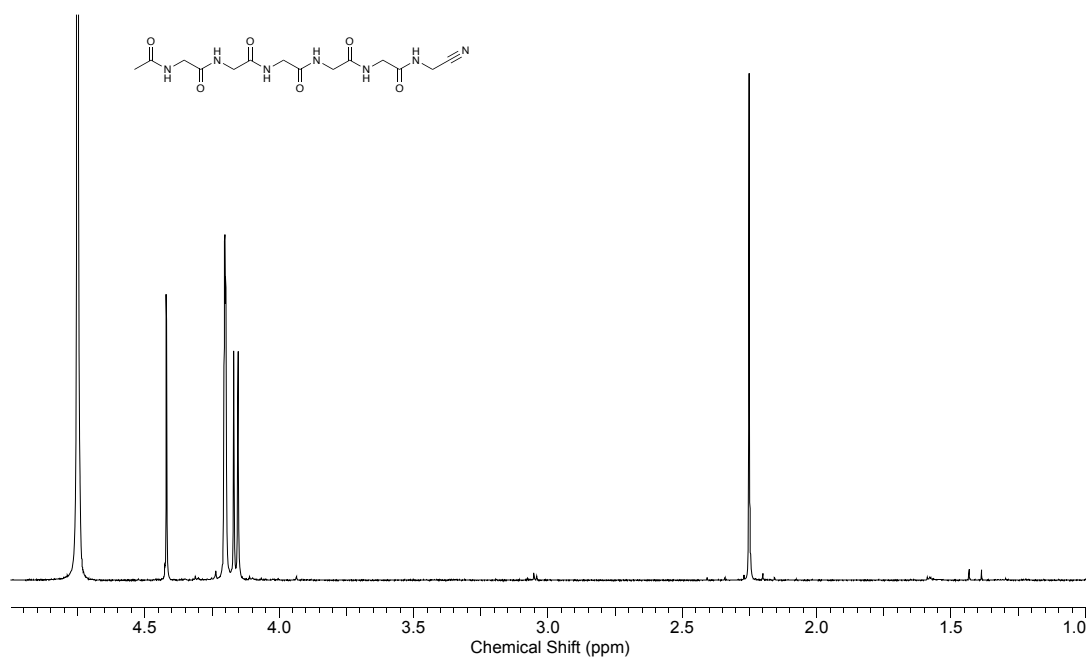
N-Acetylglycylglycylglycylglycylglycylglycine nitrile **Ac-Gly-Gly-Gly-Gly-Gly-Gly-CN**



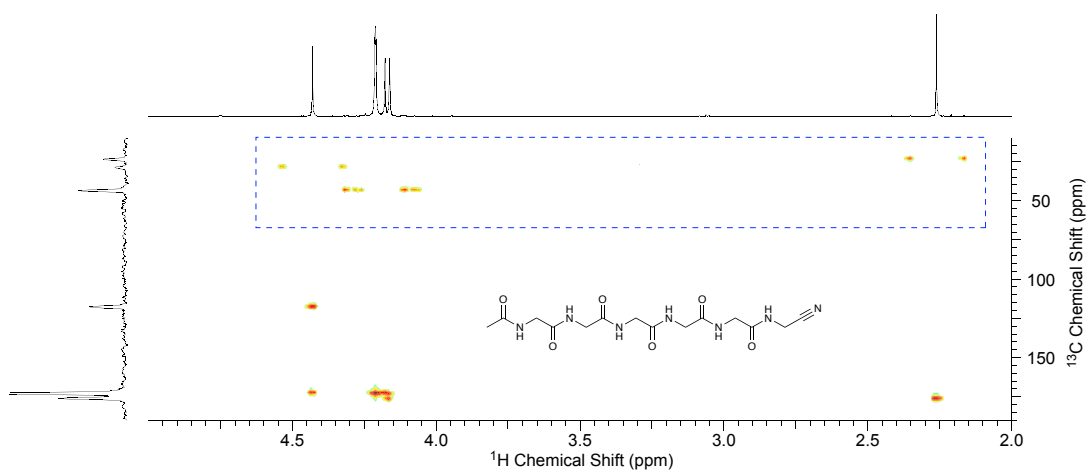
Glycylglycylglycine nitrile hydroformate **Gly₃-CN·HCO₂H** (0.025 mmol, 1 equiv.) was dissolved in D₂O (0.8 mL) and the solution was adjusted to pD 9.5 with 4 M NaOH. **Ac-Gly₃-S⁻Na⁺** (0.025 mmol, 1 equiv.) was added as a solid, the solution adjusted to pD 9.5, and the volume was adjusted to 1 mL with D₂O. Potassium hexacyanoferrate(III) (24.6 mg, 0.075 mmol) was added and the solution was stirred at room temperature for 20 min whilst maintaining the solution at pD 9.5 with 4 M NaOH. The resulting suspension was centrifuged and the supernatant removed and was analysed by ¹H NMR spectroscopy to reveal only trace amounts of **Ac-Gly₃-OH** and **Gly₃-CN** (see Supplementary Figure 225), indicating that near-quantitative conversion to **Ac-Gly₆-CN** was observed. The isolated precipitate was stirred in 8 M LiCl and analysed by ¹H (Supplementary Figure 223) and ¹H-¹³C HMBC NMR spectroscopy (Supplementary Figure 224) to reveal **Ac-Gly₆-CN** as the only peptide in the isolated precipitate. The isolated yield (>95%) was further confirmed upon conversion of **Ac-Gly₆-CN** to **Ac-Gly₁₁-OH** by ligative synthesis (see Supplementary Information page S185).

N-Acetylglycylglycylglycylglycylglycylglycine nitrile **Ac-Gly-Gly-Gly-Gly-Gly-Gly-CN** (•)

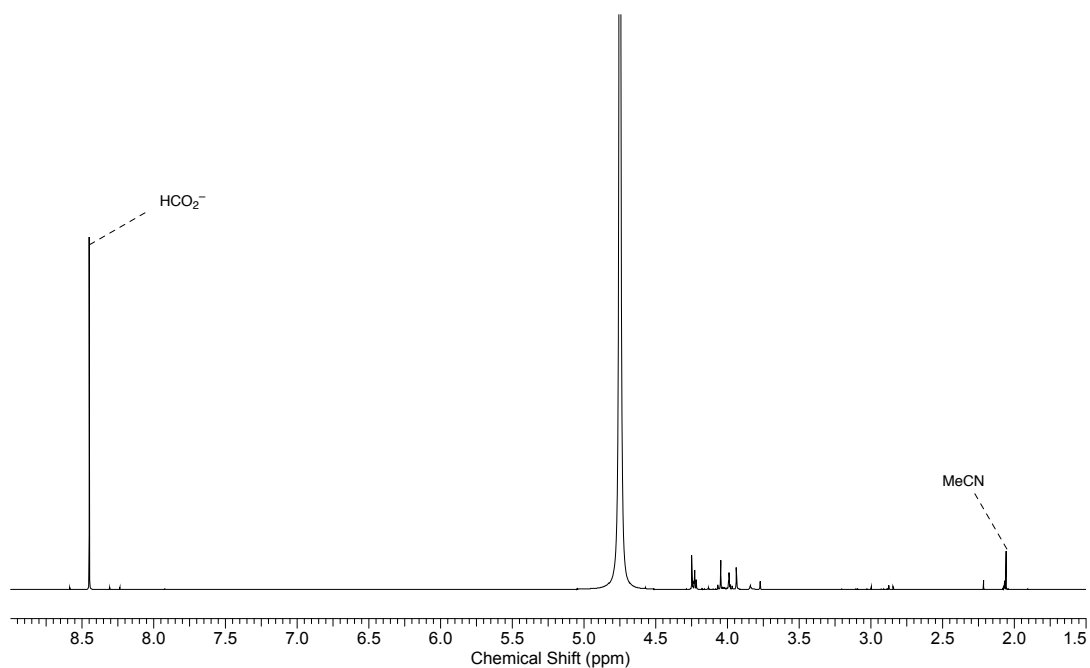
¹H NMR (700 MHz, 8 M LiCl, Supplementary Figure 223, partial assignment): δ 4.43 (s, 2H, Gly₆-(C2)-H₂), 4.21 (br. s., 6H, 3 × Gly-(C2)-H₂), 4.18 (s, 2H, Gly-(C2)-H₂), 4.16 (s, 2H, Gly₁-(C2)-H₂), 2.26 (s, 3H, COCH₃). ¹³C NMR (176 MHz, 8 M LiCl, partial assignment): δ 176.2 (COCH₃), 173.2 (Gly₁-(C1)), 172.8 (2 × Gly-(C1)), 172.7 (Gly-(C1)), 172.5 (Gly₅-(C1)), 117.9 (Gly₆-(C1)), 43.9 (Gly₁-(C2)), 43.5 (3 × Gly-(C2)), 43.4 (Gly-(C2)), 29.0 (Gly₆-(C2)), 23.6 (COCH₃).



Supplementary Figure 223. ¹H NMR (700 MHz, 8 M LiCl, noesygppr1d, 1.00–5.00 ppm) spectrum to show **Ac-Gly₆-CN**

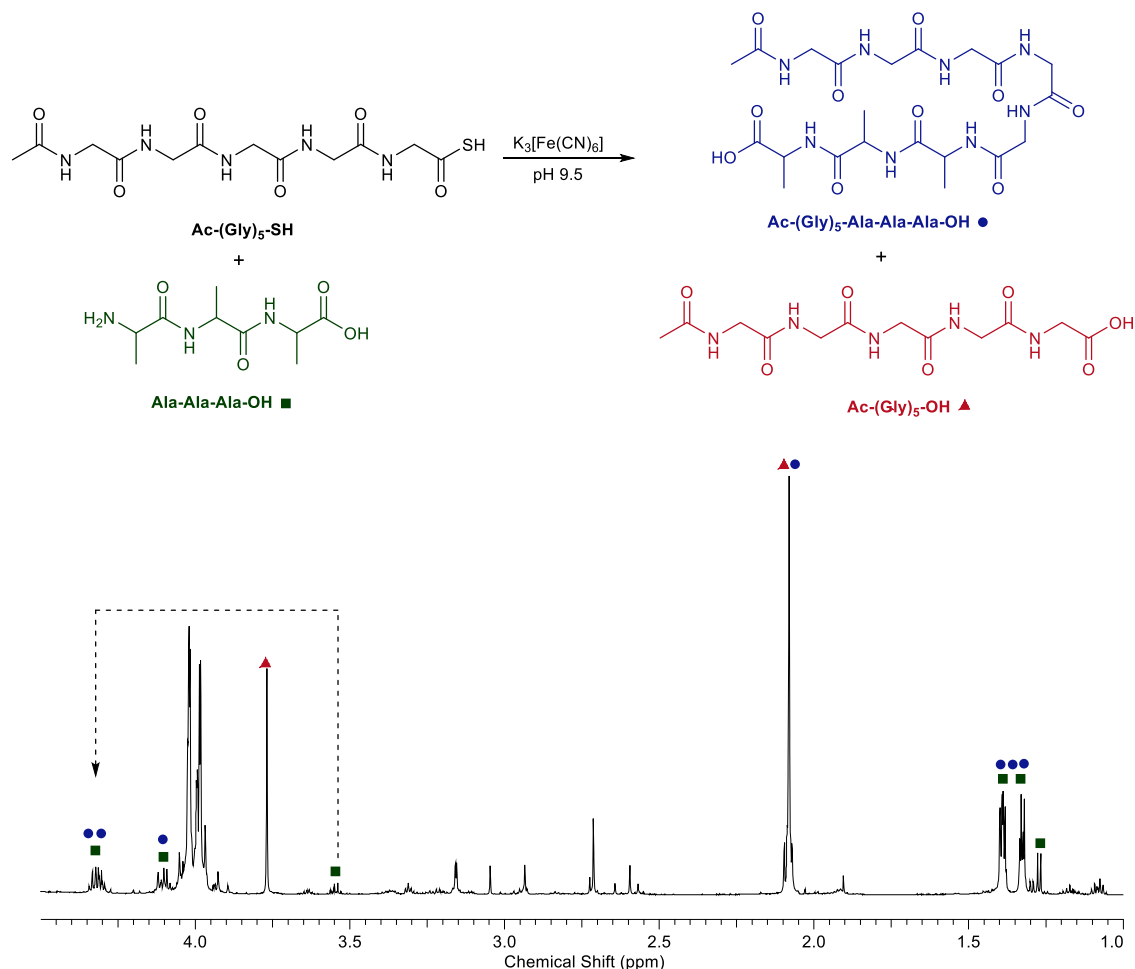


Supplementary Figure 224. ¹H–¹³C HMBC (¹H–700 MHz [2.0–5.0 ppm], ¹³C–176 MHz [10–190 ppm], H₂O) spectrum showing the diagnostic ²J_{CH} and ³J_{CH} couplings of **Ac-Gly₆-CN** and also the ¹J_{CH} correlation peaks (dotted blue box). See Supplementary Figure 223 for ¹H NMR spectrum.



Supplementary Figure 225. ^1H NMR (700 MHz, D_2O , noesygppr1d, 1.50–9.00 ppm) spectrum to show the supernatant after the reaction of **Ac-Gly₃-SH** (25 mM) with **Gly₃-CN** (25 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (75 mM) in D_2O (1 mL) at pD 9.5 and room temperature, which is enriched with the associated formate HCO_2^- counteranion of **Gly₃-CN**.

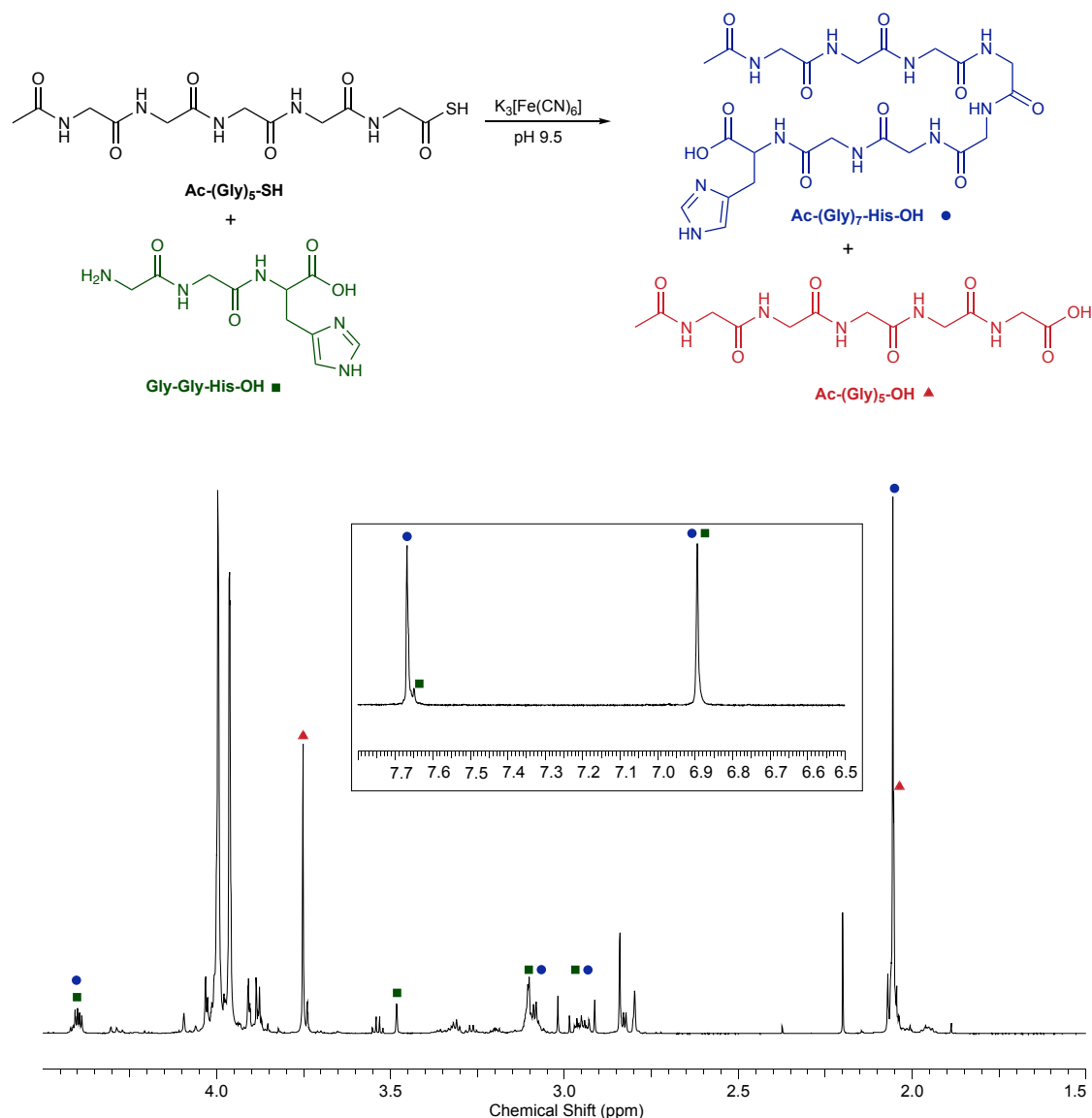
N-Acetylglycylglycylglycylglycylglycylalanylalanylalanine **Ac-Gly-Gly-Gly-Gly-Gly-Ala-Ala-Ala-OH**



Supplementary Figure 226. ¹H NMR (700 MHz, D₂O, 0.5–4.5 ppm) spectrum to show the reaction of **Ac-Gly₅-SH** (25 mM), alanylalanylalanine (**Ala-Ala-Ala-OH**; 25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D₂O (1 mL) at pD 9.5 and room temperature. The dashed lines and arrows highlight the characteristic change in chemical shift of Ala₁-(C2)-H of **Ala-Ala-Ala-OH** upon acylation by **Ac-Gly₅-SH** to give **Ac-Gly-Gly-Gly-Gly-Gly-Ala-Ala-Ala-OH** (+0.76 ppm).

¹H NMR (700 MHz, D₂O, noesygppr1d): *N*-Acetylglycylglycylglycylglycylglycylalanylalanylalanine **Ac-Gly-Gly-Gly-Gly-Gly-Ala-Ala-Ala-OH** (•) (partial assignment): δ 4.34 - 4.28 (m, 2H, Ala₆-(C2)-H, Ala₈-(C2)-H), 4.12 - 4.08 (m, 1H, Ala₇-(C2)-H), 2.08 (s, 3H, COCH₃), 1.39 (2 x d, J = 7.3 Hz, 6H, Ala₆-(C3)-H₃, Ala₈-(C3)-H₃), 1.32 (d, J = 7.3 Hz, 1H, Ala₇-(C3)-H₃). *N*-Acetylglycylglycylglycylglycylglycine **Ac-Gly₅-OH** (▲) (partial assignment): δ 3.77 (s, 2H, Gly-(C2)-H₂), 2.05, (s, 3H, COCH₃). Alanylalanylalanine **Ala-Ala-Ala-OH** (■): δ 4.33 - 4.29 (obs., 1H, Ala₃-(C2)-H), 4.14 - 4.07 (obs., 1H, Ala₂-(C2)-H), 3.54 (q, J = 7.1 Hz, 1H, Ala₁-(C2)-H), 1.40 - 1.37 (obs., 3H, Ala₃-(C3)-H₃), 1.33 (d, J = 7.2 Hz, 3H, Ala₂-(C3)-H₃), 1.27 (d, J = 7.0 Hz, 3H, Ala₁-(C3)-H₃).

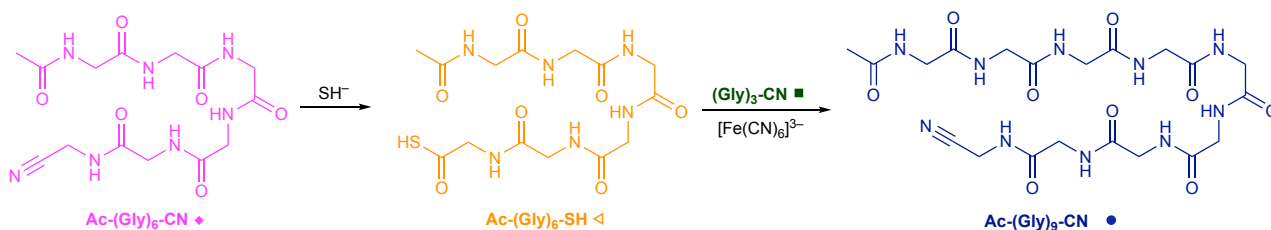
N-Acetylglycylglycylglycylglycylglycylglycylglycylhistidine **Ac-Gly-Gly-Gly-Gly-Gly-Gly-His-OH**



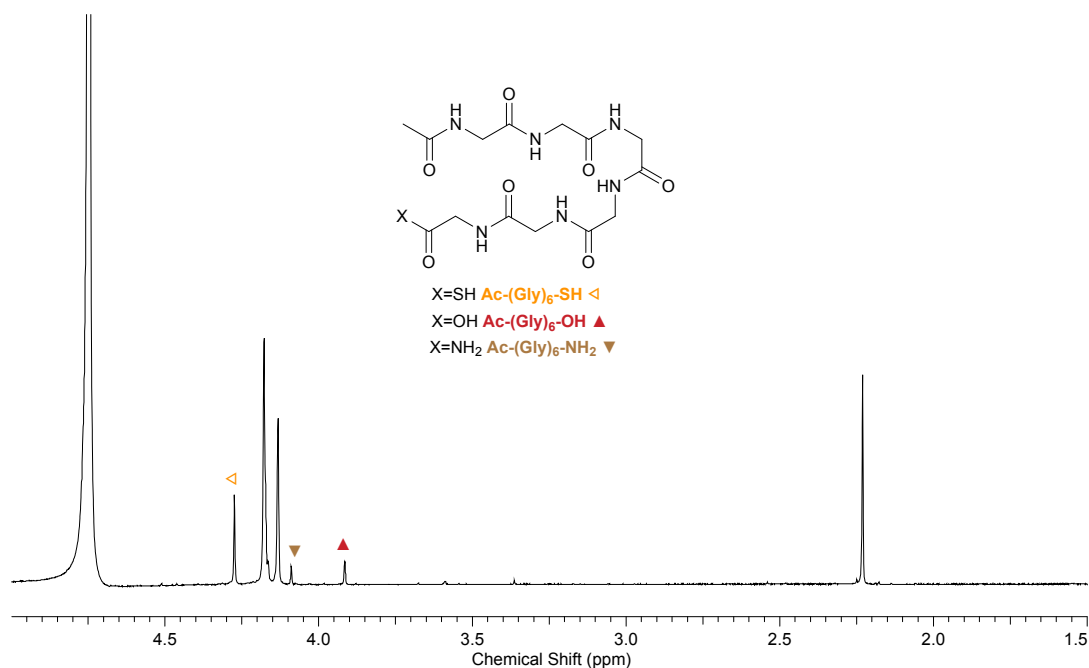
Supplementary Figure 227. ¹H NMR (700 MHz, D₂O, 1.5–4.5 ppm) spectrum to show the reaction **Ac-Gly₅-SH** (25 mM), glycylglycylhistidine (**Gly-Gly-His-OH**; 25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D₂O (1 mL) at pD 9.5 and room temperature. Inset: ¹H NMR (700 MHz, D₂O, 6.5–7.8 ppm) spectral region showing the aromatic protons.

¹H NMR (700 MHz, D₂O): *N*-Acetylglycylglycylglycylglycylglycylglycylglycylhistidine **Ac-Gly-Gly-Gly-Gly-Gly-Gly-His-OH** (•) (partial assignment): δ 7.67 (s, 1H, Ar), 6.89 (s, 1H, Ar), 4.40 (ABX, $J = 4.8, 8.8$ Hz, 1H, His-(C2)-H), 3.09 (dd, $J = 4.8, 15.1$ Hz, 1H, His-(C3)-H), 2.95 (dd, $J = 8.8, 15.1$ Hz, 1H, His-(C3)-H'), 2.06 (s, 3H, COCH₃). *N*-Acetylglycylglycylglycylglycylglycylglycine **Ac-Gly₅-OH** (▲) (partial assignment): δ 3.75 (s, 2H, Gly-(C2)-H₂), 2.06 (s, 3H, COCH₃). Glycylglycylhistidine **Gly-Gly-His-OH** (■) (partial assignment): δ 7.65 (s, 1H, Ar), 6.89 (s, 1H, Ar), 4.43 – 4.40 (obs, 1H, His-(C2)-H), 3.48 (obs, 2H, Gly₁-(C2)-H₂), 3.11 – 3.06 (obs, 1H, His-(C3)-H), 2.98 – 2.92 (obs, 1H, His-(C3)-H').

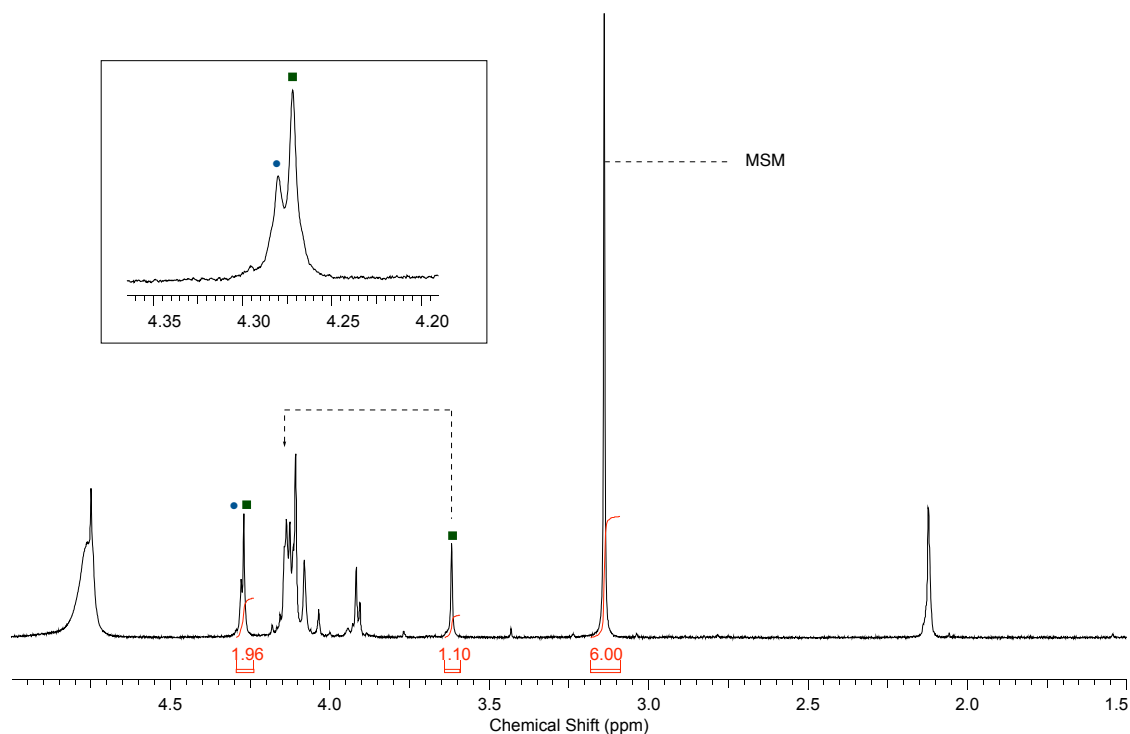
Ac-Gly₉-CN



Ac-Gly₆-CN (0.025 mmol, 1 equiv., the precipitate isolated from the quantitative ligation of **Ac-Gly₃-SH** and **Gly₃-CN**, see Supplementary Information page S176) was suspended in 8 M LiCl (1 mL) and NaSH·xH₂O (28 mg, 0.25 mmol) was added. The solution was adjusted to pH 9.0 with 6 M HCl. The reaction became homogenous within 5 min and was stirred at room temperature. The conversion of **Ac-Gly₆-CN** to **Ac-Gly₆-SH** was monitored by ¹H NMR at room temperature. **Ac-Gly₆-SNH₂** (26%), **Ac-Gly₆-SH** (53%), **Ac-Gly₆-NH₂** (13%), **Ac-Gly₆-OH** (6%) were observed after 17 h. The reaction was then heated at 60 °C until complete conversion of **Ac-Gly₆-SNH₂** to **Ac-Gly₆-SH** was observed (4 h, Supplementary Figure 228), giving **Ac-Gly₆-SH** (72%), **Ac-Gly₆-NH₂** (13%) and **Ac-Gly₆-OH** (19%). **Gly₃-CN** (0.025 mmol) and D₂O (1 mL (total volume = 2 mL; [**Ac-Gly₆-SH**] = 9.0 mM, [**Gly₃-CN**] = 12.5 mM)) was added and the solution adjusted to pH 9.5 with 6 M HCl. Potassium hexacyanoferrate(III) (329 mg, 1.0 mmol) was added and the solution was stirred rapidly at room temperature for 20 min whilst maintaining the solution at pH 9.5 with 4 M NaOH. The resulting suspension was centrifuged, and the supernatant analysed by ¹H NMR spectroscopy using MSM (0.025 mmol) as an internal NMR reference standard (Supplementary Figure 229). The yield of **Ac-Gly₉-CN** formation (43%) is over four steps from **Ac-Gly₃-SH** (**Ac-Gly₃-SH**+**Gly₃-CN**→**Ac-Gly₆-CN**→**Ac-Gly₆-SNH₂**→**Ac-Gly₆-SH** + **Gly₃-CN**→**Ac-Gly₉-CN**).



Supplementary Figure 228. ¹H NMR (700 MHz, 8 M LiCl, noesygprr1d, 1.5–5.0 ppm) spectrum to show the reaction of **Ac-Gly₆-CN** (25 mM) with NaSH (250 mM) at pH 9.0 after 17 h at room temperature followed by heating at 60 °C for 4 h to give a relative mixture of **Ac-Gly₆-SH** (72%), **Ac-Gly₆-OH** (19%) and **Ac-Gly₆-NH₂** (13%).

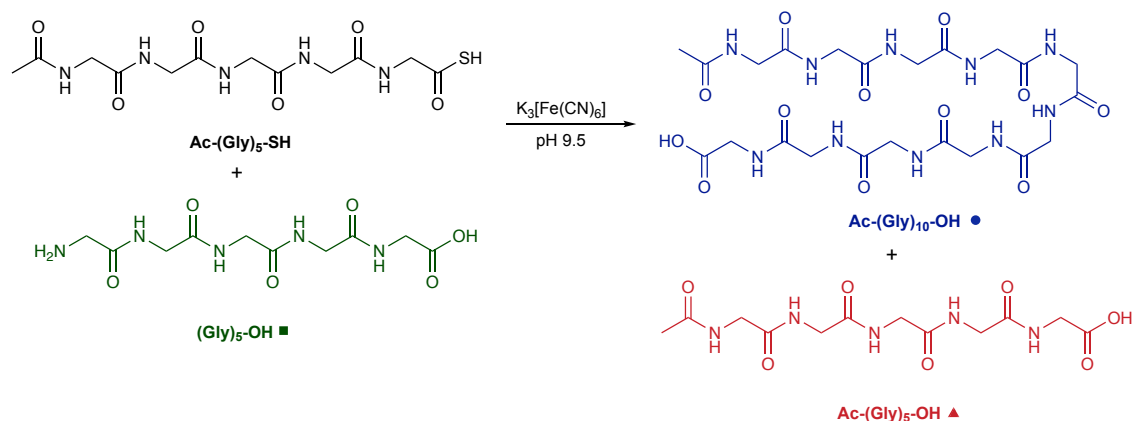


Supplementary Figure 229. ^1H NMR (600 MHz, 4 M LiCl/D₂O (1:1), noesygppr1d, 1.5–5.0 ppm) spectrum to show the supernatant after reaction of **Ac-Gly₆-SH** (9.0 mM, ◊), **Gly₃-CN** (12.5 mM, ■) and K₃[Fe(CN)₆] (1 mmol) at pH 9.5 and room temperature. A marked depletion of **Gly₃-CN** (12.5 mM→7.13 mM) was observed and quantified by integration against MSM as the internal reference NMR standard (12.5 mM). **Gly₃-CN** (5.38 mM) had reacted with **Ac-Gly₆-SH** to form **Ac-Gly₉-CN** (•). The dashed lines and arrows highlight the characteristic change in chemical shift of Gly₁-(C2)-H₂ of **Gly₃-CN** upon acylation by **Ac-Gly₆-SH** to give **Ac-Gly₉-CN** ($\approx +0.58$ ppm), which is in accordance to the chemical shift change of Gly₁-(C2)-H₂ of **Gly₃-CN** upon acylation by **Ac-Gly₃-SH** to give **Ac-Gly₆-CN** (δ_{H} Gly₄-(C2)-H = 4.1–4.21 ppm, see Supplementary Figure 223). Inset: ^1H NMR (600 MHz, 4 M LiCl/D₂O (1:1), noesygppr1d, 4.20–4.36 ppm) spectral region showing the two expected CH₂-CN resonances accounting for residual **Gly₃-CN** (Gly₃-(C2)-H₂, ■) and the ligation product **Ac-Gly₉-CN** (Gly₆-(C2)-H₂, •). No signals in this region for **Ac-Gly₆-SH** (◊) were observed by ^1H and ^1H - ^{13}C HMBC analysis.

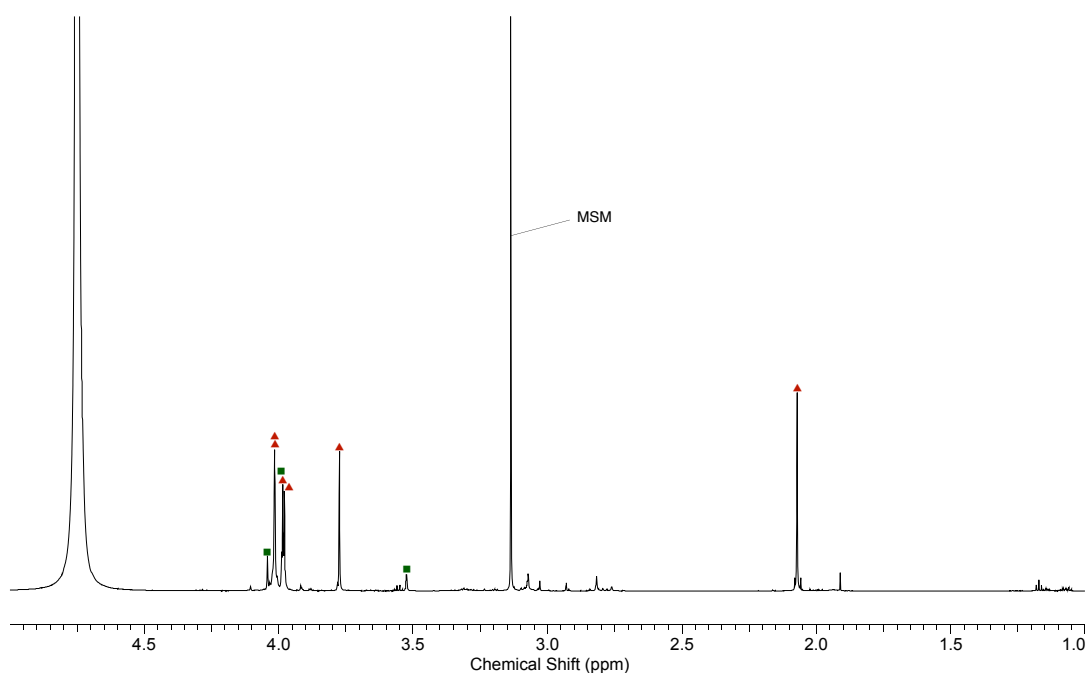
^1H NMR (700 MHz, 8 M LiCl, noesygppr1d, Supplementary Figure 228) (partial assignment): **Ac-Gly₆-SH** (◊) δ 4.28 (s, 2H, Gly₆-(C2)-H₂). **Ac-Gly₆-NH₂** (▼) δ 4.09 (s, 2H, Gly₆-(C2)-H₂). **Ac-Gly₆-OH** (▲) δ 3.91 (s, 2H, Gly₆-(C2)-H₂).

^1H NMR (600 MHz, 4 M LiCl/D₂O (1:1), noesygppr1d, Supplementary Figure 229) (partial assignment): **Ac-Gly₉-CN** (•) δ 4.28 (s, 2H, Gly₆-(C2)-H₂). **Gly₃-CN** (■) δ 4.28 (s, 2H, Gly₃-(C2)-H₂), 3.62 (s, 2H, Gly₁-(C2)-H₂).

