Communication

Gatekeeping Ketosynthases Dictate Initiation of Assembly Line Biosynthesis of Pyrrolic Polyketides

Dongqi Yi, Atanu Acharya, James C. Gumbart, Will R. Gutekunst, and Vinayak Agarwal*

Cite This: J. Am. Chem. Soc. 2021, 143, 7617−7622

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Assembly line biosynthesis of polyketide natural products involves checkpoints where identities of thiotemplated intermediates are verified before polyketide extension reactions are allowed to proceed. Determining what these checkpoints are and how they operate is critical for reprogramming polyketide assembly lines. Here we demonstrate that ketosynthase (KS) domains can perform this gatekeeping role. By comparing the substrate specificities for polyketide synthases that extend pyrrolyl and halogenated pyrrolyl substrates, we find that KS domains that need to differentiate between these two substrates exercise high selectivity. We additionally find that amino acid residues in the KS active site facilitate this selectivity and that these residues are amenable to rational engineering. On the other hand, KS domains that do not need to make selectivity decisions in their native physiological context are substrate-promiscuous. We also provide evidence that delivery of substrates to polyketide synthases by non-native carrier proteins is accompanied by reduced biosynthetic efficiency.

Type I polyketide synthases (PKSs) are multimodular enzymes that construct polyketide natural products.[1] The assembly line biosynthesis of polyketides involves repetitive steps: choice of an extender unit by the acyltransferase (AT) domain, decarboxylative Claisen condensation of this extender unit to the polyketide by the ketosynthase (KS) domain, and reductive tailoring of the β-carbonyl by ketoreductase, dehydratase, and enoyl reductase domains. Substrates for these PKS domains are thioesterified to carrier proteins (CPs). The ATs and the reductive tailoring domains were traditionally thought to determine the diversity of polyketide natural products.[2] The role of KSs in determining the polyketide diversity is relatively less well studied. Intermediary KSs in collinear PKS assembly lines are evolutionarily linked to the CP and the tailoring domains that precede them,[1,3,4] and as polyketide extension progresses, the selectivity of these intermediary KSs can constrain the extension of noncognate substrates.[5−8] Here we demonstrate that the strict selectivity of the very first KS in a collinear PKS assembly line can perform a gatekeeping function by precluding the initiation of polyketide extension of noncognate substrates and that the selectivity and promiscuity of initiating KSs are modulated in tune with the physiological natural product biosynthetic scheme.

The biosynthesis of several pyrrole-containing natural products involves polyketide extensions. Thioesterified pyrroles are delivered by a type II nonribosomal peptide synthetase (NRPS) adenylation−oxidation cascade that oxidizes the L-proline pyrrolidine to pyrrole (Figure 1A).[9−11] Thiotemplated pyrroles can be modified, as exemplified by dichlorination in the biosynthesis of pyoluteorin (1) (Figure 1B).[12] The thiotemplated dichloropyrrole is then handed off, without intermediary offloading, from the type II NRPS donor CP PltL to the type I PKS PltB. This handoff occurs by transthioesterification of the molecular cargo from the donor CP to the active-site cysteine thiol of the KS. Malonylation of

the PltB CP by the AT enables the KS to catalyze the decarboxylative chain extension. Additional polyketide extensions followed by aromatization afford 1 (Figure S1). The precursor to 4,5-dichloropyrrolyl-S-PltL, pyrrolyl-S-PltL,[12] is also available in situ to the PltB PKS module (Figure 1B). However, deschloro-1 is not produced (Figure S2). This is in contrast to the biosynthetic scheme for calcimycin (2) (Figure 1C), in which pyrrole (pyrrolyl-S-CalN3) acts as the physiological substrate to initiate polyketide extension by CalA.[13]

Why is pyrrolyl-S-PltL rejected as a substrate by PltB while pyrrolyl-S-CalN3 is accepted by CalA? To query the specificity determinants for PltB and CalA activities, we adopted an approach comparing the intermolecular recognition events that dictate polyketide extension of pyrrolic molecular cargoes. First, we determined the KS−AT−CP module boundaries for the bimodular PltB and CalA proteins by alignment with other KS−AT and CP crystal structures.[14−16] The tridomain sequences were expressed in Escherichia coli, and the PltB and CalA KS−AT−CP modules were purified (Figure S3). The CPs PltL and CalN3 were purified in their apo forms.

A library of pyrrolyl-S-pantetheines were synthesized (Scheme 1). The previously described synthesis of S-acyl pantetheines involved thioesterification of the acyl groups to give acetonide-protected pantetheine and subsequent acid deprotection to unmask the 1,3-diol.[17,18] However, we found the halogenated pyrrole derivatives to be labile toward acidic
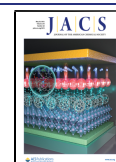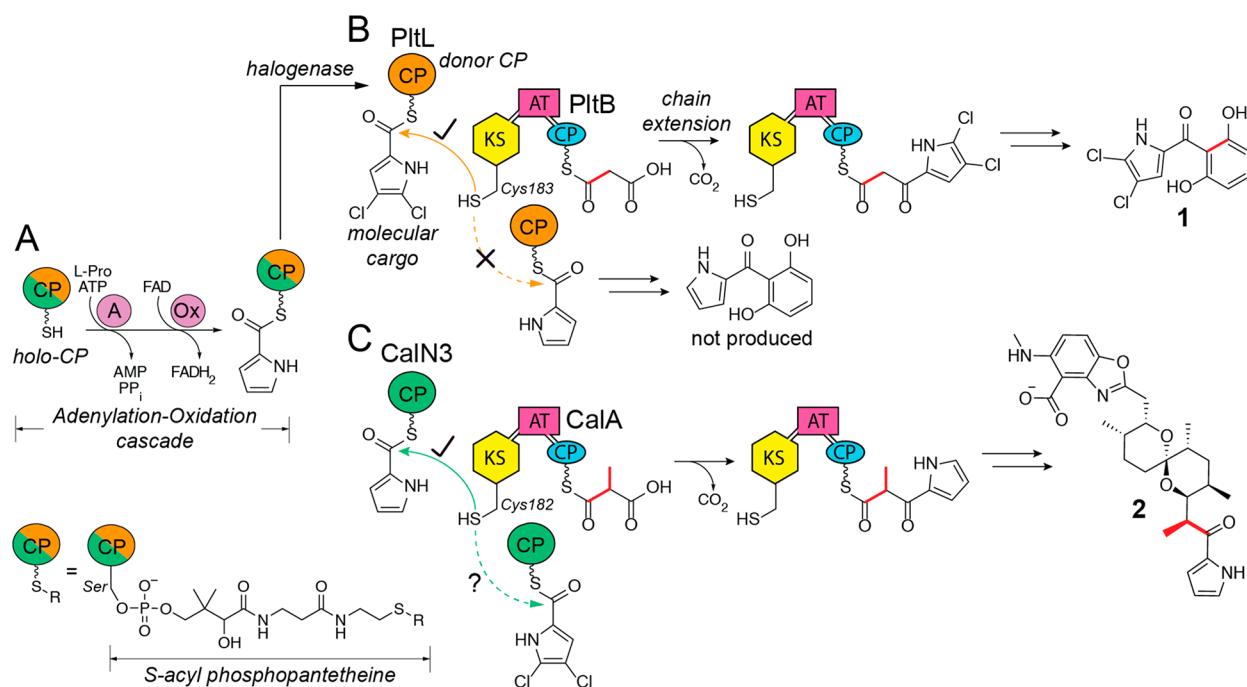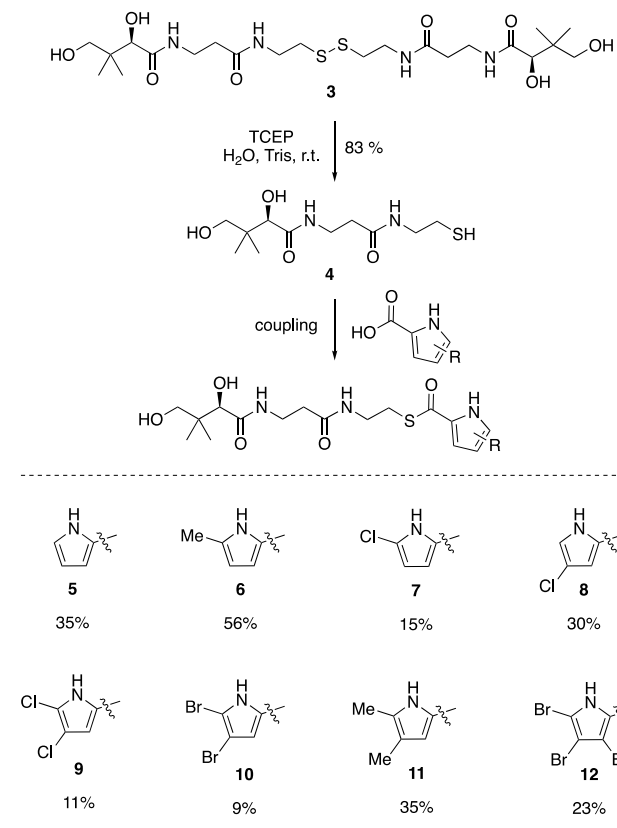
Figure 1. Biosynthesis of pyrrolic polyketides. (A) A type II NRPS furnishes the pyrrolyl-*S*-CPs. (B) Halogenation of pyrrolyl-*S*-PltL and polyketide elongation by PKS PltB (Cys183 is the KS active-site residue). (C) Polyketide extension by CalA.

**Scheme 1. Synthesis of *S*-Acyl Pantetheines 5−12**

deprotection. Instead, D-pantetheine (**4**) was accessed by disulfide reduction of D-pantethine (**3**) and used directly for ligation to pyrrolyl-2-carboxylic acids to furnish **5−11** (Figures S4−S17). Molecule **12** was accessed by tribromination of the pyrrolyl methyl ester followed by conversion to an acyl chloride, which was thioesterified to **4**.[19] Compounds **5−12**

were extended to the corresponding *S*-acyl CoAs, followed by enzymatic transfer of the *S*-acyl phosphopantetheines to the donor CPs (Figures S18−S34).[17,18,20]

With the substrates in hand, we compared the abilities of the PltB and CalA KS−AT−CP modules to extend the pyrrolic cargoes. Malonyl-CoA required for the malonylation of the PltB CP by the AT was exogenously provided. For CalA, methylmalonyl-CoA was generated *in situ* by the CoA ligase MatB (Figure S35).[21] PltB Cys183Ala and CalA Cys182Ala were used to establish negative controls.[22] Upon incubation of PltB with a 2-fold molar excess of dichloropyrrolyl-*S*-PltL and excess malonyl-CoA, we detected the formation of a diketide product; this diketide was not observed with the PltB Cys183Ala enzyme (Figure S36). Dichloropyrrolyl-*S*-PltL was completely consumed; no consumption of the substrate was observed for the PltB Cys183Ala enzyme (Figure S37).

The abundance of the diketide detected was not reproducible. Thus, we relied on quantifying the abundance of the leftover *S*-acyl donor CP as a proxy for the extent of the reaction. Compared with complete consumption of dichloropyrrolyl-*S*-PltL, the consumption of pyrrolyl-*S*-PltL was negligible (Figure 2A). Physiologically, this implies that despite its presence in the molecular milieu within the producer of **1**, pyrrolyl-*S*-PltL is not accepted as a substrate by the PltB KS, thus precluding the production of deschloro-**1**.

We tested the activity of PltB for other pyrrolyl derivatives. Both 4- and 5-chloropyrrolyl-*S*-PltL were accepted as substrates, but with a reduced preference relative to 4,5-dichloropyrrole (Figure 2A). Bromine could replace chlorine; 4,5-dibromopyrrolyl-*S*-PltL was accepted well. However, the activity was diminished for the 3,4,5-tribromo derivative. Methyls could not replace halides. Methylpyrroles are observed in aminocoumarin natural products,[23] where 5-methylpyrrole does not undergo polyketide extension but is instead esterified to afford glycosyl moieties. We observed a marked decrease in preference for 5-methylpyrrole. An even further reduction in
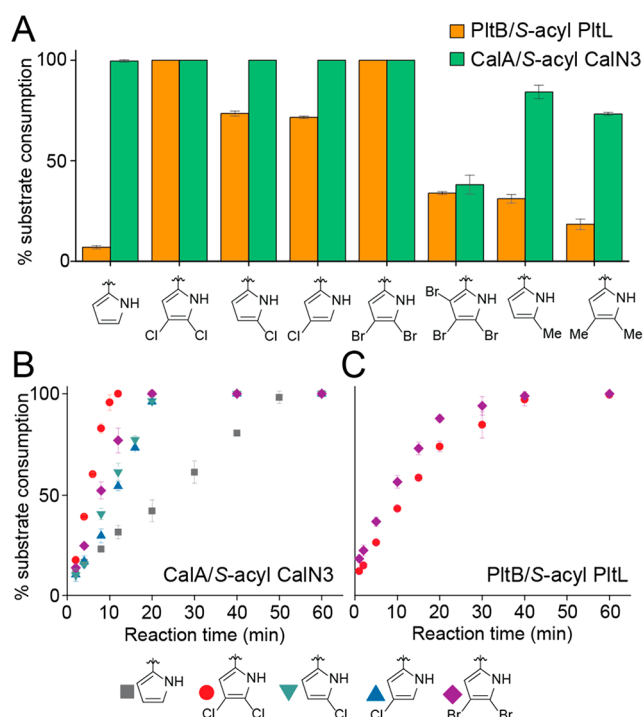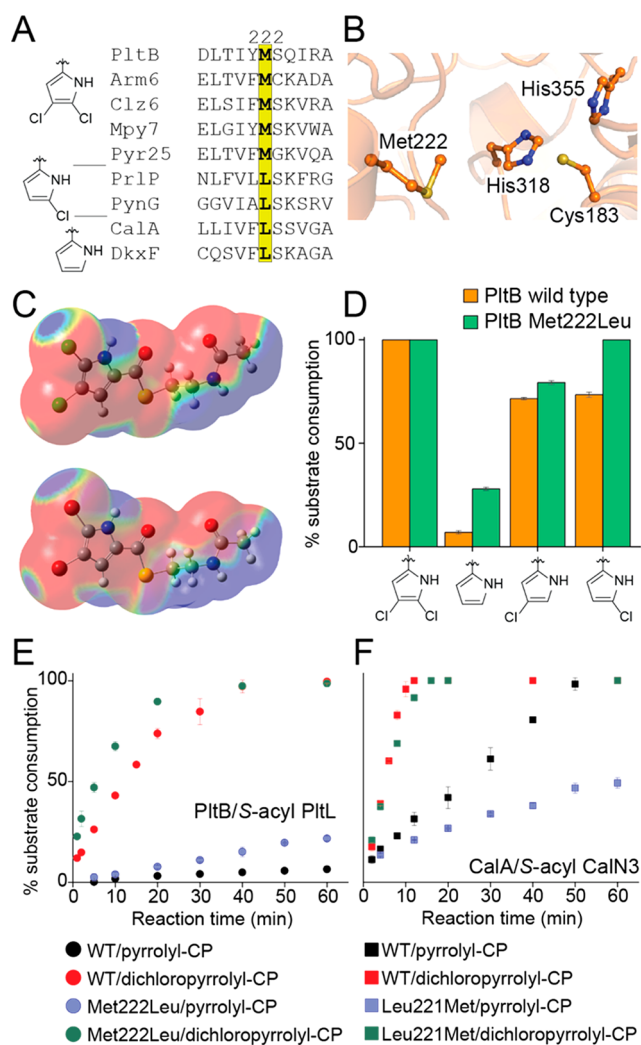
Figure 2. KS specificities. (A) Depletion of the end-point substrate *S*-acyl CP by PltB and CalA. (B, C) Time-dependent depletion of all substrates that demonstrated full consumption in the end-point assay by (B) CalA and (C) PltB.

Figure 3. KS active site. (A) KS sequence alignment. Sequences for (from top to bottom) **1**, armeniaspirol,[25] chlorizidine,[26] marinopyrrole,[27] pyrrolomycin,[28] pyralomycin,[29] pyrronazol B,[30] **2**, and DKxanthene[31] KSs are illustrated. (B) The PltB KS active site. The Cys183-His318-His355 catalytic triad[22] side chains are illustrated. (C) ESPs mapped onto the electron densities (isovalue = 0.0004) for (top) dichloropyrrolyl- and (bottom) dibromopyrrolyl-*S*-*N*-acetylcysteamines. Regions of negative potential and positive potential (>0.01 au) are displayed in red and blue, respectively. ESPs between 0.00 and 0.01 are displayed with a color gradient from red to blue. (D) Substrate specificity of the PltB Met222Leu mutant compared with the wild type. (E, F) Time-dependent consumption of pyrrolyl- and dichloropyrrolyl substrates by wild-type and mutant PltB (Met222-Leu) and CalA (Leu221Met).

preference was observed for 4,5-dimethylpyrrolyl-*S*-PltL (Figure 2A). Overall, we posit that substitutions on the pyrrole ring have a strong influence on the PltB KS activity.

Compared with PltB, the CalA KS demonstrated relaxed substrate selectivity. Complete consumption of pyrrolyl- and dichloropyrrolyl-*S*-CalN3 substrates along with other halogenated derivatives was observed (Figure 2A). As before, a reduction in substrate consumption was observed when the 3-position was brominated. Methyl substituents were less preferred than the halides; however, the distinction was less pronounced compared with PltB (Figure 2A). Querying time-dependent substrate consumption for all of the substrates that were fully depleted by CalA revealed that the rate of consumption of the physiological substrate pyrrole-*S*-CalN3 was lower than that of dichloropyrrolyl-*S*-CalN3 (Figure 2B). The faster consumption of dichloropyrrolyl-*S*-CalN3 was not due to substrate degradation (Figure S37). These data allow us to posit that the PltB KS exercises substrate selectivity, as it needs to differentiate between thiotemplated cargoes within the biosynthetic milieu. On the other hand, the CalA KS, which does not encounter differentially modified pyrrolic substrates, has relaxed substrate selectivity. Without an evolutionary pressure to maintain fidelity, relaxation in substrate selectivity is observed for primary and secondary metabolic enzymes.[24] Dichloro- and dibromopyrrole substrates were consumed by CalA and PltB at similar rates (Figure 2B,C).

Sequence alignment identified a methionine residue that was conserved for all KSs that extend dichloropyrroles (PltB Met222; Figure 3A). Homology structure of the PltB KS-AT didomain places the Met222 side chain within the KS active site (Figure 3B). The methionine thioether is a validated halogen-bonding partner.[32,33] Electrostatic potential (ESP)

maps for dichloro- and dibromopyrrole-*N*-acetylcysteamines computed at the wb97X-D/aug-cc-pVTZ level of theory demonstrated the presence of complementary halogen-bonding σ holes on the halogen substituents on the pyrrole rings (Figures 3C and S38). To probe the role of Met222 in determining the KS specificity, the substrate selectivity of the PltB Met222Leu enzyme was queried. The PltB Met222Leu mutant demonstrated a 4-fold increase in ability to deplete the pyrrolyl-*S*-PltL substrate while maintaining activity for dichloropyrrolyl-*S*-PltL (Figure 3D,E). Consumption of 4- and 5-chloropyrrolyl-*S*-PltL was also enhanced, mirroring the activity of CalA. Introducing the complementary mutation Leu221Met in CalA led to a reduction in depletion of the

pyrrolyl-*S*-CalN3 substrate while maintaining reactivity for the dichlorinated substrate (Figure 3F). Thus, the Met222/Leu221 side chains participate in the selection for the correct derivatization state of the pyrrolyl molecular cargoes in KS active sites. Active-site residues also tailor the substrate specificities of intermediary KS domains in collinear PKS pathways.[5,6] Overall, our data demonstrate that the strict substrate selectivity of the very first KS active site in collinear PKS assembly lines can perform a gatekeeping role to determine whether polyketide extension can be initiated.

Intermolecular protein−protein interactions underlie polyketide assembly.[34] CPs sequester their molecular cargoes; the structure of pyrrolyl-*S*-PltL demonstrates that the phosphopantetheine arm folds in such a way that the pyrrole ring binds at a hydrophobic patch on the PltL surface.[35] Hence, the exclusion of pyrrolyl-*S*-PltL by PltB but acceptance of pyrrolyl-*S*-CalN3 by CalA could depend on how PltL and CalN3 differentially present the pyrrolyl substrates to the KS. To query this hypothesis, we tested pyrrolyl- and dichloropyrrolyl-*S*-CalN3 as substrates for PltB. Compared with dichloropyrrolyl-*S*-PltL, the preference for dichloropyrrolyl-*S*-CalN3 was reduced. However, pyrrolyl-*S*-CalN3, just like pyrrolyl-*S*-PltL, was not depleted by PltB at all (Figure 4A). This result
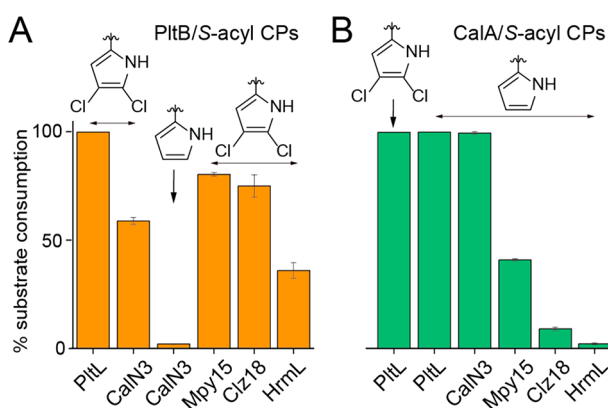
Figure 4. KS activity for noncognate donor CPs. (A) Relative depletion of dichloropyrrolyl-*S*-CPs by PltB. Depletion of pyrrolyl-*S*-CalN3, the cognate substrate for CalA, is marked by the arrow. (B) Relative depletion of pyrrolyl-*S*-CPs by CalA. Depletion of dichloropyrrolyl-*S*-PltL, the cognate substrate for PltB, is marked by the arrow.

demonstrates that while a penalty is indeed paid when the donor CP for PltB is changed from PltL to CalN3, the principal determinant for catalysis is the recognition of the molecular cargo by the KS. CalA depleted both pyrrolyl-*S*-PltL and dichloropyrrolyl-*S*-PltL (Figure 4B). The extents and rates of depletion of pyrrolyl-*S*-CalN3 and pyrrolyl-*S*-PltL by CalA were similar (Figures 4B and S39).

We also tested the ability of the type II NRPS CPs Mpy15, Clz18, and HrmL to pair with PltB and CalA. Mpy15 and Clz18 deliver thiotemplated 4,5-dichloropyrrole to PKS modules in the respective production of marinopyrroles[27] and chlorozidine,[26] and HrmL delivers 5-chloropyrrole to an NRPS module for the production of hormaomycin.[36] The activities of PltB and CalA were tested for their cognate thiotemplated substrates acylated to these different donor CPs (Figure S40−45). We observed that PltB was more accepting of different donor CPs than CalA. A marked decrease in acceptance for Mpy15 and Clz18 was observed for CalA

compared with PltB (Figure 3A,B). Both PltB and CalA demonstrated the least preference for HrmL. Phylogenetically, HrmL is more distant to PltL and CalN3 than any of the other type II NRPS CPs tested (Figure S46).

Here we have demonstrated a gatekeeper role for KSs in initiating polyketide extension. Combinatorial expansion of the chemical space explored by pyrrolyl natural products will require an understanding of the intermolecular interactions that occur between the molecular cargo, the donor CP, and the initiating PKS modules. While structural models have guided modulation of KS substrate selectivities,[5,6,37,38] this study provides evidence that certain KSs, such as CalA, are already substrate-promiscuous and should be prioritized for further exploration for diversity-generating efforts.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c02371.

> Materials and methods for protein expression and purification, chemical synthesis and spectroscopic characterization data, assay and data analysis procedures, and computational details (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Vinayak Agarwal − *School of Chemistry and Biochemistry and School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* Ⓞ orcid.org/0000-0002-2517-589X; Email: vagarwal@gatech.edu

Authors

Dongqi Yi − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*

Atanu Acharya − *School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* Ⓞ orcid.org/0000-0002-6960-7789

James C. Gumbart − *School of Chemistry and Biochemistry and School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* Ⓞ orcid.org/0000-0002-1510-7842

Will R. Gutekunst − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/jacs.1c02371

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Nivina, A.; Yuet, K. P.; Hsu, J.; Khosla, C. Evolution and diversity of assembly-line polyketide synthases. *Chem. Rev.* **2019**, *119* (24), 12524−12547.

(2) Klaus, M.; Grininger, M. Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* **2018**, *35* (10), 1070−1081.

