

Macrocyclization and Backbone Rearrangement During RiPP Biosynthesis by a SAM-Dependent Domain-of-Unknown-Function 692

Richard S. Ayikpoe, Lingyang Zhu, Jeff Y. Chen, Chi P. Ting, and Wilfred A. van der Donk*



Cite This: *ACS Cent. Sci.* 2023, 9, 1008–1018



Read Online

ACCESS |

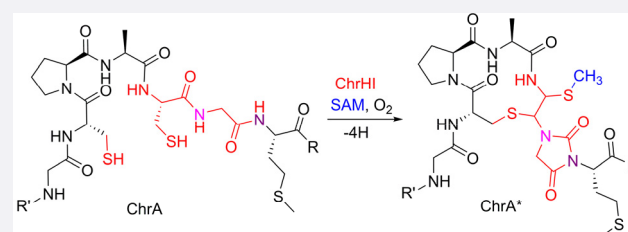
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: The domain of unknown function 692 (DUF692) is an emerging family of post-translational modification enzymes involved in the biosynthesis of ribosomally synthesized and post-translationally modified peptide (RiPP) natural products. Members of this family are multinuclear iron-containing enzymes, and only two members have been functionally characterized to date: MbnB and TglH. Here, we used bioinformatics to select another member of the DUF692 family, ChrH, that is encoded in the genomes of the *Chryseobacterium* genus along with a partner protein ChrI. We

structurally characterized the ChrH reaction product and show that the enzyme complex catalyzes an unprecedented chemical transformation that results in the formation of a macrocycle, an imidazolidinedione heterocycle, two thioaminals, and a thiomethyl group. Based on isotopic labeling studies, we propose a mechanism for the four-electron oxidation and methylation of the substrate peptide. This work identifies the first SAM-dependent reaction catalyzed by a DUF692 enzyme complex, further expanding the repertoire of remarkable reactions catalyzed by these enzymes. Based on the three currently characterized DUF692 family members, we suggest the family be called multinuclear non-heme iron dependent oxidative enzymes (MNIOs).



INTRODUCTION

Recent advances in bioinformatic tools and an explosion in the publicly available genomic data have led to the identification of many new peptide-based natural products.^{1–6} Often, these compounds possess antimicrobial, antiviral, antifungal, herbicidal, or cytotoxic properties.^{7–12} Ribosomally synthesized and post-translationally modified peptides (RiPPs) have garnered significant attention over the past decade owing to their remarkable structural diversity and biological activities.^{10,12–14} RiPPs are produced from genetically encoded precursor peptides that are modified by post-translational modification (PTM) enzymes encoded in the same biosynthetic gene clusters (BGCs). Many of the post-translational modification reactions are unique chemical transformations resulting in macrocycles and heterocycles, structures that are privileged in bioactive scaffolds. Furthermore, exciting recent developments in biocatalysis have increasingly shown the value of enzymatic transformations in preparation of compounds that are useful to human society, with many of the enzymes coming from biosynthetic pathways to natural products.^{15–18} Therefore, discovery of new chemical reactions catalyzed by enzymes and assignment of function to poorly characterized enzyme families is an important goal.

A common biosynthetic feature of most RiPPs is the ribosomal production of a precursor peptide consisting of a leader and a core sequence. The leader peptide facilitates recognition and processing by the PTM enzymes,¹⁹ with

modifications imparted to the core peptide. The modified peptide undergoes proteolytic removal of the leader peptide, and the matured peptide is in most cases exported from the cell to exert its biological function. More than 40 families of RiPPs have been classified based on the chemical transformations made to the precursor peptide.^{13,14} One emerging and underexplored group of RiPPs is generated by multinuclear iron-dependent enzymes belonging to the DUF692 family. Members of this enzyme family are structurally related to the triose-phosphate isomerase family,²⁰ and only two members of the DUF692 enzyme family have been functionally characterized thus far, MbnB^{21–23} and TglH.^{24,25}

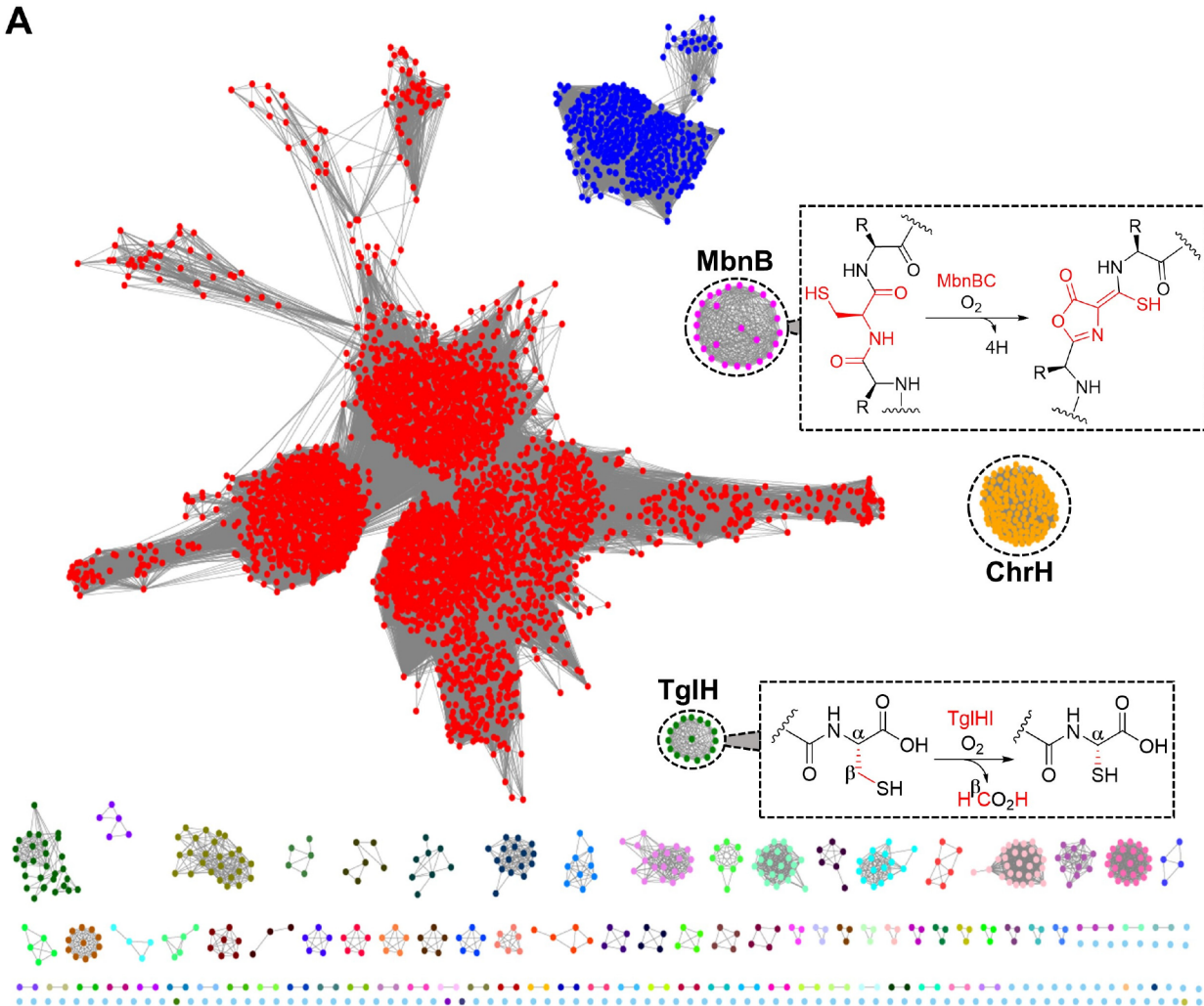
MbnB forms a heterodimeric complex with MbnC to catalyze a central step in the biosynthesis of methanobactin, a copper-chelating peptidic compound produced by methanotrophic bacteria under copper limiting conditions. During the maturation of methanobactin, MbnBC catalyzes the four-electron oxidation of Cys residues in its precursor peptide, MbnA, to an oxazolone and an adjacent thioamide group (Figure 1A).^{21–23} The second characterized member, TglH, is

