Electronic Supplementary Material (ESI) for ChemComm.

Not always an innocent bystander: the impact of stabilised phosphopantetheine moieties when studying nonribosomal peptide biosynthesis

Y. T. Candace Ho, Joe A. Kaczmarski, Julien Tailhades, Thierry Izoré, David L. Steer, Ralf B. Schittenhelm, Manuela Tosin, Colin J. Jackson and Max J. Cryle*

*Max.Cryle@monash.edu

Table of Contents

Experimental P.2

Supplementary Figures P.11

Supplementary Tables P.22

Supplementary References P.23

EXPERIMENTAL

Chemical Synthesis

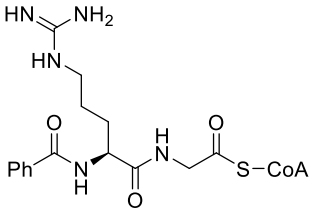
Unless specified otherwise, chemicals purchased from Sigma Aldrich, Iris Biotech, Chem-Impex International, GL Biochem and Fisher Scientific were used without further purification. Reagent grade chloroform, dichloromethane (DCM), *N,N*-dimethylformamide (DMF), methanol, acetonitrile (MeCN), diethyl ether, and water were purchased from Fisher Scientific.

All purifications were performed by using a Shimadzu high performance liquid chromatography system equipped with an SPD-M20A Prominence photo diode array detector and two LC-20AP pumps. Preparative separations were performed using a Zorbax SB-C18 column from Agilent (7 µm, 21.2 x 250 mm) with a flow rate of 10 mL/min employing a gradient of 5–5-45% ACN + 0.1% TFA in 35 min.

Probe analysis was conducted on a triple quadrupole LC-MS/MS system from Shimadzu (LCMS-8050, ESI operating in positive and negative mode); the mobile phases used were water + 0.1% FA and ACN + 0.1% FA for analytical runs. Raw data was manually analysed in LabInsight (Shimadzu), with extracted ion chromatograms to the predicted species. MS2 spectra corresponding to the predicted mass were manually characterised.

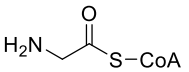
Enzyme activity analysis (control) was conducted on an HPLC-MS system from Shimadzu (LCMS-2020, ESI operating in positive and negative mode) using a Waters XBridge®Peptide BEH C18 column (300 Å, 3.5 µm, 4.6 mm × 250 mm) employing a gradient of 5–5-45% ACN + 0.1% FA in 35 min.

Tripeptidyl-CoA

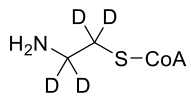


Tripeptidyl-CoA (benzoic acid-D-Arg-Gly-CoA) was synthesised, purified and characterised as previously reported. [1]

Aminoacyl-CoAs

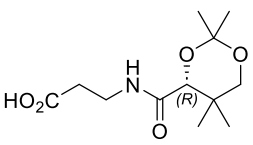


Gly-CoA was synthesised, purified, and characterised as previously reported. [1]

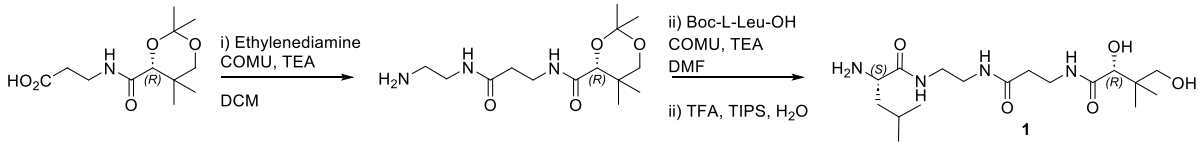


d4-Gly$_{stab}$-CoA was synthesised, purified, and characterised as previously reported. [1]

Stabilised substrate pantetheine analogues



Preparation of the protected pantothenic acid**.** Experimental protocol: pantothenic acid (1 eq), p-toluenesulfonic acid (1.2 eq), acetone, 3Å molecular sieves, RT overnight. This compound was characterised and matched the data previously reported. [2]



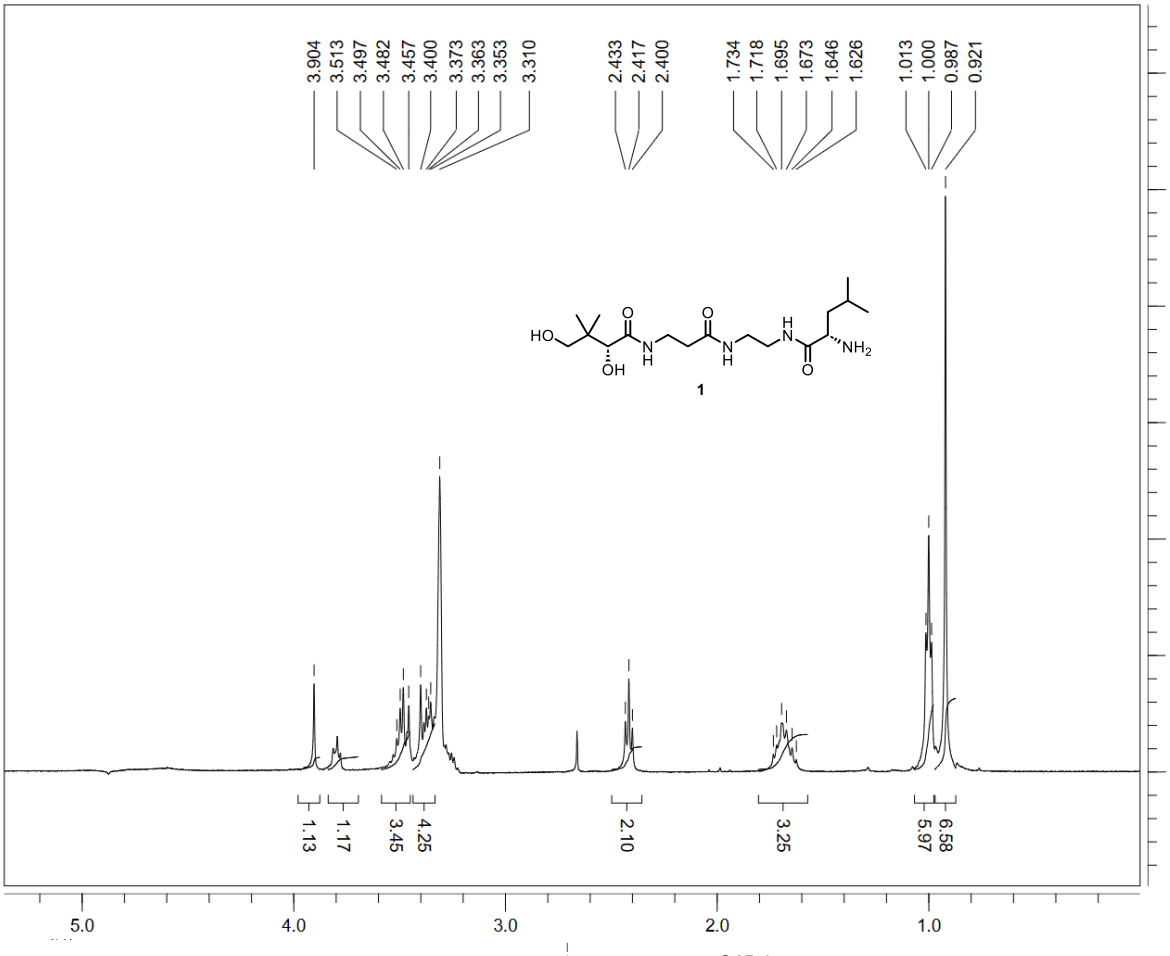
Synthesis of 1. Probe **1** was prepared following a three-step protocol concluded by RP-HPLC purification.

Step 1. The carboxylic acid of 1,3-Isopropylidene-D-pantothenic acid (1 eq.) was activated in the presence of (1-Cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium hexafluorophosphate (COMU, 1 eq.) and triethylamine (TEA, 2 eq.) in DCM at room temperature for 5 minutes. Next, a solution of ethylenediamine (10 eq.) in DCM was added slowly and the reaction was stirred for 2 h. The solution was subsequently washed with brine, dried over Na_2SO_4, filtered and concentrated *in vacuo*. NMR data was in agreement with previously reported literature.[3]

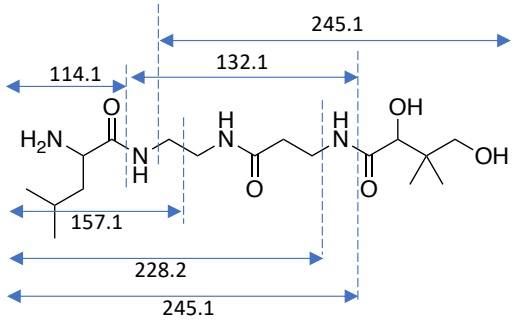
Step 2. The carboxylic acid of Boc-L-leucine (1 eq.) was activated in the presence of COMU (1 eq.) and TEA (2 eq.) at room temperature for 5 minutes and added to the compound obtained in step 1. After 2 hours, the was reaction quenched by addition of water and extracted with ethyl acetate (3 x 20 mL). The combined organic layers were washed with brine (2 x 20 mL), dried over Na_2SO_4 and the solvent removed *in vacuo* after filtration of Na_2SO_4.

Step 3. The cleavage of the protecting groups (isopropylidene, Boc) was accomplished using a solution of TFA/ TIS/ H_2O (95: 2.5: 2.5 v/ v'/ v'', 5 mL) with shaking at room temperature for 1 h. The solution was concentrated under a stream of nitrogen to ~1 mL volume and precipitated with ice-cold diethyl ether (9 mL). The product was washed using diether ether with centrifugation three times in a flame-resistant centrifuge. Probe **1** was purified by RP-HPLC and characterised by LCMS, HRMS and NMR.