Ac-Gly₁₀-OH

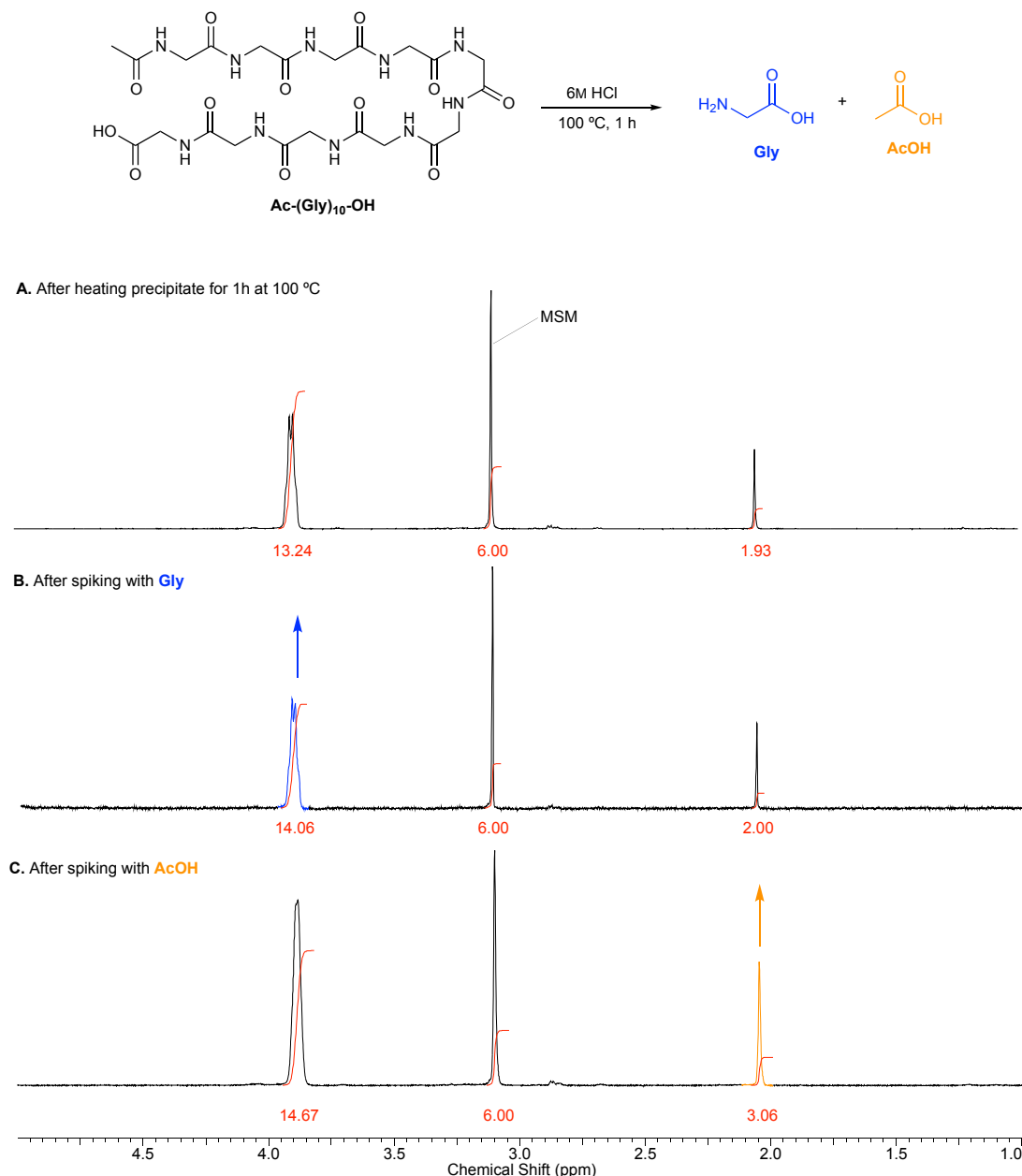


Pentaglycine **Gly₅-OH** (7.58 mg, 0.025 mmol, 1 equiv.) was suspended in H₂O (0.8 mL) and the solution adjusted to pH 9.5 with 4 M NaOH then stirred until a homogeneous solution was observed. **Ac-Gly₅-S⁻Na⁺** (0.025 mmol) was added as a solid, the solution adjusted to pH 9.5, and the volume was adjusted to 1 mL with H₂O. Potassium hexacyanoferrate(III) (24.6 mg, 0.075 mmol) was added and the solution was stirred at room temperature for 20 min whilst maintaining the solution at pH 9.5 with 4 M NaOH. The resulting suspension was centrifuged and the supernatant removed. The precipitate was re-suspended in water (1 mL), the solution adjusted to pH 10.5 whilst stirring. The suspension was centrifuged to ensure complete removal of residual **Ac-Gly₅-OH** and **Gly₅-OH**. This alkaline extraction was repeated three times. The combined washings (supernatants) were then analysed and quantified by ¹H NMR spectroscopy to reveal only residual **Ac-Gly₅-OH** and unreacted **Gly₅-OH** (21%), indicating that conversion to **Ac-Gly₁₀-OH** was 79% (Supplementary Figure 230). The isolated yield of **Ac-Gly₁₀-OH** was then quantified and confirmed by acid hydrolysis. The isolated precipitate was suspended in 6 M HCl (1 mL) and heated at 100 °C. The resultant homogenous reaction was cooled to room temperature after 60 min and analysed and quantified by ¹H NMR spectroscopy (Supplementary Figure 231a). Glycine **Gly** (66%) and acetic acid **AcOH** were the only observed products, and were quantified using MSM as an internal standard. The observed ratio of **Gly**/**AcOH** (10:1) was exactly that expected for exhaustive hydrolysis of **Ac-Gly₁₀-OH**. The presence of **Gly** and **AcOH** were both confirmed by spiking with authentic samples (Supplementary Figure 231b and c).



Supplementary Figure 230. ¹H NMR (700 MHz, D₂O, 1.0–5.0 ppm) spectrum to show the supernatant after reaction of **Ac-Gly₅-SH** (25 mM), **Gly₅-OH** (25 mM) and $K_3[Fe(CN)_6]$ (75 mM) in D₂O (1 mL) at pH 9.5 and room temperature. A marked depletion of **Gly₅-**

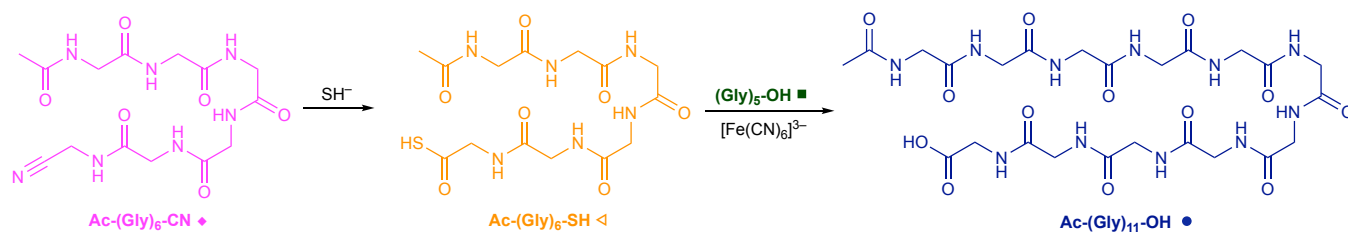
OH (5.25 mM) was observed and quantified by integration against MSM as the internal reference NMR standard. **Gly₅-OH** (19.75 mM) had reacted with **Ac-Gly₅-SH** to form **Ac-Gly₁₀-OH**, which had precipitated from the reaction (See Supplementary Figure 231).



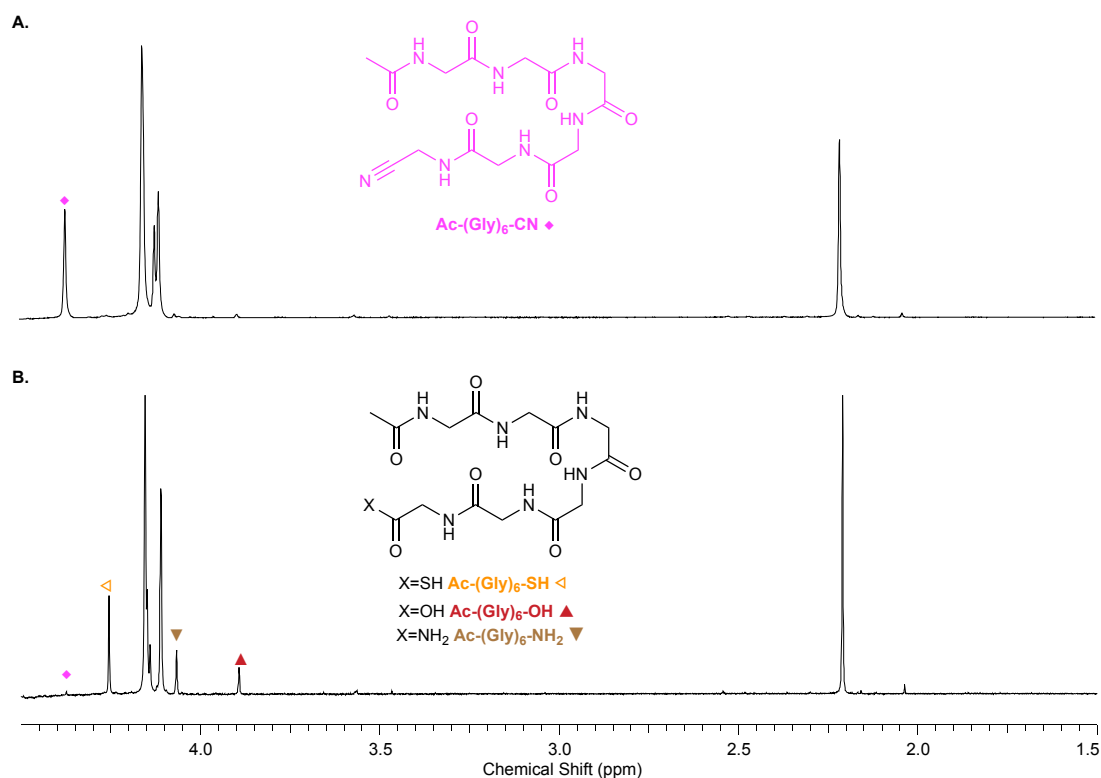
Supplementary Figure 231. ¹H NMR (400 MHz, 6 M HCl, noesygppr1d, 1.0–5.0 ppm) spectra to show the hydrolysis of **Ac-Gly₁₀-OH** precipitate obtained from the stoichiometric reaction of **Ac-Gly₅-SH** (25 mM), **Gly₅-OH** (25 mM) and K₃[Fe(CN)₆] (75 mM) in D₂O (1 mL) at pD 9.5 and room temperature. Quantification was carried out using MSM as the internal NMR reference standard ($\delta_H = 3.10$ ppm) after heating the isolated precipitate at 100 °C for 1 h in 6 M HCl. **(A)** Only glycine (**Gly**) and acetic acid (**AcOH**) were observed, and **Gly/AcOH** were observed in the expected 10:1 ratio that is obtained by exhaustive hydrolysis of **Ac-Gly₁₀-OH**. **(B)** Spiking with authentic **Gly** resulted in enhancement of the Gly-(C2)-H₂ resonance at 3.90 ppm ($\uparrow$). **(C)** Spiking with authentic **AcOH** resulted in enhancement of the CH₃CO₂H resonance at 2.05 ppm ($\uparrow$). See Supplementary Figure 230 for the supernatant from which **Ac-Gly₁₀-OH** precipitated.

¹H NMR (700 MHz, D₂O, Supplementary Figure 230): **Ac-Gly₅-OH** ($\blacktriangle$) δ 4.02 (s, 2H, Gly-(C2)-H₂), 4.01 (s, 2H, Gly-(C2)-H₂), 3.98 (s, 2H, Gly-(C2)-H₂), 3.98 (s, 2H, Gly-(C2)-H₂), 3.77 (s, 2H, Gly₅-(C2)-H₂), 2.07 (s, 3H, COCH₃). **Glycylglycylglycylglycylglycine Gly₅-OH** (partial assignment) ($\blacksquare$) δ 4.04 (s, 4H, 2 $\times$ Gly-(C2)-H₂), 3.99 (s, 2H, Gly-(C2)-H₂), 3.52 (s, 2H, Gly₁-(C2)-H₂). ¹H NMR (400 MHz, 6 M HCl, noesygppr1d, Supplementary Figure 231a): **Glycine Gly** δ 3.90 (q, $J = 5.6$ Hz, 2H, Gly-(C2)-H₂). **Acetic acid** δ 2.05 (s, 3H, CH₃CO₂H).

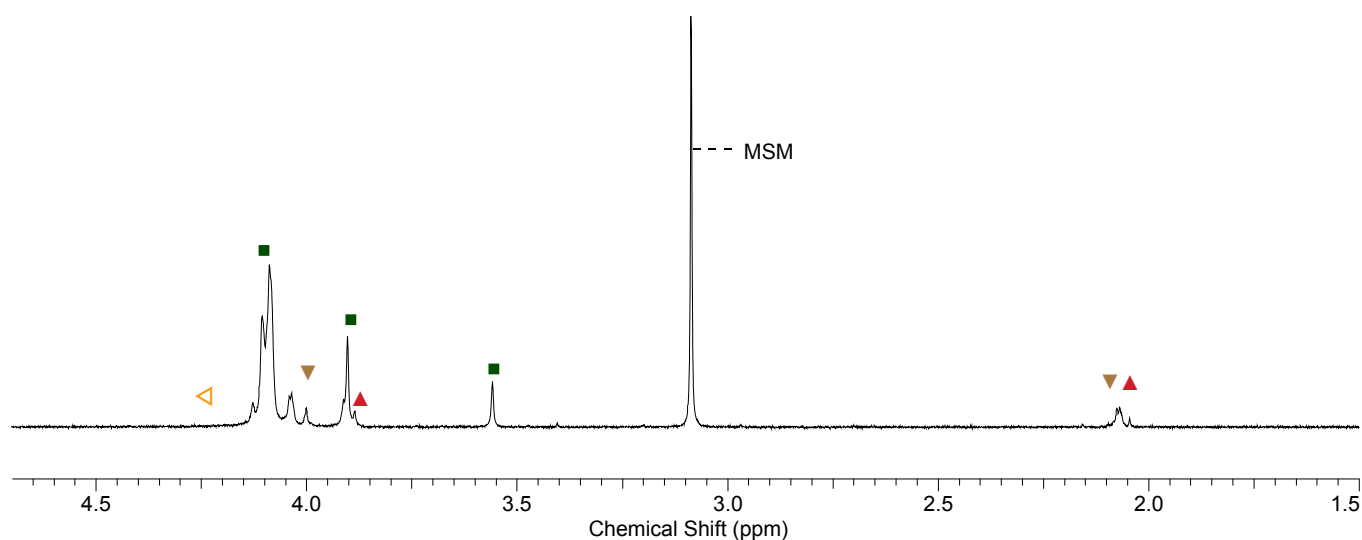
Ac-Gly₁₁-OH



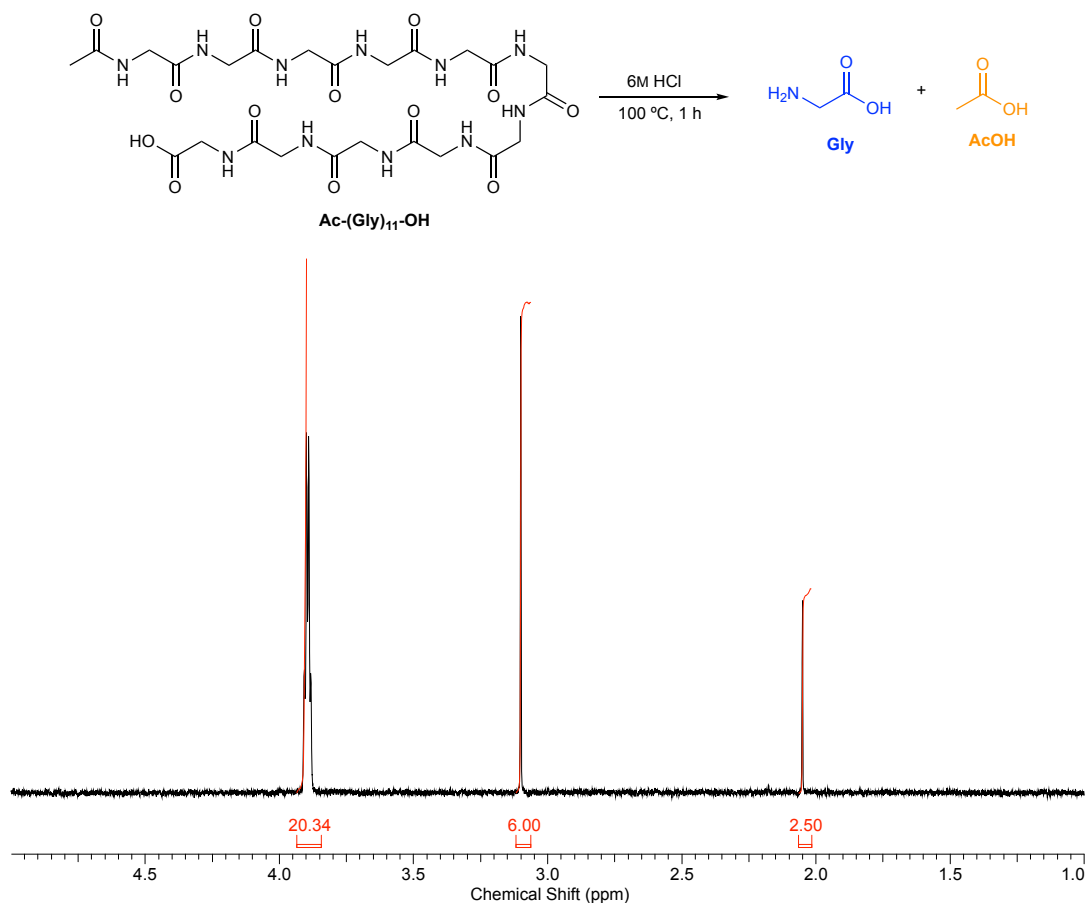
Ac-Gly₆-CN (0.025 mmol, 1 equiv., which was isolated as a precipitate from the quantitative ligation of **Ac-Gly₃-SH** and **Gly₃-CN**, see Supplementary Information page 176) was suspended in 8 M LiCl (2 mL) and NaSH·xH₂O (28 mg, 0.25 mmol) was added. The solution was adjusted to pH 9.0 with 6 M HCl. The reaction became homogenous within 5 min and was stirred at room temperature (Supplementary Figure 232a). The conversion of **Ac-Gly₆-CN** to **Ac-Gly₆-SH** was monitored by ¹H NMR at room temperature. The reaction composition was **Ac-Gly₆-SNH₂** (31%), **Ac-Gly₆-SH** (49%), **Ac-Gly₆-NH₂** (16%), **Ac-Gly₆-OH** (4%) after 2 d at room temperature. The reaction was then heated at 60 °C until complete conversion of **Ac-Gly₆-SNH₂** was observed (4 h, Supplementary Figure 232b). **Gly₅-OH** (7.58 mg, 0.025 mmol) and MSM (25 mM, internal NMR standard) were added. The reaction mixture was quantified as **Ac-Gly₆-SH** (51%, 6.25 mM), **Ac-Gly₆-OH** (16%), **Ac-Gly₆-NH₂** (33%) to further confirm the quantitative formation of **Ac-Gly₆-CN** from **Ac-Gly₃-SH** and **Gly₃-CN** in the preceding ligation reaction (See Supplementary Information page S172). The crude **Ac-Gly₆-SH** mixture was diluted with D₂O (2 mL, total volume = 4 mL; [**Ac-Gly₆-SH**] = 3.13 mM, [**Gly₅-OH**] = 6.25 mM) and the solution was adjusted to pH 9.5 with 6 M HCl. Potassium hexacyanoferrate(III) (329 mg, 1.0 mmol) was added and the solution was stirred rapidly at room temperature for 20 min whilst maintaining the solution at pH 9.5 with 4 M NaOH. The resulting suspension was centrifuged, and the supernatant analysed by ¹H NMR spectroscopy to reveal the depletion of **Gly₅-OH** (6.25 mM→3.13 mM) (Supplementary Figure 233). The isolated **Ac-Gly₁₁-OH** precipitate was then quantified by acid hydrolysis. The precipitate was suspended in 6 M HCl (1 mL) and heated at 100 °C. The resultant homogenous reaction was cooled to room temperature after 60 min and analysed by ¹H NMR spectroscopy (Supplementary Figure 234). Glycine **Gly** (92%) and acetic acid **AcOH** were the only observed products, and were quantified using MSM as an internal standard.



Supplementary Figure 232. ^1H NMR (700 MHz, 8 M LiCl, noesygppr1d, 1.5–5.0 ppm) spectra to show the reaction of **Ac-Gly₆-CN** (25 mM) with NaSH (250 mM) at pH 9.0. **A.** After 10 min. **B.** After 2 d at room temperature followed by heating at 60 °C for 4 h to give a mixture of **Ac-Gly₆-SH** (51%), **Ac-Gly₆-OH** (16%) and **Ac-Gly₆-NH₂** (33%).



Supplementary Figure 233. ^1H NMR (600 MHz, 4 M LiCl/D₂O (1:1), noesygppr1d, 1.5–4.7 ppm) spectrum to show the supernatant after reaction of **Ac-Gly₆-SH** (3.13 mM, $\triangleleft$), **Gly₅-OH** (6.25 mM, $\blacksquare$) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (1 mmol) at pH 9.5 and room temperature. A marked depletion of **Gly₅-OH** was observed and quantified by integration against MSM as the internal reference NMR standard. **Gly₅-OH** (3.13 mM) had reacted with **Ac-Gly₆-SH** to form **Ac-Gly₁₁-OH**, which had precipitated from the reaction (See Supplementary Figure 234). Residual **Ac-Gly₆-OH** ($\blacktriangle$) and **Ac-Gly₆-NH₂** ($\blacktriangledown$) is present from the preceding formation of **Ac-Gly₆-SH** ($\triangleleft$, Supplementary Figure 232b). **Ac-Gly₆-SH** has been completely consumed and the region where its Gly₆-(C2)-H₂ resonance is expected is labelled ($\triangleleft$).

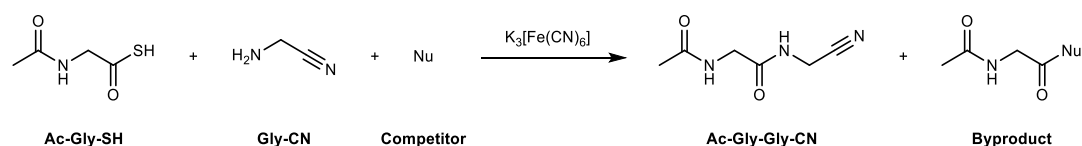


Supplementary Figure 234. ¹H NMR (700 MHz, 6 M HCl, 1.0–5.0 ppm) spectrum to show the hydrolysis of **Ac-Gly₁₁-OH** precipitate obtained from the reaction of **Ac-Gly₆-SH** (3.13 mM), **Gly₅-OH** (6.25 mM) and K₃[Fe(CN)₆] (1 mmol) at pH 9.5 and room temperature. Quantification was carried out using MSM as the internal NMR reference standard ($\delta_{\text{H}} = 3.10$ ppm) after heating the isolated precipitate at 100 °C for 1 h in 6 M HCl. Only glycine (**Gly**) and acetic acid (**AcOH**) were observed. See Supplementary Figure 233 for the supernatant from which **Ac-Gly₁₁-OH** precipitated.

¹H NMR (700 MHz, 8 M LiCl, noesygppr1d, Supplementary Figure 232b) (partial assignment): **Ac-Gly₆-SH** (◁) δ 4.27 (s, 2H, Gly₆-(C2)-H₂). **Ac-Gly₆-NH₂** (▼) δ 4.09 (s, 2H, Gly₆-(C2)-H₂). **Ac-Gly₆-OH** (▲) δ 3.91 (s, 2H, Gly₆-(C2)-H₂).

¹H NMR (600 MHz, 4 M LiCl/D₂O (1:1), noesygppr1d, Supplementary Figure 233) (partial assignment): **Ac-Gly₆-NH₂** (▼) δ 4.00 (s, 2H, Gly₆-(C2)-H₂). **Ac-Gly₆-OH** (▲) δ 3.89 (s, 2H, Gly₆-(C2)-H₂). **Gly₅-OH** (■) δ 3.56 (s, 2H, Gly₁-(C2)-H₂). ¹H NMR (400 MHz, 6 M HCl, noesygppr1d, Supplementary Figure 234): **Glycine Gly** δ 3.90 (q, $J = 5.4$ Hz, 2H, Gly-(C2)-H₂). **Acetic acid** δ 2.05 (s, 3H, CH₃CO₂H).

Coupling in the presence of prebiotic competitors

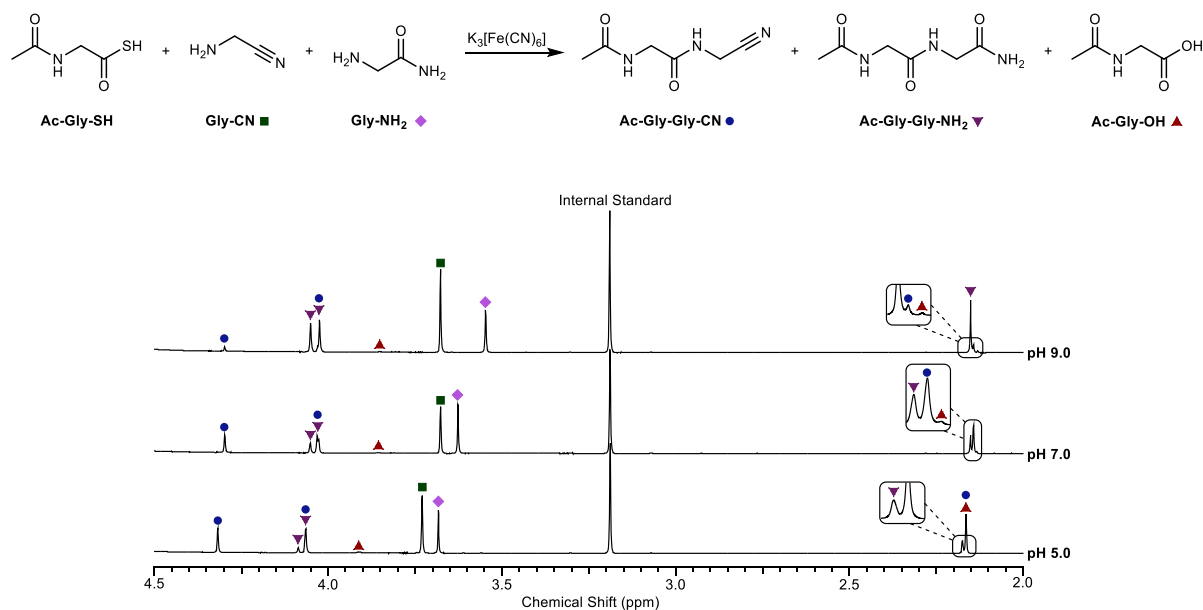


2-Acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM, 0.1 mmol) and methylsulfonylmethane (50 mM, 0.1 mmol) were dissolved in H₂O/D₂O (98:2, 1 mL). Glycine nitrile **Gly-CN** (100 mM, 0.2 mmol) and the specified competitor (100 mM, 0.2 mmol) were added, and the solution was adjusted to the specified pH value (5.0, 7.0 or 9.0). The volume of the reaction was adjusted to 2 mL with H₂O/D₂O (98:2). Potassium hexacyanoferrate(III) (98.8 mg, 0.3 mmol) was added and the reaction was stirred. After 10 min, the solution was adjusted to pH 9.0, centrifuged, and the supernatant was analysed by NMR spectroscopy. The presence of **Ac-Gly-Gly-CN** was confirmed by ¹H-³C HMBC NMR and signal amplification by spiking, or comparison with pure synthetic standards. Yields were determined by integration using methylsulfonylmethane as an internal NMR standard. Yields are given in Supplementary Table 16.

Competitor	pH	Ac-Gly-Gly-CN (%)	Byproduct (%)
 Gly-NH₂	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
NH₃	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
 	5.0	90	-
	7.0	98	-
	9.0	91	5
 	5.0	77	-
	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	-
	7.0	89	-
	9.0	92	-
 	5.0	73	-
	7.0	83	-
	9.0	90	-
 	5.0	72	-
	7.0	91	-
	9.0	84	-

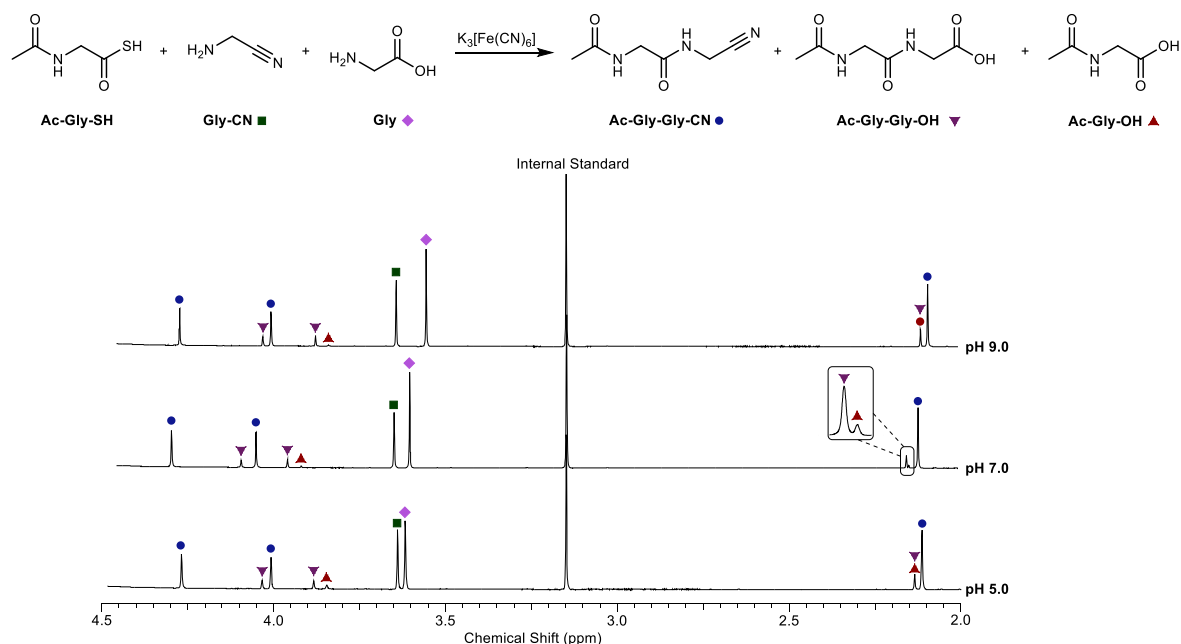
Supplementary Table 16. The formation of of *N*-acetylglycylglycine nitrile **Ac-Gly-Gly-CN** from the oxidative ligation of 2-acetamidoethanethioic S-acid **Ac-Gly-SH** (50 mM) and glycine nitrile **Gly-CN** (100 mM) with $K_3[Fe(CN)_6]$ (150 mM) in the presence of a stoichiometric competitor (100 mM) after 20 min.

Competition between glycine nitrile and glycineamide



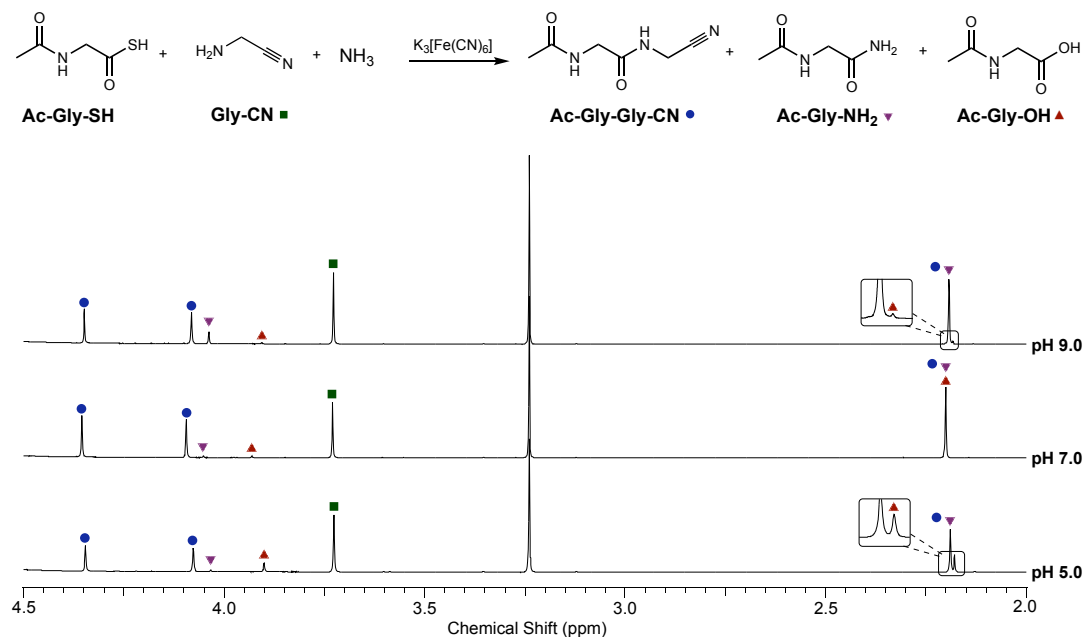
Supplementary Figure 235. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 2.00–4.50 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**; 100 mM), glycineamide (**Gly-NH₂**; 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Competition between glycine nitrile and glycine



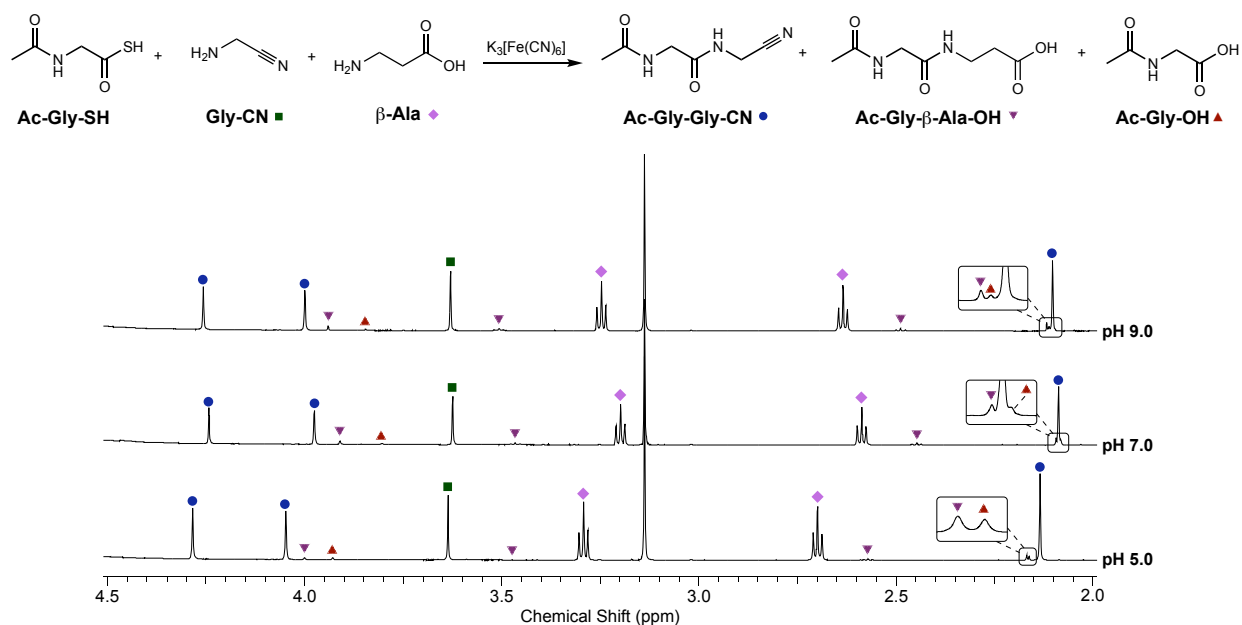
Supplementary Figure 236. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 2.00–4.50 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**; 100 mM), glycine (**Gly**, 100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Competition between glycine nitrile and ammonia



