SUPPORTING INFORMATION

Structure and substrate sequestration in the pyoluteorin peptidyl carrier protein PltL

Matt J. Jaremko, D. John Lee, Stanley J. Opella, Michael D. Burkart*

Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0358

Correspondence should be directed to mburkart@ucsd.edu.

Contents	Page
A. Synthesis of prolyl and pyrrolyl pantetheine probes	S3
A.1. Synthesis of the prolyl-N-pantetheine probe.	S3
A.2. Synthesis of the pyrrolyl-N-pantetheine probe.	S4
B. Protein NMR Studies.	S5
B.1. Site directed mutagenesis of PItL C87S	S5
B.2. Expression and purification of ¹⁵ N/ ¹³ C apo-PltL	S5
B.3. Expression and purification of ¹⁵ N/ ¹³ C <i>holo</i> -PltL	S5
B.4. Preparation of prolyl-PltL, pyrrolyl-PltL	S5
B.5. Preparation of <i>holo</i> -PltL	S6
B.6. HPLC analysis of ACP proteins	S6
B.7. Protein NMR experiments	S6
B.8. PltL Protein NMR Data Collection and Structure Calculations	S6
Supplemental Figures	Page
Table S1 - Solution NMR Structure Statistics	S8
Figure S1 - Sequence alignment of PCPs generated with MUSCLE	S9
Figure S2 - Phylogenetic tree of PCPs generated with ClustalW	S10
Figure S3 - 20% UREA Page Gel of apo, holo, prolyl, and pyrrolyl PltL	S11
Figure S4 - HPLC-MS spectra of proylyl- and pyrrolyl-PltL species	S12-S14
Figure S5 - ¹ H- ¹⁵ N HSQC of <i>holo</i> -PltL	S15
Figure S6 - ¹ H- ¹⁵ N HSQC of prolyI-N-PItL	S16
Figure S7 - ¹ H- ¹⁵ N HSQC of pyrrolyI-N-PItL	S17
Figure S8 - ¹ H- ¹⁵ N HSQC of ubiquitous ¹⁵ N-labelled <i>holo</i> -PltL	S18
Figure S9 - ¹ H- ¹³ C HSQC of <i>holo</i> -PltL	S19
Figure S10 - 'H-''C HSQC of ubiquitous ''N, ''C-labelled holo-PltL	S20
Figure S11 - 'H-''C HSQC of prolyl-PitL	S21
Figure S12 - 'H-''C HSQC of pyrrolyi-PitL	S22
Figure S13 - ¹³ C/ ¹³ N F1F2-filtered NOE spectrum of <i>holo</i> -PitL	S23
Figure S14 - "C/"N FTF2-Illered NOE spectrum of pytrolyi-N-Pill	524
Figure S15 – Interaction between Fill and 4-DMN particulatine probe	520
Figure S10 - TI-TINOE spectra of pyrioly-N-Fite locused on the protons of the pyriole	S20
Figure S17 - Ferturbations and NOEs of FitL due to pyriole	S21
Figure S19 – Structure comparison of PltL to other pentidul carrier proteins	S20
Figure S20 – Electrostatic comparison of PltL to other pentidyl carrier proteins	S30
Figure S21 - NOFs between PltL and the pyrrole	S31
Figure S22 - 1 H NMR (500 MHz) and 13 C NMR (125 MHz) spectra of 3 in CDCl ₃	S32
Figure S23 - 1 H NMR (500 MHz) and 13 C NMR (125 MHz) spectra of 4 in CD ₃ OD	S33
Figure S24 - 1 H NMR (500 MHz) and 13 C NMR (125 MHz) spectra of 6 in CDCl ₃	S34
Figure S25 - ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 7 in CDCl ₃	S35

A. Synthesis of prolyl and pyrrolyl pantetheine probes

Unless otherwise noted, all reagents and chemical compounds were purchased from Alfa Aesar, Strem Chemicals, Sigma–Aldrich or TCI and used without further purification. Amine 2 was prepared according to [Meier, J. L.; Mercer, A. C.; Rivera, H. Jr.; Burkart, M. D. J. Am. Chem. Soc. 2006, 128, 12174]. Flash chromatography was carried out on 40–63 mesh Geduran Silica Gel 60 (EMD Millipore). Thin layer chromatography (TLC) was conducted on 250 µm Silica Gel 60 F254 glass plates (EMD Millipore). NMR spectra were recorded on a Mercury Plus 400 MHz (Varian) or a VX 500 MHz equipped with XSens cold probe (Varian) spectrometer. FID files were processed using MestRenova version 8.1 (MestreLab Research). NMR spectra were referenced to residual solvent peaks according to S. Budavari, M.J. O'Neil, A. Smith, P.E. Heckelman, The Merck Index, an Encyclopedia of Chemicals, Drugs, and Biologicals, Eleventh Edition, Merck Co., Inc. Rahway, NJ, 1989. Mass spectrometric analyses were conducted on the following instruments: a LCQ Deca (ThermoFinnigan), MAT900XL (ThermoFinnigan), LTQ Orbitrap XL (ThermoScientific), or a LCT Premier (Waters) mass spectrometer. Reversed-phase HPLC separation was performed using a semi-preparative C18 Luna column (250 x 10 mm) at a flow rate of 2.5 mL/min using 600E pump (Waters) and Lambda-Max model 480 UV detector (Waters). Unless stated otherwise, anhydrous solvents were used for all chemical reactions. Reactions were conducted under Ar atmosphere in a round bottom flask or vial capped with a rubber septa and were stirred using a Teflon coated stir bar. All mixtures are provided as v:v ratios. All protein NMR experiments were conducted on a 600 MHz Avance system equipped with a cryoprobe (Bruker) at the UC San Diego Biomolecular NMR Facility. Proteins were purified by FPLC gel filtration over a HiPrep 26/60 Sephacryl S-100 HR column (GE Healthcare Life Sciences). Isotopes for NMR experiments were sourced from Cambridge Isotope Laboratories.

