A fivefold parallelized biosynthetic process secures chlorination of *Armillaria mellea* (honey mushroom) toxins

Jonas Wick[a], Daniel Heine[b], Gerald Lackner[c], Mathias Misiek[a], James Tauber[a], Hans Jagusch[a], Christian Hertweck[b], Dirk Hoffmeister[a,#]

[a]Department of Pharmaceutical Microbiology at the Hans Knöll Institute, Friedrich-Schiller-Universität Jena, Beutenbergstrasse 11a, 07745 Jena, Germany

[b]Department of Biomolecular Chemistry, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute, Beutenbergstraße 11a, 07745 Jena, Germany

[c]Institute of Microbiology, Swiss Federal Institute of Technology (ETH), Vladimir-Prelog-Weg 1-5, 8093 Zürich, Switzerland.

[*]Corresponding author. Tel.: +49 3641 949851; fax: +49 3641 949852. E-mail address: dirk.hoffmeister@leibniz-hki.de

Running title: *Armillaria mellea* (honey mushroom) halogenases

Abstract

The basidiomycetous tree pathogen *Armillaria mellea* (honey mushroom) produces a large variety of structurally related antibiotically active and phytotoxic natural products, referred to as the melleolides. During their biosynthesis, some members of the melleolide family of compounds undergo monochlorination of the aromatic moiety, whose biochemical and genetic basis was not known previously. This first study on basidiomycete halogenases presents the biochemical *in vitro* characterization of five flavin-dependent *A. mellea* enzymes (ArmH1-ArmH5) that were heterologously produced in *Escherichia coli.* We demonstrate that all five enzymes transfer a single chlorine atom to the melleolide backbone. A fivefold secured biosynthetic step during natural product assembly is unprecedented. Typically, flavin-dependent halogenases are categorized into enzymes acting on free compounds as opposed to those requiring a carrier protein-bound acceptor substrate. The enzymes characterized in this study clearly turned over free substrates. Phylogenetic clades of halogenases suggest that all fungal enzymes share a common ancestor and reflect a clear divergence between ascomycetes and basidiomycetes.

Keywords

Armillaria, biosynthesis, halogenase, melleolide, secondary metabolism, orsellinic acid

Introduction

The basidiomycete genus *Armillaria* includes numerous species that are known as notorious butt and root rot agents [1]. They are globally distributed as hardwood or conifer pathogens [2] in managed and unmanaged forests and also damage fruit trees, and grapes. Therefore, the genus is best known for its economic burden. Despite being serious plant pathogens, *Armillaria* species also play a positive environmental role as they depolymerize lignocellulose and, therefore, help maintain the carbon flux in ecosystems.

A remarkable physiological feature of *Armillaria* species is the capacity to produce melleolide natural products (Figure 1) [3-7]. These secondary metabolites feature a unique molecular scaffold, composed of an orsellinic acid (2,4-dihydroxy-6-methylbenzoic acid, Figure 1) moiety esterified to a tricyclic sesquiterpene (protoilludane) alcohol. The melleolides are intriguing as they show two distinct structure-activity relationships for their cytotoxic and antifungal bioactivities [8]. Phytotoxic activities have also been established [9,10]. Further, the melleolides represent one of the largest fungal natural product families with more than 60 structural variants. This degree of variation stems from a permutational organization of the biosynthesis that combines hydroxylation at various positions of the sesquiterpene, oxidation of the primary alcohol at C-1 to an aldehyde or carboxy group, shift or reduction of the cyclohexene double bond, and methyl ether formation at O-5'. Regioselective chlorination at C-6' also contributes to the structural diversity, reflected by about 25 described chlorinated melleolides.

None of the enzymes that modify the melleolide scaffold has been discovered to date. However, the protoilludene synthase Pro1 [11] and the orsellinic acid synthase ArmB [12] that elaborate the melleolide core structure have been characterized and described. These enzymes and those hypothesized to catalyze the above modifications are encoded in a contiguous single copy cluster of genes (Figure 2), as evident by the published genomic sequence of *A. mellea* [13]. Intriguingly, the

gene cluster does not include any halogenase gene. Covalent attachment of a halogen atom represents a frequently found modification of microbial natural products. Currently, more than 4,000 mostly chlorinated or brominated compounds of biological origin are known [14]. These include potent bioactive compounds, such as the HSP90 inhibitor radiciol from *Chaetomium chiversii* and *Metacordyceps* (*Pochonia*) *chlamydosporia* [15,16] or the cytotoxic endiyne C-1027 [17] of *Streptomyces globisporus*. Even some clinically used antibiotics and anticancer drugs show halogenation, e.g., the antifungal agent griseofulvin of *Penicilllium aethiopicum* [18] and the anticancer drug calicheamicin of *Micromonospora echinospora* spp. *calichensis* [19] that carry chlorine and iodine atoms, respectively. The gene clusters for the above biosyntheses all encode one flavin-dependent halogenase.

Using melleolide F as a model representative of this class of compounds, we provide evidence that five actively transcribed genes outside the *A. mellea* melleolide biosynthesis gene cluster code for functional flavin-dependent halogenases (ArmH1-ArmH5) that catalyze transfer of a single chlorine atom to melleolides. This first study on basidiomycete halogenases therefore reveals an unprecedented case of a biosynthetic process that is secured by a fivefold redundancy. *In vitro* characterization using heterologously produced enzymes suggest that these halogenases act on free substrates, i.e., they do not depend on carrier protein-tethered acceptor molecules.

Material and Methods

General methods, strains, and chemicals - Chemicals, organic solvents, media, reagents, and antibiotics were purchased from Becton-Dickinson, Roth, Sigma-Aldrich, and VWR, except melleolide F, which was enriched to about 90% purity from *A. mellea* mycelium, following a published procedure [20]. Molecular biology methods followed the protocols supplied with kits, enzymes, and

reagents (Fermentas, New England Biolabs, Promega, Thermo, Zymo Research). *A. mellea* DSM3731 was routinely maintained on Malt Extract Peptone Agar (MEP), pH 5.6, at 24°C. Liquid cultures were grown in glucose minimal medium (GMM) [21] at 24°C and shaken at 180 rpm for 12 - 14 days, in the dark. Stationary cultures were grown in penicillin flasks in 1 L MEP broth. For bromine and iodine feeding, KCl was replaced by KBr and KI, respectively, in the GMM.

***In silico* sequence analysis -** The published genomic DNA-sequence of *A. mellea* [13] was searched using the protein sequences of the protoilludene cyclase Pro1 (GenBank accession: AGR34199), of the polyketide synthase ArmB (GenBank accession: AFL91703), and the halogenase ArmH1 (GenBank accession: AEM76785) and as queries [11,12,22] and using the BlastP function integrated in the Joint Genome Institute's Genome Portal [23].