Received: February 7, 2023

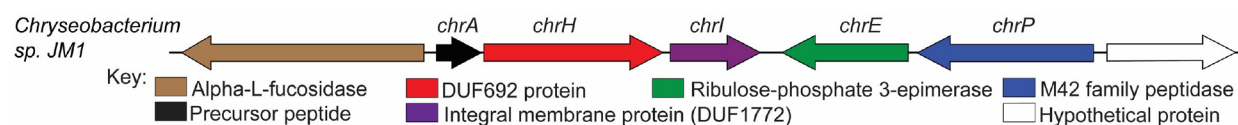
Published: April 24, 2023



A



B



C

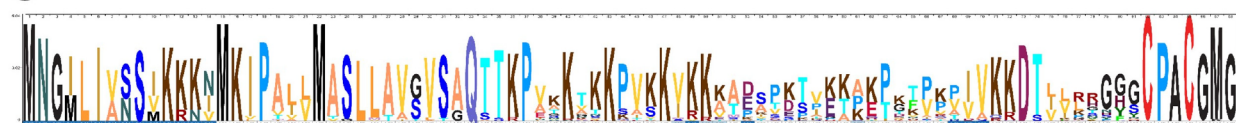


Figure 1. Bioinformatic analysis of the DUF692 family. (A) Sequence similarity network of DUF692 enzymes using an E-value of 50. The orthologs of MbnB are colored magenta and the orthologs of TgIH in green. The group of BGCs investigated in this work are colored orange. The cytoscape file for the SSN is provided in the SI. (B) Representative gene organization of the *Chryseobacterium* biosynthetic gene clusters. (C) Sequence logo of the precursor peptides from 115 BGCs showing conservation of a terminal CPACGMG motif. An Excel file containing precursor peptide sequences used to generate the sequence logo is provided in the SI.

involved in the maturation of 3-thiaglutamate, an amino acid-derived natural product biosynthesized at the C-terminus of a carrier peptide.²⁴ TgIH forms a complex with a second protein TgII to catalyze the β -carbon excision of a C-terminal Cys residue to generate a 2-mercaptoglycine residue (Figure 1A).^{24–26} Although they belong to the same enzyme family, TgIH and MbnB catalyze completely different but equally remarkable chemical transformations on their substrate peptides. This precedent motivated us to investigate additional members of the DUF692 family. Herein, we used bio-

informatics to uncover a new RiPP BGC that encodes a DUF692 homologue. This BGC is present predominantly in the genomes of several members of the *Chryseobacterium* genus. We therefore propose the name chryseobasin for the natural product of the pathway. Structural investigations of the product of the DUF692 enzyme revealed an unprecedented chemical transformation distinct from those catalyzed by the two previously characterized members.

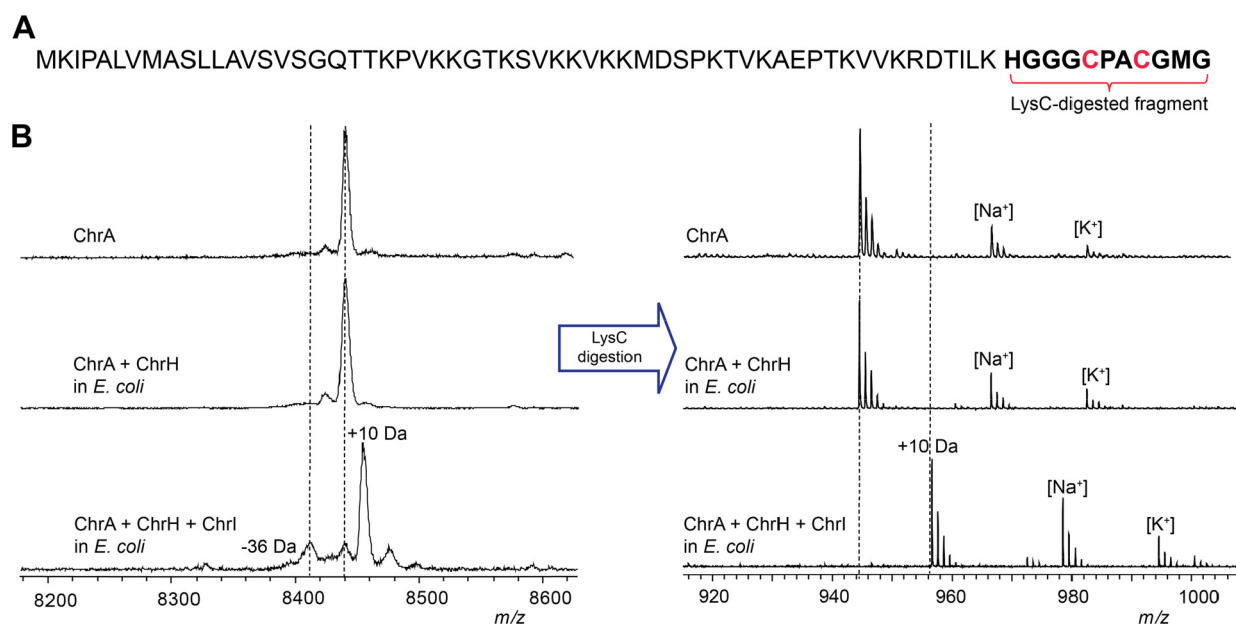


Figure 2. Modification of ChrA by ChrHI in *E. coli*. (A) Sequence of ChrA showing the C-terminal fragment after digestion with endoprotease LysC. (B) MALDI-TOF MS spectra of the unmodified and modified full-length ChrA (left) and the corresponding LysC-digested peptides (right).

RESULTS AND DISCUSSION

RiPP Biosynthetic Pathways That Encode Uncharacterized DUF692 Enzymes. To identify additional members of the DUF692 family, we generated a sequence similarity network (SSN) using the Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST).²⁷ A total of 13,108 sequences of the PF05114 family were used in the initial SSN analysis. For the SSN shown in Figure 1, an E-value of 10^{-50} and a sequence alignment score threshold of 70 were used resulting in more than 100 groups. These groups were colored based on the gene context of each cluster, and analyzed using the EFI's genome neighborhood network (GNN) tool.²⁸ The analysis identified clusters that are associated with the two characterized members TglH and MbnB (Figure 1A). In addition, the analyses also revealed additional clustered family members including a group of enzymes predominantly found in the *Chryseobacterium* genus. We selected this group of BGCs for further investigation because of the gene organization and the sequences of the putative precursor peptides (Figure 1B and C), which suggested products distinct from methanobactin and 3-thiaglutarate.

This group of BGCs encodes proteins annotated as a precursor peptide (ChrA; Uniprot ID: IW22_14840), a DUF692 enzyme (ChrH; Uniprot ID: IW22_14845), a DUF1772-containing predicted integral membrane protein (ChrI; Uniprot IDs: IW22_14850), a ribulose phosphate 3-epimerase (ChrE), and an M42 peptidase (ChrP). A sequence logo of the precursor peptides from 115 BGCs revealed a highly conserved CPACGMG motif at the C-terminus (Figure 1C). While our bioinformatic discovery was based on an enzyme SSN followed by GNN analysis, similar BGCs were also recently identified using a co-occurrence based approach.²⁹

ChrH Requires ChrI to Modify the CPACGMG Motif of ChrA. The two previously characterized members of the DUF692 family both modify cysteine residues on their substrates. We therefore anticipated that the conserved Cys residues in ChrA (Figure 2A) might be modified by ChrH. To

probe this hypothesis, we employed a heterologous coexpression approach. The precursor peptide ChrA was expressed in *Escherichia coli* with an N-terminal His₆-tag. Purification of ChrA by immobilized metal affinity chromatography (IMAC) and subsequent analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) gave the expected mass for the unmodified linear peptide (Figure 2B). When ChrA was coexpressed with ChrH in lysogeny broth (LB) supplemented with 50 μ M iron(II) citrate, no change in the mass of the purified peptide was observed.

Both MbnB and TglH require a second protein for activity.^{21,24} MbnB forms a heterodimer with MbnC,^{22,23} and TglH forms a predicted heterodimer with TglI.²⁵ Both MbnC and TglI recognize the leader peptides of their cognate substrates by formation of an antiparallel β -sheet,²² with TglI having the typical fold of a RiPP precursor peptide Recognition Element (RRE).³⁰ The integral membrane protein ChrI is not homologous in sequence to MbnC or TglI and was not bioinformatically predicted to contain an RRE,³¹ nor do any of the other enzymes in the *chr* BGC contain a predicted RRE. ChrI does contain a DUF1772 that has been speculated to play a role in protein–protein interactions.³² We therefore next included ChrI in the expression system. Following coexpression of all three proteins, a new peak consistent with a 10 Da increase in mass of ChrA was observed by MALDI-TOF MS (Figure 2B). Digestion of isolated modified and unmodified peptide with endoprotease LysC suggested the modification is localized to the C-terminus of ChrA (Figure 2B). The 10 Da mass increase was confirmed by high-resolution mass spectrometric (HRMS) analysis (Figures S1 and S2). We designate this LysC-digested peptide ChrA*.