A.1. Synthesis of the prolyl-N-pantetheine probe.



tert-butyl(2R)-2-((2-(3-((4R)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carbox-amido)propanamido)ethyl)carbamoyl)pyrrolidine-1-carboxylate (3): 1-Ethyl-3-(3-dimethyl-aminopropyl)carbodiimide (EDC) (107.7 mg, 0.694 mmol) and hydroxybenotriazole (HOBt) (93.7 mg, 0.694 mmol) were added to a mixture of N-Boc-L-proline (94.5 mg, 0.379 mmol) and amine 2 (120.0 mg, 0.316 mmol) dissolved in dry DMF (5 mL). Within 15 min, N,N-diisopropylethylamine (DIPEA) (121 uL) was added and the mixture was stirred at rt. After 12 h, the solvent was removed under reduced pressure, and the resultant oil was purified by column chromatography (EtOAc to 10% MeOH/EtOAc) to yield amide 3 as a white solid (91%). ¹H-NMR (400 MHz, CDCl₃) δ 7.44 (d), 7.04 (bs), 6.93 (d), 6.84 (bs), 5.47 (s), 4.16 (m), 4.09 (s), 3.83 (s), 3.68 (q), 3.70 (bs), 3.46 (m), 3.42 (m), 3.35 (m), 2.42 (m), 2.18 (bs), 1.97 (bs), 1.85 (m), 1.76 (quin), 1.45 (s), 0.93 (s), 0.92 (s). ¹³C-NMR (100 MHz, CDCl₃) δ 173.41, 171.45, 169.42, 160.20, 155.54, 130.15, 127.55, 113.71, 101.34, 83.92, 80.45, 78.50, 60.48, 55.34 47.32, 39.27, 38.88, 36.27, 35.14 33.09, 30.84, 28.42, 24.65, 21.86, 19.14.

(R)-N-(2-(3-((R)-2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)pyrrolidine-2carboxamide (4): Amide 3 (167.1 mg, 0.290 mmol) was dissolved in 2 mL DCM and 400 uL TFA was slowly added. After 0.5 hr at rt the solvent was removed under reduced pressure. The solid was purified by column chromatography (EtOAc to 40% MeOH/EtOAc) to yield amide 4 as a white solid (50%). ¹H-NMR (400 MHz, CD₃OD) δ 8.14 (t), 8.06 (t), 4.24 (q), 3.91 (s), 3.49-3.30 (m), 2.42 (t), 2.41 (bs), 2.06 (m), 0.93 (s). ¹³C-NMR (100 MHz, CD₃OD) δ 174.69, 172.90, 168.50, 75.80, 68.81, 59.79, 45.90, 39.19, 38.95, 38.34, 35.42, 35.03, 29.44, 23.68, 19.95, 19.45. HR-MS [M+H]+ calcd. 359.2289, found 359.2290.

A.2. Synthesis of the pyrrolyl-N-pantetheine probe.



N-(2-(3-((4R)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamido)-propan-amido)ethyl)-1H-pyrrole-2-carboxamide (6): EDC (68.6 mg, 0.442 mmol) and HOBt (59.7 mg, 0.442 mmol) were added to a mixture of 2-carboxylic acid pyrrole (27.0 mg, 0.243 mmol) and amine 2 (83.7 mg, 0.221 mmol) dissolved in dry DMF (5 mL). Within 15 min, N,N-diisopropylethyl-amine (DIPEA) (77 uL) was added and the mixture was stirred at rt. After 12 h, the solvent was

removed under reduced pressure, and the resultant oil was purified by column chromatography (EtOAc to 10% MeOH/EtOAc) to yield amide 6 as a white solid (55%).¹H-NMR (400 MHz, CDCl₃) δ 10.32 (s), 7.39 (d), 7.38 (t), 7.25 (t), 7.05 (m), 6.88 (t), 6.87 (d), 6.68 (t), 6.16 (t), 5.38 (s), 4.02 (s), 3.75 (s), 3.62 (q), 3.48 (q), 3.44-3.26 (m), 2.39 (t), 1.07 (s), 1.02 (s). ¹³C-NMR (100 MHz, CDCl₃) δ 172.49, 170.11, 162.30, 160.43, 130.31, 127.77, 125.86, 122.20, 113.93, 110.59, 109.97, 101.56, 84.06, 78.59, 55.52, 39.87, 36.29 35.60, 33.27, 22.05, 19.33. Contains residual ethyl acetate.

