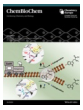
Characterization of a Tryptophan 6-Halogenase from *Streptomyces albus* and Its Regioselectivity Determinants



Jeongchan Lee,[a, b] Joonwon Kim,[b, c] Hyun Kim,[a, b] Eun-Jung Kim,[b] Hee-Jin Jeong,[f] Kwon-Young Choi,[g] and Byung-Gee Kim*[a, b, c, d, e]

Tryptophan halogenases are found in diverse organisms and catalyze regiospecific halogenation. They play an important role in the biosynthesis of halogenated indole alkaloids, which are biologically active and of therapeutic importance. Here, a tryptophan 6-halogenase (SatH) from *Streptomyces albus* was characterized by using a whole-cell reaction system in *Escherichia coli*. SatH showed substrate specificity for chloride and bromide ions, leading to regiospecific halogenation at the C6-position of L-tryptophan. In addition, SatH exhibited higher performance in bromination than that of previously reported tryptophan halogenases in the whole-cell reaction system. Through structure-based protein mutagenesis, it has been revealed that two consecutive residues, A78/V79 in SatH and G77/I78 in PyrH, are key determinants in the regioselectivity difference between tryptophan 6- and 5-halogenases. Substituting the AV with GI residues switched the regioselectivity of SatH by moving the orientation of tryptophan. These data contribute to an understanding of the key residues that determine the regioselectivity of tryptophan halogenases.

Introduction

Halogenated aromatic compounds are important materials for aromatic substitution or cross-coupling reactions.[2] Numerous halogenated natural compounds and synthetic drugs have shown that C–X bonds are beneficial in increasing their therapeutic efficacy or physicochemical properties, and efficient industrial halogenation in aromatic compounds is therefore always in high demand.[3] However, in chemical halogenation, many challenges of poor regioselectivity; harsh reaction conditions; the need for multiple protecting and activating steps; and the use of toxic compounds, including halogen gas and Lewis catalysts, still remain.[4] To overcome these problems in conventional processes, a flavin-dependent tryptophan halogenase has emerged as a promising biocatalyst for synthetic applications because it ensures not only high regiospecificity, but also green chemistry.[4,5]

Trp halogenases are classified depending on the position of halogenation, such as Trp 5-, 6-, and 7-halogenase. Protein residues that determine such regiospecificity and ion substrate specificity, such as Cl^- and Br^-, are not clearly understood in halogenases. Trp halogenases with different regioselectivity have high amino acid similarity, but also show accurate regioselectivity differences. To date, it has been revealed that Trp 7-halogenases have distinctive regioselectivity from that of Trp 5- and 6-halogenase due to the inverted orientation of the indole moiety of tryptophan in the binding sites.[6] In addition, Trp 6-halogenase was engineered to give Trp 7-halogenase, showing significantly changed regioselectivity.[6b] However, the key residues that determine the absolute difference in regioselectivity between Trp 5- and 6-halogenases are not fully understood. These two halogenases are very similar not only in the binding sites of the substrate, but also in the location of loops interacting with the amine and carboxyl groups of tryptophan. So far, only LPP/FET residues (L460/P461/P462 in Trp 6-halogenase SttH; F451/E452/T453 in Trp 5-halogenase PyrH), which interact with the amino acid backbone of tryptophan, have been reported as partial factors that determine the regioselectivity difference between these two enzymes.[7] Herein, we found a novel Trp 6-halogenase from *Streptomyces albus* by means of bioinformatics analysis. Furthermore, we identified additional regioselectivity-determining residues of Trp 6- and

[a] J. Lee, H. Kim, Prof. Dr. B.-G. Kim
School of Chemical and Biological Engineering, Seoul National University
Gwanak-ro 1, Gwanak-gu, Seoul 08826 (Republic of Korea)
E-mail: byungkim@snu.ac.kr
[b] J. Lee, Dr. J. Kim, H. Kim, E.-J. Kim, Prof. Dr. B.-G. Kim
Institute of Molecular Biology and Genetics, Seoul National University
Gwanak-ro 1, Gwanak-gu, Seoul 08826 (Republic of Korea)
[c] Dr. J. Kim, Prof. Dr. B.-G. Kim
Institute of Engineering Research, Seoul National University
Gwanak-ro 1, Gwanak-gu, Seoul 08826 (Republic of Korea)
[d] Prof. Dr. B.-G. Kim
Institute of Bioengineering in Bio-Max, Seoul National University
Gwanak-ro 1, Gwanak-gu, Seoul 08826 (Republic of Korea)
[e] Prof. Dr. B.-G. Kim
Interdisciplinary Program of Bioengineering, Seoul National University
Gwanak-ro 1, Gwanak-gu, Seoul 08826 (Republic of Korea)
[f] Prof. Dr. H.-J. Jeong
Department of Biological and Chemical Engineering, Hongik University
Sejong-ro 2639, Jochiwon-eup, Sejong 30016 (Republic of Korea)
[g] Prof. Dr. K.-Y. Choi
Department of Environmental Engineering, Ajou University
World cup-ro 206, Yeongtong-gu, Suwon 16499 (Republic of Korea)
Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201900723.

