

## Article

# Two Novel, Flavin-Dependent Halogenases from the Bacterial Consortia of *Botryococcus braunii* Catalyze Mono- and Dibromination

Pia R. Neubauer <sup>1</sup>, Olga Blifernez-Klassen <sup>2</sup> , Lara Pfaff <sup>1</sup>, Mohamed Ismail <sup>1</sup> , Olaf Kruse <sup>2</sup> and Norbert Sewald <sup>1,\*</sup> 

<sup>1</sup> Organic and Bioorganic Chemistry, Faculty of Chemistry and Centre for Biotechnology (CeBiTec), Bielefeld University, D-33615 Bielefeld, Germany; pia.neubauer@uni-bielefeld.de (P.R.N.); lara.pfaff@uni-greifswald.de (L.P.); M.Ismail10@bradford.ac.uk (M.I.)

<sup>2</sup> Algae Biotechnology and Bioenergy, Faculty of Biology and Centre for Biotechnology (CeBiTec), Bielefeld University, D-33615 Bielefeld, Germany; olga.blifernez@uni-bielefeld.de (O.B.-K.); olaf.kruse@uni-bielefeld.de (O.K.)

\* Correspondence: norbert.sewald@uni-bielefeld.de; Tel.: +49-521-106-6963

**Abstract:** Halogen substituents often lead to a profound effect on the biological activity of organic compounds. Flavin-dependent halogenases offer the possibility of regioselective halogenation at non-activated carbon atoms, while employing only halide salts and molecular oxygen. However, low enzyme activity, instability, and narrow substrate scope compromise the use of enzymatic halogenation as an economical and environmentally friendly process. To overcome these drawbacks, it is of tremendous interest to identify novel halogenases with high enzymatic activity and novel substrate scopes. Previously, Neubauer et al. developed a new hidden Markov model (pHMM) based on the PFAM tryptophan halogenase model, and identified 254 complete and partial putative flavin-dependent halogenase genes in eleven metagenomic data sets. In the present study, the pHMM was used to screen the bacterial associates of the *Botryococcus braunii* consortia (PRJEB21978), leading to the identification of several putative, flavin-dependent halogenase genes. Two of these new halogenase genes were found in one gene cluster of the *Botryococcus braunii* symbiont *Sphingomonas* sp. In vitro activity tests revealed that both heterologously expressed enzymes are active flavin-dependent halogenases able to halogenate indole and indole derivatives, as well as phenol derivatives, while preferring bromination over chlorination. Interestingly, SpH1 catalyses only monohalogenation, while SpH2 can catalyse both mono- and dihalogenation for some substrates.

**Keywords:** enzyme identification; flavin-dependent halogenases; bromination; metagenome screening; bacterial consortia of *Botryococcus braunii*; bioinformatics



**Citation:** Neubauer, P.R.; Blifernez-Klassen, O.; Pfaff, L.; Ismail, M.; Kruse, O.; Sewald, N. Two Novel, Flavin-Dependent Halogenases from the Bacterial Consortia of *Botryococcus braunii* Catalyze Mono- and Dibromination. *Catalysts* **2021**, *11*, 485. <https://doi.org/10.3390/catal11040485>

Academic Editor: Chia-Hung Kuo

Received: 1 March 2021

Accepted: 4 April 2021

Published: 10 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Halogenated organic compounds often show higher biological activities than the non-halogenated correlates. Therefore, chemical halogenation is an important methodology, but often lacks regioselectivity. Moreover, it requires relatively harsh reaction conditions and Lewis acid catalysts, and is usually performed in an organic solvent [1]. Hence, a regioselective and facile halogenation method would provide a more environmentally friendly alternative. Nature has evolved enzymatic halogenation, making use of different cofactors [2]. The enzymatic halogenation by flavin-dependent halogenases may overcome the drawbacks of chemical halogenation, since it often is highly regioselective and only requires a halide salt as a halogen source, water, oxygen, and reduced flavin-adenosine-dinucleotide (FADH<sub>2</sub>) as a cofactor [1,3–6]. In general, flavin-dependent halogenases belong to the superfamily of flavin-dependent monooxygenases, which are able to activate molecular oxygen by using reduced flavin (FADH<sub>2</sub>) [7], thus allowing diverse reactions, such as hydroxylation, epoxidation, and Baeyer–Villiger oxidation [8].

Tryptophan (Trp) halogenases are flavin-dependent halogenases, categorized according to their specific position of halogenating tryptophan. Trp halogenases have already been employed for halogenation on a gram scale, making use of enzyme immobilisation as cross-linked enzyme aggregates (CLEAs) [1]. The enzymatic halogenation approach has been combined with Suzuki–Miyaura cross-coupling [4,9–12], Sonogashira–Hagihara cross-coupling [13], and Mizoroki–Heck cross-coupling [14,15] in reaction cascades [4,10,11]. The tryptophan 7-halogenases PrnA, RebH, and KtzQ halogenate tryptophan regioselectively in the C7 position; the tryptophan 6-halogenases Thal, SttH, and Th-Hal prefer the C6 position, while the tryptophan 5-halogenases PyrH and ClaH modify the C5 position [16–24]. It has also been shown that it is possible to switch regioselectivity by exchanging amino acid residues. Following this strategy, Moritzer et al. redirected the regioselectivity of the tryptophan 6-halogenase Thal to the C7 position [25]. A combination of directed evolution, rational design, as well as site-saturation mutagenesis was also implemented for the design of a thermostable Thal variant with stronger elevated activity [26].

All known flavin-dependent halogenases require FADH<sub>2</sub> as a cofactor, whereas other flavin derivatives like riboflavin and flavin-mononucleotide (FMN) are not accepted [27]. FADH<sub>2</sub> is bound in the flavin binding site (motif GxGxxG) of the enzymes and reacts with oxygen to give a flavin hydroperoxide (FAD–OOH). This is followed by a nucleophilic attack of a halide forming hypohalous acid (HOX). The substrate binding site is positioned far away from the flavin binding site, but both sites are connected by a 10 Å long tunnel [28–30]. HOX passes through this tunnel, as verified by molecular dynamics calculations [31]. Different amino acid residues in the substrate binding site, as well in the tunnel, either interact with HOX or the substrate, while others are responsible for positioning the substrate to effect regioselective halogenation. Yeh et al. [32] elucidated the important role of a conserved lysine residue, e.g., K79 (PrnA) [28], K79 (RebH) [32], K83 (BrvH) [33], K74 (RadH) [34] that supposedly interacts with HOX. The potential formation of a long-lived chloramine (Lys–N<sup>ε</sup>H–Cl) is still under debate [32]. Alternatively, the conserved lysine residue has been postulated to stabilize and position HOX by a hydrogen bond [35]. In both cases, the electrophilic HOX species reacts with the substrate bound to the substrate binding site. This leads to a Wheland intermediate in the course of electrophilic aromatic substitution (S<sub>E</sub>Ar), which is stabilized by highly conserved glutamic acid (E346 (PrnA) [28]; E357 (RebH) [36]). The carboxylate group of the conserved glutamic acid residue interacts with HOX or the chloramine, which increases electrophilicity and aligns it for regioselective halogenation. Flavin-dependent halogenases with substrates other than tryptophan do not possess the glutamic acid residue [37]. Other compounds, such as phenols or pyrroles, might be more susceptible for halogenation by HOX or the chloramine species [35]. The conserved amino acid motif WxWxIP is located at the flavin binding domain, and is believed to block binding of the substrate near the flavin cofactor, which suppresses monooxygenase activity of the enzymes [28,29,38].

Over the last years, several new flavin-dependent halogenases have been identified and heterologously expressed in different host organisms, leading to integration of halogenated substrates in the biosynthetic production of host compounds [39–42]. PrnA was the first flavin-dependent halogenase to be identified in 1997, within the pyrrolnitrin gene cluster [42]. The tryptophan 7-halogenase RebH described in 2002 is responsible for the chlorination of the secondary metabolite rebeccamycin [18]. By employing degenerative primers for highly conserved motifs, Zehner et al. identified PyrH [23], and Fujimori et al. identified KtzR and KtzQ [19] in different bacterial genomes, while Smith et al. retrieved KrmI from a marine sponge metagenome [43]. With the advent of next-generation sequencing methods, many genomes and associated metagenomes have been sequenced, and the obtained data led to the identification of many more halogenases, elucidated based on these conserved amino acid regions [22,24,44–51]. Neubauer et al. created in 2018 a profile hidden Markov model (pHMM), which led to the identification of several putative, flavin-dependent halogenases in different metagenomic data sets [33]. One of them, BrvH, was characterized with respect to halide and substrate preference in vitro, leading to the

conclusion that it halogenates indole and indole derivatives, as well as phenol derivatives, and prefers bromination over chlorination. BrvH was crystallized and revealed a structure similar to Trp halogenases, in addition to an open substrate binding site, which might lead to the acceptance of larger substrates, such as peptides [33,52]. BrvH was the first reported flavin-dependent halogenase that accepts chloride and bromide with a preference for bromination over chlorination. Only the phenol halogenases Bmp2 and Bmp5 had been reported as brominases that do not accept chloride [46]. Ismail et al. identified three novel, flavin-dependent halogenases from *Xanthomonas campestris* that exclusively brominate the substrates indole, 7-azaindole, 5-hydroxytryptophan, tryptophol, and other heterocyclic derivatives, even in the presence of an excess of chloride [51]. Like for BrvH, the crystal structure of Xcc4156 from *Xanthomonas campestris* showed an open substrate binding site. However, crystallization with FAD and bromide resulted in disruption of the crystal, which leads to the conclusion that this binding leads to positive cooperativity and conformational change [53].

Within the present work, the metagenomes of *Botryococcus braunii* communities were screened for the presence of putative, flavin-dependent halogenases by applying the generated profile hidden Markov model (pHMM) [33]. Two of the identified novel halogenases were heterologously expressed in *E. coli* and analysed in detail.

## 2. Results and Discussion

### 2.1. Identification and Analysis of Novel, Flavin-Dependent Halogenases from the *Botryococcus braunii* Consortia

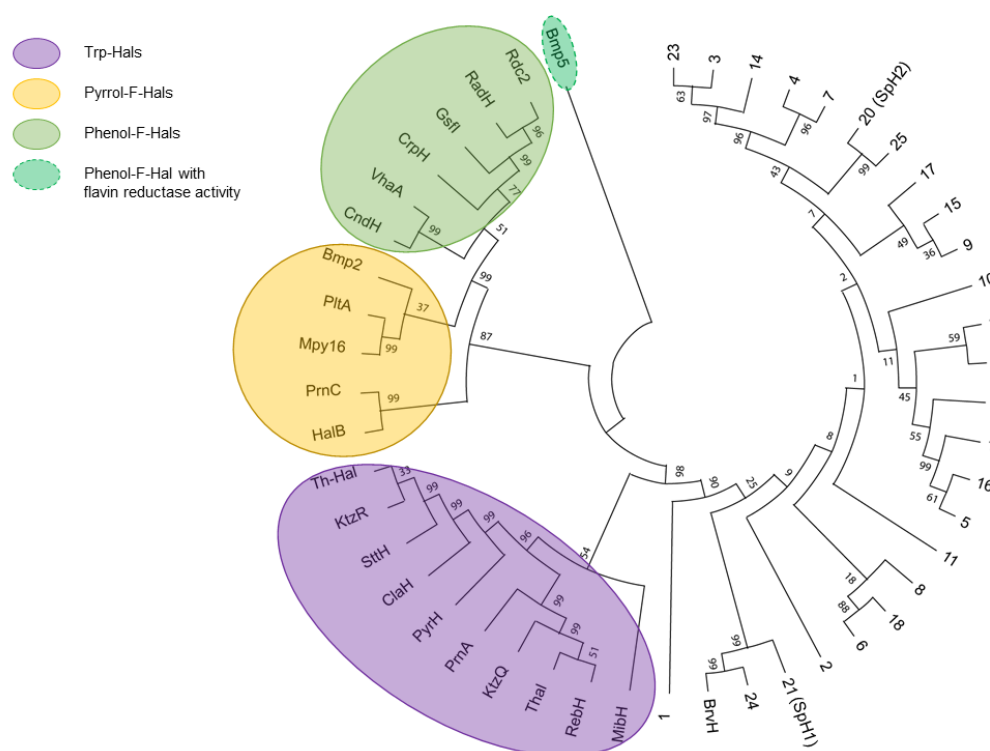
Since the utilization of hydrocarbons by microorganisms relies on a set of different monooxygenases [54], we considered the microbial consortia accompanying the hydrocarbon-secreting microalga *Botryococcus braunii* as a potential source for novel monooxygenases [55] and potential halogenases. *Botryococcus braunii* is a colony-forming green microalga belonging to the class *Treboxiophyceae*, which can be sub-divided into distinct races depending on the type of hydrocarbon synthesized [56]. *Botryococcus* readily releases large amounts of organic carbon into the extracellular medium [57], creating a phycosphere that naturally attracts many microorganisms, including various taxa known to utilize hydrocarbons [58].