Cloning of *armH1-armH5* cDNA - Total RNA was extracted from *A. mellea* DSM3731 mycelium under melleolide-producing conditions. Mycelia were filtered under vacuum, washed with dH_2O, and ground by mortar and pestle under liquid nitrogen before total RNA extraction using SV Total RNA Isolation System (Promega) following the manufacturers spin protocol. First strand synthesis of *armH1* and *armH2* was accomplished with oligonucleotide primer cHalR2 (Table 1) and ImProm-II reverse transcriptase, using described parameters [24]. Using 1 µL of the first strand synthesis reaction as template, the *armH1* cDNA was amplified by PCR using *Pfu* polymerase (94°C for 2 min, 40 cycles: 94°C for 40 s, 58°C for 30 s, 72°C for 4 min, final extension: 72°C for 10 min) and oligonucleotide primers cHalF1 and cHalR1 (Table 1). The PCR product was ligated into pGem-11Zf(-) via *Eco*RI and *Bam*HI to yield plasmid pMM1. After sequencing, a correct clone was used as a template for a second PCR under the following conditions: *Pfu*-Polymerase, primers: Hal1-28F and Hal1-28R, 25 cycles (94°C for 40 s, 55°C for 30 s, 72°C for 4 min, final extension: 72°C for 7 min). The resulting PCR product was ligated into the *Nde*I and *Bam*HI sites of vector pET28a, yielding plasmid pMM21. The *armH2* cDNA was procured as described for *armH1*, but using primer Hal2R2 for first

Applied and Environmental Microbiology

strand synthesis and oligonucleotide primers Hal2F1 and Hal2R1 for cDNA amplification (Table 1). The PCR product was ligated into the *Nde*I and *Bam*HI sites of pET28a to give pMM13.

For first strand synthesis of *armH3-armH5*, total RNA was primed with oligo $(dT)_{18}$ primers and reverse transcribed using RevertAid Reverse Transcriptase (Thermo). One µL of first-strand synthesized *armH3-armH5* cDNA was amplified in a Phusion DNA polymerase (NEB) PCR reaction using the recommended master mix from the manufacturer with a final $MgCl_2$ concentration of 2.0 mM. PCRs for *armH3-H5* were primed with oligonucleotide pairs oJT073/oHJ01, oJT040/oJT041 and oJT042/oJT043, respectively (Table 1) using the following thermal cycling parameters: initial denaturation at 98°C for 30 s; 30 cycles of 98°C for 10 s, 60°C (*armH3*), 70°C (*armH4-H5*) for 15 s, and 72°C for 2 min; and a final extension for 5 min at 72°C. Amplicons of *armH3-H5* were ligated initially to pJET1.2 using the CloneJet PCR cloning kit and subsequently cut and ligated into equally cut pET28b to produce T7-expression plasmids pHJ28 (into *Nde*I and *Bam*HI sites, *armH3*), pHJ14 (into *Nde*I and *Hin*dIII sites, *armH4*), and pHJ05 (into *Nde*I and *Eco*RI sites, *armH5*). The presence of insert integration was verified by DNA-sequencing, and all molecular cloning was done in *E. coli* XL-1 Blue, selected by carbenicillin (pJET1.2) or kanamycin (pET28b).

Cloning of the *E. coli* flavine reductase gene *fre* - The flavine reductase gene (*fre*) was amplified by PCR from genomic *E. coli* DNA using *Pfu* polymerase and primers FreN-F and FreN-R (Table 1) (30 cycles of 94°C for 40 s, 57°C for 30 s, 72°C for 4 min, final extension: 72°C for 5 min). The PCR product was cloned into the *Nde*I and *Bgl*II sites of vector pET15b to yield plasmid pMM14.

Heterologous protein production and purification - For production of ArmH1-ArmH5, *E. coli* KRX was individually transformed with pMM21 (to express *armH1*), pMM13 (*armH2*), pHJ28 (*armH3*), pHJ14

(*armH4*), or pHJ05 (*armH5*). Each transformed *E. coli* strain was grown in 2×YT medium (supplemented with kanamycin (50 µg/mL), dispensed in Erlenmeyer flasks (10 x 400 mL) at 37°C and 180 rpm to an optical density of OD_{600} = 0.4. After cooling to 16 °C, cells were grown to an OD_{600} = 0.7 and induced with L-rhamnose (0.1%, w/v). After an additional 18 h of shaking at 16 °C, cells were harvested by centrifugation, cooled to 4°C and lysed by sonication (3 × 10 s). After removal of cell debris by centrifugation the supernatants were incubated with 2 mL (each) of Ni-NTA resin for 1 h at 4°C. The resins were transferred onto gravity flow columns and washed with 100 mM phosphate buffer (pH = 7.4, 300 mM NaCl) containing increasing imidazole concentrations (10 mM, 20 mM). Elution was performed with 2.5 mL of 400 mM imidazole. The enzymes were rebuffered in 3.5 mL phosphate buffer (100 mM NaH_2PO_4, 100 mM NaCl, pH = 7.4) using PD-10 columns (GE Healthcare). For bromination and iodination assays, NaCl was substituted with KBr or KI, respectively. Enzymes were concentrated with a 50 kDa cutoff Amicon filter (Merck Millipore) and used for subsequent *in vitro* product formation assays. To produce flavin reductase, *E. coli* BL21 transformed with pMM14 was grown in liquid LB medium containing 50 µg/mL carbenicillin at 37°C and 180 rpm to an OD_{600} = 0.5 and then induced with IPTG (1 mM). Heterologous protein production proceeded over night at 16°C. Purification of the *N*-terminally hexahistidine-tagged *E. coli* flavin reductase was carried out as described above for ArmH enzymes.

***In vitro* halogenation assays -** Freshly prepared ArmH enzymes were separately incubated at 25°C over night in a total volume of 100 µL. The buffer was 100 mM NaH_2PO_4, 100 mM NaCl (or KBr or KI, respectively), pH = 7.4. The reaction also contained 2 mM NADH, 10 µM FAD, 2.4 units *E. coli* flavin reductase (Fre) to supply the respective halogenase with $FADH_2$, and the acceptor substrate (25 µM melleolide F). Fre activity was determined following a published procedure [25]. Reactions with heat-inactivated enzyme (95°C for 10 min) were run in parallel as negative controls, in particular to

exclude non-enzymatic chlorination by hypochlorite that might have formed spontaneously in the assays. To prepare the reactions for chromatography, they were extracted twice with an equal volume of ethyl acetate + 1% acetic acid (v/v), dried under reduced pressure, and redissolved in 100 µL of methanol.

Liquid chromatography and mass spectrometry - *In vitro* reactions were analyzed on an Agilent 1200 HPLC equipped with a Zorbax Eclipse XDB-C_{18} column (150 × 4.6 mm, 3.5 µm particle size). Solvent A was 0.1% trifluoroacetate in H_2O, solvent B was acetonitrile. For melleolide analysis a linear gradient from 50% B to 100% B within 16 min at a flow of 0.7 mL/min was applied, followed by a wash step (100% B) for 5 min. Routine mass spectrometry was performed on an Agilent 1260 chromatograph with a Zorbax Eclipse XDB-C_{18} column (150 × 4.6 mm, 3.5 µm particle size), coupled to an Agilent 6130 mass detector using electrospray ionization in positive and negative mode and applying the gradient described above. Diode array detection was from λ = 200-500 nm, and chromatograms were extracted at λ = 254 nm. High-resolution electrospray ionization mass spectrometry (HR-ESIMS) was performed on a Thermo Accela liquid chromatograph equipped with a Betasil C_{18} column (150 × 2.1 mm, 3 µm particle size) and fitted to an Exactive Orbitrap mass spectrometer (ThermoFisher). Solvent A was 0.1% formic acid in H_2O, solvent B was 0.1% formic acid in acetonitrile. For sample analysis, the column was washed with 95% H_2O for 1 min followed by a linear gradient (flow rate: 0.2 mL/min) from 5% B to 98% B within in 15 min, and held at this relation for 3 min.

