Biosynthesis and Biocatalysis



International Edition: DOI: 10.1002/anie.201601447 German Edition: DOI: 10.1002/ange.201601447

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

Matthew L. Hillwig⁺, Qin Zhu⁺, Kuljira Ittiamornkul, and Xinyu Liu*

Abstract: The elucidation of enigmatic enzymatic chlorination timing in ambiguine indole alkaloid biogenesis led to the discovery and characterization of AmbO5 protein as a promiscuous non-heme iron aliphatic halogenase. AmbO5 was shown capable of selectively modifying seven structurally distinct ambiguine, fischerindole and hapalindole alkaloids with chlorine via late-stage aliphatic C-H group functionalization. Cross-comparison of AmbO5 with a previously characterized aliphatic halogenase homolog WelO5 that has a restricted substrate scope led to the identification of a Cterminal sequence motif important for substrate tolerance and specificity. Mutagenesis of 18 residues of WelO5 within the identified sequence motif led to a functional mutant with an expanded substrate scope identical to AmbO5, but an altered substrate specificity from the wild-type enzymes. These observations collectively provide evidence on the evolvable nature of AmbO5/WelO5 enzyme duo in the context of hapalindole-type alkaloid biogenesis and implicate their promise for the future development of designer biocatalysis for the selective late-stage modification of unactivated aliphatic carbon centers in small molecules with halogens.

Protein-based catalysts are promising candidates for chemo-, regio- and stereoselective late-stage functionalization of sp³hybridized carbon centers with halogen groups, a challenging transformation that is poorly represented in the inventory of modern chemical transformations.^[1] Aliphatic halogenase, initially discovered and characterized in the biogenesis of polyketides and nonribosomal peptides,^[2] belongs to a subset of iron-/2-oxo-glutarate (Fe/2OG) dependent oxygenases and represents the only known enzymatic strategy for halogenation of unactivated aliphatic carbons. While the founding members of Fe/2OG halogenases, typified by SyrB2 in syringomycin biosynthesis,^[2b] served as excellent models for mechanistic studies,^[3]

they only operate on small (amino)acyl substrates covalently tethered as a thioester to the phosphopantetheine arm of their cognate carrier proteins (CPs). This proteinaceous nature of the substrates limits their development into synthetically useful biocatalysts for tailoring small molecules.

Recently, in the biosynthetic investigation of hapalindoletype indole alkaloids,^[4] we proposed the existence of a new type of aliphatic halogenase that can act on free nonchlorinated hapalindole-type molecules to introduce chlorine substitution into this family of natural products via late-stage aliphatic C–H activation.^[4a,b] We subsequently experimentally validated that Fe/2OG halogenase WelO5 can operate on 12-*epi*-fischerindole U (1) and 12-*epi*-hapalindole C (2) to give 12-*epi*-fischerindole G (1a) and 12-*epi*-hapalindole E (2a) in the context of welwitindolinone biogenesis (Figure 1a).^[4c] The discovery of WelO5 as the founding member of a new family of CP-independent aliphatic halogenases revives the opportunity to utilize Fe/2OG halogenase for biocatalyst development.^[5] To assess their potential as useful biocatalysts, knowledge on the promiscuity and evolvability



Figure 1. Representative chlorinated and deschlorinated pairs of hapalindole-type molecules from a) the welwitindolinone producers (Ref. [6]) and b) the ambiguine producers (Ref. [7]).

- [*] Dr. M. L. Hillwig,^[+] Dr. Q. Zhu,^[+] K. Ittiamornkul, Prof. Dr. X. Liu Department of Chemistry, University of Pittsburgh 219 Parkman Avenue, Pittsburgh, PA 15260 (USA) E-mail: xinyuliu@pitt.edu
- [⁺] These authors contributed equally to this work.
- Supporting information for this article (experimental details and Figures S1–S10) can be found under: http://dx.doi.org/10.1002/ anie.201601447.

of these enzymes towards small molecular substrates is required.