We screened metagenomic data sets of *Botryococcus braunii* consortia [59] for novel, flavin-dependent halogenases by applying our pHMM strategy [33], which is based on the PFAM database (<http://pfam.xfam.org/>, accessed on 8 April 2021) tryptophan-halogenase model (Trp\_halogenase, PF04820), to increase the scope of possible halogenases. The metagenomes of four *Botryococcus braunii* consortia were analysed; they contained at least 33 distinct bacterial species, as indicated by 16 S rDNA amplicon and metagenome sequencing analyses [59]. For the analysis of the metagenomes, according to Neubauer et al., the pHMM was employed, which was based on Trp halogenases and was optimised in a two-step approach [33]. The analysis led to the identification of 18 complete and seven partial putative flavin-dependent halogenase genes, which possess the known conserved amino acid regions of Trp halogenases. With one exception, all halogenases found were encoded in the genomes of alphaproteobacteria, while 16 belonged to the sphingomonads genera, such as *Sphingomonas* and *Sphingopyxis* (Table 1).

A phylogenetic tree was constructed based on amino acid sequence to further analyse the identified hits and to compare these to known flavin-dependent halogenases (Figure 1). The phylogenetic analysis was conducted in MEGAX [60] by using the neighbour-joining (NJ) method and a bootstrap of 1000 [61,62]. The phylogenetic tree is divided into five groups of flavin-dependent halogenases. The phenol, pyrrol, and Trp halogenases build their own clades, respectively.

**Table 1.** Newly found putative, flavin-dependent halogenases with the conserved amino acid motif of Trp halogenases and their taxonomic assignment, based on BLAST analyses (<https://blast.ncbi.nlm.nih.gov>; Basic Local Alignment Search Tool).

No.	Length (bp)	Complete	Contig Location	Phyla	Putative Origin
1	225	no	contig-391000007_2	Alphaproteobacteria	Porphyrobacter
2	394	no	contig-1913000010_2	Alphaproteobacteria	Sphingopyxis
3	348	no	contig-565000012_1	Alphaproteobacteria	Caulobacter
4	272	no	contig-1532000014_2	Alphaproteobacteria	Sphingopyxis
5	503	yes	contig-1000028_60	Alphaproteobacteria	Sphingomonas
6	502	yes	contig-1029000036_1	Alphaproteobacteria	Sphingopyxis
7	501	yes	contig-3201000046_41	Gammaproteobacteria	Stenotrophomonas
8	511	yes	contig-2000047_22	Alphaproteobacteria	Sphingomonas
9	501	yes	contig-2000047_33	Alphaproteobacteria	Sphingomonas
10	533	yes	contig-2000047_110	Alphaproteobacteria	Sphingomonas
11	505	yes	contig-2000047_112	Alphaproteobacteria	Sphingomonas
12	501	yes	contig-2000047_113	Alphaproteobacteria	Sphingomonas
13	418	no	contig-1419000070_1	Alphaproteobacteria	Caulobacter
14	359	no	contig-832000073_1	Alphaproteobacteria	Porphyrobacter
15	501	yes	contig-252000084_3	Alphaproteobacteria	Sphingopyxis
16	501	yes	contig-2000086_90	Alphaproteobacteria	Brevundimonas
17	443	no	contig-2264000090_1	Alphaproteobacteria	Sphingopyxis
18	521	yes	contig-16000092_39	Alphaproteobacteria	Sphingomonas
19	502	yes	contig-867000143_2	Alphaproteobacteria	Caulobacter
20	513	yes	contig-3212000146_163	Alphaproteobacteria	Sphingomonas
21	512	yes	contig-3212000146_164	Alphaproteobacteria	Sphingomonas
22	553	yes	contig-3212000146_184	Alphaproteobacteria	Sphingomonas
23	493	yes	contig-1000154_7	Alphaproteobacteria	Sphingomonas
24	522	yes	contig-981000159_36	Alphaproteobacteria	Brevundimonas
25	516	yes	contig-981000159_37	Alphaproteobacteria	Brevundimonas

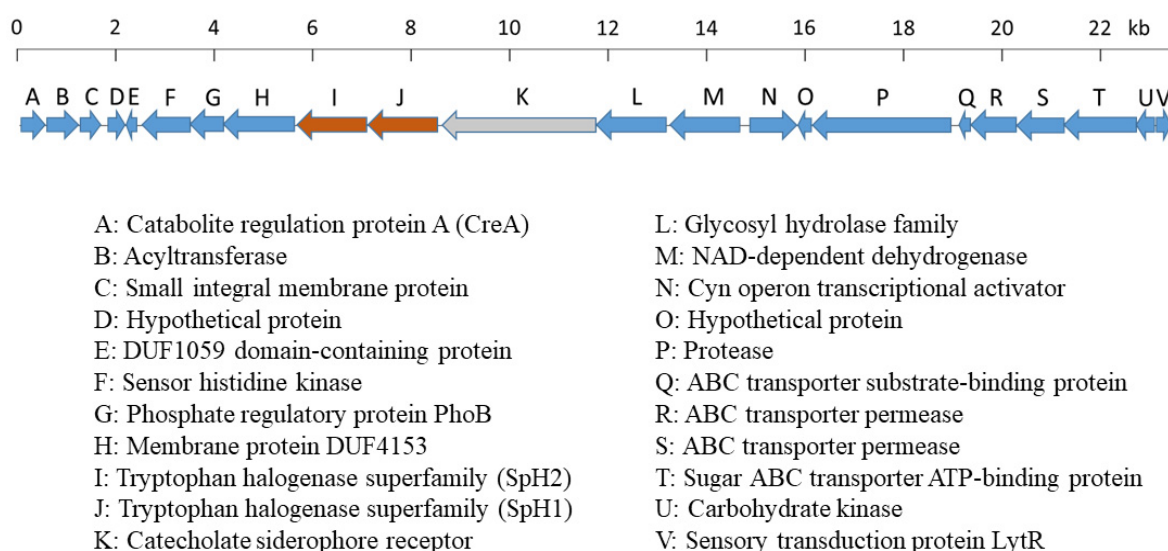


**Figure 1.** Phylogenetic tree of the hits from the *Botryococcus braunii* consortia, as well as known Trp halogenases, phenol halogenases, and pyrrol halogenases. The optimal tree, with the sum of branch lengths = 17.80254621, is shown. The newly identified hits all cluster with the flavin-dependent halogenase BrvH. We decided to further analyse hit 21 (SpH1) and hit 20 (SpH2).



Noteworthy, Bmp5 is outside the phenol halogenase clade, although it is a phenol halogenase. However, it also possesses flavin reductase activity, which might be the reason for building its own clade [46]. The newly identified hits from the *B. braunii* bacterial consortia cluster close to Trp halogenases, but build their own clade together with BrvH. BrvH was found to be very similar to hits 24 and 21 (pairwise identity of 95% and 64%, respectively), which represent flavin-dependent halogenases encoded in the metagenome-assembled genomes (MAG 10 and MAG 21, respectively [59]) of *Brevundimonas* and *Sphingomonas*. Thus, these enzymes might represent their own group within the flavin-dependent halogenases with likely similar substrate preferences. According to antiSMASH analyses (antibiotics & Secondary Metabolite Analysis Shell [63]), both MAGs encode several metabolic gene clusters for bacteriocins and antibiotic compounds [59,64]. In addition, the MAG similar to *Sphingomonas* contains a gene cluster potentially responsible for siderophore synthesis.

We decided to investigate the identified gene 21 (J) from *Sphingomonas*. However, during the gene cluster analysis, the presence of another flavin-dependent halogenase gene (20, I) located upstream to 21 (J) was detected (Figure 2). Some of the known flavin-dependent halogenases have been reported to originate from the same gene cluster, e.g., PrnA and PrnC that play a role in pyrrolnitrin biosynthesis [42]. Bmp2 and Bmp5 likewise halogenate different building blocks in the biosynthesis of natural, polybrominated marine products [41]. KtzR and KtzQ act together in the biosynthesis of kutzneride. In this case, KtzQ is responsible for the chlorination of tryptophan at C7, while KtzR catalyses the second halogenation to give 6,7-dichlorotryptophan. KtzR also catalyses chlorination of tryptophan but possesses a higher affinity to 7-chlorotryptophan [19].



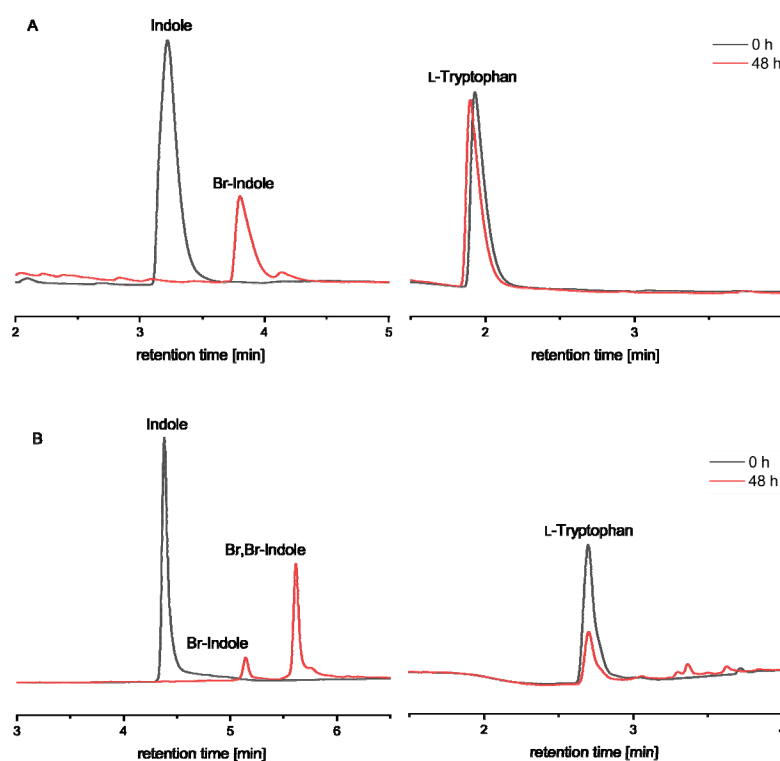
**Figure 2.** Gene cluster in the vicinity of hit 21 (J; SpH1) in the metagenome-assembled genome (MAG) 21 of the *B. braunii* bacterial consortia. Another flavin-dependent halogenase gene, *sph2* (J), was identified upstream to *sph1* (I). The genes I, J, K, and L were found to be encoded on the same contig in the metagenomics dataset.