***In silico* sequence and phylogenetic analysis -** To identify exon/intron-junctions *in silico* we used FGENESH (Softberry, Mount Kisco, NY) and Augustus software [26] as described earlier [27]. Primary sequences of FAD-dependent halogenases were aligned using the MUSCLE algorithm [28] implemented as plugin for Geneious 7.1 (Biomatters Ltd. Auckland, New Zealand). The alignment was

exported and used to construct a phylogenetic network based on the neighbor-net method implemented in the SplitsTree4 program [29]. For the construction of a phylogenetic tree, the alignment was exported to MEGA6 [30]. Based on a neighbor-joining tree the optimal model of evolution was determined to be the Le-Gascuel (LG) model [31]. A maximum likelihood analysis was run using the LG model with 1000 bootstrap replicates. A FAD-dependent *p*-hydroxybenzoate hydroxylase (PHBH, PDB accession: 1PHH) from *Pseudomonas fluorescens* was used as outgroup to root the tree. Genbank accession numbers of all proteins are compiled in Table S1. For phylogenetic analysis, the sequences of the tryptophan halogenases RebH [32], PrnA [33], and PyrH [34] were used, all of them acting on free precursor molecules. We also included SgcC3 involved in halogenation of the enediyne antitumor substance C-1027 [17], BhaA from the balhimycin biosynthesis [35] and PltA of the pyoluteorin biosynthesis [36]. All of the latter enzymes have experimentally been shown to depend on carrier protein-bound intermediates. The halogenase CndH [37] was suggested to interact with a carrier protein as well, although direct evidence was not provided. AviH from *Streptomyces viridochromogenes* was included as it chlorinates an orsellinic acid moiety at two positions during avilamycin biosynthesis [38] as well as CalO3 that transfers iodine to orsellinic acid [19]. Along with the five *A. mellea* halogenases, we included uncharacterized basidiomycete halogenases of *Heterobasidion irregulare* [39] and *Serpula lacrymans* [40] and characterized halogenases from ascomycetes to the phylogenetic analysis. From the latter we added Rdc2 [41] and RadH [16], both involved in radicicol biosynthesis, as well as GsfI of *Penicillium aethiopicum* which is encoded in the griseofulvin gene cluster [18], PtaM of the pestheic acid biosynthesis in *Pestalotiopsis fici* [42], GedL from the geodin pathway in *Aspergillus terreus* [43] and AclH of the aspirochlorine pathway in *Aspergillus oryzae* [44]. To root the tree we used the FAD-dependent *p*-hydroxybenzoate hydroxylase (PHBH, PDB accession: 1PHH) of *Pseudomonas fluorescens* [45] as outgroup.

Applied and Environmental Microbiology

Results

The melleolide gene cluster lacks a halogenase gene - In order to study the genetic basis underlying melleolide biosynthesis, we first searched the genomic sequence of *A. mellea* DSM3731 [13] using the primary sequences of the protoilludene cyclase Pro1 and polyketide synthase ArmB [11,12]. Notably, the query sequences hit loci that were separated by only nine genes. This result thus pointed towards a clustered arrangement of melleolide biosynthesis genes. This putative gene cluster includes five genes coding for P_{450}-dependent monooxygenases (Figure 2A, proteins 476, 479-482), four NAD^+-dependent oxidoreductases (proteins 468, 469, 474, 478), one flavin-dependent oxidoreductase (protein 475) and one *O*-methyltransferase (protein 473). The P_{450}-dependent enzymes may be involved in protoilludene hydroxylation to elaborate melleolides with multiple alcohol groups, such as melleolide D [3] that carries alcohol functionalities at C-4, C-5, C-10, and C-13. The role of the NAD^+-dependent enzymes remains unknown. Numerous melleolides, e.g., arnamial [6], show 5'-*O*-methylation of the aromatic moiety. This methylation step may be catalyzed by the methyltransferase encoded in the cluster. Searching for a halogenase, a potential candidate encoded in the cluster is the flavin-dependent oxidoreductase (protein 475). However, it lacks the strongly conserved motif GWxWxxPL [34] of FAD-dependent halogenases and showed homology to glucose-methanol-choline-oxidoreductase instead. Thus, we hypothesized that protein 475 does not possess halogenase activity but might represent the dehydrogenase yielding the aldehyde in position 1 of arnamial and other melleolides. In return, this suggested that the halogenase required for melleolide chlorination is not encoded in the gene cluster.

Identification of halogenase genes in the *Armillaria mellea* genome - Prior to the release of the *A. mellea* genomic sequence we have identified the sequences of *armH1* and *armH2* (GenBank accession: JF739169 and JF739170), i.e., two genes coding for putative flavin-dependent halogenases. These two genes are located on a contiguous 71.5 kb portion of the *A. mellea* genome

233 [22], which is not part of the melleolide biosynthesis gene cluster. Searching the genomic sequence,
234 we now identified three additional putative halogenase genes, hereafter referred to as *armH3*,
235 *armH4*, and *armH5* (GenBank accession: KT819179 through KT819181), which are neither located
236 within or near the melleolide gene locus, nor close to *armH1* and *armH2*. For all five halogenase
237 genes (*armH1-armH5*), cDNA was successfully procured from mycelium, confirming these are actively
238 transcribed genes. Another two gene models encoding putative halogenases were incomplete with
239 missing portions across the respective gene, which is why we assume misassembled sequence data
240 or pseudogenes.
241 The reading frames of genes *armH1*, *armH2*, *armH3*, and *armH5* are disrupted by nine, *armH4* by ten
242 Introns (Figure 2B) and encode ArmH1-ArmH5 (Table 2) that possess the canonical halogenase
243 signature sequence (GW(A/V)W(F/L)I) of hydrophobic and aromatic residues.
244
245 ***In vitro* characterization of *A. mellea* halogenases ArmH1-ArmH5 -** To test if any or all of the five *A.*
246 *mellea* halogenases function in chlorine transfer during melleolide biosynthesis, we expressed their
247 coding sequences heterologously in *E. coli* KRX so as to produce *N*-terminally hexahistidine tagged
248 fusion proteins. The enzymes were purified by metal affinity and anion exchange chromatography
249 (Figure 3) and characterized *in vitro.* Melleolide F (Figure 1) [46] is a typical representative of the
250 melleolide family of compounds that features a $\Delta^{2,3}$-protoilludene terpene and an unmodified
251 orsellinic acid moiety. Melleolide F can be procured from mycelial cultures in sufficient quantity and
252 was used as potential chlorine acceptor substrate in this study.
253 Product formation was investigated by HPLC and high-resolution electrospray mass spectrometry
254 (HR-ESIMS). When given free melleolide F as substrate, product formation could be proven the
255 chromatograms for all five ArmH enzymes by HPLC analysis. In addition to the melleolide F peak at
256 t_R= 10.8 min, formation of a chlorinated product was chromatographically verified at t_R= 11.6 min
257 (Figure 4A-4E). In the high-resolution mass spectra for ArmH1-ArmH5 reactions, the new peaks

corresponded to the mass of 6′-chloromelleolide F (Table 3) and displayed the typical pattern for isotopic abundance of stable chlorine isotopes ^{35}Cl and ^{37}Cl. Based on areas under the curve, product formation was most pronounced with ArmH4.

