# Structure and Activity of the Thermophilic Tryptophan-6 Halogenase BorH

Kazi Lingkon and John J. Bellizzi, III*[a]

Flavin-dependent halogenases carry out regioselective aryl halide synthesis in aqueous solution at ambient temperature and neutral pH using benign halide salts, making them attractive catalysts for green chemistry. BorH and BorF, two proteins encoded by the biosynthetic gene cluster for the chlorinated bisindole alkaloid borregomycin A, are the halogenase and flavin reductase subunits of a tryptophan-6-halogenase. Quantitative conversion of L-tryptophan (Trp) to 6-chlorotryptophan could be achieved using 1.2 mol% BorH and 2 mol% BorF. The optimal reaction temperature for Trp chlorination is 45 °C, and the melting temperatures of BorH and BorF are 48 and 50 °C respectively, which are higher than the thermal parameters for most other halogenases previously studied. Steady-state kinetic analysis of Trp chlorination by BorH determined parameters of $k_{cat} = 4.42\ min^{-1}$, and $K_M$ of 9.78 μM at 45 °C. BorH exhibits a broad substrate scope, chlorinating and brominating a variety of aromatic substrates with and without indole groups. Chlorination of Trp at a 100 mg scale with 52% crude yield, using 0.2 mol% BorH indicates that industrial scale biotransformations using BorH/BorF are feasible. The X-ray crystal structure of BorH with bound Trp provides additional evidence for the model that regioselectivity is determined by substrate positioning in the active site, showing C6 of Trp juxtaposed with the catalytic Lys79 in the same binding pose previously observed in the structure of Thal.

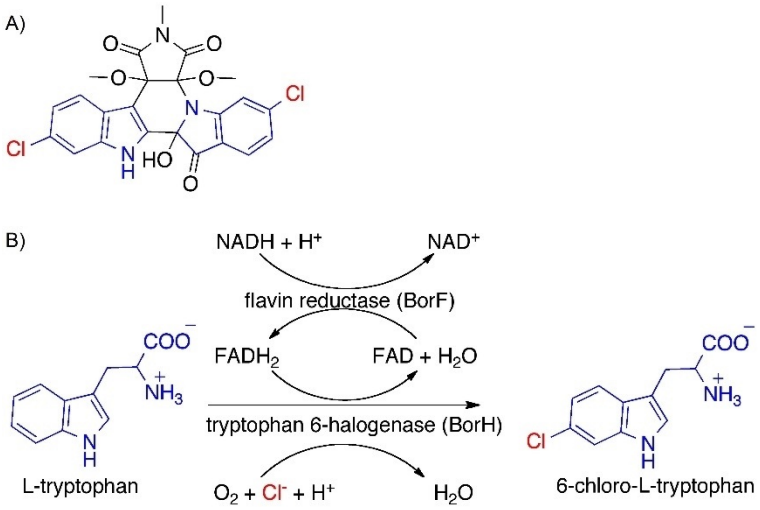
Many biologically active compounds, including natural products,[1,2] synthetic drugs,[3,4] and agricultural chemicals[5] derive potency and specificity from halogen atoms. Carbon-halogen bonds are also key components of many synthetic intermediates, such as aryl halides used in transition metal catalyzed cross-coupling reactions.[6,7] Traditional synthetic methods for the preparation of organohalogen compounds suffer from lack of regioselectivity and frequently require toxic and unsustainable reagents and solvents, though promising new "green" methods are under development.[8,9] Flavin-dependent halogenases (FDHs), which can regioselectively halogenate aromatic substrates using only halide ion and $O_2$, have emerged as a new environmentally friendly method for aryl halide synthesis.[10–15] FDHs are closely related to flavin-dependent two-component monooxygenases[16–19] and consist of two proteins: a halogenase (which converts $O_2$ and halide ion into HOCl or HOBr and water with the oxidation of $FADH_2$ to FAD) and a smaller flavin reductase component (which reduces FAD to $FADH_2$ using NADH). FDHs have been discovered in the biosynthetic pathways of bacterial and fungal natural products, where they insert chlorine or bromine into free or acyl carrier protein-bound aromatic substrates.[20] The most extensively studied halogenases (EC 1.14.19.9, 1.14.19.58, 1.14.19.59) regioselectivity halogenate L-tryptophan on the 5, 6, or 7 positions of the indole ring, though most have the ability to halogenate other substrates as well.[21–28] A newly identified viral FDH has recently been shown to have iodination activity.[29]