WelO5 halogenase has a restricted substrate scope based on our initial characterization.^[4c] This prompted us to search for related enzyme system to gain better understanding on the promiscuity and evolvability of this newly discovered enzyme family towards alternative substrates. Here we report the

5780 Wiley Online Library

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

biochemical characterization of a second CP-independent aliphatic halogenase AmbO5 involved in ambiguine alkaloid biogenesis and its in-depth substrate scope comparison with WelO5. We show AmbO5 is promiscuous towards a panel of structurally distinct ambiguine, fischerindole and hapalindole alkaloids. The sequence similarity between WelO5 and AmbO5 allowed us to further demonstrate that a change of 18 amino acid (aa) residues of the *wt*-WelO5 protein sequence readily enhances its substrate scope to that of AmbO5, providing evidence on the natural evolvability of this new family of aliphatic halogenase towards small molecules.

AmbO5 was discovered in the initial identification of the ambiguine biosynthetic gene cluster and assigned as a putative standalone aliphatic halogenase by us.^[4a] It shares 79% sequence identity with WelO5 (Figure S1 in the Supporting Information). Distinct from WelO5,^[4c] of which natural substrates (1 and 2) can be readily deduced by inspecting the structural diversity of hapalindole-type alkaloids in the welwintindolinone producer (Figure S2),^[6] those of AmbO5 are less obvious. Ambiguine producers naturally generate up to seven pairs of deschlorinated and chlorinated ambiguines (Figure S3),^[7] typified by ambiguines H/A (4/4a), C/B (5/5a), L/K (6/6 a), I/E (7/7 a) and J/D (8/8 a) that differ sequentially by 2-electron oxidation (Figure 1b). These molecular pairs (4-8/4a-8a) share the core indole monoterpenoid scaffolds of hapalindoles U/G (3/3a) (Figure 1b), suggesting they are biosynthetically related. We have previously demonstrated that AmbP3, an aromatic prenyltransferase unique to the ambiguine biosynthetic pathway, can convert **3a** to **4a**,^[3a] supporting their biogenetic relationship. Besides ambP3 and ambO5, the ambiguine biosynthetic gene cluster harbors four additional genes that encode Rieske-type oxygenases (Figure 2a, colored in orange) as potential tailoring enzymes for the sequential oxidative maturations of 4 to 8 and 4a to 8a (Figure 2b). In order to generate the wide occurrence of deschlorinated and chlorinated structural pairs observed uniquely in the ambiguine producers, two biosynthetic diversification strategies involving AmbO5 can be envisioned (Figure 2b), provided the catalytic competency of AmbO5 is less than or similar to the remaining tailoring enzymes AmbP3 and AmbO1-4. One possibility is that AmbO5 acts on 3 only, a structural isomer of 1 and 2. Under this presumption, AmbO5 with a stringent substrate preference will likely result in a mixed pool of 3 and 3a in the producer strain that can be



Figure 2. a) Organization of *amb* genes coding tailoring enzymes and b) their proposed roles in the biosynthetic maturations of ambiguines from **3**.

Angew. Chem. Int. Ed. 2016, 55, 5780–5784

subsequently converted to 4 and 4a by AmbP3, followed by sequential oxidative maturation by AmbO1-4. An alternative possibility is that AmbO5 has a more relaxed substrate specificity and can act on every deschlorinated substrate (3– 8), along the sequential maturation pathway of 3 to ambiguines 4–8. To distinguish these two functional proposals on AmbO5 in the diversification of ambiguines, investigating the biosynthetic conversions of 3/3a and 4/4a structural pairs by AmbP3 and AmbO5 constitutes a logical starting point (Figure 3a–d).



Figure 3. Elucidation of enzymatic chlorination timing in ambiguine alkaloid biogenesis. a–d) HPLC-based comparative analyses of AmbP3 and AmbO5 enzymatic activities in the biosynthetic conversions of **3**/**3 a** and **4**/**4 a**. e–h) HPLC-based analysis of AmbO5 substrate preference towards ambiguines **4–7**. Shown in y-axis for all chromatographs is the relative absorbance at 280 nm.