Bromination by ArmH4 - We selected ArmH4 as a model to gain insight into the biosynthetic capacity of *A. mellea* halogenases. As it is established that chlorinases may also catalyze bromination, at least *in vitro* [34] *in vitro* product formation assays were performed as described above, but adding the potassium salts of iodide or bromide, rather than sodium chloride, as halogen donor substrate for ArmH4. While iodination was not observed, product formation was detected by HPLC and HR-ESIMS when bromide was present (Figure 4F). Evidence for melleolide bromination came from the masses *m/z* 479.1075 and 481.1056 (Table 3) and the unique near 1:1 abundance ratio of bromine stable isotopes ^{79}Br and ^{81}Br (Figure 4F). These results indicated that a single bromine atom was introduced to melleolide F, likely to C-6' as noted above (Figure 1).

Functional categories are not applicable to fungal halogenases - It has previously been suggested that halogenases can be grouped according to their substrate requirements [37]. Group A includes enzymes that act on free substrates while group B requires carrier-bound substrates for catalytic turnover. The authors differentiate these two groups by a phenylalanine residue (Phe312 in CndH, group B) which is equivalent to a glutamate residue (Glu346 in PrnA) in group A enzymes. Following these categories, all *A. mellea* halogenases would fall into group B of carrier-protein dependent halogenases, as they show a phenylalanine residue (positions 326, 326, 327, 335, and 326 in ArmH1-ArmH5, respectively). However, our results clearly prove that all of the above halogenases act on free substrates and that the halogen-carbon bond is established as a post-PKS biosynthetic step. This is in agreement with a previous study on the ascomycetous Rdc2 [41], that was shown to chlorinate a free substrate *in vitro,* but has a phenylalanine at position 328. We therefore assume that this signature

Applied and Environmental Microbiology

281 residue - discovered with bacterial enzymes - is not applicable to differentiate halogenases of
282 basidiomycete origin.

283 To gain more insight into the evolution of FAD-dependent halogenases, we constructed a
284 phylogenetic network (Figure 5) and a maximum likelihood (ML) tree (Figure S1) of experimentally
285 characterized FAD-dependent halogenases and some putative halogenases encoded in genomes of
286 basidiomycetes. The resulting network and tree suggested that fungal halogenases form a
287 monophyletic clade. Thus, our phylogenetic analyses support the notion that categorization into
288 group A or B is not applicable to fungal enzymes. We found that the fungal halogenase tree reflects
289 the phylogenetic split between asco- and basidiomycetes. Intriguingly, ArmH1-ArmH5 are not
290 monophyletic, and only ArmH1-ArmH3 form a branch of closely related enzymes.

291 ## Discussion

292 Our findings, reported here, show that halogenases ArmH1-ArmH5 act on melleolide F, i.e., on free
293 substrates and that they do not require carrier protein-tethered intermediates. Cumulatively, we
294 identified these halogenases as melleolide biosynthesis enzymes. We cannot *a priori* exclude that a
295 position other than C-6' served as acceptor site for the halogen atom during catalysis *in vitro*.
296 However, all of the about 25 chlorinated melleolides described to date exclusively show 6'-
297 chlorination. Thus, a different halogen acceptor position would be inconsistent with all previously
298 established structures.

299 Mechanistically, different strategies have evolved independently to install halogen atoms onto
300 natural product scaffolds, making use of either the halide anion, hypochlorous acid, or a halogen
301 radical [47,48]. Recently, even a new variant of halogenating enzymes that belongs to the non-heme
302 iron dependent halogenases was reported [49]. In contrast, our current study focuses on flavin-
303 dependent halogenases which employ FAD as cofactor and use a halide anion and molecular oxygen

304 to produce hypochlorous acid as halogenating agent. Protein crystallography greatly helped
305 understand the structural basis of regioselective formation of the carbon-halogen bond [50-53]. *A.*
306 *mellea* halogenase genes encode the signature motif GW(A/V)W(F/L)I). In the three-dimensional
307 protein structure, these residues line a tunnel inside the enzyme through which the hypochloric acid
308 is routed from the flavin towards the halogen acceptor substrate binding site [50]. Also, the active
309 site lysine was found in all *A. mellea* halogenases (K77 in ArmH5, K76 in all others).

310 Based on the crystal structure of the chondromycin halogenase CndH [37], it has previously been
311 suggested that flavin-dependent halogenases fall into two functionally dissimilar groups accepting
312 free substrates (group A) or carrier protein-bound substrates (group B). Although all halogenases
313 investigated in this study as well as Rdc2 from *Pochonia chlamydosporia* seem to accept free
314 substrates, there is not enough data yet to assume that this is a general trait of fungal halogenases.
315 However, rules deduced from bacterial sequences to predict their placement in either functional
316 group are not appropriate for fungal halogenases. Alternatively, categorization of enzymes by activity
317 on free versus enzyme-tethered substrates may not be as critical as previously assumed to draw a
318 demarcation line between groups A and B. The type of substrate (tryptophan versus others) also
319 seems to play a role which would explain the major split (seen in Figure S1) between bacterial clade 1
320 and all other halogenases. Therefore, further biochemical studies are necessary to shed more light
321 onto this fascinating enzyme family, in particular, as members of basidiomycete origin have remained
322 completely uninvestigated, even though aromatic haloorganic natural products are known from
323 these fungi. Besides the melleolides, examples include the alcalinaphenols from *Mycena alcalina* [54]
324 and stephanosporin of *Stephanospora caroticolor* [55]. Hence, our work represents the first
325 biochemically and genetically characterized basidiomycete halogenases. Phylogenetically, our work
326 also revealed that ArmH1 to ArmH3 have most likely evolved by gene duplication within the *A.*
327 *mellea* lineage, and that ArmH4 and ArmH5 seem to have distinct phylogenetic histories. However, it
328 is premature to speculate if any functional clusters within the fungal clades can be deduced from the

phylogenetic tree. Again, more biochemical studies are warranted to improve functional predictions of putative halogenases unveiled by high throughput sequencing projects. In various cases, e.g., for the antibiotic vancomycin and the antitumor agent salinosporamide, an increased bioactivity has been attributed to the presence of halogen atoms [47]. However, in the case of the melleolides, the physiological reason for introducing a chlorine atom is unclear. Available studies on structure-activity relationships do not support a strongly changed bioactivity through chlorination [8,9,20]. Hence, a role in detoxification of melleolides through halogenation cannot be assumed either, and the reason remains elusive why *A. mellea* secures chlorination of its toxic natural products to the degree described herein. Although redundantly encoded pathways are known from both asco- and basidiomycetes [56-58], a fivefold parallelized biosynthetic step is unprecedented and warrants further investigation on the underlying ecological or environmental reason.

Acknowledgments

A doctoral fellowship by the Deutsche Bundesstiftung Umwelt (DBU) to M.M. is gratefully acknowledged. J.T. acknowledges support by the Collaborative Research Center ChemBioSys (SFB 1127/1). We thank Andrea Perner (Leibniz Institute for Natural Product Research and Infection Biology, Hans-Knöll-Institute, Jena, Germany) for recording high-resolution mass spectra.

References

1. **Baumgartner K, Coetzee MP, Hoffmeister D.** 2011. Secrets of the subterranean pathosystem of *Armillaria*. Mol Plant Pathol **12**: 515-534.