(R)-N-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)-1H-pyrrole-2-carboxamide (7): Amide 6 (55.7 mg, 0.157 mmol) was dissolved in 80% aq. AcOH (10 mL). After 16 h at rt, the solvent was removed under reduced pressure, and the solid was purified by column chromatography (EtOAc to 20% MeOH/EtOAc) to yield amide 7 as a white solid (69%). ¹H-NMR (400 MHz, $(CD_3)_2SO$) δ 11.43 (s), 8.06 (t), 8.02 (t), 7.70 (t), 6.81 (s), 6.72 (s), 6.04 (s), 5.39 (bs), 4.47 (23), 3.68 (s), 3.26-3.13 (m), 2.25 (t), 0.79 (s), 0.76 (s). ¹³C-NMR (125 MHz, $(CD_3)_2SO$) δ 173.29, 171.15, 161.24, 126.66, 121.69, 110.31, 108.95, 75.34, 68.40, 39.48, 39.02, 38.73, 35.64, 35.25, 21.41, 20.71. Contains residual ethyl acetate. HR-MS [M+H]+ calcd. 355.1976, found 259.1975.

B. Protein NMR Studies.

B.1. Site directed mutagenesis of PltL C87S

The pet22b-PItL plasmid was obtained from the laboratory of Christopher Walsh at Harvard University.^{S1} A PItL C87S mutation was introduced via quickchange PCR to avoid inter-protein disulfide bridge formation through the thiol of the cysteine. The quickchange PCR was performed with the following primers:

C87S_F: 5'-GGCCGAGTCCGCCCTCGAGCACCACC-3'

C87S_R: 5'-CTCGAGGGCGGACTCGGCCTTTAGTTGCTCG-3'

B.2. Expression and purification of ¹⁵N/¹³C apo-PltL

pET22b-PltL (C87S) plasmid was transformed into BL21 (DE3) cells. Uniformly labeled ¹⁵N-PltL was expressed by culturing cells in M9 minimal media (1 L) supplemented with 1 g ¹⁵N-NH₄Cl, 4 g D-glucose, and 50 mg carbenicillin. Uniformly labeled ¹³C,¹⁵N-PltL was prepared by supplementing M9 minimal media (1 L) with 4 g ¹³C-D-glucose, 1 g ¹⁵N-NH₄Cl, and 50 mg carbenicillin. Expression was induced with 0.5 mM IPTG at an OD₆₀₀ of 0.9, and the cells were incubated an additional 16 h at 16 °C. The cells were harvested by centrifugation (1000 RCF). The pelleted cells were re-suspended in 30 mL of lysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 10% glycerol) and lysed in a French pressure cell. The lysate was then centrifuged (12000 RCF) for 1 h to remove insoluble debris. The His₆-tagged proteins were purified using Ni-NTA resin (Novagen) to yield *apo*-PltL.

B.3. Expression and purification of ¹⁵N/¹³C holo-PltL

pET22b-PltL (C87S) and pRep4-Sfp (obtained from the laboratory of Christopher Walsh at Harvard University) plasmids were co-transformed into BL21 (DE3) cells. Isotopically labelled *holo*-PltL was expressed and purified as mentioned in section **B.2**. *holo*-PltL was isolated from Sfp (no his-tag) during Ni-NTA purification.

B.4. Preparation of prolyl-PltL, pyrrolyl-PltL

Loading of the prolyl and pyrrolyl mimetic probes were achieved by a previously described one pot chemoenzymatic reaction converting pantetheinamides to their corresponding CoA analog *in situ* using ATP and three of the *E. coli* CoA biosynthetic enzymes (CoaA, CoaD, CoaE), followed by loading onto *apo*-PltL by Sfp.^{S2} In this study, this was achieved by treating 1.5 ml of

400 μ M *apo*-PltL with 12.5 mM MgCl₂, 8 mM ATP, 2 mM mimetic **4** and **7**, 1.5 μ M Sfp, 0.5 μ M *E. coli* CoaA, 0.7 μ M *E. coli* CoaD, 0.6 μ M *E. coli* CoaE, 0.02% Triton X, 0.1% NaN₃, and 5 mM TCEP in a 150 mM NaCl and 50 mM Tris pH 7.4 buffer containing 10% glycerol. After incubation at 37°C for 12 h, the reactions were purified by FPLC into a 50 mM KPi 7.4 buffer. *Crypto*-PltL species eluted between 155-180 mL.

B.5. Preparation of holo-PltL

To prepare pure *holo*-PltL, *apo*-PltL from section C.3 was incubated with 50 mM Tris pH 7.5, 150 mM NaCl, 12.5 mM MgCl₂, 2 mM coenzyme A, 3 mM TCEP, 0.1% NaN₃ and 1.5 μ M Sfp at 37 °C overnight. The solution was purified by FPLC.

