CHAPTER TWELVE

In Vitro Analysis of Cyanobacterial Nonheme Iron-Dependent Aliphatic Halogenases WelO5 and AmbO5

Xinyu Liu[1]
University of Pittsburgh, Pittsburgh, PA, United States
[1]Corresponding author: e-mail address: xinyuliu@pitt.edu

Contents

Abstract

Aliphatic carbon–halogen (C–X) bonds are prevalent in modern pharmaceuticals and bioactive natural products. Three distinct chemical strategies are known in Nature to generate these structural motifs. The first is via the nucleophilic substitution at a pre-functionalized electrophilic carbon center with a halide anion (X^-), known for the *S*-adenosyl-L-methionine-dependent halogenases. The second is via the electrophilic activation of an alkene or its equivalent by a halenium ion (X^+) donor, known for the haloperoxidases and flavin-dependent halogenases. The third is via the direct functionalization of an unactivated aliphatic C–H bond with a halogen radical ($X^\bullet$) equivalent, known for the 2-oxo-glutarate and nonheme iron-dependent halogenases. Due to the ubiquitous nature of aliphatic C–H groups in organic molecules, transformations that permit *chemo*-, *regio*-, and *stereo*-selective modification(s) at an unactivated sp^3-carbon center have been a long sought-after goal in chemical science. Two nonheme iron-dependent halogenases, WelO5 and AmbO5 involved in the biogenesis of cyanobacterial hapalindole-type alkaloids, have been recently shown able to perform this type of challenging transformation. In this chapter, experimental details for the in vitro reconstitution of WelO5 and AmbO5 enzymatic activities are presented.

Methods in Enzymology, Volume 604
ISSN 0076-6879
https://doi.org/10.1016/bs.mie.2018.02.015

1. INTRODUCTION

Carbon–halogen (C–X) bonds are prevalent structural motifs in modern pharmaceuticals and agrochemicals (Smith, Eastman, & Njardarson, 2014). Natural products containing this type of functional group, once considered as the "chance products of Nature" (Petty, 1961), are now known ubiquitous across all domains of life with nearly 5000 members isolated and characterized to date (Gribble, 2010). Over the past two decades, significant progresses have been made in delineating Nature's logics for the enzymatic C–X bond formations. These advances, along with their historical perspective, have been recently systematically reviewed (Agarwal et al., 2017; Vaillancourt, Yeh, Vosburg, Garneau-Tsodikova, & Walsh, 2006) and will not be discussed in detail in this chapter. Aliphatic C–X bonds, distinct from those with halogens attached to an sp^2- or sp-hybridized carbon center, are widely present in naturally occurring halometabolites (Fig. 1; Gribble, 2010). The attachment of a halogen atom to an sp^3-hybridized carbon center profoundly alters its reactivity, notably its electrophilicity. Thus, this type of modification often plays an important role in the bioactivity of halogenated natural products (Williams et al., 2005). The logics underlying the formation of an aliphatic C–X bond in natural product biosynthesis encompass all the key chemical principles established for the enzymatic C–X bond formation (Agarwal et al., 2017; Vaillancourt et al., 2006). These principles are briefly outlined here.

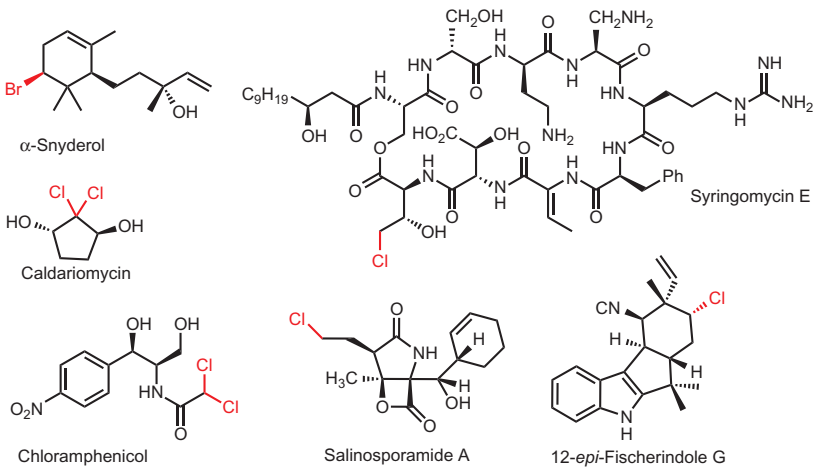

Fig. 1 Selected halogenated natural products with aliphatic carbon–halogen bonds.

Halogenated terpenoids of marine and terrestrial origins represent the largest family of naturally occurring halometabolites with aliphatic C–Cl and C–Br bonds (Gribble, 2010). This type of aliphatic C–X bonds originates primarily from an sp^2-hybridized carbon, amenable for activation by a halenium ion (X^+) donor that is generated oxidatively from a halide anion (X^-) by a haloperoxidase or a flavin-dependent halogenase. A notable example is the aliphatic C–Br bond formation in the biogenesis of α-snyderol from (*E*)-(+)-nerolidol by a vanadium-dependent bromoperoxidase (Fig. 2A; Carter-Franklin & Butler, 2004). Related mechanisms that involve the activation of an electron-rich sp^2-hybridized carbon with an enzymatically generated X^+ donor have been cited for the aliphatic C–Cl bond formations in the biosynthesis of caldariomycin and chloramphenicol, where the former requires a heme-dependent chloroperoxidase (Sundaramoorthy, Terner, & Poulos, 1995) and the latter calls for a flavin-dependent halogenase (Podzelinska et al., 2010).