2. **Shaw CG, III, Kile GA.** 1991. *Armillaria* Root Disease. United States Department of Agriculture-Forest Service Handbook No. 691. United States Department of Agriculture-Forest Service, Washington, DC.
3. **Arnone A, Cardillo R, Nasini G.** 1986. Structures of melleolides B-D, three antibacterial sesquiterpenoids from *Armillaria mellea*. Phytochemistry **25**:471-474.
4. **Bohnert M, Miethbauer S, Dahse HM, Ziemen J, Nett M, Hoffmeister D.** 2011. *In vitro* cytotoxicity of melleolide antibiotics: structural and mechanistic aspects. Bioorg Med Chem Lett **21**:2003-2006.
5. **Donnelly D, Sanada S, O´Reilly J, Polonsky J, Prangé T, Pascard CJ.** 1982. Isolation and structure (X-ray analysis) of the orsellinate of armillol, a new antibacterial metabolite from *Armillaria mellea*. J Chem Soc Chem Commun:135-137.
6. **Misiek M, Williams J, Schmich K, Hüttel W, Merfort I, Salomon CE, Aldrich CC, Hoffmeister D.** 2009. Structure and cytotoxicity of arnamial and related fungal sesquiterpene aryl esters. J Nat Prod **72**: 1888-1891.
7. **Yang JS, Chen YW, Feng XZ, Yu DQ, Liang XT.** 1984. Chemical constituents of *Armillaria mellea* mycelium I. Isolation and characterization of armillarin and armillaridin. Planta Med **50**:288-290.
8. **Bohnert M, Nützmann HW, Schroeckh V, Horn F, Dahse HM, Brakhage AA, Hoffmeister D.** 2014. Cytotoxic and antifungal activities of melleolide antibiotics follow dissimilar structure-activity relationships. Phytochemistry **105**:101-108.
9. **Kobori H, Sekiya A, Suzuki T, Choi JH, Hirai H, Kawagishi H.** 2015. Bioactive sesquiterpene aryl esters from the culture broth of *Armillaria* sp. J Nat Prod **78**:163-167.
10. **Peipp H, Sonnenbichler J.** 1992. Secondary fungal metabolites and their biological activities, II. Occurrence of antibiotic compounds in cultures of *Armillaria ostoyae* growing in the presence of an antagonistic fungus or host plant cells. *Biol Chem Hoppe-Seyler* **373**:675-683.

11. **Engels B, Heinig U, Grothe T, Stadler M, Jennewein S.** 2011. Cloning and characterization of an *Armillaria gallica* cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides. J Biol Chem **286**:6871-6878.
12. **Lackner G, Bohnert M, Wick J, Hoffmeister D.** 2013. Assembly of melleolide antibiotics involves a polyketide synthase with cross-coupling activity. Chem Biol **20**:1101-1106.
13. **Collins C, Keane TM, Turner DJ, O'Keeffe G, Fitzpatrick DA, Doyle S.** 2013. Genomic and proteomic dissection of the ubiquitous plant pathogen, *Armillaria mellea*: toward a new infection model system. J Proteome Res **12**:2552-2570.
14. **Chen X, van Pée KH.** 2008. Catalytic mechanisms, basic roles, and biotechnological and environmental significance of halogenating enzymes. Acta Biochim Biophys Sin **40**:183-193.
15. **Reeves CD, Hu Z, Reid R, Kealey JT.** 2008. Genes for the biosynthesis of the fungal polyketides hypothemycin from *Hypomyces subiculosus* and radicicol from *Pochonia chlamydosporia*. Appl Environ Microbiol **74**:5121-5129.
16. **Wang S, Xu Y, Maine EA, Wijeratne EM, Espinosa-Artiles P, Gunatilaka AA, Molnár I.** 2008. Functional characterization of the biosynthesis of radicicol, an Hsp90 inhibitor resorcylic acid lactone from *Chaetomium chiversii*. Chem Biol **15**:1328-1338.
17. **Lin S, Van Lanen SG, Shen B.** 2007. Regiospecific chlorination of (S)-beta-tyrosyl-*S*-carrier protein catalyzed by SgcC3 in the biosynthesis of the enediyne antitumor antibiotic C-1027. J Am Chem Soc **129**:12432-12438.
18. **Chooi YH, Cacho R, Tang Y.** 2010. Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. Chem Biol **17**:483-494.
19. **Ahlert J, Shepard E, Lomovskaya N, Zazopoulos E, Staffa A, Bachmann BO, Huang K, Fonstein L, Czisny A, Whitwam RE, Farnet CM, Thorson JS.** 2002. The calicheamicin gene cluster and its iterative type I enediyne PKS. Science **297**:1173-1176.

20. **Bohnert M, Scherer O, Wiechmann K, König S, Dahse HM, Hoffmeister D, Werz O.** 2014. Melleolides induce rapid cell death in human primary monocytes and cancer cells. Bioorg Med Chem **22**:3856-3861.

21. **Shimizu K, Keller NP.** 2001. Genetic involvement of a cAMP dependent protein kinase in a G protein signaling pathway regulating morphological and chemical transitions in *Aspergillus nidulans*. Genetics **157**:591-600.

22. **Misiek M, Braesel J, Hoffmeister D.** 2011. Characterisation of the ArmA adenylation domain implies a more diverse secondary metabolism in the genus *Armillaria*. Fungal Biol **115**:775-781.

23. **Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, Smirnova T, Grigoriev IV, Dubchak I.** 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res **42**:D26-31.

24. **Schneider P, Weber M, Rosenberger K, Hoffmeister D.** 2007. A one-pot chemoenzymatic synthesis to the universal precursor of antidiabetic and antiviral bis-indolylquinones. Chem Biol **14**:635-644.

25. **Matsubara T, Ohshiro T, Nishina Y, Izumi Y**. 2001. Purification, characterization, and overexpression of flavin reductase involved in dibenzothiophene desulfurization by *Rhodococcus erythropolis* D-1. Appl Environ Microbiol **67**:1179-1184.

26. **Stanke M, Morgenstern B.** 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res **33**:W465–W467.

27. **Misiek M, Hoffmeister D.** 2008. Processing sites involved in intron splicing of *Armillaria* natural product genes. Mycol Res **112**:216-224.

28. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**:1792-1797.

29. **Huson DH**. 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics **14**: 68-73.

30. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol **30**:2725-2729.

31. **Le SQ, Gascuel O.** 2008. An improved general amino acid replacement matrix. Mol Biol Evol **25**:1307-1320.

32. **Yeh E, Garneau S, Walsh CT.** 2005. Robust *in vitro* activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. Proc Natl Acad Sci USA **102**:3960-3965.

33. **Hammer PE, Hill DS, Lam ST, van Pée KH, Ligon JM.** 1997. Four genes from *Pseudomonas fluorescens* that encode the biosynthesis of pyrrolnitrin. Appl Environ Microbiol **63**:2147-2154.

34. **Zehner S, Kotzsch A, Bister B, Süssmuth, RD, Mendez C, Salas JA, van Pée KH.** 2005. A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in *Streptomyces rugosporus* LL-42D005. Chem Biol **12**:445-452.

35. **Puk O, Bischoff D, Kittel C, Pelzer S, Weist S, Stegmann E, Süssmuth RD, Wohlleben W.** 2004. Biosynthesis of chloro-beta-hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. J Bacteriol **186**:6093-6100.

36. **Dorrestein PC, Yeh E, Garneau-Tsodikova S, Kelleher NL, Walsh CT**. 2005. Dichlorination of a pyrrolyl-*S*-carrier protein by $FADH_2$-dependent halogenase PltA during pyoluteorin biosynthesis. Proc Natl Acad Sci USA **102**:13843-13848.

37. **Buedenbender S, Rachid S, Müller R, Schulz GE.** 2009. Structure and action of the myxobacterial chondrochloren halogenase CndH: a new variant of FAD-dependent halogenases. J Mol Biol **385**:520-530.