B.6. HPLC analysis of ACP proteins

Loading of probes in sections C.4 was monitored by HPLC analysis. The *crypto*-PltL proteins were passed over a C18 column (Burdick & Jackson) using 10% solvent B for 5 min, then increasing 10–50% B over 5 min, and finally 50-63.3% over 10 min (solvent A = H_2O , 0.05% trifluoroacetic acid; solvent B = CH_3CN 0.05% trifluoroacetic acid). Analyses were conducted on HP 1100 series HPLC (Agilent) equipped with a G1315A DAD detector (Agilent). The elution of the protein was monitored by absorbance at 210 nm.

B.7. Protein NMR experiments

NMR samples were prepared by first concentrating the FPLC pure proteins in 50 mM KPi pH 7.4. A 450 μ L aliquot of the appropriate PltL form was then prepared for NMR by adding 50 μ L of D₂O, 5 μ L of 10% (w/v) NaN₃ and 5 μ L of 0.5 M TCEP adjusted to pH 7.4. The final concentration of the samples evaluated was as follows: 1 mM *holo*-PltL; 1 mM prolyl-PltL; 1 mM pyrollyl-PltL. After adding the sample, each NMR tube was flushed with argon, capped, and sealed with Teflon tape. All NMR spectra were acquired at 25 °C with a 1.2 s recycle delay. HSQC spectra were collected for each sample prepared in the same buffer for CSP analysis, collecting 2048 points (R+I) in the ¹H direct dimension and 256 points (R+I) in the ¹⁵N indirect dimension. Chemical shifts were measured by peak maxima, and chemical shift perturbation was calculated using the formula CSP = ([(0.2\delta N)² + (\delta H)²]/2)^{0.5}.

Samples were prepared and subjected to an HNCACB experiment with 2048 points (R+I) in the direct ¹H dimension and 96 points (R+I) in both the ¹³C and ¹⁵N indirect dimensions. Standard backbone-assignment techniques correlating the backbone amides to their CA and CB, and those for the neighboring residue were employed to confirm our peak assignments. NMR spectra were processed using nmrPipe (NIH),^{S3} and analyzed using the Sparky (UC San Francisco) and CARA software suites.

B.8. PItL Protein NMR Data Collection and Structure Calculations

NMR spectra were collected at 25 °C using NMR spectrometers at the UC San Diego Biomolecular NMR Facility operating at ¹H frequencies of 500, 600, and 800 MHz. Backbone and sidechain assignments were made automatically assigned and manually inspected in CARA and Sparky. Initial backbone assignments were obtained through a 3D HNCACB experiment by standard backbone-assignment techniques correlating the backbone amides to their CA and CB atoms. The side chain ¹H and ¹³C assignments were obtained from ¹³C-edited NOESY-HSQC and TOCSY-HSQC experiment (mixing time 60 ms). NOE data was obtained from 3D ¹⁵N-edited and ¹³C-edited NOESY-HSQC experiments (mixing time 150 μ s). NOE peaks were picked through restricted peak picking in Sparky and inspected manually. Torsion angle restraints (ϕ and ψ) were generated from backbone chemical shift data using TALOS-N.^{S4} NOE peaks were subject to automated assignment and structure calculations with CYANA^{S5} in accordance with

the chemical shift list and TALOS-N restraints. After CYANA calculations, the family of structures were refined in discrete solvent using the Crystallography and NMR System (CNS) 1.3,^{S6,S7} and scripted by Roberto Tejero's WaterRefCNS[3]. Output files were analyzed and prepared for submission by PdbStat.^{S8} Topology, parameter, and linkage files for phosphopantetheinylated serine (PNS) and pyrrole phosphopantetheinylated serine (PYA) were constructed using a combination of known atom types and additional information from the PRODRG server.^{S9}

Assignment Statistics (%) (residues 1-90)	holo	pyrrolyl
backbone H	97.8	97.8
backbone non-H	97.4	64.4
side chain H	98.6	98.4
side chain non-H	78.5	75.7
Structure Calculations		
Distance restraints	3023	3275
short-range, <i>i-j</i> <= 1	2217	2584
medium range, 1< <i>i-j</i> <5	476	381
long-range, i-j >=5	330	310
Dihedral Angle Restraints	139	130
Structure Statistics		
Average final CYANA target function (Å ²)	1.32	1.2
Average cycle1 CYANA target function (Å ²)	40.10	64.27
CNS energies (kcal/mol)	-2949.35	-2932.7
Restraint Violations		
Max. distance restraint violation (Å)	0	0.07
Violated distance restraints > 0.1 Å	0	0
Violated dihedral angles	0	0
CYANA Ramachandran Plot		
Residues in most favored regions	82.1%	83.3%
Residues in additionally allowed regions	17.9%	15.7%
Residues in generously allowed regions	0.0%	0.9%
Residues in disallowed regions	0.0%	0.1%
CYANA RMSD (residues 4-86)		
Average backbone RMSD to mean	0.40 ± 0.09 Å	0.55 ± 0.11 Å
Average heavy atom RMSD to mean	0.84 ± 0.08 Å	1.06 ± 0.12 Å

