

Structures, mechanisms and applications of flavin-dependent halogenases

**Asisaphon Phintha^a, Kridsadakorn Prakinee^b,
and Pimchai Chaiyen^{b,*}**

^aDepartment of Biochemistry and Center for Excellence in Protein and Enzyme Technology, Faculty of Science, Mahidol University, Bangkok, Thailand

^bSchool of Biomolecular Science and Engineering, Vidyasirimedhi Institute of Science and Technology (VISTEC), Wangchan Valley, Rayong, Thailand

*Corresponding author: e-mail address: pimchai.chaiyen@vistec.ac.th

Contents

1. Introduction	328
2. Structures	331
2.1 Overall structures of FDHs	331
2.2 Flavin binding site of FDHs	331
2.3 Substrate binding site of FDHs	337
3. Structures and reactions of halogenases classified according to substrates	338
3.1 Indole halogenases	338
3.2 Pyrrole halogenases	342
3.3 Phenolic halogenases	343
3.4 Aliphatic halogenases	345
3.5 Halogenase utilizes diverse substrates	347
4. Mechanisms	349
5. Improvement of the catalytic properties of FDHs using structure-guided mutagenesis	352
5.1 Substrate scope expansion	352
5.2 Site-selective mutagenesis	353
6. Applications of FDHs	355
6.1 Scale-up biocatalytic halogenation	355
6.2 Engineered biosynthetic pathway	356
6.3 Late-stage diversification of halogenated compounds	357
7. Future perspectives	360
References	361

Abstract

Overall, this review highlights the structures, mechanisms and applications of flavin-dependent halogenases (FDHs) for future development of FDHs as potential biocatalysts. FDHs catalyze incorporation of halogen atoms into a broad range of

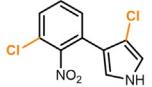
substrates. The reactions involved in the production of various halogenated natural products which are important drugs. Typical substrates for FDHs include indole, pyrrole, phenolic and aliphatic compounds. In addition to organic substrates, all FDHs utilize reduced FAD (FADH^-), oxygen and halides as co-substrates. Structural studies reveal that FDHs all have similar FAD binding sites. However, FDHs have variations between the different isotypes including different recognition residues for substrate binding and some unique loop structures and conformations. These different structural differences suggest that variations in reaction catalysis exist. However, limited knowledge of the reaction mechanisms of FDHs is currently available. Various biocatalytic applications of FDHs have been explored. Further investigation of the catalytic reactions of FDHs is essential for improving enzyme engineering work to enable FDHs catalysis of challenging reactions.



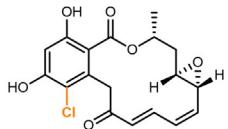
1. Introduction

FDHs are involved in the synthetic pathways of various secondary metabolites and natural products that are either important active pharmaceutical ingredients or are common drugs. FDHs can modify specific precursors such as indoles, pyrroles, and phenolic and aliphatic molecules by incorporating halogen atoms such as chlorine, bromine and iodide into these compounds [1–3]. Halogen incorporation is a highly important reaction, as it alters the molecular properties of the substrate compounds, including their bioactivities and physiological functions (Fig. 1). Natural products which have incorporated halogen atoms (halometabolites) often show anti-fungal and antibacterial activities. For example, vancomycin is an antibiotic produced by bacteria used for the treatment of infections caused by multi-drug resistant (MDR) Gram-positive bacteria in hospitals [4], while rebeccamycin can be used for the treatment of tumors and cancers (Fig. 1A) [5]. Furthermore, a few products of FDHs, including thienodolin and 6-chloro-1H-indole-3-carboxamide (LYXLF2) have been shown to be plant growth-regulating compounds (Fig. 1B) [6]. Traditional methods of halogenation are chemical reactions which generally require harsh conditions. Moreover, chemical halogenation also has the drawback of lacking control of regioselectivity [7]. Therefore, enzymatic halogenation is advantageous with regards to aspects of regioselectivity and sustainability. There are several types of halogenases reported to date, including FDHs, vanadium-dependent haloperoxidases, S-adenosyl-L-methionine-dependent halogenases and nonheme iron-dependent

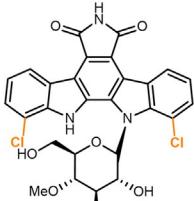
A Pharmaceuticals



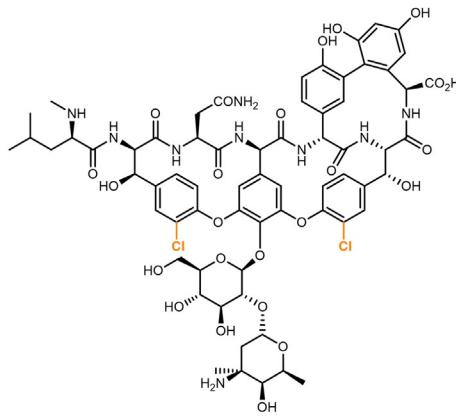
pyrrolnitrin



radicicol

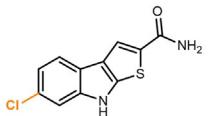


rebeccamycin

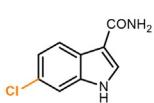


vancomycin

B Agrochemicals



thienodolin



LYXLF2

Fig. 1 Examples of valuable halometabolites resulting from halogenation reactions catalyzed by FDHs in various biosynthetic pathways. Examples of halogenated compounds which are parts of pharmaceuticals are shown in A while those used as agrochemicals are shown in B.

halogenases [8]. In this review, the structural and mechanistic aspects of FDHs and the prospects of their potential utility are discussed.

FDHs belong to the superfamily of flavin-dependent monooxygenases. FDHs are enzymes that catalyze halogenation. Most of the FDHs are two-component flavin-dependent monooxygenases in which a reduced flavin adenine dinucleotide (FADH^-) is used as one of the substrates [8]. Therefore, most of the FDHs require the reaction of flavin reductase to generate FADH^- from FAD. The only exception is the enzyme CmlS (discussed more in [Section 2.2](#)) which has FAD bound as a cofactor [9]. Despite the fact that all FDHs share the use of FADH^- , the only available crystal structure with FADH^- bound is the co-complex structure of PrnA.

Although most of FDHs use FADH^- as substrate (or in the case of chloramphenicol halogenase (CmlS) uses FAD as a cofactor), substrate utilization by FDHs is diverse. During the past decade, several types of FDHs have been discovered [1]. Most FDHs catalyze halogenation with substrate specificity, with each FDH being specific for a single group of substrates including indole, pyrrole, phenolic and aliphatic compounds [1]. The only exception was the recently discovered flavin-dependent viral halogenase (VirX1) from a cyanophage which is able to iodinate versatile substrates [3] ([Fig. 2](#)). FDHs

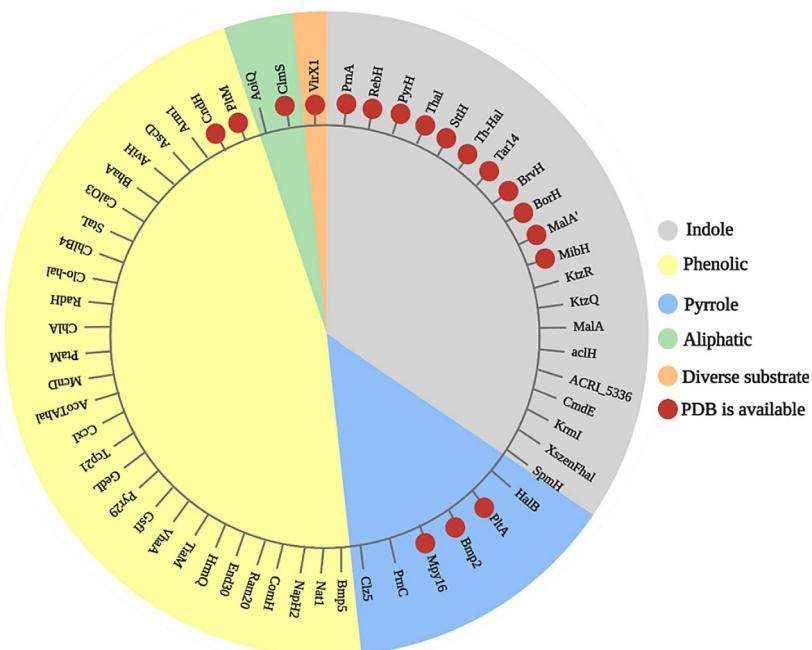
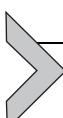


Fig. 2 FDHs discovered to date. Types of FDHs are divided based on substrate utilization. Red circles indicate that the structures of these FDHs are available.

that catalyze halogenation of indoles, phenolics and pyrroles are the majority of FDHs discovered to date (Fig. 2). Quite a few crystal structures of FDHs have been solved (Table 1). Halogenases that catalyze conversion of indoles, i.e., tryptophan halogenases are the most investigated enzymes among all types of FDHs. Crystal structures of various tryptophan halogenases have been reported [12,14,28]. Investigation of their reaction mechanisms has only been done through the reaction of RebH or tryptophan 7-halogenase [12,14,28]. Most efforts during recent years have been focused on enzyme engineering to improve enzyme stability and broaden the scope of substrates that can be used [1]. Various excellent reviews have documented recent efforts and accomplishments in FDHs engineering [1,29–31]. Therefore, our chapter aims to focus the discussion on the structure and mechanism of FDHs, and provides an overview of the current progress made with regards to their applications. Understanding the structure and mechanism of FDHs is useful for developing FDHs as potential biocatalysts for biotechnology in the future.



2. Structures

2.1 Overall structures of FDHs

Crystal structures of all FDHs consist of two separated binding sites, including a flavin binding site (N-terminus) and a substrate binding site (C-terminus) [3,9,12,25,27]. All FDHs share a common flavin or nucleotide binding architecture called a Rossman or $\beta\alpha\beta$ fold, consisting of a β -sheet flanked by helices [3,9,12,25,27] (Fig. 3). This architecture is also found in other flavoenzymes, including single-component flavin-dependent monooxygenases [1,32–34] and glutathione reductase superfamily enzymes [25]. FDHs all have similar architectures at the N-terminus. Diversity among FDHs is found more within the substrate binding site which displays more variation in length and folding [3,9,12,25,27] (Fig. 3). Therefore, the FDHs reported to date show a wide range of capabilities to catalyze the halogenation of versatile substrates.

2.2 Flavin binding site of FDHs

The flavin binding domain of all FDHs has two signature motifs including GxGxxG and WxWxIP that are highly conserved [1]. This information is useful as it has been used for identifying putative FDH genes from biosynthetic pathways. Despite the WxWxIP motif being strictly conserved in FDHs, this motif is absent in flavin-dependent brominase (Bmp5) from marine bacteria [35] and its functional role is not clearly understood.

Table 1 List of structures of FDHs currently available.