(3) Zhang, L.; Hashimoto, T.; Qin, B.; Hashimoto, J.; Kozone, I.; Kawahara, T.; Okada, M.; Awakawa, T.; Ito, T.; Asakawa, Y.; Ueki, M.; Takahashi, S.; Osada, H.; Wakimoto, T.; Ikeda, H.; Shin-ya, K.; Abe, I. Characterization of giant modular PKSs provides insight into genetic mechanism for structural diversification of aminopolyol polyketides. *Angew. Chem., Int. Ed.* **2017**, *56* (7), 1740−1745.

(4) Keatinge-Clay, A. T. Polyketide synthase modules redefined. *Angew. Chem., Int. Ed.* **2017**, *56* (17), 4658−4660.

(5) Murphy, A. C.; Hong, H.; Vance, S.; Broadhurst, R. W.; Leadlay, P. F. Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. *Chem. Commun.* **2016**, *52* (54), 8373−8376.

(6) Klaus, M.; Buyachuihan, L.; Grininger, M. Ketosynthase domain constrains the design of polyketide synthases. *ACS Chem. Biol.* **2020**, *15* (9), 2422−2432.

(7) Wu, J.; Kinoshita, K.; Khosla, C.; Cane, D. E. Biochemical analysis of the substrate specificity of the β-ketoacyl-acyl carrier protein synthase domain of module 2 of the erythromycin polyketide synthase. *Biochemistry* **2004**, *43* (51), 16301−16310.

(8) Busch, B.; Ueberschaar, N.; Behnken, S.; Sugimoto, Y.; Werneburg, M.; Traitcheva, N.; He, J.; Hertweck, C. Multifactorial control of iteration events in a modular polyketide assembly line. *Angew. Chem., Int. Ed.* **2013**, *52* (20), 5285−5289.

(9) Jaremko, M. J.; Davis, T. D.; Corpuz, J. C.; Burkart, M. D. Type II non-ribosomal peptide synthetase proteins: structure, mechanism, and protein−protein interactions. *Nat. Prod. Rep.* **2020**, *37* (3), 355−379.

(10) Corpuz, J. C.; Podust, L. M.; Davis, T. D.; Jaremko, M. J.; Burkart, M. D. Dynamic visualization of type II peptidyl carrier protein recognition in pyoluteorin biosynthesis. *RSC Chem. Biol.* **2020**, *1* (1), 8−12.

(11) Thapa, H. R.; Robbins, J. M.; Moore, B. S.; Agarwal, V. Insights into thiotemplated pyrrole biosynthesis gained from the crystal structure of flavin-dependent oxidase in complex with carrier protein. *Biochemistry* **2019**, *58* (7), 918−929.

(12) Dorrestein, P. C.; Yeh, E.; Garneau-Tsodikova, S.; Kelleher, N. L.; Walsh, C. T. Dichlorination of a pyrrolyl-S-carrier protein by FADH2-dependent halogenase PltA during pyoluteorin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (39), 13843−13848.

(13) Wu, Q.; Liang, J.; Lin, S.; Zhou, X.; Bai, L.; Deng, Z.; Wang, Z. Characterization of the biosynthesis gene cluster for the pyrrole polyether antibiotic calcimycin (A23187) in *Streptomyces chartreusis* NRRL 3882. *Antimicrob. Agents Chemother.* **2011**, *55* (3), 974−982.

(14) Tang, Y.; Kim, C.-Y.; Mathews, I. I.; Cane, D. E.; Khosla, C. The 2.7-Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (30), 11124−11129.

(15) Whicher, J. R.; Smaga, S. S.; Hansen, D. A.; Brown, W. C.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. *Chem. Biol.* **2013**, *20* (11), 1340−1351.

(16) Herbst, D. A.; Jakob, R. P.; Zähringer, F.; Maier, T. Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **2016**, *531* (7595), 533−537.

(17) Agarwal, V.; Diethelm, S.; Ray, L.; Garg, N.; Awakawa, T.; Dorrestein, P. C.; Moore, B. S. Chemoenzymatic synthesis of acyl coenzyme A substrates enables *in situ* labeling of small molecules and proteins. *Org. Lett.* **2015**, *17* (18), 4452−4455.

(18) Worthington, A. S.; Burkart, M. D. One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44−46.

(19) El Gamal, A.; Agarwal, V.; Diethelm, S.; Rahman, I.; Schorn, M. A.; Sneed, J. M.; Louie, G. V.; Whalen, K. E.; Mincer, T. J.; Noel, J. P.; Paul, V. J.; Moore, B. S. Biosynthesis of coral settlement cue tetrabromopyrrole in marine bacteria by a uniquely adapted brominase-thioesterase enzyme pair. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (14), 3797−3802.

(20) Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L. Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry* **2006**, *45* (42), 12756−12766.

(21) Hughes, A. J.; Keatinge-Clay, A. Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* **2011**, *18* (2), 165−176.

(22) Robbins, T.; Kapilivsky, J.; Cane, D. E.; Khosla, C. Roles of conserved active site residues in the ketosynthase domain of an assembly line polyketide synthase. *Biochemistry* **2016**, *55* (32), 4476−4484.

(23) Heide, L. The aminocoumarins: biosynthesis and biology. *Nat. Prod. Rep.* **2009**, *26* (10), 1241−1250.

(24) Copley, S. D. An evolutionary biochemist's perspective on promiscuity. *Trends Biochem. Sci.* **2015**, *40* (2), 72−78.

(25) Fu, C.; Xie, F.; Hoffmann, J.; Wang, Q.; Bauer, A.; Brönstrup, M.; Mahmud, T.; Müller, R. Armeniaspirol antibiotic biosynthesis: chlorination and oxidative dechlorination steps affording spiro[4.4]-non-8-ene. *ChemBioChem* **2019**, *20* (6), 764−769.

(26) Mantovani, S. M.; Moore, B. S. Flavin-linked oxidase catalyzes pyrrolizine formation of dichloropyrrole-containing polyketide extender unit in chlorizidine A. *J. Am. Chem. Soc.* **2013**, *135* (48), 18032−18035.

(27) Yamanaka, K.; Ryan, K. S.; Gulder, T. A.; Hughes, C. C.; Moore, B. S. Flavoenzyme-catalyzed atropo-selective N,C-bipyrrole homocoupling in marinopyrrole biosynthesis. *J. Am. Chem. Soc.* **2012**, *134* (30), 12434−12437.

(28) Zhang, X.; Parry, R. J. Cloning and characterization of the pyrrolomycin biosynthetic gene clusters from *Actinosporangium vitaminophilum* ATCC 31673 and *Streptomyces* sp. strain UC 11065. *Antimicrob. Agents Chemother.* **2007**, *51* (3), 946−957.

(29) Flatt, P. M.; Wu, X.; Perry, S.; Mahmud, T. Genetic insights into pyralomicin biosynthesis in *Nonomuraea spiralis* IMC A-0156. *J. Nat. Prod.* **2013**, *76* (5), 939−946.

(30) Witte, S. N. R.; Hug, J. J.; Géraldy, M. N. E.; Müller, R.; Kalesse, M. Biosynthesis and total synthesis of pyrronazol B: a secondary metabolite from *Nannocystis pusilla*. *Chem. - Eur. J.* **2017**, *23* (63), 15917−15921.

(31) Meiser, P.; Weissman, K. J.; Bode, H. B.; Krug, D.; Dickschat, J. S.; Sandmann, A.; Muller, R. DKxanthene biosynthesis−understanding the basis for diversity-oriented synthesis in myxobacterial secondary metabolism. *Chem. Biol.* **2008**, *15* (8), 771−781.

(32) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *J. Med. Chem.* **2013**, *56* (4), 1363−1388.

(33) Parisini, E.; Metrangolo, P.; Pilati, T.; Resnati, G.; Terraneo, G. Halogen bonding in halocarbon−protein complexes: a structural survey. *Chem. Soc. Rev.* **2011**, *40* (5), 2267−2278.

(34) Khosla, C.; Herschlag, D.; Cane, D. E.; Walsh, C. T. Assembly line polyketide synthases: mechanistic insights and unsolved problems. *Biochemistry* **2014**, *53* (18), 2875−2883.

(35) 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), 11546−11549.

(36) Höfer, I.; Crüsemann, M.; Radzom, M.; Geers, B.; Flachshaar, D.; Cai, X.; Zeeck, A.; Piel, J. Insights into the biosynthesis of hormaomycin, an exceptionally complex bacterial signaling metabolite. *Chem. Biol.* **2011**, *18* (3), 381−391.

(37) Jenner, M.; Frank, S.; Kampa, A.; Kohlhaas, C.; Pöplau, P.; Briggs, G. S.; Piel, J.; Oldham, N. J. Substrate specificity in ketosynthase domains from trans-AT polyketide synthases. *Angew. Chem., Int. Ed.* **2013**, *52* (4), 1143−1147.

(38) Gay, D. C.; Gay, G.; Axelrod, A. J.; Jenner, M.; Kohlhaas, C.; Kampa, A.; Oldham, N. J.; Piel, J.; Keatinge-Clay, A. T. A close look at a ketosynthase from a trans-acyltransferase modular polyketide synthase. *Structure* **2014**, *22* (3), 444−451.

Supplementary Information for:

Gatekeeping Ketosynthases Dictate Initiation of Assembly Line Biosynthesis of Pyrrolic Polyketides

Dongqi Yi,[1] Atanu Acharya,[2] James C. Gumbart,[1,2] Will R. Gutekunst,[1] Vinayak Agarwal[1,3,*]

[1]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332

[2]School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332

[3]School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332

General procedures

All chemicals, solvents and media components were obtained commercially from Sigma-Aldrich, Fisher, Chem-Impex, Acros and Alfa Aesar, and used without further purification. Reactions were monitored by thin layer chromatography (TLC) carried out on Merck silica gel (60 F_{254}) glass plates visualized by UV light. Silica gel (60, particle size 0.036-0.071 mm) was used for flash chromatography. ^{1}H and ^{13}C NMR spectra was recorded on Bruker Avance III 400 and Bruker Avance IIIHD 500 MHz instruments in the solvents indicated, and calibrated using residual undeuterated solvent as an internal reference (CDCl$_3$ δ_H 7.26, (CD$_3$)$_2$CO δ_H 2.05, MeOD δ_H 3.31 and δ_C 49.15). The splitting patterns were reported as s=singlet, d=doublet, t=triplet, q=quadruplet, m= multiplet or unresolved, br=broad signal.

Preparative HPLC

HPLC purification of acyl-*S*-pantetheines was carried out on Luna 5 µm C8(2) 100 Å LC column (250×10 mm) using Agilent 1260 Infinity HPLC system. Mixture of water (solvent A) and acetonitrile (solvent B) with 0.1 % TFA (unless otherwise stated) was used as the mobile phase. A flow rate of 2 mL·min^{-1} was used with the following gradient: 0-3 min: 5% B, 3-15 min: linear gradient to 45% B, 15-25 min: 45% B, 25-30 min: linear gradient to 100% B, 30-34 min: 100% B, 34-35 min: linear gradient to 5% B, 35-36 min: 5 % B, 36-37 min: linear gradient to 100% B, 37-38 min: 100% B, 38-39 min: linear gradient to 5% B, 39-40 min: 5 % B.

LC-MS/MS analysis

Reaction mixtures from all enzyme assays and proteolysis reactions were analyzed by Aeris 3.6 µm WIDEPORE XB-C18 LC column (250×4.6 mm) on an Agilent 1290 Infinity II UHPLC system coupled to a Bruker impact II Q-ToF mass spectrometer operating at room temperature. Mixture of water (solvent A) and acetonitrile (solvent B) with 0.1 % formic acid was used as the mobile phase. A flow rate of 0.5 mL·min^{-1} was used with the following gradient: 0-5 min: 5% B, 5-30 min: linear gradient to 75% B, 30-31 min: linear gradient to 95% B, 31-35 min: 95 % B, 35-36 min: linear gradient to 5% B, 36-38 min: 5 % B, 38-39 min: linear gradient to 95% B, 39-42 min: 95 % B, 42-43 min: linear gradient to 5% B. On the other hand, analysis of acyl-*S*-pantetheines, pyoluteorin and elongated dichloro-pyrrolyl diketides was carried out on a Kinetex 1.7 µm C18 100 Å LC column (50×2.1 mm) column using the same LC-MS operating system. All MS data were collected in the positive mode from *m/z* 50 to 2000 Da, except MS data for elongated dichloro-pyrrolyl diketides which were collected in the negative mode.

## Synthetic procedures

D-Pantetheine (4)[1]
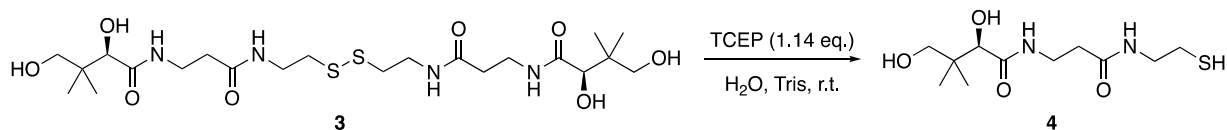
D-Pantethine (116 mg, 0.22 mmol, 1 eq.) and TCEP (68 mg, 0.24 mmol, 1.1 eq.) were dissolved in 2 mL of 10 mM Tris-HCl buffer (pH 7.0). The reaction mixture was incubated at room temperature for 2 h, and then extracted with 7:3 DCM/isopropanol (4 mL, 3×). The organic layer was dried with anhydrous Na_2SO_4. The solvent was removed under vacuum to give **4** (98 mg, 83 %) as colorless oil. ^{1}H NMR (400 MHz, MeOD) δ 3.88 (s, 1H), 3.55 – 3.32 (m, 6H), 2.59 (t, J = 6.9 Hz, 3H), 2.43 (t, J = 6.7 Hz, 2H), 0.91 (s, 6H).

Pyrrolyl acyl-*S*-pantetheine (5)
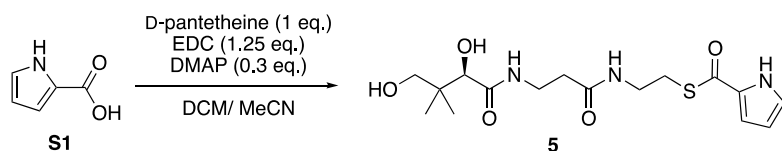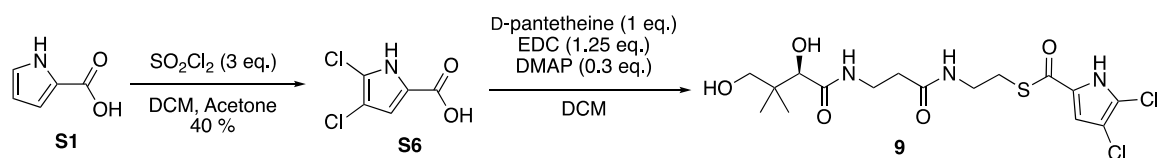

Pyrrole-2-carboxylic acid (80 mg, 0.72 mmol, 1 eq.), D-pantetheine (200.4 mg, 0.72 mmol, 1 eq.), EDC·HCl (172.6 mg, 0.9 mmol, 1.25 eq.), and DMAP (26.4 mg, 0.22 mmol, 0.3 eq.) were dissolved in 9 mL 4:1 DCM/MeCN. The reaction mixture was stirred at room temperature overnight before concentrated under vacuum. Purification by silica flash chromatography (DCM to 1:5 MeOH/DCM) gave **5** (93.8 mg, 35 %) as colorless oil. ^{1}H NMR (400 MHz, MeOD) δ 7.04 (dd, *J* = 2.5, 1.4 Hz, 1H), 6.97 (dd, *J* = 3.9, 1.4 Hz, 1H), 6.21 (dd, *J* = 3.9, 2.5 Hz, 1H), 3.89 (s, 1H), 3.56 – 3.37 (m, 6H), 3.14 (t, *J* = 6.6 Hz, 2H), 2.42 (t, *J* = 6.7 Hz, 2H), 0.92 (s, 6H). HRMS (ESI) *m/z* calculated for $C_{16}H_{25}N_3O_5S$ ($[M+H]^+$) 372.1588, found 372.1583.

5-Methyl-pyrrolyl acyl-*S*-pantetheine (6)
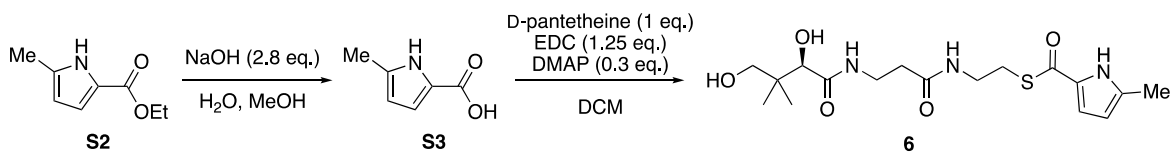
5-Methyl-1H-pyrrole-2-carboxylic acid **S3** was synthesized as previous described.[2] 5-Methyl-1H-pyrrole-2-carboxylic acid (540 mg, 3.53 mmol, 1 eq.) was dissolved in 24 mL MeOH and 14 mL H_2O, followed by

the addition of 1 M NaOH (10 mL). The reaction was stirred at 50 ˚C overnight. The reaction mixture was concentrated under vacuum to remove MeOH, followed by the addition of NH$_4$Cl until saturation. The solution was added with 1 M HCl to adjust pH to 2-3, and then extracted with EtOAc (20 mL, 4×). The organic phase was dried with anhydrous Na$_2$SO$_4$, concentrated under vacuum to give **S3** as brown solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.99 (s, 1H), 6.96 (t, *J* = 3.1 Hz, 1H), 6.00 (ddd, *J* = 3.6, 2.7, 0.8 Hz, 1H), 2.33 (s, 2H).

Compound **S3** (108 mg, 0.86 mmol, 1 eq.), D-pantetheine (240.3 mg, 0.86 mmol, 1 eq.), EDC·HCl (206.8 mg, 1.08 mmol, 1.25 eq.), and DMAP (31.6 mg, 0.26 mmol, 0.3 eq.) were dissolved in 9 mL DCM at room temperature. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by silica flash chromatography (1:4 hexane/EtOAc to 5:1 EtOAc/MeOH) to give **6** (187.6 mg, 56 %) as white solid. ^{1}H NMR (400 MHz, MeOD) δ 6.88 (d, *J* = 3.8 Hz, 1H), 5.92 (d, *J* = 3.8 Hz, 1H), 3.89 (s, 1H), 3.56 – 3.36 (m, 6H), 3.11 (t, *J* = 6.7 Hz, 2H), 2.41 (t, *J* = 6.6 Hz, 2H), 2.26 (s, 3H), 0.92 (s, 6H). ^{13}C NMR (101 MHz, MeOD) δ 181.09, 176.21, 174.07, 137.58, 130.13, 117.89, 110.23, 77.47, 70.52, 40.99, 40.52, 36.57, 36.51, 28.24, 21.47, 21.06, 12.94. HRMS (ESI) *m/z* calculated for C$_{17}$H$_{28}$N$_3$O$_5$S ([M+H]$^+$) 386.1744, found 386.1745.

5-Chloro-pyrrolyl acyl-*S*-pantetheine (7)
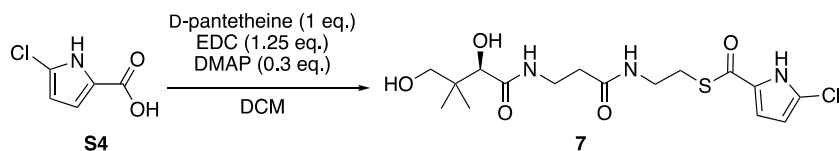
5-Chloro-1H-pyrrole-2-carboxylic acid (27 mg, 0.19 mmol, 1 eq.), D-pantetheine (51.6 mg, 0.19 mmol, 1 eq.), EDC·HCl (44.5 mg, 0.24 mmol, 1.25 eq.), and DMAP (6.8 mg, 0.06 mmol, 0.3 eq.) were dissolved in 2 mL DCM at room temperature. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by preparative HPLC (no addition of TFA) to give **7** as white solid (11.4 mg, 15 %). ^{1}H NMR (400 MHz, MeOD) δ 6.94 (d, *J* = 3.9 Hz, 1H), 6.10 (d, *J* = 4.0 Hz, 1H), 3.89 (s, 1H), 3.54 – 3.36 (m, 6H), 3.14 (t, *J* = 6.6 Hz, 2H), 2.41 (t, *J* = 6.7 Hz, 2H), 0.91 (s, 6H). ^{13}C NMR (126 MHz, MeOD) δ 181.19, 176.22, 174.11, 130.70, 123.96, 117.34, 109.66, 77.41, 70.48, 40.72, 40.52, 36.56, 36.50, 28.44, 21.48, 21.05. HRMS (ESI) *m/z* calculated for C$_{16}$H$_{25}$ClN$_3$O$_5$S ([M+H]$^+$) 406.1198, found 406.1199.

4-Chloro-pyrrolyl acyl-*S*-pantetheine (8)
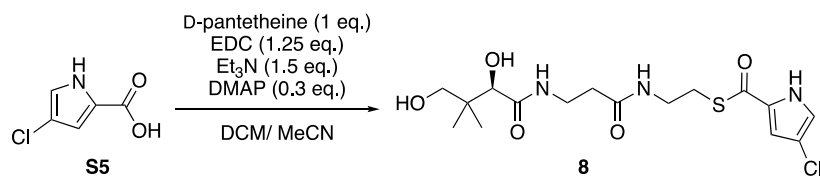
4-Chloro-1H-pyrrole-2-carboxylic acid (50 mg, 0.34 mmol, 1 eq.), D-pantetheine (95.6 mg, 0.34 mmol, 1 eq.), EDC·HCl (82.3 mg, 0.43 mmol, 1.25 eq.), and DMAP (12.6 mg, 0.10 mmol, 0.3 eq.) were dissolved in 4.5 mL 2:1 DCM/acetonitrile at room temperature, followed by the addition of Et$_3$N (52.1 mg, 0.52 mmol, 1.5 eq.). The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by silica flash chromatography (1:2 hexane/EtOAc to 5:1 EtOAc/MeOH) to give **8** (42.1 mg, 30 %) as colorless oil. ^{1}H NMR (400 MHz, MeOD) δ 7.01 (d, J = 1.6 Hz, 1H), 6.87 (d, J = 1.6 Hz, 1H), 3.90 (s, 1H), 3.56 – 3.36 (m, 6H), 3.15 (t, J = 6.6 Hz, 2H), 2.42 (t, J = 6.7 Hz, 2H), 0.92 (s, 6H). ^{13}C NMR (101 MHz, MeOD) δ 181.88, 176.17, 174.08, 130.66, 123.20, 114.93, 114.77, 77.43, 70.49, 40.60, 40.50, 36.55, 36.48, 28.59, 21.47, 21.07. HRMS (ESI) m/z calculated for C$_{16}$H$_{25}$ClN$_3$O$_5$S ([M+H]$^+$) 406.1198, found 406.1202.

4,5-Dichloro-pyrrolyl acyl-*S*-pantetheine (9)

4,5-Dichloro-1H-pyrrole-2-carboxylic acid **S6** was synthesized as previous described.[3] To a stirring solution of pyrrole-2-carboxylic acid (2 g, 18 mmol, 1 eq.) in 45 mL 4:1 dry DCM/acetone, sulfuryl chloride (7.3 g, 4.4 mL, 54 mmol, 3 eq.) was added dropwise. The reaction was stirred at room temperature for 40 min before slowly poured into water (45 mL). The layers were separated, and the aqueous phase was extracted with EtOAc (50 mL, 3×). The combined organic extracts were washed with brine, dried with anhydrous Na$_2$SO$_4$, and concentrated under vacuum. Purification by silica flash column (1:1 hexane/EtOAc to 1:9 MeOH/EtOAc) gave **S6** (1.4 g, 40 %) as dark grey solid. ^{1}H NMR (400 MHz, (CD$_3$)$_2$CO) δ 6.83 (s, 1H).

Compound **S6** (40 mg, 0.22 mmol, 1 eq.), D-pantetheine (61.9 mg, 0.22 mmol, 1 eq.), EDC·HCl (52.3 mg, 0.28 mmol, 1.25 eq.), and DMAP (8.2 mg, 0.07 mmol, 0.3 eq.) were dissolved in 2.2 mL DCM at room temperature. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by preparative HPLC (no addition of TFA) to give **9** as green solid (10.3 mg, 11 %). ^{1}H NMR (400 MHz, MeOD) δ 6.93 (s, 1H), 3.89 (s, 1H), 3.53 – 3.36 (m, 6H), 3.15 (t, J

= 6.5 Hz, 2H), 2.41 (t, J = 6.6 Hz, 2H), 0.92 (s, 6H). ^{13}C NMR (101 MHz, MeOD) δ 181.16, 176.21, 174.12, 129.13, 121.11, 115.39, 111.95, 77.45, 70.51, 40.52, 36.57, 36.49, 28.68, 21.47, 21.07. HRMS (ESI) m/z calculated for $C_{16}H_{24}Cl_2N_3O_5S$ ([M+H]$^+$) 440.0808, found 440.0802.