Borregomycin A (Scheme 1 A), a chlorinated bisindole alkaloid produced by oxidative dimerization of Trp, was discovered through metagenomic screening of soil samples from the Anza-Borrego desert and has antiproliferative activity against human colon cancer HCT116 cells and inhibits the kinase CaMKIIδ.[30] Two genes in the biosynthetic gene cluster for borregomycin A, *borF* and *borH*, were predicted by sequence homology to encode the flavin reductase and halogenase components of a flavin-dependent tryptophan halogenase.

BorH's predicted amino acid sequence (GenBank AGI62217, UniProtKB M9QSI0) contains a GxGxxG motif that places it in the flavin-dependent monooxygenase superfamily, along with a conserved WxWxIP motif unique to FDHs (Figure S1 in the Supporting Information). The crystal structure of PrnA[31] revealed that residues in the WxWxIP motif separate the flavin and substrate binding sites, preventing the direct reaction between substrate and flavin hydroperoxide intermediate that occurs in monooxygenases. Instead, the flavin hydroperoxide is thought to oxidize chloride ion to HOCl, which traverses a 10 Å tunnel to the substrate binding site, where it chlorinates the substrate in an electrophilic aromatic substitution reaction with the assistance of a conserved lysine (positioning the HOCl through hydrogen bonding[32] or reacting with HOCl to form a chloramine intermediate that in turn chlorinates the substrate[33]) and glutamate (acting as a general base). As in the two-component flavin-dependent monooxygenases, the halogenase is incapable of regenerating $FADH_2$, and depends on the flavin reductase to reduce FAD through a diffusible flavin mechanism. FAD is released by the halogenase and diffuses to the flavin reductase, which uses NADH to reduce the FAD to $FADH_2$, which diffuses back to the halogenase.[18,34–39] BorF (GenBank AGI62216, UniProtKB M9QXS1) is the flavin reductase in the *Bor* gene cluster that regenerates $FADH_2$ for BorH.

The closest characterized homolog of BorH is Thal (70% sequence identity), a tryptophan-6-halogenase involved in thienodolin biosynthesis from *Streptomyces albogriseolus*.[24,28,40] The other characterized tryptophan-6-halogenases (Th-Hal,[41] SttH,[27] KtzR,[25] and Tar14[42]) share only 37–39% identity with

[a] K. Lingkon, Dr. J. J. Bellizzi, III
Department of Chemistry and Biochemistry, The University of Toledo
2801 W. Bancroft St. MS 602, Toledo, OH 43606 (USA)
E-mail: john.bellizzi@utoledo.edu
Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201900667.




**Scheme 1.** A) The chlorinated bisindole alkaloid borregomycin A. B) Reaction catalyzed by BorH/BorF in the biosynthetic pathway of borregomycin A.

BorH and are found in a more distant clade along with the tryptophan-5-halogenases AbeH and PyrH (Figure S2). BorH is more closely related to the tryptophan-7-halogenases RebH, KtzQ, and PrnA than to the tryptophan-5-halogenases and ThHal, SttH and KtzR.

Crystal structures of the FDHs PrnA,[31] RebH,[43] PyrH,[44] and ThaI[28] with bound Trp have shown that the regioselectivity of halogenation arises from the orientation of Trp in the binding site, with the hydrogen on the carbon atom closest to the conserved Lys replaced by the halogen atom. The position of the chlorine in borregomycin A strongly suggested that BorH carries out chlorination exclusively at C6 of the indole side chain (Scheme 1 B), converting Trp to 6-chloro-L-tryptophan (6-Cl-Trp).