Enzymes	PDB id (mutant)	Ligands	Substrates	References
<i>Indole halogenases</i>				
Tryptophan-7 halogenase (PrnA)	4Z44	FAD, Cl ⁻	Trp	[10]
	4Z43 (E450K)	FAD, Cl ⁻		
	2JKC (E346D)	Trp, FAD, Cl ⁻		[11]
	2APG	FAD, Cl ⁻		[12]
	2AQJ	Trp, FAD, Cl ⁻		
	2ARD	FADH ⁻		
	2AR8	7-Cl-Trp, FAD, Cl ⁻		
Tryptophan-7 halogenase (RebH)	2O9Z	–		[13]
	2OA1	Trp, FAD, Cl ⁻ , ADN		
	2OAL	FAD, Cl ⁻		[14]
	2OAM	–		
	2E4G	Trp		
	4LU6	–		[15]
Tryptophan-5 halogenase (PyrH)	2WES (E46Q)	FAD, Cl ⁻		[16]
	2WET	Trp, FAD, Cl ⁻		
	2WEU	Trp		
Tryptophan-6 halogenase (Thal)	6H43	–		[17]
	6IB5 (V52I V82I S360T G469S S470N)	–		
	6H44	Trp		
	6SLS	FAD		[18]
	6SLT	Trp, FAD, AMP		

Table 1 List of structures of FDHs currently available.—cont'd

Enzymes	PDB id (mutant)	Ligands	Substrates	References
Tryptophan-6 halogenase (SttH)	5HY5	FAD, Cl ⁻		[19]
Tryptophan-6 halogenase (Th-Hal)	5LV9	—		[19]
Tryptophan-6 halogenase (Tar14)	6NSD	FAD		[20]
Tryptophan-6 halogenase (BorH)	6UL2	Trp		[21]
FDH from <i>Brevundimonas</i> sp. BAL3 (BrvH)	6FRL	—	Indole	[22]
FDH involved in biosynthesis of malbrancheamide (MalA')	5WGR	PM7, FAD, Cl ⁻	Premalbrancheamide	[23]
	5WGZ	IM7, FAD, Cl ⁻		
	5WGW	MB5, FAD, Cl ⁻		
	5WGX (H253A)	MB5, FAD, Cl ⁻		
	5WGY (C112S C128S)	MB5, FAD, Cl ⁻		
	5WGV (C112S C128S)	PM7, FAD, Cl ⁻		
	5WGU (E494D)	PM7, FAD, Cl ⁻		
	5WGT (H253A)	PM7, FAD, Cl ⁻		
	5WGS (H253F)	PM7, FAD, Cl ⁻		

Continued

Table 1 List of structures of FDHs currently available.—cont'd

Enzymes	PDB id (mutant)	Ligands	Substrates	References
Lathipeptide tryptophan 5-halogenase (MibH)	5UAO	FAD, Cl ⁻	Trp (within peptide)	[24]
<i>Pyrrole halogenases</i>				
Pyrrole halogenase involved in pyoluteorin biosynthesis (PltA)	5DBJ	FAD, Cl ⁻	Tethered pyrrole	[25]
Pyrrole halogenase from <i>Pseudoalteromonas luteoviolacea</i> (Bmp2)				
	5BVA	FAD	Tethered pyrrole	[26]
	5BUL (Y302S)	FAD		
	F306V			
	A345W)			
Pyrrole halogenase from <i>Streptomyces</i> sp. CNQ-418 (Myp16)	5BUK	FAD	Tethered pyrrole	[26]
<i>Phenolic halogenases</i>				
Phenolic halogenase involved in pyoluteorin biosynthesis (PltM)	6BZA	Benzene-1,3,5-triol, FAD, Cl ⁻	Free phenolic	[2]
	6BZZ	FAD		
	6BZQ	FAD, Cl ⁻ , Br ⁻		
	6BZN	—		
	6BZI	Ethyl mercury ion, Hg ²⁺		
	6BZT (L111Y)	FAD, Cl ⁻ , Br ⁻		
Myxobacterial chondrochloren halogenase (CndH)	3E1T	FAD, Cl ⁻	Tethered phenolic	[27]

Table 1 List of structures of FDHs currently available.—cont'd

Enzymes	PDB id (mutant)	Ligands	Substrates	References
<i>Aliphatic halogenase</i>				
Chloramphenicol halogenase (CmlS)	3I3L	FAD	Free aliphatic	[9]
<i>Halogenase catalyze diverse substrates</i>				
Flavin-dependent viral halogenase (VirX1)	6QGM	—	Diverse	[3]

Trp, tryptophan; Cl⁻, chloride ion; Br⁻, bromide ion; FAD, oxidized flavin adenine dinucleotide; FADH⁻, reduced flavin adenine dinucleotide; 7-Cl-Trp, 7-chlorotryptophan; ADN, adenosine; AMP, adenosine monophosphate; MES, 2-(N-morpholino)-ethanesulfonic acid; PM7, premalbrancheamide E; IM7, isomalbrancheamide B; MB5, malbrancheamide B; Hg²⁺, mercury (II) ion.

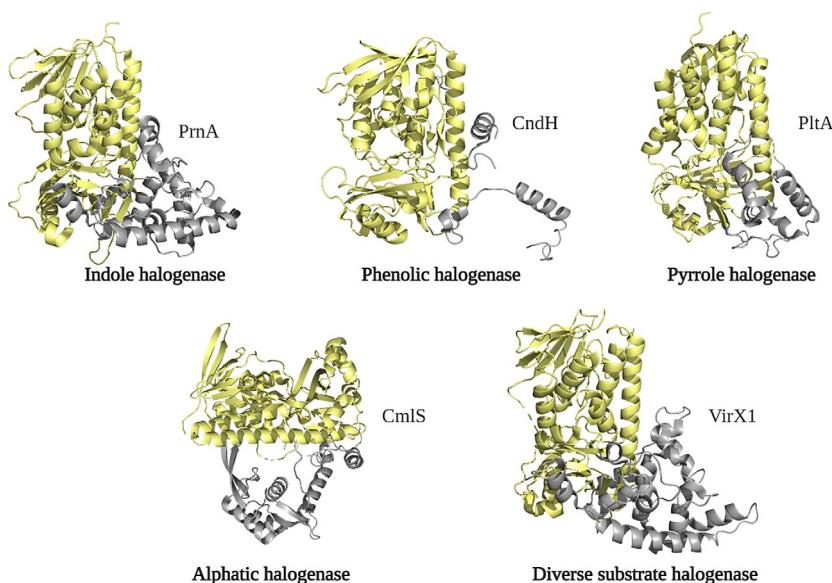


Fig. 3 Representative structures of FDHs from five groups categorized by substrate utilization. FAD binding module (pale yellow) and substrate binding module (gray). PrnA (PDB id: [2AQJ](#)), CndH (PDB id: [3E1T](#)), PltA (PDB id: [5DBJ](#)), CmlS (PDB id: [3I3L](#)) and VirX1 (PDB id: [6QGM](#)).

Dong et al. suggested that the two Trp residues might prevent substrate from binding in close proximity to the bound flavin, thus preventing the enzyme from functioning as a monooxygenase [12]. However, when these two Trp residues in PrnA were mutated to Phe, the enzyme activity was not changed [11]. Mutating Trp272 to Ala272 also did not significantly change the enzyme activity [11]. On the other hand, PrnA was totally inactive when Trp274 was changed to Ala274 [11]. Thus, the functional role of Trp272 is still unclear.

The halide binding pocket is in close proximity to the isoalloxazine ring (Fig. 4A and B). The location of halide binding at the re-face of the isoalloxazine ring is conserved in all characterized complex structures of FDHs [12,13,16,25,36] (Fig. 4). There are no obvious conserved residues that serve as a binding site for a negative ion. The clearest identification of halide is in PrnA, in which Cl^- contacts the nitrogen of Thr348 and Gly349 directly (Fig. 4B). The amide backbone of Ser357 and Gly358 in tryptophan halogenase (MibH) was identified to react with Cl^- [24]. In pyrrole halogenase (PltA), the amide backbone of residues 326–328 was identified to form contact with Cl^- [25]. Therefore, the nitrogen atoms of the amide backbone have been proposed to bind a halide ion in FDHs [12].

Pyrrole halogenase (PltA), phenolic halogenase (CndH) and aliphatic halogenase (CmlS) share a conserved loop consisting of a *cis*-peptide in

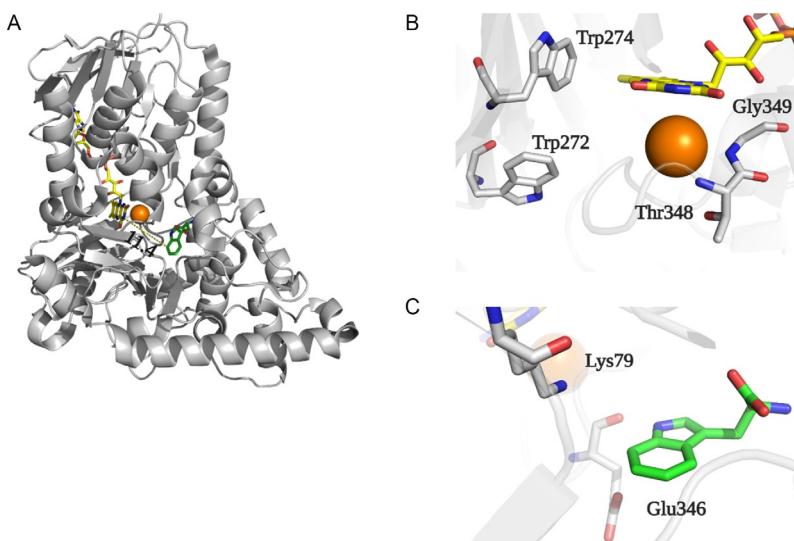


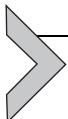
Fig. 4 The crystal structure of PrnA (PDB: 2AQJ). The overall structure of PrnA (A), FAD (yellow stick), Trp (green stick) and Cl^- (orange sphere). FAD binding sites (B). Trp binding site (C). The conserved Lys and Glu are indicated as sticks.

the FAD binding site. The location of the *cis*-peptide in PltA is between Phe39 and Glu40. This feature is absent in tryptophan halogenase. This loop in PltA is located ~10 Å away from the bound FAD, whereas the location of the similar loop in PrnA is only 6 Å from FAD [25]. Therefore, the *cis*-peptide was proposed to induce different binding conformations in PltA and PrnA [25]. However, the functional role of this *cis*-peptide in catalysis is not clearly understood and it remains to be investigated.

CmlS is the only FDH that has FAD covalently bound in the structure [8,9]. This enzyme is thus different from the other enzymes in this family in that FAD serves as a cofactor in this case. The residue Asp277 of CmlS forms a covalent bond with the 8 α carbon of FAD. Mutation of Asp277 to Asn abolishes the covalent bond [9]. This Asp is absent in tryptophan halogenases [9]. CmlS is clearly different from other FDHs with regards to its mode of binding FAD because in this case, FAD is used as a cofactor and its redox state is alternates between FAD and FADH[−]. For other FDHs, the enzyme binds to the substrate FADH[−] before the rest of the reactions proceed. Therefore, investigation of the reaction of CmlS using transient kinetics experiments should be useful to elucidate its reaction mechanisms.

2.3 Substrate binding site of FDHs

Diversity among the FDHs is mostly centered around the substrate binding site at the C-terminus, which is characterized by different lengths and topologies within the family (Fig. 3). Tryptophan 7-halogenase (PrnA) and VirX1 have well-ordered triangular C-terminus [3,12,27]. Three helices on the C-terminus of PltA are well-ordered [25]. The substrate binding site of myxobacterial chondrochloren halogenase (CndH) has a winged-helix structure [27]. The C-terminus of CmlS is largely disordered [9]. FDHs were classified into two groups based on the degree of structural order of their C-terminus, and based on their substrate utilization [9]. Group A consists of FDHs that utilize small molecule substrates (PrnA and VirX1), whereas Group B enzymes catalyze halogenation of substrates embedded in carrier proteins (CmlS and CndH) [9]. The C-terminus of group A is normally well-ordered, while the C-terminus of Group B is disordered. The exception of Group B enzymes is PltA, in which its C-terminus is well-ordered, however PltA accepts only substrates embedded in carrier proteins [25]. As the variation in FDHs is mostly in the substrate binding site, we classify our discussion of FDHs below based on the nature of their substrates.



3. Structures and reactions of halogenases classified according to substrates

3.1 Indole halogenases

Indole halogenases are halogenases that catalyze conversion of substrates containing an indole moiety. There are several substrates that can be utilized by this group of enzymes as shown in Table 1. Most of the enzymes in this group use free indole as a substrate, except the one member of the group that can use an embedded Trp in a peptide chain.