Since many flavin-dependent halogenases originating from the same gene cluster were reported to act in concert [17,19,46], both putative, flavin-dependent halogenases 21 (J) and 20 (I)—in the following named SpH1 (21) and SpH2 (20)—were heterologously expressed in *E. coli* and subjected to further analyses. Both enzymes possess the typical conserved amino acid regions of flavin-dependent halogenases like the flavin binding domain and the conserved lysine residues. However, both show 42% pairwise identity to each other, while the pairwise identity to BrvH was 64% for SpH1 and 42% for SpH2. The amino acid sequences are provided in the Supplementary Materials.

## 2.2. In Vitro Experiments with the Flavin-Dependent Halogenases SpH1 and SpH2

### 2.2.1. Determination of Halogenation Activity and Substrate Scope

In order to confirm that the identified genes are active halogenating enzymes, experiments with the novel enzymes were conducted in vitro. The genes *sph1* and *sph2* were heterologously expressed in *E. coli*, together with the chaperone system GroEL/GroES. Both enzymes were obtained in soluble form in high concentrations (Supplementary Materials). For the substrate screening, the previously established reaction conditions, with a cofactor regeneration by flavin reductase (PrnF), and *Rhodococcus ruber* alcohol dehydrogenase (RR-ADH), were used [3]. Indole and L-tryptophan were the first tested substrates, because SpH1 and SpH2 are similar to BrvH in brominating indole [33]. Similar to the activity observed for BrvH, SpH1 and SpH2 catalyse the bromination of indole, but not L-tryptophan (Figure 3), thus revealing that both novel identified enzymes represent active halogenases.



**Figure 3.** Reverse-phase high-performance liquid chromatography (RP-HPLC) profiles (detection at 280 nm) of the conversions of 1 mM indole or 1 mM L-tryptophan by SpH1 (A) and SpH2 (B) in presence of 100 mM NaBr as a halide source. *Rhodococcus ruber* alcohol dehydrogenase (RR-ADH) and flavin reductase PrnF were used for cofactor regeneration. Retention times differ due to different HPLC methods. (A) SpH1 (HPLC method 1); (B) SpH2 (HPLC method 2). SpH1 had completely halogenated indole (retention time  $t_R$ : 3.2 min) after 48 h of incubation to give bromoindole ( $t_R$ : 3.8 min), but L-tryptophan ( $t_R$ : 2 min) was not converted. SpH2 catalyses the halogenation of indole ( $t_R$ : 4.4 min) to give bromoindole ( $t_R$ : 5.1 min) and dibromoindole ( $t_R$ : 5.6 min), without converting L-tryptophan ( $t_R$ : 2.7 min).

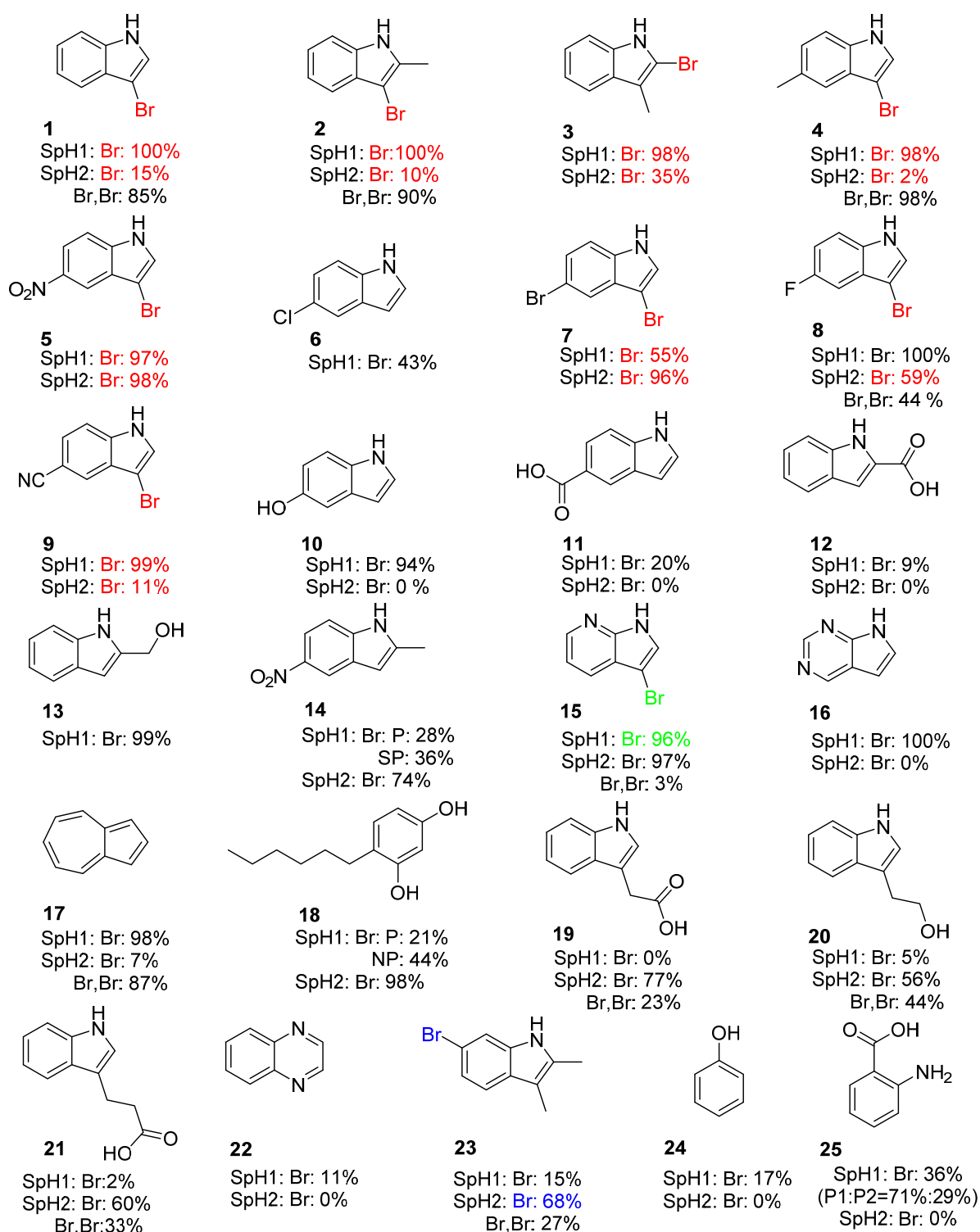
The newly identified brominating enzymes belong to the class of flavin-dependent halogenases. Both enzymes have the characteristic flavin binding module GxGxxG, the conserved lysine residue (K85 and K81, respectively), and the WxWxIP motif suppressing monooxygenase activity [33]. The reduced flavin ( $\text{FADH}_2$ ) reacts with molecular oxygen, forming the FAD(C4a)-peroxide, which in turn reacts with halide ions under HOX formation [2,5,28,29]. In contrast, heme- or vanadium-dependent haloperoxidases require hydrogen peroxide and release the halogenating species hypohalous acid HOX into the medium, which leads to unselective halogenation [2,65,66]. Flavin-dependent halogenases

can be distinguished from haloperoxidases by the monochlorodimedon assay [67]. In this assay, monochlorodimedon reacts with free HOX released by the haloperoxidase to give rise to a dihalogenated derivative, which can be monitored photometrically. This test was negative for both enzymes. The addition of catalase, which decomposes hydrogen peroxide and would hence stop the halogenation reaction by haloperoxidases, did not negatively affect the halogenation reaction, leading to the conclusion that SpH1 and SpH2 are not haloperoxidases (Supplementary Materials). FAD reconstitution is another possibility for the classification towards flavin-dependent halogenases [33,68]. SpH1 and SpH2 were incubated with FAD overnight, and then the buffer was changed to a buffer without FAD, leading to no free FAD in the solution. UV/Vis spectroscopy can be employed for the detection of bound FAD in the enzymes. For both SpH1 and SpH2, the absorption band at 350 nm shifted to 348 nm, and the absorption band at 373 nm shifted to 360 nm, showing that both enzymes bind FAD and can therefore be classified as flavin-dependent halogenases (Supplementary Materials). The SpH1 mutant where the conserved lysine residue K85 was replaced by an alanine (SpH1\_K85A; Supplementary Materials) was completely inactive. Hence, halogenase SpH1 depends on the lysine, which is known to be essential for the flavin-dependent halogenases.

A pharmacophore model was adapted from a virtual screening methodology for BrvH, and revealed 19 compounds halogenated by BrvH [52]. SpH1 is highly similar to BrvH and clusters within one clade with BrvH. Therefore, its substrate scope might be similar as well. In addition to indole and L-tryptophan, further substrates, such as tryptophan and indole derivatives, as well as other aromatic compounds, were tested for halogenation by SpH1 and SpH2 (Figure 4). Twenty of the 26 accepted substrates represent indole derivatives, but phenol derivatives (i.e., phenol (**24**), anthranilic acid (**25**), and 4-*n*-hexylresorcinol (**18**)), azulene (**17**), and quinoxaline (**22**) were also accepted. Enzymatic conversion was verified by reverse-phase high-performance liquid chromatography (RP-HPLC), as well as mass analysis (Supplementary Materials).

SpH1 and SpH2 were immobilised as CLEAs to convert higher amounts of substrate, in order to identify their regioselectivity. The halogenases and cofactor-regenerating enzymes were immobilised together using glutaraldehyde, and the CLEAs were incubated for 10 to 14 days with substrate. The obtained product was purified and analysed by NMR spectroscopy. Mass spectrometry (MS) analyses revealed the correct isotopic pattern of the brominated products (Supplementary Materials). All indole derivatives analysed so far are halogenated at C3, if this position is unsubstituted. Only when a non-hydrogen substituent is present at C3, halogenation occurs at C2 (Figure 4). The C3 and C2 positions are electronically the most favored positions for electrophilic substitution within the indole ring [69]. Likewise, PrnA and RebH halogenate indole and some derivatives in the most activated position. The regioselectivity strongly depends on the substrate and its position in the active site [3,70]. SpH2 is also able to halogenate 2,3-methylindole. In this case, it halogenates the substrate at C6, but also a minor side product halogenated at C5 was observed, leading to the conclusion that SpH2 is not strictly regioselective for 2,3-methylindole.

Furthermore, SpH1 and SpH2 show higher bromination activity compared to chlorination activity. They brominated 1 mM indole to 100% within 48 h, while SpH1 under similar conditions chlorinated only to 10%, and SpH2 to 17%. Interestingly, dichlorinated product was not found for SpH2. The brominated, as well as the chlorinated products, were verified via ESI-MS or liquid chromatography LC-MS analyses (**SpH1**: positive mode:  $[\text{C}_8\text{H}_6\text{BrN}]^+$ , calcd. 195.97 ( $^{79}\text{Br}$ ), 197.97 ( $^{81}\text{Br}$ ), obs.  $[\text{M}+\text{H}]^+$  195.55/197.53; negative mode:  $[\text{C}_8\text{H}_6\text{ClN}]^-$ , calcd. 150.02 ( $^{35}\text{Cl}$ ), 152.02 ( $^{37}\text{Cl}$ ), obs.  $[\text{M}-\text{H}]^-$  150.01/152.01; **SpH2**: negative mode:  $[\text{C}_8\text{H}_6\text{BrN}]^-$ , calcd. 193.97 ( $^{79}\text{Br}$ ), 195.97 ( $^{81}\text{Br}$ ), obs.  $[\text{M}-\text{H}]^-$  193.96/195.95;  $[\text{C}_8\text{H}_6\text{ClN}]^-$ , calcd. 150.02 ( $^{35}\text{Cl}$ ), 152.02 ( $^{37}\text{Cl}$ ), obs.  $[\text{M}-\text{H}]^-$  150.01/152.01).