Table S1. Solution NMR Structure Statistics

	C	11					α2	α3		α4	
	$\wedge \wedge$	$\wedge \wedge$		TTTT	\wedge	\sim	$\sim \sim \sim$	$\wedge \wedge$	$ \land \land \land$	$\Lambda \Lambda$	
		20)	40			60		80	•••	
Piti			PSAKEDE				MV Y L R D F - MG		KYEKDINAIS	RIVEOLKAEC	A 88
PigG		LAYLAENYLG	GECOOE	ALDTPLLEIN			VDLLRRE-AG	VTVPLKAVVP	TNEKSVOHMI	DIVERIOSES	R 82
Clz18	EGYEELLAEL	TGELEORYLP	EGESGL	OPSSPLLEWG	IL TSL S	TTEL	ISFILER - YR	VSIPPEKIVG	ANEKNLDGIT	RLVVSLRDES	A 86
HapG	MLENTL	LTYISENYLG	S DDE G F	NADTPILELN	I I DSGS	LFDL	VDLLRRE-TG	ASIPLNOVTP	GNFASVRKMV	DLVTRLOOH -	- 80
RphO	MTPONL I	FDYLADVWMD	G DA E G L	EDETPIAELN	ILIDSAG	YDL	VHYLOGE - FR	TVPLODISL	KNFRSVSTIT	ALVDRLOKAD	R 83
RedO	MTPRKL I	TDYIADAWMG	G D A E G L	EPDTPIAELN	I I DSAA	FDL	VHYLOGE - FR	VTVPLPEVTL	QNFRSVNAIC	ALVDRLRQNE	G 83
NgnN5	MTQAELTNRL	LDFIRERFLD	N D P A S D L	DGVTPLLDWG	ILNSVN	TMEL	ISFIRTE-FG	CDIPQSSVNV	RNFRNANCIA	ALVADATTEV	R 87
Bmp1	M I E E L	IDFINNELLE	G A A P D L	DQHTPLLELG	ILNSLS	MV R L	LAHVDQR-YG	AKIPEHDITP	VHFENIETLA	ALINTLSAKD	E 81
CouN5	QTREDISSOL	LTFIRESFLA	G D P E G E L	DADTPLLELG	ILNSLN	TAIL	VAHLGED-YG	VHVPLIDVTA	TTFKSVRTLS	ELVHESLSRK	- 86
PrIS	VDRSRINAAI	TDYVKVQFLD	EAEPLSE L	DQETPLLEWG	ILTSLN	TTML	LSFIRSE-LG	VPVPLTHITW	R N F Q N V R A I T	DMVCELSAQN	A 88
Mpy15	MTREAVQEQL	ADYIRRSFLD	DDTTAL	EPNTPLLEWG	ILNSMN	TAKL	LTYIREE-LN	VDVPPTHITG	S Y F K S L D T I T	DLVLSIRSET	A 86
Pyr28	ISRDEIYTKL	SAYIQEQFLD	- DSEISE L	K P D T P L L E W G	VLNSMN	TSIL	LTYIRTD-LG	VTVPPTHITG	RHFYSLDTIT	DMVYELSLQP	A 87
GrsB	E K	LAKIWEEVLG	I S Q I	GIQDNFFSLG	G - HS L K	AITL	ISRMNKE-CN	VDIPLRLLFE	APTIQEIS	NYI	- 65
TycB_1	Q K	LVAIWEQILG	V S P I	GIQDHFFTLG	G - HSLK	AIQL	ISRIQKE-CQ	ADVPLRVLFE	QPTIQALA	A Y V	- 65
FenA	K	LSQLFEDVLK	<mark>E</mark> A P I	GIHDNFFDRG	G - HSLK	ATVL	VSRIAKE-FH	VQVPLKDVFA	H P T V <mark>E</mark> G L A		- 61
Pps4	A K	LSQLFEDVLK	NGH I	GIQDNFFDNG	G - HSLK	ATVL	MSRIAKE - FH	VQVSLKDIFA	H P T V <mark>E</mark> G L A		- 62
Sky30	E	LCGLFAEVLQ	L P S - <u>-</u> V	GVDDDFFTLG	G - HSLL	ATKL	VSRVRGA-LG	VELAVRDVFE	A P T V A A L A		- 62
SnbDE	- E	LTTLFAAVLK	L P R V	GIDDNFFDLG	G - HSLL	ATRL	ISRIRTV-LG	AEITLRDLFE	A P T V A G L G	E R L	- 65
LpmD	<mark>R</mark> A	LCALFADILG	I <mark>E</mark> Q V	G A D D G F F E L G	G - HS F L	AARL	VGRIREE-MG	GELPVRSVFD	A P T P A A L A	R L L	- 65
PstD	RV	VAELFAEVLG	A A T V	GADDDFFALG	G - HS F L	AARL	VARLRER-LD	T E A P V R Y V F E	A P T V A <mark>K</mark> L A	ALL	- 65
NosD	E	LTQIWSRILK	V D K V	GVKDNFFDLG	G - HSLL	TPYL	MAQIKKQ-FG	KDIAIASLYQ	NPTIEQLA		- 61
ltuB	E Q	LICIWQDVLK	V K E I	GVKDNFFDLG	G - HSLR	GMTL	IAKIHKQ-FS	KNISLREVFQ	C P T V <mark>E E</mark> MA	QAI	- 65