Typical substrates for this group of FDHs include Trp, indole and premalbrancheamide. Tryptophan halogenases are enzymes that can catalyze regioselective halogenation of Trp [12,16,17,19] (Fig. 5). This group of enzymes are the most investigated FDHs. Several tryptophan halogenases have been investigated including tryptophan 7-halogenase (PrnA) from *Pseudomonas fluorescens* [10–12], tryptophan 7-halogenase (RebH) from *Lentzea aerocolonigenes* [13–15], tryptophan 5-halogenase (PyrH) from *Streptomyces rugosporus* [16], tryptophan 6-halogenase (Thal) from

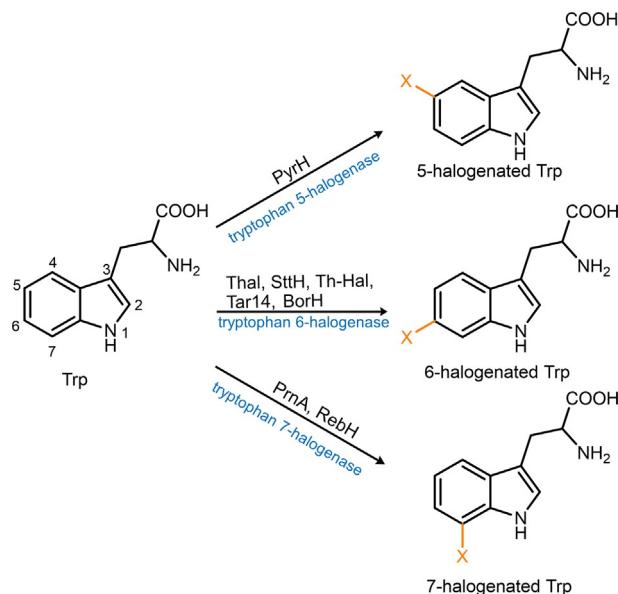


Fig. 5 Regioselective halogenation by tryptophan halogenases. X represents Cl or Br.

Streptomyces albogriseolus [17,18], tryptophan 6-halogenase (SttH) from *Streptomyces toxytricini* [19], thermophilic tryptophan 6-halogenase (Th-Hal) from *Streptomyces violaceusniger* [19], tryptophan 6-halogenase (Tar14) from *Saccharomonospora* sp. CNQ490 [20] and tryptophan 6-halogenase (BorH) from metagenomic analysis from soil [21]. Besides Trp, halogenation of indole and premalbrancheamide substrates can be catalyzed by indole halogenases, including BrvH from *Brevundimonas* sp. BAL3 and MalA' from *Malbranchea aurantiaca*, respectively. Recently, the first halogenase (MibH) that can use Trp attached to a polypeptide as a substrate has been characterized. MibH was isolated from *Microbispora* sp. ATCC PTA-5024.

The common structural similarity shared by indole halogenases is their separated flavin and substrate binding sites. The shape of the flavin binding module in tryptophan halogenases looks like a box attached to the triangular pyramid of the substrate binding domain. In tryptophan halogenases, these two binding sites are separated by about 10 Å [12,16–18] (Fig. 4). The box shape of the flavin binding module is similar for all indole halogenases [12,22–24] (Fig. 6). The majority of variation between the FDHs in this group is found in the substrate binding site which is slightly diverse among members of the group (Fig. 6).

The most well understood indole halogenase is tryptophan halogenase, which attracts a lot of attention from researchers in the field because of their promising ability to control regioselectivity. Since flavin and Trp binding sites of tryptophan halogenases are separated, it has been proposed that hypohalous acid (HOX) in the flavin binding site has to travel through a tunnel to the substrate binding site [12]. Lys and Glu in the substrate binding site are highly conserved in FDHs (Fig. 4). Mutation of this Lys to Ala can completely abolish the activity of the FDHs [12,22]. Mutation of Glu to Asp also resulted in activity reduction of PrnA [12]. Therefore, these two residues have been proposed to be key catalytic residues of indole halogenases [11,12,16,17,24]. From structural alignment of tryptophan 5-halogenase, tryptophan 6-halogenase and tryptophan 7-halogenase, the data showed that the conserved Lys is oriented directly toward the halogenated positions [16,17]. Besides Lys, there are several residues in the Trp binding site that contribute to the control of regioselectivity. Regioselectivity of tryptophan 6-halogenase (SttH) was changed from 6-halogenated tryptophan to 5-halogenated tryptophan when Leu460^{SttH}, Pro461^{SttH} and Pro462^{SttH} were mutated to be residues found in PyrH (Phe451^{PyrH}, Glu452^{PyrH} and Thr453^{PyrH}) [19]. Leu455^{Th-Hal}, Pro456^{Th-Hal} and Pro457^{Th-Hal} in thermophilic tryptophan 6-halogenase (Th-Hal)

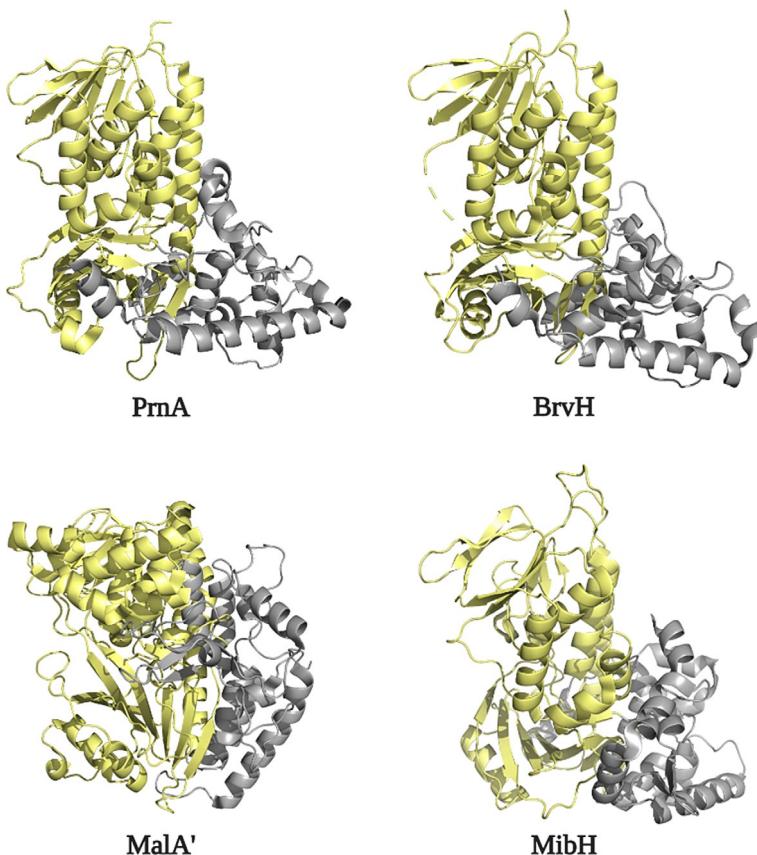


Fig. 6 Overall structures of indole halogenases. Flavin binding module (pale yellow) and substrate binding module (gray).

were proposed to be important because their orientations are similar to Leu460^{SttH}, Pro461^{SttH} and Pro462^{SttH} in SttH [37]. Regioselectivity of Thal was also altered from 6-halogenated tryptophan to 7-halogenated tryptophan when mutating Val52^{Thal}, Val82^{Thal}, Ser360^{Thal}, Gly469^{Thal} and Ser470^{Thal} to Ile52^{RebH}, Ile82^{RebH}, Asn470^{RebH}, Thr359^{RebH}, Ser469^{RebH} and Asn470^{RebH} [17]. Therefore, these residues were proposed to be key residues in controlling the regioselectivity.

The overall structure of FDH from *Brevundimonas* sp. BAL3 (BrvH) catalyzing halogenation of indole is similar to RebH [22]. However, there are some differences at the substrate binding site. Residues interacting with five- and six-membered indole rings are similar [22]. However, the significant difference is that BrvH does not have the loop which forms hydrogen bonds

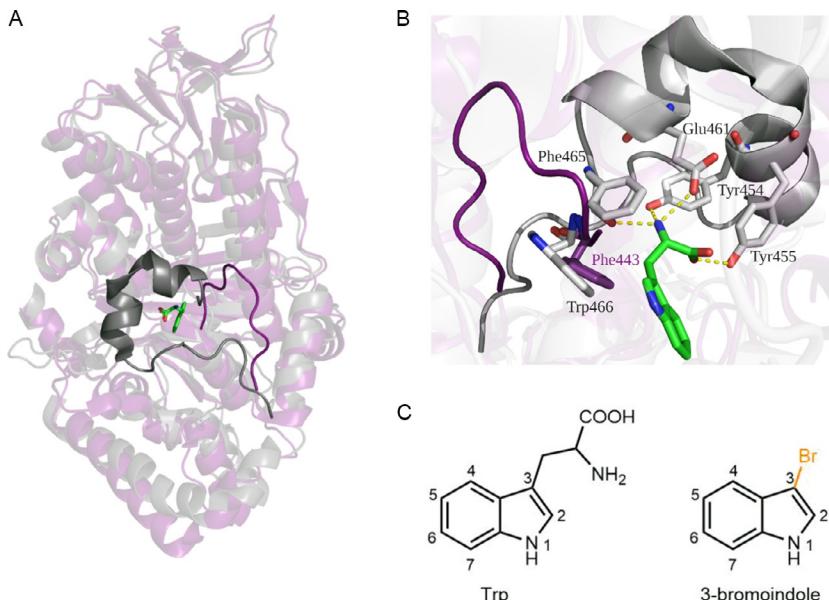


Fig. 7 Comparing structures of RebH with BrvH. Structure alignment of RebH (PDB id: 2OA1) (gray) with bound Trp (green stick) and BrvH (PDB id: 6FRL) (purple) (A). Interaction of Trp substrate with active site residues in RebH (B). Structural comparison of Trp (substrate of RebH) and 3-bromoindole (product of BrvH) (C).

with the amino and carboxy groups of the substrate as seen in RebH [22] (Fig. 7A and B). Therefore, BrvH recognizes indole as substrate instead of tryptophan. BrvH introduces Br into the C3 position of indole substrates, which is not available in tryptophan (Fig. 7C). The preference for halogenation at the C3 position may be one of the reasons why BrvH cannot use tryptophan as a substrate. Only the apo structure of BrvH has been reported to date. To clearly understand the substrate scope of BrvH, structures of BrvH co-complexed with substrate or product are needed.

In the case of MalA' and MibH which catalyze halogenation of large substrates, their substrate binding sites are much bigger than the active sites of other tryptophan halogenases so they can accommodate the binding of large compounds such as premalbrancheamide and peptide substrates [24]. MibH regioselectively catalyzes halogenation of peptide-tryptophan at the C5 position. MibH has a larger substrate binding cleft than its closest homolog PyrH. Moreover, the substrate binding cleft of MibH is lined with hydrophobic residues, supporting its preference to bind peptide-tryptophan instead of free tryptophan [23]. MalA' is a new class of FDHs because its

structure has a zinc binding site coordinated by four Cys residues. The role of zinc in MalA' catalysis is not clearly understood [23].

3.2 Pyrrole halogenases

Pyrrole halogenase catalyzes the halogenation of pyrrolic substrates either as free substrates or in the pyrrole-2-carboxyl thioester forms that are tethered to peptidyl carrier proteins [1]. Only the structures of pyrrole halogenases that catalyze halogenation of tethered substrates are available including PltA from *Pseudomonas fluorescens* (strain ATCC BAA-477/NRRL B-23932/Pf-5) and Bmp2 from *Streptomyces* and Mpy16 from *Streptomyces* sp. CNQ-418 [26]. As PltA catalyzes halogenation of substrates tethered to peptidyl carrier proteins and it has a well-ordered substrate binding site [25], PltA is an exception among the other FDHs in this regard. Most of the Group B enzymes (i.e., CmlS and CndH) have less ordered substrate binding sites [20] (Fig. 3). The crystal structure of PltA shows that the substrate binding site is in a closed state where its binding site is cut off from the bulk solvent. To accommodate large substrates, conformational changes likely take place during catalysis.

Since a structure of PltA with bound substrate is not available, the position of the catalytic Lys was used to locate the substrate binding site (Fig. 8A). The location of the substrate binding site of PltA is similar to that of PrnA, which is located at the interface of the flavin binding module and the C-terminus region (Fig. 8A and B). Catalytic Lys and Glu are strictly conserved in all tryptophan halogenases. Catalytic Glu is proposed to act as a base to deprotonate a Wheland intermediate [11,12] (Fig. 8C). However, this catalytic Glu is absent in non-trypotphan halogenases, i.e., PltA, CmlS and CndH. In these non-trypotphan halogenases, Phe is found instead of Glu. There are no other residues in proximity of the Wheland intermediate. Therefore, a nearby water bound in the active site is proposed to serve as a base in these enzymes [25].