38. **Weitnauer G, Mühlenweg A, Trefzer A, Hoffmeister D, Süssmuth RD, Jung G, Welzel K, Vente A, Girreser U, Bechthold A.** 2001. Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the *avi* biosynthetic gene cluster of *Streptomyces viridochromogenes* Tü57 and production of new antibiotics. Chem Biol **8**:569-581.

39. **Olson A, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, Canbäck B, Coutinho PM, Cullen D, Dalman K, Deflorio G, van Diepen LT, Dunand C, Duplessis S, Durling M, Gonthier P, Grimwood J, Fossdal CG, Hansson D, Henrissat B, Hietala A, Himmelstrand K, Hoffmeister D, Högberg N, James TY, Karlsson M, Kohler A, Kües U, Lee YH, Lin YC, Lind M, Lindquist E, Lombard V, Lucas S, Lundén K, Morin E, Murat C, Park J, Raffaello T, Rouzé P, Salamov A, Schmutz J, Solheim H, Ståhlberg J, Vélëz H, de Vries RP, Wiebenga A, Woodward S, Yakovlev I, Garbelotto M, Martin F, Grigoriev IV, Stenlid J.** 2012. Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. New Phytol **194**:1001-1013.

40. **Eastwood DC, Floudas D, Binder M, Majcherczyk A, Schneider P, Aerts A, Asiegbu FO, Baker SE, Barry K, Bendiksby M, Blumentritt M, Coutinho PM, Cullen D, de Vries RP, Gathman A, Goodell B, Henrissat B, Ihrmark K, Kauserud H, Kohler A, LaButti K, Lapidus A, Lavin JL, Lee YH, Lindquist E, Lilly W, Lucas S, Morin E, Murat C, Oguiza JA, Park J, Pisabarro AG, Riley R, Rosling A, Salamov A, Schmidt O, Schmutz J, Skrede I, Stenlid J, Wiebenga A, Xie X, Kües U, Hibbett DS, Hoffmeister D, Hogberg N, Martin F, Grigoriev IV, Watkinson SC.** 2011. The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science **333:**762-765.

41. **Zeng J, Zhan J.** 2010. A novel fungal flavin-dependent halogenase for natural product biosynthesis. ChemBioChem **11**:2119-2123.

42. **Xu X, Liu L, Zhang F, Wang W, Li J, Guo L, Che Y, Liu G.** 2014. Identification of the first diphenyl ether gene cluster for pestheic acid biosynthesis in plant endophyte *Pestalotiopsis fici*. ChemBioChem **15**: 284-292.

43. **Nielsen MT, Nielsen JB, Anyaogu DC, Holm DK, Nielsen KF, Larsen TO, Mortensen UH.** 2013. Heterologous reconstitution of the intact geodin gene cluster in *Aspergillus nidulans* through a simple and versatile PCR based approach. PLoS One **8**:e72871.

44. **Chankhamjon P, Boettger-Schmidt D, Scherlach K, Urbansky B, Lackner G, Kalb D, Dahse HM, Hoffmeister D, Hertweck C.** 2014. Biosynthesis of the halogenated mycotoxin aspirochlorine in koji mold involves a cryptic amino acid conversion. Angew Chem Intl Ed **53**:13409-13413.

45. **Schreuder HA, van der Laan JM, Hol WG, Drenth J.** 1988. Crystal structure of p-hydroxybenzoate hydroxylase complexed with its reaction product 3,4-dihydroxybenzoate. J Mol Biol **199**:637-648.

46. **Arnone A, Cardillo R, Di Modugno V, Nasini G.** 1988. Isolation and structure elucidation of melledonals D and E and melleolides E-H, novel sesquiterpenoid aryl esters from *Clitocybe elegans* and *Armillaria mellea*. Gazz Chim Ital **118**:517-521.

47. **Neumann CS, Fujimori DG, Walsh CT.** 2008. Halogenation strategies in natural product biosynthesis. Chem Biol **15**:99-109.

48. **van Pée KH.** 2012. Enzymatic Chlorination and Bromination. Meth Enzymol **516**:237-257.

49. **Hillwig ML, Liu X.** 2014. A new family of iron-dependent halogenases acts on freestanding substrates. Nat Chem Biol **10**:921-923.

50. **Dong C, Flecks S, Unversucht S, Haupt C, van Pée KH, Naismith JH.** 2005. Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. Science **309**:2216-2219.

51. **Bitto E, Huang Y, Bingman CA, Singh S, Thorson JS, Phillips GN. Jr.** 2008. The structure of flavin-dependent tryptophan 7-halogenase RebH. Proteins **70**:289-299.

52. **Zhu X, De Laurentis W, Leang K, Herrmann J, Ihlefeld K, van Pée KH, Naismith JH.** 2009. Structural insights into regioselectivity in the enzymatic chlorination of tryptophan. J Mol Biol **391**:74-85.

53. **Podzelinska K, Latimer R, Bhattacharya A, Vining LC, Zechel DL, Jia Z.** 2010. Chloramphenicol biosynthesis: the structure of CmlS, a flavin-dependent halogenase showing a covalent flavin-aspartate bond. J Mol Biol **397**:316-331.

54. **Peters S, Spiteller P.** 2006. Chloro- and bromophenols from cultures of *Mycena alcalina*. J Nat Prod **69**:1809-1812.

55. **Lang M, Spiteller P, Hellwig V, Steglich W.** 2001. Stephanosporin, a traceless precursor of 2-chloro-4-nitrophenol in the gasteromycete *Stephanospora caroticolor*. Angew Chem Intl Ed **40**:1704-1705.

56. **Forseth RR, Amaike S, Schwenk D, Affeldt KJ, Hoffmeister D, Schroeder FC, Keller NP.** 2013. Homologous NRPS-like gene clusters mediate redundant small-molecule biosynthesis in *Aspergillus flavus*. Angew Chem Intl Ed **52**:1590-1594.

57. **Braesel J, Götze S, Shah F, Heine D, Tauber J, Hertweck C, Tunlid A, Stallforth P, Hoffmeister D.** 2015. Three redundant synthetases secure redox-active pigment production in the basidiomycete *Paxillus involutus*. Chem Biol **22**:1325-1334.

58. **Throckmorton K, Lim FY, Kontoyiannis DP, Zheng W, Keller NP.** 2015. Redundant synthesis of a conidial polyketide by two distinct secondary metabolite clusters in *Aspergillus fumigatus*. Environ Microbiol [Epub ahead of print].

Figure legends

Fig. 1. Chemical structures of orsellinic acid and melleolide F, and its halogenated derivatives.

Fig. 2. A) Physical map of the melleolide biosynthetic gene cluster. The numbers or names of proteins are shown above the respective gene. Introns are not shown. B) Comparison of the gene structures for halogenases ArmH1-ArmH5. Lines represent introns.

Fig. 3. SDS-polyacrylamide gel of heterologously produced and purified *E. coli* flavin reductase (lane 1) and *Armillaria mellea* halogenases ArmH1-ArmH5 (lanes 2-6). Left lane: molecular weight marker. Sizes in kDa are indicated.

Fig. 4. HPLC profiles of ArmH-catalyzed *in vitro* halogenation assays. Panels A through E show results for ArmH1 through ArmH5, respectively. Within each panel, chromatogram i) shows the enzymatic reaction with melleolide F and ii) shows the negative control with heat-treated enzyme. Panel F represents the ArmH4 reaction with bromide. Asterisks (*) mark the product peaks, whose absorption values range between 4 and 15 mAU. Double asterisks (**) mark melleolide F. Insets display HR-ESIMS and UV/Vis spectra of the ArmH4-catalyzed chlorinated and brominated products, and a chromatogram of the melleolide F substrate, respectively.