4,5-Dibromo-pyrrolyl acyl-*S*-pantetheine (10)
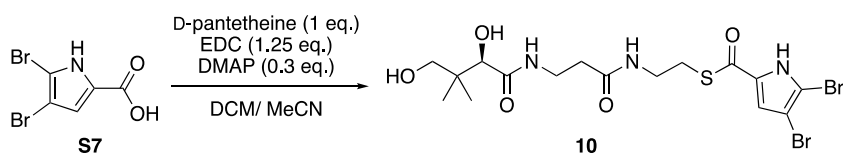
4,5-Dibromo-1H-pyrrole-2-carboxylic acid (50 mg, 0.19 mmol, 1 eq.), D-pantetheine (51.8 mg, 0.19 mmol, 1 eq.), EDC·HCl (44.6 mg, 0.23 mmol, 1.25 eq.), and DMAP (6.8 mg, 0.056 mmol, 0.3 eq.) were dissolved in 3 mL 2:1 DCM/acetonitrile at room temperature. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by preparative HPLC to give **10** (9 mg, 9 %) as white solid. ^{1}H NMR (500 MHz, MeOD) δ 6.99 (s, 1H), 3.89 (s, 1H), 3.53 – 3.36 (m, 6H), 3.15 (t, J = 6.6 Hz, 2H), 2.41 (t, J = 6.7 Hz, 2H), 0.91 (s, 6H). ^{13}C NMR (126 MHz, MeOD) δ 180.96, 176.22, 174.13, 132.53, 118.31, 110.57, 101.04, 77.41, 70.48, 40.52, 36.56, 36.48, 28.68, 21.48, 21.07. HRMS (ESI) m/z calculated for $C_{16}H_{24}Br_2N_3O_5S$ ([M+H]$^+$) 527.9798, found 527.9795.

4,5-Dimethyl-pyrrolyl acyl-*S*-pantetheine (11)
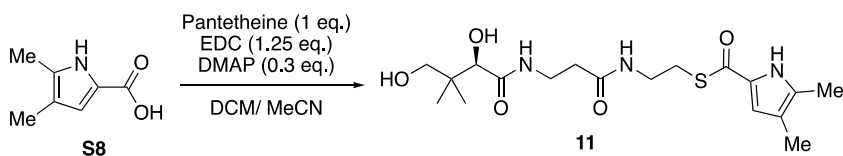
4,5-Dimethyl-1H-pyrrole-2-carboxylic acid (94 mg, 0.68 mmol, 1 eq.), D-pantetheine (188.1 mg, 0.68 mmol, 1 eq.), EDC·HCl (161.9 mg, 0.84 mmol, 1.25 eq.), and DMAP (24.8 mg, 0.20 mmol, 0.3 eq.) were dissolved in 6.8 mL DCM. The reaction was stirred at room temperature overnight. The reaction mixture was concentrated under vacuum and purified by silica flash chromatography (1:2 hexane/EtOAc to 5:1 EtOAc/MeOH) to give **11** (93.8 mg, 35 %) as pink solid. ^{1}H NMR (400 MHz, MeOD) δ 6.73 (s, 1H), 3.89 (s, 1H), 3.54 – 3.33 (m, 6H), 3.09 (t, J = 6.7 Hz, 2H), 2.41 (t, J = 6.6 Hz, 2H), 2.17 (s, 3H), 1.98 (s, 3H), 0.91 (s, 6H). ^{13}C NMR (101 MHz, MeOD) δ 180.70, 176.16, 174.02, 134.80, 128.34, 118.90, 118.52, 77.45, 70.50, 41.02, 40.50, 36.56, 36.50, 28.19, 21.47, 21.07, 11.07, 10.92. HRMS (ESI) m/z calculated for $C_{18}H_{30}N_3O_5S$ ([M+H]$^+$) 400.1901, found 400.1900.

3,4,5-Tribromo-pyrrolyl acyl-*S*-pantetheine (12)
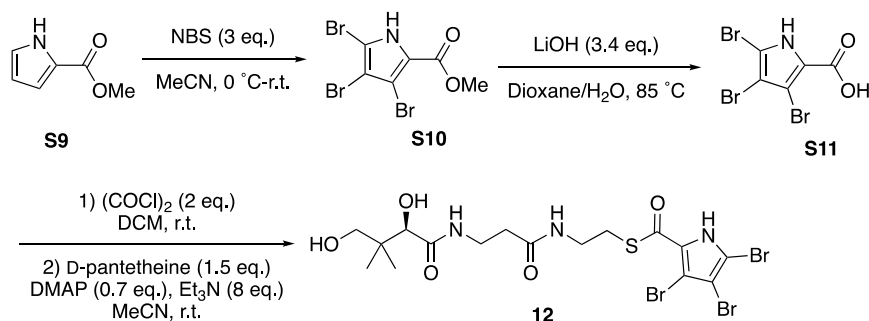
Methyl 1H-pyrrole-2-carboxylate (2 g, 15.98 mmol, 1 eq.) and *N*-Bromosuccinimide (9.9 g, 55.94 mmol, 3.5 eq.) were dissolved in 100 mL MeCN. The reaction was stirred at 0 ˚C for 1.5 h before quenched with water (100 mL). The layers were separated and the aqueous phase was extracted with EtOAc (100 mL, 3×). The organic layers were combined, washed with brine, dried with anhydrous Na_2SO_4, and concentrated under vacuum. Purification by silica flash column (hexane to 4:1 hexane/EtOAc) afforded **S10** (3.8 g, 66 %) as brown solid. ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.67 (s, 1H), 3.92 (s, 3H).

Compound **S10** (100 mg, 0.28 mmol, 1 eq.) and LiOH (40.4 mg, 0.96 mmol, 3.4 eq.) were dissolved in 3 mL 1:1 dioxane/H_2O. The reaction was stirred at 85 ˚C for 6 h before quenched with 1 M HCl (3 mL) and extracted with 7:3 DCM/isopropanol (5 mL, 3×). The organic phase was dried with anhydrous Na_2SO_4, and concentrated under vacuum. Purification by silica flash column (1:1 hexane/EtOAc to 1:9 MeOH/EtOAc) gave **S11** (91.8 mg, 93 %) as brown solid. ^{13}C NMR (101 MHz, MeOD) δ 162.46, 125.00, 106.53, 105.66, 104.76.

Molecule **S11** (75 mg, 0.22 mmol, 1 eq.) was dissolved in dry DCM (2.5 mL). Oxalyl chloride (0.04 mL, 0.44 mmol, 2 eq.) was added dropwise to the solution. The reaction was stirred at room temperature for 2 h before concentrated under vacuum to afford an orange solid. The solid (35 mg, 0.096 mmol, 1 eq.) together with D-pantetheine (40 mg, 0.143 mmol, 1.5 eq.) and DMAP (8.17 mg, 0.067 mmol, 0.7 eq.) were dissolved in 1 mL dry MeCN, followed by the addition of Et_3N (77.36 mg, 0.765 mmol, 8 eq.). The reaction was stirred at room temperature for 2 h before the removal of solvent. The residue was purified by preparative HPLC to give **12** (13.3 mg, 23%) as white solid. ^{1}H NMR (400 MHz, MeOD) δ 3.89 (s, 1H), 3.55 – 3.37 (m, 6H), 3.21 (t, *J* = 6.6 Hz, 2H), 2.42 (t, *J* = 6.7 Hz, 2H), 0.91 (s, 6H). HRMS (ESI) *m/z* calculated for $C_{16}H_{23}Br_3N_3O_5S$ ([M+H]$^+$) 605.8903, found 605.8892.

Cloning, expression and purification of holo-PKS M1

On the basis of primary sequence, PltB KS-AT-CP domain boundary was mapped to residues 1 through 1031, and CalA KS-AT-CP boundary was mapped to residues 1 through 1070. PltB M1 was amplified from the genomic DNA extracted from *Pseudomonas protegens* Pf-5 using PrimeSTAR DNA polymerase, and cloned into pET24(+) vector by NEBuilder HiFi DNA Assembly master mix. CalA M1 carried by pET28(+) vector was synthesized and cloned by Twist Bioscience. Sfp gene was amplified from pET24(+)-Sfp, a construct reported previously,[4] using Phusion DNA polymerase, and cloned into MCS1 of pCDFDuet-1 vector. Plasmids for PKS mutants were generated by site-directed mutagenesis using primers containing desired mutations.

For overexpression of holo-PKS M1 proteins, two plasmids carried PKS M1 and Sfp were co-transformed into *Escherichia coli* BL21Gold(DE3). The cell cultures were grown in terrific broth medium (2-4 L) supplemented with appropriate antibiotics and calcium pantothenate (15 mg/L) at 30 °C until OD_{600} reached 0.4-0.5. The culture temperature was reduced to 18 °C. When OD_{600} reached 0.7-0.8, protein expression was induced by the addition of 0.05-0.1 mM IPTG. The induced culture was grown at 18 °C for an addition of 18 h. All subsequent steps of protein purification were conducted at 4 °C or on ice. Cells were harvested by centrifugation, resuspended in binding buffer (20 mM Tris-HCl (pH=8.0), 500 mM NaCl), and lysed by sonication (15 s amplification, 45 s settlement per circle). The lysate was clarified by centrifugation at 18000 rpm for 45 min, and loaded on 5 mL HisTrap HP column using AKTAprime plus FPLC system. The protein-bound column was washed extensively with wash buffer (20 mM Tris-HCl (pH=8.0), 30 mM imidazole, 500 mM NaCl), and then eluted with a linear gradient to 100 % of elution buffer (20 mM Tris-HCl (pH=8.0), 250 mM imidazole, 500 mM NaCl) over 10 column volumes. Purity of eluent fractions were checked by SDS-PAGE. The fractions containing desired proteins were pooled, concentrated using 50 kDa Amicon centrifugal filters, and desalted into storage buffer (20 mM Tris-HCl (pH=8.0), 500 mM NaCl, 10 % glycerol) with PD-10 columns. Purified proteins were stored as small aliquots at -80 °C and fresh aliquots were used each time for enzyme assays.

Cloning, expression and purification of CPs and other enzymes

CoaA, CoaD, CoaE, Sfp and MatB enzymes were obtained as previously described.[4-5] CalN3, Mpy15 and Clz18 gene fragments synthesized by Twist Bioscience or Integrated DNA Technologies were used as templates for PCR, while PltL gene was amplified directly from *P. protegens* Pf-5 genomic DNA. CP DNA fragments were cloned into pET28(+) vector using NEBuilder HiFi DNA Assembly master mix. HrmL carried by pET28(+) vector was provided by Twist Bioscience. Plasmids containing CP sequences were transformed into *E. coli* BL21Gold(DE3) for expression of CPs. The cell cultures were grown in 1 L terrific broth medium supplemented with kanamycin at 30 °C until OD_{600} reached 0.4-0.5. The culture temperature

was reduced to 18 ˚C. When OD_{600} reached 0.7-0.8, protein expression was induced by the addition of 0.2 mM IPTG. The induced culture was grown at 18 ˚C for an addition of 18 h. CPs were purified following similar procedures described for PKS M1 proteins, except that the protein fractions were added with thrombin prior to dialysis in binding buffer overnight, and stored at -80 ˚C with 10% glycerol.

Preparation of pyrrolyl-*S*-CPs

Synthesized pyrrolyl acyl-*S*-pantetheines were dissolved in DMSO to a concentration of 100 mM. Assays to prepare pyrrolyl-*S*-CPs were performed as previously described.[4] In brief, the assay was composed of 1 mM pyrrolyl acyl-*S*-pantetheines, 10 mM $MgCl_2$, 50 mM HEPES-Na (pH=7.9), 200 µM CPs, 2.5 µM Sfp, and 1 µM each of CoaA, CoaD, and CoaE in a total volume of 2.5 mL. The enzymatic reaction was initiated by the addition of 9 mM ATP after incubation at 30 °C for 5 min. The assay was incubated at 30 ˚C for 4.5 h, followed by 4 ˚C overnight, and then desalted into storage buffer with PD-10 columns. Biosynthesized pyrrolyl-*S*-CPs were analyzed using LC-MS/MS described above. Protein solutions were stored as small aliquots at -80 ˚C and fresh aliquots were used each time for enzyme assay.

CP occupancy assay

PltB assay was performed in a total volume of 100 µL containing 2 mM malonyl CoA, 2 mM TCEP, 7.7 µM PltB (KS-AT-CP module1, Fig. S1), 15.4 µM pyrrolyl-*S*-CPs or substituted pyrrolyl-*S*-CPs, and 50 mM HEPES-Na (pH 7.5). For CalA assay, methylmalonyl CoA was synthesized in situ using MatB enzyme as previously described.[6] The MatB reaction comprised of 5 mM CoA-SH, 10 mM methylmalonate, 10 mM ATP, 10 mM $MgCl_2$, 5 mM TCEP, 15 % glycerol, 100 mM HEPES-Na (pH 7.5) and 5 µM MatB in a total volume of 100 µL. After incubation at 30 ˚C for 18 h, 30 µL of MatB reaction mixture was added to CalA assay to a total volume of 100 µL which contained 2 mM TCEP, 10 µM CalA (KS-AT-CP module1, Fig. S1), 20 µM pyrrolyl-*S*-CPs or substituted pyrrolyl-*S*-CPs, and 50 mM HEPES-Na (pH 7.5). Assays were incubated at 30 ˚C for 1 h. 50 µg protein samples from the assays were digested with 1 µg trypsin (proteomics grade, Sigma-Aldrich T7575) for 5 h, following the manufacturer's protocol. For PKS time-dependent assay, at each time point, 42 µL sample was taken from enzyme assays and added to 28 µL MeCN. The precipitate was removed by centrifugation, and the supernatant was used for trypsin digestion following the same protocol describe above. The proteolysis reactions were quenched by the addition of equal volume of MeCN with 5 % formic acid and analyzed by LC/MS.

In PKS assays, three different acyl states of the donor CPs were detected: (substituted)pyrrole-*S*-CPs (state **a**) corresponding to leftover donor CP substrate; holo-CPs (**b**), the product of KS

transesterification reaction which transferred pyrrolyl cargos from donor CPs and generated a free thiol; (methyl)malonyl CPs (**c**), the production of (methyl)malonyl CPs resulted from the nonspecific transacylation activity of AT domain that loaded (methyl)malonyl group onto the newly produced holo-CPs after the transthiolation of pyrrolyl cargos by KS domain. In order to demonstrate the selectivity of PKS KS domains toward different CP substrates, we calculated the percentage of donor CP substrates that were consumed by KS domain in PKS assays. After the first round of polyketide extension, the acceptor CP (acceptor CP is the CP of the KS-AT-CP PKS module) is acylated with the pyrrolic diketide. Hydrolysis of this pyrrolic diketide, the detection of which is illustrated in Figure S36, resets the acceptor CP for the next round of polyketide extension. As shown in the figure below, EICs corresponding to trypsin digested CP peptide fragments containing these three different acyl states of Ppant arms were generated (the trypsin fragments for each CP that we detected in our assays is listed in Table S1). EIC peak areas of CP peptide fragments were then integrated using Bruker DataAnalysis software. The total abundance of all CPs was represented by the sum of EIC peak areas for (substituted)pyrrolyl- (state **a**), holo- (state **b**), and (methyl)malonyl-*S*-CPs (state **c**). The CP substrate consumption was then calculated by the ratio of sum of states **b** and **c** to the sum of states **a**, **b**, and **c**. Note that state **c** is generated after the donor CP delivers the molecular cargo to the KS domain. Hence, states **b** and **c**, together, represent the amount of donor CP that was consumed in the reaction. State **a** represents the donor CP that was not consumed during the course of the reaction. The retention times for trypsin peptide fragments corresponding to states **a**, **b**, and **c** are different as illustrated in the figure below. Hence, ions detected as state **b** are not in source degradative fragments generated in the mass spectrometer.
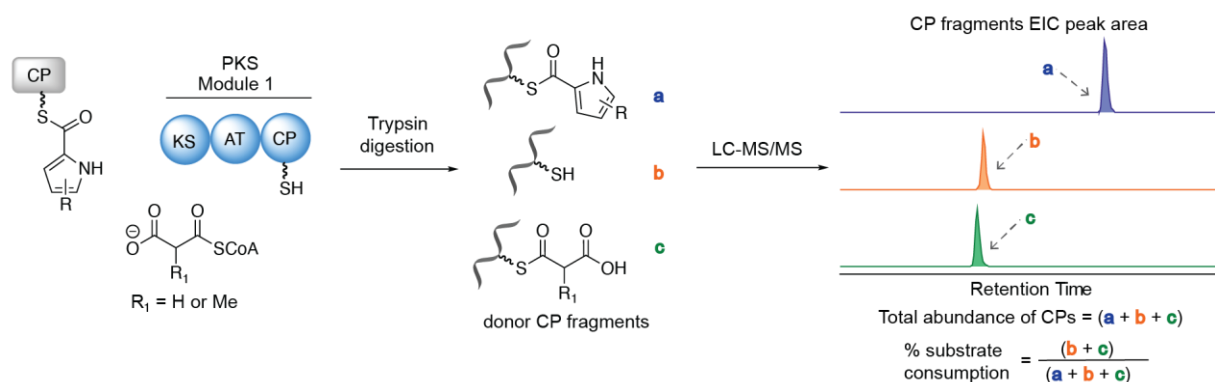
Analysis of elongated dichloro-pyrrolyl diketides

A 300 µL PltB assay was set up as described above. After incubation at 30 ˚C for 1 h, the reaction mixture was added with equal volume of 1 M KOH, incubated at 30 ˚C for another 30 min, quenched with 300 µL 1.5 M HCl, and then lyophilized to remove water. The solid was dissolved in 100 µL MeOH. The

solution was clarified by centrifugation and analyzed by LC-MS using the following conditions at a flow rate of 0.5 mL·min^{-1}: 0-5 min: 5% B, 5-20 min: linear gradient to 70% B, 20-25 min: linear gradient to 100% B, 25-28 min: 100% B, 28-29 min: linear gradient to 5% B, 29-30 min: 5 % B, 30-31 min: linear gradient to 100% B, 31-32 min: 100% B, 32-33 min: linear gradient to 5% B, 33-34 min: 5 % B

Organic extraction of *Pseudomonas protegens* Pf-5

P. protegens Pf-5 was inoculated into 10 mL LB media from glycerol stock and grown at 30 ˚C overnight. Cell pellets and liquid media was separated by centrifugation. The cell pellets were resuspended in 2 mL MeOH and sonicated for 15 min. The extracts from cell pellets were clarified by centrifugation. Otherwise, the liquid media was extracted with EtOAc (10 mL, 3×). The organic layers were combined, dried with anhydrous Na_2SO_4, and concentrated under vacuum. The concentrated residue was dissolved in 2 mL MeCN. The organic extracts from both cell pellets and liquid media were analyzed by LC-MS at a flow rate of 0.5 mL·min^{-1} using the following gradient: 0-6 min: 5% B, 6-15 min: linear gradient to 70% B, 15-25 min: linear gradient to 85% B, 25-30 min: linear gradient to 100% B, 30-34 min: 100% B, 34-35 min: linear gradient to 5% B, 35-36 min: 5 % B, 36-37 min: linear gradient to 100% B, 37-39 min: 100% B, 39-41 min: linear gradient to 5% B, 41-43 min: 5 % B.

Computation of electrostatic potential maps

The geometries of all molecules were optimized using the wb97xd functional[7] and the aug-cc-pvtz basis set with CPCM continuum solvation model.[8] We used water as the model solvent in all calculations. In each case, the minimized geometry is confirmed by the absence of imaginary frequency in computed frequencies at the optimized geometry. Electrostatic potential was computed and mapped onto the electron density surface computed plotted with an isovalue of 0.0004. All calculations were performed using Gaussian16[9] and visualization was done with Gaussview6.[10] The respective geometries of the molecules are provided below. Free rotation around C-C single bond may originate different conformation between these molecules (Fig. S38). However, the C-C single bond is away from the pyrrole ring and therefore, unlikely to significantly influence the electrostatic potential of the 5-membered pyrrole rings.

dichloro-substituted pyrrole:

 Nuclear repulsion energy before empirical dispersion term = 1432.6511293513 hartree

 C -0.2490499316 -0.6303987386 1.3485999095

C -1.3307833197 -0.2620618828 0.5784439513

N -0.9389973154 0.7233044065 -0.245664639

Cl -2.92273143 -0.8682908234 0.580105548

C 0.3861235753 1.0132329091 -0.0256173551

H -1.5061999147 1.1925444985 -0.9328119428

C 0.8392619181 0.1704306134 0.973030818

C 0.9436358678 2.0622341325 -0.8727486836

H 1.8250521049 0.0856336868 1.3994277075

Cl -0.2519900643 -1.8597063024 2.5509914368

S 2.5642541176 2.7675024428 -0.726111141

C 3.4982794077 2.0297561874 0.6416927394

C 4.2429837732 0.7566714851 0.2465411193

O 0.2644359677 2.5481077502 -1.7602373731

N 5.1413620328 0.3162807534 1.2919140014

C 4.7161027366 -0.3399137347 2.3878810218

C 5.7681004651 -0.7168479444 3.3969819967

O 3.5323390477 -0.6213348544 2.5550782184

H 2.8562266864 1.88532085 1.5055839343

H 4.2157016914 2.8085306197 0.8971784485

H 3.5401997158 -0.0432783316 0.0179089521

H 4.8346281977 0.9417205586 -0.6469242351

H 6.1134936179 0.5545655011 1.2166700846

H 6.7719271748 -0.4193483519 3.1035649941

H 5.7410396882 -1.795845485 3.5393615841

H 5.5196041889 -0.2488099462 4.3481589039

dibromo-substituted pyrrole:

Nuclear repulsion energy before empirical dispersion term = 1994.5664656622 hartree

C -0.2533992285 -0.6120349891 1.3762283125

C -1.3390448937 -0.2421285167 0.6102039543

N -0.9444091277 0.7364362351 -0.2220521925

Br -3.081240266 -0.8856822531 0.6188077645

C 0.3829936953 1.0206367566 -0.0135074381

H -1.511115102 1.2071782388 -0.9086476316

C 0.83718853 0.1814368561 0.9873354596

C 0.9393789692 2.0621906324 -0.8722322924

H 1.8259498053 0.096113208 1.4067812513

Br -0.241069165 -1.9382066419 2.694611269

S 2.5636518053 2.7610584859 -0.7431657839

C 3.5058542153 2.0261061651 0.6202897126

C 4.2404677198 0.7471703277 0.225295764

O 0.2551020268 2.5446828607 -1.7574690307

N 5.1492986285 0.3102925459 1.2630751612

C 4.7330579156 -0.3293347588 2.3722144772

C 5.7951068696 -0.703167056 3.3719344877

O 3.5492822451 -0.5984307593 2.5581205682

H 2.870832809 1.8894901622 1.4905822986

H 4.2295609963 2.8025383306 0.8650164273

H 3.5315348627 -0.0512479556 0.0104908642

H 4.8225309793 0.9233897787 -0.6762036133

H 6.1220728769 0.5404894067 1.1728077215

H 6.7979594333 -0.4215708228 3.0599648885

H 5.7579100165 -1.7794694193 3.53124059

H 5.565543383 -0.2179368177 4.3192770103

Table S1. CP tryptic fragments detected and quantified in LC/MS data.