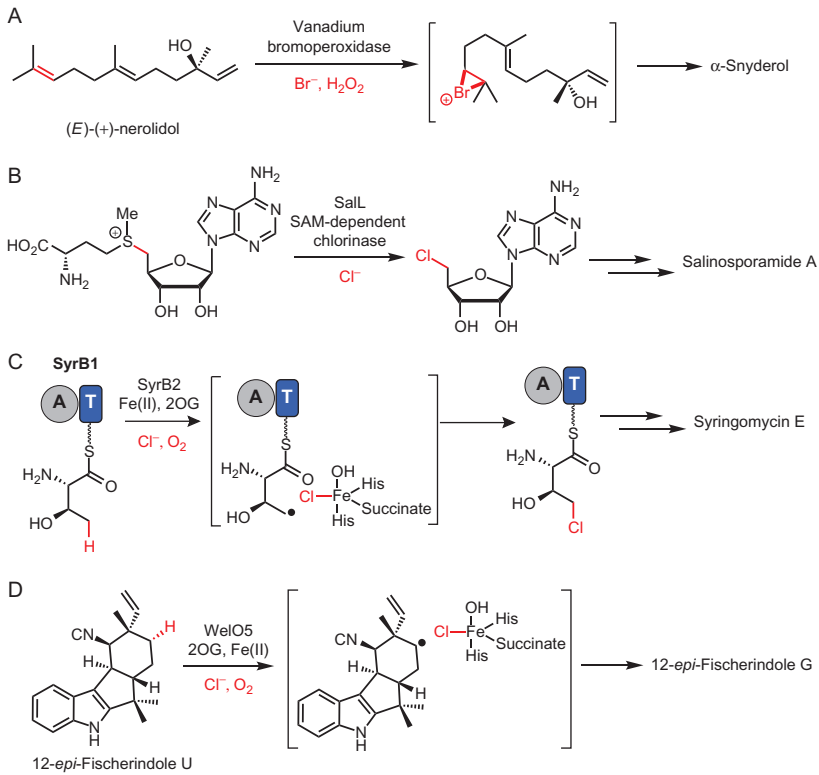

Fig. 2 Selected examples highlighting the key chemical principles underlying the formation of aliphatic carbon–halogen bonds in natural product biosynthesis.

Nucleophilic substitution by a halide anion (X^-) at an activated aliphatic carbon center of *S*-adenosyl-L-methionine (SAM), manifested by the SAM-dependent halogenases, constitutes the second major enzymatic strategy for introducing aliphatic C–X bonds to the halogenated natural products. When this substitution occurs on the activated methyl group of SAM, a halomethane, the simplest naturally occurring halometabolite, is generated and the corresponding enzyme is also known as the SAM-dependent halide methyltransferase (Wuosmaa & Hager, 1990). Alternatively, when the substitution occurs on the 5′ carbon of adenosine in SAM, a 5′-halo-5′-deoxyadenosine is generated as an intermediate for the downstream biosynthetic maturations. SalL, one of the two well-characterized SAM-dependent halogenases, catalyzes the formation of 5′-chloro-5′-deoxyadenosine en route to the biosynthesis of salinosporamide A (Fig. 2B; Eustaquio, Pojer, Noel, & Moore, 2008).

The third major enzymatic strategy for introducing aliphatic C–X bonds into halogenated natural product is via the direct modification of an aliphatic carbon–hydrogen (C–H) bond with a halogen radical ($X^\bullet$) equivalent, manifested by the nonheme iron-dependent halogenases. These halogenases constitute an important subset of a widespread family of oxygenases that combine reaction of O_2 at an H–X–D/E…H-coordinated mononuclear Fe(II) center with oxidative decarboxylation of a coordinated 2-oxo-glutarate (2OG) cofactor (Hausinger & Schofield, 2015). These enzymes generate an iron(IV)-oxo (Fe^{IV}═O) (ferryl) intermediate, a common mechanistic theme in all Fe/2OG enzymes, to initiate reaction with substrate by abstraction of a hydrogen atom ($H^\bullet$) (Krebs, Galonic Fujimori, Walsh, & Bollinger, 2007).

SyrB2, from the syringomycin biosynthetic pathway in *Pseudomonas syringae*, is the first biochemically characterized Fe/2OG halogenase (Vaillancourt, Yin, & Walsh, 2005). It monochlorinates the C4 methyl group in L-threonine covalently tethered as a thioester to the phosphopantetheine arm of its carrier protein, SyrB1 (Fig. 2C; Vaillancourt, Yin, et al., 2005). The X-ray crystal structure of Fe(II)·2OG SyrB2 with Cl^- (but in the absence of substrate, the threonine-appended carrier protein SyrB1) gave the first mechanistic insight into Fe/2OG-mediated halogenation (Blasiak, Vaillancourt, Walsh, & Drennan, 2006). Omission of the carboxylate ligand in the H-X-D/E…H motif via mutation to an Ala residue opens an Fe(II) coordination position that can be occupied by the halogen. Thus, the SyrB2 reaction proceeds by the same general pathway as Fe/2OG hydroxylases, involving a *cis*-halo-ferryl intermediate to abstract an $H^\bullet$ from C4 of the tethered threonine, then coupling the substrate carbon radical with the chlorine ligand instead of the hydroxyl ligand of the resultant Cl–Fe(III)–OH complex (Matthews, Krest, et al., 2009).

Due to the ubiquitous nature of aliphatic C–H groups in organic molecules, transformations that permit *chemo*-, *regio*-, and *stereo*-selective modification(s) at an unactivated sp^3-carbon center have been a long sought-after goal in chemical science (White, 2012). The discovery of SyrB2-type nonheme iron halogenases, widely present in the biosynthetic pathways for polyketide and nonribosomal peptide natural products (Galonic, Vaillancourt, & Walsh, 2006; Gu et al., 2009; Jiang et al., 2011; Neumann & Walsh, 2008; Ueki et al., 2006; Vaillancourt, Yeh, Vosburg, O'Connor, & Walsh, 2005), immediately raised the question whether they can be applied to biocatalytic processes. Moreover, the SyrB2 crystal structure seemingly suggests a minimal modification in the H-X-D/E…H Fe(II)-binding motif constitutes a key for the ferryl complex to promote chlorine transfer. This raised another intriguing question on whether this finding can be universally translated to the superfamily of Fe^{II}/2OG enzymes simply by deletion of the carboxylate ligand to accommodate a more diverse set of substrates. Unfortunately, neither was the case. SyrB2-type halogenases strictly require the small-molecular substrate tethered to a carried protein (Galonic et al., 2006; Gu et al., 2009; Jiang et al., 2011; Neumann & Walsh, 2008; Ueki et al., 2006; Vaillancourt, Yeh, et al., 2005; Vaillancourt, Yin, et al., 2005). Efforts to reprogram Fe/2OG hydroxylases by single-site substitution of the D/E in the H-X-D/E…H motif have not met with success (Gorres, Pua, & Raines, 2009; Grzyska, Muller, Campbell, & Hausinger, 2007). Even the restoration of the H-X-D/E…H motif in SyrB2 did not yield a competent enzyme for hydroxylation (Matthews, Neumann, et al., 2009). These data collectively suggest the challenging nature to utilize SyrB2-type halogenases for biocatalytic applications.

