

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Investigation of protein-ligand and protein-protein interactions in type II non-ribosomal peptide synthetases

Permalink

<https://escholarship.org/uc/item/9kt6k1h5>

Author

Jaremko, Matt J.

Publication Date

2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Investigating protein-ligand and protein-protein interactions in type II non-ribosomal
peptide synthetases

A dissertation submitted in partial satisfaction of the requirements for the degree of
Doctor of Philosophy

in

Chemistry

by

Matt J. Jaremko

Committee in charge:

Professor Michael D. Burkart, Chair
Professor Pieter C. Dorrestein
Professor Thomas Hermann
Professor Bradley Moore
Professor Stanley J. Opella
Professor Jerry Yang

2017

Copyright

Matt J. Jaremko, 2017

All rights reserved

This dissertation of Matt J. Jaremko is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2017

DEDICATION

To my family, who know little details of my PhD work, but provide endless love and support through the years.

EPIGRAPH

To some extent I happily don't know what I'm doing. I feel that it's an artist's responsibility to trust that.

David Byrne

TABLE OF CONTENTS

Signature Page	iii
Dedication.....	iv
Epigraph	v
Table of Contents	vi
List of Abbreviations.....	viii
List of Figures.....	ix
List of Tables.....	xv
List of Schemes	xvi
Acknowledgements	xvii
Vita	xx
Abstract of the Dissertation	xxi
Chapter 1. Type II non-ribosomal peptide synthetases.....	1
1.1 Overview of type II non-ribosomal peptide synthetases	1
1.2 Pyrrole biosynthesis in type II NRPSs	3
1.3 Structure and function of the type II PCP and other type II CPs	6
1.4 Chemoenzymatic manipulation of the peptidyl carrier protein.....	7
1.5 Interactions between the peptidyl carrier protein and adenylation domain.....	10
1.6 FADH ₂ -dependent halogenases in type II NRPSs.....	13
1.7 References	16
Chapter 2. PKS substrate sequestration in actinorhodin biosynthesis.	21
Chapter 3. Structure and Substrate Sequestration in the Pyoluteorin Type II Peptidyl Carrier Protein PltL	85
Chapter 4. Engineering protein-protein interactions in NRPS type II PCPs	126
4.1 Abstract.....	126
4.2 Introduction	126
4.3 Results and Discussion.....	127
4.4 Conclusion	133
4.5 Supporting Content.....	134
4.6 Supporting Table and Figures	142
4.7 References	159

Chapter 5. Fluorescent Mechanism-Based Probe for Aerobic Flavin-Dependent Enzyme Activity	164
--	-----

LIST OF ABBREVIATIONS

Fatty acid synthase	FAS
Polyketide synthase	PKS
Non-ribosomal peptide synthetase	NRPS
Carrier protein	CP
Acyl carrier protein	ACP
Peptidyl carrier protein	PCP
Pyoluteorin type II peptidyl carrier protein	PltL
Prodigiosin type II peptidyl carrier protein	PigG
Undecylprodigiosin type II peptidyl carrier protein	RedO
Actinorhodin polyketide synthase	Act
4' Phosphopantetheinyl transferase	PPTase
Surfactin phosphopantetheinyl transferase	Sfp
Adenylation domain	A domain
Pyoluteorin adenylation domain	PltF
Prodigiosin adenylation domain	PigI
Undecylprodigiosin adenylation domain	R
Pyoluteorin halogenase	PltA
Coenzyme A	CoA
Methyl-S-Coenzyme A	mCoA
<i>E. coli</i> Pantothenate Kinase	CoaA
<i>E. coli</i> Phosphopantetheine adenylyl transferase	CoaD
<i>E. coli</i> Dephosphocoenzyme A kinase	CoaE
Adenosine triphosphate	ATP

LIST OF FIGURES

Figure 1.1 The origin of all anti-cancer drugs from 1981 to 2010.	2
Figure 1.2 Type I vs type II non-ribosomal peptide synthetases.	3
Figure 1.3 Type II pyrrole natural products and biosynthesis.	5
Figure 1.4 Carrier protein structure and activation.	6
Figure 1.5 Manipulation of the coenzyme A (CoA) biosynthetic machinery to generate PCP intermediate species.	9
Figure 1.6 Crystal structure of EntB and EntE tightly bound by a mechanistic inhibitor	11
Figure 1.7 Pyrrole chlorination in pyoluteorin pathway.	12
Figure 2.1 Proposed biosynthesis of 1 from <i>holo</i> -actACP	22
Figure 2.2 Structures of S atom replaced <i>crypto</i> -C8-actACP	22
Figure 2.3 Structures of dual atom replaced <i>crypto</i> -C8-actACP	23
Figure 2.4 Sequestration analysis of atom replaced mimetics.	24
Figure 2.5 Structures of atom replaced <i>crypto</i> -actACP 8a, 8b, and 8c.	25
Figure 2.6 Sequestration analysis of atom replaced cyclic intermediates.	26
Figure S2.1 ¹ H NMR (400Mhz) and ¹³ C NMR (100 MHz) spectra of S1 in CDCl ₃ ..	49
Figure S2.2 ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 15 in CDCl ₃ . 50	50
Figure S2.3 ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 16 in CDCl ₃ . 51	51
Figure S2.4 ¹ H NMR (400Mhz) and ¹³ C NMR (100 MHz) spectra of 17 in CDCl ₃ ... 52	52
Figure S2.5 ¹ H NMR (500 MHz) and ¹³ C NMR (100 MHz) spectra of 19 in CDCl ₃ . 53	53
Figure S2.6 ¹ H NMR (500 MHz) and ¹³ C NMR (100 MHz) spectra of 20 in CD ₃ OD.	54
Figure S2.7 ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 21 in D ₂ O	55

Figure S2.8 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 11 in CD_3OD	56
Figure S2.9 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 24 in CDCl_3 . 57	57
Figure S2.10 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 25 in CDCl_3	58
Figure S2.11 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 26 in CD_3OD	59
Figure S2.12 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of S3 in CDCl_3	60
Figure S2.13 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of S4 in CDCl_3	61
Figure S2.14 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 27 in CDCl_3	62
Figure S2.15 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 28 in CDCl_3	63
Figure S2.16 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of S5 in CD_3OD	64
Figure S2.17 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 22 in CD_3OD	65
Figure S2.18 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 29 in CDCl_3	66
Figure S2.19 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of S6 in CD_3OD	67
Figure S2.20 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 23 in CD_3OD	68
Figure S2.21 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 35 in $\text{DMSO-}d_6$	69
Figure S2.22 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 36 in CD_3OD	70
Figure S2.23 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 30 in CD_3OD	71

Figure S2.24 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 38 in CDCl_3	72
Figure S2.25 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 39 in CDCl_3	73
Figure S2.26 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 40 in CDCl_3	74
Figure S2.27 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 41 in $\text{DMSO-}d_6$	75
Figure S2.28 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 42 in acetone- d_6	76
Figure S2.29 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 31 in CD_3OD	77
Figure S2.30 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 43 in CDCl_3	78
Figure S2.31 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 46 in acetone- d_6	79
Figure S2.32 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 47 in acetone- d_6	80
Figure S2.33 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 32 in CD_3OD	81
Figure S2.34 Docking simulations of ActACP with probes.....	82
Figure 3.1 Structures of pyrrole containing natural products.....	85
Figure 3.2 (a) HSQC overlay of ^{15}N - <i>holo</i> -PltL, ^{15}N -prolyl-N-PltL, ^{15}N -pyrrolyl-N- PltL. (b) Highlights of ^{15}N -HSQC data, illustrating perturbations of residues I65, T66, and F70. (c) CSP plots of ^{15}N -prolyl-PltL and ^{15}N -pyrrolyl-PltL relative to ^{15}N - <i>holo</i> - PltL.	86
Figure 3.3 Pyrrole NMR shift analyses.	87
Figure 3.4 Solution NMR structures of PltL species.....	87
Figure S3.1 Sequence alignment of PCPs generated with MUSCLE	97
Figure S3.2 Phylogenetic tree of PCPs generated with ClustalW	98

Figure S3.3 20% UREA Page Gel of apo, holo, prolyl, and pyrrolyl PltL	99
Figure S3.4 HPLC-MS spectra of prolyl- and pyrrolyl-PltL species	100
Figure S3.5 ^1H - ^{15}N HSQC of <i>holo</i> -PltL	103
Figure S3.6 ^1H - ^{15}N HSQC of prolyl-N-PltL	104
Figure S3.7 ^1H - ^{15}N HSQC of pyrrolyl-N-PltL	105
Figure S3.8 ^1H - ^{15}N HSQC of ubiquitous ^{15}N -labelled <i>holo</i> -PltL.....	106
Figure S3.9 ^1H - ^{13}C HSQC of <i>holo</i> -PltL.	107
Figure S3.10 ^1H - ^{13}C HSQC of ubiquitous ^{15}N , ^{13}C -labelled <i>holo</i> -PltL.....	108
Figure S3.11 ^1H - ^{13}C HSQC of prolyl-PltL	109
Figure S3.12 ^1H - ^{13}C HSQC of pyrrolyl-PltL.....	110
Figure S3.13 $^{13}\text{C}/^{15}\text{N}$ F1F2-filtered NOE spectrum of <i>holo</i> -PltL	111
Figure S3.14 $^{13}\text{C}/^{15}\text{N}$ F1F2-filtered NOE spectrum of pyrrolyl-N-PltL	112
Figure S3.15 Interaction between PltL and 4-DMN pantetheine probe	113
Figure S3.16 ^1H - ^1H NOE spectra of pyrrolyl-N-PltL focused on the protons of the pyrrole.....	114
Figure S3.17 Perturbations and NOEs of PltL due to pyrrole	115
Figure S3.18 Solution NMR structures of PltL species.....	116
Figure S3.19 Structure comparison of PltL to other peptidyl carrier proteins.	117
Figure S3.20 Electrostatic comparison of PltL to other peptidyl carrier proteins.....	118
Figure S3.21 NOEs between PltL and the pyrrole.	119
Figure S3.22 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 3 in CDCl_3	120
Figure S3.23 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 4 in CD_3OD	121
Figure S3.24 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 6 in CDCl_3	122

Figure S3.25 ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 7 in CDCl_3	123
Figure 4.1 (a) Pyrrole formation in pyoluteorin and prodiginine biosynthesis. (b) A domain activity with cognate and non-cognate PCPs.	127
Figure 4.2 The solution NMR structures and electrostatic potentials (ESPs) of PCPs.	129
Figure 4.3 Perturbations of PigG and PltL due to interaction with A domains.....	130
Figure 4.4 PCP loop 1 modification alters interactions with homologous A domains.	133
Figure S4.1 HPLC chromatograms of holo-PCP aminoacylation.....	143
Figure S4.2 Phylogenetic tree of PCPs and A domains.	144
Figure S4.3 ^1H - ^{15}N HSQC of ubiquitous ^{15}N -labelled <i>holo</i> -PigG.	145
Figure S4.4 ^1H - ^{13}C HSQC (aliphatic region) of ubiquitous ^{13}C -labelled <i>holo</i> -PigG..	146
Figure S4.5 ^1H - ^{13}C HSQC (aromatic region) of ubiquitous ^{13}C -labelled <i>holo</i> -PigG.	147
Figure S4.6 Loop 1 region in crystal structures of PA1221, EntF, and LgrA.	148
Figure S4.7 Generation of <i>S</i> -methylated CoA and <i>mholo</i> -PCPs.	149
Figure S4.8 ^1H - ^{15}N HSQC overlays of <i>mholo</i> - ^{15}N -PigG with PigI titrations.	150
Figure S4.9 ^1H - ^{15}N HSQC overlays of <i>mholo</i> - ^{15}N -PigG with PltF titrations..	151
Figure S4.10 Backbone and side-chain root mean square fluctuations of each residue of <i>holo</i> -PltL and <i>holo</i> -PigG.....	152
Figure S4.11 Sequence alignment of WT and mutant PCPs..	153
Figure S4.12 HPLC traces of mutant PCPs in adenylation assay..	154
Figure S4.13 Backbone (heavy atom) root mean square deviations (RMSD) of <i>holo</i> -PltL (red) and <i>holo</i> -PigG (orange)	155
Figure S4.14 Backbone (heavy atom) root mean square deviations (RMSD) of key secondary element of <i>holo</i> -PltL (red) and <i>holo</i> -PigG (orange).....	156

Figure S4.15	Homology model of PigI constructed as described in the main text ...	157
Figure S4.16	Superimposition of the NMR solution structure of PigG (white) and a model structure of the PigG mutant (orange) that shows altered activity	158
Figure 5.1	In-gel fluorescent SDS-PAGE analysis of protein labeling.	165
Figure 5.2	Competition experiments	166
Figure 5.3	Selectivity analysis.....	166
Figure 5.4	Labeling in PltA in lysate and live cells.....	166
Figure S5.1	Catalytic cycles of aerobic flavin-dependent enzymes.....	178
Figure S5.2	Biosynthetic pathways	179
Figure S5.3	Kinetic analyses of BpsA biosynthesis	179
Figure S5.4	Generation of a blue fluorescent <i>crypto</i> -BpsA	180
Figure S5.5	Full gel depiction of gels shown in Figure 1.....	180
Figure S5.6	Full gel depiction of gels shown in Figure 2.....	181
Figure S5.7	MALDI-TOFMS mass shift spectra.	182
Figure S5.8	Flavin disassociation by SDS-PAGE.....	182
Figure S5.9	Western blot analyses	183
Figure S5.10	¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 7 in CDCl ₃	184
Figure S5.11	¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 8 in CDCl ₃	185
Figure S5.12	¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 9 in CDCl ₃	186
Figure S5.13	¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 10 in CDCl ₃	187
Figure S5.14	¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 12 in CDCl ₃	188
Figure S5.15	¹ H (500 MHz) spectrum of 6 in CDCl ₃	189
Figure S5.16	gCOSY (500 MHz) spectrum of 6 in CDCl ₃	190
Figure S5.17	gTOCSY (600 MHz) spectrum of 6 in CDCl ₃	191

Figure S5.18 ^1H , ^{13}C HSQC (600 MHz) spectrum of 6 in CDCl_3	192
Figure S5.19 ^1H , ^{13}C HMQC (600 MHz) spectrum of 6 in CDCl_3	193

LIST OF TABLES

Table S2.1 LCMS validation of <i>crypto</i> -ACPs via PPant ejection	83
Table S3.1 Solution NMR Structure Statistics	96
Table S4.1 <i>holo</i> -PigG Solution NMR Structure Statistics.....	142
Table S5.1 NMR spectroscopic data for probe 6.....	174

LIST OF SCHEMES

Scheme S2.1 Synthesis of Tetraketide Mimetic 11	22
Scheme S2.2 Synthesis of Chain Elongated Mimetics, Heptaketide 22 and Octaketide 23	23
Scheme S2.3 Synthesis of Mimetic 30.....	25
Scheme S2.4 Synthesis of Mimetic 31.....	25
Scheme S2.5 Synthesis of Mimetic 32.....	26
Scheme S3.1 Biosynthesis of Pyoluteorin	86
Scheme S4.1 <i>holo</i> -PigG Solution NMR Structure Statistics	136
Scheme S5.1 Action of flavin-depenet enzymes	164
Scheme S5.2 Proposed mechanism for the inhibition of flavin-dependent enzymes by 1	165
Scheme S5.3 Probe 12 as prepared in 6 steps in an overall yield of 26% from ethyl 2-hydroxy-5-iodobenzoate.....	165
Scheme S5.4 Reporter-labeling of probe 12 was achieved in aqueous and anhydrous conditions.	165

ACKNOWLEDGEMENTS

I have been fortunate to have a loving family, remarkable fiancé, wonderful friends, coworkers, and mentors during my PhD tenure. My time was amazing to due these people in my life, so I would like to say thank you.

I would like to acknowledge my advisor Professor Michael D. Burkart for giving me the opportunity to work in his laboratory and the freedom to research any topic that interested me. I would like to thank Dr. James J. La Clair for his mentorship in and out of the laboratory. Also, I would like to thank Dr. Stanley J. Opella and his research group for their valuable insight which was vital to the development of my protein NMR skills.

My fiancé Kara has been an amazing partner through graduate studies. The ability to share and understand the good and bad of our time at UCSD has been special. My family has always provided immense support and a positive attitude whenever I called or visited them during my PhD tenure. Their love has always kept me level headed over these years. My friends in San Diego, including Brett and Colin, always provided a get away for a day or a weekend. They helped me realize that there is more to San Diego than my research. Even friends far away have made an effort to visit or check-in and I thank them for their thoughtfulness.

Last, but certainly not least, I would like to thank the Burkart group. All the members of the group (past and present) have made my time more enjoyable then I could have imagined. I appreciate the environment they created in and out of the laboratory.

Chapter 1, in part, is currently being prepared for submission for publication: Jaremko, M. J.; Davis, T. D.; Burkart, M. D. “Type II non-ribosomal peptide synthetases.” The dissertation author is the primary author of this manuscript. The dissertation author is the primary co-author of this manuscript along with Tony D. Davis.

Chapter 2, in full, is a reprint of the material as it appears: Shakya, G.; Rivera, H.; Lee, D. J.; Jaremko, M. J.; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdzowski, E. R.; Wong, F.; Tsai, S.-C.; Burkart, M. D. “Modeling Linear and Cyclic PKS Intermediates through Atom Replacement.” *J. Am. Chem. Soc.* 136 (2014): 16792-9. The dissertation author is a contributing author of this manuscript.

Chapter 3, in full, is a reprint of the material as it appears: Jaremko, M.J.; Lee, D. J.; Opella, S. J.; Burkart, M. D. “Structure and Substrate Sequestration in the Pyoluteorin Type II Peptidyl Carrier Protein PltL.” *J. Am. Chem. Soc.* 137 (2015): 11546-9. The dissertation author is the primary author of this manuscript.

Chapter 4, in full, is currently being prepared for submission for publication: Jaremko, M. J.; Lee, D. J.; Patel, A.; Winslow, V.; Opella, S. J.; McCammon, J. A.; Burkart, M. D. “Engineering protein-protein interactions in NRPS type II PCPs.” The dissertation author is the primary author of this manuscript.

Chapter 5, in full, is a reprint of the material as it appears: McCulloch, I. P.; La Clair, J. J.; Jaremko, M. J.; Burkart, M. D. “Fluorescent Mechanism-Based Probe

for Aerobic Flavin-Dependent Enzyme Activity.” *Chembiochem* 17 (2015): 1598-601.

The dissertation author is a contributing author of this manuscript.

VITA

- 2017 Doctor of Philosophy, Chemistry, University of California, San Diego
- 2011 Master of Science, Bioengineering, University of Hawai'i, Mānoa
- 2007 Bachelor of Arts, Biology, University of Hawai'i

PUBLICATIONS

McCulloch, I. P.; La Clair, J. J.; **Jaremko, M. J.**; Burkart, M. D. Fluorescent Mechanism-Based Probe for Aerobic Flavin-Dependent Enzyme Activity. *Chembiochem* 2016, 17 (17), 1598-601.

Jaremko, M.J.; Lee, D. J.; Opella, S. J.; Burkart, M. D. Structure and Substrate Sequestration in the Pyoluteorin Type II Peptidyl Carrier Protein PltL. *J. Am. Chem. Soc.* 2015, 137 (36), p. 11546-9.

Shakya, G.; Rivera, H.; Lee, D. J.; **Jaremko, M. J.**; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdzowski, E. R.; Wong, F.; Tsai, S.-C.; Burkart, M. D. Modeling Linear and Cyclic PKS Intermediates through Atom Replacement. *J. Am. Chem. Soc.* 2014, 136 (48), 16792-9.

ABSTRACT OF THE DISSERTATION

Investigating protein-ligand and protein-protein interactions in type II non-ribosomal
peptide synthetases

by

Matt J. Jaremko

Doctor of Philosophy in Chemistry

University of California, San Diego, 2017

Professor Michael D. Burkart, Chair

Non-ribosomal peptide synthetases (NRPSs) are responsible for the biosynthesis of many pharmaceutically relevant compounds. Type II NRPSs are an emerging subfamily of NRPSs that form hybrid pathways with type I fatty acid synthases (FAS), polyketide synthases (PKS), type I NRPSs, or others. The type II NRPSs commonly contain tailoring enzymes that generate unique substrate modifications, such as dehydrogenations and halogenation. Unlike type I NRPSs, the type II systems consist of standalone enzymes, an ideal feature for combinatorial biosynthesis and metabolic engineering. Unfortunately, engineering efforts have been

met with limited success due to lack of understanding of protein-protein interactions inherent to these pathways.

My dissertation work focuses on using structural biology to investigate type II pyrrole containing natural product pathways, specifically, the antifungal agent pyoluteorin and two prodiginine antitumor agents prodigiosin and undecylprodigiosin. Important to pyrrole formation are the peptidyl carrier protein (PCP) and the adenylation (A) domain. The PCP is post-translationally modified by a 4'-phosphopantetheine group (*holo*-PCP) at a conserved serine residue, and the terminal thiol serves as the point of attachment for all NRPS intermediates. The A domain facilitates covalent attachment of a specific amino acid to the *holo*-PCP. The PCP then shuttles the cargo from one tailoring enzyme to the next in an organized fashion (Fig. 1) The two proteins are vital to precursor incorporation into pathways and substrate alteration. In many pathways (including pyoluteorin), a FADH₂-dependent halogenase introduces chlorines to the pyrrole. Halogenation is essential for the biological activity of many natural products. Structural and chemo-enzymatic investigation of these three enzymes will aid in future engineering efforts in NRPS pathways.

In FAS and PKS pathways, the acyl carrier protein sequesters tethered substrates in a hydrophobic cleft between helix II and III for protection from undesirable reactions. Substrate sequestration in NRPS PCPs has not been demonstrated. To investigate the phenomena, we determined solution NMR structures of the type II PCP PltL, the peptidyl carrier protein from the pyoluteorin pathway (Fig. 2). Naturally, PCP and substrate are covalently attached through a thioester bond, a

labile bond known to hydrolyze in aqueous environments. Chemoenzymatic methods were used to stabilize the pyrrolyl-PltL intermediate for protein NMR studies. The structures of both the *holo*-PltL and pyrrolyl-PltL intermediates were determined as the first functionally characterized type II PCP.

The recognition between PCP and A domain is specific in NRPSs. In fact, the homologous pairs from pyoluteorin and undecylprodigiosin pathways are only active with the cognate partner. We analyzed the homologous PCP and A domain from the prodigiosin pathway and, surprisingly, the PCP PigG was a promiscuous substrate for A domains from all three pathways. We decided to structurally investigate the specificity differences between the pyoluteorin PltL and prodigiosin PigG. The solution NMR structure of *holo*-PigG was determined and compared to the structure of *holo*-PltL. The structural features of the two proteins are similar, as expected due to the distinct pyrrole PCP family. Although, dynamic simulations revealed significantly more flexibility in *holo*-PigG. NMR titration experiments revealed the loop 1 region of both PCPs that was significantly perturbed when the A domain partners were introduced (Fig. 4). Mutations to the loop 1 region of PltL and PigG significantly altered the loading activity of the A domains compared to mutations in other regions. The mutant studies further confirmed the importance of loop 1 in PCPs.

Chapter 1. Type II non-ribosomal peptide synthetases

1.1 Overview of type II non-ribosomal peptide synthetases

Natural products derived from microorganisms have many diverse biological functions and properties. The majority of drugs in the market from the past three decades are natural products or synthetic compounds inspired by natural products (Figure 1).^{1,2} The biosynthesis of these compounds have been studied extensively over the years due to the large spectrum of bioactivities exhibited by these compounds. Investigations into natural product biosynthesis allows for the discovery of novel chemical transformations and also provides a scaffold for future engineering of the pathways for enhances pharmaceutical compounds.

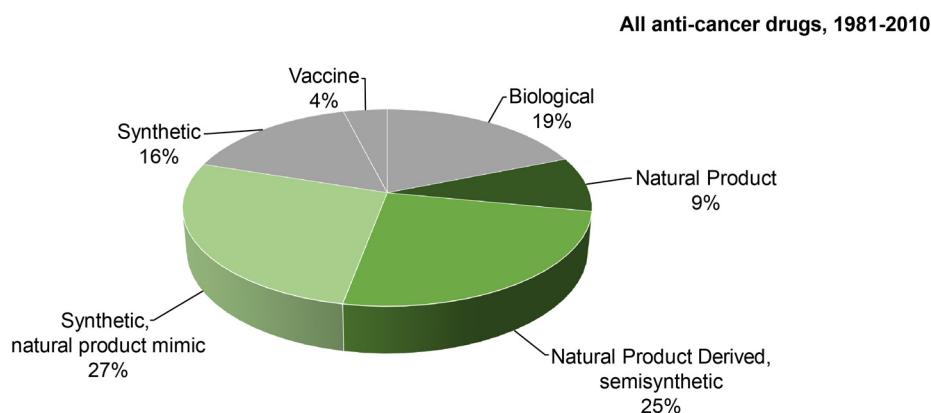


Figure 1.1 The origin of all anti-cancer drugs from 1981 to 2010.¹

Non-ribosomal peptide synthetases (NRPSs) are a large biosynthetic family that generate complex and medicinally-relevant natural products.^{3,4} As the name suggests, NRPSs incorporate precursor amino acids into the biosynthetic machinery for elongation into peptide products. The NRPSs are categorized into type I and type II

due to differences in overall architecture of the synthetases (Figure 2). The traditional type I NRPSs are large modular complexes containing all the enzymes necessary to generate a peptide product in an assembly line fashion. Type II NRPSs are standalone enzymes which coordinate to form unique amino acid derivatives (Figure 2).

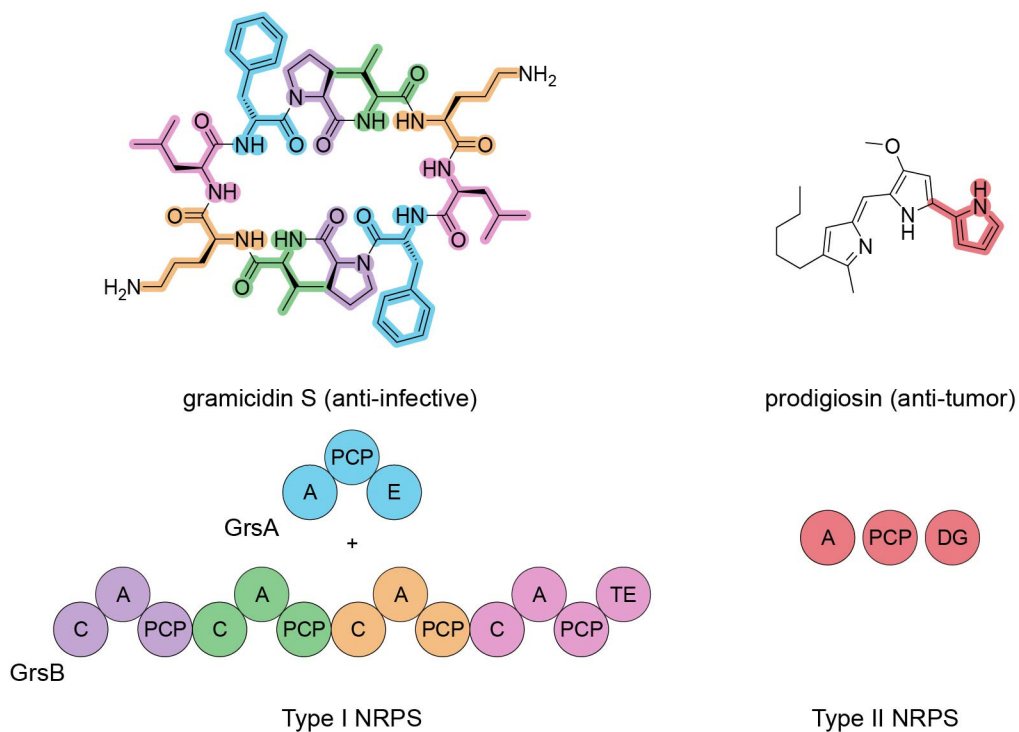


Figure 1.2 Type I vs type II non-ribosomal peptide synthetases. Gramicidin S (type I) and prodigiosin (partial type II) are used as examples of the two NRPS systems. A, adenylation domain; PCP, peptidyl carrier protein; C, condensation domain; DG, dehydrogenase.

The type II NRPSs typically operate on a single amino acid substrate and, depending on the system, may modify the substrate by hydroxylation, dehydrogenase, cyclopropanation, and many more.⁵⁻⁷ The type II derivatives are further transferred to a variety of downstream pathways to generate a final hybrid product. The downstream pathways typically include type I fatty acid synthases (FASs), polyketide synthases (PKSs), or NRPSs. The unique modifications and standalone architecture make the

type II NRPSs ideal targets for combinatorial biosynthesis and metabolic pathway engineering to generate enhanced pharmaceutical compounds. However, engineering efforts have been met with limited success due to lack of understanding of protein-protein interactions inherent to these pathways. The central protein vital to the organization of all NRPSs is the peptidyl carrier protein (PCP).

1.2 Pyrrole biosynthesis in type II NRPSs

The majority of my doctoral research was focused on pyrrole biosynthesis in type II NRPSs. Pyrroles are found in many diverse compounds, including the prodiginines, pyoluteorin, chlorizidine A, and pyralomicin (Figure 3A). The prodiginine compounds are derived from numerous species and have many biological properties.⁸ The prodiginine derivative obatoclax was recently examined in stage II clinical trials as an anticancer therapeutic for hematologic tumors.⁹ Pyoluteorin is an antifungal agent and the producing specie *Pseudomonas putida* serves as a promising biopesticide during crop production.¹⁰ The recently discovered marine *Streptomyces* sp. metabolite chlorizidine A exhibits notable anti-tumor activity targeting human α -enolase and glyceraldehyde-3-phosphate dehydrogenase within the glycolytic pathway.^{11,12} Clorobiocin and other aminocoumarin compounds inhibit bacterial DNA gyrases making the compounds highly effective antibiotic agents.¹³ The *Streptomyces* sp. metabolite hormaomycin is a structurally complex compound that exhibits highly potent activity against *Coryneform* bacteria and influences interspecies differentiation and metabolism.^{14,15} A pyrroline intermediate is generated by a type II NRPS system

in the biosynthesis of anatoxin-a. Anatoxin-a is a neurotoxin produced by cyanobacteria that contributes to water contamination and health issues.¹⁶⁻¹⁸ Overall, the pyrrole natural products are structurally and functionally diverse and display a wide range of bioactivities.

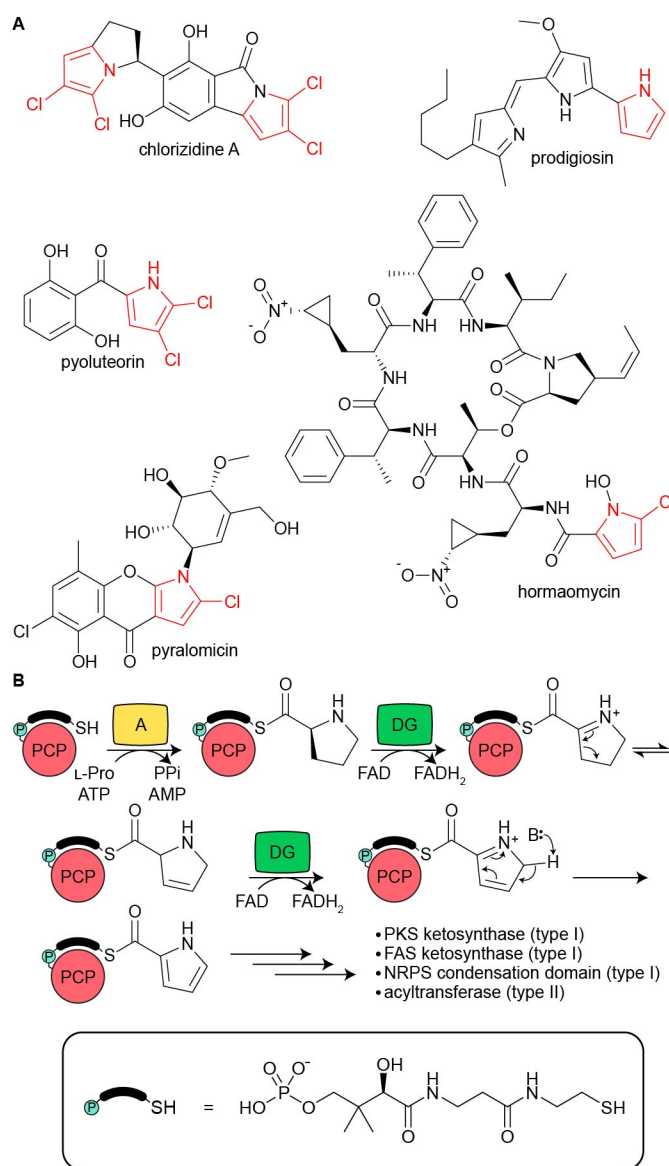


Figure 1.3 Type II pyrrole natural products and biosynthesis. (A) Natural products containing the type II pyrrole. (B) Mechanistic proposal for pyrrole biosynthesis by PCP, adenylation (A) domain, dehydrogenase (DG). The black bar and teal circle above the PCP denotes phosphopantetheine.

The first pyrrole type II enzymes to be identified were from the pyoluteorin, coumermycin, and undecylprodigiosin pathways.¹⁹⁻²¹ The core enzymes for pyrrole formation are the type II PCP, adenylation (A) domain, and FAD-dependent dehydrogenase. The continued discovery of homologous enzymes illustrates the prevalence of the family in nature.²²⁻²⁸ In the pyoluteorin pathway, pyrrole formation is initiated by the A domain PltF, which activates a L-proline to prolyl-AMP and facilitates transfer to the pantetheine arm of PltL (Figure 3B).⁵ Interestingly, the homologous A domains (PltF/RedM) and dehydrogenase (PltE/RedW) in pyoluteorin and undecylprodigiosin biosynthesis were specific for their native PCP illustrating the importance of protein-protein recognition in this family.⁵

The proline dehydrogenases perform a four electron oxidation on the tethered proline to generate pyrrole (Figure 3B). The two step catalysis was confirmed in the homologous dehydrogenases from the biosynthesis of coumermycin and clorobiocin by electrospray ionization-fourier transform mass spectroscopy (ESI-FT MS) analysis.²⁹ Recently, an analogous pyrroline type II system was discovered in the biosynthesis of anatoxin.^{30,31} In contrast to the pyrrole forming dehydrogenase, the anatoxin-a dehydrogenase AnaB selectively catalyzes a two electron oxidation.³¹ Bioinformatic analysis of AnaB indicates the enzyme belongs to the acyl-CoA dehydrogenase superfamily and, therefore, is proposed to catalyze an α/β dehydrogenation on proline and then facilitate an aza-allylic isomerization to generate (*S*)-1-pyrroline-5-carboxylic acid tethered to PCP.^{32,33} No structure of the proline dehydrogenases have been determined yet.

1.3 Structure and function of the type II PCP and other type II CPs

Carrier proteins (CPs) are small, ~10kDa proteins that consist of a four α -helical bundle (Figure 4A). In CP-dependent pathways, the carrier protein must interact with multiple enzymes to coordinate product biosynthesis. Before interacting with partner proteins, all carrier proteins must first be activated from *apo* to *holo* form (Figure 4B). The transformation, catalyzed by a phosphopantetheinyl transferase, involves the attachment of phosphopantetheine from coenzyme A to a conserved serine on the carrier protein.³⁴ The conserved serine is located at the N-terminus of helix II of the carrier protein. During biosynthesis of CP-dependent pathways, precursors and intermediates are all attached to the terminal thiol of the pantetheine arm through a thioester bond. The tethered substrates have been shown to have important functional interactions with the carrier protein.

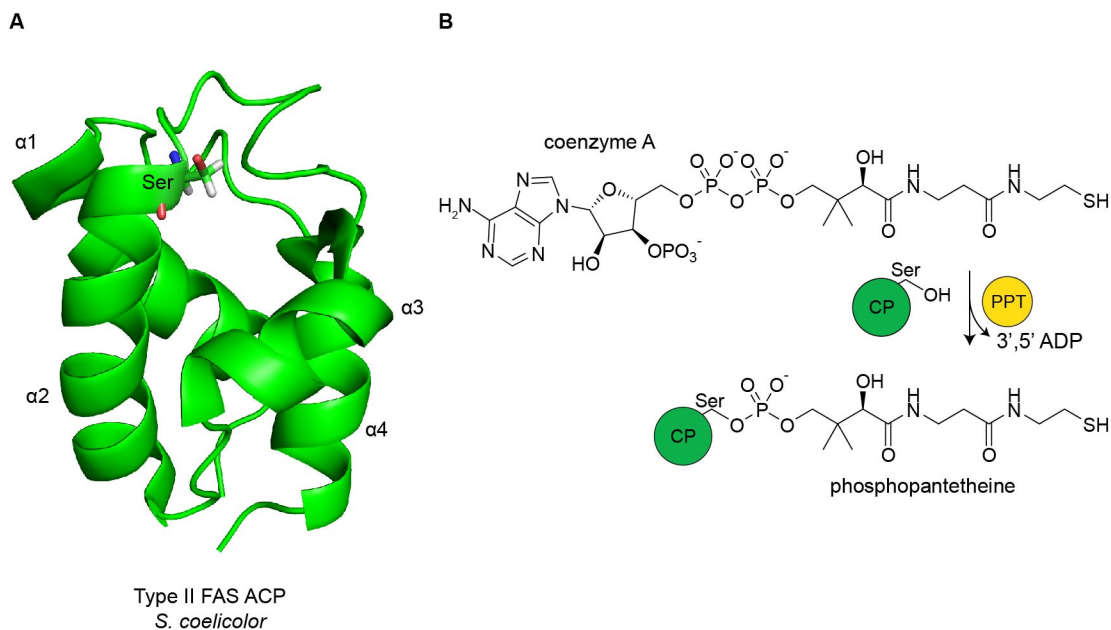


Figure 1.4 Carrier protein structure and activation. (A) The carrier protein consists of four α -helices exemplified by the type II FAS ACP from *Streptomyces coelicolor*. A conserved serine is located at the N-terminus of helix II in all carrier proteins. (B) Carrier proteins are activated by the covalent attachment of phosphopantetheine from coenzyme A to the CP conserved serine. The reaction, termed holoformation, is facilitated by a phosphopantetheinyl transferase (PPT).

In type II FAS and PKS pathways, carrier protein substrate sequestration has been recognized as an important phenomenon. Type II carrier proteins are standalone enzymes that must recognize several partner in a particular order for intermolecular reactions to occur on the carrier protein. Before interacting the partner proteins, the carrier protein has been shown to protect the tethered substrate from reactive species in the cytosol. Here the intermediate is sequestered within a hydrophobic pocket between helix II and III of the carrier protein. Both solution NMR and X-ray crystallographic structures of type II FAS carrier proteins with attached fatty acid intermediates reveal the fatty acid is buried in the hydrophobic pocket between helix II and III.³⁵⁻³⁷ NMR investigation on the type II PKS carrier protein ActACP from the actinorhodin pathway revealed that ActACP-tethered intermediates also sequester in

the pocket between helix II and III (researched discussed in chapter 2).^{38,39} While helix II and III of the PCP has been shown to be important for partner protein interactions, there has been no demonstration of PCP substrate sequestration. My dissertation work includes investigation of substrate sequestration in PltL, a type II PCP from pyoluteorin biosynthesis. The research is discussed in chapter 3.

1.4 Chemoenzymatic manipulation of the peptidyl carrier protein

In NRPS biosynthesis, the PCP-ligand and PCP-partner protein interaction are difficult to study due to labile chemical bonds and fast, transient interactions. The peptidyl carrier protein and substrate intermediates are covalently attached through a thioester bond. The labile properties of the thioester bond allows for frequent transfer of the substrate from PCP to PCP and eventually the final product is released from the final PCP. Although the thioester bond has increased stability in PCP-substrate complexes, the bond is still prone to hydrolysis in aqueous solution.^{40,41} The instability of the thioester bond impedes extensive structural analysis of the protein-ligand interactions. Therefore, chemoenzymatic methods were developed to circumvent the issue.

The biosynthesis of coenzyme A, the cofactor responsible for carrier protein holoformation, consists of five enzymes, CoAA-E,⁴²⁻⁴⁴ although an abbreviated biosynthesis has been demonstrated starting from the precursor pantetheine.^{45,46} The enzymes CoAA, CoAD, and CoAE utilize adenosine triphosphate (ATP) to convert pantetheine to coenzyme A (Figure 5). First, CoAA phosphorylates the 4'-hydroxyl

group of pantetheine, followed by CoAD adenylation at the 4'-phosphate, and lastly CoAE phosphorylates of the adenylate 3'-hydroxyl group. These enzymes were shown to be promiscuous for synthetic pantetheine analogs with modifications on the thiol terminus, including replacement of the terminal thiol with a terminal amine. The finding allowed for the generation of many coenzyme A analogs that were accepted by the phosphopantetheinyl transferase Sfp as a substrate for transfer of the mimetic phosphopantetheine to any carrier protein.

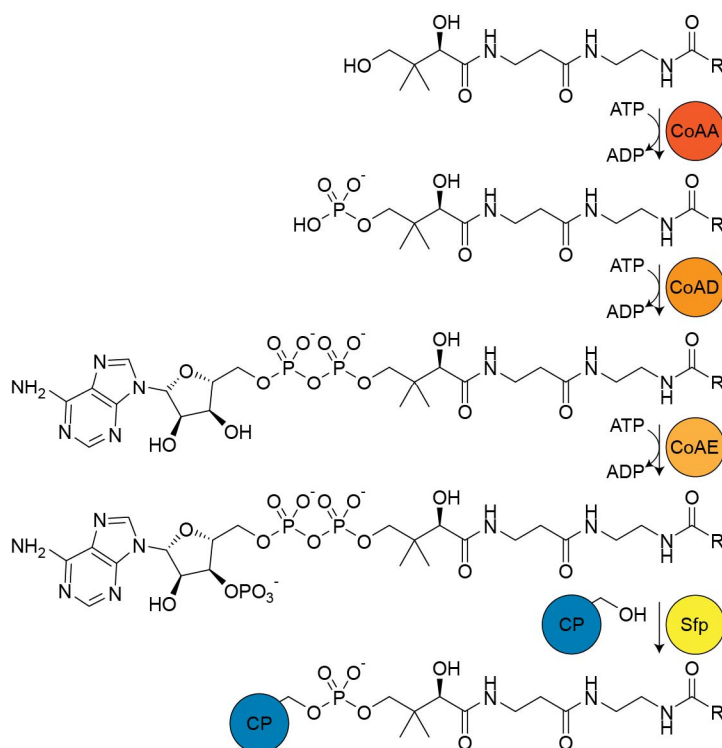


Figure 1.5 Manipulation of the coenzyme A (CoA) biosynthetic machinery to generate PCP intermediate species. The CoA enzymes CoAA, CoAD, and CoAE are able to convert a pantetheine mimetic into a CoA analog. The phosphopantetheinyl transferase Sfp is capable of transferring the phosphopantetheine mimetic from the CoA analog to a CP for further analysis.

The chemoenzymatic modification allows for generation of a stable amide bond between the carrier protein and attached substrate for more thorough analysis of their interactions. The method is utilized in chapter 2 and 3 of my thesis to investigate

substrate interactions with carrier proteins in type II PKS and NRPS pathways, respectively. Furthermore, the method can be utilized to crosslink the carrier protein and enzymatic partners. The interaction between carrier protein and partner protein is transient and therefore, difficult to capture for analysis. The introduction of an enzyme inhibitor on the CP-tethered pantetheine can mechanistically crosslink the protein partners. The technique has been utilized in our laboratory to capture and trap the otherwise transient interaction between the carrier protein and partner for further analysis.⁴⁷⁻⁴⁹ My dissertation work utilizes the technique and is discussed in Chapters 2-4.

1.5 Interactions between the peptidyl carrier protein and adenylation domain

The adenylation (A) domain is as essential to NRPS pathways as the PCP. The A domain is responsible for amino acid incorporation into the biosynthetic machinery. The A domain facilitates covalent attachment of an amino acid onto the PCP-tethered pantetheine. The A domain attaches the amino acid to the PCP in two steps: the adenylate from ATP is transferred to the carboxyl group of an amino acid as an activation step, followed by nucleophilic attack of the carboxylic acid by the pantetheine thiol. The active site of an adenylation domain is amino acid specific to the point where A domain sequence analysis can predict the preferred amino acid.^{50,51} The amino acid incorporation is controlled by the specific active site of the A domain, but just as important is the specificity of the A domain recognition of the PCP.

The interaction between PCP and A domain pairs is specific has demonstrated from early studies investigating cross species activity.^{5,52} The enterobactin A domains EntE and EntF (excised from large module) were able to recognize and aminacylate the non-cognate surfactin PCP SrfB1 (excised) and yersiniabactin Ybt PCP1 (excised), although the activity was reduced by greater than two-fold magnitude.⁵² Furthermore, the homologous pyoluteorin and undecylprodigiosin A domains and PCP were not cross compatible in aminoacylating the PCP with the identical proline substrate.⁵ These studies illustrate the specificity of A domains for their cognate PCP.

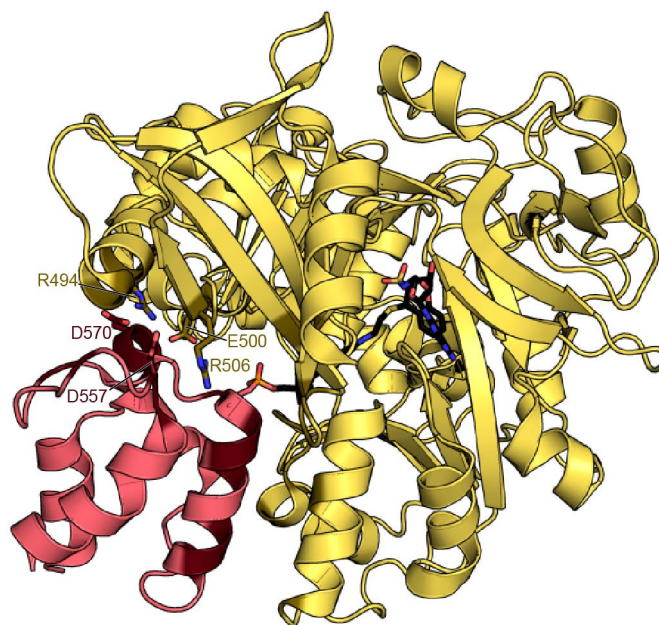


Figure 1.6 Crystal structure of EntB and EntE tightly bound by a mechanistic inhibitor. Darkened regions are important regions of contact between the PCP and A domain. The discovered interactions between the labelled residues led to mutational studies of the corresponding residues in the homologous A domain BasE to improve activity with the PCP EntB.

Mutant studies on the hypothesized interfaces between the PCPs and A domains in the siderophore pathways revealed important regions between the two proteins.⁵³⁻⁵⁵ The HMWP2 aryl-CP is a poor substrate for aminoacylation activity by

the vibriobactin A domain VibE. Several residues of HMWP2 were mutated to the corresponding residues from the PCP VibB based on a Clustal sequence alignment. The VibE activity towards HMWP2 was increased significantly with mutants S49E and H66E, two residues predicted to be located in loop 1 and helix II, respectively.⁵³ A directed evolution system was developed to mutate the vibriobactin type II PCP VibB and yersiniabactin aryl-CP PCP (excised from HMWP2) for increased recognition by the enterobactin NRPS machinery in an EntB-knockout *E. coli* strain.⁵⁴ The cells were grown in iron-limited media to select for colonies with improved enterobactin biosynthesis, an enhancement presumably due to mutant VibB or HMWP2 that increased recognition by EntB protein partners. *In vitro* studies with the evolved mutant VibB, revealed that residues in the predicted loop 1 single turn helix and helix III were important for aminoacylation activity by the adenylation domain EntE.⁵⁴

Eventually, the structure of a PCP•A domain complex was determined in an aminoacylation conformation.⁵⁵ The PCP and A domain were captured in the active state by utilizing a mechanistic-based inhibitor, which mimics the adenylation cofactor ATP to bind tightly in the A domain active site and reacts with the thiol of the PCP-tethered pantetheine.⁵⁶ The vinylsulfonamide inhibitor creates a tightly-bound, non-covalent complex. The structure revealed that loop 1 of the EntB interacts exclusively with the EntE C-terminal region (a flexible subdomain in the ANL superfamily of adenyating enzymes⁵⁷ that undergoes a $\sim 140^\circ$ rotation to facilitate thioester formation⁵⁸) with several salt bridges and other ionic interactions present between the interface (Figure 6). The other important interaction is helix II of EntB and the hairpin

turn (D476-G471) of the EntE C-terminal domain between two long β -strands, which is stabilized by hydrophobic and ionic interactions. The structure permitted mutational studies to increase aminoacylation activity of the non-homologous acinetobactin A domain BasE for the PCP EntB. Based on sequence comparison, BasE mutants M500R, Q506E, and Q512R were created that corresponded to the EntE residues important for EntB loop 1 interactions. These mutants alone and in combination increased the aminoacylation activity towards EntB confirming the importance of the interfaces. The PCP•A domain structures of PA1221 (type II NRPS), EntF (type I NRPS), and LgrA (type I NRPS) have further revealed contacts between the PCP loop 1 and A domain C-terminal subdomain.⁵⁹⁻⁶¹ Part of my dissertation studies the PCP•A domain interactions in the pyrrole-containing pyoluteorin and prodigiosin biosynthetic pathways and the work is discussed in chapter 4.

1.6 FADH₂-dependent halogenases in type II NRPSs

Many natural products contain halogens that contribute to their biological properties,^{62,63} and several type II NRPS systems are responsible for chlorination of pyrrole. In some pathways, such as pyoluteorin, the PCP-tethered pyrrole is further modified by a FAD-dependent halogenase that catalyzes chlorination or bromination at different positions of the ring (Figure 7).^{24,28,64-66} The pyoluteorin FADH₂ halogenase PltA was the first PCP dependent halogenase to be characterized enzymatically.⁶⁵ PltA catalyzes a two electron electrophilic aromatic substitution to form 5-chloropyrrolyl-PCP and then 4,5-chloropyrrolyl-PCP. The halogenases utilize

FADH₂ and O₂ to generate nascent HOCl, which is proposed to form a chloramine species with a proximal lysine residue for electrophilic pyrrole substitution.⁶⁷ The recent crystal structure of PltA highlights K73, the proposed lysine residue involved in chlorine transfer (Figure 7).⁶⁸ Furthermore, the C-terminal region of the PltA structure obstructs the halogenase active site and does not provide enough space for the pyrrolyl-PltL to access the region. The recognition between the partners must induce a conformational change, reiterating the importance of protein-protein interactions in this group. In addition, the halogenase HrmQ from the hormaomycin pathway was heterologously expressed in the clorobiocin producer *Streptomyces roseochromogenes* and, interestingly, the halogenase was able to chlorinate the naturally non-chlorinated pyrrole of clorobiocin.²⁸ The reaction presumably occurs while the pyrrole is PCP-tethered based on previous studies where PltA only chlorinates the pyrrole while attached to PltL.⁶⁵ The combinatorial biosynthetic study suggests that some of the FADH₂ halogenases are promiscuous towards other PCPs, but this has not been shown through *in vitro* studies. My dissertation worked involved development of inhibitors for FADH₂-dependent enzymes, specifically PltA, with the hopes of generating a pantetheine-inhibitor probe that would crosslink PltL and PltA. The complex would provide information on how the proteins interact. The development of an inhibitor of FADH₂-dependent enzymes is discussed in chapter 5.

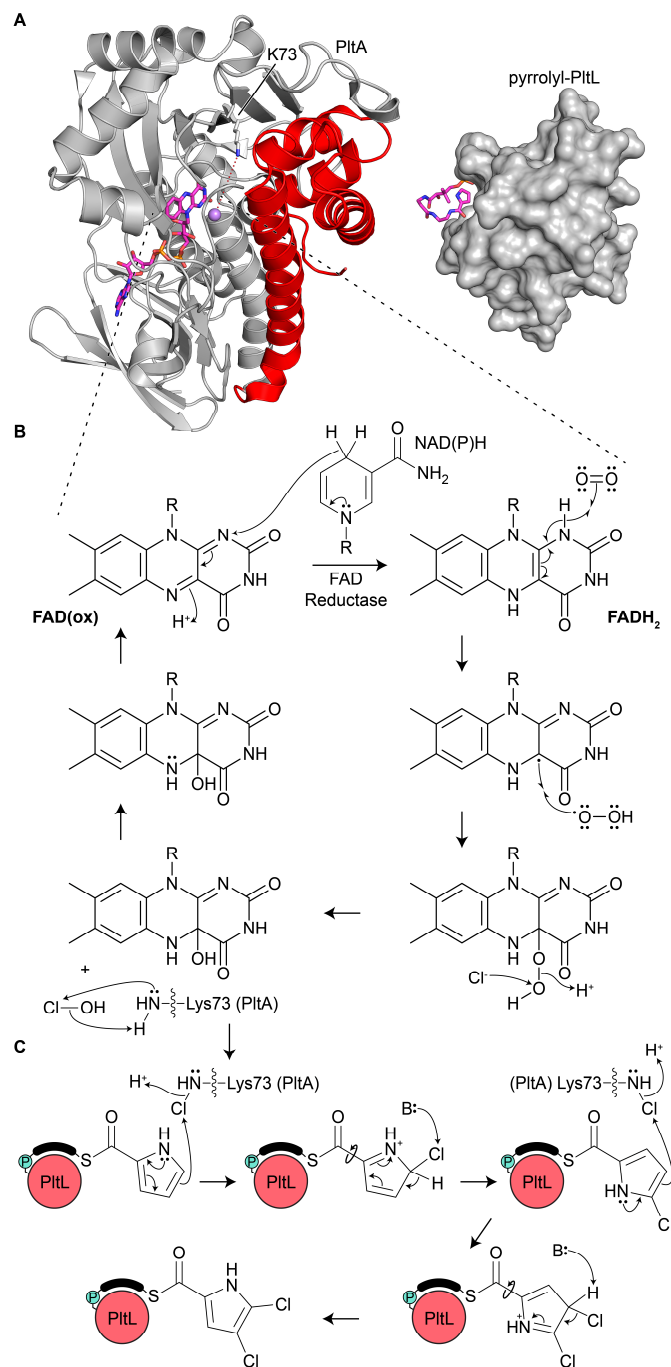


Figure 1.7 Pyrrole chlorination in pyoluteorin pathway. (A) Structures of PltA and PltL. The FAD, conserved K73, and the C-terminal region are highlighted in PltA. The ppant-pyrrole is highlighted on PltL. (B) The proposed FAD mechanism for chlorination transfer to PltL. (C) The proposed chlorination of PltL-tethered pyrrole.

Acknowledgements. The material from Chapter 1, in part, is currently being prepared for submission for publication: Jaremko, M. J.; Davis, T. D.; Burkart, M. D. “Type II non-ribosomal peptide synthetases.” manuscript. The dissertation author is the primary co-author of this manuscript. The rest of the material in this chapter is unpublished work.

1.7 References

- (1) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2012**, *75*, 311-335.
- (2) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2016**, *79*, 629-661.
- (3) Hur, G. H.; Vickery, C. R.; Burkart, M. D. *Nat. Prod. Rep.* **2012**, *29*, 1074-1098.
- (4) Finking, R.; Marahiel, M. A. *Annu Rev Microbiol* **2004**, *58*, 453-488.
- (5) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 171-184.
- (6) Galm, U.; Wendt-Pienkowski, E.; Wang, L.; George, N. P.; Oh, T.-J.; Yi, F.; Tao, M.; Coughlin, J. M.; Shen, B. *Mol. Biosyst.* **2009**, *5*, 77-90.
- (7) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'Connor, S. E.; Walsh, C. T. *Nature* **2005**, *436*, 1191-1194.
- (8) Williamson, N. R.; Fineran, P. C.; Leeper, F. J.; Salmond, G. P. C. *Nature Rev. Microbiol.* **2006**, *4*, 887-899.
- (9) Schimmer, A. D.; Raza, A.; Carter, T. H.; Claxton, D.; Erba, H.; DeAngelo, D. J.; Tallman, M. S.; Goard, C.; Borthakur, G. *Plos One* **2014**, *9*, e108694.
- (10) Hassan, M. N.; Afghan, S.; Hafeez, F. Y. *Pest. Manag. Sci.* **2011**, *67*, 1147-1154.
- (11) Alvarez-Mico, X.; Jensen, P. R.; Fenical, W.; Hughes, C. C. *Org. Lett.* **2013**, *15*, 988-991.
- (12) Alvarez-Mico, X.; Rocha, D. D.; Guimaraes, L. A.; Ambrose, A.; Chapman, E.; Costa-Lotufo, L. V.; La Clair, J. J.; Fenical, W. *ChemBiochem* **2015**, *16*, 2002-2006.

- (13) Galm, U.; Heller, S.; Shapiro, S.; Page, M.; Li, S. M.; Heide, L. *Antimicrob. Agents Chemother.* **2004**, *48*, 1307-1312.
- (14) Andres, N.; Wolf, H.; Zähler, H.; Rössner, E.; Zeeck, A.; König, W. A.; Sinnwell, V. *Helvetica Chimica Acta* **1989**, *72*, 426-437.
- (15) Andres, N.; Wolf, H.; Zähler, H. *Zeitschrift für Naturforschung C* **1990**, *45*, 851.
- (16) Edwards, C.; Beattie, K. A.; Scrimgeour, C. M.; Codd, G. A. *Toxicon* **1992**, *30*, 1165-1175.
- (17) Wood, S. A.; Selwood, A. I.; Rueckert, A.; Holland, P. T.; Milne, J. R.; Smith, K. F.; Smits, B.; Watts, L. F.; Cary, C. S. *Toxicon* **2007**, *50*, 292-301.
- (18) Cadel-Six, S.; Peyraud-Thomas, C.; Brient, L.; de Marsac, N. T.; Rippka, R.; Méjean, A. *Appl. Environ. Microbiol.* **2007**, *73*, 7605-7614.
- (19) Nowak-Thompson, B.; Chaney, N.; Wing, J. S.; Gould, S. J.; Loper, J. E. *J. Bacteriol.* **1999**, *181*, 2166-2174.
- (20) Wang, Z.-X.; Li, S.-M.; Heide, L. *Antimicrob. Agents Chemother.* **2000**, *44*, 3040-3048.
- (21) Cerdeño, A. M.; Bibb, M. J.; Challis, G. L. *Chem. Biol.* **2001**, *8*, 817-829.
- (22) Harris, A. K. P.; Williamson, N. R.; Slater, H.; Cox, A.; Abbasi, S.; Foulds, I.; Simonsen, H. T.; Leeper, F. J.; Salmond, G. P. C. *Microbiology* **2004**, *150*, 3547-3560.
- (23) Li, C.; Roege, K. E.; Kelly, W. L. *Chembiochem* **2009**, *10*, 1064-1072.
- (24) Flatt, P. M.; Wu, X.; Perry, S.; Mahmud, T. *J. Nat. Prod.* **2013**, *76*, 939-946.
- (25) Maharjan, S.; Aryal, N.; Bhattarai, S.; Koju, D.; Lamichhane, J.; Sohng, J. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 687-696.
- (26) Zhang, X.; Parry, R. J. *Antimicrob. Agents Chemother.* **2007**, *51*, 946-957.
- (27) Pojer, F.; Li, S.-M.; Heide, L. *Microbiology* **2002**, *148*, 3901-3911.
- (28) Heide, L.; Westrich, L.; Anderle, C.; Gust, B.; Kammerer, B.; Piel, J. *Chembiochem* **2008**, *9*, 1992-1999.
- (29) Garneau, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. *Biochemistry* **2005**, *44*, 2770-2780.

- (30) Méjean, A.; Mann, S.; Maldiney, T.; Vassiliadis, G.; Lequin, O.; Ploux, O. *J. Am. Chem. Soc.* **2009**, *131*, 7512-7513.
- (31) Méjean, A.; Mann, S.; Vassiliadis, G.; Lombard, B.; Loew, D.; Ploux, O. *Biochemistry* **2010**, *49*, 103-113.
- (32) Dakoji, S.; Shin, I.; Becker, D. F.; Stankovich, M. T.; Liu, H.-w. *J. Am. Chem. Soc.* **1996**, *118*, 10971-10979.
- (33) Dakoji, S.; Shin, I.; Battaile, K. P.; Vockley, J.; Liu, H.-w. *Bioorgan. Med. Chem.* **1997**, *5*, 2157-2164.
- (34) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. *Nat. Prod. Rep.* **2014**, *31*, 61-108.
- (35) Zornetzer, G. A.; Fox, B. G.; Markley, J. L. *Biochemistry* **2006**, *45*, 5217-5227.
- (36) Roujeinikova, A.; Baldock, C.; Simon, W. J.; Gilroy, J.; Baker, P. J.; Stuitje, A. R.; Rice, D. W.; Slabas, A. R.; Rafferty, J. B. *Structure* **2002**, *10*, 825-835.
- (37) Roujeinikova, A.; Simon, W. J.; Gilroy, J.; Rice, D. W.; Rafferty, J. B.; Slabas, A. R. *J. Mol. Biol.* **2007**, *365*, 135-145.
- (38) Haushalter, R. W.; Filipp, F. V.; Ko, K. S.; Yu, R.; Opella, S. J.; Burkart, M. D. *ACS Chem. Biol.* **2011**, *6*, 413-418.
- (39) Shakya, G.; Rivera, H.; Lee, D. J.; Jaremko, M. J.; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdzowski, E. R.; Wong, F.; Tsai, S.-C.; Burkart, M. D. *J. Am. Chem. Soc.* **2014**, *136*, 16792-16799.
- (40) Zornetzer, G. A.; Tanem, J.; Fox, B. G.; Markley, J. L. *Biochemistry* **2010**, *49*, 470-477.
- (41) Goodrich, A. C.; Frueh, D. P. *Biochemistry* **2015**, *54*, 1154-1156.
- (42) Leonardi, R.; Zhang, Y.-M.; Rock, C. O.; Jackowski, S. *Progress in Lipid Research* **2005**, *44*, 125-153.
- (43) Daugherty, M.; Polanuyer, B.; Farrell, M.; Scholle, M.; Lykidis, A.; de Crécy-Lagard, V.; Osterman, A. *J. Biol. Chem.* **2002**, *277*, 21431-21439.
- (44) Mishra, P. K.; Drueckhammer, D. G. *Chemical Reviews* **2000**, *100*, 3283-3310.
- (45) Nazi, I.; Koteva, K. P.; Wright, G. D. *Anal Biochem* **2004**, *324*, 100-105.

- (46) Strauss, E.; Begley, T. P. *J. Biol. Chem.* **2002**, *277*, 48205-48209.
- (47) Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Markwick, P. R. L.; Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, F.; O'Dowd, B.; McCammon, J. A.; Opella, S. J.; Tsai, S.-C.; Burkart, M. D. *Nature* **2014**, *505*, 427-431.
- (48) Blatti, J. L.; Beld, J.; Behnke, C. A.; Mendez, M.; Mayfield, S. P.; Burkart, M. D. *Plos One* **2012**, *7*, e42949.
- (49) Ishikawa, F.; Haushalter, R. W.; Burkart, M. D. *J. Am. Chem. Soc.* **2011**, *134*, 769-772.
- (50) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493-505.
- (51) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. *Nucleic Acids Res.* **2011**, *39*, W362-W367.
- (52) Ehmman, D. E.; Shaw-Reid, C. A.; Losey, H. C.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2509-2514.
- (53) Marshall, C. G.; Burkart, M. D.; Meray, R. K.; Walsh, C. T. *Biochemistry* **2002**, *41*, 8429-8437.
- (54) Zhou, Z.; Lai, J. R.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11621-11626.
- (55) Sundlov, Jesse A.; Shi, C.; Wilson, Daniel J.; Aldrich, Courtney C.; Gulick, Andrew M. *Chem. Biol.* **2012**, *19*, 188-198.
- (56) Qiao, C.; Wilson, D. J.; Bennett, E. M.; Aldrich, C. C. *J. Am. Chem. Soc.* **2007**, *129*, 6350-6351.
- (57) Gulick, A. M. *ACS Chem. Biol.* **2009**, *4*, 811-827.
- (58) Reger, A. S.; Wu, R.; Dunaway-Mariano, D.; Gulick, A. M. *Biochemistry* **2008**, *47*, 8016-8025.
- (59) Mitchell, C. A.; Shi, C.; Aldrich, C. C.; Gulick, A. M. *Biochemistry* **2012**, *51*, 3252-3263.
- (60) Drake, E. J.; Miller, B. R.; Shi, C.; Tarrasch, J. T.; Sundlov, J. A.; Leigh Allen, C.; Skiniotis, G.; Aldrich, C. C.; Gulick, A. M. *Nature* **2016**, *529*, 235-238.
- (61) Reimer, J. M.; Aloise, M. N.; Harrison, P. M.; Martin Schmeing, T. *Nature* **2016**, *529*, 239-242.

- (62) Neumann, C. S.; Fujimori, D. G.; Walsh, C. T. *Chem. Biol.* **2008**, *15*, 99-109.
- (63) Wagner, C.; El Omari, M.; König, G. M. *J. Nat. Prod.* **2009**, *72*, 540-553.
- (64) Agarwal, V.; El Gamal, A. A.; Yamanaka, K.; Poth, D.; Kersten, R. D.; Schorn, M.; Allen, E. E.; Moore, B. S. *Nat. Chem. Biol.* **2014**, *10*, 640-647.
- (65) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13843-13848.
- (66) Höfer, I.; Crüseemann, M.; Radzom, M.; Geers, B.; Flachshaar, D.; Cai, X.; Zeeck, A.; Piel, J. *Chem. Biol.* **2011**, *18*, 381-391.
- (67) Yeh, E.; Blasiak, L. C.; Koglin, A.; Drennan, C. L.; Walsh, C. T. *Biochemistry* **2007**, *46*, 1284-1292.
- (68) Pang, A. H.; Garneau-Tsodikova, S.; Tsodikov, O. V. *J. Struct. Biol.* **2015**, *192*, 349-357.

Modeling Linear and Cyclic PKS Intermediates through Atom Replacement

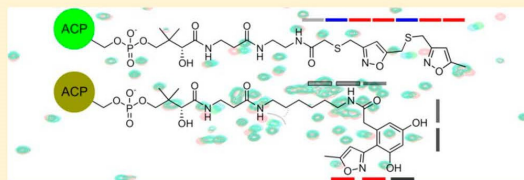
Gaurav Shakya,^{†,§} Heriberto Rivera, Jr.,^{‡,§} D. John Lee,[‡] Matt J. Jaremko,[‡] James J. La Clair,[‡] Daniel T. Fox,[‡] Robert W. Haushalter,[‡] Andrew J. Schaub,[†] Joel Bruegger,[†] Jesus F. Barajas,[†] Alexander R. White,[†] Parminder Kaur,[†] Emily R. Gwozdzowski,[‡] Fiona Wong,[†] Shiu-Chuan Tsai,^{*,†} and Michael D. Burkart^{*,‡}

[†]Departments of Molecular Biology and Biochemistry, Chemistry, and Pharmaceutical Sciences, University of California, Irvine, California 92697, United States

[‡]Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0358, United States

Supporting Information

ABSTRACT: The mechanistic details of many polyketide synthases (PKSs) remain elusive due to the instability of transient intermediates that are not accessible via conventional methods. Here we report an atom replacement strategy that enables the rapid preparation of polyketone surrogates by selective atom replacement, thereby providing key substrate mimetics for detailed mechanistic evaluations. Polyketone mimetics are positioned on the actinorhodin acyl carrier protein (actACP) to probe the underpinnings of substrate association upon nascent chain elongation and processivity. Protein NMR is used to visualize substrate interaction with the actACP, where a tetraketide substrate is shown not to bind within the protein, while heptaketide and octaketide substrates show strong association between helix II and IV. To examine the later cyclization stages, we extended this strategy to prepare stabilized cyclic intermediates and evaluate their binding by the actACP. Elongated monocyclic mimics show much longer residence time within actACP than shortened analogs. Taken together, these observations suggest ACP-substrate association occurs both before and after ketoreductase action upon the fully elongated polyketone, indicating a key role played by the ACP within PKS timing and processivity. These atom replacement mimetics offer new tools to study protein and substrate interactions and are applicable to a wide variety of PKSs.



Synthetic mimics of complex natural products are important tools to evaluate the mechanisms of natural product biosynthesis and to access syntheses that diversify the natural product templates. For example, the advent of the Stork–Eschenmoser hypothesis¹ was a defining moment in terpene biosynthesis that enabled synthetic mimicry^{2,3} to help elucidate the mechanisms of terpene biosynthesis.⁴ Comparable intermediates in aromatic polyketide biosynthesis have not been possible due to the instability of nascent polyketones, hindering advancement in the study of PKS mechanisms, timing, and processivity.⁵ Here we describe an approach for generating polyketide intermediates via atom replacement to mimic elongation intermediates in iterative (type II) polyketide synthases. We demonstrate application of these probes to evaluate the biosynthesis of actinorhodin,⁶ showing that these materials can mimic partially elongated and cyclized substrates attached to the actACP⁷ through analysis of intermediate association.

Figure 1 depicts the proposed biosynthesis of actinorhodin (1). Few intermediates within this pathway have been identified,⁵ leading to uncertainties in both the succession and process of elongation, reduction, and ring closure. Current

structural evidence suggests that the process is dictated by stabilizing residues within binding pockets of corresponding synthases,⁸ but this remains hypothetical. Important questions remain to be addressed, including: Why must chain elongation reach full length before reduction and cyclization? How is site-specific carbonyl reduction achieved? Are linear polyketones sequestered by the ACP during chain elongation? Access to elongated, polar mimics of ketide intermediates, such as 4–6 (Figure 1), could enlighten these uncertainties by their study in complex with PKS proteins for binding, stabilization, and chain transfer.

Although triketone units can be adapted synthetically, as demonstrated by Barrett^{9a–c} and Harris,^{9d,e} preparation and study of these and longer polyketones have been impeded by inherent reactivity through spontaneous intra- and intermolecular Aldol/Claisen condensations. We identified a practical solution to this problem by replacing select carbonyls with sulfur atoms, thereby thwarting spontaneous condensation and at the

Received: July 1, 2014

Published: November 19, 2014

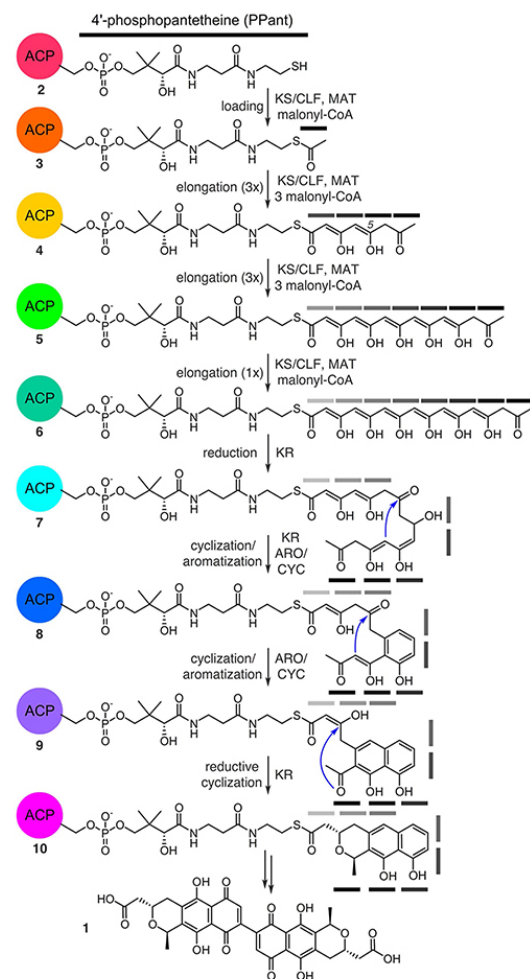


Figure 1. Proposed biosynthesis of **1** from *holo*-actACP (**2**). One depiction of the enol/ketol tautomerization states has been shown, in reality, multiple states likely exist. Gray bars denote ketide units.

same time providing probes to study the effects of length, polarity, and hydrophobicity upon the PKS elongation and cyclization process. We strategically substituted carbonyl groups nonessential for cyclization and/or reduction with sulfur atoms or isoxazole rings and validated their design via docking simulations of the probes upon actACP (Supporting Information). As shown in Figure 2, the corresponding thioethers would serve as a chemical “knockout” of the reactive carbonyl moieties. These probes differ from previously published analogs with the inclusion of both polyketide mimics and the full 4'-phosphopantetheine moiety, thus allowing us to evaluate the significance of ACP–probe interactions. In addition, this atom replacement design offers the added benefit of facile transformation to sulfoxides or sulfones, offering the ability to install greater polarity into the same parent compound.

We began by targeting an atom replacement mimetic of the ACP species bearing tetraketide unit **4** (Figure 1). Of the four carbonyl groups, the third ketide, as shown in **4a** (blue bar,

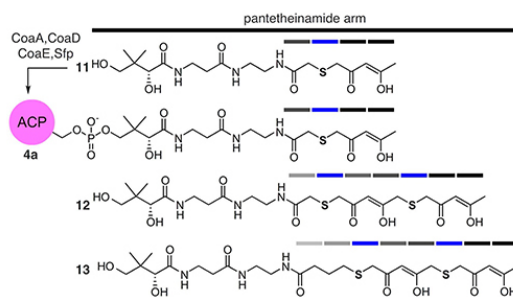
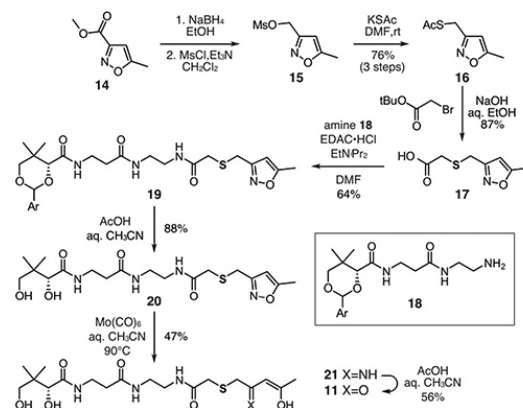


Figure 2. Structures of S atom replaced *crypto*-C8-actACP (**4a**). In this design, select carbonyl groups are replaced by S atoms, therein reducing number of potential Aldol/Claisen condensations. Access to **4a** was achieved by chemoenzymatic conversion of mimetic **11** to its corresponding CoA analog by the action of three *E. coli* enzymes CoaA, CoaD, CoaE followed by loading on the actACP by Sfp (see Supporting Information).^{10–12} Mimetics **12** and **13** and their corresponding *crypto*-loaded ACPs were not achieved due to instability. While the installation of the thioethers in **12** and **13** eliminated select Aldol/Claisen condensations, it did not ablate it. Ketide units are denoted by bars and are colored according to replacement of the carbonyl group with S atom (blue).

Figure 2), was selected for replacement. This involved the preparation of tetraketide mimetic **11** (Figure 2), which would be appended onto the actACP using chemoenzymatic methods developed in our laboratories.^{11,12} After exploring multiple approaches, we identified a route using an isoxazole as a tool to install a protected diketone unit.¹³ As shown in Scheme 1,

Scheme 1. Synthesis of Tetraketide Mimetic **11**^a



^aAr = 4-methoxybenzyl.

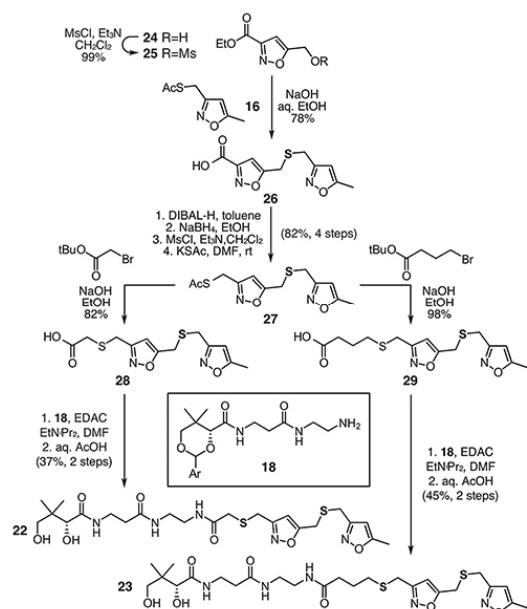
mimetic **11** was prepared through an eight-step route that began with commercially available isoxazole **14**. Reduction of **14** and subsequent mesylation afforded **15**, which was converted to thioacetate **16** over the three steps.

Thioacetate **16** was then converted into acid **17** by a three-step one-pot reaction sequence that began with deacetylation of the thioacetal, alkylation of the incipient thiolate, and ester hydrolysis (Scheme 1). Alternatively, addition of thioglycolic acid to mesylate **15** also provided acid **17** (not shown, see Supporting Information). Acid **17** was then coupled to

protected-pantetheinamine **18** using conventional EDAC coupling conditions to afford amide **19**. The synthesis of mimetic **11** was completed in three steps by removal of *p*-methoxyphenylacetal using aq. AcOH, opening of the isoxazole in **20** by treatment with Mo(CO)₆ in refluxing aq. CH₃CN, and conversion of **21** to **11** by treatment with AcOH:H₂O:CH₃CN (2:2:1). Using this route, we prepared **11** from isoxazole **14** in 10% overall yield over eight steps.

We then turned our attention to examination of mimetics bearing chain-elongated ketides as illustrated by heptaketide **12** and octaketide **13** (Figure 2). Our goal was to compare the short *crypto*-C8-actACP **4a** to the longer *crypto*-C14 or C16 actACPs as a means to validate the chain length specificity during extension of the ketide chain. We began by evaluating the synthesis of **12** and **13** (Figure 2). As depicted in Scheme 2, the

Scheme 2. Synthesis of Chain Elongated Mimetics, Heptaketide **22** and Octaketide **23**^a



^aAr = 4-methoxybenzyl.

synthesis of **12** began by converting isoxazole **24** to its mesylate **25** and coupling it to **16** using our tandem three-step coupling process, as illustrated for the elaboration of **16** to **17** (Scheme 1). The resulting bis-isoxazole-acid **26** was then converted into its corresponding thioacetate **27** in four steps. Again using the tandem three-step coupling process, we were able to access acid **28** in high yield from **27**. The resulting acid **28** was coupled to protected pantetheinamine **18**, which after deprotection, afforded heptaketide **22** in nine steps from **16** and **24** in 19% overall yield. Unfortunately, while multiple conditions and methods were screened, the conversion of **22** (Scheme 1) to **12** (Figure 2) via isoxazole ring opening underwent rapid condensation providing an intractable mixture.

Comparable complications arose in efforts to prepare mimetic **13** (Figure 2). Using thioacetate **27** as a central intermediate, preparation of **13** required only modification of the alkylating

agent. Conversion of **27** to **29** was readily achieved using *t*-butyl 4-bromobutanoate (Scheme 2). Coupling acid **29** to **18**, followed by deprotection, completed access to octaketide mimetic **23** in nine steps from **16** and **24** in 28% overall yield. Isoxazole opening in **23** again led to an intractable mixture of products. This indicated that stable mimetics required additional atom replacement.

Fortunately, the solution arose from **20** (Scheme 1), **22** (Scheme 2), and **23** (Scheme 2). Here, the isoxazole motif not only served as a synthetic tool but also provided a second type of atom replacement (Figure 3). In this dual atom replaced model,

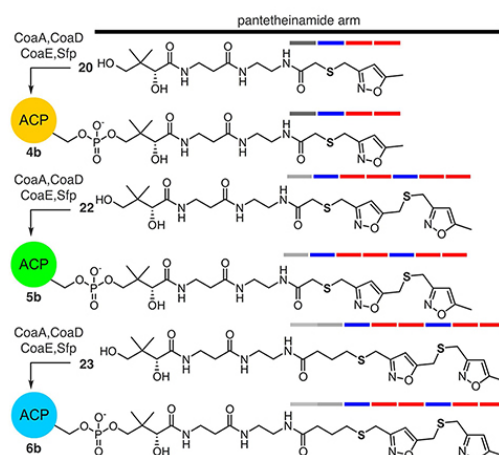


Figure 3. Structures of dual atom replaced *crypto*-C8-actACP **4b**, *crypto*-C14-actACP **5b**, and *crypto*-C16-actACP **6b**. Access to **4b**, **5b**, and **6b** was accomplished by chemoenzymatic attachment of **20** (Scheme 1), **22** (Scheme 2), and **23** (Scheme 2), respectively, to the actACP.^{10–12} Ketide units are denoted by bars and are colored according to replacement of a carbonyl with a sulfur atom (blue) or replacement of diketide unit with isoxazole (red).

both carbonyl and dicarbonyl units are replaced with thioether and isoxazole units, respectively. Unlike mimetics **12** and **13**, long chain analogs such as **22** and **23** are readily accessed and are stable, due to their inability to undergo intramolecular Aldol/Claisen condensations.