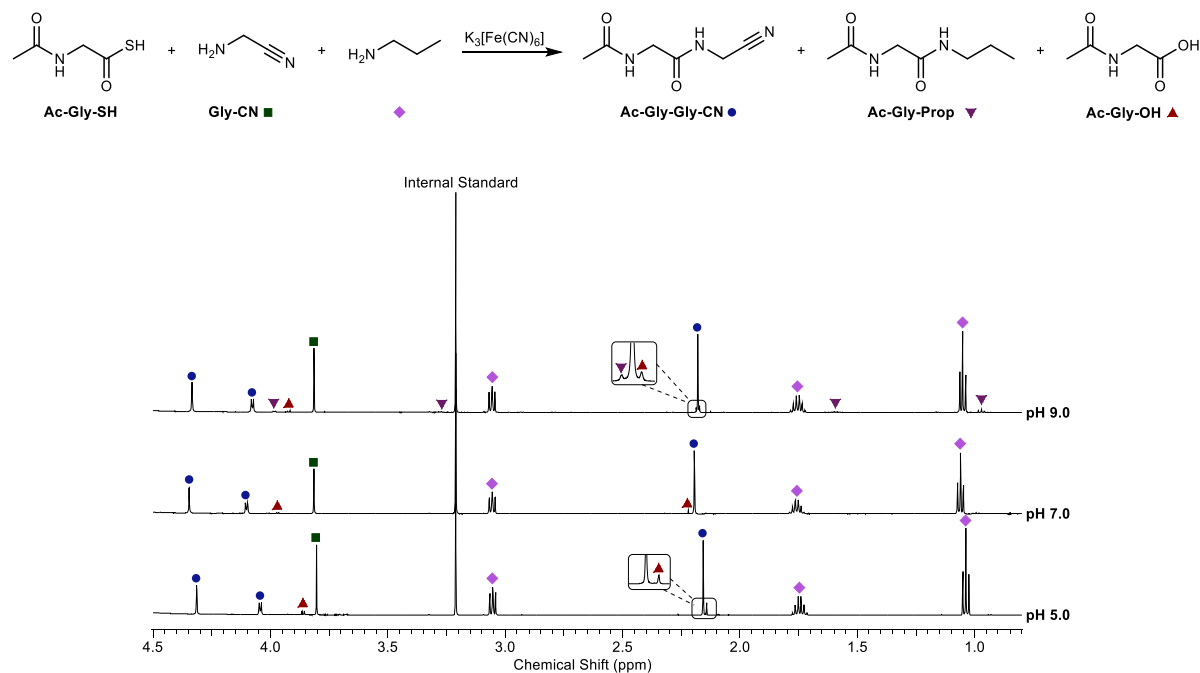
Supplementary Figure 237. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 2.00–4.50 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM), ammonia (100 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Competition between glycine nitrile and β-alanine



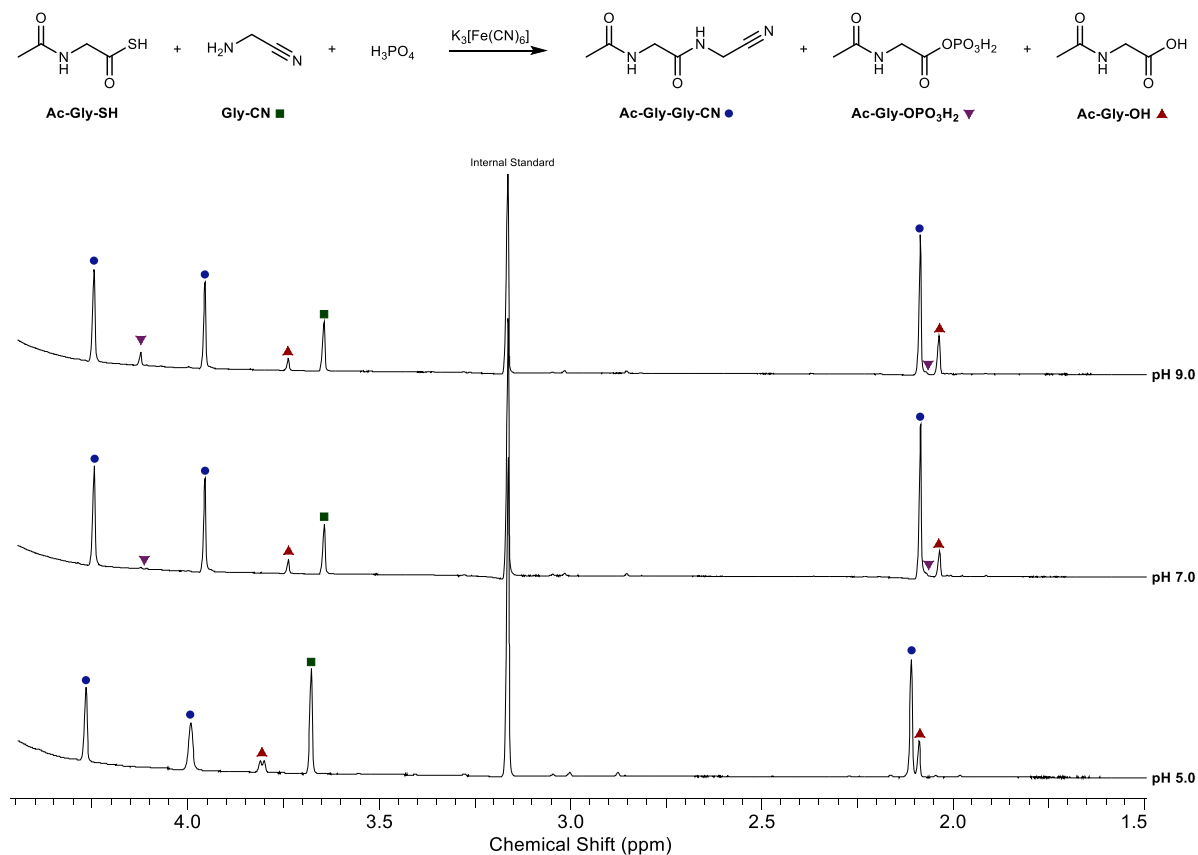
Supplementary Figure 238. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 2.00–4.50 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM), β-alanine (**β-Ala**, 100 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Competition between glycine nitrile and *n*-propylamine



Supplementary Figure 239. 1H NMR (600 MHz, H_2O/D_2O 98:2, 0.80–4.50 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM), *n*-propylamine (100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

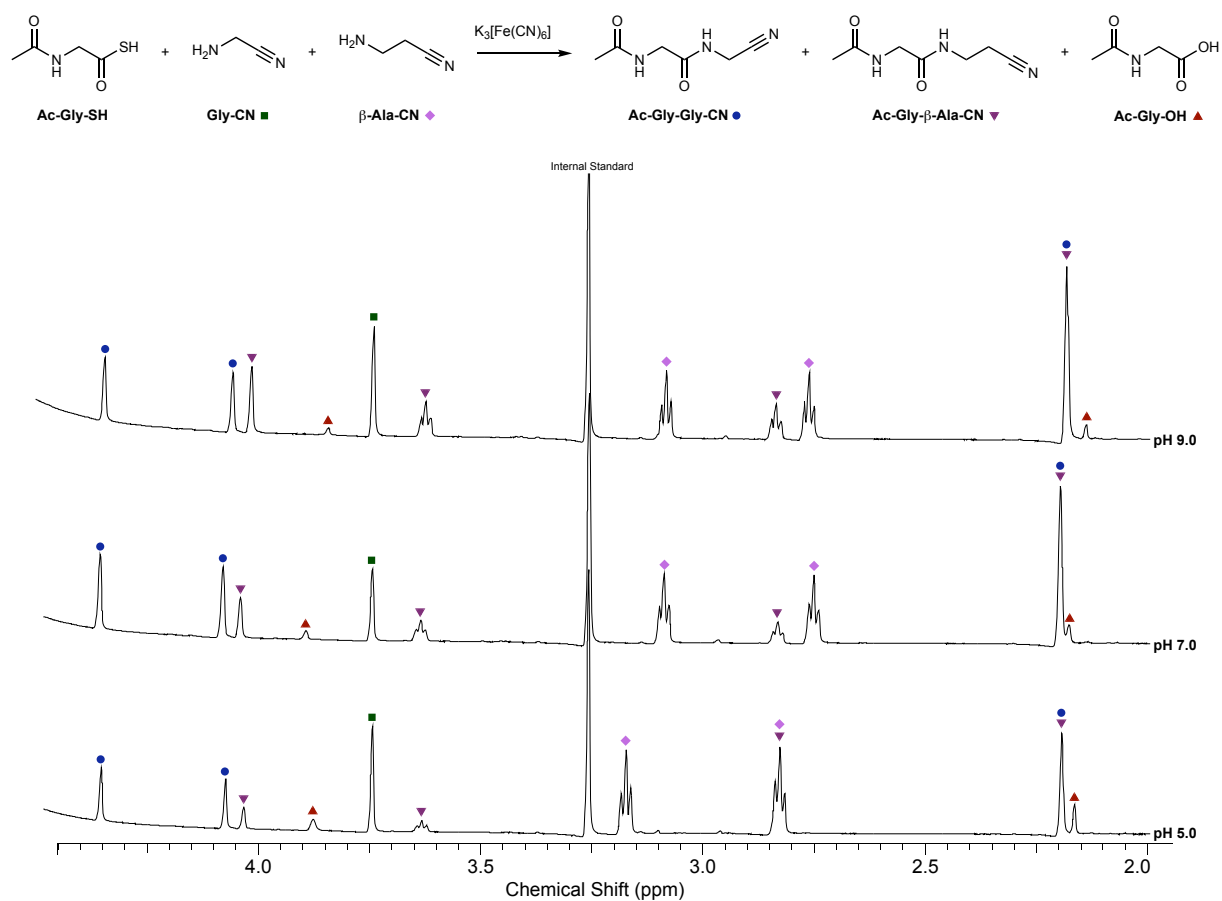
Competition between glycine nitrile and phosphate



Supplementary Figure 240. 1H NMR (600 MHz, H_2O/D_2O 98:2, 1.50–4.45 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM), phosphate (100 mM) and $K_3[Fe(CN)_6]$ (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

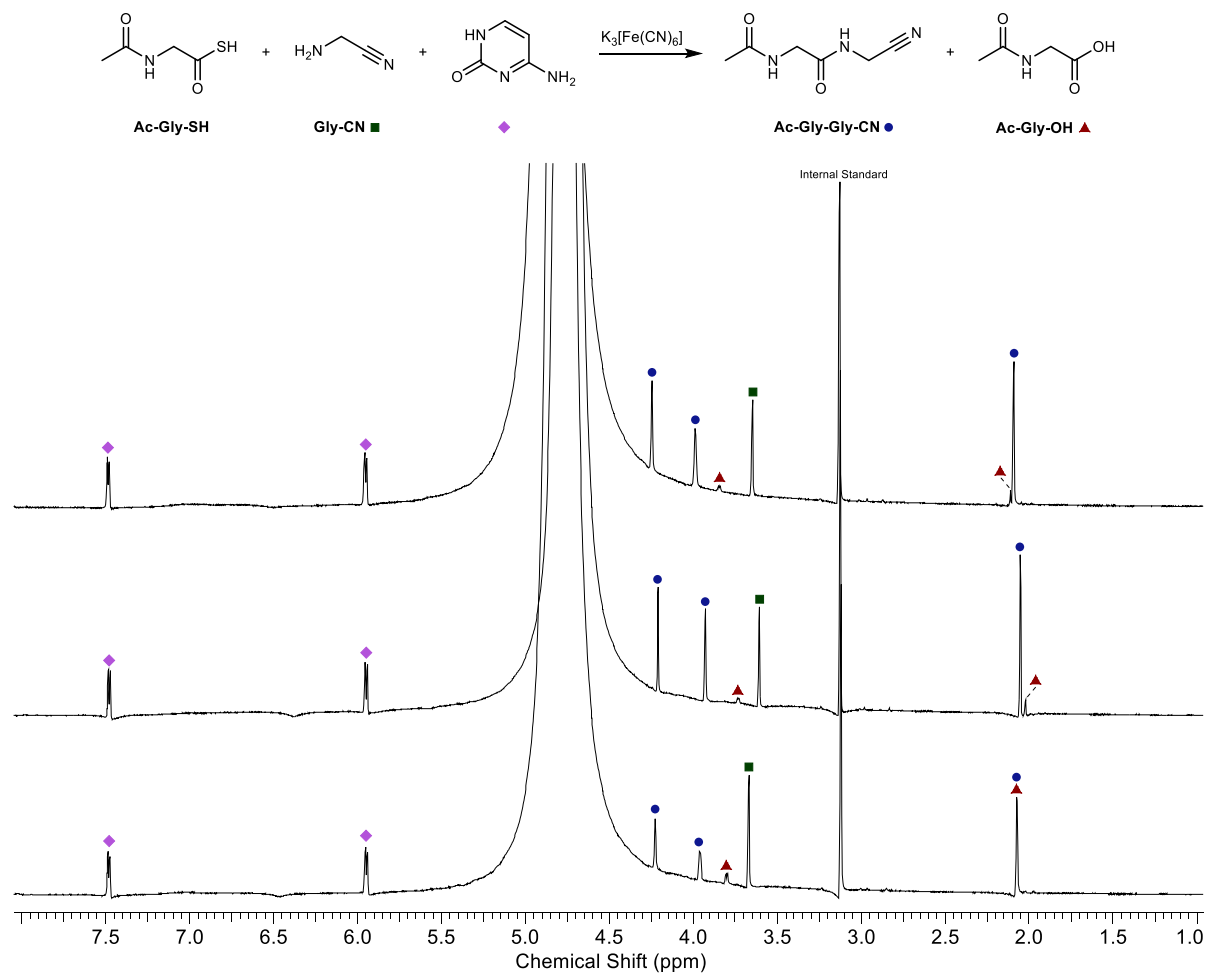
1H NMR (600 MHz, H_2O/D_2O 98:2, pH 9.0) *N*-Acetylglycylphosphate **Ac-Gly-OPO₃H₂** (▼) δ 4.12 (s, 2H, (C2)–H₂), 2.07 (s, 3H, (COCH₃)).

Competition between glycine nitrile and β -alanine nitrile



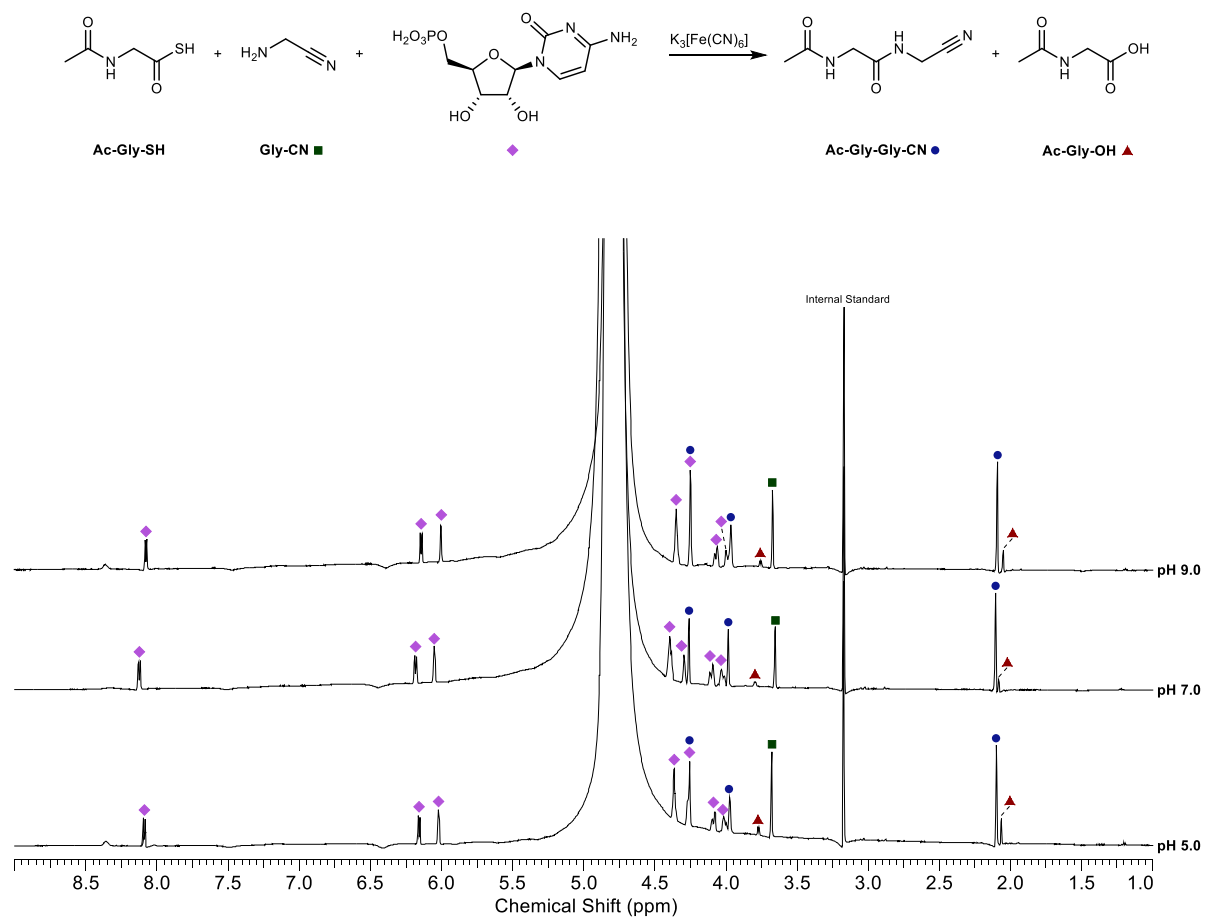
Supplementary Figure 241. ^1H NMR (600 MHz, $\text{H}_2\text{O}/\text{D}_2\text{O}$ 98:2, 2.00–4.50 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**; 100 mM), β -alanine nitrile (**β -Ala-CN**, 100 mM) and $\text{K}_3[\text{Fe}(\text{CN})_6]$ (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Competition between glycine nitrile and cytosine



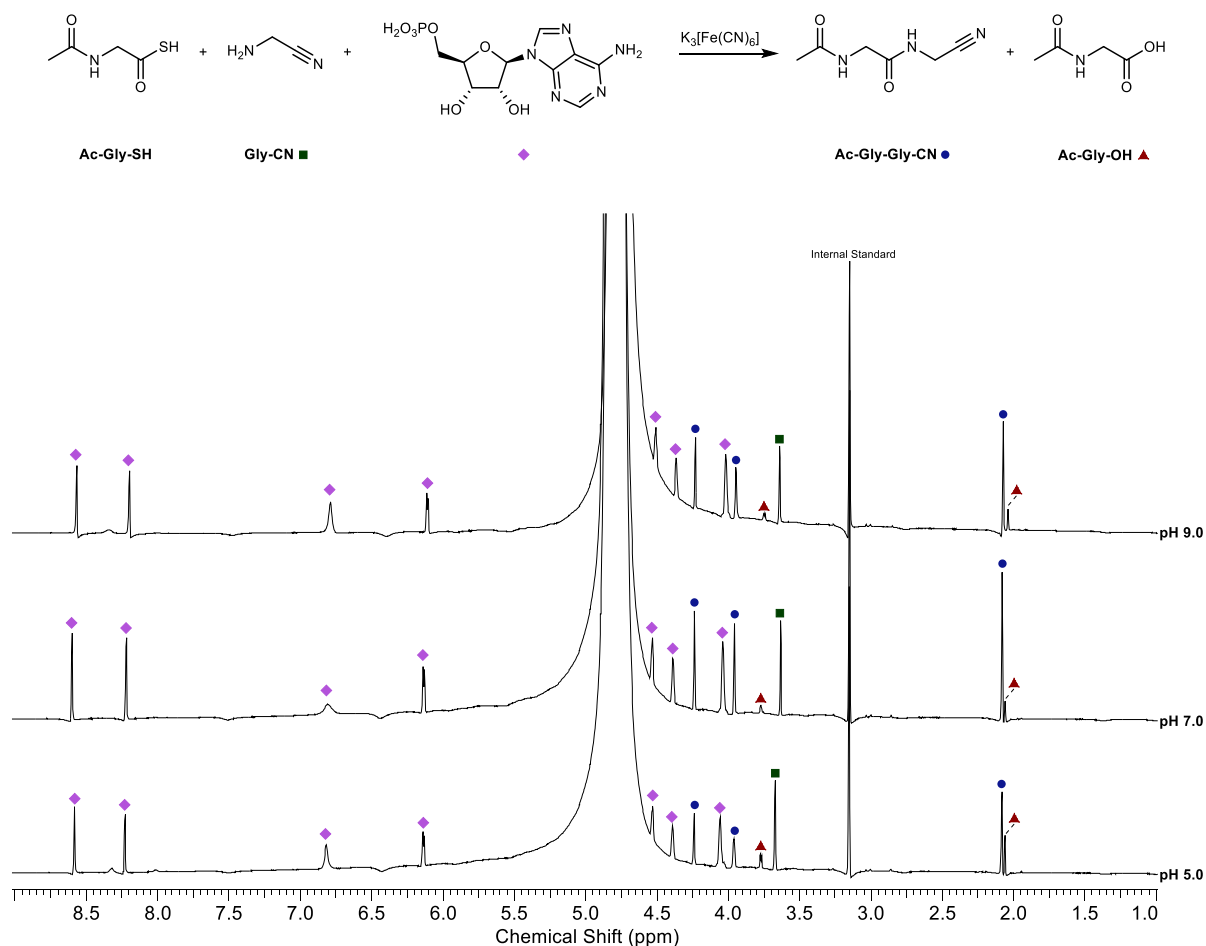
Supplementary Figure 242. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 1.00–8.00 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM), cytosine (100 mM) and K₃[Fe(CN)₆] (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Competition between glycine nitrile and cytidine-5'-monophosphate



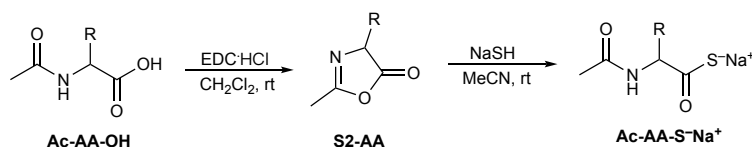
Supplementary Figure 243. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 1.00–9.00 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**; 100 mM), cytidine-5'-monophosphate (100 mM) and K₃[Fe(CN)₆] (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Competition between glycine nitrile and adenosine-5'-monophosphate



Supplementary Figure 244. ¹H NMR (600 MHz, H₂O/D₂O 98:2, 9.00 – 1.00 ppm) spectra to show the reaction of 2-acetamidoethanethioic S-acid (**Ac-Gly-SH**; 50 mM), glycine nitrile (**Gly-CN**, 100 mM), adenosine-5'-monophosphate (100 mM) and K₃[Fe(CN)₆] (150 mM) at room temperature and the specified pH (5.0, 7.0 or 9.0) after 10 min.

Preparative syntheses of α -acetamidothioacids



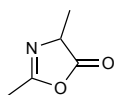
General procedure H: Syntheses of 2-methyloxazol-5(4H)-ones

N-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl; 550 mg, 2.87 mmol) was added to a suspension of *N*-acetyl amino acid (**Ac-AA-OH**; 2.11 mmol) in CH_2Cl_2 (20 mL) at 0 °C and then stirred for 1 h. The resultant homogenous solution was diluted with CH_2Cl_2 (20 mL) and washed with water (2 x 20 mL), NaHCO_3 (sat., 2 x 20 mL) and brine (20 mL). The organic layer was dried over Na_2SO_4 , filtered, and concentrated *in vacuo* to give 2-methyloxazol-5(4H)-one **S2-AA** as an oil, which was used immediately without further purification.⁵³

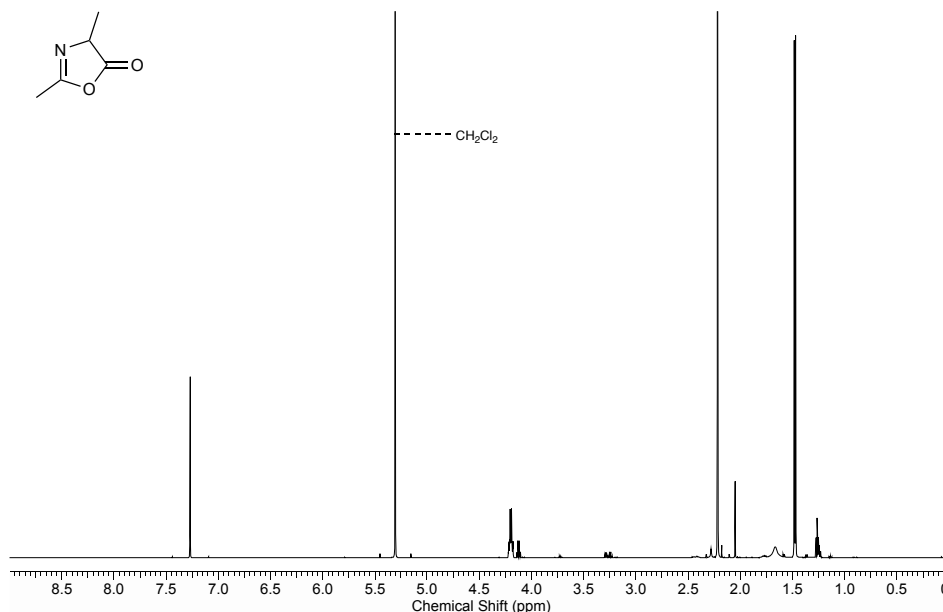
General procedure I: Syntheses of α -acetamidothioacids from 2-methyl-5(4H)-oxazolones

Finely ground sodium hydrosulfide hydrate ($\text{NaSH} \cdot x\text{H}_2\text{O}$, 2 equiv.) was added to a solution of 2-methyl-5(4H)-oxazolone **S2-AA** (1 equiv.) in anhydrous acetonitrile (0.1 M) at room temperature. The suspension was stirred under an argon atmosphere for 16 h. The resultant precipitate was isolated by centrifugation and dried *in vacuo* to afford **Ac-AA-S⁻Na⁺** as a white solid and used without further purification.

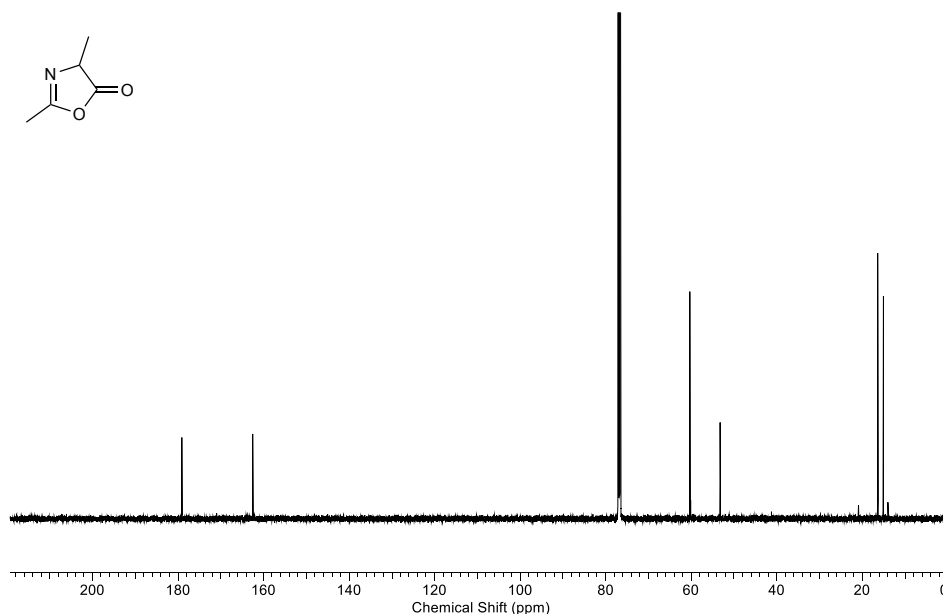
2,4-Dimethyloxazol-5(4H)-one S2-Ala



Prepared following *general procedure H* from *N*-acetyl-L-alanine **Ac-Ala-OH** (227 mg, 2.11 mmol). **S2-Ala** was isolated by careful concentration under reduced pressure (300 mbar, $T_{\text{bath}} = 0\text{ }^{\circ}\text{C}$) to give a colourless oil (424 mg). Due to its known volatility and instability,⁵⁴ **S2-Ala** was taken onto the next step without further purification. ^1H NMR (600 MHz, CDCl_3) δ 4.20 (qq, $J = 2.0, 7.6$ Hz, 1H, (C4)–H), 2.22 (d, $J = 2.0$ Hz, 3H, (C2)–CH₃), 1.47 (d, $J = 7.6$ Hz, 3H, (C4)–CH₃). ^{13}C NMR (151 MHz, CDCl_3) δ 179.1 (C5), 162.5 (C2), 60.3 (C4), 16.4 (C1'), 15.1 (CH₃). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_5\text{H}_8\text{NO}_2$ 114.0550; obs. 114.0550. IR (oil, cm^{-1}) 2985, 2939, 1822, 1683.⁵⁴

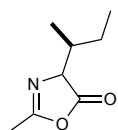


Supplementary Figure 245. ^1H NMR (600 MHz, CDCl_3 , 0–9.0 ppm) spectrum of 2,4-dimethyloxazol-5(4H)-one **S2-Ala**

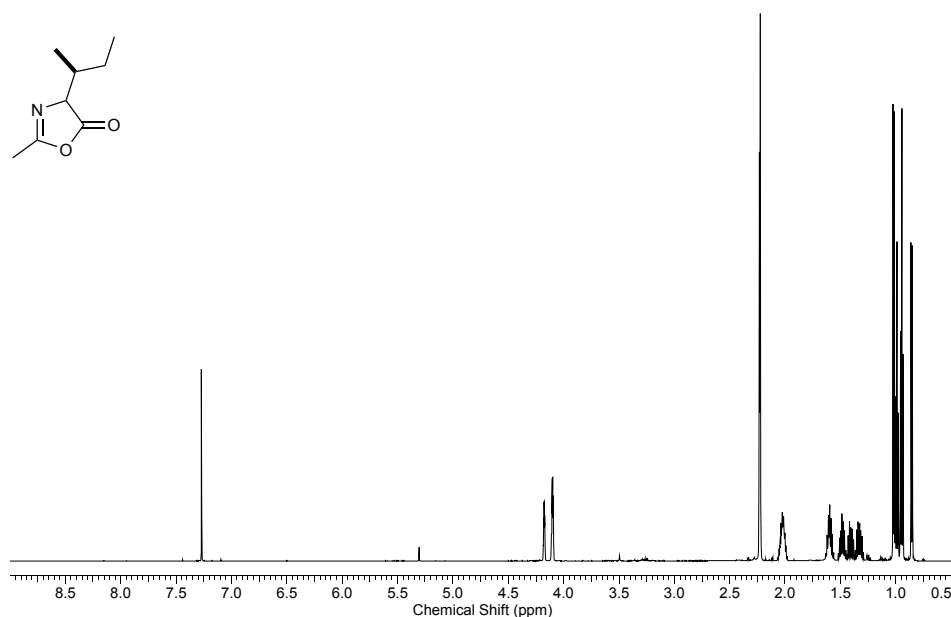


Supplementary Figure 246. ^{13}C NMR (151 MHz, CDCl_3 , 0–220 ppm) spectrum of 2,4-dimethyloxazol-5(4H)-one **S2-Ala**

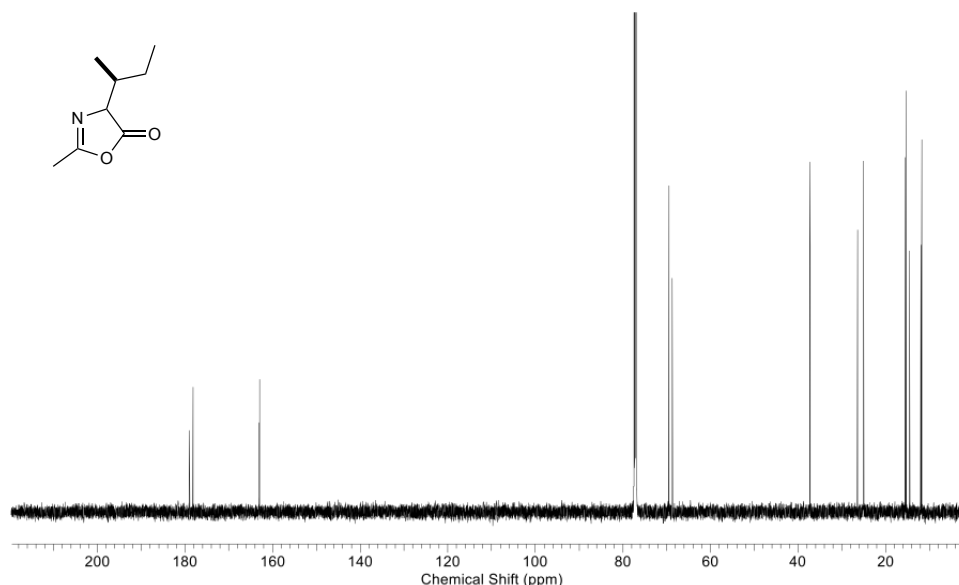
4-((*S*)-*Sec*-butyl)-2-methyloxazol-5(4*H*)-one **S2-Ile**



Prepared following *general procedure H* from *N*-acetyl-L-isoleucine **Ac-Ile-OH** (365 mg, 2.11 mmol). **S2-Ile** was isolated as a colourless oil (274 mg, 84%). ^1H NMR (600 MHz, CDCl_3 , mixture of 2 diastereoisomers (a/b 59:41)). Diastereoisomer a: δ 4.10 (qd, $J = 2.1, 4.3$ Hz, 1H, (C4)–H), 2.22 (d, $J = 2.1$ Hz, 3H, (C2)– CH_3), 2.06 - 2.00 (m, 1H, (C1')–H), 1.53 - 1.45 (m, 1H, (C2')–H), 1.36 - 1.28 (m, 1H, (C2'')–H), 1.02 (d, $J = 6.9$ Hz, 3H, (C1')– CH_3), 0.95 (t, $J = 7.4$ Hz, 3H, (C3')– H_3). Diastereoisomer b: δ 4.18 (qd, $J = 2.1, 3.9$ Hz, 1H, (C4)–H), 2.23 (d, $J = 2.1$ Hz, 3H, (C2)– CH_3), 2.06 - 2.00 (m, 1H, (C1')–H), 1.60 (m, 1H, (C2')–H), 1.44 - 1.36 (m, 1H, (C2'')–H), 0.99 (t, $J = 7.4$ Hz, 3H, (C1')– CH_3), 0.86 (d, $J = 6.8$ Hz, 3H, (C3')– H_3). ^{13}C NMR (151 MHz, CDCl_3 , mixture of 2 diastereoisomers (a/b 59:41)). Diastereoisomer a: δ 178.5 (C5), 162.6 (C2), 69.1 (C4), 36.9 (C1'), 24.6 (C2'), 14.9 (CH_3), 14.9 (CH_3), 11.5 (C3'). Diastereoisomer b: δ 177.8 (C5), 162.7 (C2), 68.3 (C4), 36.8 (C1'), 26.0 (C2'), 15.1 (CH_3), 14.1 (CH_3), 11.4 (C3'). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_8\text{H}_{14}\text{NO}_2$ 156.1024; obs. 156.1025. IR (oil, cm^{-1}) 2966, 2935, 2879, 1822, 1681.⁵⁵

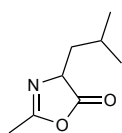


Supplementary Figure 247. ^1H NMR (600 MHz, CDCl_3 , 0–9.0 ppm) spectrum of 4-((*S*)-*sec*-butyl)-2-methyloxazol-5(4*H*)-one **S2-Ile**

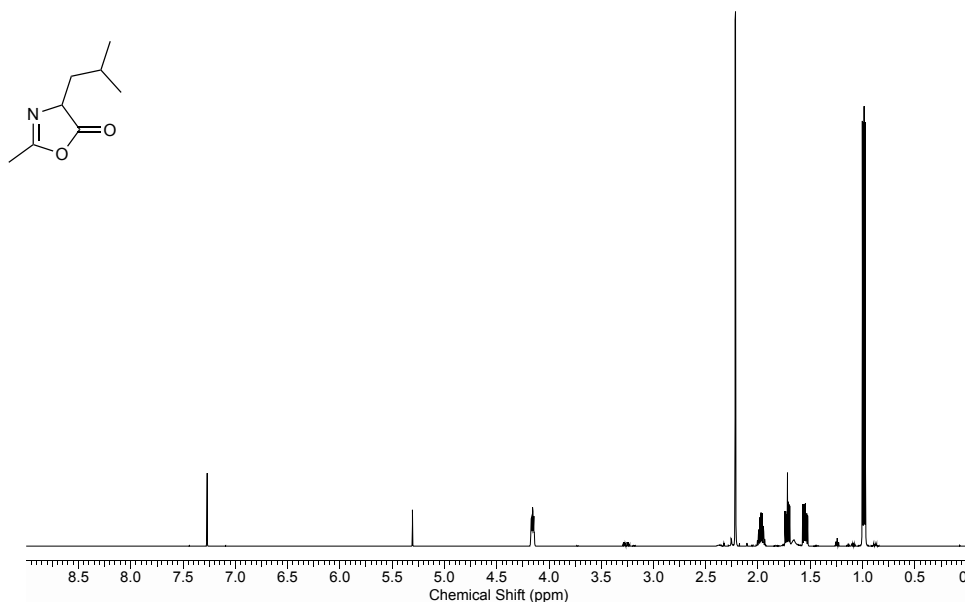


Supplementary Figure 248. ^{13}C NMR (151 MHz, CDCl_3 , 0–220 ppm) spectrum of 4-((*S*)-*sec*-butyl)-2-methyloxazol-5(4*H*)-one **S2-Ile**

