CHAPTER TEN

Chemoenzymatic Synthesis of Starting Materials and Characterization of Halogenases Requiring Acyl Carrier Protein-Tethered Substrates

Hem R. Thapa*, Andrew J. Lail*, Neha Garg*, Vinayak Agarwal*,†,1

*School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA, United States
†School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, United States
[1]Corresponding author: e-mail address: vagarwal@gatech.edu

Contents

Abstract

Flavin-adenine dinucleotide (FAD)-dependent halogenases are widespread in natural product biosynthetic gene clusters and have been demonstrated to employ small organic molecules as substrates for halogenation, as well as substrates that are tethered to carrier proteins (CPs). Despite numerous reports of FAD-dependent halogenases utilizing CP-tethered substrates, only a few have been biochemically characterized due to limited accessibility to the physiological substrates. Here, we describe a method for the preparation of acyl-*S*-CP substrates and their use in biochemical assays to query the activity of FAD-dependent halogenases. Furthermore, we describe a mass spectrometry-based method for the characterization of acyl-*S*-CP substrates and the corresponding halogenated products generated by the halogenases. Finally, we test

Methods in Enzymology, Volume 604
ISSN 0076-6879
https://doi.org/10.1016/bs.mie.2018.01.028

the substrate specificity of a physiological chlorinase and a physiological brominase from marine bacteria, and, for the first time, demonstrate the distinct halide specificity of halogenases. The methodology described here will enable characterization of new halogenases employing CP-tethered substrates.

1. INTRODUCTION

Halogenation is a modification that is routinely afforded upon natural products and often confers unique biological activities to these organic molecules, including, but not limited to desirable pharmaceutical antibiotic and cytotoxic activities. Bioactivities of small molecules can be critically dependent on halogenation, such as for the antibiotics vancomycin (**1**, Fig. 1) and chloramphenicol (**2**). Four principal classes of halogenating enzymes have been described to date, their mechanisms, and activities tailored to halogenate diverse substrates varying from electron-rich aromatic rings to unactivated alkyl centers (Agarwal et al., 2017). Of the four classes, this manuscript describes procedures to assay the activity of halogenases that employ flavin-adenine dinucleotide (FAD) as a cofactor. FAD-dependent halogenases oxidize halides (X^-, X = Cl, Br, and I) to their corresponding halenium (X^+) ions that are regiospecifically delivered either as a haloamines or as hypohalites to electron-rich centers to catalyze electrophilic substitution reactions (Flecks et al., 2008; Yeh, Blasiak, Koglin, Drennan, & Walsh, 2007). FAD-dependent halogenases require a partner enzyme, flavin

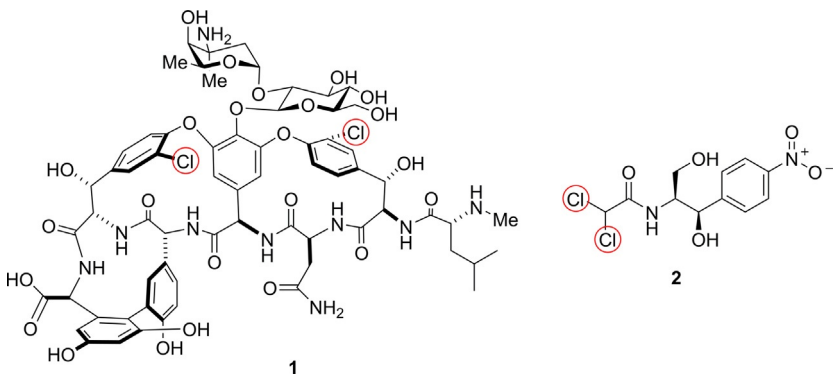

Fig. 1 Chemical structures of halogenated antibiotics **1** and **2**, biosynthesis of which involves FAD-dependent halogenases that employ CP-tethered substrates.

reductase, to deliver reduced FAD ($FADH_2$) that is required for halogenase activity. Electrons to generate $FADH_2$ from FAD are derived from NAD(P)H. Molecular oxygen acts as the final sink for these electrons, together with the electrons derived from the oxidation of the halide. The catalytic scheme for FAD-dependent halogenases is shown in Fig. 2.

The inventory of halogenated natural products involving the participation of FAD-dependent halogenases in their respective biosynthetic schemes is already impressive and continues to expand. Furthermore, characterization of their biochemical activities has advanced FAD-dependent halogenases to a stage where they can be successfully used as diagnostic elements to identify natural product biosynthetic gene clusters in genomic and metagenomics datasets (Agarwal, Blanton, et al., 2017; Agarwal, Miles, et al., 2017; Jordan & Moore, 2016). Enzymatic functionalization of carbon centers by halogenation has attracted attention toward the engineering of FAD-dependent halogenases to improve their thermal and solvent tolerance, to expand their substrate scope, and to use these enzymes in a combinatorial fashion to engineer natural product scaffolds (Brown & O'Connor, 2015; Latham, Brandenburger, Shepherd, Menon, & Micklefield, 2018; Menon et al., 2017; Payne, Andorfer, & Lewis, 2016;

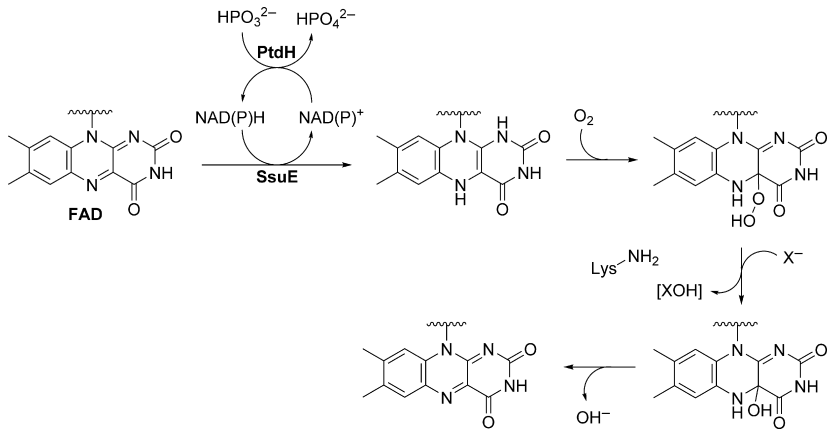

Fig. 2 Catalytic mechanism for FAD-dependent halogenases. In this study, the *Escherichia coli* flavin reductase SsuE is used to deliver $FADH_2$ to the halogenases while the phosphite dehydrogenase PtdH is used to regenerate NADPH (see Section 4.1 for experimental details). The identity of the electrophilic halogenating species generated by FAD-dependent halogenases, hypohalite ([XOH]), or a haloamine generated by the capture of the halenium by the side chain amine of a conserved Lys residue in the halogenase active site is presently not clear.

Schnepel, Minges, Frese, & Sewald, 2016). The efforts alluded to above are principally directed toward FAD-dependent halogenases that use small-molecule organic substrates, such as L-Trp. Much of our current mechanistic understanding for FAD-dependent halogenases relies on the crystal structures of and biochemical studies performed using L-Trp chlorinases (Agarwal, Miles, et al., 2017).

In contrast, the pace of characterization of FAD-dependent halogenases that employ substrates tethered to carrier proteins (CPs) has been sluggish. For instance, despite the identification of the biosynthetic gene cluster and establishment of aryl-halogenation as being critical to antibiotic efficacy, reconstitution of the bioactivity of the FAD-dependent halogenase VhaA participating in the biosynthesis of **1** has only recently been achieved (Pinchman & Boger, 2013a, 2013b; Schmartz, Zerbe, Abou-Hadeed, & Robinson, 2014; van Wageningen et al., 1998). As a second example, while the biosynthetic gene cluster has been identified and the crystal structure of the enzyme determined, a firm biochemical characterization of the FAD-dependent halogenase CmlS involved in the biosynthesis of **2** is still elusive (He, Magarvey, Piraee, & Vining, 2001; Podzelinska et al., 2010). Two challenges precipitate this knowledge gap. First, the timing of halogenation relative to other biochemical modifications in multimodular schemes involving CP-tethered substrates can be hard to decipher. Second, synthesis of CP-tethered substrates in itself can be challenging as is described in detail later.

CPs are small proteins with a conserved four-helical bundle structure. The side chain hydroxyl of a conserved serine residue in apo-CPs is post-translationally esterified to a phosphopantetheine "arm" derived from coenzyme A (CoA-SH), a transformation afforded by phosphopantetheinyl transferases to yield holo-CPs (Fig. 3) (Beld, Sonnenschein, Vickery, Noel, & Burkart, 2014). Following esterification with phosphopantetheine, the terminal thiol of holo-CPs is condensed with adenylated carboxylic acids in an ATP-dependent ligation reaction to yield acyl-*S*-CPs. Acyl-*S*-CPs serve to carry carboxylic acids through reaction schemes while the carboxylic acid is activated by the formation of the high-energy thioester bond with the phosphopantetheine arm. Following modifications of the acyl moiety, the thioester bond can be displaced by various nucleophiles, by water to yield carboxylic acids, by alcohols to yield esters and lactones, by amines to yield amides and lactams, by carbanions to yield Claisen condensation products, or transthioesterified. These modifications are extensively represented in primary metabolism and in natural product biosynthetic pathways.

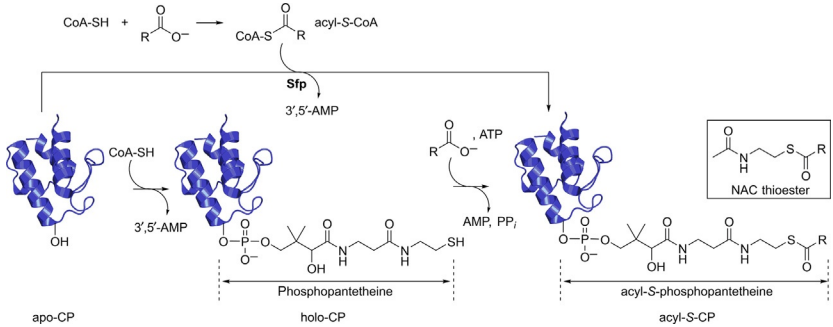

Fig. 3 Posttranslational maturation of apo-CP by transfer of the phosphopantetheine arm from CoA-SH to the side chain hydroxyl of a conserved serine residue to yield holo-CP, followed by formation of the thioester bond with a carboxylic acid to generate acyl-*S*-CPs. In some cases, acyl-*S*-CP substrates can be substituted by abbreviated NAC thioesters. An alternate route to generating acyl-*S*-CPs relies on the transfer of the acyl-*S*-phosphopantetheine from acyl-S-CoAs, chemically synthesized by the ligation of CoA-SH with carboxylic acids, by the phosphopantetheinyl transferase Sfp. The *E. coli* CP (PDB: 4ETW) is shown in cartoon representation as a representative example of the tertiary structure of CPs and the site for the conserved Ser residue (Agarwal, Lin, Lukk, Nair, & Cronan, 2012).

CPs, when expressed as recombinant proteins in heterologous hosts such as *Escherichia coli*, are posttranslationally modified in situ to their corresponding holo forms by phosphopantetheinyl transferases that are present in the heterologous host's proteome. However, these conversions usually do not reach completion and heterologously expressed CPs are purified as a mixture of apo and holo states. The discovery and characterization of the broadly substrate tolerant phosphopantetheinyl transferase Sfp has provided a straightforward, and now widely employed, route for the conversion of apo-CPs to holo-CPs (Quadri et al., 1998). However, the conversion of holo-CPs to acyl-*S*-CPs can be complicated by the nonavailability of an appropriate ligase, specifically when the carboxylic acid to be ligated to the holo-CP is an advanced biosynthetic intermediate for which a cognate ligase does not exist.