We applied these atom replacement mimetics to study actACP substrate sequestration as detected by solution-phase protein NMR.^{12,14} We appended mimetics **11**, **20**, **22**, and **23** to *apo*-actACP forming the corresponding *crypto*-C8-actACP **4a** (Figure 2), *crypto*-C8-actACP **4b** (Figure 3), *crypto*-C14-actACP **5b** (Figure 3), and *crypto*-C16-actACP **6b** (Figure 3).¹² We then examined samples of proteins **4a**, **4b**, **5b**, and **6b** by solution-phase ¹H,¹⁵N-HSQC NMR (Figure 4a,c).¹⁰ Comparison of *crypto*-C8-actACP **4a** and isoxazole-containing *crypto*-C8-actACP **4b** to *holo*-actACP via chemical shift perturbation (CSP) returned only very slight perturbations, indicative of a substrate that is mostly solvent exposed and not sequestered on the NMR time scale (Figure 4b).¹² These results compare favorably to known analog sequestration, where early intermediates (C6 and smaller) have been shown to not sequester.^{14c}

Octaketide **6b** shows ~30% larger CSPs over **5b** in these key residues, suggesting that the longer substrate, which mimics the fully mature octaketide polyketone in **6**, displays longer

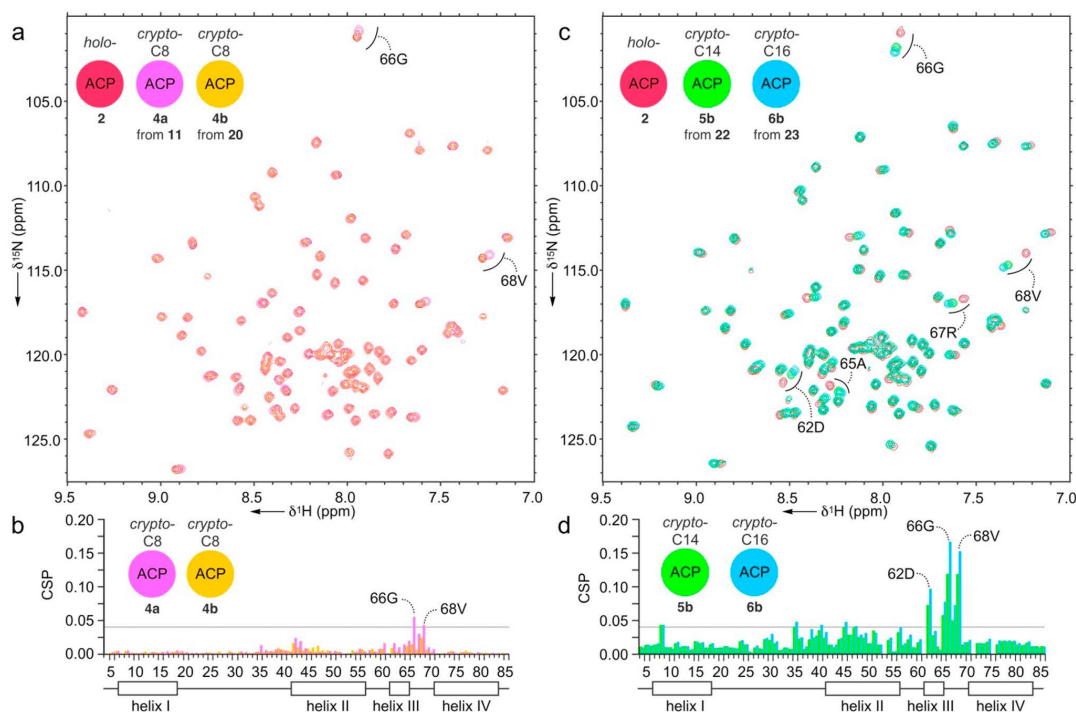


Figure 4. Sequestration analysis of atom replaced mimetics. (a) ^1H , ^{15}N -HSQC overlays depicting the *holo*-actACP (2, red) and the *crypto*-C8-actACP (4a, pink), *crypto*-C8-actACP (4b, orange) bearing 11 and 20, respectively. (b) CSP plots obtained from the spectra shown in panel (a). (c) ^1H , ^{15}N -HSQC traces depicting overlays of the *holo*-actACP (2, red), *crypto*-C14-actACP (5b, green) bearing heptaketide mimetic 22 and *crypto*-C16-actACP (6b, blue) bearing the octaketide mimetic 23. (d) CSP plots obtained from the spectra shown in panel (c). The docked actACP structure with these probes can be found in Figure S34.

residency within the hydrophobic cleft. This pocket is delineated by CSPs found in the same residues of actACP previously reported to stabilize aromatic product analogs.¹² Minimal CSPs for 4a and 4b indicate that tetraketide intermediates are not stabilized by interaction with the actACP, likely favoring further elongation by the actKS/CLF. This supports the hypothesis that sequestration of the elongating polyketone is driven by the comparative energetic stability between the ACP sequestered substrate and the ACP-KS/CLF complex. These studies suggest that the mature polyketone is stabilized through association by the hydrophobic cleft formed by movement of helix III of actACP during chain elongation.

Next we turned to evaluate the latter, cyclized intermediates, 7–10 (Figure 1), proposed in the biosynthesis of actinorhodin (1). We realized that this atom replacement methodology could be uniquely extended for the study of the cyclization events that occur in type II PKS pathways. These transformations have long been unclear, as the intractable and spontaneous cyclization of ketide intermediates (Figure 1) have prevented their preparation for detailed studies.^{8a,15} In particular, we were interested in mimetics that would allow us to probe ACP sequestration as a potential regulating force of substrate transport between elongation and initial cyclization, as illustrated by the conversion of 7 to 8 (Figure 1) by the actinorhodin ketoreductase (KR) as well as the processivity between KR and the aromatase/cyclase (ARO/CYC).^{15b}

A central mystery in type II PKS ACP arises during the processing of highly reactive polyketone intermediates to cyclized products, as exemplified by the conversion of 6 to 10 (Figure 1). This process includes steps of cyclization, reduction, and aromatization. We have hypothesized that intermediate stabilization through sequestration of elongated intermediates within the ACP may play a role during this transformation. To test this hypothesis, we recently demonstrated that actACP binds mature tricyclic product analogs between helices II and IV.¹² Crump and co-workers decisively showed that simple di- and triketide analogs are not sequestered by actACP,^{7a,16} but linear hydrocarbons do bind within the protein. Taken together, these findings indicate that identity of the intermediates likely plays a major role in substrate translocation between enzymes. As demonstrated above, the presence of polar residues such as D62 and D63 on helix III may stabilize the polar nature of polyketone intermediates and may be better suited to the environment of the actACP interior cavity. But what about the monocyclic product of the KR? Is this species directly routed to the ARO/CYC domain or does it return to the actACP pocket between enzymatic transformations? To address these questions, the preparation of elongated monocyclic probes is critical for further insight into the role of sequestration in type II PKS biosynthesis.

To this end, analogs 8a–8c (Figure 5) were proposed as probes that mimic the structural identity of putative intermediates as a means of evaluating actACP sequestration.

Analogue **8a** offered the most simplified mimetic by providing a single aromatic ring with six ketide units. Analogue **8b** and **8c** more closely resembled the putative intermediate **8** (Figure 1) with the addition of an upstream diketide, where the ketide units extended from six in **8a** and **8b** to eight in **8c**, respectively. Our first goal in this effort was the preparation of **8a–8c** for phosphopantetheine attachment to ^{15}N -enriched actACP for solution-phase protein NMR studies of substrate sequestration.

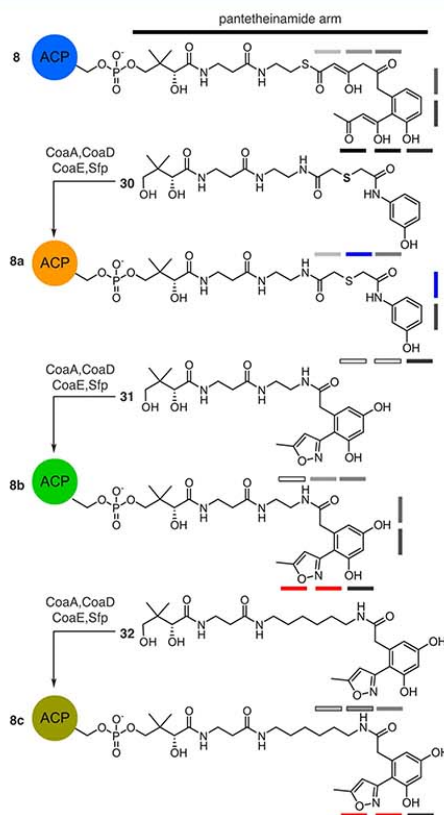
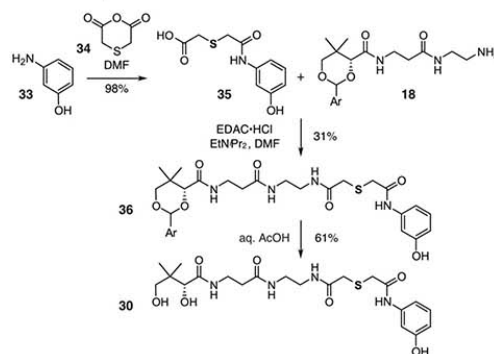


Figure 5. Structures of atom-replaced *crypt*-actACP **8a**, **8b**, and **8c**. Access to **8a**, **8b**, and **8c** was achieved by chemoenzymatic conversion of mimetics **30**, **31**, and **32**, respectively, to their corresponding CoA analogs by the action of three *E. coli* enzymes CoaA, CoaD, CoaE,¹⁰ followed by loading on the actACP by Sfp (see Supporting Information).¹¹ Mimetic **30** contains N and S atom replacement (blue bars) and bears structural truncation (black-outlined bars). In mimetics **31** and **32**, 1,3-dicarbonyl units are replaced by isoxazole (red bars). Two different length substrates were examined. Shaded gray bars denote ketide units.

We began with the preparation of **8a** by application of hexaketide mimic **30**. A general strategy for the synthesis of **30** involves the treatment of *m*-aminophenol (**33**) with thiodiglycolic anhydride (**34**) to provide aromatic acid **35** (Scheme 3). Acid **35** was then coupled to protected pantetheinamine **18** to afford **36**. Amide **36** was then deprotected by using aqueous acetic acid to deliver mimetic **30** in 19% overall yield from **33**. While **30** was readily accessed, it lacks two of the terminal ketides (black-outlined boxes, Figure 5). We next explored

Scheme 3. Synthesis of Mimetic **30**^a



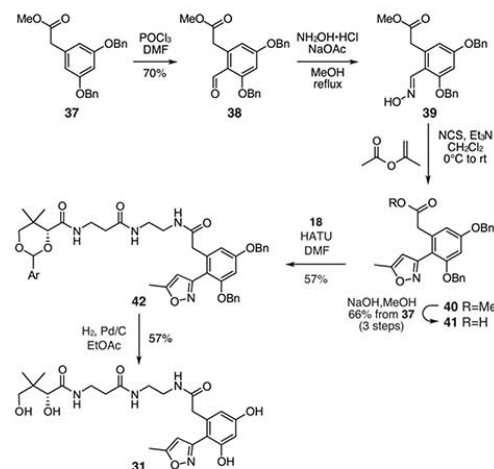
^aAr = 4-methoxyphenyl.

mimetics **31** and **32** that provide approximate isosteric placement of this additional diketide moiety. As depicted in Figure 5, these materials expanded on our use of an isoxazole as a diketide mimetic.

We then turned to the preparation of **8b** and **8c** from the respective mimetics **31** and **32** (Figure 5). The synthesis of mimetics **31** (Scheme 4) and **32** (Scheme 5) arose through preparation of acid **41** as a central intermediate. Application of a Vilsmeier–Haak formylation to **37**¹⁷ produced crystalline **38** in excellent yield. Aldehyde **38** was converted to its corresponding oxime **39**, which was then subjected to a 1,3-dipolar cycloaddition with isopropenyl acetate to afford isoxazole **40**. Preparation of acid **41** was then completed by hydrolysis under basic conditions.

Two additional steps were required to complete the synthesis of mimetic **31** (Scheme 4). Acid **41** was coupled to protected-pantetheinamine **18** to deliver amide **42**. Global deprotection under conventional hydrogenation conditions provided **31** in 15% overall yield from **37**. While mimetic **31** contained the missing diketide unit from **30** (Figure 5), its chain length did not

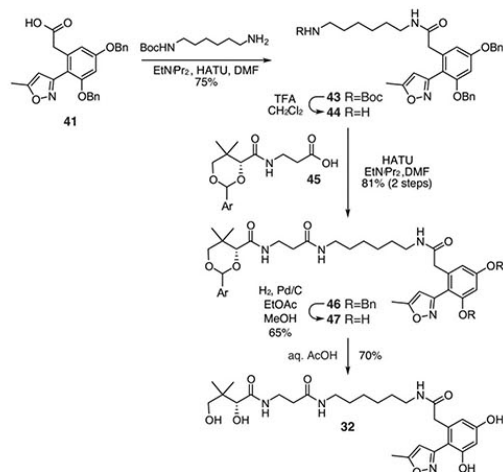
Scheme 4. Synthesis of Mimetic **31**^a



^aAr = 4-methoxyphenyl.

match that of the putative intermediate **8** (Figure 5). We therefore prepared a third mimetic **32** (Scheme 5) through the addition of a six-carbon spacer.

Scheme 5. Synthesis of Mimetic **32**^a



^aAr = 4-methoxyphenyl.

Advantageously, acid **41** could be used to prepare a chain-elongated mimetic **32** (Scheme 5). This began by preparing amide **43** by coupling *N*-Boc-1,6-hexanediamine to **41**. Boc-deprotection followed by coupling of amine **44** to protected pantoic acid **45** afforded amide **46**. A two-step sequence involving deprotection of the benzyl ethers in **46**, followed by removal of the PMP acetal in **47** under acidic conditions, provided mimetic **32** in 13% overall yield from **41**.

We then loaded mimetics **30–32** to ¹⁵N-enriched actACP chemoenzymatically as previously reported^{10–12} to generate *crypto*-actACP **8a**, **8b**, and **8c**, respectively. Next, we subjected *crypto*-actACP **8a–8c** to solution-phase protein NMR studies (Figure 6a). CSPs were observed (Figure 6b), which compare the *crypto*-actACP **8a–8c** to *holo*-actACP. We found that the *crypto*-actACP loaded with cyclized hexaketide mimetic **30** showed slight CSPs; the *crypto*-actACP loaded with cyclized mimetic **31** showed moderate CSPs; and the *crypto*-actACP loaded with cyclized mimetic **32** showed large CSPs in residues important to the interior sequestration cavity of actACP (Figure 6).¹²

On the NMR time scale, the fully elongated, cyclized octaketide **32** had much higher residence time in the actACP interior cavity than **30** or **31**. The strongest perturbations were observed in residues located at the end of helix III and the following loop, which corroborate with the sequestration residues that we previously reported in emodin-*crypto*-ACP.¹² Interestingly, only species **8c** showed moderate periodic perturbations throughout residues in helix II and those

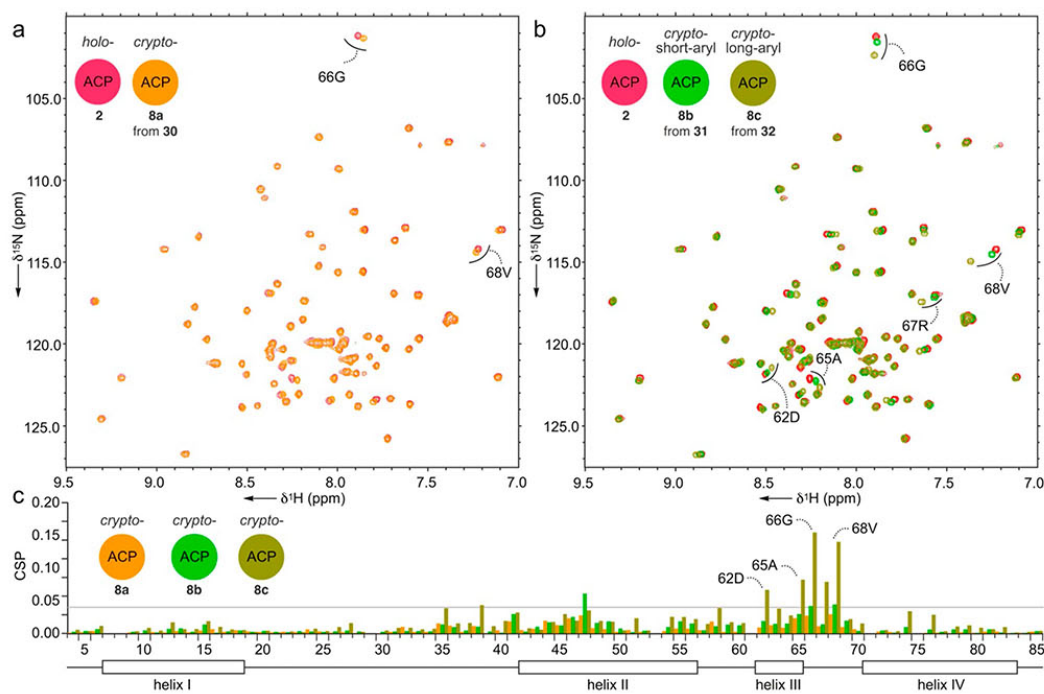


Figure 6. Sequestration analysis of atom replaced cyclic intermediates. (a) ¹H,¹⁵N-HSQC traces depicting overlays of the *holo*-actACP (**2**, red) and the *crypto*-actACP (**8a**, orange) bearing hexaketide mimetic **30** (Figure 5). (b) ¹H,¹⁵N-HSQC traces depicting overlays of the *holo*-actACP (**2**, red), *crypto*-actACP (**8b**, green) bearing hexaketide mimetic **31** and *crypto*-actACP (**8c**, brown) bearing octaketide mimetic **32**. (c) CSP plots obtained from the spectra shown in panels (a) and (b).

immediately preceding it. This information mirrored our study on linear polyketide mimetics (Figure 4), which showed the strongest sequestration of the fully elongated mimetic 32.

In this study, we have shown that selective replacement of carbonyl groups with heteroatoms facilitates access to a diverse class of polyketide mimetics that can be used to interrogate polyketide biosynthetic enzymes. Inherent to this methodology, these atom replacement mimetics can be chemoenzymatically conjugated to ACP through the corresponding CoA analogs, which provide the native scaffold to study these processes. While no probe could ever exactly mimic the properties of natural polyketone intermediates, this work demonstrates that molecules that are much more polar than fatty acids^{14c} or tricyclic aromatics¹² do associate with actACP, that longer chains experience more residence time, and that cyclic intermediates demonstrate the longest residence times. Our data, combined with docking simulation of these probes (Figure S34), strongly support that binding of both linear and cyclized polyketide mimetics is not observed unless the mimetics are of sufficient length. Taken together, these observations suggest that until the polyketide has reached its terminal length, no preferential sequestration by the actACP stabilizes the intermediate. However, at full elongation, polyketide binding with the actACP could facilitate release from the ketosynthase and assist transfer to the KR for the first cyclization and reduction steps.^{15b} Binding of the first cyclized, reduced intermediate by the actACP would then occur to facilitate release from the KR, followed by delivery to the ARO/CYC.

In this study, the goal of these atom replacement mimetics was to help understand the comparative binding and stabilization of polyketone intermediates by ACPs. These mimetics were not designed to serve as functional substrates of PKS enzymes but instead to mimic the length, polarity, and hydrophobicity found in natural intermediates. As with all mimetics, the match was not perfect. In these examples, the C–S bonds are considerably longer than those of their anticipated ketides. The actual enol-keto states of natural polyketones remain unknown, and the isoxazole moiety in part restricts conformational freedom. These caveats aside, the establishment of polyketide mimetics provides an excellent tool for interrogating the iterative processes polyketide biosynthesis.

Finally, the involvement of the ACP in these pathways adds significant complexity when compared to non-templated biosynthetic pathways, and the creation of new tools to understand this role merits special attention. Most critically, this study defines new methods of atom replacement that can be used to rapidly assemble both linear and cyclic polyketide mimetics for structural and functional applications. Studies on the protein–protein interactions of these species are ongoing.

■ ASSOCIATED CONTENT

Supporting Information

Copies of spectral data, synthetic methods, and experimental procedures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Authors

sctsai@uci.edu
mburkart@ucsd.edu

Author Contributions

[§]G.S. and H.R. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. X. Huang (UCSD), Dr. A. Mrse (UCSD), Dr. Y. Su (UCSD), and Dr. P. Dennison (ICI) for assistance with collection of NMR and MS data. We also thank Prof. C. A. Townsend (Johns Hopkins), Prof. F. Ishikawa (UCSD), Prof. C. D. Vanderwal (UCI), and Prof. A. R. Chamberlin (UCI) for helpful advice. We also thank Prof. M. P. Crump for your advices from full structural determinations as well as helpful discussions. Funding was provided by National Institutes of Health GM100305 and GM095970.

■ REFERENCES

- (1) (a) Sadler, P. A.; Eschenmoser, A.; Scheinz, H.; Stork, G. *Helv. Chem. Acta* **1957**, *40*, 2191. (b) Stork, G.; Burgstahler, A. W. *J. Am. Chem. Soc.* **1955**, *77*, 5068. (c) Ruzicka, L.; Eschenmoser, A.; Heusser, H. *Experientia* **1953**, *9*, 357. (d) Ruzicka, L. *Experientia* **1954**, *50*, 1.
- (2) Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189.
- (3) (a) Johnson, W. S.; Semmelhack, M. F.; Sultanbawa, M. U.; Dolak, L. A. *J. Am. Chem. Soc.* **1968**, *90*, 2994. (b) Schmidt, R.; Huesmann, P. L.; Johnson, W. S. *J. Am. Chem. Soc.* **1980**, *102*, 5122.
- (4) (a) Pichersky, E.; Noel, J. P.; Dudareva, N. *Science* **2006**, *311*, 808. (b) Yoder, R. A.; Johnston, J. N. *Chem. Rev.* **2005**, *105*, 4730. (c) Poralla, K. *Chem. Biol.* **2004**, *11*, 12. (d) Xu, R.; Fazio, G. C.; Matsuda, S. P. *Phytochemistry* **2004**, *65*, 261.
- (5) (a) Akey, D. L.; Gehret, J. J.; Khare, D.; Smith, J. L. *Nat. Prod. Rep.* **2012**, *29*, 1038. (b) Crawford, J. M.; Townsend, C. A. *Nat. Rev. Microbiol.* **2010**, *8*, 870. (c) Khosla, C. *J. Org. Chem.* **2009**, *74*, 6416. (d) Das, A.; Khosla, C. *Acc. Chem. Res.* **2009**, *42*, 631. (e) Bumpus, S. B.; Kelleher, N. L. *Curr. Opin. Chem. Biol.* **2008**, *12*, 475.
- (6) (a) Korman, T. P.; Hill, J. A.; Vu, T. N.; Tsai, S. C. *Biochemistry* **2004**, *43*, 14529. (b) Rudd, B. A.; Hopwood, D. A. *J. Gen. Microbiol.* **1979**, *114*, 35. (c) Fernández-Moreno, M. A.; Martínez, E.; Caballero, J. L.; Ichinose, K.; Hopwood, D. A.; Malpartida, F. *J. Biol. Chem.* **1994**, *269*, 24854.
- (7) (a) Crump, M. P.; Crosby, J.; Dempsey, C. E.; Parkinson, J. A.; Murray, M.; Hopwood, D. A.; Simpson, T. *J. Biochemistry* **1997**, *36*, 6000. (b) Revill, W. P.; Bibb, M. J.; Hopwood, D. A. *J. Bacteriol.* **1996**, *178*, 5660.
- (8) (a) Teufel, R.; Miyana, A.; Michaudel, Q.; Stull, F.; Louie, G.; Noel, J. P.; Baran, P. S.; Palfey, B.; Moore, B. S. *Nature* **2013**, *503*, 552. (b) Ames, B. D.; Lee, M. Y.; Moody, C.; Zhang, W.; Tang, Y.; Tsai, S. C. *Biochemistry* **2011**, *50*, 8392. (b1) Tsai, S. C.; Lu, H.; Cane, D. E.; Khosla, C.; Stroud, R. M. *Biochemistry* **2002**, *41*, 12598. (c) Pan, H.; Tsai, S. C.; Meadows, E. S.; Miercke, L. J.; Keatinge-Clay, A. T.; O'Connell, J.; Khosla, C.; Stroud, R. M. *Structure* **2002**, *10*, 1559.
- (9) (a) Fouché, M.; Rooney, L.; Barrett, A. G. M. *J. Org. Chem.* **2012**, *77*, 3060. (b) Calo, F.; Richardson, J.; Barrett, A. G. M. *Org. Lett.* **2009**, *11*, 4910. (c) Navarro, L.; Basset, J. F.; Hebbe, S.; Major, S. M.; Werner, T.; Howsham, C.; Bräckow, J.; Barrett, A. G. M. *J. Am. Chem. Soc.* **2008**, *130*, 10293. (d) Harris, T. M.; Murray, T. P.; Harris, C. M.; Gumulka, M. *J. Chem. Soc. Chem. Comm.* **1974**, *10*, 362. (e) Harris, T. M.; Harris, C. M. *Tetrahedron* **1977**, *33*, 2159.
- (10) (a) Worthington, A. S.; Burkart, M. D. *Org. Biomol. Chem.* **2006**, *4*, 44. (b) Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. *Nat. Methods* **2012**, *9*, 981.
- (11) (a) La Clair, J. J.; Foley, T. L.; Schegg, T. R.; Regan, C. M.; Burkart, M. D. *Chem. Biol.* **2004**, *11*, 195. (b) Meier, J. L.; Burkart, M. D. *Methods Enzymol.* **2009**, *458*, 219. (c) Foley, T. L.; Burkart, M. D. *Curr. Opin. Chem. Biol.* **2007**, *11*, 12.
- (12) Haushalter, R. W.; Filipp, F. V.; Ko, K. S.; Yu, R.; Opella, S. J.; Burkart, M. D. *ACS Chem. Biol.* **2011**, *6*, 413.
- (13) (a) Stork, G.; Hagedorn, A. A., III. *J. Am. Chem. Soc.* **1978**, *100*, 3609. (b) Stork, G.; La Clair, J. J.; Spargo, P.; Nargund, R. P.; Totah, N. *J. Am. Chem. Soc.* **1996**, *118*, 5304. (c) Wright, P. M.; Myers, A. G.

Tetrahedron **2011**, *67*, 9853. (d) Sun, C.; Wang, Q.; Brubaker, J. D.; Wright, P. M.; Lerner, C. D.; Noson, K.; Charest, M.; Siegel, D. R.; Wang, Y. M.; Myers, A. G. *J. Am. Chem. Soc.* **2008**, *130*, 17913. (e) Charest, M. G.; Siegel, D. R.; Myers, A. G. *J. Am. Chem. Soc.* **2005**, *127*, 8292.

(14) (a) Roujeinikova, A.; Simon, W. J.; Gilroy, J.; Rice, D. W.; Rafferty, J. B.; Slabas, A. R. *J. Mol. Biol.* **2007**, *365*, 135. (b) Ploskoń, E.; Arthur, C. J.; Kanari, A. L.; Wattana-amorn, P.; Williams, C.; Crosby, J.; Simpson, T. J.; Willis, C. L.; Crump, M. P. *Chem. Biol.* **2010**, *17*, 776. (c) Evans, S. E.; Williams, C.; Arthur, C. J.; Ploskoń, E.; Wattana-amorn, P.; Cox, R. J.; Crosby, J.; Willis, C. L.; Simpson, T. J.; Crump, M. P. *J. Mol. Biol.* **2009**, *389*, 511.

(15) (a) Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Markwick, P. R.; Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, F.; O'Dowd, B.; McCammon, J. A.; Opella, S. J.; Tsai, S. C.; Burkart, M. D. *Nature* **2014**, *505*, 427. (b) Javidpour, P.; Bruegger, J.; Srihahan, S.; Korman, T. P.; Crump, M. P.; Crosby, J.; Burkart, M. D.; Tsai, S. C. *Chem. Biol.* **2013**, *20*, 1225.

(16) Reville, W. P.; Bibb, M. J.; Hopwood, D. A. *J. Bacteriol.* **1996**, *178*, 5660.

(17) Marsini, M. A.; Pettus, T. R. R.; Wenderski, T. A. *Org. Lett.* **2011**, *13*, 118.

SUPPORTING INFORMATION

Modeling linear and cyclic PKS intermediates through atom replacement

Gaurav Shakya,[†] Heriberto Rivera, Jr.,[‡] D. John Lee,[‡] Matt J. Jaremko,[‡] James J. La Clair,[‡] Daniel T. Fox,[‡] Robert W. Haushalter,[‡] Andrew J. Schaub,[†] Joel Bruegger,[†] Jesus F. Barajas,[†] Alexander R. White,[†] Parminder Kaur,[†] Emily R. Gwozdzowski,[‡] Fiona Wong,[†] Shiou-Chuan Tsai,^{†*} Michael D. Burkart^{‡*}

[†]Departments of Molecular Biology and Biochemistry, Chemistry, and Pharmaceutical Sciences, University of California, Irvine, CA 92697

and

[‡]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 or sctsai@uci.edu.

Contents	Page
A. General materials and methods	S4
B. Chemical synthesis	S5-S18
B.1. Synthesis of thioacetate 16	S5
B.2. Synthesis of mimetic 11	S6-S7
B.3. Synthesis of <i>bis</i> -isoxazole acid 26	S8-S9
B.4. Synthesis of thioester 27	S9-S10
B.5. Synthesis of mimetic 22	S10-S11
B.6. Synthesis of mimetic 23	S12-S13
B.7. Synthesis of mimetic 30	S13-S14
B.8. Synthesis of mimetic 31	S14-S16
B.9. Synthesis of mimetic 32	S17-S18
C. Protein NMR studies	S19-S20
C.1. Expression, modification, and purification of <i>Pseudomonas fluorescens</i> AcpH	S19
C.2. Preparation of PfAcpH affinity resin	S19
C.3. Expression, modification, and purification of <i>apo</i> -actACP	S19
C.4. Preparation of <i>crypto</i> -actACPs 4a , 4b , 5b , 6b , 8a , 8b , and 8c .	S19
C.5. Preparation of <i>holo</i> -actACP	S20
C.6. HPLC analysis of ACP proteins	S20
C.7. Protein NMR experiments	S20
C.8. <i>In silico</i> docking experiments	S20
Supplemental Figures	Page
Figure S1. ¹ H NMR (400 MHz) and ¹³ C NMR (100 MHz) spectra of S1 in CDCl ₃	S21
Figure S2. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 15 in CDCl ₃	S22
Figure S3. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 16 in CDCl ₃	S23
Figure S4. ¹ H NMR (400 MHz) and ¹³ C NMR (100 MHz) spectra of 17 in CDCl ₃	S24
Figure S5. ¹ H NMR (500 MHz) and ¹³ C NMR (100 MHz) spectra of 19 in CDCl ₃	S25
Figure S6. ¹ H NMR (500 MHz) and ¹³ C NMR (100 MHz) spectra of 20 in CD ₃ OD	S26
Figure S7. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 21 in D ₂ O	S27
Figure S8. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 11 in CD ₃ OD	S28
Figure S9. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 24 in CDCl ₃	S29
Figure S10. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 25 in CDCl ₃	S30
Figure S11. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 26 in CD ₃ OD	S31
Figure S12. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of S3 in CDCl ₃	S32
Figure S13. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of S4 in CDCl ₃	S33
Figure S14. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 27 in CDCl ₃	S34
Figure S15. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 28 in CDCl ₃	S35
Figure S16. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of S5 in CD ₃ OD	S36
Figure S17. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 22 in CD ₃ OD	S37
Figure S18. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of 29 in CDCl ₃	S38
Figure S19. ¹ H NMR (500 MHz) and ¹³ C NMR (125 MHz) spectra of S6 in CD ₃ OD	S39

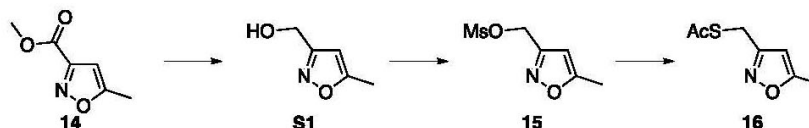
Figure S20. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 23 in CD_3OD	S40
Figure S21. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 35 in $\text{DMSO-}d_6$	S41
Figure S22. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 36 in CD_3OD	S42
Figure S23. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 30 in CD_3OD	S43
Figure S24. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 38 in CDCl_3	S44
Figure S25. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 39 in CDCl_3	S45
Figure S26. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 40 in CDCl_3	S46
Figure S27. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 41 in $\text{DMSO-}d_6$	S47
Figure S28. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 42 in $\text{acetone-}d_6$	S48
Figure S29. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 31 in CD_3OD	S49
Figure S30. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 43 in CDCl_3	S50
Figure S31. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 46 in $\text{acetone-}d_6$	S51
Figure S32. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 47 in $\text{acetone-}d_6$	S52
Figure S33. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 32 in CD_3OD	S53
Figure S34. Docking simulations of ActACP with probes.	S54
Table S1. LCMS validation of <i>crypto</i> -ACPs via PPant ejection	S55

A. General Materials and Methods.

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 **18** 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), a ECA 500 MHz (Jeol), a DMX 400 MHz (Bruker), a DMX 500 MHz (Bruker) 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. UV spectra were measured with a 1 cm cell on a DU800 (Bruker) spectrophotometer. IR spectra were obtained with a IR100 FT-IR (ThermoNicolet). 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 500 MHz VS500 spectrometer equipped with an Xsens cryoprobe (Varian) or 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.

B. Chemical Synthesis

B.1. Synthesis of thioacetate 16. A three-step procedure was used to prepare thioester **16** from commercially available **14**.



S-((5-methylisoxazol-3-yl)methyl) ethanethioate (16). The synthesis of thioester **16** was most effectively conducted in a single day through three back-to-back steps from commercially available methyl 5-methylisoxazolecarboxylate (**14**). Ester **14** (25.00 g, 0.177 mol) was dissolved in absolute EtOH (250 mL). After cooling to 0 °C, NaBH₄ (18.09 g, 0.478 mol) was added in ~0.5 g portions so that the temperature of the reaction did not exceed 5 °C. After the addition, the mixture was kept at 0 °C for 1 h and slowly warmed over 1.5 h to rt. After 1 h at rt, the flask was cooled to 0 °C with an ice bath and satd. NH₄Cl (200 mL) was added in a drop wise fashion. Once quenched, the mixture was warmed to rt, extracted with EtOAc (3×500 mL), washed with brine, dried with Na₂SO₄, and filtered through a plug of Celite (100 g) washing with EtOAc (1 L) to deliver 19.28 g (96%) of carbinol **S1**.

Carbinol **S1**: ¹H NMR (CDCl₃, 400 MHz) δ 6.02 (s, 1H), 4.68 (s, 2H), 2.92 (bs, 1H), 2.40 (s, 3H); ¹³C NMR (CDCl₃, 100 MHz) δ 170.0, 163.9, 100.6, 57.0, 12.4; HRMS *m/z* calcd. for C₅H₉NO₂ [M+H]⁺ 114.0555 found, 114.0551.

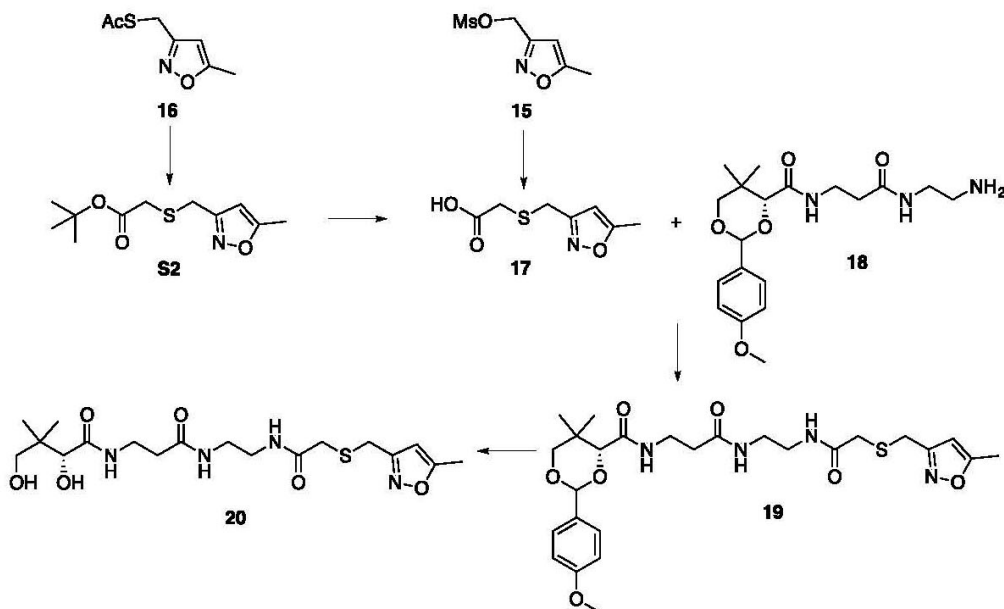
Carbinol **S1** (19.28 g, 0.170 mol) was dried by rotary evaporation of toluene (3×100 mL) and dissolved in dry CH₂Cl₂ (1 L). Freshly distilled Et₃N (31.50 mL, 0.226 mol) was added and the flask was cooled to 0 °C. MsCl (14.00 mL, 0.181 mol) was added drop wise over 20 min. The mixture was kept at 0 °C for 30 min then warmed over 1 h to rt. After 30 min at rt, brine (400 mL) was added. The resulting mixture was extracted with CH₂Cl₂ (3×500 mL), washed with brine, dried with Na₂SO₄, and filtered through a plug of SiO₂ (300 g) washing with a 1:1 CH₂Cl₂:EtOAc (800 mL) to deliver the crude mesylate **15**.

Mesylate **15**: ¹H NMR (CDCl₃, 500 MHz) δ 6.12 (s, 1H), 5.20 (s, 2H), 3.02 (s, 3H), 2.40 (s, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 171.0, 158.2, 101.3, 62.1, 38.1, 12.3; HRMS *m/z* calcd. for C₆H₁₀NO₄S [M+H]⁺ 192.0331 found, 192.0327.

Immediately after, the entire batch of crude mesylate **15** was dissolved in dry DMF (150 mL). KSAc (20.60 g, 0.180 mol) was added over 5 min as a solid. Within 5 min of the addition, the reaction mixture fused as a solid and became warm (~65 °C). This mixture was allowed to sit for 1 h at which point additional dry DMF (50 mL) was added. After 1.5 h at rt, satd. NaHCO₃ (500 mL) was added and the mixture was extracted CH₂Cl₂ (3×500 mL), washed with brine, dried with Na₂SO₄, and filtered through a plug of SiO₂ (300 g) washing with 1:1 CH₂Cl₂:EtOAc (1 L) to deliver the 24.20 g of thioester **16**, which could be used without further purification as the only impurity detectable was ~1% of mesylate **15**. If desired, this material could be purified via flash chromatography to afford 23.10 g of **16** (76% from **14**).

Thioester **16**: ¹H NMR (CDCl₃, 500 MHz) δ 5.92 (s, 1H), 4.08 (s, 2H), 2.38 (s, 6H); ¹³C NMR (CDCl₃, 125 MHz) δ 194.6, 170.2, 161.2, 101.7, 30.4, 23.9, 12.4; HRMS *m/z* calcd. for C₇H₉NO₂SNa [M+Na]⁺ 194.0245 found, 194.0246.

B.2. Synthesis of mimetic 11. The synthesis of mimetic 11 was accomplished either from mesylate 15 or thioacetate 16. The following section provides an overview of the methods and procedures used.



2-(((5-methylisoxazol-3-yl)methyl)thio)acetic acid (17). Thioester 16 (381.2 mg, 2.23 mol) was dissolved in absolute EtOH (22 mL). The solution was degassed by repeated pump and fill with Ar. A degassed solution of NaOH (1.11 mL, 5 N in H₂O, 5.56 mmol) under Ar was added via syringe to this solution at rt. Within 10 min of addition, *tert*-butylbromoacetate (0.35 mL, 2.37 mmol) was added and the mixture was kept at rt for 1 h. The pH was adjusted to 7 by the addition of acetic acid, the mixture was diluted with EtOAc (50 mL) and dried via rotary evaporation. Flash chromatography (hexanes to 2:1 hexanes:EtOAc) afforded 362.0 mg (87%) of acid 17, as a colorless oil. An additional 32.5 mg (6%) of ester S2 was also obtained.

Alternatively, acid 17 was prepared from mesylate 15. Mercaptoacetic acid (1.95 mL, 27.94 mmol) was added to mesylate 15 (4.88 g, 25.52 mmol) in DMF (60 mL). This was followed by sequential addition of NaI (3.83 g, 25.55 mmol) and Et₃N (7.85 mL, 56.27 mmol). The reaction mixture was warmed to 60 °C. After 12 h at 60 °C, the solvent was removed by rotary evaporation. The residue was dissolved in a mixture of EtOAc (100 mL) and H₂O (100 mL). The layers were separated and the water layer was further extracted with EtOAc (3×100 mL). Organic layers were combined and dried over Na₂SO₄. The solvent was removed via rotary evaporation to afford pale yellow solid. Flash chromatography (EtOAc to 9:1 EtOAc:MeOH) afforded 4.23 g (89%) of acid 17.

Acid 17: ¹H NMR (CDCl₃, 500 MHz) δ 6.03 (s, 1H), 3.83 (s, 2H), 3.23 (s, 2H), 2.41 (s, 3H); ¹³C NMR (CDCl₃, 100 MHz) δ 174.8, 170.4, 161.0, 101.6, 32.6, 26.6, 12.5; HR-ESI-TOFMS *m/z* calcd. for C₇H₉NO₃SNa [M+Na]⁺: 210.0201, found 210.0196.

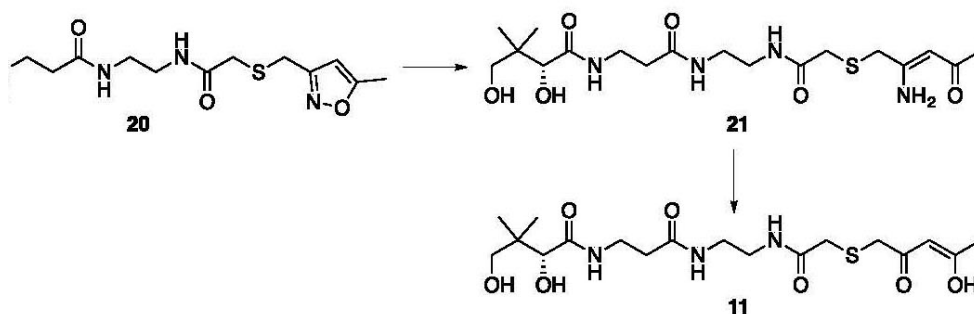
(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-*N*-(3-(((2-2-(((5-methylisoxazol-3-yl)methyl)thio)-acetamido)ethyl)amino)-3-oxopropyl)-1,3-dioxane-4-carboxamide (19): EDAC·HCl (768.00 mg, 4.01 mmol) was added to a mixture of acid 17 (300.00 mg, 1.60 mmol) and amine 18 (608.00 mg, 1.60 mmol) dissolved in dry DMF (10 mL). Within 5 min, Et₃N (1.40 mL, 8.04 mmol) was added and the mixture was stirred at rt. After 12 h at rt, the solvent was removed by

e 19, as colorless oil.

NMR (CDCl₃, 500 MHz) δ 7.42 (d, J = 8.7 Hz, 2H), 7.06 (t, J = 6.1 Hz, 2H), 6.90 (t, 2H), 5.99 (d, J = 1.1 Hz, 1H), 5.45 (s, 1H), 4.07 (s, 1H), 3.81 (s, 3H), 3.70 (d, J = 3.66 (s, 2H), 3.63 (d, J = 11.3 Hz, 1H), 3.55 (p, J = 6.2 Hz, 2H), 3.34 (m, 4H), 3.07 (t, J = 6.2 Hz, 2H), 2.40 (s, 3H), 1.10 (s, 3H), 1.07 (s, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 170.6, 169.6, 169.1, 161.0, 160.3, 130.3, 127.7, 113.8, 101.6, 101.5, 84.0, 78.6, 71.8, 36.2, 35.1, 35.1, 33.2, 26.4, 22.0, 19.3, 12.5; HR-ESI-TOFMS m/z calcd. for **1a** [M+Na]⁺: 571.2202, found 571.2202.

roxy-3,3-dimethyl-N-(3-((2-(2-((5-methylisoxazol-3-yl)methyl)thio)ethyl)amino)-3-oxopropyl)butanamide (20): Amide **19** (567.00 mg, 1.03 mmol) in 80% aq. AcOH (10 mL). After 16 h at rt, the solvent was removed via rotary evaporation assisted by azeotropic removal of toluene. Flash chromatography (EtOAc to 4:1 EtOAc:MeOH) afforded 390.0 mg (88%) of mimetic **20**, as a colorless oil.

¹H NMR (CD₃OD, 500 MHz) δ 6.17 (d, J = 1.2 Hz, 1H), 3.89 (s, 1H), 3.76 (s, 2H), 3.46 (d, J = 11.1 Hz, 1H), 3.38 (d, J = 11.0 Hz, 1H), 3.35 (s, 2H), 3.30 (m, 4H), 2.43 (t, J = 6.7 Hz, 2H), 2.41 (d, J = 1.1 Hz, 3H), 0.91 (s, 6H); ¹³C NMR (CD₃OD, 125 MHz) δ 176.0, 174.2, 171.9, 171.7, 162.9, 102.7, 77.3, 70.3, 49.8, 40.4, 39.9, 36.6, 36.4, 20.9, 12.0; HRMS m/z calcd. for C₁₈H₃₀N₄O₆SNa [M+Na]⁺: 453.1784, found 453.1784.



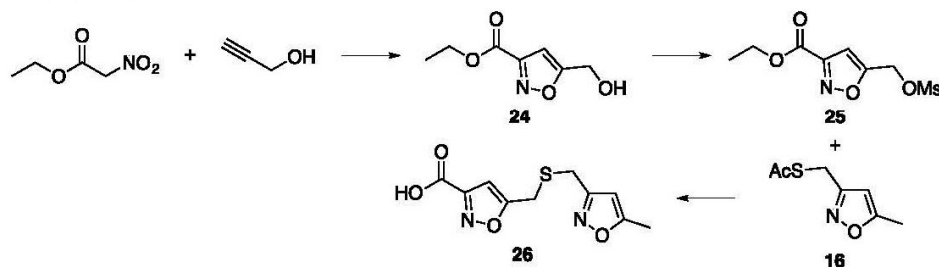
-(2-((2-amino-4-oxopent-2-en-1-yl)thio)acetamido)ethyl)amino)-3-oxopropyl)- γ -3,3-dimethylbutanamide (21): Fresh Mo(CO)₆ (84.1 mg, 0.32 mmol) was dissolved in a 25% aq. MeCN (20 mL) at rt. The reaction was warmed to reflux. After 3 h, the reaction was cooled and the solvent removed via rotary evaporation. Flash chromatography (EtOAc to 3:1 EtOAc:MeOH) afforded 64.3 mg (47%) of pale-yellow oil.

21: ¹H NMR (D₂O, 500 MHz) δ 3.96 (s, 1H), 3.52 (m, 2H), 3.49 (d, J = 11.5 Hz, J = 11.3 Hz, 1H), 3.33 (s, 2H), 3.30 (s, 2H), 3.29 (s, 2H), 3.28 (m, 4H), 2.47 (t, J = 6.03 (s, 3H), 0.90 (s, 3H), 0.87 (s, 3H); ¹³C NMR (D₂O, 125 MHz) δ 201.8, 177.8, 165.7, 78.4, 71.0, 51.5, 41.8, 41.3, 41.2, 38.6, 38.1, 38.0, 37.8, 30.9, 23.2, 21.8; HRMS m/z calcd. for C₁₈H₃₂N₄O₆SNa [M+Na]⁺: 455.1935, found 455.1936.

hydroxy-N-(3-((2-(2-((4-hydroxy-2-oxopent-3-en-1-yl)thio)acetamido)ethyl)amino)-3-oxopropyl)-3,3-dimethylbutanamide (11): Intermediate **21** (10.7 mg, 0.025 mmol) in 2:2:1 AcOH:H₂O:CH₃CN (1 mL) and stirred at rt. After 3.5 h, the solvent was removed via rotary evaporation. Flash chromatography (1:5 MeOH:EtOAc to 3:5 MeOH:EtOAc) afforded 6.0 mg (56%) of mimetic **11**, as pale-yellow oil.

Mimetic **21**: ^1H NMR (CD_3OD , 500 MHz) δ 4.04 (s, 1H), 3.59 (m, 2H), 3.57 (d, $J = 11.3$ Hz, 1H), 3.45 (d, $J = 11.3$ Hz, 1H), 3.40 (m, 4H), 3.37 (s, 2H), 3.36 (s, 2H), 3.31 (s, 2H, minor conformer), 3.30 (s, 2H, minor conformer), 2.55 (m, 2H), 2.35 (s, 1H), 2.11 (s, 3H, minor conformer), 2.08 (s, 3H, minor conformer), 2.04 (s, 3H), 0.98 (s, 3H), 0.94 (s, 3H); ^{13}C NMR (CDCl_3 , 125 MHz) δ 177.5, 176.1, 174.2, 172.7, 77.3, 70.3, 40.3, 40.0, 38.9, 37.3, 36.6, 36.4, 21.4, 21.0; HR-ESI-TOF-MS m/z calcd. for $\text{C}_{18}\text{H}_{31}\text{N}_3\text{O}_7\text{SNa}$ $[\text{M}+\text{Na}]^+$: 456.1780 found 456.1775.

B.3. Synthesis of bis-isoxazole acid 26. A three-step procedure was developed to access gram quantities of compound **26**. This method was particularly viable for scale due to the fact that acid **26** could be obtained in pure form through acid-base extraction and did not require chromatographic purification.



Ethyl 5-(hydroxymethyl)isoxazole-3-carboxylate (24). Ethylnitroacetate (15.90 mL, 143.23 mmol) and propargyl alcohol (4.35 mL, 75.42 mmol) were dissolved in absolute EtOH (200 mL). DABCO (0.846 g, 7.54 mmol) was added and the mixture was warmed to 80 °C over 30 min. After 72 h at 80 °C, the mixture was dried via rotary evaporation. Flash chromatography (1:1 hexanes: CH_2Cl_2 to CH_2Cl_2 to 10:1 CH_2Cl_2 :MeOH) afforded 12.5 g (97%) of carbinol **24**.

Carbinol **24**: ^1H NMR (CDCl_3 , 500 MHz) δ 6.60 (s, 1H), 4.75 (s, 2H), 4.37 (q, $J = 7.1$ Hz, 2H), 4.00 (bs, 1H), 1.35 (t, $J = 7.2$ Hz, 3H); ^{13}C NMR (CDCl_3 , 125 MHz) δ 173.9, 159.9, 156.2, 102.5, 62.4, 56.1, 14.1; HRMS m/z calcd. for $\text{C}_7\text{H}_{10}\text{NO}_4$ $[\text{M}+\text{H}]^+$: 172.0610, found. 171.0608.

Ethyl 5-(((methylsulfonyl)oxy)methyl)isoxazole-3-carboxylate (25). Carbinol **24** (6.41 g, 37.45 mmol) was dried by rotary evaporation of toluene (3×25 mL) and then dissolved in dry CH_2Cl_2 (320 mL). Freshly distilled Et_3N (7.8 mL, 55.91 mol) was added and the flask was cooled to 0 °C. MsCl (3.48 mL, 44.96 mol) was added drop wise over 20 min. The mixture was kept at 0 °C for 30 min then warmed over 1 h to rt. After 30 min at rt, brine (50 mL) was added. The resulting mixture was extracted with CH_2Cl_2 (3×150 mL), washed with brine, dried with Na_2SO_4 , and filtered through a plug of SiO_2 (300 g) washing with 1:1 CH_2Cl_2 :EtOAc (500 mL) to deliver the crude 9.31 g (99%) of mesylate **25**, which was used without further purification.

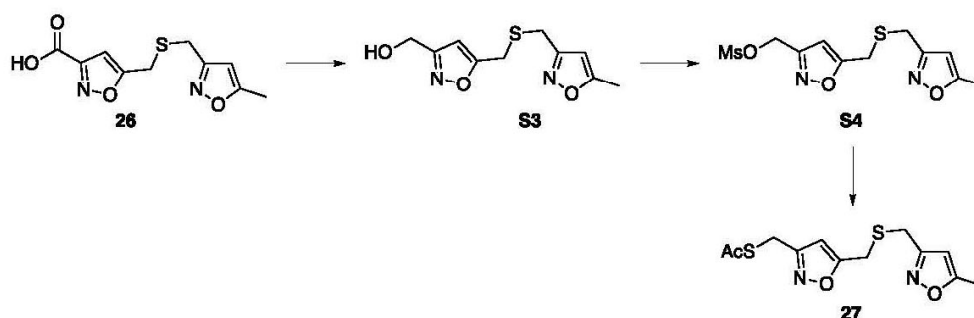
Mesylate **25**: ^1H NMR (CDCl_3 , 500 MHz) δ 6.73 (s, 1H), 5.22 (s, 2H), 4.45 (q, $J = 7.2$ Hz, 2H), 2.13 (s, 3H), 1.42 (t, $J = 7.2$ Hz, 3H); ^{13}C NMR (CDCl_3 , 125 MHz) δ 166.4, 159.3, 156.9, 106.1, 62.7, 59.7, 38.6, 14.2; HRMS m/z calcd. for $\text{C}_8\text{H}_{11}\text{NO}_6\text{SNa}$ $[\text{M}+\text{Na}]^+$: 272.0199, found 272.0199.

Sodium 5-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazole-3-carboxylate (26). Thioester **16** (2.30 g, 13.43 mmol) and mesylate **25** (3.04 g, 12.20 mmol) were dissolved in absolute EtOH (140 mL). The solution was degassed by repeated pump and fill with Ar. A degassed solution of NaOH (4.48 mL, 6 N in H_2O , 26.88 mmol) under Ar was added via syringe to this solution at rt over 10 min. The mixture was vigorously stirred for 1.5 h at rt during which time a white precipitate appeared. The resulting mixture was extracted with EtOAc (200 mL). The pH of the aqueous layer was adjusted to 3 by the addition of 1 N HCl and the resulting mixture was extracted with EtOAc (3×100 mL). The combined organic phases were dried with

Na₂SO₄ and concentrated via rotary evaporation to afford 2.63 g (78%) of the *bis*-isoxazole acid salt **26**, as a colorless wax. Samples of **26** contained traces of NaOMs (¹H NMR (CD₃OD, 500 MHz) δ 6.47 (s, 1H) and ¹³C NMR (CDCl₃, 125 MHz) δ 39.4) under these conditions, the presence of this material did not interfere with further processing.

Bis-isoxazole acid salt **26**: ¹H NMR (CD₃OD, 500 MHz) δ 6.47 (s, 1H), 6.12 (d, *J* = 1.0 Hz, 1H), 3.81 (s, 2H), 3.74 (s, 2H), 2.70 (s, 2H), 2.40 (d, *J* = 0.8 Hz, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 171.8, 171.2, 163.2, 162.8, 104.1, 102.6, 39.4, 26.9, 26.5, 12.0; HRMS *m/z* calcd. for C₁₀H₁₁N₂O₄S [M+H]⁺: 255.0177, found 255.0177.

B.4. Synthesis of thioester 27. The preparation of thioacetate **27** was most effectively conducted in through a four-step two-day procedure from acid **26**.



(5-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazol-3-yl)methanol (S3). Acid **26** (2.63 g, 9.53 mmol) was dissolved in toluene (530 mL). After cooling to -80 °C, DIBAL-H (25.86 mL, 31.03 mmol, 1.2 M) was added carefully in a drop wise fashion over 30 min so that the temperature of the reaction did not exceed -70 °C. The mixture was kept at -80 °C for 1 h at which point it was warmed to 0 °C over 1.5 h. After cooling to -20 °C, EtOAc (100 mL) was added carefully, followed satd. solution of Rochelle's salt (200 mL). The mixture was warmed to rt and vigorously stirred for ~4 h when the solution became clear. The resulting mixture was extracted with EtOAc (3×200 mL). The combined organic phases were washed with brine, dried with Na₂SO₄ and concentrated via rotary evaporation to afford a mixture of the corresponding aldehyde and carbinol **S3**. The crude material was dissolved in EtOH (65 mL) and cooled to 0 °C. NaBH₄ (784.0 mg, 20.72 mmol) was added in portions over 30 min. The resulting mixture was warmed to rt over 1 h and then extracted with EtOAc (3×200 mL). The combined organic layers were washed with brine, dried with Na₂SO₄ and concentrated via rotary evaporation. Flash chromatography (hexanes to 1:1 hexanes:EtOAc) provided 2.10 g (92%) of **S3**, as a clear wax.

Bis-isoxazole carbinol **S3**: ¹H NMR (CDCl₃, 500 MHz) δ 6.26 (s, 1H), 5.98 (s, 1H), 4.67 (s, 2H), 3.67 (s, 2H), 3.61 (s, 2H), 3.34 (bs, 1H), 2.38 (s, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 170.3, 168.9, 164.0, 161.0, 102.1, 101.4, 56.6, 26.1, 25.5, 12.4; HRMS *m/z* calcd. for C₁₀H₁₂N₂O₃SNa [M+Na]⁺: 263.0461, found 264.0459.

(5-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazol-3-yl)methyl methane-sulfonate (S4). *Bis*-isoxazole carbinol **S3** (572.3 mg, 2.38 mmol) was dried by rotary evaporation of toluene (3×10 mL) and then dissolved in dry CH₂Cl₂ (25 mL). Freshly distilled Et₃N (0.51 mL, 3.66 mmol) was added and the flask was cooled to 0 °C. MsCl (0.22 mL, 2.84 mmol) was added drop wise over 15 min. The mixture was kept at 0 °C for 30 min then warmed over 1 h to rt. After 30 min at rt, brine (30 mL) was added. The resulting mixture was extracted with CH₂Cl₂ (3×50

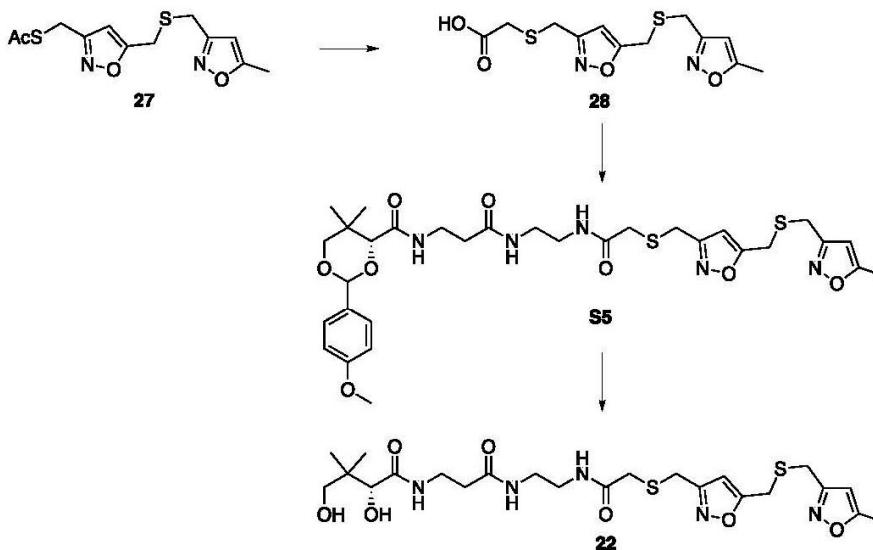
mL), washed with brine, dried with Na₂SO₄, and filtered through a plug of SiO₂ (50 g) washing with 1:1 CH₂Cl₂:EtOAc (200 mL) to deliver the crude *bis*-isoxazole mesylate **S4**.

Bis-isoxazole mesylate **S4**: ¹H NMR (CDCl₃, 500 MHz) δ 6.36 (s, 1H), 5.98 (dd, *J* = 0.9, 0.7 Hz, 1H), 5.24 (s, 2H), 3.71 (d, *J* = 0.7 Hz, 2H), 3.64 (s, 2H), 3.06 (s, 3H), 2.39 (d, *J* = 0.9, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 170.3, 170.3, 160.8, 158.2, 102.6, 101.3, 61.9, 38.1, 26.2, 25.4, 12.3; HRMS *m/z* calcd. for C₁₁H₁₄N₂O₅SNa [M+Na]⁺: 341.0232, found 341.0232.

S-(((5-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazol-3-yl)methyl) ethane-thioate (27). The crude *bis*-isoxazole mesylate **S4** was dissolved in dry DMF (5 mL). KSAc (0.38 g, 3.33 mmol) was added over 5 min as a solid. Within 5 min of the addition, the reaction mixture fused as a solid wax and became warm (45-50 °C). This mixture was allowed to sit for 1 h at which point DMF (2.5 mL) was added and the mixture was allowed to sit for 1.5 h until stirring became possible. Satd. NaHCO₃ (25 mL) was added and the mixture was extracted with CH₂Cl₂ (3×50 mL), washed with brine, and dried with Na₂SO₄. Flash chromatography (hexanes to 1:1 hexanes:EtOAc) provided 631.8 mg (89%) of **27**, as a clear crystalline solid.

Bis-isoxazole thioacetate **27**: ¹H NMR (CDCl₃, 500 MHz) δ 6.15 (s, 1H), 5.98 (d, *J* = 1.0 Hz, 1H), 4.11 (s, 2H), 3.68 (s, 2H), 3.66 (s, 2H), 2.42 (d, *J* = 0.9 Hz, 3H), 2.39 (s, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 194.4, 170.3, 169.4, 161.2, 161.0, 103.1, 101.4, 30.4, 26.4, 25.7, 23.8, 12.5; HRMS *m/z* calcd. for C₁₂H₁₅N₂O₃S₂ [M+H]⁺: 299.0524, found 299.0523.

B.5 Synthesis of mimetic 22. A three-step procedure was used to prepare mimetic **22** from *bis*-isoxazole thioester **27**.



2-(((5-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazol-3-yl)methyl)thio)acetic acid (28). Thioester **27** (200.5 mg, 0.67 mmol) and *tert*-butylbromoacetate (105.0 μL, 0.71 mmol) were dissolved in absolute EtOH (12 mL). The solution was degassed by repeated pump and fill with Ar. A degassed solution of NaOH (0.34 mL, 6 N in H₂O, 2.04 mmol) under Ar was added via syringe to this solution at rt. The resulting mixture was kept at rt for 1.5 h at which point it was extracted with EtOAc (100 mL). The pH of the aqueous layer was adjusted to 3 by the addition of 1 N HCl and the resulting mixture was extracted with EtOAc (3×60 mL). The

combined organic phases were dried with Na₂SO₄ and concentrated via rotary evaporation to afford 172.4 mg (82%) of *bis*-isoxazole acid **28**, as a colorless oil.

Bis-isoxazole acid **28**: ¹H NMR (CDCl₃, 500 MHz) δ 6.33 (s, 1H), 6.07 (s, 1H), 3.84 (s, 2H), 3.72 (s, 2H), 3.64 (s, 2H), 3.25 (s, 2H), 2.43 (s, 3H); ¹³C NMR (CDCl₃, 100 MHz) δ 173.5, 170.9, 169.1, 161.7, 161.3, 103.7, 101.9, 34.0, 27.8, 25.7, 25.5, 12.8; HRMS *m/z* calcd. for C₁₂H₁₃N₂O₄S₂ [M-H]⁺: 313.0322, found 313.0325.

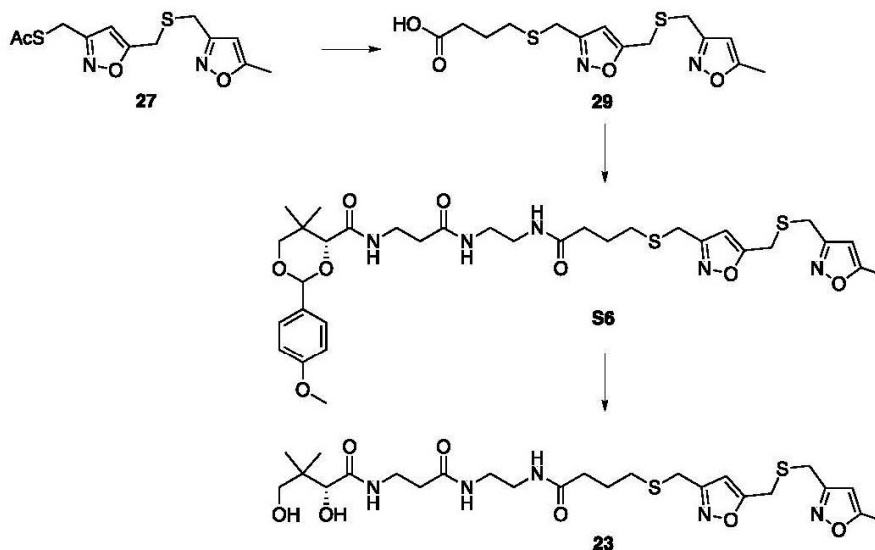
(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-*N*-(3-((2-(2-(((5-methylisoxazol-3-yl)methylthio)methyl)isoxazol-3-yl)methylthio)acetamido)ethyl)amino)-3-oxopropyl)-1,3-dioxane-4-carboxamide (S5). EDAC·HCl (324.60 mg, 1.69 mmol) was added to a mixture of acid **28** (213.00 mg, 0.68 mmol) and amine **18** (257.10 mg, 0.68 mmol) dissolved in dry DMF (10 mL). Within 5 min, Et₃N (0.60 mL, 3.44 mmol) was added and the reaction mixture was stirred at rt. After 12 h, the solvent was removed by rotary evaporation. Flash chromatography (EtOAc to 4:1 EtOAc:MeOH) afforded 209.0 mg (46%) of amide **S5**, as a colorless oil.

Heptaketide amide **S5**: ¹H NMR (CD₃OD, 500 MHz) δ 7.98 (s, 1H, NH), 7.65 (t, *J* = 6.0 Hz, 1H, NH), 7.44 (d, *J* = 8.7 Hz, 2H), 6.92 (d, *J* = 8.7 Hz, 2H), 6.33 (s, 1H), 6.13 (s, 1H), 5.53 (s, 1H), 4.14 (s, 1H), 3.80 (s, 3H), 3.76 (s, 2H), 3.76 (s, 2H), 3.72 (d, *J* = 11.1 Hz, 1H), 3.71 (s, 2H), 3.67 (d, *J* = 11.3 Hz, 1H), 3.45 (d, *J* = 6.6 Hz, 2H), 3.25 (m, 4H), 3.14 (s, 2H), 2.99 (s, 1H), 2.86 (s, 1H), 2.42 (t, *J* = 6.8 Hz, 2H), 2.40 (s, 3H), 1.10 (s, 3H), 1.03 (s, 3H); ¹³C NMR (CD₃OD, 125 MHz) 174.1, 171.8, 171.6, 171.5, 171.4, 162.9, 162.8, 161.7, 131.8, 128.8, 114.5, 103.8, 102.6, 85.1, 79.3, 55.8, 40.4, 39.9, 36.4, 36.4, 36.3, 35.8, 34.0, 27.3, 26.9, 26.4, 22.2, 19.7, 12.1; HR-ESI-TOFMS *m/z* calcd. for C₃₁H₄₁N₅O₈S₂Na [M+Na]⁺: 698.2294, found 698.2269.

(*R*)-2,4-dihydroxy-3,3-dimethyl-*N*-(3-((2-(2-(((5-methylisoxazol-3-yl)methylthio)methyl)isoxazol-3-yl)methylthio)acetamido)ethyl)amino)-3-oxopropyl)butanamide (22). Amide **S5** (22.8 mg, 0.034 mmol) was dissolved in 80% aq. AcOH (1 mL). After 16 h at rt, the solvent was removed via rotary evaporation assisted by azeotropic removal of toluene. Flash chromatography (EtOAc to 4:1 EtOAc:MeOH) afforded 15.3 mg (81%) of mimetic **22**, as a colorless oil.

Mimetic **22**: ¹H NMR (CD₃OD, 500 MHz) δ 6.35 (s, 1H), 6.14 (s, 1H), 3.89 (s, 1H), 3.79 (s, 4H), 3.74 (s, 2H), 3.49 (q, *J* = 6.5 Hz, 2H), 3.46 (d, *J* = 10.8 Hz, 1H), 3.39 (d, *J* = 10.9 Hz, 1H), 3.35 (s, 2H), 3.31 (m, 4H), 3.18 (s, 2H), 2.44 (t, *J* = 6.7 Hz, 2H), 2.41 (d, *J* = 1.0 Hz, 3H), 0.92 (s, 6H); ¹³C NMR (CD₃OD, 125 MHz) δ 176.0, 174.2, 171.9, 171.8, 171.4, 163.0, 162.9, 103.8, 102.6, 77.3, 70.3, 49.8, 40.4, 39.9, 36.6, 36.4, 35.8, 27.3, 26.9, 26.4, 21.4, 20.9, 12.1; HRMS *m/z* calcd. for C₂₃H₃₅N₅O₇S₂Na [M+Na]⁺: 580.1876, found 580.1880.

B.6. Synthesis of mimetic 23. A three-step procedure was used to prepare mimetic **23** from *bis*-isoxazole thioester **27**.



4-(((5-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazol-3-yl)methyl)thio)-butanoic acid (29**).** Thioester **27** (279.0 mg, 0.94 mmol) and *tert*-butyl 4-bromobutanoate (313.0 mg, 1.40 mmol) were dissolved in absolute EtOH (18 mL). The solution was degassed by repeated pump and fill with Ar. A degassed solution of NaOH (0.47 mL, 5 N in H₂O, 2.35 mmol) under Ar was added via syringe to this solution at rt. The resulting mixture was kept at rt for 1.5 h at which point it was extracted with EtOAc (100 mL). The pH of the aqueous layer was adjusted to 3 by the addition of 1 N HCl and the resulting mixture was extracted with EtOAc (3×60 mL). The combined organic phases were dried with Na₂SO₄ and concentrated via rotary evaporation to afford 312.5 mg (97%) of *bis*-isoxazole acid **29**, as a colorless oil.

Bis-isoxazole acid **29**: ¹H NMR (CDCl₃, 500 MHz) δ 6.26 (s, 1H), 6.02 (d, *J* = 1.1 Hz, 1H), 3.68 (s, 2H), 3.67 (s, 2H), 3.66 (s, 2H), 2.53 (t, *J* = 7.2 Hz, 2H), 2.45 (t, *J* = 7.2 Hz, 2H), 2.42 (d, *J* = 0.9 Hz, 3H), 1.90 (p, *J* = 7.2 Hz, 2H); ¹³C NMR (CDCl₃, 125 MHz) δ 177.9, 170.4, 169.1, 162.0, 161.1, 102.9, 101.5, 32.6, 30.7, 26.2, 26.1, 25.7, 23.8, 12.5; HR-ESI-TOFMS *m/z* calcd. for C₁₄H₁₇N₂O₄S₂ [M-H]⁺: 341.0635, found 341.0640.

(4*R*)-2-(4-methoxyphenyl)-5,5-dimethyl-*N*-(3-((2-(4-(((5-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazol-3-yl)methyl)thio)butanamido)ethyl)amino)-3-oxopropyl)-1,3-dioxane-4-carboxamide (S6**).** EDAC·HCl (28.0 mg, 0.15 mmol) was added to an ice-cold mixture of acid **29** (20.0 mg, 0.06 mmol) and amine **18** (22.1 mg, 0.06 mmol) dissolved in dry CH₂Cl₂ (1 mL). Within 5 min, Et₃N (0.05 mL, 0.29 mmol) was added and the reaction mixture was stirred at rt. After 12 h, the solvent was removed by rotary evaporation. Flash chromatography (EtOAc to 4:1 EtOAc:MeOH) afforded 31.8 mg (77%) of **S6**, as a colorless oil.

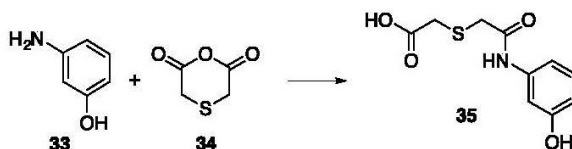
Amide **S6**: ¹H NMR (CD₃OD, 500 MHz) δ 7.94 (s, 1H, NH), 7.65 (t, *J* = 6.1 Hz, 1H, NH), 7.45 (d, *J* = 8.7 Hz, 2H), 6.92 (d, *J* = 8.8 Hz, 2H), 6.31 (s, 1H), 6.13 (s, 1H), 5.53 (s, 1H), 5.49 (s, 1H), 4.15 (s, 1H), 3.80 (s, 3H), 3.76 (s, 2H), 3.72 (d, *J* = 11.7 Hz, 1H), 3.71 (s, 2H), 3.68 (s, 2H), 3.67 (d, *J* = 11.6 Hz, 1H), 3.48 (m, 3H), 3.35 (s, 3H), 3.23 (m, 3H), 2.47 (t, *J* = 7.2 Hz, 2H), 2.41 (d, *J* = 1.0 Hz, 3H), 2.39 (t, *J* = 6.5 Hz, 2H), 2.25 (t, *J* = 7.4 Hz, 2H), 1.84 (p, *J* = 7.4 Hz, 2H), 1.10 (s,

3H), 1.03 (s, 3H); ^{13}C NMR (CD_3OD , 125 MHz) δ 175.5, 174.1, 171.8, 171.6, 171.2, 163.6, 162.8, 161.7, 131.8, 128.8, 114.5, 103.8, 102.6, 102.5, 85.1, 79.3, 55.7, 49.8, 40.0, 36.4, 36.4, 35.8, 35.8, 34.0, 31.8, 26.9, 26.5, 26.4, 26.1, 22.2, 19.7, 12.1; HR-ESI-TOFMS m/z calcd. for $\text{C}_{33}\text{H}_{45}\text{N}_5\text{O}_8\text{S}_2\text{Na}$ $[\text{M}+\text{Na}]^+$: 726.2607, found 726.2596.

(R)-2,4-dihydroxy-3,3-dimethyl-N-(3-((2-(4-(((5-methylisoxazol-3-yl)methyl)thio)methyl)isoxazol-3-yl)methyl)thio)butanamido)ethyl)amino)-3-oxopropyl)butanamide (23). Amide **S6** (5.8 mg, 8.24 μmol) was dissolved in 80% aq. AcOH (0.2 mL). After 16 h at rt, the solvent was removed via rotary evaporation assisted by azeotropic removal of toluene. Flash chromatography (EtOAc to 3:1 EtOAc:MeOH) afforded 2.8 mg (58%) of mimetic **23**, as a colorless oil.

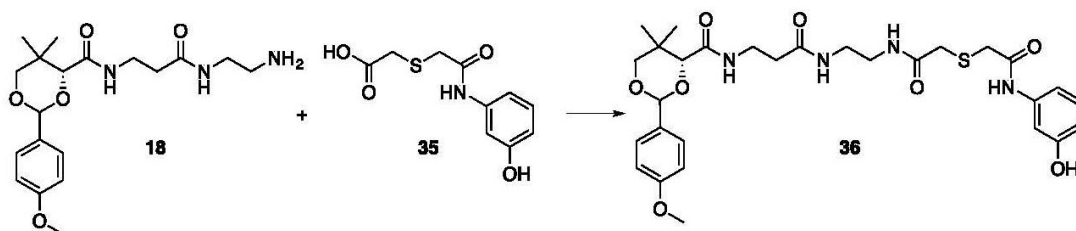
Mimetic **23**: ^1H NMR (CD_3OD , 500 MHz) δ 6.33 (s, 1H), 6.14 (s, 1H), 5.49 (s, 1H), 3.89 (s, 1H) 3.78 (s, 2H), 3.73 (s, 2H), 3.70 (s, 2H), 3.48 (q, $J = 6.7$ Hz, 2H), 3.46 (d, $J = 10.6$ Hz, 1H), 3.39 (d, $J = 11.0$ Hz, 1H), 3.35 (s, 2H), 3.27 (m, 3H), 2.50 (t, $J = 7.2$ Hz, 2H), 2.42 (t, $J = 6.6$ Hz, 2H), 2.41 (s, 3H), 2.29 (t, $J = 7.4$ Hz, 2H), 1.87 (p, $J = 7.5$ Hz, 2H), 0.92 (s, 6H); ^{13}C NMR (CD_3OD , 125 MHz) δ 176.1, 175.6, 174.1, 171.8, 171.2, 163.6, 162.9, 103.8, 102.6, 77.3, 70.3, 54.8, 49.8, 40.4, 40.1, 40.0, 36.6, 36.4, 35.8, 31.8, 26.9, 26.5, 26.4, 26.1, 21.4, 20.9, 12.1; HRMS m/z calcd. for $\text{C}_{25}\text{H}_{39}\text{N}_5\text{O}_7\text{S}_2\text{Na}$ $[\text{M}+\text{Na}]^+$: 608.2189, found 608.2200.

B.7. Synthesis of mimetic 30. The first mimetic **30** as shown in Scheme 3 of the manuscript were prepared through a three to four step sequence. The following section provides experimental procedures for these steps.



2-(2-(3-Hydroxyphenylamino)-2-oxoethylthio)acetic acid (35). A mixture of *m*-aminophenol (**33**) (1.14 g, 10.46 mmol) and thiodiglycolic anhydride (**34**) (1.32 g, 10 mmol) was dissolved in dry DMF (5 mL). After 16 h at rt, the solvent was removed via rotary evaporation to afford pale yellow solid residue. Flash chromatography (1:1 hexanes:EtOAc to 9:1 EtOAc:MeOH) afforded 2.37 g (97%) of acid **35**.

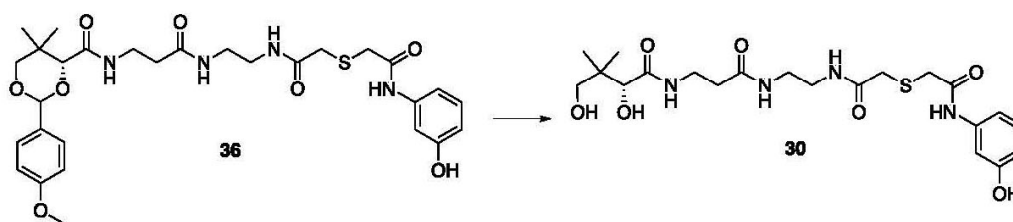
Acid **35**: TLC (2:1 EtOAc/MeOH): $R_f = 0.6$; ^1H NMR ($\text{DMSO}-d_6$, 400 MHz) δ 12.63 (brs, 1H), 9.95 (s, 1H), 9.37 (s, 1H), 7.16 (t, $J = 2.1$ Hz, 1H), 7.06 (t, $J = 8.1$ Hz, 1H), 6.92 (ddd, $J = 1.0, 1.9, 8.0$ Hz, 1H), 6.45 (ddd, $J = 1.0, 2.2, 8.1$, 1H), 3.42 (s, 2H), 3.39 (s, 2H); ^{13}C NMR ($\text{DMSO}-d_6$, 100 MHz) δ 171.1, 167.3, 157.6, 140.0, 129.4, 110.5, 109.9, 106.3, 35.0, 33.7; HR-ESI-TOFMS m/z calcd. for $\text{C}_{10}\text{H}_{10}\text{NO}_4\text{S}$ $[\text{M}-\text{H}]^-$: 240.0331, found 240.0334.



(4R)-N-(3-(2-(2-(2-(3-Hydroxyphenylamino)-2-oxoethylthio)acetamido)ethyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (36). EDAC·HCl

(1.77 g, 9.23 mmol) was added to a mixture of acid **35** (916 mg, 3.8 mmol) and amine **18** (1.44 g, 3.8 mmol) dissolved in dry DMF (10 mL). Et₃NPr₂ (3.22 mL, 18.47 mmol) was added at rt. After 12 h at rt, the solvent was removed by rotary evaporation. Flash chromatography (EtOAc to 9:1 EtOAc:MeOH) provided 720.0 mg (32%) of amide **36**, as colorless oil.

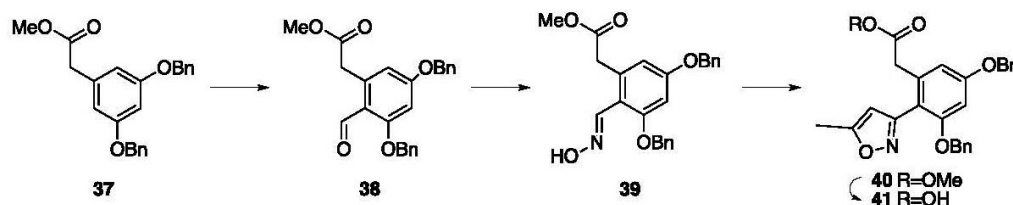
Amide **36**: TLC (9:1 EtOAc/MeOH): R_f = 0.6; ¹H NMR (CD₃OD, 500 MHz) δ 7.43 (d, *J* = 8.7 Hz, 2H), 7.17 (t, *J* = 2.2 Hz, 1H), 7.09 (t, *J* = 8.1 Hz, 1H), 6.94 (ddd, *J* = 1.0, 2.0, 8.1 Hz, 1H), 6.90 (d, *J* = 8.7 Hz, 2H), 6.53 (ddd, *J* = 1.0, 2.4, 8.1 Hz, 1H), 5.50 (s, 1H), 4.12 (s, 1H), 3.77 (s, 3H), 3.69 (d, *J* = 11.5 Hz, 1H), 3.65 (d, *J* = 11.4 Hz, 1H), 3.46 (t, *J* = 6.6 Hz, 2H), 3.37 (s, 2H), 3.34 (s, 2H), 3.30 (m, 4H), 3.24 (m, 2H), 2.38 (t, *J* = 6.6 Hz, 2H), 1.08 (s, 3H), 1.01 (s, 3H); ¹³C NMR (CD₃OD, 125 MHz) δ 174.1, 172.0, 171.6, 170.0, 161.7, 159.0, 140.6, 131.8, 130.6, 128.8, 114.5, 112.4, 112.3, 108.4, 102.6, 85.1, 79.3, 55.7, 49.6, 40.4, 39.8, 37.6, 36.7, 36.5, 36.3, 34.0, 22.2, 19.7; HR-ESI-TOFMS *m/z* calcd. for C₂₉H₃₈N₄O₈SNa [M+Na]⁺: 625.2308, found 625.2319.



(R)-2,4-Dihydroxy-N-(3-(2-(2-(2-(3-hydroxyphenylamino)-2-oxoethylthio)-acetamido)-ethylamino)-3-oxopropyl)-3,3-dimethylbutanamide (30): Amide **36** (50.0 mg, 0.082 mmol) was dissolved in 80% aq. AcOH (1 mL). After 16 h at rt, the solvent was removed via rotary evaporation assisted by azeotropic removal of toluene. Flash chromatography (EtOAc to 4:1 EtOAc:MeOH) afforded 24.6 mg (62%) mimetic **30**, as a colorless oil.

Mimetic **30**: TLC (1:2 MeOH/EtOAc): R_f = 0.6; ¹H NMR (CD₃OD, 500 MHz) δ 7.17 (t, *J* = 2.1 Hz, 1H), 7.10 (t, *J* = 8.1 Hz, 1H), 6.95 (dd, *J* = 2.6, 7.9 Hz, 1H), 6.54 (dd, *J* = 2.4, 8.2 Hz, 1H), 3.89 (s, 1H), 3.48 (m, 2H), 3.45 (d, *J* = 11.4 Hz, 1H), 3.41 (s, 2H), 3.37 (d, *J* = 10.9 Hz, 1H), 3.35 (s, 2H), 3.28 (m, 4H), 2.40 (t, *J* = 6.6 Hz, 2H), 0.91 (s, 6H); ¹³C NMR (CD₃OD, 125 MHz) δ 176.0, 174.2, 172.1, 170.1, 158.9, 140.6, 130.6, 112.5, 112.3, 108.4, 77.3, 70.3, 40.4, 40.3, 39.9, 37.7, 36.7, 36.6, 36.4, 21.4, 21.0; HR-ESI-TOFMS *m/z* calcd. for C₂₁H₃₂N₄O₇SNa [M+Na]⁺: 507.1889, found 507.1871.

B.8. Synthesis of mimetic 31. The synthesis of **31** began with the preparation of isoxazole intermediate **41** in 4-steps from ester **37**, as shown in Scheme 4 of the manuscript. A three-step procedure was then used to convert **41** into mimetic **31**. The following section provides experimental procedures for these steps.



Methyl 2-(3,5-bis(benzyloxy)-2-formylphenyl)acetate (38). Ester **37** (3.0 g, 8.23 mmol) was dissolved in DMF (4 mL) and heated to 50 °C. POCl₃ (1.16 mL, 12.4 mmol) was added and the

reaction mixture was heated to 100 °C. After 10 min, the reaction was cooled to rt. After 16 h at rt, the mixture was cooled to 0 °C and 10% aq. NaOAc (20 mL) was added. After 3 h, the reaction was filtered to give a tan solid. Recrystallization from 95% EtOH afforded 2.25 g (70%) of aldehyde **38**, as a white solid.