Normally, FDHs incorporate less than four halogen atoms into their substrate. Bmp2 is the first pyrrole tetrahalogenase characterized and it can incorporate up to four halides into its substrates (REF). Comparison of the structures of Bmp2 and its homolog Mpy16 (pyrrole dihalogenase) indicates that the positions of the catalytic Lys of both enzymes are superimposable. The major difference is that Bmp2 lacks the three residues (Ser325, Val329 and Trp369) that are strictly conserved in pyrrole halogenases catalyzing single or double halogenation, such as Mpy16.

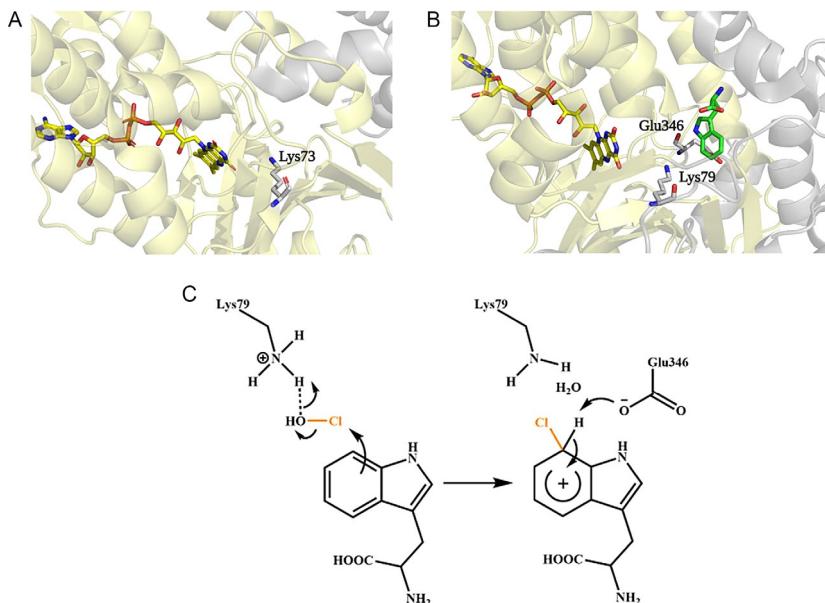


Fig. 8 Comparing substrate binding sites of PltA and PrnA. Substrate binding sites of PltA (PDB id: 5DBJ) (A). Substrate binding site of PrnA with bound Trp substrate (PDB id: 2AQJ) (B). Proposed mechanism of Wheland intermediate deprotonation by Glu346 in PrnA (C).

The triple variant of Bmp2 (Tyr302Ser, Phe306Val and Ala345Trp) cannot produce tretrahalogenated pyrrole product. These data reveal that the degree of halogenation may be controlled by residues lining the substrate binding site.

3.3 Phenolic halogenases

Phenolic halogenases catalyze the halogenation of phenolic compounds. There are three types of phenolic halogenases that have been discovered to date, including phenolic halogenases catalyzing the halogenation of free phenolics (class A) and phenolics embedded into carrier proteins (class B), and also phenolic halogenases that can catalyze decarboxylation in addition to halogenation (class C) [1]. Only the structures of phenolic halogenases class A and B have been reported including the structures of PltM from *Pseudomonas fluorescens* (strain ATCC BAA-477/NRRL B-23932/Pf-5) [2] and CndH from *Chondromyces crocatus* [27], respectively. PltM is the first reported FDH that can iodinate substrates [2].

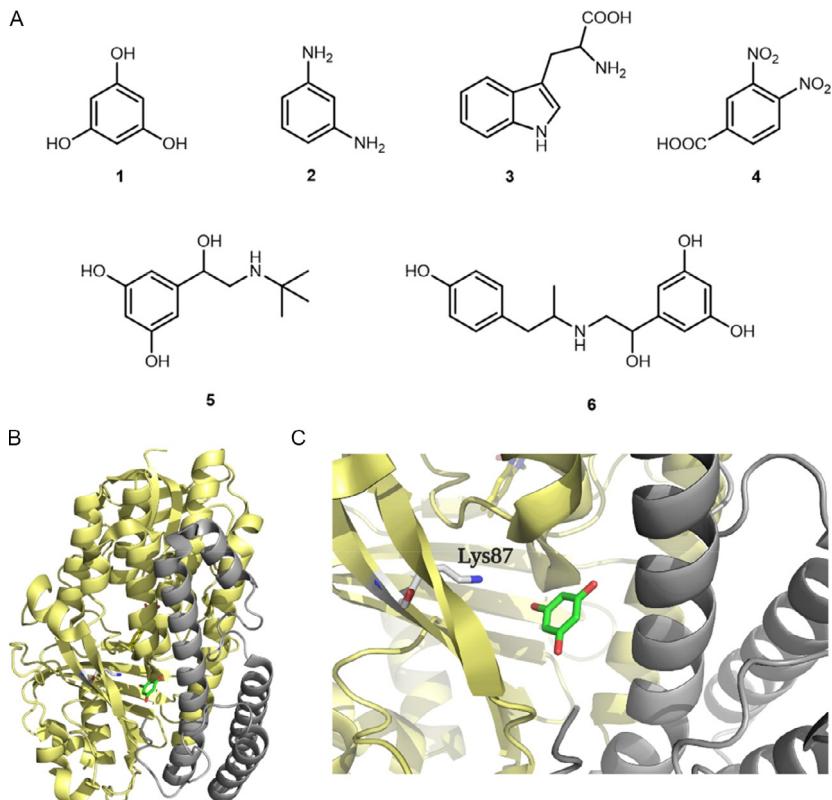


Fig. 9 Substrate scope and structure of PltM (PDB id: [6BZA](#)). Substrate scope of PltM (A). Overall structure of PltM (B). Substrate binding site of PltM (C). FAD binding module (pale yellow), substrate binding site (gray). The conserved Lys is indicated as sticks. Phenolic substrate [1] is indicated in green stick.

PltM catalyzed halogenation of versatile small molecule substrates including phenolic (**1**) and aniline (**2**) derivatives (Fig. 9A). Interestingly, bulky substrates including terbutaline (**5**) and fenoterol (**6**) which are FDA-approved drugs can also be halogenated by PltM (Fig. 9A). However, tryptophan (**3**) and nitrobenzene (**4**) cannot be used as substrates by PltM (Fig. 9A). The substrate binding site of PltM is close to the protein surface which may allow the access of large substrates (Fig. 9B). Therefore, fenoterol and terbutaline can bind to PltM. However, the reason why tryptophan cannot be used as a substrate is still unclear. PltM has a conserved Lys in the substrate binding site which is the same as others FDHs (Fig. 9C). Mutation of this Lys87 to Ala completely abolished the activity of PltM. Further, there are charged and polar residues in the substrate binding site,

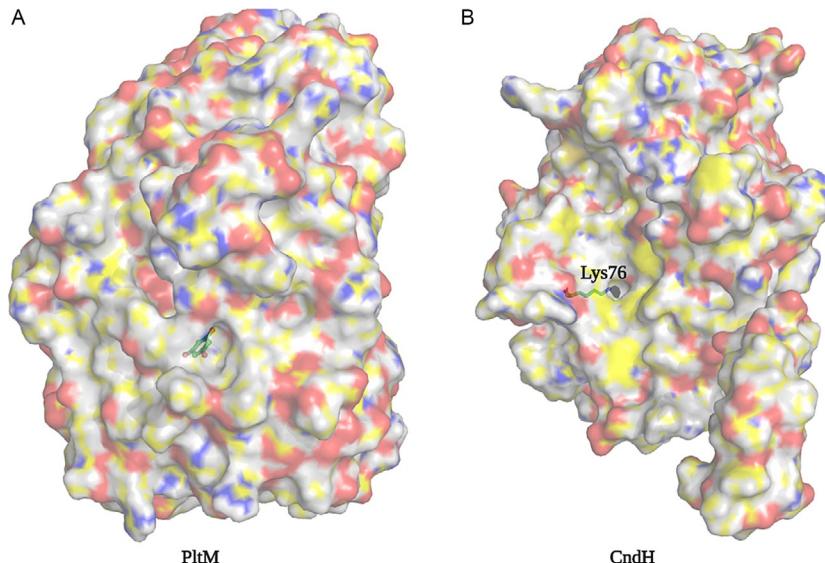


Fig. 10 Surface hydrophobicity of PltM (PDB id: [6BZA](#)) and CndH (PDB id: [3E1T](#)). Carbon, nitrogen, oxygen and hydrogen are colored as yellow, blue, red, and white, respectively. Phenolic substrate bound in PltM was indicated in green stick (A). Lys76 of CndH was indicated in green stick (B).

which may interact with the hydroxy and amine groups of substrates including phenolic and aniline substrates, respectively (Fig. 9A).

Comparing the structures of CndH with PltM reveal that the structure of CndH has a hydrophobic patch on the surface for catching a carrier protein substrate [27] as represented by hydrophobicity of carbon atom (yellow) (Fig. 10A and B). The C-terminus of CndH is disordered, which is the signature of Group B FDHs catalyzing halogenation of substrates embedded in carrier proteins. The active site of CndH contains the conserved catalytic Lys (Lys76). However, CndH lacks the catalytic base Glu. Unfortunately, some parts of the reported CndH structure are missing; this may be the result of denaturation upon crystal packing [27]. To fully understand the enzyme structure and function, a complete structure of CndH is needed.

3.4 Aliphatic halogenases

There are a few halogenases that have been reported to be capable of halogenating aliphatic substrates. Only one structure of the enzymes in this group has been solved, which is CmlS from *Streptomyces venezuelae* [9]. CmlS is involved in chloramphenicol biosynthesis. The FAD binding domain of

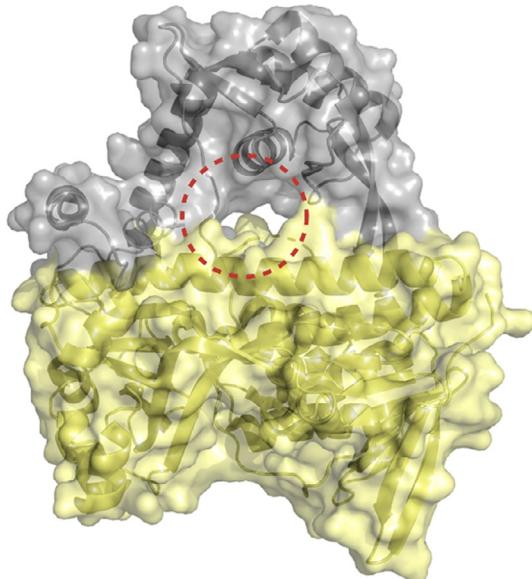


Fig. 11 Structure of CmlS. FAD binding module (yellow) and substrate binding module (gray). Red circle indicates a tunnel.

CmlS is similar to that of other FDHs, but the C-terminus is unique. The C-terminus of CmlS forms a winged-helix over the FAD binding module which creates a tunnel leading to the halogenation site (Fig. 11). Surprisingly, the structure of CmlS reveals a covalently bound FAD [9]. This feature has not been found in others FDHs discovered to date.

The structure of CmlS reveals that the C-terminal tail blocks the access of substrate to the active site. Therefore, catalysis is proposed to involve conformational dynamics. Although chloramphenicol biosynthesis involves an aminoacyl intermediate bound to a carrier protein (ClmP), limited access of the substrate binding site of CmlS suggests that CmlS acts on a free substrate containing an acyl group. However, it is still a major question which form of acyl is used. CmlS lacks the conserved Glu which is found in tryptophan halogenases [9]. There are water molecules observed in the tunnel forming hydrogen-bond networks beginning at Glu44 and running to Lys71 [9]. This feature is proposed to function as proton transfer machinery as seen in *p*-hydroxybenzoate hydroxylase (PHBH) [38]. This Glu44 is strictly conserved in FDHs. Mutation of this residue is needed to clearly elucidate its functional role.

As the substrate of this group is significantly different from other FDHs, information is needed to explain how this group of enzymes recognize aliphatic substrates. The unique features of the CmlS structure including the covalently bound FAD, the lack of the catalytic Glu conserved in tryptophan halogenases and the chain of water molecules makes the catalytic mechanism of this enzyme intriguing, leaving much to be investigated.