Strategies to address this challenge have been devised. Numerous biosynthetic enzymes that employ acyl-*S*-CoA and acyl-*S*-CP substrates have been reported to use abbreviated substrates, such as *N*-acetylcysteamine thioesters (NAC thioesters, Fig. 3). Strategies for the elaboration of NAC thioesters were recently reviewed (Franke & Hertweck, 2016). An alternate strategy is based on the substrate tolerance of Sfp. The crystal structure of Sfp in complex with CoA-SH reveals that the enzyme makes extensive contacts with

the 3′-phosphoadenosine and the pyrophosphate moieties of CoA-SH. However, the cysteamine moiety is not actively engaged by the enzyme (Reuter, Mofid, Marahiel, & Ficner, 1999). Thus, Sfp can directly catalyze the production of acyl-*S*-CPs starting from apo-CPs by the transfer of the acyl-*S*-phosphopantetheine moiety from acyl-*S*-CoAs (Fig. 3). This strategy requires the elaboration of acyl-*S*-CoAs by the synthetic ligation of carboxylic acids to CoA-SH, a reaction that proceeds in low yields while using typically high-priced CoA-SH as a substrate, and requires a chromatographic purification of the acyl-*S*-CoA product (Mishra & Drueckhammer, 2000). Hence, alternate routes for the elaboration of acyl-*S*-CoAs have been sought (Peter, Vogeli, Cortina, & Erb, 2016). By adopting the substrate promiscuity of the CoA-SH biosynthetic enzymes together with the broad substrate tolerance of Sfp, we have recently reported a chemoenzymatic scheme for the elaboration of acyl-*S*-CoAs and acyl-*S*-CPs at a multimilligram scale (Agarwal et al., 2015; Nazi, Koteva, & Wright, 2004; Strauss & Begley, 2002). These procedures are outlined in Section 3 together with mass spectrometric assays to query the acylation states of CPs. In Section 4, we describe procedures to assay for the activity of FAD-dependent halogenases using acyl-*S*-CP substrates together with the mass spectrometric characterization of the assay products. Using these procedures, for the first time, we experimentally validate the distinction between physiological chlorinases and brominases.

2. EQUIPMENTS AND CONSUMABLES

Equipments

1. Temperature controlled incubators and shakers for bacterial growth
2. Refrigerated centrifuges
3. Sonicator
4. SDS-PAGE gel casting accessories and electrophoresis chamber
5. GE Healthcare Amersham Imager 600
6. White light transilluminator
7. Spectrophotometer
8. GE Healthcare Aktaprime plus chromatography system
9. Agilent 1260 infinity quaternary LC coupled to impact II ToF mass spectrometer (Bruker Daltonics)

Consumables

1. Chromatography: His-Trap FF 5 mL (GE Healthcare); Hi-Trap Q 5 mL (GE Healthcare); Hi-Load 16/600 Superdex 75pg (GE Healthcare);

Aeris 3.6 μm WIDEPORE XB-C18 200 Å 250 × 4.6 mm^2 (Phenomenex).

2. Antibiotics: Kanamycin monosulfate and ampicillin trihydrate (Fisher).
3. Chemicals and reagents: All chemicals and reagents were purchased from Fisher Scientific unless noted otherwise. Tris base, sodium chloride, imidazole, glycerol, HEPES, potassium chloride, potassium bromide, sodium phosphite dibasic pentahydrate, magnesium chloride tetrahydrate, isopropyl β-D-1-thiogalactopyranoside (IPTG), adenosine 5′-triphosphate disodium salt dihydrate (ATP) (Sigma-Aldrich), FAD (Sigma-Aldrich), nicotinamide adenine dinucleotide phosphate hydrate ($NADP^+$) (Sigma-Aldrich), and tris(2-carboxyethyl)phosphine hydrochloride (Sigma-Aldrich), protein assay dye reagent (Bio-Rad), Brilliant Blue R-250 (Bio-Rad).
4. Solvents: Chromatography grade acetonitrile, water, and formic acid (Fisher).
5. Media: Luria-Bertani (LB) broth and terrific broth (TB) (Fisher).
6. Centrifugal filters: Amicon ultra-15 centrifugal filters with MWCO of 3, 10, and 30 kDa (EMD Millipore).

3. PREPARATION OF Acyl-*S*-CPs

3.1 Discussion

The biosynthesis of the halogenated pyrrole containing marine natural products pentabromopseudilin (**3**, Fig. 4), marinopyrrole A (**4**), and chlorizidine (**5**) each involves halogenation of a pyrrolyl-*S*-CP intermediate by the FAD-dependent halogenases Bmp2, Mpy16, and Clz5, respectively (Agarwal et al., 2014; Mantovani & Moore, 2013; Yamanaka, Ryan, Gulder, Hughes, & Moore, 2012). As described above, ligases that append pyrrole-2-carboxylic acid to the corresponding holo-*S*-CPs (Bmp1, Mpy15, and Clz18, respectively) do not exist, since pyrrolyl-*S*-CPs are advanced biosynthetic intermediates that are derived by the enzymatic oxidation of prolyl-*S*-CPs (Fig. 4) (Agarwal et al., 2014; Thomas, Burkart, & Walsh, 2002; Walsh, Garneau-Tsodikova, & Howard-Jones, 2006). It should be noted that the synthesis and halogenation of pyrrolyl-*S*-CPs are not restricted to marine natural products. The discovery of gene clusters encoding production of **3**–**5** was inspired by the prior characterization of the molecular logic for the construction of the dichlorinated pyrrole moiety in the terrestrial natural product pyoluteorin (**6**) (Dorrestein, Yeh, Garneau-Tsodikova, Kelleher, & Walsh, 2005; Thomas et al., 2002).

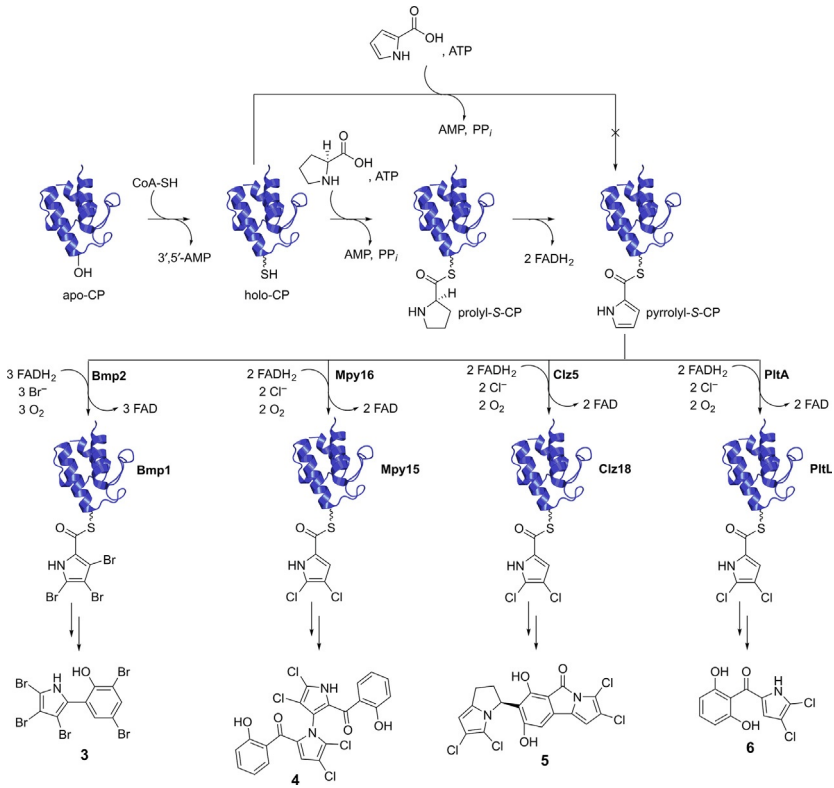

Fig. 4 Molecular logic for the construction of halogenated pyrrole moiety in **3**–**6**. The phosphopantetheine arm of holo-CPs is abbreviated for clarity.

As ATP-dependent ligases that ligate pyrrole-2-carboxylic acid to holo-CPs are not available, alternate strategies have been employed for the synthesis of pyrrolyl-*S*-CP substrates in order to assay the activities of the FAD-dependent halogenases as illustrated in Fig. 4. In a strategy that mimics the natural product biosynthetic scheme, ATP-dependent ligases are used to ligate L-Pro to holo-CPs to generate prolyl-*S*-CPs, followed by oxidation of the prolyl heterocycle to the aromatic pyrrolyl ring by FAD-dependent dehydrogenases (Dorrestein et al., 2005; Thomas et al., 2002). This approach is complicated by the FAD-dependent dehydrogenase being difficult to purify, and which does not yield stoichiometric conversion of the prolyl-*S*-CP substrate to the pyrrolyl-*S*-CP product (Agarwal et al., 2014; Dorrestein et al., 2005; Thomas et al., 2002). Based on the prior demonstration of the broad substrate tolerance for the coenzyme A biosynthetic enzymes CoaA, CoaD, and CoaE, and the phosphopantetheinyl transferase

Sfp, we have described a chemoenzymatic scheme to generate pyrrolyl-*S*-CPs, the procedures for which are described here (Agarwal et al., 2015; El Gamal et al., 2016; Nazi et al., 2004; Quadri et al., 1998; Strauss & Begley, 2002; Worthington & Burkart, 2006). The strategy described affords multimilligram quantities of highly pure pyrrolyl-*S*-CPs and does not employ the FAD-dependent prolyl-*S*-CP dehydrogenases. The chemoenzymatic scheme, as illustrated in Fig. 5 involves pyrrolyl-*S*-pantetheine (**7**), the synthesis of which has been described in detail previously (Agarwal et al., 2015).

3.2 Procedure

Buffers and solvents

TB media prepared according to manufacturer's directions and sterile autoclaved prior to use.

LB media prepared according to manufacturer's directions and sterile autoclaved prior to use.

Resuspension buffer: 20 m*M* Tris–HCl (pH 8.0), 500 m*M* NaCl, 10% glycerol.

Wash buffer: 20 m*M* Tris–HCl (pH 8.0), 1 *M* NaCl, 30 m*M* imidazole.

Elution buffer: 20 m*M* Tris–HCl (pH 8.0), 1 *M* NaCl, 250 m*M* imidazole.

Elution buffer (#2): 20 m*M* Tris–HCl (pH 8.0), 1 *M* NaCl, 250 m*M* imidazole, 10% glycerol.

Dialysis buffer: 20 m*M* Tris–HCl (pH 8.9), 50 m*M* KCl.

Ion-exchange buffer A: 20 m*M* Tris–HCl (pH 8.9).

Ion-exchange buffer B: 20 m*M* Tris–HCl (pH 8.9), 1 *M* KCl.

Storage buffer: 20 m*M* HEPES (pH 7.5), 50 m*M* KCl, 10% glycerol.

Assay quenching buffer: MeCN + 0.1% formic acid.

HPLC solvent A: Water + 0.1% formic acid.

HPLC solvent B: MeCN + 0.1% formic acid.

Bacterial strains

E. coli DH5α for plasmid maintenance and cloning.

E. coli BL21Gold(DE3) for overexpression of recombinant proteins.

3.2.1 Preparation of Plasmid DNA

Table 1 lists the plasmids that are used in the procedures described in this study. All recombinant proteins are appended with fusion tags to assist in purification of recombinant proteins.