WelO5, from the welwitindolinone biosynthetic pathway in *Hapalosiphon welwitschii* (Hillwig, Fuhrman, et al., 2014; Hillwig, Zhu, & Liu, 2014), is the first biochemically characterized Fe/2OG halogenase that can act on free-standing small molecules independent of carrier proteins (Hillwig & Liu, 2014). WelO5 preferentially monochlorinates C13 of 12-*epi*-fischerindole U to yield 12-*epi*-fischerindole G (Fig. 2D) and is also able to convert 12-*epi*-hapalindole C to 12-*epi*-hapalindole E at a reduced efficiency (Fig. 3). A related halogenase AmbO5 (79% sequence identical to WelO5), from the ambiguine biosynthetic pathway in *Fischerella ambigua* (Hillwig, Zhu, et al., 2014), has been shown able to accommodate seven structurally distinct hapalindole-type alkaloids (Fig. 3; Hillwig, Zhu, Ittiamornkul, & Liu, 2016). In addition, the restoration of the H-X-D…H motif in WelO5 generates a catalytically competent hydroxylase WelO5(G166D) (Fig. 4A), suggesting a unique active site dynamics may be operant that is distinct from SyrB2. The cocrystal structure of Fe(II)·2OG WelO5 with Cl^- and 12-*epi*-fischerindole U (Fig. 4B;

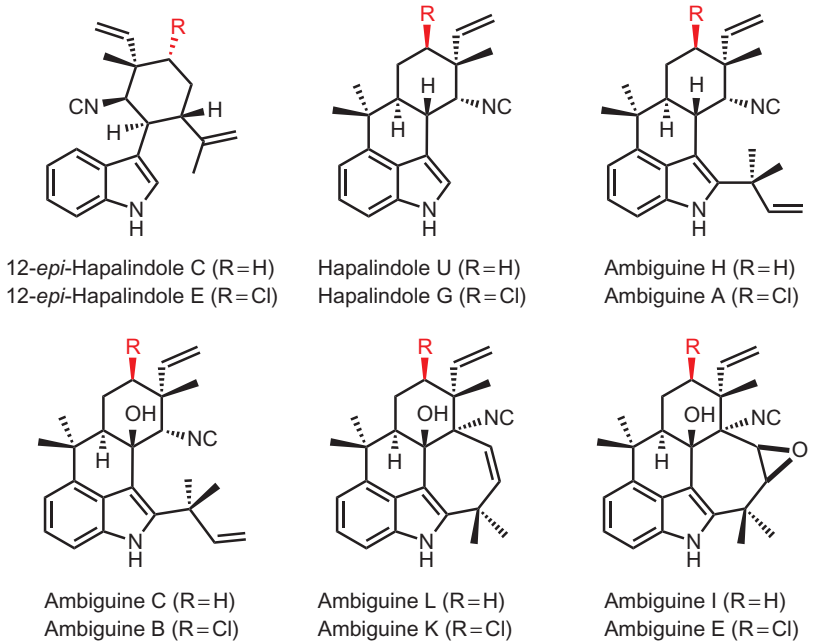

Fig. 3 Additional hapalindole-type alkaloids (R=H) that are validated substrates for WelO5 and/or AmbO5 aliphatic halogenases.

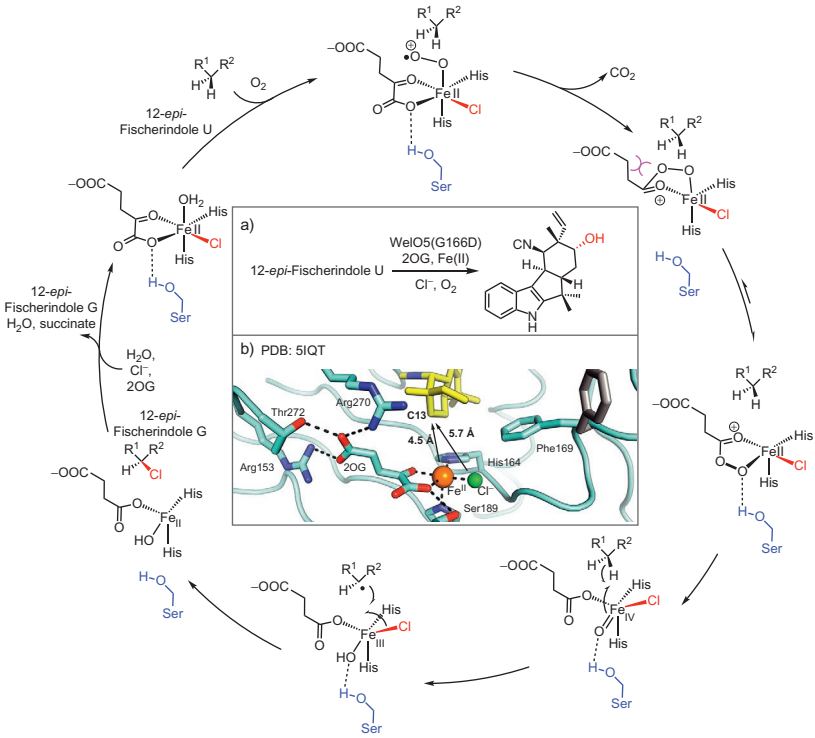

Fig. 4 Proposed reaction mechanism of WelO5 halogenase derived from its X-ray crystal structure in complex with 12-*epi*-fischerindole U, 2OG, Cl^-, and Fe(II) (inlet B) that accounts for the observed selective chlorination activity by the wild-type enzyme and the distinct hydroxylation activity by the WelO5(G166D) mutant (inlet A).