Fig. 5. Phylogenetic network of 26 verified and putative halogenases, inferred by the neighbor-net method. Fungal halogenases are monophyletic and reflect the phylogenetic split between ascomycetes and basidiomycetes. The functional categories A (halogenases accepting free substrates) and group B (carrier protein-bound substrates) are indicated by bold letters in parentheses. These categories do not strictly correlate with phylogenetic clades. The scale bar

537 indicates the uncorrected pairwise distance. The asterisk indicates a postulated requirement for
538 carrier protein-bound substrates [37].

539

540

541

542

543 **Table 1:** Oligonucleotide primers used during this study.

Oligonucleotide	Sequence (5' to 3')
cHalF1	CATCGGATCCCATCATGGAAGAAC
cHalR1	CTTGAAATAGAATTCGGTTTGCC
cHalR2	CACAGCTCTTCATTATAGCCTAC
FreN-F	AGAAAGCATATGACAACCTTA
FreN-R	AGCTTTTTAGATCTCAGATAAATGC
Hal1-28F	ATCCCCATATGGAAGAACAAGTG
Hal1-28R	CTCGGATCCGGTTTGCCGCTAAG
Hal2F1	CAGTTCCATATGGTTACACAAGTGCC
Hal2R1	CCCTTTGGATCCAGTCTAAAAATCTACG
Hal2R2	AATTCATAACCCTTTCGTTACAGTC
oJT040	AGCCACCATATGAGCTCACTATTGCC
oJT041	CTACGTAAGCTTAGTGTACGTTTTTCAGCC
oJT042	ATACCTCATATGCCTTCCGAATACGTGC
oJT043	AAAAGAATTCTTAGGCCCTGACCAATCCAAGC
oJT073	TTTTTCATATGGAAGCGCAAGTGCC
oHJ01	AAAAGGATCCCTAAGCACACACGAG

544

545

546

547

548

549

550 **Table 2:** Active halogenases/halogenase genes of *Armillaria mellea*.

Protein name	Length (aa)	Protein ID*	GenBank accession	Number of introns in gene	Sequence similarity**
ArmH1	522	10956	JF739169	9	-
ArmH2	516	2897	JF739170	9	68.3
ArmH3	504	9219	KT819179	9	79.7
ArmH4	533	168	KT819180	10	56.8
ArmH5	523	2211	KT819181	9	49.7

551 *refers to the *Armillaria mellea* genome project as provided through the Joint Genome Institute and
552 Collins et al. [13].

553 ** percentage of identical amino acids relative to ArmH1.

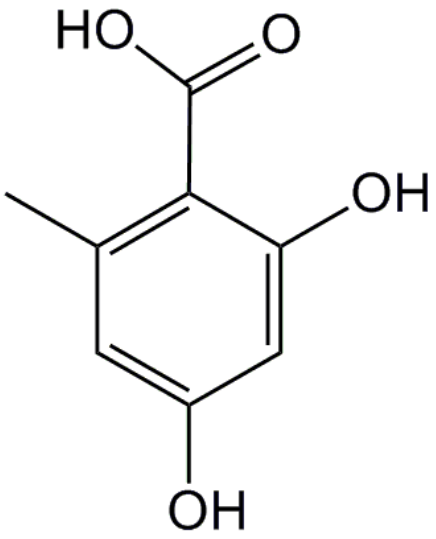
554

555 **Table 3:** High-resolution mass spectrometry data of *in vitro* chlorinated or brominated melleolide F,
556 using ArmH1-ArmH5.

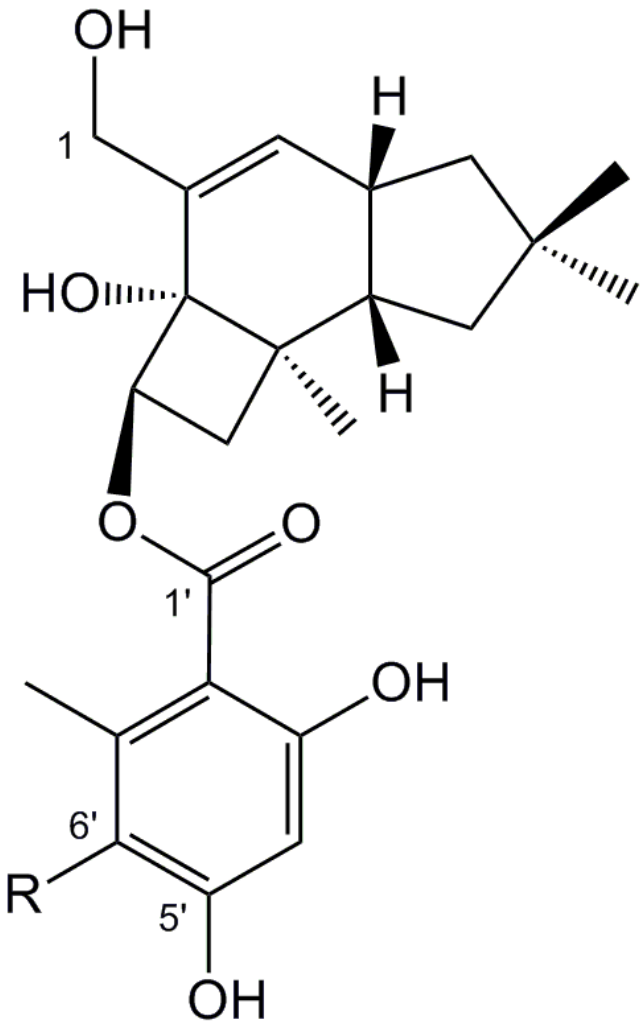
Halogenase	found ***m/z*** $[M - H]^-$	calculated ***m/z*** $[M - H]^-$	
ArmH1	435.1581	435.1580*	chlorination
ArmH2	435.1595	435.1580*	
ArmH3	435.1600	435.1580*	
ArmH4	435.1584	435.1580*	
ArmH5	435.1579	435.1580*	
ArmH4	479.1075, 481.1056	479.1069, 481.1049**	bromination

557 *for $C_{23}H_{28}O_6Cl$

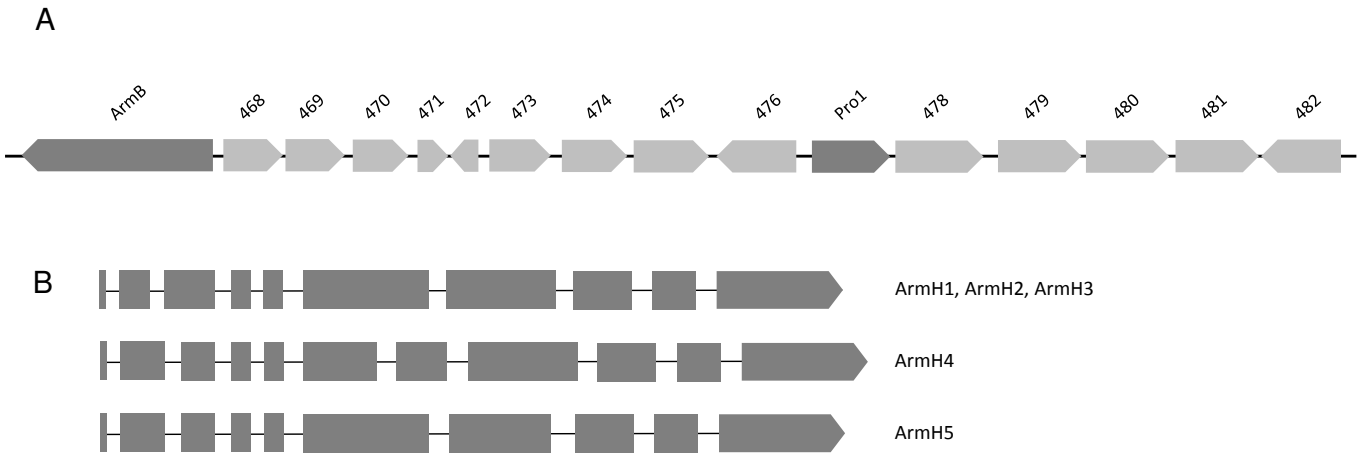
558 ** for $C_{23}H_{28}O_6{}^{79}Br$ and $C_{23}H_{28}O_6{}^{81}Br$, respectively.