**Figure 4.** Halogenated substrates by SpH1 and SpH2. The conversions were calculated by determining the ratio of the peak areas of substrates and products. The position of halogenation was determined for 10 substrates by NMR analysis (red: SpH1 and SpH2, green: SpH1, and blue: SpH2). SP: different unidentified side products; P1, P2: two monobrominated products at a ratio of P1/P2.; NP: side product.

SpH1 and SpH2 are thus similar to BrvH and the Xcc halogenases Xcc4156, Xcc1333, and Xcc4345, which all prefer bromination over chlorination [33,51]. Noteworthy, the Xcc halogenases, SpH1, and SpH2 originate from terrestrial habitats, but favour bromination over chlorination. The Xcc halogenases were identified in *Xanthomonas campestris*, which



is a plant pathogen [51], and SpH1 and SpH2 originate from *B. braunii* consortia, which mainly inhabit fresh water [57]. Usually, it is presumed that because of the higher bromide concentration in sea water, bromination is preferred in marine habitats, while halogenases from terrestrial habitat would favour chlorination [71].

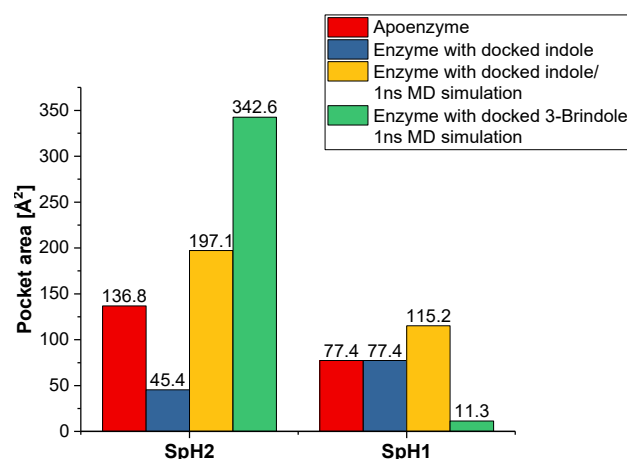
In 2019, the Lewis group identified many flavin-dependent halogenases via family-wide activity profiling, and most of them preferred bromination over chlorination [72]. The group postulated that bromination activity is more widespread than chlorination activity of flavin-dependent halogenases, which matches our data. However, even RebH, which chlorinates its natural substrate tryptophan, was shown in competition experiments to prefer bromination [72]. Therefore, it is also possible that SpH1 and SpH2 chlorinate their natural substrate.

However, the natural substrates of SpH1 and SpH2 are yet unknown. SpH2 accepts larger substrates, such as (3-indolyl)acetic acid (19), 3-(3-indolyl)propionic acid (21), and tryptophol (20). This possible substrate group could be represented by auxins, pivotal plant hormones (also called phytohormones), with (3-indolyl)acetic acid (IAA) (19) as one of the best characterised class of regulators [73,74]. The chlorinated form of auxin, (4-chloro-3-indolyl)acetic acid (4-Cl-IAA), is a highly active hormone that is thought to play a key role in early pericarp growth [75] and fruit development in peas [76]. Auxins also represent a key modulator of plant–bacteria interactions with the bacterial ability to synthesise, e.g., IAA being an attribute for both promoting plant growth and phytopathogenic effects [77]. So far, however, there is no evidence in the literature about brominated auxins in nature, although the chemical synthetic bromination of (3-indolyl)acetic acid has been reported [78]. Another interesting hint towards possible natural substrates is represented by the gene located downstream next to *sph1* (Hit 21; J), which encodes for a catecholate siderophore receptor (Figure 2). Since the MAG similar to *Sphingomonas* also contains at least one metabolic gene cluster potentially responsible for siderophore synthesis (Supplementary Materials), siderophores may represent another potential substrate for the halogenases SpH1 and SpH2. Siderophores possess phenol moieties, which might be halogenated by SpH1 and SpH2, since it was shown that these are able to halogenate 4-*n*-hexylresorcinol (18), phenol (24), and anthranilic acid (25) (Figure 4). In addition to iron scavenging, siderophores can have many alternative functions, including non-iron-metal transport, signaling, and antibiotic activity functions [79]. The halogenation of siderophores is common in nature; for instance, two *Streptomyces* siderophores, chlorocatechelins A and B, have been reported to contain chlorinated catecholate groups [80].

### 2.2.2. Investigation of Mono- and Dibromination Activity

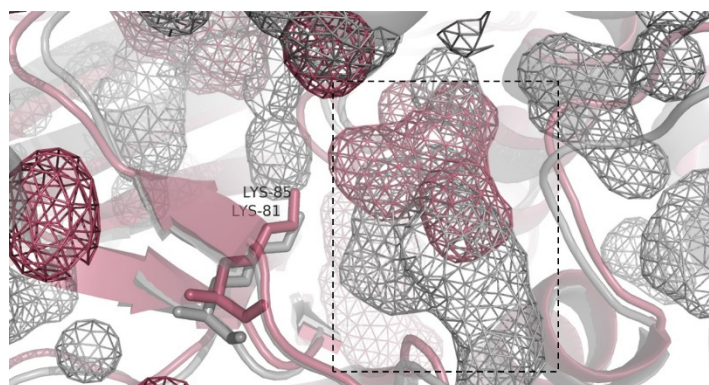
Interestingly, SpH1 only catalyses monobromination, while SpH2 is able to catalyse both mono- and dibromination of indole and its derivatives (Figure 4). Kinetic measurements showed that SpH2 has a negligibly higher specific activity towards indole ( $4.2 \pm 0.03$  mU/mg) than towards 3-bromoindole ( $3.8 \pm 0.45$  mU/mg). In comparison, the specific activity of SpH1 towards indole is  $3.0 \pm 0.7$  mU/mg. The synergistic cooperation of two enzymes in the halogenation process is known in the literature [17,19,46]. PrnA and PrnC have been identified as catalysing the chlorination of tryptophan and monodechloroaminopyrrolnitrin in the biosynthesis cluster of pyrrolnitrin. PrnC catalyses the chlorination of monodechloroaminopyrrolnitrin, halogenated by PrnA in the first step [17,28]. Another example is the biosynthesis of kutzneride, which includes KtzQ and KtzR [19].

SpH1 and SpH2 were evaluated in silico for unveiling the dibrominating ability exhibited by SpH2. Since no crystal structure was available for either of the two enzymes, the homology model was built using YASARA structure ([www.yasara.org](http://www.yasara.org); 2020) for in silico study of the novel halogenases. SpH2 exhibited a wider active site pocket with an area of 136.8 Å<sup>2</sup>, compared to SpH1 with 77.39 Å<sup>2</sup> (Figure 5 and Supplementary Materials).



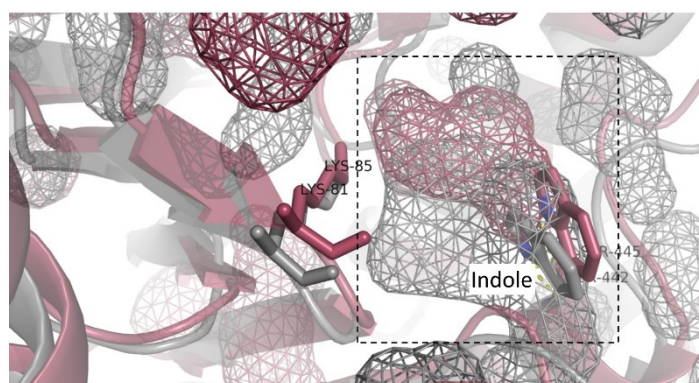
**Figure 5.** Active site pocket area for the apo-, indole-, and 3-bromoindole-docked SpH2 and SpH1 halogenases, showing the flexibility of SpH2 in accommodating 3-bromoindole, unlike SpH1.

The active site pockets of both halogenases differ in their geometry, as does their depth within the pocket for reaching the conserved amino acid residues (Lys and Glu). The SpH1 active site pocket is deeper compared to SpH2, with the loop containing the conserved glutamic acid residue (Glu348) being displaced backwards. On the other hand, the SpH2 active site pocket is wider and not as deep as that of SpH1; still, the pocket is extended to allow the substrate to reach the conserved Lys residue (Figures 6 and 7, and Supplementary Materials).



**Figure 6.** Superimposed homology model of SpH2 (grey) and SpH1 (red) after 1ns molecular dynamics (MD) simulation, showing the active site cavity in wireframe (in dashed-line box). The conserved active site lysine residues 81 and 85 in SpH2 and SpH1, respectively, are shown.

Indole was docked into both models to indicate its flexibility and expansion within the active site pocket of SpH2. The docked substrates were then subjected to MD simulation to further understand the changes that might occur as a result of indole binding to the active site. Obvious differences were detected in the size of the active site pockets containing indole after MD simulation, where the SpH2 active site was expanded to allow the binding of bulkier substrates, thus showing that the active site is flexible. However, SpH1 was not largely expanded, indicating less flexibility to accommodate larger substrates. Both cases showed H-bond formation with the indole NH group (with Ser445 in SpH1 and Ser442 in SpH2; Figures 5 and 7, and Supplementary Materials). Furthermore, the binding affinity of the docked indole towards both enzymes revealed a slight preference for SpH2 over SpH1 (Table 2), which corresponds to the specific activity calculated for SpH2 towards indole compared to SpH1.

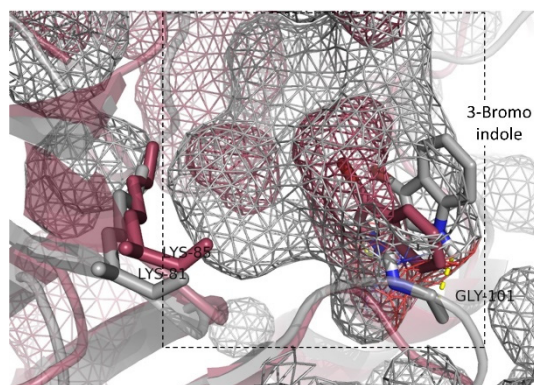


**Figure 7.** Superimposed homology model of SpH2 (grey) and SpH1 (red) with the docked indole after 1 ns MD simulation. The active site cavity is shown in wireframe (surrounded by dashed-line box). In both cases, an H-bond formed between indole NH and Ser442 and Ser445 in SpH2 and SpH1, respectively.

**Table 2.** Binding affinity of indole and 3-bromoindole to the active site of SpH2 and SpH1.

Enzyme/Substrate	Mode	Binding Affinity [kJ/mol]
SpH2/indole	3	22.9
SpH1/indole	9	20.1
SpH2/3-bromoindole	5	22.5
SpH1/3-bromoindole	11	19.4

In addition, 3-bromoindole was docked into the homology models. Higher affinity towards 3-bromoindole was detected for SpH2 over SpH1, which provides further evidence for the dibromination ability of SpH2 (Table 2). In addition, the size of the active site pocket in the case of SpH2 was significantly expanded compared to the active site pocket of SpH1 (Figure 5 and Figure SI9), which is more rigid, and hence would prohibit further bromination of the 3-bromoindole (Figure 8).