Because of the potential synthetic applications for enzyme-catalyzed regioselective aryl halide synthesis, many groups are engineering FDHs with improved stability and catalytic efficiency and with altered substrate specificity or regioselectivity for synthetic applications and generation of novel natural product analogues.[14,15,45–58] The goal of developing biocatalysts with the necessary stability for industrial scale bioconversions has motivated the identification of a halogenase from a thermophilic bacterium,[41] as well as the engineering of thermostable RebH mutants.[59]

To further expand the FDH toolkit, we have expressed and purified the halogenase BorH along with its flavin reductase BorF, verified its regioselectivity for Trp halogenation, and shown that it is able to chlorinate and brominate a variety of aromatic substrates. BorH displays a higher reaction temperature and thermal stability than most previously characterized FDHs, and a pilot study showed that it is capable of bioconversion on a preparative scale. We also report the crystal structure of the BorH-Trp complex, which verified the structural basis for the observed regioselectivity.

## BorH and BorF have tryptophan halogenase and flavin reductase activity

BorH and BorF were overexpressed in *Escherichia coli* and purified to homogeneity using affinity chromatography and size exclusion chromatography (Figure S3). The yield of purified BorH was 4.50 mg $L^{-1}$ culture and the yield of purified BorF was 4.25 mg $L^{-1}$ culture. BorF copurified with bound FAD (referred to as holo-BorF), as indicated by yellow color and confirmed by electrospray ionization mass spectrometry (ESI-MS) analysis. Apo-BorF could be prepared by washing column-immobilized holo-BorF with urea and KBr.[60]

BorF's flavin reductase activity was observed by monitoring oxidation of NADH to $NAD^+$ spectrophotometrically at 340 nm (Figure 1). No depletion of NADH was observed with FAD alone or apo-BorF alone, but when apo-BorF and FAD were added, oxidation of NADH was observed. A small decrease in NADH was observed in the presence of holo-BorF, presumably due to a single turnover of NADH reducing the FAD copurified with BorF.

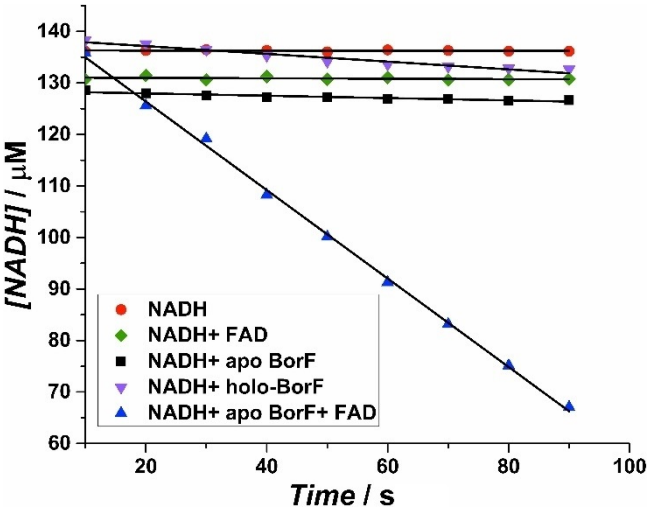



**Figure 1.** BorF catalyzes the FAD-dependent oxidation of NADH. NADH concentration was monitored at 340 nm and decreased in the presence of 0.2 µM apo-BorF + 25 µM FAD (blue triangles), but not when BorF or FAD were absent (◆ or ■). Slow oxidation was evident with holo-BorF (gray inverted triangles).

BorH's Trp halogenase activity was confirmed by detection of Cl-Trp produced in an in vitro halogenation reaction (Figure 2 A). Reactions containing Trp, BorH, holo-BorF and NADH were analyzed using reversed-phase (RP) HPLC to monitor (by 280 nm absorbance) disappearance of Trp ($t_R$ = 10.3 min) and appearance of a product ($t_R$ = 11.4 min) with a retention time matching that of a 6-Cl-Trp standard. The product was not observed in the absence of BorF, indicating that halogenation requires the $FADH_2$ produced by BorF.