Mitchell et al., 2016) provided structural corroborations with the in vitro biochemical characteristics of WelO5 and AmbO5 (Hillwig & Liu, 2014; Hillwig et al., 2016; Zhu & Liu, 2017). This led to the identification of an external α-helical motif that defines the small-molecular substrate-binding pocket. Moreover, the WelO5 structure revealed that the target C13 of 12-*epi*-fischerindole U is naturally poised for hydroxylation in the active site (Fig. 4B), and the selective chlorine transfer is likely dictated by a previously unknown second sphere serine residue that helps control the geometry of the reactive Cl–Fe(IV)═O and Cl–Fe(III)–OH intermediates in the WelO5 catalytic cycle (Fig. 4). The structural characterization of WelO5 further led to the identification of an Fe/2OG hydroxylase (SadA) that can be readily engineered to gain halogenase activity (Mitchell et al., 2017). Together with the identification of additional WelO5-type halogenase that may mediate late-stage chlorinations beyond the hapalindole-type scaffolds (Moosmann et al., 2017), a promising future can be forecasted for engineering this family of enzymes for selective functionalization of aliphatic C–H groups with halogens in a diverse set of molecular scaffolds.

2. IN VITRO ANALYSIS OF WelO5 AND AmbO5 HALOGENASES

As for any given in vitro enzyme reconstitution, its success requires three key procedural components that are indispensable from each other: (1) the ability to procure sufficient quantity of the target enzyme, (2) the ability to procure sufficient quantity of the presumptive substrate(s) for the target enzyme, and (3) analytical procedures to assess the turnover of the substrate(s) by the target enzyme and to define the molecular nature of the enzymatic product. The following subsections detail my laboratory's experience on these three parts for WelO5 and AmbO5 halogenases.

2.1 Purification of WelO5 and AmbO5 Proteins

Genes (870 bp) encoding WelO5 and AmbO5 proteins can be readily amplified by PCR from the genomic DNA of *H. welwitschii* UTEX B 1830 and *F. ambigua* UTEX 1903, respectively. Both strains are available at the Culture Collection of Algae at the University of Texas at Austin (https://utex.org). Alternatively, these sequences (GenBank ID: AHI58816 for WelO5 and AKP23998 for AmbO5) can be synthesized by any reputable company offering a synthetic gene service. Codon optimization is not required, as the native sequences lead to an appreciable amount of soluble proteins in *Escherichia coli*

(vide infra), sufficient for in vitro biochemical and structural studies. The pQTEV plasmid, available at Addgene (https://www.addgene.org/31291/), has been the vector of choice by my laboratory to clone the *wel*O5 and *amb*O5 genes. These constructs gave the N-His_7-tagged WelO5 and AmbO5 proteins upon overexpression in *E. coli*. Any other conventional cloning vectors that carry N-terminal His-Tag sequence, such as pET22 or pET28 plasmids, should perform equally well as the pQTEV. Site-directed mutagenesis can be routinely performed on the WelO5 and AmbO5 proteins readily once their expression constructs are generated using a PCR-based approach, of which details will not be included here. The following protocol describes a purification procedure for the WelO5 protein, which is equally applicable to the AmbO5 protein as well as their mutant proteins.

Day 1: Transform a pQTEV plasmid with the *wel*O5 gene cloned between the *Bam*HI and *Not*I sites to electro- or chemically competent C43 (DE3) or BL21(DE3) *E. coli* cells.

Day 2 evening: Pick a single colony of C43 (DE3) or BL21(DE3) *E. coli* cells with the pQTEV::*wel*O5 plasmid and prepare a starter culture by inoculating the picked colony to a 5-mL LB culture overnight at 30°C.

Day 3 morning: Inoculate 1 mL of the starter culture to a 2.8-L Fernbach flask containing 1 L of LB medium with 100 μg/mL ampicillin. Grow the cultures at 37°C to an optical density of approximately 0.6, then drop the temperature to 16°C. Add isopropyl β-D-thiogalactopyranoside (IPTG) (1 m*M*, final concentration) to induce protein expression.

Day 4 morning: Pellet the cells from the 1-L overnight culture by centrifugation at 8800 × *g* for 20 min. Resuspend the cell pellets in a 10-mL protein lysis buffer (50 m*M* Tris/pH 7.4, 500 m*M* NaCl, 20 m*M* imidazole, 10 m*M* β-mercaptoethanol, and 0.1% Tween-20) and perform cell lysis by sonicating on ice in a 10-s interval for 3 min and then pellet the cell debris by centrifugation at 20,000 × *g* and 4°C for 30 min. Transfer the cleared lysate to a 2-mL prewashed Ni-NTA bead (Qiagen) slurry and incubate for 1 h at 4°C. Thereafter, load the Ni-NTA beads on to a column and wash with 60 mL of binding buffer (50 m*M* Tris/pH 7.4, 500 m*M* NaCl, 0.1% Tween-20, 20 m*M* imidazole, 10 m*M* β-mercaptoethanol). The His-tagged WelO5 protein will be eluted by adding 5 mL of elution buffer (50 m*M* Tris/pH 7.4, 500 m*M* NaCl, 100 then 250 m*M* imidazole, β-mercaptoethanol) to the bead bed. The eluted protein will be subjected to a dialysis at 4°C using a 30-kDa molecular weight cutoff membrane (Spectrum Laboratory Products) with a buffer containing 50 m*M* HEPES (pH 7.4), 10 m*M* EDTA, 10 m*M* NaCl, 10% glycerol, and 0.5 m*M* DTT. This initial dialysis can be

carried overnight and ensures the complete removal of metal ions in the eluted protein sample post-Ni-NTA column.

Day 5 morning: The WelO5 protein sample postdialysis against EDTA will be subjected to a second round of dialysis with a buffer containing 50 m*M* HEPES (pH 7.4), 10 m*M* NaCl, 10% glycerol, and 0.5 m*M* DTT to remove EDTA. This dialysis step to remove EDTA is recommend to be repeated one more time before the sample is concentrated with an Amicon MWCO 10 centrifugal filter. The WelO5 protein purified by the procedure described earlier is sufficiently pure for in vitro reconstitution experiment. Its purity can be judged by routine SDS-PAGE analysis, and its concentration can be estimated either using a UV–vis spectrometer or the colormetric Bradford assay.