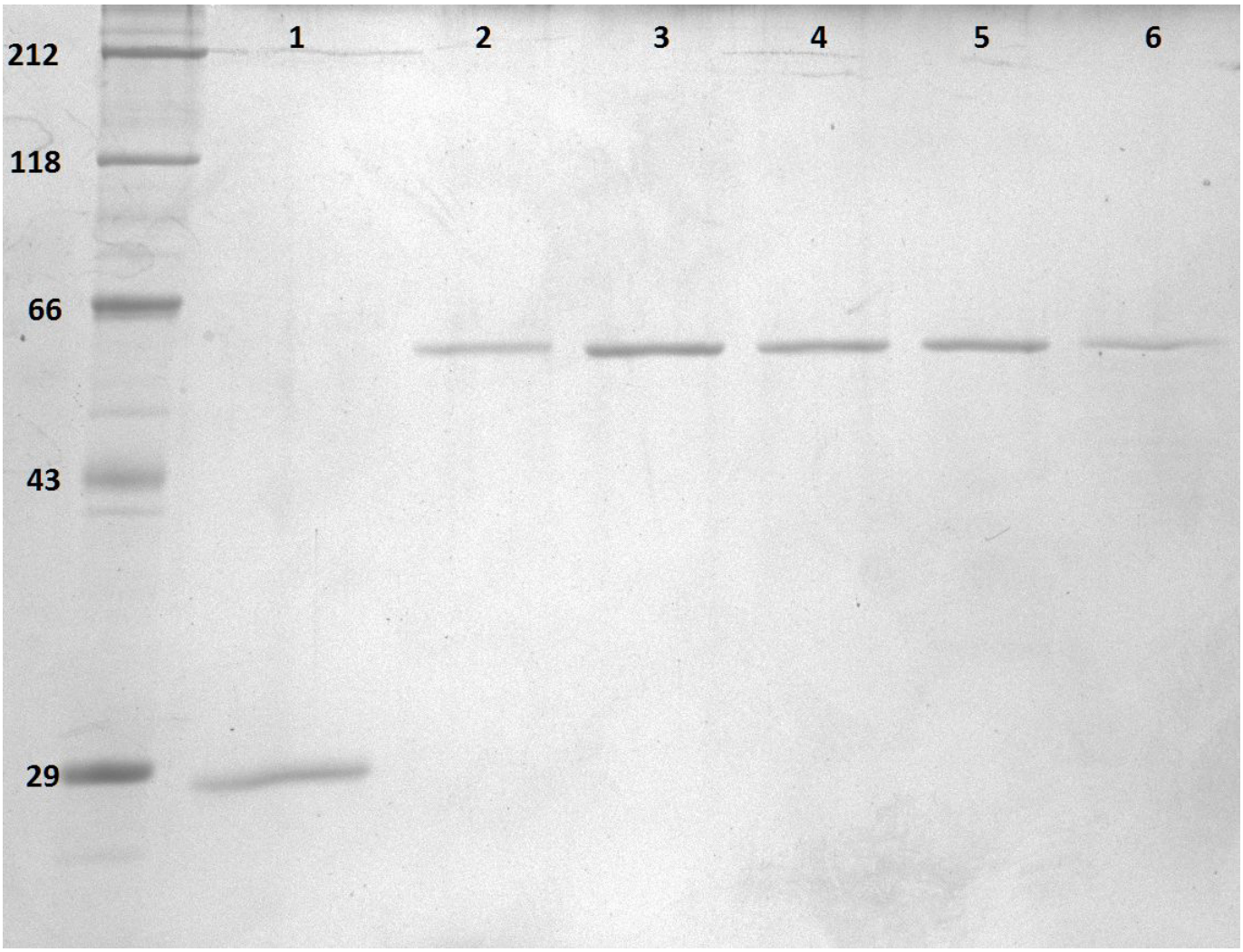


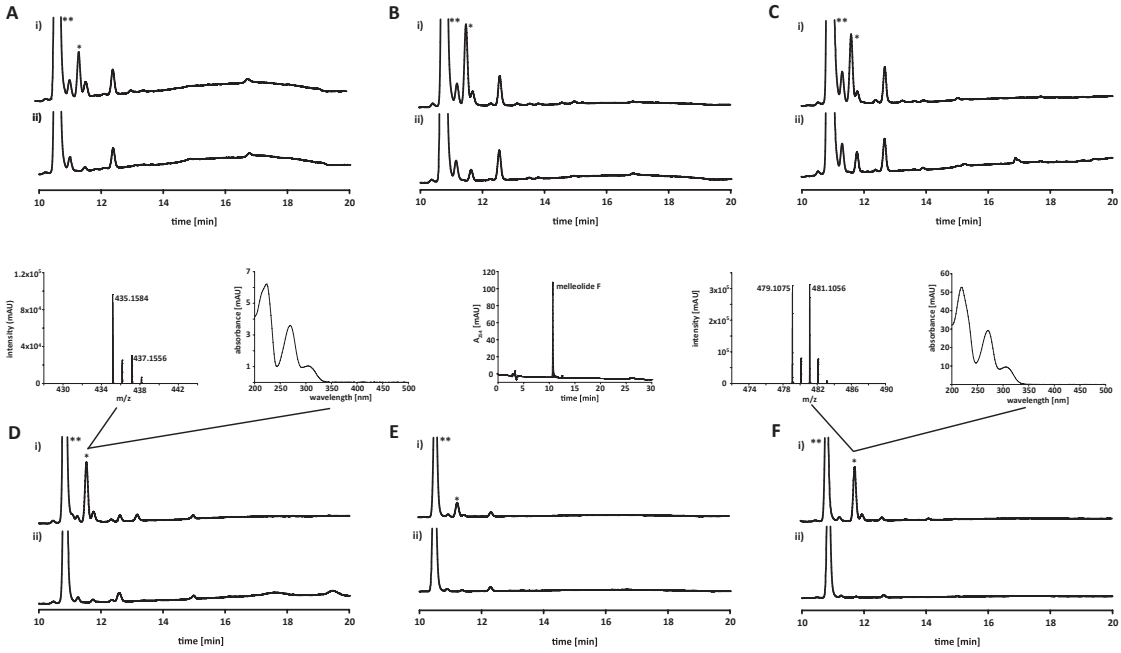
Orsellinic acid



Melleolide F: R = H
6'-Chloromelleolide F: R = Cl
6'-Bromomelleolide F: R = Br







Applied and Environmental Microbiology