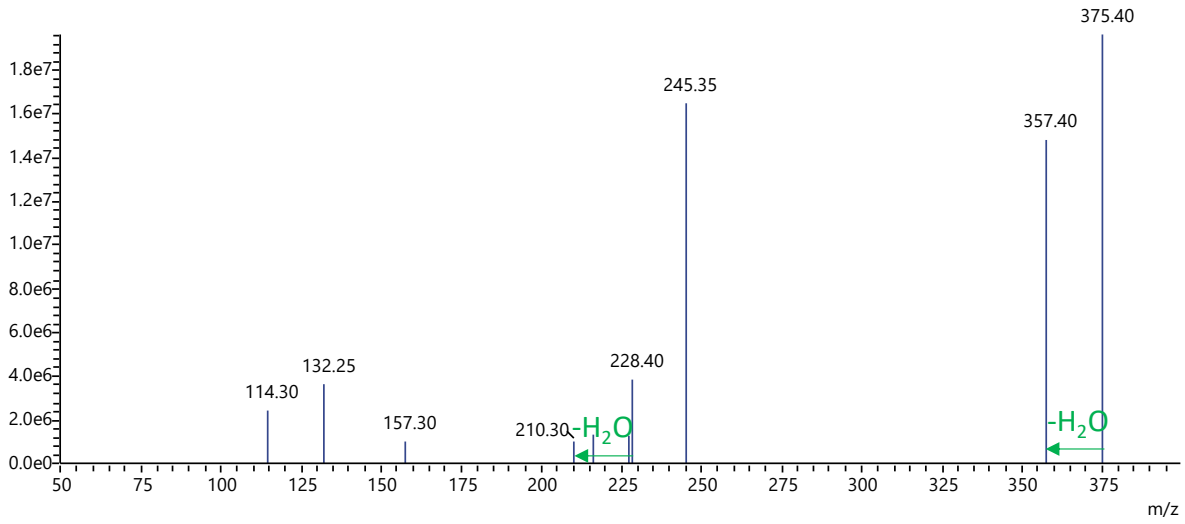
Characterisation of 1. ^{1}H NMR (400 MHz, *CD_3OD*): δ 3.90 (s, 1H), 3.79 (t, *J* = 7 Hz, 1H), 3.51 – 3.45 (m, 3H), 3.40 – 3.31 (m, 4H), 2.41 (t, *J* = 7 Hz, 2H), 1.73 – 1.62 (m, 3H), 1.00 (t, *J* = 5 Hz, 6H), 0.92 (s, 6H). HRMS (ESI): calculated for $C_{17}H_{35}N_4O_5^+$ $[M+H]^+$: 375.2602, found: 375.2609.

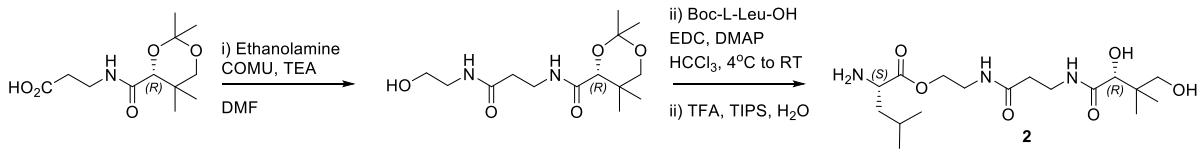


1



Product Ion(+)[375.50] CE:-10.0 RT:0.909 1.96e7





Synthesis of 2. Probe **2** was prepared following a three-step protocol concluded by RP-HPLC purification.

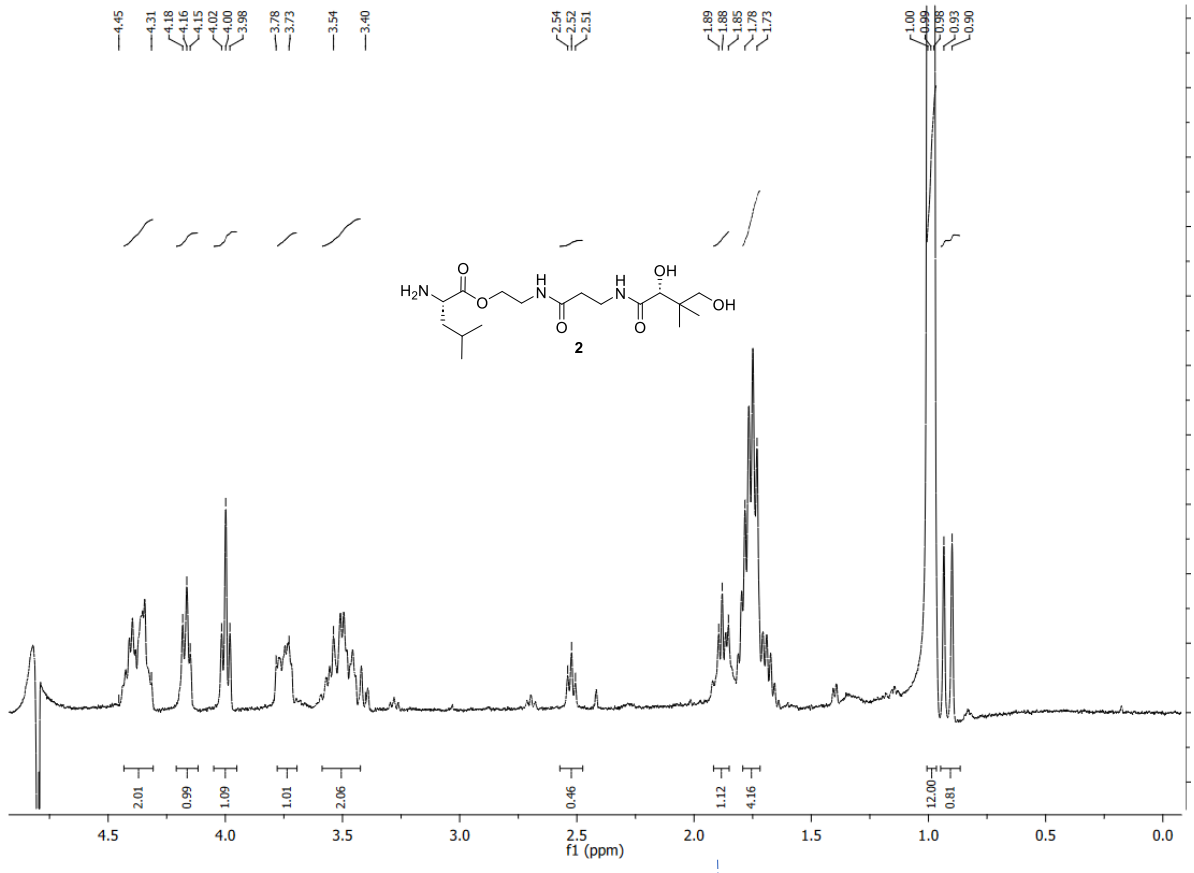
Step 1. The carboxylic acid of 1,3-Isopropylidene-D-pantothenic acid (1 eq.) was activated in DMF with COMU (1 eq.) and TEA (2 eq.) at room temperature for 5 minutes. Ethanolamine (3 eq.) was then added forming an amide.*

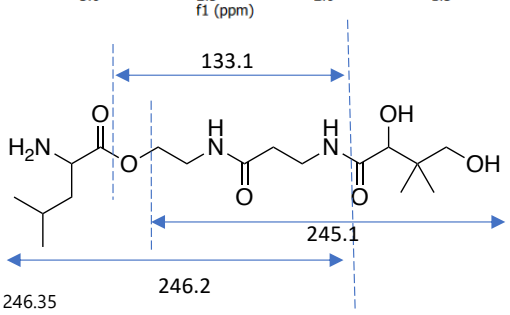
Step 2. The carboxylic acid of Boc-L-leucine (2 eq.) was activated in the presence of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC, 1 eq.) and 4-Dimethylaminopyridine (DMAP, 0.1 eq.) in chloroform (4 ^{o}C, 15 min) and added to the compound obtained in step 1 dissolved in chloroform. After 4 hours, the reaction was quenched by addition of water and extracted with ethyl acetate (3 x 20 mL). The combined organic layers were washed with brine (2 x 20 mL), dried over Na_2SO_4 and the solvent removed *in vacuo* after filtration of Na_2SO_4.

Step 3. The cleavage of the protecting groups (isopropylidene, Boc) was accomplished using a solution of TFA/ TIS/ H_2O (95: 2.5: 2.5 v/ v'/ v'', 5 mL) with shaking at room temperature for 1 h. The solution was concentrated under a stream of nitrogen to ~1 mL volume and precipitated with ice-cold diethyl ether (9 mL). The product was washed using diether ether with centrifugation three times in a flame-resistant centrifuge. Probe **2** was purified by RP-HPLC and characterised by LCMS, HRMS and NMR.