To determine if the conserved Cys residues were modified, we first derivatized ChrA and ChrA* with *N*-ethylmaleimide (NEM) to probe for free thiols. Analysis of the derivatized peptides by MALDI-TOF MS showed that ChrHI-processed ChrA did not react with NEM (Figure S3), suggesting that the thiols have been modified. To corroborate this conclusion, the

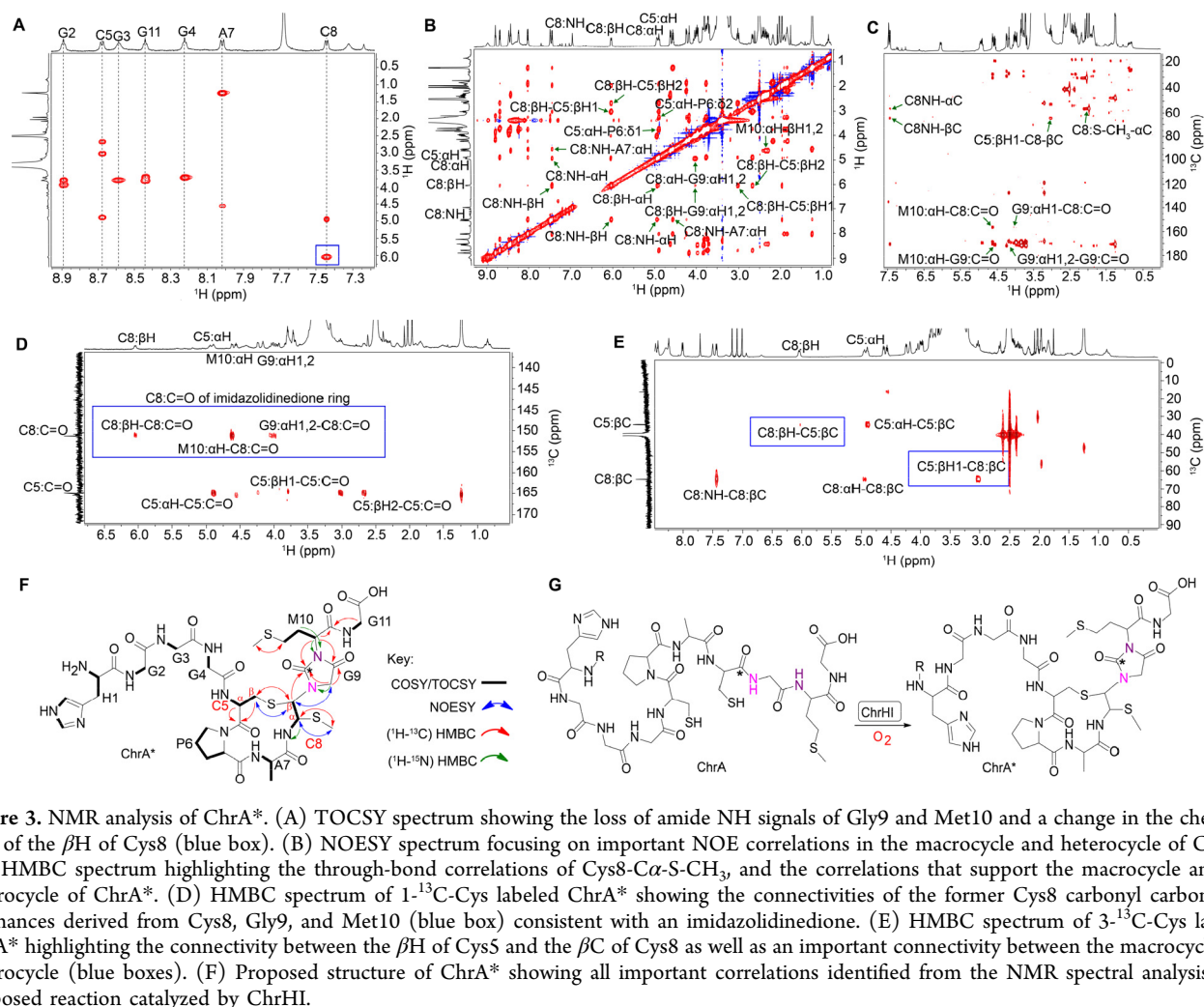


Figure 3. NMR analysis of ChrA*. (A) TOCSY spectrum showing the loss of amide NH signals of Gly9 and Met10 and a change in the chemical shift of the β H of Cys8 (blue box). (B) NOESY spectrum focusing on important NOE correlations in the macrocycle and heterocycle of ChrA*. (C) HMBC spectrum highlighting the through-bond correlations of Cys8-C α -S-CH $_3$, and the correlations that support the macrocycle and the heterocycle of ChrA*. (D) HMBC spectrum of 1- 13 C-Cys labeled ChrA* showing the connectivities of the former Cys8 carbonyl carbon with resonances derived from Cys8, Gly9, and Met10 (blue box) consistent with an imidazolidinedione. (E) HMBC spectrum of 3- 13 C-Cys labeled ChrA* highlighting the connectivity between the β H of Cys5 and the β C of Cys8 as well as an important connectivity between the macrocycle and heterocycle (blue boxes). (F) Proposed structure of ChrA* showing all important correlations identified from the NMR spectral analysis. (G) Proposed reaction catalyzed by ChrHI.

two conserved Cys residues in ChrA, Cys63 and Cys66, were mutated to Ser residues. Neither single nor double mutants resulted in formation of the product that had increased by 10 Da, suggesting that both Cys residues are modified by ChrH (Figure S4). The ChrA-C63S variant resulted in partial formation of a product with a decrease in mass of 36 Da that will be further discussed below. Taken together, these results suggest that ChrH, like TglH and MbnB, requires a complex with a second protein to install post-translational modifications on ChrA and that this transformation involves the thiols of the two conserved Cys residues of ChrA.

Structure Elucidation of ChrA*. To gain more detailed structural information regarding the reaction product of ChrHI, both unmodified and modified ChrA were produced on a larger scale and subjected to detailed one-dimensional and two-dimensional nuclear magnetic resonance (NMR) analysis after endoproteinase LysC digestion to reduce the size to an 11mer peptide. Analysis of TOCSY data for LysC-digested ChrA identified all 11 amino acids including nine amide protons (Figure S5 and Table S1), whereas only seven spin systems were identified for ChrA* associated with just seven amide protons (Figures 3A, S6, and Table S2). Significantly, the amide protons of Gly9 and Met10 (Gly67 and Met68 of full-length peptide) were missing in ChrA*, suggesting that these two amides were modified (Figures 3A and S6). Subsequent ^1H - ^{13}C -HSQC analysis of substrate and product

also revealed a new cross peak with chemical shifts at 1.96 (^1H) and 10 ppm (^{13}C) (Figure S6). Interestingly, this new peak integrated in the ^1H spectrum to three protons and appeared as a singlet suggesting the possible introduction of a methyl group. The ^1H and ^{13}C chemical shifts of this methyl group suggest attachment to either carbon or a heteroatom like sulfur that is not highly electronegative.

In addition, the HSQC data identified the disappearance of the CH $_2$ group at the β position of the former Cys8. Instead, a new CH cross peak was observed in ChrA* at 6.04 ppm, a significant change in the chemical shift compared to the original β -protons of Cys8 of ChrA (3.36 and 2.96 ppm; Figure S6 and Tables S1 and S2). As we will show, this new signal arises from one of the β -protons of the former Cys8, and we will refer to it as such in the remainder of the discussion. The chemical shift of the α H of the former Cys8 also changed from 4.44 ppm in ChrA to 4.94 ppm in ChrA* (Figures S6 and Tables S1 and S2). The assignments of the new signals originating from the α and β protons of the former Cys8 in ChrA* were supported by the ^1H - ^1H dqCOSY spectrum, which showed cross-peaks between the proton at 4.94 ppm and the NH of the former Cys8, and between the two protons at 4.94 and 6.04 ppm (Figure S6).