Note: Literature describing the in vitro reconstitution of several carrier protein-dependent nonheme iron halogenases (Galonic et al., 2006; Vaillancourt, Yeh, et al., 2005; Vaillancourt, Yin, et al., 2005), including a protocol chapter in this book series (van Pee, 2012), has indicated that these proteins need to be purified in the *absence* of air. These observations are counterintuitive to the facts that these proteins are oxygenases and require molecular oxygen for their activities. My laboratory and others have independently validated (Hillwig & Liu, 2014; Matthews, Krest, et al., 2009), as the protocol described earlier, that these halogenases can be routinely purified in an atmospheric environment. The key to obtain consistently active proteins is to ensure the metal ions are completely stripped off from the protein initially by a dialysis against EDTA.

2.2 Procurement of Substrates for WelO5 and AmbO5 Proteins

Seven hapalindole-type alkaloids (Figs. 2D and 3) have been validated as the native substrates for WelO5 and/or AmbO5 halogenases (Hillwig & Liu, 2014; Hillwig et al., 2016). These alkaloids can be procured via three different routes. The first is via isolation from their native cyanobacterial producers (Smitka et al., 1992; Stratmann et al., 1994), which have been practiced by my laboratory for several years and a sample protocol will be provided here. The second is via de novo organic synthesis. Synthetic routes that consist of less than 10 steps from commercially available building blocks have been established for 12-*epi*-fischerindole U, 12-*epi*-hapalindole C, and hapalindole U (Baran, Maimone, & Richter, 2007; Baran & Richter, 2004). These synthetic advancements rendered the custom synthesis as a viable third option to procure these molecules. For example, 12-*epi*-hapalindole

C was known accessible in an estimated cost of 5000 USD/g via custom synthesis (2015 data, personal communication with the Biocatalysis group of Novartis AG, Basel, CH).

Described below is an exemplary protocol for the isolation of 12-*epi*-ficherindole U from *H. welwitschii* UTEX B 1830.

Step 1: Culture *H. welwitschii* UTEX B 1830 at a 20-L scale at room temperature in BG-11 medium (for recipe, see https://utex.org/products/bg-11-medium) with a 12 h/12 h light/dark cycle for approximately 5 weeks.

Step 2: Harvest the algal cells by filtration and then lyophilize. This typically yields 5–6 g of freeze-dried material.

Step 3: Extract the lyophilized tissue with CH_2Cl_2/MeOH (v/v = 1:1, 50 mL) overnight with stirring at room temperature. The crude organic extract postfiltration will be dried under reduced pressure and redissolved in methanol (15 mL). This methanol solution will be filtered through an organic solvent compatible syringe filter and injected on to a Phenomenex Luna C18 21.5 × 20 mm reversed phase preparative HPLC column for initial crude separation of hapalindole-type compounds using a gradient from 50% to 100% methanol in water over 100 min.

Step 4: The fractions that contain 12-*epi*-ficherindole U from the preparative HPLC run will need to be pooled, dried, and redissolved in methanol and subjected to another round of fine purification by HPLC. This second round HPLC is typically done via a C18 Luna Phenomenex 250 × 4.6 mm column with a gradient from 50% to 80% acetonitrile in water over 35 min and a UV detector wavelength set at 280 nm. This procedure typically results in 0.2–1 mg of 12-*epi*-ficherindole U from a 20-L cyanobacterial culture. The authenticity of 12-*epi*-ficherindole U will need to be verified by a ^{1}H NMR analysis in comparison with literature data (Stratmann et al., 1994). This NMR analysis step is critical, in particular for the first-time practitioner of this protocol, as *H. welwitschii* UTEX B 1830 produces several isomers of 12-*epi*-ficherindole U that cannot be readily distinguished by UV and mass spec analyses alone.

2.3 Activity Assay for WelO5 and AmbO5 Proteins

The protocol described here is a standard procedure for the WelO5 wild-type protein using 12-*epi*-ficherindole U as a substrate. This method can be readily applied to the AmbO5 protein as well as in combination with other substrates.

The analytical scale assay (100 μL) is carried out in a 1.8-mL borosilicate glass vial, rather than an Eppendorf tube made of plastic. The reason for this

is that my laboratory has found hapalindole-type compounds are inherently hydrophobic and tend to "stick" to the plastic surface of some Eppendorf tubes. This makes the subsequent enzymatic product recovery challenging. The metal cofactor Fe(II) is provided in the form of $(NH_4)_2Fe(SO_4)_2$. It is recommended the stock solution of $(NH_4)_2Fe(SO_4)_2$ to be made freshly to minimize its deterioration in the air, unless the practitioner's lab has a glove box to keep this solution.

Step 1: All assay components, except the WelO5 protein and Fe(II) salt, are premixed with water to a final volume of 90 μL with their final concentrations at 100 μL adjusted to the following values: 12-*epi*-ficherindole U (1 m*M*), 2OG (2 m*M*), NaCl (10 m*M*), and HEPES (pH 7.4, 20 m*M*). This mixture should be briefly purged with nitrogen gas.

Step 2: To this initially prepared solution, the WelO5 protein (20 μ*M*, final concentration) and $(NH_4)_2Fe(SO_4)_2$ (500 μ*M*, final concentration) are added with a small amount of water to adjust the final volume to 100 μL. The mixture will be briefly incubated under a nitrogen atmosphere before being purged with air and incubated aerobically at 30°C for 1 h.

Step 3: The enzymatic assay is stopped by extracting with ethyl acetate (0.5 mL × 2) and centrifuged. Ethyl acetate extracts are combined, dried under a stream of nitrogen gas, and redissolved in methanol (50 μL). A portion of the methanol solution is injected into a C18 Luna Phenomenex 250 × 4.6 mm HPLC column using a mobile phase gradient 50%–80% acetonitrile in water over 35 min with UV absorption monitoring set at 280 nm.