This article is part of a joint Special Collection dedicated to the Biotrans 2019 symposium. To view the complete collection, visit our homepage

5-halogenase near the catalytic key residues (K76/E359 in SatH; K75/E354 in PyrH) of the halogenases. AV/GI variants (A78G or V79I) of SatH revealed that, upon changing the two residues in the halogenase, its regioselectivity in chlorination and bromination of L-tryptophan changed from Trp C6-specific to C5-specific.

Results and Discussion

Bioinformatics analysis to seek a novel Trp 6-halogenase from *S. albus*

To discover a novel Trp 6-halogenase, *Streptomyces* species that produced natural compounds containing 6-chloroindole as a substructure were searched for in StreptomeDB 2.0,[8] the *Streptomyces*-specific biocompound database. In the results, borregomycins (Figure 1B) were identified in *S. albus*. However, notably, borregomycin was not produced by *S. albus* itself, but by the *bor* gene cluster isolated from desert soil metagenomes, which were successfully and heterologously expressed in *S. albus*.[1] In the *bor* gene cluster, one Trp 6-halogenase (BorH) is present and its characterization has been reported.[9] Serendipitously, we found one potential Trp halogenase in *S. albus* through a BLAST search, and it was named SatH (*S. albus* Trp halogenase).

A phylogenetic tree was drawn using amino acid sequences of SatH and previously known Trp halogenases (Figure 1A). SatH showed high identity with Th-Hal (85%), SttH (75%), and KtzR (73%), which were characterized as Trp 6-halogenases. It also showed considerable identity with PyrH (55%) and RebH (38%), which were characterized as Trp 5- and 7-halogenases, respectively. In addition, multiple sequence alignment of SatH with the reported halogenases showed that flavin adenine dinucleotide (FAD)-binding motifs, halogenase catalytic key residues, and the residues known to interact with tryptophan were identical among the Trp halogenases (Figure 2). However, the Trp halogenases displayed clear differences in the residues that determined the regioselectivity in halogenation at C5 or C6 of the indole ring in tryptophan.[7] The corresponding residues were F451/E452/T453 (FET) in PyrH, L456/P457/P458 in SatH, L455/P456/P457 in Th-Hal, and L460/P461/P462 (LPP) in SttH (square A in Figure 2). Based on the above multiple sequence analysis, SatH had the characteristic residues of Trp 6-halogenases and was potentially categorized as a Trp 6-halogenase. To verify its regioselectivity, we cloned and expressed SatH in *Escherichia coli* for further analysis.

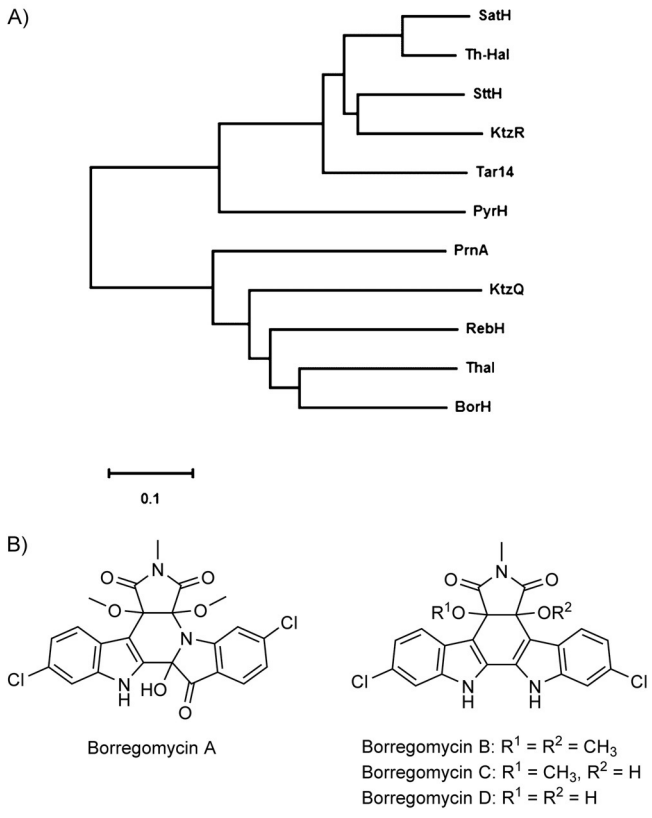
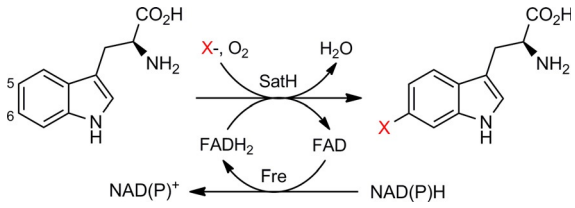