To probe the chemical environment of the new signals further, ^1H - ^{13}C -HMBC and ^1H - ^1H NOESY experiments were carried out on ChrA*. The NOESY spectra showed direct

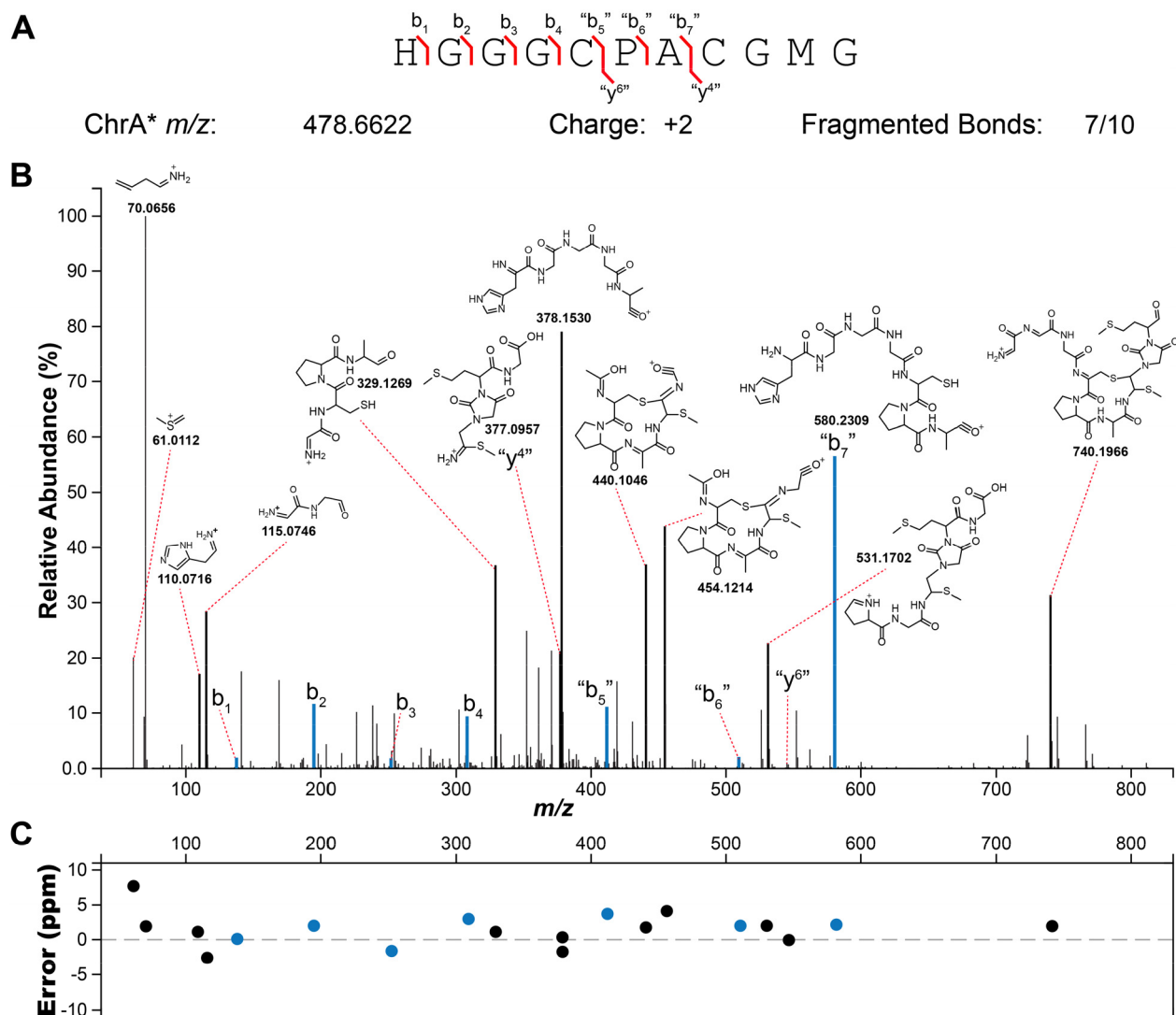


Figure 4. HR-ESI tandem mass spectrometry analysis of ChrA* digested with LysC. (A) Sequence of ChrA* and observed fragment ions. (B) Tandem mass spectrum showing ions and their corresponding fragments. (C) A graph of the ppm errors for each identified ion in panel B.³³

correlations between the new methyl group and the αH of the former Cys8. In addition, a second NOE correlation was observed from the βH of Cys5 to the signal at 6.04 ppm associated with the βH of the former Cys8 (Figures 3B and S6), as well as a weaker NOE between the βH of the former Cys5 and the signal at 4.94 ppm associated with the αH of the former Cys8. The HMBC data also showed connectivity (via 3-bond correlations) between the new methyl group and the αC of the former Cys8 and vice versa. Another connectivity between the βC of Cys5 and the ^1H peak at 6.04 ppm (former βH of Cys8) was observed, suggesting the formation of a macrocycle. The HMBC spectrum also revealed direct connectivities between the carbonyl of the former Cys8 (156.5 ppm) and the αH of Gly9 and αH of Met10, indicating a significant rearrangement in ChrA* (Figure 3C).

These preliminary experiments suggested that Cys8 is significantly altered. To probe the fate of the individual carbon atoms of Cys5 and Cys8, we first prepared ChrA* generated from ChrA that was selectively labeled with $1\text{-}^{13}\text{C}$ -Cys (label at the carbonyl carbon) by heterologous expression in *E. coli* for HMBC NMR analysis. As expected, two carbon peaks, one at 170.0 ppm for Cys5 and the other at 156.5 ppm for the former

Cys8, were observed in the ^{13}C NMR spectrum (Figure S7), corroborating a significant rearrangement of the position of the carbonyl carbon of the former Cys8. The HMBC spectrum also revealed direct correlations between the βH of the former Cys8 to the carbonyl of the former Cys8, the αH of Met10 to the carbonyl carbons of Gly9 as well as the former Cys8, and from the αH of Gly9 to the carbonyl carbons of Gly9 and the former Cys8 (Figures 3D and S7). When ChrA* was labeled with $3\text{-}^{13}\text{C}$ -Cys (label at the β -carbon), the acquired HMBC data showed through-bond connectivities from the βC of Cys5 to the βH of the former Cys8 and vice versa, suggesting that the sulfur of Cys5 is directly connected to the $\text{C}\beta$ of the former Cys8 (Figures 3E and S8). Collectively, all collected NMR data are consistent with the structure shown in Figure 3F. The stereochemistry at the two new stereogenic centers of the former Cys8 is currently not known. However, the $^3J_{\text{H-H}}$ coupling constant between the protons at 4.94 and 6.04 ppm is 11.6 Hz, suggesting that the relative stereochemistry involves a *trans* arrangement.

To corroborate the findings from the NMR data, ChrA* was next analyzed by high-resolution electrospray ionization tandem mass spectrometry (HR-ESI MS/MS). The HR-ESI

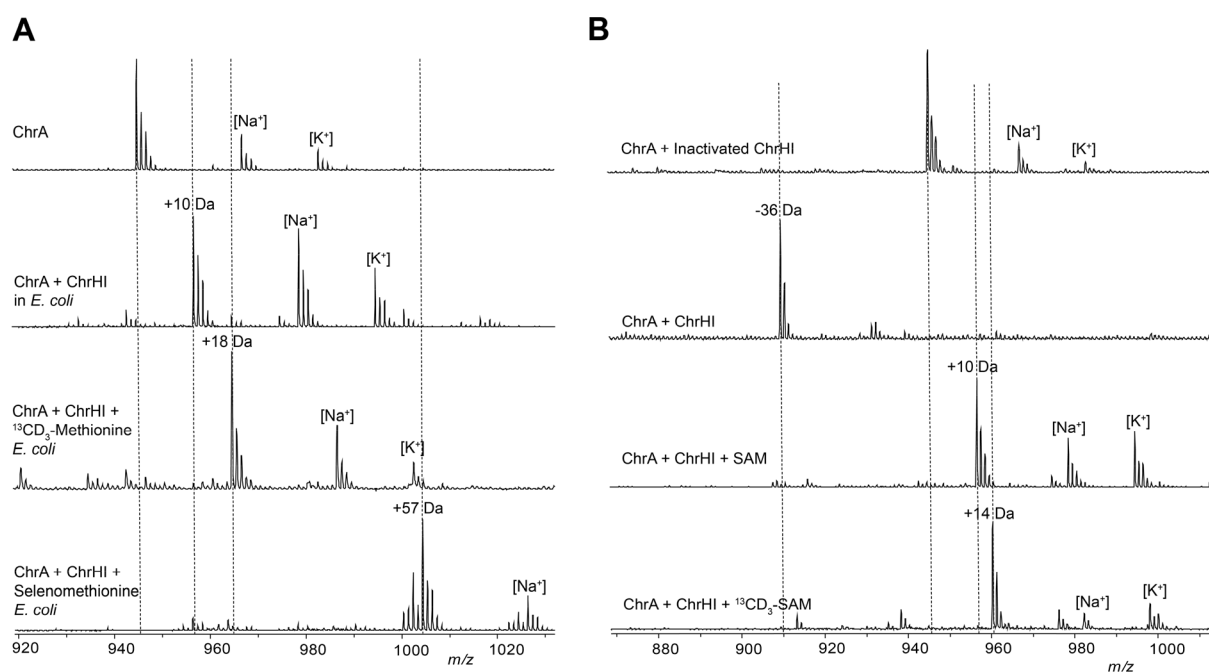


Figure 5. Determination of the methyl source and *in vitro* reconstitution of ChrHI activity. (A) MALDI-TOF MS data of ChrA* produced by heterologous coexpression of ChrA with ChrHI in M9 media supplemented with either ¹³CD₃-Met or selenomethionine and treatment of the purified peptide with endoproteinase LysC. (B) MALDI-TOF MS of LysC-digested *in vitro* reaction products of ChrHI in the presence and absence of SAM. The dashed lines indicate the expected *m/z* of modified and unmodified peptides.