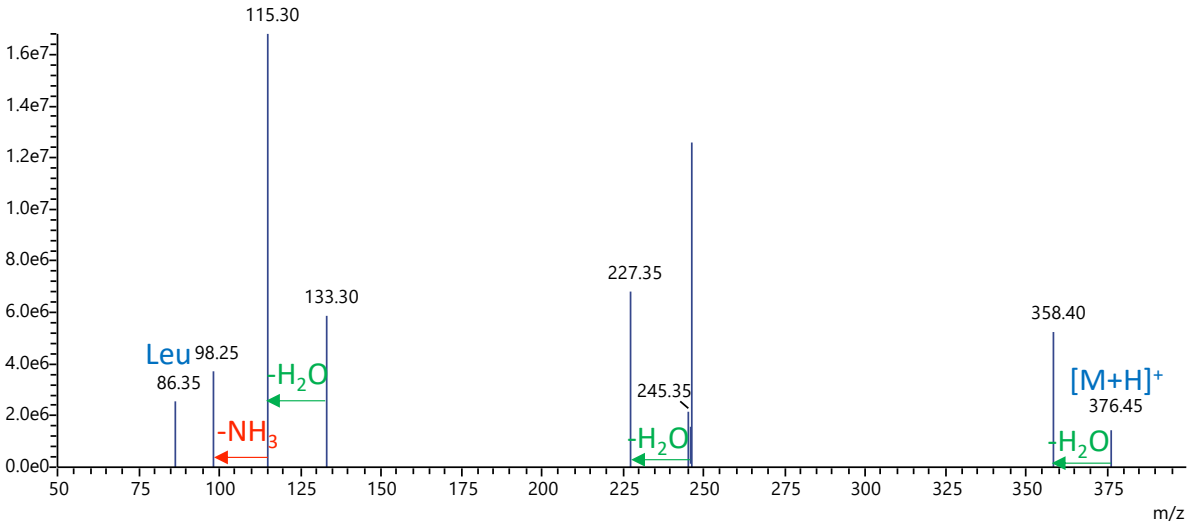
Characterisation of 2. ^{1}H NMR (400 MHz, *D_2O*): δ 4.43 – 4.30 (m, 2H), 4.16 (t, *J* = 6.4 Hz, 1H), 4.00 (t, *J* = 7 Hz, 1H), 3.77 – 3.69 (m, 1H), 3.58 – 3.42 (m, 2H), 2.52 (t, *J* = 6.6 Hz, 1H), 1.91 – 1.85 (m, 1H), 1.78 – 1.73 (m, 4H), 0.99 (t, *J* = 5 Hz, 12H), 0.92 (d, *J* = 14 Hz, 1H). HRMS (ESI): calculated for $C_{17}H_{35}N_4O_5^+$ $[M+H]^+$: 376.2445, found: 376.2445.

*O-to-N acyl transfer drives the formation of the amide as it is more thermodynamically favoured than the ester.





Product Ion(+)[376.50] CE:-20.0 RT:1.433 1.68e7



Construct Cloning

PCP_2-C_3 constructs. Wild-type gene PCP_2-C_3 encoding the desired region of FscG (UniProt ID Q47NR9) and PCP_2-C_3 R2577G mutant were amplified and cloned into the pOPINS vector as previously reported.[1]

PCP_2-C_3 SpyCatcher and PCP_3 SpyTag constructs. The PCP_2-C_3 SpyCatcher construct (pOPINS vector) and PCP_3 SpyTag construct (pHIS17 vector) were obtained in a previous study.[1] PCP_2-C_3 R2577G SpyCatcher mutant was generated using standard Quick-Change site-directed mutagenesis procedures using the following primers:

Fwd - 5'-CACTTCCAGCGGGGGCTCTGGCTCA-3'

Rev - 5'-TGAGCCAGAGCCCCCGCTGGAAGTG-3'

Protein Expression and Purification

Both wild-type and R2577G mutant PCP_2-C_3 proteins (with and without Spycatcher/Spytag) were expressed and purified as previously described.[1]

Sfp (R4-4 mutant) was expressed and purified as previously reported.[4]

Substrate Loading onto PCP-domains

All proteins containing PCP-domains were expressed and purified in their *apo* form, which were converted into the loaded *holo* form using the phosphopantetheinyl transferase Sfp (R4-4 mutant) and desired peptidyl-CoAs.[5]

For the tripeptidyl-CoA loading - the loading reaction utilised a 1 : 2 : 0.1 molar ratio of the PCP domain, peptidyl-CoA and Sfp (R4-4 mutant), respectively. Peptidyl-CoA (200 μM) was loaded onto PCP-containing construct (100 μM), for 1 h at 30 °C via the activity of Sfp (10 μM) in PCP-loading buffer (50 mM HEPES, pH 7.0; 50 mM NaCl; 10 mM $MgCl_2$).

For substrate pantetheine analogues loading – **1**/**2** (500 μM), PCP-containing construct (250 μM), PanK (2.5 μM), PPAT(5 μM), ATP (1 mM) and Sfp (R4-4 mutant, 25 μM) were mixed in PCP-loading. The reaction was incubated for 1 hour at 30 °C.

For aminoacyl-CoA loading – the desired aminoacyl-CoA (200 μM), PCP-containing construct (100 μM), and Sfp (R4-4 mutant, 10 μM) were mixed in PCP-loading buffer. The reaction was incubated for 1 hour at 30 °C.

Following the loading reaction, remaining peptidyl-CoA/ substrate pantetheine analogue was removed by three sequential concentration / dilution steps (Amicon® Ultra-0.5 mL centrifugal filters, 50 kDa MWCO, Merck-Millipore) using PCP-loading buffer.

Enzyme Activity Assays

Peptide loaded PCP_2-C_3 Spy-Catcher constructs were incubated with unloaded PCP_3 Spy-tag construct (both 100 μM) for 10 min at 30 °C, followed by loading of the desired substrate pantetheine analogues on the PCP_3 as described above. The reaction was then incubated for an additional 1 h at 30 °C to allow for the condensation reaction to occur.

For leucinyl-amide **1** tethered loaded PCP_3 substrates, reaction mixtures were directly analysed using nano LC ESI MS.

For leucinyl-ester **2** tethered loaded PCP_3 substrates, substrates were chemically cleaved by an addition of a 40% methylamine solution in water (0.5 M) to liberate the methylamide peptides; reaction mixtures were incubated for 15 min at room temperature. After chemical cleavage, 850 µL of 0.1% FA in water was added into the reaction to adjust the pH to ~7. Purification was performed using solid phase extraction (SPE) column (Bond Elut Plexa 30 mg/mL, Agilent Technologies) that had been activated with 0.1% FA in MeOH (1 mL) and equilibrated with 0.1% FA in water (1 mL). The neutralised reaction mixture was applied to the equilibrate SPE column via gravity flow, washed with 0.1% FA in water (1 mL) and eluted with 1 mL of 0.1% FA in ACN / H_2O (50:50). Samples were then dried by freeze dryer at -50 °C and analysed by HRMS.

PPant Ejection

Mass spectrometry measurements were performed on a MicroTOFq mass spectrometer (Bruker Daltonics) coupled online to a 1200 series capillary/nano-LC (Agilent Technologies) *via* a Bruker nano ESI source. Proteins were separated on a 150 mm reverse-phase column (ZORBAX 300SB-C18, 3.5 µm, 0.075 x 150 mm; Agilent Technologies) after binding to a trap column (ZORBAX 300SB-C18, 5 µm, 0.30 x 5 mm cartridges; Agilent Technologies). Elution was performed on-line with a gradient from 4% MeCN to 60% MeCN in 0.1% FA over 30 min at 300 nL / min. Proteins greater than 20 kDa were separated on a MabPac SEC-1 5 μm 300 Å 50 x 4 mm (Thermo Scientific) column with an isocratic gradient of 50% MeCN, 0.05% TFA and 0.05% FA at a flow rate of 50 µL / minute. The protein is eluted over a 20-minute run-time monitored by UV detection at 254 nm. After 20 minutes the flow path was switched to infuse low concentration tune mix (Agilent Technologies, Santa Clara, CA, USA) to calibrate the spectrum post acquisition. The eluent was nebulised and ionised using the Bruker electrospray source with a capillary voltage of 4500 V dry gas at 180 °C, flow rate of 4 L / minute and nebulizer gas pressure at 0.6 bar. MS^2 spectra were acquired by manual selection of isolation mass and isolation width with a collision energy of 32. The spectra were extracted and deconvoluted using Data explorer software version 3.4 build 192 (Bruker Daltonics, Bremen, Germany).

HRMS and MS^2 Measurements

High resolution mass spectrometry measurements were performed on an Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled online to a nano-LC (Ultimate 3000 RSLCnano; Thermo Scientific) *via* a nanospray source. Peptides were separated on a 50 cm reverse-phase column (Acclaim PepMap RSLC, 75 µm × 50 cm, nanoViper, C18, 2 µm, 100 Å; Thermo Scientific) after binding to a trap column (Acclaim PepMap 100, 100 µm × 2 cm, nanoViper, C18, 5 µm, 100 Å; Thermo Scientific). Elution was performed on-line with a gradient from 6% MeCN to 30% MeCN in 0.1% formic acid over 30 min at 250 nL min^{-1}. Full scan MS was performed in the Orbitrap at 60 000 nominal resolution, with targeted MS^2 scans of peptides of interest acquired at 15 000 nominal resolution in the Orbitrap using HCD with stepped collision energy (24 ± 5% NCE). Raw data was manually analysed in XCalibur QualBrowser (Thermo Scientific), with extracted ion chromatograms to the predicted species generated

with 6 ppm mass tolerance. Predicted MS^2 fragments were generated with MS-Product (ProteinProspector v5.22.1, UCSF) and manually assigned to spectra.

Crystallisation of PCP_2-C_3 Proteins

Stabilised leucinyl-CoAs were loaded onto PCP_2-C_3 affording the *holo* form of PCP_2-C_3 and concentrated to a final concentration of 10 mg/mL in gel-filtration buffer. Crystals of PCP_2-C_3 grew overnight at RT at a concentration of 10 mg / mL in a 3:2:1 ratio (v/v/v) with the crystallisation solution and seed stock (1.8 µL drops) composed of 18–22% v/v PEG 3350 and 0.17 – 0.3 M magnesium formate). Crystals were cryoprotected by transferring in a drop composed of reservoir solution supplemented with glycerol (to a final concertation of 30% v/v). Crystals (WT PCP_2-C_3 Leu_{Stab}N and R2577G PCP_2-C_3 Leu_{Stab}O) were collected in cryo loops and flash frozen in liquid nitrogen.

Data Collection and Structure Determination

All data set was collected using the MX2 beamline[6] at the Australian Synchrotron (Clayton, Victoria, Australia) equipped with an Eiger detector (Dectris) at 100 K[7]. Data processing was performed using XDS[8] and AIMLESS as implemented in CCP4[9]. Phases were obtained via a molecular replacement experiment using the PHENIX in-built Phaser module[10] with the starting model PDB 7KW0. The crystals belonged to the $P2_12_12_1$ space group, with the unit cell comprising 2 highly similar copies of the PCP_2-C_3 construct. Structural models were built in COOT[11] and refined using PHENIX-refine[10]. All graphics were generated using Pymol (Schrödinger LLC).