In the presence of dimethylallylpyrophosphate (DMAPP) (0.5 mм), AmbP3 (2 µм) readily converted 3 (0.5 mм) to a new product (Figure 3a), of which structure was rigorously proved to be 4 by a combined NMR and MS analysis (Figure S4). AmbP3-mediated conversion of 3 to 4 appeared as efficient as that of **3a** to **4a** under identical assay conditions, shown by HPLC analysis (Figure 3a,b). This was further supported by the comparable apparent k_{cat} value of AmbP3 towards **3** ($6.5 \pm 0.8 \text{ min}^{-1}$) and **3a** ($5.3 \pm 0.5 \text{ min}^{-1}$), indicating 3 and 3a are equally competent substrates for AmbP3. Recombinant AmbO5 halogenase, overexpressed and purified in an identical manner as WelO5,^[4c] was able to act on 3 and 4 to give new products of which retention times (Figure 3c,d) and high resolution (HR) MS data (Figure S5-S6) matched the authentic 3a and 4a, respectively, by HPLC and LC-MS analysis. Both assays, carried out under identical conditions with 20 µm of AmbO5 and 500 µm of substrate (3 or 4), also revealed the endpoint turnover of 4 to 4a is approximately 7-fold greater than that of 3 to 3a, indicating 4 is a preferred substrate for AmbO5.

The comparative functional analysis of AmbP3 and AmbO5 provides initial evidence that the ambiguine biosynthetic maturation starts with 3, proceeds preferentially with prenylation to give 4 that is a more competent substrate for subsequent chlorination. The ability of AmbO5 to act on both 3 and 4 with the latter as a preferred substrate indicates AmbO5 is likely a more promiscuous aliphatic halogenase and can process a wide range of ambiguines (Figure 2b). To validate this hypothesis, we procured all remaining representative deschlorinated ambiguines (5, 6, 7) in each distinct oxidation state (Figure 1b), except 8 that is not accessible either by isolation or synthesis in our hands. When ambiguines 5/6/7 were subjected individually to AmbO5 under the identical assay conditions applied to 4, the generation of monochlorinated adducts was readily observed (Figure 3eg). The authenticity of each enzymatic product was validated to be 5a, 6a and 7a, by either comparative HPLC and HRMS analysis with the standards (for 5a and 7a) (Figures S7 and S8) or by ¹H NMR and HRMS analysis of the purified enzymatic product (for 6a) (Figure S9 and S10). To gain insight on the relative preference of AmbO5 towards different ambiguines, a series of competition experiments with equimolar amounts of 4 and 5 or 6 or 7 (0.25 mm each) were carried out with AmbO5 (20 µM) (Figure 3 h). Subsequent HPLC analysis revealed that the turnover of 4 to 4a under these experimental conditions is nearly identical with those of 5 to 5a, 6 to 6a and 7 to 7a (Figure 3h), confirming that ambiguines 4-7 are equally competent substrates for AmbO5 (vide infra). These observations collectively define the overall enzymatic timing of chlorination in ambiguine biogenesis and provide conclusive evidence that the observed structural diversity of deschlorinated and chlorinated ambiguines is due to the promiscuous AmbO5 halogenase that can act on the deschlorinated intermediates, typified by ambiguines 4-7, parallel to their biosynthetic maturation by sequential prenylation and oxidations.

The promiscuous nature of AmbO5 towards small molecular substrates prompted us to compare its substrate scope with WelO5. We collected all seven known hapalindole-type alkaloids (1-7) that are validated substrates for either WelO5^[4c] or AmbO5 and subjected them (0.5 mM) to parallel assay with each halogenase protein (20 µM) under identical assay conditions. The comparative HPLC analysis revealed WelO5 was only able to act on 1 and 2, reasserting its narrow substrate scope (Figure 4 a–g, bottom traces). Surprisingly, AmbO5 also processed 1 and 2 as efficiently as WelO5 (Figure 4 a,b). To establish the substrate preference of AmbO5, we measured its apparent k_{cat} and that of WelO5 towards substrates 1-7 (Figure 4 inset). These data show that 1 is processed most efficiently by AmbO5, followed by 2 and 4–7 being the second best group of substrates and 3 being the least preferred substrate.