3.5 Halogenase utilizes diverse substrates

Recently, an FDH that is capable of regioselective halogenation of versatile substrates has been discovered using bioinformatics-based approaches [3]. Genome mining of halogenases using a previously unnoticed motif (Fx.Px.Sx.G) reveals a new FDH named VirX1. VirX1 is an FDH from an oceanic virus cyanophage [3]. It shows 29% sequence similarity to PrnA [3,12]. VirX1 can use a broad range of halides including Cl^- , Br^- , and I^- [3]. VirX1 shows *in vitro* iodination activity toward 6-azaindole [3]. VirX1 can catalyze halogenation of sterically and electronically diverse substrates that are less activated [3]. Interestingly, VirX1 iodinates at chemically less reactive positions in some substrates (Fig. 12).

The apo structure of VirX1 shows a similar flavin binding site to other FDHs (Fig. 3). The positions of Lys79 and Glu358 are similar to those of PrnA. The residues K79, F353, S359 and P356 were identified as important residues in PrnA because the variants K79A, K79R, F353A, S359A and P356A were completely inactive [3]. Although the structure of VirX1 is similar to PrnA, VirX1 has a higher degree of substrate flexibility and halide utilization. A key difference between VirX1 and PrnA is that the α -helical lid (Thr435–Trp455) in PrnA is absent in VirX1 as indicated in black circle of Fig. 13A. The absence of an α -helical lid in VirX1 results in a wider

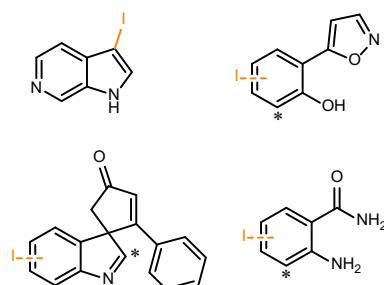


Fig. 12 Iodinated products of VirX1. The asterisk symbol indicates the chemically favorable position [3].

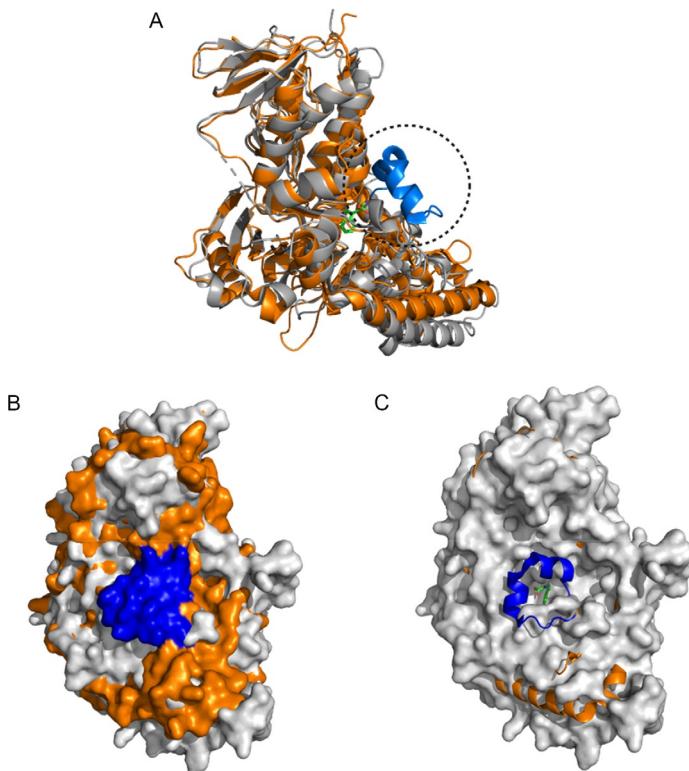
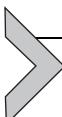


Fig. 13 Structural comparison of VirX1 (PDB id: [6QGM](#)) with PrnA (PDB id: [2AQJ](#)). The α -helical (blue) loop in PrnA (orange) is absent in VirX1 (gray) (A). Surface representation of VirX1 aligned with PrnA (B). Alignment of the substrate pocket of VirX1 with that of PrnA and bound Trp substrate (green stick) (C).

substrate pocket (Fig. 13B and C). Therefore, the active site of VirX1 can be accessed by a broader range of substrates and the enzyme can also accommodate I^- which has a large van der Waals radius.

The current finding implies that there are many interesting FDHs which remain to be discovered. Structures of VirX1 in complex with halides, substrates and products are needed to explain the unique regioselectivity of VirX1 and its ability to use less chemically reactive positions. As the native VirX1 has the ability to catalyze halogenation of a wide variety of substrates, the enzyme should be improved by enzyme engineering to further develop VirX1 for catalyzing more challenging reactions.



4. Mechanisms

To date, only one transient kinetics investigation of FDHs (RebH) has been reported. Studies of RebH by the Walsh group using stopped-flow spectrophotometry indicate that the RebH-bound FADH⁻ reacts with molecular oxygen (O₂) to form a C4a-hydroperoxy flavin intermediate (C4aOOH-FAD). This intermediate can react with halide ions (Cl⁻ or Br⁻) to form hypohalous acids (HOX) and presumably C4a-hydroxy flavin intermediate (C4aOH-FAD) [28]. The report showed that addition of Trp does not affect the flavin reaction, implying that halogenation occurs after the reaction of reduced flavin with oxygen. Structural studies of tryptophan halogenases reveal that FDHs have two separated active sites and the FAD and Trp binding sites are separated by about 10 Å [12]. The combined knowledge of mechanistic and structural data indicates that HOX should pass through a tunnel connecting the FAD binding site to the Trp binding site to react with the substrate [12,14,28].

The functional role of the conserved Lys which has been proposed to be a key catalytic residue in the FDH reaction is still under debate. A few models for explaining the mechanistic role of Lys with HOX are possible. The Walsh group proposed that HOX reacts with Lys to form a chloramine intermediate (Fig. 14) [14]. On the other hand, the Van Pee group proposed that HOX can form hydrogen bonding with Lys instead (Fig. 14, inset) [12]. All these proposals are based on mutation results, indicating that the change of Lys to Ala causes the enzyme to be inactive [12]. In addition, the conserved Glu in the Trp binding site is also proposed to be important for catalysis because mutation of Glu to Asp or Gln causes a decrease in activity [11,12]. It was proposed that the interaction of Glu with HOX or a chloramine derivative may increase the electrophilicity of Cl⁻ such that it can easily react with Trp *via* electrophilic aromatic substitution [11].

The importance of catalytic Lys79 and Glu346 residues in PrnA was explored using quantum mechanical/molecular mechanical (QM/MM) analysis, employing density functional theory (DFT) for QM region [39]. A hydrogen bond between the hydrogen of Lys79 and oxygen of HOCl was observed in the optimized reactant complex structure (Trp, HOCl, Lys79^{PrnA} and Glu346^{PrnA}). This result supports the proposed function of the catalytic Lys [12]. Calculation results indicate that a positive charge was transferred from Lys to Glu346, indicating that Glu346 deprotonates

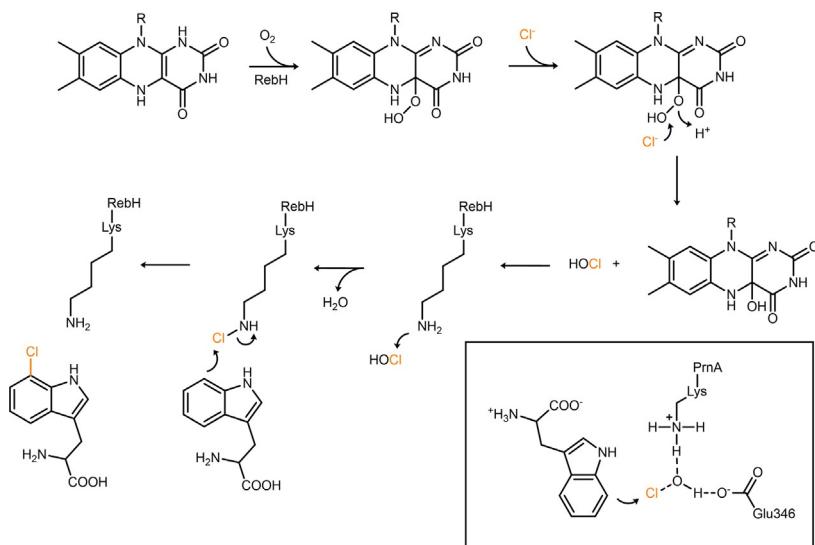


Fig. 14 Proposed mechanisms of tryptophan halogenases. Formation of a reactive HOCl by interaction of C4aOOH-FAD with Cl⁻. HOCl may react with Lys to form a chloramine intermediate (proposed by Walsh's group). A different halogenation mechanism has been proposed by Van Pee's group (inset).

the Wheland intermediate (Fig. 8C). Calculations from initial halogenation, Wheland intermediate formation and Wheland intermediate deprotonation revealed that the Wheland intermediate deprotonation is the rate-limiting step of the halogenation step. The results of this calculation are in greater support with the mechanism proposed by the Van Pee group.

Besides tryptophan halogenases, the mechanism of MalA' was investigated using MD simulations and QM by employing DFT calculations [23]. Site-directed mutagenesis was done to probe functional roles of interested residues. MalA' catalyzes halogenation at C8 and C9 of the premalbrancheamide substrate (Fig. 15). A catalytic residue, Lys108, is present in the active site of MalA'. MD simulations showed that a carboxylic group of Glu494 interacts with an indole ring (N-H) of premalbrancheamide, positioning the substrate to be close to the catalytic Lys108. MD simulations reveal key interactions of Lys108 with a backbone carbonyl group of Asp109, positioning Lys108 to be close to C8 and C9 positions of the substrate. Mutation of Asp109Ala resulted in lower activity. MD simulations of this mutant showed that Lys108 is not close to the C8 or C9 positions as seen in the wild-type enzyme, indicating that Asp109 contributes to regioselective control. DFT calculations using methyl chloramine

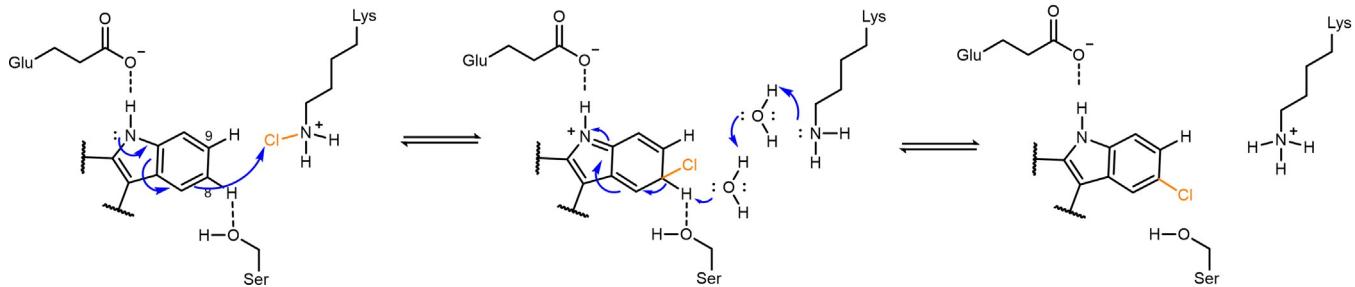
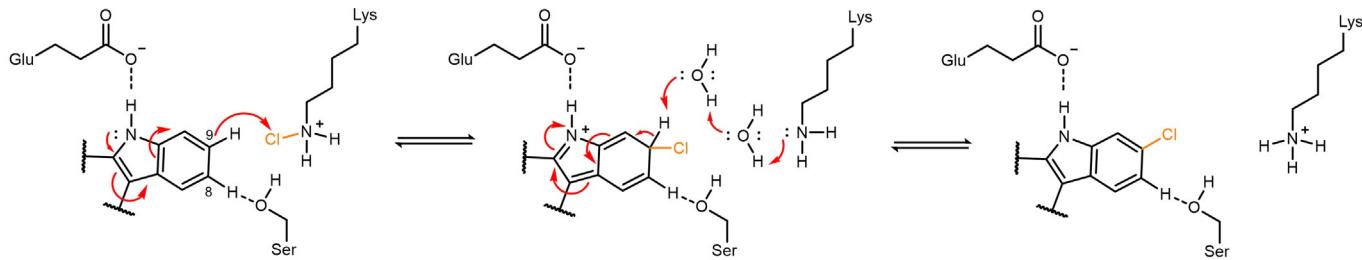
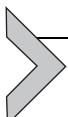


Fig. 15 Proposed mechanisms of MalA' [23]. Chlorination of premalbraneamide at the C9 position (A). Chlorination of premalbraneamide at the C8 position (B).

and indole rings as a starting model suggest that an intrinsic rate-limiting step should be an initial step of halogenation. An energy barrier is reduced when methanol (representative of Ser129) or water molecules are included in calculation models. In addition, MD simulations along the mechanism proposed by Walsh's group starting from the chloramine intermediate showed that a distance between the methanol side chain (O) with C8-H is close enough to form a hydrogen bond (Fig. 15B). This interaction may accelerate the rate of the halogenation step by increasing the nucleophilicity of C8 and C9 of the substrate. Since Asp109 is not in close proximity with the halogenated sites, water molecules and Lys108 are proposed to deprotonate the Wheland intermediate (Fig. 15). To elucidate and understand more of the reaction of MalA', transient kinetic experiments are still needed.