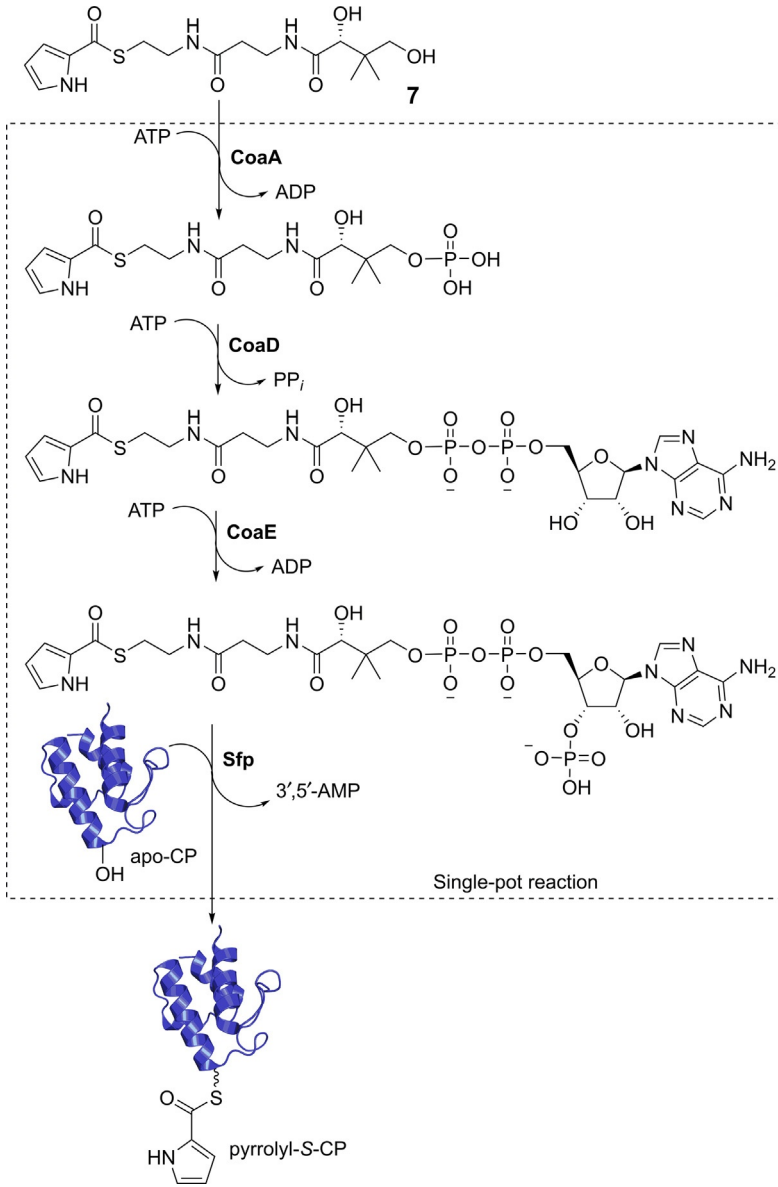

Fig. 5 Route for the enzymatic synthesis of pyrrolyl-*S*-CPs starting from **7** with a sequential action of CoaA, CoaD, CoaE, and Sfp enzymes. The reaction can be performed as a "single-pot" reaction in which all four enzymes can be incubated with **7**, apo-CPs, and ATP to yield pyrrolyl-*S*-CPs, as is described.

Table 1 Plasmids Used in This Study

Construct	Fusion Tag	Plasmid Backbone	Source Organism	Reference
Bmp2	N-terminal His_6	pET-28b(+)	*Marinomonas mediterranea* MMB-1	Agarwal et al. (2014)
Mpy16	N-terminal His_6 followed by maltose-binding protein (MBP)	pET-28b(+)	*Streptomyces* sp. CNQ-418	El Gamal et al. (2016)
Bmp1[a]	N-terminal His_6	pET-28b(+)	*M. mediterranea* MMB-1	Agarwal et al. (2014)
Mpy15	N-terminal His_6	pET-28b(+)	*Streptomyces* sp. CNQ-418	El Gamal et al. (2016)
SsuE	N-terminal His_6	pET-28b(+)	*E. coli*	Agarwal et al. (2014)
PtdH	N-terminal His_6	pET-15b(+)	*Pseudomonas stutzeri*	Johannes, Woodyer, and Zhao (2007)
Sfp	C-terminal His_6	pET-24b(+)	*Bacillus subtilis*	Agarwal et al. (2014)
CoaA	N-terminal His_6	pET-28b(+)	*E. coli*	Agarwal et al. (2015)
CoaD	N-terminal His_6 followed by MBP	pET-28b(+)	*E. coli*	Agarwal et al. (2015)
CoaE	N-terminal His_6 followed by MBP	pET-28b(+)	*E. coli*	Agarwal et al. (2015)

[a]Bmp1 is a didomain protein with an N-terminal CP domain and a C-terminal thioesterase domain. In this study, for clarity, Bmp1 refers to the Bmp1 CP domain (Agarwal et al., 2014; El Gamal et al., 2016).

3.2.2 Expression and Purification of Recombinant Proteins

All proteins described in this study were recombinantly expressed and purified in an identical manner unless otherwise noted.

Experimental protocol

1. Transform plasmid DNA into competent BL21Gold(DE3) cells and grow transformants on LB agar media with antibiotic selection for 14–16 h at 37°C.
2. Pick individual colonies and grow for 14–16 h in 5 mL TB media with antibiotic selection with vigorous shaking at 37°C.
3. Use 5 mL of the starter culture to innoculate 1 L of TB media with antibiotic selection and grow with vigorous shaking at 37°C till optical density (measured at 600 nm) reaches 0.6.

4. Reduce the incubation temperature to 18°C and induce protein expression by the addition of 0.3 m*M* IPTG (isopropyl-β-D-1-thiogalactopyranoside). Grow for additional 18 h with vigorous shaking at 18°C.
5. Collect biomass by centrifugation at $5000 \times g$ for 15 min at 4°C. Resuspend the pellet in 50 mL resuspension buffer that is prechilled to 4°C.
6. Lyse bacterial cells by sonication. Care should be exercised to maintain the temperature close to 4°C during sonication.
7. Clarify the lysate by centrifugation at $30{,}000 \times g$ for 1 h at 4°C. Collect the supernatant. The pellet can be discarded.
8. Apply the supernatant to a 5 mL His-Trap chromatography column using an AKTA Prime FPLC system operating at 4°C.
9. After the supernatant has been applied to the column, wash the column extensively with wash buffer.
10. Elute the bound protein by a continuous linear gradient to 100% elution buffer across 20 column volumes and collect the eluate in 5 mL fractions. Analyze the fractions by SDS-PAGE with Coomassie staining. *Note*: Elute SsuE using elution buffer (#2). Collect fractions with pure SsuE protein and proceed directly to step 18. Do not concentrate the SsuE protein as it is prone to aggregation.
11. Analyze the fractions by SDS-PAGE with Coomassie staining.
12. Pool the fractions with pure protein and dialyze the protein sample overnight using dialysis buffer.
13. Apply the dialyzed protein sample to a 5 mL Hi-Trap Q chromatography column using an AKTA Prime FPLC system operating at 4°C.
14. Wash the column with two column volumes of 95% ion-exchange buffer A and 5% ion-exchange buffer B.
15. Elute the bound protein by a continuous linear gradient to 100% ion-exchange buffer B across 20 column volumes and collect the eluate in 5 mL fractions.
16. Analyze the fractions by SDS-PAGE with Coomassie staining.
17. Pool the fractions with pure protein and concentrate to 2.5 mL final volume using centrifugal filters with 10 kDa molecular weight cut-off (MWCO). For CPs, use centrifugal filters with 3 kDa MWCO.
18. Desalt the concentrated protein into storage buffer using gravity flow desalting columns.

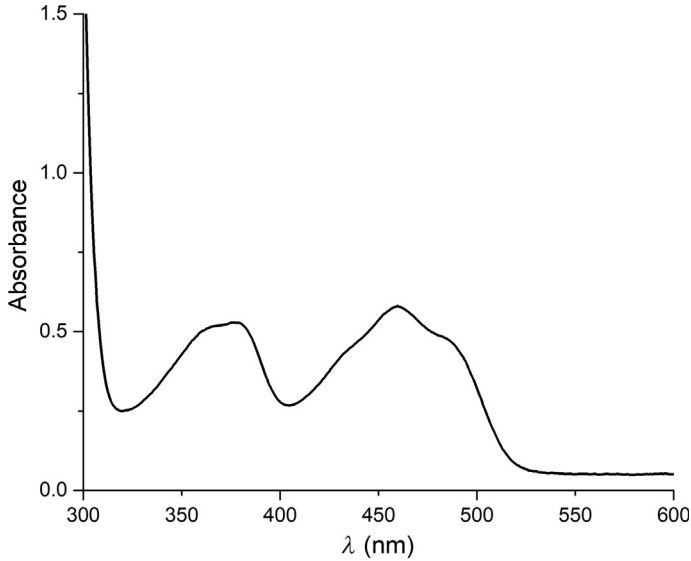

Fig. 6 UV–vis absorbance spectra for Bmp2 demonstrating the presence of FAD cofactor.

19. For halogenases, record an absorbance spectra from 300 to 600 nm. The absorbance spectra should demonstrate the presence of FAD. Typical absorbance spectra for Bmp2 halogenase are shown in Fig. 6.
20. Measure protein concentration using Bradford assay. Store purified proteins at −80°C in 100 μL aliquots. Store SsuE protein at 4°C and use without further modifications. *Note*: Procedures described here typically yield greater than 30 mg of purified CPs and greater than 50 mg of all other proteins. SsuE protein, as purified above, retains activity for ~2 days at 4°C. A fresh purification of SsuE is recommended after that.

3.2.3 Chemoenzymatic Synthesis of Pyrrolyl-S-CPs

The scheme for the synthesis of pyrrolyl-*S*-CPs is illustrated in Fig. 5. Briefly, the four catalysts—CoaA, CoaD, CoaE, and Sfp—are incubated together with the substrates **7**, ATP/Mg^{2+}, and purified apo-CPs as listed in Table 2. The reaction is conducted in a "single-pot" without the requirement of purifying the intermediates. After incubation at 30°C, the reaction is purified by size exclusion chromatography and the pyrrolyl-*S*-CP is collected. Size exclusion chromatography removes the residual ATP and other

Table 2 Final Concentration of Assay Components for Synthesis of Pyrrolyl-*S*-CPs

Component	Concentration in Assay
apo-CP	250 μ*M*
7	1 m*M*
CoaA[a]	1 μ*M*
CoaD[a]	1 μ*M*
CoaE[a]	1 μ*M*
Sfp[a]	1 μ*M*
ATP	9 m*M*
$MgCl_2$	10 m*M*
HEPES-Na (pH 7.9)	50 m*M*
Water	

[a]CoaA, CoaD, CoaE, and Sfp can be premixed as a single "enzyme cocktail" if multiple assays are being set up in parallel.

small-molecule products generated during the course of the reaction and yields pure pyrrolyl-*S*-CP in a defined buffer. At this stage, the pyrrolyl-*S*-CPs can be concentrated and stored at −80°C in small aliquots for use in halogenation assays described below. As shown in Fig. 5, this procedure generates pyrrolyl-*S*-CoA in situ which then acts as a substrate for the Sfp enzyme for transfer of the pyrrolyl-*S*-phosphopantetheine moiety to the side chain hydroxyl of the apo-CP. As Sfp requires the CP to be in the apo state for the transfer of the pyrrolyl-*S*-phosphopantetheine moiety, holo-CPs are not responsive to this methodology. The procedure illustrated in Fig. 5 is fundamentally different from the biosynthetic scheme as shown in Fig. 4 in which a prolyl ligase is used to ligate L-pro to the phosphopantetheine thiol of a holo-CP and the prolyl side chain is subsequently oxidized to generate pyrrolyl-*S*-CP.