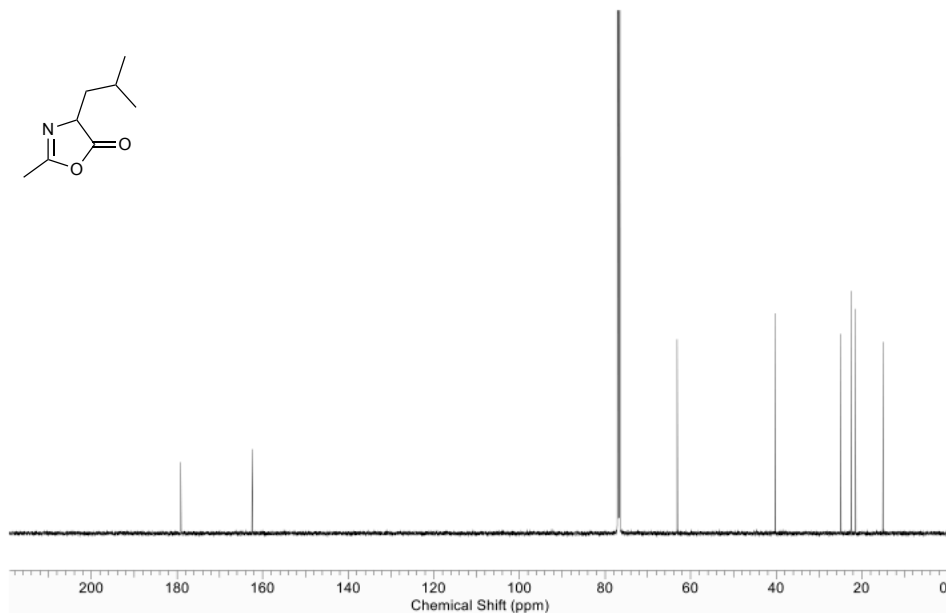
4-Isobutyl-2-methyloxazol-5(4H)-one S2-Leu



Prepared following *general procedure H* from *N*-acetyl-L-leucine **Ac-Leu-OH** (365 mg, 2.11 mmol). **S2-Leu** was isolated as a colourless oil (306 mg, 94%). ^1H NMR (600 MHz, CDCl_3) δ 4.16 (ddq $J = 2.0, 5.5, 9.5$ Hz, 1H, (C4)-H), 2.22 (d, $J = 2.0$ Hz, 3H, (C2)-CH₃), 2.01 - 1.93 (m, 1H, (C2')-H), 1.72 (ddd, $J = 5.5, 8.3, 13.8$ Hz, 1H, (C1')-H), 1.55 (ddd, $J = 6.0, 9.5, 13.8$ Hz, 1H, (C1'')-H), 1.00 (d, $J = 6.7$ Hz, 3H, (C3')-H₃), 0.98 (d, $J = 6.7$ Hz, 3H, (C3'')-H₃). ^{13}C NMR (151 MHz, CDCl_3) δ 179.1 (C5), 162.4 (C2), 63.1 (C4), 40.2 (C1'), 24.9 (C2'), 22.5 (C3'), 21.5 (C3''), 15.1 (CH₃). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_8\text{H}_{14}\text{NO}_2$ 156.1024; obs. 156.1021. IR (oil, cm^{-1}) 2958, 2935, 2872, 1824, 1683.⁵⁶

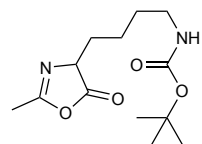


Supplementary Figure 249. ^1H NMR (600 MHz, CDCl_3 , 0–9.0 ppm) spectrum of 4-isobutyl-2-methyloxazol-5(4H)-one **S2-Leu**

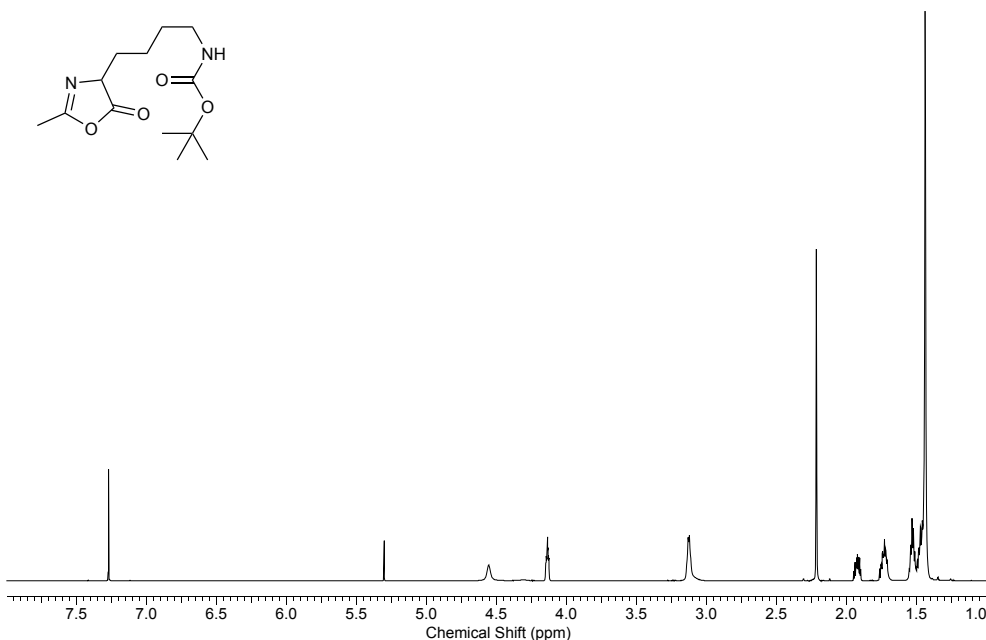


Supplementary Figure 250. ^{13}C NMR (151 MHz, CDCl_3 , 0–220 ppm) spectrum of 4-isobutyl-2-methyloxazol-5(4H)-one **S2-Leu**

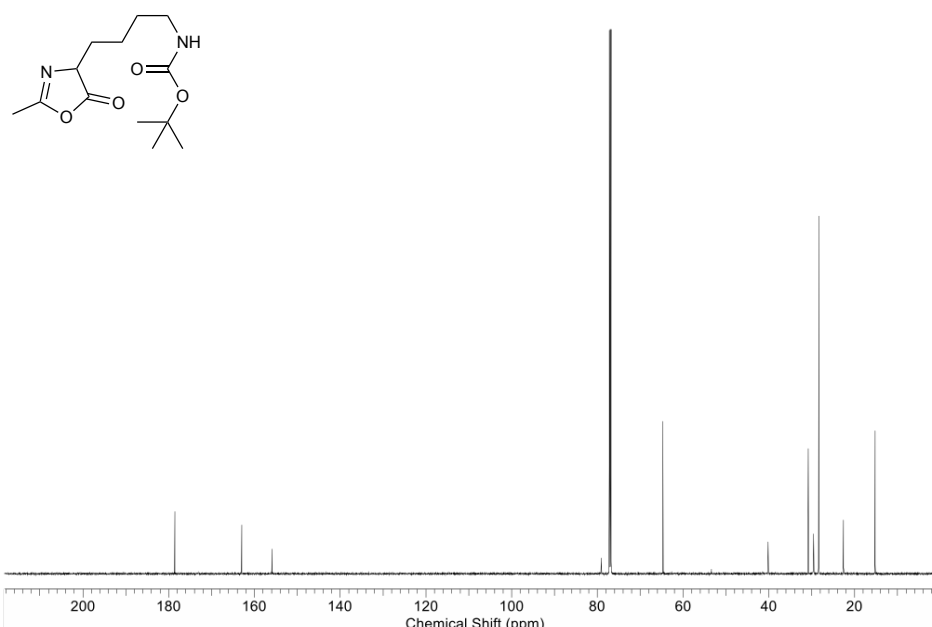
4-(4-Boc-aminobutyl)-2-methyloxazol-5(4H)-one **S2-Lys(Boc)**



Prepared following *general procedure H* from *N*₂-acetyl-*N*₆-Boc-L-lysine (608 mg, 2.11 mmol). **S2-Lys(Boc)** was isolated as a colourless oil (532 mg, 98%). ¹H NMR (700 MHz, CDCl₃, mixture of two rotamers a/b; 85:15) δ 4.56 (br. s., 0.85H, NH, rotamer a), 4.34 (br. s., 0.15H, NH rotamer b), 4.13 (qdd, *J* = 2.0, 5.7, 7.7 Hz, 1H, (C4)–H), 3.18 - 3.00 (m, 2H, (C4')–H₂), 2.21 (d, *J* = 2.0 Hz, 3H, (C2)–H₃), 1.94 - 1.88 (m, 1H, (C1')–H), 1.73 (dddd, *J* = 5.7, 7.7, 9.7, 13.7 Hz, 1H, (C1'')–H), 1.55 - 1.49 (m, 2H, (C3)–H₂), 1.49 - 1.45 (m, 2H, (C2)–H₂), 1.44 - 1.40 (m, 9H, (CO)OC(CH₃)₃). ¹³C NMR (176 MHz, CDCl₃) δ 178.7 (C5), 163.1 (C2), 156.1 ((CO)OC(CH₃)₃), 79.3 ((CO)OC(CH₃)₃), 64.9 (C4), 40.3 (C4'), 31.0 (C1'), 29.8 (C3'), 28.5 ((CO)OC(CH₃)₃), 22.8 (C2'), 15.4 ((C2)–CH₃). HRMS-ESI [M+H]⁺ calc. for C₁₃H₂₃N₂O₄ 271.1652; obs. 271.1652. IR (oil, cm⁻¹) 3356, 2976, 2933, 2866, 1824, 1683, 1516.

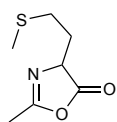


Supplementary Figure 251. ¹H NMR (700 MHz, CDCl₃, 0–8.0 ppm) spectrum of 4-(4-Boc-aminobutyl)-2-methyloxazol-5(4H)-one **S2-Lys(Boc)**



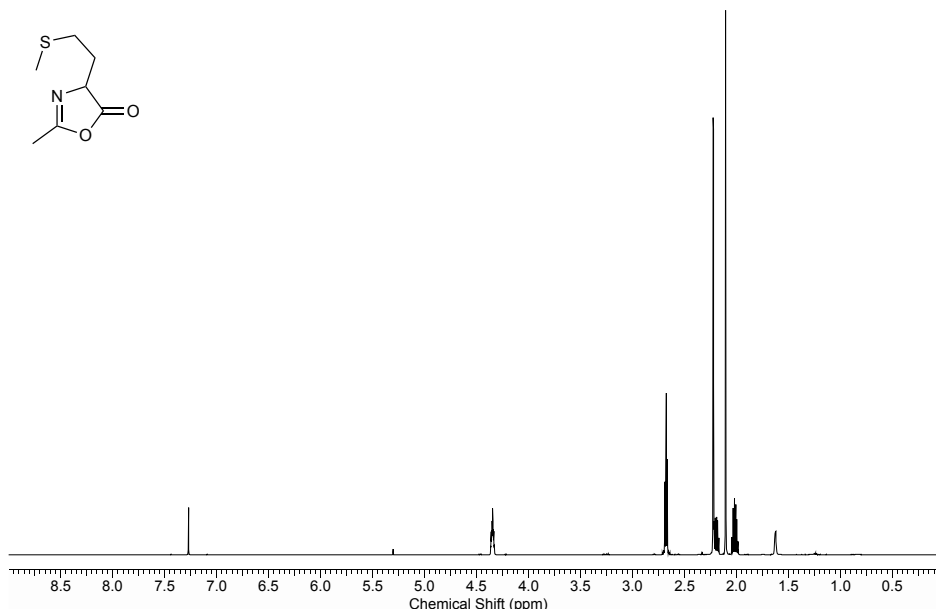
Supplementary Figure 252. ¹³C NMR (176 MHz, CDCl₃, 0–220 ppm) spectrum of 4-(4-Boc-aminobutyl)-2-methyloxazol-5(4H)-one **S2-Lys(Boc)**

2-Methyl-4-(2-(methylthio)ethyl)oxazol-5(4H)-one S2-Met

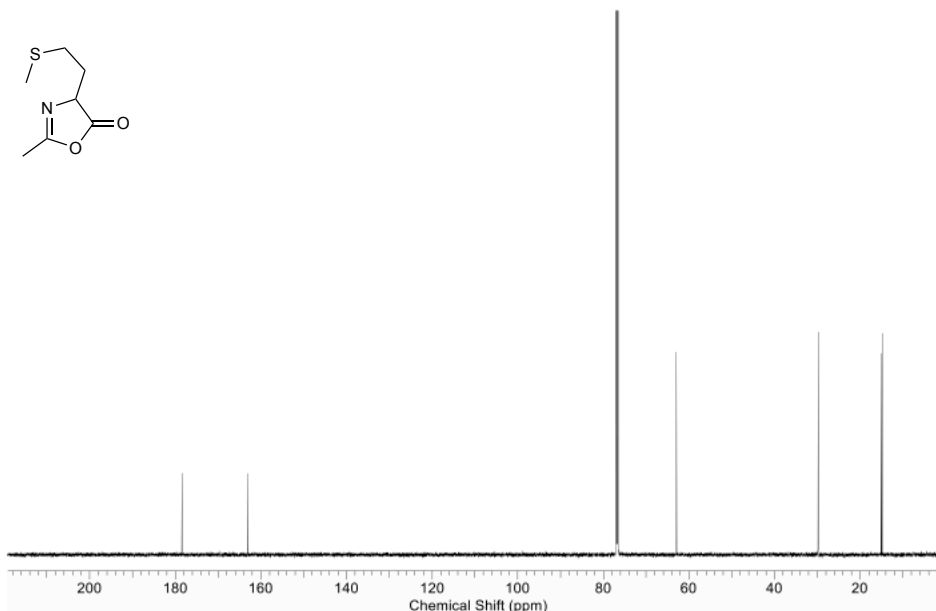


Prepared following *general procedure H* from *N*-acetyl-*rac*-methionine **Ac-Met-OH** (404 mg, 2.11 mmol).

S2-Met was isolated as a colourless oil (327 mg, 89%). ^1H NMR (600 MHz, CDCl_3) δ 4.37 - 4.32 (m, 1H, (C4)-H), 2.67 (app. t, J = 6.9 Hz, 2H, (C2')-H₂), 2.22 (d, J = 2.1 Hz, 3H, (C2)-CH₃), 2.22 - 2.16 (m, 1H, (C1')-H), 2.11 (s, 3H, S-CH₃), 2.04 - 1.98 (m, 1H, (C1'')-H). ^{13}C NMR (151 MHz, CDCl_3) δ 178.7 (C5), 163.3 (C2), 63.2 (C4), 30.1 (C2'), 30.0 (C1'), 15.3 (CH₃), 15.1 (S-CH₃). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_7\text{H}_{12}\text{NO}_2\text{S}$ 174.0589; obs. 174.0591. IR (oil, cm^{-1}) 2920, 1820, 1681, 1433.

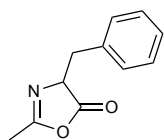


Supplementary Figure 253. ^1H NMR (600 MHz, CDCl_3 , 0–9.0 ppm) spectrum of 2-methyl-4-(2-(methylthio)ethyl)oxazol-5(4H)-one **S2-Met**

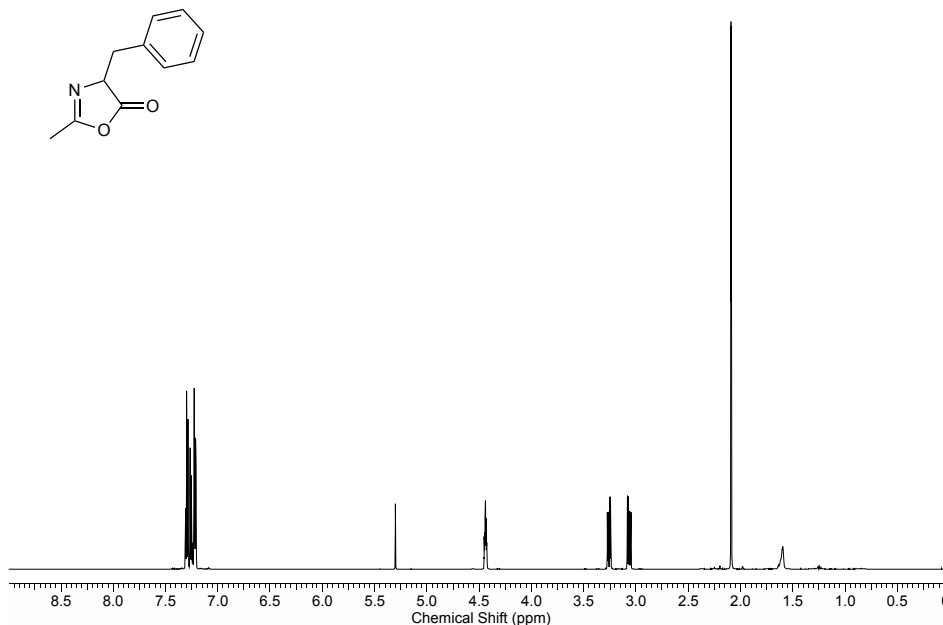


Supplementary Figure 254. ^{13}C NMR (151 MHz, CDCl_3 , 0–220 ppm) spectrum of 2-methyl-4-(2-(methylthio)ethyl)oxazol-5(4H)-one **S2-Met**

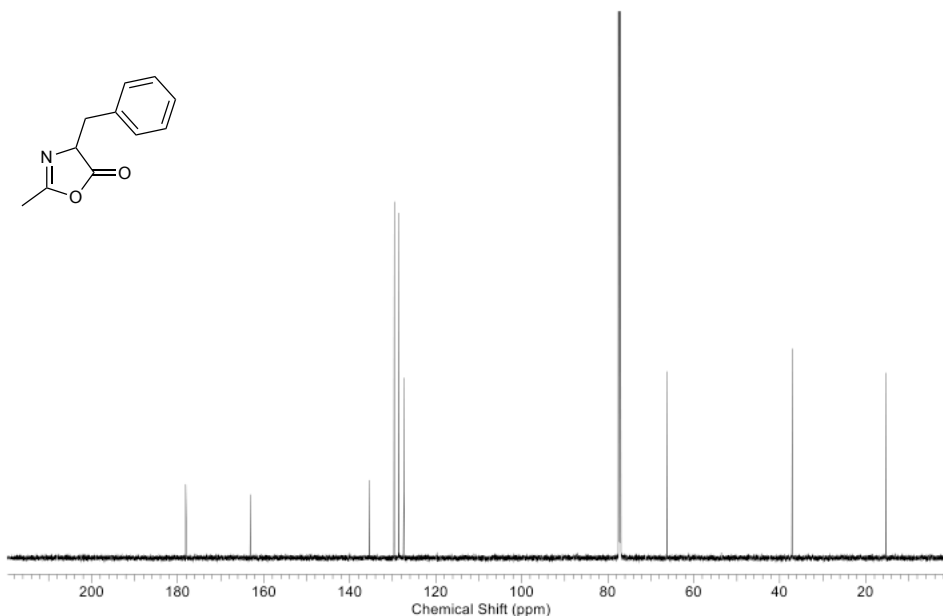
4-Benzyl-2-methyloxazol-5(4H)-one **S2-Phe**



Prepared following *general procedure H* from *N*-acetyl-D-phenylalanine **Ac-Phe-OH** (437 mg, 2.11 mmol). **S2-Phe** was isolated as a colourless oil (399 mg, 99%). ^1H NMR (600 MHz, CDCl_3) δ 7.32 - 7.28 (m, 2H, Ar), 7.27 - 7.24 (m, 1H, Ar), 7.24 - 7.20 (m, 2H, Ar), 4.46 - 4.42 (ddq, J = 2.0, 4.8, 6.8, 1H, (C4)-H), 3.26 (ABX, J = 4.8, 14.0 Hz, 1H, (C1')-H), 3.07 (ABX, J = 6.8, 14.0 Hz, 1H, (C1'')-H), 2.09 (d, J = 2.0 Hz, 3H, (C2)-CH₃). ^{13}C NMR (151 MHz, CDCl_3) δ 177.7 (C5), 162.7 (C2), 135.0 (Ar), 129.2 (Ar $\times$ 2), 128.2 (Ar $\times$ 2), 127.0 (Ar), 65.8 (C4), 36.7 (C1'), 14.8 (CH₃). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_{11}\text{H}_{12}\text{NO}_2$ 190.0868; obs. 190.0863. IR (oil, cm^{-1}) 3059, 3039, 2941, 1816, 1681, 1604.⁵⁷

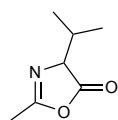


Supplementary Figure 255. ^1H NMR (600 MHz, CDCl_3 , 0–9.0 ppm) spectrum of 4-benzyl-2-methyloxazol-5(4H)-one **S2-Phe**

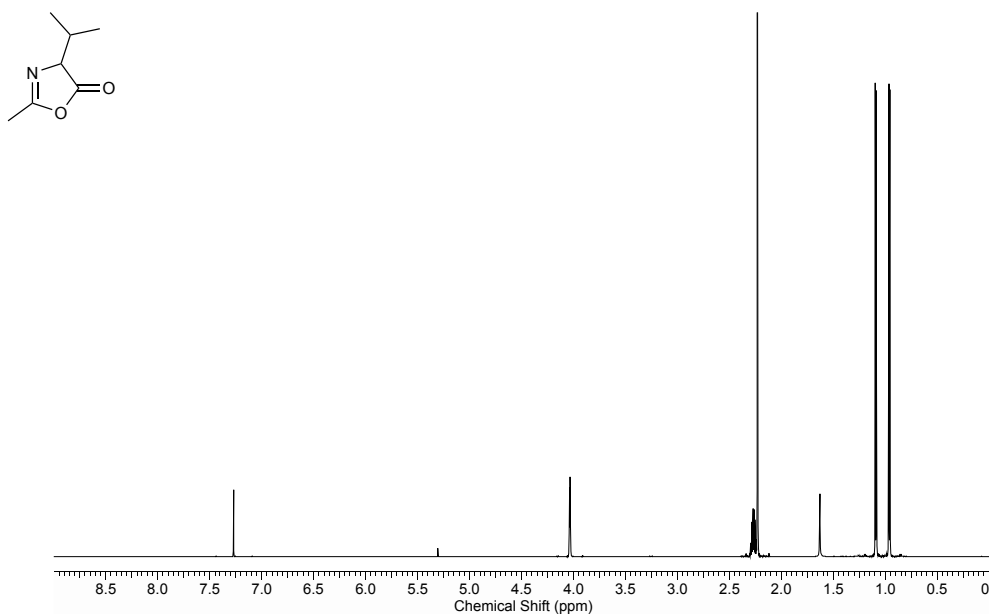


Supplementary Figure 256. ^{13}C NMR (151 MHz, CDCl_3 , 0–220 ppm) spectrum of 4-benzyl-2-methyloxazol-5(4H)-one **S2-Phe**

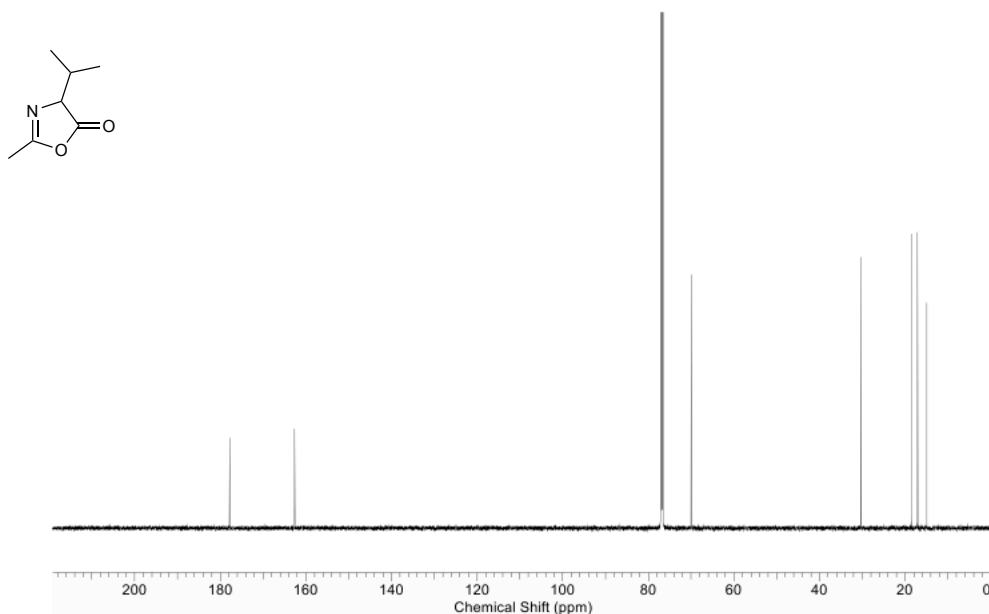
4-Isopropyl-2-methyloxazol-5(4H)-one S2-Val



Prepared following *general procedure H* from *N*-acetyl-L-valine **Ac-Val-OH** (336 mg, 2.11 mmol). **S2-Val** was isolated as a colourless oil (253 mg, 85%). ^1H NMR (600 MHz, CDCl_3) δ 4.03 (dq, $J = 2.1, 4.4$ Hz, 1H, (C4)–H), 2.31 - 2.24 (m, 1H, (C1')–H), 2.23 (d, $J = 2.1$ Hz, 3H, (C2)–CH₃), 1.09 (d, $J = 6.9$ Hz, 3H, (C2')–H₃), 0.96 (d, $J = 6.9$ Hz, 3H, (C2'')–H₃). ^{13}C NMR (151 MHz, CDCl_3) δ 177.8 (C5), 162.7 (C2), 70.0 (C4), 30.4 (C1'), 18.5 (C2'), 17.1 (C2''), 14.9 (CH₃). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_7\text{H}_{12}\text{NO}_2$ 142.0870; obs. 142.0868. IR (oil, cm^{-1}) 2966, 2935, 2877, 1820, 1681.⁵⁸

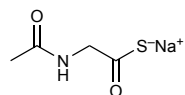


Supplementary Figure 257. ^1H NMR (600 MHz, CDCl_3 , 0–9.0 ppm) spectrum of 4-isopropyl-2-methyloxazol-5(4H)-one **S2-Val**

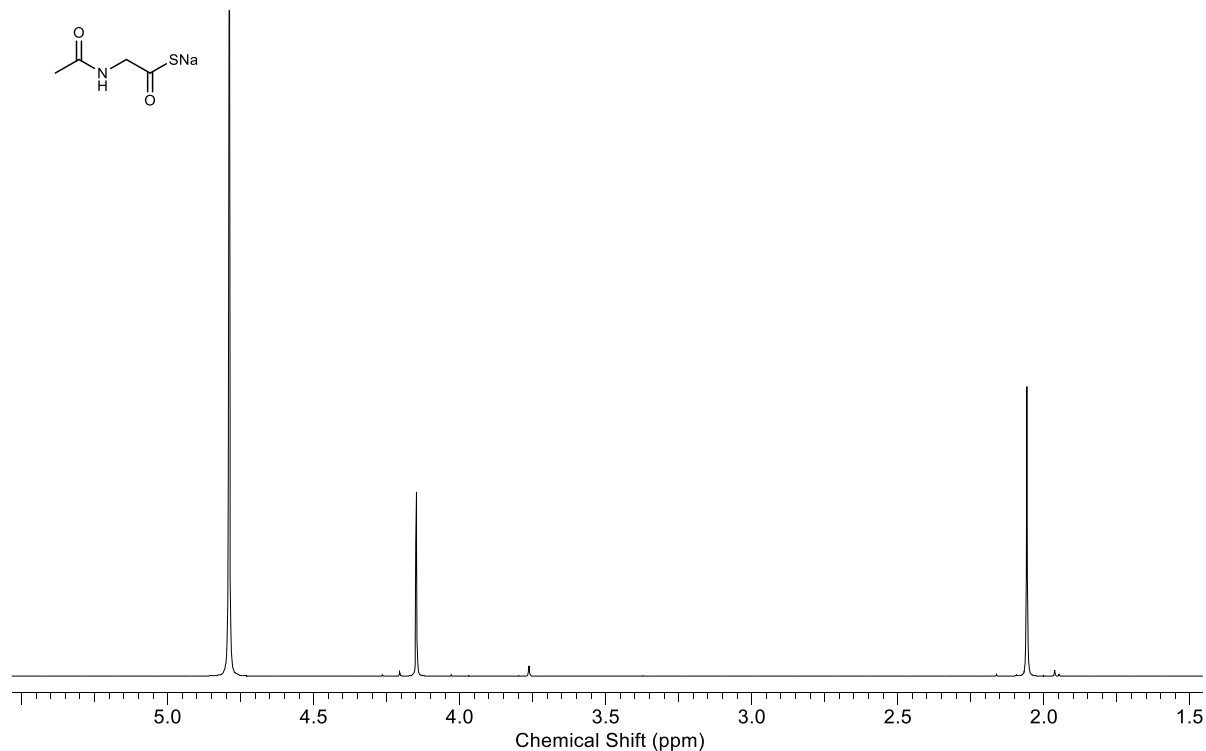


Supplementary Figure 258. ^{13}C NMR (151 MHz, CDCl_3 , 0–220 ppm) spectrum of 4-isopropyl-2-methyloxazol-5(4H)-one **S2-Val**

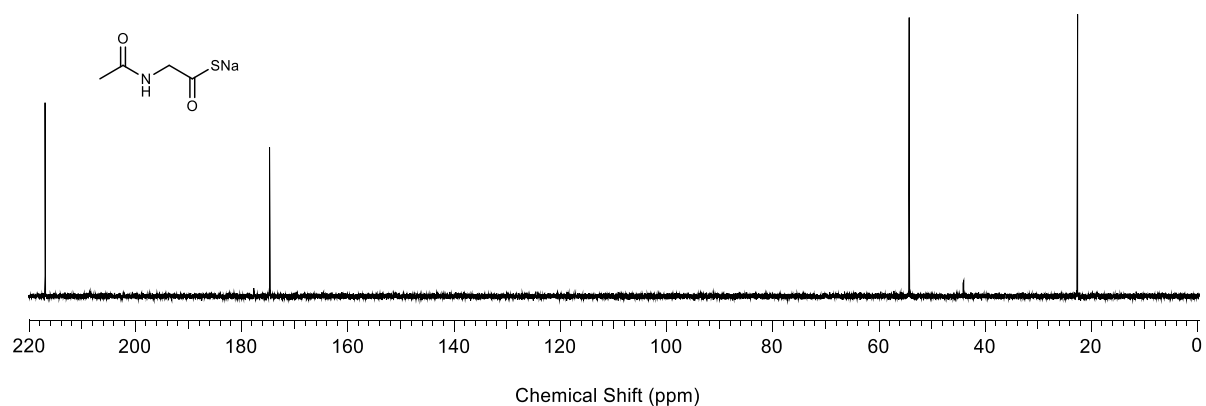
Sodium 2-acetamidoethanethioate **Ac-Gly-S⁻Na⁺**



Sodium hydrosulfide hydrate (NaSH·xH₂O; 807 mg, 6 mmol) was dissolved in degassed H₂O (50 mL), and the solution was adjusted to pH 9.0. *N*-Acetylglycine pentafluorophenyl ester (1.42 g, 5 mmol) was added portion-wise whilst the solution was maintained at pH 9.0 through the addition of NaOH. The reaction was vigorously stirred at room temperature for 1 h. The solution was lyophilised and the resultant residue was washed with diethyl ether (5 x 50 mL) and dried *in vacuo* to afford sodium 2-acetamidoethanethioate **Ac-Gly-S⁻Na⁺** as a white solid (1.53 g), which was used without further purification. ¹H NMR (600 MHz, D₂O): δ 4.15 (s, 2H, (C2)–H₂), 2.06 (s, 3H, (COCH₃)). ¹³C NMR (151 MHz, D₂O): δ 217.0 (C1), 174.7 (COCH₃), 54.3 (C2), 22.6 (COCH₃). HRMS-ESI [M–H][–] calcd. for [C₄H₇NO₂S–H][–]: 132.0119; observed: 132.0117. IR (solid): 3403, 3254, 2985, 1629, 1568, 1518 cm^{–1}.

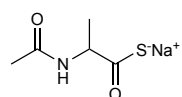


Supplementary Figure 259. ¹H NMR (600 MHz, D₂O, 1.50–5.50 ppm) spectrum to show sodium 2-acetamidoethanethioate **Ac-Gly-S⁻Na⁺**

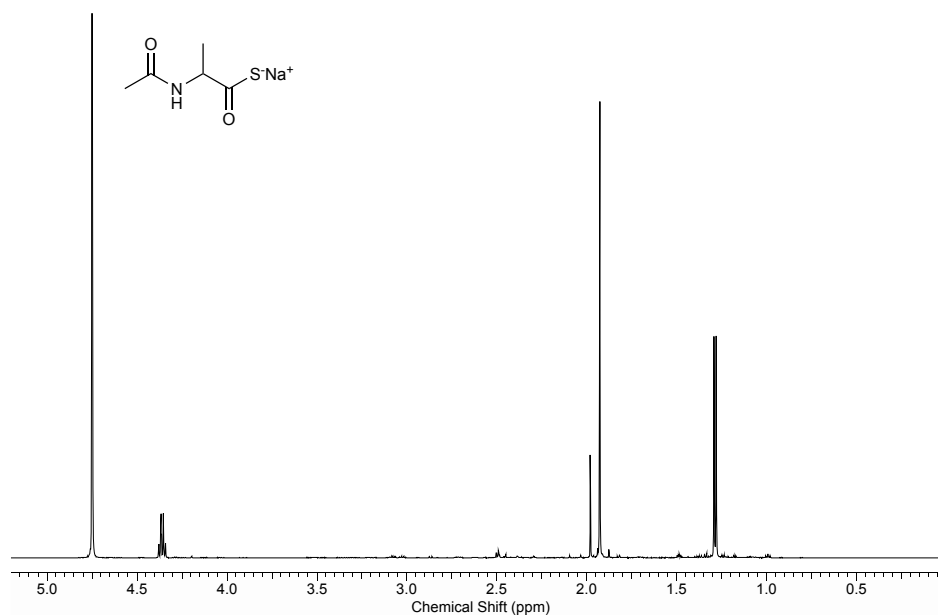


Supplementary Figure 260. ¹³C NMR (151 MHz, D₂O, 0–220 ppm) spectrum to show sodium 2-acetamidoethanethioate **Ac-Gly-S⁻Na⁺**

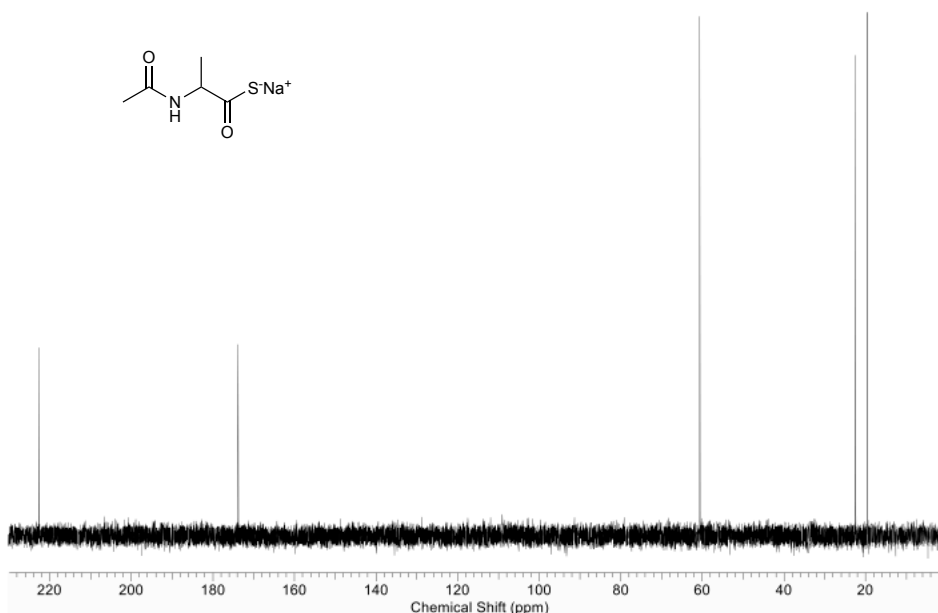
Sodium 2-acetamidopropanethioate **Ac-Ala-S⁻Na⁺**



Prepared following a modification of *general procedure I*. Finely ground sodium hydrosulfide hydrate ($\text{NaSH} \cdot x\text{H}_2\text{O}$, 211 mg, 3.77 mmol) was added to a solution of 2,4-dimethyloxazol-5-*4H*-one **S2-Ala** (212 mg, 1.88 mmol) in anhydrous acetonitrile (10 mL) at room temperature. The suspension was stirred under an argon atmosphere for 16 h. The resultant solution was concentrated *in vacuo* to afford sodium 2-acetamidopropanethioate **Ac-Ala-S⁻Na⁺** as a white hygroscopic foam (290 mg) that was used without further purification. ^1H NMR (600 MHz, D_2O) δ 4.36 (q, $J = 7.2$ Hz, 1H, (C2)–H), 1.93 (s, 3H, (COCH₃), 1.29 (d, $J = 7.2$ Hz, 3H, (C3)–H₃). ^{13}C NMR (151 MHz, D_2O) δ 222.6 (C1), 173.8 (COCH₃), 60.6 (C2), 22.5 (COCH₃), 19.5 (C3). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_5\text{H}_9\text{NO}_2\text{SNa}$ 170.0252; obs. 170.0245. IR (solid, cm^{-1}) 3284, 2978, 2931, 1631, 1508.