5. Improvement of the catalytic properties of FDHs using structure-guided mutagenesis

Targeted mutagenesis has been used to probe the mechanism of and alter the catalytic properties of several FDHs. The first FDH to be structurally characterized was PrnA [11,12]. The mutational studies provide insights into the mechanism and regioselectivity of flavin-dependent halogenases. In this section, targeted mutations to improve catalytic properties and alter substrate specificity based on the structure of FDHs are discussed.

5.1 Substrate scope expansion

Natural product derivatization is a key strategy for drug discovery and development. O'Connor and coworkers used site-directed mutagenesis to change RebH specificity from L-tryptophan to tryptamine, a direct precursor to many alkaloid natural products, including approximately 3000 monoterpene indole alkaloids [29]. The researchers created a series of 17 mutations based on 6 residues aligning specifically in the tryptophan-binding site. The Y455W mutant is the only variant that can chlorinate tryptamine. It was found that in the presence of equivalent amount of native substrate, the variant can chlorinate tryptamine with a 30-fold higher preference over tryptophan (Fig. 16). It was postulated that the introduction of steric hindrance in Y455W redesigns a binding site to prefer tryptamine in which the carboxylic acid moiety is absent relative to tryptophan.

Structure-guided mutagenesis is a straightforward way to improve the catalytic efficiency of enzymes in general. However, it requires the

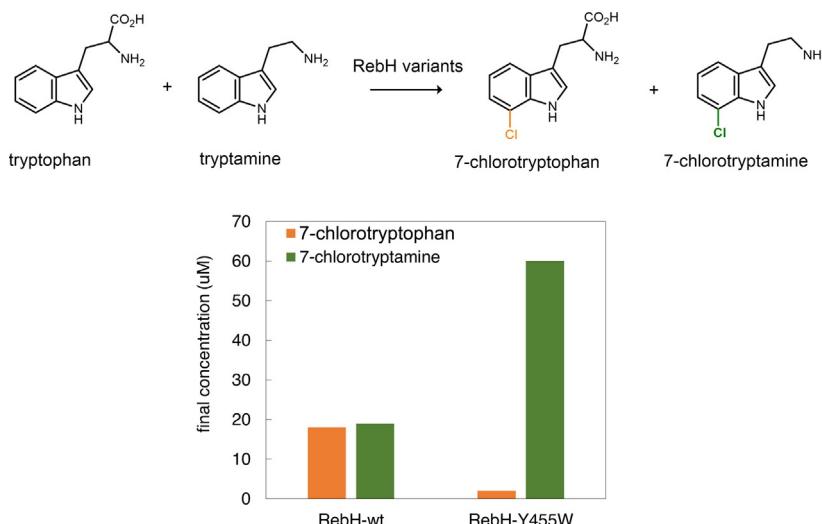


Fig. 16 Competition assays of RebH-wildtype (RebH-wt) and RebH-Y455W with a mixture of tryptophan and tryptamine (1:1).

availability of crystal structures of the enzymes. Alternatively, directed evolution can be applied to explore substrate specificity or expansion without knowing the enzyme structures. Error-prone PCR was frequently used to generate a large library of variants for high-throughput screening (HTS). The Lewis group reported directed evolution of RebH and demonstrated that two variants generated, RebH-G112S/N470S and RebH-A442V have a larger binding site than the wild-type enzyme and can accept bulky groups distal to the indole moiety of these unnatural substrates [30]. This random mutagenesis combined with HTS was also applied to RadH, a phenolic halogenase to explore its catalytic efficiency [31]. The selected variants showed improved activity toward coumarin molecules and the reaction was integrated into the biosynthetic pathway. It can be seen that the key value of random mutagenesis is discovering variants that are difficult to predict. However, to gain insights into the role of each functional group of the residues found to have an effect in improving catalytic efficiency, further efforts in protein crystallography or further site-selective mutations are required.

5.2 Site-selective mutagenesis

To construct FDHs with improved regioselectivity, the approach of using structure-guided mutagenesis is more straightforward. Most FDHs have

been characterized to be regioselective toward one substrate [12,40,41]. However, some FDHs can perform multiple-site halogenation with their native small molecule substrates. Bmp2, a pyrrole halogenase identified from marine bacteria can incorporate a total of four bromine atoms into a single pyrrole molecule [26]. The Moore group used structure-guided mutagenesis based on the proposed substrate-binding pocket of the previously reported structure of Mpy16, an analog FDH participating in the dibromination of pyrrole to reduce the degree of halogenation. The resulting triple variant, Bmp2-TM (Y302S/F306V/A345W) showed a significant reduction in the ability to halogenate from tetra- to monobromination (Fig. 17A).

Despite the high regioselectivity of FDHs toward their native substrates, when catalyzing the reaction with non-natural substrates, this property is found to be diminished. The Lewis group focused on engineering a RebH and explored its halogenation of tryptamine as an alkaloid precursor. The products were found to be a mixture of three different mono-chlorinated tryptamines. The group used a total of six rounds of mutation, consisting of site-saturated and site-directed mutagenesis for creating three selective variants with respective regioselective chlorination (Fig. 17B) [39a].

Improvement in substrate scope and regioselectivity are key properties that will enable the effective use of FDHs biocatalysts. Other improvements of catalytic properties, particularly, the turnover rate of the enzyme should be further optimized. Such improvement will be difficult to achieve by random-mutagenesis approaches, and can only be achieved through rational-based approaches which leverage the knowledge of the enzyme structure and mechanism.

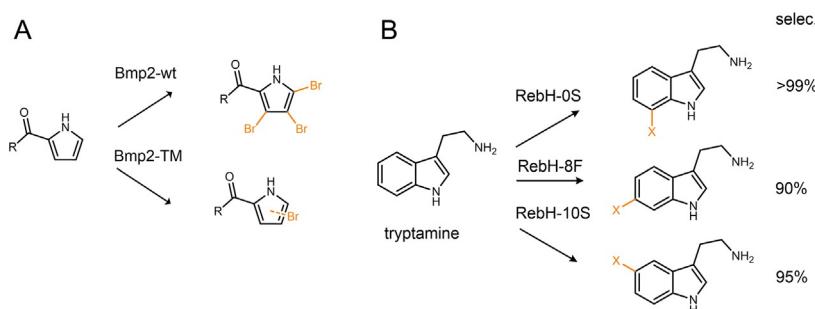


Fig. 17 Regioselectivity control of FDHs through structure-guided mutagenesis. Bmp2-TM generates monobrominated products (A). RebH variants with different selectivity determined by NMR analysis of purified chlorinated tryptamine isomers (B).



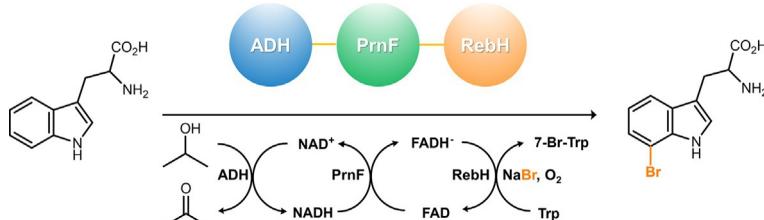
6. Applications of FDHs

In recent decades, several FDHs from both bacterial and fungal biosynthetic pathways have been characterized and identified for their roles in natural product biosynthesis [12,40,41]. These enzymes are involved in the biosynthesis of halogenated natural products as secondary metabolites. These indicate their potential to be applied as biocatalysts for biocatalysis under green and mild conditions. Therefore, engineering of FDHs to broaden their potential is key to achieving their successful utility in these synthetic applications.

6.1 Scale-up biocatalytic halogenation

Normally, aromatic halogenations require harsh reaction conditions and have poor regioselectivity. Enzymatic-catalyzed regioselective halogenation offers a promising alternative green and efficient method for aromatic halogenation for improved large-scale production of halogenated compounds.

In 2013, the Lewis group demonstrated a successful preparative halogenation system which enabled chlorination of 100 mg of tryptophan using crude cell lysate from bacteria expressing RebH and a cofactor-regenerating system as the catalyst [42]. However, the conversion yield was much lower than from the reaction using purified enzymes. The preparative efficiency of aromatic halogenation was further optimized by Sewald's group using a developed strategy to immobilize halogenase and auxiliary enzymes by precipitation and cross-linking simultaneously to obtain one solid biocatalyst, so-called combiCLEAs [43]. In this method, the RebH halogenase was immobilized together with its cofactors and regenerating enzymes (flavin reductase, PrnF and alcohol dehydrogenase, ADH) through glutaraldehyde linkage ([Scheme 1](#)). The RebH combiCLEAs system can be recycled at least



Scheme 1 Preparative bromination of L-tryptophan by biocatalysts composed of RebH cross-linked with flavin reductase (PrnF) and alcohol dehydrogenase (ADH).

10 times in a batchwise manner with an average conversion yield of 81%. Using this approach, a complete bromination of L-tryptophan was achieved after 8 days, and 1.81 g of L-7-bromotryptophan was obtained in a salt form, making it the first report of gram-scale biosynthesis of halogenated compounds.

6.2 Engineered biosynthetic pathway

In the field of drug discovery and bioactive compound exploration, generation of natural product analogs is important for improvement of the bioactivity and physicochemical properties of the compounds. However, natural product biosynthesis is limited by the availability of precursor compounds that are usable substrates, thus limiting diversification of the products.

Heterologous expression of halogenases has been developed to synthesize new analogs of available drugs. In 2010, the Goss group successfully incorporated PrnA originally from pyrrolnitrin biosynthesis into a pacidamycin-producing bacterium (*Streptomyces coeruleorubidus*) to generate chlorinated pacidamycins (Fig. 18) [44]. The isolated yields of chloropacidamycin were approximately 1 mg per liter culture with concomitant production of non-halogenated pacidamycins.

Halogenase genes have also been introduced into a medicinal plant metabolic pathway to produce novel plant products. Two of FDHs, PyrH and RebH from soil bacteria, which catalyze different regioselective halogenation reactions were incorporated into an alkaloid-producing plant Madagascar periwinkle (*Catharanthus roseus*), resulting in the production of a series of halogenated monoterpene indole alkaloids [45]. However, when a gene for a halogenase was introduced into the pathway, the decarboxylation of the respective 7-chlorotryptophan was found to be a bottleneck in the engineered pathway, resulting in a 15-fold decrease in the yield of total natural alkaloids produced compared to the native plant

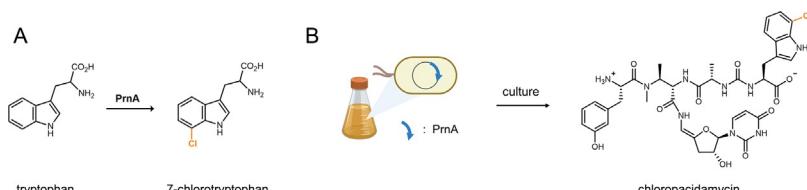


Fig. 18 Chloropacidamycin biosynthesis. 7-Chlorination of tryptophan catalyzed by PrnA (A). *In vivo* biosynthesis of chloropacidamycin by *S. coeruleorubidus* containing the PrnA gene (B).