Experimental protocol

1. Thaw aliquots of apo-CPs (Bmp1 and Mpy15), Sfp, PtdH, CoaA, CoaD, and CoaE on ice. Centrifuge briefly using a table-top centrifuge.
2. Prepare a fresh aliquot of 200 m*M* ATP in prechilled 50 m*M* HEPES-Na (pH 7.9) buffer.
3. For each CP, in duplicate, add the reaction components to the final concentrations delineated in Table 2. The reactions should be set up

in duplicate, with one of the reactions serving as the negative control (see below).

4. Incubate the reaction at 30°C for 5 min.
5. Initiate one of the two reactions by adding ATP. Adjust the final reaction volume to 2.5 mL with water.
6. For the other reaction, add water in place of ATP. This reaction is henceforth referred to as "negative control."
7. Incubate the reactions at 30°C for 3 h.
8. Centrifuge the reaction tubes at 4°C to remove any aggregated proteins.
9. Inject the reactions on a Superdex 75 16/200 size exclusion chromatography column equilibrated in 20 m*M* Tris–HCl (pH 8.0) using an AKTA Prime FPLC system operating at 4°C. Monitor the absorbance of the eluate at 280 nm.
10. For reactions in which ATP is added, two major peaks will be observed on the absorbance chromatogram. The first peak corresponds to the modified CPs, and the second peak corresponds to the residual ATP, and ADP and 3′,5′-AMP that are generated during the course of the reaction. The second peak is not observed for the negative control reaction.
11. Collect the eluate corresponding the first peak of the absorbance chromatogram. The presence of CP can be verified by SDS-PAGE and Coomassie staining as shown in Fig. 7.
12. Withdraw 100 μL aliquot of the eluate. Inject 10 μL of this sample on a HPLC–MS/MS system and elute using the following solvent gradient

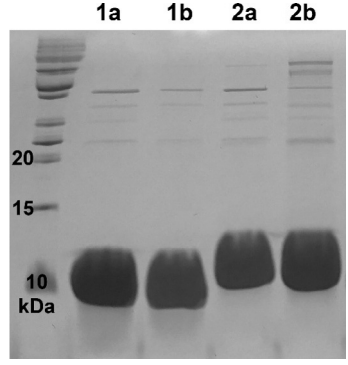

Fig. 7 SDS-PAGE demonstrating 1a: 20 μg apo-Bmp1; 1b: 20 μg pyrrolyl-*S*-Bmp1; 2a: 20 μg apo-Mpy15; and 2b: 20 μg pyrrolyl-*S*-Mpy15.

at a flow rate of 0.5 mL/min using an Aeris 3.6 μm widepore XB-C18 200 Å 250 × 4.6 mm^2 column:

a. 0–5 min: 5% HPLC solvent B
b. 5–30 min: linear gradient to 70% HPLC solvent B
c. 30–31 min: linear gradient to 95% HPLC solvent B
d. 31–35 min: 95% HPLC solvent B
e. 35–36 min: linear gradient to 5% HPLC solvent B
f. 36–38 min: 5% HPLC solvent B
g. 38–39 min: linear gradient to 95% HPLC solvent B
h. 39–42 min: 95% HPLC solvent B
i. 42–43 min: linear gradient to 5% HPLC solvent B
j. 43–46 min: 5% HPLC solvent B

Note: The elution gradient described here includes reintroduction of high concentration of the organic solvent in the mobile phase which affords a rigorous regeneration of the stationary phase.

13. Acquire mass spectrometry data between 5 and 40 min in the positive ionization mode in the range of 100–2000 m/z. Perform an external calibration with ESI-L Low Concentration Tuning Mix (Agilent Technologies) prior to data collection and use internal calibrant hexakis (1H,1H,3H–tetrafluoropropoxy)phosphazene (SynQuest Laboratories) throughout the runs. Use the following instrument parameters: capillary voltage 4500 V; nebulizer gas (nitrogen) pressure 2 bar; ion source temperature 200°C; dry gas flow 9 L/min, spectral rate 3 Hz for MS^1, and 6 Hz for MS^2. Select the three most intense ions per MS^1 for acquiring MS/MS fragmentation and use collision-induced dissociation (CID) energies listed in Table 3. Use basic stepping function to fragment ions at 50% and 125% of the CID calculated for each m/z from Table 3 with timing of 50% for each step. Similarly, employ basic stepping of collision RF of 550 and 800 Vpp with a timing of 50% for each step and transfer time stepping of 57 and 90 μs with a timing of 50% for each step. Set MS/MS active exclusion parameter to "3" and release after 30 s. Exclude the mass of internal calibrant from the MS/MS list using a mass range of m/z 921.5–924.5.
14. The data analysis procedure is described in Section 3.3.
15. Concentrate the eluate from the size exclusion chromatography column using a 3 kDa MWCO centrifugal filter and measure the protein concentration using the Bradford assay. Store the protein in small aliquots at −80°C.

Table 3 Collision-Induced Energies for Tandem MS Data Collection

Type	Mass	Width	Collision	Charge State
Base	100	4	22	1
Base	100	4	22	2
Base	300	5	27	1
Base	300	5	27	2
Base	500	6	35	1
Base	500	6	30	2
Base	1000	8	45	1
Base	1000	8	35	2
Base	2000	10	50	1
Base	2000	10	50	2

3.3 Data Analysis

The acylation states of CPs can be differentiated based on the mass observed in the MS^1 mass spectra as well as characteristic phosphopantetheine and acyl-*S*-phosphopantetheine ejection ions that is observed in the MS^2 spectra for holo- and acyl-*S*-CPs, respectively, as illustrated in Fig. 8 (Dorrestein et al., 2006; Dorrestein & Kelleher, 2006). Under positive ionization mode, mining for the masses of these ejection ions in the MS^2 spectra allow for deconvolution of the acylation states of the CPs which can be further corroborated by calculating the difference in the MS^1 masses as illustrated below for apo-Bmp1 and pyrrolyl-*S*-Bmp1 (Figs. 9 and 10). As mentioned previously, expression of CPs in heterologous hosts, *E. coli* in this study, will lead to partial conversion of apo-CPs to holo-CPs, which will also be deconvoluted as described below.

Data analysis protocol

1. Generate MS^2 extracted ion chromatograms (EICs) for $[M+H]^{1+}=261.13$ and $[M+H]^{1+}=354.15$ corresponding to the phosphopantetheine and pyrrolyl-*S*-phosphopantetheine ejection ions, respectively, as shown in Fig. 8. The EICs for the reactions with and without ATP as described in Section 3.2.3 for Bmp1 are shown in Fig. 9.
2. Using the retention times of the MS^2 ejection ions shown in Fig. 9, determine the retention time of MS^1 spectra corresponding to holo- and pyrrolyl-*S*-Bmp1 and generate their corresponding MS^1 spectra.

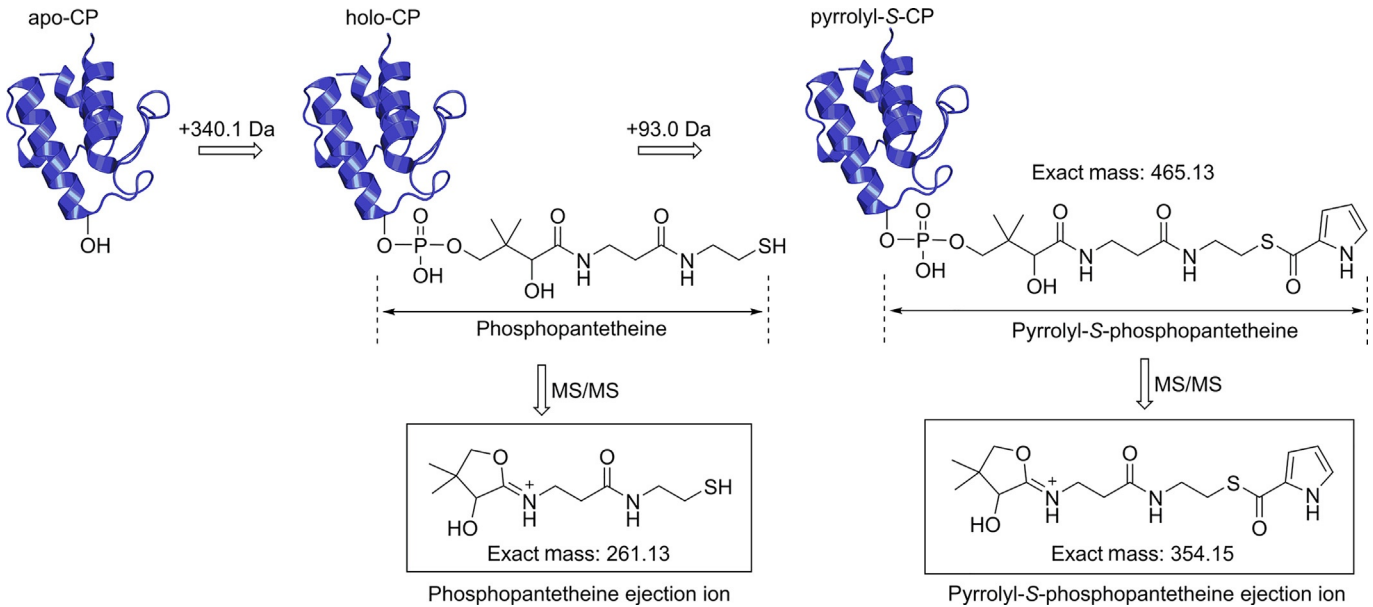

Fig. 8 Under MS/MS conditions in a mass spectrometer, holo-CPs generate a characteristic pantetheine ejection ion of $[M+H]^{1+} = 261.13$ that can be detected in the mass spectrometry dataset to signify the presence of holo-CPs. Similarly, acyl-*S*-CPs, such as pyrrolyl-*S*-CPs, also generate the corresponding ejection ions that can be used to identify acyl-*S*-CPs in mass spectrometry datasets. Apo-CPs do not generate these characteristic ejection ions. The theoretical parent mass difference between the apo-CPs, holo-CPs, and pyrrolyl-*S*-CPs is shown.

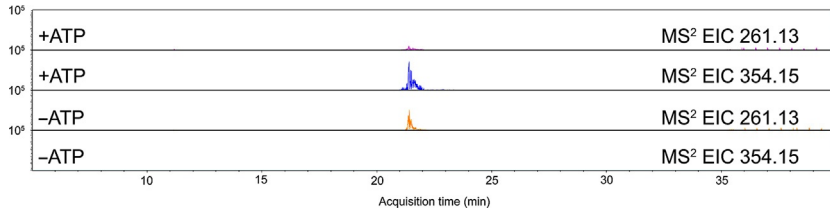

Fig. 9 EICs for Bmp1 reactions with and without ATP as described in Section 3.2.3. The MS^2 pyrrolyl-*S*-phosphopantetheine ejection ion (354.15 Da) is observed only when ATP is present in the reaction. However the phosphopantetheine ejection ion (261.13 Da), which corresponds to the holo-Bmp1 is observed in both reactions. As mentioned before, during recombinant expression in heterologous hosts, apo-CPs are partially converted to their holo state by the activity of the host encoded phosphopantetheinyl transferase(s).

These MS^1 spectra are shown in Fig. 10. The MS^1 spectra corresponding to the apo state has to be identified empirically.