It is noteworthy that the small molecular substrate scope of AmbO5 is significantly broadened from its close homolog WelO5 (Figure 4h). From a mechanistic perspective, it is unexpected that AmbO5 can maintain chemo-, regio- and stereoselective chlorination on a range of substrates (1–7) that are structurally distinct from each other in both functional group types/densities and carbon skeleton connectiv-



Figure 4. HPLC-based comparative analyses of AmbO5 and WelO5 substrate scopes. Shown in *y*-axis for all chromatographs is the relative absorbance at 280 nm. Inset: summary of AmbO5 and WelO5 k_{cat} values on substrates 1–7.

ities/conformations. Previous study on SyrB2, the founding member of Fe/2OG dependent aliphatic halogenase, showed a subtle chemical modification of its native substrate Lthreonine (such as to L-2-aminobutyric acid) tethered to the CP domain of SyrB1 readily derailed the chemoselectivity of the enzyme and led to a mixture of hydroxylated and chlorinated products.^[3d] AmbO5 and WelO5 share identical first sphere iron-coordinating residues (H164, G166, H259) (Figure S1), analogous to those in CP-dependent aliphatic halogenase SyrB2.^[3a] Assuming they both use *cis*-haloferryl reactive intermediate in their catalytic cycles for hydrogen abstraction as observed for CP-dependent halogenase,^[3b,c] these newly discovered halogenation enzymes are likely equipped with novel structural elements to control the selective chlorine rebound to the target carbon radical.

From a biocatalyst development perspective, the current observation is equally noteworthy. The intersected nature of AmbO5 and WelO5 in both protein sequences (Figure S1) and substrate scopes (Figure 4h) implicates the halogenase duo may be naturally evolved from each other or a common ancestor via divergent evolution. This provides a logical entry point to study the protein elements in this newly discovered enzyme family to confer small molecular substrate tolerance in the absence of a substrate-bound protein crystal structure.

To provide insights on the inter-evolvable nature of AmbO5 and WelO5 on substrates 1–7, we first generated a pair of chimeric proteins by fusing the N-terminal WelO5 (aa1-145) with C-terminal AmbO5 (aa146-290) to give WelO5AmbO5 and vice versa to give AmbO5WelO5 (Figure 5 a). Dividing both proteins at 145-6 residues was based on a statistical consideration as the WelO5/AmbO5 protein sequences differ by 62 aa residues (Figure S1) with 30 of those at their N-termini (aa1-145) and 32 at their C-termini (aa146-290). The enzymatic activities of these chimeras were initially assessed using **5**, selected from the second best group of substrates for AmbO5 that cannot be processed by *wt*-WelO5 for chlorination. As shown in Figure 5 to **5a**, albeit at a lower







Figure 5. Evolvable nature of WelO5 and AmbO5 halogenation activities on substrates 1–7. a) Generation of AmbO5 and WelO5 chimeras; b) Protein sequence difference of AmbO5 and WelO5 at residues 213–235; c) HPLC-based comparative analyses of AmbO5WelO5, WelO5AmbO5, WelO5AmbO5. WelO5AmbO5. NeelO5AmbO5-18 chimeric activities on ambiguine **5** relative to *wt*-AmbO5. Inset: summary of WelO5AmbO5(-18) k_{cat} values on substrates 1–7.