Carrier protein	CP fragment sequence with active site Ser	Exacted ions $[M+nH]^{n+}$
pyrrolyl-*S*-CalN3		1063.2930
dichloro-pyrrolyl-*S*-CalN3		1080.2736
monochloro-pyrrolyl-*S*-CalN3		1071.7833
5-methylpyrrolyl-*S*-CalN3		1066.7969
dibromo-pyrrolyl-*S*-CalN3		1102.7483
dimethyl-pyrrolyl-*S*-CalN3	FLDGDPQGELEETTPLLELGVLNSLNTVVLLAHIR	1070.3009
tribromo-pyrrolyl-*S*-CalN3		1121.7259
holo-CalN3		1040.0377
malonyl-*S*-CalN3		1061.5378
methylmalonyl-*S*-CalN3		1065.0417
pyrrolyl-*S*-PltL		1106.7764
dichloro-pyrrolyl-*S*-PltL		1123.7569
monochloro-pyrrolyl-*S*-PltL		1115.2666
5-methylpyrrolyl-*S*-PltL		1110.2803
dibromo-pyrrolyl-*S*-PltL	YIMEDLIGPSAKEDELDDQTPLLEWGILNSMNIVK	1145.7316
dimethyl-pyrrolyl-*S*-PltL		1113.7842
tribromo-pyrrolyl-*S*-PltL		1165.2092
holo-PltL		1083.5210
malonyl-*S*-PltL		1105.0211
methylmalonyl-*S*-PltL		1108.5250
pyrrolyl-*S*-Mpy15		1214.2260
dichloro-pyrrolyl-*S*-Mpy15		1236.8667
holo-Mpy15	SFLDDDTTALEPNTPLLEWGILNSMNTAK	1183.2189
malonyl-*S*-Mpy15		1211.8857
methylmalonyl-*S*-Mpy15		1216.5578
pyrrolyl-*S*-Clz18		1103.3033
dichloro-pyrrolyl-*S*-Clz18		1120.2839
holo-Clz18	YLPEGESGLQPSSPLLEWGILTSLSTTELISFILER	1080.0480
malonyl-*S*-Clz18		1101.5481
methylmalonyl-*S*-Clz18		1105.0520
pyrrolyl-*S*-Clz18		1159.8416
dichloro-pyrrolyl-*S*-HrmL	LLDLGILDSLAVLSVVSEIEQTLDLEFPETEIVATNFR	1176.8221
holo-HrmL		1136.5862

malonyl-*S*-HrmL		1158.0863
methylmalonyl-*S*-HrmL		1161.5902
holo-PltB CP		757.0495
malonyl-PltB CP	GFFELGLS**S**ISLVEFKR	785.7163
dichloro-pyrrolyl-diketide acylated to PltB CP		824.7008
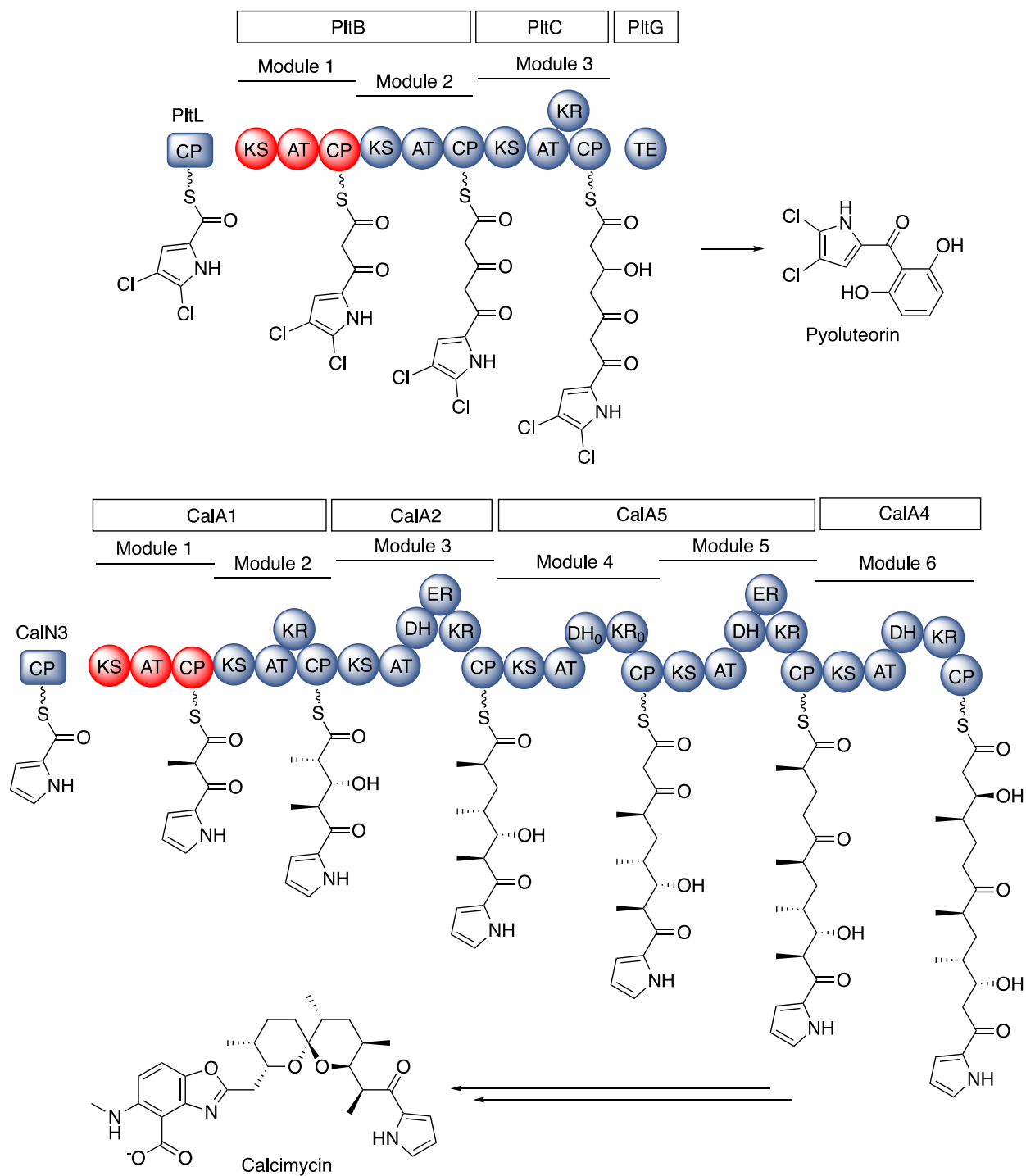
Figure S1: Pyoluteorin (top) and calcimycin (bottom) PKS pathways. PltB and CalA1 are bimodular PKSs. The PltB and CalA1 KS-AT-CP modules investigated in this study are highlighted in red.
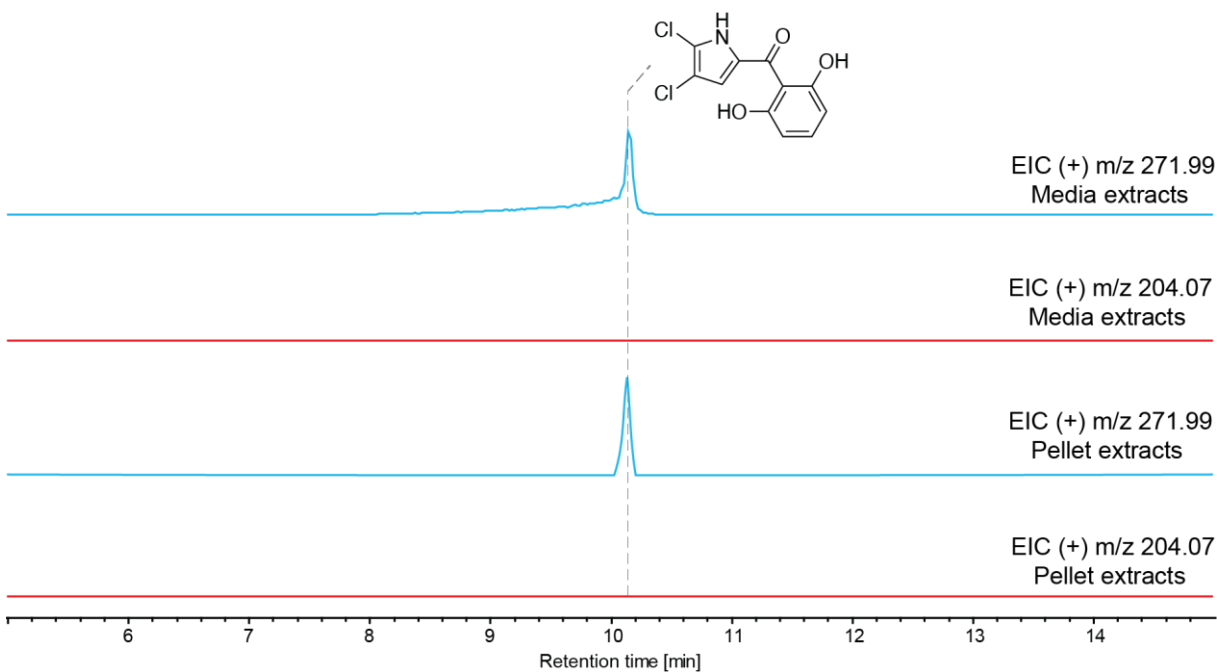
Figure S2: Organic extracts of *P. protegens* Pf-5 (cell pellets and liquid media) analyzed by LC-MS. Pyoluteorin (**1**), as denoted by extracted ion chromatograms (EICs) corresponding to *m/z* 271.99 Da, was detected in culture supernatant and the cell pellet. However, deschloro-pyoluteorin, denoted EICs corresponding to *m/z* 204.07 Da, was not detected in either the culture supernatant or the cell pellet.
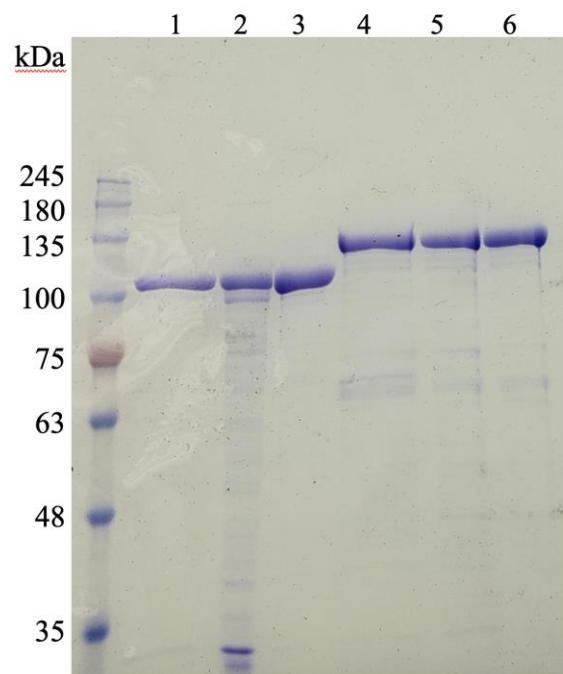
Figure S3: SDS-PAGE demonstrating purified PKS modules (from left to right): ladder, (lane 1) PltB (111.5 kDa), (lane 2) PltB Cys183Ala, (lane 3) PltB Met222Leu, (lane 4) CalA (113.4 kDa), (lane 5) CalA Cys182Ala mutant, and (lane 6) CalA Leu221Met.
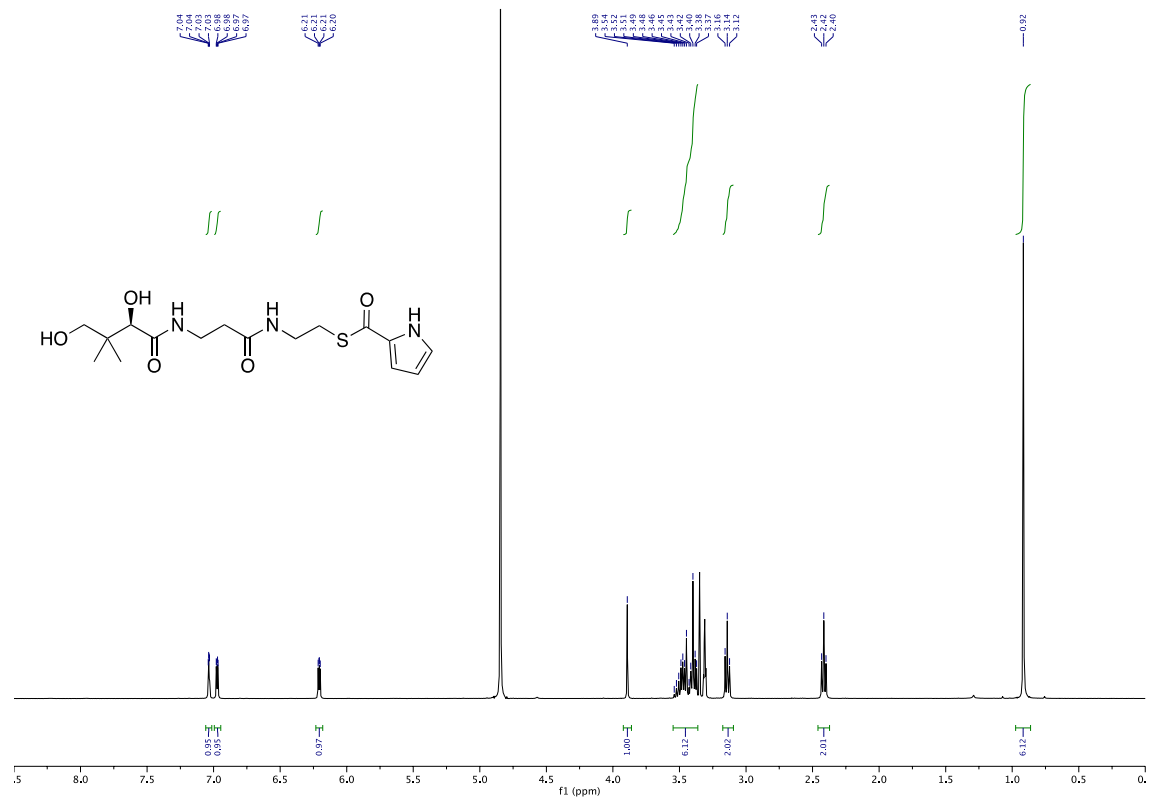
Figure S4: ^{1}H NMR spectrum (400 MHz, MeOD) of compound **5**
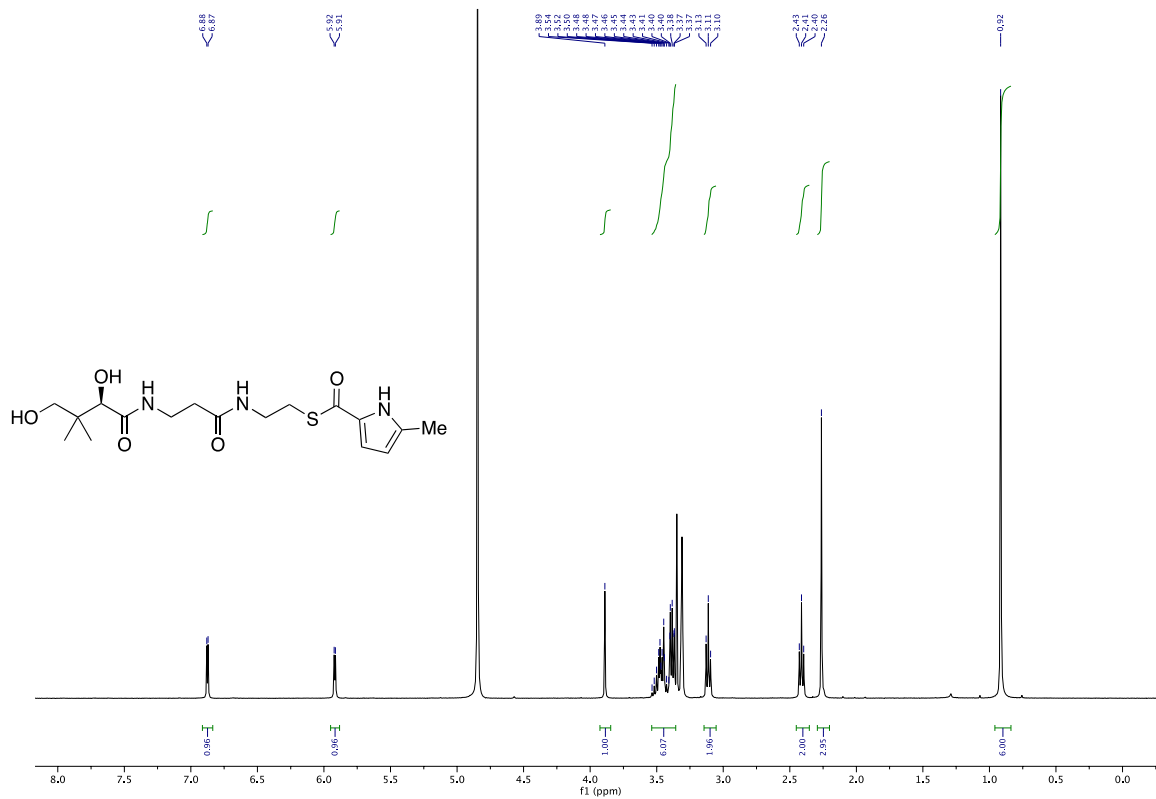
Figure S5: ^{1}H NMR spectrum (400 MHz, MeOD) of compound **6**
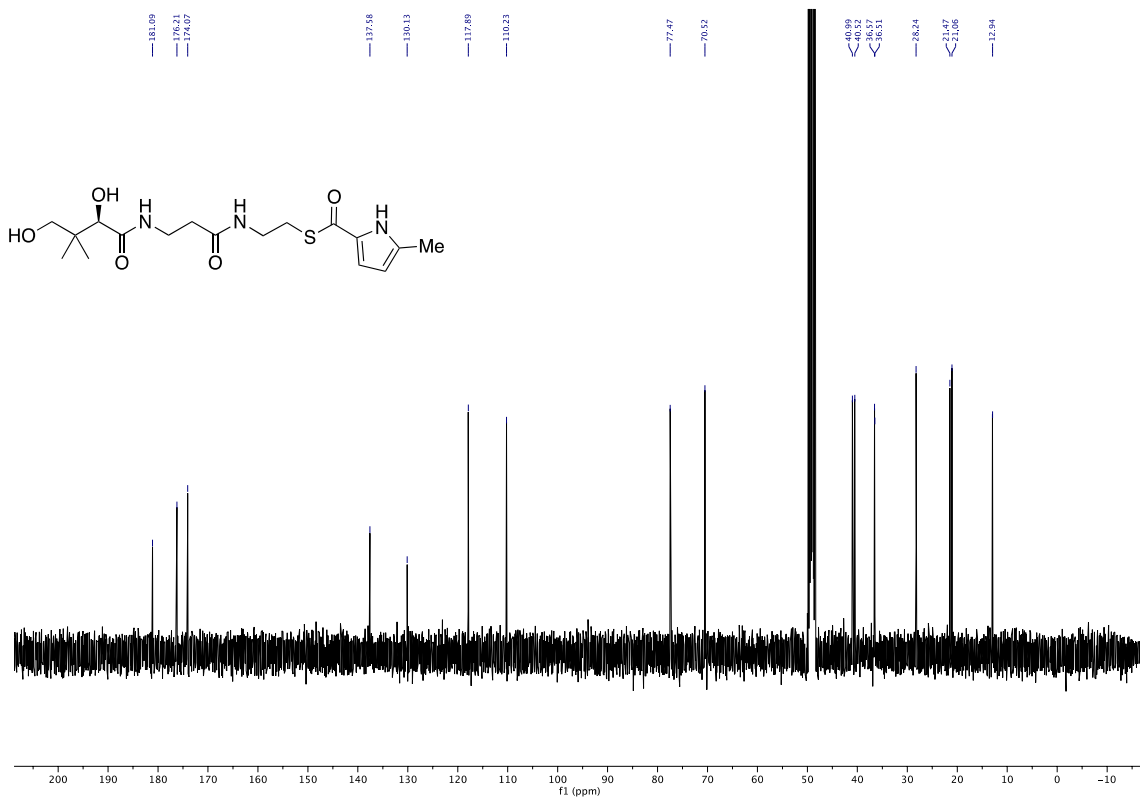
Figure S6: ^{13}C NMR spectrum (101 MHz, MeOD) of compound **6**
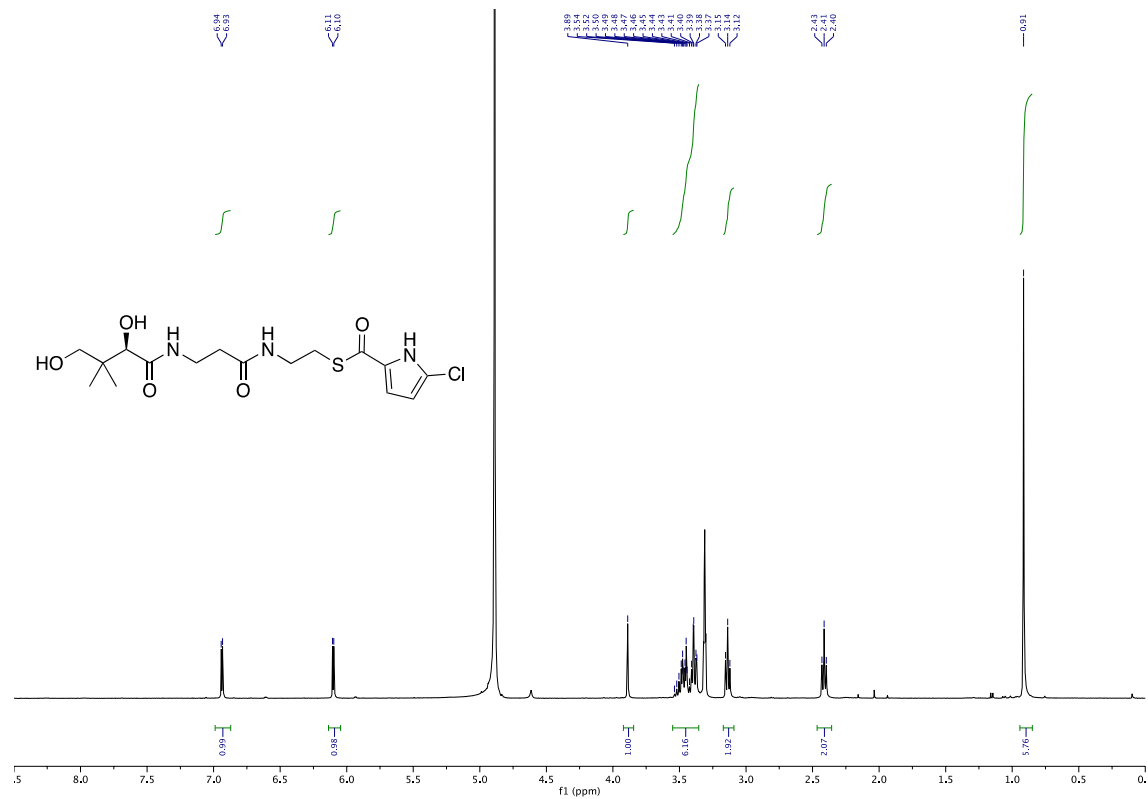
Figure S7: ^{1}H NMR spectrum (400 MHz, MeOD) of compound **7**
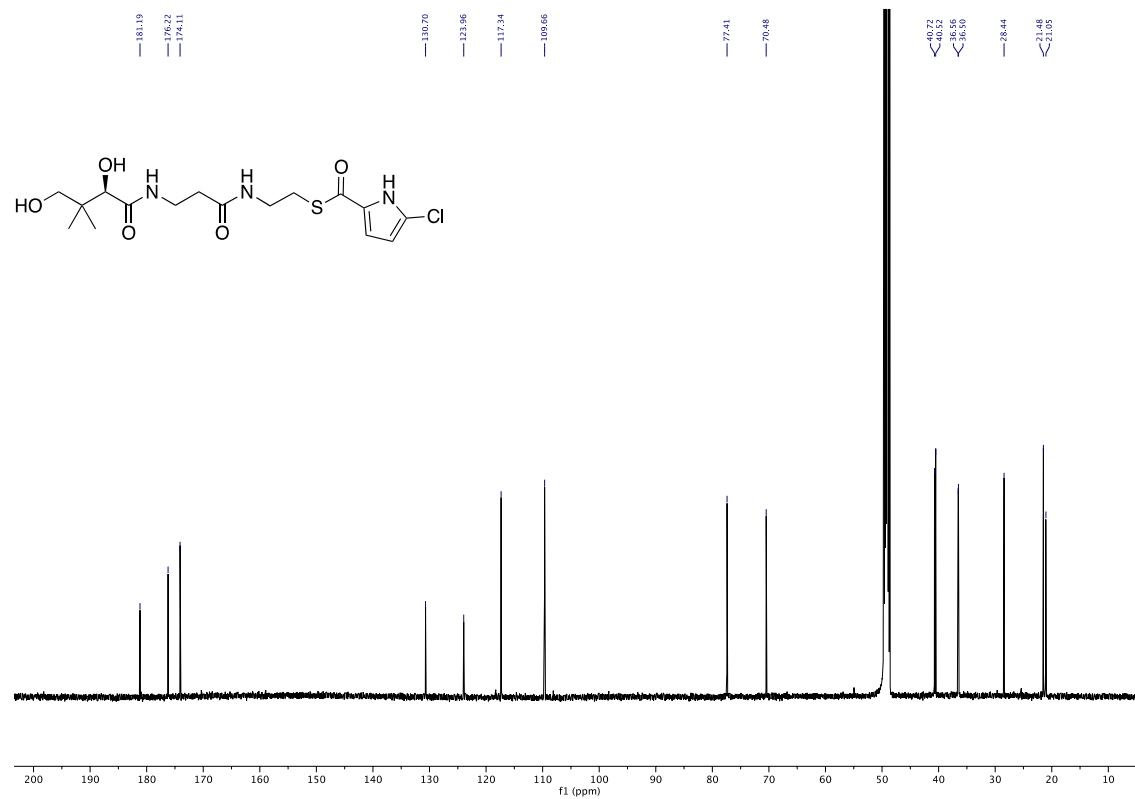
Figure S8: ^{13}C NMR spectrum (126 MHz, MeOD) of compound **7**
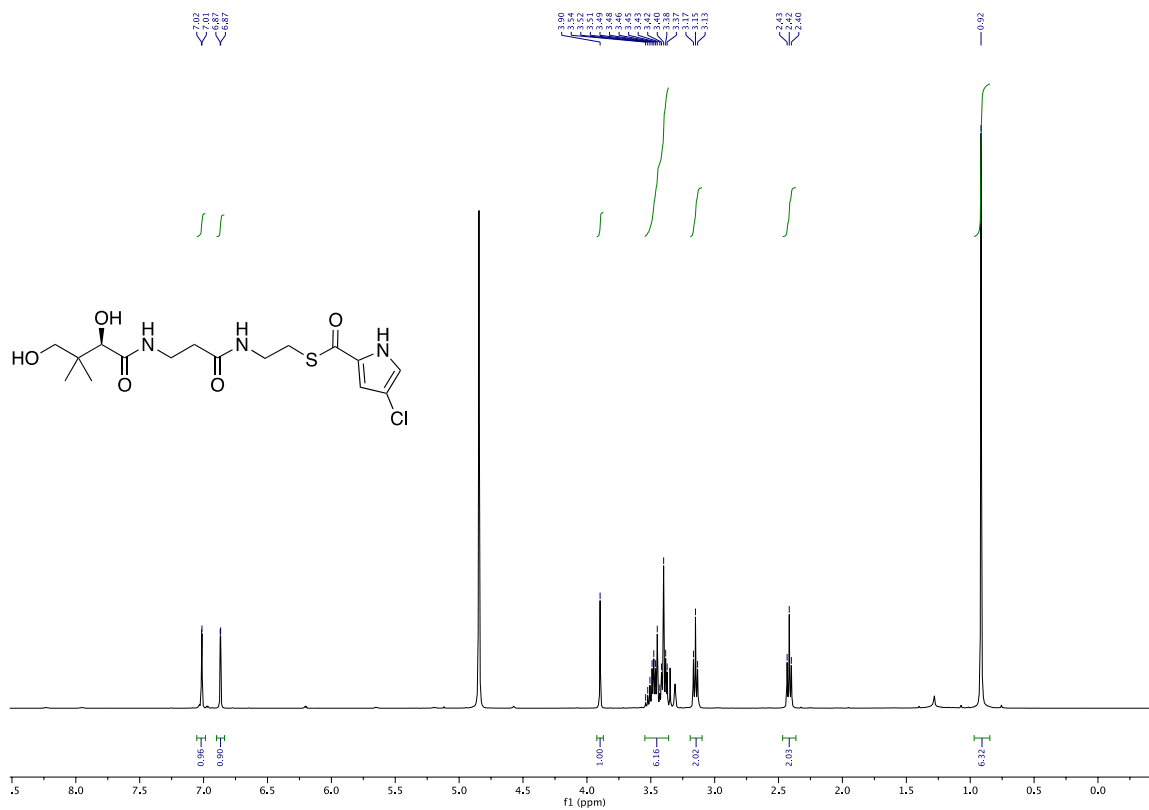
Figure S9: ^{1}H NMR spectrum (400 MHz, MeOD) of compound **8**
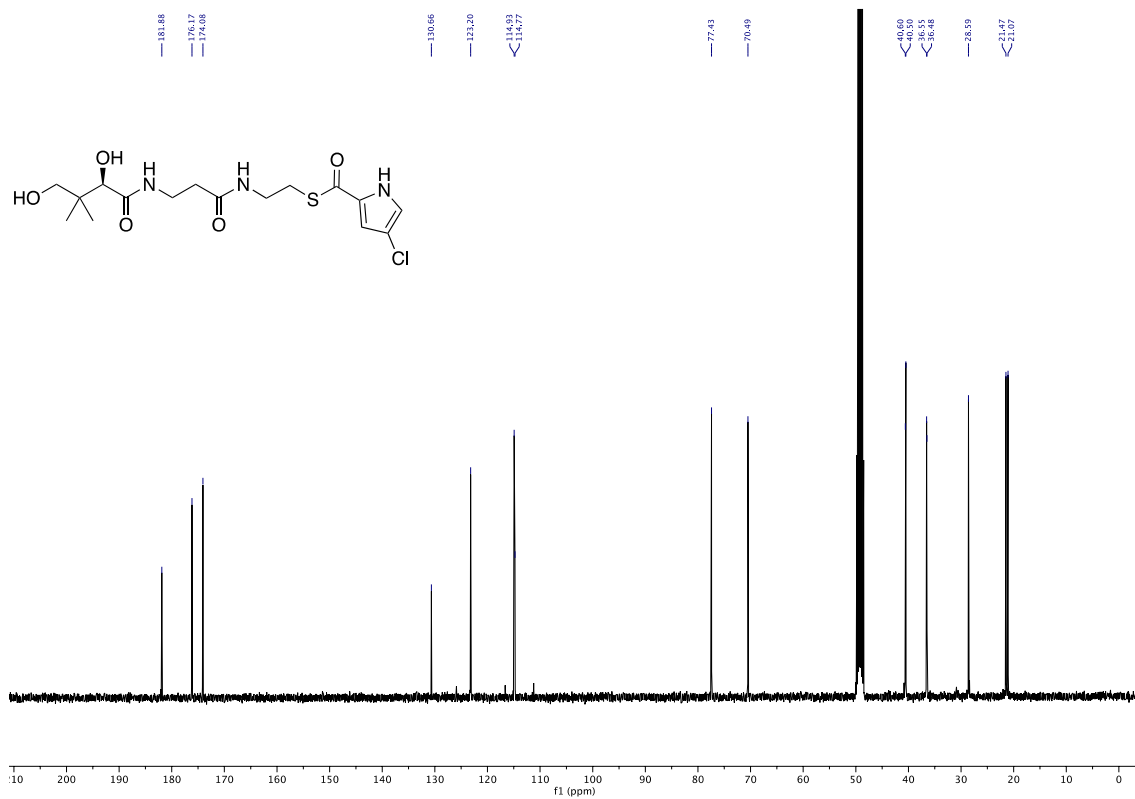
Figure S10: ^{13}C NMR spectrum (101 MHz, MeOD) of compound **8**
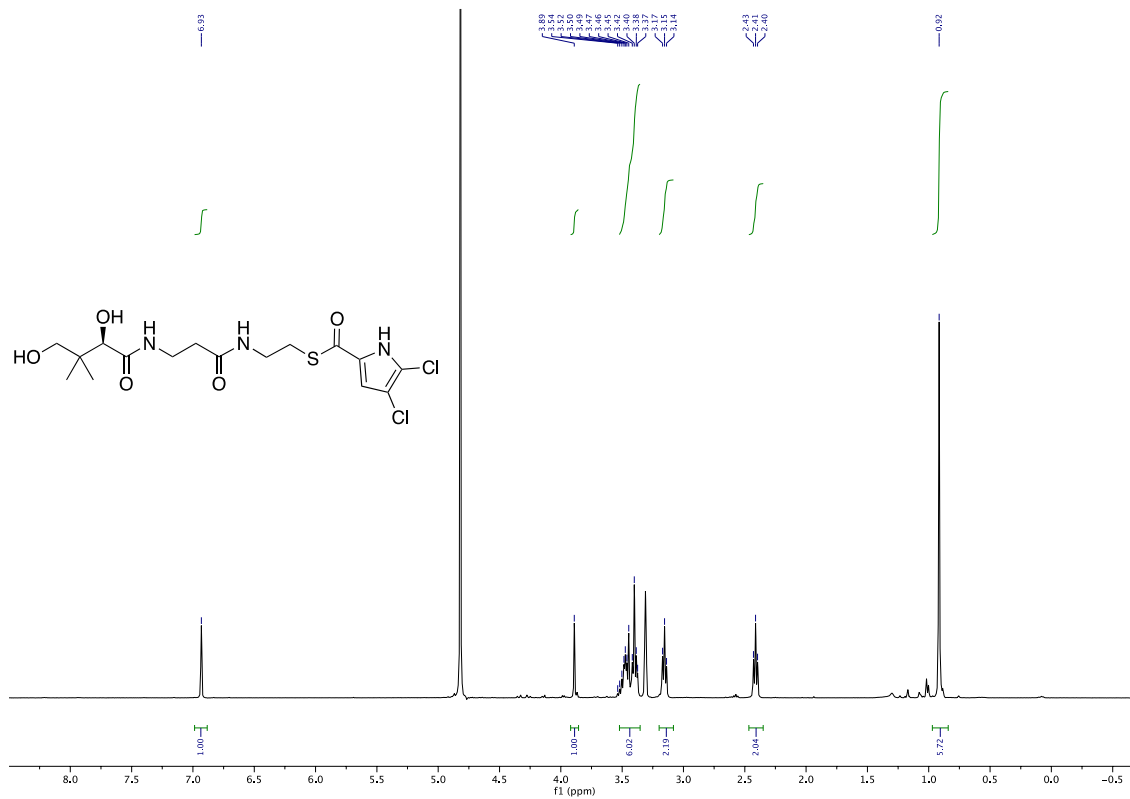
Figure S11: ^{1}H NMR spectrum (400 MHz, MeOD) of compound **9**