MS/MS spectrum localized the modifications to the C-terminal CGMG peptide as no fragments corresponding to b-ions could be observed beyond this point, whereas all amino acids in the N-terminal segment were unmodified (Figure 4). The observation of fragment ions indicated as “y4” and “b7” (since the structure is no longer an α -amino acid peptide these fragments are not true y and b ions) is explained by cleavage of the thioaminal linkages in the proposed structure (Table S3). In addition, various internal fragment ions consistent with the proposed structure of ChrA* were observed, especially fragmentation in the two thioether bonds as well as the amide bonds of the imidazolidinedione. Taken together, the HR-ESI MS/MS and the NMR data suggest that ChrHI catalyzes an unprecedented multistep chemical transformation on ChrA: macrocyclization to form a cross-link between the thiol of Cys5 and the β C of the former Cys8, migration of the carbonyl carbon of the former Cys8 to become inserted between the amide nitrogens of Gly9 and Met10 to form an imidazolidinedione, and methylation and migration of the thiol from C β to C α of the former Cys8. Thus, the net +10 Da transformation from ChrA observed by MS involves a -4 Da change resulting from two oxidative ring formations in which the amide protons of Gly9 and Met10 are removed as well as the thiol proton of Cys5 and one of the β -protons of the former Cys8, and a +14 Da change from the methylation event.

In Vitro Reconstitution of ChrHI Activity and Identification of S-Adenosylmethionine as the Methyl Donor. The proposed structure of ChrA* generated in *E. coli* contains a new methyl group that was not present in the ribosomally generated precursor peptide. This finding raised a question regarding the source of the methyl group. To determine its origin, isotope feeding experiments were carried out. ChrA was coexpressed with ChrHI in *E. coli* in M9 minimal media supplemented with ¹³CD₃-Met.³⁴ Isolation of the modified peptide and subsequent MALDI-TOF MS

analysis of the full-length and LysC-digested peptide revealed the incorporation of two methyl groups from ¹³CD₃-Met into ChrA*, one in Met10 and one in the new thiomethyl group (Figures 5A and S9). When selenomethionine was substituted for ¹³CD₃-methionine, the isolated product contained four selenium atoms in the full-length modified ChrA consistent with the four Met residues in ChrA and only one selenium atom in ChrA* as evidenced by the isotope distribution pattern (Figures 5A and S10).³⁵ These results suggest that only the methyl group (and not the thiomethyl group) of methionine is incorporated into ChrA* and suggests that either S-adenosylmethionine (SAM) or possibly a methionine residue of the modifying enzyme ChrHI could be the methyl donor.

SAM is the most common methyl donor in biology.³⁶ We next aimed to reconstitute the activity of ChrHI *in vitro* to probe for the possibility of SAM as the methyl source. ChrH was heterologously expressed in and aerobically purified from *E. coli* Rosetta-pLysS cells as an N-terminally His₆-tagged protein (Figure S11A). The purified protein exhibited a purple color. The amount of iron was quantified using the ferene assay,³⁷ and the as-isolated enzyme contained 1.9 equiv of iron per monomer of ChrH. Since the coexpression studies revealed the requirement of ChrI for ChrH activity, the predicted integral membrane protein ChrI was also heterologously expressed in *E. coli* as an N-terminally hexahistidine-tagged protein and partially purified in the presence of detergent (0.01% *n*-dodecyl- β -D-maltopyranoside) (Figure S11A). Reactions were carried out aerobically by incubating ChrA with ChrH, ChrI, and DTT in the presence and absence of SAM. DTT was included because the two Cys residues in the substrate peptide readily form a disulfide (Figures S1 and S2). When the reaction without SAM was analyzed by MALDI-TOF MS, the +10 Da product was not observed, but instead, an ion indicating the loss of -36 Da from ChrA was observed

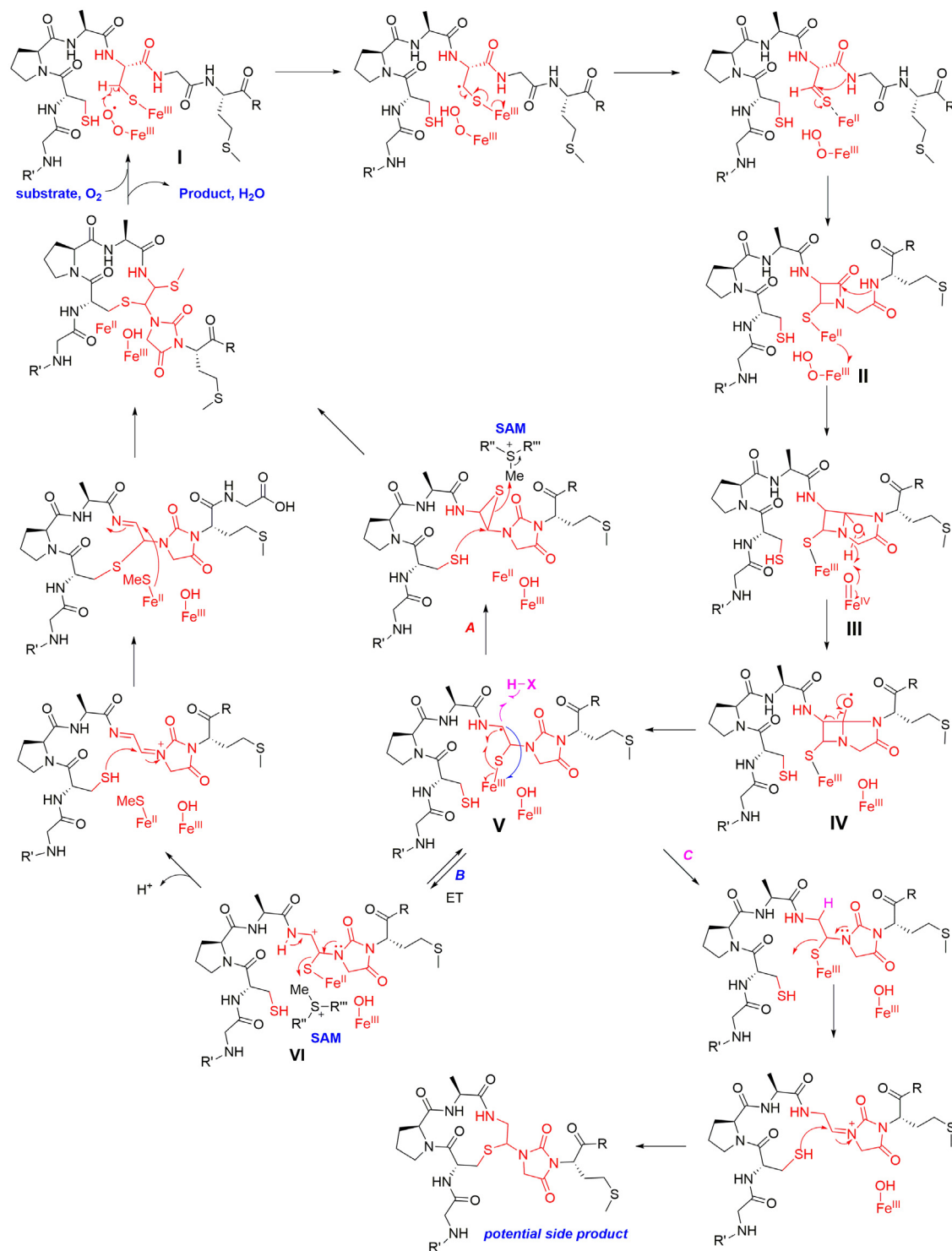


Figure 6. Proposed mechanism for ChrHI catalysis. The order of some steps can be inverted (e.g., the two C–S bond forming steps), and the timing of methylation is not known and could happen much earlier (e.g., in intermediate II or III). For some alternative mechanisms, see Figure S14.