MycB	E Q	LVLIWKEVLK	L E Q V	GVKDNFFDLG	G - HSLR	GMTL	VGKIHKQ-FN	KTISLREVFQ	G P T I E E MA	KVI	- 65
Lps1	AQ	LQELVGQVLH	K P S H T I	PLDEDLFRLG	T - DSLT	AMTE	ATAARG CG	WEVSVPIIFQ	HSRLSDLA		- 66
SypA		LAALWCEVLS	VERV	GRODHFFELG	G - HSLL	AVKL	IERMRQ VG	LSAUVRVLFG	QPT		- 55
EntB	PAPIPASKAA		ESDEPF	D-DDNLTDYG	L - DSV R	MMAL	AARWRK VH	GDTDFVMLAK	NPIIDAWW	KLLSREVK	- 78
SyrB1	VKGLSEQEHF	VLKVWSEDLG	LKNI	GVNDDFFDSG	G - ISLA	LIKS	LSKLKIH-YK	TNEDPGTEAD	GATAKVLA	DHITKSLVQA	H 81
CmaD			KPV	ENDEDIFULG		ATQL	TGQVNEA-FG	ANTINMEQFFL		ACLOVAAAAD	K /5
MapC2		VAETWODVLG	VDS C	VEDATELELN	C OSIT	AVEL	VND LEFE LC	VVVDICTIT	DPDLAAFI	REVLAARGAE	6 01
SecC2	TAACSSIEEK		TSC	FEGATELEL	c deve		ATRICEE ID				
BA1221			CEV C	AADODEELCC		AVAI	VALCOSACAC				G 00
BarA	VVPOTETEKO		VETI	SIEDREMDIG	GINSLIN		VNOVTEK-MG	VSIKARWEED	KHKSTIAFIS	KELDAVREOT	L 13
BlmIV	W			DPDANEFALG	G TSP V	ALTE	VTRIFARALA	VEVELARIED		FTL	- 64
BlmVI	FRPRTATERT			GTHDNI EDIG	G - DSL T	VTOF	HSRVVFF-FA	VDIPVRRVYO		VTVDDERRRA	F 83
TycB 3	AAPOTELEOL	LAGIWADVLG		GTODNEFELG	G - DSLK	ALOV	STRLNAS G	WTLAMKELFO	YPTIEEAA		E 80
TvcA	HPPRTETESI	LVSIWONVLG	EK	GIRDNEYSLG	G - DSI O	AIOV	VARLHS YO	LKLETKDLLN	YPTIEOVA	LEVKSTTRKS	D 80
EntF	RAPKAGSETI	IAAAFSSLLG	C DV Q	DADADFFALG	G - HSLL	AMKL	AAOLSRO-VA	ROVTPGOVMV	ASTVAKLA	TIDAEEDST	R 81
BlmV	RPPRTEAERI	LAGVWEEVLE	T G G I	GADDDYFALG	G - DSV H	AIVI	VAKAROA G	LALTAHDLFE	ARTLAAVA	R R A A P A G P A E	P 80
TycC 4	VAPRNATEQQ	LAAIWQEVLG	V <mark>E</mark> P I	GITDQFFELG	G - HSLK	ATLL	IAKVYEY-MQ	IELPLNLIFQ	Y P T I <mark>E K</mark> V A	DFITHKRFES	R 81
TycC 2	AAPQNEIEAK	LAEIWQQVLG	I S Q V	GIHDDFFDLG	G - HSLK	AMTV	VFQVSKA-LE	VELPVKALFE	H P T V A E L A	RFLSRSEKTE	Y 81
TycC 5	VAPRSVWEAR	LAQVWEQVLN	V P Q V	GALDDFFALG	G - HSLR	AMRV	LSSMHNE - YQ	VDIPLRILFE	KPTIQELA	AFIEETAKGN	V 81
TycB_2	V A P A T E V E A K	LVAIWENALG	I S G V	GVLDHFFELG	G - HSLK	AMTV	VAQVHRE-FQ	IDLLLKQFFA	A P T I R D L A	RLIEHSEQAA	G 81
TycC_3	VAAQNDTEAK	LQQIWQEVLG	I P A I	GIHDNFFEIG	G - HSLK	AMNV	I T Q V H K T - F Q	VELPLKALFA	T P T I H E L A	AHIAESAFEQ	F 81
TycC_6	VAPTNAVESK	LAEIWERVLG	V S G I	GILDNFFQIG	G - HSLK	AMAV	AAQVHRE-YQ	VELPLKVLFA	QPTIKALA	Q Y V A T S G K E T	Y 81
SrfA-C	IGPRNEMEET	IAQIWSEVLG	R K Q I	GIHDDFFALG	G - HSLK	AMTA	V P H - Q Q E - L G	IDLPVKLLFE	APTIAGIS	A Y L K N G G S D G	L 80
TycC_1	VAPRNEWEAK	LAAIWESVLG	V <mark>E</mark> P I	G V H D H F F E L G	G - HSLK	AMHV	ISLLQRS-FQ	VDVPLKVLFE	S P T I A G L A	P L V A A A <mark>R K</mark> G T	Y 81
SpoT2	EGPATEREKL	LLSLFADVLG	V <mark>D R</mark> I	GVNDGFFDLG	G - QS L Q	AMRL	ISRIRTT-LG	VDLTISDFFD	A P T V A D L D	AHLTLRRADR	S 81
Conconntian		· · · ·		a second set					-		
Conservation				I.I.I.IIIII			Resserved and	Indianal.		all and an	_