Step 4: 12-*epi*-ficherindole G is less polar than 12-*epi*-ficherindole U and can be readily identified by examining the HPLC chromatograph. For a first-time practitioner, the standard 12-*epi*-ficherindole G can be isolated from *H. welwitschii* UTEX B 1830 using the protocol described in Section 2.2, which can be used as a coelution standard in this assay. The HPLC analysis described in step 3 can be coupled to a MS detector to ascertain the chlorine transfer. To be absolutely certain on the structure of the chlorinated adduct, the enzymatic reaction has to be scaled up proportionally to at least 3 mL to allow for the enzymatic product to be isolated by semipreparative HPLC for NMR analysis.

3. OUTLOOK

Methods that enable selective intermolecular aliphatic C–H bond functionalization reactions have been long considered a Holy Grail of modern chemical science (Arndtsen, Bergman, Mobley, & Peterson, 1995). Despite the significant progress in this scientific frontier over the past two

decades (Crabtree & Lei, 2017; Davies & Morton, 2017; Goldberg & Goldman, 2017; Hartwig, 2017; He, Wasa, Chan, Shao, & Yu, 2017; McIntosh, Farwell, & Arnold, 2014; Shang, Ilies, & Nakamura, 2017), broadly applicable chemical or enzymatic transformations that allow for the selective modification of an aliphatic C–H bond to a C–X bond in a diverse set of molecules have not yet emerged. The discovery of WelO5-type aliphatic halogenases provides a new platform and opportunity to explore this unmet need. While this statement is likely subject to rebuttals due to the limited currently known substrates of WelO5 and AmbO5 (Hillwig & Liu, 2014; Hillwig et al., 2016), several layers of evidence have accumulated over the past 3 years that augments this family of halogenases as a privileged system for biocatalysis development. These evidences address two important issues that often concern practitioners in this field: one is whether this family of halogenase is evolvable and can accept small-molecular substrates beyond hapalindole-type alkaloids; the other is if they are evolvable, then what are the rules? First of all, there is strong evidence that WelO5-type halogenases are naturally evolvable for alternating substrate preferences within the related structural class as well as for completely different structural scaffolds. By characterizing three WelO5-type halogenases related to the biosynthesis of hapalindole alkaloids (Hillwig & Liu, 2014; Hillwig et al., 2016; Zhu & Liu, 2017), my laboratory has identified a conserved C-terminal amino acid sequence that typically resides at 210–240 residues and is critically important for their substrate scope and tolerance of the identified hapalindole-type substrates. This observation corroborates with the substrate-bound X-ray crystal structure of WelO5 (Mitchell et al., 2016). More recently, a WelO5-type halogenase (MstM) was identified in the biosynthetic pathway of an *ent*-sterol-containing meroterpenoid in the cyanobacterium *Scytonema* sp. PCC10023 (Moosmann et al., 2017). Although the halogenation timing for MstM remains to be elucidated, the substrate for MstM is likely a steroid-like molecule, distinct from the hapalindole's indole monoterpene scaffold. Notably, the sequence of MstM diverges most notably from WelO5 and AmbO5 in the previously identified C-terminal motif that defines the substrate-binding pocket. Second, the X-ray crystal structure characterization of WelO5 discovered a novel structural element for controlling the selective halogen transfer by this family of 2OG/Fe(II) enzymes (Fig. 4) and laid the conceptual ground on how to engineer the broadly available 2OG/Fe(II) hydroxylases to halogenases. As a proof of principle, my laboratory has shown the *N*-suc-L-leucine hydroxylase SadA (Hibi et al., 2012) can readily gain chlorination activity by a rational engineering approach (Mitchell et al., 2017). This enzyme is now ready

for lab-based directed evolution to further enhance its halogenation activity. It also offers a quick entry point for many other groups to pursue this goal simultaneously.

Overall, with growing microbial genomic data, in particular those related to cyanobacterial organisms (Shih et al., 2013), it is safe to envision more WelO5-type small-molecular aliphatic halogenases will be discovered. The comparative biochemical and structural characterizations of these new enzymes will provide us a better and more concrete picture on Nature's principle to evolve this family of enzymes for alternative structural scaffolds. This knowledge, in combination with those that will soon be learned from the directed evolution of SadA mutant protein to gain halogenation activity, will undoubtedly guide us synergistically toward a broadly applicable enzymatic system that enables late-stage aliphatic halogenations in small-molecular pharmaceuticals, agrichemicals, and bioactive natural products.

ACKNOWLEDGMENTS

The author is indebted to Dr. Matthew L. Hillwig, who established the protocols in the author's lab for assessing the enzymatic activity of WelO5 halogenase as described in this chapter. The initial discovery work associated with the WelO5 and AmbO5 halogenases that constitutes the basis for this chapter was supported by the unrestricted fund from Department of Chemistry, University of Pittsburgh.

REFERENCES

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.

Arndtsen, B. A., Bergman, R. G., Mobley, T. A., & Peterson, T. H. (1995). Selective intermolecular carbon-hydrogen bond activation by synthetic metal complexes in homogeneous solution. *Accounts of Chemical Research*, *28*(3), 154–162. https://doi.org/10.1021/ar00051a009.

Baran, P. S., Maimone, T. J., & Richter, J. M. (2007). Total synthesis of marine natural products without using protecting groups. *Nature*, *446*(7134), 404–408. https://doi.org/10.1038/nature05569.

Baran, P. S., & Richter, J. M. (2004). Direct coupling of indoles with carbonyl compounds: Short, enantioselective, gram-scale synthetic entry into the hapalindole and fischerindole alkaloid families. *Journal of the American Chemical Society*, *126*(24), 7450–7451. https://doi.org/10.1021/ja047874w.

Blasiak, L. C., Vaillancourt, F. H., Walsh, C. T., & Drennan, C. L. (2006). Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature*, *440*(7082), 368–371. https://doi.org/10.1038/nature04544.

Carter-Franklin, J. N., & Butler, A. (2004). Vanadium bromoperoxidase-catalyzed biosynthesis of halogenated marine natural products. *Journal of the American Chemical Society*, *126*(46), 15060–15066. https://doi.org/10.1021/ja047925p.