**Figure 8.** Superimposed homology model of SpH2 (grey) and SpH1 (red) with the docked 3-bromoindole after 1 ns MD simulation. The active site cavity is shown in wireframe (surrounded by a dashed-line box). In both cases, an H-bond is formed between indole NH and Gly442 in SpH2 and Ser445 in SpH1.

The *in silico* simulation of the SpH1 and SpH2 homology models underlines our findings that SpH2 is able to catalyse dihalogenation. Both SpH1 and SpH2 originate from one gene cluster, and therefore could act in concert and interact with each other. SpH1 has a deeper but less flexible active site pocket, which is likely to be suitable for the binding of smaller substrates like indole, while SpH2 features a wider and more flexible active site

pocket, likely accommodating bulkier substrates, as well as the dibromination observed with some substrates.

### 3. Materials and Methods

Solvents and chemicals were, unless otherwise specified, of analytical grade (p.a.) and purchased from commercial suppliers. Chemicals such as *iso*-propanol, acetonitrile, sodium chloride/bromide, ethanol, and methanol were purchased from VWR Chemicals (Darmstadt, Germany) and L-arabinose, NADH/NAD, FAD and glutaraldehyde from Carl Roth (Karlsruhe, Germany). Antibiotics were obtained from Fisher Scientific (Waltham, MA, USA; ampicillin sodium salt), AppliChem (Darmstadt, Germany; chloramphenicol) and Carl Roth (Karlsruhe, Germany; kanamycin sulfate).

Chemicals for substrate assays were acquired at Acros Organics (Fair Lawn, NJ, USA; indole, indazole, 5-cyanoindole, 3-(3-indolyl)propionic acid, indole-3-carbaldehyde, tryptophan, tryptamine, tryptophol), Sigma Aldrich (Darmstadt, Germany; 2-methylindole, 4-hydroxybenzoic acid, 5-aminoindole, 5-bromoindole, indole-2-carboxylic acid, phenylalanine, phenol, 7H-pyrido[2,3-d]pyrimidine, quinoxaline), Alfa Aesar (Kandel, Germany; 2,3-dimethylindole, 3-methylindole, 4-*n*-hexylresorcinol, 5-hydroxyindole, 5-fluoroindole, 5-methylindole, azulene, indole-3-acetonitrile, (3-indolyl)acetic acid, indole-5-carbaldehyde, indole-5-carboxylic acid, 5-hydroxytryptophan hydrate, naphthylacetic acid), Maybridge (Altrincham, UK; 5-nitroindole), Carl Roth (Karlsruhe, Germany; 7-azaindole) and Fluorochem (Derbyshire, UK; benzoxazole).

The plasmid pCIBhis-prnF encoding the flavin reductase PrnF from *Pseudomonas fluorescens*, was donated by Prof. Dr. Karl-Heinz van Pée, Technical University Dresden, and the plasmid vector pET-21-adh encoding the alcohol dehydrogenase from *Rhodococcus ruber* (RR-ADH) was kindly provided by Prof. Dr. Werner Hummel, Bielefeld University. The plasmid pGro7 coding for the chaperone system GroEL–GroES was purchased from TaKaRa Bio Inc. *E. coli* DH5 and *E. coli* BL21 (DE3) were purchased from Novagen.

#### 3.1. Analytics

##### 3.1.1. Analytical, Reverse-Phase High-Performance Liquid Chromatography (RP-HPLC)

For analytical, reverse-phase high performance liquid chromatography, two different methods were employed.

With method 1, reactions were monitored using a Thermo Scientific (Waltham, MA, USA) Accela 600 with NUCLEOSHELL RP 18 column 18.5  $\mu$ m from Macherey-Nagel (Düren, Germany; 150  $\times$  2.1 mm, eluent A: H<sub>2</sub>O/CH<sub>3</sub>CN/TFA = 95.0:5.0:0.1; eluent B: H<sub>2</sub>O/CH<sub>3</sub>CN/TFA = 5.0:95.0:0.1, flow rate 900  $\mu$ L/min; 0–1 min 100% eluent A, 1–6 min linear gradient to 100% eluent B, 6–7 min 0% eluent A, 7–7.5 min back to 100% eluent A, 7.5–9 min 100% eluent A).

With method 2 for analytical RP-HPLC analysis a Shimadzu (Duisburg, Germany) Nexera XR chromatography system was employed (Luna C18(2) column (3  $\mu$ M, 100 Å, LC column, 100  $\times$  2 mm) from Phenomenex (Aschaffenburg, Germany; eluent A: H<sub>2</sub>O/CH<sub>3</sub>CN/TFA = 95.0:5.0:0.1; eluent B: H<sub>2</sub>O/CH<sub>3</sub>CN/TFA = 5.0:95.0:0.1; flow rate 650  $\mu$ L/min; 0–5.5 min linear gradient to 95% eluent B, 5.5–6.0 min 95% eluent B, 6.0–6.1 min back to 95% eluent A, 6.1–9.0 min 95% eluent A).

##### 3.1.2. Preparative, Reverse-Phase High-Performance Liquid Chromatography (RP-HPLC)

For the isolation and purification of products, a Merck-Hitachi (Darmstadt, Germany) LaChrom RP-HPLC system was employed. Separation of the products was conducted using a Hypersil Gold C18 column (8  $\mu$ M, 250  $\times$  21.2 mm) from Thermo Scientific (Waltham, MA, USA). Absorbance was measured simultaneously at 220 nm, 254 nm, and 280 nm. The setup was as follows: eluent A, H<sub>2</sub>O/CH<sub>3</sub>CN/TFA = 95.0:5.0:0.1; eluent B, H<sub>2</sub>O/CH<sub>3</sub>CN/TFA = 5.0:95.0:0.1; flow rate = 10 mL/min; 0–45 min linear gradient to 100% eluent B; 45–50 min 100% eluent B; 50–55 min 100% eluent A.

### 3.1.3. High-Performance Liquid Chromatography–Mass Spectrometry (HPLC-MS)

For the identification of products by electrospray ionisation, a high-performance liquid chromatography–mass spectrometry (HPLC-MS) system consisting of an Agilent 1200 HPLC system and a 6220 TOF mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) was used, with the following setup: solvent A, H<sub>2</sub>O/CH<sub>3</sub>CN/FA = 95.0:5.0:0.1; solvent B, 5.0:95.0:0.1; flow rate = 0.3 mL/min, with a linear gradient from 0% to 98% B over 10 min, 1 min at 98% B, and back to 0% B for 5 min.

### 3.1.4. Gas Chromatography–Mass Spectrometry (GC-MS)

For gas chromatography–mass spectrometry (GC-MS) analysis, a system consisting of a Trace GC Ultra gas chromatograph (Thermo Scientific, Waltham, MA, USA) and an ITG900 mass spectrometer from Thermo Finnigan (20 measurements per minute,  $50 \pm 750$  *m/z*) was applied. Separation of the products was conducted using a VF-5 column (0.25  $\mu$ M, 30 m  $\times$  0.25 mm, 5% diphenylsiloxan, 95% dimethylsiloxan) from Thermo Scientific (Waltham, MA, USA). Helium was used as the mobile phase, with a temperature gradient of 5 °C/min from 80 °C to 325 °C.

### 3.1.5. Nuclear Magnetic Resonance (NMR) Spectroscopy

NMR spectra were recorded on an Avance 500 (<sup>1</sup>H: 500 MHz, <sup>13</sup>C: 126 MHz) or a DRX-500 spectrometer (<sup>1</sup>H: 500 MHz, <sup>13</sup>C: 126 MHz) (Bruker, Billerica, MA, USA). Chemical shifts are reported relative to residual solvent peaks (DMSO-*d*<sub>6</sub>: <sup>1</sup>H: 2.5 ppm; <sup>13</sup>C: 39.5 ppm).

## 3.2. Bioinformatic Analysis

### 3.2.1. Metagenomic Analysis for the Detection of Flavin-Dependent Halogenases

Metagenomic datasets of *Botryococcus braunii* consortia, deposited under the BioProjectID PRJEB26344 [59], were used for the search of novel flavin-dependent halogenases. The analyses were performed according to the previously described pHMM strategy [33], which is based on the PFAM (<http://pfam.xfam.org>, 2016) tryptophan–halogenase model (Trp\_halogenase, PF04820). The dataset of the MAG21 (similar to the genus *Sphingomonas*) can be found under the deposited BioProject ID PRJEB26345 [59], and was used for the antiSMASH (<https://antismash.secondarymetabolites.org>, 2020) analysis (using default settings) [63]. Amino acid sequence alignments were performed using NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov), accessed on 8 April 2021, Bethesda, MD, USA, 2020) NR [81] and MultAlin [82]. The assignment of the active sites of the enzymes was accomplished on the basis of the published information for RebH and PrnA [29,32,36].

### 3.2.2. Phylogenetic Analysis

Phylogenetic analyses were performed with MEGA7 ([www.megasoftware.net](http://www.megasoftware.net), accessed on 8 April 2021) [60], based on protein alignment (MUSCLE) [83], neighbour joining (NJ) method, and a bootstrap of 1000 [61,62].

## 3.3. In Vitro Enzyme Assays

### 3.3.1. Vector Construction and Molecular Cloning

Gene synthesis of the putative halogenase genes *spH1* and *spH2* was ordered from ThermoFisher GeneArt (Waltham, MA, USA). The genes pMA-T-spH1 and pMA-T-spH2 were codon-optimised for *E. coli* and provided with restriction sites (BamHI and NdeI (New England BioLabs, Frankfurt am Main, Germany)). The synthetic genes were cloned into pET28a expression vectors and transformed in *E. coli* BL21 (DE3) pGro7.

### 3.3.2. Heterologous Protein Expression and Purification

A total of 1.5 L of *Luria-Bertani* medium, containing 60  $\mu$ g/mL kanamycin and 50  $\mu$ g/mL chloramphenicol, was inoculated with 2% of an overnight culture of *E. coli* BL21 (DE3) pGro7 pET28a-spH1 or pET28a-spH2, and incubated at 37 °C until reaching



an optical density ( $OD_{600}$ ) of 0.4. The temperature was decreased to 25 °C for 30 min, and overexpression was induced by the addition of isopropyl- $\beta$ -D-thiogalactopyranoside (0.1 mM) and L-arabinose (2 g/L). After cultivation for a further 22 h at 150 rpm, cells were harvested by centrifugation ( $4200 \times g$ , 1 h, 4 °C), washed with 100 mM  $Na_2HPO_4$  buffer (pH 7.4), and stored at −20 °C.

Expression of alcohol dehydrogenase RR-ADH and flavin reductase PrnF was performed as previously described [3].

The cells from 1.5 L cultivation were resuspended in 30 mL of 100 mM  $Na_2HPO_4$  buffer (pH 7.4) and lysed by French Press (three times, 1000 psig). Cell debris was removed by centrifugation ( $10,000 \times g$ , 30 min, 4 °C), and soluble protein was filtered through a 0.2  $\mu$ M Whatman filter. The purification of the hexahistidin-tagged proteins was performed via a HisTALON matrix. The lysate was loaded on HisTALON agarose affinity resin (TaKaRa Bio Inc., Saint-Germain-en-Laye, France) and washed with  $10 \times$  column volume (CV) with 100 mM  $Na_2HPO_4$  buffer (pH 7.4), and  $10 \times$  CV 50 mM  $Na_2HPO_4$  buffer (pH 7.4) with 300 mM NaBr and 10 mM imidazole. The proteins were eluted with 50 mM  $Na_2HPO_4$  buffer (pH 7.4), 300 mM NaBr, and 300 mM imidazole and collected in 0.75 mL fractions. After determination of the protein concentration by NanoDrop UV spectroscopy (Thermo Scientific, Waltham, MA, USA), fractions with the purified proteins were pooled and desalted via a HiTrap Desalting column, with 50 mM  $Na_2HPO_4$  buffer (pH 7.4) and 50 mM NaBr.