efficiency comparing with wt-AmbO5, whereas AmbO5-WelO5 showed no measurable activity towards 5 under identical assay conditions. The observation that WelO5 can gain activity on 5 by changing 32 aa residues at its C-terminus is significant; as it represents the first example of successfully engineering an Fe/2OG halogenase for selectively chlorinating an alternative small molecular substrate that is otherwise unprocessable by the wt enzyme. This result further implicates the C-terminal residues of AmbO5 may play a role on its broadened substrate scope and prompted us to search for additional unique sequence motif(s). We noticed among the 32 residues that differ between AmbO5 and WelO5 at their Ctermini, 18 of them cluster densely between residues 213 and 235 (Figures 5C and S1), distinct from the remaining differences that spread more evenly throughout the entire protein sequence. To test the functional role of this sequence motif, we generated a third chimeric construct by mutating 18 residues of WelO5 between aa213 and aa235 to those of AmbO5 to give WelO5AmbO5-18. Comparative HPLCbased analysis showed that this new chimera could act on 5 (Figure 5b, trace 4) with similar efficiency as WelO5AmbO5 chimera. The abilities of WelO5AmbO5(-18) chimeras to process substrates 1-7 were individually assessed and the corresponding apparent k_{cat} values were deducted (Figure 5, inset). Cross-comparing the turnover rates of substrates 1-7 by WelO5AmbO(-18) chimeras and wt-WelO5/wt-AmbO5 (Figure 5, inset), revealed several notable features. First, WelO5AmbO(-18) gained activities on ambiguines 4-7, which are not substrates for wt-WelO5, albeit with lower (ca. 8-fold) k_{cat} values in comparison with those of *wt*-AmbO5. Second, WelO5AmbO(-18) showed diminished activities on 1-2, of which k_{cat} values are reduced by ca. 10-fold from those of wt-WelO5 and wt-AmbO5. Third, WelO5AmbO5(-18) could process 3 as efficiently as wt-AmbO5, a substrate that cannot be processed by wt-WelO5. Last, WelO5AmbO5(-18) chimeras shared the aforementioned features as their apparent k_{cat} towards substrates 1-7 are nearly identical. Thus, a modification of as few as 18 residues of WelO5 sequence, not only led to an expanded substrate scope from wt-WelO5, but also a completely altered substrate specificity profile from wt-AmbO5 and wt-WelO5. To corroborate the apparent kinetic data of WelO5AmbO(-18) on substrates 1-7, AmbO5WelO5 chimera was profiled analogously (see Method section in the Supporting Information). AmbO5WelO5 retained catalytic competency on **1** with a k_{cat} of $1.9 \pm 0.2 \text{ min}^{-1}$ but showed diminished activity on **2** with a k_{cat} of $0.16 \pm 0.01 \text{ min}^{-1}$, ca. 4.5-fold reduction from wt- WelO5 and wt-AmbO5, and no activity on 3-7. These data collectively provides compelling evidence on the evolvable nature of this newly discovered halogenase enzyme family towards small molecules 1-7 and the C-terminal sequence, marked by the aa213-235 motif, is likely an important contributor to the observed substrate tolerance and specificity. The fact that the AmbO5/WelO5 chimeras no longer share the most preferred substrate, different from what was observed for wt-WelO5 and wt-AmbO5, suggests a divergent evolution from a common ancestor protein may be operant.

In summary, we discovered and characterized AmbO5 protein in ambiguine biosynthetic pathway as a promiscuous Fe/2OG aliphatic halogenase with a significantly broadened substrate scope than its close homolog WelO5. AmbO5 can selectively modify the C13 aliphatic carbon center of seven structurally distinct ambiguine, fischerindole and hapalindole alkaloids with chlorine. This in-depth characterization allows us to conclusively define the timing of enzymatic chlorination in ambiguine alkaloid biosynthesis as parallel to prenylation and oxidative maturations by Rieske-type oxygenases. Furthermore, the generation and characterization of three AmbO5 and WelO5 chimeric proteins provided evidence that a C-terminal 23aa-sequence motif plays a role in substrate tolerance and specificity. These data supports the evolvable nature of this halogenase duo for small molecular substrates in the context of hapalindole-type alkaloid biogenesis. Currently, no chemical method is available for selective chlorinating unactivated aliphatic carbon centres in complex molecular scaffolds such as 1-7. The results shown in this work constitute a promising lead for future engineering and application of this newly discovered enzyme family as selective biohalogenation catalysts for tailoring bioactive small molecules. This enzymatic approach is expected to be complementary to the parallel evolving chemical methods for selective halogenation at sp³-carbon centers in less complex molecular scaffolds.^[8]

Acknowledgements

We thank the University of Pittsburgh for supporting this work and Prof. Phil Baran and Dr. Yoshihiro Ishihara (The Scripps Research Institute) for generously providing **3**.