Figure 1. A) Phylogenetic tree of flavin-dependent tryptophan halogenases, including SatH (WP_078654696.1), Th-Hal (A0A1L1QK36), SttH (E9P162), KtzR (A8CF74), Tar14 (W5VG40), PyrH (A4D0H5), PrnA (P95480), KtzQ (A8CF75), RebH (Q8KHZ8), Thal (A1E280), and BorH (M9QSI0). B) Structure of borregomycins. The borregomycin gene cluster was found in desert soil metagenomes and successfully expressed in *S. albus*.[1]

Characterization and product identification of SatH in *E. coli* whole-cell reaction

To verify the suggested regioselectivity of SatH based on its protein sequence analysis, reaction products of a whole-cell reaction system by using *E. coli* Δ*tnaA* expressing SatH with flavin reductase (Fre; SatH + Fre; Table S3 in the Supporting Information) were analyzed (Scheme 1). The *tnaA* gene in *E. coli* was knocked out to prevent tryptophan degradation into indole (Table S3). Product identification was performed by using mass spectrometry and ^{1}H NMR spectroscopy. The SatH + Fre system converted tryptophan only into 6-chlorotryptophan (6-Cl-Trp), which confirmed that SatH was a Trp 6-halogenase with high regioselectivity (Figures 3A and S1). However, the bromination reaction of SatH did not show such high regioselectivity, in which both 5-Br-Trp (9%) and 6-Br-Trp (91%) were produced together (Figures 3B and S2). In the case of iodination, no substantial activity was detected (data not shown); this means that SatH does not accept an iodide ion in the active site. The ion substrate specificity of SatH was almost the same as that of known Trp halogenases.



Scheme 1. SatH-catalyzed halogenation of tryptophan. Fre is coexpressed for the regeneration of FAD to $FADH_2$. X^-: Cl^- or Br^-; $NAD(P)^+$: nicotinamide adenine dinucleotide (phosphate).