Aldehyde **38**: $^1\text{H NMR}$ (CDCl_3 , 500 MHz) δ 10.53 (s, 1H), 7.41 (t, $J = 4.6$ Hz, 4H), 7.37 (m, 6H), 6.59 (d, $J = 2.3$ Hz, 1H), 6.46 (d, $J = 2.3$ Hz, 1H), 5.11 (s, 2H), 5.09 (s, 2H), 3.96 (s, 2H), 3.72 (s, 3H); $^{13}\text{C NMR}$ (CDCl_3 , 125 MHz) δ 190.1, 171.6, 164.5, 163.9, 139.3, 135.8, 135.8, 128.2, 128.8, 128.5, 128.4, 127.7, 127.4, 117.4, 111.2, 99.2, 70.7, 70.4, 52.0, 40.7; HRMS m/z calcd. for $\text{C}_{24}\text{H}_{22}\text{O}_5\text{Na}$ $[\text{M}+\text{Na}]^+$: 413.1350, found 413.1360.

Methyl (E)-2-(3,5-bis(benzyloxy)-2-((hydroxyimino)methyl)phenyl)acetate (39). NaOAc (189.5 mg, 2.31 mmol) and $\text{NH}_2\text{OH}\cdot\text{HCl}$ (97.9 mg, 1.41 mmol) were added sequentially to a suspension of aldehyde **38** (500 mg, 1.28 mmol) in MeOH (9 mL). The resulting reaction mixture was heated to reflux for 1 h. The reaction was allowed to cool to rt and the solvent was removed by rotary evaporation. The resulting solid was dissolved in CH_2Cl_2 (15 mL) and washed with H_2O (2×15 mL), brine (15 mL) and the organic phase was dried over Na_2SO_4 . The solvent was removed to yield oxime **39**, which was used without further purification. Analytical samples were purified by flash column chromatography (hexanes to 4:1 hexanes:EtOAc).

Oxime **39**: $^1\text{H NMR}$ (CDCl_3 , 500 MHz) δ 8.57 (s, 1H), 7.68 (bs, 1H), 7.39 (m, 6H), 7.35 (m, 4H), 6.56 (d, $J = 2.3$ Hz, 1H), 6.49 (d, $J = 2.3$ Hz, 1H), 5.40 (s, 4H), 3.91 (s, 2H), 3.67 (s, 3H); $^{13}\text{C NMR}$ (CDCl_3 , 125 MHz) δ 190.1, 171.6, 164.5, 163.9, 139.3, 135.8, 128.8, 128.5, 128.4, 127.7, 127.4, 117.4, 111.2, 99.2, 70.7, 70.4, 52.0, 40.7.

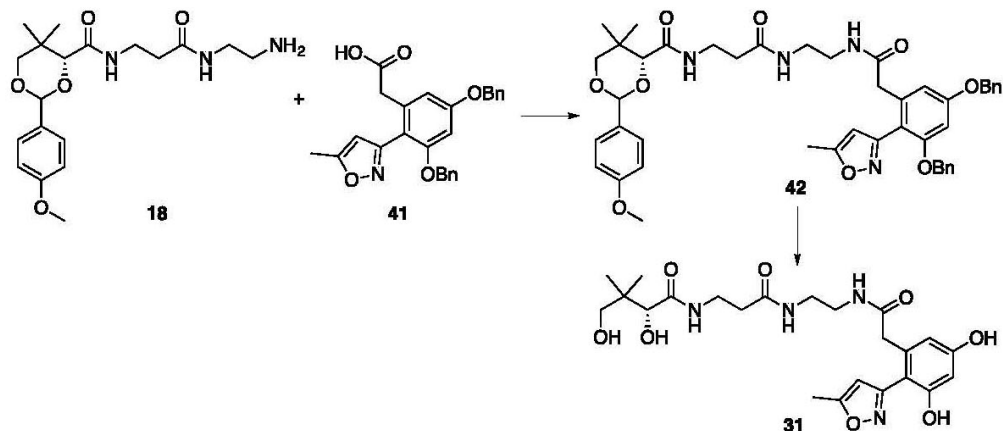
Methyl 2-(3,5-bis(benzyloxy)-2-(5-methylisoxazol-3-yl)phenyl)acetate (40). Et_3N (0.20 mL, 1.41 mmol) was added to a solution of oxime **39** in CHCl_3 (6 mL) and the solution was cooled to 0 °C. *N*-Chlorosuccinimide (88.3 mg, 1.41 mmol) was added. After 1.5 h at 0 °C, isopropenyl acetate (1.41 mL, 12.8 mmol) was added and the reaction mixture was heated to 50 °C for 18 h. The reaction mixture was allowed to cool, washed with water (3×5 mL), brine (5 mL) and dried over Na_2SO_4 . The solvent was removed via rotary evaporation to afford isoxazole **40**, which was used without further purification. Analytical samples were purified by flash column chromatography (20:1 hexanes:EtOAc to 3:1 hexanes:EtOAc).

Isoxazole **40**: $^1\text{H NMR}$ (CDCl_3 , 500 MHz) δ 7.41 (m, 6H), 7.32 (m, 4H), 6.61 (d, $J = 2.4$ Hz, 1H), 6.59 (d, $J = 2.4$ Hz, 1H), 6.13 (d, $J = 1.1$ Hz, 1H), 5.05 (s, 2H), 5.05 (s, 2H), 3.73 (s, 2H), 3.64 (s, 3H), 2.44 (d, $J = 1.0$ Hz, 2H); $^{13}\text{C NMR}$ (CDCl_3 , 125 MHz) δ 171.9, 168.3, 160.3, 159.3, 158.2, 136.7, 136.5, 136.1, 128.7, 128.6, 128.2, 127.8, 127.7, 126.9, 112.6, 108.7, 104.6, 100.0, 70.5, 70.2, 52.1, 39.6, 12.4; HRMS m/z calcd. for $\text{C}_{27}\text{H}_{25}\text{N O}_5\text{Na}$ $[\text{M}+\text{Na}]^+$: 466.1625, found 466.1626.

2-(3,5-Bis(benzyloxy)-2-(5-methylisoxazol-3-yl)phenyl)acetic acid (41). A 1 M solution of NaOH (7 mL) was added to isoxazole **40** in MeOH (7 mL) and the mixture was heated to 50 °C. After 2 h, the mixture was allowed to cool to rt and the MeOH was removed by rotary evaporation. The resulting aqueous reaction mixture was neutralized with 1 M HCl and extracted with EtOAc (3×15 mL). The organic layers were combined, dried over Na_2SO_4 and concentrated via rotary evaporation. The crude mixture was purified by flash column chromatography (100:1 CHCl_3 :MeOH to 25:1 CHCl_3 :MeOH) to afford 365.0 mg (65% over 3 steps from **37**) of acid **41**, as a white solid.

Acid **41**: $^1\text{H NMR}$ ($\text{DMSO}-d_6$, 500 MHz) δ 12.16 (bs, 1H), 7.46 (m, 2H), 7.40 (m, 2H), 7.35 (m, 5H), 7.29 (m, 1H), 6.78 (d, $J = 2.3$ Hz, 1H), 6.68 (d, $J = 2.3$ Hz, 1H), 6.24 (d, $J = 1.0$ Hz, 1H), 5.13 (s, 2H), 5.11 (s, 2H), 3.52 (s, 2H), 2.41 (d, $J = 1.0$ Hz, 2H); $^{13}\text{C NMR}$ ($\text{DMSO}-d_6$, 125 MHz)

δ 172.0, 168.1, 159.7, 158.9, 157.4, 136.9, 136.7, 136.7, 128.5, 128.5, 128.0, 128.0, 127.8, 127.2, 111.8, 109.5, 104.5, 99.3, 69.6, 69.5, 11.8; HRMS m/z calcd. for $C_{26}H_{23}NO_5Na$ $[M+Na]^+$: 452.1468, found 452.1471.



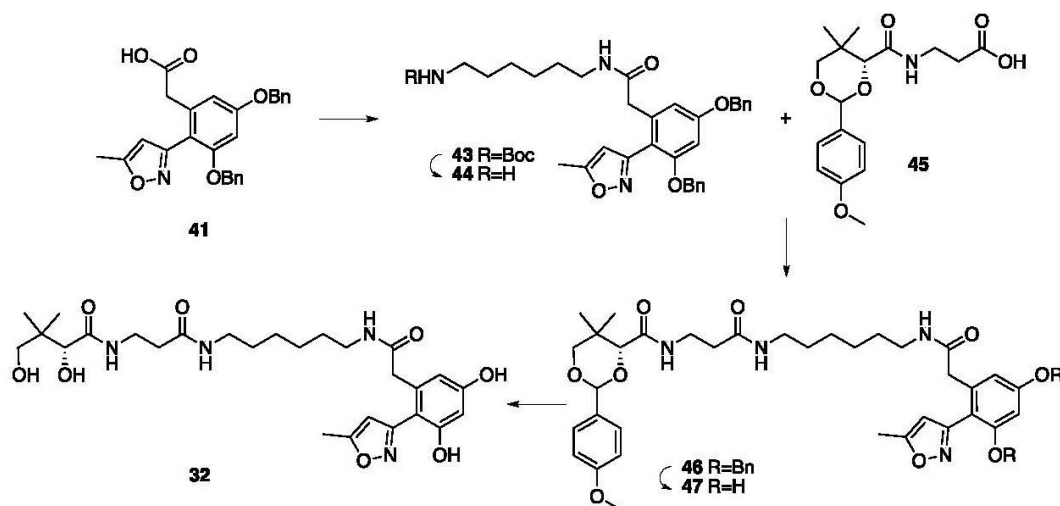
(4*R*)-*N*-(3-((2-(2-(3,5-Bis(benzyloxy)-2-(5-methylisoxazol-3-yl)phenyl)acetamido)thyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (42). EtNPr₂ (0.085 mL, 0.49 mmol) was added to a solution of acid **41** (70.0 mg, 0.16 mmol) and HATU (93.2 mg, 0.25 mmol) in DMF (1 mL). After 10 min at rt, amine **18** (74.4 mg, 0.196 mmol) was added to the resulting reaction mixture. After 18 h at rt, the solvent was removed by rotary evaporation. The crude material was purified by flash column chromatography (100:1 CH₂Cl₂:MeOH to 33:1 CH₂Cl₂:MeOH) to afford 73.3 mg (57%) of amide **42**, as a white solid.

Amide **42**: ¹H NMR (CDCl₃, 500 MHz) δ 7.50 (m, 2H), 7.44 (m, 4H), 7.38 (m, 6H), 7.31 (t, J = 1.7 Hz, 1H), 7.29 (m, 1H), 7.22 (m, 1H), 6.91 (t, J = 8.8 Hz, 2H), 6.77 (d, J = 2.4 Hz, 1H), 6.76 (d, J = 2.3 Hz, 1H), 6.29 (d, J = 0.9 Hz, 1H), 5.51 (s, 1H), 5.15 (s, 4H), 4.05 (s, 1H), 3.79 (s, 3H), 3.69 (d, J = 11.0 Hz, 1H), 3.63 (d, J = 11.2 Hz, 1H), 3.42 (s, 2H), 3.41 (m, 2H), 3.21 (m, 4H), 2.45 (d, J = 0.9 Hz, 3H), 2.34 (t, J = 6.5 Hz, 2H), 1.07 (s, 3H), 1.03 (s, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 171.9, 171.0, 169.4, 169.4, 161.4, 161.0, 160.4, 159.1, 138.7, 138.1, 131.9, 129.5, 129.4, 129.3, 128.9, 128.7, 128.6, 128.6, 128.1, 114.2, 112.8, 109.8, 105.7, 102.0, 100.3, 84.6, 78.9, 71.0, 70.7, 55.6, 42.3, 39.9, 36.4, 35.7, 33.7, 22.2, 19.6, 12.2; HRMS m/z calcd. for $C_{45}H_{50}N_4O_9$ $[M+Na]^+$: 813.4270, found 813.3472.

(*R*)-*N*-(3-((2-(2-(3,5-Dihydroxy-2-(5-methylisoxazol-3-yl)phenyl)acetamido)ethyl)amino)-3-oxopropyl)-2,4-dihydroxy-3,3-dimethylbutanamide (31). Amide **42** (38.0 mg, 0.048 mmol) was dissolved in MeOH (2 mL) and 5% Pd on C (10.2 mg, 0.0048 mmol) was added as a powder. The reaction mixture degassed and charged with a H₂ atmosphere. After 4.5 h at rt, the mixture was filtered through a pad of Celite (1 g) and concentrated by rotary evaporation. The resulting tan solid was washed sequentially with hexanes (2 mL) and CH₂Cl₂ (2 mL) providing 13.5 mg (57%) of mimetic **31**, which used without further purification.

Mimetic **30**: ¹H NMR (CDCl₃, 500 MHz) δ 6.36 (d, J = 2.2 Hz, 1H), 6.34 (d, J = 2.3 Hz, 1H), 6.25 (d, J = 0.9 Hz, 1H), 3.87 (s, 1H), 3.45 (m, 2H), 3.43 (d, J = 12.2 Hz, 1H), 3.41 (s, 2H), 3.38 (d, J = 11.0 Hz, 1H), 3.24 (m, 4H), 2.45 (d, J = 0.9 Hz, 3H), 2.40 (t, J = 6.6 Hz, 2H), 0.91 (s, 3H), 0.90 (s, 3H); ¹³C NMR (CDCl₃, 125 MHz) δ 174.5, 172.9, 172.7, 168.8, 160.1, 159.1, 157.3, 136.3, 109.1, 108.0, 104.4, 101.3, 79.5, 68.9, 48.5, 40.8, 38.9, 38.5, 35.2, 34.9, 19.9, 19.7, 10.7; HRMS m/z calcd. for $C_{23}H_{32}N_4O_8$ $[M-H]^-$: 491.2147, found 491.2151.

B.9. Synthesis of mimetic 32. A five-step procedure was used to convert acid intermediate **41** into mimetic **32**. The following section provides experimental procedures for these steps.



tert-Butyl (6-(2-(3,5-bis(benzyloxy)-2-(5-methylisoxazol-3-yl)phenyl)acetamido)hexyl)-carbamate (43). EtNⁱPr₂ (0.24 mL, 1.40 mmol) was added to a solution of acid **41** (200.0 mg, 0.47 mmol) and HATU (265.8 mg, 0.70 mmol) in DMF (2 mL). After 10 min at rt, *N*-boc-1,6-hexanediamine (110.8 mg, 0.51 mmol) was added. After 18 h at rt, the solvent was removed by rotary evaporation. The crude material was purified by flash column chromatography (10:1 hexanes:EtOAc to 1:1 hexanes:EtOAc) to afford 220.3 mg (75%) of Boc-amide **43**, as a white solid.

Boc-amide **43**: ¹H NMR (CDCl₃, 500 MHz) δ 7.61 (m, 1H), 7.42 (m, 2H), 7.38 (m, 2H), 7.34 (m, 4H), 7.29 (m, 2H), 6.82 (d, *J* = 2.3 Hz, 1H), 6.57 (d, *J* = 2.3 Hz, 1H), 6.18 (d, *J* = 1.1 Hz, 1H), 5.06 (s, 2H), 5.04 (s, 2H), 3.45 (s, 2H), 3.18 (q, *J* = 6.6 Hz, 2H), 3.06 (q, *J* = 6.7 Hz, 2H), 2.46 (d, *J* = 0.9 Hz, 3H), 1.47 (m, 4H), 1.43 (s, 9H), 1.29 (m, 4H); ¹³C NMR (CDCl₃, 125 MHz) δ 171.0, 168.4, 160.6, 159.6, 158.2, 156.1, 138.3, 136.5, 136.5, 128.7, 128.6, 128.3, 128.0, 127.9, 127.0, 111.2, 107.1, 105.1, 100.0, 79.1, 70.4, 70.3, 42.2, 40.5, 39.5, 30.0, 29.3, 28.5, 26.5, 26.4, 12.4; HRMS *m/z* calcd. for C₃₇H₄₅N₃O₆Na [M+Na]⁺: 650.3199, found 650.3201

(4*R*)-*N*-(3-((6-(2-(3,5-bis(benzyloxy)-2-(5-methylisoxazol-3-yl)phenyl)acetamido)-hexyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (46). Boc-amide **43** (60.0 mg, 0.096 mmol) was dissolved in CH₂Cl₂ (2 mL) and 4 M HCl in dioxanes (100 μL) was added at rt. After 3 h at rt, the solvent was removed to afford amine **44**. EtNⁱPr₂ (0.075 mL, 0.430 mmol) was added to a solution of acid **45** (32.4 mg, 0.096 mmol) and HATU (54.4 mg, 0.14 mmol) in DMF (1 mL). After 10 min at rt, amine **44** (54.2 mg, 0.096 mmol) was added as solution in DMF (1 mL). After 18 h at rt, the solvent was removed by rotary evaporation. The crude material was purified by flash column chromatography (100:1 CH₂Cl₂:MeOH to 33:1 CH₂Cl₂:MeOH) to yield 65.4 mg (80%) of amide **46**, as a white foam.

Amide **46**: ¹H NMR (CDCl₃, 500 MHz) δ 7.49 (m, 2H), 7.45 (d, *J* = 8.7 Hz, 2H), 7.37 (m, 7H), 7.30 (m, 1H), 7.26 (t, *J* = 5.9 Hz, 1H), 7.16 (t, *J* = 5.8 Hz, 1H), 6.91 (d, *J* = 8.7 Hz, 2H), 6.76 (d, *J* = 2.1 Hz, 1H), 6.75 (d, *J* = 2.4 Hz, 1H), 6.29 (d, *J* = 1.0 Hz, 1H), 5.54 (s, 1H), 5.15 (s, 2H), 5.13 (s, 2H), 4.08 (s, 1H), 3.79 (s, 3H), 3.69 (d, *J* = 11.4 Hz, 1H), 3.64 (d, *J* = 11.3 Hz, 1H), 3.48 (qd, *J* = 6.5, 13.0 Hz, 1H), 3.42 (s, 2H), 3.37 (qd, *J* = 6.3, 12.8 Hz, 1H), 3.11 (qd, *J* = 6.8, 12.8 Hz,

2H), 2.44 (d, $J = 0.9$ Hz, 3H), 2.33 (t, $J = 6.5$ Hz, 2H), 1.41 (m, 4H), 1.27 (m, 4H), 1.07 (s, 3H), 1.03 (s, 3H); ^{13}C NMR (CDCl_3 , 125 MHz) δ 171.3, 170.3, 169.2, 169.2, 161.2, 160.9, 160.3, 159.0, 139.1, 137.9, 131.8, 129.3, 129.2, 128.8, 128.5, 128.0, 114.1, 112.6, 109.2, 105.7, 101.8, 100.0, 84.4, 78.8, 70.9, 70.6, 55.5, 42.0, 39.5, 39.4, 36.0, 35.6, 33.6, 26.9, 22.1, 19.5, 12.1; HRMS m/z calcd. for $\text{C}_{49}\text{H}_{56}\text{N}_4\text{O}_9$ $[\text{M}+\text{Na}]^+$: 869.4096, found 869.4093.

(4R)-N-(3-((6-(2-(3,5-dihydroxy-2-(5-methylisoxazol-3-yl)phenyl)acetamido)hexyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (47). 5% Pd on C (5 mg, 0.002 mmol) was added as a powder to a solution of **46** (20 mg, 0.024 mmol) in EtOAc (1 mL) and MeOH (0.1 mL). The reaction mixture degassed and charged with a H_2 atmosphere. After 5 h at rt, the mixture was filtered through a pad of Celite (1 g) and concentrated by rotary evaporation to afford 10.2 mg (65%) of amide **47**, which was used without further purification.

Amide **47**: ^1H NMR (acetone- d_6 , 500 MHz) δ 8.95 (s, 1H), 7.45 (m, 2H), 7.34 (t, $J = 6.0$ Hz, 1H), 7.30 (t, $J = 5.6$ Hz, 1H), 6.91 (m, 2H), 6.52 (d, $J = 2.4$ Hz, 1H), 6.43 (d, $J = 0.9$ Hz, 1H), 6.42 (d, $J = 2.4$ Hz, 1H), 5.55 (s, 1H), 4.12 (s, 1H), 3.79 (s, 3H), 3.71 (dd, $J = 0.9, 11.3$ Hz, 1H), 3.65 (d, $J = 11.2$ Hz, 1H), 3.50 (m, 1H), 3.45 (m, 1H), 3.42 (s, 2H), 3.13 (m, 4H), 2.46 (d, $J = 0.9$ Hz, 3H), 2.39 (t, $J = 6.6$ Hz, 2H), 1.41 (m, 4H), 1.25 (m, 4H), 1.08 (s, 3H), 1.04 (s, 3H); ^{13}C NMR (acetone- d_6 , 125 MHz) δ 171.3, 170.3, 169.2, 169.2, 161.2, 160.9, 160.3, 159.0, 139.1, 137.9, 131.8, 129.3, 128.8, 128.5, 128.0, 114.1, 112.6, 109.2, 105.7, 101.8, 100.0, 84.4, 78.8, 70.9, 70.6, 55.5, 42.0, 39.5, 39.4, 36.0, 35.6, 33.6, 26.9, 22.1, 19.5, 12.1; HRMS m/z calcd. for $\text{C}_{35}\text{H}_{46}\text{N}_4\text{O}_9$ $[\text{M}+\text{Na}]^+$: 689.3157, found 689.3159.

(R)-N-(3-((2-(2-(3,5-dihydroxy-2-(5-methylisoxazol-3-yl)phenyl)acetamido)ethyl)amino)-3-oxopropyl)-2,4-dihydroxy-3,3-dimethylbutanamide (32). Amide **47** (4.0 mg, 0.0060 mmol) was dissolved in 80% aq. AcOH (1 mL). After 3.5 h, the reaction was concentrated by rotary evaporation. The resulting light-brown solid was washed with sequentially hexanes (2 mL), CH_2Cl_2 (2 mL) and EtOAc (2 mL) to afford 2.3 mg (63%) of mimetic **32**, as a tan solid.

Mimetic **32**: ^1H NMR (CD_3OD , 500 MHz) δ 6.36 (d, $J = 2.3$ Hz, 1H), 6.32 (d, $J = 2.3$ Hz, 1H), 6.26 (d, $J = 0.9$ Hz, 1H), 3.88 (s, 1H), 3.50 (m, 2H), 3.46 (d, $J = 10.9$ Hz, 1H), 3.42 (s, 2H), 3.38 (d, $J = 11.0$ Hz, 1H), 3.13 (m, 4H), 2.46 (d, $J = 0.9$ Hz, 3H), 2.40 (t, $J = 6.7$ Hz, 2H), 1.46 (m, 4H), 1.30 (m, 4H), 0.91 (s, 3H), 0.91 (s, 3H); ^{13}C NMR (CD_3OD , 125 MHz) δ 176.1, 173.8, 173.6, 169.9, 161.4, 160.5, 158.7, 138.2, 109.9, 109.3, 105.9, 102.5, 77.2, 70.3, 42.0, 40.4, 36.4, 36.4, 30.8, 30.2, 24.1, 23.8, 21.3, 20.9, 12.0; HRMS m/z calcd. for $\text{C}_{27}\text{H}_{40}\text{N}_4\text{O}_8$ $[\text{M}+\text{Na}]^+$: 571.2743, found 571.2739.

C. Protein NMR Studies.

The following methods were used to collect the NMR data presented in Fig. 4 and Fig. 6 of the manuscript.

C.1. Expression, modification, and purification of *Pseudomonas fluorescens* AcpH. pET29-PfAcpH plasmid was transformed into BL21 (DE3) cells [Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. *Nat. Methods*. 2012, 9, 981]. The construct was grown in Luria-Bertani (LB) media containing 50 mg/L kanamycin. Expression was induced with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (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 at 1000 relative centrifugal force (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).

C.2. Preparation of PfAcpH affinity resin. Coupling of *Plasmodium falciparum* PfAcpH to Affi-gel 15 (Bio-Rad) resin was conducted in 0.1 M MOPS pH 7.5 and the resin capped with 0.1 M glycine ethyl ester according to the manufacturer's instructions. By Bradford analysis the resin obtained from this procedure contained 4 mg/mL of PfAcpH.

C.3. Expression, modification, and purification of *apo*-actACP. pET28-actACP (C17S) plasmid was transformed into BL21 (DE3) cells [Haushalter, R. W.; Filipp, F. V.; Ko, K. S.; Yu, R.; Opella, S. J.; Burkart, M. D. *ACS Chem. Biol.* 2011, 6, 413]. Uniformly labeled ¹⁵N-actACP was expressed by culturing cells in M9 minimal media (1 L) supplemented with 1 g ¹⁵N-NH₄Cl, 4 g D-glucose, and 50 mg of kanamycin. Uniformly labeled ¹³C, ¹⁵N-actACP was prepared by supplementing M9 minimal media (1 L) with 4 g ¹³C-D-glucose, 1 g ¹⁵N-NH₄Cl, and 50 mg kanamycin. 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 a mixture of *apo*-actACP and *holo*-actACP due to endogenous phosphopantetheinyl transferase activity. It was necessary to "apofy" the actACP mixture by tumbling 10 mL of 300 μ M ACP in 150 mM NaCl, 50 mM Tris pH 8.0, 15 mM MgCl₂, 1 mM MnCl₂, 3 mM tris(2-carboxyethyl)phosphine (TCEP), and 250 μ L of Affi-gel 15 (Bio-Rad) containing 4 mg/mL covalently-attached PfAcpH discussed in section C.2 at 37 °C for 12 h [Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. *Nat. Methods*. 2012, 9, 981]. The resin was spun down by gentle centrifugation (~200 RCF) and pure *apo*-actACP was isolated from the supernatant.

C.4. Preparation of *crypto*-actACPs 4a, 4b, 5b, 6b, 8a, 8b, and 8c. Loading of the polyketide mimetics 11, 20, 22, 23, 30, 31 or 32 was achieved by a previously described one pot chemoenzymatic reaction converting 11, 20, 22, 23, 30, 31 or 32 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*-actACP by Sfp [Worthington, A. S.; Burkart M. D. *Org. Biomol. Chem.* 2006, 4, 44]. In this study, this was achieved by treating 1.5 ml of 400 μ M *apo*-actACP with 12.5 mM MgCl₂, 8 mM ATP, 2 mM mimetic 11, 20, 22, 23, 30, 31 or 32, 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*-actACPs eluted between 155-180 mL.

C.5. Preparation of *holo*-actACP. To prepare pure *holo*-actACP, *apo*-actACP 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 μM Sfp at 37 °C overnight. The solution was purified by FPLC, as described in section C.4.

C.6. HPLC analysis of ACP proteins. Loading of polyketide analogs in sections C.4 and C.5 was monitored by HPLC analysis. The *crypto*-actACP 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% of solvent B over 10 min (solvent A = H₂O, 0.05% trifluoroacetic acid; solvent B = CH₃CN 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.

C.7. Protein NMR experiments. NMR samples were prepared by first concentrating the FPLC pure proteins in 50 mM KPi pH 7.4. A 450 μL aliquot of the appropriate actACP form was then prepared for NMR by adding 50 μL of D₂O, 5 μL of 10% (w/v) NaN₃ and 5 μL of 0.5 M TCEP adjusted to pH 7.3. The final concentration of the *crypto*-labeled actACP evaluated was as: 1 mM for **4a**, **4b**, **5b**, **6b**, **8a**, **8b**, and **8c**. After adding the sample, each NMR tube was flushed with argon, capped, and sealed with Teflon tape. All NMR spectra were acquired at 37 °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 observed for the *holo*-actACP were in good agreement with the literature [Haushalter, R. W.; Filipp, F. V.; Ko, K. S.; Yu, R.; Opella, S. J.; Burkart, M. D. *ACS Chem. Biol.* **2011**, *6*, 413]. Chemical shifts were measured by peak maxima, and chemical shift perturbation was calculated using the formula $CSP = \{[(0.2\delta N)^2 + (\delta H)^2]/2\}^{0.5}$.

Because significant but similar peak perturbations were observed for **5b** and **6b**, a ¹³C,¹⁵N-labeled sample of **5b** was 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), and analyzed using the Sparky (UC San Francisco) and CARA software suites.

C.8. In Silico Docking experiments. The docking program GOLD was used for docking between the ActACP and the phosphopantetheine-tethered atom replacement probes. Both protein and ligands were prepared for docking by removing waters, adding hydrogens, charges, and converting the pdb files to Mol2 files using the program Chimera 2. The Act ACP ligand-binding pocket was defined as residues within 10 Å of the hydroxyl group of the active site residue Ser 42. A 1.8 Å restraint was placed between the phosphate group from the pantetheine moiety and the active site Ser 42. Docking was performed using the default settings with 100 docking trials. The docking solutions were ranked using the ChemPLP scoring functions.

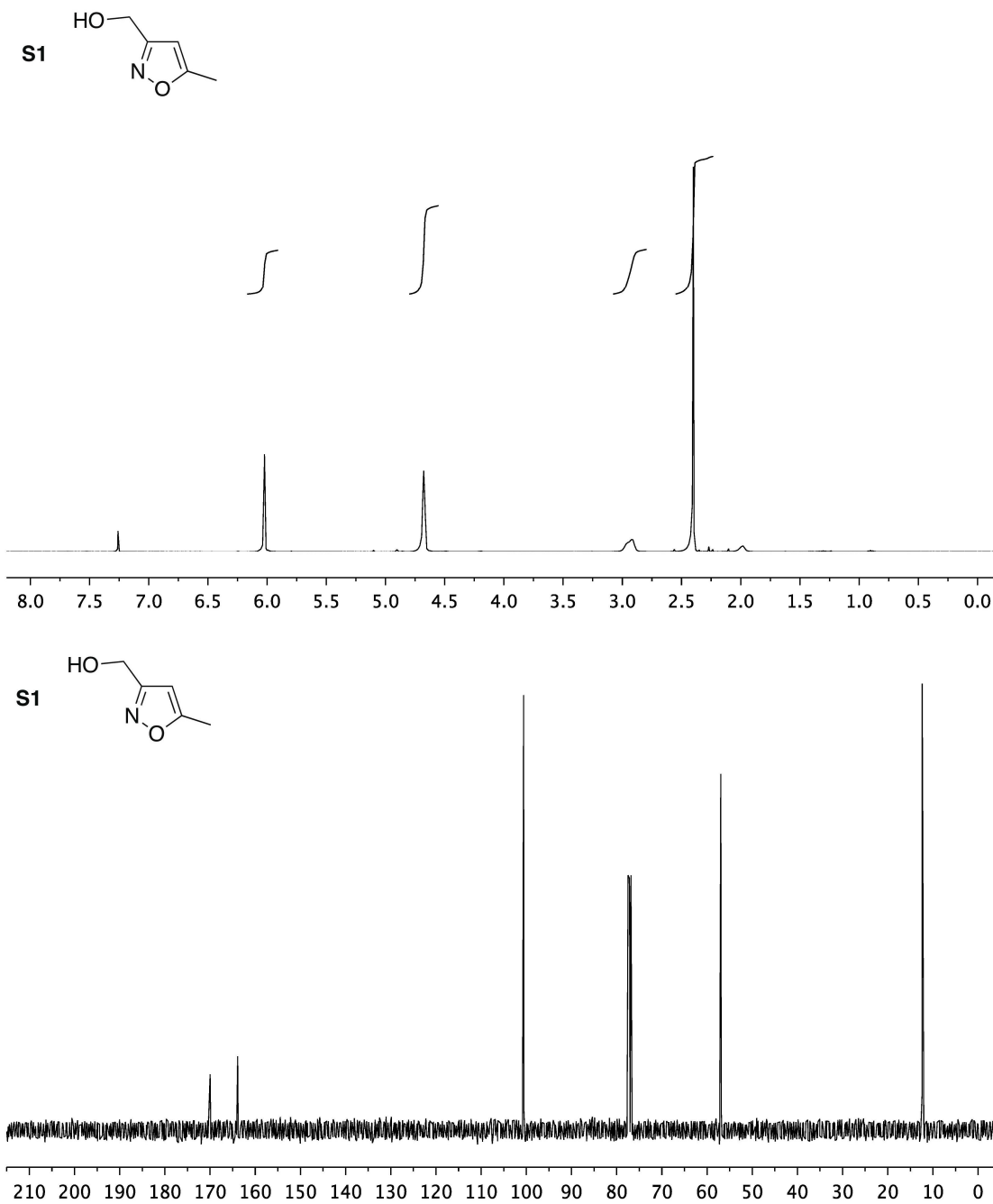
Figure S1. ^1H NMR (400 Mhz) and ^{13}C NMR (100 MHz) spectra of **S1** in CDCl_3 

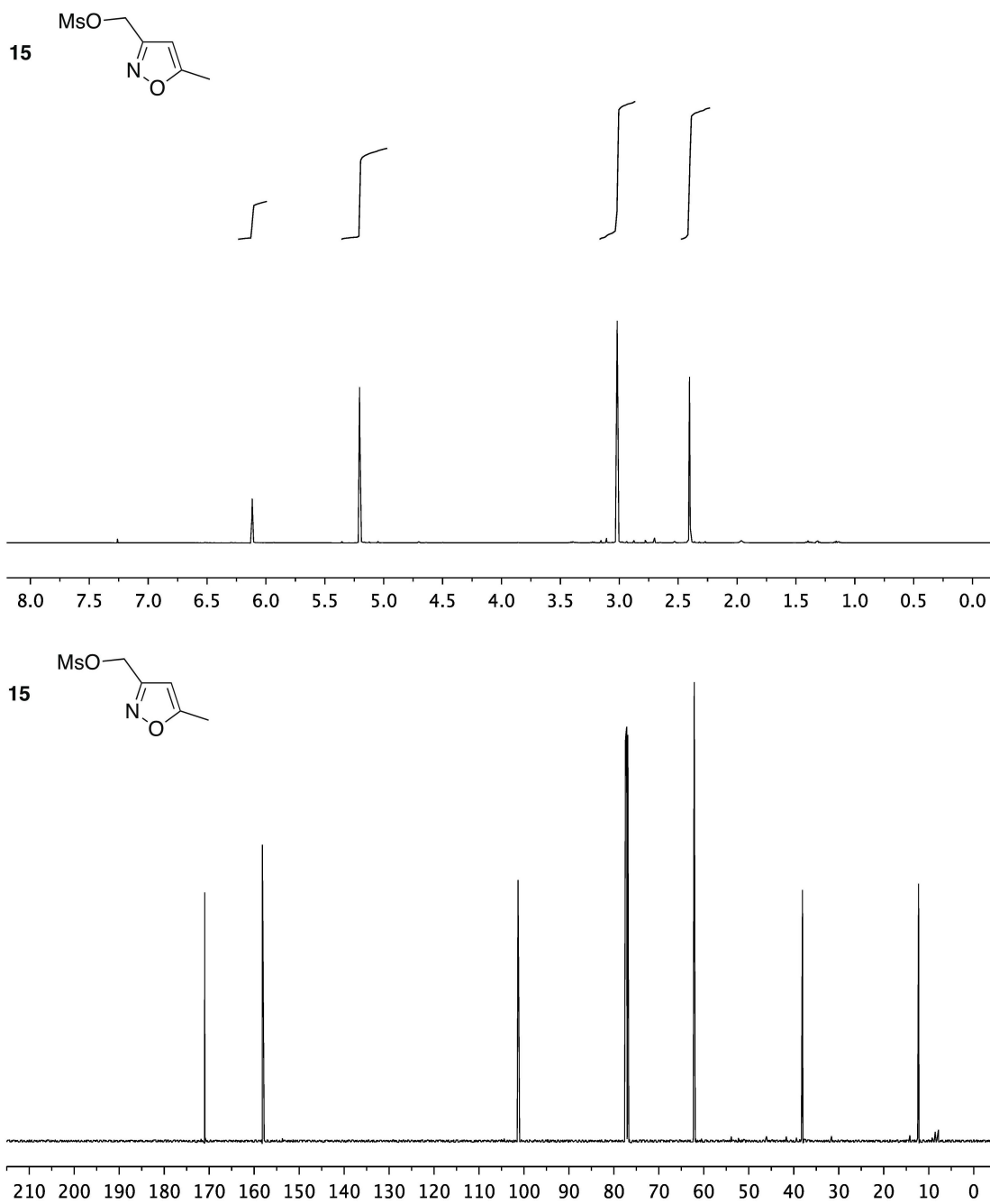
Figure S2. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **15** in CDCl_3 

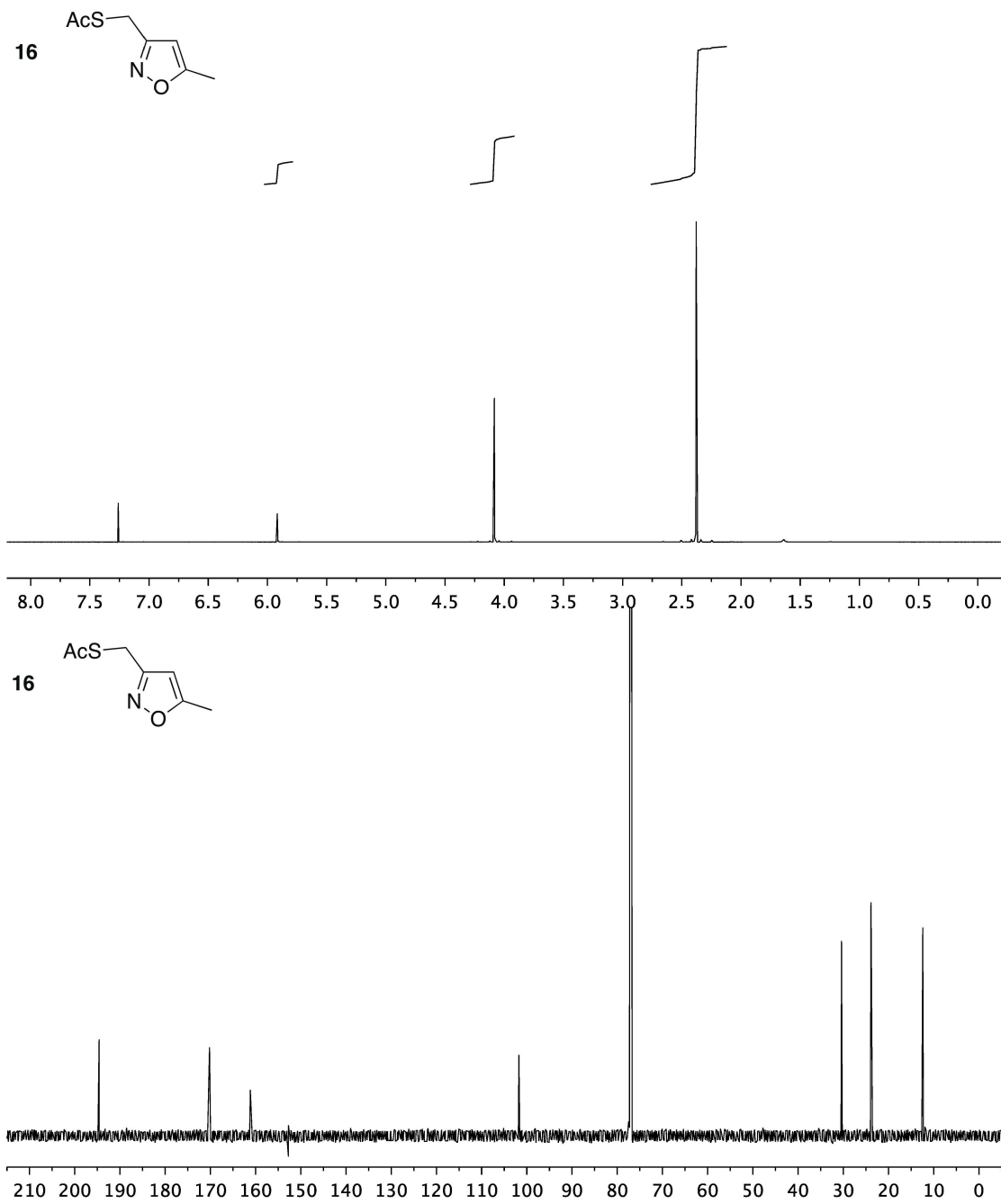
Figure S3. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **16** in CDCl_3 

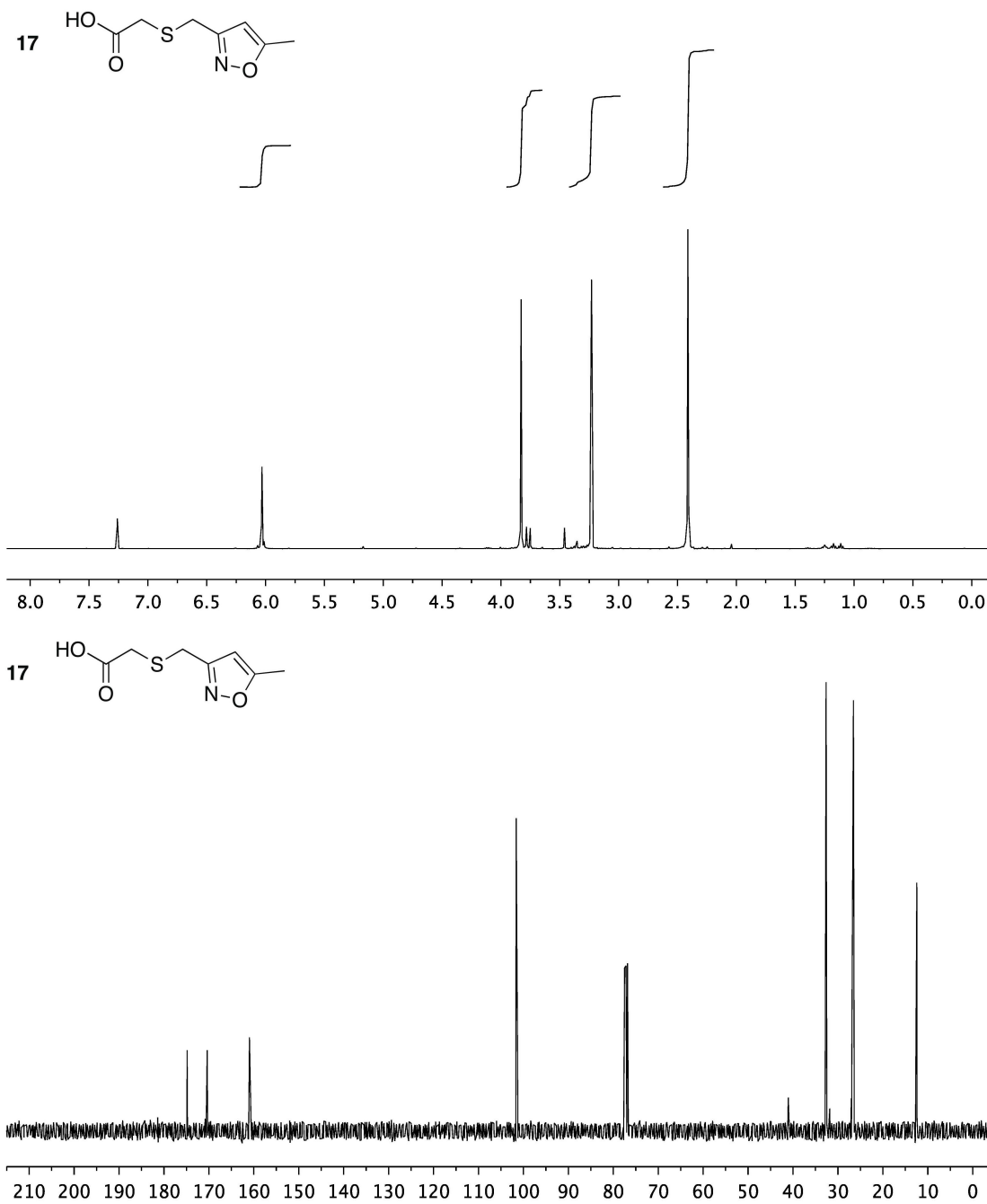
Figure S4. ^1H NMR (400 Mhz) and ^{13}C NMR (100 MHz) spectra of **17** in CDCl_3 

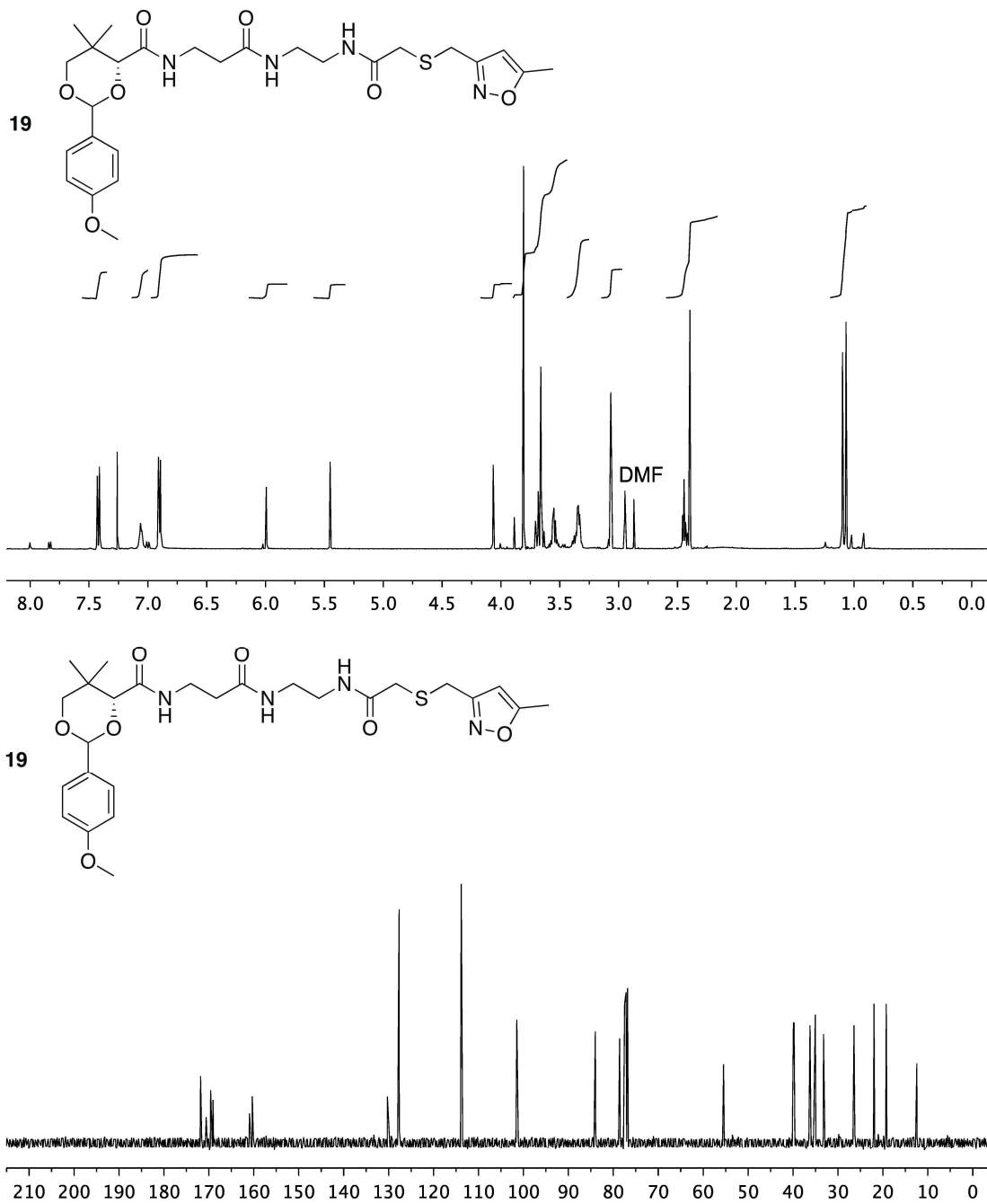
Figure S5. ^1H NMR (500 Mhz) and ^{13}C NMR (100 MHz) spectra of **19** in CDCl_3 

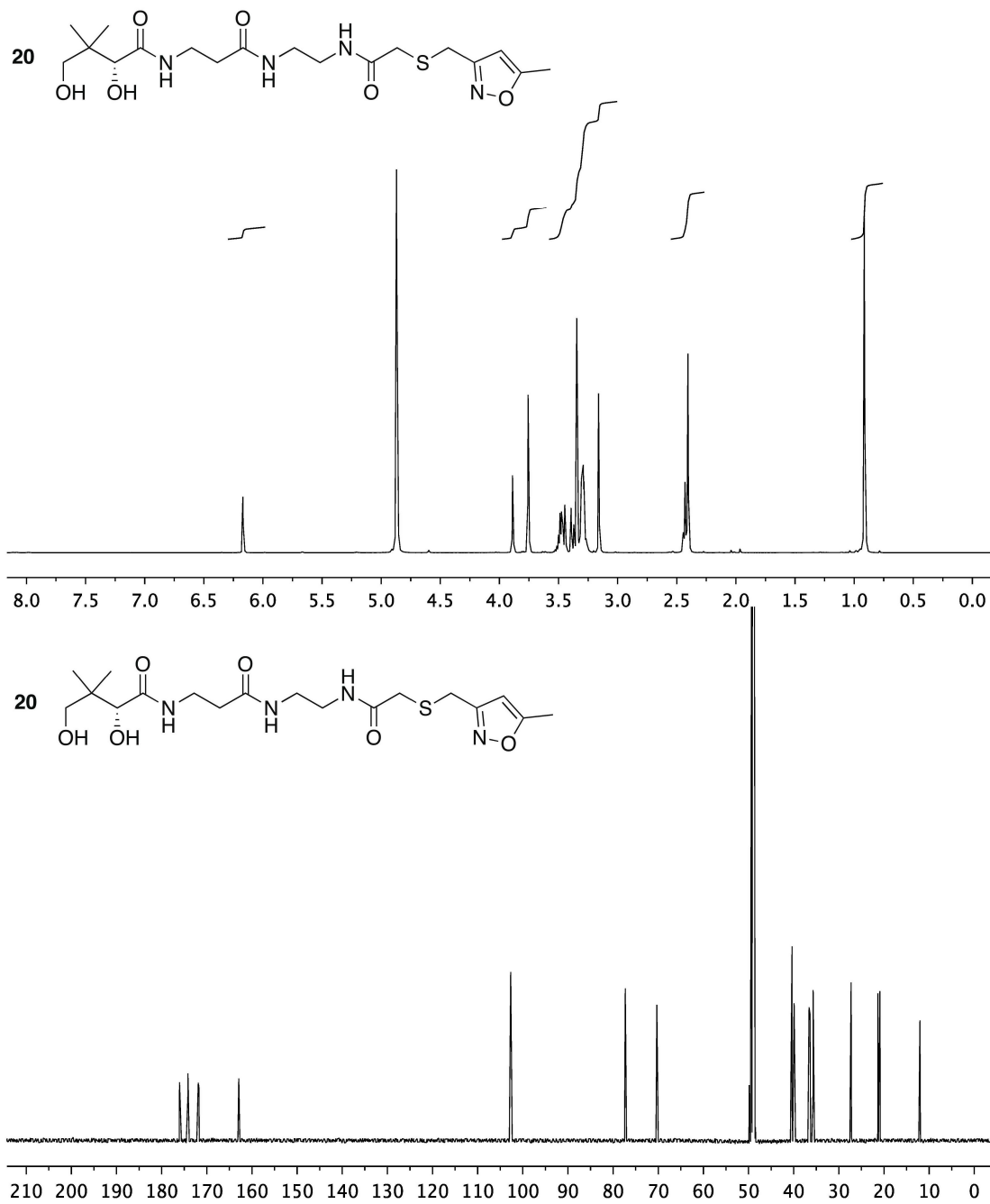
Figure S6. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **20** in CD_3OD 

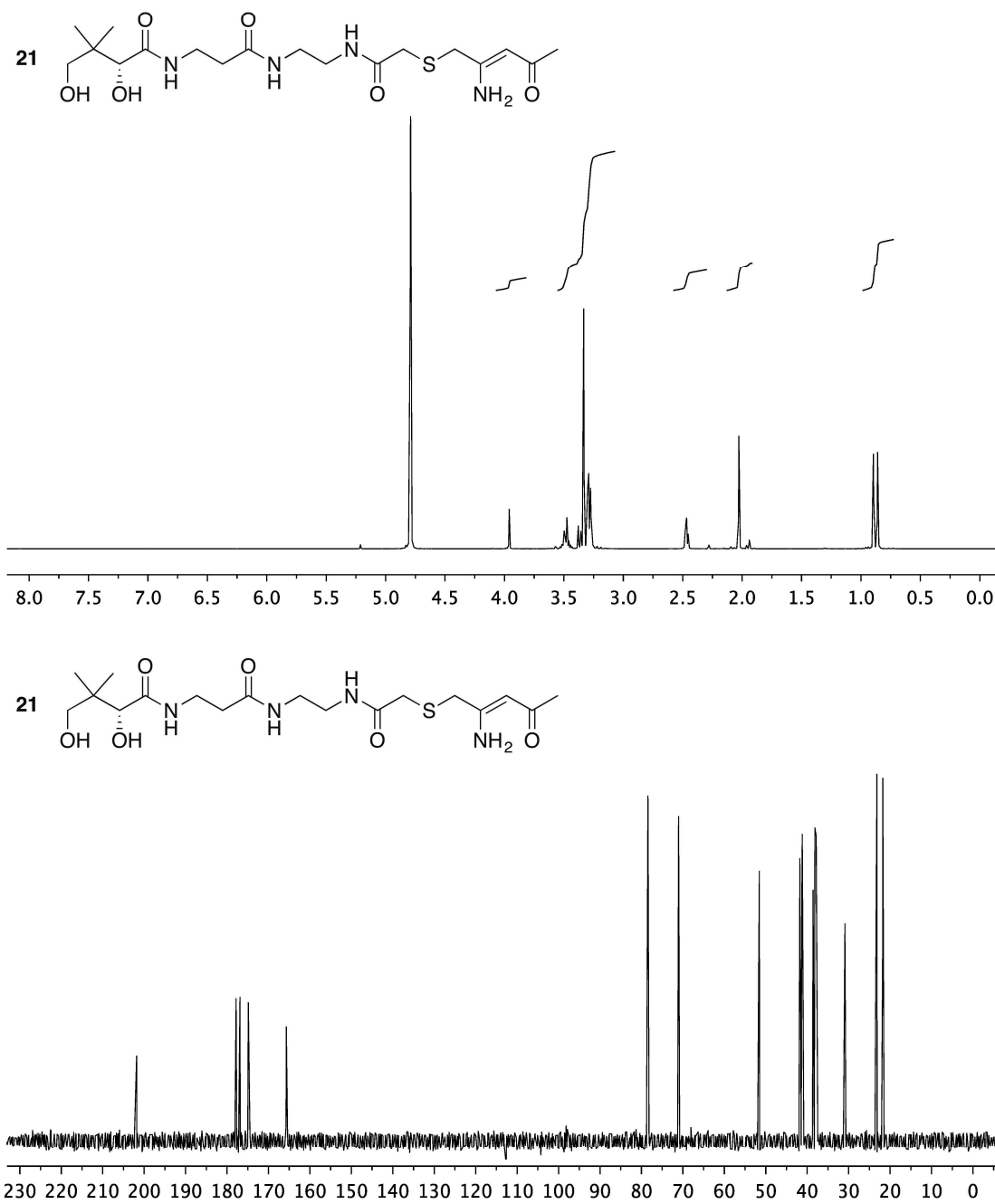
Figure S7. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **21** in D_2O 

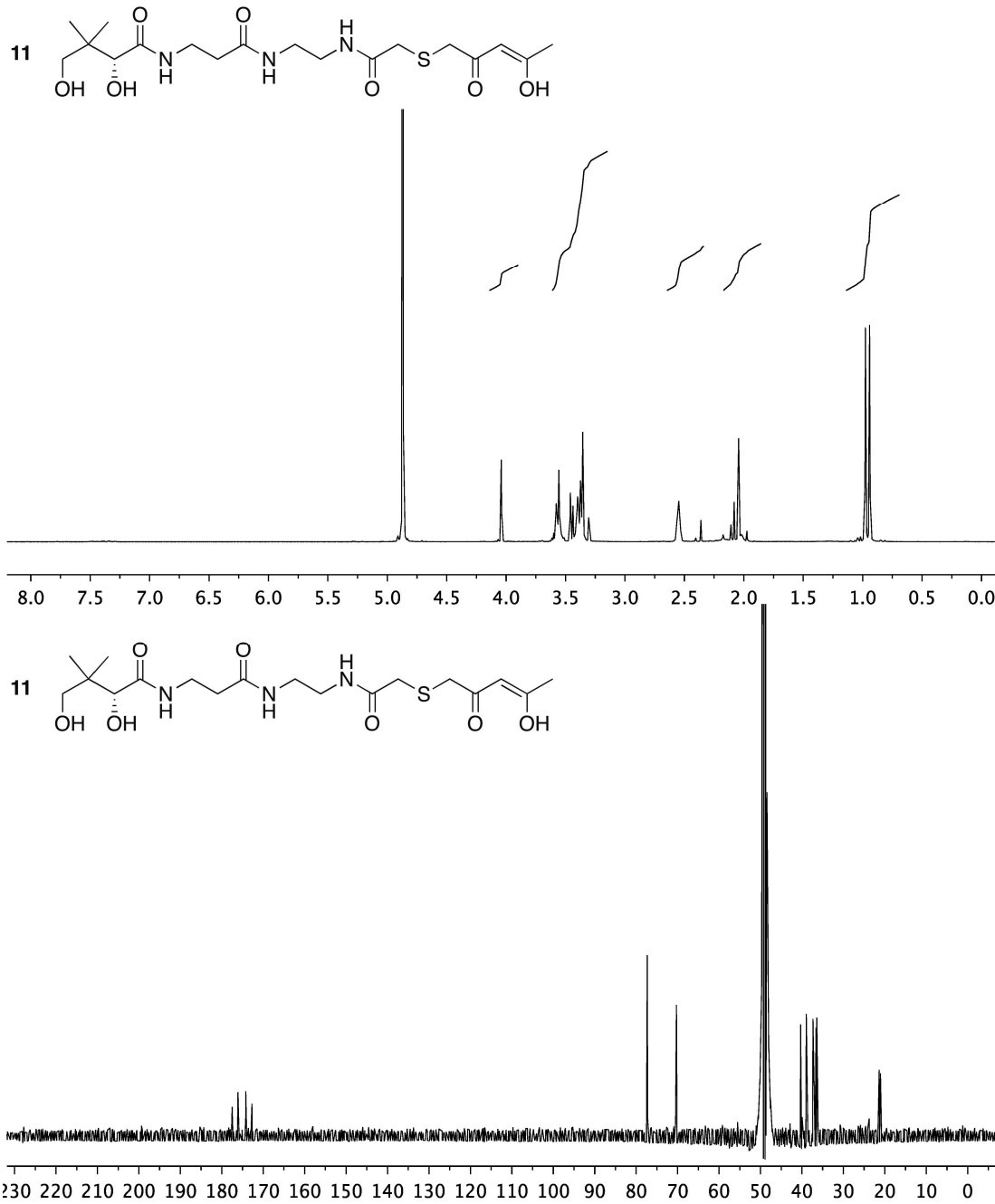
Figure S8. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **11** in CD_3OD 

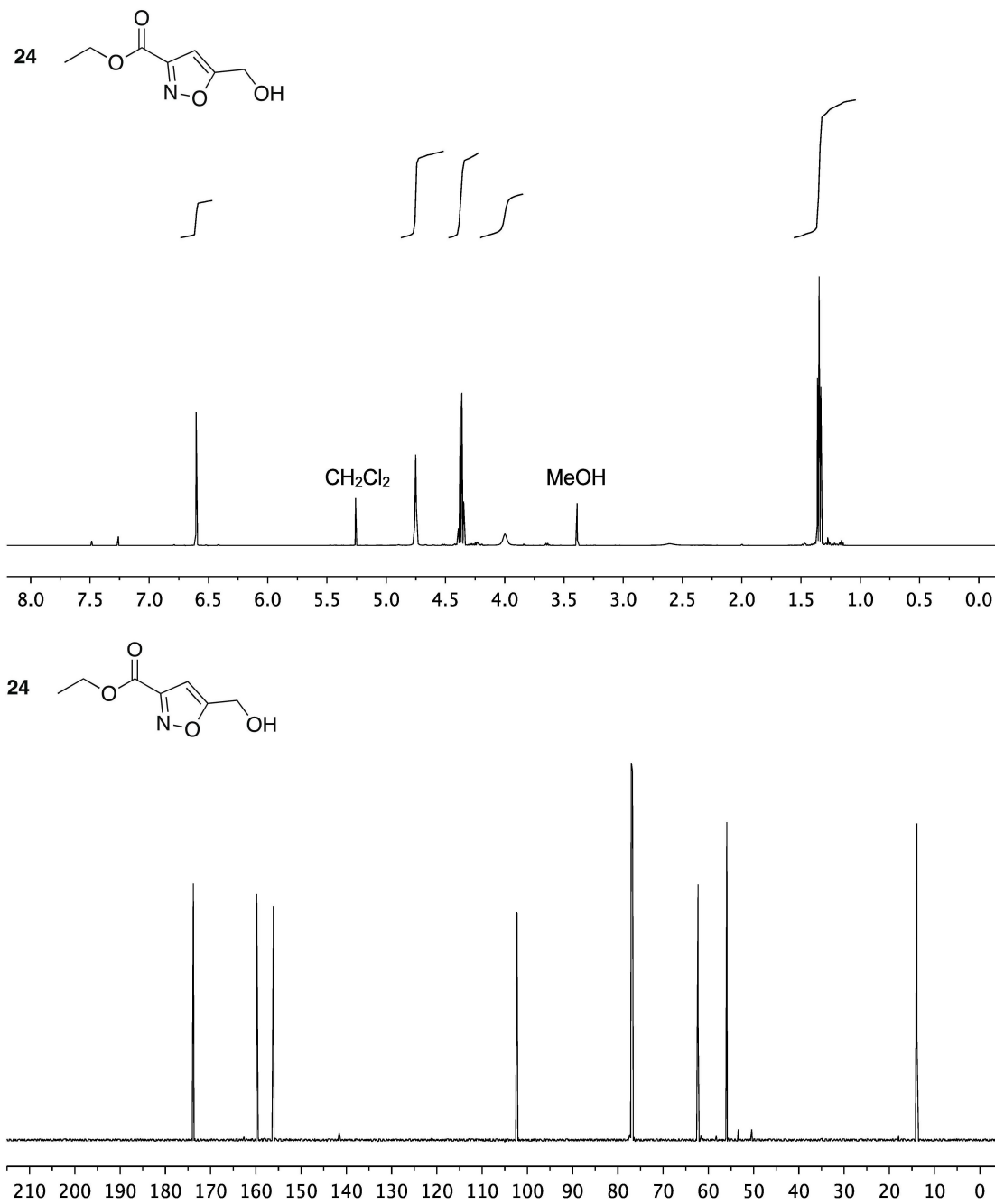
Figure S9. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **24** in CDCl_3 

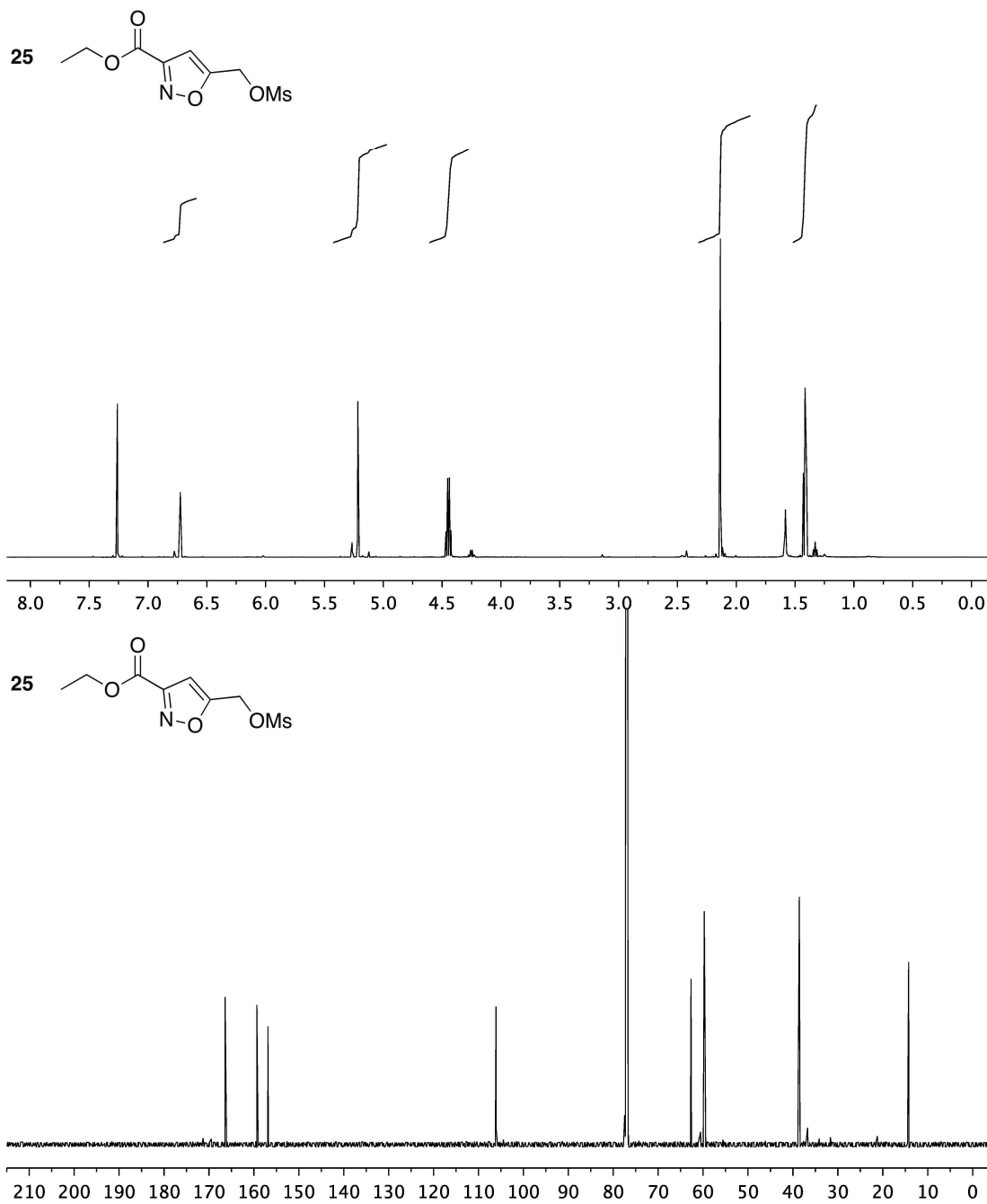
Figure S10. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **25** in CDCl_3 

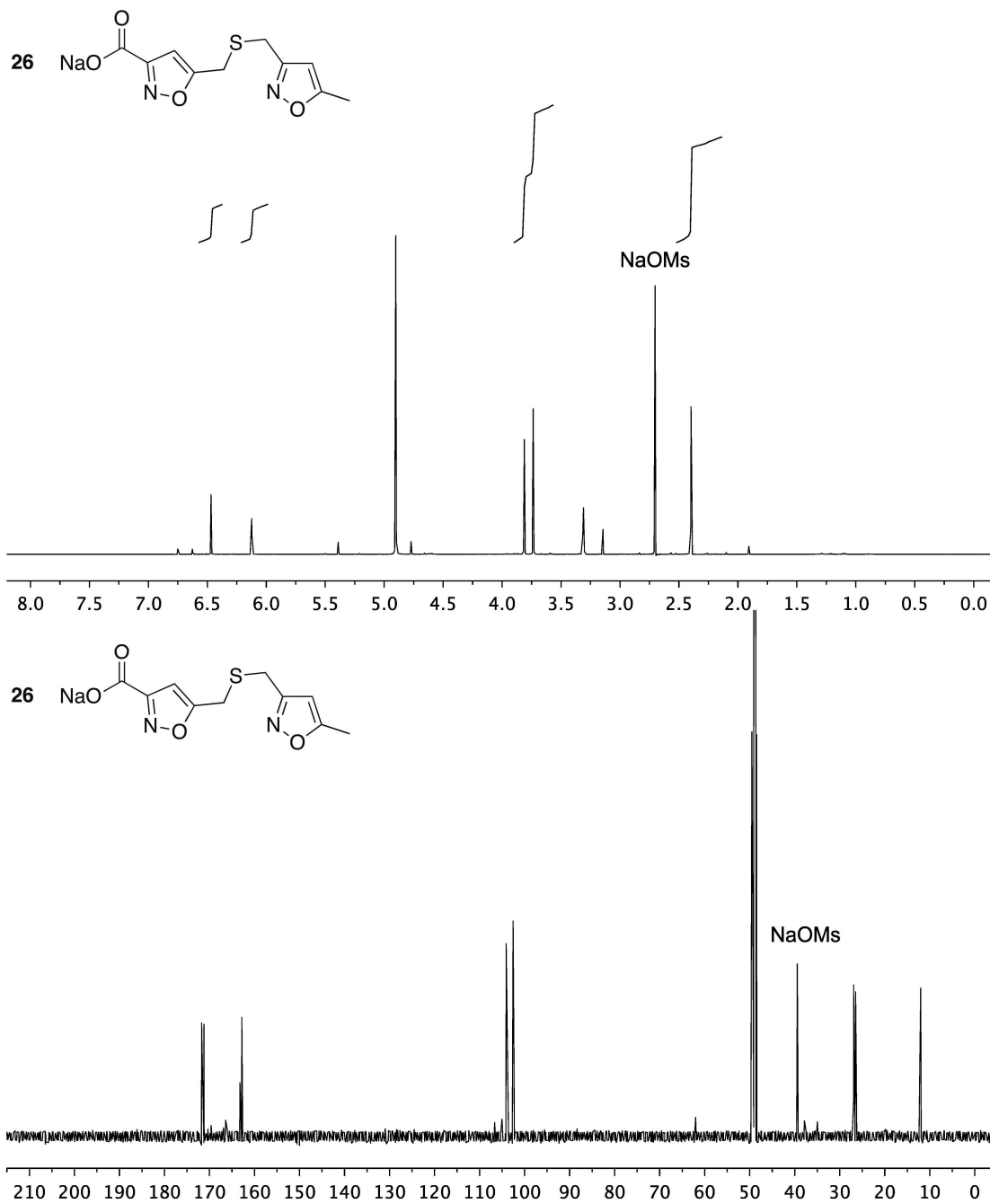
Figure S11. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **26** in CD_3OD 

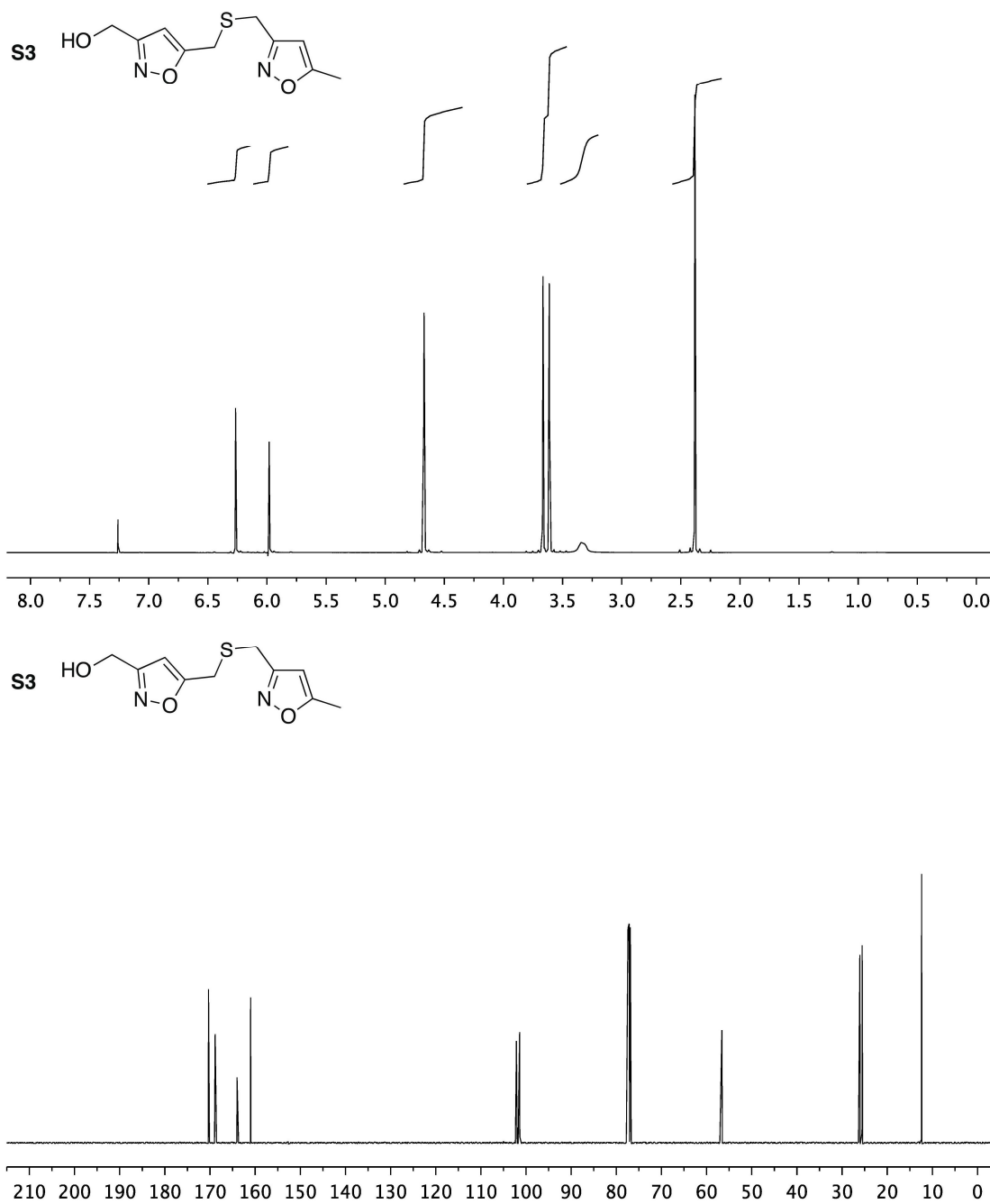
Figure S12. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **S3** in CDCl_3 

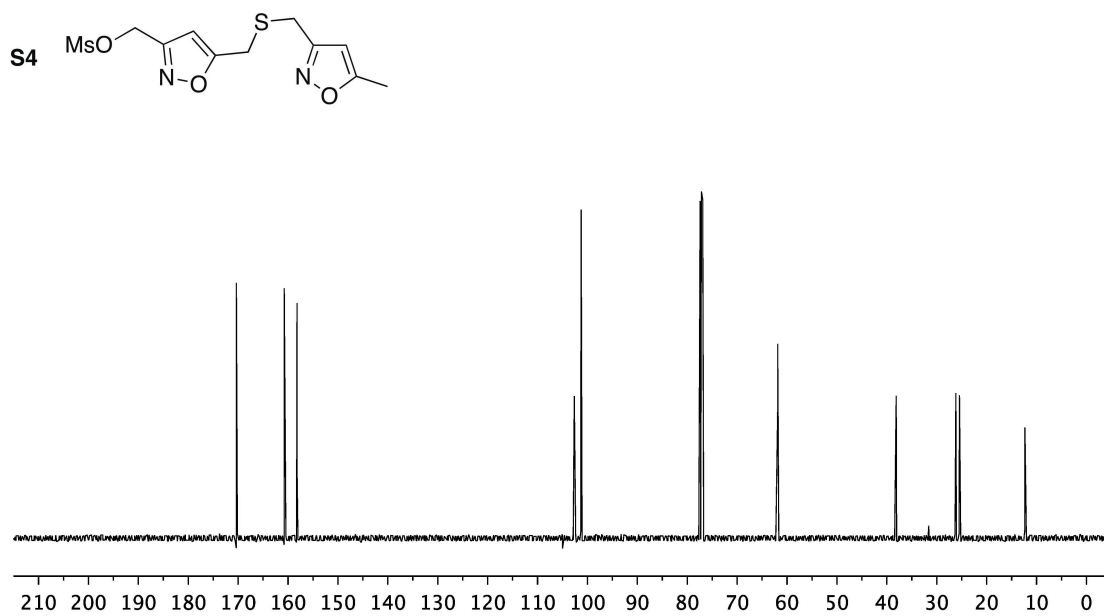
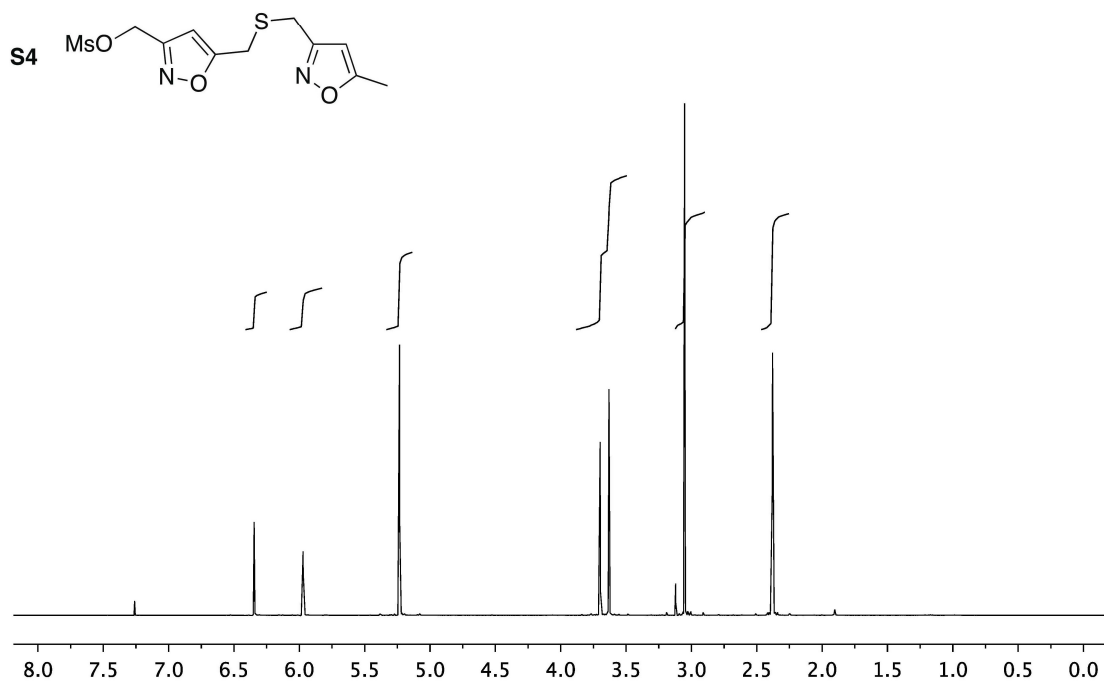
Figure S13. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **S4** in CDCl_3 

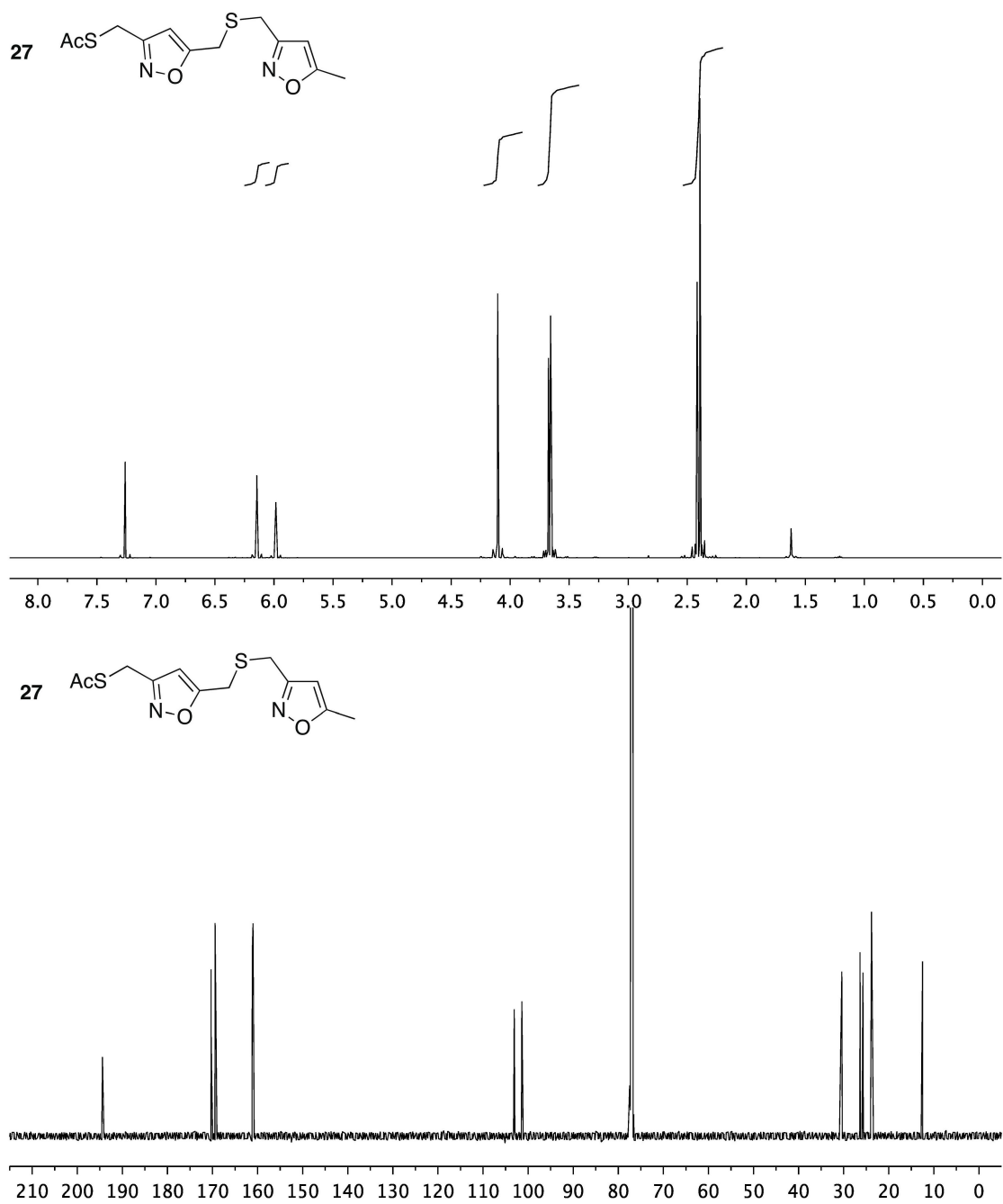
Figure S14. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of 27 in CDCl_3 