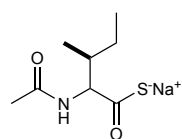


Supplementary Figure 261. ^1H NMR (600 MHz, D_2O , 0–5.2 ppm) spectrum of sodium 2-acetamidopropanethioate **Ac-Ala-S⁻Na⁺**

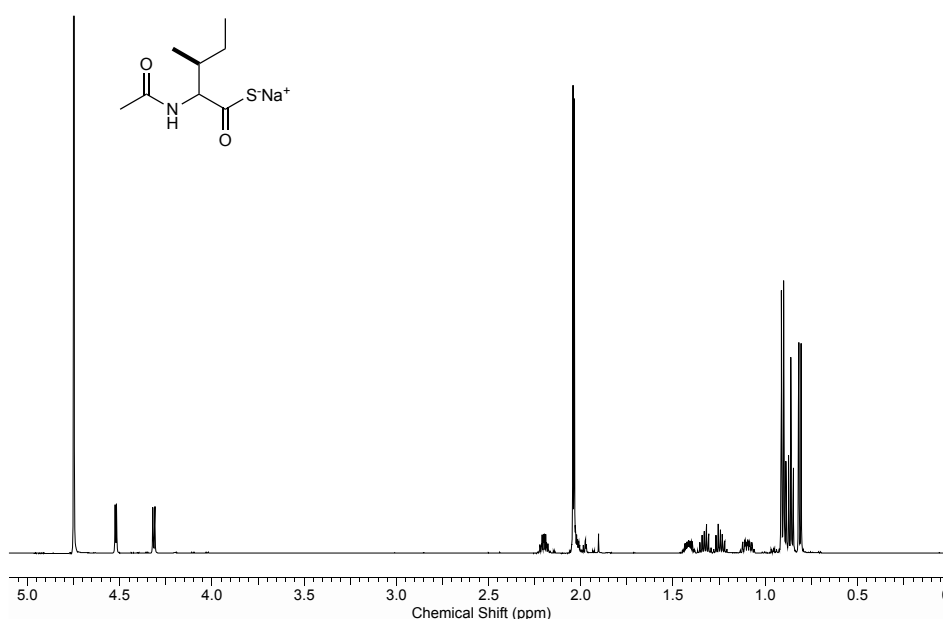


Supplementary Figure 262. ^{13}C NMR (151 MHz, D_2O , 0–230 ppm) spectrum of sodium 2-acetamidopropanethioate **Ac-Ala-S⁻Na⁺**

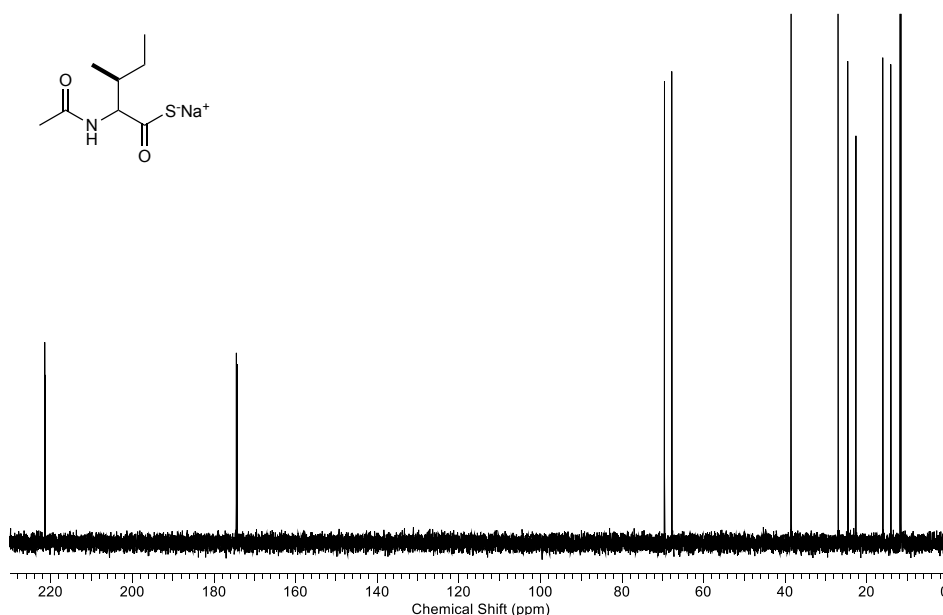
Sodium (3S)-2-acetamido-3-methylpentanethioate **Ac-Ile-S⁻Na⁺**



Prepared following *general procedure I* from 4-((S)-sec-butyl)-2-methyloxazol-5(4H)-one **S2-Ile** (240 mg, 1.55 mmol). **Ac-Ile-S⁻Na⁺** was isolated as a white powder (263 mg). ¹H NMR (600 MHz, D₂O, mixture of 2 diastereoisomers (a/b 51:49)). Diastereoisomer a: δ 4.52 (d, *J* = 4.7 Hz, 1H, (C2)–H), 2.23 – 2.17 (m, 1H, (C3)–H), 2.04 (s, 3H, (COCH₃), 1.37 – 1.30 (m, 1H, (C4)–H), 1.29 – 1.20 (m, 1H, (C4')–H), 0.86 (t, *J* = 7.4 Hz, 3H, (C3)–CH₃), 0.81 (d, *J* = 7.0 Hz, 3H, (C5)–H₃). Diastereoisomer b: δ 4.31 (d, *J* = 6.4 Hz, 1H, (C2)–H), 2.04 (s, 3H, (COCH₃), 2.03 – 2.00 (m, 1H, (C3)–H), 1.45 – 1.38 (m, 1H, (C4)–H), 1.15 – 1.05 (m, 1H, (C4)–H), 0.91 (d, 3H, *J* = 7.6 Hz, (C3)–CH₃), 0.89 – 0.91 (obs., 3H, (C5)–H₃). ¹³C NMR (151 MHz, D₂O, mixture of 2 diastereoisomers (a/b 51:49)). Diastereoisomer a: δ 221.5 (C1), 174.5 (COCH₃), 67.7 (C2), 38.5 (C3), 27.0 (C4), 22.7 (COCH₃), 14.1 ((C2)–CH₃), 11.6 (C5). Diastereoisomer b: δ 221.4, (C1) 174.3 (COCH₃), 69.6 (C2), 38.5 (C3), 24.6 (C4), 22.7 (COCH₃), 16.1 ((C2)–CH₃), 11.8 (C5). HRMS-ESI [M+H]⁺ calcd. for C₈H₁₅NO₂SNa 212.0721; obs. 212.0721. IR (solid, cm⁻¹) 3298, 2960, 2931, 1656, 1544, 1531.

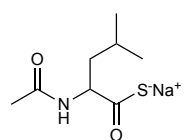


Supplementary Figure 263. ¹H NMR (600 MHz, D₂O, 0–5.2 ppm) spectrum of sodium (3S)-2-acetamido-3-methylpentanethioate **Ac-Ile-S⁻Na⁺**

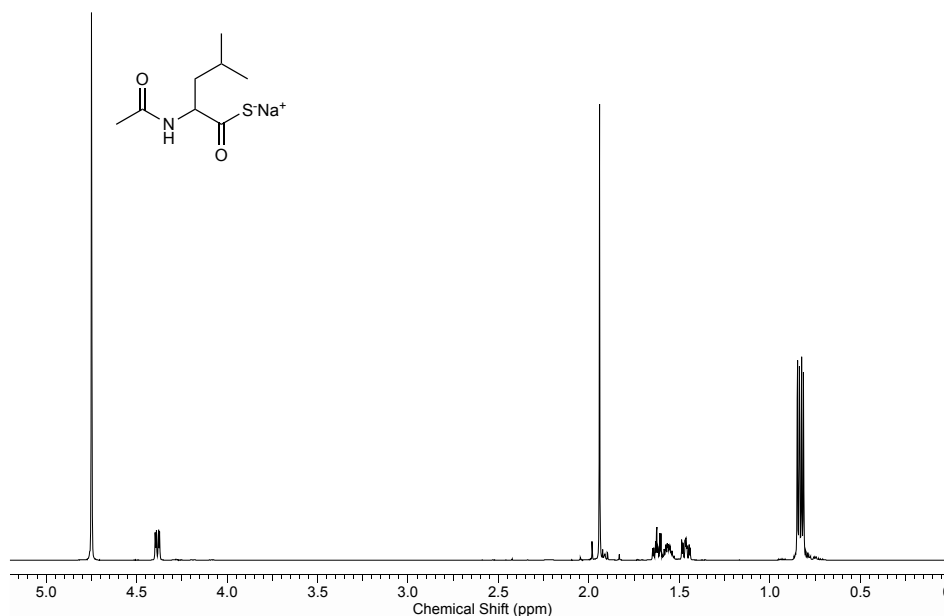


Supplementary Figure 264. ¹³C NMR (151 MHz, D₂O, 0–230 ppm) spectrum of sodium (3S)-2-acetamido-3-methylpentanethioate **Ac-Ile-S⁻Na⁺**

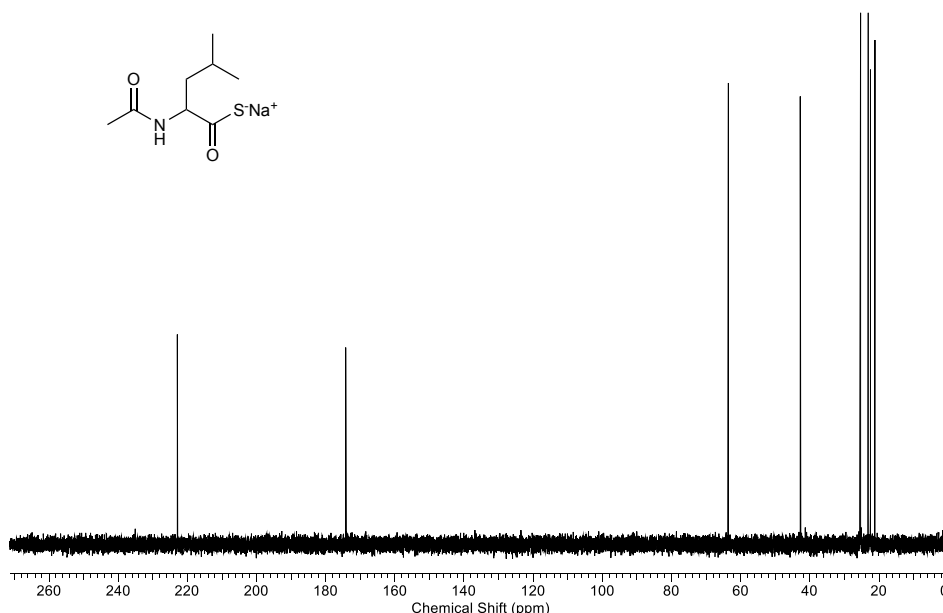
Sodium 2-acetamido-4-methylpentanethioate **Ac-Leu-S⁻Na⁺**



Prepared following *general procedure I* from 4-isobutyl-2-methyloxazol-5(4*H*)-one **S2-Leu** (240 mg, 1.55 mmol). **Ac-Leu-S⁻Na⁺** was isolated as a white solid (322 mg). ¹H NMR (600 MHz, D₂O) δ 4.38 (dd, *J* = 3.8, 10.7 Hz, 1H, (C2)–H), 1.94 (s, 3H, (COCH₃)), 1.65 - 1.60 (m, 1H, (C3)–H), 1.59 - 1.53 (m, 1H, (C4)–H), 1.50 - 1.43 (m, 1H, (C3')–H), 0.84 (d, *J* = 6.5 Hz, 3H, (C5)–H₃), 0.82 (d, *J* = 6.4 Hz, 3H, (C5')–H₃). ¹³C NMR (151 MHz, D₂O) δ 222.9 (C1), 174.1 (COCH₃), 63.6 (C2), 42.8 (C3), 25.4 (C4), 23.2 (C5), 22.5 (COCH₃), 21.3 (C5'). HRMS-ESI [*M*+*H*]⁺ calcd. for C₈H₁₅NO₂SNa 212.0721; obs. 212.0720. IR (solid, cm⁻¹) 3300, 2964, 2956, 2870, 1654, 1537.

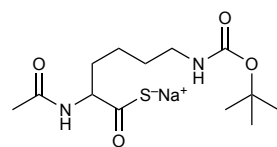
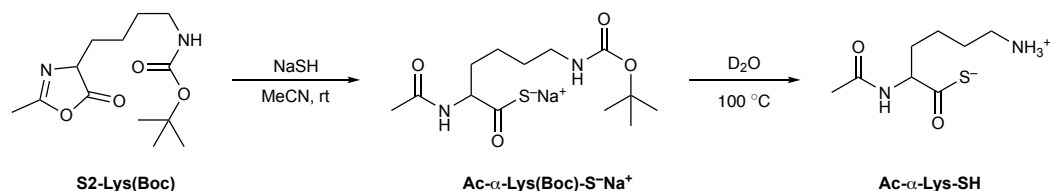


Supplementary Figure 265. ¹H NMR (600 MHz, D₂O, 0–5.2 ppm) spectrum of sodium 2-acetamido-4-methylpentanethioate **Ac-Leu-S⁻Na⁺**



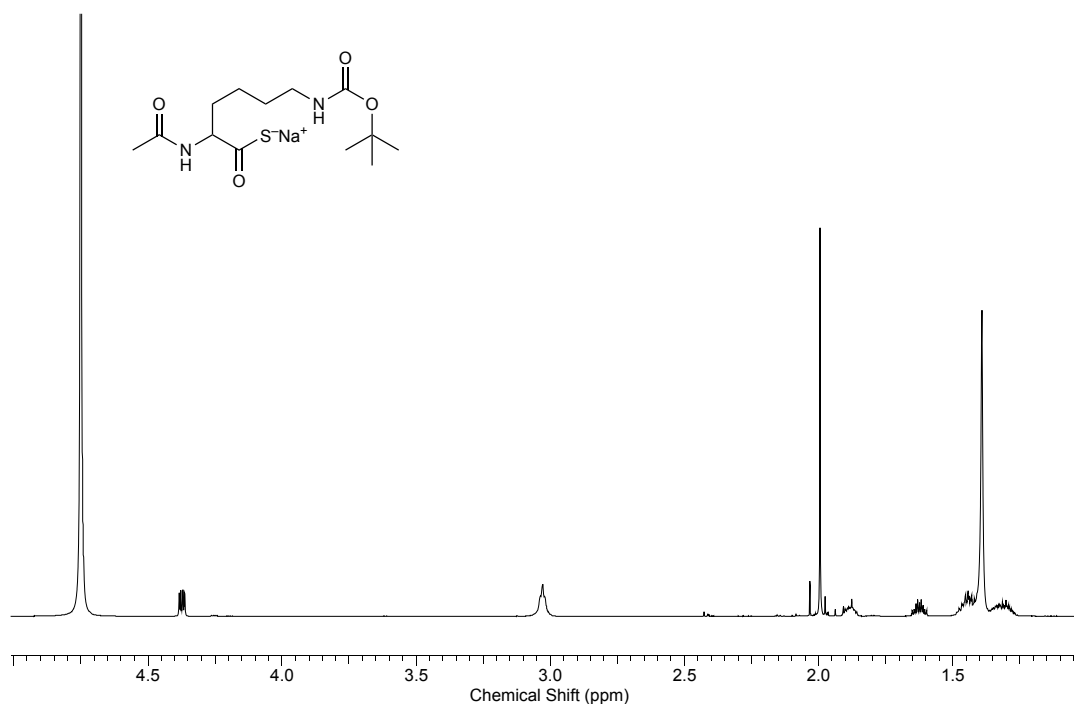
Supplementary Figure 266. ¹³C NMR (151 MHz, D₂O, 0–280 ppm) spectrum of sodium 2-acetamido-4-methylpentanethioate **Ac-Leu-S⁻Na⁺**

2-Acetamido-6-ammoniohexanethioate **Ac- α -Lys-S⁻Na⁺**

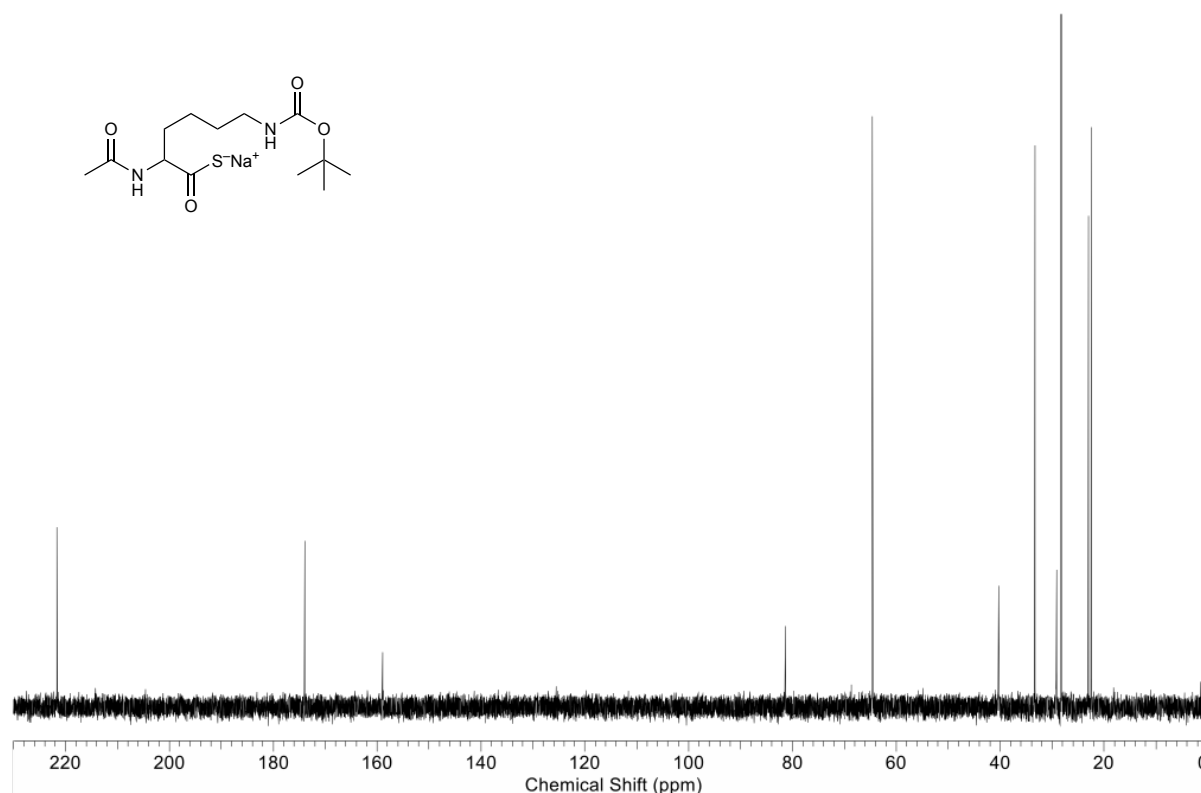


Finely ground sodium hydrosulfide hydrate (236 mg, 4.22 mmol) was added to a solution of 4-(4-Boc-aminobutyl)-2-methyloxazol-5(4*H*)-one **S2-Lys(Boc)** (534 mg, 1.98 mmol) in anhydrous acetonitrile (8 mL) at room temperature. The suspension was stirred under an argon atmosphere for 16 h. The resultant solution was concentrated *in vacuo* to afford sodium

2-acetamido-6-((*tert*-butoxycarbonyl)amino)hexanethioate **Ac- α -Lys(Boc)-S⁻Na⁺** as a white hygroscopic foam (705 mg) that was used without further purification. ¹H NMR (700 MHz, D₂O) δ 4.38 (dd, *J* = 4.4, 9.3 Hz, 1H, (C2)–H), 3.04 (br. t, *J* = 5.5 Hz, 2H, (C6)–H₂), 2.01 (s, 3H, COCH₃), 1.93 - 1.85 (m, 1H, (C3)–H), 1.63 (m, 1H, (C3')–H), 1.50 - 1.42 (m, 2H, (C5)–H₂), 1.40 (s, 9H, C(CH₃)₃), 1.37 - 1.27 (m, 2H, (C4)–H₂). ¹³C NMR (176 MHz, D₂O) δ 221.6 (C1), 174.0 (COCH₃), 158.9 ((CO)OC(CH₃)₃), 81.3 ((CO)OC(CH₃)₃), 64.7 (C2), 40.3 (C6), 33.3 (C3), 29.2 ((CO)OC(CH₃)₃), 28.3 (C5), 23.0 (C4), 22.5 (COCH₃). HRMS-ESI-[M+H]⁺ calc. for C₁₃H₂₅N₂O₄S 305.1530; obs. 305.1528. IR (solid, cm⁻¹) 3284, 2972, 2926, 2862, 1687, 1651, 1516.



Supplementary Figure 267. ¹H NMR (700 MHz, D₂O, 1.0–5.0 ppm) spectrum of 2-acetamido-6-((*tert*-butoxycarbonyl)amino)hexanethioate **Ac- α -Lys(Boc)-S⁻Na⁺**

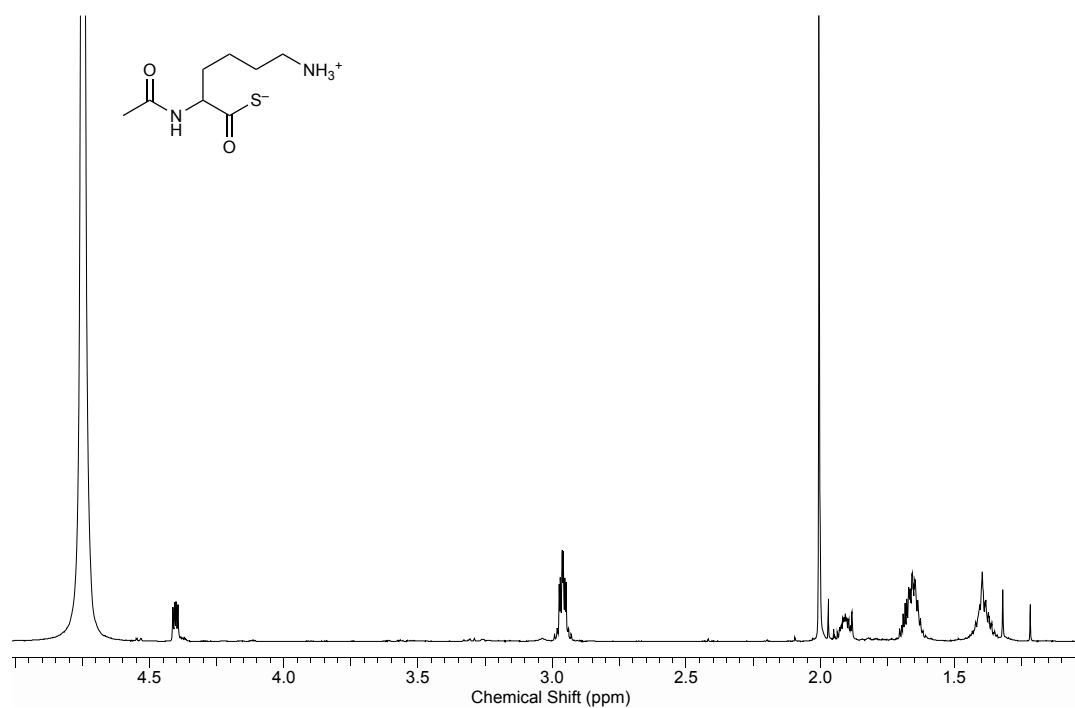


Supplementary Figure 268. ^{13}C NMR (176 MHz, D_2O , 0–230 ppm) spectrum of 2-acetamido-6-((tert-butoxycarbonyl)amino)hexanethioate **Ac- α -Lys(Boc)-S-Na $^+$**

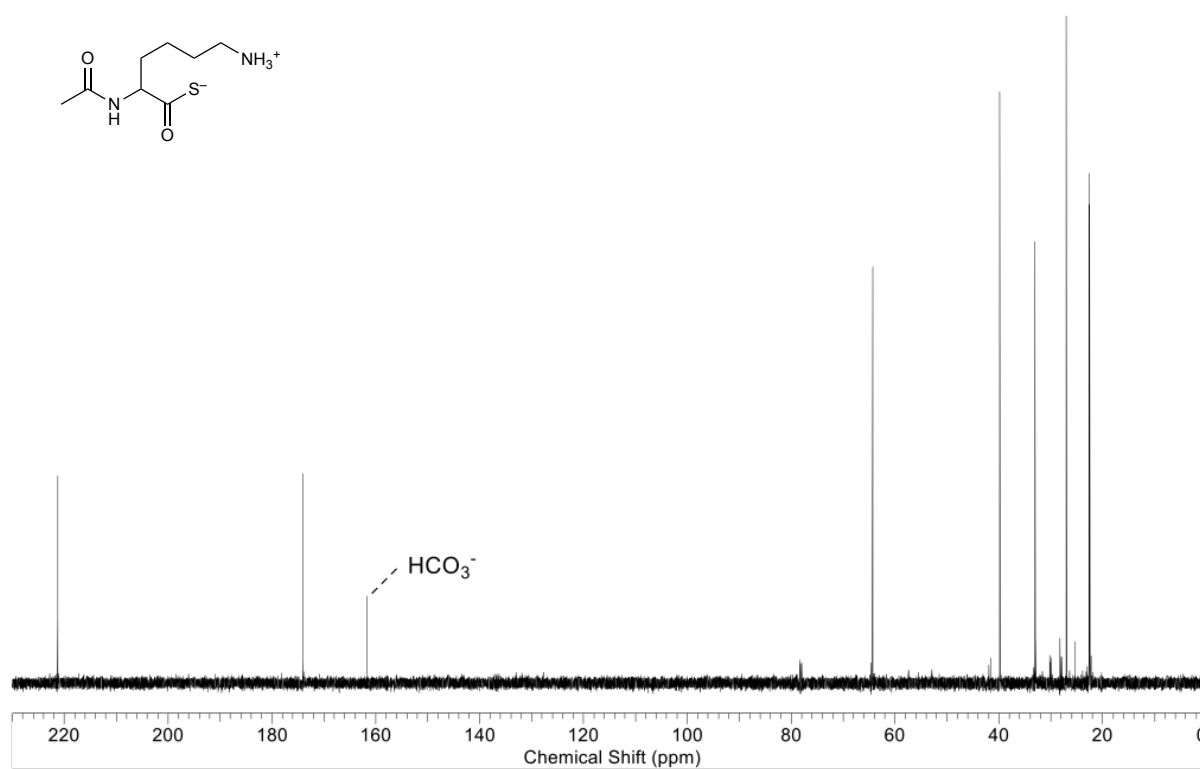
Ac- α -Lys(Boc)-S-Na $^+$ (16.4 mg, 0.05 mmol) was dissolved in D_2O (1 mL) and the solution was heated at 100 $^\circ\text{C}$ in a sealed NMR tube.⁵⁹ The reaction was periodically cooled to room temperature and analysed by ^1H NMR spectroscopy. Heating was continued until complete conversion of **Ac- α -Lys(Boc)-S-Na $^+$** was observed. The reaction mixture was allowed to cool to room temperature and the aqueous layer extracted with CDCl_3 (4 x 2.5 mL), to remove *N*-acetyl-lysine- ϵ -lactam **c-Ac- α -Lys** (see Supplementary Page S67) and *t*-butanol by-products, giving 2-acetamido-6-aminothiohexanoic S-acid **Ac- α -Lys-SH** (58%,^x determined using potassium hydrogen phthalate as an internal NMR standard) and *N*- α -acetyllysine **Ac- α -Lys-OH** (2-acetamido-6-aminothiohexanoic S-acid **Ac- α -Lys-SH**: *N*- α -acetyllysine **Ac- α -Lys-OH**; 90:10) ^1H NMR (700 MHz, D_2O) δ = 4.40 (dd, J = 4.7, 8.8 Hz, 1H, (C2)–H), 3.00 - 2.92 (m, 2H, (C6)–H₂), 2.00 (s, 3H, COCH₃), 1.95 - 1.88 (m, 1H, (C3)–H), 1.71 - 1.61 (m, 3H, (C3)–H', (C5)–H₂), 1.45 - 1.34 (m, 2H, (C4)–H₂). ^{13}C NMR (176 MHz, D_2O) δ = 221.2 (C1), 174.0 (COCH₃), 64.3 (C2), 39.9 (C5), 33.0 (C3), 27.0 (C5), 22.7 (COCH₃), 22.5 (C4).^{xi} HRMS-ESI $[\text{M}+\text{H}]^+$ calc. for $\text{C}_8\text{H}_{17}\text{N}_2\text{O}_2\text{S}^+$ 205.1005; obs. 205.1008.

^x The Boc-deprotection of **Ac- α -Lys(Boc)-S-Na $^+$** was observed to be a capricious reaction: yield varied (58–71%) and length of heating time to achieve full starting material consumption (8–24 h). Yields were sufficient for our requirements and no further effort was made to optimise these Boc-deprotections.

^{xi} HCO_3^- is a by product of Boc-deprotection **Ac- α -Lys(Boc)-S-Na $^+$** was observed in the ^{13}C NMR spectrum at 161.7 ppm. See Supplementary Reference 59 for further details.

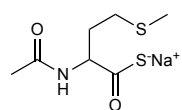


Supplementary Figure 269. ^1H NMR (700 MHz, D_2O , 1.0–5.0 ppm) spectrum of 2-acetamido-6-aminohexanethioic S-acid **Ac-α-Lys-SH**

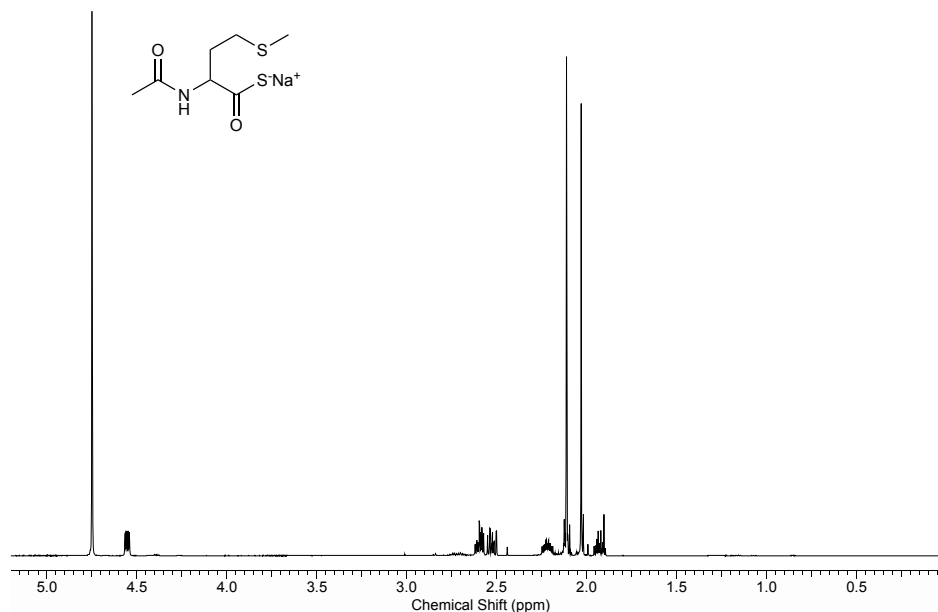


Supplementary Figure 270. ^{13}C NMR (176 MHz, D_2O , 0–230 ppm) spectrum of 2-acetamido-6-aminohexanethioic S-acid **Ac-α-Lys-SH**

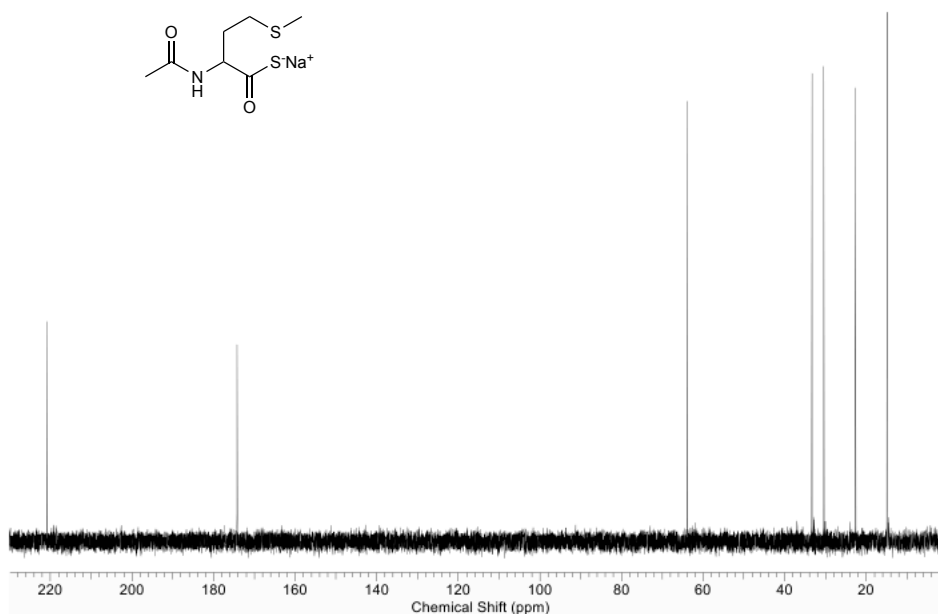
Sodium 2-acetamido-4-(methylthio)butanethioate **Ac-Met-S⁻Na⁺**



Prepared following *general procedure I* from 2-methyl-4-(2-(methylthio)ethyl)oxazol-5(4*H*)-one **S2-Met** (252 mg, 1.46 mmol). **Ac-Met-S⁻Na⁺** was isolated as a white foam (320 mg). ¹H NMR (600 MHz, D₂O) δ 4.55 (dd, *J* = 4.3, 9.2 Hz, 1H, (C2)-H), 2.63 - 2.57 (m, 1H, (C3)-H), 2.56 - 2.50 (m, 1H, (C3')-H), 2.25 - 2.18 (m, 1H, (C4)-H), 2.11 (s, 3H, S-CH₃), 2.03 (s, 3H, (COCH₃), 1.96 - 1.90 (m, 1H, (C4')-H). ¹³C NMR (151 MHz, D₂O) δ 220.8 (C1), 174.2 (COCH₃), 63.8 (C2), 33.2 (C4), 30.3 (C3), 22.6 (COCH₃), 14.8 (S-CH₃). HRMS-ESI [M+H]⁺ calcd. for C₇H₁₃NO₂S₂Na 230.0285; obs. 230.0275. IR (solid, cm⁻¹) 3294, 2968, 2916, 1635, 1517.