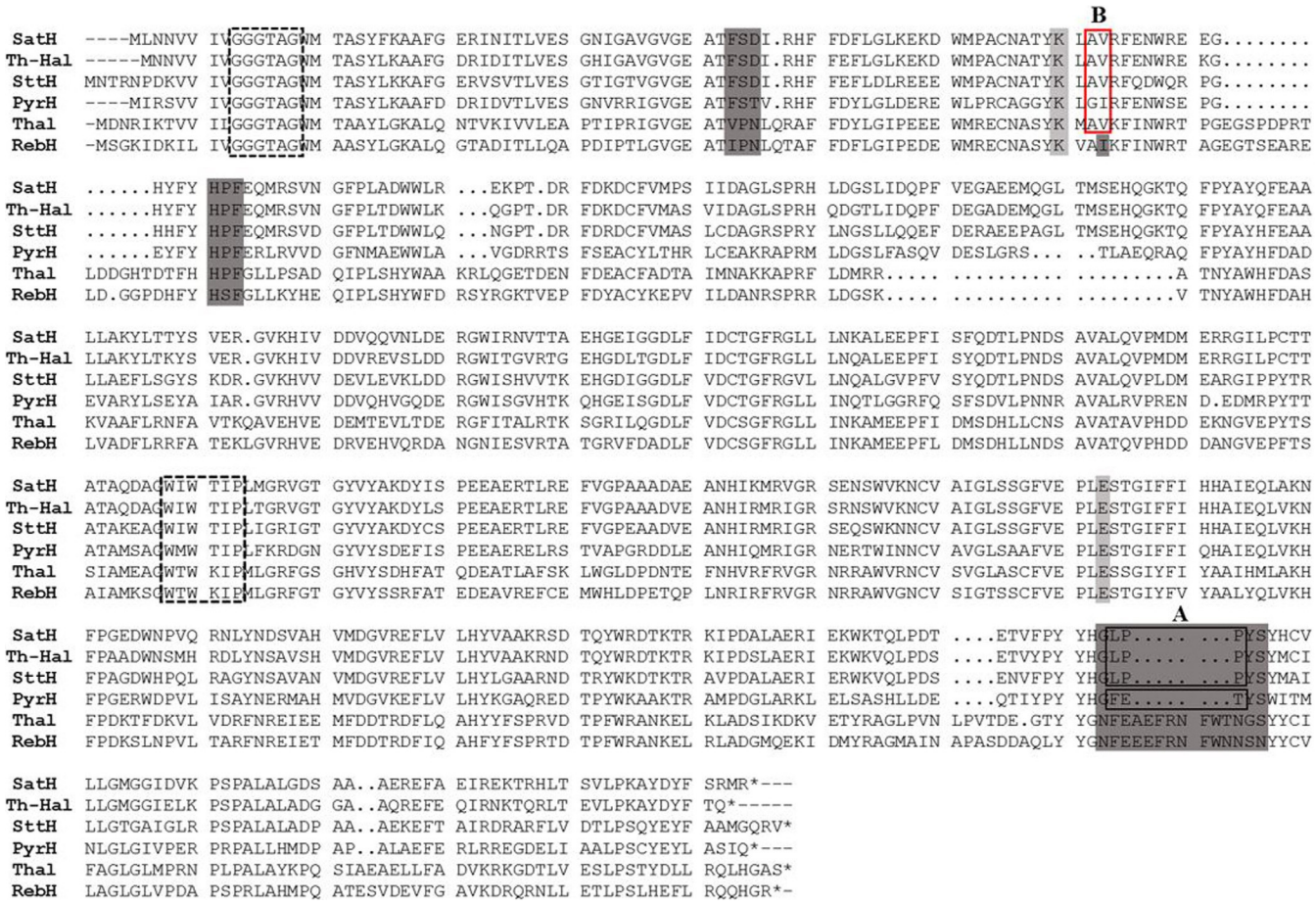


Figure 2. Multiple sequence alignment of SatH with Th-Hal, SttH, PyrH, Thal, and RebH. Residues in the box with a black dashed line indicate FAD-binding motifs. Light-gray shaded residues indicate key catalytic residues. Dark-gray shaded residues are known to interact with tryptophan. Residues in the box with a black rigid line (A) are known as key residues to affect tryptophan halogenase regioselectivity between C5 and C6 of the indole ring in tryptophan. Residues in the box with a red rigid line (B) are regioselectivity determinants between C5 and C6 halogenation found in this study.

Next, the regioselectivity of SatH was compared with that of other reported halogenases. Th-Hal, SttH, and PyrH (Table S2) were selected for comparison, owing to their high regioselectivity toward tryptophan.[10] The halogenases were coexpressed with Fre in *E. coli ΔtnaA* (Table S3) and their expression levels were identified (Figure S4). They were evaluated by using the whole-cell reaction supplemented with tryptophan and sodium chloride or sodium bromide. In the chlorination reaction, SatH showed a lower product yield of 6-Cl-Trp (0.86 mM) than that (1.08 mM) with Th-Hal, but high regioselectivity (Figure 3A). In the bromination reaction, SatH exhibited the highest product yield (0.76 mM Br-Trp) among all halogenases examined (Figure 3B). However, SatH produced both 5-Br- (9%) and 6-Br-Trp (91%), and Th-Hal showed similar characteristics. In contrast, SttH produced only 6-Br-Trp specifically, even though it had 75% identity with SatH. This result suggests that the key interacting residues of halogenases involved in chlorination and bromination reactions would be somewhat different. Here, through the analysis of the structure of SatH and subsequent protein mutagenesis, regioselectivity-determining residues of SatH were investigated further in detail.

Tryptophan binding poses in the active site identified the two key residues responsible for regioselective halogenation

To obtain a clue for finding the regioselectivity determinants of SatH, previously reported residues involved in regioselectivity of Trp 6- and 5-halogenases were analyzed. The LPP or FET (L460/P461/P462 in SttH; F451/E452/T453 in PyrH) residues were identified as the partial determinants for regioselectivity between Trp 6- and 5-halogenases, respectively.[7] Because the LPP/FET residues are located distant from the key catalytic residues (K79/E363 in SttH; K75/E354 in PyrH), which are located near the indole ring of tryptophan, LPP/FET are more likely to affect the substrate binding and coordination by interacting with the amino acid backbone of tryptophan. We assumed that there would be more influential residues near the catalytic key residues, which were not interacting with the amino acid backbone of tryptophan, but with the indole ring and directly affecting the regioselectivity of SatH.

The homology model of SatH was docked with tryptophan, and it was compared with the crystal structure of PyrH bound to tryptophan (PDB ID: 2WEU). In the active sites of PyrH and

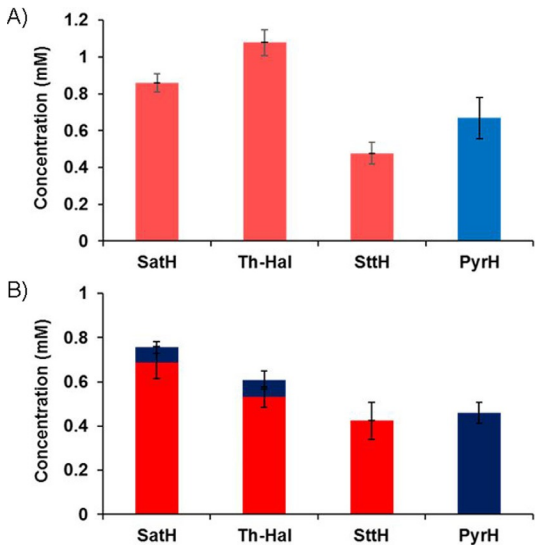


Figure 3. Production of A) chloro- and B) bromotryptophan from the reaction of SatH, Th-Hal, SttH, and PyrH; substitution at the 5 (■ ■) or 6 (■ ■) position. The reactions were performed with individual *E. coli* whole cells. The halogenases were coexpressed with Fre for the regeneration of FAD. Reaction conditions: 50 mM sodium phosphate buffer at pH 8.0, cell OD_{600} = 20, Trp 1.5 mM, NaCl or NaBr 300 mM, glucose 0.6% (*w*/*v*), 24 h, 30 °C, 200 rpm (*n* = 3). The halogenases were coexpressed with Fre for the regeneration of FAD.