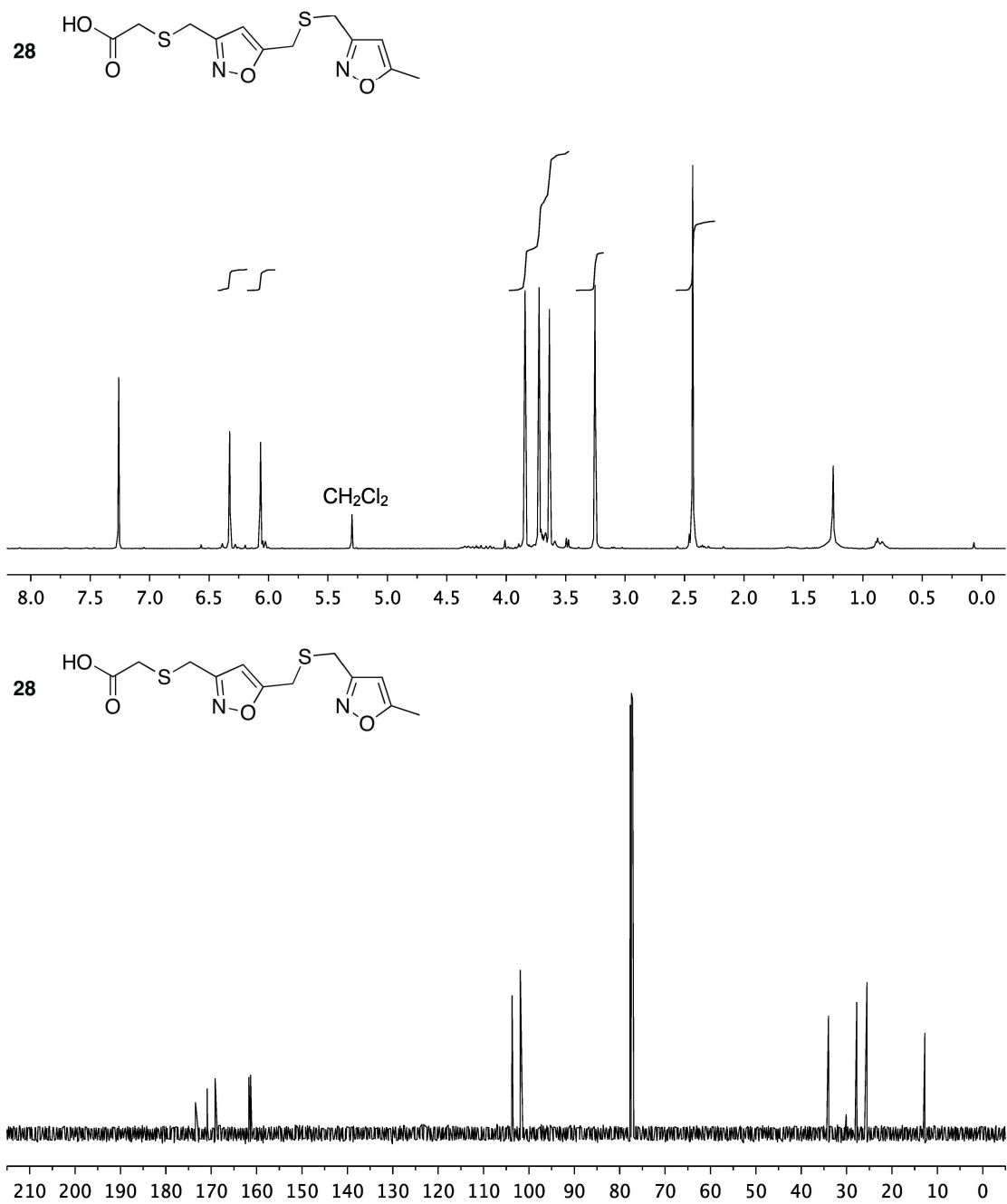
Figure S15. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **28** in CDCl_3 

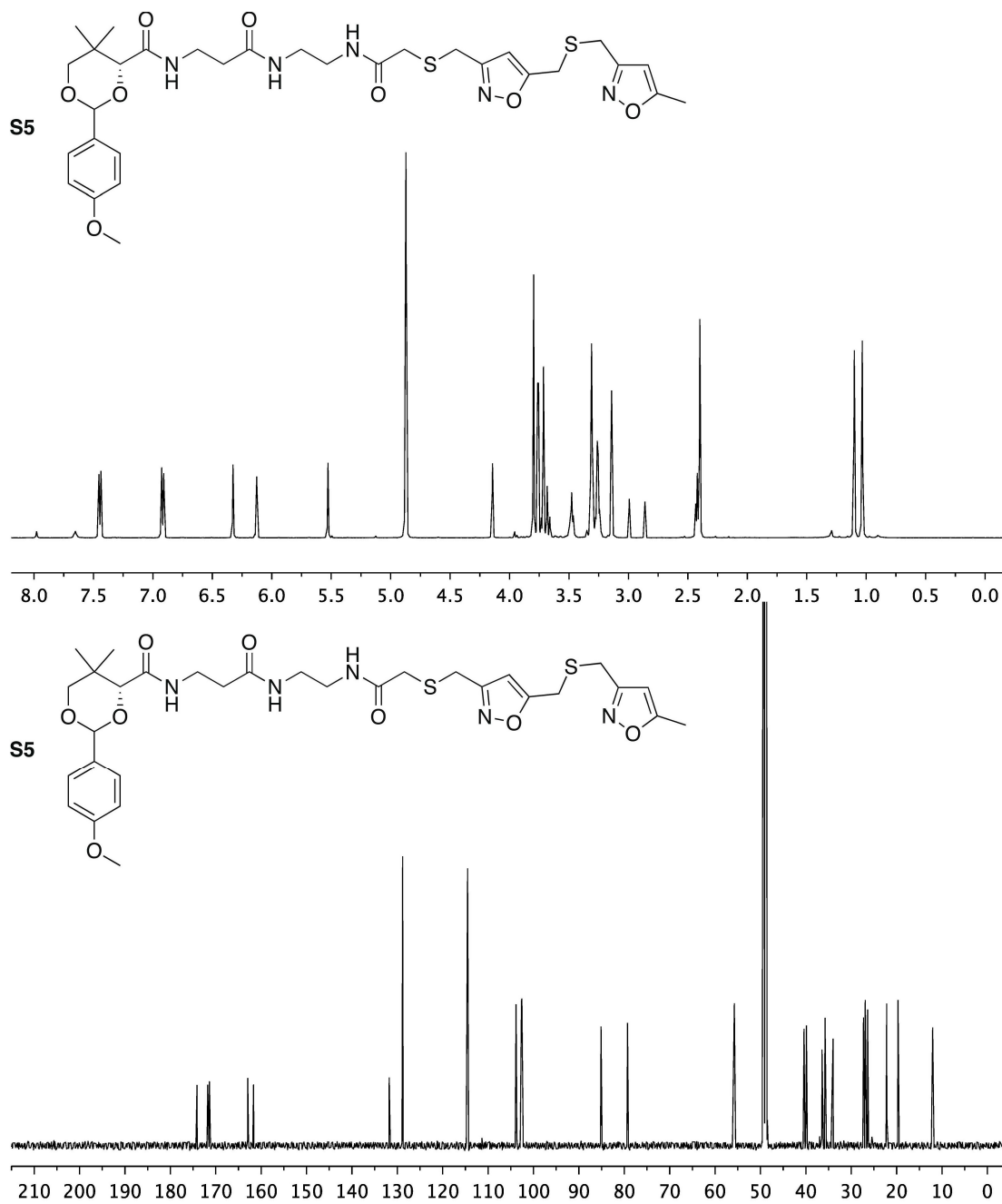
Figure S16. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **S5** in CD_3OD 

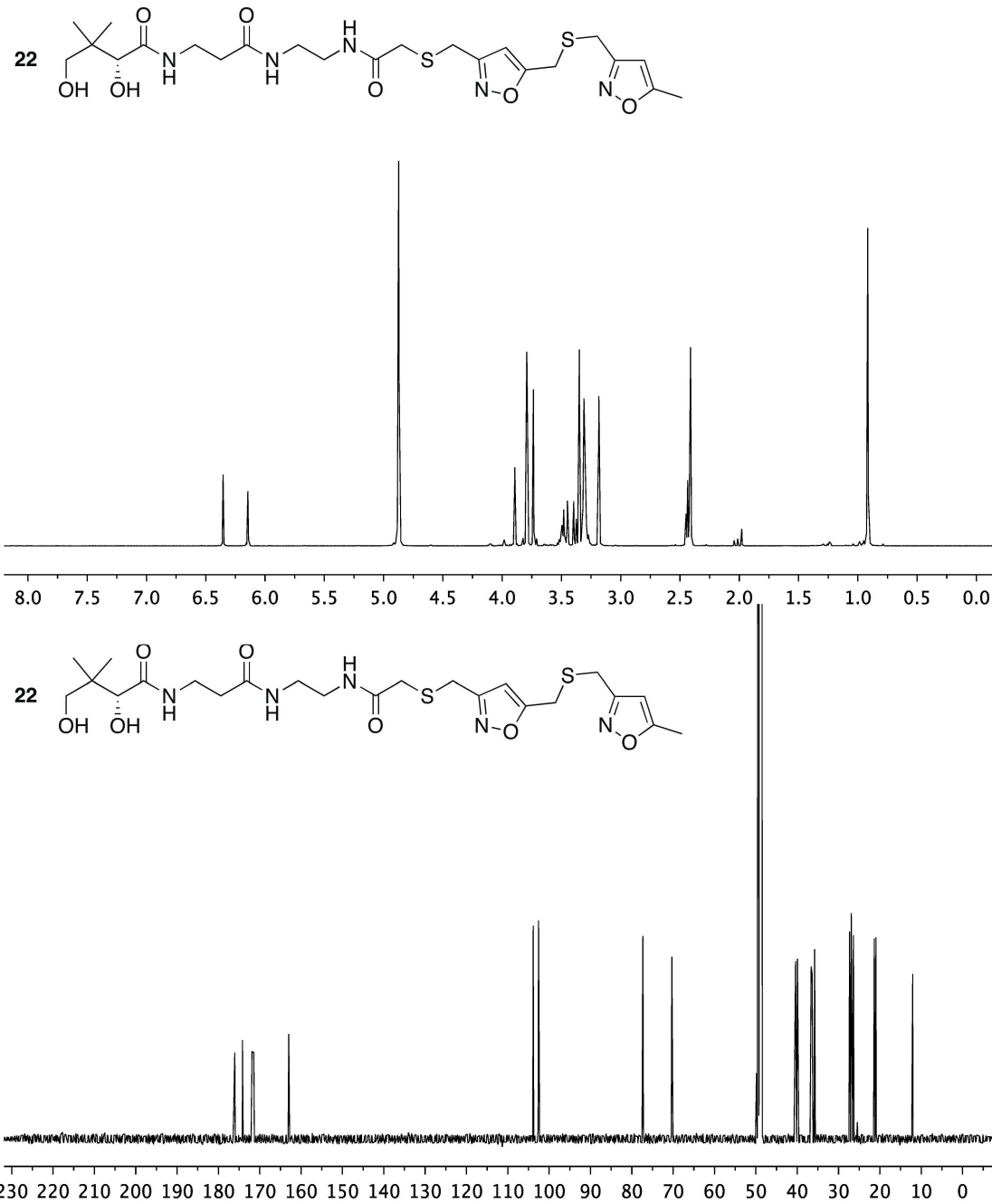
Figure S17. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **22** in CD_3OD 

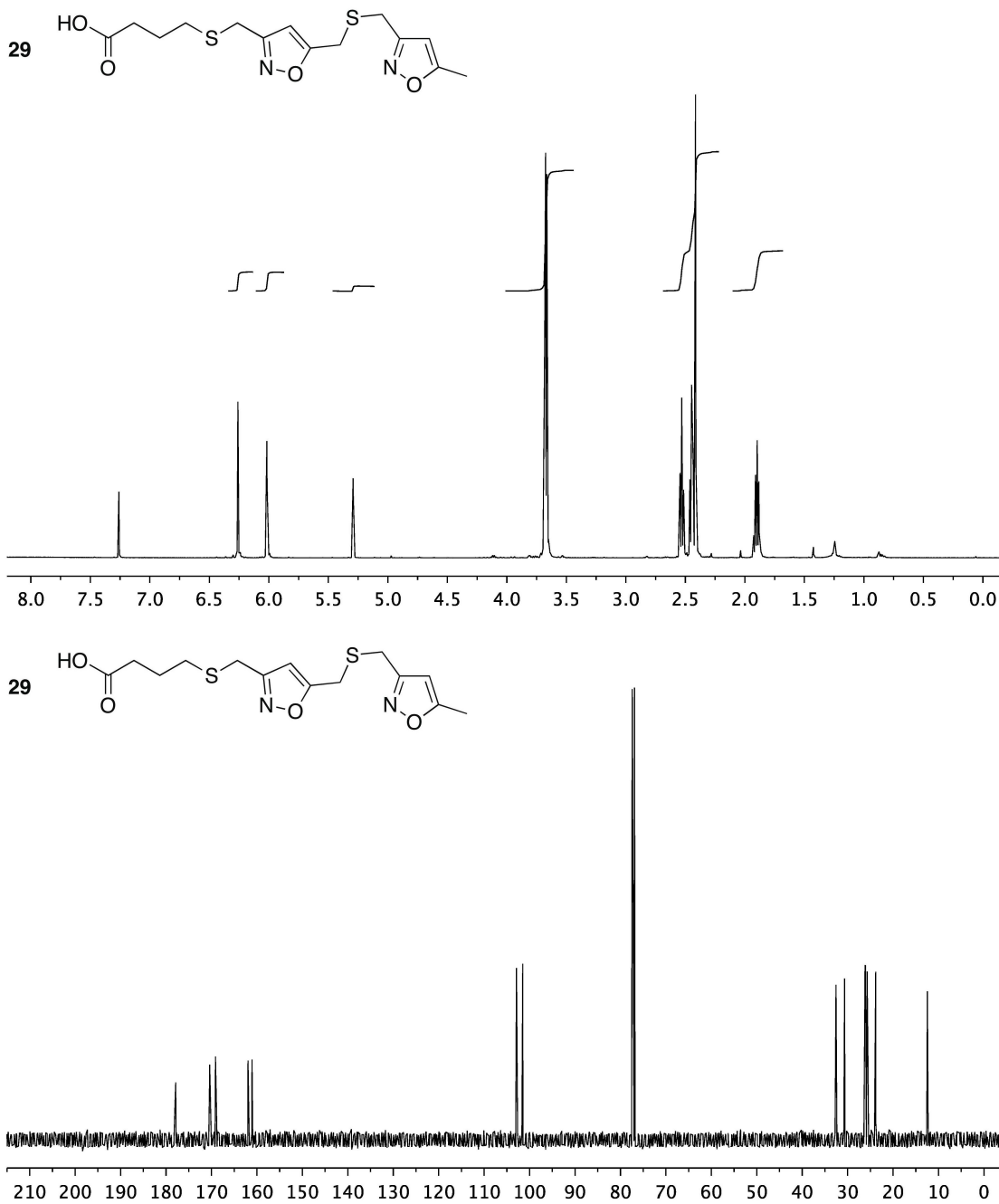
Figure S18. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **29** in CDCl_3 

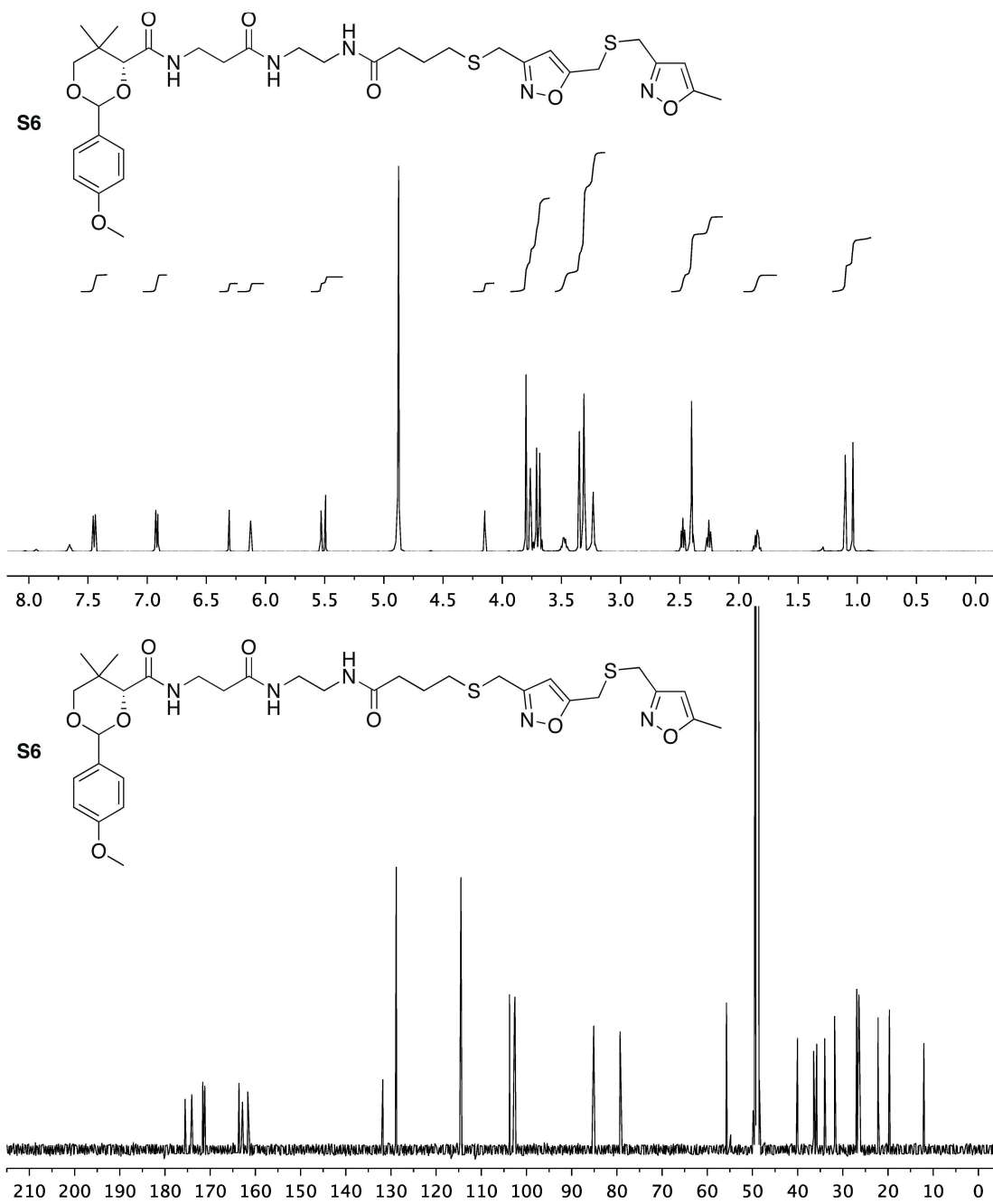
Figure S19. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **S6** in CD_3OD 

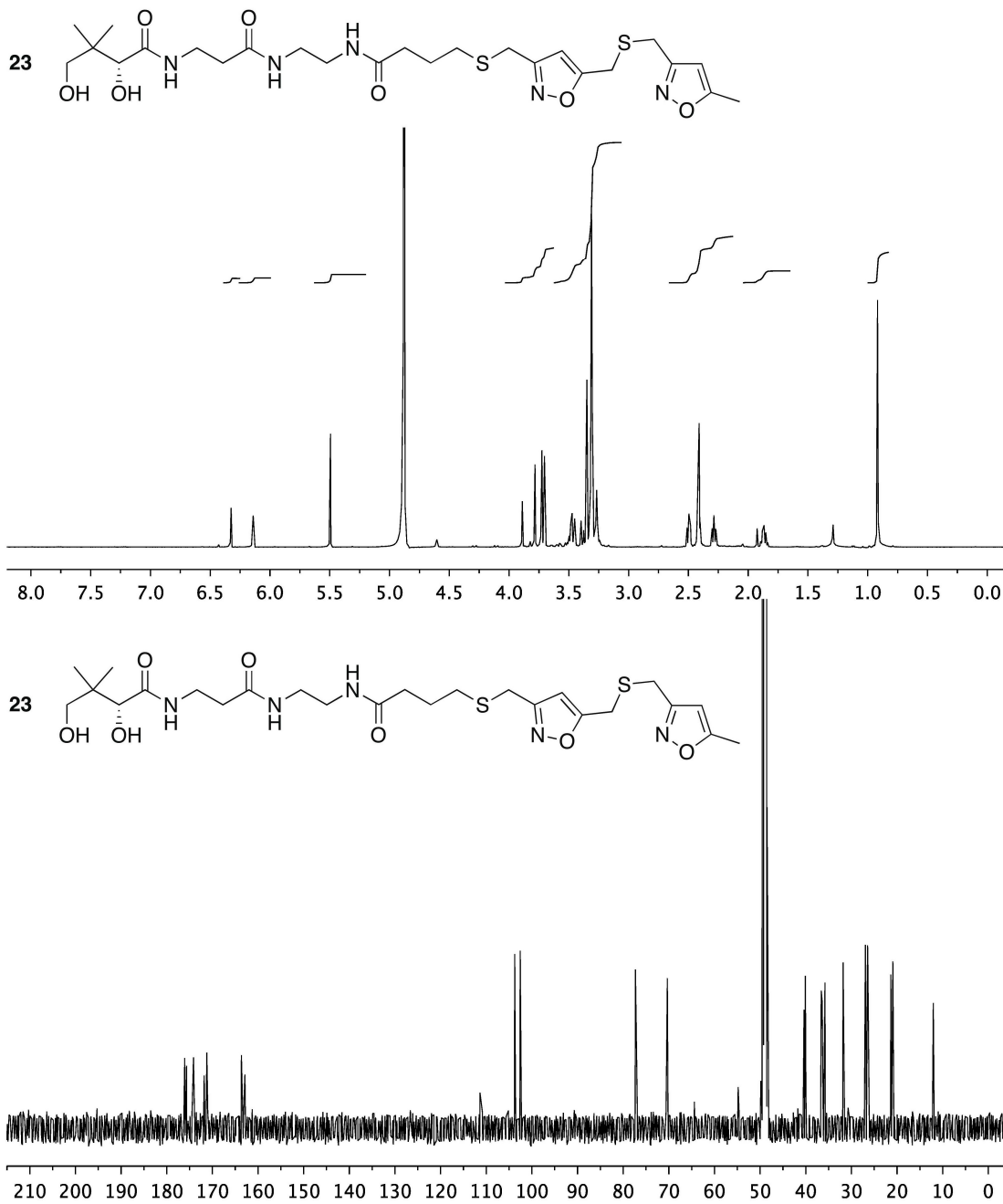
Figure S20. ^1H NMR (500 Mhz) and ^{13}C NMR (125 MHz) spectra of **23** in CD_3OD 

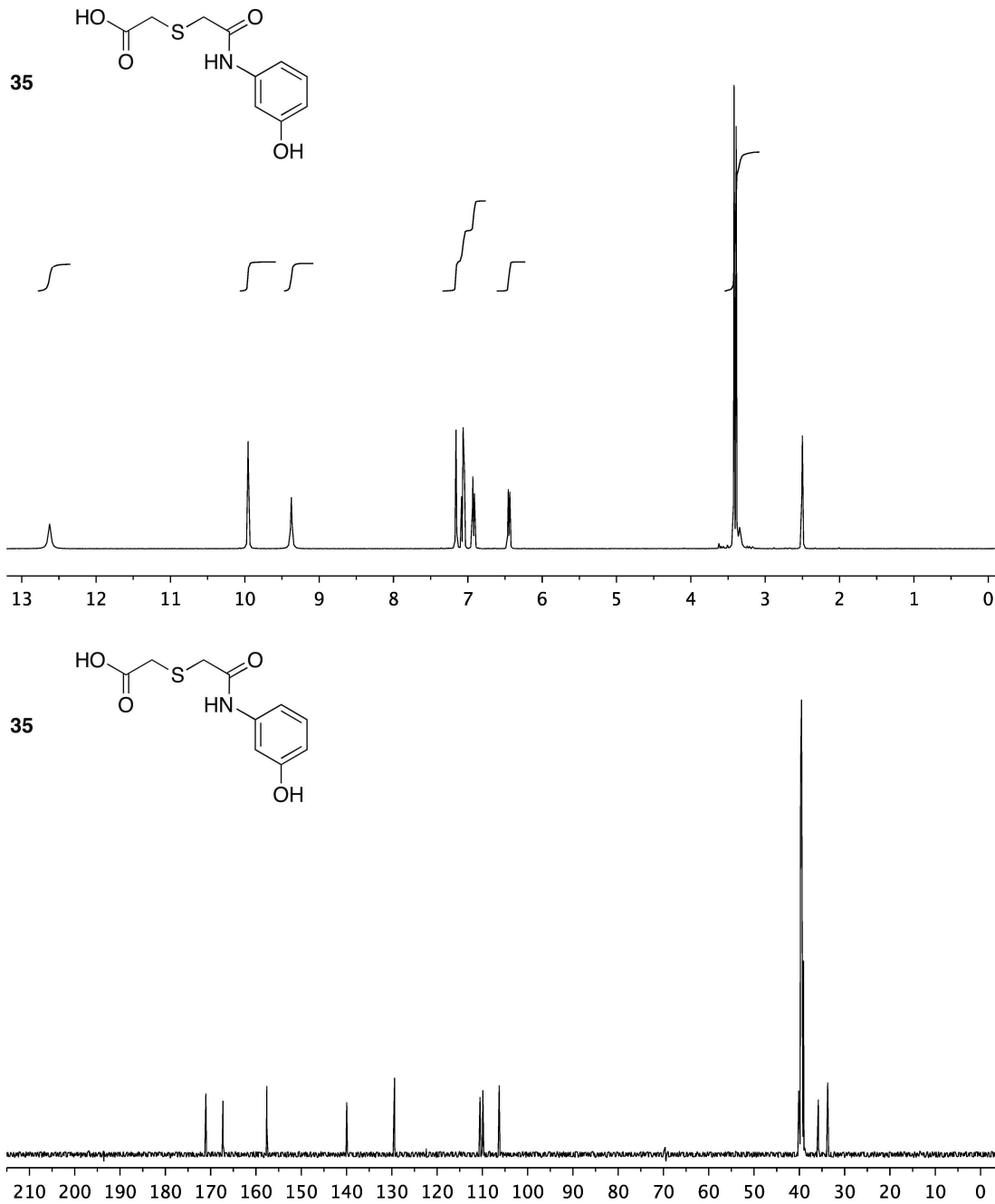
Figure S21. $^1\text{H-NMR}$ (400 MHz) and $^{13}\text{C-NMR}$ (100 MHz) spectra of **35** in $\text{DMSO-}d_6$ 

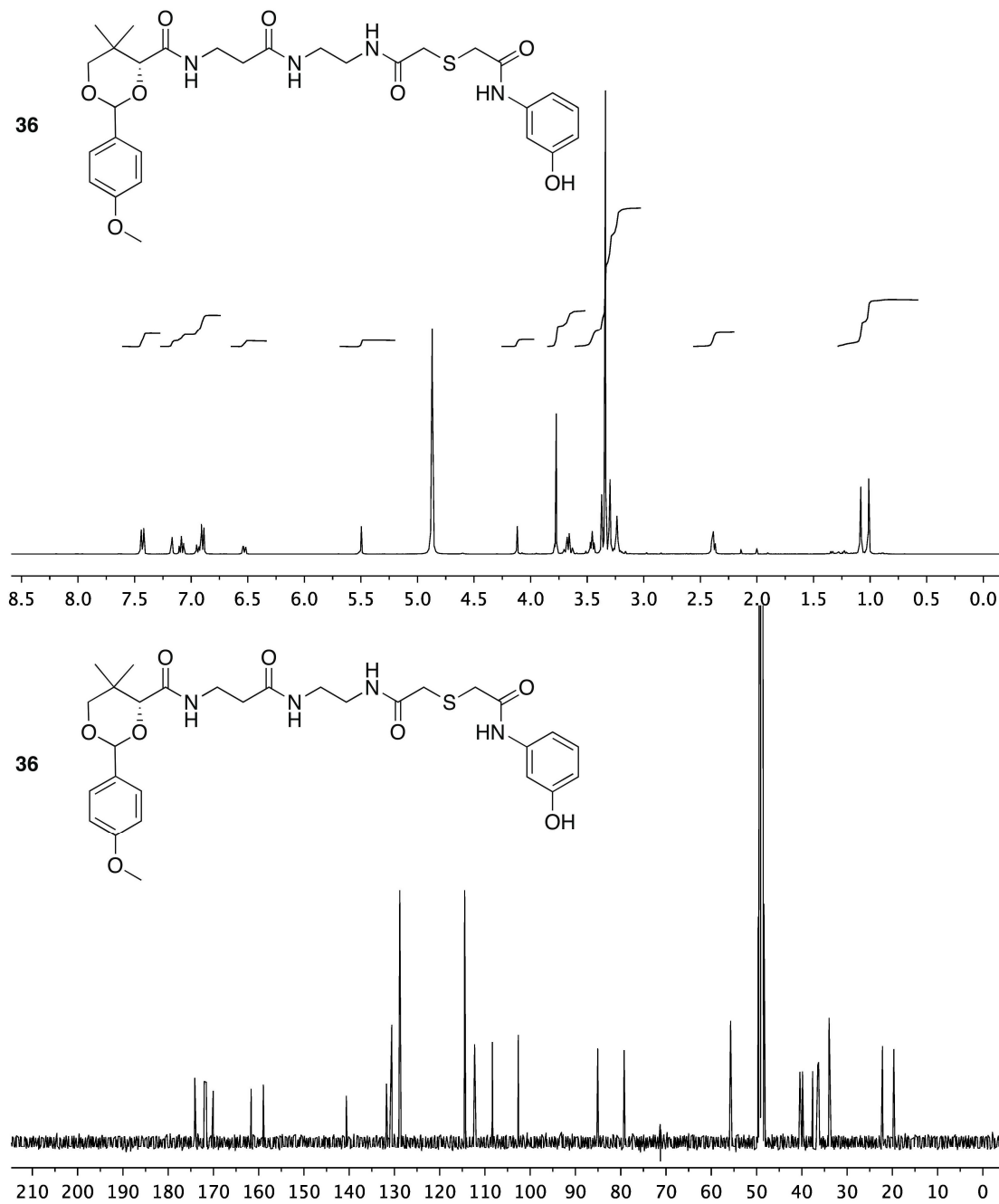
Figure S22. $^1\text{H-NMR}$ (400 MHz) and $^{13}\text{C-NMR}$ (100 MHz) spectra of **36** in CD_3OD 

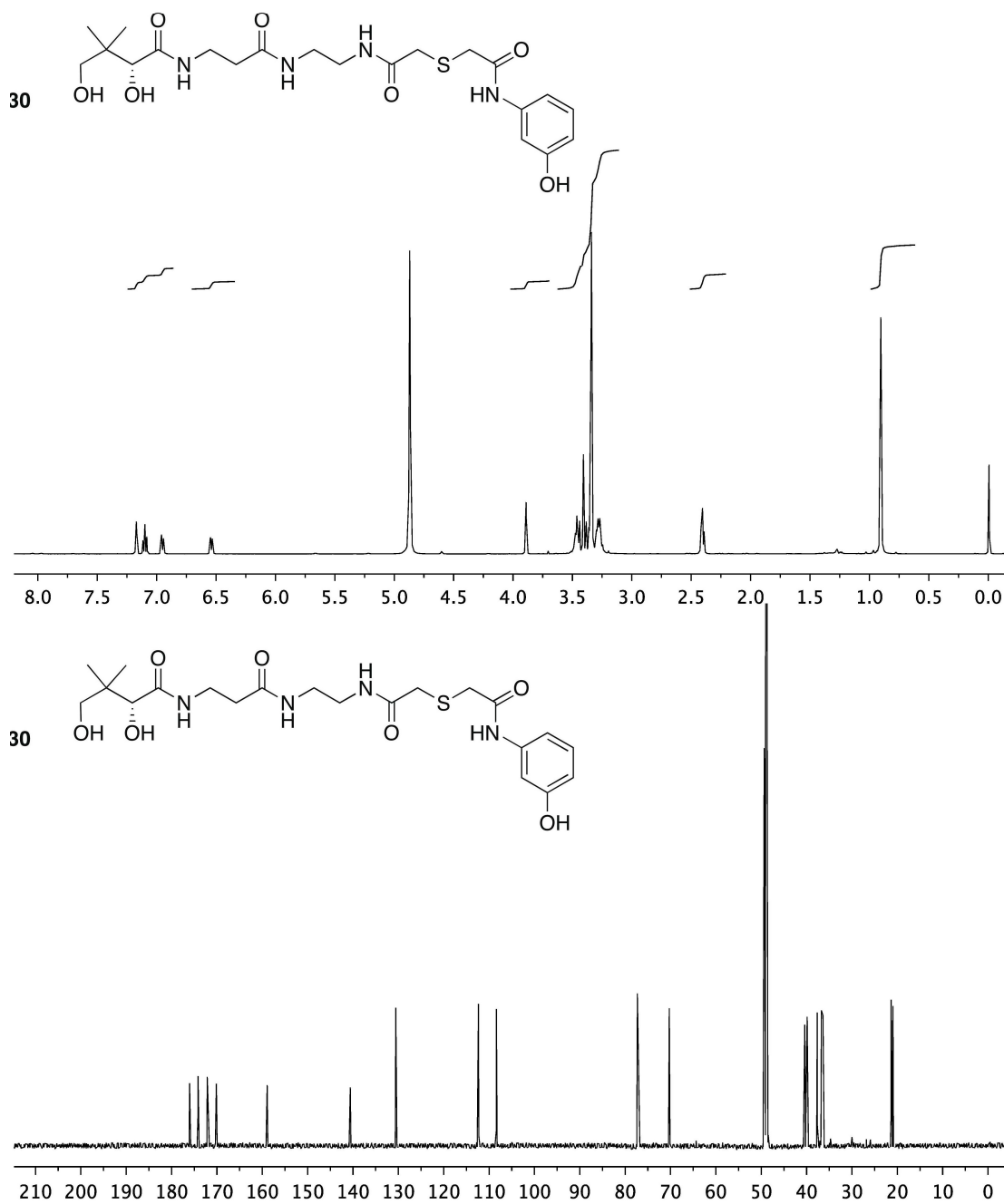
Figure S23. $^1\text{H-NMR}$ (500 MHz) and $^{13}\text{C-NMR}$ (125 MHz) spectra of **30** in CD_3OD 

Figure S24. ^1H -NMR (500 MHz) and ^{13}C -NMR (125 MHz) spectra of **38** in CDCl_3

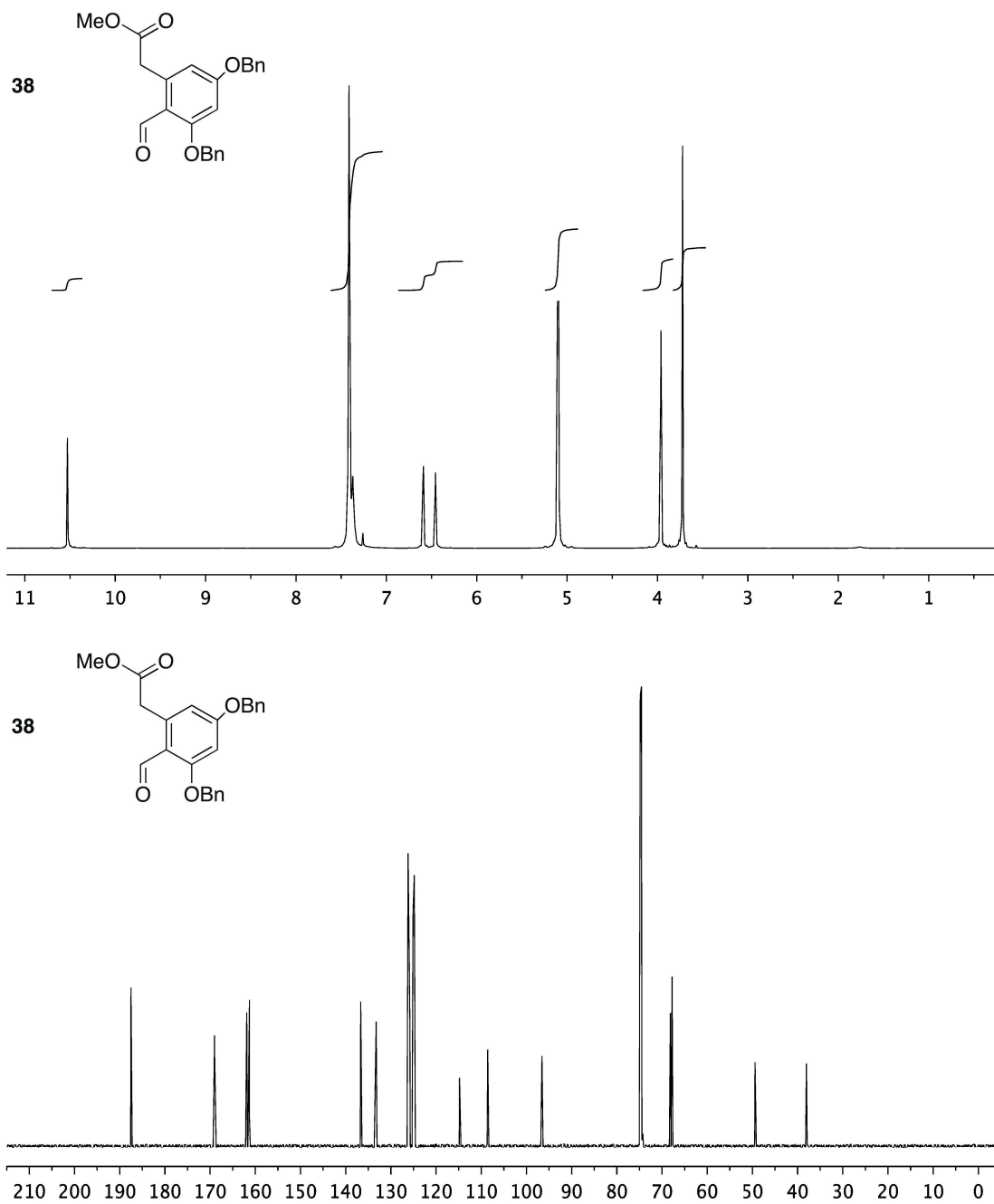


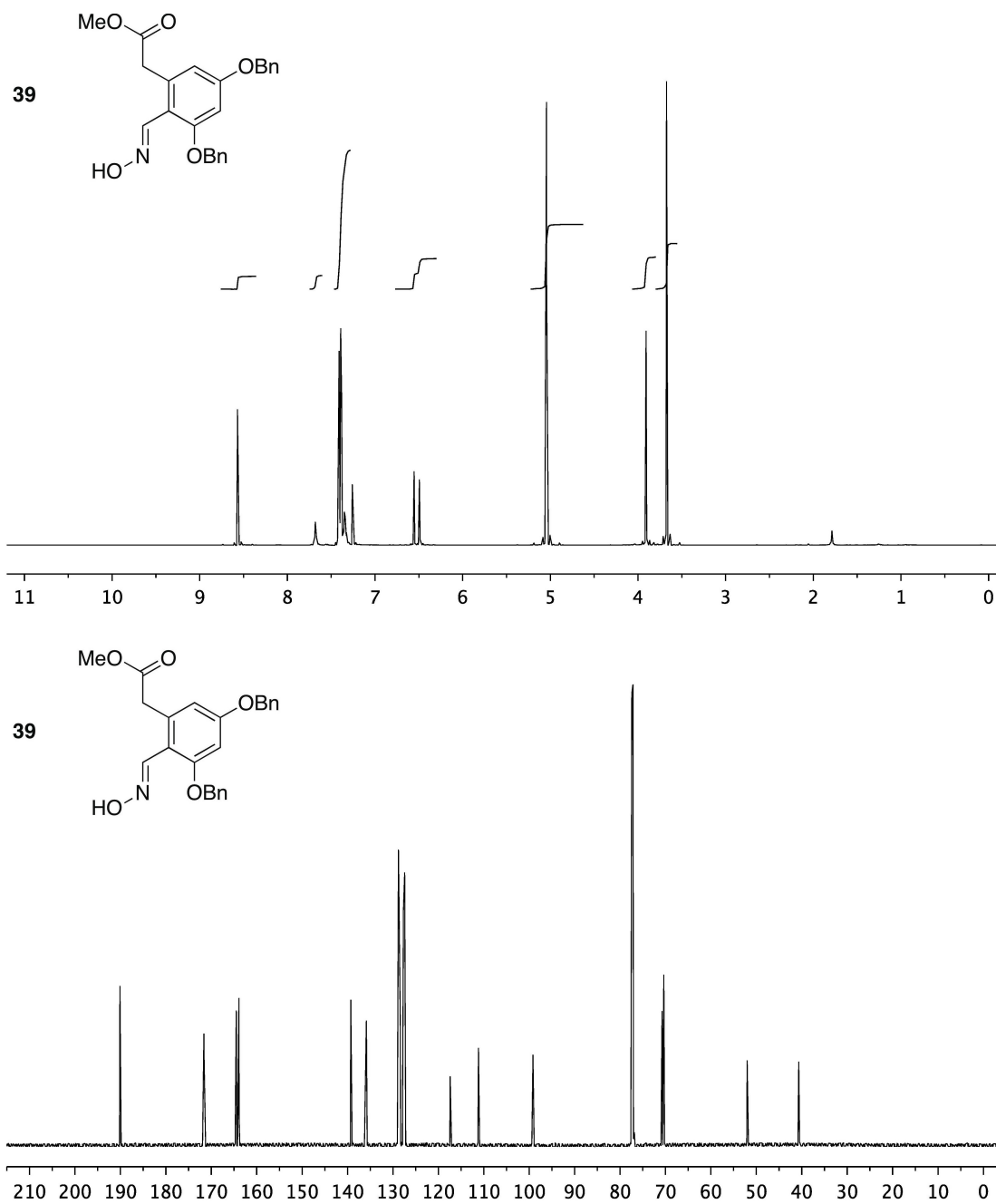
Figure S25. $^1\text{H-NMR}$ (500 MHz) and $^{13}\text{C-NMR}$ (125 MHz) spectra of **39** in CDCl_3 

Figure S26. ^1H -NMR (500 MHz) and ^{13}C -NMR (125 MHz) spectra of **40** in CDCl_3

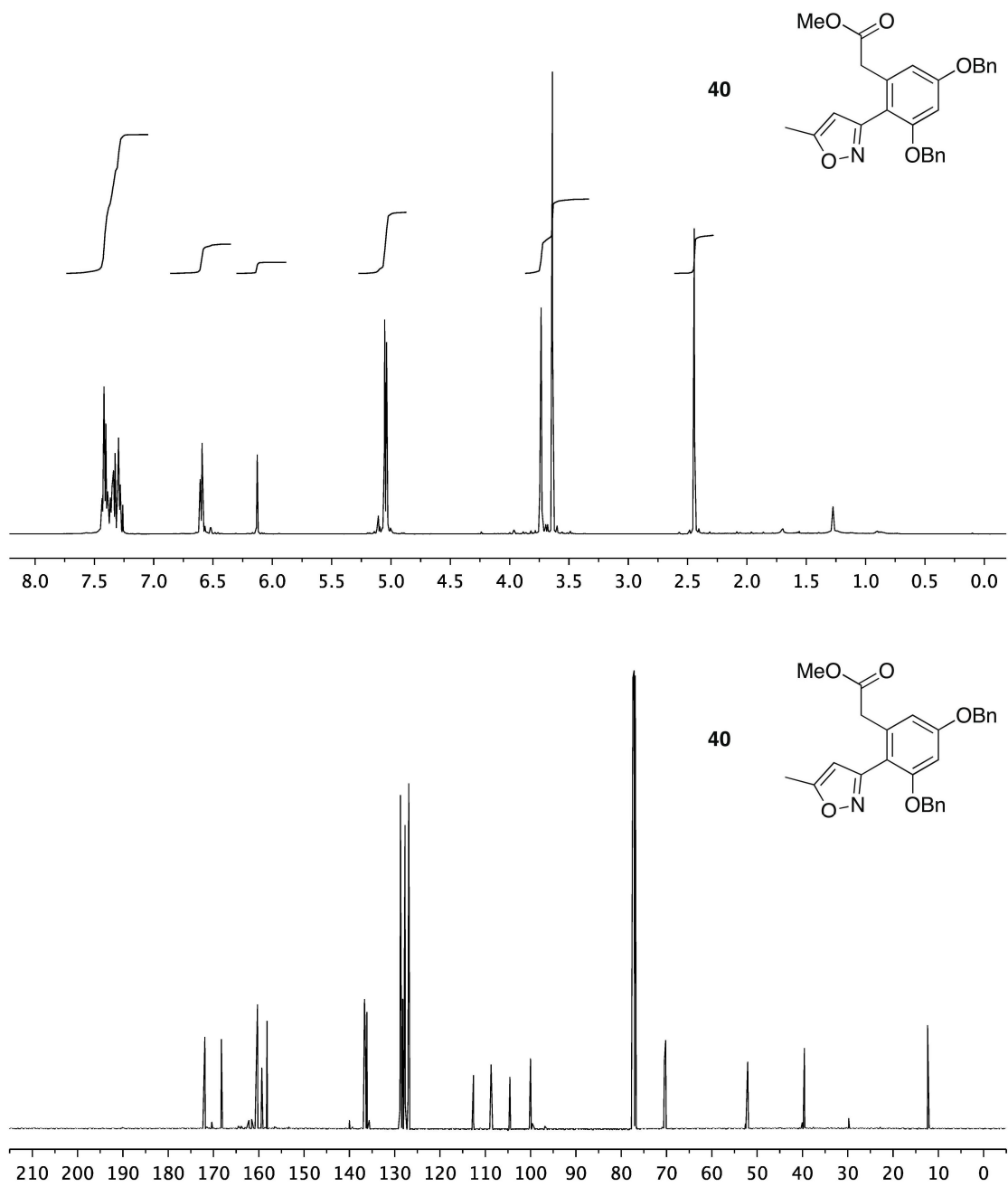


Figure S27. $^1\text{H-NMR}$ (500 MHz) and $^{13}\text{C-NMR}$ (125 MHz) spectra of **41** in $\text{DMSO-}d_6$

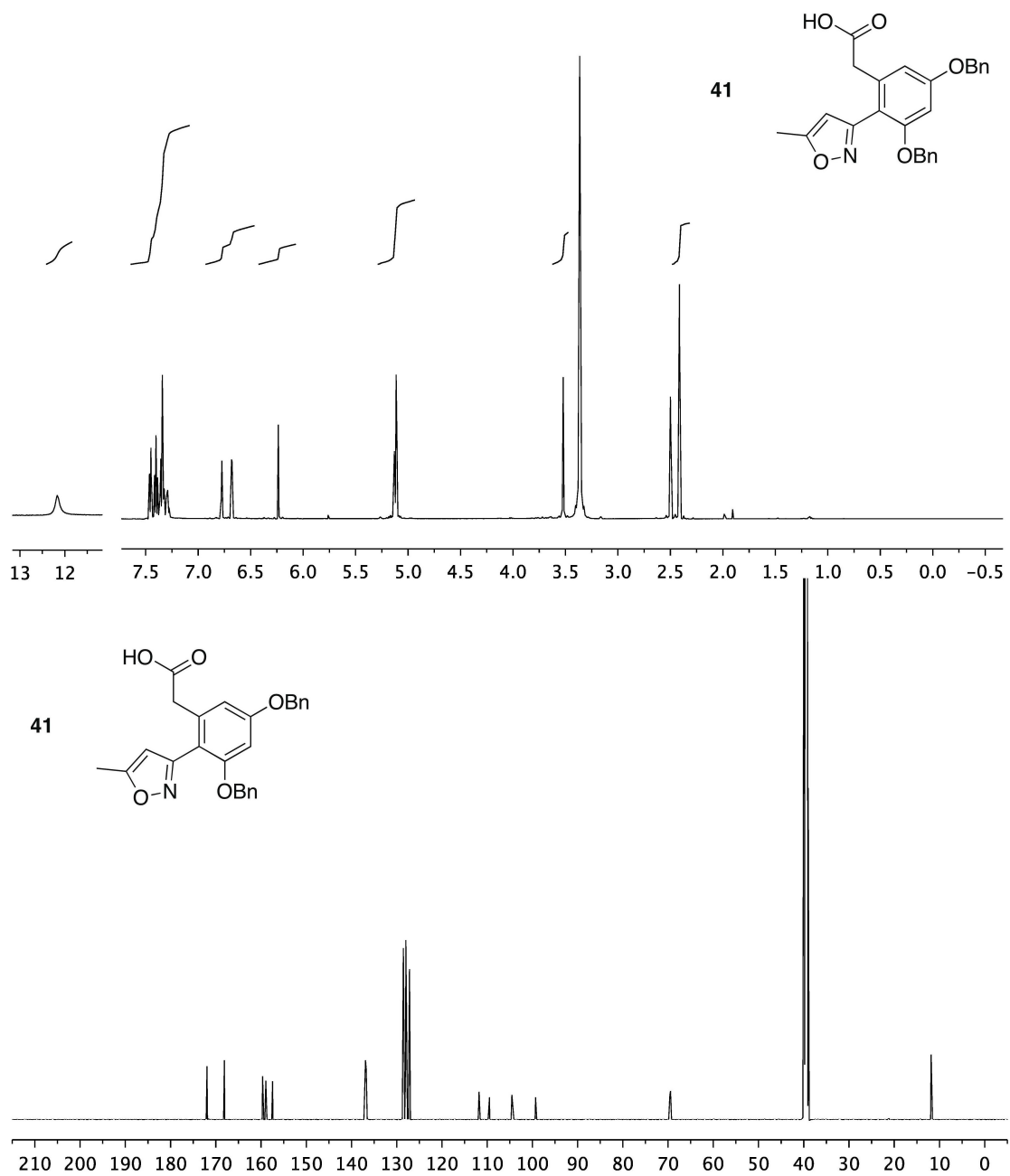


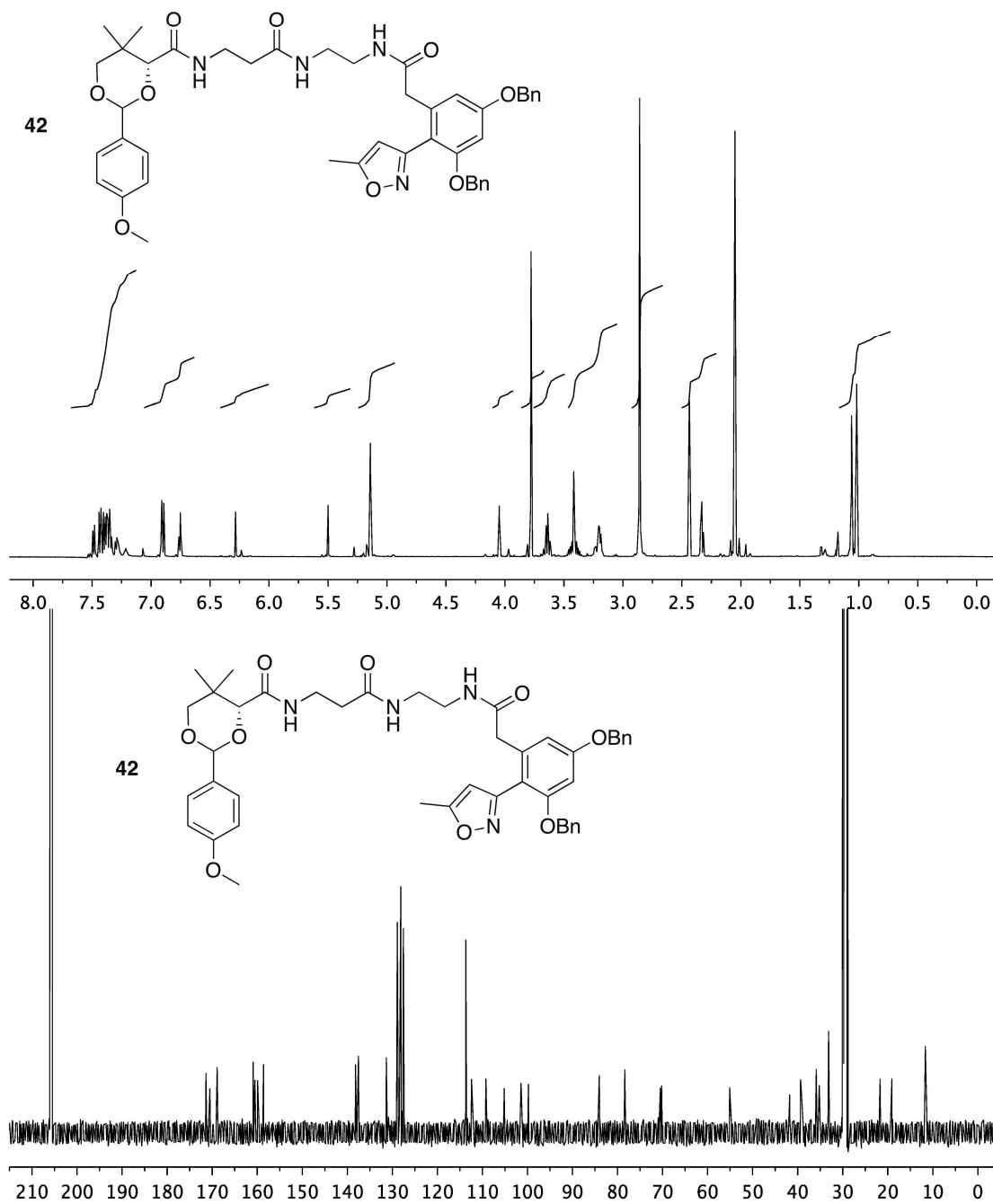
Figure S28. $^1\text{H-NMR}$ (500 MHz) and $^{13}\text{C-NMR}$ (125 MHz) spectra of **42** in acetone- d_6 

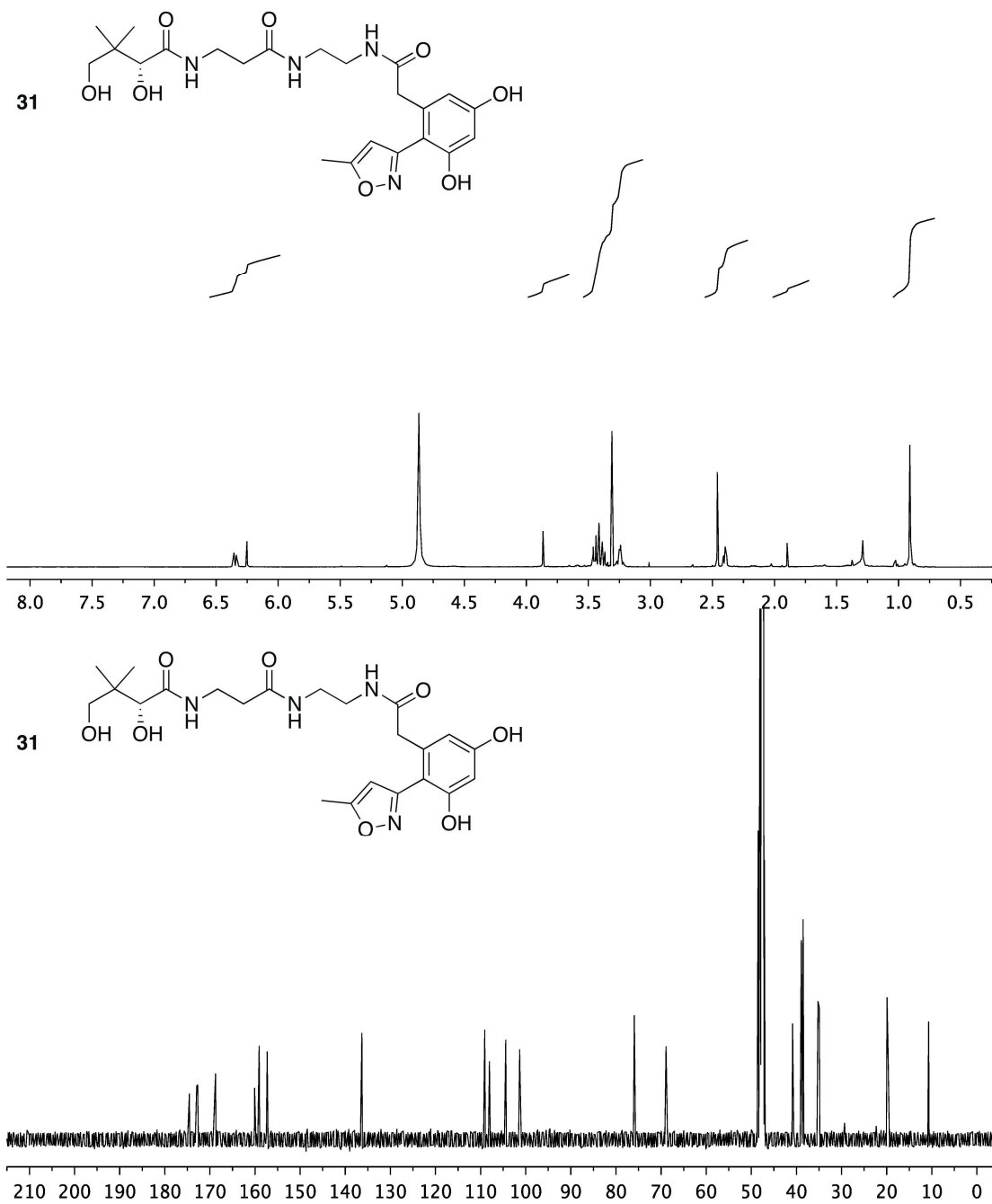
Figure S29. $^1\text{H-NMR}$ (500 MHz) and $^{13}\text{C-NMR}$ (125 MHz) spectra of mimetic **31** in CD_3OD 

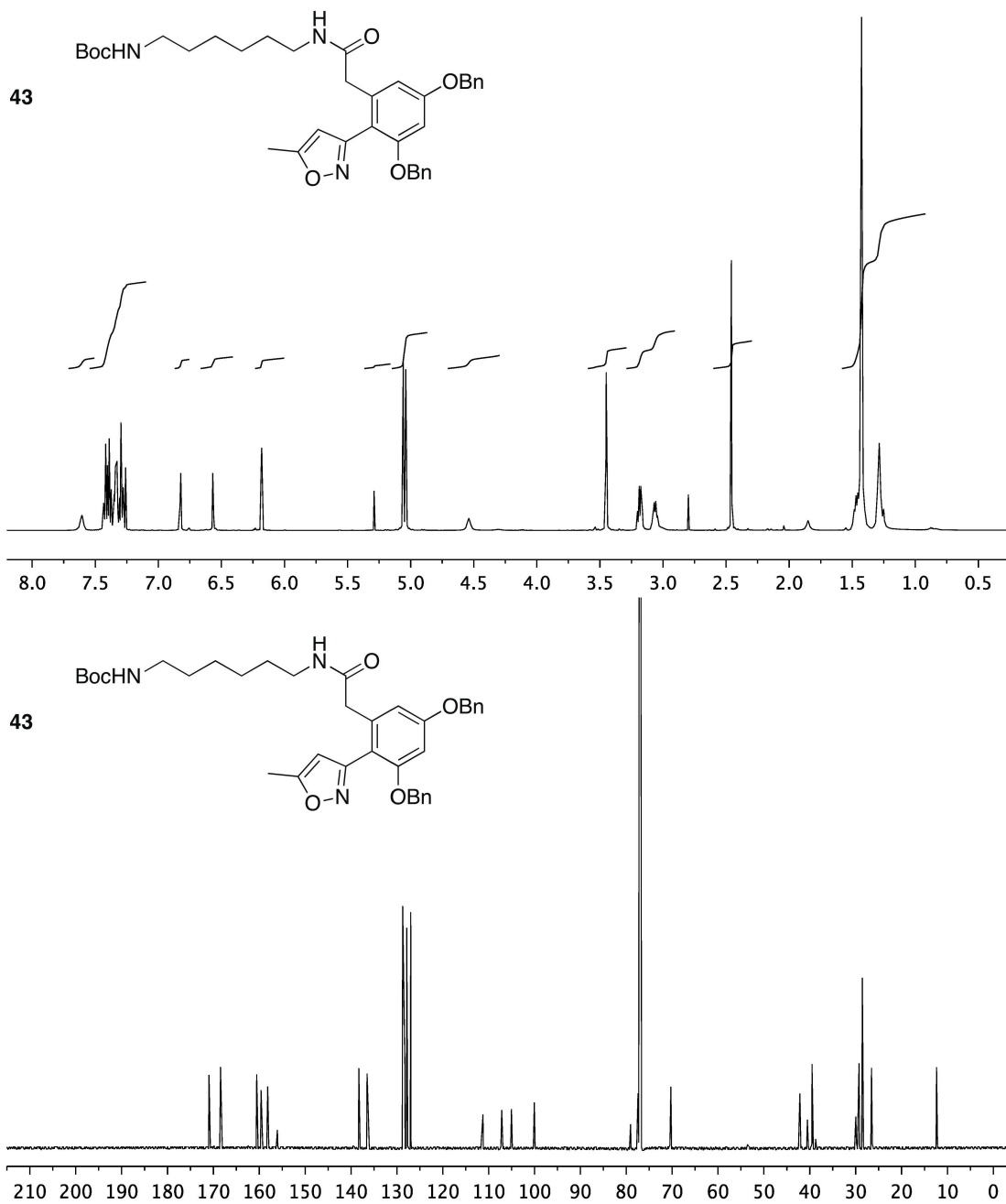
Figure S30. $^1\text{H-NMR}$ (500 MHz) and $^{13}\text{C-NMR}$ (125 MHz) spectra of **43** in CDCl_3 

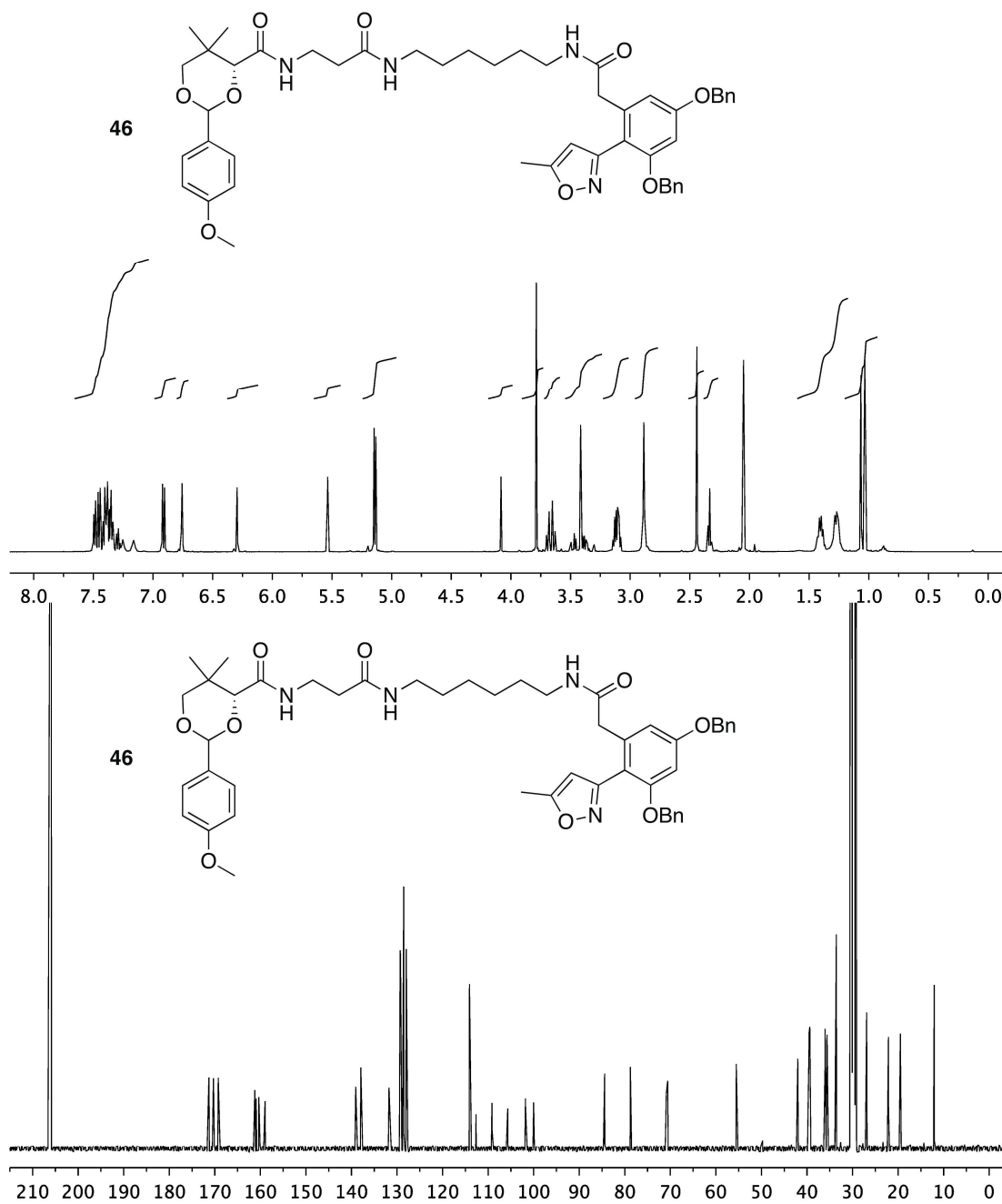
Figure S31. $^1\text{H-NMR}$ (500 MHz) and $^{13}\text{C-NMR}$ (125 MHz) spectra of **46** in acetone- d_6 

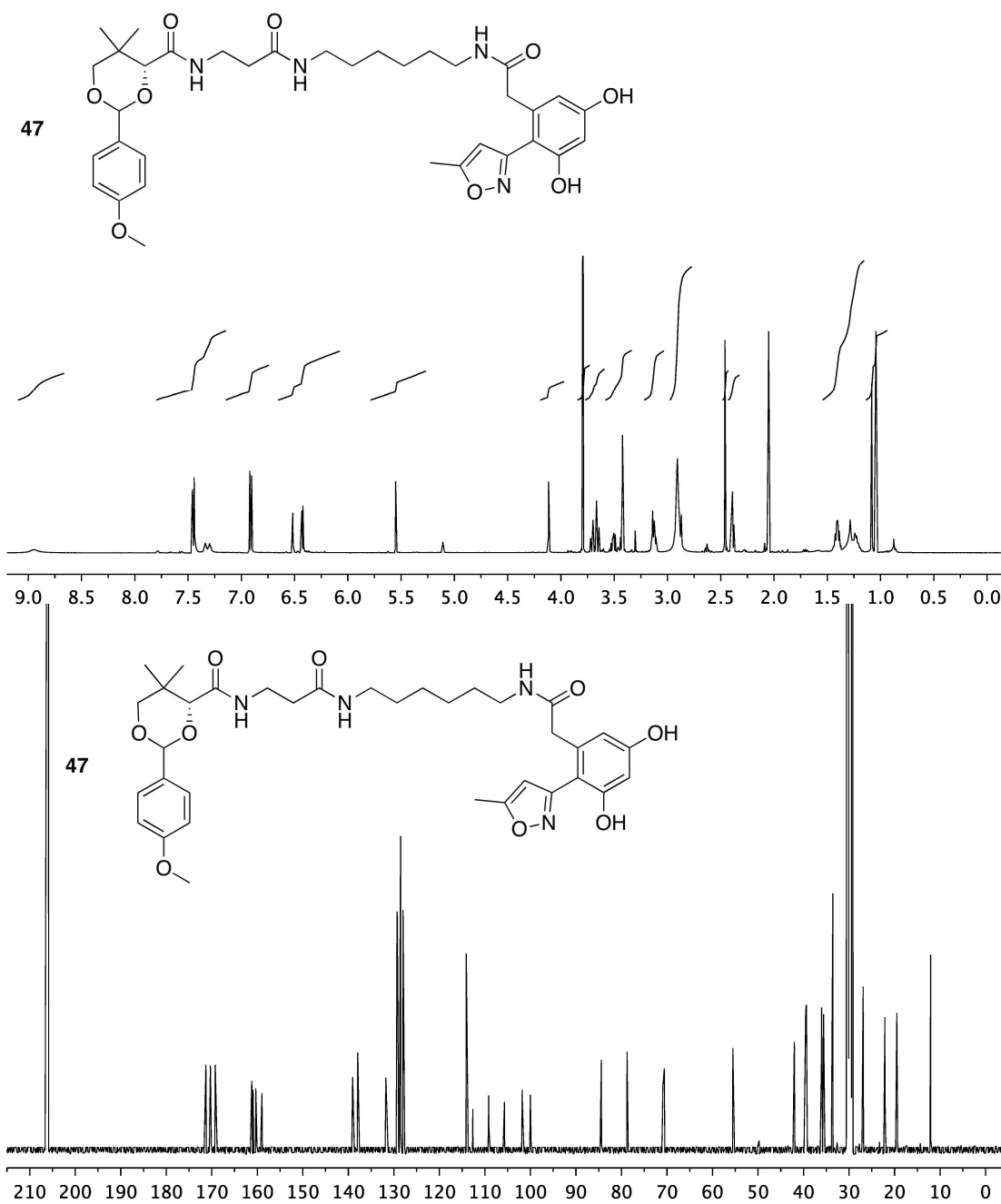
Figure S32. ^1H -NMR (500 MHz) and ^{13}C -NMR (125 MHz) spectra of **47** in acetone- d_6 

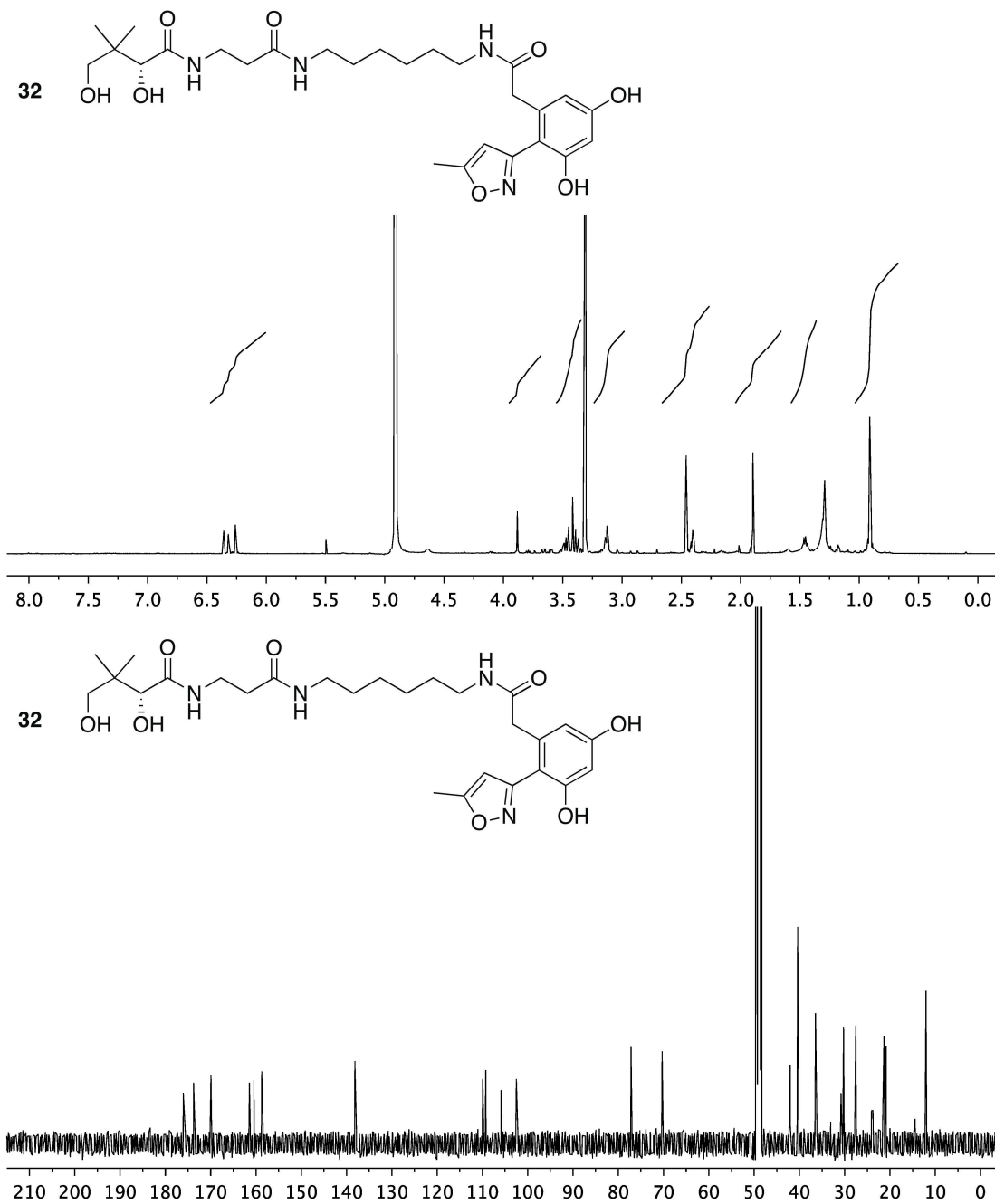
Figure S33. ^1H -NMR (500 MHz) and ^{13}C -NMR (125 MHz) spectra of **32** in CD_3OD 

Figure S34. Docking simulations of ActACP with probes.

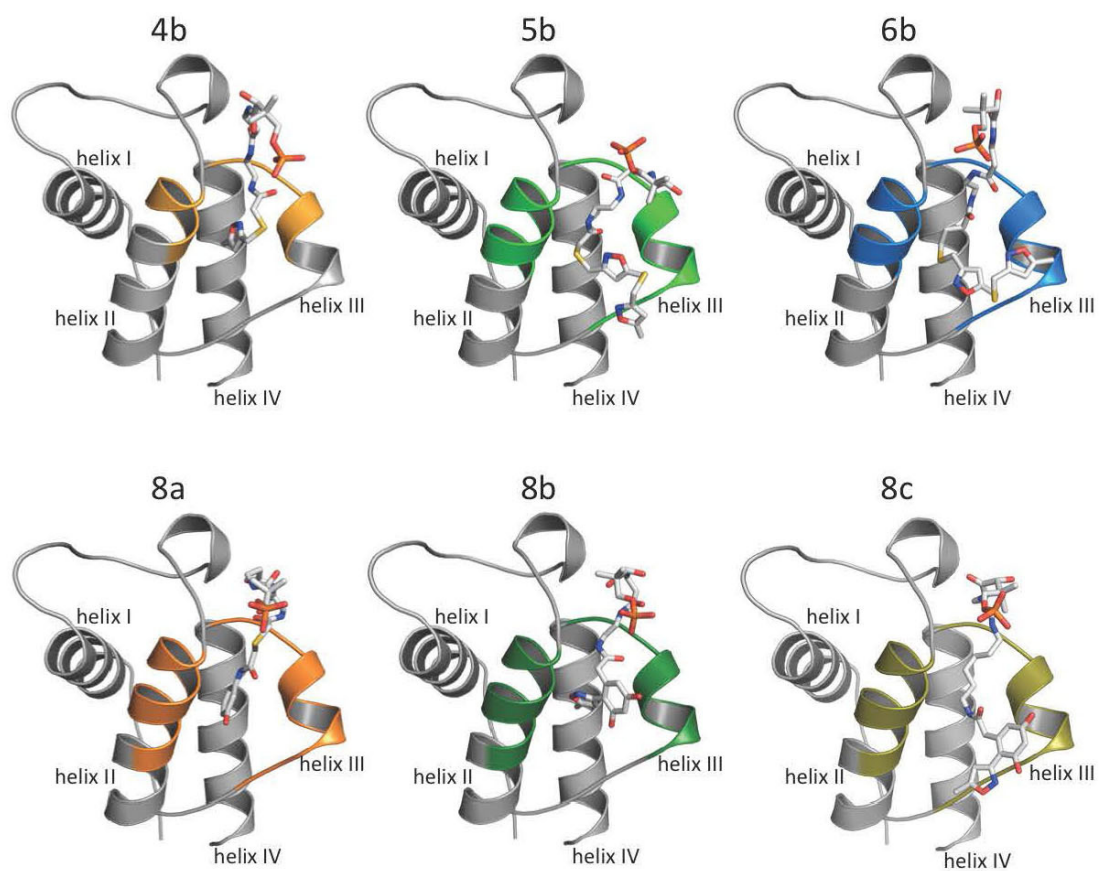
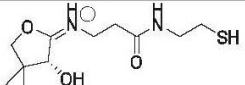
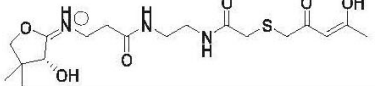
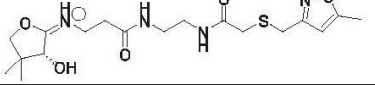
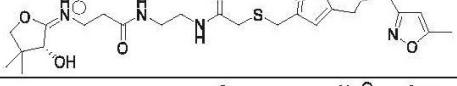
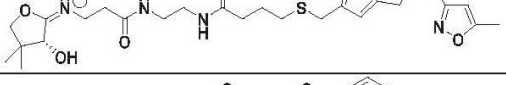
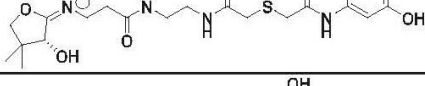
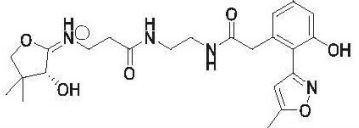
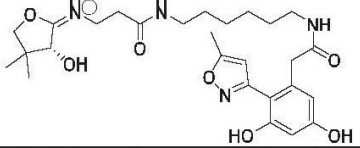
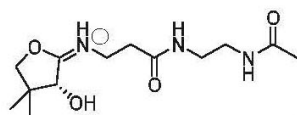


Table S1. LCMS validation of *crypto*-ACPs via PPant ejection.

ACP Species	Ejected Molecule	Expected Mass (Da)	Observed Mass (Da)
<i>apo</i> -ACP	None expected	N/A	None observed
2 (<i>holo</i> -ACP)		261.36	261.48
4a		416.51	286.56*
4b		413.51	413.57
5b		540.67	540.56
6b		568.73	568.62
8a		467.56	467.61
8b		475.52	475.56
8c		531.63	531.60

*Ejected mass under high voltage conditions yields decomposed ion:



(Expected mass 286.35)

Loaded *crypto*-ActACP samples were subjected to LCMS analysis using high-voltage settings to eject their cargo by lactone formation (Meier JL, et al., *J Proteome Res.* 2011, 10, 320-9.) The observed small molecule peaks are reported here, to further confirm loading of the actACP (in addition to characteristic S42 shifts observed by NMR).

Acknowledgements: Chapter 2, in full, is a reprint of the material as it appears: Shakya, G.; Rivera, H.; Lee, D. J.; Jaremko, M. J.; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdzowski, E. R.; Wong, F.; Tsai, S.-C.; Burkart, M. D. “Modeling Linear and Cyclic PKS Intermediates through Atom Replacement.” *J. Am. Chem. Soc.* 136 (2014): 16792-9. The dissertation author is a contributing author of this manuscript.

Chapter 3. Structure and Substrate Sequestration in the Pyoluteorin Type II

Peptidyl Carrier Protein PtlL

Structure and Substrate Sequestration in the Pyoluteorin Type II Peptidyl Carrier Protein PtlL

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

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

Supporting Information

ABSTRACT: Type II nonribosomal peptide synthetases (NRPS) generate exotic amino acid derivatives that, combined with additional pathways, form many bioactive natural products. One family of type II NRPSs produce pyrrole moieties, which commonly arise from proline oxidation while tethered to a conserved, type II peptidyl carrier protein (PCP), as exemplified by PtlL in the biosynthesis of pyoluteorin. We sought to understand the structural role of pyrrole PCPs in substrate and protein interactions through the study of pyrrole analogs tethered to PtlL. Solution-phase NMR structural analysis revealed key interactions in residues of helix II and III with a bound pyrrole moiety. Conservation of these residues among PCPs in other pyrrole containing pathways suggests a conserved mechanism for formation, modification, and incorporation of pyrrole moieties. Further NOE analysis provided a unique pyrrole binding motif, offering accurate substrate positioning within the cleft between helices II and III. The overall structure resembles other PCPs but contains a unique conformation for helix III. This provides evidence of sequestration by the PCP of aromatic pyrrole substrates, illustrating the importance of substrate protection and regulation in type II NRPS systems.

Type II nonribosomal peptide synthetases (NRPSs) commonly combine with fatty acid and/or polyketide synthases (FAS/PKS) generating structurally diverse secondary metabolites with activities ranging from anti-infective to anticancer agents. The products of type II NRPSs are modified amino acids and are frequently aromatic. For example, proline, tyrosine, and salicylic acid are shuttled through type II NRPSs to form pyrrole, hydroxytyrosine, and methoxy-chlorosalicylic acid before incorporation into downstream pathways.¹ These moieties are electron-rich rings that can participate in hydrogen bonding and π -stacking interactions that can contribute to the bioactivity of their respective product. The pyrrole containing compounds prodigiosin and chlorizidine A exhibit promising antitumor activity,² for which a prodigiosin derivative recently entered stage II clinical trials (Figure 1).³ Vancomycin (hydroxytyrosine) and chlorothricin (methoxy-chlorosalicylic acid) are known antibacterial and cholesterol reducing compounds, respectively.⁴ All type II NRPS components in these compounds are generated while covalently attached to a peptidyl carrier protein (PCP).

As shown in Scheme 1, the PCP is responsible for shuttling peptide substrates to partner enzymes for type II NRPS catalyzed

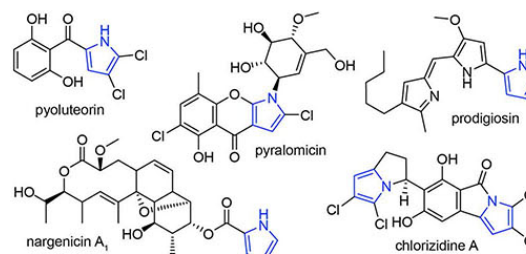


Figure 1. Structures of pyrrole containing natural products.

modifications. To prime the *apo*-PCP, a 4'-phosphopantetheine (PPant) arm is first coupled to a conserved serine by a 4'-phosphopantetheinyl transferase (PPTase), generating *holo*-PCP. In systems that incorporate a pyrrole, an adenylation domain serves to load proline (Pro) onto the *holo*-PCP. The resulting prolyl-PCP is then recognized by a flavoprotein dehydrogenase, which catalyzes a four-electron oxidation of Pro to the corresponding pyrrolyl-PCP.⁵ Further downstream processing can occur while tethered to the PCP, as illustrated by pyrrole halogenation in pyoluteorin biosynthesis (Scheme 1). In this system, the modified pyrrole is then transferred to a three-module type I PKS system that facilitates elongation and reduction, followed by cyclization to pyoluteorin.⁶

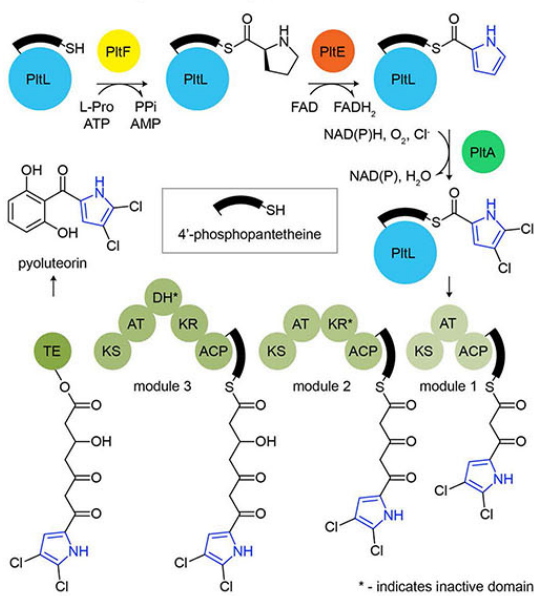
Given that PCP attachment is essential for product formation, protein-protein interactions between the PCP and partner enzymes are presumed to be critical, although the structural details of these binding events remain unknown. In pyoluteorin biosynthesis, Pro is loaded onto the PCP PtlL by PtlF and then oxidized to pyrrole by PtlE (Scheme 1). The homologous adenylation and dehydrogenase enzymes in undecylprodigiosin biosynthesis cannot catalyze pyrrole formation with PtlL,⁵ illustrating the importance of PCP identity, substrate demonstration, and protein-protein interactions.