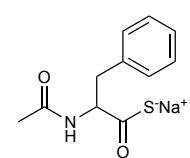


Supplementary Figure 271. ¹H NMR (600 MHz, D₂O, 0–5.2 ppm) spectrum for sodium 2-acetamido-4-(methylthio)butanethioate **Ac-Met-S⁻Na⁺**

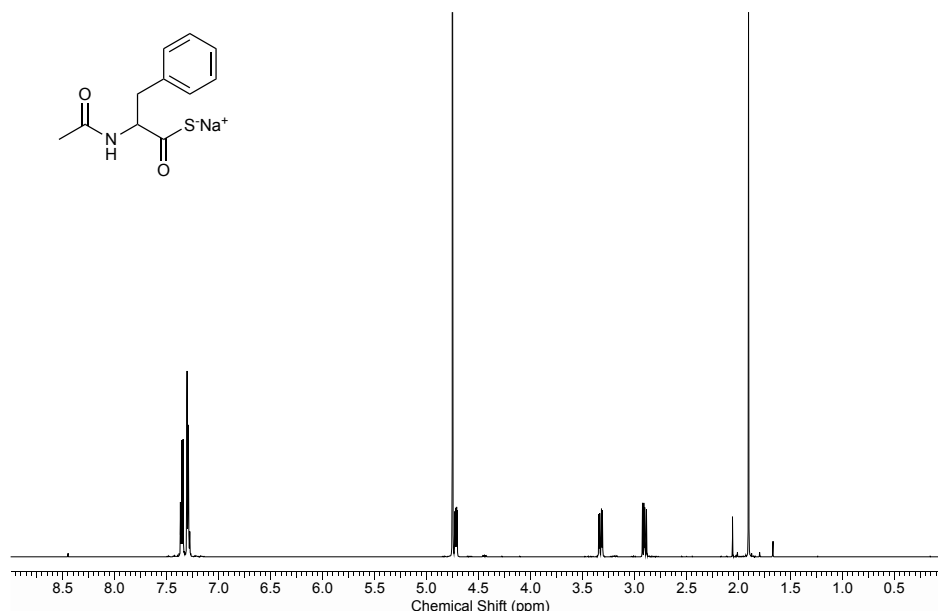


Supplementary Figure 272. ¹³C NMR (151 MHz, D₂O, 0–230 ppm) spectrum for sodium 2-acetamido-4-(methylthio)butanethioate **Ac-Met-S⁻Na⁺**

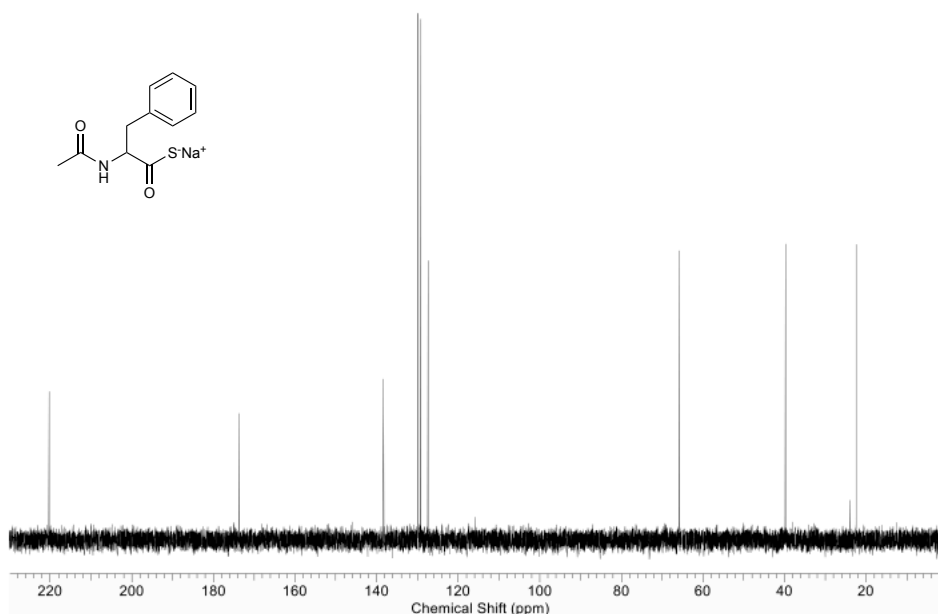
Sodium 2-acetamido-3-phenylpropanethioate **Ac-Phe-S⁻Na⁺**



Prepared following *general procedure I* from 4-benzyl-2-methyloxazol-5(4*H*)-one **S2-Phe** (189 mg, 1.00 mmol). **Ac-Phe-S⁻Na⁺** was isolated as a white solid (245 mg). ¹H NMR (600 MHz, D₂O) δ 7.38 - 7.33 (m, 2H, Ar), 7.32 - 7.27 (m, 3H, Ar), 4.72 (dd, *J* = 4.8, 9.3 Hz, 1H, (C2)-H), 3.33 (ABX, *J* = 4.8, 14.0 Hz, 1H, (C3)-H), 2.90 (ABX, *J* = 9.3, 14.0 Hz, 1H, (C3')-H), 1.90 (s, 3H, (COCH₃)). ¹³C NMR (151 MHz, D₂O) δ 220.3 (C1), 173.8 (COCH₃), 138.3 (Ar), 129.9 (Ar × 2), 129.2 (Ar × 2), 127.4 (Ar), 65.9 (C2), 39.8 (C3), 22.5 (COCH₃). HRMS-ESI [M+H]⁺ calcd. for C₁₁H₁₃NO₂Na 246.0565; obs. 246.0560. IR (solid, cm⁻¹) 3298, 1653, 1535.

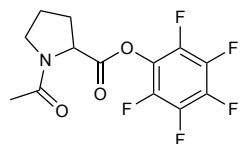
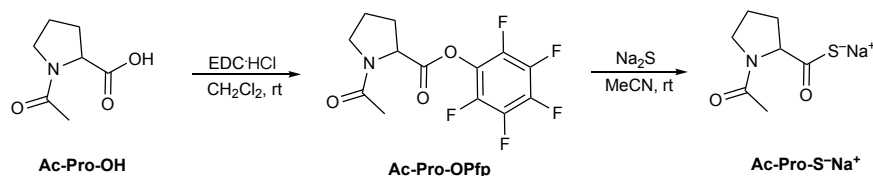


Supplementary Figure 273. ¹H NMR (600 MHz, D₂O, 0–9.0 ppm) spectrum of sodium 2-acetamido-3-phenylpropanethioate **Ac-Phe-S⁻Na⁺**



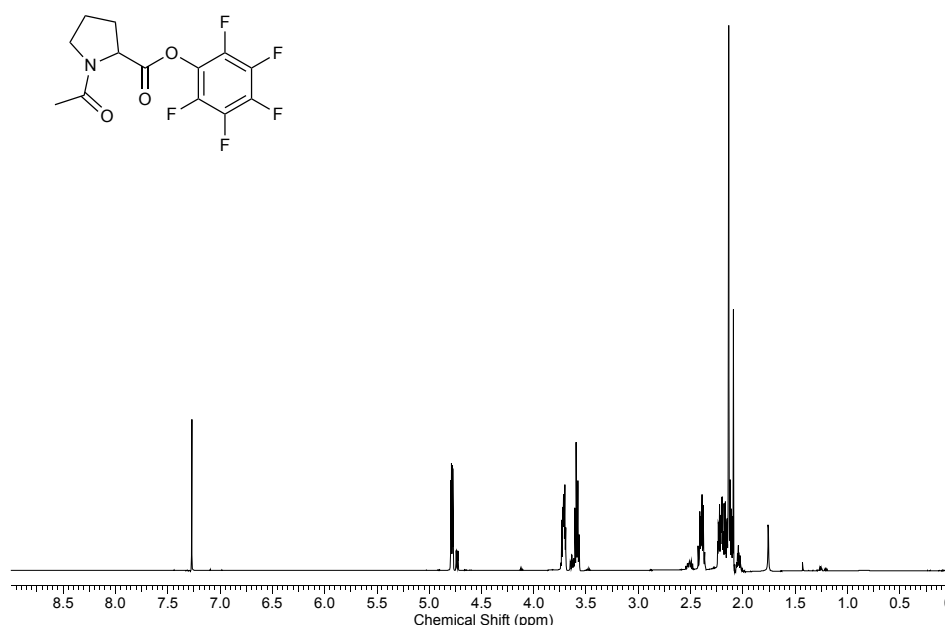
Supplementary Figure 274. ¹³C NMR (151 MHz, D₂O, 0–230 ppm) spectrum of sodium 2-acetamido-3-phenylpropanethioate **Ac-Phe-S⁻Na⁺**

Sodium *N*-acetylpyrrolidine-2-carbothioate **Ac-Pro-S⁻Na⁺**

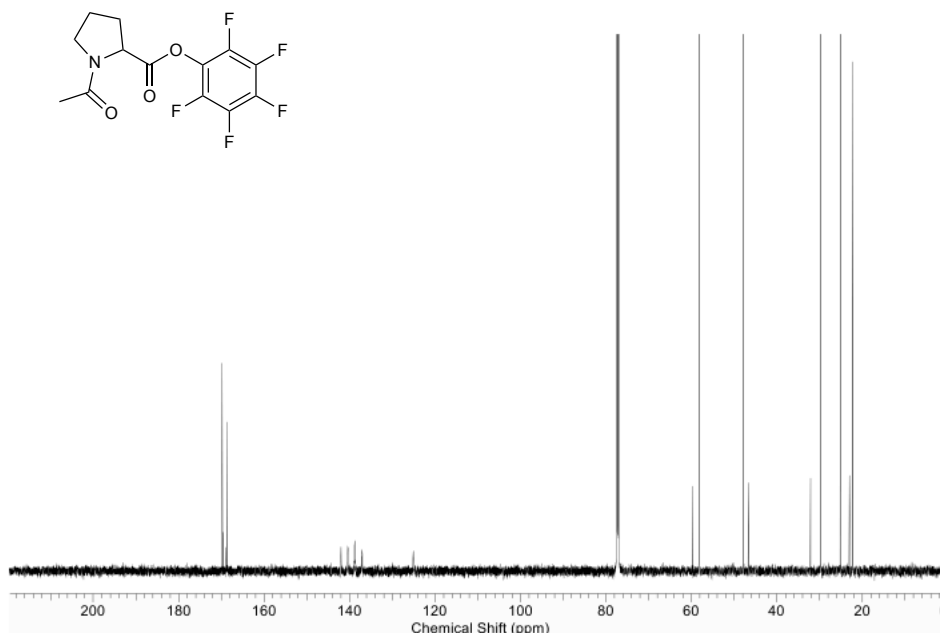


N-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC·HCl) (3.45 g, 18.00 mmol) was added to a stirring solution of *N*-acetyl-L-proline **Ac-Pro-OH** (943 mg, 6.00 mmol) and pentafluorophenol (1.22 g, 6.60 mmol) in CH₂Cl₂ (25 mL) at 0 °C. The reaction mixture was allowed to warm to room temperature and stirred for 16 h before concentrating *in vacuo*. The

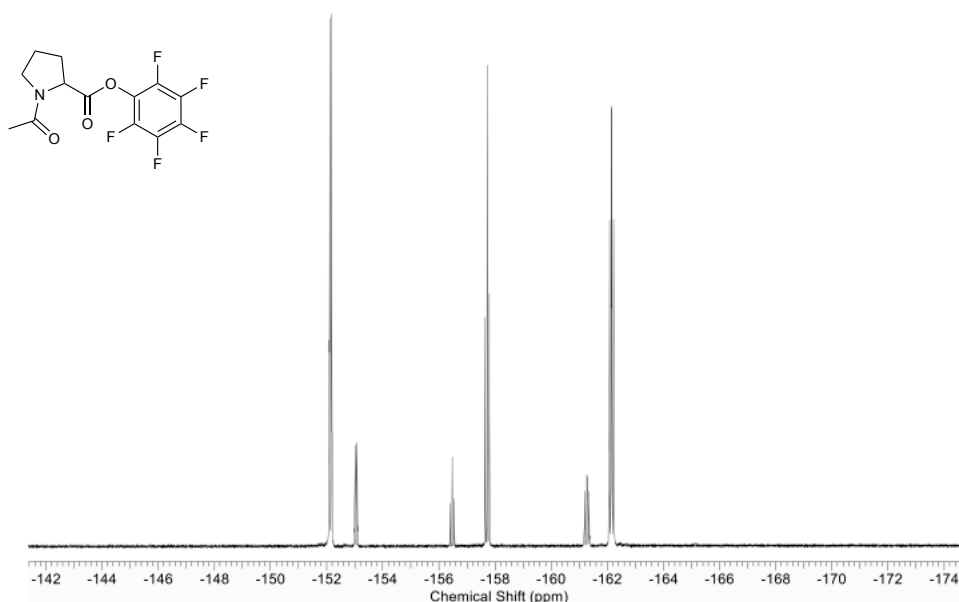
crude mixture was triturated with diethyl ether (10 mL), the supernatant immobilized onto silica, and purified by flash column chromatography (SiO₂; petroleum ether (40-60)/diethyl ether (100:0 to 0:100)) to give pentafluorophenyl-*N*-acetyl-prolinate **Ac-Pro-OPfp** (1.73 g, 89%) as a colourless oil. *R*_f = 0.33 (100% diethyl ether). ¹H-NMR (600 MHz, CDCl₃, mixture of 2 rotamers (a/b 85:15)) Rotamer a: δ 4.78 (dd, *J* = 4.0, 8.8 Hz, 1H, (C2)-H), 3.75 - 3.68 (m, 1H, (C5)-H), 3.61 - 3.55 (m, 1H, (C5')-H), 2.43 - 2.34 (m, 1H, (C3)-H), 2.25 - 2.20 (m, 1H, (C3')-H), 2.19 - 2.15 (m, 1H, (C4)-H), 2.13 (s, 3H, COCH₃), 2.13 - 2.10 (m, 1H, (C4')-H). Rotamer b: δ 4.73 (dd, *J* = 2.6, 8.8 Hz, 1H, (C2)-H), 3.76 - 3.68 (m, 1H, (C5)-H), 3.67 - 3.61 (m, 1H, (C5')-H), 2.56 - 2.47 (m, 1H, (C3)-H), 2.44 - 2.37 (m, 1H, (C3')-H), 2.25 - 2.18 (m, 1H, (C4)-H), 2.09 (s, 3H, COCH₃), 2.08 - 2.00 (m, 1H, (C4')-H). ¹³C NMR (151 MHz, CDCl₃) Rotamer a: δ 169.8 (COCH₃), 168.7 (C1), 141.9-136.8 (4 x Ar), 124.8 (Ar), 58.2 (C2), 47.8 (C5), 29.7 (C3), 25.1 (C4), 22.3 (COCH₃). Rotamer b: δ 169.6 (COCH₃), 168.9 (C1), 141.9-136.8 (4 x Ar), 124.8 (Ar), 59.8 (C2), 46.6 (C5), 32.2 (C3), 23.0 (C4), 22.4 (COCH₃). ¹⁹F-NMR (282 MHz, CDCl₃) Rotamer a: δ -152.3 (2F, m), -157.9 (1F, t, *J* = 21.7 Hz), -162.4 (2F, m). Rotamer b: δ -153.3 (2F, m), -156.7 (1F, t, *J* = 17 Hz), -161.5 (2F, m). HRMS-ESI [*M*+H]⁺ calcd. for C₁₃H₁₁F₅NO₃ 324.0659; obs. 324.0647. IR (oil, cm⁻¹) 2980, 2887, 2667, 1793, 1724, 1602, 1533, 1512, 1471, 1456, 1425.



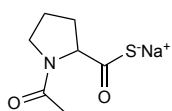
Supplementary Figure 275. ¹H NMR (600 MHz, CDCl₃, 0–9.0 ppm) spectrum of pentafluorophenyl-*N*-acetyl-prolinate **Ac-Pro-OPfp**



Supplementary Figure 276. ^{13}C NMR (151 MHz, CDCl_3 , 0–220 ppm) spectrum of pentafluorophenyl-*N*-acetyl-prolinate **Ac-Pro-OPfp**

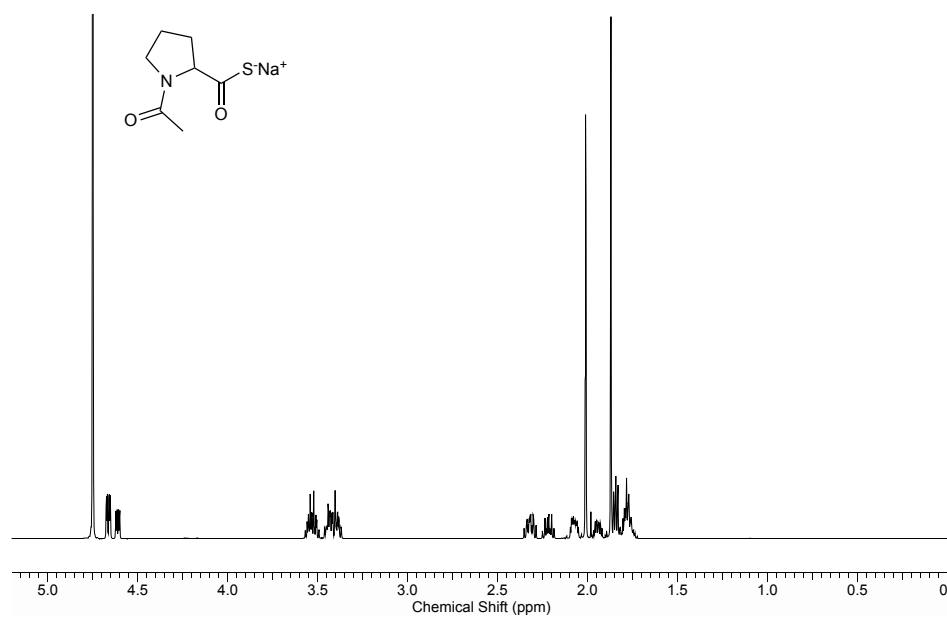


Supplementary Figure 277. ^{19}F NMR (282 MHz, CDCl_3 , -174– -142 ppm) spectrum of pentafluorophenyl-*N*-acetyl-prolinate **Ac-Pro-OPfp**

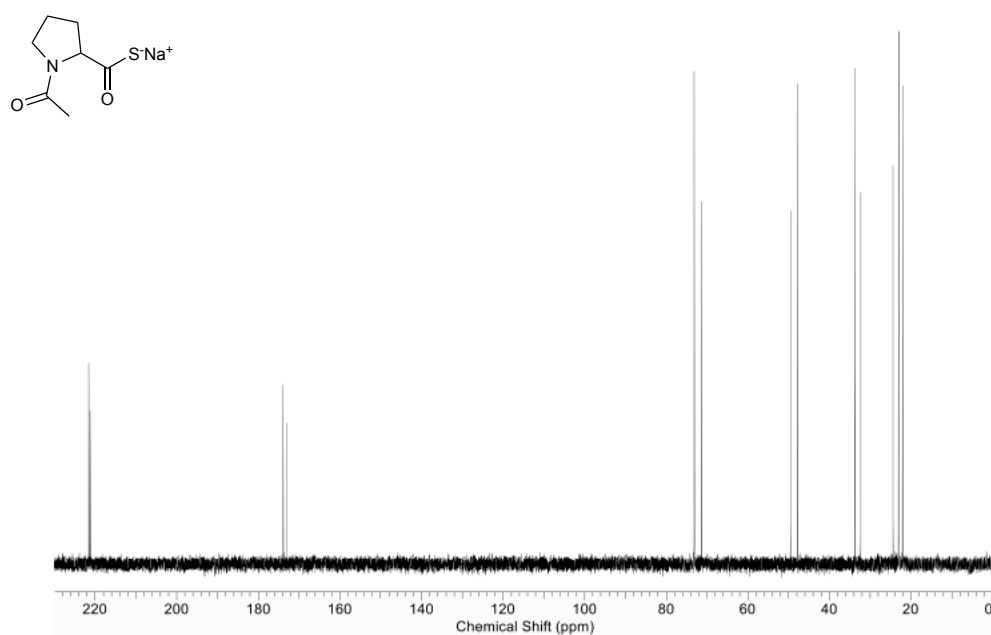


To a solution of pentafluorophenyl-*N*-acetyl-prolinate **Ac-Pro-OPfp** (162 mg, 0.50 mmol) in anhydrous acetonitrile (2.5 mL) was added finely ground anhydrous sodium sulfide (43 mg, 0.55 mmol) at room temperature. The suspension was stirred under an argon atmosphere for 16 h. The resultant precipitate was isolated by centrifugation, washed with diethyl ether (1 mL) and dried *in vacuo* to afford sodium *N*-acetylpyrrolidine-2-carbithioate **Ac-Pro-S-Na⁺** as a white solid (94 mg) that was used without further purification. ^1H NMR (600 MHz, D_2O , mixture of 2 rotamers (a/b 57:43)). Rotamer a: δ 4.66 (dd, J = 3.4, 8.9 Hz, 1H, (C2)–H), 3.47 - 3.36 (m, 2H, (C5)–H₂), 2.35 - 2.29 (m, 1H, (C3)–H), 2.10 - 2.04 (m, 1H, (C3')–H), 1.87 (s, 3H, COCH_3), 1.81 - 1.74 (m, 2H, (C4)–H₂). Rotamer b: δ 4.61 (dd, J = 4.1, 8.9 Hz, 1H, (C2)–H), 3.59 - 3.48 (m, 2H, (C5)–H₂), 2.25 - 2.19 (m, 1H, (C3)–H), 2.01 (s, 3H, COCH_3), 1.97 - 1.91 (m, 1H, (C3')–H), 1.86 - 1.82 (m, 2H, (C4)–H₂). ^{13}C NMR (151 MHz, D_2O , mixture of 2 rotamers (a/b 57:43)). Rotamer a: δ 221.5 (C1), 173.9 (COCH_3), 73.1 (C2), 47.8 (C5), 33.7 (C3), 23.0 (C4), 22.0 (COCH_3). Rotamer b: δ 221.2 (C1), 173.0 (COCH_3), 71.4 (C2), 49.5 (C5), 32.3 (C3), 24.4 (C4), 22.1 (COCH_3).

HRMS-ESI $[M+H]^+$ calcd. for $C_7H_{11}NO_2SNa$ 195.0330; obs. 195.0333. IR (solid, cm^{-1}) 2974, 2947, 2873, 1635, 1608, 1527, 1500.

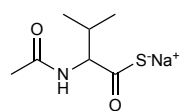


Supplementary Figure 278. ¹H NMR (600 MHz, D₂O, 0–5.2 ppm) spectrum of sodium *N*-acetylpyrrolidine-2-carbothioate **Ac-Pro-S⁻Na⁺**

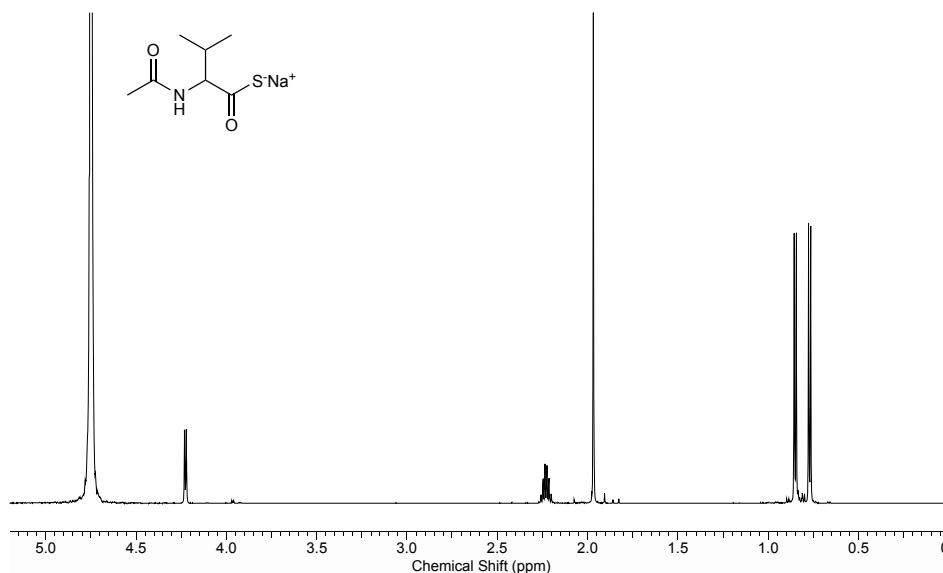


Supplementary Figure 279. ¹³C NMR (151 MHz, D₂O, 0–230 ppm) spectrum of sodium *N*-acetylpyrrolidine-2-carbothioate **Ac-Pro-S⁻Na⁺**

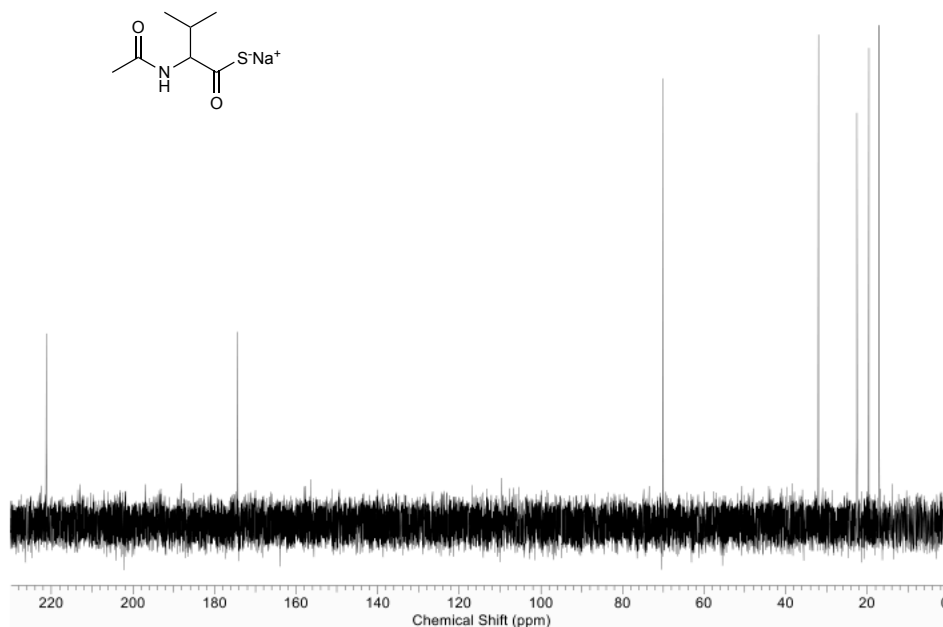
Sodium 2-acetamido-3-methylbutanethioate **Ac-Val-S⁻Na⁺**



Prepared following *general procedure I* from 4-isobutyl-2-methyloxazol-5(4*H*)-one **S2-Val** (44 mg, 0.31 mmol) and isolated as a white solid (61 mg). ¹H NMR (600 MHz, D₂O) δ 4.31 (d, *J* = 5.9 Hz, 1H, (C2)–H), 2.35 - 2.26 (m, 1H, (C3)–H), 2.05 (s, 3H, (COCH₃), 0.93 (d, *J* = 6.9 Hz, 3H, (C4)–H₃), 0.86 (d, *J* = 6.9 Hz, 3H, (C4')–H₃). ¹³C NMR (151 MHz, D₂O) δ 221.2 (C1), 174.3 (COCH₃), 70.1 (C2), 32.0 (C3), 22.6 (COCH₃), 19.6 (C4), 17.1 (C4'). HRMS-ESI [M+H]⁺ calcd. for C₇H₁₃NO₂SNa 198.0565; obs. 198.0566. IR (solid, cm⁻¹). 3309, 2962, 2929, 1653, 1624, 1597, 1548, 1519, 1467, 1442.



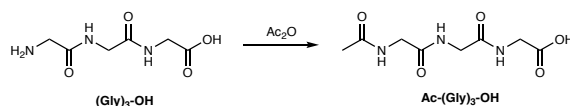
Supplementary Figure 280. ¹H NMR (600 MHz, D₂O, 0–5.2 ppm) spectrum of sodium 2-acetamido-3-methylbutanethioate **Ac-Val-S⁻Na⁺**



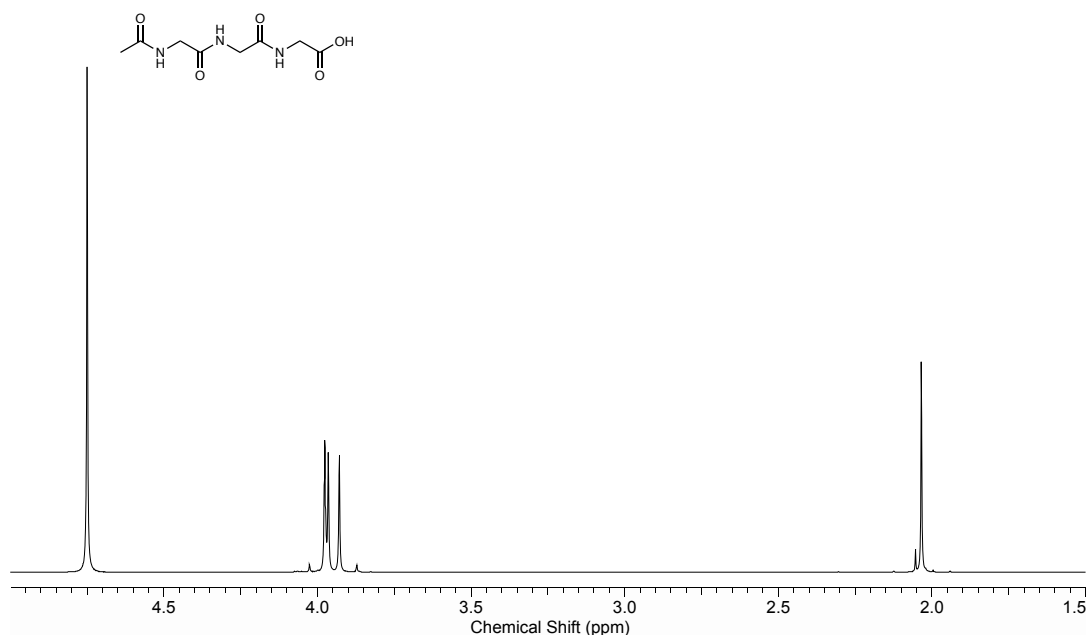
Supplementary Figure 281. ¹³C NMR (151 MHz, D₂O, 0–230 ppm) spectrum of sodium 2-acetamido-3-methylbutanethioate **Ac-Val-S⁻Na⁺**

Synthesis of 2-(2-(2-acetamidoacetamido)acetamido)ethanethioic S-acid **Ac-(Gly)₃-SH**

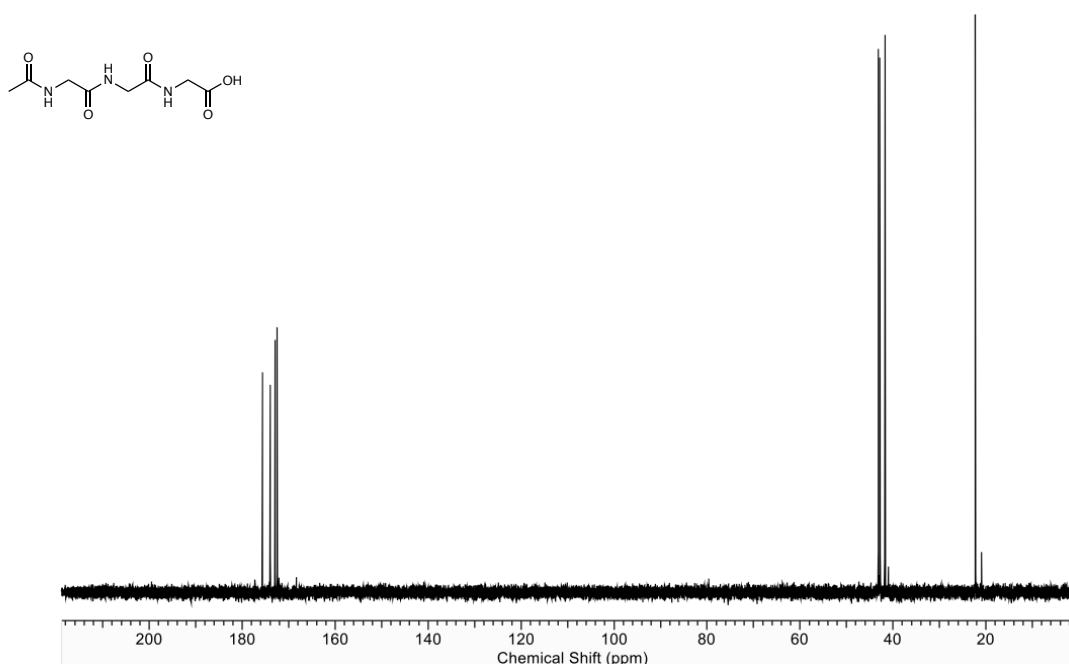
N-Acetylglycylglycylglycine **Ac-(Gly)₃-OH**



Triglycine (**(Gly)₃-OH**) (1.00 g, 5.29 mmol) and acetic anhydride (1.00 mL) were sonicated in water (7 mL) for 10 min at room temperature. The reaction mixture was concentrated *in vacuo* and co-evaporated with toluene (25 mL), followed by diethyl ether (25 mL) to give *N*-acetylglycylglycylglycine **Ac-(Gly)₃-OH** (1.22 g, quantitative) as a white powder. ^1H NMR (700 MHz, D_2O) δ 3.98 (s, 2H, Gly₁-(C2)-H₂), 3.97–3.93 (2 × s, 4H, Gly₂-(C2)-H₂, Gly₃-(C2)-H₂), 2.03 (s, 3H, (COCH₃)). ^{13}C NMR (176 MHz, D_2O) δ 175.7 (COCH₃), 174.0 (Gly₁-(C1)), 172.9, 172.5 (Gly₂-(C1), Gly₃-(C1)), 43.2, 42.8 (Gly₂-(C2), (Gly₃-(C2))), 41.7 (Gly₁-(C2)), 22.3 (COCH₃). HRMS-ESI $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_8\text{H}_{14}\text{N}_3\text{O}_5$ 232.0928; obs. 232.0927.

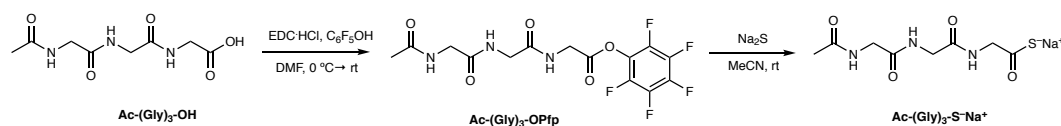


Supplementary Figure 282. ^1H NMR (700 MHz, D_2O , 0.00 – 9.00 ppm) spectrum to show *N*-acetylglycylglycylglycine **Ac-(Gly)₃-OH**.

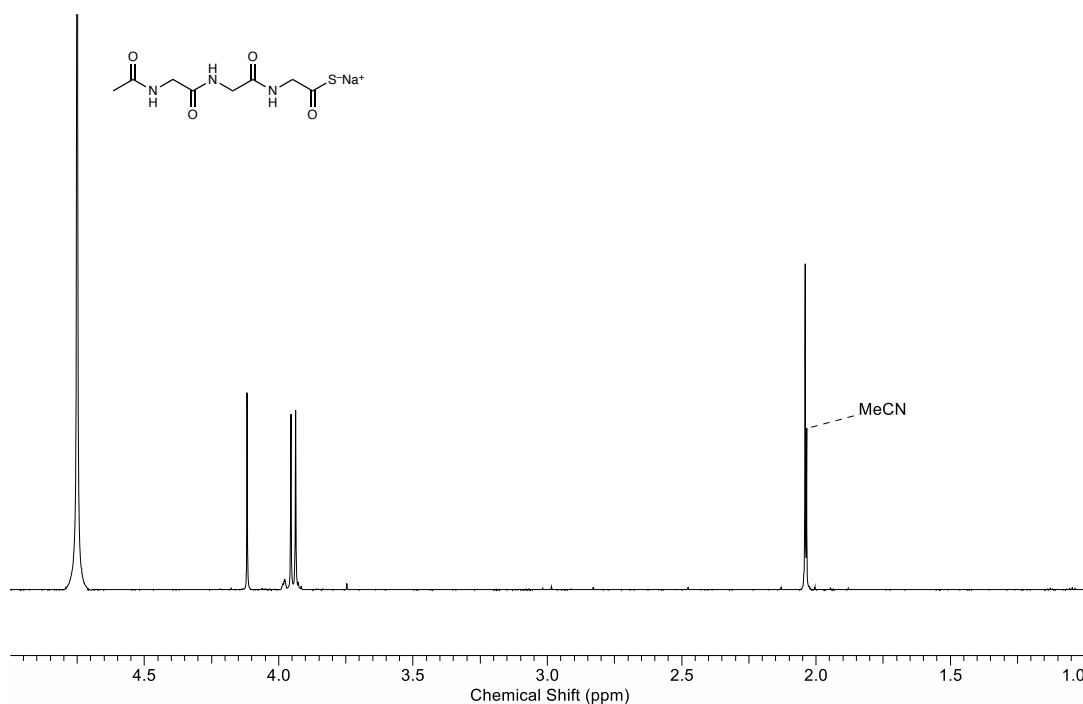


Supplementary Figure 283. ^{13}C NMR (176 MHz, D_2O , 0 – 220 ppm) spectrum to show *N*-acetylglycylglycylglycine **Ac-(Gly)₃-OH**.