### 3.3.3. Enzymatic Halogenation/Enzyme Assay with Purified Protein on an Analytical Scale

The activities of PrnF and ADH were determined as published by Frese et al. [3]. The halogenation activities of SpH1 and SpH2 towards different substrates were carried out for 48 h at 25 °C and 500 rpm in a total volume of 0.5 mL, containing 1.25 mg/mL enzyme, 1 mM substrate, 1  $\mu$ M FAD, 100  $\mu$ M NAD, 100 mM NaBr/NaCl, RR-ADH (2 U/mL) PrnF (2.5 U/mL), 5% (*v/v*) *iso*-propanol, and 100 mM  $Na_2HPO_4$  buffer (pH 7.4) [33,52]. Samples were quenched by adding an equal volume of methanol. Reactions were analysed by analytical RP-HPLC and HPLC-MS.

### 3.3.4. Determination of Specific Activity

Halogenation reactions, as described in Section 3.3.3., were carried out with 0.05 mM substrate and 10  $\mu$ M enzyme. Samples were taken at 5 min intervals over a 20 min time period, and then the enzyme reaction was stopped with 1:1 methanol/water (*v/v*). The determination was carried out in triplicate, reactions were analysed by analytical RP-HPLC, and conversion rates (*v*) were calculated. With this in hand, the specific activity can be calculated using the following:

$$\text{specific activity} \left[ \frac{mU}{mg} \right] = \frac{v}{c_E \cdot M_E} \times 10^9 \quad (1)$$

where  $c_E$  is the enzyme concentration (10  $\mu$ M) and  $M_E$  is the enzyme mass (g/mol).

### 3.3.5. Catalase Activity Assay

Catalase from bovine liver lyophilised powder (2000–5000 units/mg protein; Sigma Aldrich, Darmstadt, Germany) at a final concentration of 100 U/mL was additionally added to the reaction mixture, and the procedure continued as described above to determine the FAD dependency of SpH1 and SpH2 [33].

### 3.3.6. Enzymatic Halogenation with Immobilised Protein on a Preparative Scale

For the isolation and analysis of halogenated products on a preparative scale, cross-linked enzyme aggregates (CLEAs), as described by Frese et al. [1], were used.

The *E. coli* BL21 (DE3) pGro7 cells from a 1.5 L cultivation containing SpH1 and SpH2 were resuspended in 30 mL 100 mM  $Na_2HPO_4$  buffer (pH 7.4) and lysed by French Press

(three times, 1000 psig). After centrifugation (30 min,  $10,000\times g$ ,  $4\text{ }^{\circ}\text{C}$ ), the cell lysate was divided in two equal parts, and PrnF (2.5 U/mL) and RR-ADH (1 U/mL) were added. Precipitation of the proteins was carried out by adding ammonium sulfate (8.1 g, 95% saturation), followed by incubation for 1 h at  $4\text{ }^{\circ}\text{C}$  in a tube rotator. Glutaraldehyde (1.5 mL, 0.5% (*w/v*)) was added for cross-linking, and the mixture was further kept for 2 h at  $4\text{ }^{\circ}\text{C}$ . The CLEAs were centrifuged and washed three times with 30 mL of 100 mM  $\text{Na}_2\text{HPO}_4$  buffer (pH 7.4). The pH of the reaction buffer for the biocatalysis with CLEAs containing, 1.5 mM substrate, 1  $\mu\text{M}$  FAD, 100  $\mu\text{M}$  NAD, 30 mM NaBr/NaCl, and 5% (*v/v*) and 15 mM  $\text{Na}_2\text{HPO}_4$ , was adjusted with phosphoric acid to pH 7.4. Finally, CLEAs were added to reaction buffer in a final volume of 500 mL, and incubated for up to 10 days (until no further conversion was visible at the analytical RP-HPLC) at  $25\text{ }^{\circ}\text{C}$  and 150 rpm. The reaction mixture was filtered and extracted three times with 250 mL dichloromethane. The organic phase was dried over magnesium sulfate and evaporated. Before the characterisation of the halogenated products with HPLC-MS, analytical HPLC, and NMR, the remaining substance was purified over preparative RP-HPLC and lyophilised.

### 3.3.7. FAD Reconstitution

For FAD reconstitution, the buffer of the purified proteins was exchanged to FAD reconstitution buffer (50 mM  $\text{Na}_2\text{HPO}_4$ , 150 mM NaBr, 1 mM FAD, pH 7.4) using a HiTrap-Desalting column and incubated overnight at  $4\text{ }^{\circ}\text{C}$  on a rocking shaker. The proteins were washed three times with 50 mM  $\text{Na}_2\text{HPO}_4$  (pH 7.4) and 150 mM NaBr to remove free FAD, and concentrated to a final volume of 400  $\mu\text{L}$  in an Amicon Ultra-4 centrifugal filter device with a 50 kDa cutoff. To determine the degree of occupation, UV/Vis spectra of the proteins were measured in a quartz cuvette (Suprasil, Hellma, Müllheim, Germany) with a path length of 1 cm, using a Shimadzu (Duisburg, Germany) UV-2450 spectrometer [33].

### 3.4. In Silico Study Using YASARA

The homology models of SpH1 and SpH2 were built using YASARA structure ([www.yasara.org](http://www.yasara.org), 2020) [84]. The built models were further energy minimised before MD simulation and the docking study [85]. The structures of indole and 3-bromoindole were built with YASARA structure and energy-minimised using YASARA 2 force field before using in the docking experiments.

MD simulations were run for the homology models of both enzymes for 1 ns using YASARA structure [86]. The simulation cell was extended by 10 Å around the whole model and filled with randomly-oriented water molecules with a density of 0.99 g/mL. The force field AMBER14, with 10 Å for cut off and particle mesh Ewald (PME) for long-range electrostatic, was used for running the simulation. The stability of the homology models was monitored by the root mean square deviation (RMSD) of the alpha carbon of the models throughout the simulation time.

Docking of either indole or 3-bromoindole was performed by YASARA structure, using Autodock for the MD simulated models of SpH1 and SpH2 [87]. The simulation cell was defined around the whole homology model of both enzymes. Docking results were analysed based on the binding energy (B-factor) calculated by YASARA structure and the interaction of the substrate with the amino acid residues of the active site pocket.

The docked models with indole or 3-bromoindole were further subjected to MD simulation using the previously mentioned parameters.

The size of the active site pocket cavity was determined using the Computed Atlas of Surface Topography of proteins (CASTp) (<http://sts.bioe.uic.edu/castp/>, 2020) for the free and docked homology models before and after MD simulation [88].

## 4. Conclusions

Flavin-dependent halogenases are the focus of research because of their enormous potential for the regioselective halogenation of organic compounds, leading to products with profound biological activity. In the present study, using our hidden Markov model

(pHMM), we screened bacterial associates of the *B. braunii* consortia (PRJEB21978) and identified 25 complete and partial putative, flavin-dependent halogenase genes. Interestingly, all newly identified genes were found to form a distinct clade, together with previously characterised BrvH, potentially indicating a similar substrate preference. Two selected flavin-dependent halogenases (SpH1 and SpH2), derived from one gene cluster of *Sphingomonas* sp., were subjected to a screening with different compounds, and 26 substrates were found to become halogenated (Figure 4). In accordance with recent findings, as well as with the activity performance of BrvH, both enzymes were observed to prefer bromination over chlorination. Notably, SpH2 is capable of dibrominating substrates, while SpH1 can only monobrominate. Moreover, the simulation in silico using YASARA homology models revealed that SpH1 possesses a deeper but less flexible active site pocket, while SpH2 features a wider and more flexible active site pocket. This explains the ability of SpH2 to halogenate bulkier substrates, as well as catalyse dibromination. Both were identified in one gene, which suggests that the newly identified flavin-dependent halogenases could act in concert and interact with one another. Likewise, other known flavin-dependent halogenases originating from one gene cluster, such as the correlates PrnA/PrnC, Bmp2/Bmp5, or KtzR/KtzQ, act in concert for the biosynthesis of multiple halogenated compounds. SpH1 and SpH2 accepted, in total, 26 different indole and phenol derivatives. Interestingly, both accepted 4-*n*-hexylresorcinol (18), tryptophol (20), and 3-(3-indolyl)propionic acid (21). However, the natural substrates of SpH1 and SpH2 remain unknown so far. In the gene cluster of both enzymes, a catecholate siderophore receptor was detected, which might give a hint towards their substrate. This study shows that screening with the HMM may discover many novel flavin-dependent halogenases in natural metagenomes. The algorithm detects mostly flavin-dependent halogenases from a clade similar to BrvH. The phylogenetic tree demonstrates that the detected novel, putative, flavin-dependent halogenases cluster within the clade of BrvH.

**Supplementary Materials:** Further data is available online at <https://www.mdpi.com/article/10.3390/catal11040485/s1>. Amino acid sequences of the novel identified F-Hals SpH1 and SpH2. SDS-PAGE with heterologously expressed SpH1 and SpH2; RP-HPLC (280 nm) of the enzymatic halogenation of SpH1 (A) and SpH2 (B) of indole to bromoindole in the presence of catalase. Absorption bands of FAD bound in SpH1 and SpH2 during FAD reconstitution. HPLC trace of the enzyme reaction with SpH1\_K81A and indole. Characterization of the brominated substrates by SpH1 and SpH2 via mass spectrometry. Determination of regioselective halogenation site via NMR analysis. List of detected secondary metabolite biosynthetic gene clusters within the MAG 21. Area and volume of the active site pocket for Apo SpH2 and SpH1, docked with indole and 3-bromoindole before and after MD simulation.

**Author Contributions:** Conceptualization, P.R.N., O.B.-K. and N.S.; data curation, P.R.N.; formal analysis, P.R.N.; funding acquisition, O.K. and N.S.; investigation, P.R.N., O.B.-K., L.P. and M.I.; project administration, O.K. and N.S.; resources, O.K. and N.S.; supervision, O.K. and N.S.; validation, P.R.N., O.B.-K., L.P. and M.I.; visualization, M.I.; writing—original draft, P.R.N.; writing—review and editing, P.R.N., O.B.-K., L.P., M.I. and N.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partly funded by the Deutsche Forschungsgemeinschaft (SFB 1416/1-2020) and European Community's Seventh Programme for research, technological development, and demonstration under grant agreement no. FP7-311956 (relating to the project "SPLASH—Sustainable PoLymerS from Algae Sugars and Hydrocarbons"). The authors acknowledge the financial support of the German Research Foundation (DFG) and the Open Access Publication Fund of Bielefeld University for the article processing charge.