Substrate sequestration in carrier proteins has been recognized as an important phenomenon in type II fatty acid and polyketide synthases.⁷ Type II carrier proteins, as opposed to type I, are stand-alone enzymes that must recognize up to five partner proteins in a particular order. When not interacting with a partner, the carrier protein has been shown to protect the extending substrate from reactive compounds in the cytosol. Here the substrate localizes in a hydrophobic cleft between helix II and III

Received: April 30, 2015

Published: September 4, 2015

Scheme 1. Biosynthesis of Pyoluteorin



of the acyl carrier protein (ACP). NMR structural studies of acylated *E. coli* fatty acid ACP reveals sequestration of chain lengths greater than C₄,^{7d,e} which shields the growing metabolite from nucleophiles other than the cognate partner protein.⁸ Similar interactions are also seen in type II PKS, where helix II and III of actinorhodin ACP were shown to interact with cyclic and linear intermediate analogs.^{7f,g} While PCPs have been shown to interact with partner enzymes at the helix II/III interface,⁹ no studies have demonstrated PCP substrate sequestration. The overall helical structure of PCPs is similar in helix II and III; however, amino acid distributions are notably diverse, creating significant electrostatic differences.¹⁰ These surfaces may be key to protection of various intermediates to discriminate for reactivity with proper partner enzymes. Structural studies of carrier proteins can provide insight into the regulation of substrate modification. Currently, there are no structures of a peptidyl-bound PCP, and there are only two structures of stand-alone type II PCPs, BlmI and A3404, for which functions are unknown.^{10,11}

Here we show that pyrrole PCPs actively sequester their pyrrole cargo. Bioinformatic analysis of PtlL and other pyrrole PCPs revealed a secluded PCP subfamily with several conserved residues. Using chemoenzymatic techniques developed in our laboratory,¹² we attached PPant-Pro and PPant-pyrrole mimetics to PtlL to study interactions between the PCP and substrate. NMR analysis revealed direct interactions of the pyrrole with residues in helix II and III not observed in other PCPs.¹³ Through NOE techniques, the pyrrole acyl substrate was localized in a 3D solution NMR structure of PtlL. The results shed light on how type II NRPS PCPs regulate peptidyl substrates and for the first time reveal direct interactions between the substrate and PCP.

Due to the unique nature of pyrrole biosynthesis, we first hypothesized that pyrrole PCPs could encompass a subfamily of PCPs. Sequence homology, generated with MUSCLE,¹⁴ was used to compare PCPs that tether pyrrole, Pro, and other amino acids as final products (Figure S1). The comparison highlights conserved residues and polarity in pyrrole PCPs, particularly at

the helix II N-terminus and in helix III. A phylogenetic tree generated with ClustalW¹⁵ further illustrates the seclusion of pyrrole PCPs (Figure S2). By comparison, PCPs that load Pro as a final product are dispersed among PCPs that load other amino acids, emphasizing the unique conservations in pyrrole PCPs. Some of the conserved residues play an important part in substrate stabilization and possibly protein–protein recognition.

To analyze the interaction between PtlL and peptidyl substrates, NMR experiments were considered. However, the natural thioester between the pantetheine and substrate is susceptible to hydrolysis in aqueous solution, and this instability is further aggravated under conditions for NMR studies.^{7a,16} Therefore, we prepared pantetheine mimetics with Pro and pyrrole for attachment onto PtlL employing an amide linkage in lieu of the thioester (SI A.1, A.2). These Pro and pyrrole probes were then loaded onto PtlL via a one-pot chemo-enzymatic strategy as previously described.¹² Posttranslational modification was verified by urea PAGE and LCMS (Figures S3, S4). To locate which PtlL residues are perturbed by addition of the Pro and pyrrole moieties, ¹H–¹⁵N HSQC experiments were conducted on ¹⁵N-holo-PtlL, ¹⁵N-prolyl-PtlL, and ¹⁵N-pyrrolyl-PtlL, and chemical shift perturbations (CSPs) were calculated (Figure 2). Analysis revealed several key residues that were altered with the

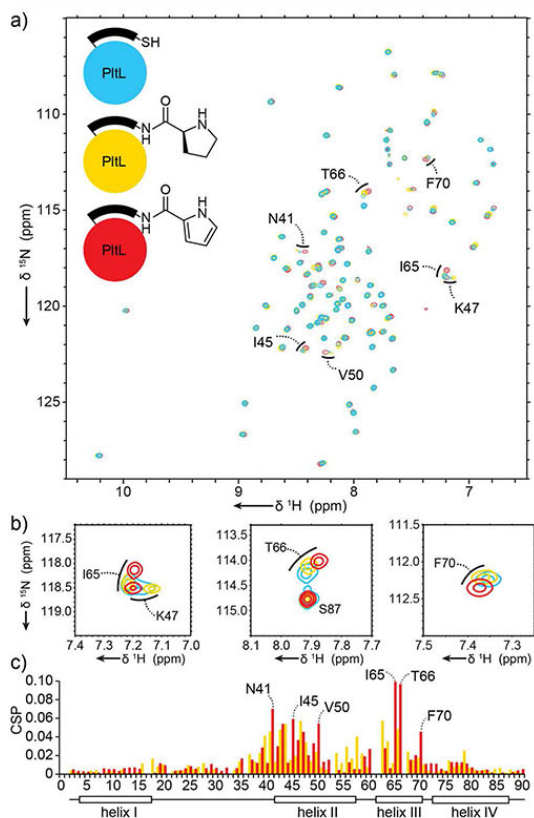


Figure 2. (a) HSQC overlay of ¹⁵N-holo-PtlL, ¹⁵N-prolyl-N-PtlL, ¹⁵N-pyrrolyl-N-PtlL. (b) Highlights of ¹⁵N-HSQC data, illustrating perturbations of residues I65, T66, and F70. (c) CSP plots of ¹⁵N-prolyl-PtlL and ¹⁵N-pyrrolyl-PtlL relative to ¹⁵N-holo-PtlL.

attachment of the Pro and pyrrole moieties. Residues in helix II and proximal to Ser42 (covalently attached to PPant) showed significant movement, specifically Asn41, Ile45, Lys47, and Val50. However, the largest movements were seen in helix III, including residues Ile65, Thr66, and Phe70 (Figure 2). The side chains of Asn41, Lys47, and Val50 do not protrude toward the helix II/III cleft and may indicate general movement of helix II, although other residues, as mentioned below, have direct contact with the substrates. The CSP results of PltL coincide with other carrier protein–substrate interactions, where the majority of the interactions occur at helix II and III. Overall, the perturbations were larger for pyrrolyl-PltL compared to prolyl-PltL. This may indicate further sequestration or stabilization of the more hydrophobic pyrrole moiety compared to Pro.

In conjunction with the PltL perturbations, movement was also observed in protons of the pyrrole (Figure 3). Due to the unique chemical shifts of pyrrole protons, a ^1H NMR experiment could be used to observe the protons in PltL-bound pyrrole. Three isolated peaks of the pyrrole shift upfield when pyrrole-*N*-pantetheine is tethered to PltL, compared to the probe alone, or in solution but unattached to PltL (Figure 3). This indicates that the pyrrole is located in a more electron-rich environment, in this case, the hydrophobic cleft between helix II and III. PltL was also shown to interact with the solvatochromatic compound 4-DMN (Figure S15).¹⁷ The NMR and fluorescent data suggest that the pyrrole is solvent protected. A ^1H - ^1H NOESY experiment was then performed on pyrrolyl-PltL to observe any specific proton interactions between the pyrrole and PltL residues. The pyrrole proton NOESY strips contained several peaks corresponding to

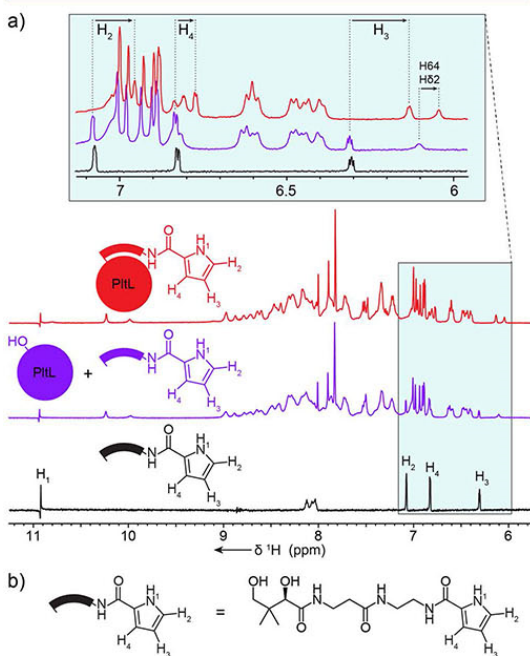


Figure 3. Pyrrole NMR shift analyses. (a) ^1H NMR experiment of pyrrolyl-*N*-pantetheine probe isolated (red), with *apo*-PltL (blue), and covalently attached to PltL (green). The enlarged spectra reveal perturbations of pyrrole protons, suggesting pyrrole-PltL interactions. (b) Structure of pyrrolyl-*N*-pantetheine probe.

Leu35 $\text{H}\delta^*$, Ile45 $\text{H}\delta^*$, Ile65 $\text{H}\gamma^*/\delta^*$, and Thr66 $\text{H}\alpha$, all of which are located in the cleft of helix II and III (Figure S16). Correspondingly, pyrrole proton shifts can also be visualized within NOE strips of hydrophobic residues of the PltL cleft (Figure S17).

To further visualize the interactions between PltL and peptidyl substrates, the solution NMR structures of both *holo*-PltL (PDB ID: 2NSH) and pyrrolyl-PltL (PDB ID: 2NSI) were determined with Cyana 3.97 (Figures 4a–b, S18).¹⁸ The overall structure of PltL consists of four bundled helices as seen with all carrier protein structures (Figure S19).^{9–11} PltL has both positive and negative electrostatic surfaces similar to other PCPs (Figure S20), but different from FAS and PKS ACPs, which have primarily negative surfaces.¹⁰ Unlike the multiple states observed for the type I excised PCP TycC3,¹⁹ a single state was observed for both PltL structures similar to that seen in other PCPs.^{10,13,20} Further similarities with PCPs include a linker region and helical turn between helix I and II and a Pro situated on the N-terminus of helix III.^{10,13} However, PltL contains Gly67 that interrupts the center of helix III, consequently positioning neighboring residues closer to helix II (Figure 4c). For example, Ile65 and Thr66 are only in position to interact with peptidyl substrates due to the helical interruption. Several other pyrrole carrier proteins contain Gly and/or Pro in the middle of helix III, likely interrupting these

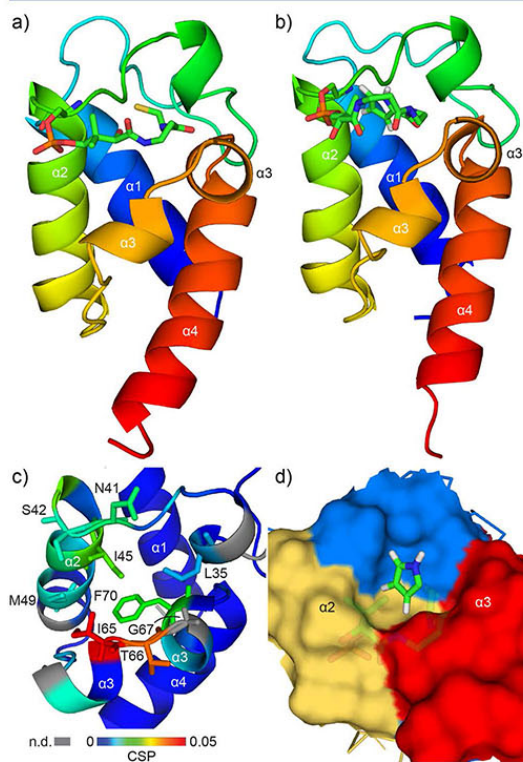


Figure 4. Solution NMR structures of PltL species. (a) *holo*-PltL. (b) pyrrolyl-PltL. (c) Expanded view of pyrrolyl-PltL with colors corresponding to CSP values from Figure 2. Key residues are labeled. CSPs not determined are shown in gray. (d) Expanded view of surface-filled pyrrolyl-PltL structure illustrating pyrrole sequestration between the helix II/III cleft.

helices (Figure S1), while other PCPs do not contain an interrupter (Figure S19).

Spectroscopic observation of the PPant arm of *holo*-PtlL was facilitated by ^{15}N and ^{13}C isotope labeling of the tethered molecule (Figures S8, S10). This was achieved by cotransformation of PtlL and Sfp constructs for *in vivo* production of *holo*-PtlL (SI B.3). NOEs from the PPant to the protein were observed on both ends, but only intra-NOEs were observed in the middle of the PPant arm. These NOEs provided constraints between the thiol end and residues in both helix II and III, yielding a loop conformation with the thiol positioned between helix II and III.

The solution NMR structure of pyrrolyl-PtlL was determined containing the solvent protected pyrrole localized between helix II and III (Figures 4d, S21). The structure of pyrrolyl-PtlL differs slightly with the *holo*-PtlL, where the hydrophobic cleft between helix II and III expands to accommodate the pyrrole. The PPant in *holo*-PtlL has more constraints than in pyrrolyl-PtlL, possibly due to the pyrrole displacing the PPant further into solution. The position of the pyrrole protects the reactive 4 and 5 positions of the aromatic ring, which are eventually chlorinated (Scheme 1).²¹ In this position, the pyrrole is inaccessible to the halogenase PtlA and, therefore, the interaction between PtlL and PtlA must induce a conformational change in PtlL to display the pyrrole to PtlA. These structural data highlight substrate sequestration by PtlL in the biosynthesis of pyrrole and point to a similar role for other PCPs in type II NRPS pathways.

In this report, the structure of the first pyrrole carrier protein was determined, and the protein's interaction with substrate mimetics is revealed. This work illustrates the features of PCPs that are responsible for binding and protection of substrate intermediates and suggests the existence of similar PCP-dependent regulation in other type II NRPS biosynthetic pathways. For example, the methoxy-chlorosalicylic acid in chlorothricin is generated while attached to the type II PCP ChlB2,^{1a} which has high sequence identity to PtlL (36%) and therefore may have similar sequestration events. These studies also provide important information toward understanding the protein interfaces that facilitate specificity within pyrrole biogenesis. Future work with type II NRPS enzymes will provide a more complete understanding of protein-protein recognition and offer the prospect of enhanced bioactive compounds via synthetic biology.

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b04525.

NMR structures were deposited to the PDB under IDs 2NSH (*holo*PtlL) and 2NSI (pyrrolyl-PtlL) (PDF)

■ AUTHOR INFORMATION

■ Corresponding Author

*mburkart@ucsd.edu

■ Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Funding was provided from NIH GM095970 and NIH/NCI T32 CA009523. We thank J. Beld for the solvachromatic probe, C. T. Walsh for the plasmid constructs, Drs. X. Huang and A. Mrse for NMR assistance, and Dr. Y. Su for MS services.

■ REFERENCES

- (1) (a) Jia, X.-Y.; Tian, Z.-H.; Shao, L.; Qu, X.-D.; Zhao, Q.-F.; Tang, J.; Tang, G.-L.; Liu, W. *Chem. Biol.* **2006**, *13*, 575. (b) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 171. (c) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13843. (d) Hubbard, B. K.; Walsh, C. T. *Angew. Chem., Int. Ed.* **2003**, *42*, 730.
- (2) (a) Ho, T.-F.; Ma, C.-J.; Lu, C.-H.; Tsai, Y.-T.; Wei, Y.-H.; Chang, J.-S.; Lai, J.-K.; Cheuh, P.-J.; Yeh, C.-T.; Tang, P.-C.; Tsai Chang, J.; Ko, J.-L.; Liu, F.-S.; Yen, H. E.; Chang, C.-C. *Toxicol. Appl. Pharmacol.* **2007**, *225*, 318. (b) Alvarez-Mico, X.; Jensen, P. R.; Fenical, W.; Hughes, C. C. *Org. Lett.* **2013**, *15*, 988.
- (3) Parikh, S. A.; Kantarjian, H.; Schimmer, A.; Walsh, W.; Asatiani, E.; El-Shami, K.; Winton, E.; Verstovsek, S. *Clin. Lymphoma Myeloma Leuk.* **2010**, *10*, 285.
- (4) (a) Williams, D. H.; Bardsley, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1172. (b) Kawashima, A.; Nakamura, Y.; Ohta, Y.; Akama, T.; Yamagishi, M.; Hanada, K. *J. Antibiot.* **1992**, *45*, 207.
- (5) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 171.
- (6) Gross, H.; Loper, J. E. *Nat. Prod. Rep.* **2009**, *26*, 1408.
- (7) (a) Zornetzer, G. A.; Fox, B. G.; Markley, J. L. *Biochemistry* **2006**, *45*, 5217. (b) Roujeinikova, A.; Baldock, C.; Simon, W. J.; Gilroy, J.; Baker, P. J.; Stuitje, A. R.; Rice, D. W.; Slabas, A. R.; Rafferty, J. B. *Structure* **2002**, *10*, 825. (c) Mercer, A. C.; Burkart, M. D. *Nat. Prod. Rep.* **2007**, *24*, 750. (d) Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. *Nat. Methods* **2012**, *9*, 981. (e) Roujeinikova, A.; Simon, W. J.; Gilroy, J.; Rice, D. W.; Rafferty, J. B.; Slabas, A. R. *J. Mol. Biol.* **2007**, *365*, 135. (f) Haushalter, R. W.; Filipp, F. V.; Ko, K. S.; Yu, R.; Opella, S. J.; Burkart, M. D. *ACS Chem. Biol.* **2011**, *6*, 413. (g) Shakya, G.; Rivera, H.; Lee, D. J.; Jaremko, M. J.; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdziowski, E. R.; Wong, F.; Tsai, S.-C.; Burkart, M. D. *J. Am. Chem. Soc.* **2014**, *136*, 16792.
- (8) Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Markwick, P. R. L.; Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, F.; O'Dowd, B.; McCammon, J. A.; Opella, S. J.; Tsai, S.-C.; Burkart, M. D. *Nature* **2013**, *505*, 427.
- (9) (a) Lai, J. R.; Fischbach, M. A.; Liu, D. R.; Walsh, C. T. *J. Am. Chem. Soc.* **2006**, *128*, 11002. (b) Lai, J. R.; Fischbach, M. A.; Liu, D. R.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 5314. (c) Zhou, Z.; Lai, J. R.; Walsh, C. T. *Chem. Biol.* **2006**, *13*, 869.
- (10) Lohman, J. R.; Ma, M.; Cuff, M. E.; Bigelow, L.; Bearden, J.; Babbitt, G.; Joachimiak, A.; Phillips, G. N.; Shen, B. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 1210.
- (11) Allen, C. L.; Gulick, A. M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 1718.
- (12) Worthington, A. S.; Burkart, M. D. *Org. Biomol. Chem.* **2006**, *4*, 44.
- (13) Tufar, P.; Rahighi, S.; Kraas, F. L.; Kirchner, D. K.; Löhr, F.; Henrich, E.; Köpke, J.; Dikic, I.; Güntert, P.; Marahiel, M. A.; Dötsch, V. *Chem. Biol.* **2014**, *21*, 552.
- (14) Edgar, R. C. *Nucleic Acids Res.* **2004**, *32*, 1792.
- (15) 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.
- (16) Goodrich, A. C.; Frueh, D. P. *Biochemistry* **2015**, *54*, 1154.
- (17) Beld, J.; Cang, H.; Burkart, M. D. *Angew. Chem., Int. Ed.* **2014**, *53*, 14456.
- (18) Güntert, P. *Eur. Biophys. J.* **2009**, *38*, 129.
- (19) (a) Koglin, A.; Mofid, M. R.; Löhr, F.; Schäfer, B.; Rogov, V. V.; Blum, M.-M.; Mittag, T.; Marahiel, M. A.; Bernhard, F.; Dötsch, V. *Science* **2006**, *312*, 273. (b) Koglin, A.; Lohr, F.; Bernhard, F.; Rogov, V. V.; Frueh, D. P.; Strieter, E. R.; Mofid, M. R.; Guntert, P.; Wagner, G.; Walsh, C. T.; Marahiel, M. A.; Dötsch, V. *Nature* **2008**, *454*, 907.
- (20) Haslinger, K.; Redfield, C.; Cryle, M. J. *Proteins: Struct., Funct., Genet.* **2015**, *83*, 711.
- (21) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13843.

SUPPORTING INFORMATION

Structure and substrate sequestration in the pyoluteorin peptidyl carrier protein PItL

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

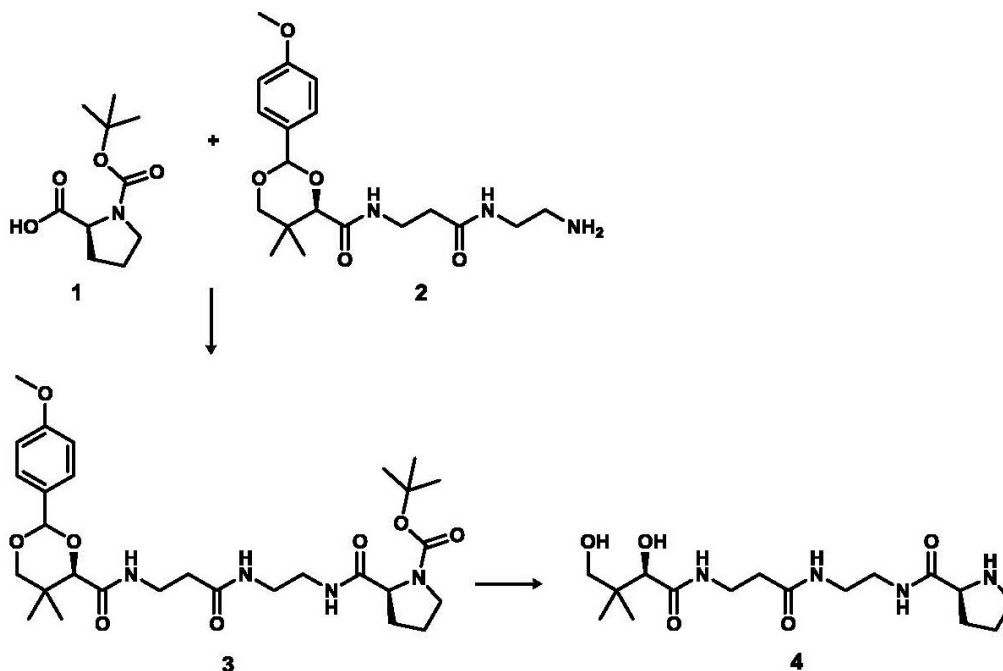
Correspondence should be directed to mburkart@ucsd.edu.

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 PtlL C87S	S5
B.2. Expression and purification of $^{15}\text{N}/^{13}\text{C}$ <i>apo</i> -PtlL	S5
B.3. Expression and purification of $^{15}\text{N}/^{13}\text{C}$ <i>holo</i> -PtlL	S5
B.4. Preparation of prolyl-PtlL, pyrrolyl-PtlL	S5
B.5. Preparation of <i>holo</i> -PtlL	S6
B.6. HPLC analysis of ACP proteins	S6
B.7. Protein NMR experiments	S6
B.8. PtlL 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 ClustaW	S10
Figure S3 - 20% UREA Page Gel of apo, holo, prolyl, and pyrrolyl PtlL	S11
Figure S4 - HPLC-MS spectra of prolyl- and pyrrolyl-PtlL species	S12-S14
Figure S5 - ^1H - ^{15}N HSQC of <i>holo</i> -PtlL	S15
Figure S6 - ^1H - ^{15}N HSQC of prolyl-N-PtlL	S16
Figure S7 - ^1H - ^{15}N HSQC of pyrrolyl-N-PtlL	S17
Figure S8 - ^1H - ^{15}N HSQC of ubiquitous ^{15}N -labelled <i>holo</i> -PtlL	S18
Figure S9 - ^1H - ^{13}C HSQC of <i>holo</i> -PtlL	S19
Figure S10 - ^1H - ^{13}C HSQC of ubiquitous ^{15}N , ^{13}C -labelled <i>holo</i> -PtlL	S20
Figure S11 - ^1H - ^{13}C HSQC of prolyl-PtlL	S21
Figure S12 - ^1H - ^{13}C HSQC of pyrrolyl-PtlL	S22
Figure S13 - $^{13}\text{C}/^{15}\text{N}$ F1F2-filtered NOE spectrum of <i>holo</i> -PtlL	S23
Figure S14 - $^{13}\text{C}/^{15}\text{N}$ F1F2-filtered NOE spectrum of pyrrolyl-N-PtlL	S24
Figure S15 - Interaction between PtlL and 4-DMN pantetheine probe	S25
Figure S16 - ^1H - ^1H NOE spectra of pyrrolyl-N-PtlL focused on the protons of the pyrrole	S26
Figure S17 - Perturbations and NOEs of PtlL due to pyrrole	S27
Figure S18 - Solution NMR structures of PtlL species	S28
Figure S19 - Structure comparison of PtlL to other peptidyl carrier proteins	S29
Figure S20 - Electrostatic comparison of PtlL to other peptidyl carrier proteins	S30
Figure S21 - NOEs between PtlL and the pyrrole	S31
Figure S22 - ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 3 in CDCl_3	S32
Figure S23 - ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 4 in CD_3OD	S33
Figure S24 - ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 6 in CDCl_3	S34
Figure S25 - ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of 7 in CDCl_3	S35
Supplemental References	S36

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.

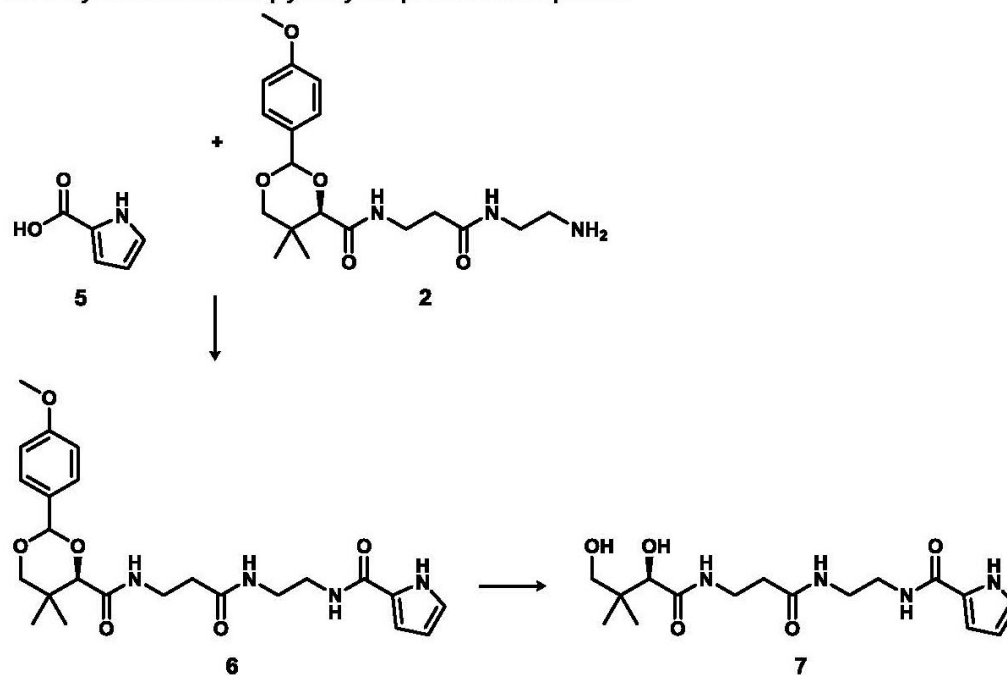
A.1. Synthesis of the prolyl-N-pantetheine probe.



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 μ L) 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%). $^1\text{H-NMR}$ (400 MHz, CDCl_3) 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). $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) 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-2-carboxamide (4): Amide 3 (167.1 mg, 0.290 mmol) was dissolved in 2 mL DCM and 400 μ L 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%). $^1\text{H-NMR}$ (400 MHz, CD_3OD) 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). $^{13}\text{C-NMR}$ (100 MHz, CD_3OD) 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 $[\text{M}+\text{H}]^+$ 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-diisopropylethylamine (DIPEA) (77 μ L) 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%). $^1\text{H-NMR}$ (400 MHz, CDCl_3) 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). $^{13}\text{C-NMR}$ (100 MHz, CDCl_3) 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%). $^1\text{H-NMR}$ (400 MHz, $(\text{CD}_3)_2\text{SO}$) 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). $^{13}\text{C-NMR}$ (125 MHz, $(\text{CD}_3)_2\text{SO}$) 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 $[\text{M}+\text{H}]^+$ calcd. 355.1976, found 259.1975.

B. Protein NMR Studies.

B.1. Site directed mutagenesis of PitL 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'-GGCCGAGTCCGCCCTCGAGCACCACCACC-3'

C87S_R: 5'-CTCGAGGGCGGACTCGGCCTTTAGTTGCTCG-3'

B.2. Expression and purification of $^{15}\text{N}/^{13}\text{C}$ apo-PitL

pET22b-PitL (C87S) plasmid was transformed into BL21 (DE3) cells. Uniformly labeled ^{15}N -PitL was expressed by culturing cells in M9 minimal media (1 L) supplemented with 1 g $^{15}\text{N-NH}_4\text{Cl}$, 4 g D-glucose, and 50 mg carbenicillin. Uniformly labeled ^{13}C , ^{15}N -PitL was prepared by supplementing M9 minimal media (1 L) with 4 g ^{13}C -D-glucose, 1 g $^{15}\text{N-NH}_4\text{Cl}$, and 50 mg carbenicillin. Expression was induced with 0.5 mM IPTG at an OD_{600} 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-PitL.

B.3. Expression and purification of $^{15}\text{N}/^{13}\text{C}$ holo-PitL

pET22b-PitL (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-PitL was expressed and purified as mentioned in section B.2. holo-PitL was isolated from Sfp (no his-tag) during Ni-NTA purification.

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

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-PitL by Sfp.^{S2} In this study, this was achieved by treating 1.5 ml of

400 μM *apo*-PitL with 12.5 mM MgCl_2 , 8 mM ATP, 2 mM mimetic 4 and 7, 1.5 μM Sfp, 0.5 μM *E. coli* CoaA, 0.7 μM *E. coli* CoaD, 0.6 μM *E. coli* CoaE, 0.02% Triton X, 0.1% NaN_3 , 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*-PitL species eluted between 155-180 mL.

B.5. Preparation of *holo*-PitL

To prepare pure *holo*-PitL, *apo*-PitL from section C.3 was incubated with 50 mM Tris pH 7.5, 150 mM NaCl, 12.5 mM MgCl_2 , 2 mM coenzyme A, 3 mM TCEP, 0.1% NaN_3 and 1.5 μM 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*-PitL 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 μL aliquot of the appropriate PitL form was then prepared for NMR by adding 50 μL of D_2O , 5 μL of 10% (w/v) NaN_3 and 5 μL of 0.5 M TCEP adjusted to pH 7.4. The final concentration of the samples evaluated was as follows: 1 mM *holo*-PitL; 1 mM prolyl-PitL; 1 mM pyrrolyl-PitL. 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 ^{15}N indirect dimension. Chemical shifts were measured by peak maxima, and chemical shift perturbation was calculated using the formula $\text{CSP} = ((0.2\delta\text{N})^2 + (\delta\text{H})^2/2)^{0.5}$.

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 ^{13}C and ^{15}N 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.

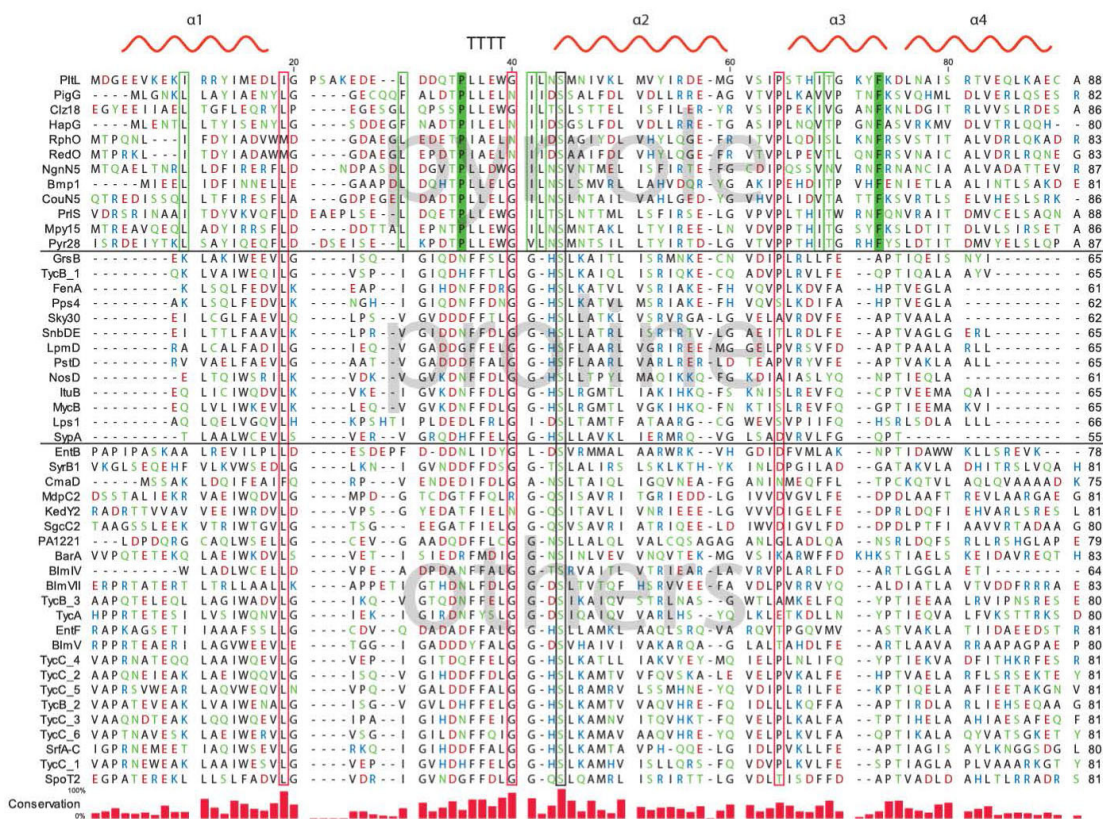
B.8. PitL Protein NMR Data Collection and Structure Calculations

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 ^{13}C assignments were obtained from ^{13}C -edited NOESY-HSQC and TOCSY-HSQC experiment (mixing time 60 ms). NOE data was obtained from 3D ^{15}N -edited and ^{13}C -edited NOESY-HSQC experiments (mixing time 150 μs). 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}

Table S1. Solution NMR Structure Statistics

Assignment Statistics (%) (residues 1-90)	<i>holo</i>	<i>pyrrolyl</i>
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-j \leq 1$	2217	2584
medium range, $1 < i-j < 5$	476	381
long-range, $ i-j \geq 5$	330	310
Dihedral Angle Restraints	139	130
Structure Statistics		
Average final CYANA target function (\AA^2)	1.32	1.2
Average cycle1 CYANA target function (\AA^2)	40.10	64.27
CNS energies (kcal/mol)	-2949.35	-2932.7
Restraint Violations		
Max. distance restraint violation (\AA)	0	0.07
Violated distance restraints $> 0.1 \text{\AA}$	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	$0.40 \pm 0.09 \text{\AA}$	$0.55 \pm 0.11 \text{\AA}$
Average heavy atom RMSD to mean	$0.84 \pm 0.08 \text{\AA}$	$1.06 \pm 0.12 \text{\AA}$



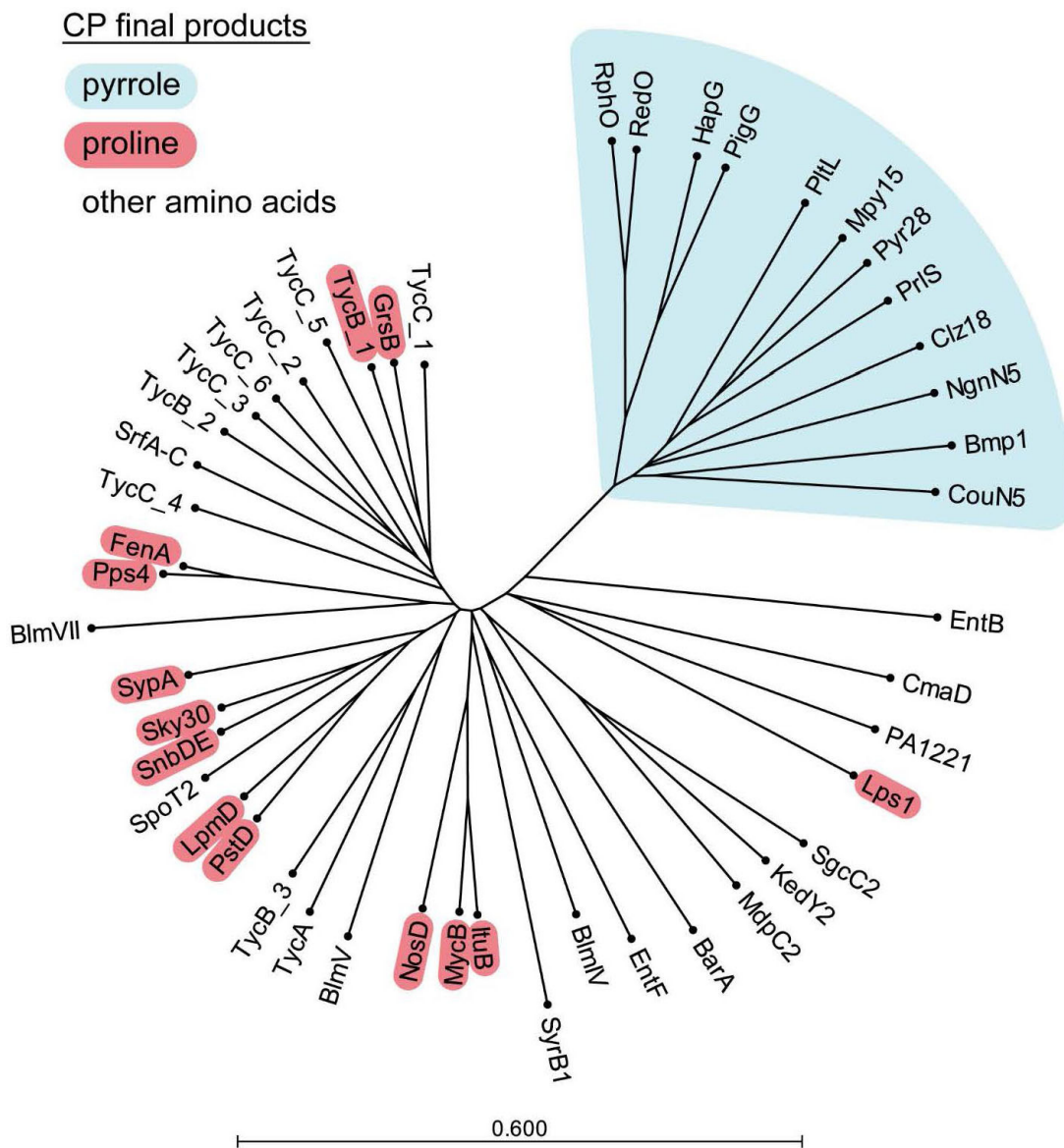


Figure S2 - Phylogenetic tree of PCPs generated with ClustalW.^{S12}

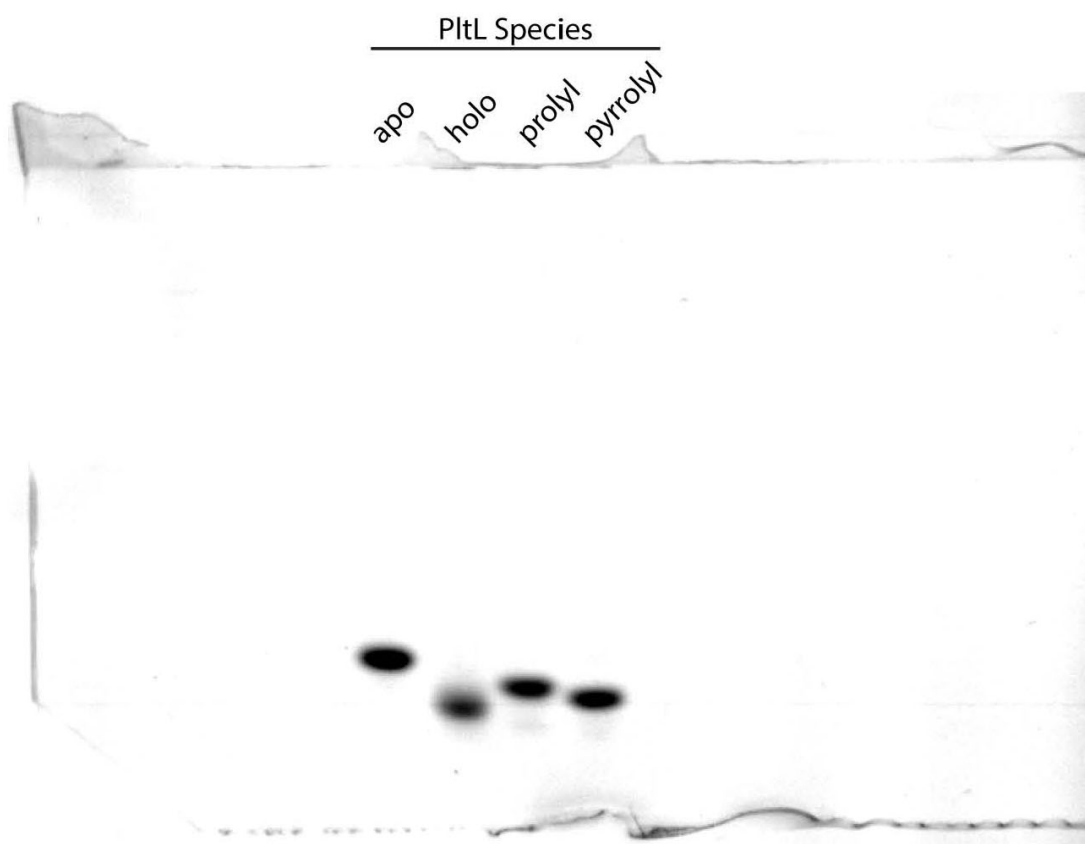


Figure S3 - 20% UREA Page Gel of apo, holo, prolyl, and pyrrolyl PitL. The PitL Species were stained with coomassie dye. To note, *apo* PitL runs slower in the gel possible due to having a more positive charge (does not contain the Ppant) and/or a more relaxed conformation, which is common for carrier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

Figure S4A – HPLCMS data for prolyl-PitL.

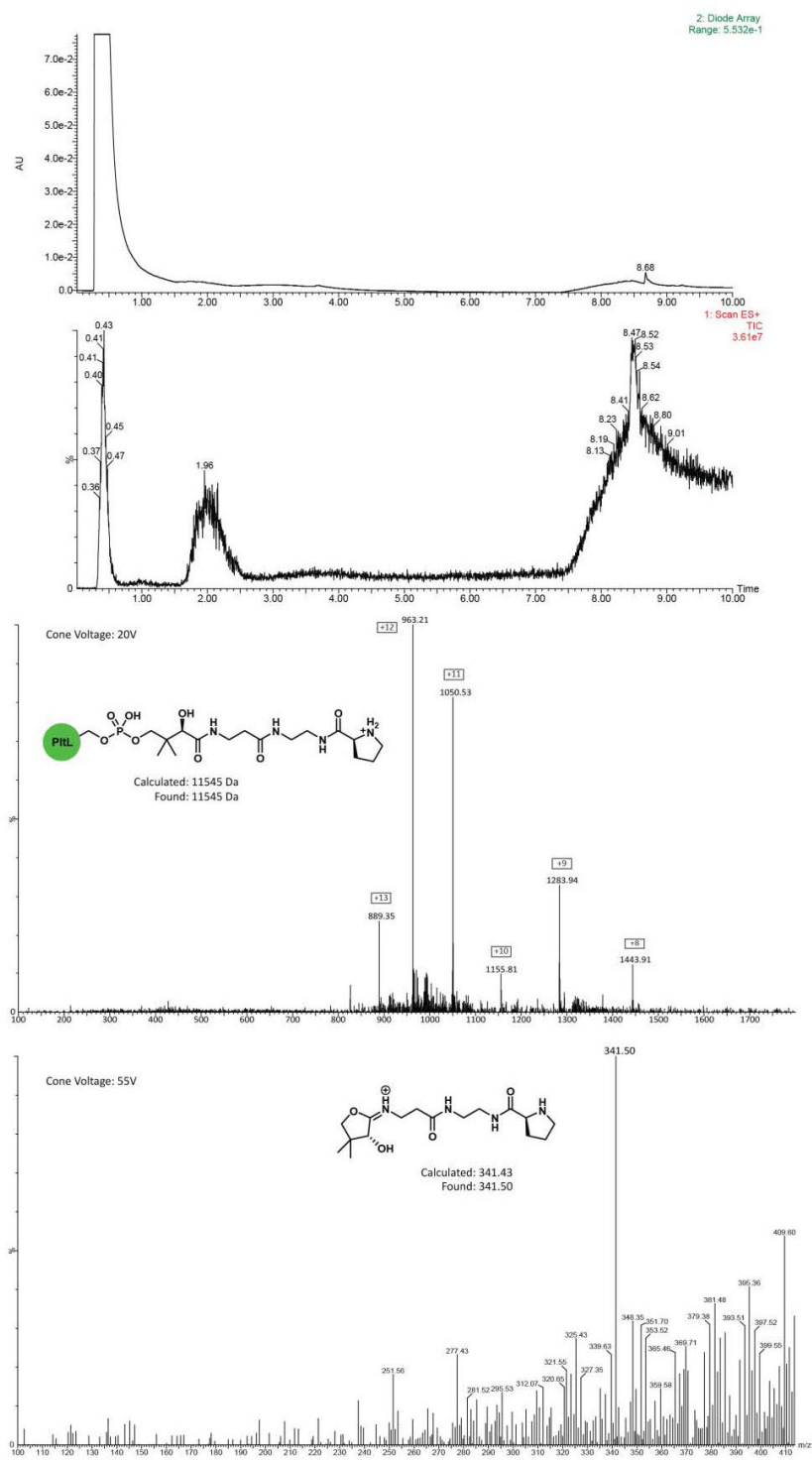


Figure S4B – HPLCMS data for pyrrolyl-PIIL.

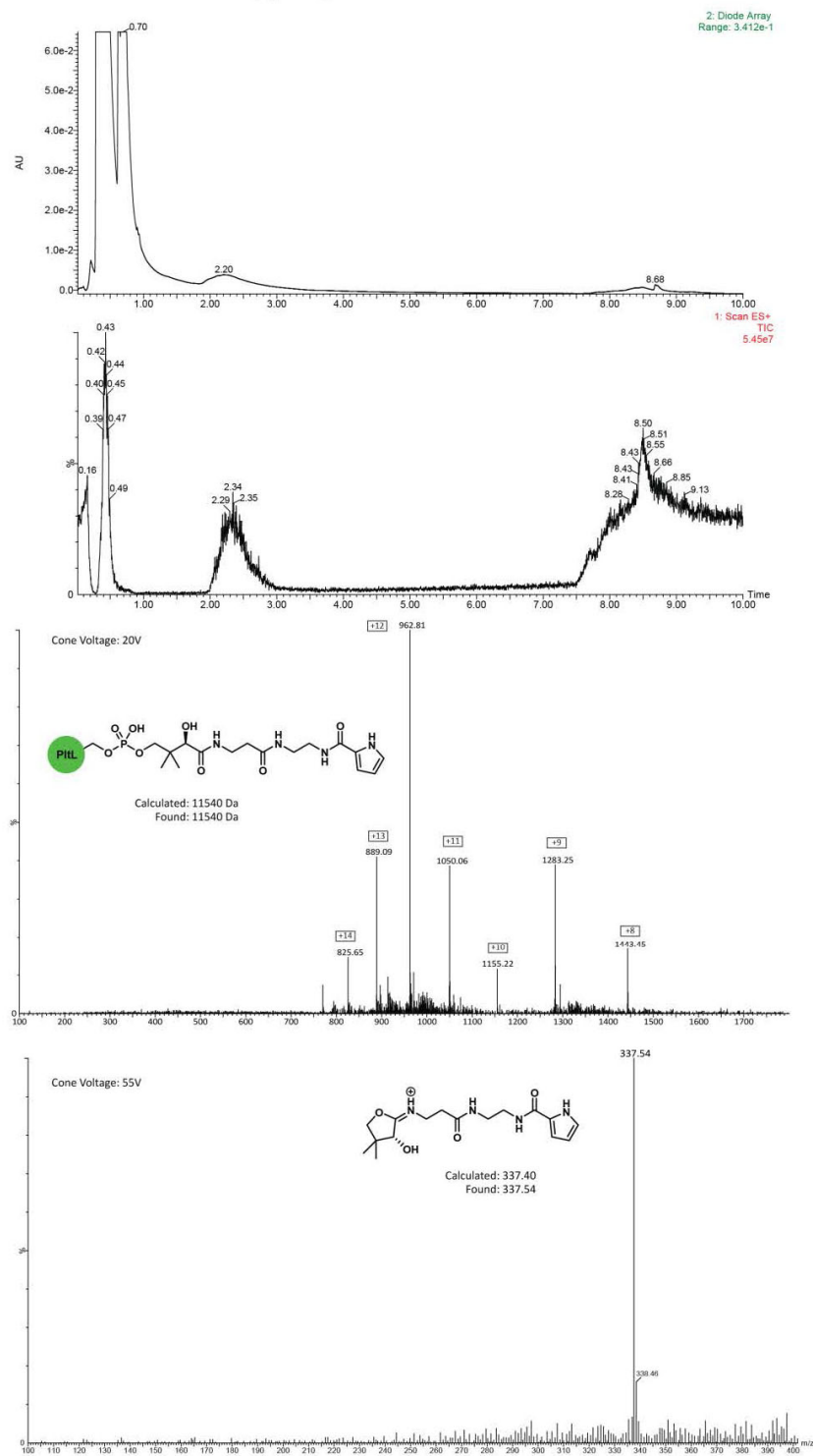


Figure S4 – HPLC-MS spectra of prolyl- and pyrrolyl-PitL species. PitL species from in vitro loading (see section C.4) were analyzed by HPLC-MS. HPLC-MS analysis was performed on a Waters Acquity system using an Acquity binary solvent manager, an Acquity column manager, a 2777c robotic autosampler, an Acquity TUV detector and an SQ detector. Samples in buffer (2 μ l) were separated on a Waters C18 Xbridge (3.0 x 50mm, 2.5 μ) 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 re-equilibrating 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/prot-param/>) 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-PitL. The mass spectra 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-PitL. 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.

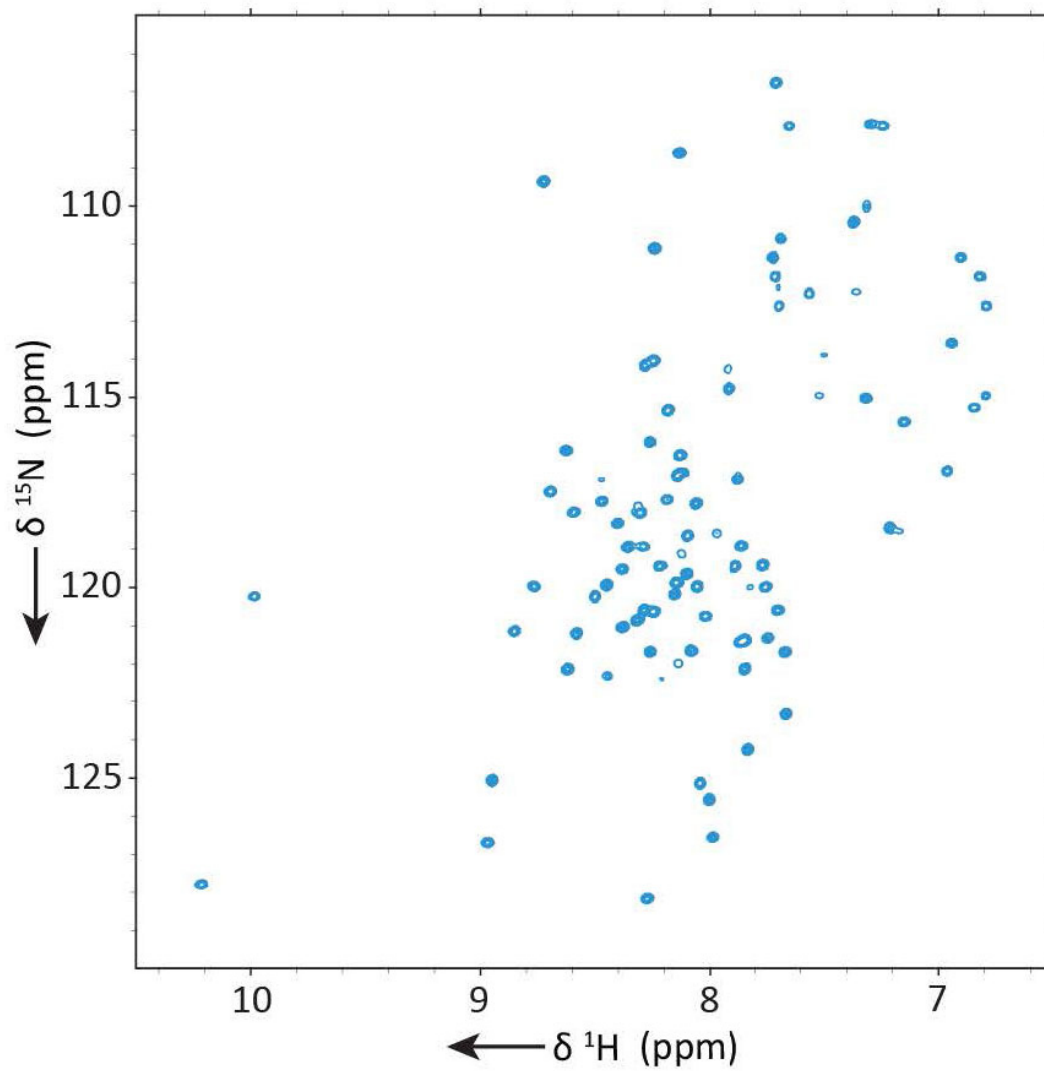


Figure S5 - ^1H - ^{15}N HSQC of *holo*-PitL.

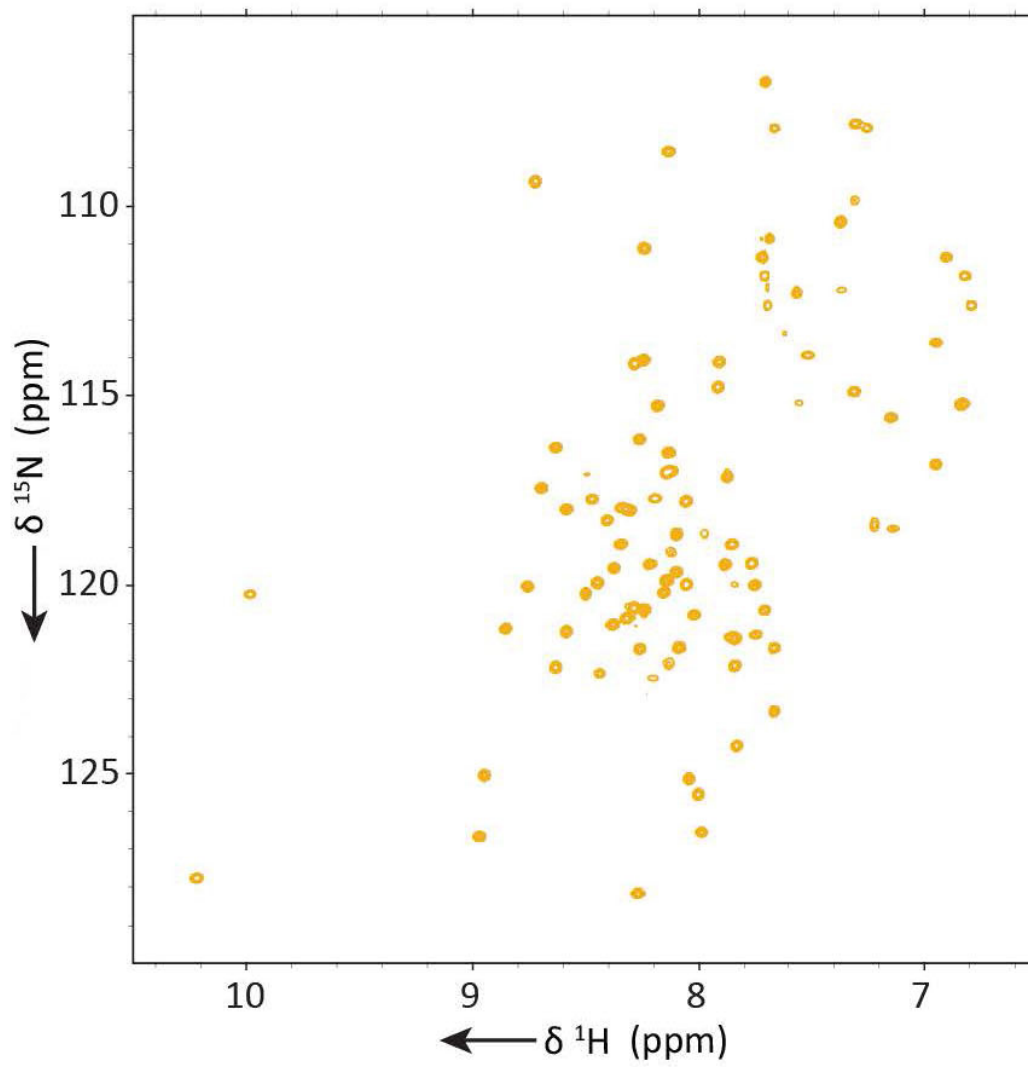


Figure S6 - ^1H - ^{15}N HSQC of poly(l-N-PiL).

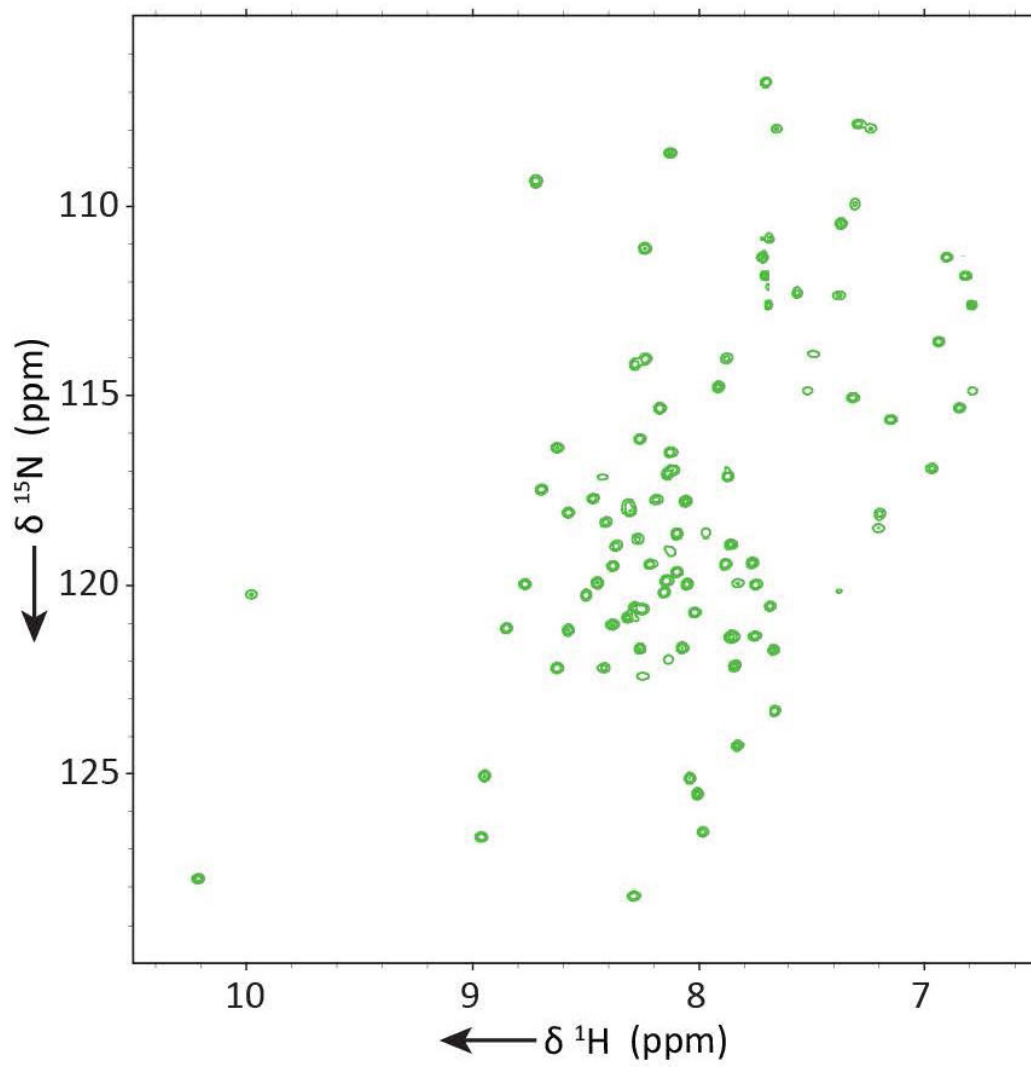


Figure S7 - ^1H - ^{15}N HSQC of pyrrolyl-N-PitL.

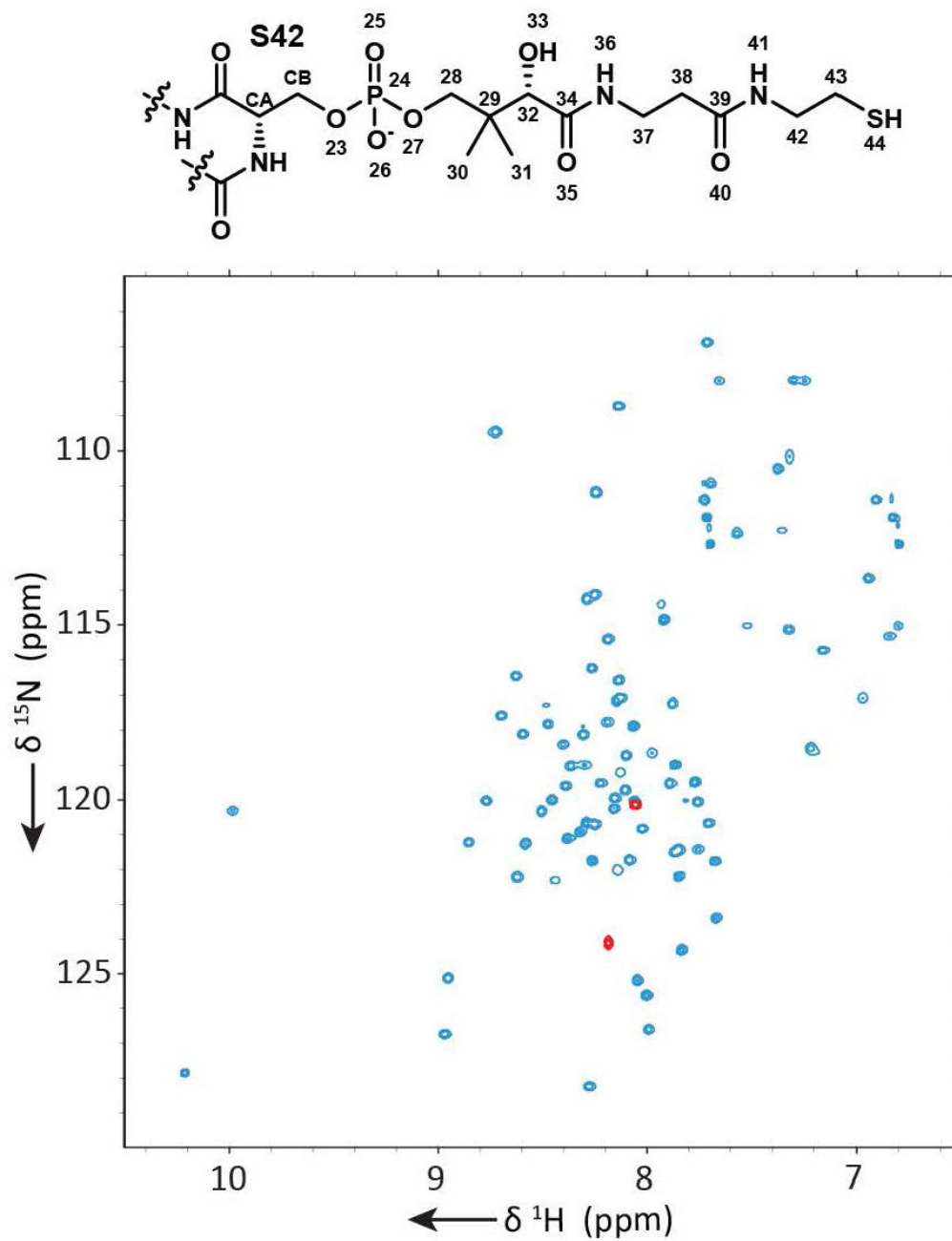


Figure S8 - ^1H - ^{15}N HSQC of ubiquitous ^{15}N -labelled *holo*-PitL. The red peaks correspond to the phosphopantetheine amides (H-N36: δ 8.05, 120.1; H-N41: δ 8.19, 123.9). The phosphopantetheine arm was ^{15}N -labelled by co-expressing PitL and Sfp in BL21 cultured in ^{15}N minimal media.

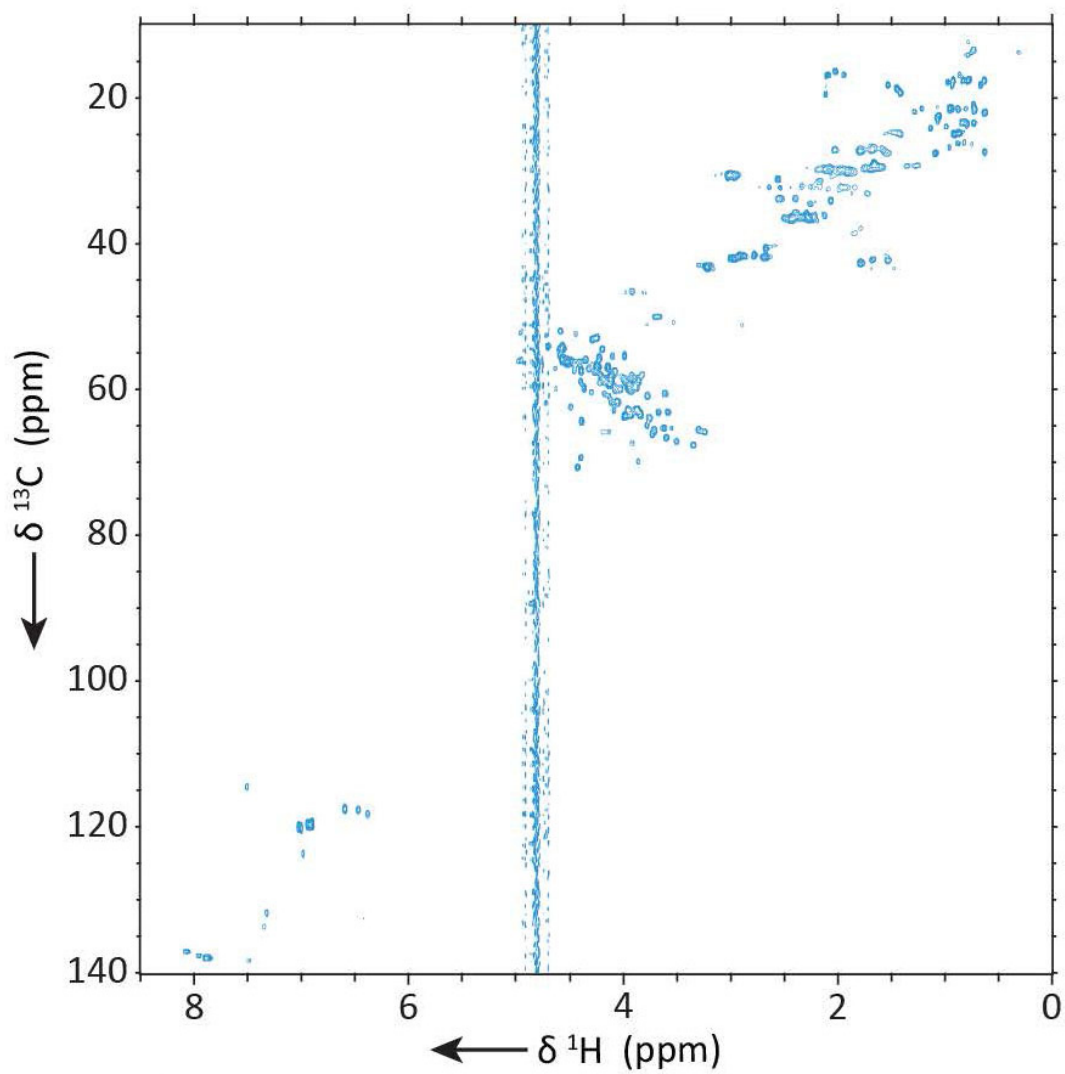


Figure S9 - ^1H - ^{13}C HSQC of *holo*-PitL.

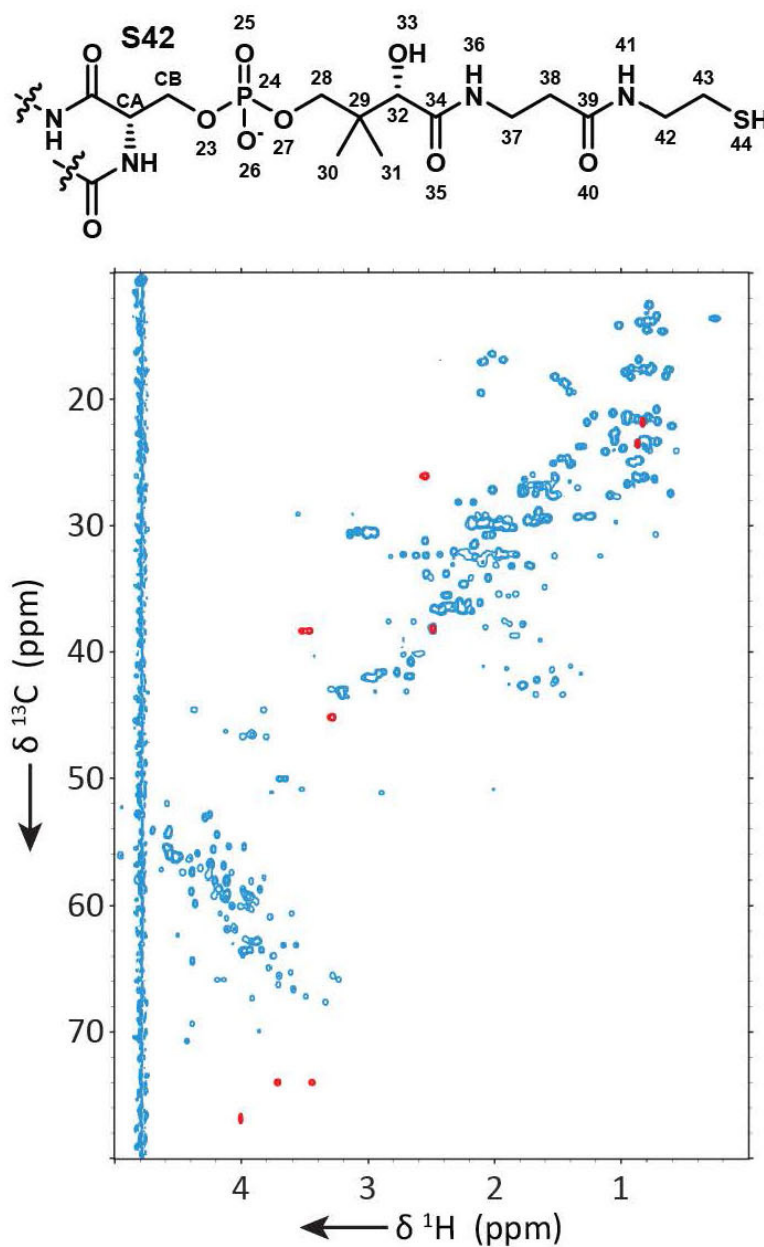


Figure S10 - ^1H - ^{13}C HSQC of ubiquitous ^{15}N , ^{13}C -labelled *holo*-PltL. The red peaks correspond to the phosphopantetheine aliphatic proton-carbon pairs (H-H-C28: δ 8.45, 3.72, 73.9; H-C30: δ 0.83, 21.8; H-C31: δ 23.53, 0.87; H-C32: δ 4.00, 76.9; H-H-C37: δ 3.47, 3.53, 38.3; H*-C38: δ 2.49,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

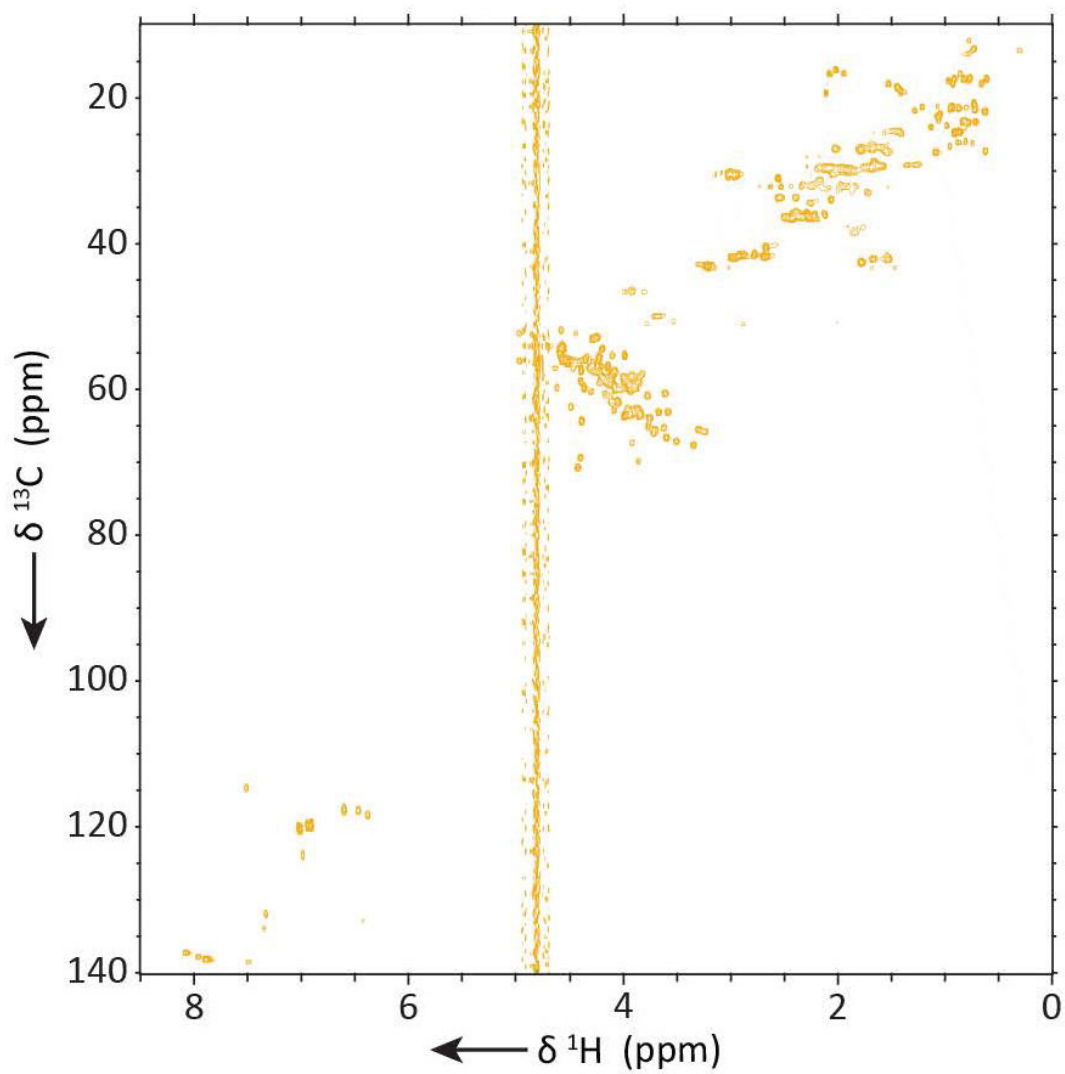


Figure S11 - ^1H - ^{13}C HSQC of prolyl-PitL.

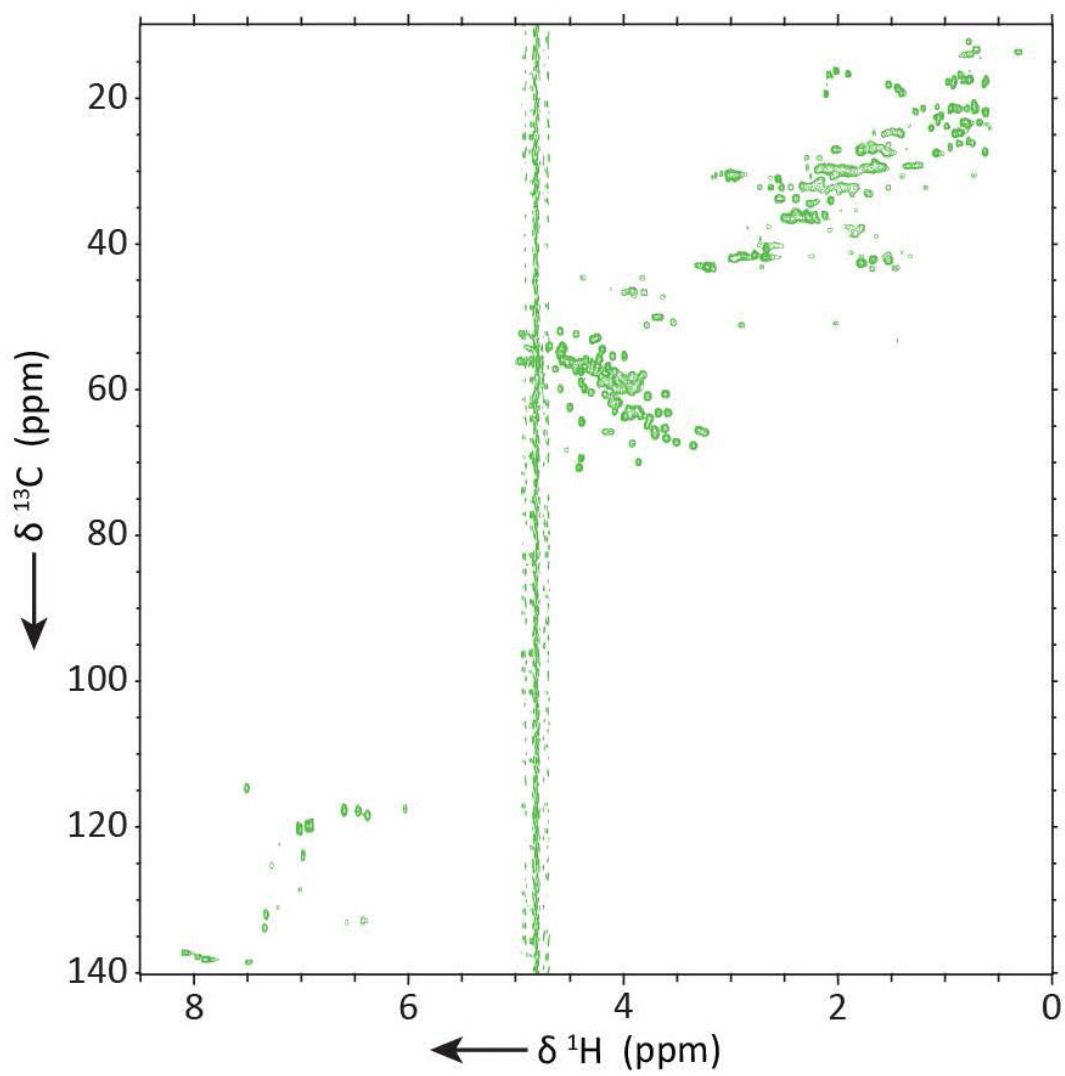


Figure S12 - ^1H - ^{13}C HSQC of pyrrolyl-PitL.