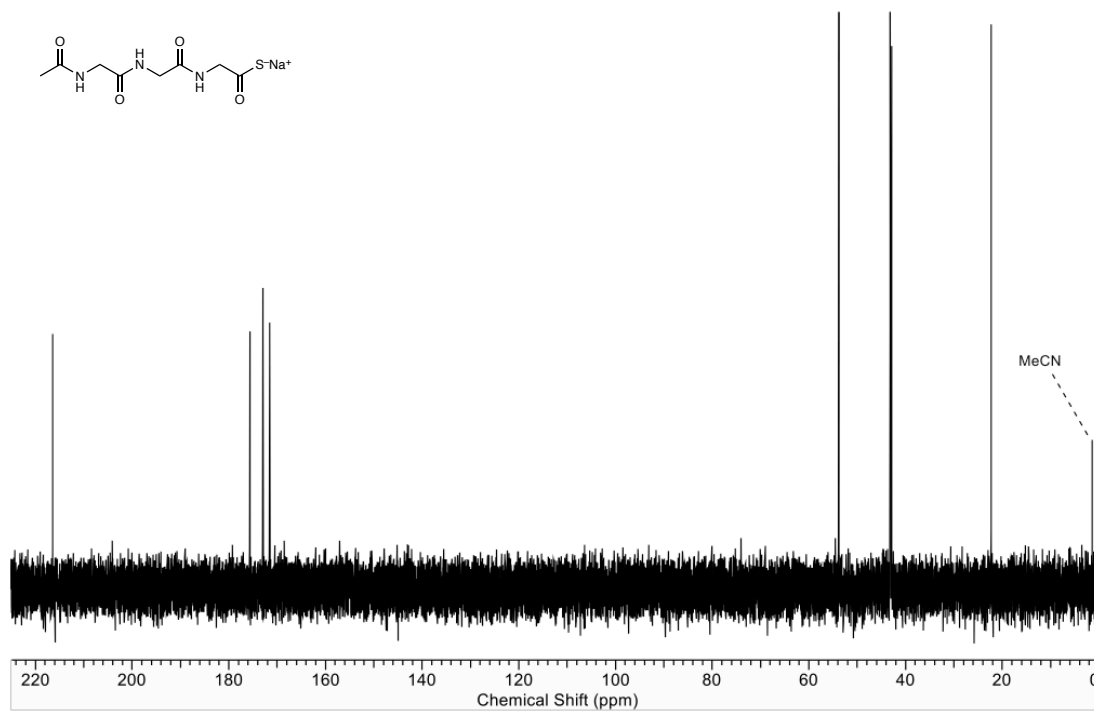
Sodium 2-(2-(2-acetamidoacetamido)acetamido)ethanethioate **Ac-(Gly)₃-S⁻Na⁺**



N-Acetylglycylglycylglycine (244 mg, 1.06 mmol) and EDC·HCl (521 mg, 2.72 mmol) were stirred in anhydrous DMF (15 mL) in an ice bath for 1 h. Pentafluorophenol (466 mg, 2.40 mmol) was added and the reaction was allowed to warm to room temperature and stirred vigorously for 16 h. The resulting suspension was concentrated *in vacuo* and co-evaporated with toluene. The residue was triturated with dry acetonitrile (2 × 10 mL) and diethyl ether (10 mL) and dried *in vacuo*. The crude pentafluorophenol ester **Ac-(Gly)₃-OPfp** was resuspended in dry acetonitrile (15 mL) and stirred vigorously with anhydrous sodium sulfide (91 mg, 1.16 mmol) for 24 h. The precipitate was isolated by filtration, washed with diethyl ether (2 × 10 mL), and dried *in vacuo* to give sodium 2-(2-(2-acetamidoacetamido)acetamido)ethanethioate **Ac-(Gly)₃-S⁻Na⁺** (227 mg) as a light pink powder, which was used without further purification. ¹H NMR (700 MHz, D₂O) δ 4.12 (s, 2H, Gly₃-(C2)-H₂), 3.96 (s, 2H, Gly₂-(C2)-H₂), 3.94 (s, 2H, Gly₁-(C2)-H₂), 2.04 (s, 3H, (COCH₃)). ¹³C NMR (176 MHz, D₂O) δ 216.4 (Gly₃-(C1)), 175.6 (COCH₃), 172.9 (Gly₂-(C1)), 171.6 (Gly₁-(C1)), 53.8 (Gly₃-(C2)), 43.2, 42.9 (Gly₁-(C2), Gly₂-(C2)), 22.3 (COCH₃). HRMS-ESI [M+Na]⁺ calcd. for C₈H₁₂N₃O₄SNa 270.0525; obs. 232.0518.



Supplementary Figure 284. ¹H NMR (700 MHz, D₂O, 1.00 – 5.00 ppm) spectrum to show sodium 2-(2-(2-acetamidoacetamido)acetamido)ethanethioate **Ac-(Gly)₃-S⁻Na⁺**.



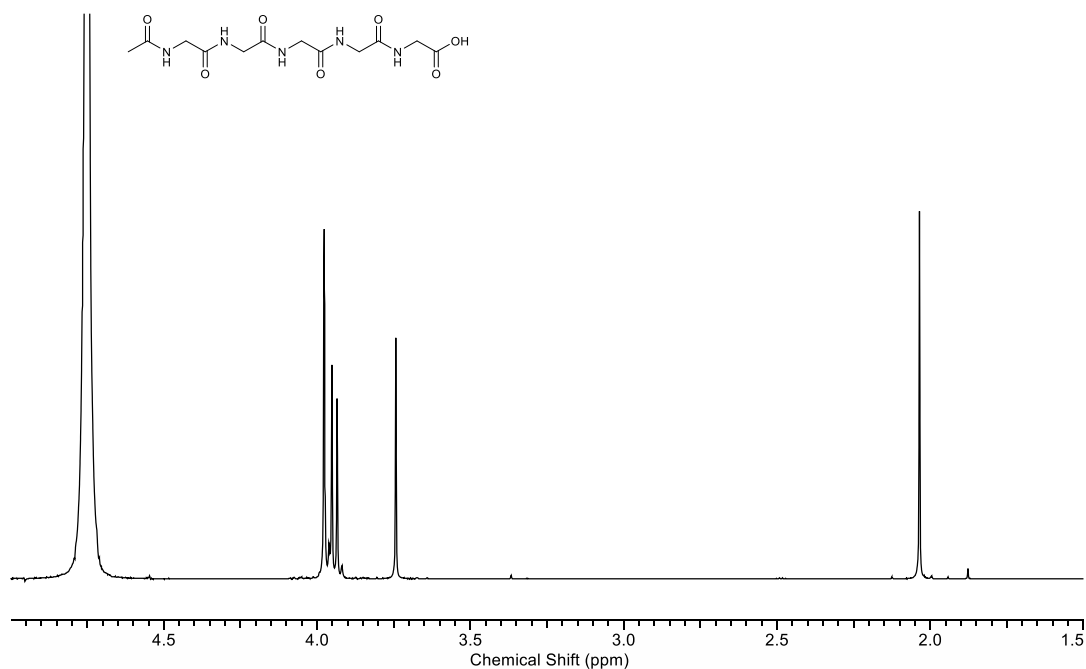
Supplementary Figure 285. ¹³C NMR (176 MHz, D₂O, 0 – 225 ppm) spectrum to show sodium 2-(2-(2-acetamidoacetamido)acetamido)ethanethioate **Ac-(Gly)₃-S⁻Na⁺**.

Synthesis of **Ac-(Gly)₅-SH**

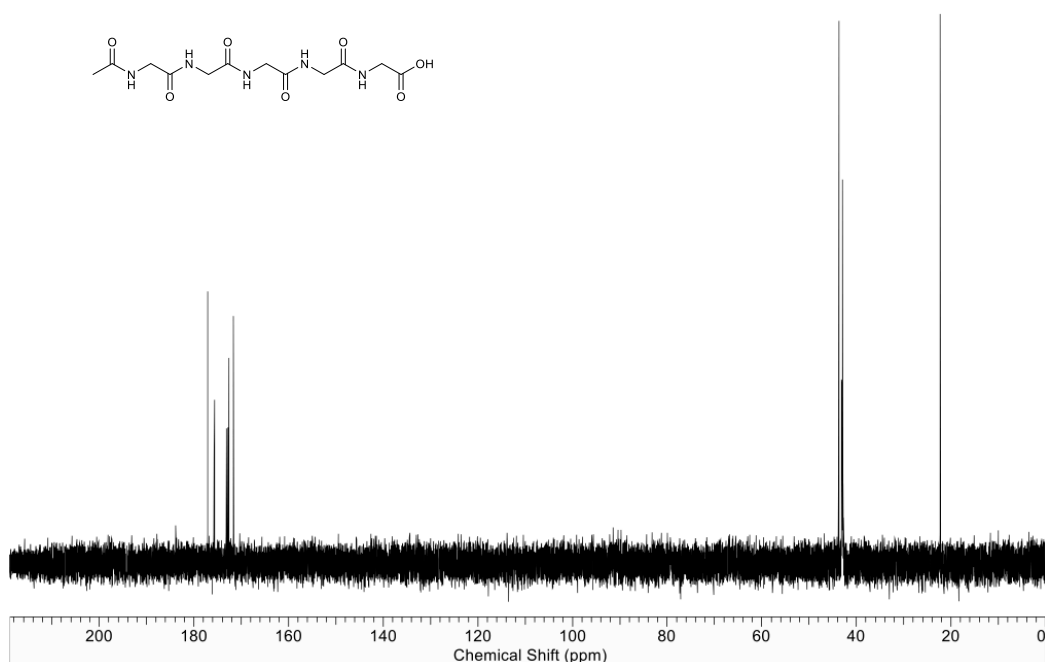
Ac-Gly₅-OH



Pentaglycine (320 g, 1.06 mmol) was stirred in acetic anhydride (2 mL) and water (1.5 mL) for 10 min. The suspension was then sonicated for 20 mins. The suspension was then heated to 80 °C for 30 min, cooled to 40 °C and stirred for 48 h before concentrating *in vacuo* and co-evaporating with toluene (20 mL) to give **Ac-(Gly)₅-OH** as a white amorphous solid (364 mg) that was used without further purification. ¹H NMR (700 MHz, D₂O, pD 12) δ 3.98 (s, 4H), 3.95 (s, 2H), 3.93 (s, 2H), 3.74 (s, 2H), 2.06 - 1.97 (m, 3H). ¹³C NMR (176 MHz, D₂O, pD 12) δ 177.1, 175.7, 173.1, 172.9, 172.6, 171.6, 43.7, 43.2, 43.1, 43.0, 43.0, 22.3. HRMS-ESI [M+Na]⁺ calcd. for C₁₂H₁₈N₅O₇Na 368.1182; obs. 368.1177.

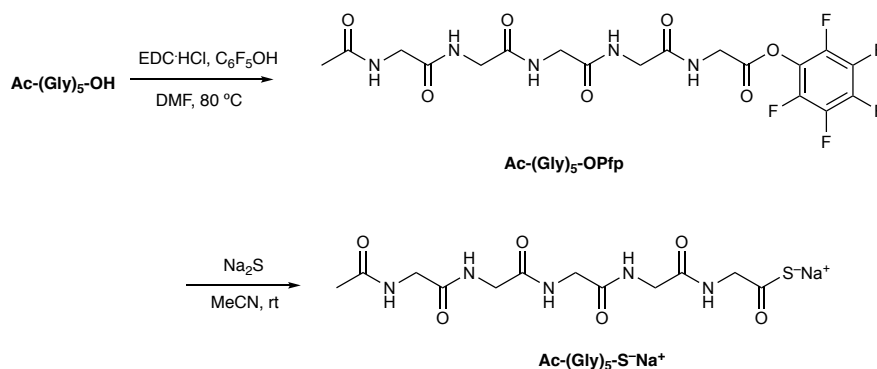


Supplementary Figure 286. ¹H NMR (700 MHz, D₂O, pD 12, 1.50 – 5.00 ppm) spectrum to show **Ac-(Gly)₅-OH**.

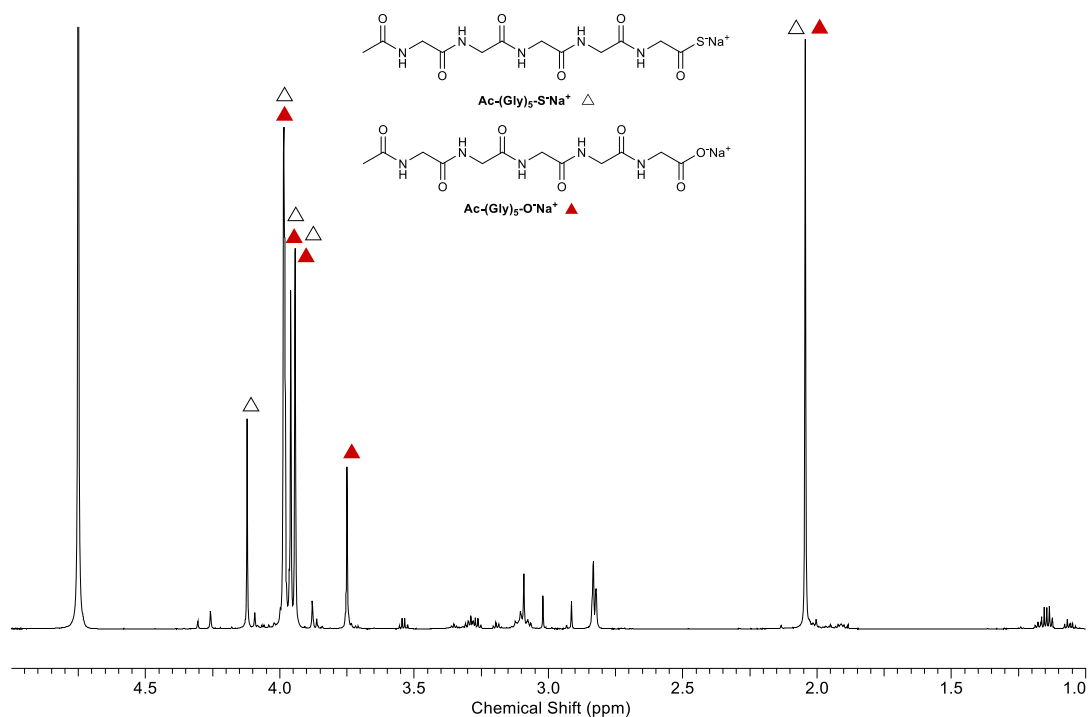


Supplementary Figure 287. ¹³C NMR (176 MHz, D₂O, pD 12, 0 – 220 ppm) spectrum to show **Ac-(Gly)₅-OH**.

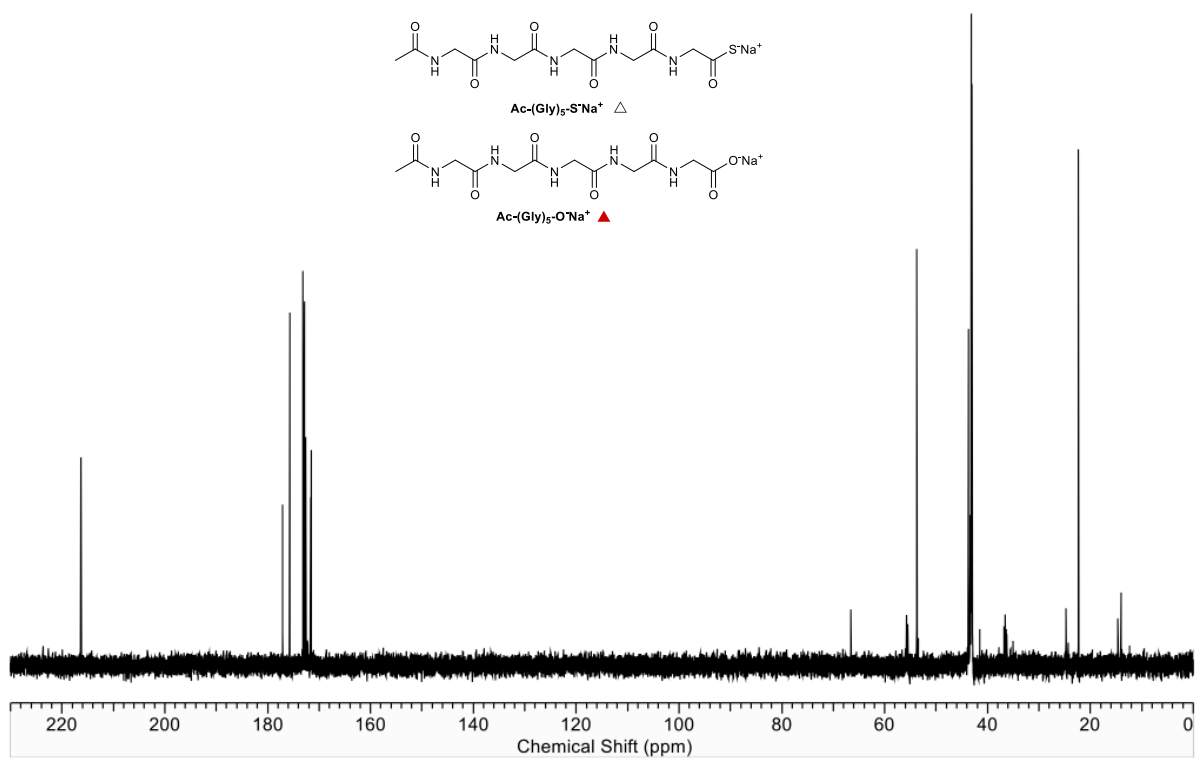
Ac-(Gly)₅-SH



Crude **Ac-(Gly)₅-OH** (364 mg), pentafluorophenol (466 mg, 2.40 mmol) and EDC·HCl (521 mg, 2.72 mmol) were stirred as a slurry in anhydrous DMF (15 mL) at 80 °C for 17 h. The slurry was cooled, concentrated *in vacuo* and co-evaporated with toluene (20 mL). The residue was re-suspended in dichloromethane (10 mL), centrifuged, and the supernatant discarded. This process was repeated twice, and again with diethyl ether (2 x 10 mL). The residue was dried *in vacuo* to give an off-white amorphous solid (484 mg). The crude pentafluorophenol ester **Ac-(Gly)₅-OPfp** was resuspended in anhydrous acetonitrile (15 mL) and stirred vigorously with anhydrous sodium sulfide (80 mg, 1.03 mmol) for 20 h. The slurry was centrifuged and the supernatant removed. The precipitate was re-suspended in MeCN (10 mL), centrifuged, and the supernatant removed again. This process was repeated with dichloromethane (2 x 10 mL) and diethyl ether (2 x 10 mL). The residue was dried *in vacuo* to give a white amorphous powder (438 mg). NMR spectroscopic analysis against MSM as an internal reference standard revealed that **Ac-(Gly)₅-S⁻Na⁺** (△) constituted 33% of the bulk material, alongside **Ac-(Gly)₅-O⁻Na⁺** (29%, ▲). The crude **Ac-(Gly)₅-S⁻Na⁺** was used without further purification. ¹H NMR (700 MHz, D₂O) **Ac-(Gly)₅-S⁻Na⁺** (△) δ 4.12 (s, 2H), 3.99 (s, 4H), 3.96 (s, 2H), 3.94 (s, 2H), 2.04 (s, 3H). **Ac-(Gly)₅-O⁻Na⁺** (▲) δ 3.99 (s, 4H), 3.96 (s, 2H), 3.94 (s, 2H), 3.75 (s, 2H), 2.04 (s, 3H). ¹³C NMR (176 MHz, D₂O, characterised as a mixture, partial assignment given) δ 216.3 (△, Gly₅-(C1)), 177.1 (▲, Gly₅-(C1)), 175.7 (CO), 173.1 (CO × 2), 172.9 (CO × 2), 172.6 (CO × 2), 172.6 (CO), 171.6 (CO), 171.5 (CO), 53.8 (△, Gly₅-(C2)), 43.7 (▲, Gly₅-(C2)), 43.4 (Gly-(C2)), 43.3 (Gly-(C2)), 43.2 (Gly-(C2)), 43.1 (Gly-(C2)), 43.1 (Gly-(C2)), 43.0 (Gly-(C2)), 43.0 (Gly-(C2)), 42.9 (Gly-(C2)), 22.3 (COCH₃), 22.3 (COCH₃). HRMS-ESI [M+Na]⁺ calcd. for C₁₂H₁₈N₅O₆SNa 384.0954; obs. 384.0934.



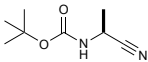
Supplementary Figure 288. ^1H NMR (700 MHz, D_2O , 1.00 – 5.00 ppm) spectrum to show $\text{Ac-(Gly)}_5\text{-S}^-\text{Na}^+$.

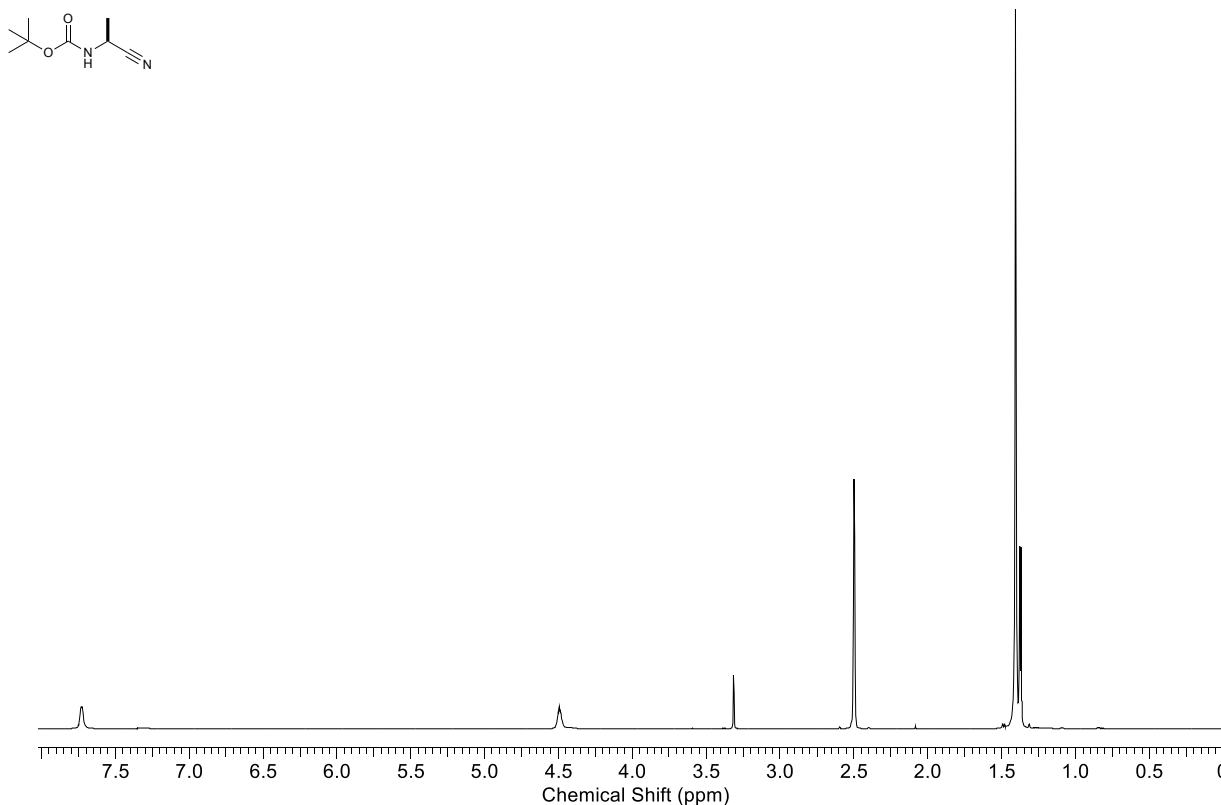


Supplementary Figure 289. ^{13}C NMR (176 MHz, D_2O , 0 – 230 ppm) spectrum to show $\text{Ac-(Gly)}_5\text{-S}^-\text{Na}^+$.

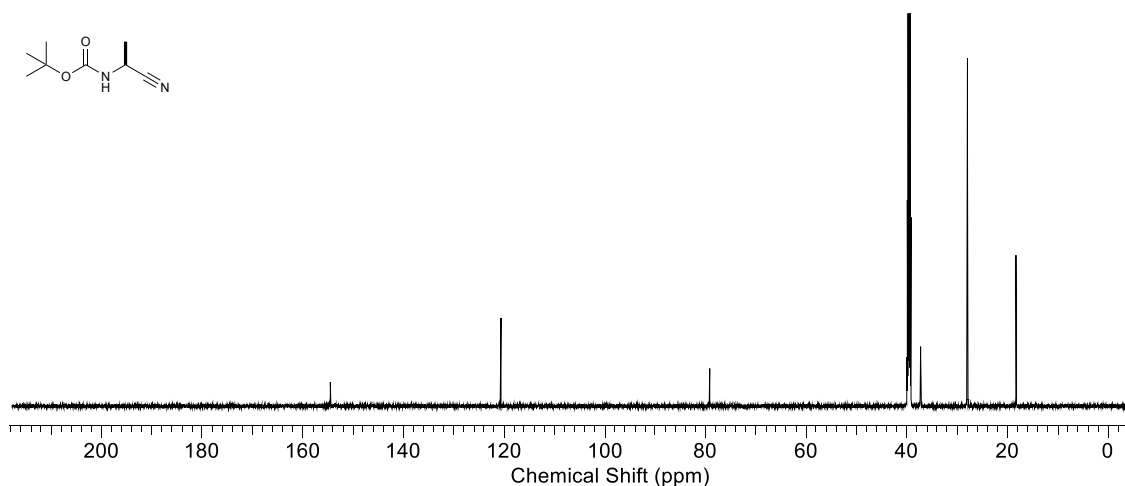
Stereoretentive prebiotic α -amidothioacid synthesis and ligation

N-Boc-L-alanine nitrile **Boc-L-Ala-CN**

 To a solution of L-alaninamide hydrochloride **L-Ala-NH₂-HCl** (5.00 g, 40.14 mmol) and triethylamine (20.3 g, 201 mmol, 28 mL) in THF (400 mL) 0 °C was added di-*tert*-butyl dicarbonate (13.1 g, 60.2 mmol) portion-wise over 10 minutes. The reaction was stirred at 0 °C for 2 h. Trifluoroacetic anhydride (25.3 g, 120.4 mmol, 17 mL) was then added dropwise over 30 min. The reaction was stirred at 0 °C for 1 h, and then at room temperature for 2 h. The reaction was concentrated to dryness *in vacuo*, dissolved water (200 mL) and extracted with diethyl ether (5 × 200 mL). The combined organic layers were washed with 0.1 M HCl (3 × 200 mL), saturated potassium carbonate (3 × 200 mL) and brine (200 mL). The organic layer was concentrated *in vacuo*. The crude residue was recrystallized from diethyl ether to give *N*-Boc-L-alanine nitrile **Boc-L-Ala-CN** as a white crystalline solid (4.87 g, 28.62 mmol, 71%). ¹H NMR (700 MHz, DMSO-*d*₆): δ 7.73 (br s, 1H, NH), 4.49 (br m, 1H, (C2)-H), 1.40 (s, 9H, 3 × (C)-CH₃), 1.37 (d, *J* = 7.1 Hz, 3H, (C2)-CH₃). M.p. 106.1 – 106.8 °C. ¹³C NMR (176 MHz, DMSO-*d*₆): δ 154.5 (CO), 120.7 (C1), 79.2 (C(CH₃)₃), 37.3 (C2), 28.0 (C(CH₃)₃), 18.3 (C2)-CH₃). LRMS (ESI⁺): 193.1 (100, [M + Na]⁺). HRMS (ESI⁺): calcd. for [C₈H₁₄N₂O₂Na]⁺: 193.0947; observed: 193.0945. IR (solid): 3339, 3316, 3005, 2978, 2946, 2192, 1737, 1704, 1677, 1524 cm⁻¹. M.p. 111.8 – 112.5 °C.



Supplementary Figure 290. ¹H NMR (700 MHz, DMSO-*d*₆, 0.00 – 8.00 ppm) spectrum to show *N*-Boc-L-alanine nitrile **Boc-L-Ala-CN**.

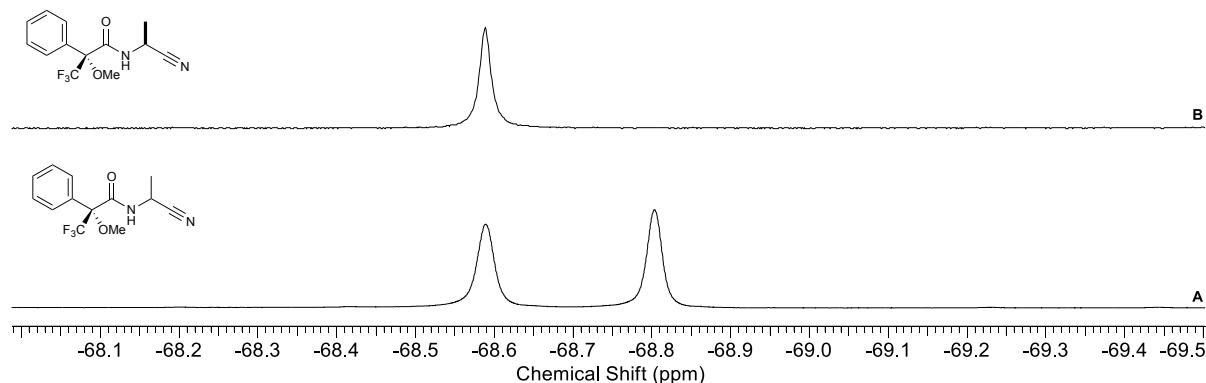


Supplementary Figure 291. ^{13}C NMR (176 MHz, $\text{DMSO}-d_6$, 0 – 220 ppm) spectrum to show *N*-Boc-L-alanine nitrile **Boc-L-Ala-CN**.

L-Alanine nitrile hydrochloride **L-Ala-CN·HCl**

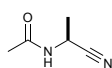
N-Boc-L-alanine nitrile **Boc-L-Ala-CN** (2.00 g, 11.75 mmol) was stirred in formic acid (50 mL) at room temperature for 1 h. The reaction was concentrated *in vacuo* and purified by flash column chromatography (SiO_2 ; EtOAc/petroleum ether 0:100 $\rightarrow$ 100:0 + 0.1% MeOH). Fractions containing L-alanine nitrile **L-Ala-CN** were collected and stirred at -78°C . 2M HCl in Et_2O (6 mL, 12.0 mmol) was added, and the resultant precipitate was filtered, washed with Et_2O (3×100 mL), and dried *in vacuo* to give L-alanine nitrile hydrochloride **L-Ala-CN·HCl** as a white powder (712 mg, 6.68 mmol, 51%). All NMR spectral data were found to be identical to those reported for *rac*-alanine nitrile hydrochloride **Ala-CN·HCl**.

The enantiomeric purity of L-alanine nitrile hydrochloride **L-Ala-CN·HCl** was quantified by derivatisation with (*R*)-(+)- α -methoxy- α -(trifluoromethyl)phenylacetyl chloride (Mosher's acid chloride) according to the following procedure. To a solution of L-alanine nitrile hydrochloride **L-Ala-CN·HCl** (10.7 mg, 0.10 mmol) and triethylamine (24.5 mg, 0.24 mmol, 33.5 μL) in DMSO (1 mL) was added (*R*)-(+)- α -methoxy- α -(trifluoromethyl)phenylacetyl chloride (126 mg, 0.50 mmol, 93.6 μL). The solution was stirred at room temperature for 10 min, and ^{19}F NMR data were acquired. The procedure was repeated using *rac*-alanine nitrile hydrochloride **Ala-CN·HCl** for comparison. The enantiomeric ratio of L-alanine nitrile hydrochloride **Ala-CN·HCl** was found to be >99:1, and that of *rac*-alanine nitrile hydrochloride **rac-Ala-CN·HCl** to be 50:50.



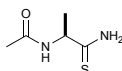
Supplementary Figure 292. ^{19}F NMR (376 Hz, $\text{DMSO}-d_6$, -68.0 – -69.5 ppm) to show the reaction of triethylamine (240 mM) and (*R*)-(+)- α -methoxy- α -(trifluoromethyl)phenylacetyl chloride (500 mM) at room temperature with **A**: *rac*-alanine nitrile hydrochloride **rac-Ala-CN·HCl** (100 mM); **B**: L-alanine nitrile hydrochloride **L-Ala-CN·HCl** (100 mM).

N-Acetyl-L-alanine nitrile **Ac-L-Ala-CN**

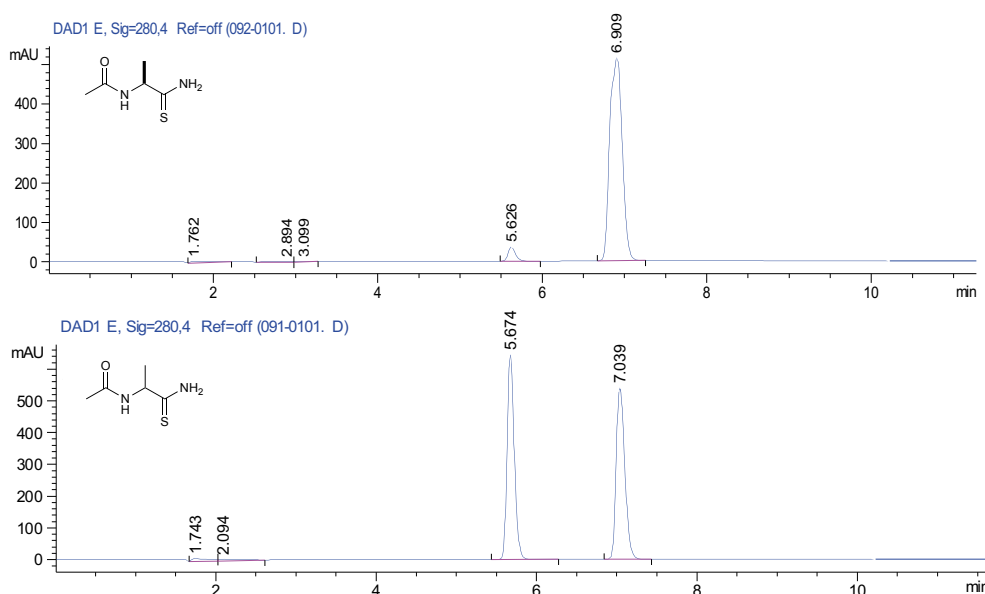


N-acetyl-L-alanine nitrile **Ac-L-Ala-CN** was prepared following *general procedure B* using L-alanine nitrile hydrochloride **L-Ala-CN·HCl** (1.07 g, 10 mmol). The reaction mixture was concentrated *in vacuo* and the residue purified by flash column chromatography (SiO₂; EtOAc/petroleum ether 0:100 → 100:0) to afford *N*-acetyl-L-alanine nitrile **Ac-L-Ala-CN** as a white crystalline solid (473 mg 4.22 mmol, 42%). All spectral data were found to be identical to those obtained for the racemic *N*-acetyl-L-alanine nitrile **Ac-Ala-CN**.

N-Acetyl-L-alanine thioamide **Ac-L-Ala-SNH₂**

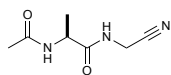


N-Acetyl-L-alanine thioamide **Ac-L-Ala-SNH₂** was prepared following *general procedure C* using *N*-acetyl-L-alanine nitrile (280.3 mg, 2.50 mmol). The reaction afforded *N*-acetyl-L-alanine thioamide **Ac-L-Ala-SNH₂** as a white solid (277.1 mg, 1.90 mmol, 76%). All NMR spectral data were found to be identical to those obtained for racemic *N*-acetyl-alanine thioamide **Ac-Ala-SNH₂**. The ratio of enantiomers was determined to be 98:2 by chiral HPLC analysis (Supplementary Figure 4) (YMC Chiral Cellulose-SB Classical Analytical S-5 column (250 x 4.6 mm); ethanol (100%), flow rate: 1.0 mL/min, UV detection λ = 280 nm).

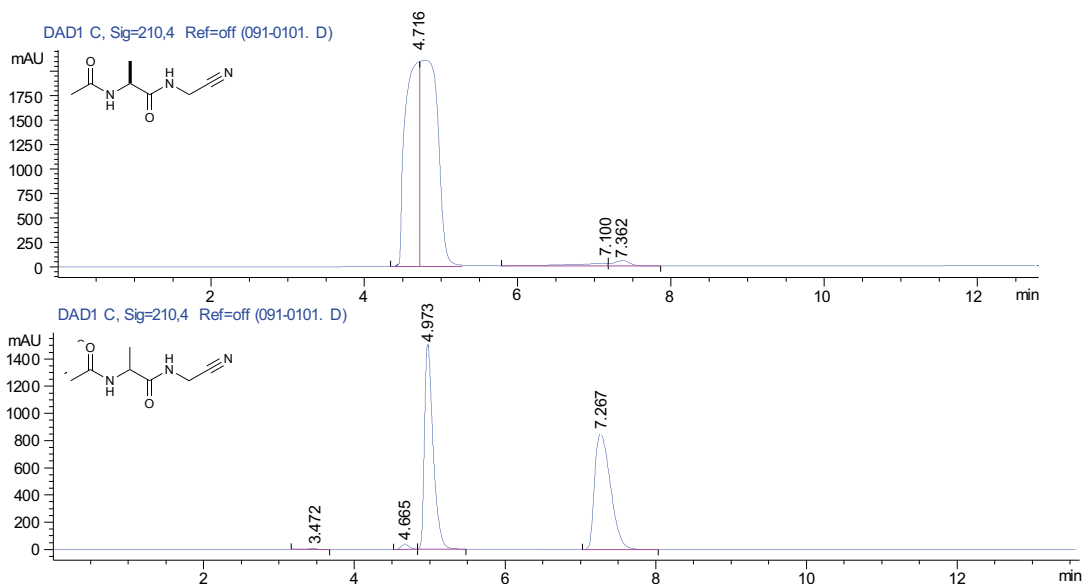


Supplementary Figure 293. Chiral HPLC traces for *rac*-*N*-acetyl-alanine thioamide **Ac-Ala-SNH₂** (bottom) and *N*-acetyl-L-alanine thioamide **Ac-L-Ala-SNH₂** (top) obtained from prebiotic thiolysis of the corresponding *N*-acetyl aminonitrile following *general procedure C*.