Figure S1 - Sequence alignment of PCPs generated with MUSCLE.^{S11} The sequences are separated by the natural final product that tethers to the PCP, in this case, pyrrole, proline, and other amino acids. The three groups are separated by horizontal black lines and the final peptidyl product is listed in the background in large font. The secondary structure of PltL is depicted above the sequences. Red, blue, green, and black font indicate negative, positive, polar, and nonpolar amino acids, respectively. We included many proline CPs to compare with the pyrrole CPs, since the pyrrole CPs load a proline also. We added various other PCPs in which the structure is known (SrfA-C, PA1221, EntF, etc...) and which are standalone type II PCPs (KedY2, SgcC2, CmaD, etc...).



Figure S2 - Phylogenetic tree of PCPs generated with ClustalW.^{S12}



Figure S3 - 20% UREA Page Gel of apo, holo, prolyl, and pyrrolyl PltL. The PltL Species were stained with coomassie dye. To note, *apo* PltL runs slower in the gel possible due to having a more positive charge (does not contain the Ppant) and/or a more relaxed conformation, which is common for carrier proteins (a wider species will travel slower in the gel).

Figure S4A – HPLCMS data for prolyl-PltL.



Figure S4B – HPLCMS data for pyrrolyl-PltL.



Figure S4 – HPLC-MS spectra of provivil- and pyrrolyl-PltL species. PltL species from in vitro loading (see section C.4) were analyzed by HPLC-MS. HPLC-MS analysis was performed on a Waters Acquity system using an Acquity binary solvent manager, an Acquity column manager, a 2777c robotic autosampler, an Acquity TUV detector and an SQ detector. Samples in buffer (2 µl) were separated on a Waters C18 Xbridge (3.0 x 50mm. 2.5µ) at 0.8 ml/min using a linear gradient of 95% H₂O containing 0.1% formic acid to 30% acetonitrile, containing 0.1% formic acid, in 30 seconds, followed by a linear gradient to 50% acetonitrile (containing 0.1% formic acid) in 7 minutes, followed by cleaning and reequilibrating the column (total run time of 10 minutes). The column was maintained at 55 °C. Protein molecular weights were calculated using Expasy (http://web.expasy.org/protparam/) and ESI-MS (operated in positive mode) data deconvoluted using Spectraphile (http://home.iprimus.com.au/pakholt/lcms/spectraphile.html). A) UV absorbance chromatogram at 273 nm and total ion count (TIC) MS chromatogram of prolyl-PltL. The mass specta corresponds to the peak at 1.96 minutes. The top and bottom spectra corresponds to the PItL fragmentation pattern at a cone voltage of 20V and the MS-ejection^{S13} of the phosphopantetheine probe at a cone voltage of 55V, respectively. B) UV absorbance chromatogram at 273 nm and total ion count (TIC) MS chromatogram of pyrrolyl-PltL. The mass spectrum corresponds to the peak at 2.34 minutes. The top and bottom spectra corresponds to the PItL fragmentation pattern at a cone voltage of 20V and the MS-ejection^{s7} of the phosphopantetheine probe at a cone voltage of 55V, respectively.



Figure S5 - ¹H-¹⁵N HSQC of *holo*-PltL.



Figure S6 - ¹H-¹⁵N HSQC of prolyl-N-PltL.



Figure S7 - ¹H-¹⁵N HSQC of pyrrolyl-N-PltL.



Figure S8 - ¹H-¹⁵N HSQC of ubiquitous ¹⁵N-labelled *holo*-PltL. The red peaks correspond to the phosphopantetheine amides (H-N36: δ 8.05, 120.1; H-N41: δ 8.19, 123.9). The phosphopantetheine arm was ¹⁵N-labelled by co-expressing PltL and Sfp in BL21 cultured in ¹⁵N minimal media.



Figure S9 - ¹H-¹³C HSQC of *holo*-PltL.



Figure S10 - ¹H-¹³C HSQC of ubiquitous ¹⁵N,¹³C-labelled *holo*-PltL. The red peaks correspond to the phosphopantetheine aliphatic proton-carbon pairs (H-H-C28: δ 8.45, 3.72, 73.9; H-C30: δ 0.83, 21.8; H-C31: δ 23.53, 0.87; H-C32: δ 4.00, 76.9; H-H-C37: δ 3.47, 3.53, 38.3; H*-C38: δ 2.49, 38.1; H*-C42: δ 3.29, 45.2; H*-C43: δ 2.55, 26.1). The phosphopantetheine arm was ¹³C-labelled by co-expressing PltL and Sfp in BL21 cultured in ¹³C minimal media. * denotes pseudoatoms.



Figure S11 - ¹H-¹³C HSQC of prolyl-PltL.



Figure S12 - ¹H-¹³C HSQC of pyrrolyl-PltL.