(Figures 5B and S12). Such an ion was also observed in small amounts when ChrA was heterologously coexpressed with ChrHI in *E. coli* (Figures 2B and S3). The difference in mass between the –36 Da species and +10 Da product is 46 Da, consistent with a thiomethyl substituent. Further attempts to isolate the –36 species for detailed structural characterization proved unsuccessful due to degradation under the acidic conditions required for ChrA purification. We next analyzed

the *in vitro* reaction in the presence of SAM resulting in the formation of the +10 Da peptide as the main product together with the –36 Da species (Figures 5B and S12). These results suggest that SAM is indeed the source of the methyl group and show that *in vitro* and in *E. coli* the same product is formed, strongly suggesting that the observed transformation is the native function of ChrHI. We next synthesized ¹³CD₃-SAM enzymatically^{38–40} and used the isotopically labeled product in

the *in vitro* reaction with ChrA and ChrHI. Analysis of the full-length and LysC-digested reaction products revealed the incorporation of one $^{13}\text{CD}_3$ -methyl group from $^{13}\text{CD}_3$ -SAM into ChrA* but no label incorporation in the -36 Da product. Taken together, these results suggest that SAM is the methyl donor in the post-translational modification of ChrA by ChrHI, thus representing the first example of a DUF692-mediated reaction to use SAM and iron as cofactors.

Possible Mechanism of ChrHI Catalysis. With the structure of ChrA* as well as the source of the methyl group established, a proposed mechanism for the ChrHI-catalyzed transformation of ChrA is shown in Figure 6. DUF692 enzymes have been shown to contain two or three iron ions in their active sites (Protein Database accession numbers 3BWW, 7DZ9, 7FC0, 7TCR, 7TCX, 7TCU, and 7TCW).^{21–24} In the structure of MbnB bound to its partner MbnC and its substrate MbnA, the side chain sulfur atom of a Cys in the substrate is liganded to one of the iron atoms.²² The protein ligands to bind two or three iron ions are conserved in ChrH and present in the active site of an AlphaFold model of the enzyme (Figure S13). Akin to TglH²⁴ and MbnB,^{22,23} we suggest that the active form of ChrH requires at least two iron ions, one of which is in the Fe(II) form that is used for catalysis and a second (or third) ion that is in the Fe(III) oxidation state that is important for substrate binding and positioning. Such use of iron ions in different oxidation states for different roles is an emerging feature of multinuclear mixed-valent non-heme iron enzymes.^{41–45} It is possible that the two Cys thiols of ChrA coordinate to different irons in the active site, facilitated by the Pro-Ala sequence that separates the two Cys residues that is an inducer of β -turn formation. Because we currently do not have any structural data on the coordination of the substrate to the metal ions, the mechanism in Figure 6 shows just one of the Cys side chains liganded to Fe as in the MbnABC cocrystal structure.²²

The Fe^{II} ion in the active site of ChrH is proposed to react with molecular oxygen to form a superoxo-Fe^{III} intermediate I,^{23,46,47} which initiates the reaction by abstracting a hydrogen atom from the β -carbon of Cys8 as in the proposed mechanisms for TglH and MbnB.^{21–24} The highly reducing thioketyl radical formed during this abstraction is expected to transfer an electron to one of the Fe(III) ions to generate a thioaldehyde and Fe(II). As drawn in Figure 6, this is a different iron than the iron that activates O₂, but it could be the same iron (e.g., some of the mechanisms shown in Figure S14). Subsequent attack of the amide nitrogen of Gly9 onto the thioaldehyde forms a β -lactam II, possibly facilitated by the iron ion functioning as a Lewis acid. Formation of a similar β -lactam was also proposed for MbnBC catalysis as well as other mononuclear non-heme iron enzymes such as isopenicillin N synthase.^{21–23,48–50} At this point, the reaction of ChrHI diverges from previous enzymes in that the proposed model involves a second nucleophilic attack of the amide nitrogen of the downstream Met10 onto the carbonyl of the β -lactam, which would result in the formation of a bicyclic intermediate III. For MbnBC, the β -lactam is proposed to be attacked by the oxygen atom of the upstream amide,^{21,22} possibly with the intermediacy of an enzyme bound intermediate.²³

The transfer of an electron from Fe(II) in intermediate II to the ferric hydroperoxo intermediate results in generation of a ferryl (Fe^{IV}-oxo) species. The exact timing of the formation of this highly reactive intermediate is interesting, as it could be favorable to delay its formation until after (or concomitant

with) the formation of the bicycle in intermediate III. One possible means to delay its formation could be by having the reducing equivalent reside on an iron atom that is not bound to the peroxide as shown in Figure 6. Once formed, the ferryl could initiate a proton-coupled electron transfer from the adjacent O–H bond to generate an oxygen radical IV. A well-precedented β -scission reaction⁵¹ would result in the opening of the four-membered ring to form a resonance stabilized carbon-based radical V. Although the context and enzyme classes are quite different, β -scission by oxidation of an alcohol by a ferryl was also proposed as a key step in other non-heme iron dependent enzymes with experimental and computational support.^{42,52,53} We propose that this intermediate radical V could generate the product via formation of an episulfide followed by nucleophilic ring opening by Cys5 and methylation of the thiol (pathway A). Alternatively, oxidation of radical V to a cation VI (pathway B) followed by S-methylation of the thioaminal, migration to the C β of the former Cys, and addition of the thiol of Cys5 to the C α of the former Cys8 would complete the formation of ChrA*. We note that many of these steps could occur in a different order, that several steps are expected to require acid–base catalysis, and that other mechanisms can be drawn to arrive at the final product (e.g., Figure S14).

We currently do not know whether ChrH or ChrI is responsible for catalyzing the methylation event. An AlphaFold-multimer⁵⁴ model of the complex of ChrHI and ChrA did not predict a Rossmann fold that is typically used to bind SAM (Figure S15). ChrH and its orthologs in the genomes are ~ 80 – 100 amino acids longer than MbnB, which may be the origin of the long unstructured loop in the AlphaFold model. It is possible that the additional amino acids form a SAM binding domain in the presence of substrate. Analysis of ChrI by DeepTMHMM⁵⁵ suggests the protein is mostly composed of transmembrane helices, and indeed the AlphaFold model predicts a helical bundle for the protein. The predicted interfaces between ChrH and ChrI are quite different from that observed in the structures of MbnBC,^{22,23} and so is the manner by which ChrA is predicted to engage with ChrHI with the interaction almost entirely between ChrA and ChrH (Figure S16) compared to extensive interactions of MbnA with both MbnB and MbnC.²² Despite the major differences, which may or may not be accurate as they are based on a theoretical model, the Cys residues of ChrA are in close proximity to the iron ligands in the ChrAHI complex (Figure S16).

The mechanism in Figure 6 (and alternatives in Figure S14) may also explain the formation of the -36 Da species detected in our assays. Unlike MbnBC and TglHI, ChrHI also uses SAM as a cofactor. *In vitro* reconstitution of ChrHI revealed that in the absence of SAM, only a -36 Da species accumulates as evidenced by MALDI-TOF MS (Figures 5 and S12). This compound cannot be generated by elimination of methanethiol from ChrA*, as that would lead to a product that would be 34 Da lighter than ChrA. We propose that in the absence of SAM, radical V does not form the episulfide (pathway A) or the cation VI (pathway B). Instead, the radical V could be reduced leading to the product shown in pathway C (Figure 6). Regardless of the exact mechanism and the actual structure of the -36 Da product, methylation appears critical to arrive at the structure of ChrA*.

The mechanisms in Figures 6 and S14 can potentially also explain the outcome with the ChrA variants C63S and C66S. Since the proposed chemistry is initiated at Cys66,

replacement with a Ser residue abolishes all activity. In contrast, replacement of Cys63 with Ser can still result in the initial oxidation steps on Cys66. If macrocycle formation involving Cys63 is required for either methylation or thiomethyl migration to the former α -carbon then once again a shunt product can be formed that is 36 Da decreased in mass from the starting peptide (Figure S17).

CONCLUSIONS

Using bioinformatics, we identified many groups of uncharacterized enzymes that belong to the DUF692 enzyme family. By combining biochemical assays, mass spectrometric and spectroscopic experiments, we demonstrate that one group of such enzymes catalyzes a peptide backbone rearrangement to form an imidazolidinedione heterocycle adding to the current repertoire of heterocycles formed in RiPPs.¹³ In addition to heterocycle formation, ChrHI installs a thioether macrocycle from two conserved Cys residues and methylates a thiohemiaminal using SAM. Future spectroscopic studies and possibly structural information on the observed side products may provide further insights into the mechanism of the remarkable overall process.