Crabtree, R. H., & Lei, A. (2017). Introduction: CH activation. *Chemical Reviews*, *117*(13), 8481–8482. https://doi.org/10.1021/acs.chemrev.7b00307.

Davies, H. M. L., & Morton, D. (2017). Collective approach to advancing C–H functionalization. *ACS Central Science*, *3*(9), 936–943. https://doi.org/10.1021/acscentsci.7b00329.

Eustaquio, A. S., Pojer, F., Noel, J. P., & Moore, B. S. (2008). Discovery and characterization of a marine bacterial SAM-dependent chlorinase. *Nature Chemical Biology*, *4*(1), 69–74. https://doi.org/10.1038/nchembio.2007.56.

Galonic, D. P., Vaillancourt, F. H., & Walsh, C. T. (2006). Halogenation of unactivated carbon centers in natural product biosynthesis: Trichlorination of leucine during barbamide biosynthesis. *Journal of the American Chemical Society*, *128*(12), 3900–3901. https://doi.org/10.1021/ja060151n.

Goldberg, K. I., & Goldman, A. S. (2017). Large-scale selective functionalization of alkanes. *Accounts of Chemical Research*, *50*(3), 620–626. https://doi.org/10.1021/acs.accounts.6b00621.

Gorres, K. L., Pua, K. H., & Raines, R. T. (2009). Stringency of the 2-His-1-Asp active-site motif in prolyl 4-hydroxylase. *PLoS One*, *4*(11), e7635. https://doi.org/10.1371/journal.pone.0007635.

Gribble, G. W. (2010). Naturally occurring organohalogen compounds—A comprehensive update. *Progress in the Chemistry of Organic Natural Products*, *91*, 1–613.

Grzyska, P. K., Muller, T. A., Campbell, M. G., & Hausinger, R. P. (2007). Metal ligand substitution and evidence for quinone formation in taurine/alpha-ketoglutarate dioxygenase. *Journal of Inorganic Biochemistry*, *101*(5), 797–808. https://doi.org/10.1016/j.jinorgbio.2007.01.011.

Gu, L., Wang, B., Kulkarni, A., Geders, T. W., Grindberg, R. V., Gerwick, L., et al. (2009). Metamorphic enzyme assembly in polyketide diversification. *Nature*, *459*(7247), 731–735. https://doi.org/10.1038/nature07870.

Hartwig, J. F. (2017). Catalyst-controlled site-selective bond activation. *Accounts of Chemical Research*, *50*(3), 549–555. https://doi.org/10.1021/acs.accounts.6b00546.

Hausinger, R., & Schofield, C. (2015). In R. Hausinger & C. Schofield (Eds.), *2-Oxoglutarate-dependent oxygenases*: The Royal Society of Chemistry.

He, J., Wasa, M., Chan, K. S. L., Shao, Q., & Yu, J.-Q. (2017). Palladium-catalyzed transformations of alkyl C–H bonds. *Chemical Reviews*, *117*(13), 8754–8786. https://doi.org/10.1021/acs.chemrev.6b00622.

Hibi, M., Kawashima, T., Kasahara, T., Sokolov, P. M., Smirnov, S. V., Kodera, T., et al. (2012). A novel Fe(II)/alpha-ketoglutarate-dependent dioxygenase from Burkholderia ambifaria has beta-hydroxylating activity of N-succinyl l-leucine. *Letters in Applied Microbiology*, *55*(6), 414–419. https://doi.org/10.1111/j.1472-765X.2012.03308.x.

Hillwig, M. L., Fuhrman, H. A., Ittiamornkul, K., Sevco, T. J., Kwak, D. H., & Liu, X. (2014). Identification and characterization of a welwitindolinone alkaloid biosynthetic gene cluster in the stigonematalean Cyanobacterium Hapalosiphon welwitschii. *Chembiochem*, *15*(5), 665–669. https://doi.org/10.1002/cbic.201300794.

Hillwig, M. L., & Liu, X. (2014). A new family of iron-dependent halogenases acts on free-standing substrates. *Nature Chemical Biology*, *10*(11), 921–923. https://doi.org/10.1038/nchembio.1625.

Hillwig, M. L., Zhu, Q., Ittiamornkul, K., & Liu, X. (2016). Discovery of a promiscuous non-heme iron halogenase in ambiguine alkaloid biogenesis: Implication for an evolvable enzyme family for late-stage halogenation of aliphatic carbons in small molecules. *Angewandte Chemie (International Ed. in English)*, *55*(19), 5780–5784. https://doi.org/10.1002/anie.201601447.

Hillwig, M. L., Zhu, Q., & Liu, X. (2014). Biosynthesis of ambiguine indole alkaloids in cyanobacterium Fischerella ambigua. *ACS Chemical Biology*, *9*(2), 372–377. https://doi.org/10.1021/cb400681n.

Jiang, W., Heemstra, J. R., Jr., Forseth, R. R., Neumann, C. S., Manaviazar, S., Schroeder, F. C., et al. (2011). Biosynthetic chlorination of the piperazate residue in kutzneride biosynthesis by KthP. *Biochemistry*, *50*(27), 6063–6072. https://doi.org/10.1021/bi200656k.

Krebs, C., Galonic Fujimori, D., Walsh, C. T., & Bollinger, J. M., Jr. (2007). Non-heme Fe (IV)-oxo intermediates. *Accounts of Chemical Research*, *40*(7), 484–492. https://doi.org/10.1021/ar700066p.

Matthews, M. L., Krest, C. M., Barr, E. W., Vaillancourt, F. H., Walsh, C. T., Green, M. T., et al. (2009). Substrate-triggered formation and remarkable stability of the C-H bond-cleaving chloroferryl intermediate in the aliphatic halogenase, SyrB2. *Biochemistry*, *48*(20), 4331–4343. https://doi.org/10.1021/bi900109z.

Matthews, M. L., Neumann, C. S., Miles, L. A., Grove, T. L., Booker, S. J., Krebs, C., et al. (2009). Substrate positioning controls the partition between halogenation and hydroxylation in the aliphatic halogenase, SyrB2. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(42), 17723–17728. https://doi.org/10.1073/pnas.0909649106.