Figure S13 - ¹³C/¹⁵N *F*1F2-filtered NOE spectrum of *holo*-PltL. Only intramolecular NOEs of the unlabeled phosphopantetheine are visible.



Figure S14 - ¹³C/¹⁵N *F*1F2-filtered NOE spectrum of pyrrolyl-N-PltL. Only intramolecular NOEs of the unlabeled phosphopantetheine-N-pyrrolyl probe are visible.



Figure S15 - The interaction of PltL and attached solvatochromic 4-DMN (4-N,N-dimethylamino-1,8-naphthalimide) pantetheine probe. We recently developed solvatochromatic pantetheine probes to observe interactions between attached substrates and carrier proteins. 4-DMN is fluorescent when exposed to a hydrophobic environment. Fluorsecence is only seen when the 4-DMN pantetheine probe is attached to PltL (red) and not fluorescence when the 4-DMN is alone in solution (yellow) or in a solution with apo PltL (blue). For methods, see Beld et. al 2014 publication.^{S14}



Figure S16 - ¹H-¹H NOE spectra of pyrrolyl-N-PltL focused on the protons of the pyrrole. a, Pyrrolyl-N-PltL dissolved in H₂O/KPi buffer. Dotted lines indicate the resonances of the pyrrole protons (δ 6.93, 6.76, 6.11). b, NOE strips of the pyrrole protons with peaks corresponding to atoms in PltL labelled.



Figure S17 – Pertrubations and NOEs of PltL due to pyrrole. a, Chemical shift perturbations due to pyrrole attachment of key atoms in PltL. Shown are methyl (top) and alpha (bottom) regions of overlaid *holo* (blue) and pyrrolyl (red) ¹H-¹³C HSQCs. b, 1D slices of L35 H δ 2*, I45 H δ 1*, M49 H ϵ *, and I65 γ 2* NOE strips. Dotted lines indicate the resonances of the pyrrole protons (δ 6.93, 6.76, 6.11).



Figure S18 - Solution NMR structures of PltL species, represented as a family of structures with the 20 lowest CYANA target functions. a, *holo*-PltL. b, PyrrolyI-PltL.



Figure S19 - Structure comparison of PltL to other peptidyl carrier proteins. The top row consists of all type II PCPs structure determined. PDB IDs: PltL (2n5h), 4neo (Blml), 4hkg (A3404), 2mr8 (Tcp7 PCP), 3tej (EntF PCP), 2vsq (SrfA-C PCP), 4mrt (TycC3 PCP), 2jgp (TycC5 PCP), 4dg9 (PA1221). The conserved PPant serine is highlighted in magenta.



Figure S20 - Electrostatic comparison of PltL to other peptidyl carrier proteins. The top row consists of all type II PCPs structure determined. Carrier proteins are in the same orientation and order as Fig. S19. Red, white, and blue are negative, neutral, and positive charged, respectively.



Figure S21 – NOEs between PltL and the pyrrole. a, Key NOEs are shown between the pyrrole and PltL. b, Pyrrole sequestration into the helix II/III hydrophobic pocket of PltL. 20 pyrrole rings are shown from the 20 lowest target function structures from CYANA.

Figure S22. ¹H NMR (500 MHz) and ¹³C NMR (125 MHz) spectra of 3 in CDCl₃



Figure S23. ¹H NMR (500 MHz) and ¹³C NMR (125 MHz) spectra of **4** in CD₃OD.



Figure S24. ¹H NMR (500 MHz) and ¹³C NMR (125 MHz) spectra of 6 in CDCl₃



Figure S25. ¹H NMR (500 MHz) and ¹³C NMR (125 MHz) spectra of 7 in CDCl₃



Supplemental References

(S1) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. Chem. Biol. 2002, 9, 171.

(S2) Worthington, A. S.; Burkart, M. D. Org. Biomol. Chem. 2006, 4, 44.

(S3) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.

(S4) Shen, Y.; Bax, A. J. Biomol. NMR 2013, 56, 227.

(S5) Güntert, P. Eur. Biophys. J 2009, 38, 129.

(S6) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr D Biol Crystallogr* **1998**, *54*, 905.

(S7) Brunger, A. T. *Nat. Protocols* **2007**, *2*, 2728.

(S8) Locally installed, found at <u>http://biopent.uv.es/~roberto/Index.php?sec=water</u>

(S9) Bhattacharya, A.; Tejero, R.; Montelione, G. T. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 778.

(S10) Schuttelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallographica Section D* **2004**, *60*, 1355.

(S11) Edgar, R. C. Nucleic Acids Res. 2004, 32, 1792.

(S12) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J.; Higgins, D. G. *Bioinformatics* **2007**, *23*, 2947.

(S13) Meier, J. L.; Patel, A. D.; Niessen, S.; Meehan, M.; Kersten, R.; Yang, J. Y.; Rothmann, M.; Cravatt, B. F.; Dorrestein, P. C.; Burkart, M. D.; Bafna, V. *J. Proteome Res.* **2010**, *10*, 320.

(S14) Beld, J.; Cang, H.; Burkart, M. D. Angew. Chem. Int. Ed 2014, 53, 14456.