**Acknowledgments:** The authors thank Karl-Heinz van Pée, TU Dresden, for donation of the plasmid pClBhis-prnF encoding the flavin reductase PrnF from *Pseudomonas fluorescens*, and Prof. Dr. Werner Hummel, Bielefeld University, for the plasmid vector pET-21 ADH encoding the alcohol dehydrogenase from *Rhodococcus ruber* (RR-ADH).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Frese, M.; Sewald, N. Enzymatic halogenation of tryptophan on a gram scale. *Angew. Chem. Int. Ed.* **2015**, *54*, 298–301. [\[CrossRef\]](#)
2. Schnepel, C.; Sewald, N. Enzymatic halogenation: A timely strategy for regioselective C–H activation. *Chem. Eur. J.* **2017**, *23*, 12064–12086. [\[CrossRef\]](#)
3. Frese, M.; Guzowska, P.H.; Voss, H.; Sewald, N. Regioselective Enzymatic Halogenation of Substituted Tryptophan Derivatives using the FAD-Dependent Halogenase RebH. *ChemCatChem* **2014**, *6*, 1270–1276.
4. Frese, M.; Schnepel, C.; Minges, H.; Voß, H.; Feiner, R.; Sewald, N. Modular Combination of Enzymatic Halogenation of Tryptophan with Suzuki-Miyaura Cross-Coupling Reactions. *ChemCatChem* **2016**, *8*, 1799–1803. [\[CrossRef\]](#)
5. Dachwitz, S.; Widmann, C.; Frese, M.; Niemann, H.H.; Sewald, N. Enzymatic halogenation: Enzyme mining, mechanisms, and implementation in reaction cascades. *Amino Acids Pept. Proteins* **2021**, *44*, 1.
6. Minges, H.; Sewald, N. Recent Advances in Synthetic Application and Engineering of Halogenases. *ChemCatChem* **2020**, *12*, 4450–4470. [\[CrossRef\]](#)
7. Mascotti, M.L.; Ayub, M.J.; Furnham, N.; Thornton, J.M.; Laskowski, R.A. Chopping and changing: The evolution of the flavin-dependent monooxygenases. *J. Mol. Biol.* **2016**, *428*, 3131–3146. [\[CrossRef\]](#)
8. Badiyan, S.; Bach, R.D.; Sobrado, P. Mechanism of N-hydroxylation catalyzed by flavin-dependent monooxygenases. *J. Org. Chem.* **2015**, *80*, 2139–2147. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Kemker, I.; Schnepel, C.; Schröder, D.C.; Marion, A.; Sewald, N. Cyclization of RGD Peptides by Suzuki-Miyaura Cross-Coupling. *J. Med. Chem.* **2019**, *62*, 7417–7430. [\[CrossRef\]](#)
10. Durak, L.J.; Payne, J.T.; Lewis, J.C. Late-stage diversification of biologically active molecules via chemoenzymatic C–H functionalization. *ACS Catal.* **2016**, *6*, 1451–1454. [\[CrossRef\]](#)
11. Latham, J.; Henry, J.-M.; Sharif, H.H.; Menon, B.R.K.; Shepherd, S.A.; Greaney, M.F.; Micklefield, J. Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C–H activation. *Nat. Commun.* **2016**, *7*, 1–8. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Kemker, I.; Schröder, D.C.; Feiner, R.C.; Müller, K.M.; Marion, A.; Sewald, N. Tuning the Biological Activity of RGD Peptides with Halotryptophans. *J. Med. Chem.* **2021**, *64*, 586–601. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Corr, M.J.; Sharma, S.V.; Pubill-Ulldemolins, C.; Bown, R.T.; Poirot, P.; Smith, D.R.M.; Cartmell, C.; Abou-Fayad, A.; Goss, R.J.M. Sonogashira diversification of unprotected halotryptophans, halotryptophan containing tripeptides; and generation of a new to nature bromo-natural product and its diversification in water. *Chem. Sci.* **2017**, *8*, 2039–2046. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Größ, H.; Sewald, N. Late-Stage Diversification of Tryptophan-Derived Biomolecules. *Chem. Eur. J.* **2020**, *26*, 5328–5340. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Pubill-Ulldemolins, C.; Sharma, S.V.; Cartmell, C.; Zhao, J.; Cárdenas, P.; Goss, R.J.M. Heck Diversification of Indole-Based Substrates under Aqueous Conditions: From Indoles to Unprotected Halo-tryptophans and Halo-tryptophans in Natural Product Derivatives. *Chem. Eur. J.* **2019**, *25*, 10866–10875. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Keller, S.; Wage, T.; Hohaus, K.; Hölzer, M.; Eichhorn, E.; van Pée, K.-H. Purification and partial characterization of tryptophan 7-halogenase (PrnA) from *Pseudomonas fluorescens*. *Angew. Chem. Int. Ed.* **2000**, *39*, 2300–2302. [\[CrossRef\]](#)
17. Kirner, S.; Hammer, P.E.; Hill, D.S.; Altmann, A.; Fischer, I.; Weislo, L.J.; Lanahan, M.; van Pée, K.-H.; Ligon, J.M. Functions encoded by pyrrolnitrin biosynthetic genes from *Pseudomonas fluorescens*. *J. Bacteriol. Res.* **1998**, *180*, 1939–1943. [\[CrossRef\]](#)
18. Sánchez, C.; Butovich, I.A.; Braña, A.F.; Rohr, J.; Méndez, C.; Salas, J.A. The biosynthetic gene cluster for the antitumor rebeccamycin: Characterization and generation of indolocarbazole derivatives. *Chem. Biol.* **2002**, *9*, 519–531. [\[CrossRef\]](#)
19. Heemstra, J.R., Jr.; Walsh, C.T. Tandem action of the O<sub>2</sub>- and FADH<sub>2</sub>-dependent halogenases KtzQ and KtzR produce 6, 7-dichlorotryptophan for kutzneride assembly. *J. Am. Chem. Soc.* **2008**, *130*, 14024–14025. [\[CrossRef\]](#)
20. Seibold, C.; Schnerr, H.; Rumpf, J.; Kunzendorf, A.; Hatscher, C.; Wage, T.; Ernyei, A.J.; Dong, C.; Naismith, J.H.; van Pée, K.-H. A flavin-dependent tryptophan 6-halogenase and its use in modification of pyrrolnitrin biosynthesis. *Biocatal. Biotransform.* **2006**, *24*, 401–408. [\[CrossRef\]](#)
21. Zeng, J.; Zhan, J. Characterization of a tryptophan 6-halogenase from *Streptomyces toxytricini*. *Biotechnol. Lett.* **2011**, *33*, 1607–1613. [\[CrossRef\]](#)
22. Xu, L.; Han, T.; Ge, M.; Zhu, L.; Qian, X. Discovery of the new plant growth-regulating compound LYXLF2 based on manipulating the halogenase in *Amycolatopsis orientalis*. *Curr. Microbiol.* **2016**, *73*, 335–340. [\[CrossRef\]](#)
23. Zehner, S.; Kotzsch, A.; Bister, B.; Süßmuth, R.D.; Méndez, C.; Salas, J.A.; van Pée, K.-H. A regioselective tryptophan 5-halogenase is involved in pyrroindomycin biosynthesis in *Streptomyces rugosporus* LL-42D005. *Chem. Biol.* **2005**, *12*, 445–452. [\[CrossRef\]](#)
24. Ryan, K.S. Biosynthetic gene cluster for the cladoniamides, bis-indoles with a rearranged scaffold. *PLoS ONE* **2011**, *6*, e23694. [\[CrossRef\]](#)
25. Moritzer, A.-C.; Minges, H.; Prior, T.; Frese, M.; Sewald, N.; Niemann, H.H. Structure-based switch of regioselectivity in the flavin-dependent tryptophan 6-halogenase Thal. *J. Biol. Chem.* **2019**, *294*, 2529–2542. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Minges, H.; Schnepel, C.; Böttcher, D.; Weiß, M.S.; Sproß, J.; Bornscheuer, U.T.; Sewald, N. Targeted Enzyme Engineering Unveiled Unexpected Patterns of Halogenase Stabilization. *ChemCatChem* **2020**, *12*, 818–831. [\[CrossRef\]](#)
27. Van Pée, K.-H. Enzymatic chlorination and bromination. *Meth. Enzymol.* **2012**, *516*, 237–257.
28. Dong, C.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pée, K.-H.; Naismith, J.H. Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. *Science* **2005**, *309*, 2216–2219. [\[CrossRef\]](#)