3. Deconvolute the mass of the apo-, holo-, and pyrrolyl-*S*-Bmp1 using the formula:

   Deconvoluted mass = (monoisotopic mass × z) − (z × mass of H^+).

   Thus:

   For $[M + 13H]^{13+}$ species for apo-Bmp1 as shown in Fig. 10, the calculated deconvoluted mass is = (817.35 × 13) − (13 × 1.01) = 10,612.42 Da

   For $[M + 13H]^{13+}$ species for holo-Bmp1 as shown in Fig. 10, the calculated deconvoluted mass is = (843.50 × 13) − (13 × 1.01) = 10,952.37 Da

   For $[M + 13H]^{13+}$ species for pyrrolyl-*S*-Bmp1 as shown in Fig. 10, the calculated deconvoluted mass is = (850.66 × 13) − (13 × 1.01) = 11,045.45 Da.

4. Mass difference between holo-Bmp1 and apo-Bmp1 (10,952.37–10,612.42 = 339.95 Da) and between pyrrolyl-*S*-Bmp1 and holo-Bmp1 (11,045.45–10,952.37 = 93.08 Da) is in good agreement with the theoretical mass differences as shown in Fig. 8.

Note: It is not necessary that identical $[M + nH]^{n+}$ species be used for parent mass deconvolution across all MS^1 spectra as shown in Fig. 10. Clusters of ions corresponding to different charge states should be manually examined, and the charge state that offers the best resolution between the isotopic peaks should be selected for parent mass deconvolution. Different charge states should lead to identical, or closely matching parent ion masses (depending on the resolution of the mass spectrometer, see below). The resolution of the isotopic peaks will decrease as the charge state increases. However, the

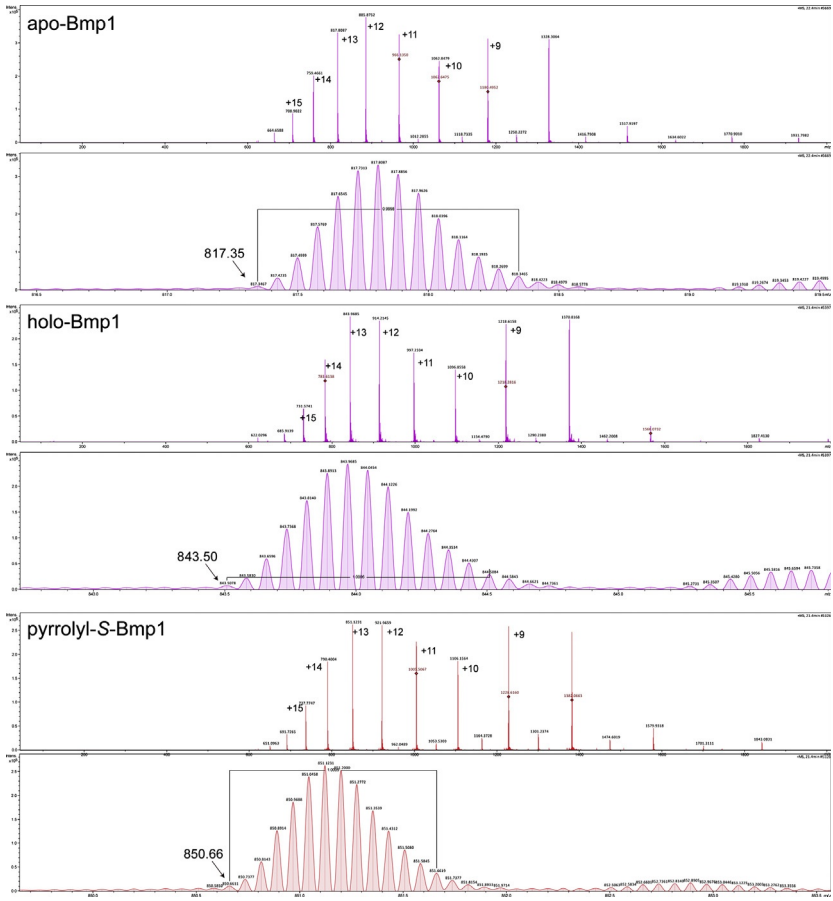

Fig. 10 MS^1 spectra and zoomed in view of the $[M+13H]^{13+}$ cluster of ions for apo-Bmp1 (*top*), holo-Bmp1 (*middle*), and pyrrolyl-*S*-Bmp1 (*bottom*). In the MS^1 spectra, multiple charge states for peptides are typically observed. Charge states decrease toward the right of the spectra. The charge state for a particular cluster of ions can be calculated by counting the number of isotopic peaks that have to be traversed for generating a difference of 1 Da starting from the monoisotopic peak.

intensity of lower charge states might be too low for the monoisotopic peak to be identified accurately. This tradeoff between resolution and abundance must be empirically balanced for each MS^1 spectra. For relatively high resolution mass spectrometers, such as the Bruker Impact II Q-ToF (resolution ~40,000) and the Agilent 6530 Q-ToF (resolution ~20,000), the isotopic resolution is usually sufficient that monoisotopic peaks can be accurately identified even for low abundance charge states. For lower resolution mass

spectrometers, such as the Bruker amaZon SL ion-trap (resolution ~2000), only higher abundance charge states can be used to accurately deconvolute parent ion peptide masses. In our experience, ToF and ion-trap instruments both yield identical ejection ions for acyl-*S*-CP species. As has been reported previously, attention must be paid to fragmentation energies that are employed to generate the MS^2 spectra (Agarwal et al., 2015). At low fragmentation energies, ejection ions will not be observed in sufficient abundance to generate MS^2 EICs. At high fragmentation energies, acyl-*S*-pantetheine ejection ions can degrade to pantetheine ejection ions due to cleavage of the labile thioester bond. The appropriate fragmentation energy must also be empirically determined.

4. HALOGENATION OF Acyl-*S*-CPs

4.1 Discussion

FAD-dependent halogenases acting on substrates tethered to CPs are mechanistically identical to FAD-dependent halogenases acting on small-molecule substrates such as L-tryptophan (Fig. 2). Halogenases acting on CP-tethered substrates routinely catalyze multiple halogenations on their CP-tethered aromatic substrates. For instance, the halogenase Bmp2 brominates all four carbon atoms of a pyrrole ring while Mpy16, Clz5, and PltA dichlorinate the pyrrole ring (Dorrestein et al., 2005; El Gamal et al., 2016; Mantovani & Moore, 2013) (Fig. 4). Note, while Bmp2 tetrabrominates the pyrrole ring, a tetrabrominated product will not be produced with the absence of a partner thioesterase enzyme (El Gamal et al., 2016). In the absence of the thioesterase, a tribrominated product, as shown in Fig. 4, will predominate. The halogenase HrmQ, involved in the biosynthesis of hormaomycin, only monochlorinates the pyrrole ring (Hofer et al., 2011). Similarly, the FAD-dependent halogenase AltN likely mono- and dibrominates the phenyl ring en route production of the alterochromides (Ross, Gulland, Dorrestein, & Moore, 2015), while the chlorinase SgcC3 only monochlorinates the phenyl ring during the biosynthesis of the enediyne C-1027 (Lin, Van Lanen, & Shen, 2007). FAD-halogenases that work on small organic molecules exclusively catalyze monohalogenations in a regiospecific fashion which is perhaps a result of the highly ordered binding of the substrate in the halogenase active site (Agarwal, Miles, et al., 2017; Bitto et al., 2008; Dong et al., 2005; Zhu et al., 2009). While a crystal structure of FAD-dependent halogenase with a CP-tethered substrate is not available yet, the active sites of Bmp2, Mpy16, and PltA seem

to be more accommodating of substrates of different steric sizes (El Gamal et al., 2016; Pang, Garneau-Tsodikova, & Tsodikov, 2015). Recent discovery of the bifunctional methyltransferase-halogenase AoiQ and the MalA/MalA' indole halogenases, together with the fungal halogenase CazI, however challenge this notion and is likely to further expand the mechanistic repertoire of FAD-dependent halogenases (Chankhamjon et al., 2016; Fraley et al., 2017; Sato et al., 2016).

In this section, we focus on the description of the unique activities of FAD-dependent halogenases that affect multiple halogenations on their aromatic substrates. We describe the procedures to assay for the activity of the pyrrolyl-*S*-Bmp1 tribrominase Bmp2 and the pyrrolyl-*S*-dichlorinase Mpy16. FAD-dependent halogenases require a partner flavin reductase that provides $FADH_2$ to the halogenase (Fig. 2). We employ the *E. coli* flavin reductase SsuE for this purpose (Eichhorn, van der Ploeg, & Leisinger, 1999). Electrons required for the reduction of FAD by SsuE are derived from NADPH. In the assay described below, NADPH is regenerated by the thermostable phosphite dehydrogenase PtdH in situ (Fig. 2) (Johannes et al., 2007). Using these procedures, we experimentally demonstrate, for the first time, that physiological chlorinases as defined by Drennan and coworkers can chlorinate as well as brominate (Blasiak & Drennan, 2009). However, the brominase Bmp2 can only brominate and is incapable of chlorination activity, thus providing a firm experimentally validated distinction between physiological chlorinases and brominases.

4.2 Procedure

The purification of halogenases followed the protein purification protocol described in Section 3.2.2, and synthesis of pyrrolyl-*S*-CP substrates is described in Section 3.

Experimental protocol

1. Thaw aliquots of halogenases, pyrrolyl-*S*-CPs, and PtdH on ice. Centrifuge briefly using a table-top centrifuge.
2. The pyrrolyl-*S*-CPs can be analyzed by HPLC–MS/MS as described in Section 3.2.3 to ensure that no degradation has occurred. In our experience, pyrrolyl-*S*-Bmp1 and pyrrolyl-*S*-Mpy15 do not undergo degradation during long-term storage at −80°C.
3. For halogenation of pyrrolyl-*S*-Mpy15 by Mpy16, keep the chloride concentration in the assay constant at 100 m*M* and vary the concentration of bromide as 0, 25, 50, 100, and 200 m*M*. Add assays components

Table 4 Concentration of Assay Components for Halogenation of Pyrrolyl-*S*-Mpy15 by Mpy16

Component	Concentration in Assay
Pyrrolyl-*S*-Mpy15	50 μ*M*
Mpy16	20 μ*M*
TCEP[a]	5 m*M*
Na-phosphite	10 m*M*
FAD	20 μ*M*
$NADP^+$	1 m*M*
PtdH	10 μ*M*
SsuE	50 n*M*
KCl	100 m*M*
KBr	0, 5, 25, 50, 100, and 200 m*M*
HEPES-Na (pH 7.9)	20 m*M*
10% glycerol[b]	

[a]Tris(2-carboxyethyl)phosphine.
[b]10% glycerol was used in place of water to adjust the final volume of the assay to 100 μL.

according to Table 4 to a final volume of 100 μL. In case of halogenation of pyrrolyl-*S*-Bmp1 by Bmp2, keep the bromide concentration in the assay constant at 100 m*M* and vary the concentration of chloride as 0, 25, 50, 100, and 200 m*M*. Add assays components according to Table 5 to a final volume of 100 μL.