181.16
176.21
174.12
129.13
121.11
115.39
111.95
77.45
70.51
40.52
36.57
36.49
28.68
21.47
21.07

f1 (ppm)
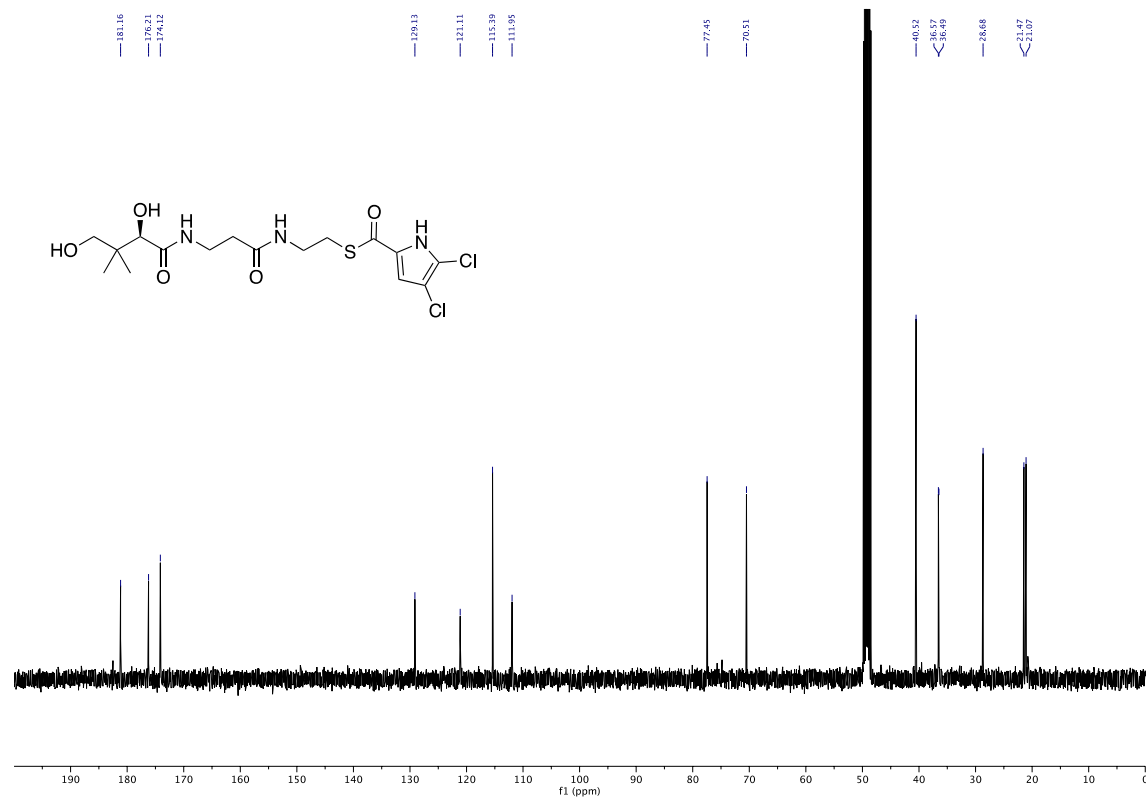
Figure S12: ^{13}C NMR spectrum (101 MHz, MeOD) of compound **9**
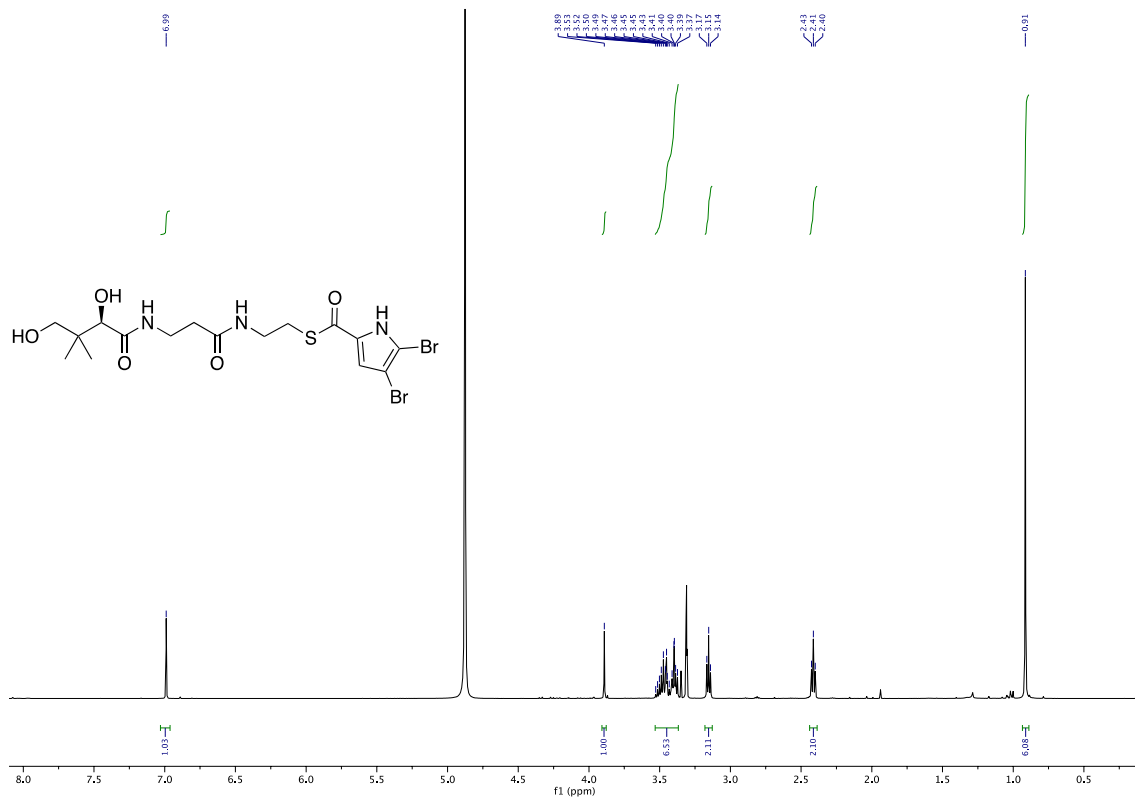
Figure S13: ^{1}H NMR spectrum (500 MHz, MeOD) of compound **10**
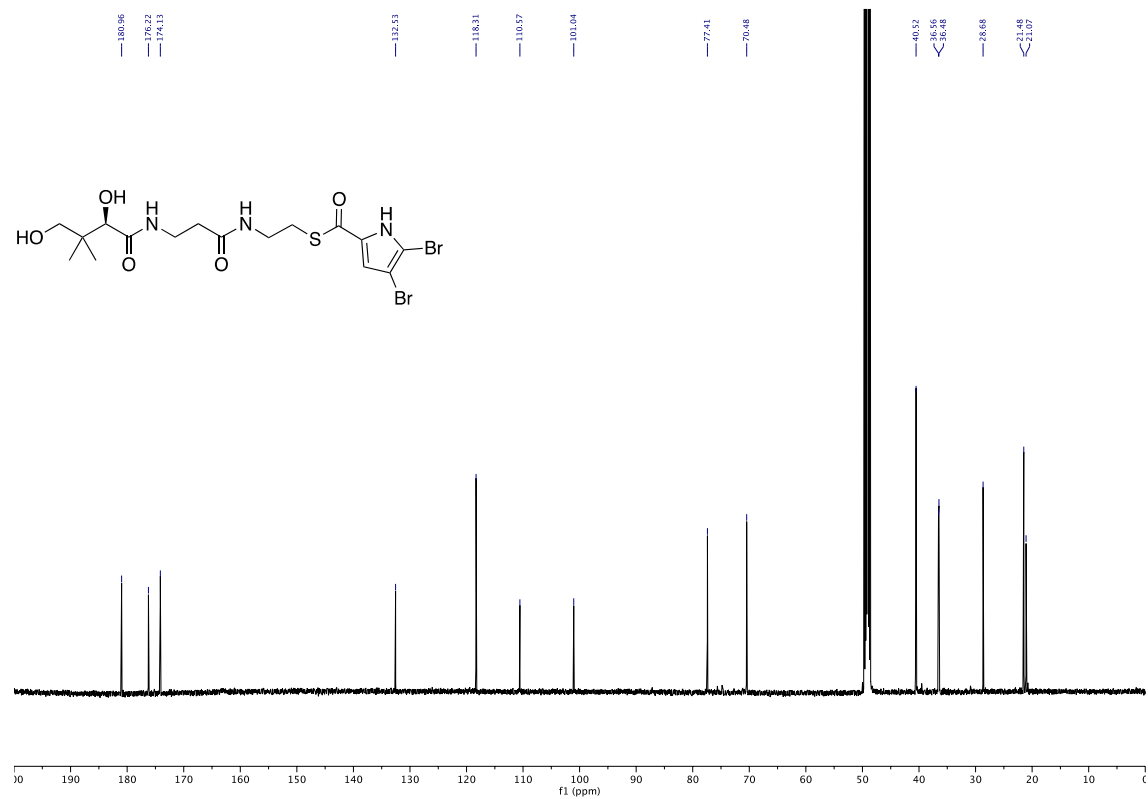
Figure S14: ^{13}C NMR spectrum (126 MHz, MeOD) of compound **10**
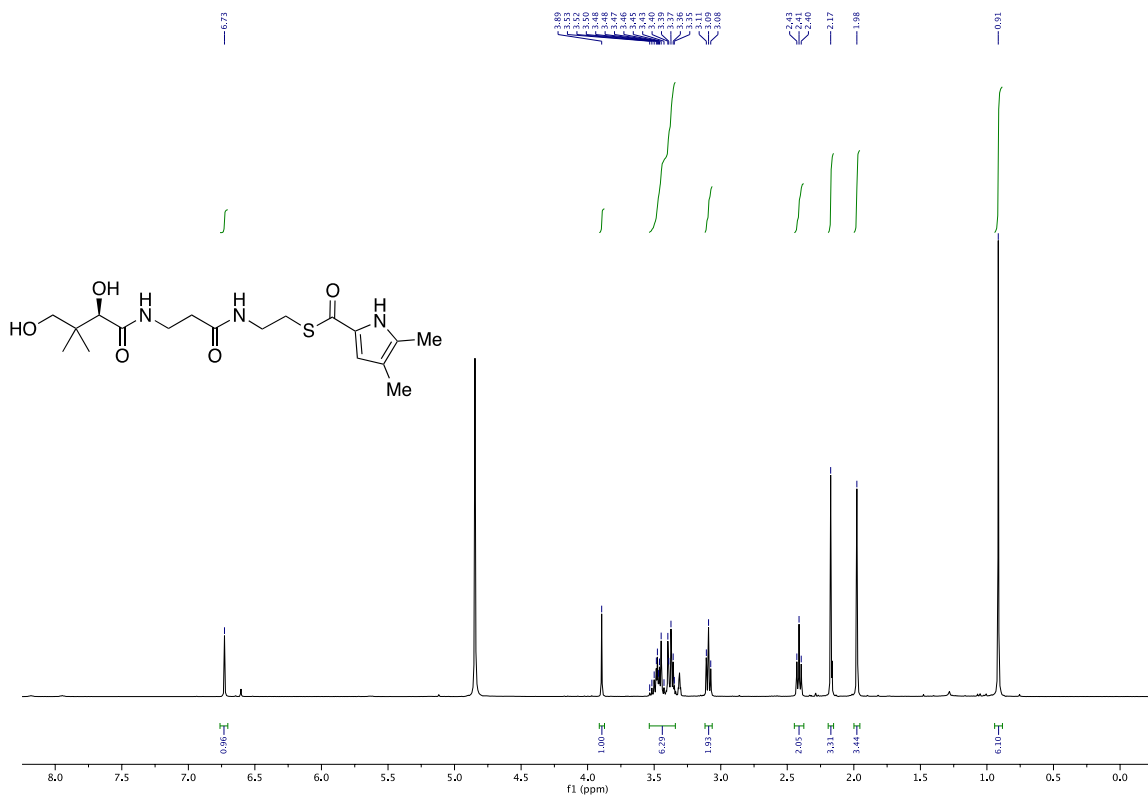
Figure S15: ^{1}H NMR spectrum (400 MHz, MeOD) of compound **11**
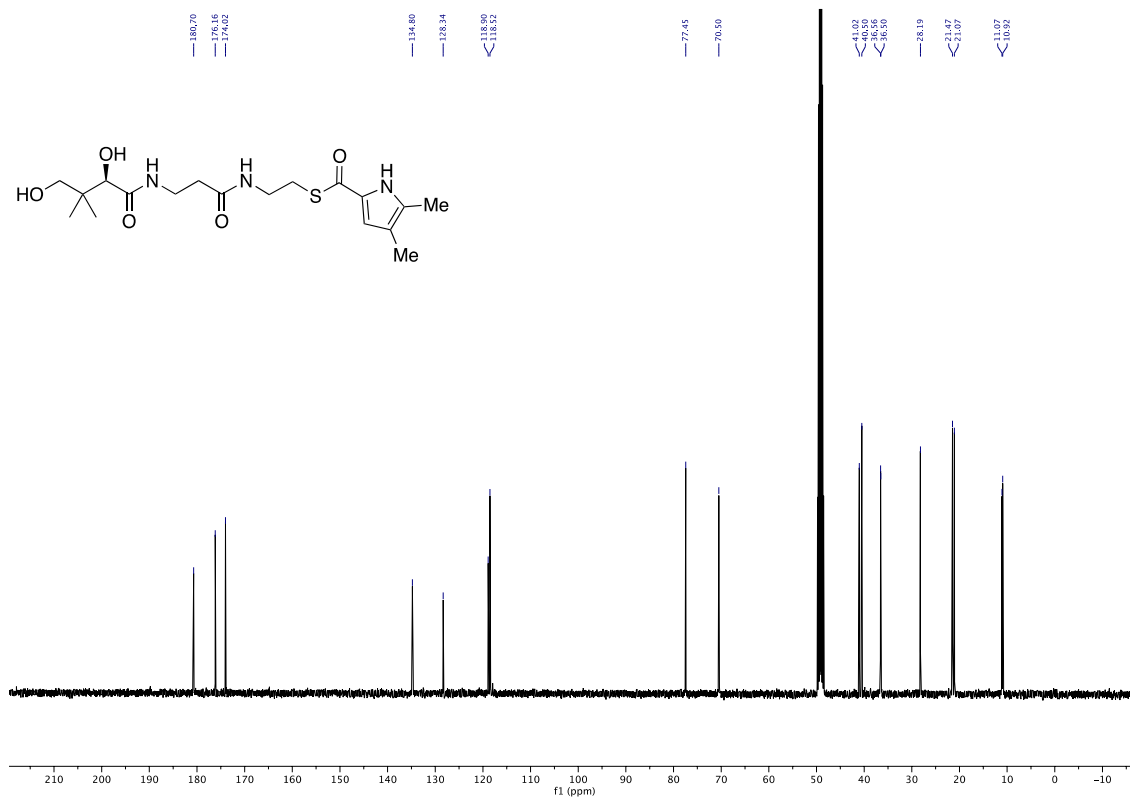
Figure S16: ^{13}C NMR spectrum (101 MHz, MeOD) of compound **11**
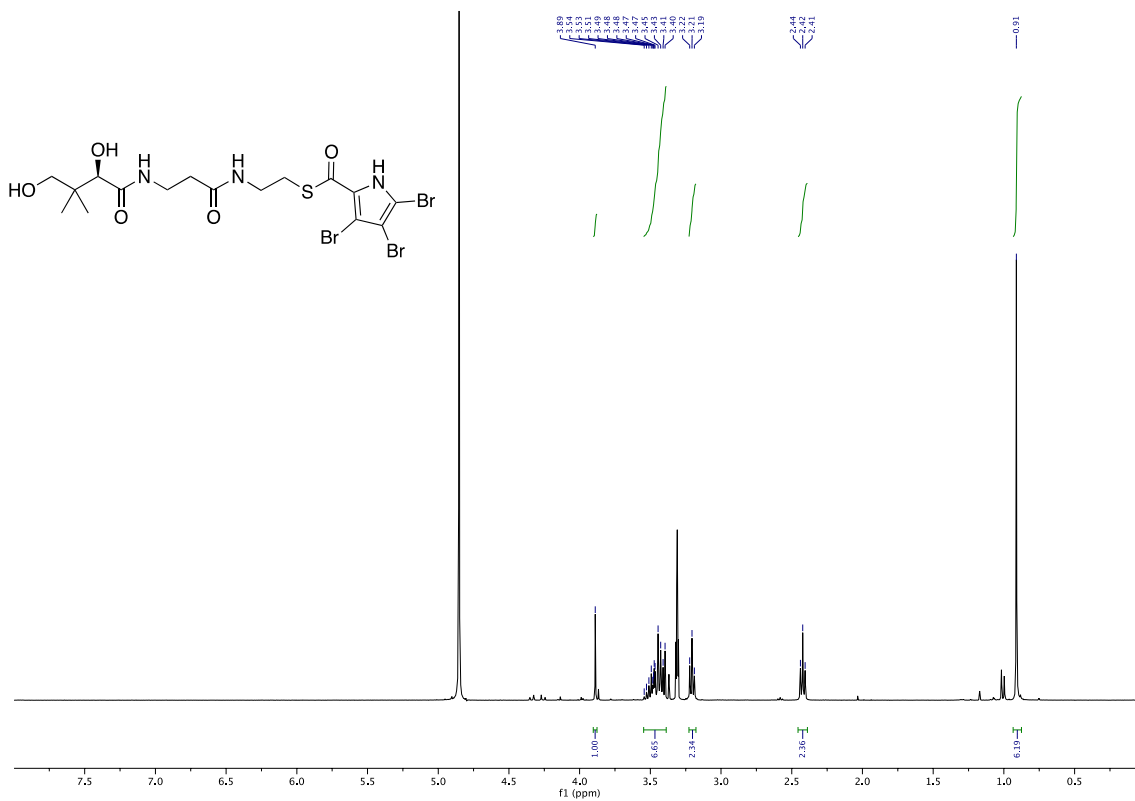
Figure S17: ^{1}H NMR spectrum (400 MHz, MeOD) of compound **12**
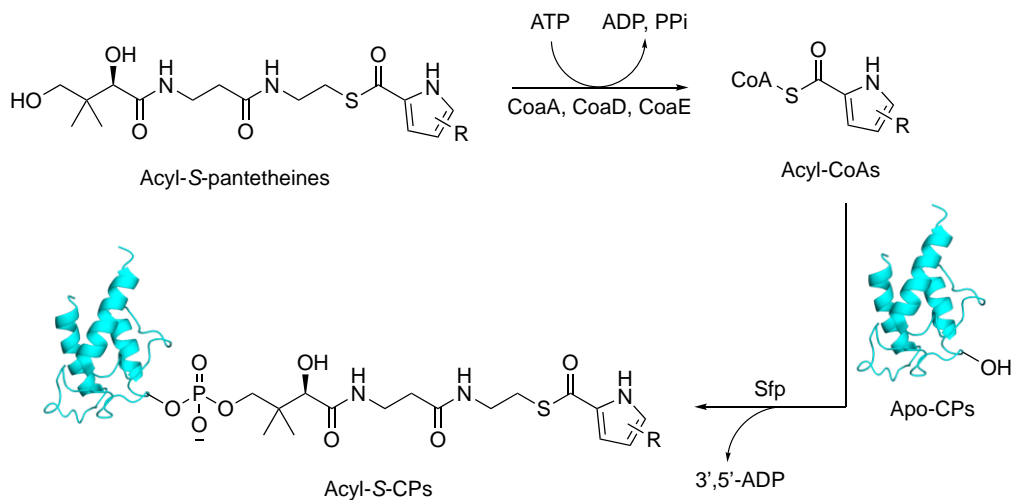
Figure S18: Reaction scheme for the one-pot enzymatic synthesis of pyrrolyl-*S*-CPs starting from pyrrolyl acyl-*S*-pantetheines using CoaA, CoaD, CoaE and Sfp enzymes.
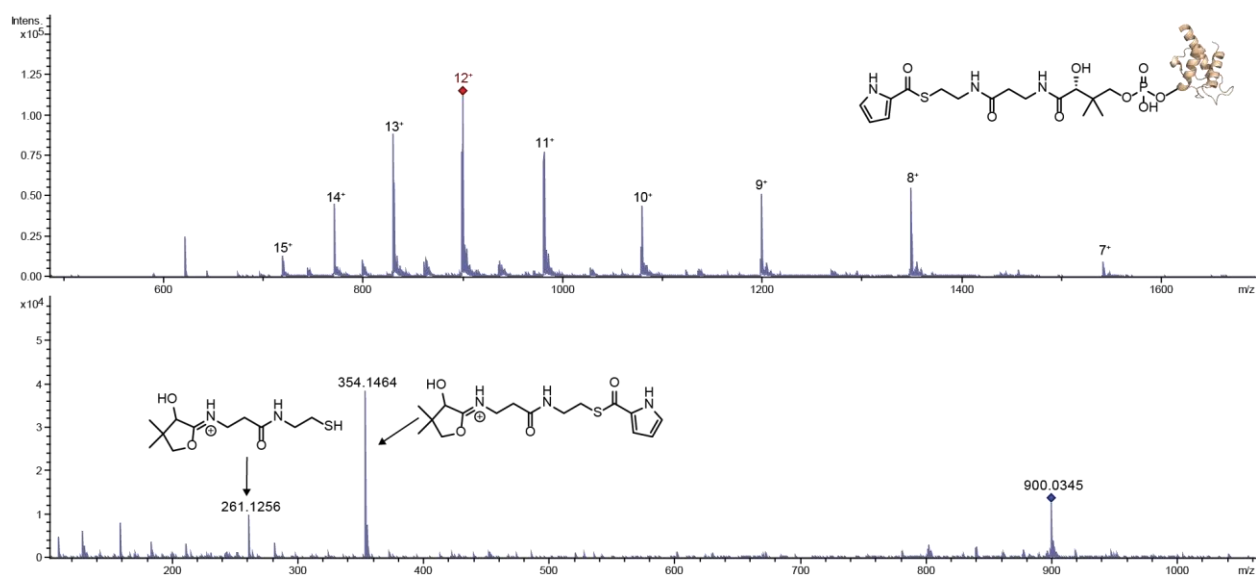
Figure S19: MS[1] (top) and MS[2] (bottom) spectra demonstrating pyrrolyl-*S*-PltL. In the MS[1] spectra, multiple charge states for peptides are observed. The [M+12H][12+] ion generates the characteristic pyrrolyl-*S*-cyclopantetheine MS[2] ejection ion (354.15 Da) and cyclopantetheine MS[2] ejection ion (261.13 Da).
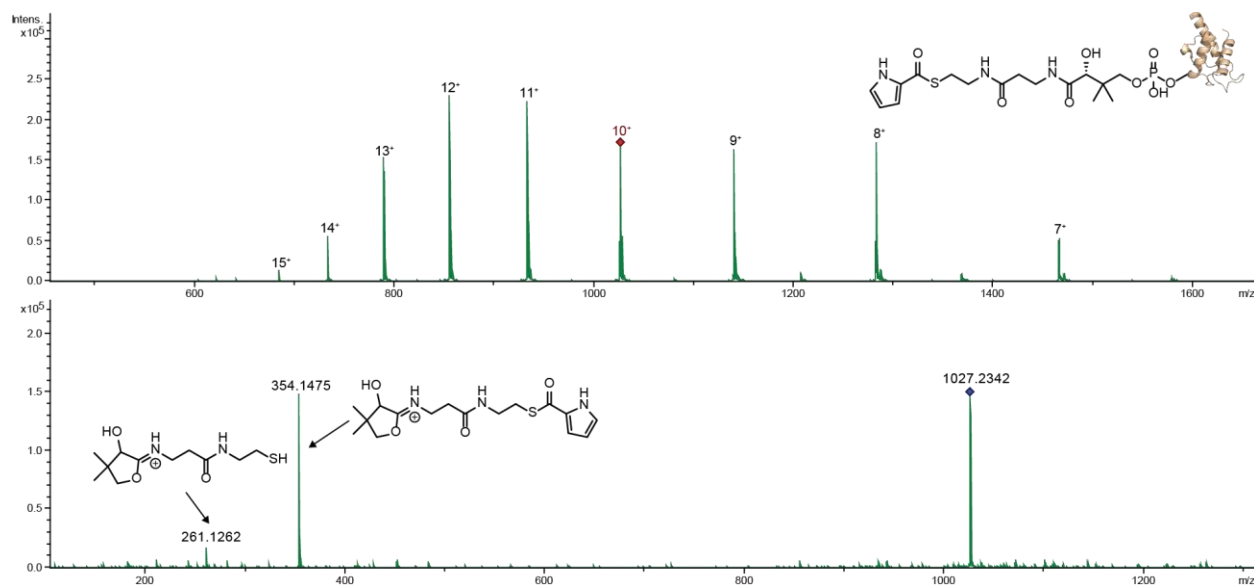
Figure S20: MS^1 (top) and MS^2 (bottom) spectra demonstrating 5-chloro-pyrrolyl-*S*-PltL. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic 5-chloro-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (388.11 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
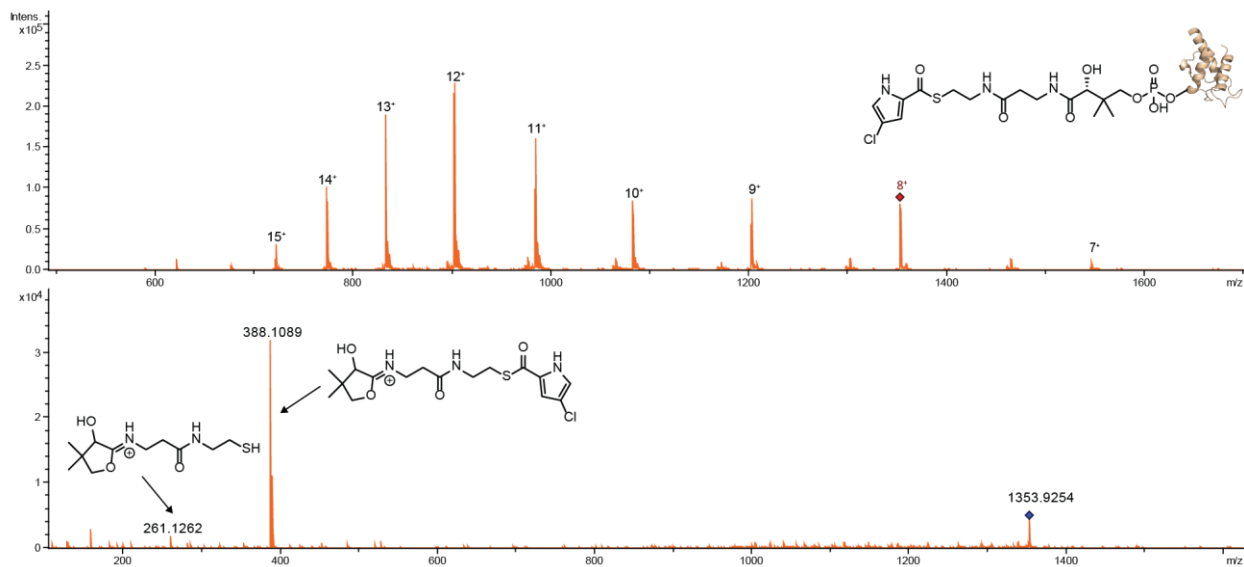
Figure S21: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4-chloro-pyrrolyl-*S*-PltL. In the MS^1 spectra, multiple charge states for peptides are observed. The $[M+8H]^{8+}$ ion generates the characteristic 4-chloro-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (388.11 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
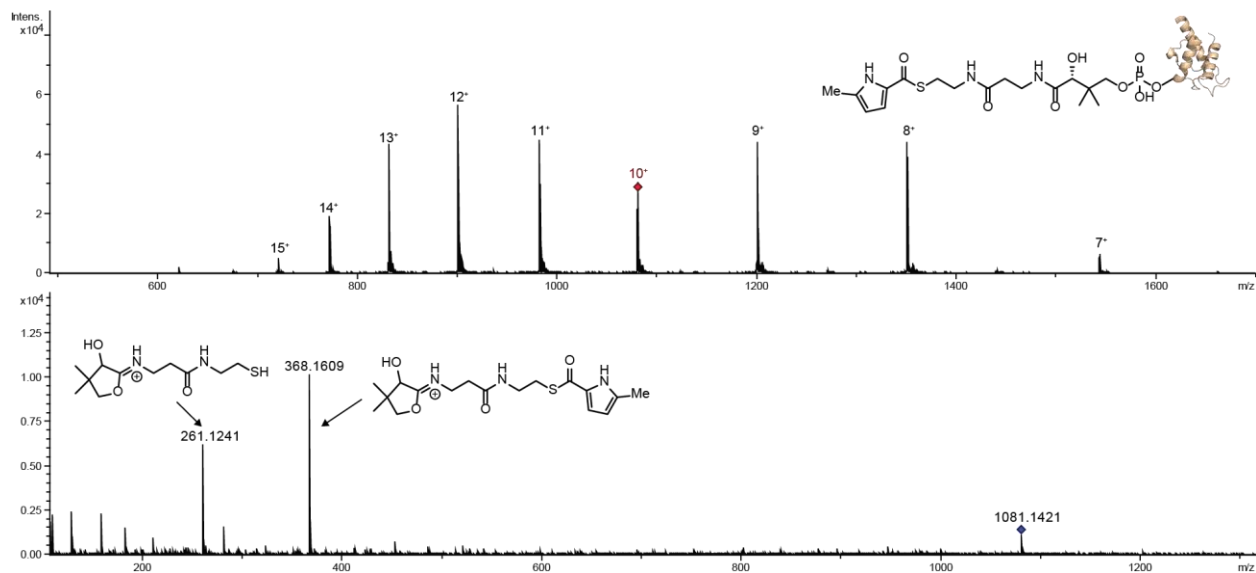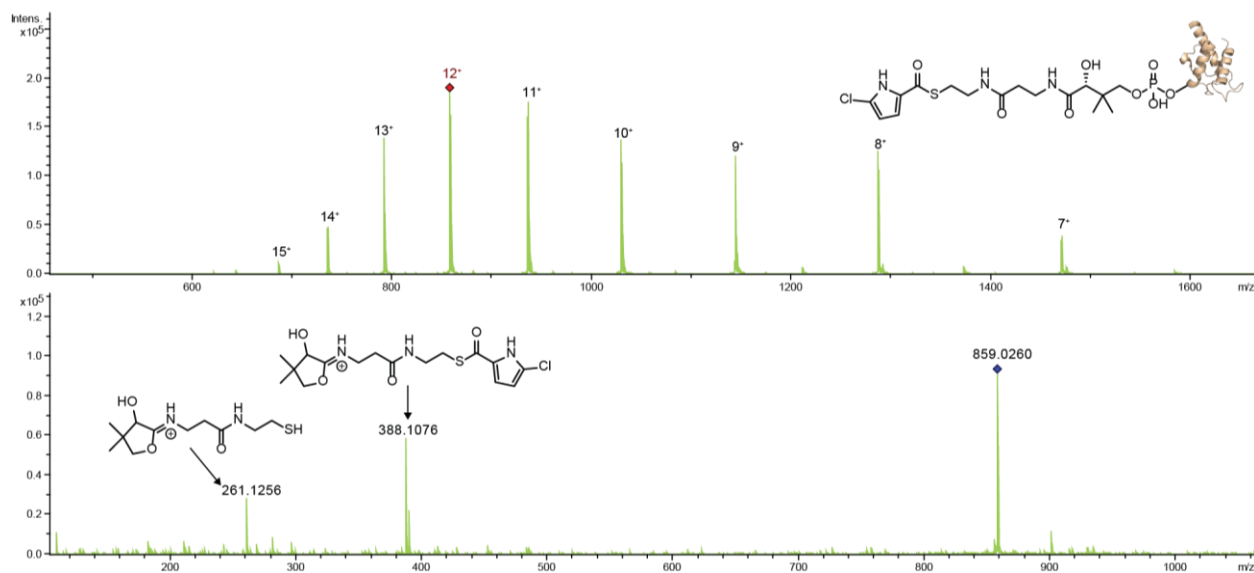
Figure S22: MS^1 (top) and MS^2 (bottom) spectra demonstrating 5-methyl-pyrrolyl-*S*-PltL. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+10H]$^{10+}$ ion generates the characteristic 5-methyl-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (368.16 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