With three different DUF692 reactions now characterized, some commonalities are starting to emerge. All three proteins (MbnB, TglH, and ChrH) catalyze four-electron oxidations of their peptide substrates, apparently without requiring reducing equivalents. All three proteins also act on Cys residues, although whether this will hold for the entire family remains to be established. All three enzymes catalyze rearrangements that result in attachment of the sulfur atom of Cys and an amide nitrogen to the same carbon, the former β -carbon of Cys in the case of MbnBC and the former α -carbon of Cys for TglHI and ChrHI. For MbnB and ChrH, the net transformation involves removal of four hydrogens, likely with the formation of two water molecules (oxidase), whereas for TglH oxygenation chemistry is presumably involved in generating formate; whether formate generation involves monooxygenase or dioxygenase chemistry is at present not known. All three proteins seem to bind at least two iron ions and have the ligand set for binding three irons. Presently only for MbnB has activity been correlated with the amount of Fe(II) and Fe(III) in the protein,²³ but it is likely that all three enzymes will be utilizing a mixed-valent state for catalysis. While the reactions catalyzed by the majority of DUF692 enzymes remain to be determined (Figure 1), the characterized transformations are all oxidations of the substrates using O₂ catalyzed by multinuclear iron enzymes. Because these enzymes differ in fold from other families of iron-dependent enzymes,⁴² we suggest to replace the designation domain-of-unknown function 692 with multinuclear non-heme iron dependent oxidative enzymes (MNIOs).

The final structure of chryseobasin remains to be determined and will require reconstitution of the remaining two enzymes encoded in the BGC, a protease ChrP and a putative epimerase ChrE, which has not yet been achieved to date. Their characterization, or that of orthologs, would facilitate investigation into the biological function of chryseobasin. Collectively, the data in this study expand the scope of post-translational modifications in RiPP biosynthesis, demonstrate another unexpected and complex reaction catalyzed by a homologue of MbnB and TglH, and lay the foundation toward understanding the chemistry of additional members of the enzyme family formerly known as DUF692.

ASSOCIATED CONTENT

Data Availability Statement

The authors declare that the data supporting the findings of this study are available within the paper and its Supporting Information files, and at Mendeley Data, V1, doi: 10.17632/dn37cj6z5m.1, as well as from the corresponding author upon reasonable request.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscentsci.3c00160>.

Supporting data (XLSX)

Cytoscape file for Figure 1A (ZIP)

Materials and Methods, including supporting figures S1–S17 and tables S1–S6 (PDF)

Transparent Peer Review report available (PDF)

AUTHOR INFORMATION

Corresponding Author

Wilfred A. van der Donk – Department of Chemistry and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana 61801 Illinois, United States; Howard Hughes Medical Institute at the University of Illinois at Urbana–Champaign, Urbana 61801 Illinois, United States; orcid.org/0000-0002-5467-7071; Email: vddonk@illinois.edu

Authors

Richard S. Ayikpoe – Department of Chemistry and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana 61801 Illinois, United States; orcid.org/0000-0001-6698-4020

Linyang Zhu – School of Chemical Sciences NMR Laboratory, University of Illinois at Urbana–Champaign, Urbana 61801 Illinois, United States; orcid.org/0000-0002-6657-271X

Jeff Y. Chen – Department of Chemistry and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana 61801 Illinois, United States; orcid.org/0000-0002-8507-8215

Chi P. Ting – Department of Chemistry and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana 61801 Illinois, United States; orcid.org/0000-0003-1386-1902

Complete contact information is available at: <https://pubs.acs.org/10.1021/acscentsci.3c00160>

Funding

This work was supported by the National Institutes of Health (F32 GM140621 to RSA and R37 GM058822 to WAV).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is dedicated to the memory of Prof. Christopher T. Walsh whose pioneering studies on enzyme mechanisms and natural product biosynthesis have been an inspiration to the authors. We thank D. T. Nguyen for his assistance with HR-MS/MS experiments. We also thank Dr. P. Yau and Dr. J. Arrington in the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana–Champaign for assistance with MS/MS data acquisition.

REFERENCES

- (1) Hetrick, K. J.; van der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Curr. Opin. Chem. Biol.* **2017**, *38*, 36–44.
- (2) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47*, W81–W7.
- (3) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H. C.; Zakai, U. I.; Mitchell, D. A. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* **2017**, *13*, 470–8.
- (4) Ren, H.; Shi, C.; Zhao, H. Computational tools for discovering and engineering natural product biosynthetic pathways. *iScience* **2020**, *23*, 100795.
- (5) Kloosterman, A. M.; Medema, M. H.; van Wezel, G. P. Omics-based strategies to discover novel classes of RiPP natural products. *Curr. Opin. Biotechnol.* **2021**, *69*, 60–7.
- (6) Baltz, R. H. Genome mining for drug discovery: progress at the front end. *J. Ind. Microbiol. Biotechnol.* **2021**, *48*, kuab044.
- (7) Wang, G.; Li, X.; Wang, Z. APD3: the antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res.* **2016**, *44*, D1087–93.
- (8) Dang, T.; Süßmuth, R. D. Bioactive peptide natural products as lead structures for medicinal use. *Acc. Chem. Res.* **2017**, *50*, 1566–76.
- (9) Sarkar, T.; Chetia, M.; Chatterjee, S. Antimicrobial peptides and proteins: from nature's reservoir to the laboratory and beyond. *Front. Chem.* **2021**, *9*, 691532.
- (10) Cao, L.; Do, T.; Link, A. J. Mechanisms of action of ribosomally synthesized and posttranslationally modified peptides (RiPPs). *J. Ind. Microbiol. Biotechnol.* **2021**, *48*, kuab005.
- (11) Lahiri, D.; Nag, M.; Dutta, B.; Sarkar, T.; Pati, S.; Basu, D.; Abdul Kari, Z.; Wei, L. S.; Smaoui, S.; Wen Goh, K.; et al. Bacteriocin: A natural approach for food safety and food security. *Front. Bioeng. Biotechnol.* **2022**, *10*, 1005918.
- (12) Ongpipattanakul, C.; Desormeaux, E. K.; DiCaprio, A.; van der Donk, W. A.; Mitchell, D. A.; Nair, S. K. Mechanism of action of ribosomally synthesized and post-translationally modified peptides. *Chem. Rev.* **2022**, *122*, 14722–814.
- (13) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; et al. New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* **2021**, *38*, 130–239.
- (14) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; et al. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013**, *30*, 108–60.
- (15) Rosenthal, K.; Bornscheuer, U. T.; Lütz, S. Cascades of evolved enzymes for the synthesis of complex molecules. *Angew. Chem., Int. Ed.* **2022**, *61*, e202208358.
- (16) Kaspar, F.; Schallmeyer, A. Chemo-enzymatic synthesis of natural products and their analogs. *Curr. Opin. Biotechnol.* **2022**, *77*, 102759.
- (17) Heath, R. S.; Ruscoe, R. E.; Turner, N. J. The beauty of biocatalysis: sustainable synthesis of ingredients in cosmetics. *Nat. Prod. Rep.* **2022**, *39*, 335–88.
- (18) Benítez-Mateos, A. I.; Roura Padrosa, D.; Paradisi, F. Multistep enzyme cascades as a route towards green and sustainable pharmaceutical syntheses. *Nat. Chem.* **2022**, *14*, 489–99.
- (19) Oman, T. J.; van der Donk, W. A. Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol.* **2010**, *6*, 9–18.
- (20) Nagano, N.; Orengo, C. A.; Thornton, J. M. One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J. Mol. Biol.* **2002**, *321*, 741–65.
- (21) Kenney, G. E.; Dassama, L. M. K.; Pandelia, M. E.; Gizzi, A. S.; Martinie, R. J.; Gao, P.; DeHart, C. J.; Schachner, L. F.; Skinner, O. S.; Ro, S. Y.; et al. The biosynthesis of methanobactin. *Science* **2018**, *359*, 1411–6.
- (22) Dou, C.; Long, Z.; Li, S.; Zhou, D.; Jin, Y.; Zhang, L.; Zhang, X.; Zheng, Y.; Li, L.; Zhu, X.; et al. Crystal structure and catalytic mechanism of the MbnBC holoenzyme required for methanobactin biosynthesis. *Cell Res.* **2022**, *32*, 302–14.
- (23) Park, Y. J.; Jodts, R. J.; Slater, J. W.; Reyes, R. M.; Winton, V. J.; Montaser, R. A.; Thomas, P. M.; Dowdle, W. B.; Ruiz, A.; Kelleher, N. L.; et al. A mixed-valent Fe(II)Fe(III) species converts cysteine to an oxazolone/thioamide pair in methanobactin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, e2123566119.
- (24) Ting, C. P.; Funk, M. A.; Halaby, S. L.; Zhang, Z.; Gonen, T.; van der Donk, W. A. Use of a scaffold peptide in the biosynthesis of amino acid-derived natural products. *Science* **2019**, *365*, 280–4.
- (25) McLaughlin, M. I.; Yu, Y.; van der Donk, W. A. Substrate recognition by the peptidyl-(S)-2-mercaptoglycine synthase TglHI during 3-thiaglutamate biosynthesis. *ACS Chem. Biol.* **2022**, *17*, 930–40.
- (26) Yu, Y.; van der Donk, W. A. Biosynthesis of 3-thia- α -amino acids on a carrier peptide. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, e2205285119.
- (27) Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sakhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta* **2015**, *1854*, 1019–37.
- (28) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI web resource for genomic enzymology tools: Leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **2019**, *58*, 4169–82.
- (29) Clark, K. A.; Seyedsayamdost, M. R. Bioinformatic atlas of radical SAM enzyme-modified RiPP natural products reveals an isoleucine-tryptophan crosslink. *J. Am. Chem. Soc.* **2022**, *144*, 17876–88.
- (30) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* **2015**, *11*, 564–70.
- (31) Kloosterman, A. M.; Shelton, K. E.; van Wezel, G. P.; Medema, M. H.; Mitchell, D. A. RRE-finder: a genome-mining tool for class-independent RiPP discovery. *mSystems* **2020**, *5*, e00267-20.
- (32) Ehrlich, K. C.; Li, P.; Scharfenstein, L.; Chang, P. K. HypC, the anthrone oxidase involved in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **2010**, *76*, 3374–7.
- (33) Brademan, D. R.; Riley, N. M.; Kwiczen, N. W.; Coon, J. J. Interactive peptide spectral annotator: A versatile web-based tool for proteomic applications. *Mol. Cell. Proteom.* **2019**, *18*, S193–S201.
- (34) Bumbak, F.; Keen, A. C.; Gunn, N. J.; Gooley, P. R.; Bathgate, R. A. D.; Scott, D. J. Optimization and $^{13}\text{C}_3$ methionine labeling of a signaling competent neurotensin receptor 1 variant for NMR studies. *Biochim. Biophys. Acta. Biomembr.* **2018**, *1860*, 1372–83.
- (35) Hu, J.; Liu, F.; Feng, N.; Ju, H. Selenium-isotopic signature toward mass spectrometric identification and enzyme activity assay. *Anal. Chim. Acta* **2019**, *1064*, 1–10.
- (36) Walsh, C. *Enzymatic reaction mechanisms*; WH Freeman & Co: Oxford, 1979.
- (37) Hennessy, D. J.; Reid, G. R.; Smith, F. E.; Thompson, S. L. Ferene—a new spectrophotometric reagent for iron. *Can. J. Chem.* **1984**, *62*, 721–4.
- (38) Dippe, M.; Brandt, W.; Rost, H.; Porzel, A.; Schmidt, J.; Wessjohann, L. A. Rationally engineered variants of S-adenosylmethionine (SAM) synthase: reduced product inhibition and synthesis of artificial cofactor homologues. *Chem. Commun.* **2015**, *51*, 3637–40.
- (39) McCarty, R. M.; Krebs, C.; Bandarian, V. Spectroscopic, steady-state kinetic, and mechanistic characterization of the radical SAM enzyme QueE, which catalyzes a complex cyclization reaction in the biosynthesis of 7-deazapurines. *Biochemistry* **2013**, *52*, 188–98.
- (40) Zhong, A.; Lee, Y.-H.; Liu, Y.-n.; Liu, H.-w. Biosynthesis of oxetanocin-A includes a B₁₂-dependent radical SAM enzyme that can