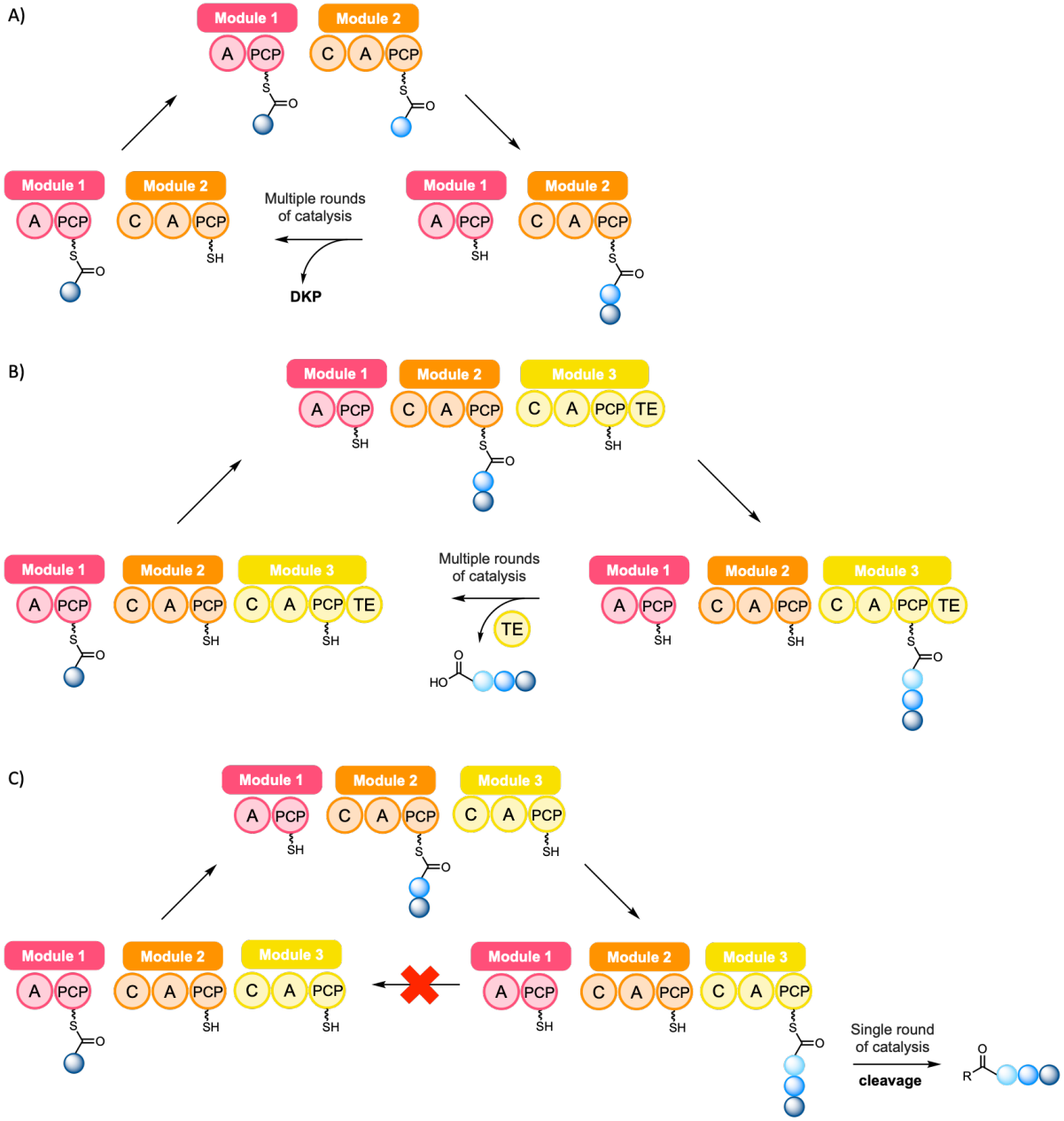
Molecular Dynamics (MD) Simulations

MD simulations were performed using Desmond (Schrödinger Release 2022-1: Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2021.) and the OPLS4 forcefield.[12]

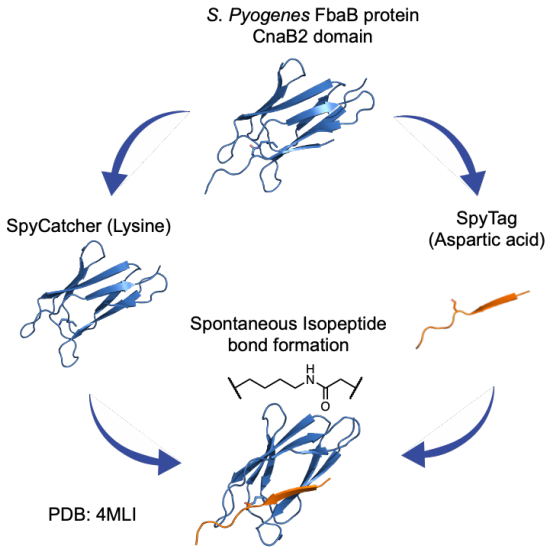
Simulations of the ester-linked and amide-linked substrate were initiated from models containing a single chain of the condensation domain (chain A of the crystal structure in each case) and substrate-activated PPant (from chain B of the crystal structures). Models of these complexes were generated from the symmetry mates within the crystal structure. For simplicity, the PCP domains (residues up to and including residue 2556) were removed, leaving a terminal phosphate group on the PPant substrate. Small molecules (e.g. GOL, NA) and any crystallographic waters further than 8 Å from the protein chain were removed. The resulting model was prepared using the default settings in the Protein Preparation Workflow tool (bond orders were assigned, hydrogens were added, termini were capped, missing side chains filled in, hydrogen bonds were optimized using PROPKA[13] at pH 7.4 and the structure was minimized using OPLS4 until heavy atoms converged to an RMSD of 0.3 Å). For the thioester-liked substrate system, the backbone oxygen in the ester-linked model was changed to a sulfur atom, followed by local minimization of the structure (within 10 Å of the sulfur atom) in Maestro (Schrödinger Release 2022-1: Maestro, Schrödinger, LLC, New York, NY, 2021.).

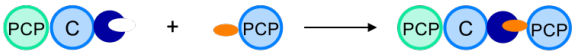
For each system, the protein-PPant complex was solvated in a dodecahedron with a 10 Å buffer around the complex. Enough sodium ions to neutralize the system were then added. Each system was then simulated in triplicate (random-seed triplicate), for 200 ns using OPLS4 and the default settings in Desmond. Prior to these production runs, each system was relaxed using the default "relax model" protocol in Desmond. Simulations were analysed in Maestro (Schrödinger Release 2022-1: Maestro, Schrödinger, LLC, New York, NY, 2021.) using the "Simulation Interactions Diagram" tool. The presence of residue-residue interactions in each frame of the simulations were determined according to the standard definition of interactions in Schrödinger: H-bond interaction were defined by a H-A distance less than 2.8 Å, D-H-A angle greater than 120°, and a H-A-X angle greater than 90°; water bridge H-A distance less than 2.7 Å, a D-H-A angle greater than 110° and a H-A-X angle greater than 80°. Structural figures were prepared in the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

SUPPLEMENTARY FIGURES



Supplementary Figure S1. Single turnover experiments are representative of NRPS function when no chain termination (via diketopiperazine (DKP) formation or enzymatic cleavage) is included in the assay. A) DKP formation and offloading by self-cyclisation can occur at module 2 of an NRPS and has been exploited widely for biochemical characterisation of systems comprising two NRPS modules. B) Peptide is hydrolysed (or cyclised) by a thioesterase (TE) domain in the terminal module (shown here in module 3), leading to multiple cycles of catalysis. C) When no TE domain is included, the machinery is effectively stalled at module 3 and hence the product needs to be chemically cleaved for analysis; this also means that such assembly lines are capable of *single turnovers only*.