McIntosh, J. A., Farwell, C. C., & Arnold, F. H. (2014). Expanding P450 catalytic reaction space through evolution and engineering. *Current Opinion in Chemical Biology*, *19*, 126–134. https://doi.org/10.1016/j.cbpa.2014.02.001.

Mitchell, A. J., Dunham, N. P., Bergman, J. A., Wang, B., Zhu, Q., Chang, W. C., et al. (2017). Structure-guided reprogramming of a hydroxylase to halogenate its small molecule substrate. *Biochemistry*, *56*(3), 441–444. https://doi.org/10.1021/acs.biochem.6b01173.

Mitchell, A. J., Zhu, Q., Maggiolo, A. O., Ananth, N. R., Hillwig, M. L., Liu, X., et al. (2016). Structural basis for halogenation by iron- and 2-oxo-glutarate-dependent enzyme WelO5. *Nature Chemical Biology*, *12*(8), 636–640. https://doi.org/10.1038/nchembio.2112.

Moosmann, P., Ueoka, R., Grauso, L., Mangoni, A., Morinaka, B. I., Gugger, M., et al. (2017). Cyanobacterial ent-sterol-like natural products from a deviated ubiquinone pathway. *Angewandte Chemie (International Ed. in English)*, *56*(18), 4987–4990. https://doi.org/10.1002/anie.201611617.

Neumann, C. S., & Walsh, C. T. (2008). Biosynthesis of (-)-(1S,2R)-allocoronamic acyl thioester by an Fe(II)-dependent halogenase and a cyclopropane-forming flavoprotein. *Journal of the American Chemical Society*, *130*(43), 14022–14023. https://doi.org/10.1021/ja8064667.

Petty, M. A. (1961). An introduction to the origin and biochemistry of microbial halometabolites. *Bacteriological Reviews*, *25*, 111–130.

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.

Shang, R., Ilies, L., & Nakamura, E. (2017). Iron-catalyzed C–H bond activation. *Chemical Reviews*, *117*(13), 9086–9139. https://doi.org/10.1021/acs.chemrev.6b00772.

Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., et al. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(3), 1053–1058. https://doi.org/10.1073/pnas.1217107110.

Smith, B. R., Eastman, C. M., & Njardarson, J. T. (2014). Beyond C, H, O, and N: Analysis of the elemental composition of U.S. FDA approved drug architectures. *Journal of Medicinal Chemistry*, *57*(23), 9764–9773. https://doi.org/10.1021/jm501105n.

Smitka, T. A., Bonjouklian, R., Doolin, L., Jones, N. D., Deeter, J. B., Yoshida, W. Y., et al. (1992). Ambiguine isonitriles, fungicidal hapalindole-type alkaloids from three genera of blue-green algae belonging to the Stigonemataceae. *Journal of Organic Chemistry*, *57*(3), 857–861. https://doi.org/10.1021/jo00029a014.

Stratmann, K., Moore, R. E., Bonjouklian, R., Deeter, J. B., Patterson, G. M. L., Shaffer, S., et al. (1994). Welwitindolinones, unusual alkaloids from the blue-green algae *Hapalosiphon welwitschii* and *Westiella intricata*. Relationship to fischerindoles and hapalinodoles. *Journal of the American Chemical Society*, *116*(22), 9935–9942. https://doi.org/10.1021/ja00101a015.

Sundaramoorthy, M., Terner, J., & Poulos, T. L. (1995). The crystal structure of chloroperoxidase: A heme peroxidase—Cytochrome P450 functional hybrid. *Structure*, *3*(12), 1367–1377.

Ueki, M., Galonic, D. P., Vaillancourt, F. H., Garneau-Tsodikova, S., Yeh, E., Vosburg, D. A., et al. (2006). Enzymatic generation of the antimetabolite gamma, gamma-dichloroaminobutyrate by NRPS and mononuclear iron halogenase action in a streptomycete. *Chemistry & Biology*, *13*(11), 1183–1191. https://doi.org/10.1016/j.chembiol.2006.09.012.

Vaillancourt, F. H., Yeh, E., Vosburg, D. A., Garneau-Tsodikova, S., & Walsh, C. T. (2006). Nature's inventory of halogenation catalysts: Oxidative strategies predominate. *Chemical Reviews*, *106*(8), 3364–3378. https://doi.org/10.1021/cr050313i.

Vaillancourt, F. H., Yeh, E., Vosburg, D. A., O'Connor, S. E., & Walsh, C. T. (2005). Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis. *Nature*, *436*(7054), 1191–1194. https://doi.org/10.1038/nature03797.

Vaillancourt, F. H., Yin, J., & Walsh, C. T. (2005). SyrB2 in syringomycin E biosynthesis is a nonheme Fe^{II} alpha-ketoglutarate- and O2-dependent halogenase. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(29), 10111–10116. https://doi.org/10.1073/pnas.0504412102.

van Pee, K. H. (2012). Enzymatic chlorination and bromination. *Methods in Enzymology*, *516*, 237–257. https://doi.org/10.1016/b978-0-12-394291-3.00004-6.

White, M. C. (2012). Adding aliphatic C-H bond oxidations to synthesis. *Science*, *335*(6070), 807–809. https://doi.org/10.1126/science.1207661.

Williams, P. G., Buchanan, G. O., Feling, R. H., Kauffman, C. A., Jensen, P. R., & Fenical, W. (2005). New cytotoxic salinosporamides from the marine Actinomycete Salinispora tropica. *Journal of Organic Chemistry*, *70*(16), 6196–6203. https://doi.org/10.1021/jo050511+.

Wuosmaa, A. M., & Hager, L. P. (1990). Methyl chloride transferase: A carbocation route for biosynthesis of halometabolites. *Science*, *249*(4965), 160–162.

Zhu, Q., & Liu, X. (2017). Characterization of non-heme iron aliphatic halogenase WelO5* from Hapalosiphon welwitschiiIC-52-3: Identification of a minimal protein sequence motif that confers enzymatic chlorination specificity in the biosynthesis of welwitindolelinones. *Beilstein Journal of Organic Chemistry*, *13*, 1168–1173. https://doi.org/10.3762/bjoc.13.115.