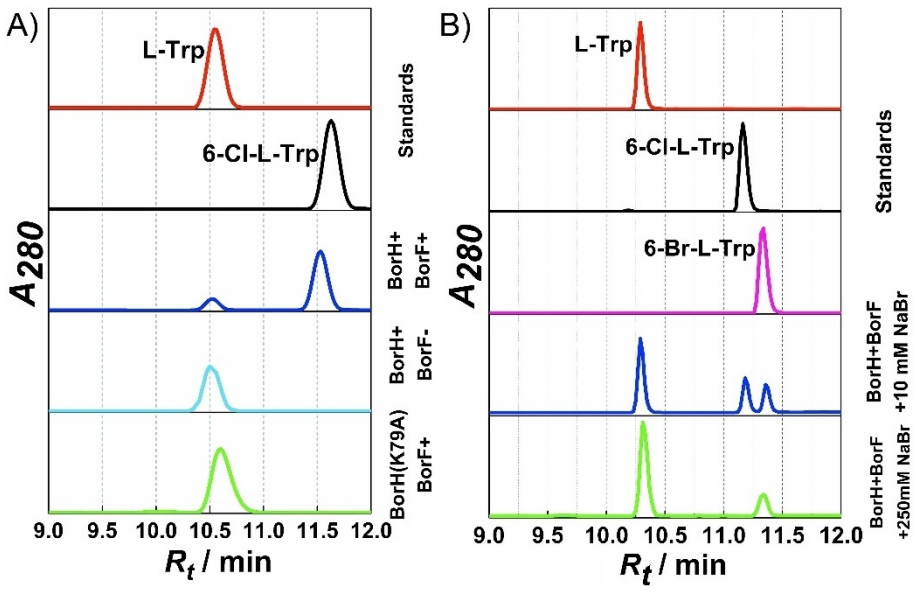



**Figure 2.** BorH is a tryptophan halogenase. A) BorH converts Trp into 6-Cl-Trp; 100 μL reactions containing 3 μM BorH, 5 μM holo-BorF, 5 mM NADH, 60 mM NaCl and 0.5 mM Trp in 10 mM HEPES pH 8 were incubated for 1 h, quenched, and analyzed by RP-HPLC. A peak matching the $R_t$ of a 6-Cl-Trp standard was formed in reactions containing BorH and holo-BorF, but not when BorF was absent or when BorH was replaced with BorH(K79A), which lacks the catalytic lysine. B) BorH converts Trp into 6-Br-Trp. Using the same reaction conditions except replacing NaCl with 10 mM NaBr, peaks matching both 6-Cl-Trp and 6-Br-Trp standards appeared (due to the presence of residual chloride in the BorF storage buffer). Increasing NaBr to 250 mM led to suppression of the 6-Cl-Trp peak.

In RebH and other previously characterized halogenases, a conserved lysine is essential for halogenation activity[31] Sequence homology predicts Lys79 of BorH to be the essential lysine, and this was confirmed by the absence of product detected from reactions containing a BorH(K79A) mutant.

ESI-MS analysis of the product identified ions at *m/z* 239.0574 and 241.0559 in a 3:1 intensity ratio, which matched the mass of monochlorinated Trp and the isotopic ratio of Cl (Figure S4 A). To identify the regioisomer of Cl-Trp formed, 0.8 mg of product was isolated and analyzed by $^1$H NMR. The results identified the product as 6-Cl-Trp (Figure S5), confirming the prediction that BorH is a tryptophan-6-halogenase.

Substituting NaBr for NaCl in the in vitro halogenation reaction led to decreased production of 6-Cl-Trp and the presence of a new peak ($t_R = 11.6$ min) which matched the retention time for a 6-Br-Trp standard (Figure 2 B). The residual 6-Cl-Trp peak results from the presence of NaCl in the BorF storage buffer (final concentration 10 mM in the reaction). Increasing NaBr to 250 mM led to a complete disappearance of the 6-Cl-Trp peak. The product of the bromination reaction was analyzed by ESI-MS and showed two peaks with *m/z* 283.0084 and 285.0048 in a 1:1 ratio, matching the mass and isotopic ratio for monobrominated tryptophan (Figure S4 B).