4. Incubate the reactions at 30°C for 3 h.
5. After 3 h incubation, add 8 μL of trypsin (1 mg/mL) and incubate at 30°C for 1 h.
6. Add 108 μL of assay quenching buffer and centrifuge at 16,000 × *g* for 30 min. Carefully withdraw the supernatant into an appropriate vial.
7. Inject 10 μL of this sample on a HPLC–MS/MS system and elute using the following solvent gradient at a flow rate of 0.5 mL/min using an Aeris 3.6 μm widepore XB-C18 200 Å 250 × 4.6 mm^2 column:
   a. 0–5 min: 5% HPLC solvent B
   b. 5–30 min: linear gradient to 70% HPLC solvent B
   c. 30–31 min: linear gradient to 95% HPLC solvent B
   d. 31–35 min: 95% HPLC solvent B
   e. 35–36 min: linear gradient to 5% HPLC solvent B

Table 5 Concentration of Assay Components for Halogenation of Pyrrolyl-*S*-Bmp1 by Bmp2

Component	Concentration in Assay
Pyrrolyl-*S*-Bmp1	50 μ*M*
Bmp2	20 μ*M*
TCEP	5 m*M*
Na-phosphite	10 m*M*
FAD	20 μ*M*
$NADP^+$	1 m*M*
PtdH	10 μ*M*
SsuE	50 n*M*
KBr	100 m*M*
KCl	0, 5, 25, 50, 100, and 200 m*M*
HEPES-Na (pH 7.9)	20 m*M*
10% glycerol	

f. 36–38 min: 5% HPLC solvent B
g. 38–39 min: linear gradient to 95% HPLC solvent B
h. 39–42 min: 95% HPLC solvent B
i. 42–43 min: linear gradient to 5% HPLC solvent B
j. 43–46 min: 5% HPLC solvent B

8. Acquire mass spectrometry data between 5 and 40 min in the positive ionization mode. The data analysis procedure is described below.

4.3 Data Analysis

Data analysis for the halogenation of pyrrolyl-*S*-Mpy15 by the FAD-dependent halogenase Mpy16 is described in detail here using Figs. 11 and 12. The data analysis procedures for the halogenation of pyrrolyl-*S*-Bmp1 by Bmp2 are identical.

1. Generate MS^2 EICs, as described in Section 3.3, for $[M+H]^{1+} = 388.11$ and $[M+H]^{1+} = 422.07$ corresponding to the monochloropyrrolyl-*S*-phosphopantetheine and dichloropyrrolyl-*S*-phosphopantetheine ejection ions, respectively. These EICs correspond to the monochlorinated and dichlorinated products generated by Mpy16 using pyrrolyl-*S*-Mpy15 as substrate (Fig. 11). As Mpy16 can also accept bromide in

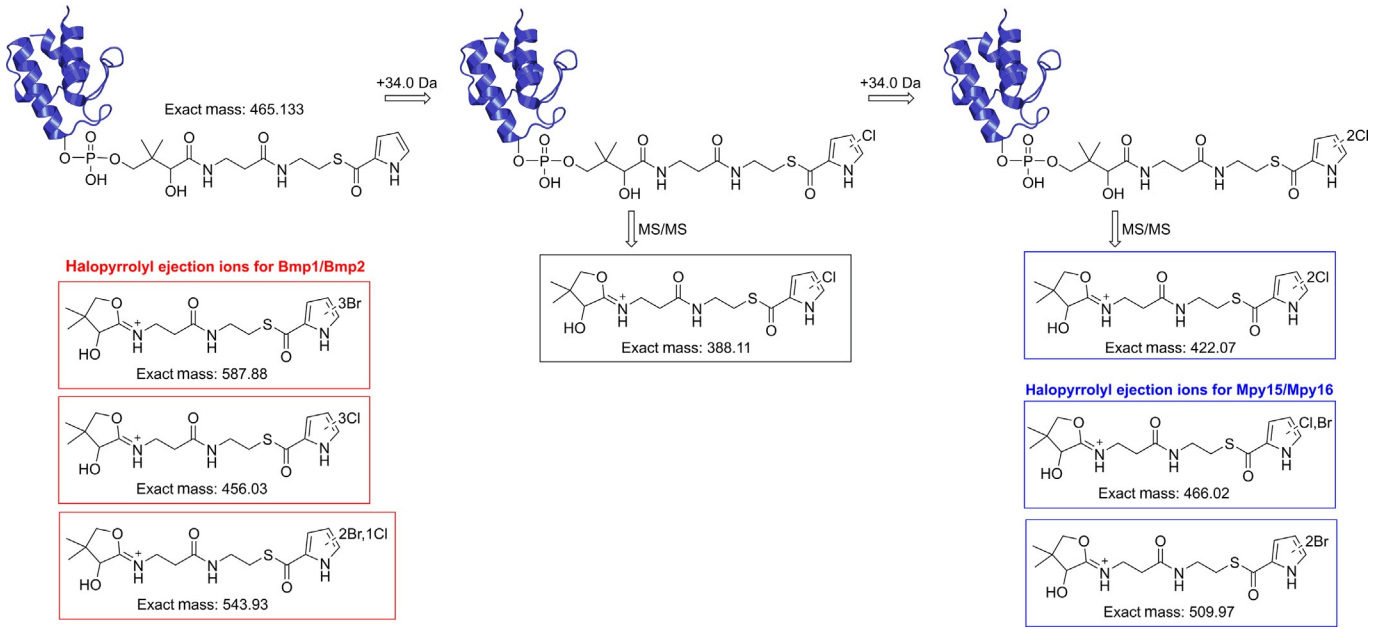

Fig. 11 MS^2 ejection ions for hypothetical dihalogenated products generated by halogenation of pyrrolyl-*S*-Mpy15 by Mpy16 (in *blue boxes*) and by trihalogenated products generated by the halogenation of pyrrolyl-*S*-Bmp1 by Bmp2 (in *red boxes*).

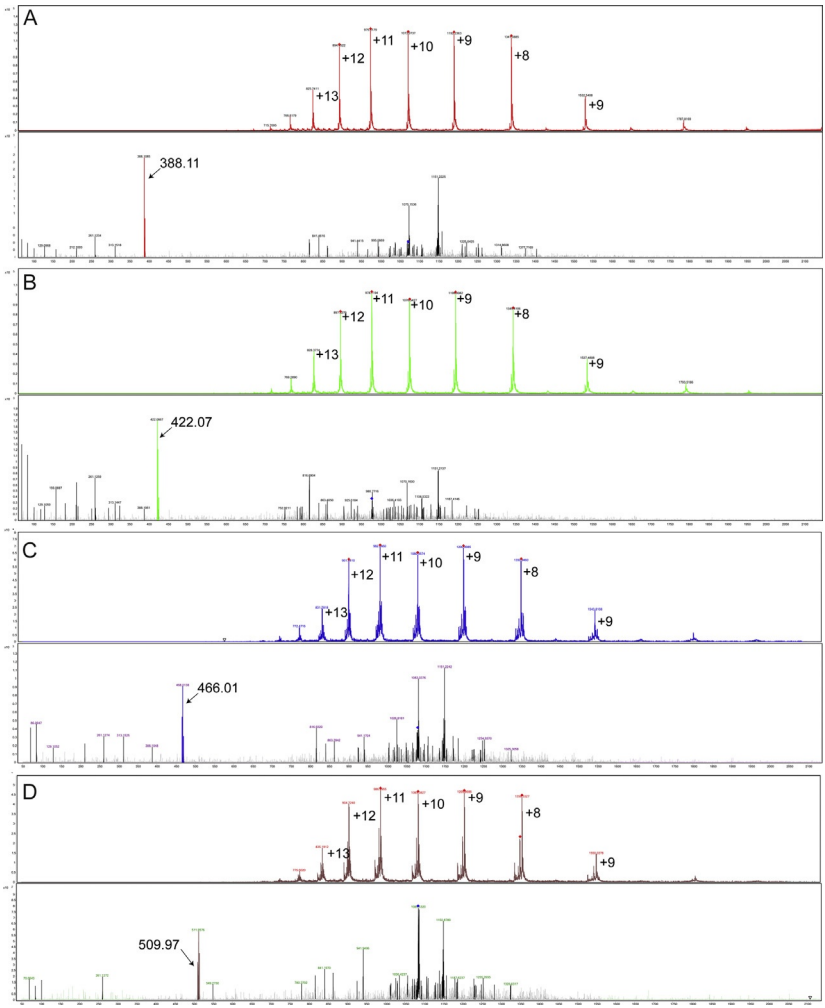

Fig. 12 MS^1 (*top*) and MS^2 (*bottom*) spectra corresponding to (A) monochloropyrrolyl-*S*-Mpy15, (B) dichloropyrrolyl-*S*-Mpy15, (C) mixed monochloro,monobromopyrrolyl-*S*-Mpy15, and (D) dibromopyrrolyl-*S*-Mpy15 products generated by Mpy16. The different peptide charge states in the MS^1 spectra and the ejection ions in the MS^2 spectra are illustrated.

place of chloride (vide infra), generate MS^2 EICs for $[M+H]^{1+} = 466.02$ and $[M+H]^{1+} = 509.97$ corresponding to the monochloro,monobromopyrrolyl-*S*-phosphopantetheine and dibromopyrrolyl-*S*-phosphopantetheine ejection ions, respectively. These EICs correspond

to the mixed monochlorinated–monobrominated and dibromintaed products generated by Mpy16 (Fig. 11).

2. Using the retention times of the four MS^2 ejection ions illustrated in Fig. 11, determine the MS^1 spectra corresponding to monochlorinated-, dichlorinated-, mixed monochlorinated–monobrominated-, and dibromintaed-pyrrolyl-*S*-Mpy15. The MS^1 spectra for the four possible products generated by Mpy16, along with the MS^2 spectra showing the corresponding ejection ions are illustrated in Fig. 12.
3. Using the procedures described in Section 3.3, select an appropriate ionization state of the acyl-*S*-Mpy15 peptide to deconvolute the parent masses. The thus-calculated mass differences should be in good agreement with the theoretical mass differences as calculated in Fig. 11.
4. To determine the halide specificity of Mpy16, extract MS^2 EICs for all reactions listed in Table 4 for the following ejection ions: $[M+H]^{1+}=422.07$, $[M+H]^{1+}=466.02$, and $[M+H]^{1+}=509.97$. As illustrated in Fig. 11 (blue boxes), these MS^2 ions correspond, respectively, to the dichlorinated, dibrominated, and mixed monobrominated–monochlorinated products that can possibly be generated by Mpy16.
5. Similarly, to determine the halide specificity of Bmp2, extract MS^2 EICs for all reactions listed in Table 5 for the following ejection ions: $[M+H]^{1+}=587.88$, $[M+H]^{1+}=456.03$, and $[M+H]^{1+}=543.93$. As illustrated in Fig. 11 (red boxes), these MS^2 ions correspond, respectively, to the tribrominated, trichlorinated, and mixed dibrominated–monochlorinated products that can hypothetically be generated by Bmp2.

   Note: It has been demonstrated that Mpy16 specifically catalyzes dihalogenation of the pyrrole ring, while in the absence of the thioesterase, Bmp2 catalyzes trihalogenation of the pyrrole ring (El Gamal et al., 2016).
6. The thus-generated EICs for all reactions listed in Table 4 are plotted in Fig. 13A, and EICs for all reactions listed in Table 5 are plotted in Fig. 13B.