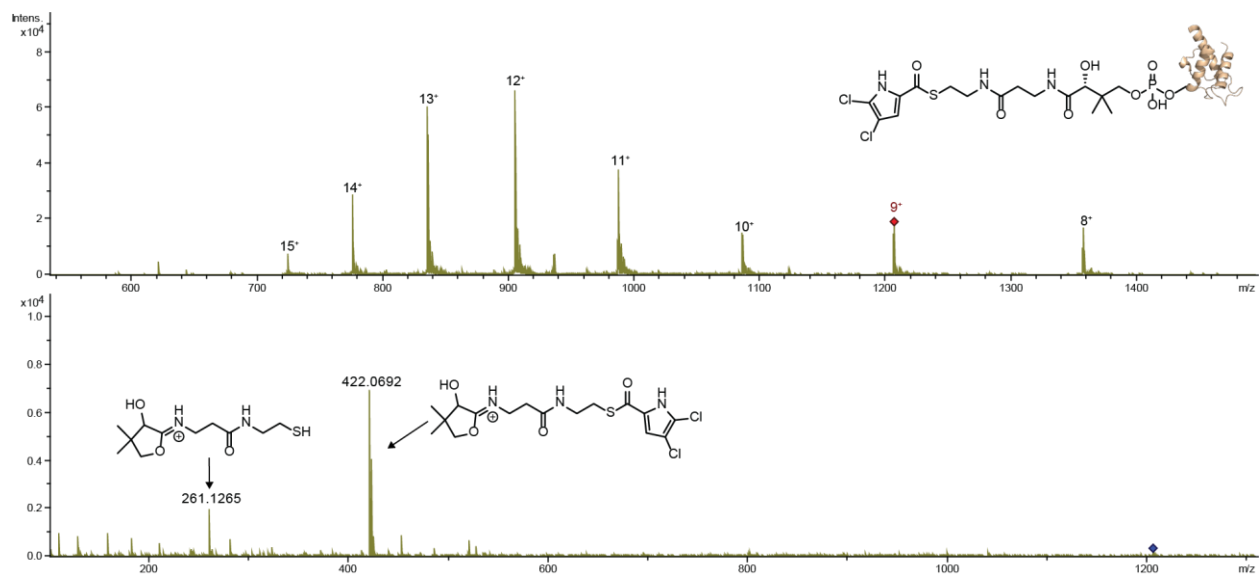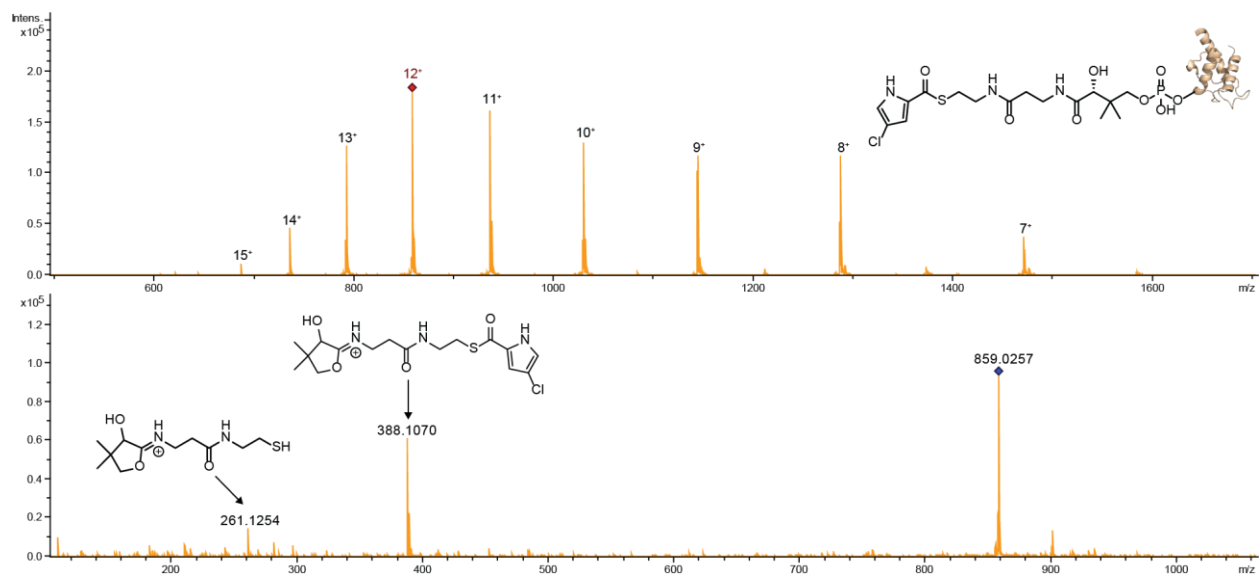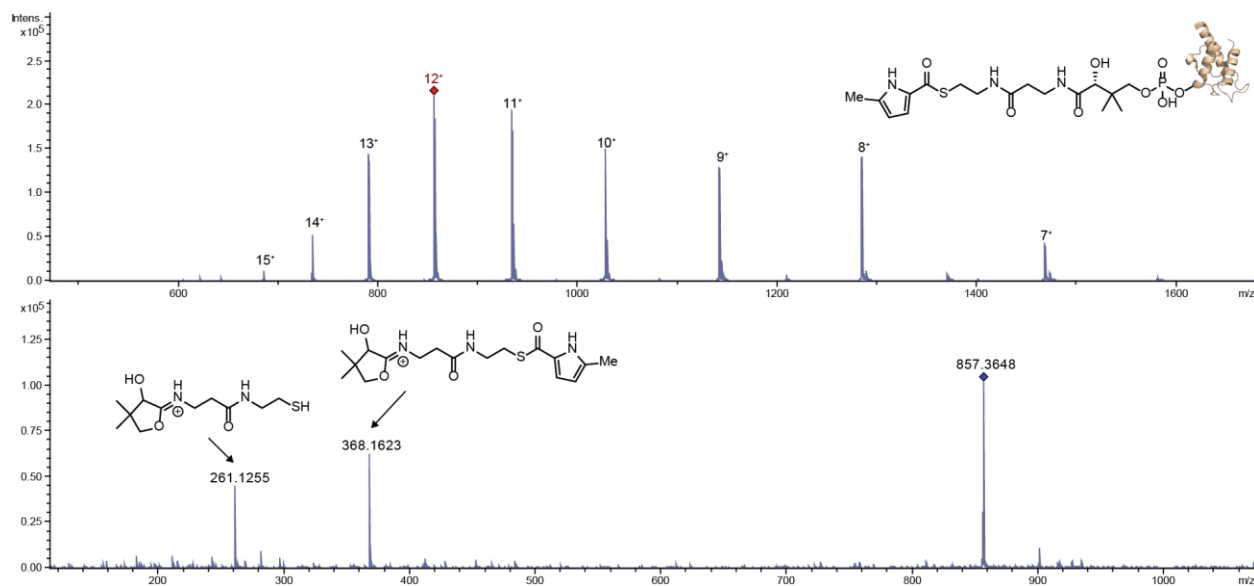
Figure S23: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dichloro-pyrrolyl-*S*-PltL. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+9H]$^{9+}$ ion generates the characteristic 4,5-dichloro-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (422.07 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
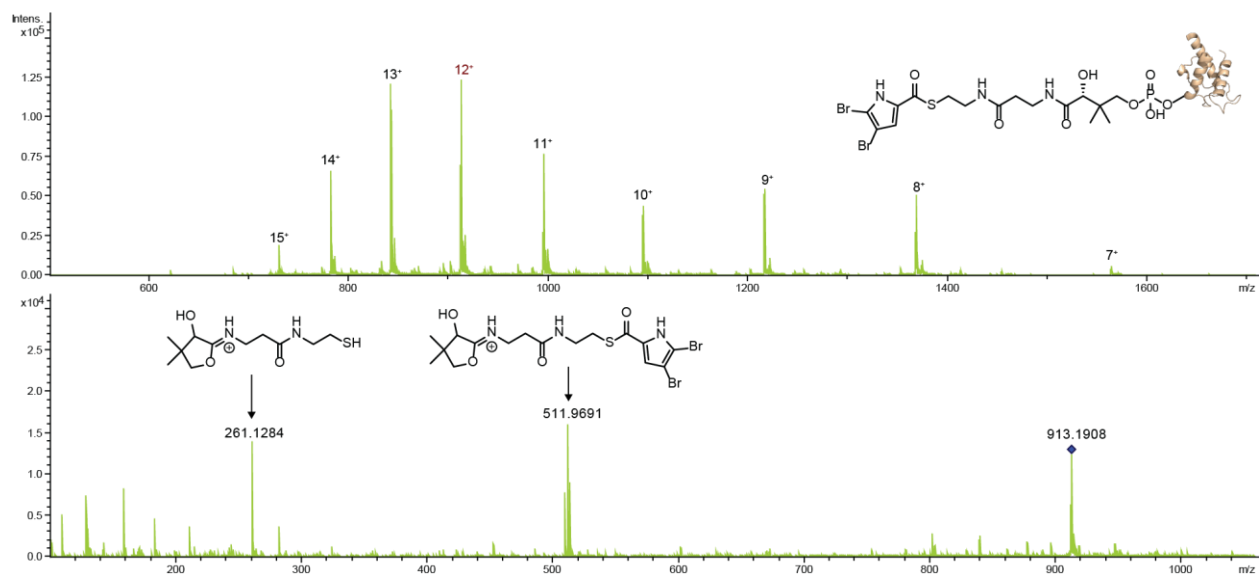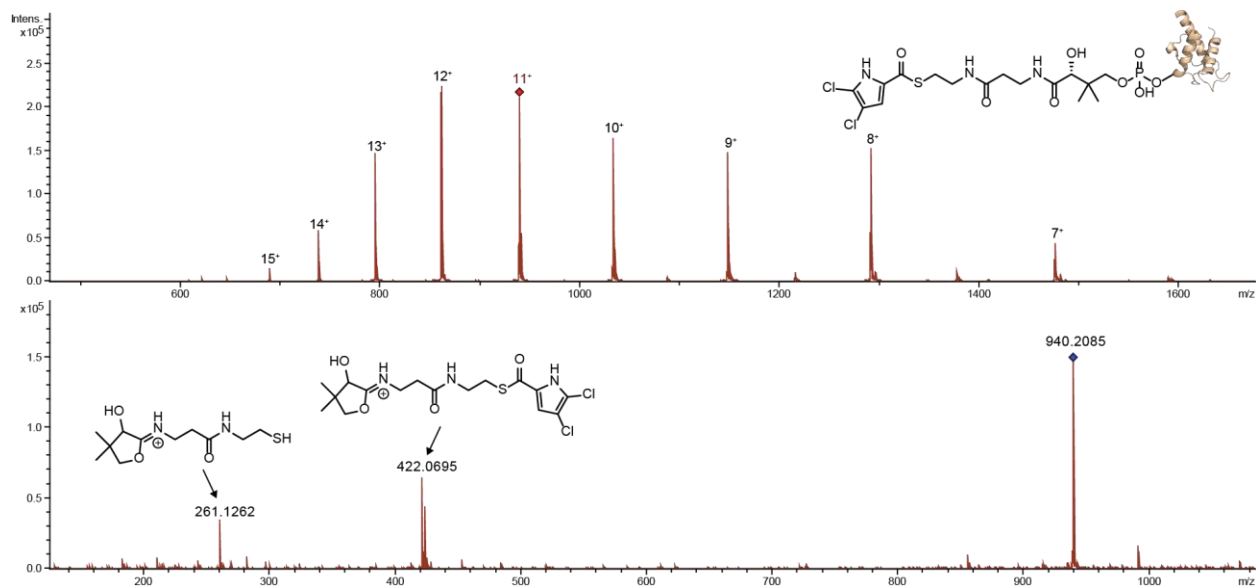
Figure S24: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dibromo-pyrrolyl-*S*-PltL. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic 4,5-dibromo-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (509.97 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
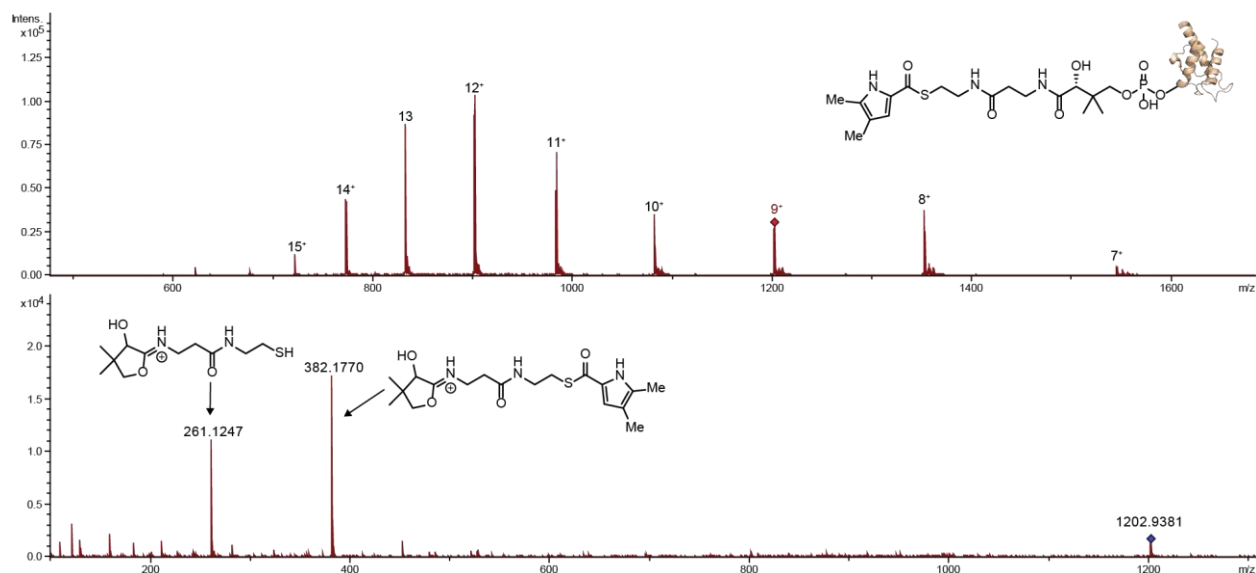
Figure S25: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dimethyl-pyrrolyl-*S*-PltL. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+9H]$^{9+}$ ion generates the characteristic 4,5-dimethyl-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (382.18 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
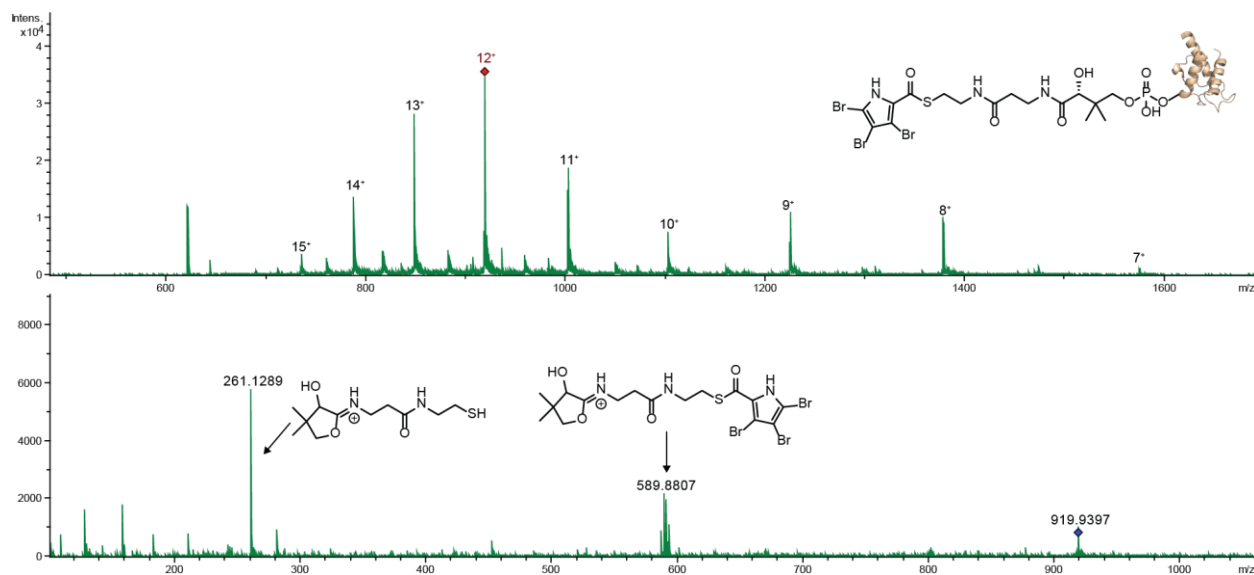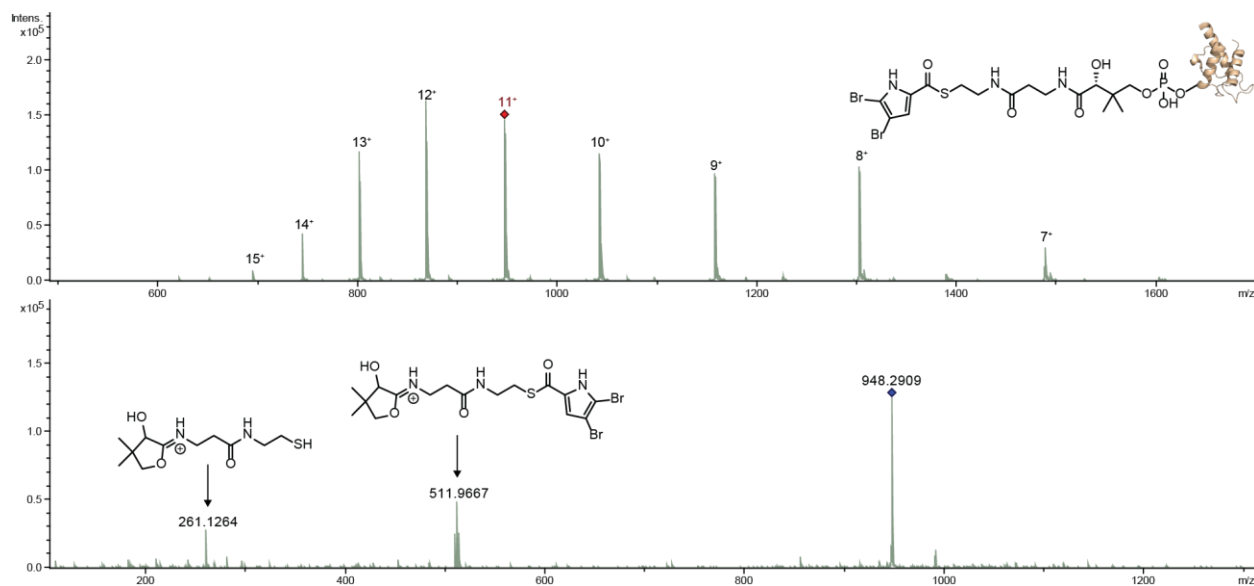
Figure S26: MS^1 (top) and MS^2 (bottom) spectra demonstrating 3,4,5-tribromo-pyrrolyl-*S*-PltL. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic 3,4,5-tribromo-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (587.88 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S27: MS^1 (top) and MS^2 (bottom) spectra demonstrating pyrrolyl-*S*-CalN3. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+10H]$^{10+}$ ion generates the characteristic pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (354.15 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
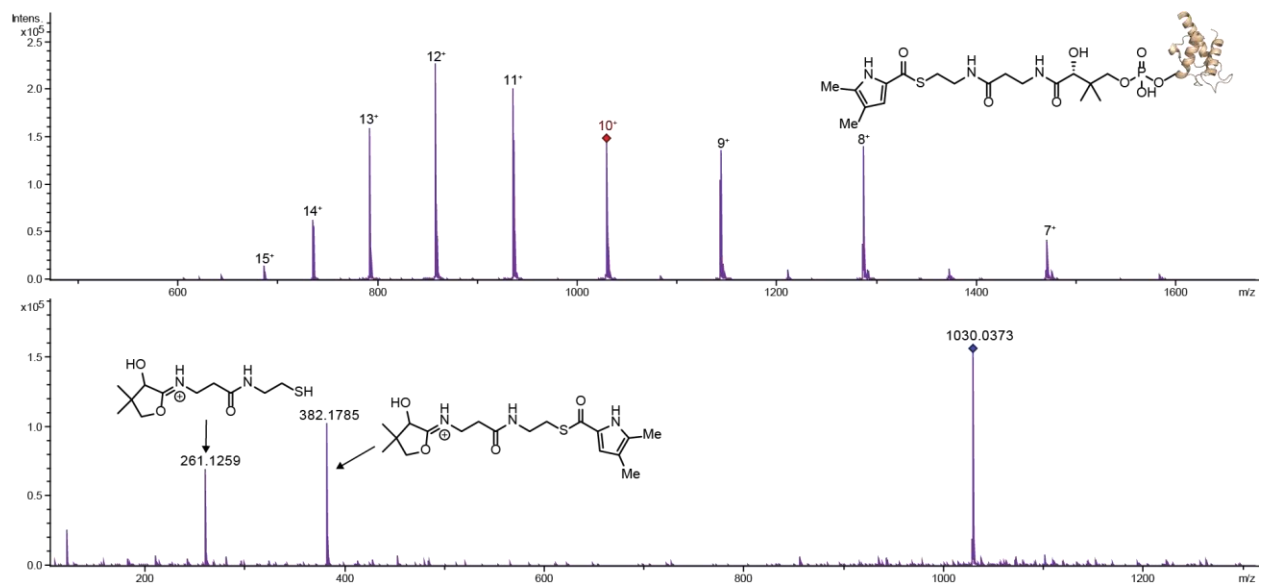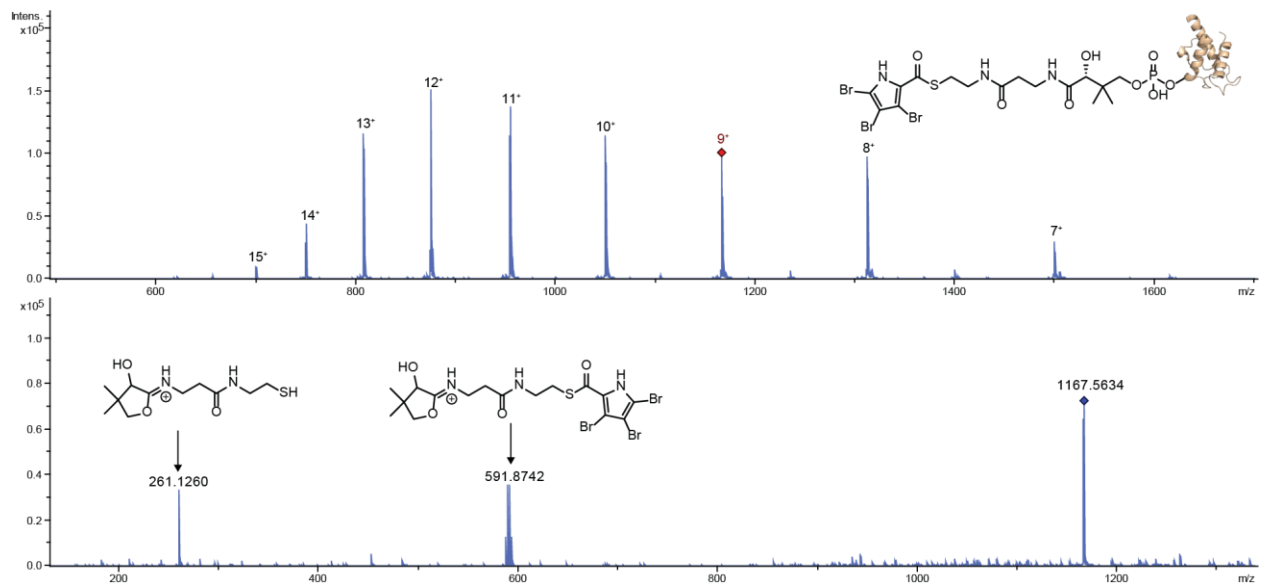
Figure S28: MS^1 (top) and MS^2 (bottom) spectra demonstrating 5-chloro-pyrrolyl-*S*-CalN3. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic 5-chloro-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (388.11 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S29: MS[1] (top) and MS[2] (bottom) spectra demonstrating 4-chloro-pyrrolyl-*S*-CalN3. In the MS[1] spectra, multiple charge states for peptides are observed. The $[M+12H]^{12+}$ ion generates the characteristic 4-chloro-pyrrolyl-*S*-cyclopantetheine MS[2] ejection ion (388.11 Da) and cyclopantetheine MS[2] ejection ion (261.13 Da).