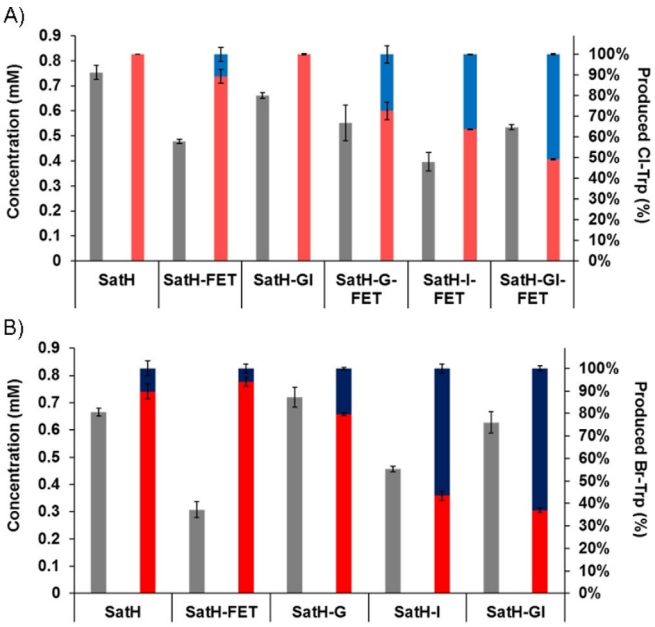


Figure 4. The total production (mM, ■) and ratio of substitution at the 5 (■ ■) to 6 (■ ■) position of A) chloro- and B) bromotryptophan obtained from the reaction of SatH and its mutants. The reactions were performed with individual *E. coli* whole cells. The halogenases were coexpressed with Fre for the regeneration of FAD. Reaction conditions: 50 mM sodium phosphate buffer at pH 8.0, cell OD_{600} = 20, Trp 1.5 mM, NaCl or NaBr 300 mM, glucose 0.6% (*w*/*v*), 24 h, 30 °C, 200 rpm (*n* = 3).

SatH, all residues within 5 Å of the tryptophan were identical, except for the LPP/FET residues and two other distinguishable residues, G77 and I78 in PyrH and A78 and V79 in SatH. To determine whether such differences were commonly found in other halogenases, amino acid sequences of Trp 5-halogenases (W2EQU4, A0A2D3E318, A4D0H5, and W1J423) and Trp 6-halogenases (E9P162, WP_078654696.1, A0A1L1QK36, W5VG40, M9QSI0, A1E280, and A8CF74) were aligned and compared (Figure S3). In the corresponding region, Trp 5-halogenases again had G and I, whereas Trp 6-halogenases, except KtzR (A8CF74), had A and V; this suggested that the two residues were conserved for their regioselectivity determination.

First, to confirm that the LPP residues in SatH and the FET residues in PyrH really determined their 6- and 5-halogenase regioselectivity, the LPP residues of SatH were mutated to FET residues, yielding mutant SatH-FET (L456F/P457E/P458T). The mutant regioselectivity was evaluated through the whole-cell reactions, as previously described. In the chlorination reaction, although wild-type SatH produced only 6-Cl-Trp (Figure 4A), the SatH-FET mutant, exhibiting somewhat decreased total halogenase activity (ca. 63%), produced 11% 5-Cl-Trp and 89% 6-Cl-Trp. However, in the bromination reaction, the SatH-FET mutant did not show any changes in its regioselectivity, as desired. It produced 6% 5-Br-Trp and 94% 6-Br-Trp, which was almost the same as that of the production profile of the wild type (Figure 4B). In the case of SatH, mutation from the LPP to FET residues (SatH-FET) affected the regioselectivity of chlorination, but not bromination.

Next, to examine the functional role of AV/GI residues (A78/V79 in SatH; G77/I78 in PyrH) in determining the regioselectivity, the A78/V79 residues of SatH were mutated to G/I, yielding mutant SatH-GI (A78G/V79I). In the chlorination reaction, SatH-GI mutant produced only 6-Cl-Trp, with slightly decreased total halogenase activity compared with that of the wild type. However, if the LPP residues of SatH-GI mutant were mutated to FET, yielding mutant SatH-GI-FET (A78G/V79I/L456F/P457E/P458T), the regioselectivity was changed to produce 50% 5-Cl-Trp and 50% 6-Cl-Trp, with a slightly increased total activity relative to that of the SatH-FET mutant (Figure 4A). Moreover, in the case of bromination, the regioselectivity of the SatH-GI mutant, without FET residues, showed a preference for 5- (63%) over 6-bromination (37%), with a similar total halogenase activity to that of the wild type (Figure 4B). The result indicates that the mutations from the AV to GI residues have a substantial impact on the regioselectivity of both chlorination and bromination of SatH, even compared with mutations from the LPP to FET residues; this suggests that AV/GI residues are key determinants of Trp 6- and 5-halogenases.