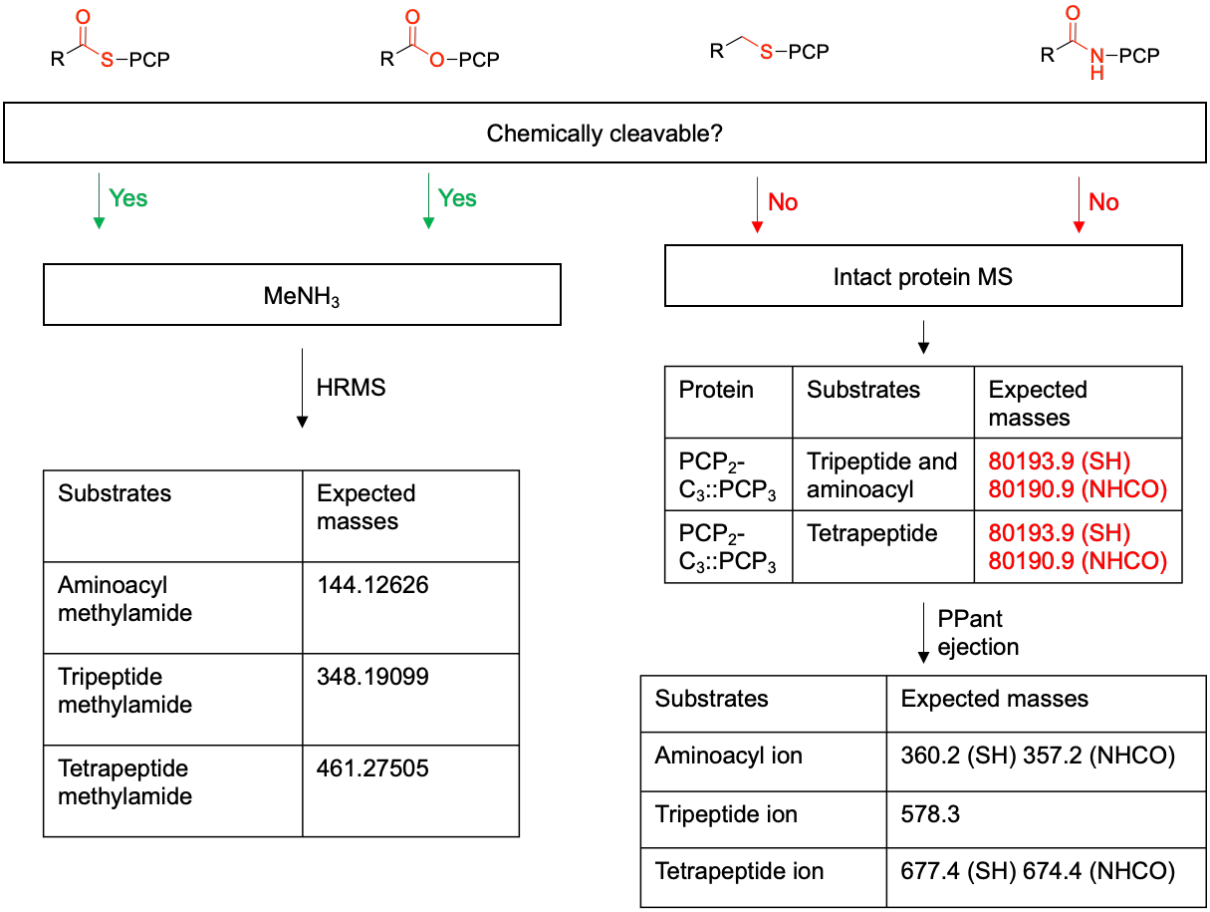


Tetrapeptide conversion

Acceptor substrate	Wild-type	
	Without SpyCatcher	With SpyCatcher
Gly (Natural substrate)	1.4 ± 1.4%	93.9%
L-Leu	-	74.6 ± 7.6%

Izoré and Ho *et al*, **Nat. Commun.** (**2021**), *12*, 2511

Supplementary Figure S2. The irreversible conjugation of SpyCatcher and SpyTag is used to reconstitute covalent NRPS complexes to enable C-domain mediated activity, with the condensation assay result from a previous study showing the essential nature of the covalent linkage of the PCP to the C-domain for C-domain mediated amide bond formation in this case.[1]

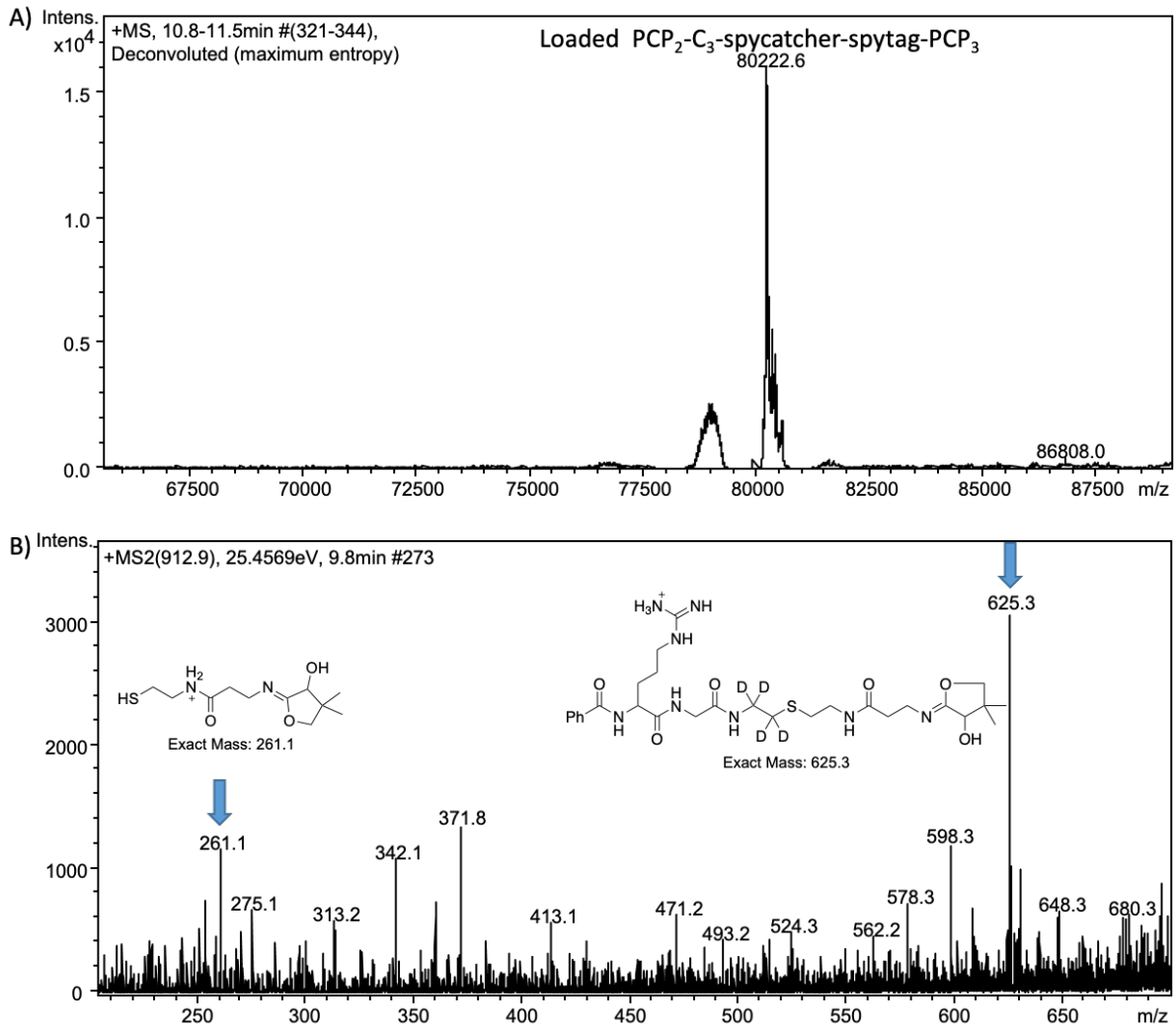


Substrates	Expected masses
Aminoacyl methylamide	144.12626
Tripeptide methylamide	348.19099
Tetrapeptide methylamide	461.27505

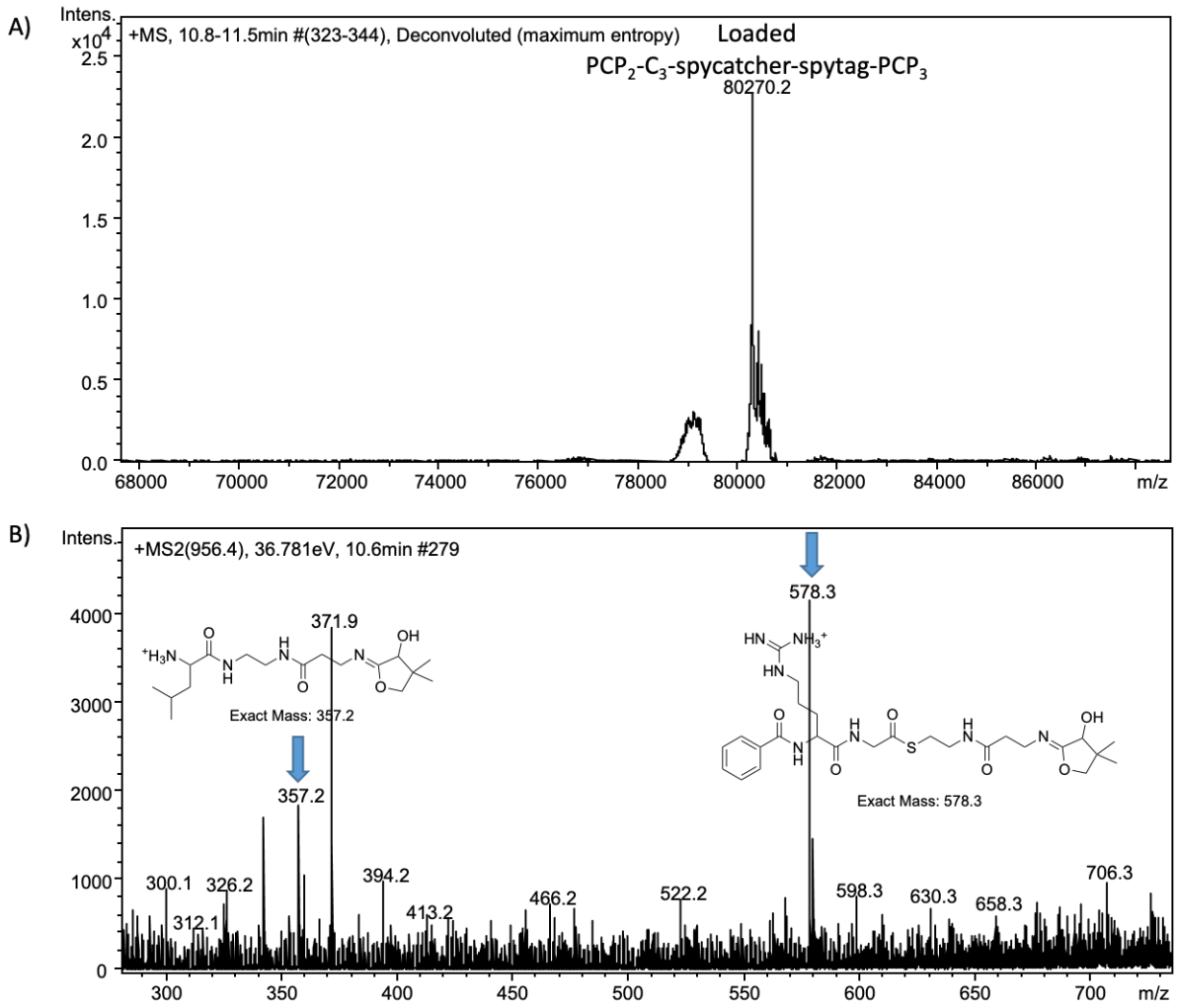
Protein	Substrates	Expected masses
PCP_2-C_3::PCP_3	Tripeptide and aminoacyl	80193.9 (SH) 80190.9 (NHCO)
PCP_2-C_3::PCP_3	Tetrapeptide	80193.9 (SH) 80190.9 (NHCO)