Keywords: alkaloid biogenesis \cdot C–H activation \cdot evolution \cdot halogenase \cdot non-heme iron enzyme

How to cite: Angew. Chem. Int. Ed. 2016, 55, 5780–5784 Angew. Chem. 2016, 128, 5874–5878





- W. J. Chung, C. D. Vanderwal, Angew. Chem. Int. Ed. 2016, 55, 4396-4434; Angew. Chem. 2016, 128, 4470-4510.
- [2] a) F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. E. O'Connor, C. T. Walsh, Nature 2005, 436, 1191-1194; b) F. H. Vaillancourt, J. Yin, C. T. Walsh, Proc. Natl. Acad. Sci. USA 2005, 102, 10111-10116; c) D. P. Galonić, F. H. Vaillancourt, C. T. Walsh, J. Am. Chem. Soc. 2006, 128, 3900-3901; d) M. Ueki, D. P. Galonic, F. H. Vaillancourt, S. Garneau-Tsodikova, E. Yeh, D. A. Vosburg, F. C. Schroeder, H. Osada, C. T. Walsh, Chem. Biol. 2006, 13, 1183-1191; e) C. S. Neumann, C. T. Walsh, J. Am. Chem. Soc. 2008, 130, 14022-14023; f) L. Gu, B. Wang, A. Kulkarni, T. W. Geders, R. V. Grindberg, L. Gerwick, K. Hakansson, P. Wipf, J. L. Smith, W. H. Gerwick, D. H. Sherman, Nature 2009, 459, 731-735; g) W. Jiang, J. R. Heemstra, Jr., R. R. Forseth, C. S. Neumann, S. Manaviazar, F. C. Schroeder, K. J. Hale, C. T. Walsh, Biochemistry 2011, 50, 6063-6072; h) S. M. Pratter, J. Ivkovic, R. Birner-Gruenberger, R. Breinbauer, K. Zangger, G. D. Straganz, ChemBioChem 2014, 15.567-574.
- [3] a) L. C. Blasiak, F. H. Vaillancourt, C. T. Walsh, C. L. Drennan, *Nature* 2006, 440, 368–371; b) D. P. Galonić, E. W. Barr, C. T. Walsh, J. M. Bollinger, Jr., C. Krebs, *Nat. Chem. Biol.* 2007, 3, 113–116; c) M. L. Matthews, C. M. Krest, E. W. Barr, F. H. Vaillancourt, C. T. Walsh, M. T. Green, C. Krebs, J. M. Bollinger, *Biochemistry* 2009, 48, 4331–4343; d) M. L. Matthews, C. S. Neumann, L. A. Miles, T. L. Grove, S. J. Booker, C. Krebs, C. T.

Walsh, J. M. Bollinger, Jr., Proc. Natl. Acad. Sci. USA 2009, 106, 17723-17728.

- [4] a) M. L. Hillwig, Q. Zhu, X. Liu, ACS Chem. Biol. 2014, 9, 372–377; b) M. L. Hillwig, H. A. Fuhrman, K. Ittiamornkul, T. J. Sevco, D. H. Kwak, X. Liu, ChemBioChem 2014, 15, 665–669; c) M. L. Hillwig, X. Liu, Nat. Chem. Biol. 2014, 10, 921–923; d) X. Liu, M. L. Hillwig, L. I. M. Koharudin, A. M. Gronenborn, Chem. Commun. 2016, 52, 1737–1740.
- [5] S. Brown, S. E. O'Connor, ChemBioChem 2015, 16, 2129.
- [6] K. Stratmann, R. E. Moore, R. Bonjouklian, J. B. Deeter, G. M. L. Patterson, S. Shaffer, C. D. Smith, T. A. Smitka, J. Am. Chem. Soc. 1994, 116, 9935.
- [7] a) T. A. Smitka, R. Bonjouklian, L. Doolin, N. D. Jones, J. B. Deeter, W. Y. Yoshida, M. R. Prinsep, R. E. Moore, G. M. L. Patterson, *J. Org. Chem.* **1992**, *57*, 857; b) A. Raveh, S. Carmeli, *J. Nat. Prod.* **2007**, *70*, 196.
- [8] a) W. Liu, J. T. Groves, J. Am. Chem. Soc. 2010, 132, 12847–12849; b) V. A. Schmidt, R. K. Quinn, A. T. Brusoe, E. J. Alexanian, J. Am. Chem. Soc. 2014, 136, 14389–14392; c) R. K. Quinn, Z. A. Konst, S. E. Michalak, Y. Schmidt, A. R. Szklarski, A. R. Flores, S. Nam, D. A. Horne, C. D. Vanderwal, E. J. Alexanian, J. Am. Chem. Soc. 2016, 138, 696–702.

Received: February 10, 2016 Published online: March 30, 2016