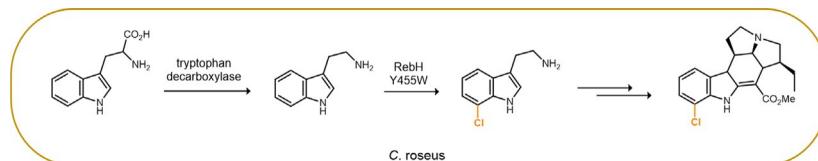


Fig. 19 Production of chlorinated monoterpenoid indole alkaloid by cultures of *C. roseus* containing RebH-Y455W.

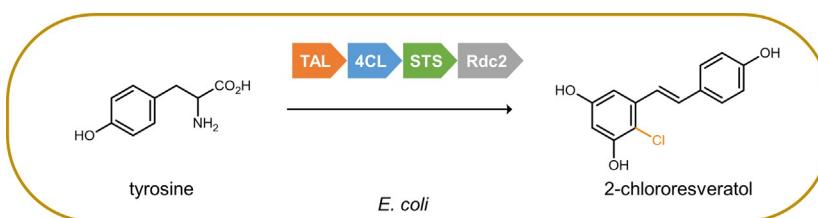


Fig. 20 A reconstituted resveratrol biosynthetic pathway in *E. coli* for production of 2-chlororesveratrol. Rdc2 was co-expressed with tyrosine ammonia-lyase (TAL), 4-coumarate:CoA ligase (4CL), and stilbene synthase (STS).

tissue culture. Later on, the engineered RebH-Y455W mentioned in Section 6.1 was reconstituted into this pathway to overcome the metabolic bottleneck (Fig. 19) [29]. However, the total production of alkaloid could not be calculated accurately because the final product was only slowly produced *in vitro*. Nevertheless, this *de novo* biosynthesis of unnatural halogenated alkaloids demonstrated a successful metabolic engineering approach to overcome the bottleneck in the reconstituted natural pathway using structure-guided mutagenesis.

Further work was expanded to generate novel molecules. As halogenated compounds account for a variety of drugs and several more currently under development, the use of halogenases to modify currently available drugs will expand the library for bioactivity screening. Halogenated resveratrol was generated by co-expression of the phenolic flavin-dependent halogenase (Rdc2) gene and the resveratrol biosynthetic gene cluster in *E. coli* (Fig. 20). The obtained product was then identified as 2-chlorinated resveratrol which is a new drug analog.

6.3 Late-stage diversification of halogenated compounds

The site-specific introduction of halogen atoms into complex natural products provides a convenient moiety for further functionalization and

diversification. Cross-coupling reactions that are of interest can introduce regioselective and regiodivergent functionalizations.

The cross-coupling application of biosynthetic halogenation can be achieved by carrying out a subsequent Pd-catalyzed Suzuki-Miyaura cross-coupling (SMC) reaction following FDH halogenation (Fig. 21). A number of unnatural products could be obtained by this approach [44]. An engineered variant of rebeccamycin halogenase (RebH) was chosen to generate regioselective-halogenated tryptoline which are bioactive compounds [46]. The haloaromatic compounds were subsequently subjected to cross-coupling reactions using a Pd catalyst. The coupling reaction could install aryl, amine, and ether substituents at the halogenation site. This chemoenzymatic cross-coupling reaction demonstrates the advantage of having specific halogenation at the 7- and 6-positions of indoles by allowing the halogenation to proceed efficiently with more than 90% conversion in crude extract. However, the respective Pd-catalyst coupled reactions resulted in final isolated yields of 33–99%. The decrease in final yield was due to the subsequent steps for combination between two distinct enzymatic and chemical catalysis.

The above combined cascades of chemo- and enzymatic reactions cannot be carried out in one pot because the conditions for the two types of catalysts are not compatible. Thus, the halogenase-SMC cascade was further developed using polydimethylsiloxane membranes (PDMS) as a compartment for separating the enzyme from the Pd catalyst (Fig. 22) [47]. The compartmentalization allows nonpolar arylhalide to diffuse freely between the compartments, which later is then functionalized *via* the SMC reaction at the chemocatalyst phase. Taking advantage of different regioselective FDHs, the developed method was capable of being expanded to a series of regioselective and regiodivergent arylation, vinylation, and heteroarylation reactions done in one pot.

Therefore, the stability of FDHs can be successfully improved and their compatibility with chemical reactions of interest has been demonstrated. Introduction of stability-enhancing functions can generate a recyclable halogenase and the compatible combination of two different types of catalysts has expanded the possibilities for biosynthesis. This progress will contribute to the advancement of chemical and pharmaceutical industries and will enable greener production of these valuable compounds. However, challenges remaining for utilization of FDHs in industries are their low catalytic efficiencies as well as their requirement for cofactor-regenerating systems. Further development to overcome these drawbacks are required for real-life implementation of FDHs in large scale and commercial synthesis applications.

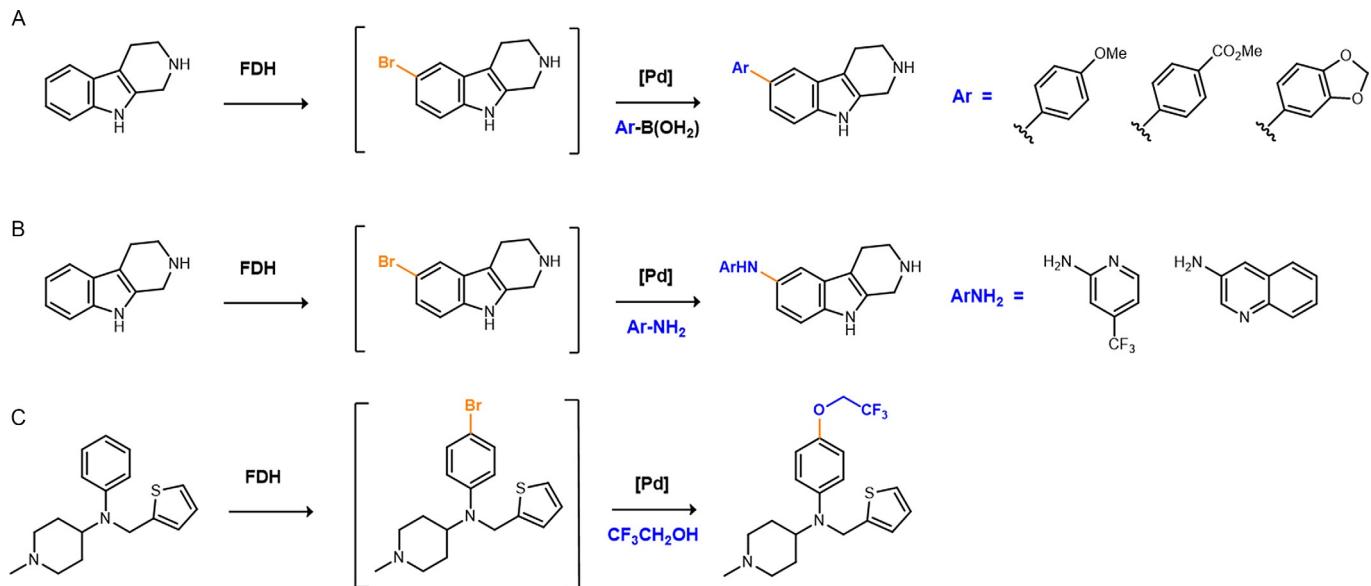


Fig. 21 Regioselective biohalogenation and subsequent Suzuki-Miyaura cross-coupling reactions. C—C bond arylation (A). Amination (B). Alkoxylation (C).

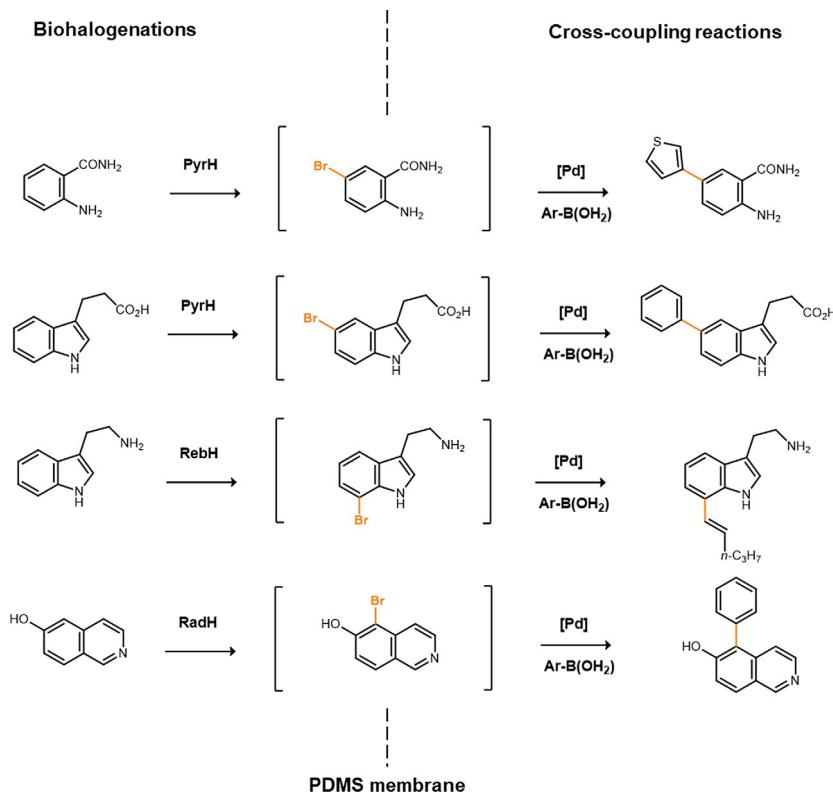


Fig. 22 Different regioselective biohalogenations and subsequent Suzuki-Miyaura cross-coupling reactions with PDMS compartmentalization.

7. Future perspectives

FDHs have the ability to halogenate a broad spectrum of substrates ranging from aliphatic to complex substrates. The enzymes have great potential in biocatalytic applications, primarily in making the halogenation more environmental friendly and green. Several recent investigations were involved in identification of new FDHs and on enzyme engineering. Several crystal structures of FDHs are now available. However, some key structures such as the ternary complex of the enzyme with reduced FAD and substrate is not currently available.

To date, there is only one transient kinetics report of RebH [14,28]. In-depth kinetic investigation should be valuable to investigate some intriguing FDHs such as CmlS which has a FAD covalently bound in the

structure and VirX1 which can use a wide variety of substrates including catalyzing iodination. It is interesting to learn about reduction and oxygenation cycles and steps involved in the catalysis of these systems.

The issue regarding the role of the conserved Lys is still debatable as to whether it would form the chloramine intermediate or the residue only closely interacts with HOX. Up to now, the site-directed mutagenesis studies of the catalytic Lys was only investigated in two enzymatic system and both reports only changed the catalytic Lys to Ala [2,3,12,22] and Lys to Arg [3]. It would be interesting to learn more about the role of this Lys by changing it to other residues.

Although FDHs are recognized as useful enzymes for synthetic applications, their real applications in industries are still limited. Comprehensive understanding of the structures and mechanisms of FDHs should be useful for future work on enzyme engineering to expand their substrate scope or to gain properties required such as improved turnovers and thermostability.