Substrates	Expected masses
Aminoacyl ion	360.2 (SH) 357.2 (NHCO)
Tripeptide ion	578.3
Tetrapeptide ion	677.4 (SH) 674.4 (NHCO)

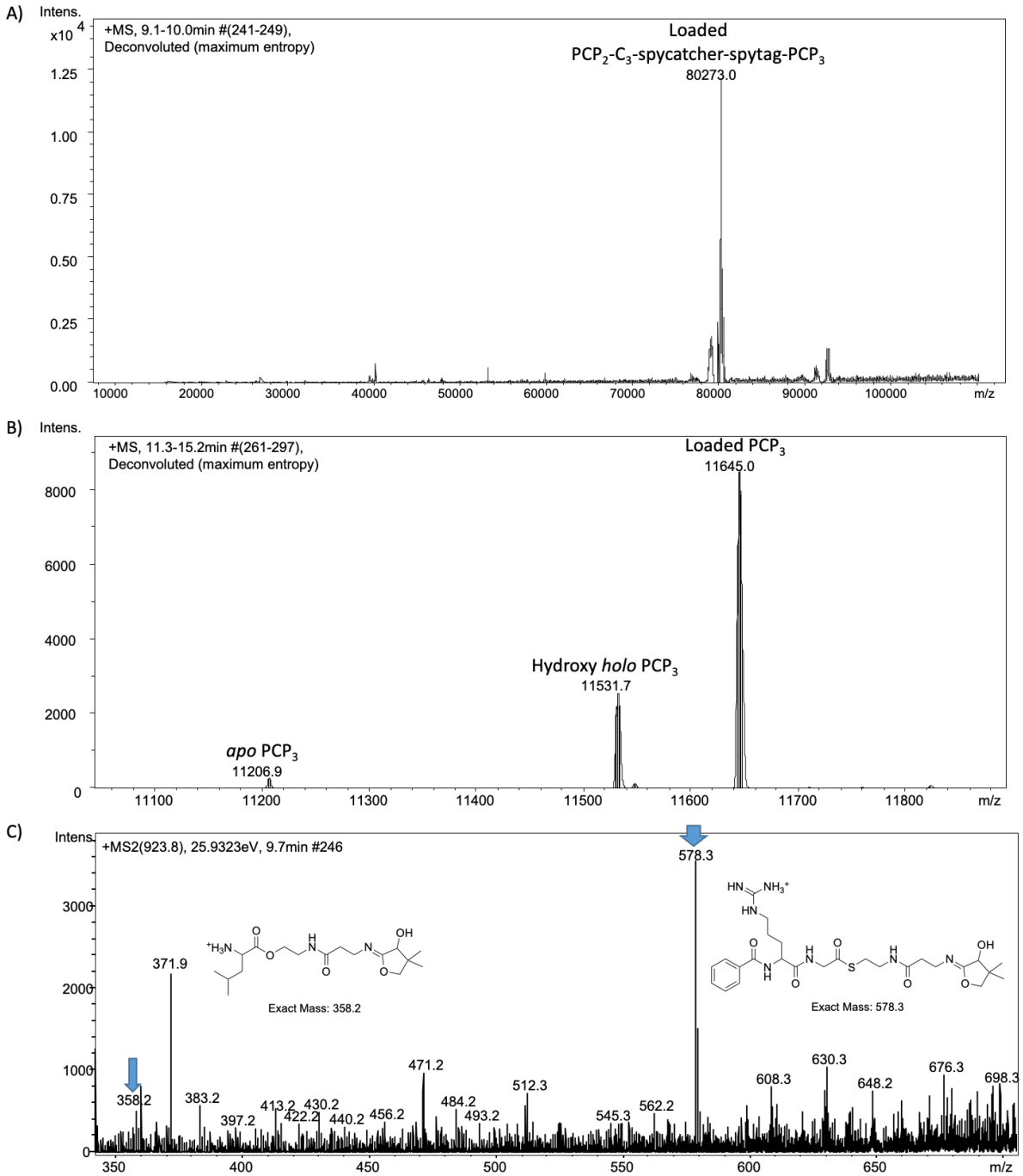
Supplementary Figure S3. Flow chart of the methods available to analyse the substrates bound to native and stabilised PPant linkers using either chemical cleavage (left) or intact protein MS (right). Intact protein MS values are for the BA-Arg-Gly donor peptide and the $Leu_{Stab}N$ **1** acceptor substrate; masses for thioether stabilisation are not shown, but contained in individual figure legends where such controls were used (see **SI Figure S4**).



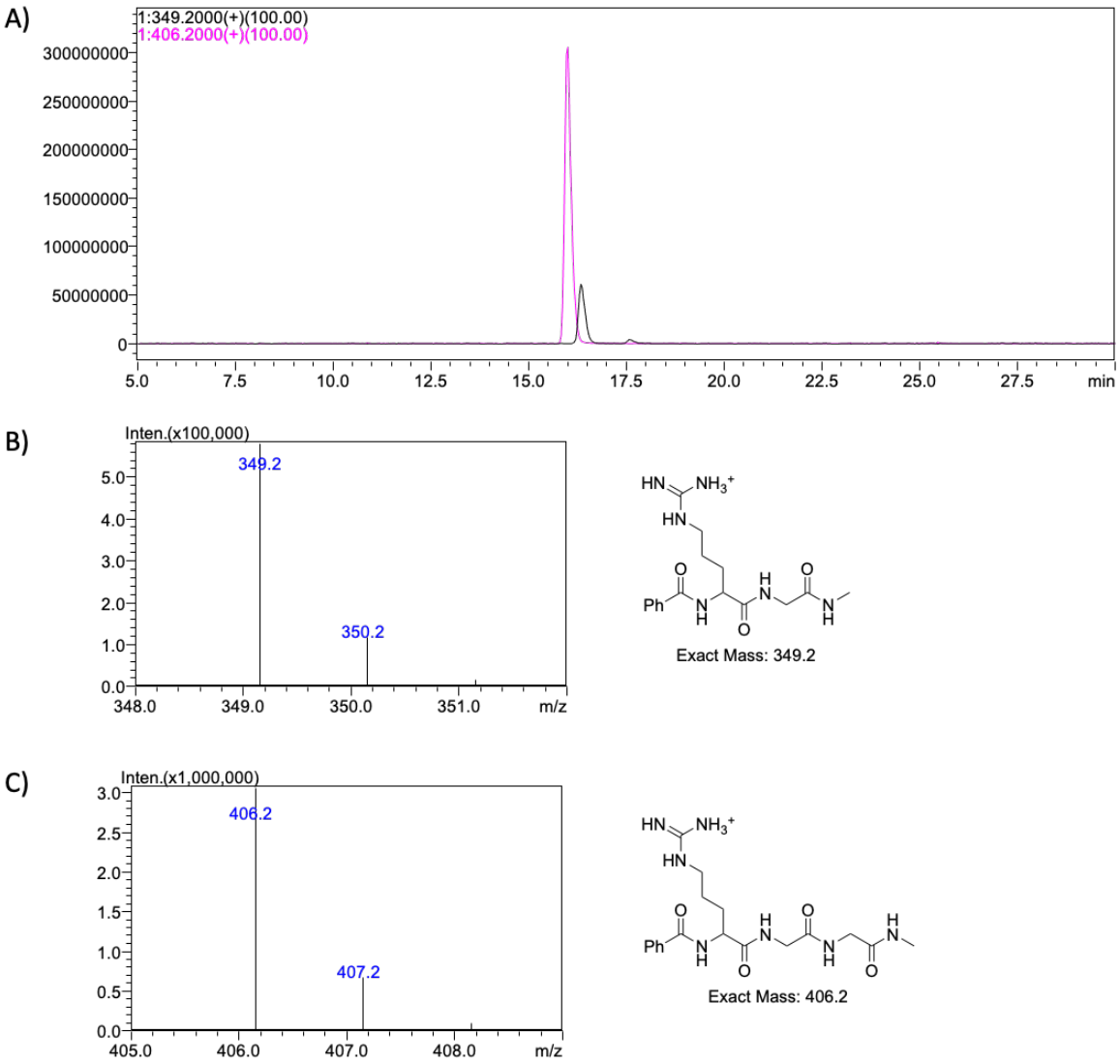
Supplementary Figure S4. PPant ejection results showing extension of *d4*-Gly_{stab}. NanoLC-MS analyses indicating the reconstitution of PCP_2-C_3::PCP_3 WT for BA-D-Arg-Gly donor peptide together with *d4*-Gly_{stab}-CoA as an acceptor substrate. A) Deconvoluted spectrum showing the $[M]^+$ masses observed. C) MS^2 spectrum of the 88+ charged ion 912 (tetrapeptide ion - calculated for $C_{28}H_{41}D_4N_8O_6S^+$ $[M+H]^+$: 625.3, found: 625.3). This result was characterised as previously reported. [1]