To determine the effects of each residue, SatH-FET A78G and V79I were constructed, yielding mutants SatH-G-FET and SatH-I-FET, respectively, and evaluated for chlorination (Figure 4A). The V79I mutation had a greater effect on the regioselectivity change from C6 to C5 compared with that of the A78G mutation because the SatH-I-FET mutant produced more 5-Cl-Trp (36%) than that of the SatH-G-FET mutant (27%). Furthermore, these characteristics of single mutations, A78G and V79I, were analyzed in the case of bromination. The LPP residues of SatH was not mutated to FET residues because they did not affect the bromination regioselectivity, as previously described. Instead, only SatH A78G and V79I were constructed, yielding mu-

tants SatH-G and SatH-I, respectively, and evaluated for their bromination (Figure 4B). Similar to chlorination, the V79I mutation had a greater effect on the regioselectivity in bromination because SatH-I showed a high yielder of 5-Br-Trp (57%) than that with SatH-G (20%). Although the effect of the A78G mutation on its regioselectivity was small, this mutation also slightly enhanced C5-bromination. In addition, the A78G mutation contributed to enhanced total halogenase activity, as inferred from the results that SatH-G-FET and SatH-GI-FET mutants produced higher yields of Cl-Trp than those of SatH-FET and SatH-I-FET, respectively (Figure 4A). The enhanced total halogenase activity in bromination as a result of the A78G mutation was also identified based on the results that SatH-G and SatH-GI mutants produced higher yields of Br-Trp than those with SatH and SatH-I, respectively (Figure 4B). Therefore, the two residues appeared to have a synergistic effect on determining the regioselectivity of SatH.

To better understand how these residues affect the regioselectivity, the distances between C5 or C6 of the indole ring of tryptophan and the key catalytic residues (K76/E359 in SatH; K75/E354 in PyrH), and the distances between C5 or C6 and the β carbon of V79 in SatH or that of I78 in PyrH were compared (Figure 5). The distances from the β carbon atoms to the indole ring were almost the same for both PyrH and SatH. However, the distances from the ε-amino group of the catalytic Lys residue to C5 or C6 of the indole ring were distinctively different between PyrH and SatH (3.7 and 4.4 Å in PyrH; 4.5 and 3.7 Å in SatH), even though the residues in the halogenases around the indole ring within 5 Å were the same, except for the AV residues in SatH and the GI residues in PyrH (Figure S5). The different distances are possibly caused by the hydrophobic interaction between the indole ring and the longer side chain of isoleucine (I78 in PyrH) compared with that of valine (V79 in SatH), which may cause a slight change in the position of the substrate. Because the Lys residue (K75 in PyrH; K76 in SatH) is the main factor that determines the position of a halogenating agent, the distances between Lys and C5/C6 of the indole ring will determine the regiospecificity of the halogenase. As a result, the slight shift of the position of the indole ring, owing to the hydrophobic interaction with the AV/GI residues, would change the distances, which would influence the regioselectivity of C6- and C5-halogenation.

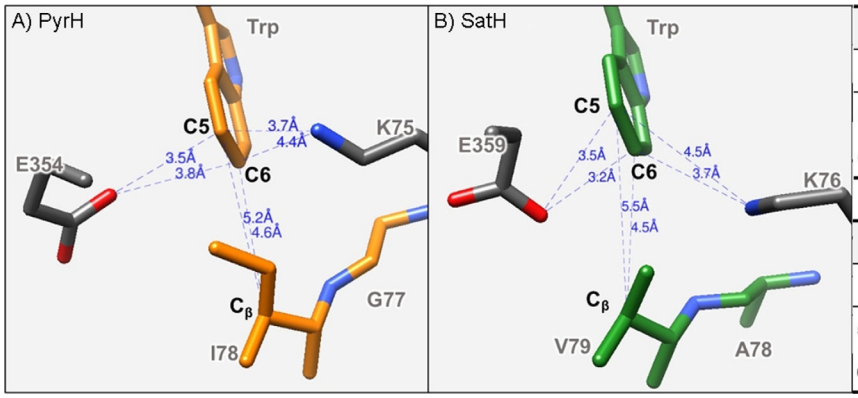
Conclusions

A novel Trp 6-halogenase from *S. albus N*-16041, SatH, was characterized by using an *E. coli* whole-cell reaction system. Except for the LPP/FET motifs previously known as the regioselectivity determinants of Trp 6- and 5-halogenases, the extra residues affecting coordination of substrate binding were identified: Ala-Val in Trp 6-halogenase and Gly-Ile in Trp 5-halogenase (A78/V79 in SatH; G77/I78 in PyrH), which might be critical for the addition of a halogen atom from the catalytic site Lys residue to the indole ring of tryptophan, and hence, be important in determining regioselectivity between C6- and C5-halogenation. These two residues appear to change the coordination of tryptophan binding and result in the shift of the halogenated position between C5 and C6 of tryptophan. Identification of these residues could further be used for the prediction of the regiospecificity of putative halogenases and for that of natural compound structures based on the genome sequence.