References

- [1] J. Latham, E. Brandenburger, S.A. Shepherd, B.R.K. Menon, J. Micklefield, Development of halogenase enzymes for use in synthesis, *Chem. Rev.* 118 (2018) 232–269.
- [2] S. Mori, A.H. Pang, N. Thamban Chandrika, S. Garneau-Tsodikova, O.V. Tsodikov, Unusual substrate and halide versatility of phenolic halogenase PltM, *Nat. Commun.* 10 (2019) 1255.
- [3] D.S. Gkotsi, H. Ludewig, S.V. Sharma, J.A. Connolly, J. Dhaliwal, Y. Wang, W.P. Unsworth, R.J.K. Taylor, M.M.W. McLachlan, S. Shanahan, J.H. Naismith, R.J.M. Goss, A marine viral halogenase that iodinates diverse substrates, *Nat. Chem.* 11 (2019) 1091–1097.
- [4] P.C. Schmartz, K. Zerbe, K. Abou-Hadeed, J.A. Robinson, Bis-chlorination of a hexapeptide-PCP conjugate by the halogenase involved in vancomycin biosynthesis, *Org. Biomol. Chem.* 12 (2014) 5574–5577.
- [5] C. Sánchez, I.A. Butovich, A.F. Braña, J. Rohr, C. Méndez, J.A. Salas, The biosynthetic gene cluster for the antitumor rebeccamycin: characterization and generation of indolocarbazole derivatives, *Chem. Biol.* 9 (2002) 519–531.
- [6] L. Xu, T. Han, M. Ge, L. Zhu, X. Qian, Discovery of the new plant growth-regulating compound LYXLF2 based on manipulating the halogenase in *Amycolatopsis orientalis*, *Curr. Microbiol.* 73 (2016) 335–340.
- [7] S. Keith, A.E.-H. Gamal, Regioselective control of electrophilic aromatic substitution reactions, *Curr. Org. Synth.* 1 (2004) 253–274.
- [8] V. Agarwal, Z.D. Miles, J.M. Winter, A.S. Eustáquio, A.A. El Gamal, B.S. Moore, Enzymatic halogenation and dehalogenation reactions: pervasive and mechanistically diverse, *Chem. Rev.* 117 (2017) 5619–5674.
- [9] K. Podzelinska, R. Latimer, A. Bhattacharya, L.C. Vining, D.L. Zechel, Z. Jia, Chloramphenicol biosynthesis: the structure of CmlS, a flavin-dependent halogenase showing a covalent flavin-aspartate bond, *J. Mol. Biol.* 397 (2010) 316–331.
- [10] S.A. Shepherd, C. Karthikeyan, J. Latham, A.-W. Struck, M.L. Thompson, B.R.K. Menon, M.Q. Styles, C. Levy, D. Leys, J. Micklefield, Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes, *Chem. Sci.* 6 (2015) 3454–3460.

- [11] S. Flecks, E.P. Patallo, X. Zhu, A.J. Ernyei, G. Seifert, A. Schneider, C. Dong, J.H. Naismith, K.-H. van Pee, New insights into the mechanism of enzymatic chlorination of tryptophan, *Angew. Chem. Int. Ed. Engl.* 47 (2008) 9533–9536.
- [12] C. Dong, S. Flecks, S. Universcht, C. Haupt, K.-H. van Pee, J.H. Naismith, Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination, *Science* 309 (2005) 2216–2219.
- [13] E. Bitto, Y. Huang, C.A. Bingman, S. Singh, J.S. Thorson, G.N. Phillips Jr., The structure of flavin-dependent tryptophan 7-halogenase RebH, *Proteins: Struct., Funct., Bioinf.* 70 (2008) 289–293.
- [14] E. Yeh, L.C. Blasiak, A. Koglin, C.L. Drennan, C.T. Walsh, Chlorination by a long-lived intermediate in the mechanism of flavin-dependent halogenases, *Biochemistry* 46 (2007) 1284–1292.
- [15] C.B. Poor, M.C. Andorfer, J.C. Lewis, Improving the stability and catalyst lifetime of the halogenase RebH by directed evolution, *Chembiochem* 15 (2014) 1286–1289.
- [16] X. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K.-H. van Pee, J.H. Naismith, Structural insights into regioselectivity in the enzymatic chlorination of tryptophan, *J. Mol. Biol.* 391 (2009) 74–85.
- [17] A.-C. Moritzer, H. Minges, T. Prior, M. Frese, N. Sewald, H.H. Niemann, Structure-based switch of regioselectivity in the flavin-dependent tryptophan 6-halogenase Thal, *J. Biol. Chem.* 294 (2019) 2529–2542.
- [18] A.-C. Moritzer, H.H. Niemann, Binding of FAD and tryptophan to the tryptophan 6-halogenase Thal is negatively coupled, *Protein Sci.* 28 (2019) 2112–2118.
- [19] S.A. Shepherd, B.R.K. Menon, H. Fisk, A.-W. Struck, C. Levy, D. Leys, J. Micklefield, A structure-guided switch in the regioselectivity of a tryptophan halogenase, *Chembiochem* 17 (2016) 821–824.
- [20] H. Luhavaya, R. Sigrist, J.R. Chekan, S.M.K. McKinnie, B.S. Moore, Biosynthesis of l-4-chlorokynurenone, an antidepressant prodrug and a non-proteinogenic amino acid found in lipopeptide antibiotics, *Angew. Chem. Int. Ed.* 58 (2019) 8394–8399.
- [21] K. Lingkon, J.J. Bellizzi Iii, Structure and activity of the thermophilic tryptophan-6 halogenase BorH, *ChemBioChem* 21 (2019) 1121–1128.
- [22] P.R. Neubauer, C. Widmann, D. Wibberg, L. Schröder, M. Frese, T. Kottke, J. Kalinowski, H.H. Niemann, N. Sewald, A flavin-dependent halogenase from metagenomic analysis prefers bromination over chlorination, *PLoS One* 13 (2018) e0196797.
- [23] A.E. Fraley, M. Garcia-Borràs, A. Tripathi, D. Khare, E.V. Mercado-Marin, H. Tran, Q. Dan, G.P. Webb, K.R. Watts, P. Crews, R. Sarpong, R.M. Williams, J.L. Smith, K.N. Houk, D.H. Sherman, Function and structure of MalA/MalA', iterative halogenases for late-stage C–H functionalization of indole alkaloids, *J. Am. Chem. Soc.* 139 (2017) 12060–12068.
- [24] M.A. Ortega, D.P. Cogan, S. Mukherjee, N. Garg, B. Li, G.N. Thibodeaux, S.I. Maffioli, S. Donadio, M. Sosio, J. Escano, L. Smith, S.K. Nair, W.A. van der Donk, Two flavoenzymes catalyze the post-translational generation of 5-chlorotryptophan and 2-aminovinyl-cysteine during NAI-107 biosynthesis, *ACS Chem. Biol.* 12 (2017) 548–557.
- [25] A.H. Pang, S. Garneau-Tsodikova, O.V. Tsodikov, Crystal structure of halogenase PltA from the pyoluteorin biosynthetic pathway, *J. Struct. Biol.* 192 (2015) 349–357.
- [26] A. El Gamal, V. Agarwal, S. Dietel, I. Rahman, M.A. Schorn, J.M. Snead, G.V. Louie, K.E. Whalen, T.J. Mincer, J.P. Noel, V.J. Paul, B.S. Moore, Biosynthesis of coral settlement cue tetrabromopyrrole in marine bacteria by a uniquely adapted brominase–thioesterase enzyme pair, *Proc. Natl. Acad. Sci. U. S. A.* 113 (2016) 3797.
- [27] S. Buedenbender, S. Rachid, R. Müller, G.E. Schulz, Structure and action of the myxobacterial chondrochlore halogenase CndH: a new variant of FAD-dependent halogenases, *J. Mol. Biol.* 385 (2009) 520–530.

- [28] E. Yeh, L.J. Cole, E.W. Barr, J.M. Bollinger, D.P. Ballou, C.T. Walsh, Flavin redox chemistry precedes substrate chlorination during the reaction of the flavin-dependent halogenase RebH, *Biochemistry* 45 (2006) 7904–7912.
- [29] W.S. Glenn, E. Nims, S.E. O'Connor, Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor, *J. Am. Chem. Soc.* 133 (2011) 19346–19349.
- [30] J.T. Payne, C.B. Poor, J.C. Lewis, Directed evolution of RebH for site-selective halogenation of large biologically active molecules, *Angew. Chem. Int. Ed.* 54 (2015) 4226–4230.
- [31] B.R.K. Menon, E. Brandenburger, H.H. Sharif, U. Klemstein, S.A. Shepherd, M.F. Greaney, J. Micklefield, RadH: a versatile halogenase for integration into synthetic pathways, *Angew. Chem. Int. Ed. Engl.* 56 (2017) 11841–11845.
- [32] P. Chenprakhon, T. Wongnate, P. Chaiyen, Monooxygenation of aromatic compounds by flavin-dependent monooxygenases, *Protein Sci.* 28 (2019) 8–29.
- [33] E. Romero, J.R. Gómez Castellanos, G. Gadda, M.W. Fraaije, A. Mattevi, Same substrate, many reactions: oxygen activation in flavoenzymes, *Chem. Rev.* 118 (2018) 1742–1769.
- [34] W.J.H. van Berkel, N.M. Kamerbeek, M.W. Fraaije, Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts, *J. Biotechnol.* 124 (2006) 670–689.
- [35] V. Agarwal, A.A. El Gamal, K. Yamanaka, D. Poth, R.D. Kersten, M. Schorn, E.E. Allen, B.S. Moore, Biosynthesis of polybrominated aromatic organic compounds by marine bacteria, *Nat. Chem. Biol.* 10 (2014) 640–647.
- [36] T. Heine, W.J.H. van Berkel, G. Gassner, K.-H. van Pee, D. Tischler, Two-component FAD-dependent monooxygenases: current knowledge and biotechnological opportunities, *Biology (Basel)* 7 (2018) 42.
- [37] B.R.K. Menon, J. Latham, M.S. Dunstan, E. Brandenburger, U. Klemstein, D. Leys, C. Karthikeyan, M.F. Greaney, S.A. Shepherd, J. Micklefield, Structure and biocatalytic scope of thermophilic flavin-dependent halogenase and flavin reductase enzymes, *Org. Biomol. Chem.* 14 (2016) 9354–9361.
- [38] K.K. Frederick, D.P. Ballou, B.A. Palley, Protein dynamics control proton transfers to the substrate on the His72Asn mutant of p-hydroxybenzoate hydroxylase, *Biochemistry* 40 (2001) 3891–3899.
- [39] T.G. Karabencheva-Christova, J. Torras, A.J. Mulholland, A. Lodola, C.Z. Christov, Mechanistic insights into the reaction of chlorination of tryptophan catalyzed by tryptophan 7-halogenase, *Sci. Rep.* 7 (2017) 17395.
- [39a] M.C. Andorfer, H.J. Park, J. Vergara-Coll, J.C. Lewis, Directed evolution of RebH for catalyst-controlled halogenation of indole C–H bonds, *Chem. Sci.* 7 (6) (2016) 3720–3729. <https://doi.org/10.1039/c5sc04680g>.
- [40] C. Seibold, H. Schnerr, J. Rumpf, A. Kunzendorf, C. Hatscher, T. Wage, A.J. Ernyei, C. Dong, J.H. Naismith, K.-H. Van Pee, A flavin-dependent tryptophan 6-halogenase and its use in modification of pyrrolnitrin biosynthesis, *Biocatal. Biotransformation* 24 (2006) 401–408.
- [41] S. Zehner, A. Kotzsch, B. Bister, R.D. Süssmuth, C. Méndez, J.A. Salas, K.-H. van Pee, A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in *Streptomyces rugosporus* LL-42D005, *Chem. Biol.* 12 (2005) 445–452.
- [42] J.T. Payne, M.C. Andorfer, J.C. Lewis, Regioselective Arene halogenation using the FAD-dependent halogenase RebH, *Angew. Chem. Int. Ed.* 52 (2013) 5271–5274.
- [43] M. Frese, N. Sewald, Enzymatic halogenation of tryptophan on a gram scale, *Angew. Chem. Int. Ed.* 54 (2015) 298–301.
- [44] A.D. Roy, S. Grüs Chow, N. Cairns, R.J.M. Goss, Gene expression enabling synthetic diversification of natural products: chemogenetic generation of pacidamycin analogs, *J. Am. Chem. Soc.* 132 (2010) 12243–12245.

- [45] W. Runguphan, X. Qu, S.E. O'Connor, Integrating carbon–halogen bond formation into medicinal plant metabolism, *Nature* 468 (2010) 461–464.
- [46] L.J. Durak, J.T. Payne, J.C. Lewis, Late-stage diversification of biologically active molecules via chemoenzymatic C–H functionalization, *ACS Catal.* 6 (2016) 1451–1454.
- [47] J. Latham, J.-M. Henry, H.H. Sharif, B.R.K. Menon, S.A. Shepherd, M.F. Greaney, J. Micklefield, Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C–H activation, *Nat. Commun.* 7 (2016) 11873.