Figure S30: MS^1 (top) and MS^2 (bottom) spectra demonstrating 5-methyl-pyrrolyl-*S*-CalN3. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic 5-methyl-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (368.16 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
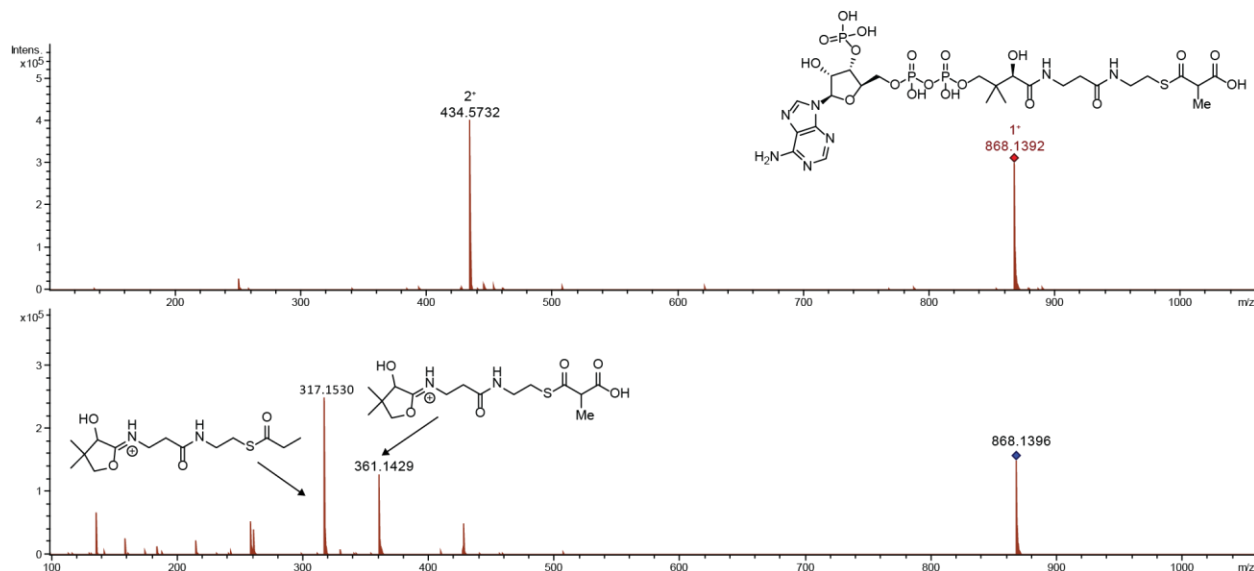
Figure S31: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dichloro-pyrrolyl-*S*-CalN3. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+11H]$^{11+}$ ion generates the characteristic 4,5-dichloro-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (422.07 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S32: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dibromo-pyrrolyl-S-CalN3. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+11H]$^{11+}$ ion generates the characteristic 4,5-dibromo-pyrrolyl-S-cyclopantetheine MS^2 ejection ion (509.97 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S33: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dimethyl-pyrrolyl-*S*-CalN3. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+10H]$^{10+}$ ion generates the characteristic 4,5-dimethyl-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (382.18 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S34: MS^1 (top) and MS^2 (bottom) spectra demonstrating 3,4,5-tribromo-pyrrolyl-*S*-CalN3. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+9H]$^{9+}$ ion generates the characteristic 3,4,5-tribromo-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (587.88 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S35: MS^1 (top) and MS^2 (bottom) spectra of methylmalonate CoA generated using enzyme MatB and methylmalonate (sodium salt), ATP/Mg^{2+}, and CoA-SH as substrates. In the MS^1 spectra, the [M+H]$^+$ (868.14 Da) and [M+2H]$^{2+}$ (434.57 Da) ions are observed that correspond to molecular formula C$_{25}$H$_{41}$N$_7$O$_{19}$P$_3$S. The [M+H]$^+$ ion generates the characteristic methylmalonyl-*S*-cyclopantetheine MS^2 ejection ion (361.14 Da) and decarboxylated methylmalonyl-*S*-cyclopantetheine MS^2 ejection ion (317.15 Da).
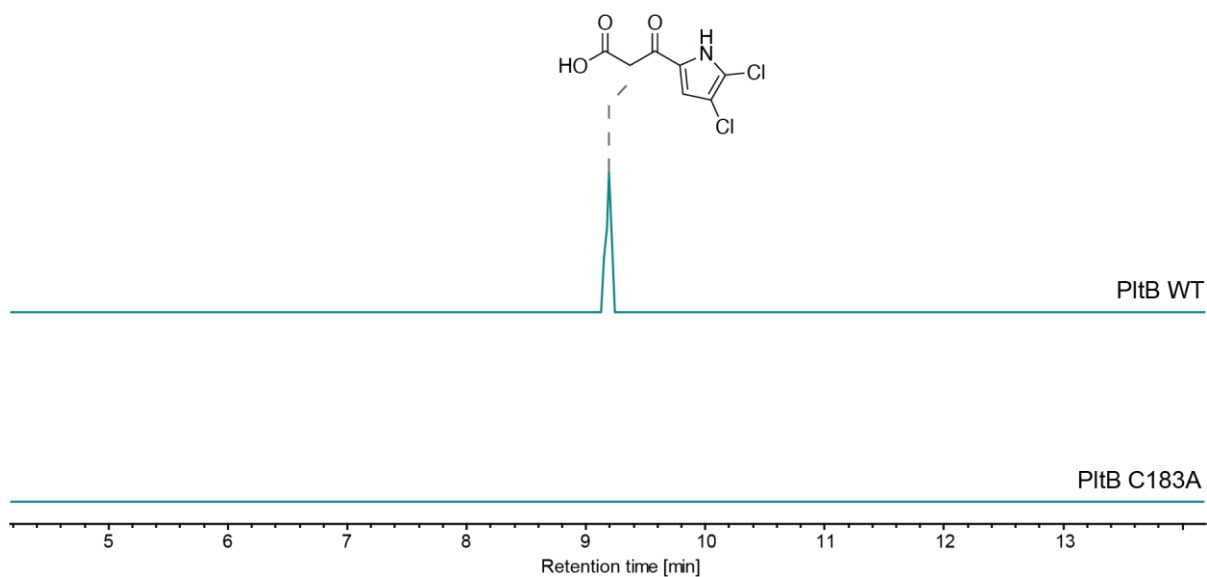
Figure S36: Elongated dichloro-pyrrolyl diketide from PltB assay analyzed by LC-MS. Dichloro-pyrrolyl-*S*-PltL was incubated with PltB WT or C183A mutant as well as malonylCoA. EIC for *m/z* 219.96 Da corresponding to [M-H]⁻ of elongated dichloro-pyrrolyl diketide was observed in the assay preformed with PltB WT, while EIC for *m/z* 219.96 was not detected in the assay with PltB C183A mutant.
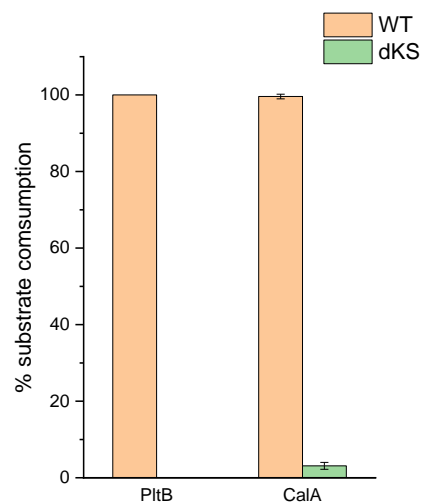
Figure S37: Comparison of substrate consumption by wild type and Cys-to-Ala mutant (dKS) PltB and CalA under identical reaction conditions. Physiological substrates were tested, dichloropyrrolyl-*S*-PltL for PltB and pyrrolyl-*S*-CalN3 for CalA. Lack of substrate consumption by the mutant enzymes is suggestive of the substrate stability and lack of substrate degradation during the course of the reaction.
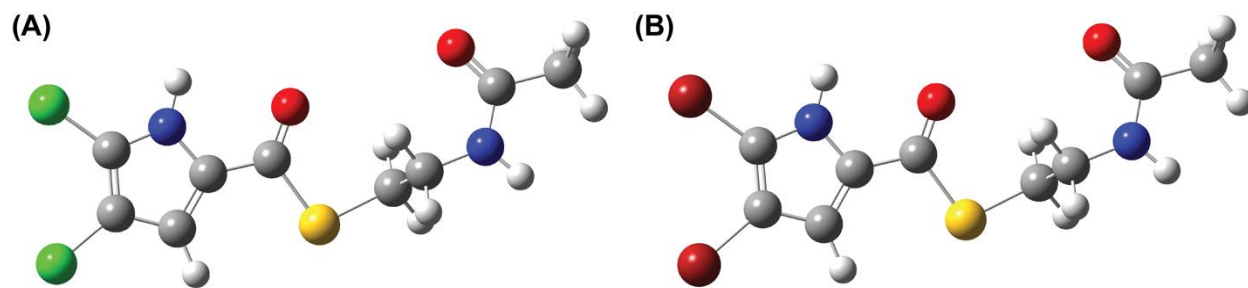
Figure S38: Optimized geometries of (**A**) dichloropyrrole, and (**B**) dibromopyrrole. Carbon, nitrogen, oxygen, sulfur, hydrogen, chlorine, and bromine atoms are shown in grey, blue, red, yellow, white, green, and brown spheres, respectively.
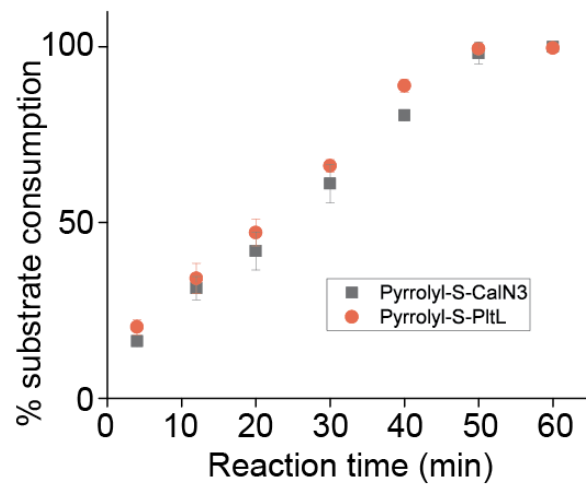
Figure S39: Time dependent depletion of pyrrolyl-*S*-CalN3 and pyrrolyl-*S*-PltL by CalA.
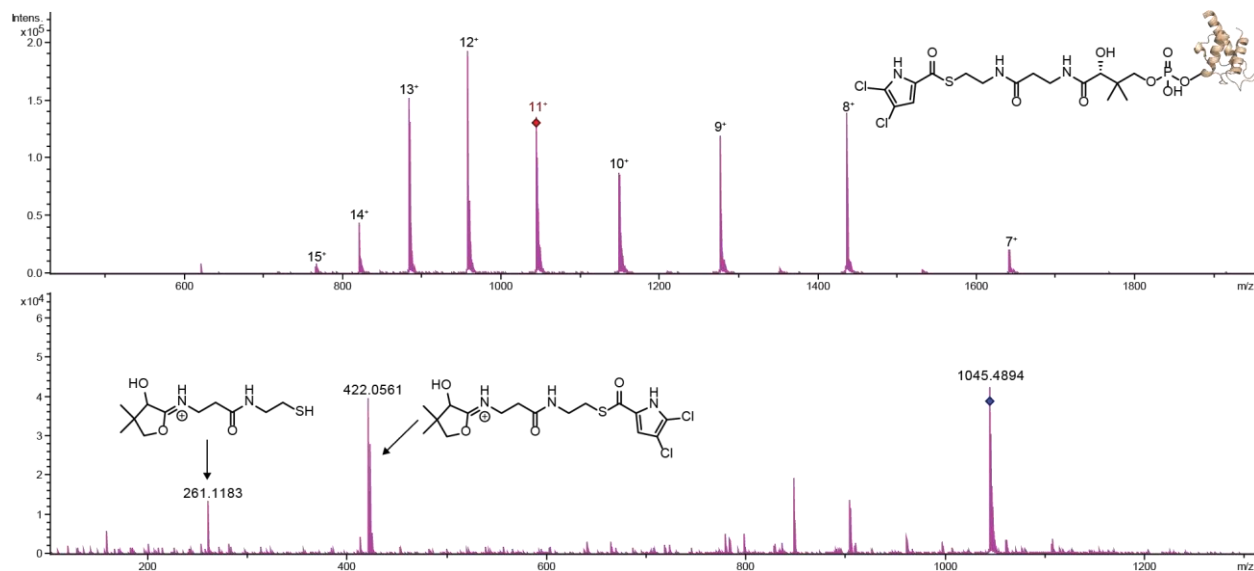
Figure S40: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dichloro-pyrrolyl-*S*-Clz18. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+11H]$^{11+}$ ion generates the characteristic 4,5-dichloro-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (422.07 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
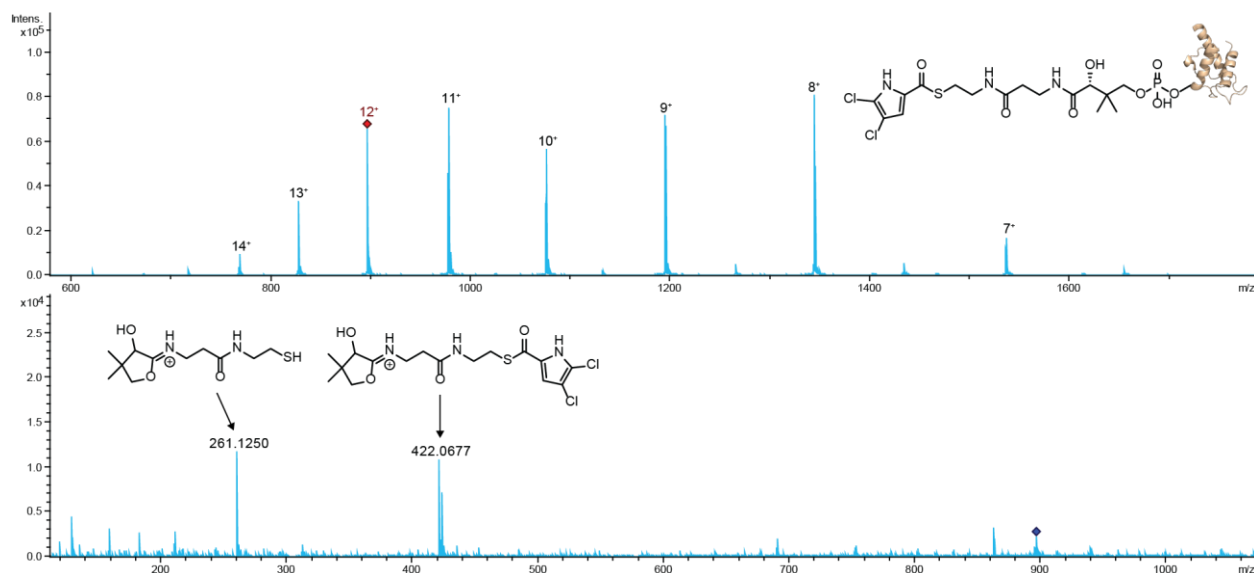
Figure S41: MS^1 (top) and MS^2 (bottom) spectra demonstrating 4,5-dichloro-pyrrolyl-*S*-Mpy15. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic 4,5-dichloro-pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (422.07 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
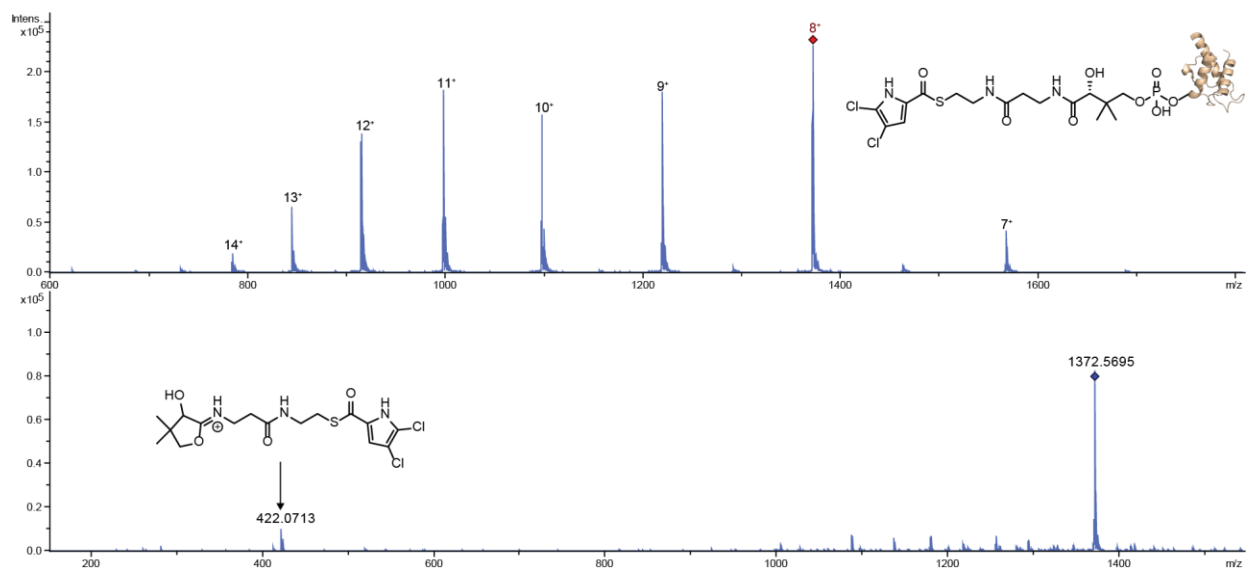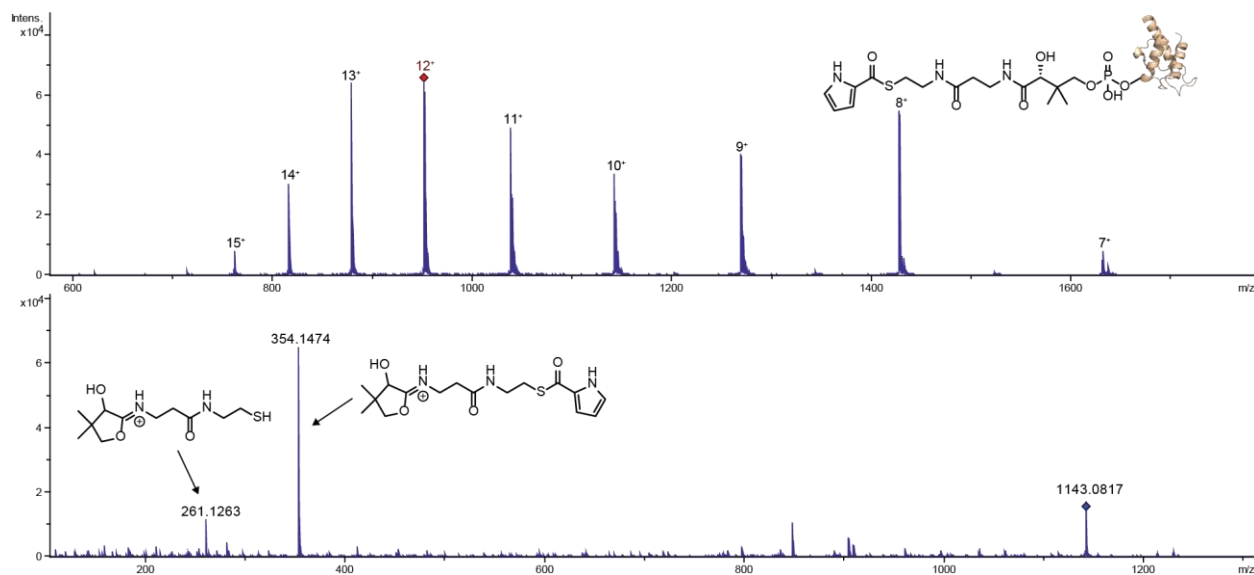
Figure S42: MS[1] (top) and MS[2] (bottom) spectra demonstrating 4,5-dichloro-pyrrolyl-*S*-HrmL. In the MS[1] spectra, multiple charge states for peptides are observed. The [M+8H][8+] ion generates the characteristic 4,5-dichloro-pyrrolyl-*S*-cyclopantetheine MS[2] ejection ion (422.07 Da) and cyclopantetheine MS[2] ejection ion (261.13 Da).
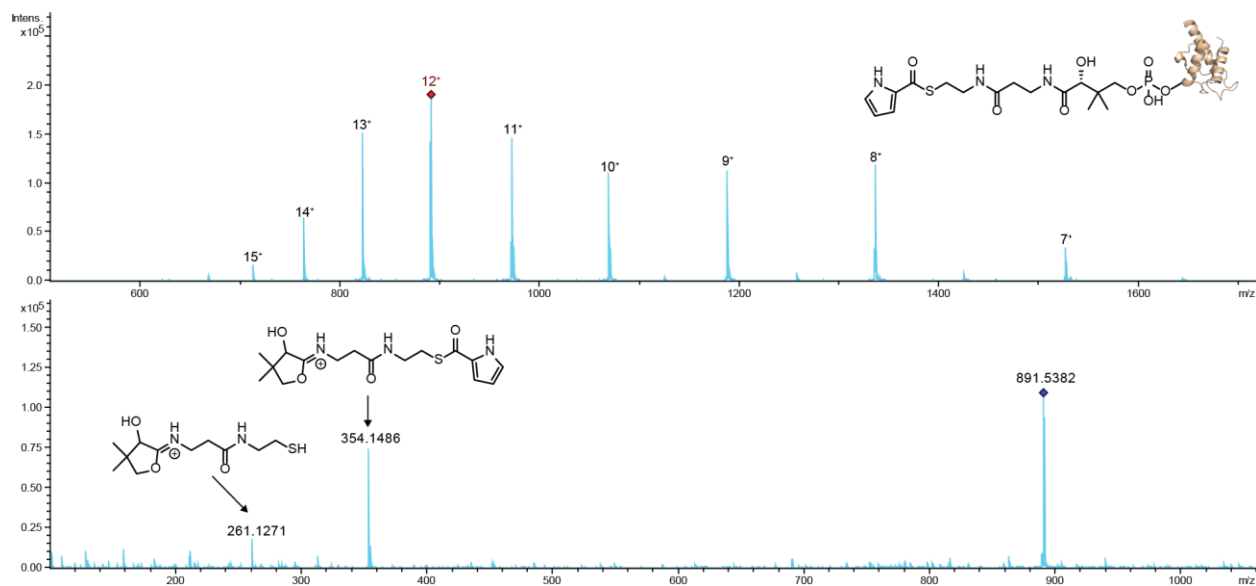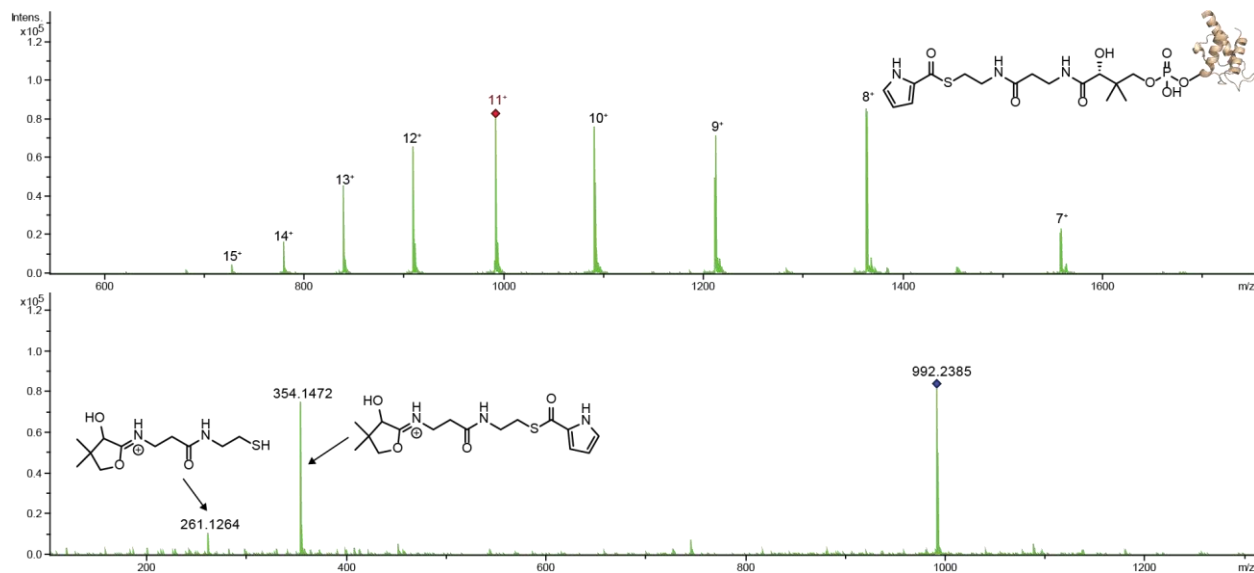
Figure S43: MS^1 (top) and MS^2 (bottom) spectra demonstrating pyrrolyl-*S*-Clz18. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (354.15 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S44: MS^1 (top) and MS^2 (bottom) spectra demonstrating pyrrolyl-*S*-Mpy15. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+12H]$^{12+}$ ion generates the characteristic pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (354.15 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).
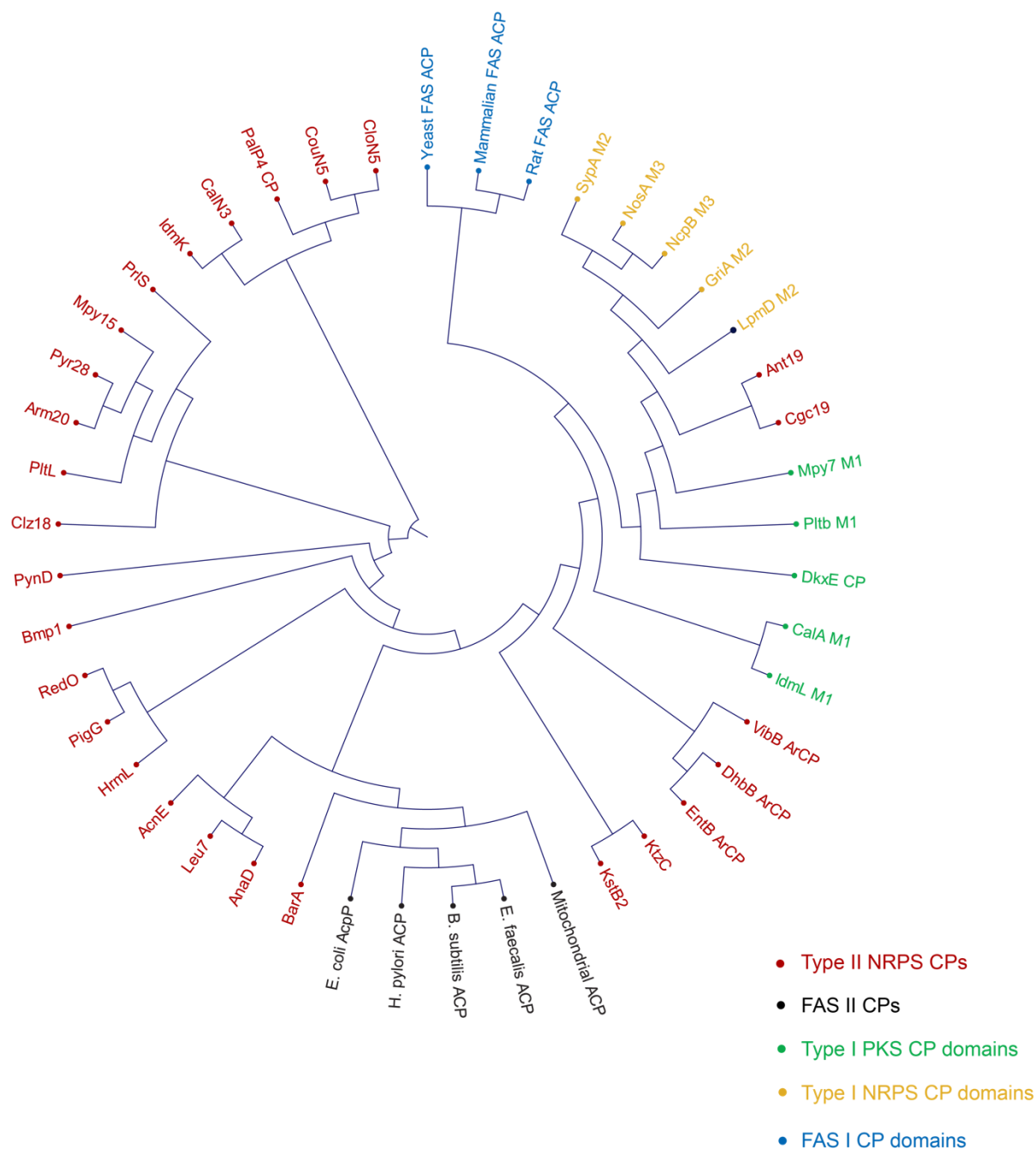
Figure S45: MS^1 (top) and MS^2 (bottom) spectra demonstrating pyrrolyl-*S*-HrmL. In the MS^1 spectra, multiple charge states for peptides are observed. The [M+11H]$^{11+}$ ion generates the characteristic pyrrolyl-*S*-cyclopantetheine MS^2 ejection ion (354.15 Da) and cyclopantetheine MS^2 ejection ion (261.13 Da).