Experimental Section

Materials: Lysogeny broth (LB) was purchased from BD bioscience. Tryptophan was purchased from Sigma–Aldrich. Chloro- and bromotryptophans were purchased from Santa Cruz Biotechnology or Tokyo Chemical Industry. Oligomers and sequencing were purchased and performed through Bioneer (Seoul, South Korea) and Bionics (Seoul, South Korea), respectively. Enzymes involved in the restriction reaction, ligation, and PCR were purchased from Thermo Scientific, Promega, and Novagens, respectively

Bioinformatic analysis: Natural compounds with 6-chloroindole as a substructure were searched in StreptomeDB 2.0. The structure of 6-chloroindole was drawn and submitted to the database. The results contained the structure of natural products and *Streptomyces* species producing the relevant compounds. The availability of *Streptomyces* species in the results was searched in bioresource centers. *S. albus* was the only strain available at public bioresource centers and *S. albus* N-16041 was purchased from the ARS culture collection. BLAST was used to seek Trp 6-halogenase in *S. albus* by



A) PyrH			
Unit: Å	K75	E354	C β (I78)
5-Carbon	3.7	3.5	5.2
6-Carbon	4.4	3.8	4.6

B) SatH			
Unit: Å	K76	E359	C β (V79)
5-Carbon	4.5	3.5	5.5
6-Carbon	3.7	3.2	4.5

Figure 5. Residues of the active site around the indole ring of tryptophan and distances from the residues to C5 or C6 of the indole ring. A) The crystal structure of PyrH (Trp 5-halogenase) with tryptophan (PDB ID: 2WEU) and B) the homologous model of SatH (Trp 6-halogenase). Tryptophan was docked into the SatH structure by using AutoDock Vina.

using the amino acid sequence of SttH with nucleotide database sets of *S. albus*, including nucleotide collection and whole-genome shotgun contigs. The protein sequences of the halogenase (WP_078654696.1) and previously characterized halogenases were collected for phylogenetic analysis. A molecular phylogenetic analysis was conducted through MEGA6 by using the maximum likelihood method.[11] Multiple sequence alignments of Trp 5- and 6-halogenases were performed by using Clustal Omega.[12]

Construction of plasmids and strains: Primers and plasmids used in this study are listed in Tables S1 and S2, respectively. The halogenase gene was amplified from the genomic DNA of *S. albus* N-16041 by using primers SatH_F_NdeI and SatH_R_HindIII, and cloned into pET28a, yielding pET28a::satH. The Fre gene (*fre*) in *E. coli* was amplified from the genomic DNA of *E. coli* BL21(DE3) by using primers Fre_F_BamHI and Fre_R_XhoI and cloned into pACYC, yielding pACYC::fre. For the regioselectivity study, mutations were generated by using pET28a::satH as a template. Primers AV_GI_F and AV_GI_R were used to generate the A78G/V79I mutation. Primers LPP_FET_F and LPP_FET_R were used to generate the L456F/P457E/P458T mutation. Primers AV_GV_F and AV_GV_R were used to generate the A78G mutation. Primers AV_AI_F and AV_AI_R were used to generate V79I mutation. Strains were based on *E. coli* strain BL21 containing the DE3 prophage. The *tnaA* gene was knocked out of the genome by using a previously reported method.[13]

Protein expression: All expression of enzymes was conducted by using *E. coli* BL21 (DE3). Each of the constructed vectors was transformed into the competent cell and selected on a LB agar plate supplemented with appropriate antibiotics. A single colony was inoculated into LB broth with the corresponding antibiotic and grown overnight at 37 °C. The seed culture (500 μL) was inoculated into LB (50 mL) and grown at 37 °C until the cell density reached OD_{600} = 0.6–0.8. Expression of halogenases was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG; 0.5 mM) with further overnight incubation at 18 °C. The cells were harvested by centrifugation at 3200 *g* for 15 min and washed with phosphate-buffered saline (PBS) for further whole-cell reaction. For SDS-PAGE analysis, the cells were resuspended in Tris buffer solution containing 50 mM Tris·HCl (pH 8.0), 200 mM NaCl, and 10% glycerol to OD_{600} = 20. The cell soup was disrupted by ultrasonication in ice-chilled water for 8 min (10 s on, 12 s off). The soluble fraction was collected by centrifugation of the cell lysate at 13000 *g* for 30 min at 4 °C. After the denaturation process, the total and soluble fractions of lysates were loaded into 12% sodium dodecyl sulfate polyacrylamide gel at 50 V for 15 min and 130 V for 60 min.