N-Acetyl-L-alanylglycine nitrile **Ac-L-Ala-Gly-CN**



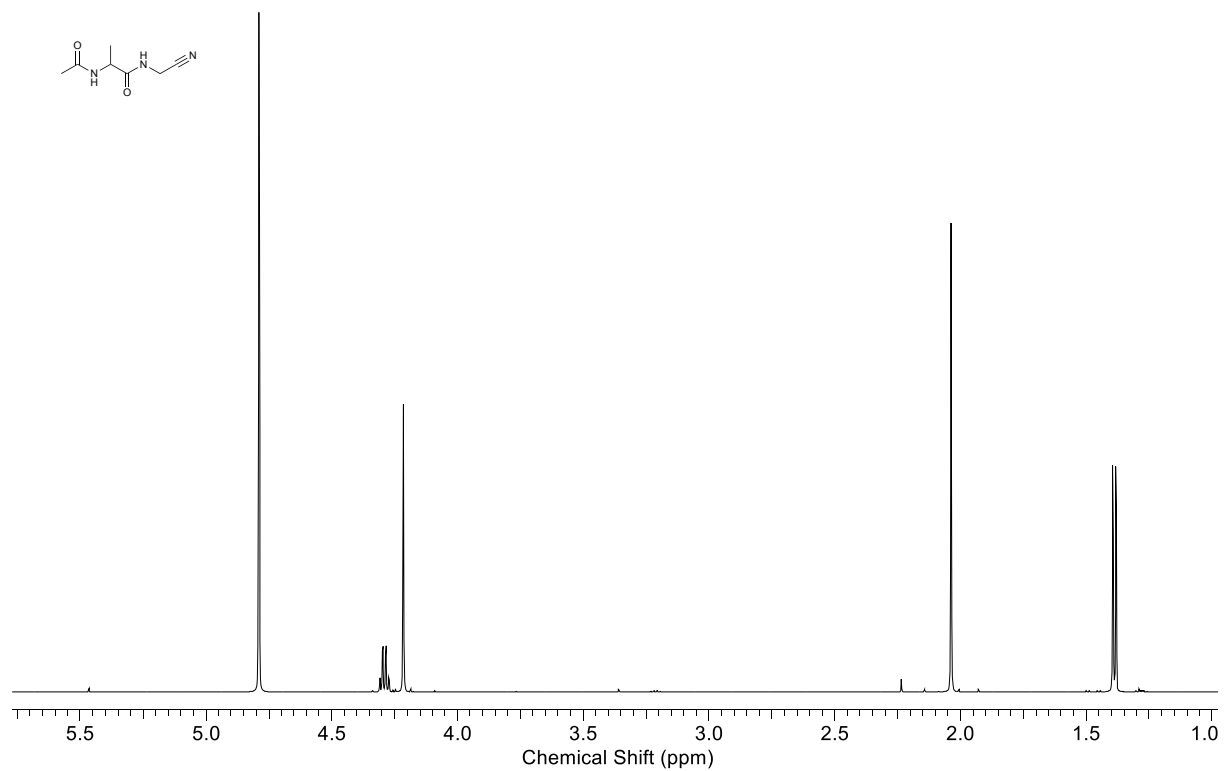
Sodium L-2-acetamidopropanethioate **Ac-L-Ala-S⁻Na⁺** was prepared following *general procedure D* using *N*-acetyl-L-alanine thioamide **Ac-L-Ala-SNH₂** (146 mg, 1.00 mmol). The crude **Ac-L-Ala-S⁻Na⁺** was ligated with glycine nitrile hydrochloride (**Gly-CN·HCl**) (46.2 mg, 5.00 mmol) following *general procedure F*. The reaction was concentrated *in vacuo* and the residue purified by flash column chromatography (SiO₂; MeOH/CH₂Cl₂ 0:100→5:95) to give a white crystalline solid (14.6 mg, 0.09 mmol). All spectral data were found to be identical to those obtained for the racemic *N*-acetyl-alanylglycine nitrile **Ac-Ala-Gly-CN**. The enantiomeric was determined to be 98:2 by chiral HPLC analysis (YMC Chiral Cellulose-SB Classical Analytical S-5 column (250 x 4.6 mm); ethanol (100%), flow rate: 0.7 mL/min, UV detection λ = 210 nm).



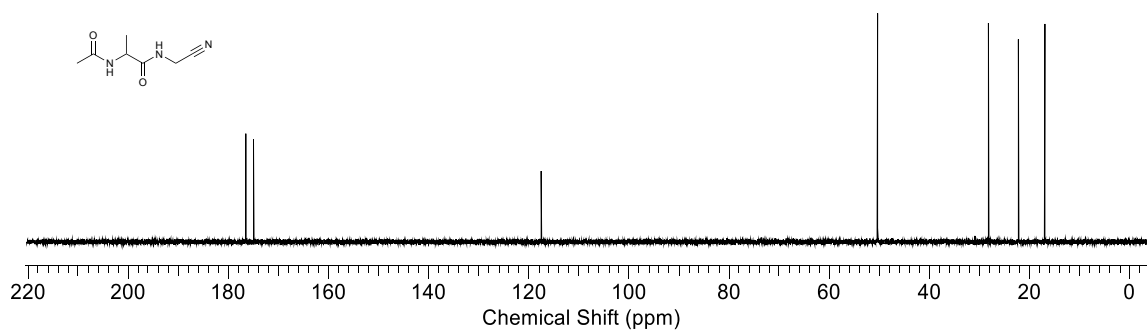
Supplementary Figure 294. Chiral HPLC traces for a synthetic standard of racemic *N*-acetyl-alanylglycine nitrile **Ac-Ala-Gly-CN** (bottom), and *N*-acetyl-L-alanylglycine nitrile **Ac-L-Ala-Gly-CN** obtained from the oxidative prebiotic ligation of **Ac-L-Ala-S-Na⁺** with glycine nitrile **Gly-CN** in the presence of $K_3[Fe(CN)_6]$ at pH 9.0, room temperature and in water (top).

N-Acetyl-alanylglycine nitrile **Ac-Ala-Gly-CN**

Racemic *N*-acetyl-alanylglycine nitrile **Ac-Ala-Gly-CN** was synthesised using conventional peptide synthesis to be used as an authentic chiral HPLC standard. *N*-Acetyl-alanine **Ac-Ala-OH** (1.00 g, 7.63 mmol) and triethylamine (3.86 g, 38.13 mmol, 5.32 mL) were dissolved in dry dichloromethane (50 mL) and stirred at 0 °C under argon. 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (2.96 g, 19.08 mmol) was added portion-wise over 5 min, and the reaction was stirred for 10 min at 0 °C. Glycine nitrile hydrochloride **Gly-CN**·HCl (1.77 g, 19.08 mmol) was added, and the reaction was allowed to warm up to room temperature whilst stirring for 2 h. The reaction was concentrated *in vacuo* and the residue washed dissolved in water (20 mL), and extracted with dichloromethane (5 × 50 mL). The organic layer was concentrated *in vacuo*. The crude residue was purified by flash column chromatography (SiO₂; MeOH/DCM 0:100→5:95) to give *N*-acetyl-alanylglycine nitrile **Ac-Ala-Gly-CN** as a white solid (942 mg, 5.57 mmol, 74%). ¹H NMR (600 MHz, D₂O): δ 4.29 (q, *J* = 7.3 Hz, 1H, Ala-(C2)-H), 4.21 (s, 2H, Gly-(C2)-H₂), 2.04 (s, 3H, (CO)CH₃), 1.39 (d, *J* = 7.3 Hz, 3H, Ala-(C3)-H₃). ¹³C NMR (151 MHz, D₂O): δ 176.5 (COCH₃), 174.9 (Ala-(C1)), 117.5 (Gly-(C1)), 50.4 (Ala-(C2)), 28.3 (Gly-(C2)), 22.3 (COCH₃), 17.0 (Ala-(C3)). LRMS (ES⁺): 170.1 (100, [M + H]⁺). HRMS (ES⁺): calcd. for [C₇H₁₂N₃O₂]⁺: 170.0924; observed: 170.0927. IR (solid): 3291, 3212, 3063, 2989, 2945, 2812, 2246, 1746, 1666, 1630, 1566, 1538 cm⁻¹. M.p. 156.0 – 157.1 °C. The enantiomeric was determined to be 50:50 by chiral HPLC analysis (Supplementary Figure 5) (YMC Chiral Cellulose-SB Classical Analytical S-5 column (250 × 4.6 mm); ethanol (100%), flow rate: 0.7 mL/min, UV detection λ = 210 nm).



Supplementary Figure 295. ¹H NMR (600 MHz, D₂O, 1.00 – 5.75 ppm) spectrum to show *N*-acetyl-alanylglycine nitrile **Ac-Ala-Gly-CN**



Supplementary Figure 296. ¹³C NMR (151 MHz, D₂O, 0 – 220 ppm) spectrum to show *N*-acetyl-alanylglycine nitrile **Ac-Ala-Gly-CN**

Supplementary References

1. Covington, A. K., Paabo, M., Robinson, R. A. & Bates, R. G. Use of the glass electrode in deuterium oxide and the relation between the standardized pD (paD) scale and the operational pH in heavy water. *Anal. Chem.* **40**, 700-706 (1968).
2. Leman, L., Orgel, L. & Ghadiri, M. R. Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **306**, 283-286 (2004).
3. Maurel, M.-C. & Orgel, L. E. Oligomerization of α -thioglutamic acid. *Orig. Life Evol. Biosph.* **30**, 423-430 (2000).
4. 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).
5. Weber, A. L. & Pizzarello, S. The peptide-catalyzed stereospecific synthesis of tetroses: A possible model for prebiotic molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12713–12717 (2006).
6. Adamala, K. & Szostak, J. W. Competition between model protocells driven by an encapsulated catalyst. *Nat. Chem.* **5**, 495–501 (2013).
7. Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. A self-replicating peptide. *Nature* **382**, 525–528 (1996).
8. Lipmann, F. Attempts to map a process evolution of peptide biosynthesis. *Science* **173**, 875–884 (1971).
9. De Duve, C. Blueprint for a cell– the nature and origin of life (N. Patterson, Burlington, NC) (1991).
10. 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).
11. Leman, L. J. & Ghadiri, M. R. Potentially prebiotic synthesis of α -amino thioacids in water. *Synlett* **28**, 68–72 (2017).
12. Kajihara, Y. *et al.* Regioselective α -peptide bond formation through oxidation of amino thioacids. *Biochemistry* **58**, 1672–1678 (2019).
13. Bowler, F. R. *et al.* Prebiotically plausible oligoribonucleotide ligation facilitated by chemoselective acetylation. *Nat. Chem.* **5**, 383–389 (2013).
14. 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).
15. Semenov, S. N. *et al.* Autocatalytic, bistable, oscillatory networks of biologically relevant organic reactions. *Nature* **537**, 656–660 (2016).
16. 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).
17. Stairs, S. *et al.* Divergent prebiotic synthesis of pyrimidine and 8-oxo-purine ribonucleotides. *Nat. Commun.* **8**, 15270 (2017).
18. Islam, S., & Powner, M. W. Prebiotic systems chemistry: Complexity overcoming clutter. *Chem* **2**, 470-501 (2017).
19. 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).
20. Liu, R. & Orgel, L. E. Oxidative acylation using thioacids. *Nature* **389**, 52–54 (1997).
21. 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).
22. 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).
23. Denkwalter, R. G. *et al.* The controlled synthesis of peptides in aqueous medium. I. The use of α -amino acid N-carboxyanhydrides. *J. Am. Chem. Soc.* **88**, 3163-3164 (1966).

24. Dewey, R. S. *et al.* Synthesis of peptides in aqueous medium. V. Preparation and use of 2,5-thiazolidinediones (NTA's). Use of the ^{13}C -H nuclear magnetic resonance signal as internal standard for quantitative studies. *J. Am. Chem. Soc.* **90**, 3254-3255 (1968).
25. Hirschmann, R. F. Controlled synthesis of peptides in aqueous medium. III. Use of Leuch's anhydrides in the synthesis of di-peptides. Mechanism and control of side reactions. *J. Org. Chem.* **32**, 3415-3425 (1967).
26. Hirschmann, R. *et al.* Synthesis of peptides in aqueous medium. VII. Preparation and use of 2,5-thiazolidinediones in peptide synthesis. *J. Org. Chem.* **36**, 49-59 (1971).
27. Hirschmann, R. *et al.* Controlled synthesis of peptides in aqueous medium. VIII. Preparation and use of novel .alpha.-amino acid N-carboxyanhydrides. *J. Am. Chem. Soc.* **93**, 2746-2754 (1971).
28. Kopple, K. D. The reaction of amines with amino acid N-carboxy anhydrides. *J. Am. Chem. Soc.* **79**, 662-664 (1957).
29. Losse, G. & Weddige, H. Neue möglichkeiten zur knüpfung der peptidbindung. *Justus Liebigs Annalen der Chemie* **636**, 144-149 (1960).
30. Losse, G. & Weddige, H. Neue möglichkeiten zur knüpfung der peptidbindung. *Angew. Chem.* **72**, 323-323 (1960).
31. Higashimura, T., Suzuoki, K., Kato, H. & Okamura, S. Condensation polymerization of N-dithiocarbonyl alkoxy-carbonyl-amino acids. Part. II. Condensation polymerization of N-dithiocarbonyl ethoxycarbonyl-DL- α -alanine in organic solvents. *Die Makromolekulare Chemie* **90**, 254-261 (1966).
32. Kato, H., Higashimura, T. & Okamura, S. Condensation polymerization of N-dithiocarbonyl alkoxy-carbonyl amino acids. Part V. Studies on reaction mechanism. *Die Makromolekulare Chemie* **109**, 9-21 (1967).
33. Memeger, W. A Novel synthesis of AB-Aromatic polyamides from aromatic amino acids and carbon disulfide. *Macromolecules* **9**, 195-199 (1976).
34. Lindenmann, A., Khan, N. H. & Hofmann, K. Studies on polypeptides. IV.1 Remarks regarding the use of the phenylthiocarbonyl protecting-group in peptide synthesis. *J. Am. Chem. Soc.* **74**, 476-480 (1952).
35. Leman, L. J., Huang, Z.-Z. & Ghadiri, M. R. Peptide bond formation in water mediated by carbon disulfide. *Astrobiology* **15**, 709-716 (2015).
36. Greenwald, J., Friedmann, M. P. & Riek, R. Amyloid aggregates arise from amino acid condensations under prebiotic conditions. *Angew. Chem. Int. Ed.* **55**, 11609-11613 (2016).
37. Cairns, T. L., Larchar, A. W. & McKusick, B. C. Reactions of hydrogen sulfide with various organic compounds at high pressures. *J. Org. Chem.* **18**, 748-752 (1953).
38. Elias, C., Gelpi, M. E. & Cadenas, R. A. Reaction of peracylated sugars with nitriles catalyzed by lewis acids. *J. Carbohydr. Chem.* **14**, 1209-1216 (1995).
39. Paventi, M. & Edward, J. T. Preparation of α -aminothioamides from aldehydes. *Can. J. Chem.* **65**, 282-289 (1987).
40. Shirai, T. & Kurashige, S. Hydrolysis of α,α -dialkyl substituted α -N-acylaminonitriles. Studies on amino acids. V. *J. Synth. Org. Chem Jpn.* **30**, 1019-1023 (1972).
41. Lipshutz, B. H., Hungate, R. W. & McCarthy, K. E. Heterocycles as masked diamide/dipeptide equivalents. Formation and reactions of substituted 5-(acylamino)oxazoles as intermediates en route to the cyclopeptide alkaloids. *J. Am. Chem. Soc.* **105**, 7703-7713 (1983).
42. Chitale, S., Derasp, J. S., Hussain, B., Tanveer, K. & Beauchemin, A. M. Carbohydrates as efficient catalysts for the hydration of α -amino nitriles. *Chem. Commun.* **52**, 13147-13150 (2016).
43. Zeng, P., Hu, Y. & Hu, H. A Convenient reduction of N-(2-substituted-1-cyanoethenyl) acetamides with NaTeH. *Synth. Commun.* **27**, 939-944 (1997).
44. Paul, R., Williams, R. P. & Cohen, E. Synthesis of three substituted aminochloropropanes. *J. Org. Chem.* **40**, 1653-1656 (1975).
45. Bailey, D. M., Bowers, J. E. & Gutsche, C. D. Ring enlargements. X. The reactions of cycloalkanones carrying two-carbon and four-carbon side chains containing potential diazoalkyl moieties. *J. Org. Chem.* **28**, 610-614 (1963).

46. Lowe, G. & Yuthavong, Y. Kinetic specificity in papain-catalysed hydrolyses. *Biochem. J.* **124**, 107-115 (1971).
47. Culbertson, T. P., Domagala, J. M., Peterson, P., Bongers, S. & Nichols, J. B. New 7-substituted quinolone antibacterial agents. The synthesis of 1-ethyl-1,4-dihydro-4-oxo-7-(2-thiazolyl and 4-thiazolyl)-3-quinolinecarboxylic acids. *J. Heterocycl. Chem.* **24**, 1509-1520 (1987).
48. Ressler, C., Nagarajan, G. R., Kirisawa, M. & Kashelkar, D. V. Synthesis and properties of α -cyanoamino acids. α -Cyanoglycine, L- β -cyano- β -alanine, and L- γ -cyano- γ -aminobutyric acid. *J. Org. Chem.* **36**, 3960-3966 (1971).
49. Bordier, F. et al. Large α -aminonitrilase activity screening of nitrilase superfamily members: Access to conversion and enantiospecificity by LC–MS. *J. Mol. Catal. B: Enzym.* **107**, 79-88 (2014).
50. Ritson, D. J. & Sutherland, J. D. Synthesis of aldehydic ribonucleotide and amino acid precursors by photoredox chemistry. *Angew. Chem. Int. Ed.* **52**, 5845-5847 (2013).
51. Kotova, D. L., Beilina, D. S., Selemenev, V. F. & Shepeleva, O. A. Hydration characteristics of aromatic amino acids. *Pharm. Chem. J.* **35**, 221-222 (2001).
52. Claridge, T. D. W. High-resolution NMR techniques in organic chemistry (3rd edition). Ch 3, 61-132 (Elsevier, Boston, 2016).
53. Danger, G. et al. 5(4H)-Oxazolones as intermediates in the carbodiimide- and cyanamide-promoted peptide activations in aqueous Solution. *Angew. Chem. Int. Ed.* **52**, 611-614 (2013).
54. Melhado, A. D., Luparia, M. & Toste, F. D. Au(I)-catalyzed enantioselective 1,3-dipolar cycloadditions of münchnones with electron-deficient alkenes. *J. Am. Chem. Soc.* **129**, 12638-12639 (2007).
55. Khatskevich, I. É., Kalnin', I. K., Karpeiskaya, E. I. & Klabunovskii, E. I. Asymmetric synthesis of amino acids by catalytic reduction of azlactones of substituted acylaminoacrylic acids. *Russ. Chem. Bull.* **32**, 330-335 (1983).
56. Hershenson, F. M. & Pavia, M. R. Synthesis of N-substituted pyrroles from azlactones via 1,3-oxazolium 5-oxides. *Synthesis*, 999-1001 (1988).
57. Wende, R. C. et al. The enantioselective Dakin–West reaction. *Angew. Chem. Int. Ed.* **55**, 2719-2723 (2016).
58. Vallin, K. S. A. et al. N-1-alkyl-2-oxo-2-aryl amides as novel antagonists of the TRPA1 receptor. *Bioorg. Med. Chem. Lett.* **22**, 5485-5492 (2012).
59. Wang, J., Liang, Y.-L. & Qu, J. Boiling water-catalyzed neutral and selective N-Boc deprotection. *Chem. Commun.*, 5144-5146 (2009).

Supplementary Table of Figures

SUPPLEMENTARY FIGURE 1.....	6
SUPPLEMENTARY FIGURE 2.....	9
SUPPLEMENTARY FIGURE 3.....	9
SUPPLEMENTARY FIGURE 4.....	10
SUPPLEMENTARY FIGURE 5.....	10
SUPPLEMENTARY FIGURE 6.....	11
SUPPLEMENTARY FIGURE 7.....	11
SUPPLEMENTARY FIGURE 8.....	12
SUPPLEMENTARY FIGURE 9.....	12
SUPPLEMENTARY FIGURE 10.....	13
SUPPLEMENTARY FIGURE 11.....	13
SUPPLEMENTARY FIGURE 12.....	14
SUPPLEMENTARY FIGURE 13.....	14
SUPPLEMENTARY FIGURE 14.....	15
SUPPLEMENTARY FIGURE 15.....	15
SUPPLEMENTARY FIGURE 16.....	16
SUPPLEMENTARY FIGURE 17.....	17
SUPPLEMENTARY FIGURE 18.....	18
SUPPLEMENTARY FIGURE 19.....	19
SUPPLEMENTARY FIGURE 20.....	20
SUPPLEMENTARY FIGURE 21.....	22
SUPPLEMENTARY FIGURE 22.....	22
SUPPLEMENTARY FIGURE 23.....	23
SUPPLEMENTARY FIGURE 24.....	23
SUPPLEMENTARY FIGURE 25.....	24
SUPPLEMENTARY FIGURE 26.....	24
SUPPLEMENTARY FIGURE 27.....	25
SUPPLEMENTARY FIGURE 28.....	25
SUPPLEMENTARY FIGURE 29.....	26
SUPPLEMENTARY FIGURE 30.....	26
SUPPLEMENTARY FIGURE 31.....	27
SUPPLEMENTARY FIGURE 32.....	27
SUPPLEMENTARY FIGURE 33.....	28
SUPPLEMENTARY FIGURE 34.....	28
SUPPLEMENTARY FIGURE 35.....	29
SUPPLEMENTARY FIGURE 36.....	29
SUPPLEMENTARY FIGURE 37.....	30
SUPPLEMENTARY FIGURE 38.....	30
SUPPLEMENTARY FIGURE 39.....	32
SUPPLEMENTARY FIGURE 40.....	32
SUPPLEMENTARY FIGURE 41.....	33
SUPPLEMENTARY FIGURE 42.....	33
SUPPLEMENTARY FIGURE 43.....	34

SUPPLEMENTARY FIGURE 44.....	34
SUPPLEMENTARY FIGURE 45.....	35
SUPPLEMENTARY FIGURE 46.....	35
SUPPLEMENTARY FIGURE 47.....	36
SUPPLEMENTARY FIGURE 48.....	36
SUPPLEMENTARY FIGURE 49.....	37
SUPPLEMENTARY FIGURE 50.....	37
SUPPLEMENTARY FIGURE 51.....	38
SUPPLEMENTARY FIGURE 52.....	38
SUPPLEMENTARY FIGURE 53.....	40
SUPPLEMENTARY FIGURE 54.....	41
SUPPLEMENTARY FIGURE 55.....	42
SUPPLEMENTARY FIGURE 56.....	43
SUPPLEMENTARY FIGURE 57.....	44
SUPPLEMENTARY FIGURE 58.....	45
SUPPLEMENTARY FIGURE 59.....	46
SUPPLEMENTARY FIGURE 60.....	55
SUPPLEMENTARY FIGURE 61.....	55
SUPPLEMENTARY FIGURE 62.....	56
SUPPLEMENTARY FIGURE 63.....	56
SUPPLEMENTARY FIGURE 64.....	58
SUPPLEMENTARY FIGURE 65.....	59
SUPPLEMENTARY FIGURE 66.....	60
SUPPLEMENTARY FIGURE 67.....	61
SUPPLEMENTARY FIGURE 68.....	62
SUPPLEMENTARY FIGURE 69.....	63
SUPPLEMENTARY FIGURE 70.....	65
SUPPLEMENTARY FIGURE 71.....	66
SUPPLEMENTARY FIGURE 72.....	67
SUPPLEMENTARY FIGURE 73.....	68
SUPPLEMENTARY FIGURE 74.....	69
SUPPLEMENTARY FIGURE 75.....	70
SUPPLEMENTARY FIGURE 76.....	71
SUPPLEMENTARY FIGURE 77.....	72
SUPPLEMENTARY FIGURE 78.....	73
SUPPLEMENTARY FIGURE 79.....	74
SUPPLEMENTARY FIGURE 80.....	76
SUPPLEMENTARY FIGURE 81.....	77
SUPPLEMENTARY FIGURE 82.....	78
SUPPLEMENTARY FIGURE 83.....	78
SUPPLEMENTARY FIGURE 84.....	79
SUPPLEMENTARY FIGURE 85.....	79
SUPPLEMENTARY FIGURE 86.....	80
SUPPLEMENTARY FIGURE 87.....	80

SUPPLEMENTARY FIGURE 88.....	81
SUPPLEMENTARY FIGURE 89.....	82
SUPPLEMENTARY FIGURE 90.....	82
SUPPLEMENTARY FIGURE 91.....	83
SUPPLEMENTARY FIGURE 92.....	83
SUPPLEMENTARY FIGURE 93.....	85
SUPPLEMENTARY FIGURE 94.....	86
SUPPLEMENTARY FIGURE 95.....	86
SUPPLEMENTARY FIGURE 96.....	87
SUPPLEMENTARY FIGURE 97.....	87
SUPPLEMENTARY FIGURE 98.....	88
SUPPLEMENTARY FIGURE 99.....	88
SUPPLEMENTARY FIGURE 100.....	89
SUPPLEMENTARY FIGURE 101.....	89
SUPPLEMENTARY FIGURE 102.....	90
SUPPLEMENTARY FIGURE 103.....	90
SUPPLEMENTARY FIGURE 104.....	91
SUPPLEMENTARY FIGURE 105.....	92
SUPPLEMENTARY FIGURE 106.....	93
SUPPLEMENTARY FIGURE 107.....	93
SUPPLEMENTARY FIGURE 108.....	94
SUPPLEMENTARY FIGURE 109.....	94
SUPPLEMENTARY FIGURE 110.....	95
SUPPLEMENTARY FIGURE 111.....	95
SUPPLEMENTARY FIGURE 112.....	97
SUPPLEMENTARY FIGURE 113.....	98
SUPPLEMENTARY FIGURE 114.....	98
SUPPLEMENTARY FIGURE 115.....	99
SUPPLEMENTARY FIGURE 116.....	100
SUPPLEMENTARY FIGURE 117.....	101
SUPPLEMENTARY FIGURE 118.....	101
SUPPLEMENTARY FIGURE 119.....	102
SUPPLEMENTARY FIGURE 120.....	102
SUPPLEMENTARY FIGURE 121.....	103
SUPPLEMENTARY FIGURE 122.....	103
SUPPLEMENTARY FIGURE 123.....	104
SUPPLEMENTARY FIGURE 124.....	104
SUPPLEMENTARY FIGURE 125.....	105
SUPPLEMENTARY FIGURE 126.....	106
SUPPLEMENTARY FIGURE 127.....	107
SUPPLEMENTARY FIGURE 128.....	107
SUPPLEMENTARY FIGURE 129.....	108
SUPPLEMENTARY FIGURE 130.....	108
SUPPLEMENTARY FIGURE 131.....	109

SUPPLEMENTARY FIGURE 132.....	109
SUPPLEMENTARY FIGURE 133.....	110
SUPPLEMENTARY FIGURE 134.....	110
SUPPLEMENTARY FIGURE 135.....	111
SUPPLEMENTARY FIGURE 136.....	112
SUPPLEMENTARY FIGURE 137.....	112
SUPPLEMENTARY FIGURE 138.....	113
SUPPLEMENTARY FIGURE 139.....	113
SUPPLEMENTARY FIGURE 140.....	114
SUPPLEMENTARY FIGURE 141.....	114
SUPPLEMENTARY FIGURE 142.....	115
SUPPLEMENTARY FIGURE 143.....	116
SUPPLEMENTARY FIGURE 144.....	116
SUPPLEMENTARY FIGURE 145.....	117
SUPPLEMENTARY FIGURE 146.....	118
SUPPLEMENTARY FIGURE 147.....	119
SUPPLEMENTARY FIGURE 148.....	120
SUPPLEMENTARY FIGURE 149.....	121
SUPPLEMENTARY FIGURE 150.....	122
SUPPLEMENTARY FIGURE 151.....	122
SUPPLEMENTARY FIGURE 152.....	124
SUPPLEMENTARY FIGURE 153.....	124
SUPPLEMENTARY FIGURE 154.....	125
SUPPLEMENTARY FIGURE 155.....	125
SUPPLEMENTARY FIGURE 156.....	126
SUPPLEMENTARY FIGURE 157.....	127
SUPPLEMENTARY FIGURE 158.....	127
SUPPLEMENTARY FIGURE 159.....	128
SUPPLEMENTARY FIGURE 160.....	128
SUPPLEMENTARY FIGURE 161.....	129
SUPPLEMENTARY FIGURE 162.....	129
SUPPLEMENTARY FIGURE 163.....	130
SUPPLEMENTARY FIGURE 164.....	131
SUPPLEMENTARY FIGURE 165.....	131
SUPPLEMENTARY FIGURE 166.....	132
SUPPLEMENTARY FIGURE 167.....	133
SUPPLEMENTARY FIGURE 168.....	133
SUPPLEMENTARY FIGURE 169.....	134
SUPPLEMENTARY FIGURE 170.....	135
SUPPLEMENTARY FIGURE 171.....	136
SUPPLEMENTARY FIGURE 172.....	137
SUPPLEMENTARY FIGURE 173.....	138
SUPPLEMENTARY FIGURE 174.....	139
SUPPLEMENTARY FIGURE 175.....	139

SUPPLEMENTARY FIGURE 176.....	140
SUPPLEMENTARY FIGURE 177.....	140
SUPPLEMENTARY FIGURE 178.....	141
SUPPLEMENTARY FIGURE 179.....	141
SUPPLEMENTARY FIGURE 180.....	142
SUPPLEMENTARY FIGURE 181.....	142
SUPPLEMENTARY FIGURE 182.....	143
SUPPLEMENTARY FIGURE 183.....	143
SUPPLEMENTARY FIGURE 184.....	144
SUPPLEMENTARY FIGURE 185.....	144
SUPPLEMENTARY FIGURE 186.....	145
SUPPLEMENTARY FIGURE 187.....	145
SUPPLEMENTARY FIGURE 188.....	146
SUPPLEMENTARY FIGURE 189.....	147
SUPPLEMENTARY FIGURE 190.....	148
SUPPLEMENTARY FIGURE 191.....	149
SUPPLEMENTARY FIGURE 192.....	150
SUPPLEMENTARY FIGURE 193.....	151
SUPPLEMENTARY FIGURE 194.....	152
SUPPLEMENTARY FIGURE 195.....	153
SUPPLEMENTARY FIGURE 196.....	154
SUPPLEMENTARY FIGURE 197.....	154
SUPPLEMENTARY FIGURE 198.....	155
SUPPLEMENTARY FIGURE 199.....	155
SUPPLEMENTARY FIGURE 200.....	156
SUPPLEMENTARY FIGURE 201.....	157
SUPPLEMENTARY FIGURE 202.....	157
SUPPLEMENTARY FIGURE 203.....	158
SUPPLEMENTARY FIGURE 204.....	159
SUPPLEMENTARY FIGURE 205.....	159
SUPPLEMENTARY FIGURE 206.....	160
SUPPLEMENTARY FIGURE 207.....	161
SUPPLEMENTARY FIGURE 208.....	161
SUPPLEMENTARY FIGURE 209.....	163
SUPPLEMENTARY FIGURE 210.....	164
SUPPLEMENTARY FIGURE 211.....	164
SUPPLEMENTARY FIGURE 212.....	165
SUPPLEMENTARY FIGURE 213.....	165
SUPPLEMENTARY FIGURE 214.....	166
SUPPLEMENTARY FIGURE 215.....	166
SUPPLEMENTARY FIGURE 216.....	168
SUPPLEMENTARY FIGURE 217.....	169
SUPPLEMENTARY FIGURE 218.....	170
SUPPLEMENTARY FIGURE 219.....	171

SUPPLEMENTARY FIGURE 220.....	172
SUPPLEMENTARY FIGURE 221.....	173
SUPPLEMENTARY FIGURE 222.....	174
SUPPLEMENTARY FIGURE 223.....	176
SUPPLEMENTARY FIGURE 224.....	176
SUPPLEMENTARY FIGURE 225.....	177
SUPPLEMENTARY FIGURE 226.....	178
SUPPLEMENTARY FIGURE 227.....	179
SUPPLEMENTARY FIGURE 228.....	180
SUPPLEMENTARY FIGURE 229.....	181
SUPPLEMENTARY FIGURE 230.....	182
SUPPLEMENTARY FIGURE 231.....	183
SUPPLEMENTARY FIGURE 232.....	185
SUPPLEMENTARY FIGURE 233.....	185
SUPPLEMENTARY FIGURE 234.....	186
SUPPLEMENTARY FIGURE 235.....	188
SUPPLEMENTARY FIGURE 236.....	188
SUPPLEMENTARY FIGURE 237.....	189
SUPPLEMENTARY FIGURE 238.....	189
SUPPLEMENTARY FIGURE 239.....	190
SUPPLEMENTARY FIGURE 240.....	190
SUPPLEMENTARY FIGURE 241.....	191
SUPPLEMENTARY FIGURE 242.....	192
SUPPLEMENTARY FIGURE 243.....	193
SUPPLEMENTARY FIGURE 244.....	194
SUPPLEMENTARY FIGURE 245.....	196
SUPPLEMENTARY FIGURE 246.....	196
SUPPLEMENTARY FIGURE 247.....	197
SUPPLEMENTARY FIGURE 248.....	197
SUPPLEMENTARY FIGURE 249.....	198
SUPPLEMENTARY FIGURE 250.....	198
SUPPLEMENTARY FIGURE 251.....	199
SUPPLEMENTARY FIGURE 252.....	199
SUPPLEMENTARY FIGURE 253.....	200
SUPPLEMENTARY FIGURE 254.....	200
SUPPLEMENTARY FIGURE 255.....	201
SUPPLEMENTARY FIGURE 256.....	201
SUPPLEMENTARY FIGURE 257.....	202
SUPPLEMENTARY FIGURE 258.....	202
SUPPLEMENTARY FIGURE 259.....	203
SUPPLEMENTARY FIGURE 260.....	203
SUPPLEMENTARY FIGURE 261.....	204
SUPPLEMENTARY FIGURE 262.....	204
SUPPLEMENTARY FIGURE 263.....	205

SUPPLEMENTARY FIGURE 264.....	205
SUPPLEMENTARY FIGURE 265.....	206
SUPPLEMENTARY FIGURE 266.....	206
SUPPLEMENTARY FIGURE 267.....	207
SUPPLEMENTARY FIGURE 268.....	208
SUPPLEMENTARY FIGURE 269.....	209
SUPPLEMENTARY FIGURE 270.....	209
SUPPLEMENTARY FIGURE 271.....	210
SUPPLEMENTARY FIGURE 272.....	210
SUPPLEMENTARY FIGURE 273.....	211
SUPPLEMENTARY FIGURE 274.....	211
SUPPLEMENTARY FIGURE 275.	212
SUPPLEMENTARY FIGURE 276.....	213
SUPPLEMENTARY FIGURE 277.....	213
SUPPLEMENTARY FIGURE 278.....	214
SUPPLEMENTARY FIGURE 279.....	214
SUPPLEMENTARY FIGURE 280.....	215
SUPPLEMENTARY FIGURE 281.....	215
SUPPLEMENTARY FIGURE 282.....	216
SUPPLEMENTARY FIGURE 283.....	216
SUPPLEMENTARY FIGURE 284.....	217
SUPPLEMENTARY FIGURE 285.....	218
SUPPLEMENTARY FIGURE 286.....	219
SUPPLEMENTARY FIGURE 287.....	219
SUPPLEMENTARY FIGURE 288.....	221
SUPPLEMENTARY FIGURE 289.....	221
SUPPLEMENTARY FIGURE 290.....	222
SUPPLEMENTARY FIGURE 291.....	223
SUPPLEMENTARY FIGURE 292.....	223
SUPPLEMENTARY FIGURE 293.....	224
SUPPLEMENTARY FIGURE 294.....	225
SUPPLEMENTARY FIGURE 295.....	226
SUPPLEMENTARY FIGURE 296.....	226