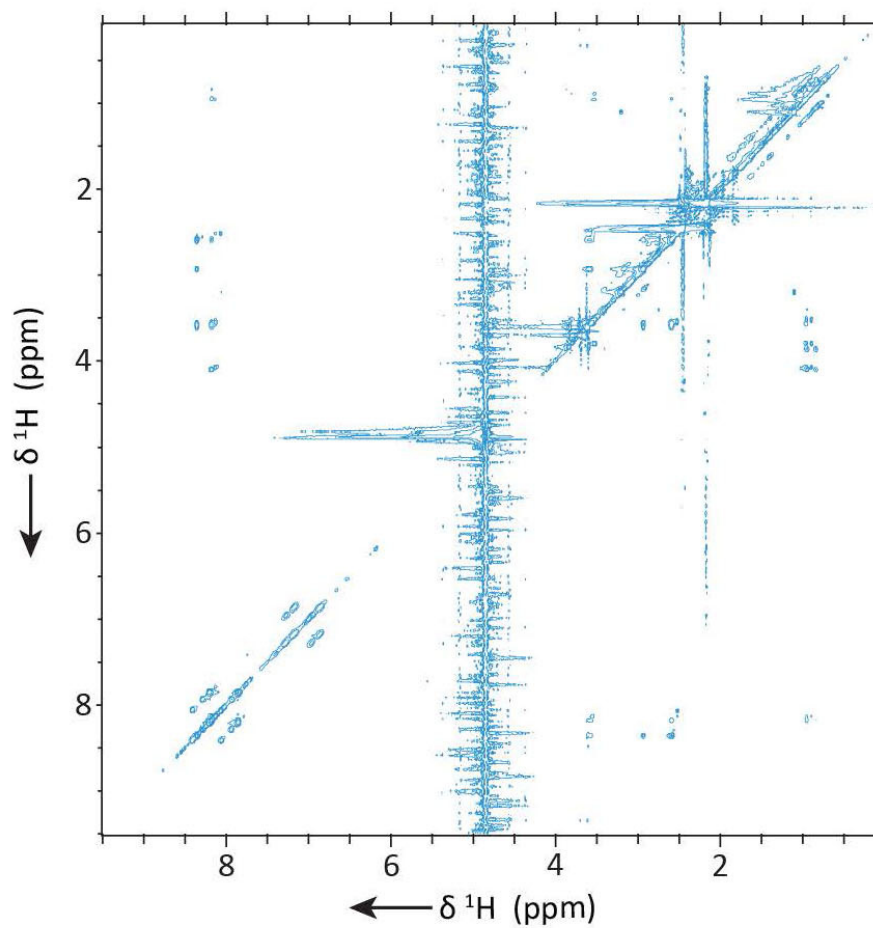


Figure S13 - $^{13}\text{C}/^{15}\text{N}$ F1F2-filtered NOE spectrum of *holo*-PitL. Only intramolecular NOEs of the unlabeled phosphopantetheine are visible.

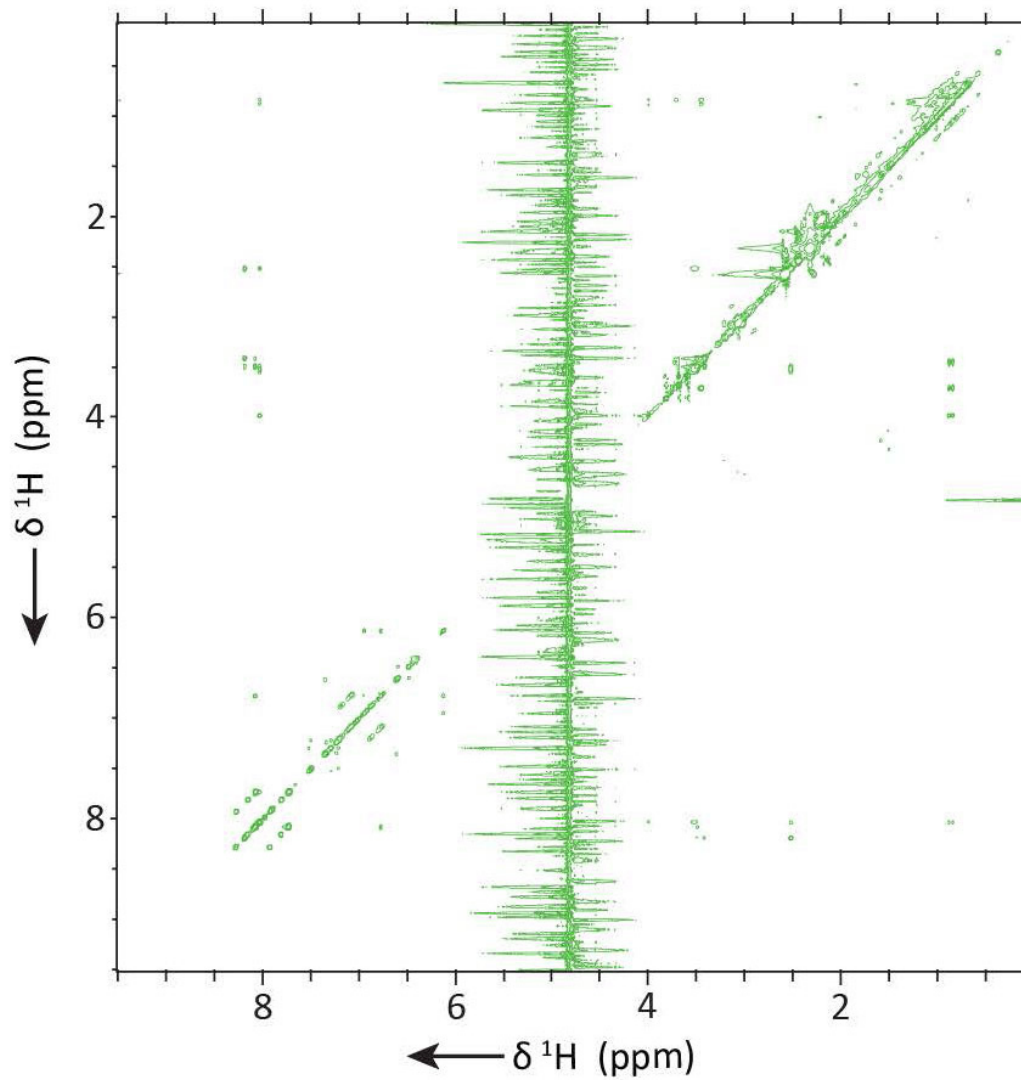


Figure S14 - $^{13}C/^{15}N$ F1F2-filtered NOE spectrum of pyrrolyl-N-PitL. Only intramolecular NOEs of the unlabeled phosphopantetheine-N-pyrrolyl probe are visible.

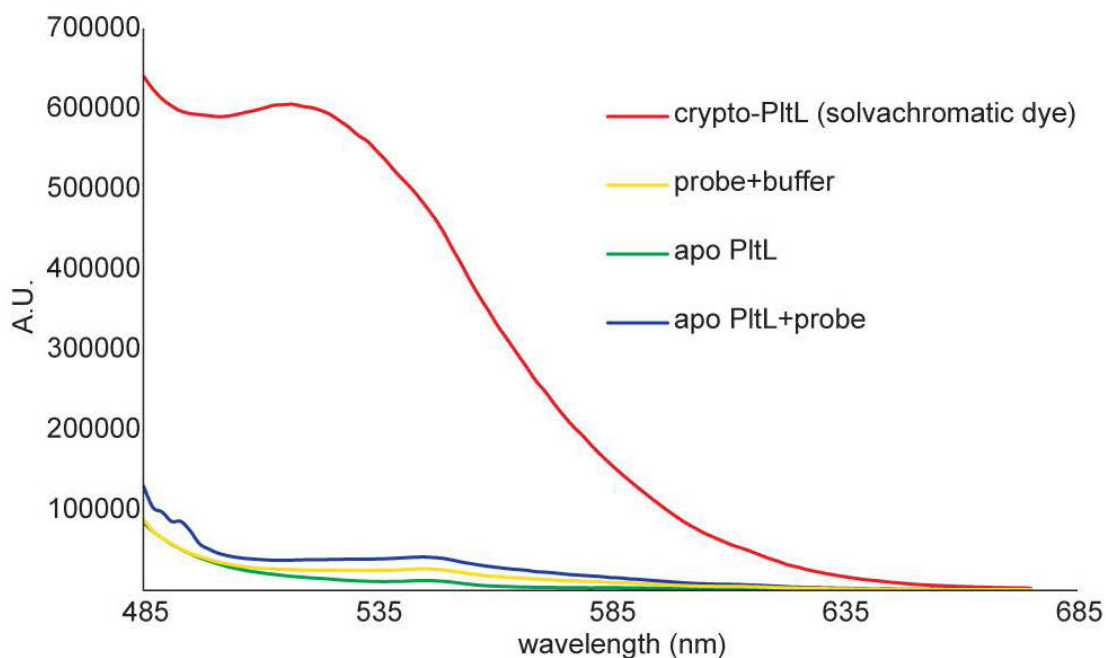


Figure S15 - The interaction of PitL and attached solvatochromic 4-DMN (4-N,N-dimethyl-amino-1,8-naphthalimide) pantetheine probe. We recently developed solvatochromatic pantetheine probes to observe interactions between attached substrates and carrier proteins. 4-DMN is fluorescent when exposed to a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

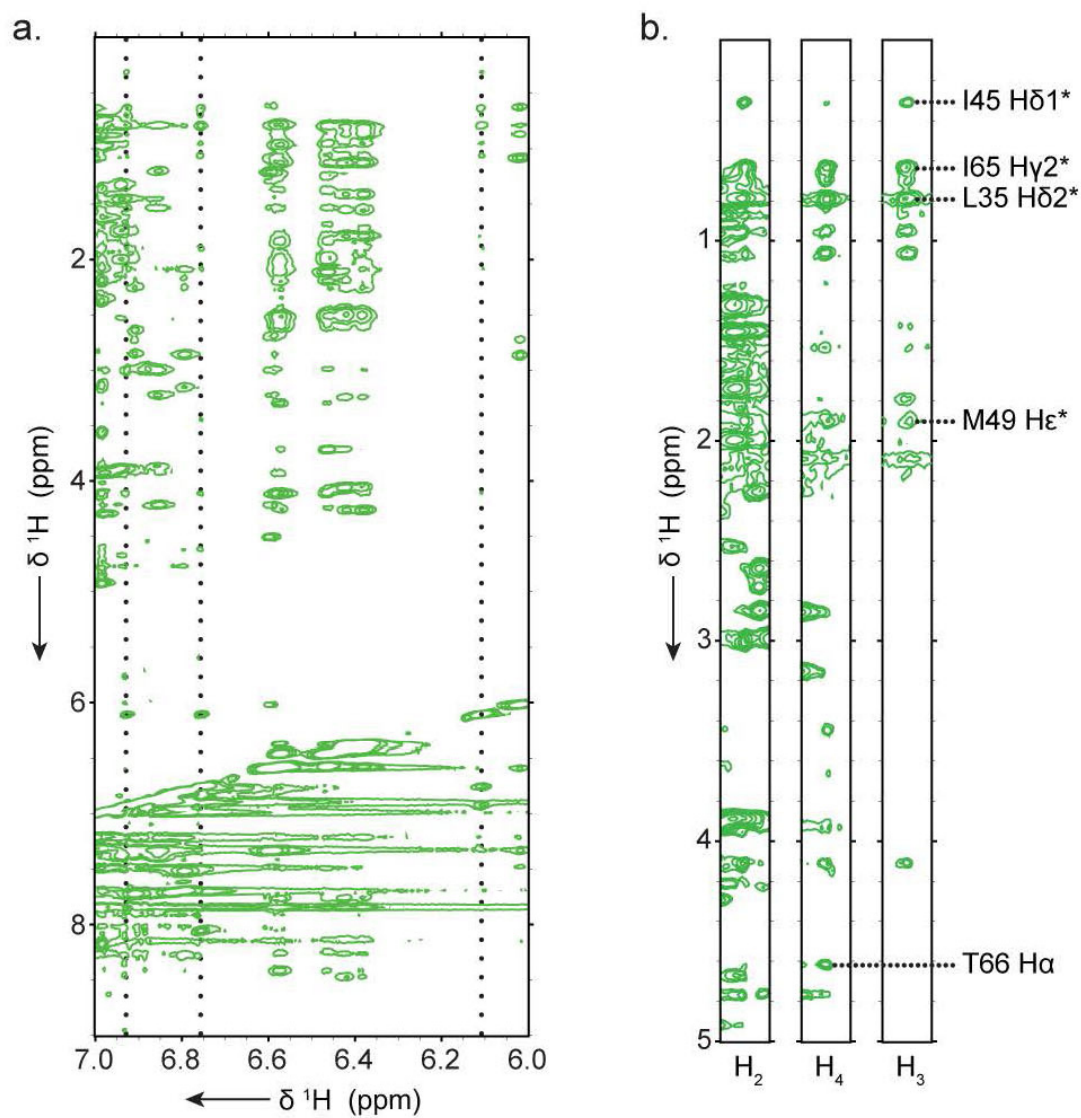


Figure S16 - ^1H - ^1H NOE spectra of pyrrolyl-N-PitL focused on the protons of the pyrrole. a, Pyrrolyl-N-PitL dissolved in $\text{H}_2\text{O}/\text{KPi}$ buffer. Dotted lines indicate the resonances of the pyrrole protons (δ 6.93, 6.76, 6.11). b, NOE strips of the pyrrole protons with peaks corresponding to atoms in PitL labelled.

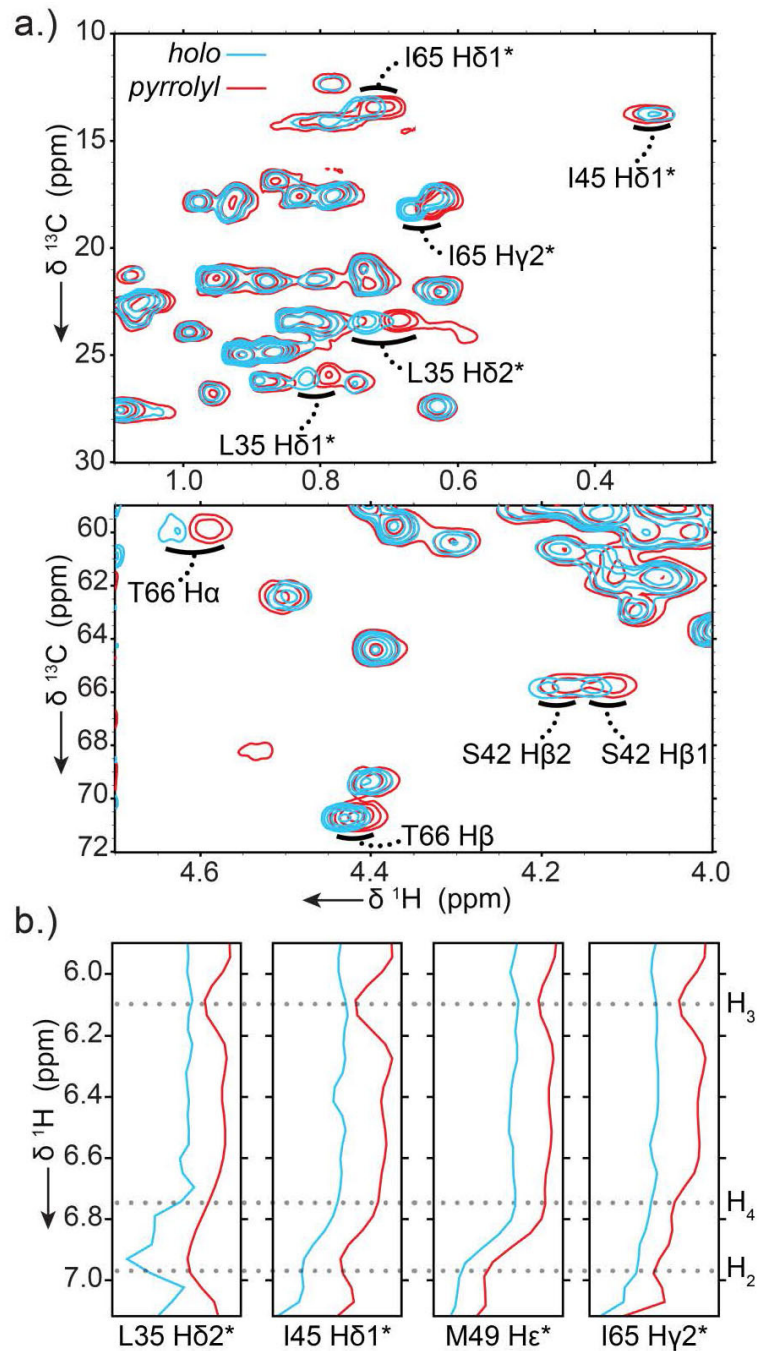


Figure S17 – Perturbations and NOEs of PitL due to pyrrole. a, Chemical shift perturbations due to pyrrole attachment of key atoms in PitL. Shown are methyl (top) and alpha (bottom) regions of overlaid *holo* (blue) and *pyrrolyl* (red) ^1H - ^{13}C HSQCs. b, 1D slices of L35 H δ 2*, I45 H δ 1*, M49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

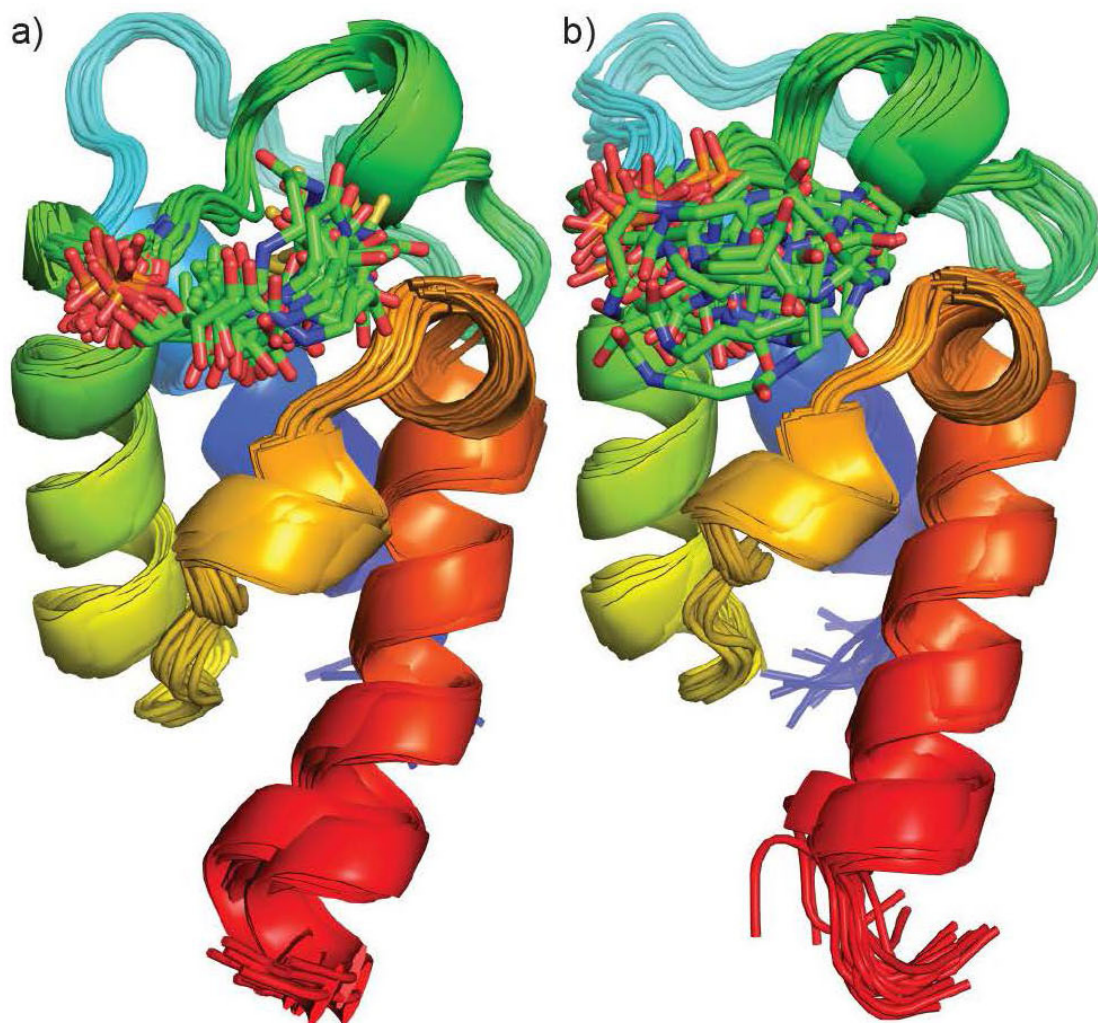


Figure S18 - Solution NMR structures of PtlL species, represented as a family of structures with the 20 lowest CYANA target functions. a, *holo*-PtlL. b, Pyrrolyl-PtlL.

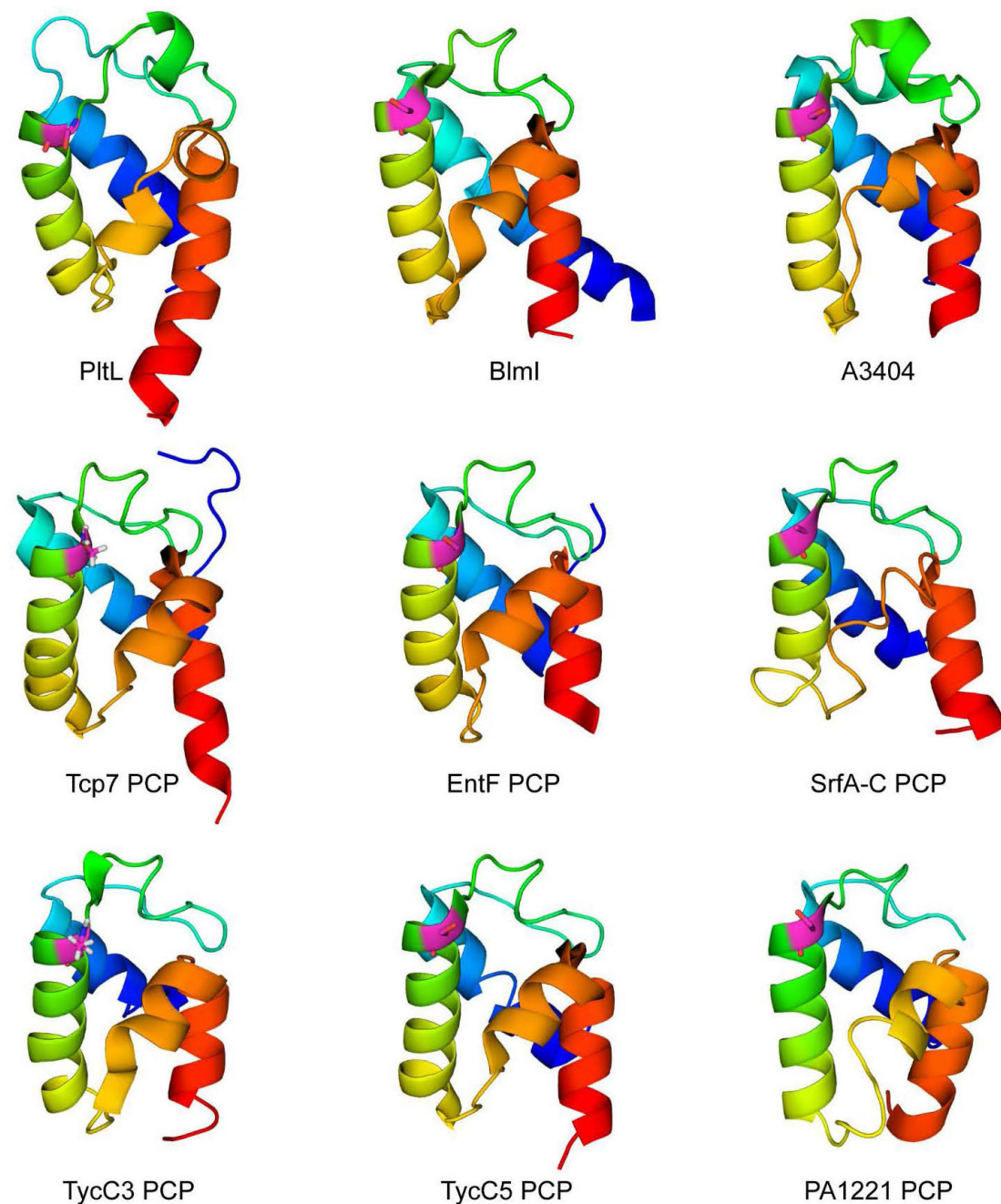


Figure S19 - Structure comparison of PltL to other peptidyl carrier proteins. The top row consists of all type II PCPs structure determined. PDB IDs: PltL (2n5h), 4neo (BlmI), 4hkg (A3404), 2mr8 (Tcp7 PCP), 3tej (EntF PCP), 2vsq (SrfA-C PCP), 4mrt (TycC3 PCP), 2jgp (TycC5 PCP), 4dg9 (PA1221). The conserved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

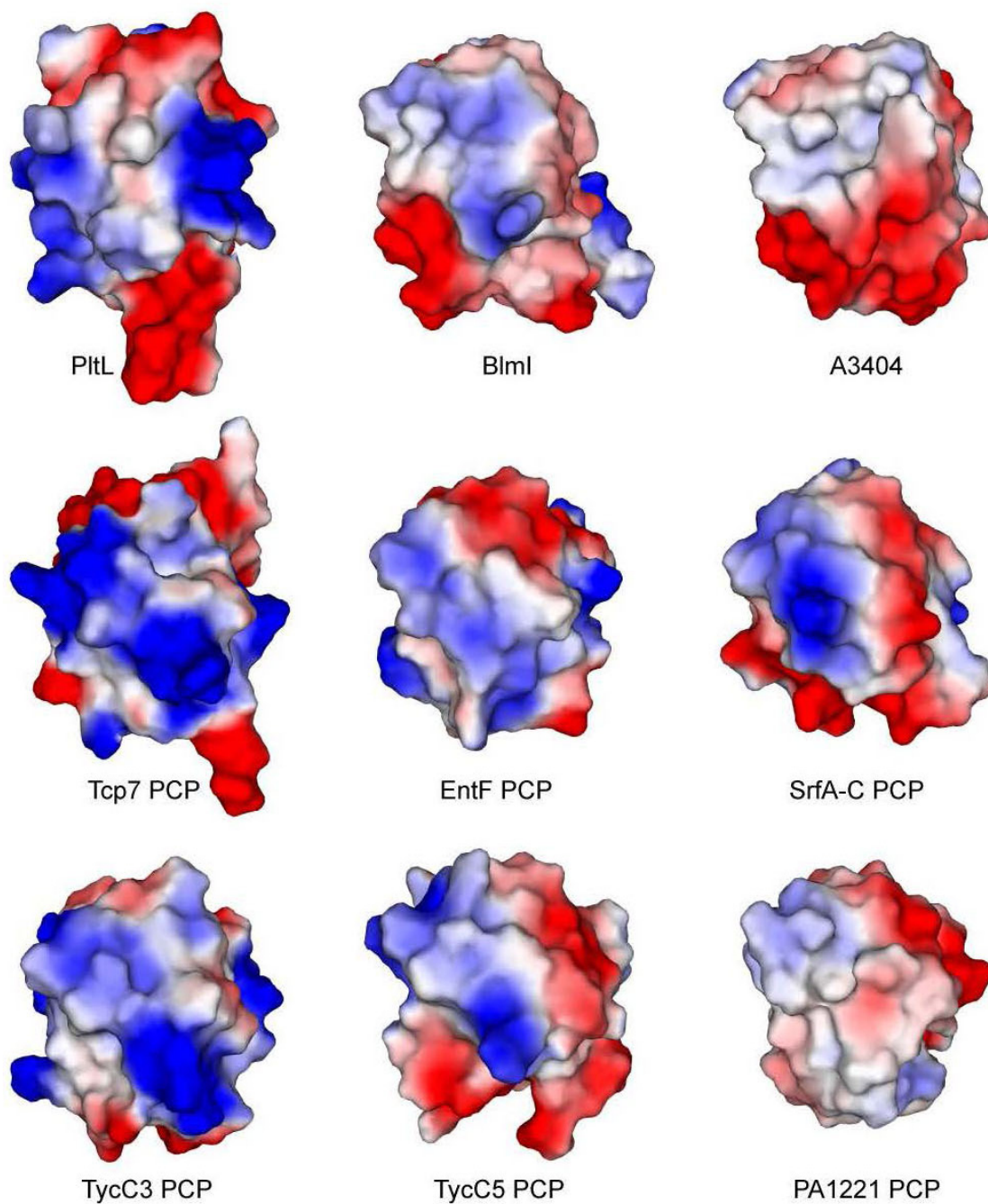


Figure S20 - Electrostatic comparison of PltL to other peptidyl carrier proteins. The top row consists of all type II PCPs structure determined. Carrier proteins are in the same orientation and order as Fig. S19. Red, white, and blue are negative, neutral, and positive charged, respectively.

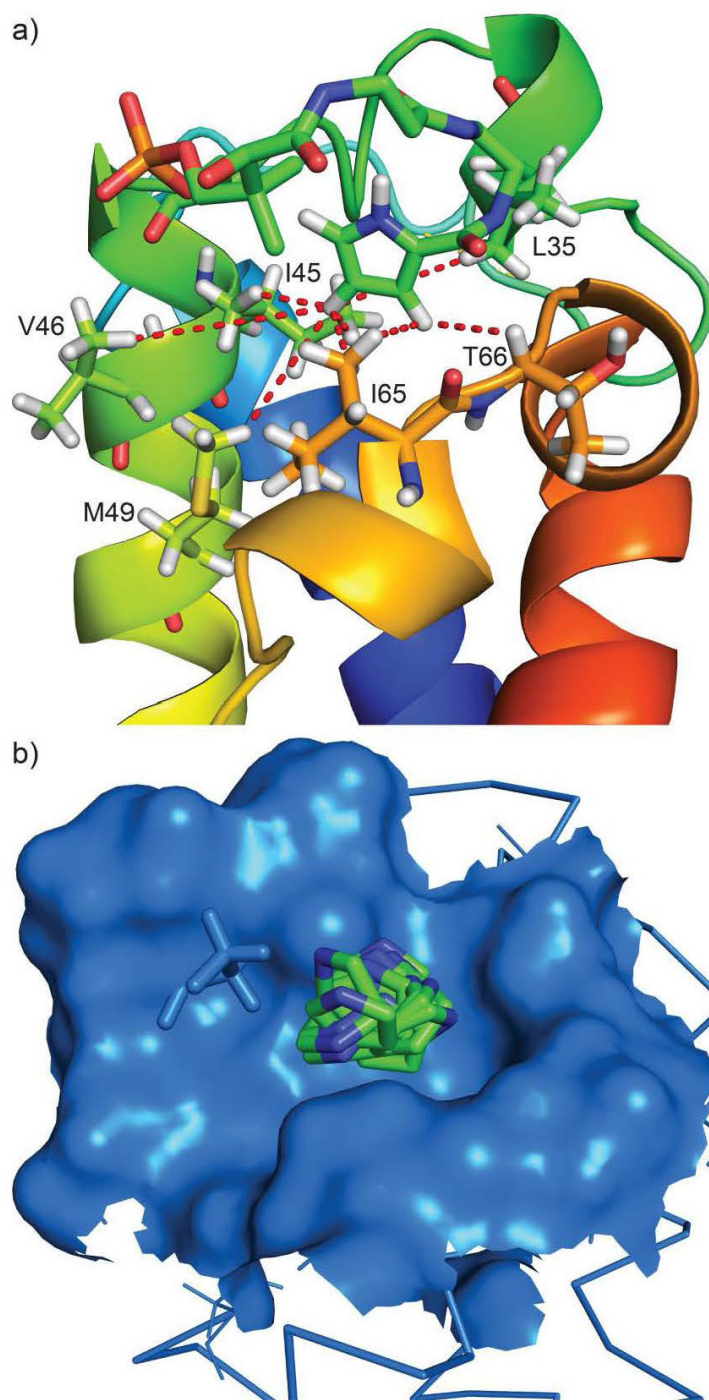


Figure S21 – NOEs between PitL and the pyrrole. a, Key NOEs are shown between the pyrrole and PitL. b, Pyrrole sequestration into the helix II/III hydrophobic pocket of PitL. 20 pyrrole rings are shown from the 20 lowest target function structures from CYANA.

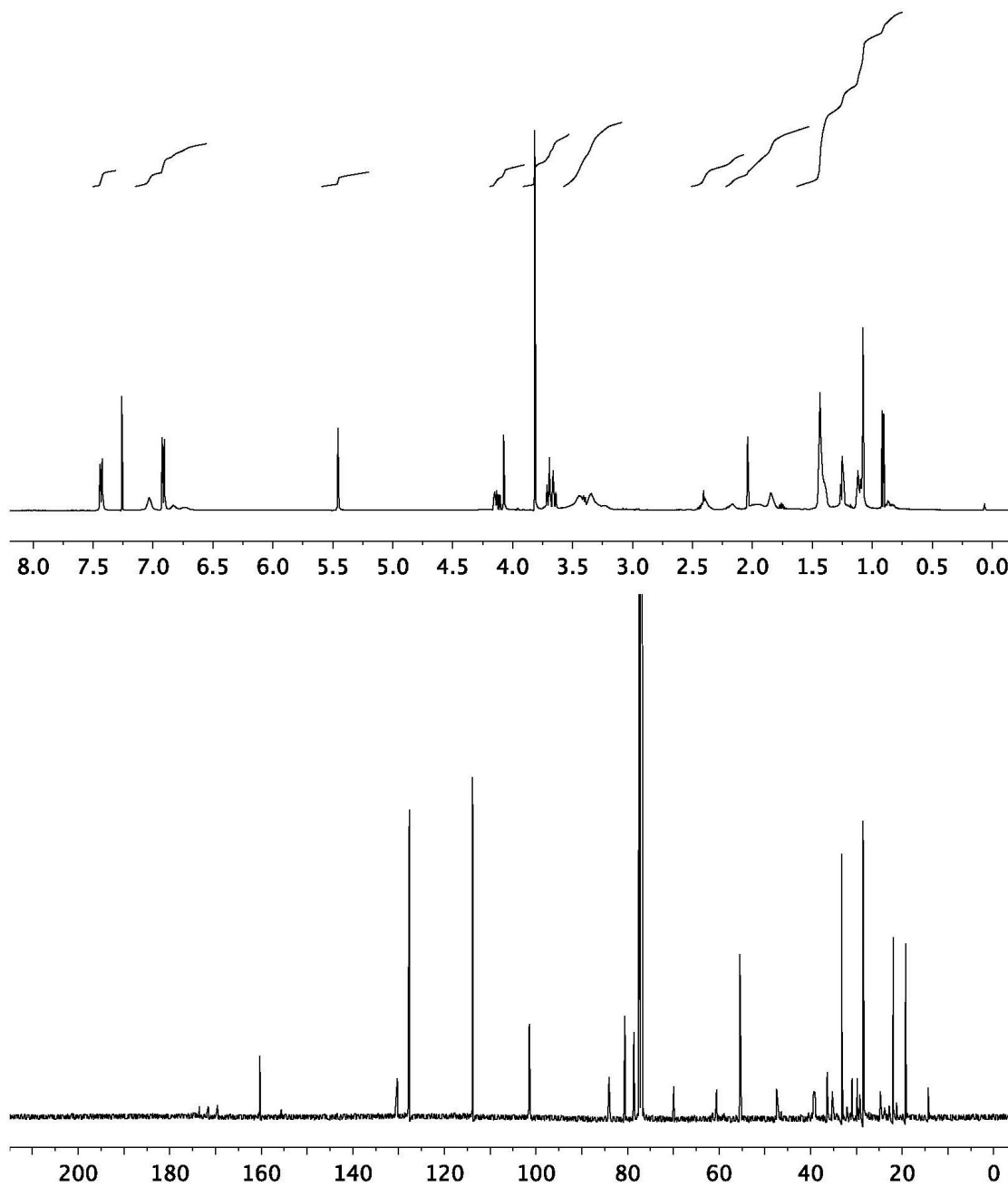
Figure S22. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **3** in CDCl_3 

Figure S23. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **4** in CD_3OD .

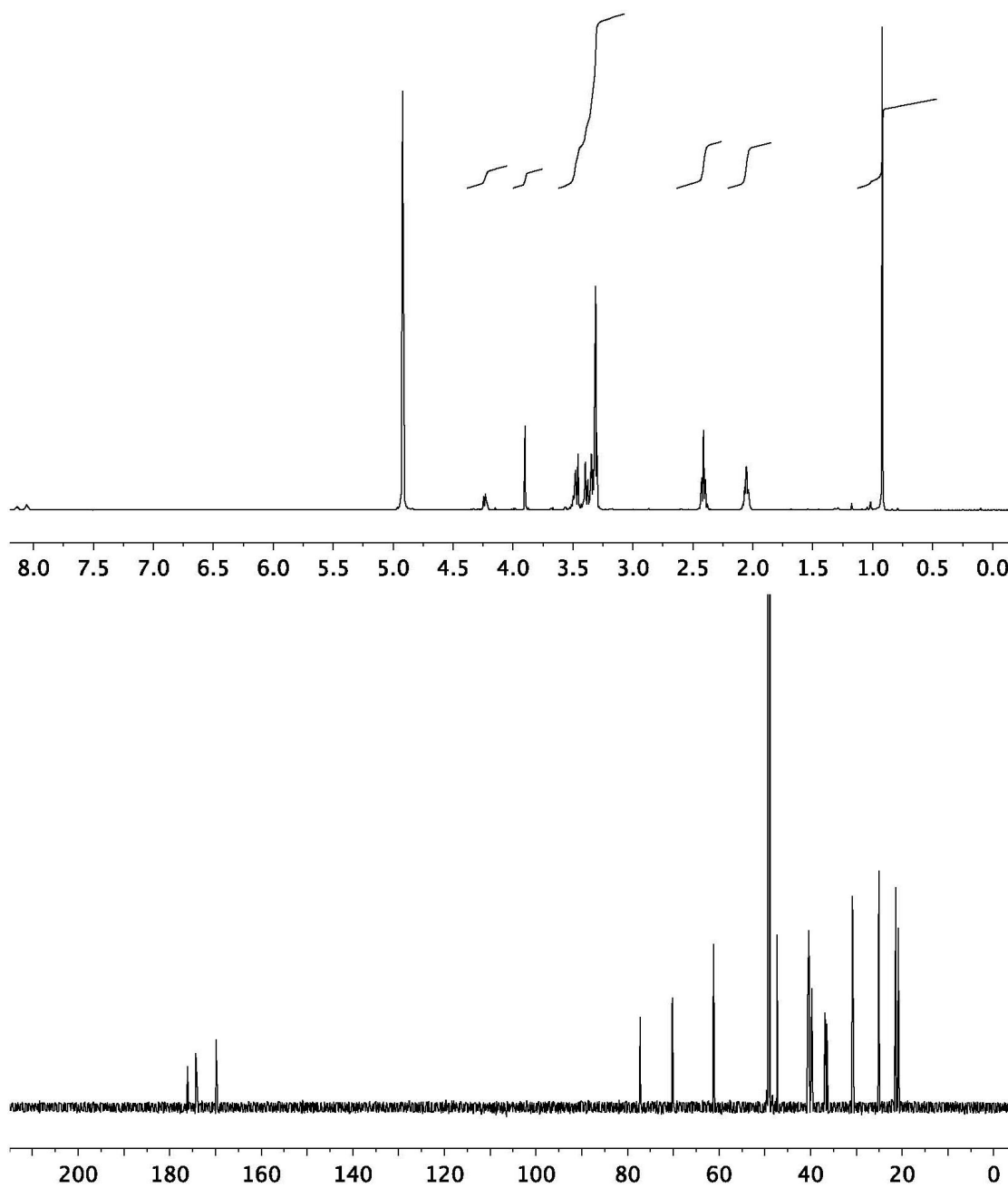


Figure S24. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **6** in CDCl_3

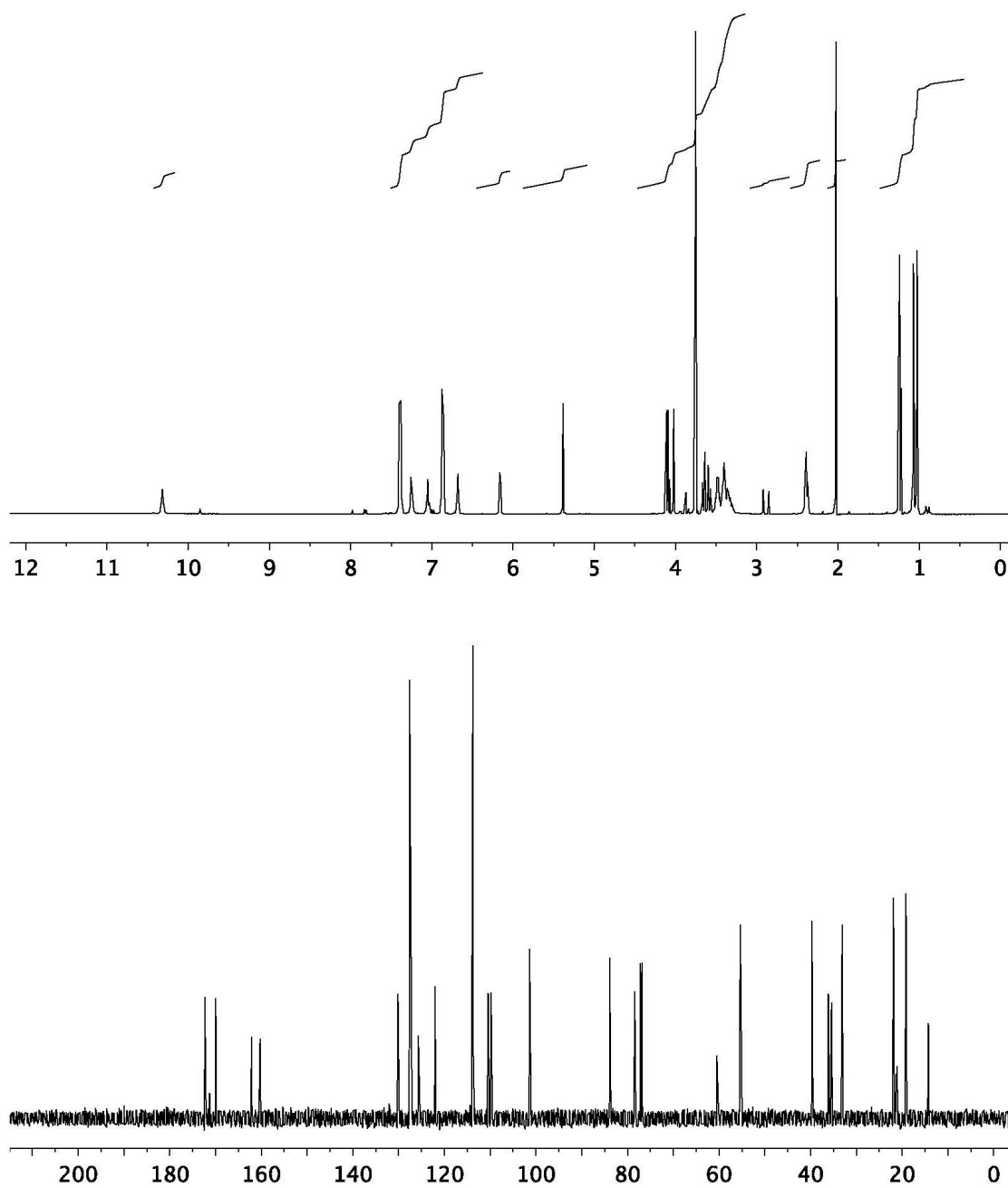
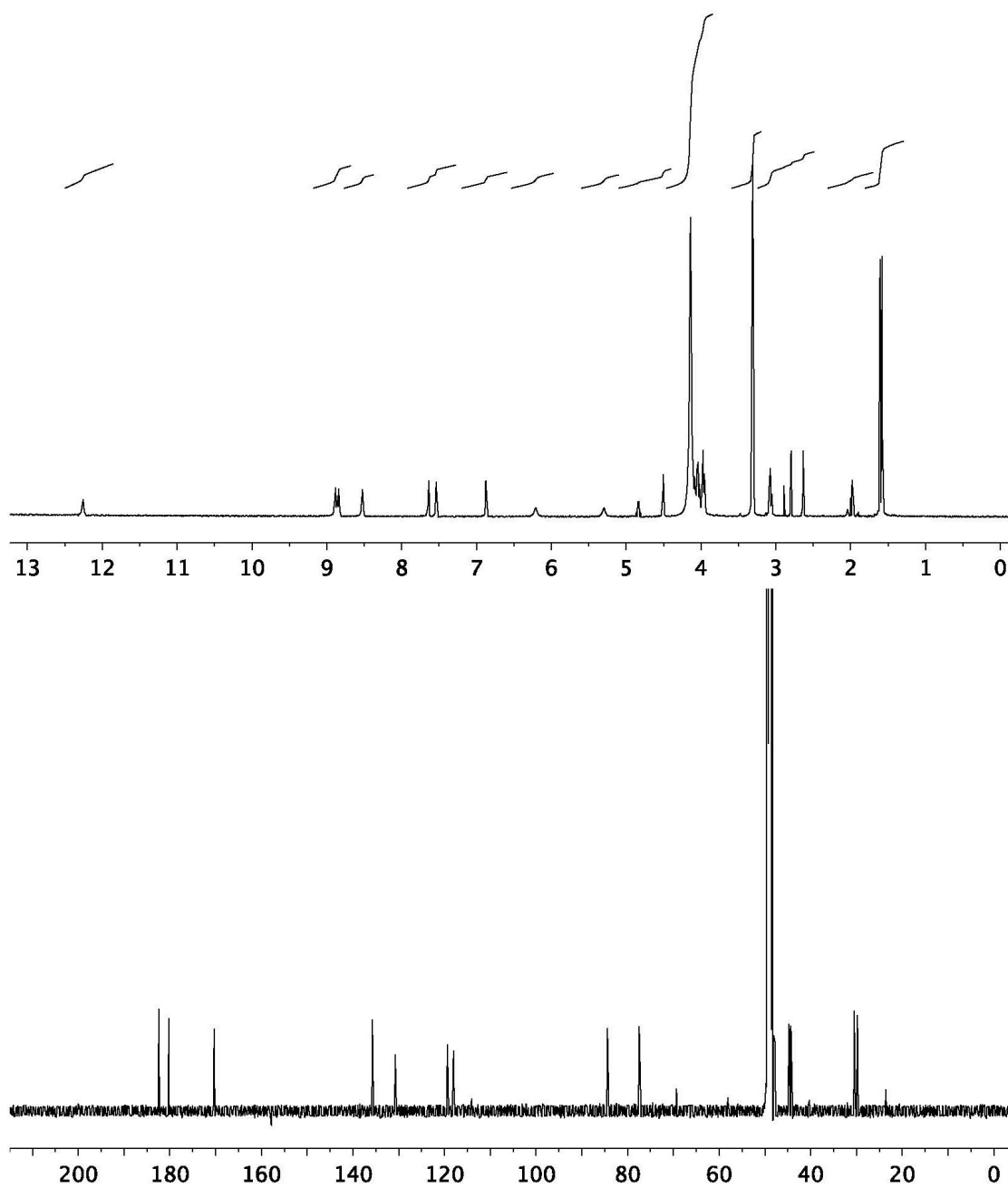


Figure S25. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **7** in CDCl_3



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 <http://biopent.uv.es/~roberto/Index.php?sec=water>
- (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.

Acknowledgements: Chapter 3, in full, is a reprint of the material as it appears: Jaremko, M.J.; Lee, D. J.; Opella, S. J.; Burkart, M. D. “Structure and Substrate Sequestration in the Pyoluteorin Type II Peptidyl Carrier Protein PltL.” *J. Am. Chem. Soc.* 137 (2015): 11546-9. The dissertation author is the primary author of this manuscript.

Chapter 4. Engineering protein-protein interactions in NRPS type II

PCPs

4.1 Abstract

In an effort to elucidate and engineer peptidyl carrier protein (PCP)/adenylation (A) domain interactions in type II nonribosomal peptide synthetases (NRPS)s, we analyzed biomolecular recognition of proteins involved in identical adenylation reactions in the related prodigiosin and pyoluteorin pathways using NMR spectroscopy. Of the PCPs discussed herein, only one, PigG, showed cross-species activity. NMR titration experiments demonstrated that biomolecular recognition perturbed the chemical shifts of residues in loop 1 of the PCPs, suggesting this loop is an important recognition element in type II PCPs. Characterization of the hitherto unknown NMR solution structure of PigG and molecular dynamic (MD) simulations of these PCPs revealed differences in the PCPs' structures and dynamics. Mutational studies supported the role of loop 1 in molecular recognition, as mutations to this region of the PCPs modulated their activities.

4.2 Introduction

Type II non-ribosomal peptide synthetases (NRPSs) participate in hybrid biosynthetic pathways with fatty acid synthases, polyketide synthases, and type I NRPSs.¹⁻³ These hybrid systems generate complex natural products that often demonstrate valuable bioactivities.⁴⁻⁶ Type II NRPSs typically modify amino acids

through oxidations, hydroxylations, and chlorinations,⁷⁻¹¹ and these precursors provide unique features and diversity to the complete natural product. For example, the terminal pyrrole in the prodiginines is generated by a type II NRPS system. The family is known to have many antitumor and antimalarial activities.^{12,13} Another type II NRPS system hydroxylates and chlorinates a tyrosine that is incorporated into the antitumor agent C-1027.¹⁴ Given their makeup of primarily stand-alone proteins, type II NRPSs are ideal targets for metabolic engineering efforts, however the ability to design even simple NRPS systems remains challenging.

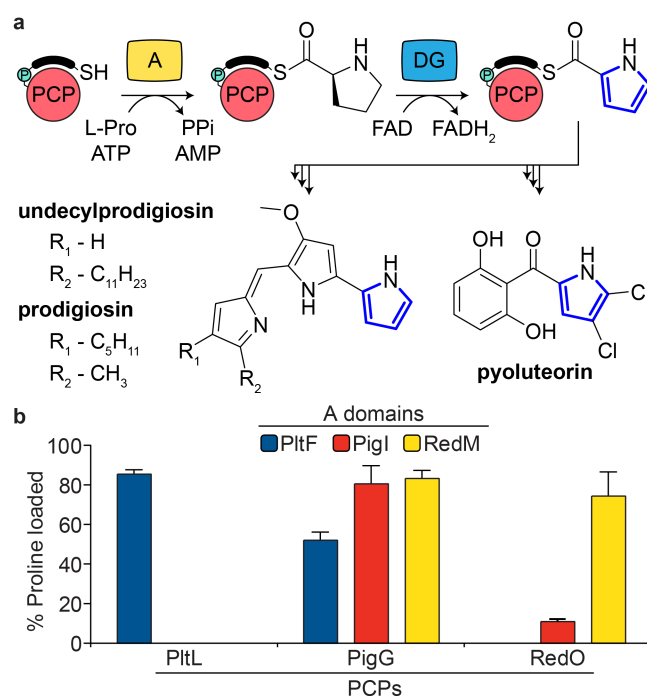


Figure 4.1. (a) Pyrrole formation in pyoluteorin and prodiginine biosynthesis. PCP, peptidyl carrier protein; A, adenylation domain; DG, dehydrogenase. The black bar and teal circle above the PCP denotes phosphopantetheine. (b) A domain activity with cognate and non-cognate PCPs.

4.3 Results and Discussion

Previously, the homologous PCPs and A domains in the undecylprodigiosin and pyoluteorin pathways were shown to be specific for their cognate partner (Figure

1b, S1).⁷ We further examined the activity of PigG and PigI with these homologous pairs by monitoring the aminoacylated PCP product. The assay confirmed that PltF and RedM load only their cognate PCP, PltL and RedO, respectively (Figure 1b). Although, RedM has slight recognition for PigG. Interestingly, PigG was acylated by all three A domains. The promiscuity of PigG is surprising, especially for PltF, considering the low sequence identity (25.0%) between PigG and PltL. The identity between the prodiginine PCPs PigG and RedO is significantly higher (39.8%) (Figure S2). The results reveal varying specificities between partners and provide a model to study the important interfaces for protein-protein recognition in type II NRPS systems. Given PigG's promiscuity, we set out to investigate the differences between its structure and that of PltL.

We began by determining the NMR solution structure of *holo*-PigG in order to compare this structure with that of *holo*-PltL (Figure 2, S3-S5, Table S1). Our prior elucidation of the 3D NMR structure of PltL in *holo* and pyrrolyl forms demonstrated that type II PCPs have the capacity to sequester their substrates analogous to type II fatty acid and polyketide synthases.^{2,15,16} Both *holo*-PigG and *holo*-PltL possess a unique interruption in helix III,² and the N-terminal portion of helix II in both PCPs has slight positive potential that could form electrostatic interactions with an A domain, as this PCP region is proximal to the A domain in the recent crystal structures (Figure 2c-d).¹⁷⁻¹⁹ In contrast, loop 1 of these PCPs varies significantly between the two structures (Figure 2 c-d). The loop 1 N-terminal region of PigG (residues 15-34) has a strong negative potential, while the same region of loop 1 of PltL (residues 19-

40) has a weak positive potential. Interestingly, the PCP loop 1 region of the A•PCP crystal structures is in close proximity to the Asub domain, a flexible C-terminal region of the A domain (Figure S6).¹⁸

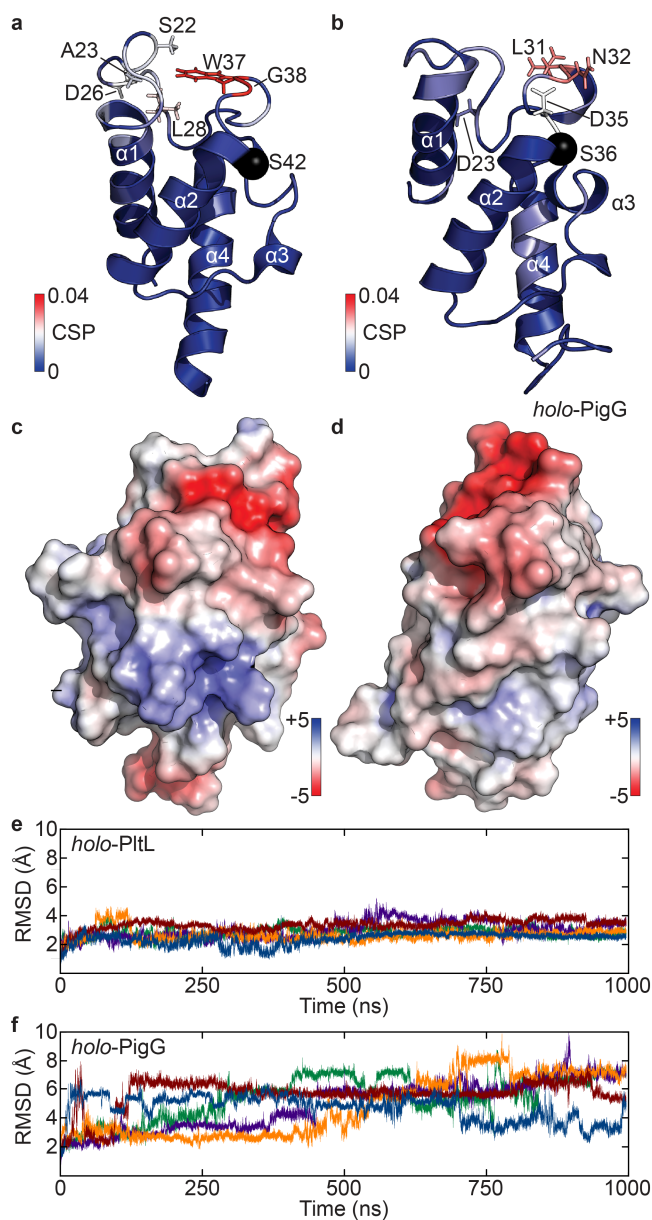


Figure 4.2. The solution NMR structures and electrostatic potentials (ESPs) of PCPs. (a,b) Structure of *holo-PltL* and *holo-PigG*. Structure color maps represent the CSP in Figure 3. (c,d) ESPs of *holo-PltL* and *holo-PigG*. The color map ranges from $-5 \text{ kT } e^{-1}$ and $5 \text{ kT } e^{-1}$. (e,f) Backbone (heavy atom) root mean square deviations of *holo-PltL* (e) and *holo-PigG* (f). Data from 5 independent cMD simulations are colored uniquely.

Conventional and Gaussian accelerated MD simulations of the two PCPs were performed. Analysis of the simulation data revealed significantly larger backbone heavy root mean square deviations (RMSDs) in *holo*-PigG (suggesting), indicating that PigG is less well-ordered than PltL. The greater flexibility of *holo*-PigG may allow it to sample conformations that can be recognized by noncognate adenylation domains (Figure 2e-f).

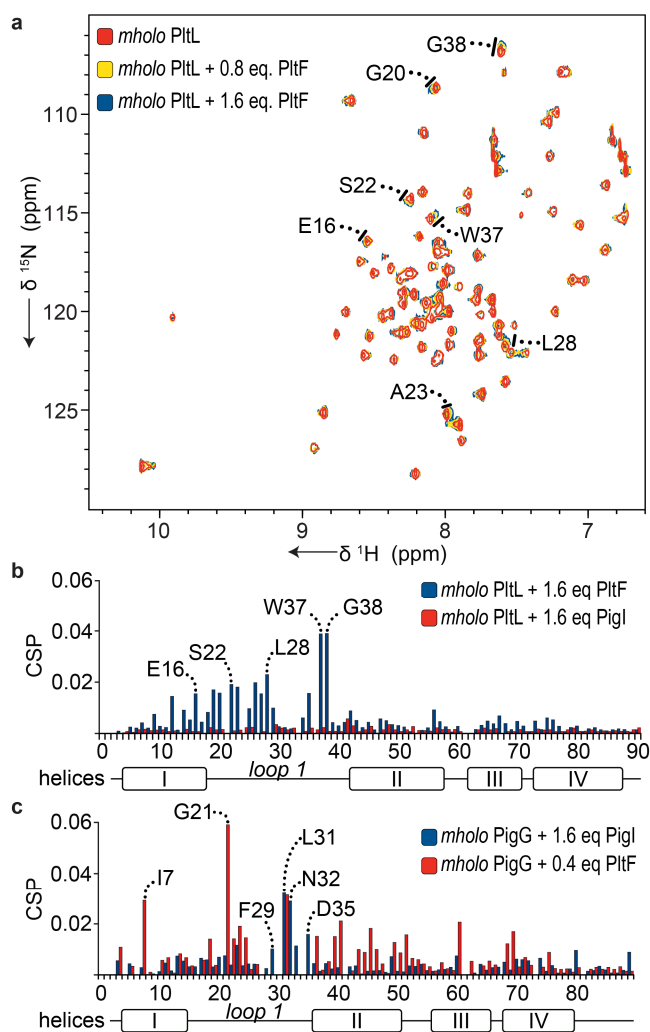


Figure 4.3. Perturbations of PigG and PltL due to interaction with A domains. (a) ^1H - ^{15}N HSQC overlays of *mholo*- ^{15}N -PltL with increasing PltF concentrations. (b) CSP plots of *mholo*- ^{15}N -PltL with PltF or PigI relative to *mholo*- ^{15}N -PltL alone. (c) CSP plots of *mholo*- ^{15}N -PigG with PigI or PltF relative to *mholo*- ^{15}N -PigG alone.

NMR titration experiments were next performed for each PCP with the A domains PltF and PigI to inform the interface residues required for product formation. To capture the productive interaction between carrier protein and A domain, we methylated the thiol on the *holo*-PCP, as a conservative modification, to prevent aminoacylation. The S-methylated *holo*-PCP (*mholo*-PCP) was generated by synthesizing a methylated coenzyme A (CoA) (Figure S6), which was subsequently phosphopantetheinylated onto *apo*-PCP by the action of Sfp.^{20,21} ^1H - ^{15}N HSQC spectra were collected using *mholo*- ^{15}N -PltL or *mholo*- ^{15}N -PigG solutions with PCP•A domain molar ratios ranging from 0 to 1.6 and the chemical shift perturbations (CSPs) of the assigned backbone peaks were calculated (Figure 3, S7-S8). For PltL titrations, significant linear peak perturbations were observed as PltF concentrations increased, while no significant perturbations were observed when PigI was introduced (Figure 2a-b, 3 b-c), indicating a lack of functional interaction. For PigG titrations, significant peak perturbations were observed as both PigI and PltF concentrations were increased, which agrees with the PigG promiscuity seen in product formation assays (Figure 1b). In fact, PigG NMR signal was effectively lost when PltF exceeded 0.4 molar equivalence, indicating that PigG imparts a significantly stronger affinity for PltF than PigI. Interestingly, when both PltL and PigG interacted with an A domain, significant peak perturbations were observed in residues situated in loop 1 between helix I and II (Figures 2a-b, 3b-c). Furthermore, the diminished residues of PigG at 0.4 equivalence PltF were located in loop 1 (residues 20-21, 28-29, 32-33, 35) of PigG (Figure S8), an indication of binding of this region to the significantly larger A domain. The dramatic

loss of PigG signal in presence of PltF agrees with the decreased product formation of the A domain compared to PigI (Figure 1b), where the affinity between the enzymes may delay the release for proline loading on subsequent PigG species.

Further inspection of the simulation data with respect to the loops of these PCPs suggest subtle differences in their flexibilities. The largest loop fluctuations in the PltL occur at the beginning of the loop; whereas the end of loop undergoes greater fluctuations in PigG (Figure S9). To our knowledge, loop 1 has never been demonstrated to play a role in the function of carrier proteins in any pathway. Given this important finding, we next performed mutational analysis on the PCPs to further assess the importance of the loop 1 region.

Based on the difference in dynamics of the PCPs and the activity in loop 1 in the presence of A domain, mutational studies were performed on the PCPs in three regions: residues on loop 1 (Figure 4, red), residues underneath loop 1 that interact with the region (Figure 4, orange), and residues on helix I and II that interact to hold the helices together via intramolecular interactions (Figure 4, blue). Based on polarity differences and length of the loop, residues of PigG were mutated to the corresponding residues on PltL in each region and vice-versa (Figure S10). Each mutant PCP species was analyzed in the adenylation assay with either PltF, PigI, or RedM to determine the influence of the different PCP regions. PigG mutant species 1 lost significant recognition with all three A domains compared to mutant species 2 and 3 (Figure 4c, S11). PltL mutant species did not significantly alter recognition with the A domains, although the PltL mutant species 1 did gain minimal function with RedM. The

activities gained or lost for these mutants with the A domain illustrates the importance of loop 1 in the pyoluteorin and prodigiosin type II PCPs and highlights a region of carrier proteins that has not been identified before as critical to activity.

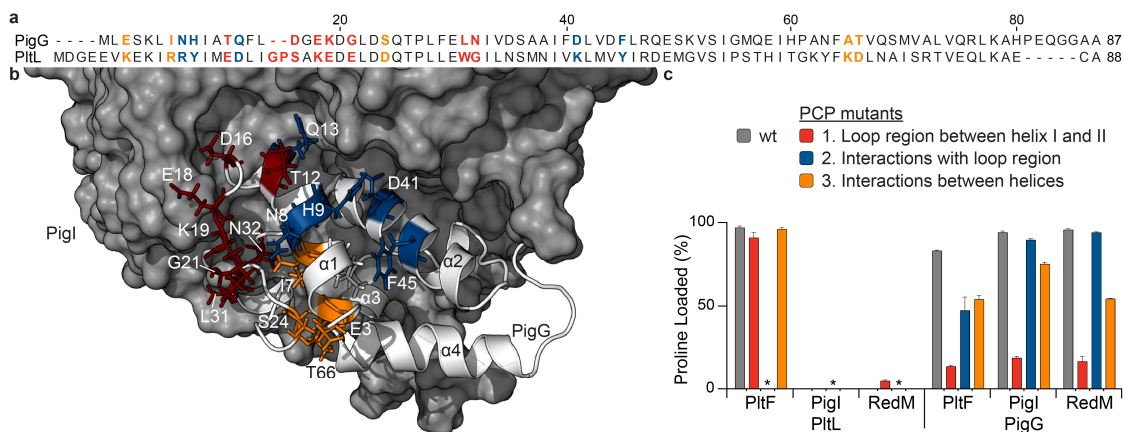


Figure 4.4. PCP loop 1 modification alters interactions with homologous A domains. Residues were mutated in loop 1 (red), underneath loop 1 (yellow), and between helix I and II (blue). (a) Sequence alignment of PigG and PltL using MUSCLE.²² (b) *holo*-PigG NMR structure docked to model structure of PigI. Mutated residues are highlighted in red, blue, and orange and are explained in (c). (c) Adenylation assays with mutated PigG and either PigI or PltF. Astericks indicate PltL mutant 2 variants were not analyzed due to instability.

4.4 Conclusion

The type II NRPS systems produce unique moieties in natural product biosynthesis, and their architecture offers potential for the development of engineered biosynthetic pathways. Here we demonstrated that type II PCPs in homologous pathways present different specificity profiles with regard to adenylation domain partners. Solution structures, MD simulations, and NMR titration experiments were used to identify a loop region of the carrier proteins that has not been previously scrutinized. Mutational studies demonstrated the significance of the loop 1 interface by altering A domain recognition and validates this location as a potential hot spot for A domain selectivity. NRPS pathways rely on the fidelity of protein-protein

interactions between each A domain and PCP cognate pair for the proper loading of starter units. A fundamental understanding of the interactions between these partners should allow for engineering and eventual control over the identity of amino acids incorporated into pathways. The PCP loop 1 should also be considered when investigating other unique type II partners, including halogenases and cyclopropanases.^{23,24} It has become clear that control over PCP and partner enzyme recognition will contribute to the future engineering of carrier protein dependent pathways.

4.5 Supporting Content

Expression and purification of proteins

pET22b-PigG, pet22b-PltL, pet37b-RedO, pet22b-PigI, pet37b-PltF, and pet37b plasmids were obtained from the laboratory of Christopher Walsh at Harvard University^{7,25}. The plasmids were transformed into BL21 (DE3) cells. The BL21 cells were grown in Luria-Bertani (LB) media containing 50 mg/L kanamycin (pet37 constructs) or 100 mg/L ampicillin (pet22b constructs). Expression was induced with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (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 at 1000 relative centrifugal force (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).

Uniformly labeled ^{15}N -PigG and ^{15}N -PltL were expressed by culturing cells in M9 minimal media (1 L) supplemented with 1 g ^{15}N - NH_4Cl , 4 g D-glucose, and 100 mg of ampicillin. ^{15}N -PigG and ^{15}N -PltL were expressed and purified as in above.

pET22b-PigG and pRep4-Sfp (obtained from the laboratory of Christopher Walsh at Harvard University) plasmids were co-transformed into BL21 (DE3) cells. The BL21 containing both constructs were grown in M9 minimal media (1 L) with 4 g ^{13}C -D-glucose, 1 g ^{15}N - NH_4Cl , and 100 mg of ampicillin. Uniformly labeled ^{13}C , ^{15}N *holo*-PigG was expressed and purified as mentioned in section **B.2**. ^{13}C , ^{15}N *holo*-PigG was isolated from Sfp (no his-tag) during Ni-NTA purification.

Protein NMR Studies

For preparation of *mholo*- ^{15}N -PigG and *mholo*- ^{15}N -PltL, the thiol of coenzyme A (CoA) was methylated as previously described for thiol group in cysteamine (Figure S6).²⁶ The methylated CoA was transferred to ^{15}N *apo*-PigG and ^{15}N *apo*-PltL by the phosphopantetheinyl transferase Sfp (Figure S6).²⁷ In this study, this was achieved by treating 1.5 mL of 400 μM ^{15}N *apo*-PigG or ^{15}N *apo*-PltL with 12.5 mM MgCl_2 , 2 mM methylated CoA, 4 μM Sfp, 0.1% NaN_3 , and 5 mM TCEP in a 150 mM NaCl and 50 mM Tris pH 7.4 buffer. The *mholo*-PCPs were purified on a Superdex 75 into a 50 mM KPi 7.4 buffer.

For *holo*-PigG NMR Data Collection and Structure calculations, the ^{13}C , ^{15}N *holo*-PigG sample was prepared by first concentrating the FPLC pure protein in 50 mM KPi pH 7.4. A 450 μL aliquot of the sample was then prepared for NMR by adding 50 μL of D_2O , 5 μL of 10% (w/v) NaN_3 and 5 μL of 0.5 M TCEP adjusted to

pH 7.4. The final concentration of ^{13}C , ^{15}N *holo*-PigG was 1.5 mM. 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 NMRFAM-SPARKY²⁸ utilizing the PINE-SPARKY extension²⁹. 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 ^{13}C assignments were obtained from ^{13}C -edited NOESY-HSQC and TOCSY-HSQC experiment (mixing time 60 ms). NOE data was obtained from 3D ^{15}N -edited and ^{13}C -edited NOESY-HSQC experiments (mixing time 150 μs). 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-N³⁰. NOE peaks were subject to automated assignment and structure calculations with CYANA³¹ 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^{32,33}, and scripted by Roberto Tejero's WaterRefCNS[3]. Output files were analyzed and prepared for submission by PdbStat³⁴. Topology, parameter, and linkage files for phosphopantetheinylated serine (PNS) and were constructed using a combination of known atom types and additional information from the PRODRG server³⁴.

For NMR titration experiments, the proteins *mholo*- ^{15}N -PigG, *mholo*- ^{15}N -PltL, PigI, and PltF were prepared by first concentrating the FPLC pure proteins in 50 mM

KPi pH 7.4. The NMR titration samples contained 82.2 μM ^{15}N -PCP, A domain (0, 32.9, 65.8, 98.6, 131.5, or 164.4 μM), 10 % D_2O , 0.1% (w/v) NaN_3 and 5mM TCEP. 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 ^{15}N indirect dimension. Chemical shifts were measured by peak maxima, and chemical shift perturbation was calculated using the formula $\text{CSP} = [((0.2\delta\text{N})^2 + (\delta\text{H})^2)/2]^{0.5}$.

Mutational Studies

For site-directed mutagenesis of PltL and PigG, the PltL and PigG mutants were introduced via quickchange PCR to determine the important PCP interfaces for interaction with the A domain. Sequences of mutated PigG and PltL are listed in Figure S16. The mutant PCPs were purified as the unlabeled enzymes above.

For the adenylation assay, the PCPs were converted from *apo*-PCP to *holo*-PCP by the Sfp^{20,21,27}. In this study, this was achieved by treating 1 mL of 50 μM PCP with 12.5 mM MgCl_2 , 2 mM CoA, 4 μM Sfp, 0.1% NaN_3 , and 5 mM TCEP in a 150 mM NaCl and 50 mM Tris pH 7.4 buffer. The *mholo*-PCPs were purified on a Superdex 75 into a 50 mM Tris pH 7.4, 150 mM NaCl, and 5% glycerol buffer.

The adenylation reaction (15 μL) contained 50 mM Tris pH 7.4, 12.5 mM MgCl_2 , and 2 mM TCEP, 5 mM ATP, 5 mM L-proline, 15 μM *holo*-PCP, 0.27 μM A domain. The reactions were incubated at 25°C for 5 min (WT experiment) or 1 h

(mutant experiment) then quenched with 3 μ L of 10% formic acid. The solution was centrifuged for 10 min at 13,000 rpm. The supernatants were passed over a C18 column (Ascentis Express Peptide ES-C18) using 5% solvent B for 2 min, then increasing 5–40% B over 8 min, and finally 40–44% over 10 min (solvent A = H₂O, 0.05% trifluoroacetic acid; solvent B = CH₃CN 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.

Computation methodology

Initial coordinates for simulations of PltL and PigG, were prepared using the NMR solution structures reported previously² and herein, respectively. All molecular dynamics (MD) simulations were performed using the AMBER ff14SB force field.³⁵ The carrier proteins were simulated in their *holo* forms, which features a conserved serine residue is phosphopantetheinylated post-translationally. In this work, that modified serine was treated as a nonstandard residue. The Restrained Electrostatic Potential (RESP) methodology³⁶ was used to generate partial charges for the atoms in this residue using a HF/6-31G(d) electrostatic potential. *Ab initio* computations were performed using Gaussian 09.³⁷ The ANTECHAMBER³⁸ program of AmberTools14³⁹ was used to generate AMBER or GAFF⁴⁰ type parameters. The H++ webserver⁴¹⁻⁴⁴ was used to determine the protonation states of all titratable residues of *holo*-PltL and *holo*-PigG, assuming an external dielectric constant of 80, a salinity of 0.150 M, and a pH of 7.4. The protonation states of histidine residues were confirmed ‘by hand’. The resulting structures were solvated with TIP3P⁴⁵ water, such that no proteinogenic atom

was less than 10 Å from any box edge, and were neutralized using the TLEAP program of AmberTools15.³⁹ Na⁺ and Cl⁻ ions were added to the system to mimic physiological conditions (ca. 0.150 M).

All conventional (cMD) and Gaussian accelerated molecular dynamics⁴⁶⁻⁴⁹ (GaMD) simulations were performed using the GPU-compatible PMEMD engine available in a modified version of Amber14³⁹. Simulations utilized the SHAKE⁵⁰ algorithm to constrain all bonds between nonpolar and hydrogen atoms; a 2.0 fs time-step was used in these simulations.⁵¹ Long-range electrostatics were calculated using the Particle Mesh Ewald (PME) method with a 10.0 Å cutoff was used for nonbonded interactions. In all simulations, the Langevin thermostat (with $\gamma = 1.0 \text{ ps}^{-1}$) was used to maintain temperature control.⁵² After a two-step minimization process, in which solvent molecules were allowed to relax before the entire system was minimized, the system was slowly heated to 305 K over 0.4 ns in an NVT simulation. The systems were then equilibrated for 1 ns by performing isothermal-isobaric (NPT) simulations at 305 K using a Monte Carlo barostat. Five 1 μs NPT production-grade cMD simulations of each system were performed.

Five independent production GaMD simulations for each protein were performed as follows: The system threshold energy was set as $E = V_{\text{max}}$; maximum (V_{max}), minimum (V_{min}), average (V_{av}), and the standard deviation (σ_v) of the system potential were determined from a 2 ns nonbiased NVT simulation. GaMD simulations included a 2 ns equilibration period; these portions of each independent simulation has

been excluded from analysis. Production-grade GaMD simulations were performed using NPT ensemble.

The electrostatic potentials of PltL and PigG were computed using the Adaptive Boltzmann-Poisson Solver⁵³⁻⁵⁷ (ABPS) webserver, using default values. The pdb2pqr^{58,59} server was used to generate pqr files, which serve as inputs for the ABPS. PROPKA⁶⁰ was used to determine protonation states of titratable residues in PltL and PigG, assuming a pH of 7. The partial charges, radii, and internal naming schemes for the atoms in these systems were derived using the AMBER94 force field.⁶¹ The electrostatic potential of PltL and PigG were mapped onto the respective surfaces of these proteins using the ABPS plugin for Pymol (v1.8.6). The color range of the electrostatic potential maps shown in Figure 2 span from $-5 \text{ kT } e^{-1}$ to $5 \text{ kT } e^{-1}$.

Simulation frames were written every 5 ps for analysis. MD trajectories were analyzed using the CPPTRAJ⁶² module of AmberTools14, PYTRAJ, a Python packing binding to CPPTRAJ,⁵³ and MDTraj⁶³. Gaussview5,⁶⁴ Pymol v1.8,⁶⁵ and VMD⁶⁶ were used throughout the course to prepare and to visualize structures. All images of protein structures depicted herein were rendered using PyMOL.

The structure depicted in Figure 4 of the main text was constructed by generating a structure model through comparative modeling and that was then subjected to protein-protein docking simulations with PigG (NMR solution structure). The homology model of PigI was generated using Prime⁶⁷⁻⁶⁹ and Maestro⁷⁰ modules of Schrodinger 2017-1.⁷¹ Comparative modeling was performed using the EntF structure (PDB code: 5T3D)¹⁹ as a template structure. Although this template structure shows

low similarity with PigI (ca. 22%), it exists in a PCP-bound conformation. Sequence alignment was performed using Prime's alignment program, STA, which has been developed to align low sequence identity (~25 to ~50 %) templates. A model structure of PigI – based upon this sequence alignment – was constructed using an energy-based method that utilized the OPLS2005⁷²⁻⁷⁴ force field.

Protein-protein docking simulations were performed using the RosettaDock webserver.⁷⁵⁻⁷⁷ RosettaDock performs a local search of configuration space and consequently, a structure of a “guess” complex is required as an input. This input structure was prepared by aligning PigG and PigI homology model to the relevant domains of EntF using Pymol. The highest scoring model generated in this was subjected to further refinement using the webserver. The lowest energy ‘refined’ model of the PigI-PigG complex is depicted in Figure 4 of the main text. Note that these structures are crude and are used only in an illustrative fashion.

4.6 Supporting Table and Figures

Table S4.1 *holo*-PigG Solution NMR Structure Statistics

Assignment Statistics (%) (residues 1-88)	
backbone H	94.8
backbone non-H	95.5
side chain H	98.4
Structure Calculations	
Distance restraints	2084
short-range, $ i-j \leq 1$	1671
medium range, $1 < i-j < 5$	339
long-range, $ i-j \geq 5$	258
Dihedral Angle Restraints	144
Structure Statistics	
Average final CYANA target function (\AA^2)	1.82
Average cycle1 CYANA target function (\AA^2)	98.22
CNS energies (kcal/mol)	-2718.33
Restraint Violations	
Max. distance restraint violation (\AA)	0.5
Violated distance restraints $> 0.2 \text{\AA}$	4
Violated dihedral angles	0
CYANA Ramachandran Plot	
Residues in most favored regions	87.6%
Residues in additionally allowed regions	12.4%
Residues in generously allowed regions	0.0%
Residues in disallowed regions	0.0%
CYANA RMSD (residues 4-82)	
Average backbone RMSD to mean	$0.46 \pm 0.09 \text{\AA}$
Average heavy atom RMSD to mean	$0.97 \pm 0.09 \text{\AA}$

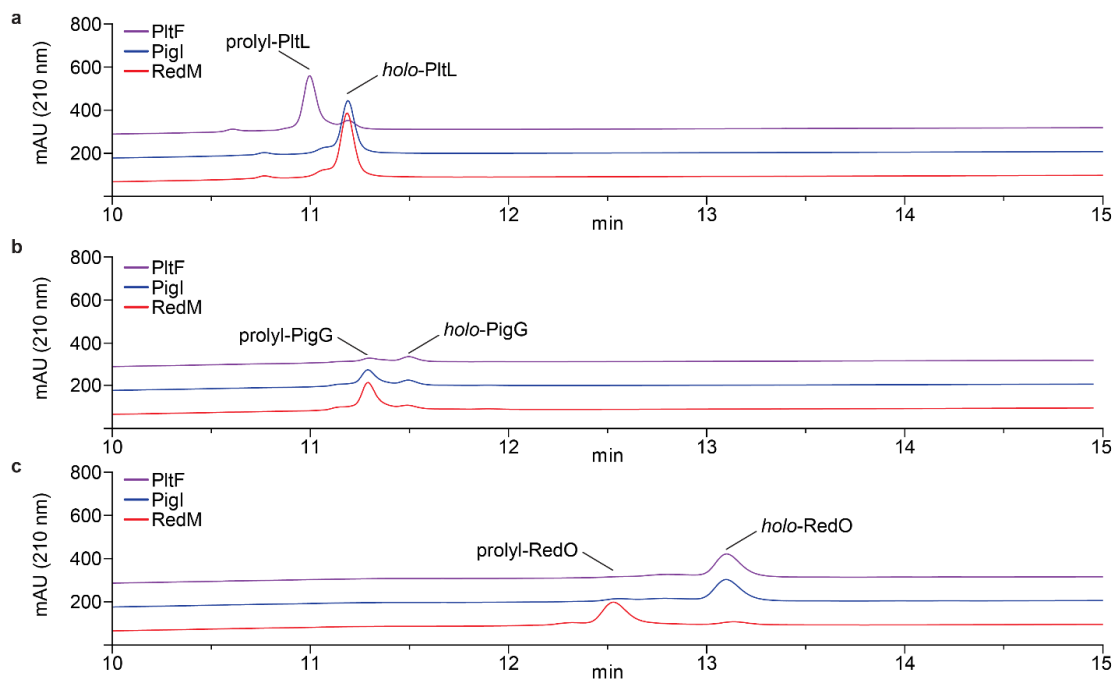


Figure S4.1. HPLC chromatograms of holo-PCP aminoacylation. (a) *holo-PigG* aminoacylation reaction with PltF, PigI, and RedM. (b) *holo-PltL* aminoacylation reaction with PltF, PigI, and RedM. (c) *holo-RedO* aminoacylation reaction with PltF, PigI, and RedM. The HPLC gradient utilized was 5% solvent B for 2 min, then increasing 5–46% B over 7 min, and finally 46–52% over 6 min (solvent A = H₂O, 0.05% trifluoroacetic acid; solvent B = CH₃CN 0.05% trifluoroacetic acid).

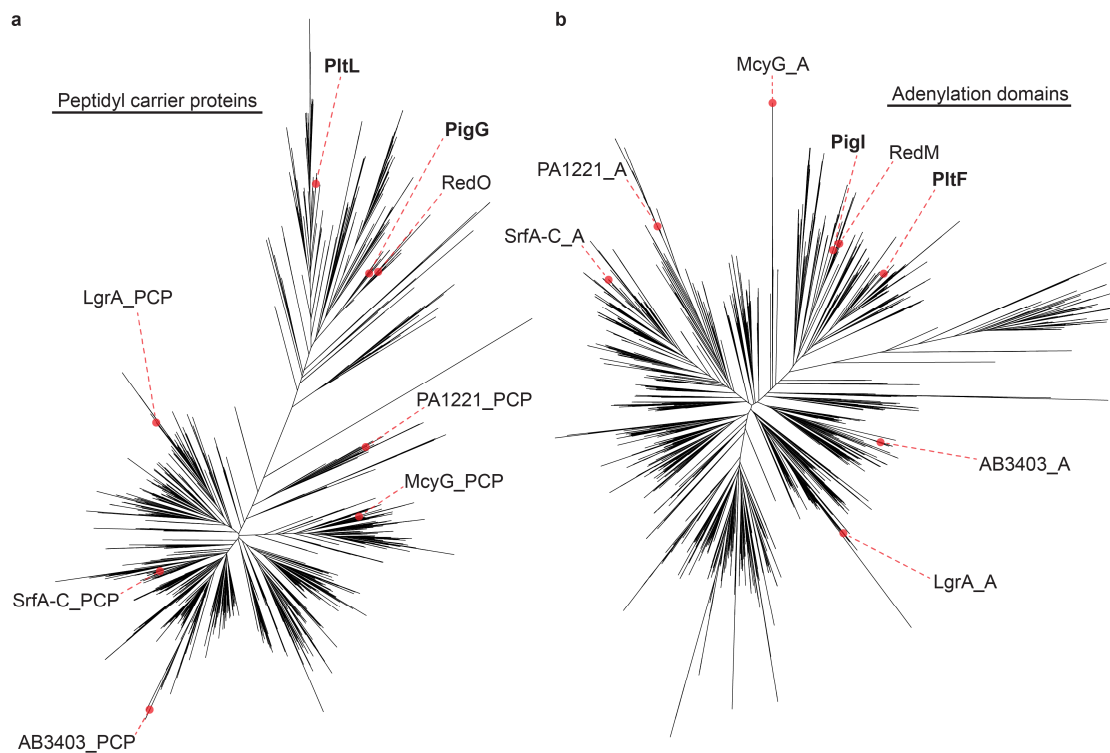


Figure S4.2. Phylogenetic tree of PCPs and A domains. (a) PCP phylogenetic tree with PCP selected that were used in this study or that are structurally characterized with the A domain partner. (b) A domain phylogenetic tree with A domains selected that were used in this study or that are structurally characterized with the PCP partner. The sequences were collected in Blast⁷⁸ and aligned with Muscle²². The phylogenetic tree was generated with Fasttree⁷⁹ and visualized with FigTree.