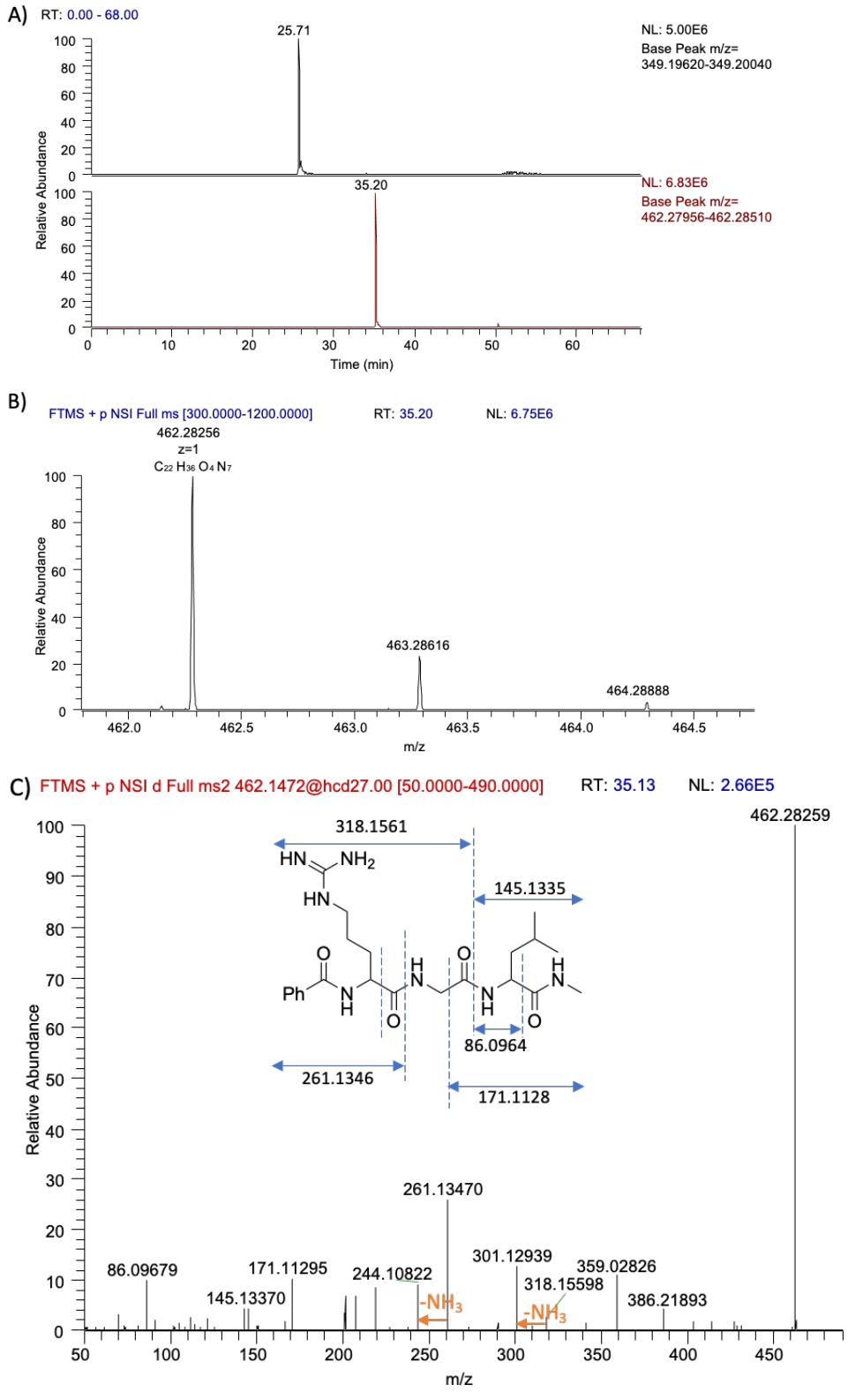
Supplementary Figure S5. PPant ejection results showing no extension of **1**. NanoLC-MS analyses indicating the reconstitution of PCP_2-C_3::PCP_3 WT for BA-D-Arg-Gly donor peptide together with **1** as an acceptor substrate. A) Deconvoluted spectrum showing the $[M]^+$ masses observed. B) MS^2 spectrum of the 84+ charged ion 956, two ejected starting material ions were found (aminoacyl ion - calculated for $C_{17}H_{33}N_4O_4^+$ $[M+H]^+$: 357.2, found: 357.2; tripeptide ion - calculated for $C_{26}H_{40}N_7O_6S^+$ $[M+H]^+$: 578.3, found: 578.3; tetrapeptide ion – calculated for $C_{32}H_{52}N_9O_7^+$ $[M+H]^+$: 674.4, not found).



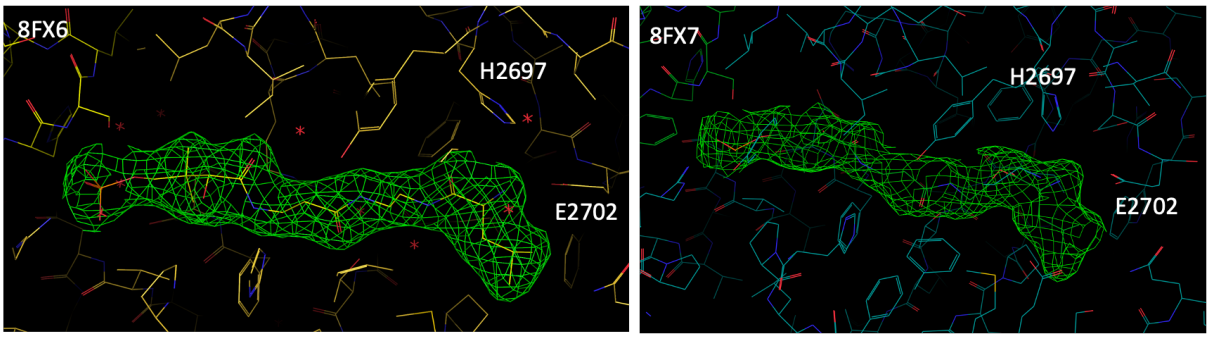
Supplementary Figure S6. PPant ejection results of **2**. NanoLC-MS analyses of the reconstitution of PCP_2-C_3::PCP_3 WT for BA-D-Arg-Gly donor peptide together with **2** as an acceptor substrate. A) Deconvoluted spectrum showing the $[M]^+$ masses of PCP_2-C_3::PCP_3 observed. B) Deconvoluted spectrum showing the $[M]^+$ masses of PCP_3 observed. C) MS^2 spectrum of the 87+ charged ion 923, two ejected starting material ions were found (aminoacyl ion - calculated for $C_{17}H_{32}N_3O_5^+$ $[M+H]^+$: 358.2, found: 358.2; tripeptide ion - calculated for $C_{26}H_{40}N_7O_6S^+$ $[M+H]^+$: 578.3, found: 578.3; tetrapeptide ion – calculated for $C_{32}H_{52}N_9O_7^+$ $[M+H]^+$: 675.4). The results show that PCP_3 has a 1:3 ratio of hydroxy *holo* PCP_3 to loaded PCP_3, the presence of which implies that the elongated substrate may have been hydrolysed during MS, which would influence the result of elongation. This was the reason for switching to chemical cleavage from intact protein MS.



Supplementary Figure S7. LC-MS for reconstitution of PCP_2-C_3::PCP_3 WT using SpyCatcher and SpyTag. Experiments utilised BA-D-Arg-Gly as the donor substrate and Gly as the acceptor substrate. A) Extracted ion chromatograms for masses corresponding to the donor tripeptide (black) and product tetrapeptide (pink) ($[M+H]^+$). B) Accurate mass and isotopic distribution of BA-D-Arg-Gly product. C) Accurate mass and isotopic distribution of BA-D-Arg-Gly-Gly product.



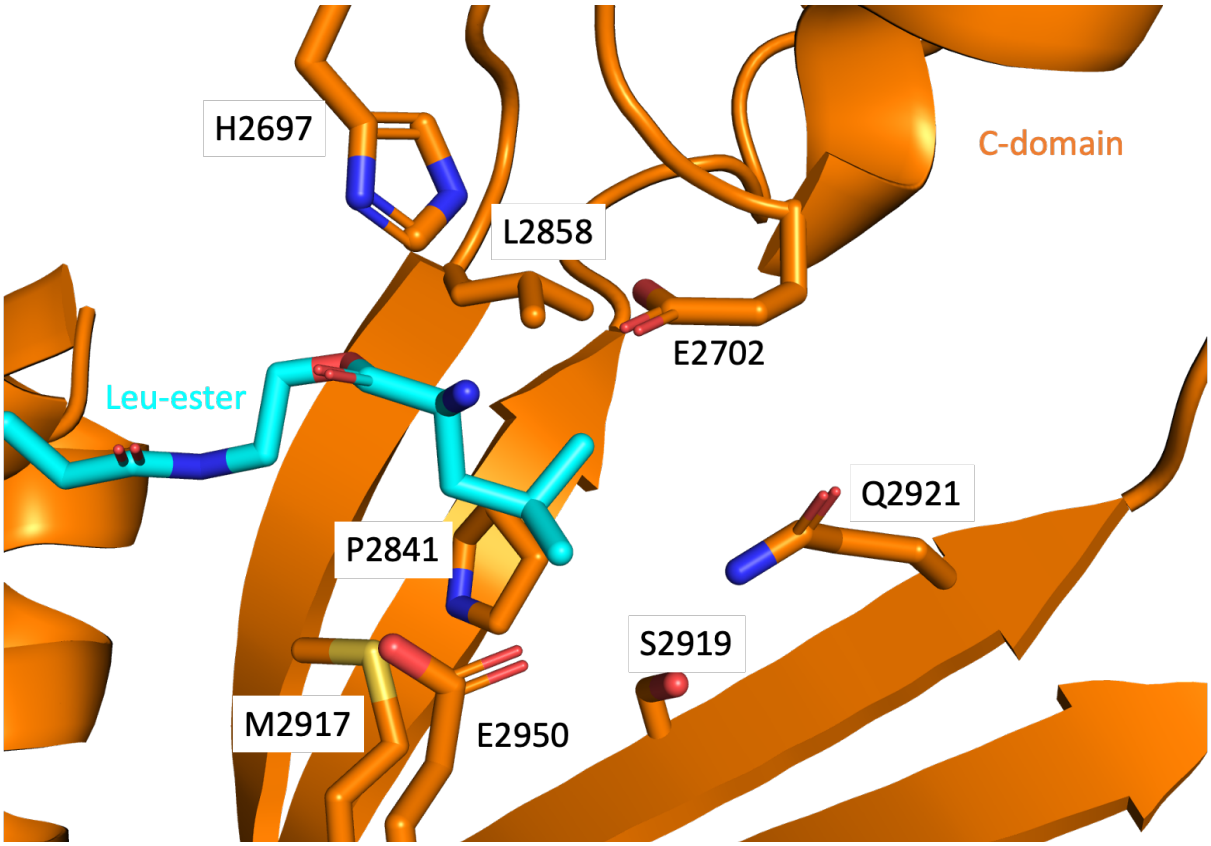
Supplementary Figure S8. LC-HRMS analyses for reconstitution of PCP_2-C_3::PCP_3 WT using SpyCatcher and SpyTag. Experiments utilised BA-D-Arg-Gly as the donor substrate and **2** as the acceptor substrate. A) Extracted ion chromatograms for masses corresponding to the donor tripeptide (upper) and product tetrapeptide (lower) (Orbitrap Fusion, $[M+H]^+$). B) Accurate mass and isotopic distribution of BA-D-Arg-Gly-Leu product. C) MS^2 fragmentation for BA-D-Arg-Gly-Leu product.