Figure S46: Phylogenetic analysis of CPs. The amino acid sequences of CPs were aligned using CLC Sequence Viewer with gap open cost 10.0 and gap extension 1.0. The alignment was then used for phylogenetic analysis in which Neighbor Joining was used as tree constructing method and Jukes-Cantor as protein distance measurements.

References

1. Sapkota, K.; Huang, F., Efficient one-pot enzymatic synthesis of dephospho coenzyme A. *Bioorg Chem* **2018,** *76*, 23-27.

2. Thiede, S.; Wosniok, P. R.; Herkommer, D.; Debnar, T.; Tian, M.; Wang, T.; Schrempp, M.; Menche, D., Total synthesis of leupyrrins A1 and B1, highly potent antifungal agents from the myxobacterium *Sorangium cellulosum*. *Chem Eur J* **2017,** *23* (14), 3300-3320.

3. Wang, M. Z.; Xu, H.; Liu, T. W.; Feng, Q.; Yu, S. J.; Wang, S. H.; Li, Z. M., Design, synthesis and antifungal activities of novel pyrrole alkaloid analogs. *Eur J Med Chem* **2011,** *46* (5), 1463-1472.

4. Agarwal, V.; Diethelm, S.; Ray, L.; Garg, N.; Awakawa, T.; Dorrestein, P. C.; Moore, B. S., Chemoenzymatic synthesis of acyl coenzyme A substrates enables *in situ* labeling of small molecules and proteins. *Org Lett* **2015,** *17* (18), 4452-4455.

5. Thapa, H. R.; Lin, Z.; Yi, D.; Smith, J. E.; Schmidt, E. W.; Agarwal, V., Genetic and biochemical reconstitution of bromoform biosynthesis in *Asparagopsis* lends insights into seaweed reactive oxygen species enzymology. *ACS Chem Biol* **2020,** *15* (6), 1662-1670.

6. Hughes, A. J.; Keatinge-Clay, A., Enzymatic extender unit generation for *in vitro* polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem Biol* **2011,** *18* (2), 165-176.

7. Chai, J.-D.; Head-Gordon, M., Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys Chem Chem Phys* **2008,** *10* (44), 6615-6620.

8. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V., Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J Comput Chem* **2003,** *24* (6), 669-681.

9. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.;

Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

10. Dennington, R.; Keith, T. A.; Millam, J. M. *GaussView Ver 6*, Shawnee Mission, KS, 2016.