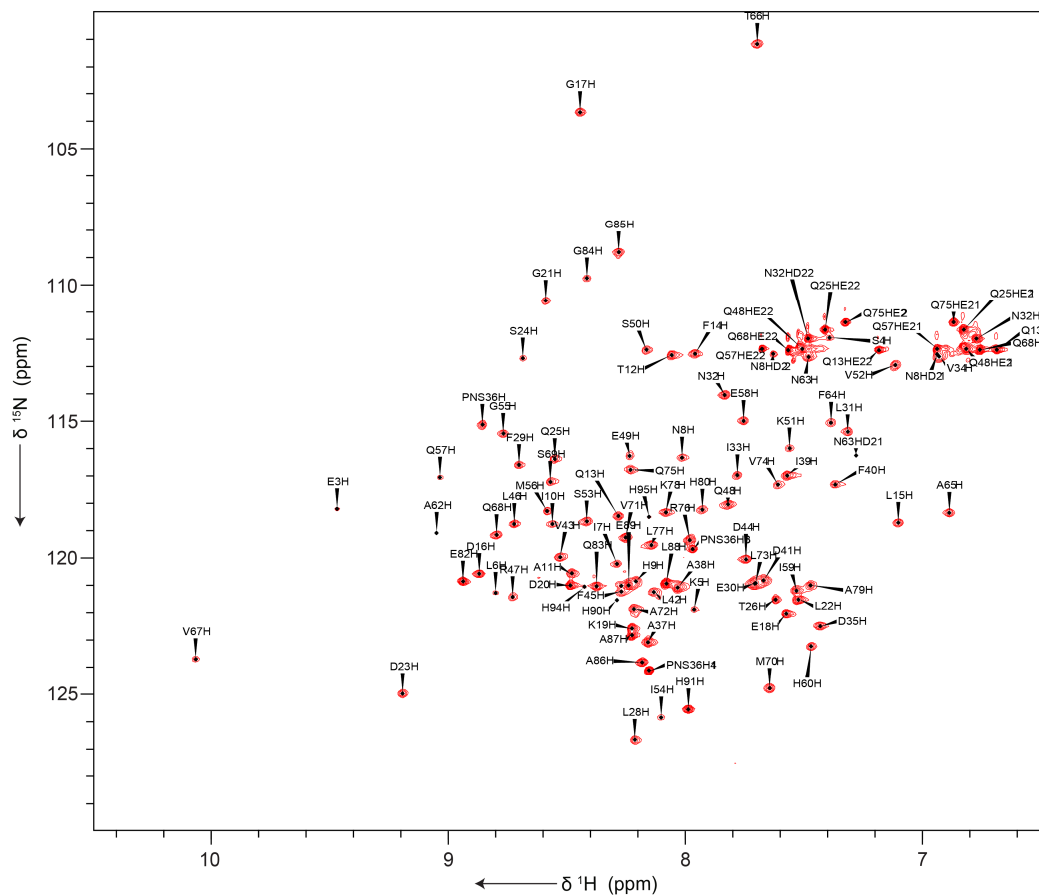


Figure S4.3. ^1H - ^{15}N HSQC of ubiquitous ^{15}N -labelled *holo*-PigG. The phosphopantetheine arm was ^{15}N -labelled by co-expressing PigG and Sfp in BL21 cultured in ^{15}N minimal media.

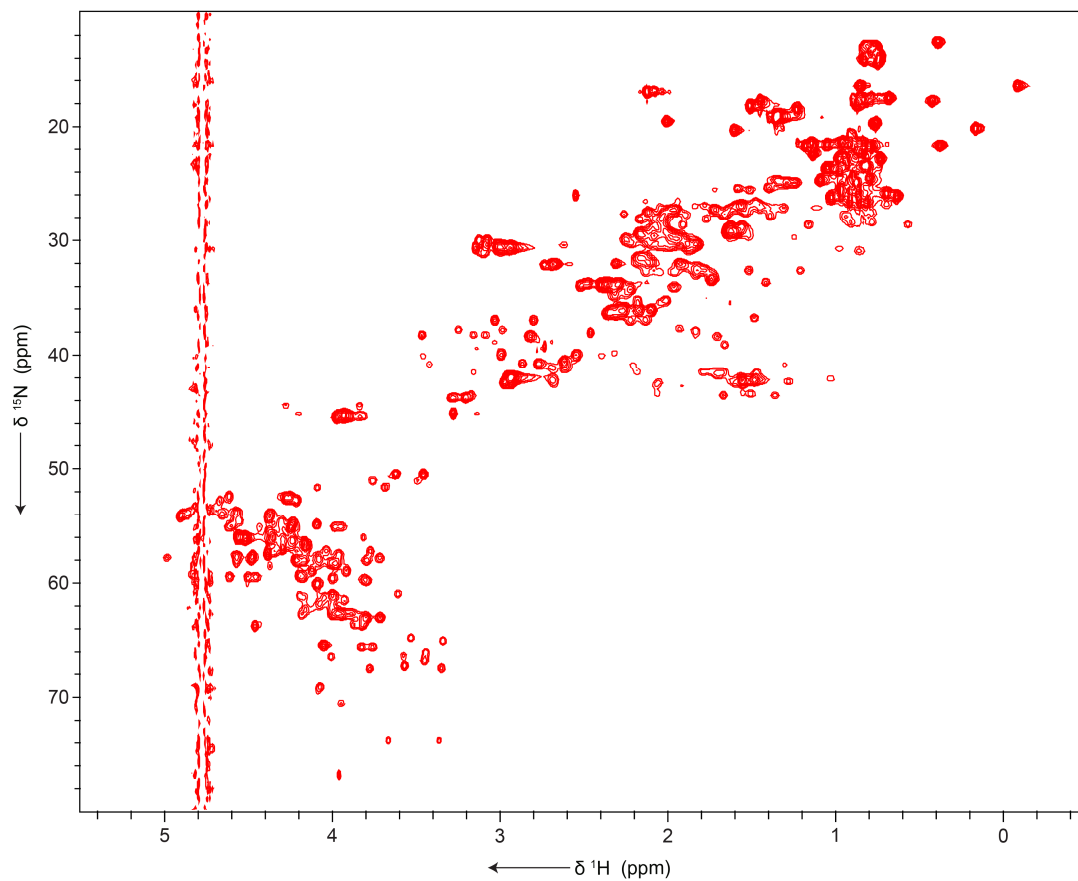


Figure S4.4. ^1H - ^{13}C HSQC (aliphatic region) of ubiquitous ^{13}C -labelled *holo*-PigG. The phosphopantetheine arm was ^{13}C -labelled by co-expressing PigG and Sfp in BL21 cultured in ^{13}C minimal media.

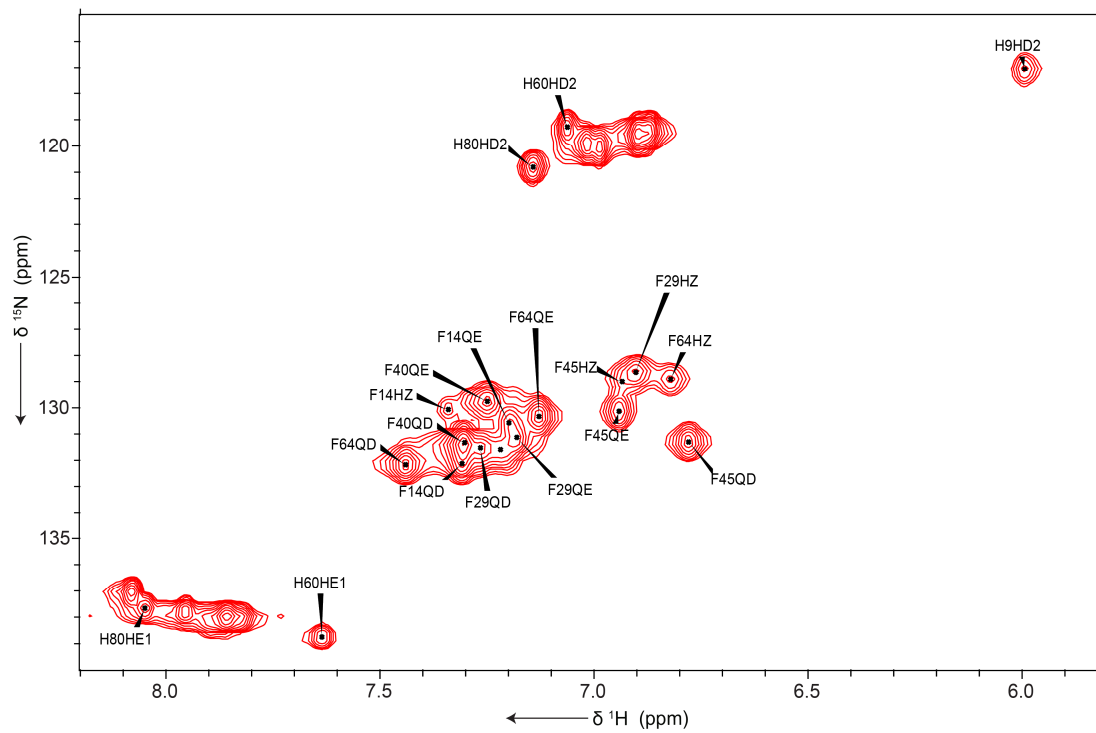


Figure S4.5. ^1H - ^{13}C HSQC (aromatic region) of ubiquitous ^{13}C -labelled *holo*-PigG.

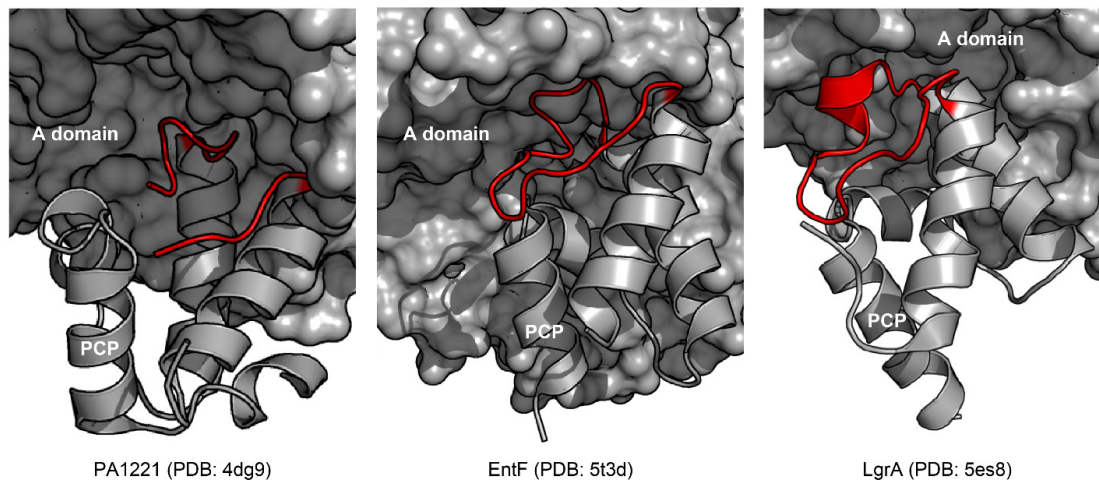


Figure S4.6. Loop 1 region in crystal structures of PA1221,¹⁷ EntF,¹⁹ and LgrA.¹⁸ The view is focused on the PCP and highlights the loop 1 region in red. The loop 1 region is part of the PCP•A domain interface in all three structures. A portion of the PA1221 loop 1 is unresolved in the crystal structure.

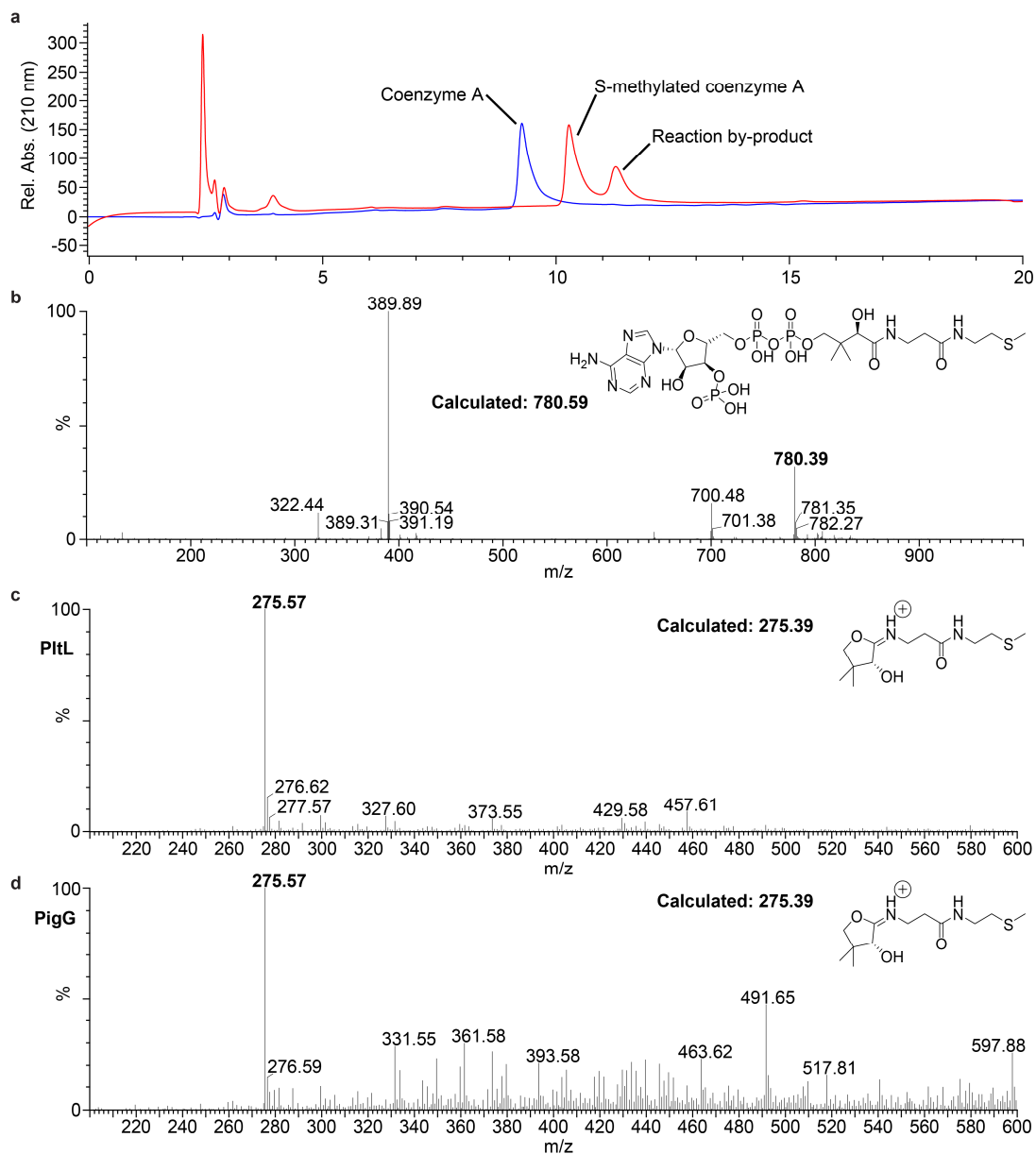


Figure S4.7. Generation of *S*-methylated CoA and *mholo*-PCPs. (a) HPLC chromatograms of coenzyme A (blue) and 2 hr *S*-methylation reaction with *mholo*-PCPs. (b) MS of *S*-methylated coenzyme A. (c) MS-ejection of the phosphopantetheine probe (ref) from *mholo*-PltL at a cone voltage of 55V. (d) MS-ejection of the phosphopantetheine probe (ref) from *mholo*-PigG at a cone voltage of 55V.

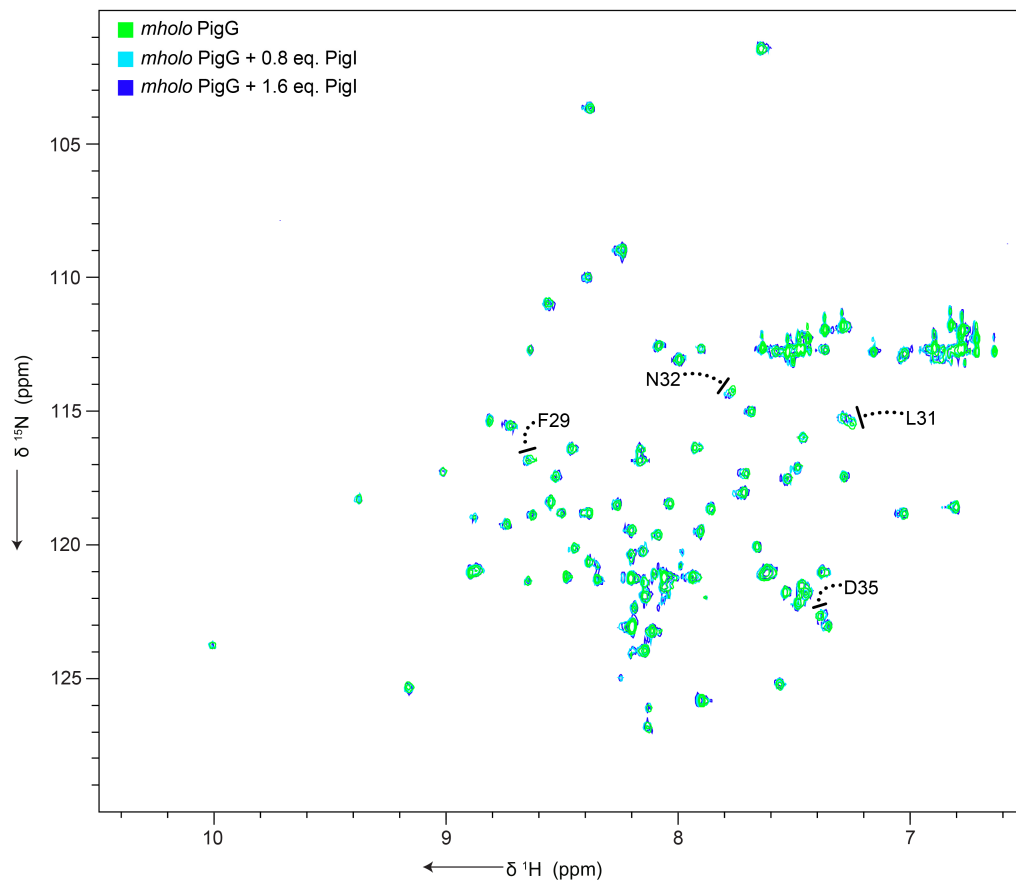


Figure S4.8. ^1H - ^{15}N HSQC overlays of *mholo*- ^{15}N -PigG with PigI titrations. Residues with significant perturbations are highlighted.

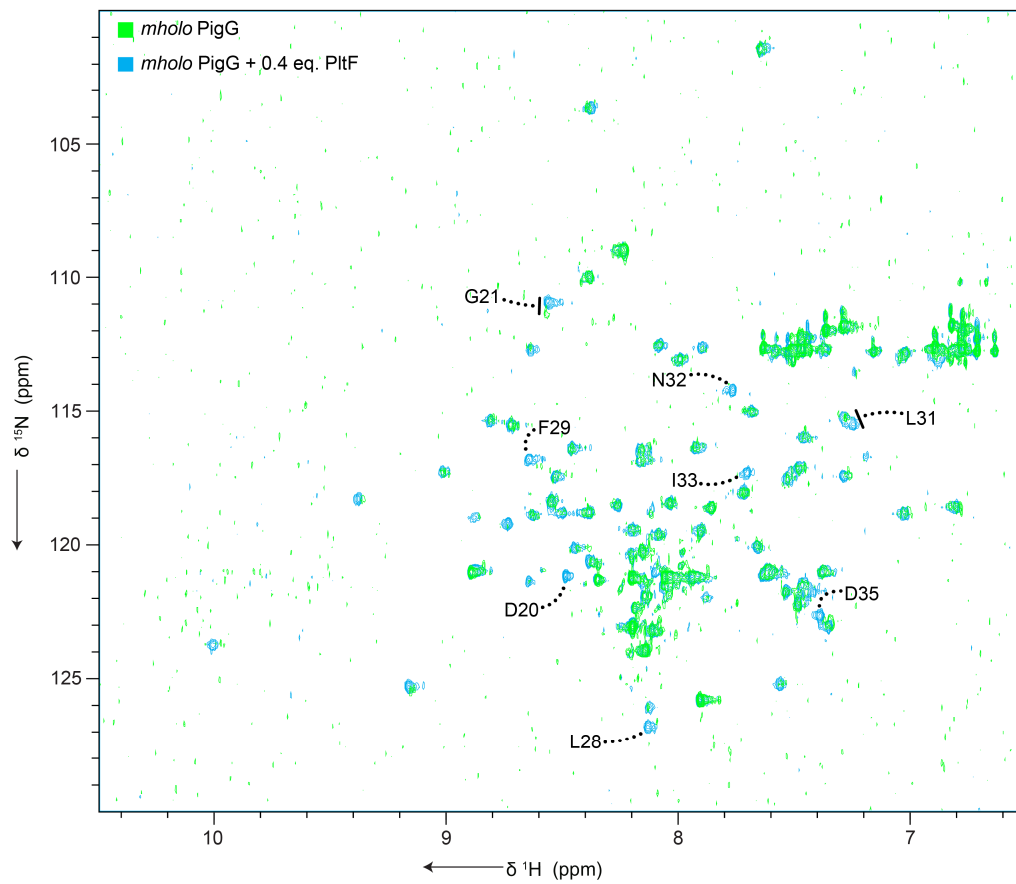


Figure S4.9. ^1H - ^{15}N HSQC overlays of *mholo*- ^{15}N -PigG with PltF titrations. Residues with significant perturbations or loss of signal are highlighted.

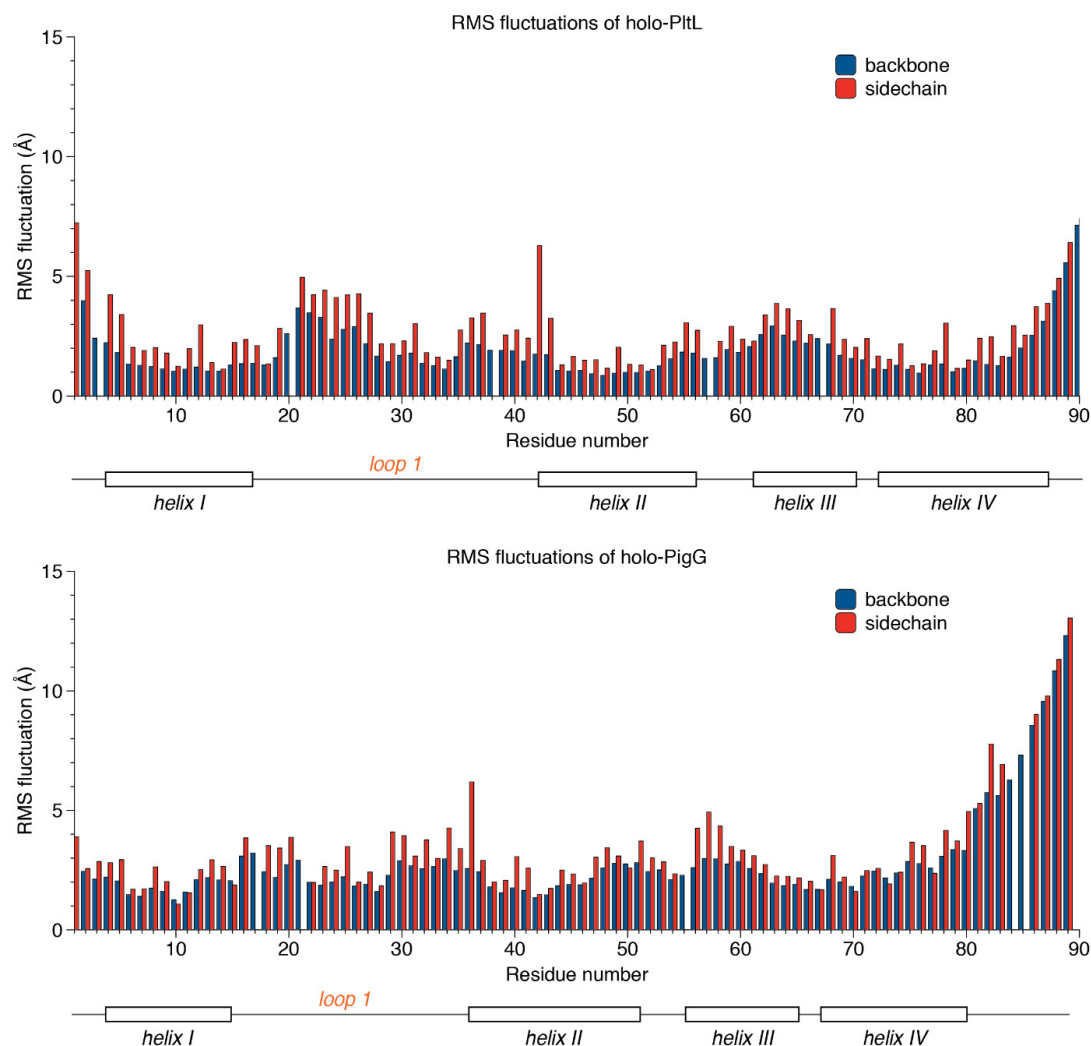


Figure S4.10. Backbone (blue) and side-chain (red) root mean square fluctuations of each residue of *holo*-PltL and *holo*-PigG. Loop 1 and the helical regions (based upon NMR solution structures of PltL and PigG the carrier protein) are labeled under both graphs. Data from all 5 conventional MD simulations of each carrier protein were used to these RMS fluctuations. The root mean square fluctuations of PltL and PigG are shown below. Loop 1 of both carrier proteins are relatively flexible regions of the protein. Unlike type II fatty acid ACPs, helix 3 in PltL and PigG fluctuate significantly more than the other helices in the protein. This can be attributed to a proline residue in both structures that likely reduces the stability of this helix.

```

                20          40          60          80          100
                |          |          |          |          |
PigG wt      - - - -MLESKLINHIATGFL--DGEKDGLDSQTPLFELNIVDSAAIFDLVDFLRQESKVSIGMQEIHPANFATVQSMVALVQRLKAHPEOGGAAWEHHHHHHH 87
PigG Mutant 1 - - - -MLESKLINHIAEQFLGPSGKEDELDSQTPLFEWGIVDSAAIFDLVDFLRQESKVSIGMQEIHPANFATVQSMVALVQRLKAHPEOGGAAWEHHHHHHH 97
PigG Mutant 2 - - - -MLESKLRNHIATGFL--DGEKDGLDDQTPLFELNIVDSAAIFDLVDFLRQESKVSIGMQEIHPANFKDVQSMVALVQRLKAHPEOGGAAWEHHHHHHH 95
PigG Mutant 3 - - - -MLESKLIRYIATDFL--DCEKDGLDSQTPLFELNIVDSAAIFKLVDYLRQESKVSIGMQEIHPANFATVQSMVALVQRLKAHPEOGGAAWEHHHHHHH 95
PitL wt      MDGEEVKEKIRRYIMEDLIGPSAKEDELDDQTPLLEWGILNSMNIVKLMVYIRDEMGVSIPSTHITGKYFKDLNAISRTVEQLKAE- - - -SALEHHHHHHH 88
PitL Mutant 1 MDGEEVKEKIRRYIMTDLIGPSAKEDELDDQTPLLEWGILNSMNIVKLMVYIRDEMGVSIPSTHITGKYFKDLNAISRTVEQLKAE- - - -SALEHHHHHHH 94
PitL Mutant 2 MDGEEVKEKIRYIMEDLIGPSAKEDELDSQTPLLEWGILNSMNIVKLMVYIRDEMGVSIPSTHITGKYFATLNAISRTVEQLKAE- - - -SALEHHHHHHH 96
PitL Mutant 3 MDGEEVKEKIRNHIMEQLIGPSAKEDELDDQTPLLEWGILNSMNIVDLMVFIRDEMGVSIPSTHITGKYFKDLNAISRTVEQLKAE- - - -SALEHHHHHHH 96

```

Figure S4.11. Sequence alignment of WT and mutant PCPs. The alignment was performed with MUSCLES.²²

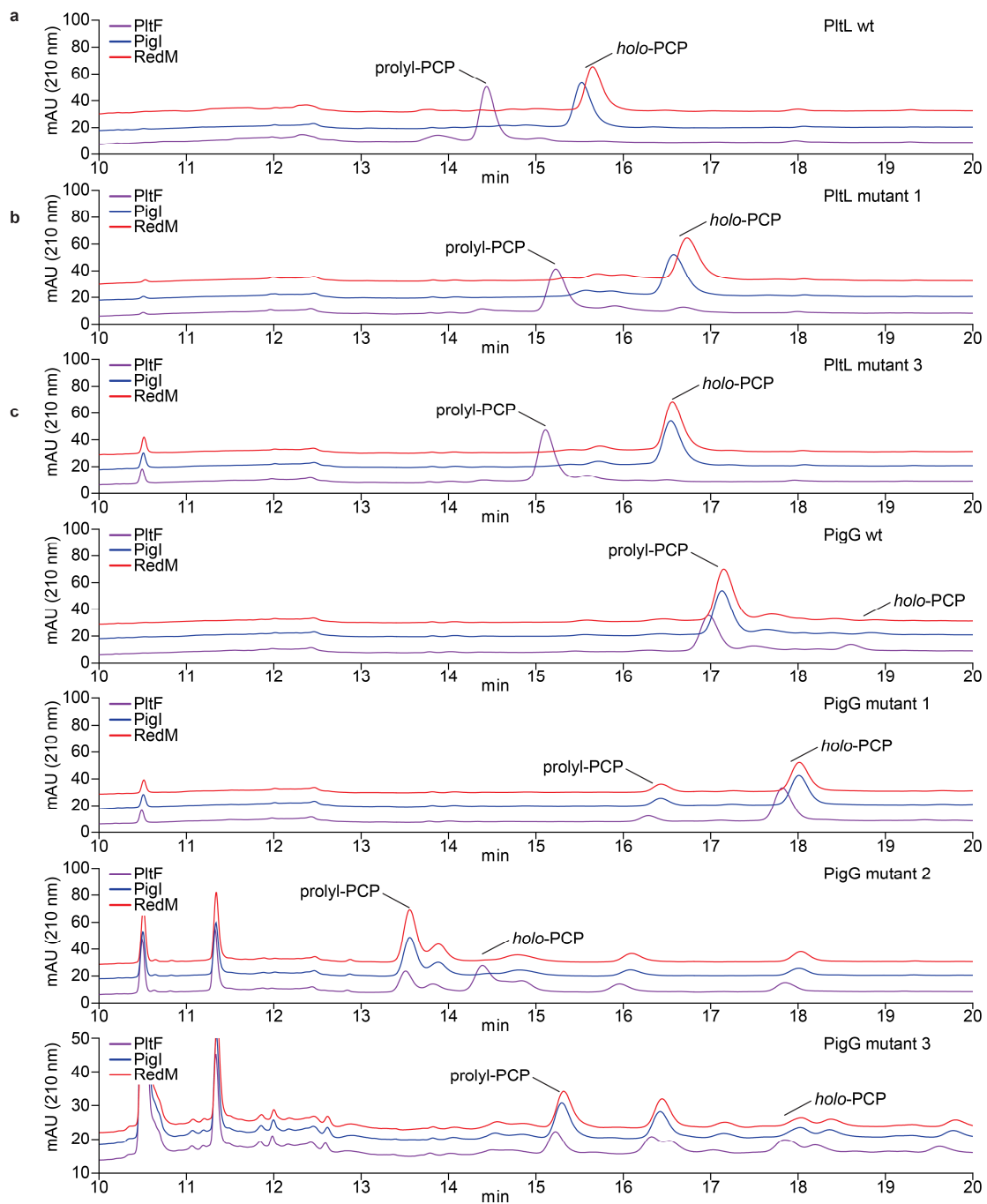


Figure S4.12. HPLC traces of mutant PCPs in adenylation assay.

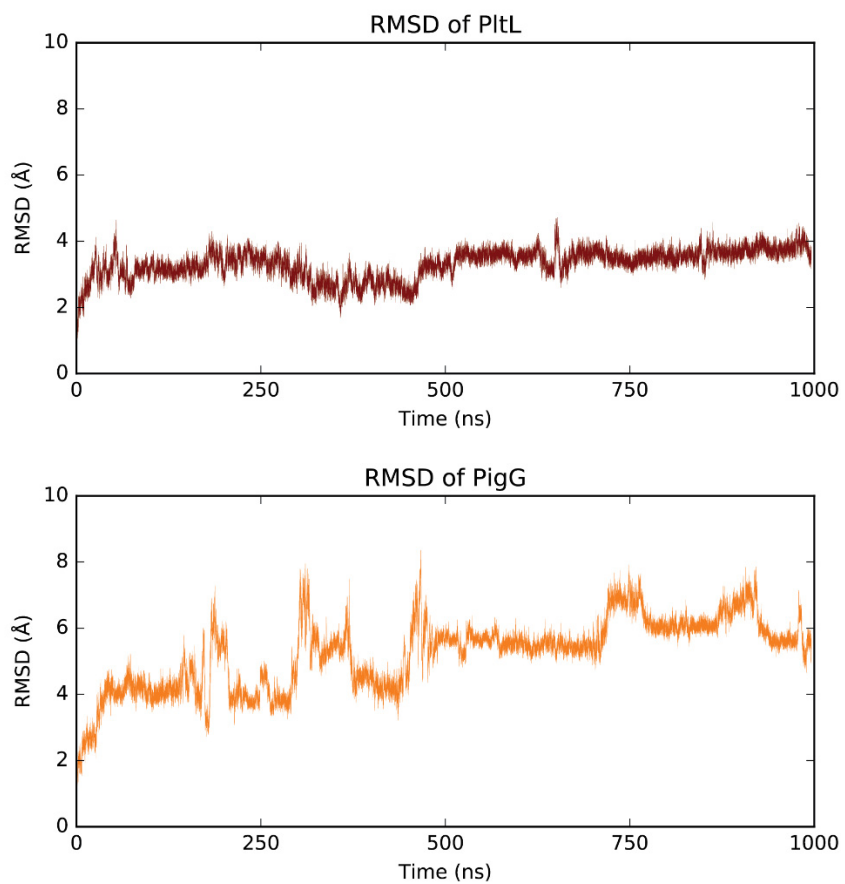


Figure S4.13. Backbone (heavy atom) root mean square deviations (RMSD) of *holo*-PItL (red) and *holo*-PigG (orange). Analysis performed on data collected from Gaussian accelerated molecular dynamics. GaMD simulation yield the same qualitative conclusion as do the cMD simulations: *holo*-PigG is more flexible than *holo*-PItL.

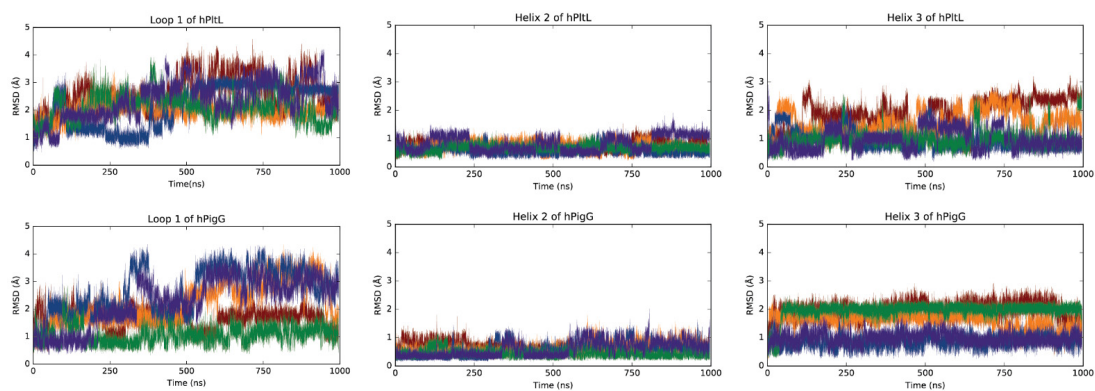


Figure S4.14. Backbone (heavy atom) root mean square deviations (RMSD) of key secondary element of *holo-PilL* (red) and *holo-PigG* (orange). Analysis performed on data collected from conventional molecular dynamics. For both PCPs, helix III is much more disordered than helix II likely due to the presence of proline residue in this region of the proteins.

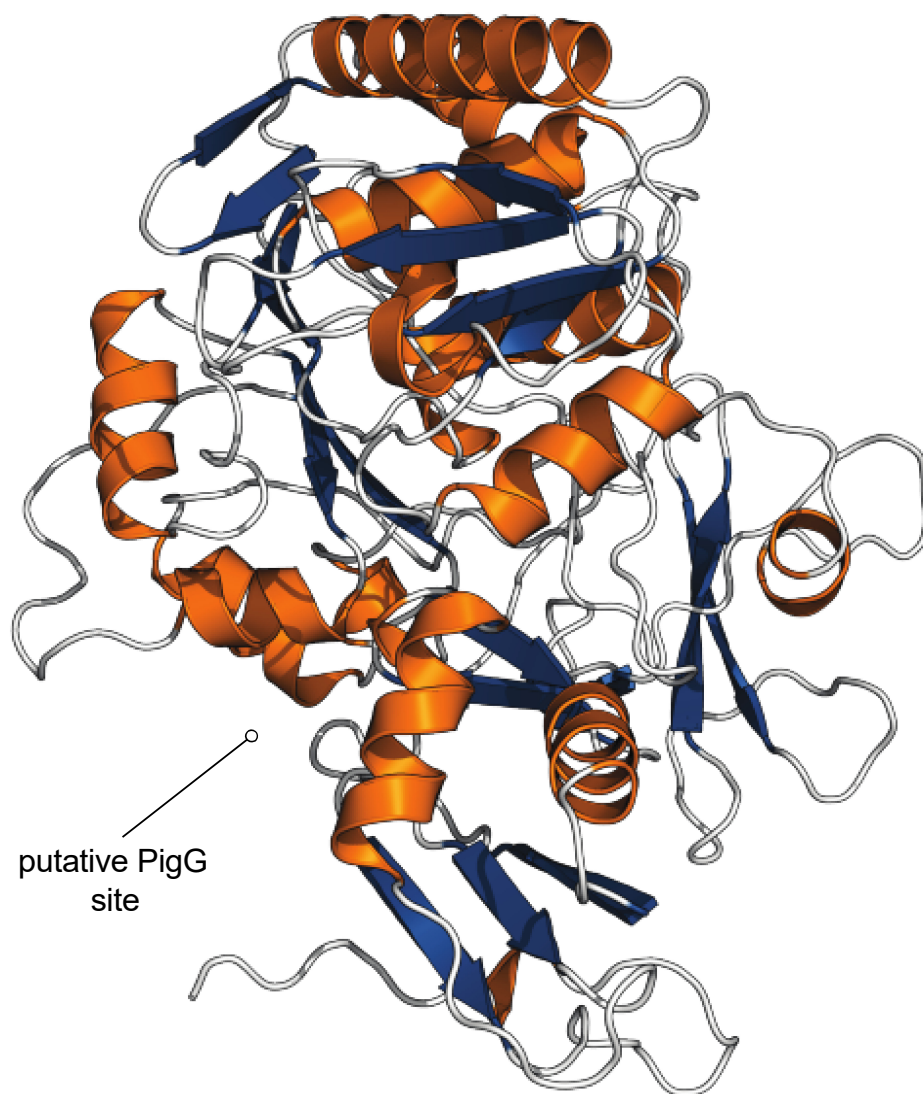


Figure S4.15. Homology model of PigI constructed as described in the main text. This structure was generated in order to generate the structural model of the PigG-PigI complex shown in Figure 4. Helices and sheets are colored in orange and blue in the image. The putative binding pocket (based on alignment of this structure with the EntF (5TAJ) for PigG) is labeled. The coordinates of this model complex and that of PigI have uploaded in pdb format as additional supplementary files.



Figure S4.16. Superimposition of the NMR solution structure of PigG (white) and a model structure of the PigG mutant (orange) that shows altered activity. The portion of the 1st loop of mutant-PigG extended by addition of glycine and proline residue to the PigG primary sequence. Conserved serine residue that is post-translationally phosphopantetheinylated is a white sphere. Ordered helices are also labeled.

Acknowledgements. The material from Chapter 4, in full, is currently being prepared for submission for publication: Jaremko, M. J.; Lee, D. J.; Patel, A.; Winslow, V.; Opella, S. J.; McCammon, J. A.; Burkart, M. D. “Engineering protein-protein interactions in NRPS type II PCPs.” The dissertation author is the primary author of this manuscript.

4.7 References

- (1) Mantovani, S. M.; Moore, B. S. *J. Am. Chem. Soc.* **2013**, *135*, 18032-18035.
- (2) Jaremko, M. J.; Lee, D. J.; Opella, S. J.; Burkart, M. D. *J. Am. Chem. Soc.* **2015**, *137*, 11546-11549.
- (3) Guenzi, E.; Galli, G.; Grgurina, I.; Gross, D. C.; Grandi, G. *J. Biol. Chem.* **1998**, *273*, 32857-32863.
- (4) Williams, P. G.; Buchanan, G. O.; Feling, R. H.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. *J. Org. Chem.* **2005**, *70*, 6196-6203.
- (5) Alvarez-Mico, X.; Jensen, P. R.; Fenical, W.; Hughes, C. C. *Org. Lett.* **2013**, *15*, 988-991.
- (6) Liou, J.-W.; Hung, Y.-J.; Yang, C.-H.; Chen, Y.-C. *Plos One* **2015**, *10*.
- (7) Thomas, M. G.; Burkart, M. D.; Walsh, C. T. *Chem. Biol.* **2002**, *9*, 171-184.
- (8) Galm, U.; Wendt-Pienkowski, E.; Wang, L.; George, N. P.; Oh, T.-J.; Yi, F.; Tao, M.; Coughlin, J. M.; Shen, B. *Mol. Biosyst.* **2009**, *5*, 77-90.
- (9) Lin, S.; Van Lanen, S. G.; Shen, B. *J. Am. Chem. Soc.* **2008**, *130*, 6616-6623.
- (10) Lin, S.; Van Lanen, S. G.; Shen, B. *J. Am. Chem. Soc.* **2007**, *129*, 12432-12438.
- (11) Galonić, D. P.; Vaillancourt, F. H.; Walsh, C. T. *J. Am. Chem. Soc.* **2006**, *128*, 3900-3901.
- (12) Williamson, N. R.; Fineran, P. C.; Leeper, F. J.; Salmond, G. P. C. *Nature Rev. Microbiol.* **2006**, *4*, 887-899.
- (13) Papireddy, K.; Smilkstein, M.; Kelly, J. X.; Shweta; Salem, S. M.; Alhamadsheh, M.; Haynes, S. W.; Challis, G. L.; Reynolds, K. A. *J Med Chem* **2011**, *54*, 5296-5306.
- (14) Liu, W.; Christenson, S. D.; Standage, S.; Shen, B. *Science* **2002**, *297*, 1170-1173.
- (15) Beld, J.; Cang, H.; Burkart, M. D. *Angew. Chem. Int. Ed* **2014**, *53*, 14456-14461.
- (16) Haushalter, R. W.; Filipp, F. V.; Ko, K. S.; Yu, R.; Opella, S. J.; Burkart, M. D. *ACS Chem. Biol.* **2011**, *6*, 413-418.
- (17) Mitchell, C. A.; Shi, C.; Aldrich, C. C.; Gulick, A. M. *Biochemistry* **2012**, *51*, 3252-3263.

- (18) Reimer, J. M.; Aloise, M. N.; Harrison, P. M.; Martin Schmeing, T. *Nature* **2016**, *529*, 239-242.
- (19) Drake, E. J.; Miller, B. R.; Shi, C.; Tarrasch, J. T.; Sundlov, J. A.; Leigh Allen, C.; Skiniotis, G.; Aldrich, C. C.; Gulick, A. M. *Nature* **2016**, *529*, 235-238.
- (20) Mootz, H. D.; Finking, R.; Marahiel, M. A. *J. Biol. Chem.* **2001**, *276*, 37289-37298.
- (21) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. *Nat. Prod. Rep.* **2014**, *31*, 61-108.
- (22) Edgar, R. C. *Nucleic Acids Res.* **2004**, *32*, 1792-1797.
- (23) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13843-13848.
- (24) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'Connor, S. E.; Walsh, C. T. *Nature* **2005**, *436*, 1191-1194.
- (25) Garneau-Tsodikova, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. *J. Am. Chem. Soc.* **2006**, *128*, 12600-12601.
- (26) Magata, Y.; Saji, H.; Tokui, T.; Ohmomo, Y.; Yamada, Y.; Hirata, M.; Konishi, J.; Yokoyama, A. *Ann. Nucl. Med.* **1993**, *7*, 173-177.
- (27) Worthington, A. S.; Burkart, M. D. *Org. Biomol. Chem.* **2006**, *4*, 44-46.
- (28) Lee, W.; Tonelli, M.; Markley, J. L. *Bioinformatics* **2014**, *31*, 1325-1327.
- (29) Lee, W.; Westler, W. M.; Bahrami, A.; Eghbalnia, H. R.; Markley, J. L. *Bioinformatics* **2009**, *25*, 2085-2087.
- (30) Shen, Y.; Bax, A. *J. Biomol. NMR* **2013**, *56*, 227-241.
- (31) Güntert, P. *Eur. Biophys. J.* **2009**, *38*, 129-143.
- (32) 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. Sect. D-Biol. Crystallogr.* **1998**, *54*, 905-921.
- (33) Brunger, A. T. *Nat. Protocols* **2007**, *2*, 2728-2733.
- (34) Bhattacharya, A.; Tejero, R.; Montelione, G. T. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 778-795.

- (35) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *Journal of Chemical Theory and Computation* **2015**, *11*, 3696-3713.
- (36) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *Journal of Computational Chemistry* **1995**, *16*, 1357-1377.
- (37) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09 Revision D.01*.
- (38) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *Journal of molecular graphics & modelling* **2006**, *25*, 247-260.
- (39) Case, D. A.; Babin, V.; Berryman, J.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham Iii, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossvry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. **2014**.
- (40) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *Journal of Computational Chemistry* **2004**, *25*, 1157-1174.
- (41) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219--10225.
- (42) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368-W371.
- (43) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. *Proteins: Structure, Function and Bioinformatics* **2006**, *63*, 928-938.
- (44) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. *Nucleic Acids Res.* **2012**, *40*, W537-W541.
- (45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *Journal of Chemical Physics* **1983**, *79*, 926-935.

- (46) Hamelberg, D.; Mongan, J.; McCammon, J. A. *Journal of Chemical Physics* **2004**, *120*, 11919-11929.
- (47) Pierce, L. C. T.; Salomon-Ferrer, R.; de Oliveira, C. A. F.; McCammon, J. A.; Walker, R. C. *Journal of Chemical Theory and Computation* **2012**, *8*, 2997-3002.
- (48) Salomon-Ferrer, R.; Goetz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. *Journal of Chemical Theory and Computation* **2013**, *9*, 3878-3888.
- (49) Miao, Y.; Feher, V. A.; McCammon, J. A. *Journal of Chemical Theory and Computation* **2015**, *11*, 3584-3595.
- (50) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *Journal of Computational Physics* **1977**, *23*, 327-341.
- (51) Darden, T.; York, D.; Pedersen, L. *Journal of Chemical Physics* **1993**, *98*, 10089-10092.
- (52) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. *Biopolymers* **1992**, *32*, 523--535.
- (53) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proceedings of the National Academy of the Sciences, U. S. A.* **2001**, *98*, 10037-10041.
- (54) Holst, M.; Saied, F. *Journal of Computational Chemistry* **1993**, *14*, 105--113.
- (55) Holst, M. J.; Saied, F. *Journal of Computational Chemistry* **1995**, *16*, 337--364.
- (56) Holst, M. *Advances in Computational Mathematics* **2001**, *15*, 139--191.
- (57) Bank, R. E.; Holst, M. *SIAM review* **2003**, *45*, 291--323.
- (58) Dolinsky, T. J.; Czodrowski, P.; Li, H.; Nielsen, J. E.; Jensen, J. H.; Klebe, G.; Baker, N. A. *Nucleic Acids Res.* **2007**, *35*, W522--525.
- (59) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. *Nucleic Acids Res* **2004**, *32*, W665-667.
- (60) Olsson, M. H. M.; Sndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *Journal of Chemical Theory and Computation* **2011**, *7*, 525--537.
- (61) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179--5197.
- (62) Roe, D. R.; Cheatham, T. E. *Journal of Chemical Theory and Computation* **2013**, *9*, 3084--3095.

- (63) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernandez, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S. *Biophys J* **2015**, *109*, 1528-1532.
- (64) GaussView Version 5. 2016
- (65) The PyMOL Molecular Graphics System, Version 1.7, Schrödinger, LLC. November 2015
- (66) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33-38.
- (67) 2nd ed. International Union of Crystallography: Chester, England April 1, 2012
- (68) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. *J. Mol. Biol.* **2002**, *320*, 597--608.
- (69) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A. *Proteins: Structure, Function and Bioinformatics* **2004**, *55*, 351--367.
- (70) Schrödinger Release 2017-1: Prime, Schrödinger, LLC, New York, NY, 2017.
- (71) Biologics Suite 2017-1, Schrödinger, LLC, New York, NY, 2017.
- (72) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657--1666.
- (73) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225--11236.
- (74) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. *Journal of Chemical Theory and Computation* **2010**, *6*, 1509--1519.
- (75) Lyskov, S.; Gray, J. J. *Nucleic Acids Res.* **2008**, *36*, W233-238.
- (76) Lyskov, S.; Chou, F.-C.; Conchir, S.; Der, B. S.; Drew, K.; Kuroda, D.; Xu, J.; Weitzner, B. D.; Renfrew, P. D.; Sripakdeevong, P.; Borgo, B.; Havranek, J. J.; Kuhlman, B.; Kortemme, T.; Bonneau, R.; Gray, J. J.; Das, R. *Plos One* **2013**, *8*, e63906.
- (77) Chaudhury, S.; Berrondo, M.; Weitzner, B. D.; Muthu, P.; Bergman, H.; Gray, J. J. *Plos One* **2011**, *6*, e22477--e22477.
- (78) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J Mol Biol* **1990**, *215*, 403-410.
- (79) Price, M. N.; Dehal, P. S.; Arkin, A. P. *Plos One* **2010**, *5*, e9490.

Chapter 5. Fluorescent Mechanism-Based Probe for Aerobic Flavin-Dependent Enzyme Activity

Fluorescent Mechanism-Based Probe for Aerobic Flavin-Dependent Enzyme Activity

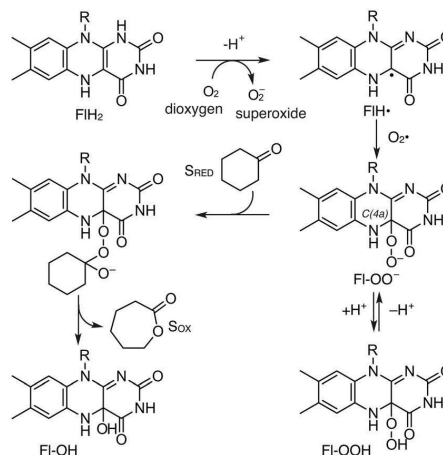
Ian P. McCulloch, James J. La Clair, Matt J. Jaremko, and Michael D. Burkart^{*[a]}

Diversity in non-ribosomal peptide and polyketide secondary metabolism is facilitated by interactions between biosynthetic domains with discrete monomer loading and their cognate tailoring enzymes, such as oxidation or halogenation enzymes. The cooperation between peptidyl carrier proteins and flavin-dependent enzymes offers a specialized strategy for monomer selectivity for oxidization of small molecules from within a complex cellular milieu. In an effort to study this process, we have developed fluorescent probes to selectively label aerobic flavin-dependent enzymes. Here we report the preparation and implementation of these tools to label oxidase, monooxygenase, and halogenase flavin-dependent enzymes.

Natural product biosynthesis occurs through a cascade of enzymatic reactions that combine monomeric building blocks into complex chemical scaffolds.^[1] Over the last decade, we have participated in a global effort to develop methods for the discrete labeling of enzymatic domains responsible for fatty acid, polyketide, and non-ribosomal peptide biosynthesis.^[2] We have reported tools to selectively label domains for the activities of acyl and peptidyl carrier proteins,^[3] ketosynthases,^[4] dehydratases,^[5] and thioesterases.^[6] We now report efforts to deliver probes that label aerobic flavin-dependent enzymes (flavoenzymes) as a first step in delivering selective probes for oxidative enzymes in carrier-protein-dependent pathways.

Aerobic flavoenzymes are oxygen-dependent and catalyze a variety of reactions in natural product biosynthesis, including desaturations, mono-oxygenations, and halogenations.^[7] These enzymes are subcategorized catalytic role into three major classes: oxidases,^[7] monooxygenases,^[8] and halogenases^[9] (Figure S1 in the Supporting Information). Each cycle starts with the O₂ oxidation of reduced flavin (FIH₂) by a radical mechanism, thereby forming the hydroperoxy-containing flavin (FI-OOH) at the C(4a) position (Scheme 1).^[10] The fate of the FI-OOH species is resolved differently by each of the three flavo-enzyme classes (Figure S1).

Mechanism-based inhibitors have been developed for the monoamine oxidase subset of flavin-dependent enzymes: irreversible inhibitors containing *N*-propargylamine moieties, such as pargyline and clorgiline, covalently react with the flavin by oxidation and Michael addition.^[11] These inhibitors have been used to probe substrate specificities in the active sites of these



Scheme 1. Action of a flavin-dependent enzyme: Baeyer-Villiger oxidation of cyclohexanone (S_{RED}) to caprolactone (S_{OX}).^[10]

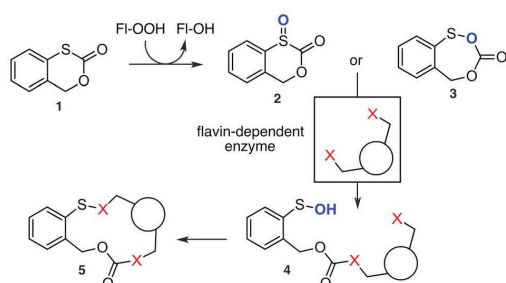
enzymes, and fluorescent derivatives have been used as probes for targeting enzymatic activity.^[12] However, no similar probes have been reported for monooxygenases or halogenases.

In 1987, Latham and Walsh reported mechanism-based inhibition of an *Acinetobacter*-derived cyclohexanone monooxygenase (CHMO) by a cyclic thiocarbonate.^[13] They proposed that, upon oxidation with the FI-OOH species, this motif becomes an electrophilic warhead for enzyme inactivation, but further studies were not pursued. In order to examine the potential of these agents, we turned our attention to **1** as an aromatic thiocarbonate core, for the installation of reporter motifs.^[14] Exposure to FI-OOH would result in the generation of either acyl sulfoxide **2** or sulfenic anhydride **3** (Scheme 2). These reactive intermediates would then be susceptible to nucleophilic attack in the active site of the corresponding flavin-dependent enzyme, thereby leading to covalently linked **4** or fully crosslinked **5**.

We began by testing the inhibitory properties of **1** with a model flavin-dependent enzyme. We chose the BpsA oxidase (Ox) domain for a convenient colorimetric readout, as previously employed for other non-ribosomal peptide synthetase (NRPS)-associated enzymes.^[15] Ox domains within NRPSs catalyze the desaturation or dimerization of carrier-protein-tethered substrates. BpsA Ox catalyzes the desaturation of a piperidinedione intermediate, thus leading to the formation of indigoidine, a blue pigment (Figure S2).^[16] In vitro inhibition assays of BpsA by **1** revealed the IC₅₀ value of 5.6 ± 3.0 μM, K_i = 47 ±

[a] I. P. McCulloch, Dr. J. J. La Clair, M. J. Jaremko, Prof. M. D. Burkart
Department of Chemistry and Biochemistry, University of California
9500 Gilman Drive, San Diego, La Jolla, CA 92093-0358 (USA)
E-mail: mburkart@ucsd.edu

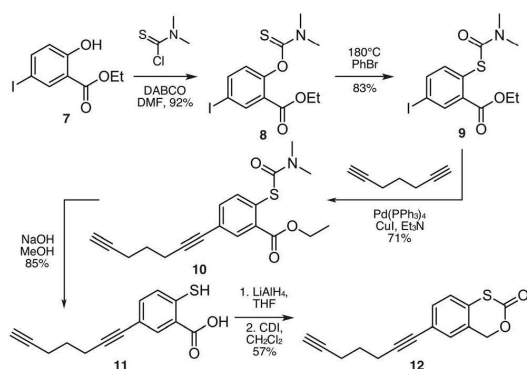
Supporting information and ORCIDs from the authors for this article can be found under <http://dx.doi.org/10.1002/cbic.201600275>.



Scheme 2. Proposed mechanism for the inhibition of flavin-dependent enzymes by **1**. The sphere represents a flavin-dependent enzyme, X (red) denotes nucleophilic residues within a flavin-dependent enzyme (in box) and O (blue) denotes oxygen atoms incorporated by the action of a flavin-dependent enzyme.

3 nm, and $k_{\text{inact}} = 5.6 \pm 0.5 \text{ ms}^{-1}$ (Figure S3). This oxidase inhibition (along with the CHMO inhibition demonstrated by Latham and Walsh) highlights the potential of this probe to target other aerobic flavoenzyme activities. We next turned our attention to adapting **1** into a fluorescent reporter. Given the proposed mechanism and structure of **1** (Scheme 2), we modified the aryl ring of **1** so as to regioselectively install a reporter at a position remote from the active moiety.

The synthesis of **6** began by converting ethyl 2-hydroxy-5-iodobenzoate (**7**) into its corresponding dimethylcarbamothioate **8**. Heating **8** in bromobenzene afforded an effective Miyazaki–Newman–Kwart rearrangement,^[17] thereby providing **9** (76% yield from **7**). After exploring various options for Sonogashira coupling, we found that hepta-1,6-diyne offered a facile route

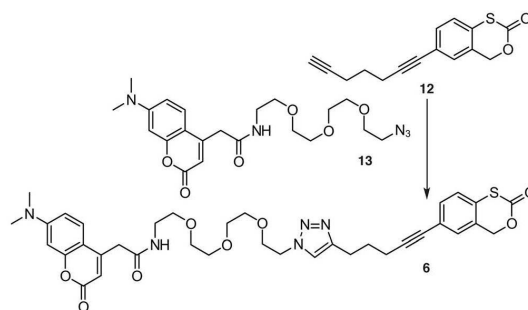


Scheme 3. Probe **12** as prepared in 6 steps in an overall yield of 26% from ethyl 2-hydroxy-5-iodobenzoate (**7**).

to functionalization, thus yielding **10** (Scheme 3) with a terminal mono-substituted alkyne handle for reporter installation. The conversion of **10** into **12** required considerable optimization. The most effective route to install the 1,3-oxathiin-2-one motif

was hydrolysis of **10** to afford **11**, which was immediately reduced and treated with 1,1'-carbonyldiimidazole (CDI) to afford **12** in three back-to-back operations.^[18]

We then explored the incorporation of a reporter tag to **12** by conventional 1,3-dipolar cycloaddition.^[19] We found that the blue fluorescent immunoaffinity azide tag **13**^[20] could be installed to afford **6** under either aqueous conditions with 5% molar equivalents of CuSO_4 and aqueous sodium ascorbate or anhydrous conditions with CuI and Et_3N , respectively, in toluene



Scheme 4. Reporter-labeling of probe **12** was achieved in aqueous (0.05 equiv CuSO_4 in 0.10 equiv sodium ascorbate) and anhydrous conditions (CuI , Et_3N , toluene).

(Scheme 4).

holo-BpsA was generated in vitro by incubation with CoA and a 4'-phosphopantetheinyl transferase, Sfp,^[21] (Figure S4). Both *holo*- and *apo*-BpsA were labeled with **6** (Figure 1A). Labeling of *holo*-BpsA occurred regardless of the presence of ATP and glutamine (Figure 1A), thus indicating that BpsA was

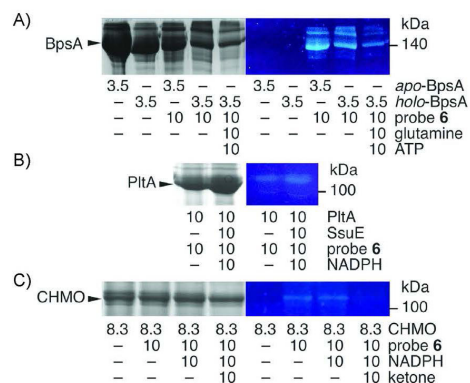


Figure 1. In-gel fluorescent SDS-PAGE analysis of protein labeling. Coomassie-stained (left) and fluorescent (right) SDS-PAGE gels depicting labeling of BpsA, PItA, or CHMO with **6**. A) Labeling of *apo* or *holo*-BpsA, with and without glutamine and ATP. B) Labeling of PItA, with and without SsuE and NADPH. C) Labeling of CHMO, with and without cyclohexanone (ketone) and NADPH. Concentrations are in μM . Full images of the gels are provided in Figure S5.

labeled without substrate dependency or 4'-phosphopantetheinylation.

Next, we investigated the activity with other oxidative flavoenzymes to include the Baeyer–Villiger CHMO (*Rhodococcus* sp. HI-31), which can convert cyclohexanone to caprolactone (Scheme 1),^[8] and the NRPS halogenase PltA (*Pseudomonas fluorescens*) responsible for dichlorination of a pyrrolyl-loaded carrier protein (pyrrolyl-PltL) substrate.^[22]

We found that **6** effectively labeled PltA (Figure 1 B). Addition of SsuE (a flavin reductase compatible with PltA)^[23] and NADPH increased the labeling efficiency (Figure 1 B), but were not essential. MALDI-MS analysis revealed labeling of PltA with **6** occurred in a 1:1 stoichiometry (Figure S7), thus suggesting labeling at a single site on PltA. Multiple attempts were made to identify the modified residue in PltA using tryptic digestion followed by LC-MS-MS analysis. Unfortunately, we were not able to identify peptides bearing modifications of **6** or **1**, although 83% peptide coverage was obtained (data not shown). This observation was consistent with a previous study that used similar inhibitors,^[13] thus suggesting that covalent linkages were not stable to either digestion or LC MS-MS analysis.

CHMO was labeled with **6** under the same conditions (Figure 1 C). Interestingly, addition of cyclohexanone and cofactor NADPH diminished labeling, thus suggesting that **6** is a weak competitive inhibitor of CHMO under the assayed conditions. Interestingly, all three enzymes were labeled without the need of cofactors. This might be attributable to relative abundance of these proteins in the C(4a)-hydroperoxide state, which is thought to predominate in the cell,^[24] studies have demonstrated its relatively long-lasting stability by spectroscopic and crystallographic means.^[24,25]

To further validate labeling specificity, we examined if labeling was selective for the 1,3-oxathiin-2-one motif. The labeling of all three enzymes by **6** decreased in a concentration-dependent manner when the enzymes were pre-incubated with **1** (two- to fourfold excess of **1** blocked labeling; Figure 2 A–C), thus suggesting a single site for inhibition.

Next, we tested selectivity of the probe for each flavoenzyme versus bovine serum albumin (BSA), a protein known to bind to a variety small molecules including lipids, primary metabolites, natural products, synthetic drug leads, and pharmaceuticals.^[26] Probe **6** preferentially labeled BpsA, PltA, and CHMO over BSA (Figure 3 A–C, respectively); no labeling of BSA was detected, thus demonstrating selectivity for flavoenzyme. Furthermore labeling by the 1,3-oxathiin-2-one motif did not occur outside the flavin cofactor pocket (e.g., through release of H₂O₂ or HOCl), as would have resulted from nonspecific labeling.

We then applied our probe to labeling PltA in *Escherichia coli* lysate as a model of in vivo applications. Western blot analysis confirmed labeling of PltA in *E. coli* lysate after addition of 10 μM PltA, 15 μM **6**, and 15 μM NADPH (Figure 4 A–C).^[20] We also conducted these experiments in whole cell lysate. Incubation of 15 μM **6** with an *E. coli* strain engineered to heterologously express PltA resulted in labeled PltA, thus indicating in vivo utility (Figure 4 D). Finally, we examined the role of the flavin cofactor. Under the denaturing and non-reductive condi-

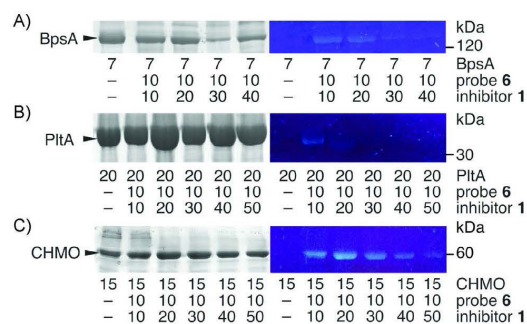


Figure 2. Competition experiments. Coomassie stained (left) and fluorescent (right) SDS-PAGE gels depicting the competition of **1** and **6** in labeling of: A) BpsA, B) PltA, or C) CHMO. Proteins were treated with **1** for 30 min then subjected to labeling with **6**. Concentrations are in μM. Full images of the gels are provided in Figure S6.

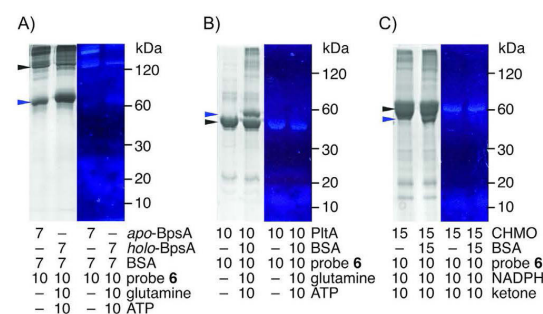


Figure 3. Selectivity analysis. Coomassie-stained (left) and fluorescent (right) SDS-PAGE gels of labeled flavoenzymes BpsA, PltA, or CHMO (black arrows) in the presence of BSA (blue arrows). Gels depicting the labeling of: A) apo-BpsA/holo-BpsA; B) PltA; or C) CHMO. Concentrations are in μM. Additional observed protein bands are due to aggregates of protein from solution or background protein from purification.

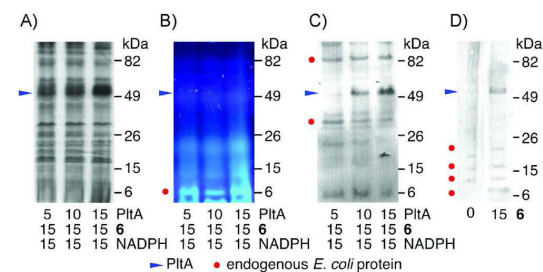


Figure 4. Labeling in PltA in lysate and live cells. A)–C) SDS-PAGE gels depicting 20 μL *E. coli* lysate (1.0 mg mL⁻¹ total protein) supplemented with 5–15 μM PltA. Samples were incubated with **6** and NADPH for 30 min prior to analysis via SDS-PAGE. A) Coomassie-stained gel depicting total protein in the sample. B) Fluorescent image ($\lambda_{exc,max} = 365$ nm) prior to Coomassie staining. C) Western blot from the gel in (A) generated after development with an anti-IAF tag XRI-TF35 mAb (primary) and development by using anti-mouse IgG alkaline phosphatase conjugate antibody (secondary). D) A 5 mL sample of the *E. coli* strain engineered to overexpress PltA (8×10^9 cells per mL) was treated with either 0 or 15 μM **6** for 30 min at 25 °C. The cells were lysed and subjected to SDS-PAGE and western blotting as in (C). Western blot controls for wild-type *E. coli* and PltA expression levels are provided in Figure S9. Concentrations are provided in μM.

tions of SDS-PAGE, the flavin cofactors of PltA, CHMO and BpsA disassociated to discrete, green auto-fluorescent bands (Figure S8; not the blue fluorescence from the tag in 6). In contrast to other flavin-dependent probes,^[12] 6 labeled PltA, CHMO and BpsA regardless of flavin covalency.

We demonstrate here the development of fluorescent-based probes containing a 1,3-oxathiin-2-one motif that can be used to selectively label aerobic flavin-dependent enzymes. This tool allows, through a modular synthetic preparation (Scheme 4), the incorporation of a variety of molecular markers to target these enzymes.^[19b] This development can be further modified to explore the array of oxidative enzymes associated with primary and secondary metabolism, with a focus on substrate selectivity and specificity found within modular syntheses.

Overall, this study identifies an oxidatively-activated warhead for probing aerobic flavin-dependent activity.^[27] The appendage of 1,3-oxathiin-2-one motif onto a protein or related biomolecule offers a tool to selectively probe biomolecular interactivity. With this advance, the inhibitor motif may also be positioned as a tool for crosslinking flavoenzyme domains with cognate partner domains such as carrier proteins, in order to study the crucial protein-protein interactions that guide substrate processing.^[28]

Acknowledgements

Funding was provided from NIH GM095970. We thank Prof. Albert M. Berghuis for the CHMO plasmid construct, Prof. David F. Ackerley for the BpsA plasmid construct, Prof. Christopher T. Walsh for the PltA plasmid construct, Drs. Xuemei Huang and Anthony Mrse for assistance with NMR data collection, Dr. Yongxuan Su for MS services, and Prof. Joris Beld for proofreading support.

Keywords: biosynthesis · flavin-dependent enzymes · fluorescent probes · molecular probes · polyketides · synthase

- [1] a) C. Hertweck, *Angew. Chem. Int. Ed.* **2009**, *48*, 4688; *Angew. Chem.* **2009**, *121*, 4782; b) J. M. Crawford, T. P. Korman, J. W. Labonte, A. L. Vagstad, E. A. Hill, O. Kamari-Bidkorpeh, S.-C. Tsai, C. A. Townsend, *Nature* **2009**, *461*, 1139; c) M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468; d) G. H. Hur, C. R. Vickery, M. D. Burkart, *Nat. Prod. Rep.* **2012**, *29*, 1074.
- [2] a) J. L. Meier, M. D. Burkart, *Curr. Opin. Chem. Biol.* **2011**, *15*, 48; b) J. L. Meier, S. Niessen, H. S. Hoover, T. L. Foley, B. F. Cravatt, M. D. Burkart, *ACS Chem. Biol.* **2009**, *4*, 948.
- [3] J. J. La Clair, T. L. Foley, T. R. Schegg, C. M. Regan, M. D. Burkart, *Chem. Biol.* **2004**, *11*, 195.
- [4] A. S. Worthington, D. F. Porter, M. D. Burkart, *Org. Biomol. Chem.* **2010**, *8*, 1769.
- [5] F. Ishikawa, R. W. Haushalter, M. D. Burkart, *J. Am. Chem. Soc.* **2012**, *134*, 769.
- [6] J. L. Meier, A. C. Mercer, M. D. Burkart, *J. Am. Chem. Soc.* **2008**, *130*, 5443.
- [7] C. T. Walsh, T. A. Wencewicz, *Nat. Prod. Rep.* **2013**, *30*, 175.
- [8] B. J. Yachnin, T. Sprules, M. B. McEvoy, P. C. K. Lau, A. M. Berghuis, *J. Am. Chem. Soc.* **2012**, *134*, 7788.
- [9] F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. Garneau-Tsodikova, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3364.
- [10] a) H. R. Ellis, *Arch. Biochem. Biophys.* **2010**, *497*, 1; b) J. Sucharitakul, R. Tinnikul, P. Chaiyen, *Arch. Biochem. Biophys.* **2014**, *555*, 33; c) S. Ghisla, V. Massey, *Eur. J. Biochem.* **1989**, *181*, 1.
- [11] a) C. Binda, F. Hubálek, M. Li, Y. Herzig, J. Sterling, D. E. Edmondson, A. Mattevi, *J. Med. Chem.* **2004**, *47*, 1767; b) L. De Colibus, M. Li, C. Binda, A. Lustig, D. E. Edmondson, A. Mattevi, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12684.
- [12] a) L. Li, C.-W. Zhang, J. Ge, L. Qian, B.-H. Chai, Q. Zhu, J.-S. Lee, K.-L. Lim, S. Q. Yao, *Angew. Chem. Int. Ed.* **2015**, *54*, 10821; *Angew. Chem.* **2015**, *127*, 10971; b) M. D. Mertens, S. Hinz, C. E. Müller, M. Gütschow, *Bioorg. Med. Chem.* **2014**, *22*, 1916; c) J. M. Krysiak, J. Kreuzer, P. Macheroux, A. Hermetter, S. A. Sieber, R. Breinbauer, *Angew. Chem. Int. Ed.* **2012**, *51*, 7035; *Angew. Chem.* **2012**, *124*, 7142; d) J. Wang, D. E. Edmondson, *Biochemistry* **2011**, *50*, 2499.
- [13] J. A. Latham, C. Walsh, *J. Am. Chem. Soc.* **1987**, *109*, 3421.
- [14] H. Meier, A. Mayer, *Synthesis* **1996**, 327.
- [15] a) H. Takahashi, T. Kumagai, K. Kitani, M. Mori, Y. Matoba, M. Sugiyama, *J. Biol. Chem.* **2007**, *282*, 9073; b) J. G. Owen, J. N. Copp, D. F. Ackerley, *Biochem. J.* **2011**, *436*, 709.
- [16] a) D. Yu, F. Xu, J. Valiente, S. Wang, J. Zhan, *J. Ind. Microbiol. Biotechnol.* **2013**, *40*, 159; b) S. Reverchon, C. Rouanet, D. Expert, W. Nasser, *J. Bacteriol.* **2002**, *184*, 654.
- [17] a) K. Miyazaki, *Tetrahedron Lett.* **1968**, *9*, 2793; b) M. S. Newman, H. A. Karnes, *J. Org. Chem.* **1966**, *31*, 3980; c) H. Kwart, E. R. Evans, *J. Org. Chem.* **1966**, *31*, 410.
- [18] S. Kamila, O. Kahn, H. Zhang, E. R. Biehl, *Synth. Commun.* **2006**, *36*, 1419.
- [19] a) K. Horisawa, *Front. Physiol.* **2014**, *5*, 457; b) J. Martell, E. Weerapana, *Molecules* **2014**, *19*, 1378.
- [20] W.-I. Yu, G. Guizzunti, T. L. Foley, M. D. Burkart, J. J. La Clair, *J. Nat. Prod.* **2010**, *73*, 1659.
- [21] A. S. Worthington, M. D. Burkart, *Org. Biomol. Chem.* **2006**, *4*, 44.
- [22] P. C. Dorrestein, E. Yeh, S. Garneau-Tsodikova, N. L. Kelleher, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13843.
- [23] C. M. Driggers, P. V. Dayal, H. R. Ellis, P. A. Karplus, *Biochemistry* **2014**, *53*, 3509.
- [24] K. C. Jones, D. P. Ballou, *J. Biol. Chem.* **1986**, *261*, 2553.
- [25] A. Alfieri, E. Malito, R. Orru, M. W. Fraaije, A. Mattevi, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6572.
- [26] A. M. Merlot, D. S. Kalinowski, D. R. Richardson, *Front. Physiol.* **2014**, *5*, 299.
- [27] a) M. J. Niphakis, B. F. Cravatt, *Annu. Rev. Biochem.* **2014**, *83*, 341; b) N. C. Sadler, A. T. Wright, *Curr. Opin. Chem. Biol.* **2015**, *24*, 139; c) D. Hunerdosse, D. K. Nomura, *Curr. Opin. Biotechnol.* **2014**, *28*, 116; d) J.-S. Lee, Y.-H. Yoo, C. N. Yoon, *BMB Rep.* **2014**, *47*, 149.
- [28] C. Nguyen, R. W. Haushalter, D. J. Lee, P. R. L. Markwick, J. Bruegger, G. Caldara-Festin, K. Finzel, D. R. Jackson, F. Ishikawa, B. O'Dowd, J. A. McCammon, S. Opella, S.-C. Tsai, M. D. Burkart, *Nature* **2014**, *505*, 427.

Manuscript received: May 13, 2016

Accepted article published: June 8, 2016

Final article published: August 5, 2016

CHEM**BIO**CHEM

Supporting Information

Fluorescent Mechanism-Based Probe for Aerobic Flavin-Dependent Enzyme Activity

Ian P. McCulloch, James J. La Clair, Matt J. Jaremko, and Michael D. Burkart*^[a]

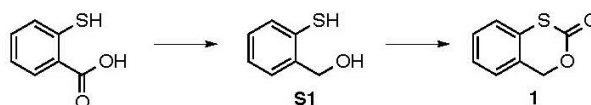
cbic_201600275_sm_miscellaneous_information.pdf

Contents

A. General Experimental Methods	S2
B. Probe Synthesis	S2-S5
C. Protein Expression and Purification	S7
D. Preparation of <i>holo</i> -BpsA	S7
E. Preparation of <i>crypto</i> -BpsA	S7
F. BpsA Inhibition	S7-S8
G. In Vitro Labeling Experiments	S8
H. Competitive in Vitro Labeling Experiments	S8
I. Selective in Vitro Labeling Experiments	S8
J. MALDI-TOFMS Mass Shift Experiments	S8
K. PltA Labeling in <i>E. coli</i> Lysate	S8-S9
L. PltA Labeling in live <i>E. coli</i>	S9
M. Additional References	S9
Table S3 NMR spectroscopic data on probe 6 .	S6
Figure S1. Catalytic cycles of aerobic flavin-dependent enzymes.	S10
Figure S2. Biosynthetic pathways.	S11
Figure S3. Kinetic analyses of BpsA biosynthesis.	S11
Figure S4. Generation of a blue fluorescent <i>crypto</i> -BpsA.	S12
Figure S5. Full gel depiction of gels shown in Figure 1	S12
Figure S6. Full gel depiction of gels shown in Figure 2	S13
Figure S7. MALDI-TOFMS mass shift spectra.	S14
Figure S8. Flavin disassociation by SDS-PAGE.	S14
Figure S9. Western blot analyses.	S15
Figure S10. ¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 7 in CDCl ₃	S16
Figure S11. ¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 8 in CDCl ₃	S17
Figure S12. ¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 9 in CDCl ₃	S18
Figure S13. ¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 10 in CDCl ₃	S19
Figure S14. ¹ H (500 MHz) and ¹³ C NMR (125 MHz) spectra of 12 in CDCl ₃	S20
Figure S15. ¹ H (500 MHz) spectrum of 6 in CDCl ₃	S21
Figure S16. gCOSY (500 MHz) spectrum of 6 in CDCl ₃	S22
Figure S17. TOCSY (600 MHz) spectrum of 6 in CDCl ₃	S23

A. General Experimental Methods. Chemical reagents were purchased from Acros, Fluka, Sigma-Aldrich, or TCI, and used as is. Deuterated NMR solvents were obtained from Cambridge Isotope Laboratories. All reactions were conducted with dried anhydrous solvents that were obtained by passage through a solvent column composed of activated A1 alumina. Triethylamine (Et₃N) was dried over Na and freshly distilled. All reactions were performed under positive pressure of Ar in oven-dried glassware sealed with septa, with stirring from a Teflon coated stir bars using an IKAMAG RCT-basic mechanical stirrer (IKA GmbH). Solutions were heated in a silicon oil bath. Analytical Thin Layer Chromatography (TLC) was performed on Silica Gel 60 F₂₅₄ precoated glass plates (EM Sciences). Visualization was achieved with UV light and/or an appropriate stain (I₂ on SiO₂, KMnO₄, bromocresol green, dinitrophenylhydrazine, ninhydrin, and ceric ammonium molybdate). Flash chromatography was conducted using 40-63 mesh Geduran Silica Gel 60 (EM Biosciences). Yields and characterization data correspond to isolated, chromatographically, and spectroscopically homogeneous materials. ¹H NMR and ¹³C NMR spectra were recorded on a Varian VX500 spectrometer equipped with an XSens Cold probe or a 600 MHz Bruker Avance III spectrometer with a 1.7 mm microcoil cryoprobe. All ¹³C NMR spectra were recorded with complete proton decoupling. Chemical shifts were referenced to the reported values of Gottlieb,^{S1} using the signal from the residual solvent for ¹H spectra, or to the ¹³C signal from the deuterated solvent. Chemical shift δ values for ¹H and ¹³C spectra are reported in parts per million (ppm) relative to these referenced values, and multiplicities are abbreviated as s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br = broad. FID files were processed using MestReNova 10.0.1. (Mestrelab Research). Electrospray Ionization (ESI) mass spectrometric analyses were performed using a LCQ Deca spectrometer (Thermo Scientific) and high-resolution analyses were conducted using a MAT900XL mass spectrometer (Thermo Scientific) with electron impact (EI) ionization. A LTQ Orbitrap XL mass spectrometer (Thermo Scientific) was used for high-resolution electrospray ionization mass spectrometry analysis (HR-ESI-MS).

B. Probe Synthesis. The following sections provide methods and procedures for the syntheses of inhibitor 1, probe 6 and probe 12. Copies of select NMR spectra have been provided for each intermediate at the end of this document.



4H-Benzo[d][1,3]oxathiin-2-one (1). A solution of 2-mercaptobenzoic acid (0.5 g, 3.2 mmol) in THF (10 mL) was added in a drop wise fashion a suspension of LiAlH₄ (0.3 g, 7.9 mmol) in THF (20 mL) at 0°C. The solution was allowed to warm to rt. After 3 h at rt, the mixture was cooled to 0 °C and H₂O (10 mL) and 6 M HCl (5 mL) were added sequentially in drop wise fashion. The resulting mixture was poured into additional H₂O (50 mL) and extracted with EtOAc (3 × 100 mL). The organic layer was dried over anhydrous Na₂SO₄, filtered, and concentrated via rotary evaporation to afford carbinol **S1**, as a brown oil (0.44 g, 98%). Crude **S1** was dissolved in CH₂Cl₂ (50 mL) and carried on without purification. Over the course of 1 min, *N,N'*-carbonyldiimidazole (0.56 g, 3.44 mmol) was added in a portion wise fashion. After 14 h at rt, H₂O (50 mL) was added and the aqueous layer was extracted with CH₂Cl₂ (2 × 50 mL). The organic layers were collected, dried over anhydrous Na₂SO₄, filtered, and concentrated via rotary evaporation. Purification by flash chromatography eluting with CH₂Cl₂ afforded 0.39 g (74% yield) of 4H-benzo[d][1,3]oxathiin-2-one (**1**), as a white crystalline solid.

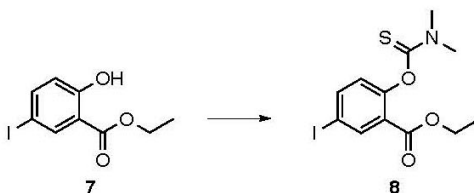
Inhibitor 1: m.p. 58-60°C; R_f = 0.63 (CH₂Cl₂); ¹H NMR (500 MHz, CDCl₃) δ 7.30-7.41 (m, 4H), 5.29 (s, 2H); ¹³C NMR (125 MHz, CDCl₃) δ 166.6, 131.2, 129.9, 128.9, 127.8, 126.8, 126.5, 72.0; FT-IR: 1687, 1581, 1471, 1447, 1388, 1248, 1218, 1161, 1137, 1108, 1062, 1000, 971,

939, 870, 837, 751 cm^{-1} ; HRMS (ESI-TOF) m/z $[\text{M}+\text{H}]^+$ calcd. for $\text{C}_8\text{H}_7\text{O}_2\text{S}$, 167.0167; found, 167.0156.