catalyze both oxidative ring contraction and the demethylation of SAM. *Biochemistry* **2021**, *60*, 537–46.

(41) Rajakovich, L. J.; Pandelia, M. E.; Mitchell, A. J.; Chang, W. C.; Zhang, B.; Boal, A. K.; Krebs, C.; Bollinger, J. M., Jr. A new microbial pathway for organophosphonate degradation catalyzed by two previously misannotated non-heme-iron oxygenases. *Biochemistry* **2019**, *58*, 1627–47.

(42) Rajakovich, L. J.; Zhang, B.; McBride, M. J.; Boal, A. K.; Krebs, C.; Bollinger, J. M., Jr. In *Comprehensive Natural Products III: Chemistry and Biology*, Liu, H.-w.; Begley, T. P., Eds.; Elsevier: Amsterdam, 2020; pp 215–50.

(43) Xing, G.; Barr, E. W.; Diao, Y.; Hoffart, L. M.; Prabhu, K. S.; Arner, R. J.; Reddy, C. C.; Krebs, C.; Bollinger, J. M., Jr. Oxygen activation by a mixed-valent, diiron(II/III) cluster in the glycol cleavage reaction catalyzed by *myo*-inositol oxygenase. *Biochemistry* **2006**, *45*, 5402–12.

(44) van Staalduinen, L. M.; McSorley, F. R.; Schiessl, K.; Séguin, J.; Wyatt, P. B.; Hammerschmidt, F.; Zechel, D. L.; Jia, Z. Crystal structure of PhnZ in complex with substrate reveals a di-iron oxygenase mechanism for catabolism of organophosphonates. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 5171–6.

(45) Wörsdörfer, B.; Lingaraju, M.; Yennawar, N. H.; Boal, A. K.; Krebs, C.; Bollinger, J. M., Jr.; Pandelia, M. E. Organophosphonate-degrading PhnZ reveals an emerging family of HD domain mixed-valent diiron oxygenases. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 18874–9.

(46) Brown, P. M.; Caradoc-Davies, T. T.; Dickson, J. M.; Cooper, G. J.; Loomes, K. M.; Baker, E. N. Crystal structure of a substrate complex of *myo*-inositol oxygenase, a di-iron oxygenase with a key role in inositol metabolism. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15032–7.

(47) Xing, G.; Diao, Y.; Hoffart, L. M.; Barr, E. W.; Prabhu, K. S.; Arner, R. J.; Reddy, C. C.; Krebs, C.; Bollinger, J. M., Jr. Evidence for C-H cleavage by an iron-superoxide complex in the glycol cleavage reaction catalyzed by *myo*-inositol oxygenase. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 6130–5.

(48) Lundberg, M.; Siegbahn, P. E.; Morokuma, K. The mechanism for isopenicillin N synthase from density-functional modeling highlights the similarities with other enzymes in the 2-His-1-carboxylate family. *Biochemistry* **2008**, *47*, 1031–42.

(49) Burzlaff, N. I.; Rutledge, P. J.; Clifton, I. J.; Hensgens, C. M.; Pickford, M.; Adlington, R. M.; Roach, P. L.; Baldwin, J. E. The reaction cycle of isopenicillin N synthase observed by X-ray diffraction. *Nature* **1999**, *401*, 721–4.

(50) Tamanaha, E.; Zhang, B.; Guo, Y.; Chang, W. C.; Barr, E. W.; Xing, G.; St Clair, J.; Ye, S.; Neese, F.; Bollinger, J. M., Jr.; et al. Spectroscopic evidence for the two C-H-cleaving intermediates of *Aspergillus nidulans* isopenicillin N synthase. *J. Am. Chem. Soc.* **2016**, *138*, 8862–8874.

(51) Headlam, H. A.; Davies, M. J. Beta-scission of side-chain alkoxy radicals on peptides and proteins results in the loss of side-chains as aldehydes and ketones. *Free Radic. Biol. Med.* **2002**, *32*, 1171–84.

(52) Peck, S. C.; Wang, C.; Dassama, L. M.; Zhang, B.; Guo, Y.; Rajakovich, L. J.; Bollinger, J. M., Jr.; Krebs, C.; van der Donk, W. A. O-H activation by an unexpected ferryl intermediate during catalysis by 2-hydroxyethylphosphonate dioxygenase. *J. Am. Chem. Soc.* **2017**, *139*, 2045–52.

(53) Born, D. A.; Ulrich, E. C.; Ju, K. S.; Peck, S. C.; van der Donk, W. A.; Drennan, C. L. Structural basis for methylphosphonate biosynthesis. *Science* **2017**, *358*, 1336–9.

(54) Evans, R.; O'Neill, M.; Pritzel, A.; Antropova, N.; Senior, A.; Green, T.; Židek, A.; Bates, R.; Blackwell, S.; Yim, J. et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*; 2021.10.04.463034; 2021.

(55) Hallgren, J.; Tsigirgos, K. D.; Pedersen, M. D.; Almagro Armenteros, J. J.; Marcatili, P.; Nielsen, H.; Krogh, A.; Winther, O. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv*; 2022.04.08.487609; 2022.