Note: For the halogenation of pyrrolyl-*S*-Bmp1 by Bmp2, under no relative concentration of chloride and bromide in the assay can a chlorinated product be observed (Fig. 13B). A different trend is observed for Mpy16. At high relative concentration of chloride in the assay, Mpy16 generates a predominantly dichlorinated product (Fig. 13A). However, as the relative

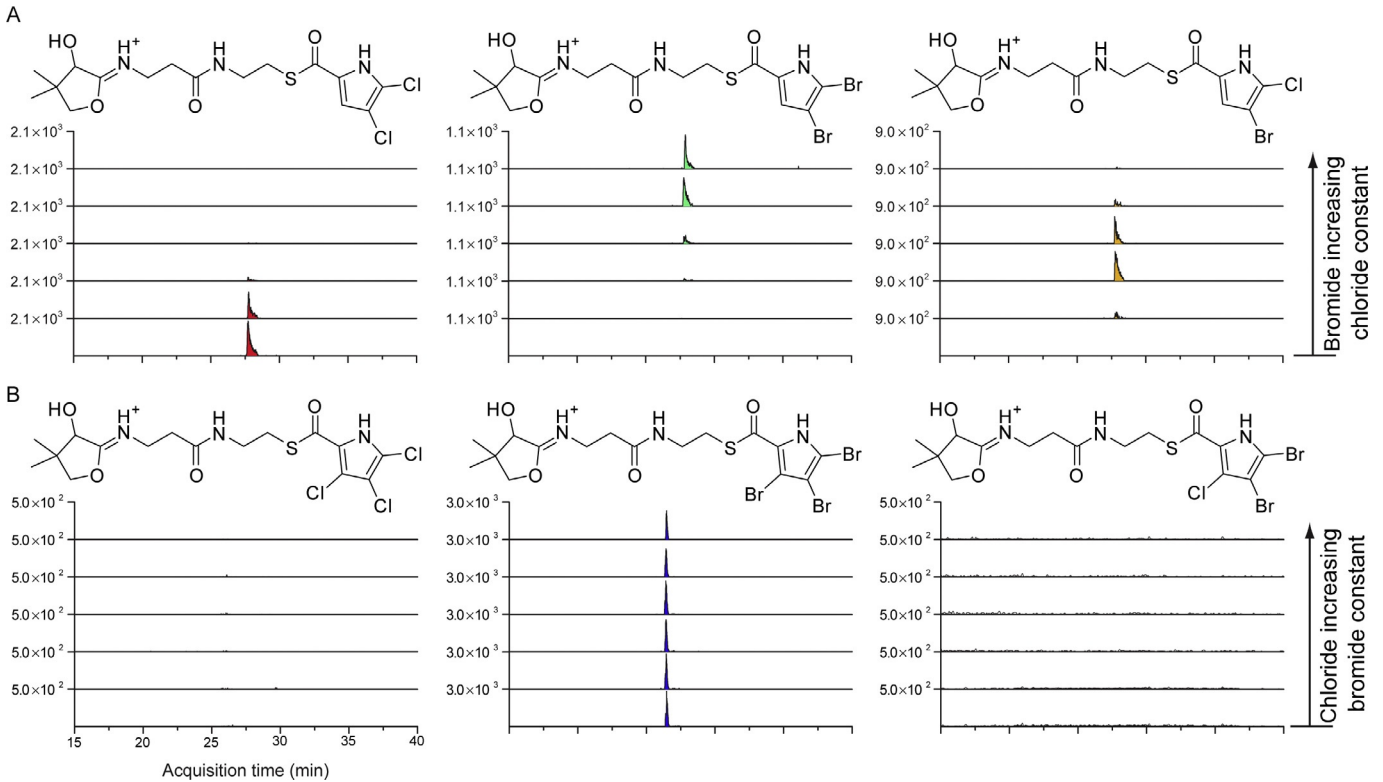

Fig. 13 MS^2 EICs (A) for all reactions listed in Table 4 for ejection ions corresponding to the dichlorinated, dibrominated, and mixed monobrominated–monochlorinated products that can possibly be generated by Mpy16. (B) MS^2 EICs for all reactions listed in Table 5 for ejection ions corresponding to trichlorinated, tribrominated, and mixed dibrominated–monochlorinated products that can hypothetically be generated by Bmp2.

concentration of bromide in the assay increases, ejection ion corresponding to a mixed brominated–chlorinated product is observed. At high relative concentration of bromide in the assay, the chlorinating activity of Mpy16 is completely abolished and only the ejection ion corresponding to dibrominated product is observed. Hence, the activity of a chlorinase is fundamentally different from a brominase, in that, a chlorinase (such as Mpy16) is an effective brominating catalyst depending on the relative concentration of halides in the assay. However, a brominase (such as Bmp2) is specific for incorporating bromide. To the best of our knowledge, this is the first experimental demonstration of the halide specificity for a chlorinase and a brominase and is aligned with the theoretical distinction between chlorinases and brominases (Blasiak & Drennan, 2009).

5. CONCLUSIONS

Access to starting material in sufficient quantity and purity is a major obstacle to investigating the reactions catalyzed by halogenase that require CP-tethered substrates. Crystal structures of halogenases in complex with their small-molecule organic substrates and products, such as for the Trp halogenases, have been critical in advancing our mechanistic understanding of FAD-dependent halogenases and their engineering for novel applications (Agarwal, Miles, et al., 2017; Latham et al., 2018). However, a cocrystal structure of a FAD-dependent halogenase in complex with a CP-tethered substrate has so far been elusive. In addition to the halogenase, cocrystallization requires access to large quantities of highly pure substrates, and the lack of the above-mentioned crystal structure is perhaps reflective of lack of access to these substrates needed to drive these efforts forward. The procedures describe in this study will enable these efforts in the future.

Mechanistically, FAD-dependent halogenases that employ CP-tethered substrates are presumed to be identical to FAD-dependent halogenases that employ small-molecule substrates such as L-Trp. The structural engagement of the FAD cofactor and the positioning of key catalytic resides for the two subclasses of FAD-dependent halogenases is highly similar. However, some key biochemical differences remain to be addressed. Primary among these is the absence of a catalytic Glu reside in acyl-*S*-CP halogenases (Fig. 14). To facilitate a electrophilic aromatic substitution by the halenium, the side chain of this Glu residue is postulated to serve as the catalytic base to abstract a proton from the substrate molecule, a role that is supported by kinetic and structural studies (Dong et al., 2005; Flecks et al., 2008). Thus, the

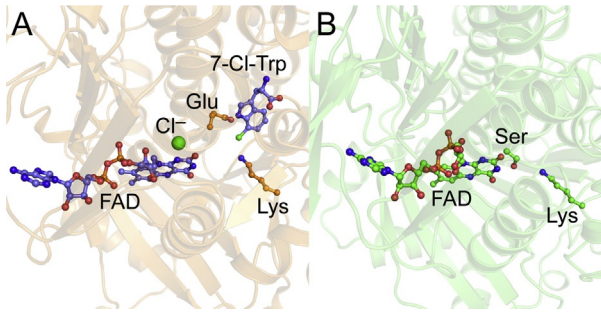

Fig. 14 Comparison of active sites of (A) L-Trp chlorinase PrnA (Dong et al., 2005) and (B) pyrrolyl-*S*-Mpy15 chlorinase Mpy16 (El Gamal et al., 2016). While the cofactor FAD and the catalytic Lys side chain are positioned analogously for the two structures, Mpy16, like other acyl-*S*-CP halogenases, is missing the key Glu residue which is replaced by a Ser. PrnA is complexed with a Cl^- anion and the product, 7-chlorotryptophan.

identity of the analogous catalytic base is still elusive for acyl-*S*-CP halogenases. Addressing these mechanistic questions will undoubtedly be aided by availability of a cocrystal structure of a FAD-dependent halogenase in complex with its cognate acyl-*S*-CP substrate. While this study specifically deals with FAD-dependent halogenases, these procedures described herein can be extended to other classes of halogenases as well. Most pertinently, numerous α-ketoglutarate-dependent nonheme Fe(II) (NHFe) halogenases have been described in literature that require substrates that are tethered to CPs (Agarwal, Miles, et al., 2017). However, the activities of several of these enzymes have not been reconstituted in vitro, perhaps due to lack of access to CP-tethered substrates.

Owing to high abundance of halides in seawater, the marine environment is an exceptionally rich source of halogenated natural products, specifically, of brominated natural products that are exclusively identified from marine sources (Blunt, Copp, Keyzers, Munro, & Prinsep, 2017). The existence of brominated natural products implies that brominases, such as Bmp2, have devised a "molecular filter" that rejects chloride being used as a substrate by the enzyme. This filter is extraordinarily selective given the ~600-fold higher concentration of chloride in seawater as compared to bromide. The chlorinases such as Mpy16, on the other hand, do not exert a similar halide selectivity filter. Chlorinases are equally adept at incorporating chloride and bromide. Hence, it is likely that the presence of chlorinated natural products in the marine environment, as by extension in the terrestrial environment as well, is a consequence of greater abundance and availability of chloride as compared to bromide.

ACKNOWLEDGMENTS

We gratefully acknowledge financial support from the NIH (R00ES026620) and Georgia Institute of Technology.

REFERENCES

Agarwal, V., Blanton, J. M., Podell, S., Taton, A., Schorn, M. A., Busch, J., et al. (2017). Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. *Nature Chemical Biology*, *13*(5), 537–543. https://doi.org/10.1038/nchembio.2330.

Agarwal, V., Diethelm, S., Ray, L., Garg, N., Awakawa, T., Dorrestein, P. C., et al. (2015). Chemoenzymatic synthesis of acyl coenzyme A substrates enables *in situ* labeling of small molecules and proteins. *Organic Letters*, *17*(18), 4452–4455. https://doi.org/10.1021/acs.orglett.5b02113.

Agarwal, V., El Gamal, A. A., Yamanaka, K., Poth, D., Kersten, R. D., Schorn, M., et al. (2014). Biosynthesis of polybrominated aromatic organic compounds by marine bacteria. *Nature Chemical Biology*, *10*(8), 640–647. https://doi.org/10.1038/nchembio.1564.

Agarwal, V., Lin, S., Lukk, T., Nair, S. K., & Cronan, J. E. (2012). Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(43), 17406–17411. https://doi.org/10.1073/pnas.1207028109.

Agarwal, V., Miles, Z. D., Winter, J. M., Eustaquio, A. S., El Gamal, A. A., & Moore, B. S. (2017). Enzymatic halogenation and dehalogenation reactions: Pervasive and mechanistically diverse. *Chemical Reviews*, *117*(8), 5619–5674. https://doi.org/10.1021/acs.chemrev.6b00571.

Beld, J., Sonnenschein, E. C., Vickery, C. R., Noel, J. P., & Burkart, M. D. (2014). The phosphopantetheinyl transferases: Catalysis of a post-translational modification crucial for life. *Natural Product Reports*, *31*(1), 61–108. https://doi.org/10.1039/c3np70054b.

Bitto, E., Huang, Y., Bingman, C. A., Singh, S., Thorson, J. S., & Phillips, G. N., Jr. (2008). The structure of flavin-dependent tryptophan 7-halogenase RebH. *Proteins*, *70*(1), 289–293. https://doi.org/10.1002/prot.21627.

Blasiak, L. C., & Drennan, C. L. (2009). Structural perspective on enzymatic halogenation. *Accounts of Chemical Research*, *42*(1), 147–155. https://doi.org/10.1021/ar800088r.

Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H. G., & Prinsep, M. R. (2017). Marine natural products. *Natural Product Reports*, *34*(3), 235–294. https://doi.org/10.1039/c6np00124f.

Brown, S., & O'Connor, S. E. (2015). Halogenase engineering for the generation of new natural product analogues. *ChemBioChem*, *16*(15), 2129–2135. https://doi.org/10.1002/cbic.201500338.

Chankhamjon, P., Tsunematsu, Y., Ishida-Ito, M., Sasa, Y., Meyer, F., Boettger-Schmidt, D., et al. (2016). Regioselective dichlorination of a non-activated aliphatic carbon atom and phenolic bismethylation by a multifunctional fungal flavoenzyme. *Angewandte Chemie (International ed. in English)*, *55*(39), 11955–11959. https://doi.org/10.1002/anie.201604516.