29. Zhu, X.; de Laurentis, W.; Leang, K.; Herrmann, J.; Ihlefeld, K.; Van Pée, K.-H.; Naismith, J.H. Structural insights into regioselectivity in the enzymatic chlorination of tryptophan. *J. Mol. Biol.* **2009**, *391*, 74–85. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Dong, C.; Kotzsch, A.; Dorward, M.; van Pée, K.-H.; Naismith, J.H. Crystallization and X-ray diffraction of a halogenating enzyme, tryptophan 7-halogenase, from *Pseudomonas fluorescens*. *Acta Cryst.* **2004**, *60*, 1438–1440.
31. Ainsley, J.; Mulholland, A.J.; Black, G.W.; Sparagano, O.; Christov, C.Z.; Karabencheva-Christova, T.G. Structural insights from molecular dynamics simulations of tryptophan 7-halogenase and tryptophan 5-halogenase. *ACS Omega* **2018**, *3*, 4847–4859. [\[CrossRef\]](#)
32. Yeh, E.; Blasiak, L.C.; Koglin, A.; Drennan, C.L.; Walsh, C.T. Chlorination by a long-lived intermediate in the mechanism of flavin-dependent halogenases. *Biochemistry* **2007**, *46*, 1284–1292. [\[CrossRef\]](#)
33. Neubauer, P.R.; Widmann, C.; Wibberg, D.; Schröder, L.; Frese, M.; Kottke, T.; Kalinowski, J.; Niemann, H.H.; Sewald, N. A flavin-dependent halogenase from metagenomic analysis prefers bromination over chlorination. *PLoS ONE* **2018**, *13*, e0196797. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Menon, B.R.K.; Brandenburger, E.; Sharif, H.H.; Klemstein, U.; Shepherd, S.A.; Greaney, M.F.; Micklefield, J. RadH: A Versatile Halogenase for Integration into Synthetic Pathways. *Angew. Chem. Int. Ed.* **2017**, *56*, 11841–11845. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Flecks, S.; Patallo, E.P.; Zhu, X.; Ernyei, A.J.; Seifert, G.; Schneider, A.; Dong, C.; Naismith, J.H.; van Pée, K.-H. New insights into the mechanism of enzymatic chlorination of tryptophan. *Angew. Chem. Int. Ed.* **2008**, *47*, 9533–9536. [\[CrossRef\]](#)
36. Bitto, E.; Huang, Y.; Bingman, C.A.; Singh, S.; Thorson, J.S.; Phillips, G.N. The structure of flavin-dependent tryptophan 7-halogenase RebH. *Proteins* **2008**, *70*, 289–293. [\[CrossRef\]](#)
37. Podzelinska, K.; Latimer, R.; Bhattacharya, A.; Vining, L.C.; Zechel, D.L.; Jia, Z. Chloramphenicol biosynthesis: The structure of CmlS, a flavin-dependent halogenase showing a covalent flavin-aspartate bond. *J. Mol. Biol.* **2010**, *397*, 316–331. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Roy, A.D.; Grischow, S.; Cairns, N.; Goss, R.J.M. Gene expression enabling synthetic diversification of natural products: Chemogenetic generation of pacidamycin analogs. *J. Am. Chem. Soc.* **2010**, *132*, 12243–12245. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Glenn, W.S.; Nims, E.; O'Connor, S.E. Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor. *J. Am. Chem. Soc.* **2011**, *133*, 19346–19349. [\[CrossRef\]](#)
40. Rungtaphan, W.; Qu, X.; O'Connor, S.E. Integrating carbon–halogen bond formation into medicinal plant metabolism. *Nature* **2010**, *468*, 461–464. [\[CrossRef\]](#)
41. Veldmann, K.H.; Dachwitz, S.; Risse, J.M.; Lee, J.-H.; Sewald, N.; Wendisch, V.F. Bromination of L-tryptophan in a Fermentative Process with *Corynebacterium glutamicum*. *Front. Bioeng. Biotechnol.* **2019**, *7*, 219. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Hammer, P.E.; Hill, D.S.; Lam, S.T.; Van Pée, K.-H.; Ligon, J.M. Four genes from *Pseudomonas fluorescens* that encode the biosynthesis of pyrrolnitrin. *Appl. Environ. Microbiol.* **1997**, *63*, 2147–2154. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Smith, D.R.M.; Uriá, A.R.; Helfrich, E.J.N.; Milbredt, D.; van Pée, K.-H.; Piel, J.; Goss, R.J.M. An unusual flavin-dependent halogenase from the metagenome of the marine sponge *Theonella swinhoei* WA. *ACS Chem. Biol.* **2017**, *12*, 1281–1287. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Menon, B.R.K.; Latham, J.; Dunstan, M.S.; Brandenburger, E.; Klemstein, U.; Leys, D.; Karthikeyan, C.; Greaney, M.F.; Shepherd, S.A.; Micklefield, J. Structure and biocatalytic scope of thermophilic flavin-dependent halogenase and flavin reductase enzymes. *Org. Biomol. Chem.* **2016**, *14*, 9354–9361. [\[CrossRef\]](#)
45. Ortega, M.A.; Cogan, D.P.; Mukherjee, S.; Garg, N.; Li, B.; Thibodeaux, G.N.; Maffioli, S.I.; Donadio, S.; Sosio, M.; Escano, J. Two flavoenzymes catalyze the post-translational generation of 5-chlorotryptophan and 2-aminovinyl-cysteine during NAI-107 biosynthesis. *ACS Chem. Biol.* **2017**, *12*, 548–557. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Agarwal, V.; el Gamal, A.A.; Yamanaka, K.; Poth, D.; Kersten, R.D.; Schorn, M.; Allen, E.E.; Moore, B.S. Biosynthesis of polybrominated aromatic organic compounds by marine bacteria. *Nat. Chem. Biol.* **2014**, *10*, 640. [\[CrossRef\]](#)
47. Zeng, J.; Zhan, J. A Novel Fungal Flavin-Dependent Halogenase for Natural Product Biosynthesis. *ChemBioChem* **2010**, *11*, 2119–2123. [\[CrossRef\]](#)
48. Chooi, Y.-H.; Cacho, R.; Tang, Y. Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem. Biol.* **2010**, *17*, 483–494. [\[CrossRef\]](#)
49. Wang, S.; Xu, Y.; Maine, E.A.; Wijeratne, E.K.; Espinosa-Artiles, P.; Gunatilaka, A.L.; Molnár, I. Functional characterization of the biosynthesis of radicicol, an Hsp90 inhibitor resorcylic acid lactone from *Chaetomium chiversii*. *Chem. Biol.* **2008**, *15*, 1328–1338. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Bayer, K.; Scheuermayer, M.; Fieseler, L.; Hentschel, U. Genomic mining for novel FADH 2-dependent halogenases in marine sponge-associated microbial consortia. *Mar. Biotechnol.* **2013**, *15*, 63–72. [\[CrossRef\]](#)
51. Ismail, M.; Frese, M.; Patschkowski, T.; Ortseifen, V.; Niehaus, K.; Sewald, N. Flavin-Dependent Halogenases from *Xanthomonas campestris* pv. *campestris* B100 Prefer Bromination over Chlorination. *Adv. Synth. Catal.* **2019**, *361*, 2475–2486. [\[CrossRef\]](#)
52. Neubauer, P.R.; Pienkny, S.; Wessjohann, L.A.; Brandt, W.; Sewald, N. Predicting the substrate scope of the flavin-dependent halogenase BrvH. *ChemBioChem* **2020**, *21*, 3282–3288. [\[CrossRef\]](#)
53. Widmann, C.; Ismail, M.; Sewald, N.; Niemann, H.H. Structure of apo flavin-dependent halogenase Xcc4156 hints at a reason for cofactor-soaking difficulties. *Acta Cryst.* **2020**, *76*, 687–697. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Rojo, F. Degradation of alkanes by bacteria. *Environ. Microbiol.* **2009**, *11*, 2477–2490. [\[CrossRef\]](#) [\[PubMed\]](#)



55. Löwe, J.; Blifernez-Klassen, O.; Baier, T.; Wobbe, L.; Kruse, O.; Gröger, H. Type II flavoprotein monooxygenase PsFMO\_A from the bacterium *Pimelobacter* sp. Bb-B catalyzes enantioselective Baeyer-Villiger oxidations with a relaxed cofactor specificity. *J. Biotechnol.* **2019**, *294*, 81–87. [\[CrossRef\]](#)
56. Metzger, P.; Largeau, C. *Botryococcus braunii*: A rich source for hydrocarbons and related ether lipids. *Appl. Microbiol. Biotechnol.* **2005**, *66*, 486–496. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Blifernez-Klassen, O.; Chaudhari, S.; Klassen, V.; Wördenweber, R.; Steffens, T.; Cholewa, D.; Niehaus, K.; Kalinowski, J.; Kruse, O. Metabolic survey of *Botryococcus braunii*: Impact of the physiological state on product formation. *PLoS ONE* **2018**, *13*, e0198976. [\[CrossRef\]](#) [\[PubMed\]](#)
58. Chirac, C.; Casadevall, E.; Largeau, C.; Metzger, P. Bacterial influence upon growth and hydrocarbon production of the green alga *Botryococcus braunii* 1. *J. Phycol.* **1985**, *21*, 380–387. [\[CrossRef\]](#)
59. Blifernez-Klassen, O.; Klassen, V.; Wibberg, D.; Cebeci, E.; Henke, C.; Rückert, C.; Chaudhari, S.; Rupp, O.; Blom, J.; Winkler, A. *Phytoplankton consortia* as a blueprint for mutually beneficial eukaryote-bacteria ecosystems based on the biocoenosis of *Botryococcus consortia*. *Sci. Rep.* **2021**, *11*, 1–13. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [\[CrossRef\]](#)
61. Saitou, N.; Nei, M. The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425.
62. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Blin, K.; Wolf, T.; Chevrete, M.G.; Lu, X.; Schwalen, C.J.; Kautsar, S.A.; Suarez Duran, H.G.; de los Santos, E.L.C.; Kim, H.U.; Nave, M. antiSMASH 4.0—Improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45*, W36–W41. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Cotter, P.D.; Ross, R.P.; Hill, C. Bacteriocins—A viable alternative to antibiotics? *Nat. Rev. Microbiol.* **2013**, *11*, 95–105. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Ligtenbarg, A.G.J.; Hage, R.; Feringa, B.L. Catalytic oxidations by vanadium complexes. *Coord. Chem. Rev.* **2003**, *237*, 89–101. [\[CrossRef\]](#)
66. Crichton, R.R. *Biological Inorganic Chemistry: A New Introduction to Molecular Structure and Function*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2012; ISBN 9780444537829. reprinted.
67. Hager, L.P.; Morris, D.R.; Brown, F.S.; Eberwein, H. Chloroperoxidase: II. Utilization of halogen anions. *J. Biol. Chem.* **1966**, *241*, 1769–1777. [\[CrossRef\]](#)
68. Schroeder, L.; Frese, M.; Müller, C.; Sewald, N.; Kottke, T. Photochemically Driven Biocatalysis of Halogenases for the Green Production of Chlorinated Compounds. *ChemCatChem* **2018**, *10*, 3336–3341. [\[CrossRef\]](#)
69. Sundberg, R.J. Electrophilic Substitution Reactions of Indoles. In *Heterocyclic Scaffolds II: Topics in Heterocyclic Chemistry*; Gribble, G.W.W., Ed.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 47–115. [\[CrossRef\]](#)
70. Hölzer, M.; Burd, W.; Reißig, H.-U.; van Pée, K.-H. Substrate specificity and regioselectivity of tryptophan 7-halogenase from *Pseudomonas fluorescens* BL915. *Adv. Synth. Catal.* **2001**, *343*, 591–595. [\[CrossRef\]](#)
71. Van Pée, K.H. Microbial biosynthesis of halometabolites. *Arch. Microbiol.* **2001**, *175*, 250–258. [\[CrossRef\]](#)
72. Fisher, B.F.; Snodgrass, H.M.; Jones, K.A.; Andorfer, M.C.; Lewis, J.C. Site-Selective C–H Halogenation Using Flavin-Dependent Halogenases Identified via Family-Wide Activity Profiling. *ACS Cent. Sci.* **2019**, *5*, 1844–1856. [\[CrossRef\]](#)
73. Kunkel, B.N.; Harper, C.P. The roles of auxin during interactions between bacterial plant pathogens and their hosts. *J. Exp. Bot.* **2018**, *69*, 245–254. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Woodward, A.W.; Bartel, B. Auxin: Regulation, action, and interaction. *Ann. Bot.* **2005**, *95*, 707–735. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Ozga, J.A.; Reinecke, D.M.; Ayele, B.T.; Ngo, P.; Nadeau, C.; Wickramarathna, A.D. Developmental and hormonal regulation of gibberellin biosynthesis and catabolism in pea fruit. *Plant Physiol.* **2009**, *150*, 448–462. [\[CrossRef\]](#)
76. Tivendale, N.D.; Davidson, S.E.; Davies, N.W.; Smith, J.A.; Dalmais, M.; Bendahmane, A.I.; Quittenden, L.J.; Sutton, L.; Bala, R.K.; Le Signor, C. Biosynthesis of the halogenated auxin, 4-chloroindole-3-acetic acid. *Plant Physiol.* **2012**, *159*, 1055–1063. [\[CrossRef\]](#)
77. Duca, D.; Lorv, J.; Patten, C.L.; Rose, D.; Glick, B.R. Indole-3-acetic acid in plant–microbe interactions. *Antonie Leeuwenhoek* **2014**, *106*, 85–125. [\[CrossRef\]](#)
78. Murphy, K. Applying Green Chemistry Principles in the Electrophilic Bromination of Indole-3-Acetic Acid. *Undergrad. Rev.* **2014**, *10*, 111–115.
79. Kramer, J.; Özkaya, Ö.; Kümmerli, R. Bacterial siderophores in community and host interactions. *Nat. Rev. Microbiol.* **2020**, *18*, 152–163. [\[CrossRef\]](#)
80. Kishimoto, S.; Nishimura, S.; Hattori, A.; Tsujimoto, M.; Hatano, M.; Igarashi, M.; Kakeya, H. Chlorocatechins A and B from *Streptomyces* sp.: New siderophores containing chlorinated catecholate groups and an acylguanidine structure. *Org. Lett.* **2014**, *16*, 6108–6111. [\[CrossRef\]](#) [\[PubMed\]](#)
81. O’Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufu, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [\[CrossRef\]](#)

- 
82. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881–10890. [[CrossRef](#)] [[PubMed](#)]
  83. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)]
  84. Krieger, E.; Vriend, G. YASARA View—Molecular graphics for all devices—From smartphones to workstations. *Bioinformatics* **2014**, *30*, 2981–2982. [[CrossRef](#)] [[PubMed](#)]
  85. Krieger, E.; Joo, K.; Lee, J.; Lee, J.; Raman, S.; Thompson, J.; Tyka, M.; Baker, D.; Karplus, K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **2009**, *77*, 114–122. [[CrossRef](#)] [[PubMed](#)]
  86. Krieger, E.; Vriend, G. New ways to boost molecular dynamics simulations. *J. Comput. Chem.* **2015**, *36*, 996–1007. [[CrossRef](#)] [[PubMed](#)]
  87. Trott, O.; Olson, A.J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [[CrossRef](#)]
  88. Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* **2018**, *46*, W363–W367. [[CrossRef](#)]