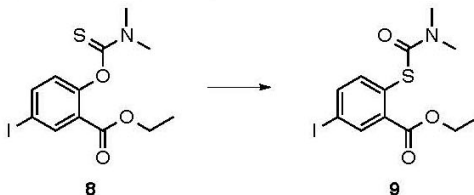
Ethyl 2-hydroxy-5-iodobenzoate (7): Chloramine-T (16.4 g, 0.072 mol) was added over 5 min in three portions to a solution of NaI (10.5 g, 0.07 mol) and ethyl salicylate (10.0 g, 0.60 mol) in anhydrous DMF (50 mL) at 0°C . The solution was allowed to warm to rt. After 3 h at rt, the solution was then poured into H_2O (300 mL) and the pH was adjusted to 3 by the addition of 2 N HCl. The solution was filtered and the solid was washed with H_2O (100 mL) and 10% $\text{Na}_2\text{S}_2\text{O}_3$. Purification by flash chromatography eluting with CH_2Cl_2 afforded 12.9 g (74% yield) of ester 7, as a white crystalline solid.

Ester 7: m.p. $61\text{--}62^\circ\text{C}$; $R_f = 0.84$ (CH_2Cl_2); ^1H NMR (500 MHz, CDCl_3) δ 10.82 (s, 1H), 8.12 (d, $J = 2.3$ Hz, 1H), 7.67 (dd, $J = 8.8$ Hz, 2.3 Hz, 1H), 6.75 (d, $J = 8.8$ Hz, 1H), 4.40 (q, $J = 7.1$ Hz, 2H), 1.41 (t, $J = 7.2$ Hz, 3H); ^{13}C NMR (125 MHz, CDCl_3) δ 169.1, 161.4, 144.0, 138.3, 120.1, 114.8, 80.1, 62.0, 14.3; HRMS (ESI-TOF) m/z $[\text{M}-\text{H}]^-$ calcd. for $[\text{C}_9\text{H}_8\text{IO}_3]^-$, 290.9524; found, 290.9528.



Ethyl 2-((dimethylcarbamothioyl)oxy)-5-iodobenzoate (8): Dimethylthiocarbamoyl chloride (7.5 g, 0.075 mol) was added over 3 min to a solution of ester 7 (8.8 g, 0.030 mol) and DABCO (8.5 g, 0.085 mol) in anhydrous DMF (250 mL). After stirring for 18 h at rt, the solution was poured into H_2O (500 mL) and the pH was adjusted to 4 by the addition of 3 N HCl. The off-white precipitate was collected and dried under vacuum. Purification by flash chromatography eluting with hexanes/EtOAc (95:5) afforded 10.5 g (92% yield) of *S*-thiocarbamate 8, as a white crystalline solid.

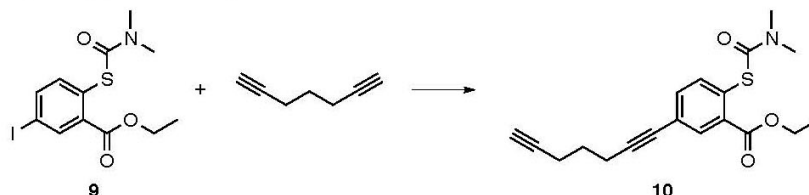
***S*-thiocarbamate 8:** m.p. $105\text{--}107^\circ\text{C}$; $R_f = 0.50$ (CH_2Cl_2); ^1H NMR (500 MHz, CDCl_3) δ 8.29 (d, $J = 2.3$ Hz, 1H), 7.83 (dd, $J = 8.5$, 2.3 Hz, 1H), 6.85 (d, $J = 8.5$ Hz, 1H), 4.29 (q, $J = 7.2$ Hz, 2H), 3.43 (s, 3H), 3.36 (s, 3H), 1.32 (t, $J = 7.2$ Hz, 3H); ^{13}C NMR (125 MHz, CDCl_3) δ 186.9, 163.1, 153.4, 142.1, 140.2, 126.9, 126.3, 89.8, 61.5, 43.3, 39.0, 14.2; HRMS (ESI-TOF) m/z $[\text{M}+\text{Na}]^+$ calcd. for $[\text{C}_{12}\text{H}_{14}\text{INO}_3\text{SNa}]^+$, 401.9631; found, 401.9630.



Ethyl 2-((dimethylcarbamoylthio)-5-iodobenzoate (9): A solution of *S*-thiocarbamate 8 (2.0 g, 5.3 mmol) was dissolved in bromobenzene (5 mL) and heated to 165°C for 24 h in a pressure sealed flask (Sigma-Aldrich). The mixture was concentrated to afford a crude product.

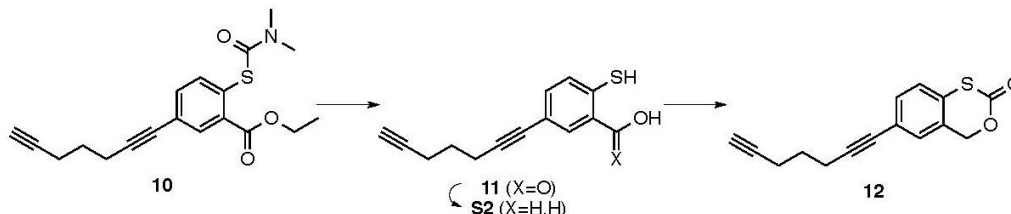
Purification by flash chromatography eluting with hexanes/EtOAc (95:5) afforded 1.7 g (83% yield) of *O*-thiocarbamate **9**, as a white crystalline solid.

***O*-Thiocarbamate 9**: m.p. 82-83°C; $R_f = 0.26$ (CH_2Cl_2); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 8.17 (d, $J = 2.0$ Hz, 1H), 7.76 (dd, $J = 8.3, 2.1$ Hz, 1H), 7.28 (d, $J = 8.4$ Hz, 1H), 4.32 (q, $J = 7.1$ Hz, 2H), 3.09 (bs, 3H), 2.98 (bs, 3H), 1.35 (t, $J = 7.2$ Hz, 3H), $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 165.5, 165.2, 140.2, 139.1, 138.6, 136.7, 129.6, 94.5, 61.6, 37.0, 14.2, HRMS (ESI-TOF) m/z $[\text{M}+\text{Na}]^+$ calcd for $[\text{C}_{12}\text{H}_{14}\text{INO}_3\text{SNa}]^+$, 401.9631; found, 401.9632.



Ethyl 2-((dimethylcarbamoyl)thio)-5-(hepta-1,6-diyn-1-yl)benzoate (10). *O*-thiocarbamate **9** (600.0 mg, 1.58 mmol) was added to a suspension of CuI (30.1 mg, 0.158 mmol) and $\text{Pd}(\text{PPh}_3)_4$ (91.4 mg, 0.079 mmol) in Et_3N (15 mL). After degassing, hepta-1,6-diyne (292.0 mg, 3.16 mmol) was added as a neat liquid. After 3 h at rt, the crude product was obtained by rotary evaporation. Purification by flash chromatography eluting with hexanes/EtOAc (95:5) afforded 385.0 mg (71% yield) of *bis*-alkyne **10**, as a yellow oil.

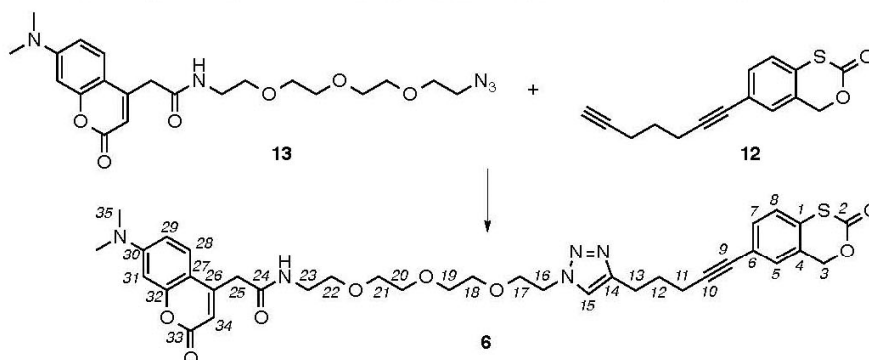
***Bis*-alkyne 10**: $R_f = 0.18$ (CH_2Cl_2); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.88 (d, $J = 1.9$ Hz, 1H), 7.51 (d, $J = 8.1$ Hz, 1H), 7.45 (dd, $J = 8.1, 1.9$ Hz, 1H), 4.33 (q, $J = 7.1$ Hz, 2H), 3.11 (bs, 3H), 3.00 (bs, 3H), 2.55 (t, $J = 7.0$ Hz, 2H), 2.37 (td, $J = 7.0, 2.6$ Hz, 2H), 1.99 (t, $J = 2.6$ Hz, 1H), 1.83 (p, $J = 7.0$ Hz, 2H), 1.37 (t, $J = 7.1$ Hz, 3H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 166.2, 166.0, 137.2, 135.2, 134.1, 133.6, 129.0, 124.8, 91.6, 83.5, 80.1, 69.1, 61.5, 37.1, 27.5, 18.6, 17.7, 14.3; HRMS (ESI-TOF) m/z $[\text{M}+\text{H}]^+$ calcd for $[\text{C}_{19}\text{H}_{22}\text{NO}_3\text{S}]^+$, 344.1315; found, 344.1314.



6-(Hepta-1,6-diyn-1-yl)-4*H*-benzo[*d*][1,3]oxathiin-2-one (12). *Bis*-alkyne **10** (150 mg, 0.44 mmol) was dissolved in 2.5 M NaOH (10 mL) and warmed to 100°C in a sealed pressure flask (Sigma-Aldrich). After 14 h, the mixture was cooled and poured into cold H_2O (50 mL) and extracted with EtOAc (50 mL) to remove organic impurities. The aqueous layer was adjusted to pH 3 by the addition of 6 N HCl . The resulting mixture was extracted with EtOAc (3×100 mL), dried over Na_2SO_4 , filtered, and concentrated via rotary evaporation. The resulting product **11** (150 mg, 0.61 mmol) was dissolved in THF (5 mL), and added in a drop wise manner to a slurry of LiAlH_4 (70 mg, 1.84 mmol) in THF (10 mL) at 0°C. The mixture was allowed to warm to rt. The progress of the reduction was followed by TLC and worked up once the **11** was consumed, typically ~3 h. The mixture was cooled to 0°C and ~2 mL H_2O was added in a drop wise manner. The resulting solution was adjusted to pH 4 by the addition 3N HCl . The resulting mixture was extracted with EtOAc (3×100 mL), dried over Na_2SO_4 , filtered, and concentrated via rotary evaporation to afford carbinol **S2**, as a brown oil. Crude **S2** dissolved in anhydrous CH_2Cl_2 (100 mL) and treated with *N,N'*-carbonyl diimidazole (299 mg, 1.84 mmol). After 3 h, the

mixture was concentrated via rotary evaporation. Purification by flash chromatography eluting with hexanes/EtOAc (95:5) afforded 65.0 mg (57% yield) of probe **12**, as a wax.

Probe 12: $R_f = 0.59$ (CH_2Cl_2); $^1\text{H NMR}$ (500 MHz, CDCl_3) δ 7.39 (dd, $J = 8.1, 1.8$ Hz, 1H), 7.34 (d, $J = 1.6$ Hz, 1H), 7.22 (d, $J = 8.1$ Hz, 1H), 5.24 (s, 2H), 2.55 (t, $J = 7.0$ Hz, 2H), 2.36 (td, $J = 7.0, 2.6$ Hz, 2H), 2.00 (t, $J = 2.6$ Hz, 1H), 1.82 (p, $J = 7.0$ Hz, 2H); $^{13}\text{C NMR}$ (125 MHz, CDCl_3) δ 166.0, 132.7, 130.3, 129.5, 128.7, 126.2, 123.5, 91.1, 83.5, 80.0, 71.6, 69.2, 27.5, 18.6, 17.7; HRMS (ESI-TOF) m/z $[\text{M}+\text{Na}]^+$ calcd for $[\text{C}_{15}\text{H}_{12}\text{O}_2\text{SNa}]^+$, 279.0456; found, 279.0944.



2-(7-(dimethylamino)-2-oxo-2H-chromen-4-yl)-N-(2-(2-(2-(2-(4-(5-(2-oxo-4H-benzo[d][1,3]oxathiin-6-yl)pent-4-yn-1-yl)-1H-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethoxy)ethyl)acetamide (6). Samples of probe **6** were prepared *in situ* and used immediately after preparation without purification. Aliquots of these products were collected and checked prior by NMR (copies of selected NMR spectra are within this file) and LC-MS analysis (HRMS ESI-TOF m/z $[\text{M}+\text{H}]^+$ calcd for $[\text{C}_{36}\text{H}_{42}\text{N}_5\text{O}_8\text{S}]^+$, 704.2749; found, 704.2777). Two methods, an aqueous and organic procedure, were used to prepare **6**, and both provided comparable results. The following provide exemplary procedures.

Aqueous protocol: An aliquot of 2 mM $\text{CuSO}_4 \cdot 5 \text{H}_2\text{O}$ (100 μL , 0.20 μmol) and 4 mM sodium ascorbate (100 μL , 0.40 μmol) was added to mixture of tag **13**^{S2} (2.0 mg, 4.47 μmol) and inhibitor **12** (1.0 mg, 3.90 μmol) in 2:1 butanol/ H_2O (1.5 mL). After 24 h at rt, the mixture was concentrated *via* rotary evaporation and used as is.

Organic protocol: CuI (0.5 mg, 2.6 μmol) and Et_3NiPr_2 (122 μL , 11.7 μmol) were added sequentially to a mixture of tag **13**^{S2} (2.0 mg, 4.47 μmol) and inhibitor **12** (1.0 mg, 3.90 μmol) anhydrous toluene (1.5 mL). After 24 h at rt, the mixture was concentrated *via* rotary evaporation and used as is. If desired, increased purity can be obtained by HPLC purification or purified by preparative TLC, however, we found that the crude product provided comparable labeling and hence purification was not required.

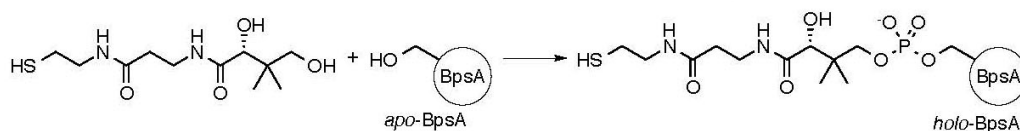
Probe 6: $R_f = 0.16$ (10% MeOH in CH_2Cl_2); $^1\text{H NMR}$ (500 MHz, CDCl_3) and $^{13}\text{C NMR}$ (125 MHz, CDCl_3) are provided in Table S1; HRMS ESI-TOF m/z $[\text{M}+\text{H}]^+$ calcd for $[\text{C}_{36}\text{H}_{42}\text{N}_5\text{O}_8\text{S}]^+$, 704.2749; found, 704.2777).

Table S1. NMR spectroscopic data for probe 6.^a

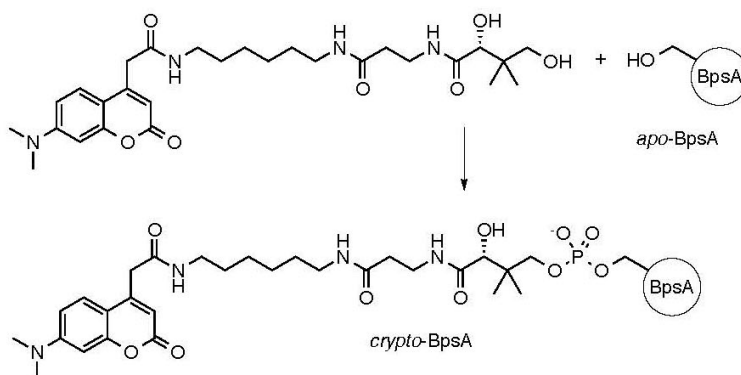
No.	δ C, type	δ H (multiplicity, <i>J</i> in Hz)	COSY
1	131.6, C		
2	167.1, C		
3	71.6, CH ₂	5.25 (s)	
4	130.3, C		
5	128.8, CH	7.36 (d, 1.6)	7
6	124.7, C		
7	132.8, CH	7.39 (dd, 1.8, 8.1)	5,8
8	126.4, CH	7.22 (d, 8.1)	7
9	81.5, C		
10	92.7, C		
11	19.1, CH ₂	2.49 (t, 6.9)	12
12	28.2, CH ₂	2.00 (p, 6.8)	11,13
13	24.6, CH ₂	2.89 (t, 7.5)	12
14	147.8, C		
15	126.1, CH	7.51 (bs)	
16	50.7, CH ₂	4.51 (t, 4.4)	17
17	69.5, CH ₂	3.89 (t, 5.2)	16
18	70.7, CH ₂	3.62 (m)	n.a.
19	70.6, CH ₂	3.58 (m)	n.a.
20	70.4, CH ₂	3.50 (m)	n.a.
21	70.4, CH ₂	3.50 (m)	n.a.
22	70.4, CH ₂	3.50 (m)	n.a.
23	39.8, CH ₂	3.44 (m)	n.a.
24	151.4, C		
25	40.7, CH ₂	3.65 (s)	
26	110.1, C		
27	110.2, C		
28	122.8, CH	7.53 (d, 9.0)	29
29	109.6, CH	6.62 (dd, 2.6, 9.0)	28,31
30	157.3, C		
31	98.6, CH	6.50 (d, 2.6)	29
32	154.2, C		
33	169.4, C		
34	110.5, CH	6.07 (s)	
35	40.4, CH ₃	3.06 (s)	
NH		6.58 (bs)	

^a NMR data was collected in CDCl₃ at 23 °C using ¹H NMR (500 MHz) and ¹³C NMR (125 MHz). n.a. denotes overlapping or unresolved peaks.

C. Protein Expression and Purification. The plasmid constructs were obtained from prior studies as given by BpsA with a *N*-terminal and *C*-terminal His tag,^{S3} PltA with a *N*-terminal His tag^{S4} and CHMO with an *N*-terminal His tag.^{S5} BpsA and PltA were expressed in *E. coli* BL21(DE3). CHMO was expressed in *E. coli* Rosetta (pLysS, DE3). Starter cultures of each strain were grown in 5 mL of LB supplemented with 50 µg/mL kanamycin for BpsA and PltA, respectively, or supplemented with 100 µg/mL ampicillin and 25 µg/mL chloramphenicol for CHMO. Starter cultures were inoculated into 1L of LB expression cultures containing 50 µg/mL kanamycin for BpsA and PltA or 100 µg/mL ampicillin and 25 µg/mL chloramphenicol for CHMO. After growth at 37°C to an OD₆₀₀ of 0.5-0.6, cells were induced with 1 mM IPTG and further cultured at 16°C. After 16 h at 16°C, the cells were harvested by centrifugation (1000×g, 30 min). The resulting cell pellets were resuspended in binding buffer (50 mM Tris·HCl, pH 8.0, 250 mM NaCl, 10% (v/v) glycerol), incubated with lysozyme for 30 min on ice, and lysed by passage through a French press. The lysate was clarified by centrifugation (12,000×g, 30 min, 4°C), and Ni-NTA resin (1.0 mL) was added. After batch binding with resin for 1 h at 4°C, the resin was washed with wash buffer (20 mM imidazole, 50 mM Tris·HCl, pH 8.0, 250 mM NaCl, 10% (v/v) glycerol) and eluted with elution buffer (250 mM imidazole, 50 mM Tris·HCl, pH 8.0, 250 mM NaCl, 10% (v/v) glycerol). Eluted fractions were concentrated and dialyzed into storage buffer (50 mM Tris·HCl, pH 8.0, 250 mM NaCl, 10% (v/v) glycerol). Proteins were flash frozen and stored at -80°C until thawed for usage.



D. Preparation of *holo*-BpsA. Conversion of *apo*-BpsA into *holo*-BpsA was accomplished by incubation of 3.5 µM *apo*-BpsA, 0.25 µM Sfp (a phosphopantetheinyl transferase from *Bacillus subtilis*),^{S6} 10 mM MgCl₂, 100 µM coenzyme A (CoA) in 50 mM sodium phosphate buffer pH 7.8. Reactions were conducted at 1000 µL volume for 30 min at 37°C and 100 µL aliquots were used immediately in inhibition and labeling experiments.



E. Preparation of *crypto*-BpsA. Conversion of *apo*-BpsA into *crypto*-BpsA was accomplished by incubation of 2.3 µM *apo*-BpsA, 0.25 µM Sfp, 10 mM MgCl₂, 100 µM pantetheine analog **S3**, 0.25 µM CoaA, 0.25 µM CoaD, and 0.25 µM CoaE in 50 mM sodium phosphate buffer pH 7.8 using the one-pot method.^{S6} Reactions were conducted at 34 µL volume for 30 min at 37°C and 10 µL aliquots were assayed by SDS-PAGE.

F. BpsA inhibition. IC₅₀, K_i, and K_{inact} values were determined across a serial dilution series of inhibitor 1 from 50 µM to 78 nM. For this procedure, reactions were prepared at 150 µL volume

containing 0.30 μM *holo*-BpsA (see Section C), 1 mM ATP, 5 mM glutamine, 10 mM MgCl_2 in 50 mM sodium phosphate pH 7.8. The inhibitor was from stocks in CH_3CN such that samples had <5% (v/v) CH_3CN . Each reaction was run in triplicate in a 96 well microplate. Indigoidine production was measured by absorbance at 590 nm for 15 min by a HTS 7000 Bio Assay Reader (Perkin Elmer). The initial velocity of activity as a fraction of the initial velocity at a zero inhibitor concentration (V_i/V_o) was calculated and plotted against the logarithm of inhibitor concentration for each sample using Prism 6 (GraphPad Software), as shown in Figure S3a. An IC_{50} value of $5.6 \pm 3.0 \mu\text{M}$ was obtained by a nonlinear least-squares fit of the log plot. To determine K_i and k_{inact} values, the rate constants (k_{obs}) were determined for each condition and plotted against the inhibitor concentration (Figure S3b), fitting to a hyperbolic curve where $k_{\text{obs}} = k_{\text{inact}} \cdot [I] / (K_i + [I])$. ($K_i = 47 \pm 3 \text{ nM}$ and $k_{\text{inact}} = 5.6 \pm 0.5 \text{ ms}^{-1}$).

G. *In vitro* Labeling Experiments. A 20 μL aliquot of 3.5 μM BpsA, 10 μM PltA, or 8.3 μM CHMO in 50 mM potassium phosphate buffer pH 7.6 was incubated with 10 μM probe 6 from by the addition of 1 μL of 500 μM 6 in CH_3CN . For *holo*-BpsA optimization conditions, glutamine and ATP were each added to a final concentration of 10 μM . For PltA optimization conditions, SsuE (from *Escherichia coli*)^{S4} and NADPH were each added to a final concentration of 10 μM . For CHMO optimization conditions, NADPH and cyclohexanone were each added to a final concentration of 10 μM . After 30 min at rt, the sample was diluted with 20 μL SDS-containing loading dye (100 mM Tris·HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, and 0.2% (w/v) bromophenol blue) boiled at 95°C for 5 min, and subjected to SDS-PAGE analysis followed by fluorescence imaging and staining with Coomassie Brilliant Blue (0.05% (w/v) Coomassie Brilliant Blue, 50% (v/v) methanol, 10% (v/v) glacial HOAc, 40% H_2O).

H. Competitive *In vitro* Labeling Experiments. A 20 μL aliquot of 7 μM BpsA, 20 μM PltA, or 15 μM CHMO was treated 10 μM to 50 μM inhibitor 1 by the addition of a 1 μL stock in CH_3CN . After 30 min at rt, the solution was treated with 10 μM probe 6 by the addition of 1 μL of 500 μM 6 in CH_3CN . After 30 min at rt, the sample was diluted with 20 μL SDS-containing loading dye (100 mM Tris·HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, and 0.2% (w/v) bromophenol blue) boiled at 95°C for 5 min, and subjected to SDS-PAGE analysis followed fluorescence imaging and staining with Coomassie Brilliant Blue.

I. Selective *In vitro* Labeling Experiments. A 20 μL aliquot containing 7 μM BpsA and 7 μM bovine serum albumin (BSA), 10 μM PltA and 10 μM BSA, and 15 μM CHMO and 15 μM BSA in 50 mM potassium phosphate buffer pH 7.6 was incubated with 10 μM probe 6 from by the addition of 1 μL of 500 μM 6 in CH_3CN . After 30 min at rt, the sample were was diluted with 20 μL SDS-containing loading dye (100 mM Tris·HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, and 0.2% (w/v) bromophenol blue) boiled at 95°C for 5 min, and subjected to SDS-PAGE analysis followed fluorescence imaging and staining with Coomassie Brilliant Blue.

J. MALDI-TOFMS Mass Shift Experiments. Two 20 μL samples of 10 μM PltA were either incubated with CH_3CN (control) or 10 μM probe 6 at rt for 1 h. Samples were desalted and purified by C4 ZipTip pipette tips (EMD Millipore) eluting with 0.1% TFA in 1:1 $\text{CH}_3\text{CN}/\text{H}_2\text{O}$. Samples and a protein molecular weight standard were added to caffeic acid matrix solution and dried on a matrix until crystalline. MALDI-TOFMS (Bruker Biflex IV, 337 nm laser) was used to detect ionized protein masses, normalized by protein standards (Fig. S7): [PltA] = 52797.38 m/z ; [PltA + Probe 6] = 53497.83 m/z . A mass shift of 700.45 m/z was observed (703.27 m/z expected) corresponding to the mass of probe 6.

K. PltA Labeling in *E. coli* Lysate. Aliquots containing 20 μL of lysate from: a 1 mg/mL net protein from *E. coli* (BL21) supplemented with either 5 μM , 10 μM , or 15 μM PltA or 1 mg/mL net protein from *E. coli* (BL21) lysate were added to a 250 μL Eppendorf tube. Stocks of NADPH and probe 6 were then added to final concentration of 15 μM . Samples were incubated for 30

min at rt, diluted with 20 μ L SDS-PAGE loading buffer (100 mM Tris-HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, and 0.2% (w/v) bromophenol blue), boiled at 95°C for 5 min, and subjected to SDS-PAGE analysis followed by fluorescence imaging or Western blotting. Western blotting was conducted using the following procedure: a sample containing 10-30 μ g of total protein was loaded per well of a NuPage 4-12% Bis-Tris SDS-PAGE gel (ThermoFisher Scientific) and run using MOPS running buffer (ThermoFisher Scientific). Gels were wet transferred onto a polyvinylidene fluoride (PVDF) membrane and blocked in 5% dry milk. After blocking, the membrane was incubated overnight at 4°C with 50 μ L of a 5 mg/mL stock of mouse TF-35 anti-IAF mAb (Xenobe Research Institute) diluted 1:1000 in Tris-buffered saline containing 0.01% tween (TBST) and 5% dry milk. The membrane was washed four times with an equal volume of TBST. The blot was then incubated for 1 h in 5% dry milk-TBST with an anti-mouse IgG alkaline phosphatase conjugate antibody (Promega) at 1:10000. After washing three times with TBST, detection was performed using NBT/BCIP (34042, Pierce) according to manufacturer's instructions.

L. PltA Labeling in live *E. coli*. Cultures containing 5 mL LB and *E. coli* inoculates bearing a PltA plasmid were grown to OD₆₀₀ of 0.6 at 37 °C, induced with IPTG at 37 °C, grown to OD₆₀₀ of 1.0 ($\sim 8 \times 10^8$ cells/mL) at 37°C, harvested by centrifugation (1000 \times g, 30 min), and resuspended in 1 mL buffer (50 mM sodium phosphate, 250 mM NaCl, pH 7.2). Samples were treated without probe or with 15 μ M probe **6** for 30 min at 25 °C, lysed by sonication, and clarified by centrifugation (12,000 \times g, 30 min, 4°C). Aliquots containing 15 μ L of lysate were then diluted with 15 μ L SDS-containing loading dye (100 mM Tris-HCl pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, and 0.2% (w/v) bromophenol blue) boiled for 5 min, and subjected to SDS-PAGE followed by Western blotting. Western blotting was conducted using the following procedure: a sample containing 10-30 μ g of total protein was loaded per well of a NuPage 4-12% Bis-Tris SDS-PAGE gel (ThermoFisher Scientific) and run using MOPS running buffer (ThermoFisher Scientific). Gels were wet transferred onto a polyvinylidene fluoride (PVDF) membrane and blocked in 5% dry milk. After blocking, the membrane was incubated overnight at 4°C with 50 μ L of a 5 mg/mL stock of mouse TF-35 anti-IAF mAb (Xenobe Research Institute) diluted 1:1000 in Tris-buffered saline containing 0.01% tween (TBST) and 5% dry milk. The membrane was washed four times with an equal volume of TBST. The blot was then incubated for 1 h in 5% dry milk-TBST with an anti-mouse IgG alkaline phosphatase conjugate antibody (Promega) at 1:10000. After washing three times with TBST, detection was performed using NBT/BCIP (34042, Pierce) according to manufacturer's instructions.

M. Additional References.

- (S1) H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512.
- (S2) J. L. Meier, A. C. Mercer, H. Rivera, M. D. Burkart, *J. Am. Chem. Soc.* **2006**, *128*, 12174-84.
- (S3) J. G. Owen, J. N. Copp, D. F. Ackerley, *Biochem. J.* **2011**, *436*, 709-717.
- (S4) P. C. Dorrestein, E. Yeh, S. Garneau-Tsodikova, N. L. Kelleher, C. T. Walsh, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13843-13848.
- (S5) B. J. Yachnin, T. Sprules, M. B. McEvoy, P. C. Lau, A. M. Berghuis, *J. Am. Chem. Soc.* **2012**, *134*, 7788-7795.
- (S6) A. S. Worthington, M. D. Burkart, *Org. Biomol. Chem.* **2006**, *4* (1), 44-6.
- (S7) C. T. Walsh, T. A. Wenczewicz, *Nat. Prod. Rep.* **2013**, *30*, 175-200.
- (S8) H. Takahashi, T. Kamaqai, K. Kitani, M. Mori, *J. Biol. Chem.* **2007**, *282*, 9073-81.

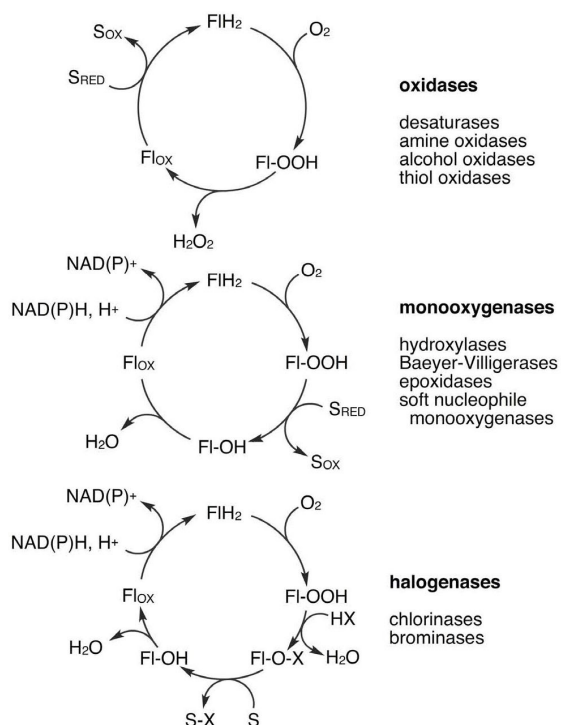


Figure S1. Catalytic cycles of aerobic flavin-dependent enzymes. Oxidative flavin-dependent enzymes typically fall within three classes as given by oxidase, monooxygenase, or halogenase activity.^{S7} Each of these classes operates through a common oxidation/reduction cycle. Structures of the flavin intermediates shown in each cycle are provided in Scheme 1. X = Cl or Br naturally and chemoenzymatically X=F, Cl or Br.

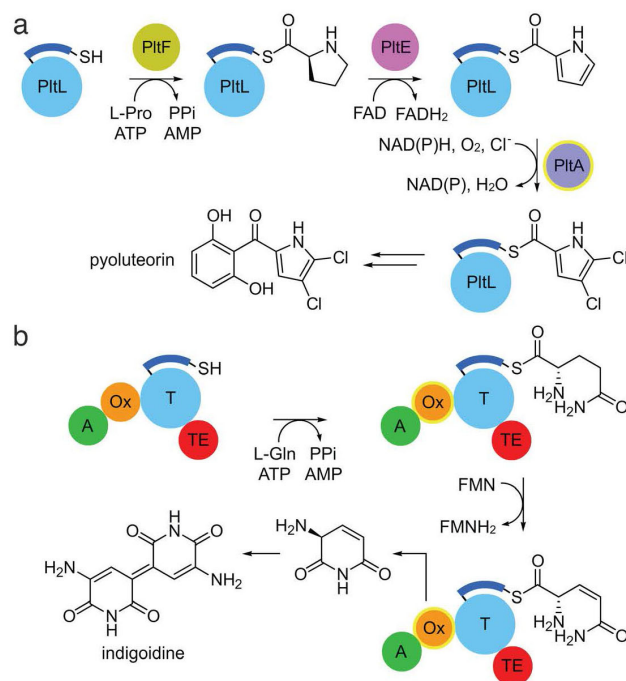


Figure S2. Biosynthetic pathways. **(a)** Biosynthetic pathway of pyoluteorin depicting the role of PitA as a halogenase.^{S4} **(b)** Biosynthetic pathway of indigoidine depicting the multidomain activity of BpsA, including (A) adenylation, (Ox) oxidation, (T) thiolation and (TE) thioester domains.^{S8} The oxidation domain under investigation are highlighted in yellow.

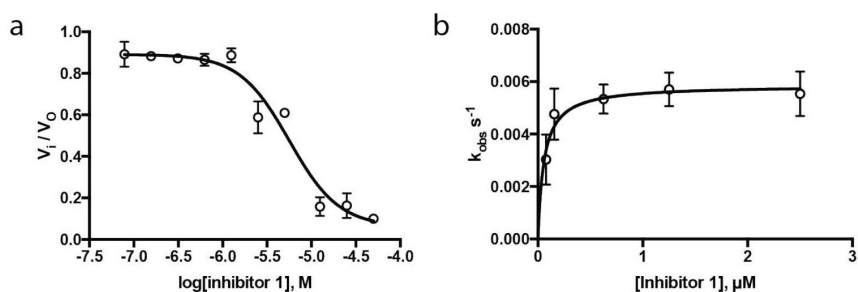


Figure S3. Kinetic analyses of BpsA inhibition. **(a)** Inhibition of BpsA by inhibitor 1 plotted by the logarithm of inhibitor concentration against V_i/V_o , revealing an IC_{50} value of $5.6 \pm 3.0 \mu\text{M}$. **(b)** Plot of k_{obs} by inhibitor 1 concentration, where $K_i = 47 \pm 3 \text{ nM}$ and $k_{\text{inact}} = 5.6 \pm 0.5 \text{ ms}^{-1}$.

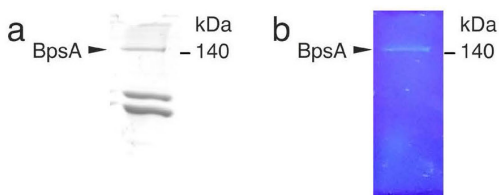


Figure S4. Generation of a blue fluorescent *crypto*-BpsA. **(a)** Coomassie stained SDS-PAGE and **(b)** fluorescent image of the corresponding gel containing fluorescently-tagged *crypto*-BpsA.

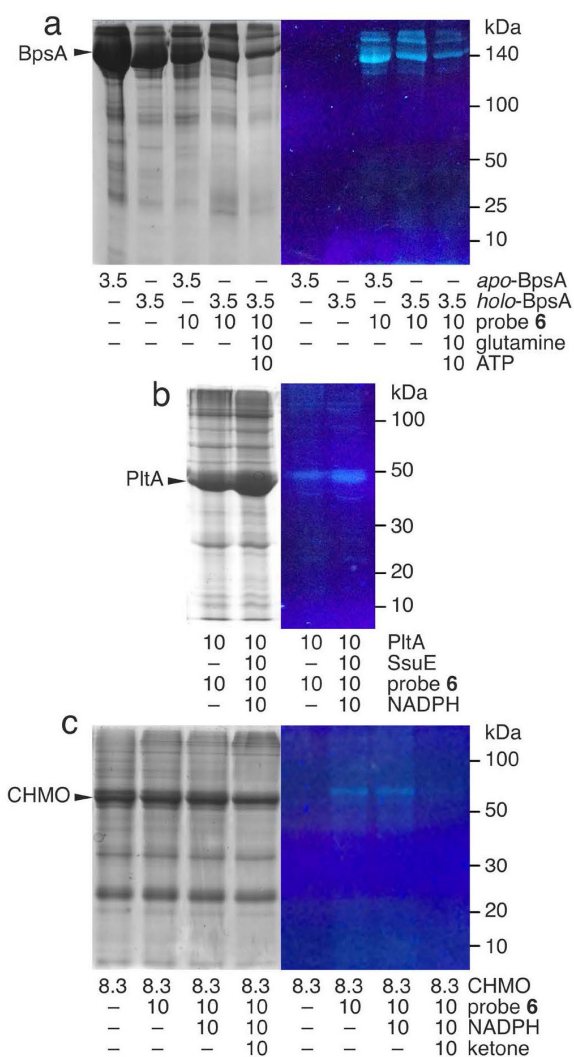


Figure S5. Full gel depiction of gels shown in Figure 2.

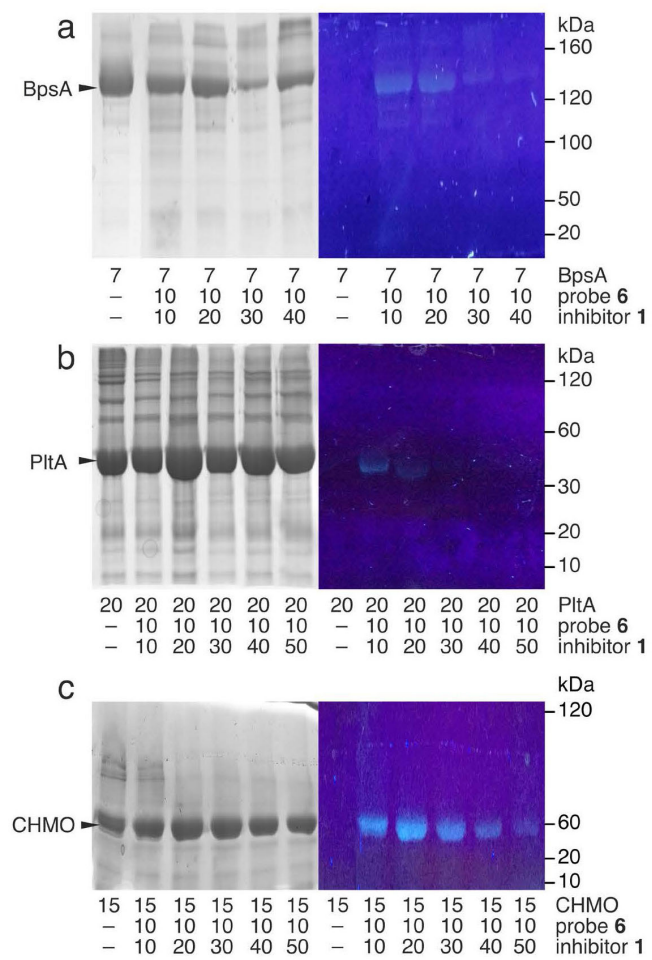


Figure S6. Full gel depiction of gels shown in Figure 3.

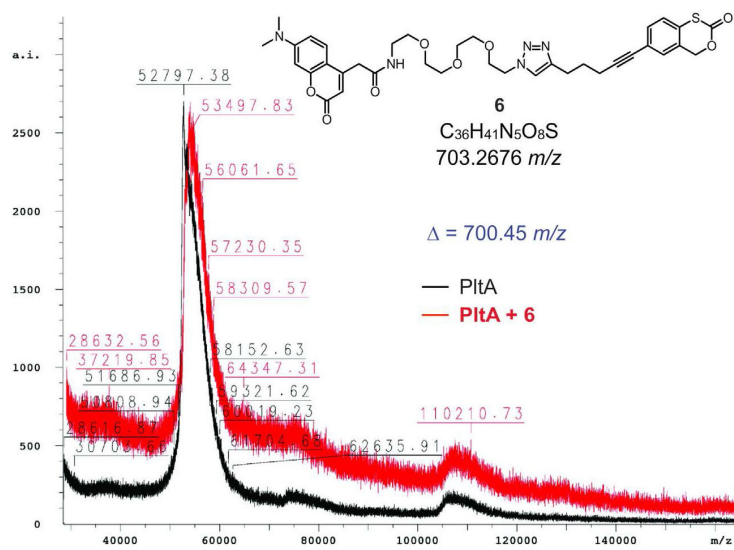


Figure S7. MALDI-TOFMS spectra. Overlaid spectra of PItA (black) and PItA+ 6 (red) display a mass shift of 700.45 m/z, corresponding to covalent addition of probe 6 to PItA.

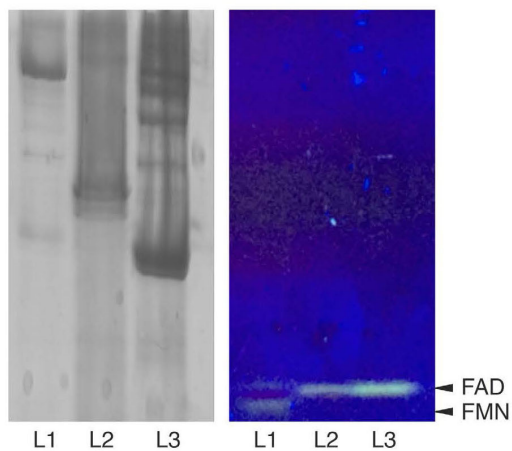


Figure S8. Flavin disassociation by SDS-PAGE. Lanes are given by BpsA (L1), CHMO (L2) and PItA (L3). Flavin mononucleotide (FMN) was detected after denaturation of BpsA. Flavin adenine dinucleotide (FAD) was obtained from both CHMO and PItA. Fluorescent bands were obtained with excitation at λ_{\max} of 365 nm.

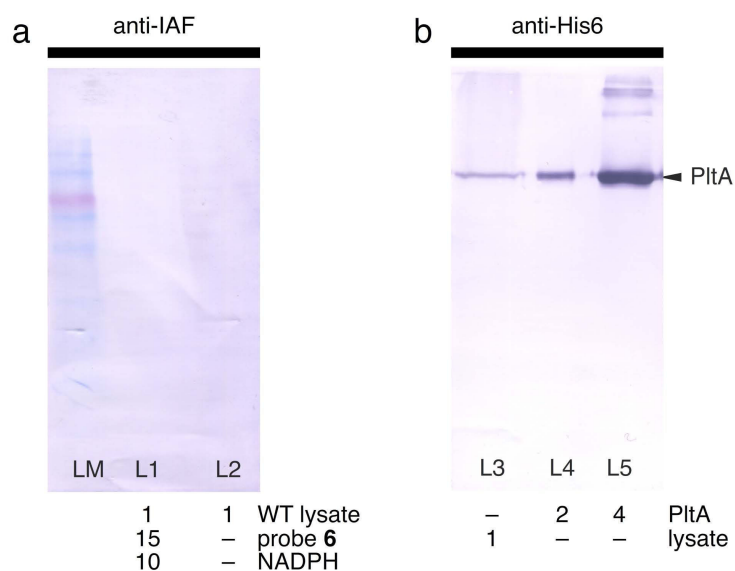


Figure S9. Western blot analyses. **a)** Western blot analysis using the anti-IAF mAb to demonstrate the lack of bands present in wild type (WT) *E. coli* BL21. LM denotes lane marker. L1-L2 are from bands containing 1 mg/mL of *E. coli* lysate that was either treated with 15 μ M probe 6 (L1) and 10 μ M NADPH or untreated (L2). These studies show that the IAF mAb did not have non-specific interactions with proteins in the *E. coli* lysate. Approximately 10 μ g of total protein was loaded per lane. **b)** Western blot analysis using an anti-His6 mAb to show the levels of PltA expression in the *E. coli* expression strain (L3) as compared to standardized concentrations of PltA in μ M. The level of PltA was approximately 1 μ M based on densitometry measurements. These blots were run under the same procedures as used section L (above).

Figure S10. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **7** in CDCl_3

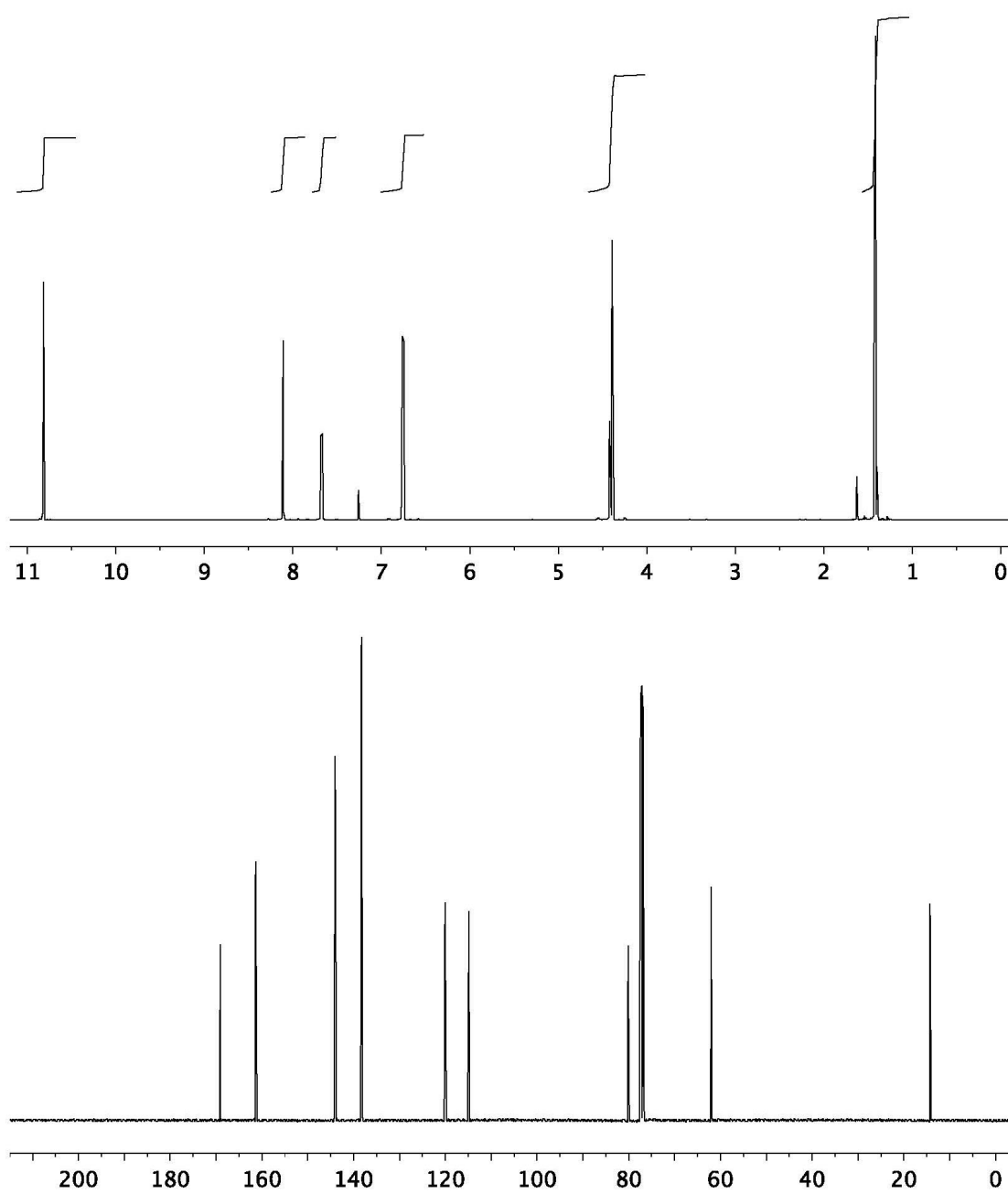


Figure S11. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **8** in CDCl_3

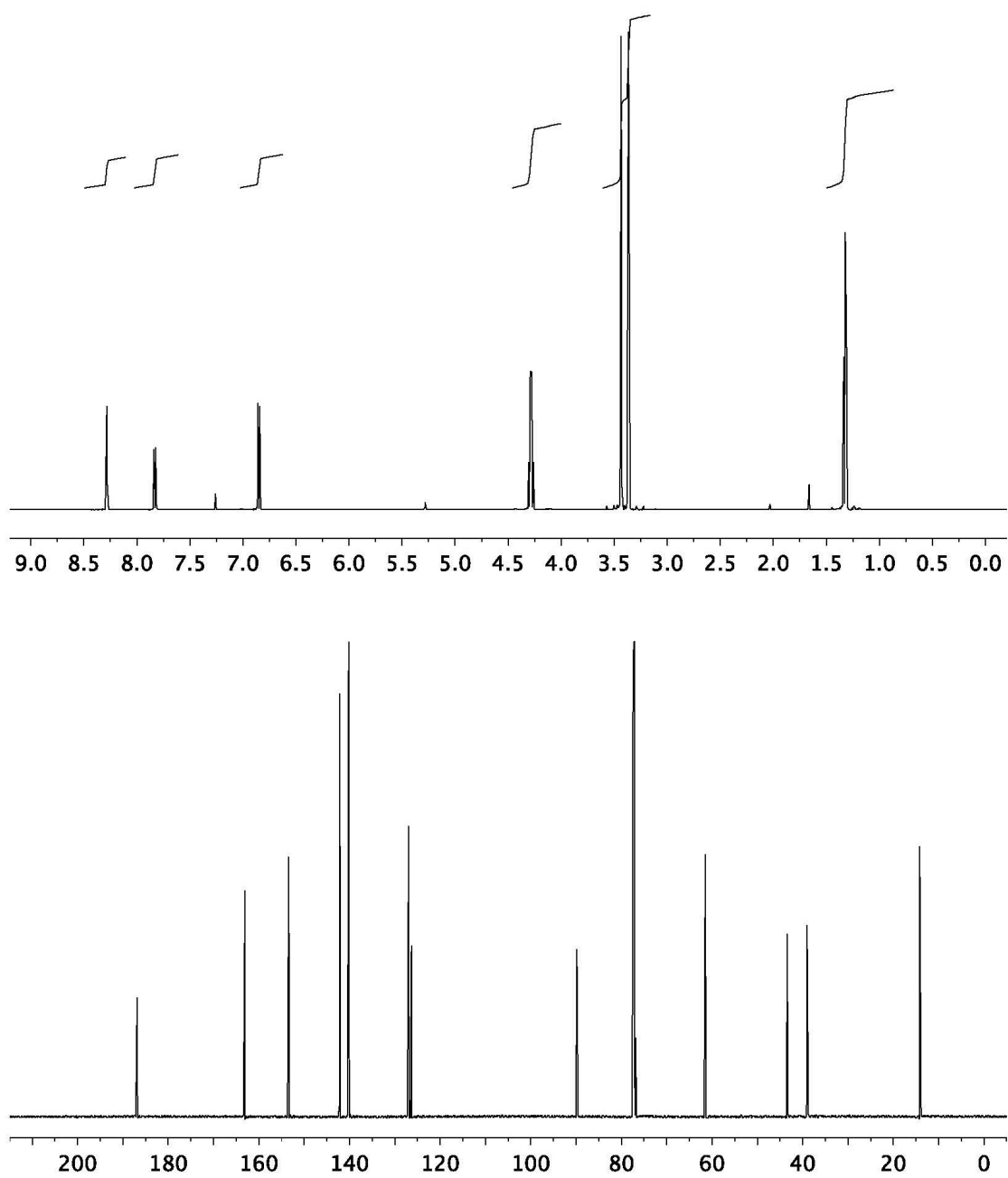


Figure S12. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **9** in CDCl_3

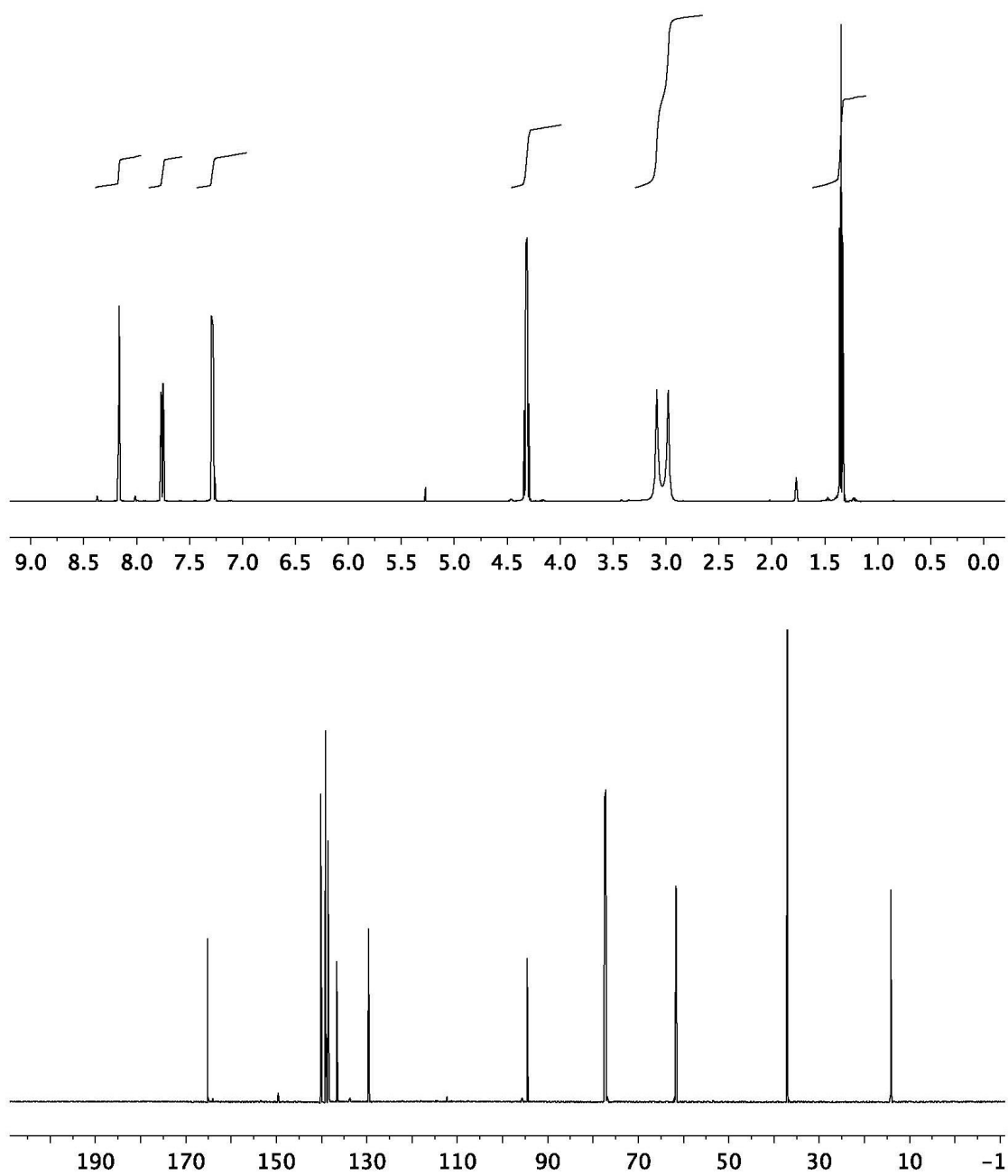


Figure S13. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **10** in CDCl_3

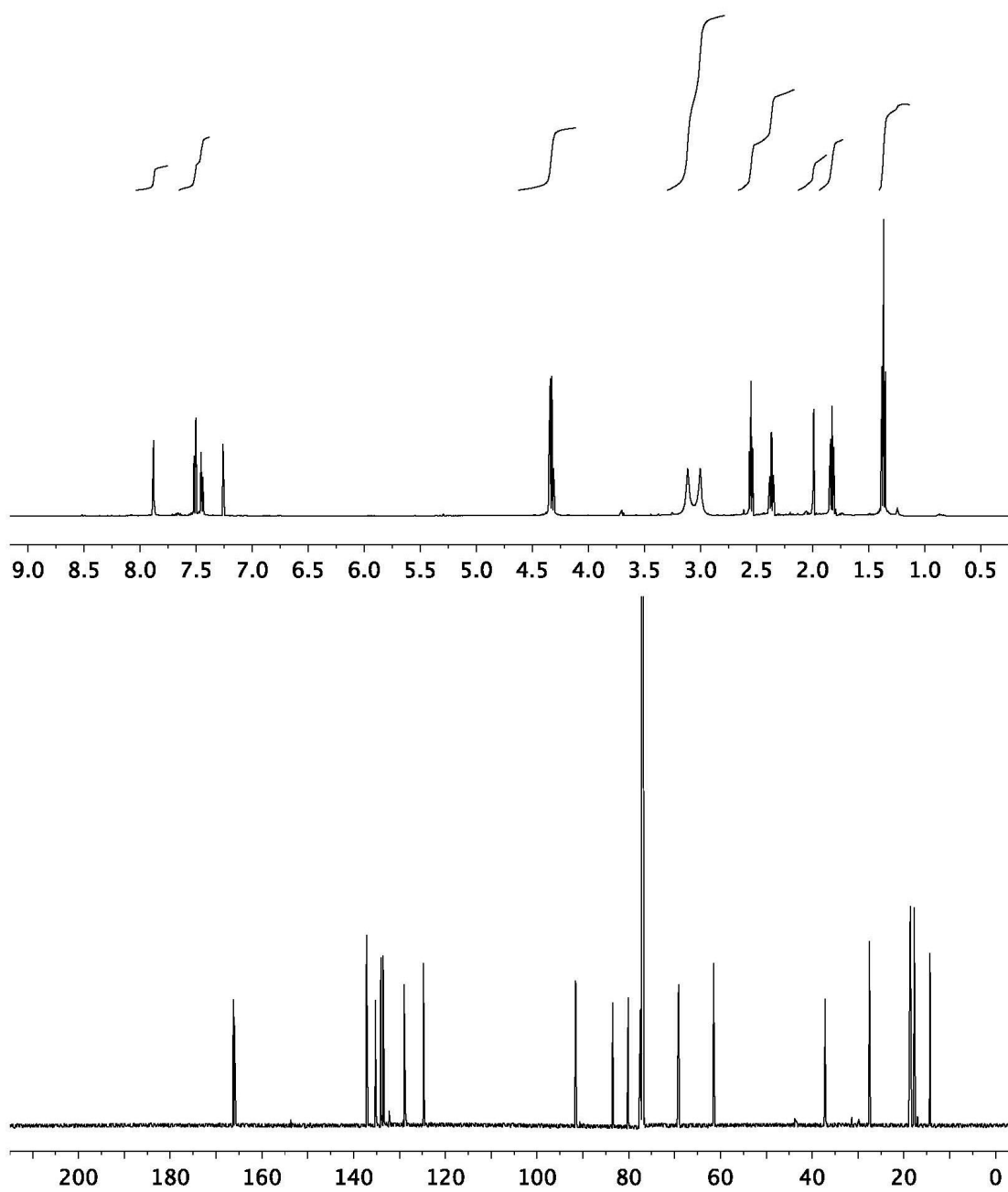


Figure S14. ^1H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra of **12** in CDCl_3

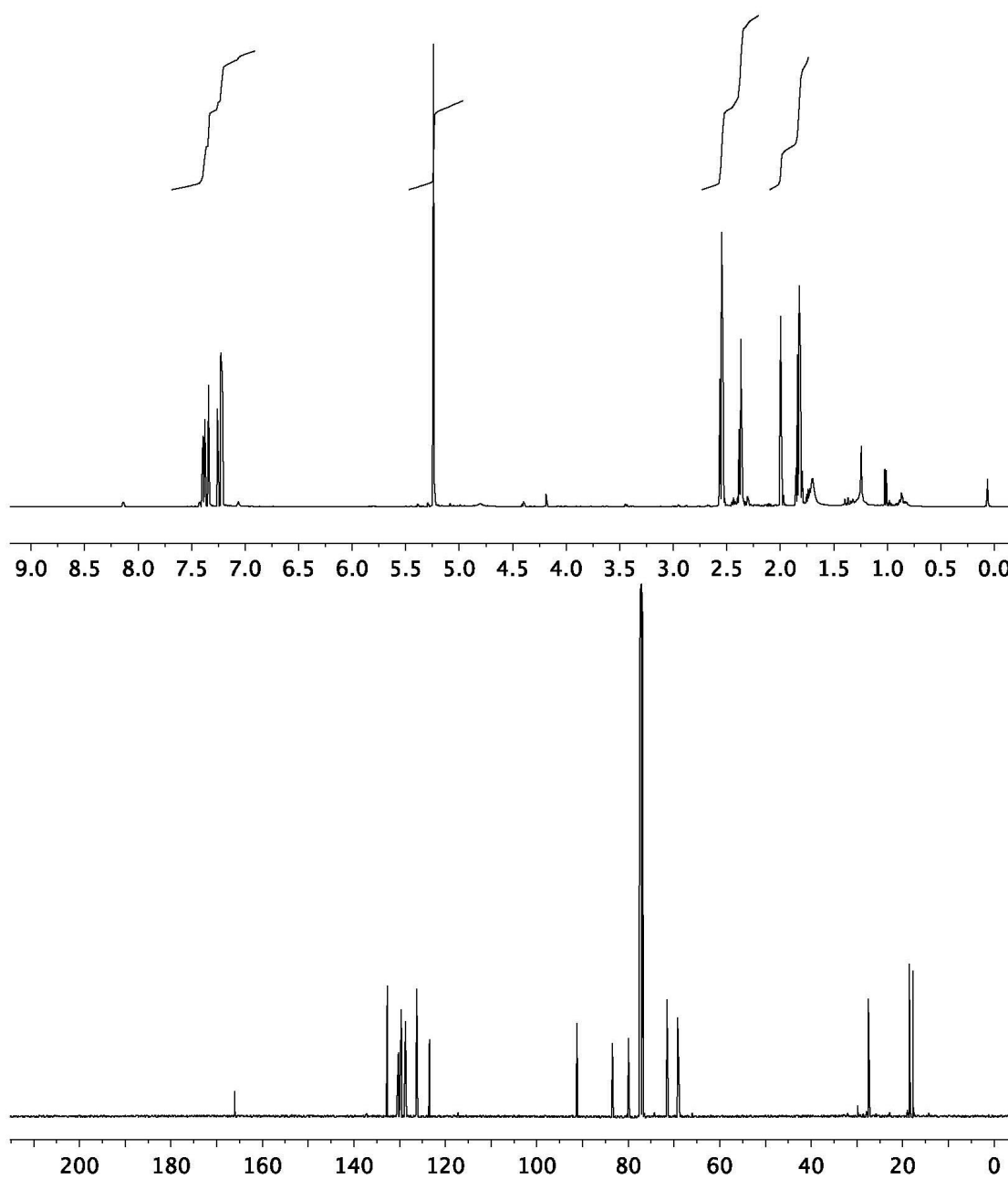


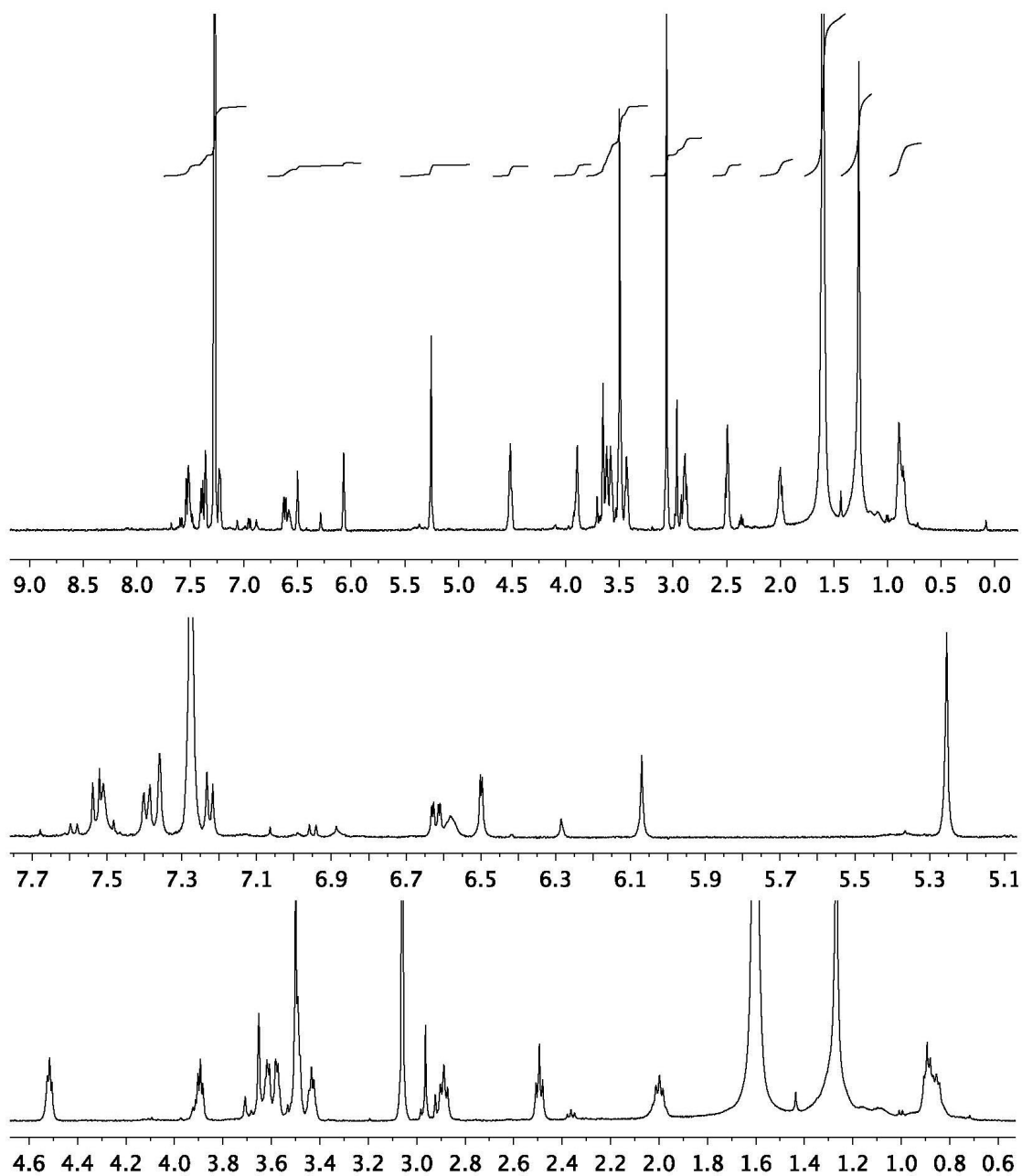
Figure S15. ^1H NMR (500 MHz) spectrum of **6** in CDCl_3 

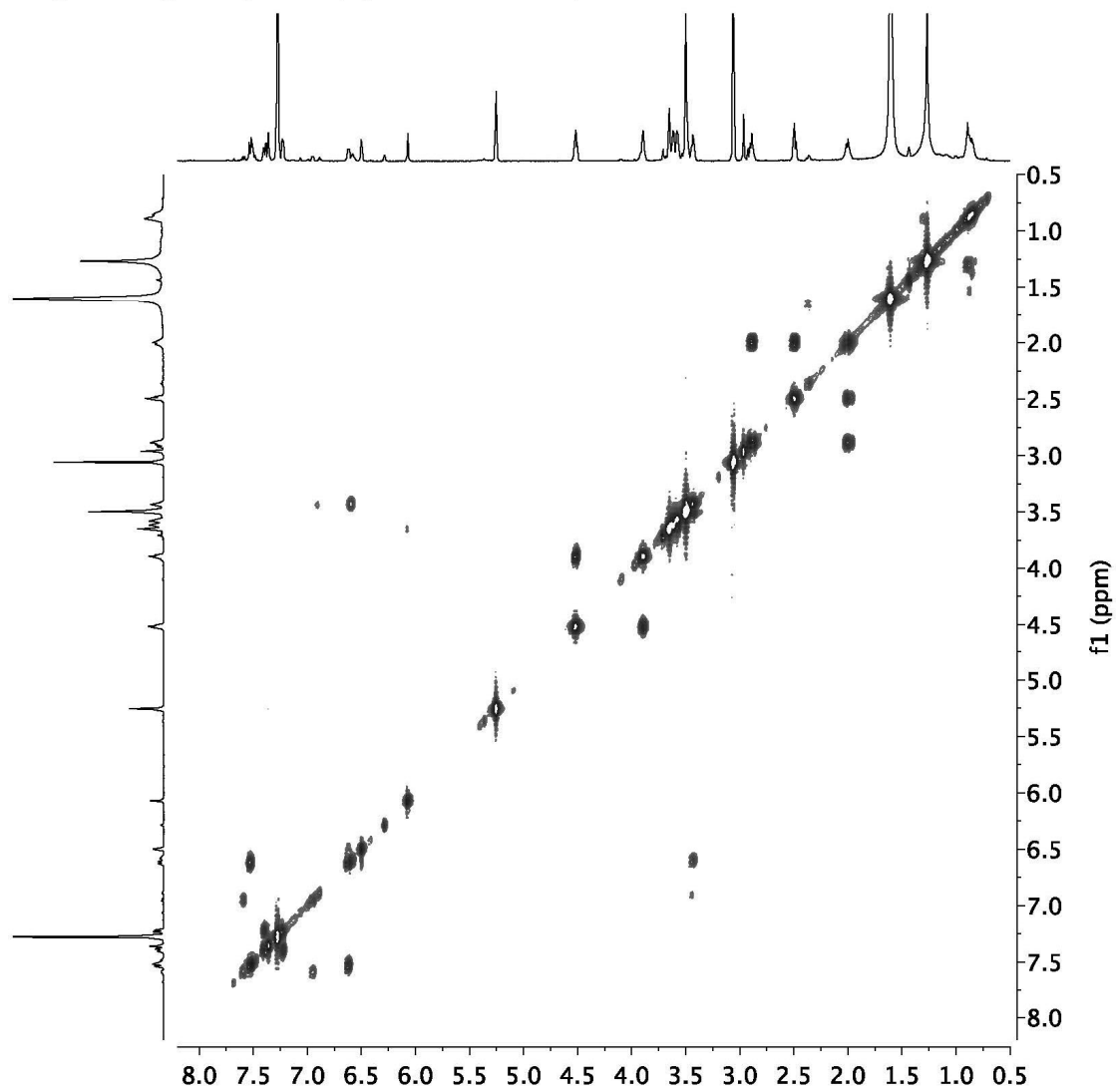
Figure S16. gCOSY (500 MHz) spectrum of **6** in CDCl₃

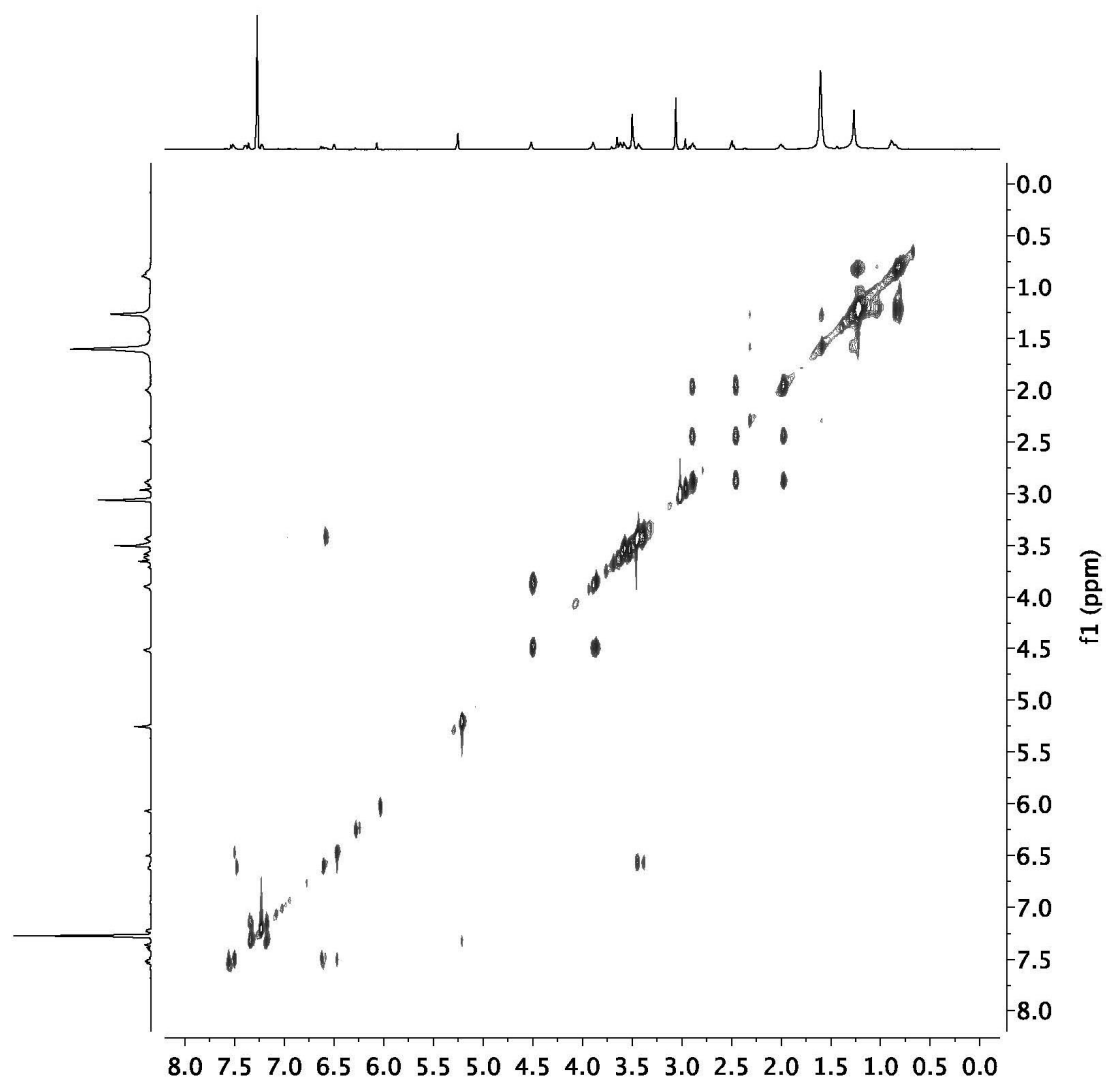
Figure S17. TOCSY (600 MHz) spectrum of **6** in CDCl_3 

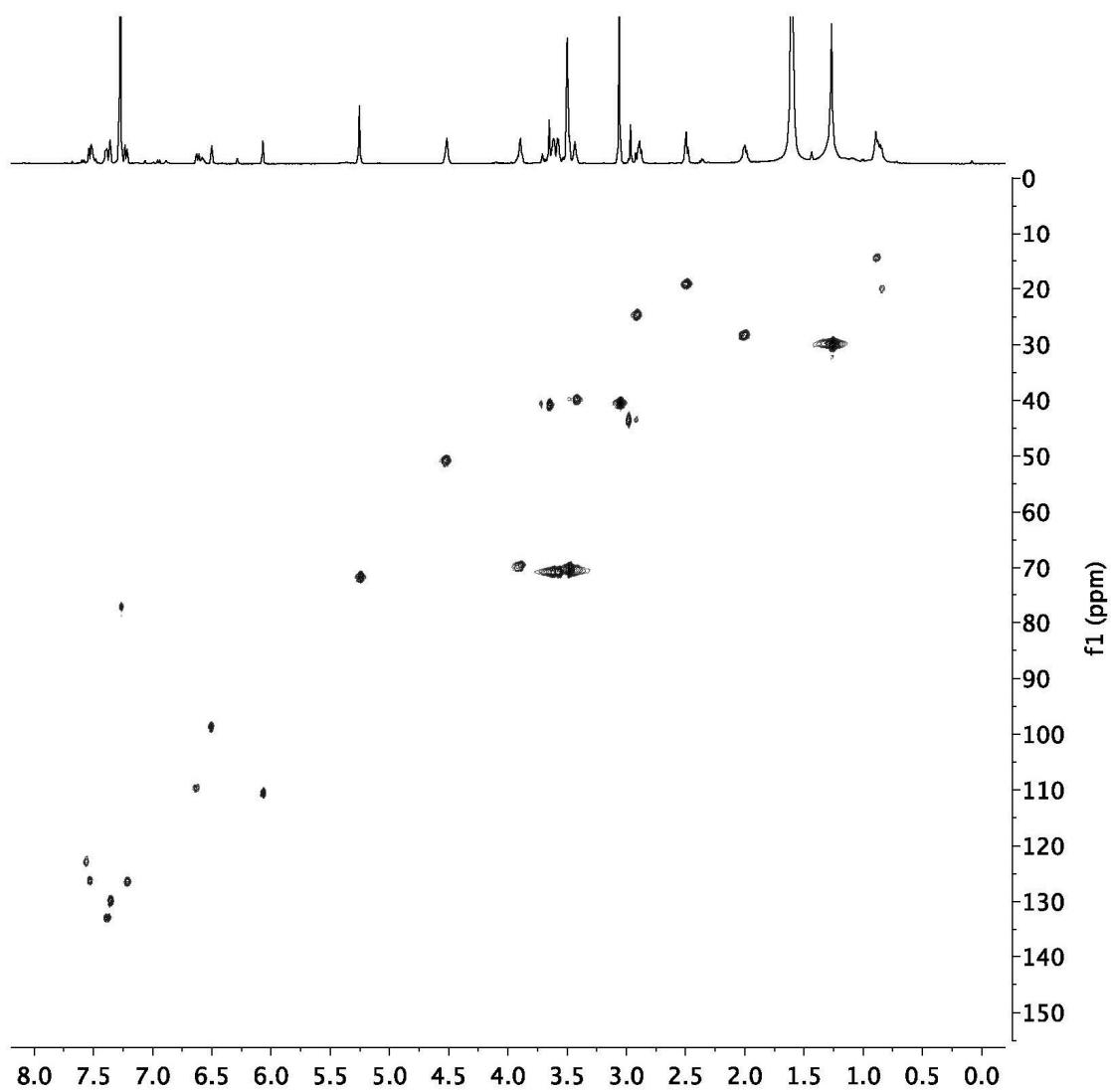
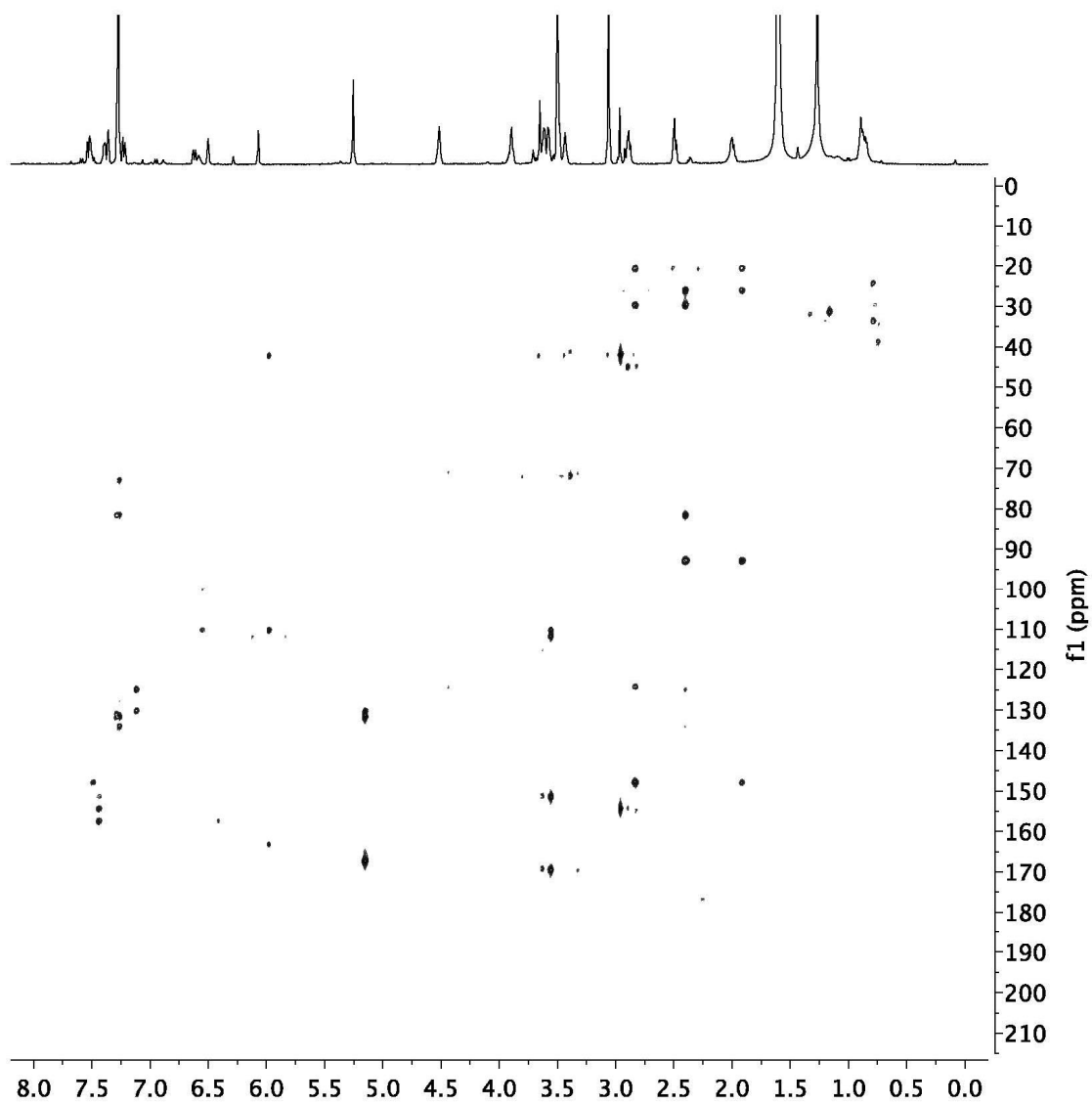
Figure S18. ^1H , ^{13}C HSQC (600 MHz) spectrum of **6** in CDCl_3 

Figure S19. ^1H , ^{13}C HMQC (600 MHz) spectrum of **6** in CDCl_3 

Acknowledgements: Chapter 5, in full, is a reprint of the material as it appears: McCulloch, I. P.; La Clair, J. J.; Jaremko, M. J.; Burkart, M. D. “Fluorescent Mechanism-Based Probe for Aerobic Flavin-Dependent Enzyme Activity.” *ChemBiochem* 17 (2015): 1598-601. The dissertation author is a contributing author of this manuscript.