Dong, C., Flecks, S., Unversucht, S., Haupt, C., van Pee, K. H., & Naismith, J. H. (2005). Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. *Science (New York, NY)*, *309*(5744), 2216–2219. https://doi.org/10.1126/science.1116510.

Dorrestein, P. C., Bumpus, S. B., Calderone, C. T., Garneau-Tsodikova, S., Aron, Z. D., Straight, P. D., et al. (2006). Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry*, *45*(42), 12756–12766. https://doi.org/10.1021/bi061169d.

Dorrestein, P. C., & Kelleher, N. L. (2006). Dissecting non-ribosomal and polyketide biosynthetic machineries using electrospray ionization Fourier-Transform mass spectrometry. *Natural Product Reports*, *23*(6), 893–918. https://doi.org/10.1039/b511400b.

Dorrestein, P. C., Yeh, E., Garneau-Tsodikova, S., Kelleher, N. L., & Walsh, C. T. (2005). Dichlorination of a pyrrolyl-S-carrier protein by FADH2-dependent halogenase PltA during pyoluteorin biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(39), 13843–13848. https://doi.org/10.1073/pnas.0506964102.

Eichhorn, E., van der Ploeg, J. R., & Leisinger, T. (1999). Characterization of a two-component alkanesulfonate monooxygenase from *Escherichia coli*. *Journal of Biological Chemistry*, *274*(38), 26639–26646.

El Gamal, A., Agarwal, V., Diethelm, S., Rahman, I., Schorn, M. A., Sneed, J. M., et al. (2016). Biosynthesis of coral settlement cue tetrabromopyrrole in marine bacteria by a uniquely adapted brominase–thioesterase enzyme pair. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(14), 3797–3802. https://doi.org/10.1073/pnas.1519695113.

Flecks, S., Patallo, E. P., Zhu, X., Ernyei, A. J., Seifert, G., Schneider, A., et al. (2008). New insights into the mechanism of enzymatic chlorination of tryptophan. *Angewandte Chemie (International ed. in English)*, *47*(49), 9533–9536. https://doi.org/10.1002/anie.200802466.

Fraley, A. E., Garcia-Borras, M., Tripathi, A., Khare, D., Mercado-Marin, E. V., Tran, H., et al. (2017). Function and structure of MalA/MalA', iterative halogenases for late-stage C–H functionalization of indole alkaloids. *Journal of the American Chemical Society*, *139*(34), 12060–12068. https://doi.org/10.1021/jacs.7b06773.

Franke, J., & Hertweck, C. (2016). Biomimetic thioesters as probes for enzymatic assembly lines: Synthesis, applications, and challenges. *Cell Chemical Biology*, *23*(10), 1179–1192. https://doi.org/10.1016/j.chembiol.2016.08.014.

He, J., Magarvey, N., Piraee, M., & Vining, L. C. (2001). The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. *Microbiology*, *147*(Pt. 10), 2817–2829. https://doi.org/10.1099/00221287-147-10-2817.

Hofer, I., Crusemann, M., Radzom, M., Geers, B., Flachshaar, D., Cai, X., et al. (2011). Insights into the biosynthesis of hormaomycin, an exceptionally complex bacterial signaling metabolite. *Chemistry and Biology*, *18*(3), 381–391. https://doi.org/10.1016/j.chembiol.2010.12.018.

Johannes, T. W., Woodyer, R. D., & Zhao, H. (2007). Efficient regeneration of NADPH using an engineered phosphite dehydrogenase. *Biotechnology and Bioengineering*, *96*(1), 18–26. https://doi.org/10.1002/bit.21168.

Jordan, P. A., & Moore, B. S. (2016). Biosynthetic pathway connects cryptic ribosomally synthesized posttranslationally modified peptide genes with pyrroloquinoline alkaloids. *Cell Chemical Biology*, *23*(12), 1504–1514. https://doi.org/10.1016/j.chembiol.2016.10.009.

Latham, J., Brandenburger, E., Shepherd, S. A., Menon, B. R. K., & Micklefield, J. (2018). Development of halogenase enzymes for use in synthesis. *Chemical Reviews*, *118*(1), 232–269. https://doi.org/10.1021/acs.chemrev.7b00032.

Lin, S., Van Lanen, S. G., & Shen, B. (2007). Regiospecific chlorination of (S)-beta-tyrosyl-S-carrier protein catalyzed by SgcC3 in the biosynthesis of the enediyne antitumor antibiotic C-1027. *Journal of the American Chemical Society*, *129*(41), 12432–12438. https://doi.org/10.1021/ja072311g.

Mantovani, S. M., & Moore, B. S. (2013). Flavin-linked oxidase catalyzes pyrrolizine formation of dichloropyrrole-containing polyketide extender unit in chlorizidine A. *Journal of the American Chemical Society*, *135*(48), 18032–18035. https://doi.org/10.1021/ja409520v.

Menon, B. R. K., Brandenburger, E., Sharif, H. H., Klemstein, U., Shepherd, S. A., Greaney, M. F., et al. (2017). RadH: A versatile halogenase for integration into synthetic pathways. *Angewandte Chemie (International ed. in English)*, *56*(39), 11841–11845. https://doi.org/10.1002/anie.201706342.

Mishra, P. K., & Drueckhammer, D. G. (2000). Coenzyme A analogues and derivatives: Synthesis and applications as mechanistic probes of coenzyme A ester-utilizing enzymes. *Chemical Reviews*, *100*(9), 3283–3310.

Nazi, I., Koteva, K. P., & Wright, G. D. (2004). One-pot chemoenzymatic preparation of coenzyme A analogues. *Analytical Biochemistry*, *324*(1), 100–105.

Pang, A. H., Garneau-Tsodikova, S., & Tsodikov, O. V. (2015). Crystal structure of halogenase PltA from the pyoluteorin biosynthetic pathway. *Journal of Structural Biology*, *192*(3), 349–357. https://doi.org/10.1016/j.jsb.2015.09.013.

Payne, J. T., Andorfer, M. C., & Lewis, J. C. (2016). Engineering flavin-dependent halogenases. *Methods in Enzymology*, *575*, 93–126. https://doi.org/10.1016/bs.mie.2016.03.024.

Peter, D. M., Vogeli, B., Cortina, N. S., & Erb, T. J. (2016). A chemo-enzymatic road map to the synthesis of CoA esters. *Molecules*, *21*(4), 517. https://doi.org/10.3390/molecules21040517.

Pinchman, J. R., & Boger, D. L. (2013a). Investigation into the functional impact of the vancomycin C-ring aryl chloride. *Bioorganic and Medicinal Chemistry Letters*, *23*(17), 4817–4819. https://doi.org/10.1016/j.bmcl.2013.06.080.

Pinchman, J. R., & Boger, D. L. (2013b). Probing the role of the vancomycin e-ring aryl chloride: Selective divergent synthesis and evaluation of alternatively substituted E-ring analogues. *Journal of Medicinal Chemistry*, *56*(10), 4116–4124. https://doi.org/10.1021/jm4004494.

Podzelinska, K., Latimer, R., Bhattacharya, A., Vining, L. C., Zechel, D. L., & Jia, Z. (2010). Chloramphenicol biosynthesis: The structure of CmlS, a flavin-dependent halogenase showing a covalent flavin-aspartate bond. *Journal of Molecular Biology*, *397*(1), 316–331. https://doi.org/10.1016/j.jmb.2010.01.020.

Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., & Walsh, C. T. (1998). Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry*, *37*(6), 1585–1595. https://doi.org/10.1021/bi9719861bi9719861.

Reuter, K., Mofid, M. R., Marahiel, M. A., & Ficner, R. (1999). Crystal structure of the surfactin synthetase-activating enzyme sfp: A prototype of the 4'-phosphopantetheinyl transferase superfamily. *EMBO Journal*, *18*(23), 6823–6831. https://doi.org/10.1093/emboj/18.23.6823.

Ross, A. C., Gulland, L. E., Dorrestein, P. C., & Moore, B. S. (2015). Targeted capture and heterologous expression of the *Pseudoalteromonas* alterochromide gene cluster in *Escherichia coli* represents a promising natural product exploratory platform. *ACS Synthetic Biology*, *4*(4), 414–420. https://doi.org/10.1021/sb500280q.

Sato, M., Winter, J. M., Kishimoto, S., Noguchi, H., Tang, Y., & Watanabe, K. (2016). Combinatorial generation of chemical diversity by redox enzymes in chaetoviridin biosynthesis. *Organic Letters*, *18*(6), 1446–1449. https://doi.org/10.1021/acs.orglett.6b00380.

Schmartz, P. C., Zerbe, K., Abou-Hadeed, K., & Robinson, J. A. (2014). Bis-chlorination of a hexapeptide-PCP conjugate by the halogenase involved in vancomycin biosynthesis. *Organic & Biomolecular Chemistry*, *12*(30), 5574–5577. https://doi.org/10.1039/c4ob00474d.

Schnepel, C., Minges, H., Frese, M., & Sewald, N. (2016). A high-throughput fluorescence assay to determine the activity of tryptophan halogenases. *Angewandte Chemie (International ed. in English)*, *55*(45), 14159–14163. https://doi.org/10.1002/anie.201605635.

Strauss, E., & Begley, T. P. (2002). The antibiotic activity of N-pentylpantothenamide results from its conversion to ethyldethia-coenzyme A, a coenzyme A antimetabolite. *Journal of Biological Chemistry*, *277*(50), 48205–48209. https://doi.org/10.1074/jbc.M204560200M204560200.

Thomas, M. G., Burkart, M. D., & Walsh, C. T. (2002). Conversion of L-proline to pyrrolyl-2-carboxyl-S-PCP during undecylprodigiosin and pyoluteorin biosynthesis. *Chemistry and Biology*, *9*(2), 171–184. https://doi.org/10.1016/S1074-5521(02)00100-X.

van Wageningen, A. M., Kirkpatrick, P. N., Williams, D. H., Harris, B. R., Kershaw, J. K., Lennard, N. J., et al. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chemistry and Biology*, *5*(3), 155–162.

Walsh, C. T., Garneau-Tsodikova, S., & Howard-Jones, A. R. (2006). Biological formation of pyrroles: Nature's logic and enzymatic machinery. *Natural Product Reports*, *23*(4), 517–531. https://doi.org/10.1039/b605245m.

Worthington, A. S., & Burkart, M. D. (2006). One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Organic & Biomolecular Chemistry*, *4*(1), 44–46. https://doi.org/10.1039/b512735a.

Yamanaka, K., Ryan, K. S., Gulder, T. A., Hughes, C. C., & Moore, B. S. (2012). Flavoenzyme-catalyzed atropo-selective N,C-bipyrrole homocoupling in marinopyrrole biosynthesis. *Journal of the American Chemical Society*, *134*(30), 12434–12437. https://doi.org/10.1021/ja305670f.

Yeh, E., Blasiak, L. C., Koglin, A., Drennan, C. L., & Walsh, C. T. (2007). Chlorination by a long-lived intermediate in the mechanism of flavin-dependent halogenases. *Biochemistry*, *46*(5), 1284–1292. https://doi.org/10.1021/bi0621213.

Zhu, X., De Laurentis, W., Leang, K., Herrmann, J., Ihlefeld, K., van Pee, K. H., et al. (2009). Structural insights into regioselectivity in the enzymatic chlorination of tryptophan. *Journal of Molecular Biology*, *391*(1), 74–85. https://doi.org/10.1016/j.jmb.2009.06.008.