Whole-cell reaction and product analysis: The cells for the resting-cell reaction were prepared as previously described. For a resting-cell system with tryptophan as a substrate, *E. coli* BL21(DE3) *ΔtnaA* was used. The cells were resuspended in 50 mM sodium phosphate buffer at pH 8.0. The cell was diluted to OD_{600} = 20 with the addition of NaCl (300 mM) and glucose (0.6%) in a glass vial and incubated in a shaking incubator in 200 rpm at 30 °C for 24 h. An aliquot (500 μL) of the reaction soup was quenched with methanol (500 μL). The quenched sample was centrifuged at 10000 *g* for 20 min followed by a centrifugal filtration through an MWCO 3 kDa filter. An aliquot (200 μL) of the filtered-out sample was further used in HPLC. Tryptophan derivatives were separated by using the HPLC sample fraction for LC/MS/MS analysis. For quantitative and structural analysis of halogenated tryptophans, HPLC and ^{1}H NMR spectroscopic analyses were performed as follows. The halogenated products were separated by means of HPLC (Young-Lin, South Korea) equipped with a PFP reversed-phase column (2.5 μm particle size, 4.6×100 mm, Thermo Scientific), with an isocratic mobile phase of water/CH_3CN/trifluoroacetic acid (60:40:0.1% *v/v/v*) at a flow rate of 0.7 mL min^{-1}. The compounds were monitored at λ = 280 nm. For structural analysis, halogenated products were harvested from a sample fraction obtained by means of HPLC. Tryptophan derivative powders were prepared by freeze-drying. For ^{1}H NMR spectroscopic analysis, the tryptophans were dissolved in D_2O.

Homology modeling of halogenases: The homology model of SatH was prepared by using SWISS-MODEL.[14] The structure of ThHal (PDB ID: 5LV9) was used as a template for the SatH model. A docking simulation of L-tryptophan to the structure of the SatH homology model was performed by using Autodock Vina.[15] Figures of the structures were generated by using UCSF Chimera.[16]

Acknowledgements

We would like to express our appreciation to the Agricultural Research Service for the supply of S. albus *N-16041 strain. This research was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, and Future Planning; the Korean government (NRF-2017R1E1A1A01073523, NRF-2016S1A5A2A03926786, MSIT-2018R1C1B5044988); and the Hongik University new faculty research support fund.*

Conflict of Interest

The authors declare no conflict of interest.

Keywords: mutagenesis · proteins · regioselectivity · tryptophan halogenases · whole-cell reactions

[1] F. Y. Chang, S. F. Brady, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2478–2483.
[2] a) W. J. Chung, C. D. Vanderwal, *Angew. Chem. Int. Ed.* **2016**, *55*, 4396–4434; *Angew. Chem.* **2016**, *128*, 4470–4510; b) J. Yamaguchi, A. D. Yamaguchi, K. Itami, *Angew. Chem. Int. Ed.* **2012**, *51*, 8960–9009; *Angew. Chem.* **2012**, *124*, 9092–9142.
[3] a) V. Agarwal, Z. D. Miles, J. M. Winter, A. S. Eustaquio, A. A. El Gamal, B. S. Moore, *Chem. Rev.* **2017**, *117*, 5619–5674; b) M. R. Scholfield, C. M. Zanden, M. Carter, P. S. Ho, *Protein Sci.* **2013**, *22*, 139–152.
[4] C. Schnepel, N. Sewald, *Chem. Eur. J.* **2017**, *23*, 12064–12086.
[5] L. C. Blasiak, C. L. Drennan, *Acc. Chem. Res.* **2009**, *42*, 147–155.
[6] a) X. F. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Lhiefeld, K. H. van Pee, J. H. Naismith, *J. Mol. Biol.* **2009**, *391*, 74–85; b) A. C. Moritzer, H. Minges, T. Prior, M. Frese, N. Sewald, H. H. Niemann, *J. Biol. Chem.* **2019**, *294*, 2529–2542.
[7] S. A. Shepherd, B. R. Menon, H. Fisk, A. W. Struck, C. Levy, D. Leys, J. Micklefield, *ChemBioChem* **2016**, *17*, 821–824.
[8] D. Klementz, K. Doring, X. Lucas, K. K. Telukunta, A. Erxleben, D. Deubel, A. Erber, I. Santillana, O. S. Thomas, A. Bechthold, S. Gunther, *Nucleic Acids Res.* **2016**, *44*, D509–514.
[9] K. Lingkon, J. J. Bellizzi III, *ChemBioChem* **2020**, *21*, 1121–1128.
[10] B. R. Menon, J. Latham, M. S. Dunstan, E. Brandenburger, U. Klemstein, D. Leys, C. Karthikeyan, M. F. Greaney, S. A. Shepherd, J. Micklefield, *Org. Biomol. Chem.* **2016**, *14*, 9354–9361.
[11] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, *Mol. Biol. Evol.* **2013**, *30*, 2725–2729.
[12] F. Sievers, D. G. Higgins, *Protein Sci.* **2018**, *27*, 135–145.
[13] a) K. A. Datsenko, B. L. Wanner, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6640–6645; b) T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, *Mol. Syst. Biol.* **2006**, *2*, 2006.0008.

[14] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, *Nucleic Acids Res.* **2018**, *46*, W296 – W303.
[15] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455 – 461.
[16] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605 – 1612.

Manuscript received: November 29, 2019
Accepted manuscript online: January 8, 2020
Version of record online: February 27, 2020