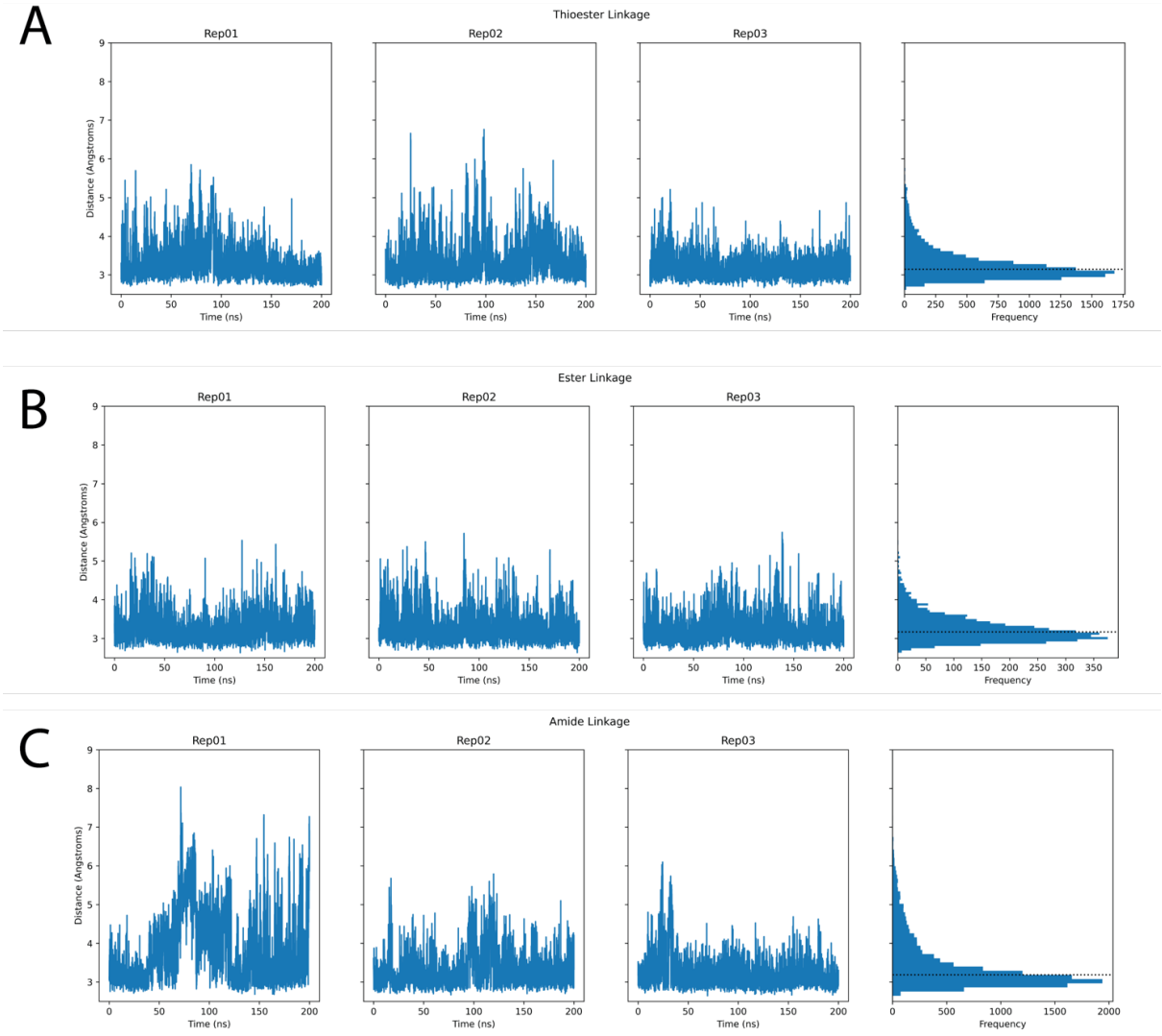
	8FX6	8FX7
CC(1,2)	0.6383	0.5106
CC(1,3)	0.9406	0.8699
CC(2,3)	0.5450	0.5489

Map 1: calculated Fobs with ligand
Map 2: calculated Fobs without ligand
Map 3: real Fobs data

Supplementary Figure S9. Polder maps showing the electron density for the expected PCP-substrates. CC(1,3) is more significant than CC(1,2) and CC(2,3), which supports that the density in the polder map represents the substrate that loaded onto the PCP. CC - Correlation coefficients, 8FX6 - $Leu_{Stab}N$, 8FX7 - $Leu_{Stab}O$.



Supplementary Figure S10. Enlarged view of the C-domain acceptor site pocket residues that are nearby the side chain of the acceptor substrate. Structure shown is from the $Leu_{Stab}O$ **2** complex (PDB 8FX7).



Supplementary Figure S11. Plots showing the distance between His2697 NE2 and the terminal nitrogen of the substrate vs time from molecular dynamics simulations of the thioester-linked (A), ester-linked (B) and amide-linked (C) substrates. Histograms on the right show distribution of distances across the three replicates for each system, and the dotted horizontal line represents the median distance.

SUPPLEMENTARY TABLES

Supplementary Table S1. Data collection and refinement statistics (molecular replacement)

	WT PCP_2-C_3 **$Leu_{Stab}N$** (PDB 8FX6)	**R2577G PCP_2-C_3** **$Leu_{Stab}O$** (PDB 8FX7)
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	104.7, 105.0, 108.3	104.3, 104.7, 106.5
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	48.12 – 2.2 (2.279 – 2.2)	47.49 – 2.2 (2.279 – 2.2)
R_{merge}	0.02 (0.44)	0.3 (0.22)
R_{pim}	0.02 (0.44)	0.03 (0.22)
$I / \sigma I$	8.4 (1.6)	7.7 (2.8)
$CC_{1/2}$	0.99 (0.69)	0.99 (0.62)
Completeness (%)	99 (100)	99 (99)
Redundancy	7.1 (6.2)	7.0 (6.0)
Refinement		
Resolution (Å)	48.1 – 2.2	47.4 – 2.2
No. reflections	61349	59848
R_{work} / R_{free}	0.20 / 0.24	0.18 / 0.23
No. atoms		
Protein	7929	7973
PPant/Leu_{Stab}N/O	58	58
Ion (SO_4)	-	-
Water	207	346
B-factors		
Protein	57.61	42.58
Ppant/Leu_{Stab}N/O	52.31	56.48
Ion (SO_4)	-	-
Water	51.80	45.12
RMS deviations		
Bond lengths (Å)	0.002	0.003
Bond angles (°)	0.44	0.53

[a] Number of crystals = 1

SUPPLEMENTARY REFERENCES

1. T. Izore, Y. T. Candace Ho, J. A. Kaczmarski, A. Gavriilidou, K. H. Chow, D. L. Steer, R. J. A. Goode, R. B. Schittenhelm, J. Tailhades, M. Tosin, G. L. Challis, E. H. Krenske, N. Ziemert, C. J. Jackson and M. J. Cryle, *Nat Commun*, 2021, **12**, 2511.
2. P. D. Walker, C. Williams, A. N. M. Weir, L. Wang, J. Crosby, P. R. Race, T. J. Simpson, C. L. Willis and M. P. Crump, *Angew Chem Int Ed Engl*, 2019, **58**, 12446-12450.
3. M. Strieker, E. M. Nolan, C. T. Walsh and M. A. Marahiel, *J Am Chem Soc*, 2009, **131**, 13523-13530.
4. J. Yin, A. J. Lin, D. E. Golan and C. T. Walsh, *Nat Protoc*, 2006, **1**, 280-285.
5. M. Sunbul, N. J. Marshall, Y. Zou, K. Zhang and J. Yin, *J Mol Biol*, 2009, **387**, 883-898.
6. D. Aragao, J. Aishima, H. Cherukuvada, R. Clarken, M. Clift, N. P. Cowieson, D. J. Ericsson, C. L. Gee, S. Macedo, N. Mudie, S. Panjikar, J. R. Price, A. Riboldi-Tunnicliffe, R. Rostan, R. Williamson and T. T. Caradoc-Davies, *J Synchrotron Radiat*, 2018, **25**, 885-891.
7. T. M. McPhillips, S. E. McPhillips, H. J. Chiu, A. E. Cohen, A. M. Deacon, P. J. Ellis, E. Garman, A. Gonzalez, N. K. Sauter, R. P. Phizackerley, S. M. Soltis and P. Kuhn, *J Synchrotron Radiat*, 2002, **9**, 401-406.
8. W. Kabsch, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 133-144.
9. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr D Biol Crystallogr*, 2011, **67**, 235-242.
10. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 213-221.
11. P. Emsley and K. Cowtan, *Acta Crystallogr D Biol Crystallogr*, 2004, **60**, 2126-2132.
12. C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner and E. D. Harder, *J Chem Theory Comput*, 2021, **17**, 4291-4300.
13. M. H. Olsson, C. R. Sondergaard, M. Rostkowski and J. H. Jensen, *J Chem Theory Comput*, 2011, **7**, 525-537.