### BorH and BorF display elevated thermal stability and optimal reaction temperature

Because thermostability is a desirable characteristic in a biocatalyst,[41,59] and as the *Bor* gene cluster was isolated from a desert soil sample,[30] we investigated the optimal reaction temperature as well as the melting temperatures of the BorH/BorF system. BorH displayed maximum activity for Trp chlorination at 45 °C (Figures 3 and S6), which is higher than the optimal temperature reported for other flavin-dependent halogenases (SttH 40 °C,[26] RebH 30–35 °C,[59] PrnA 30 °C,[41] Th-Hal 30 °C[41]).

The melting temperature ($T_m$) of BorH was determined to be 48.0 °C using a ThermaFluor fluorescent dye-binding thermal stability assay[61] (Figure S7 A, C). This is similar to the reported

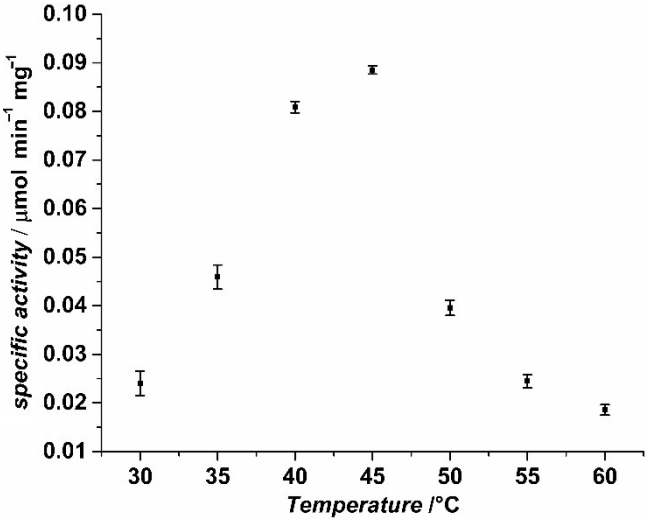



**Figure 3.** Temperature dependence of BorH/BorF-catalyzed chlorination of Trp. Reactions containing 3 μM BorH, 5 μM holo-BorF with 5 mM NADH and 0.5 mM Trp in 10 mM HEPES pH 8.0, 60 mM NaCl in a total volume of 100 μL were incubated for 25 min at the indicated temperatures, quenched, and analyzed by RP-HPLC.

$T_m$ of Th-Hal (47.8 °C), a halogenase from the thermophile *Streptomyces violaceusniger* SPC6, and significantly higher than the $T_m$ of RebH, PrnA, PyrH, SttH, and KtzR reported in the Th-Hal study (30–40 °C),[41] though the $T_m$ of RebH has been reported elsewhere as 52.4 °C.[59] The $T_m$ of BorF was determined to be 50.0 °C (Figure S7B, D), which is similar to the thermophilic flavin reductase Th-Fre from *Bacillus subtilis* WU-S2 which was identified in conjunction with Th-Hal.[41]

Reaction of 0.5 mM Trp with chloride using 0.6 mol% BorH and 1 mol% holo-BorF at 45 °C reached a maximum conversion of 84% after 25 minutes (Figure S8). The specific activity under these conditions was 0.089 μmol $min^{-1}$ $mg^{-1}$. The specific activity for bromination under these conditions using 250 mM NaBr is 0.019 μmol $min^{-1}$ $mg^{-1}$ (21% of the specific activity for chlorination). In comparison, 0.5 mol% of a thermostabilized RebH mutant halogenates 0.5 mM Trp with ≈60% yield in 2 h at 49 °C[59] and 0.5 mol% Th-Hal halogenates 0.5 mM Trp with ≈75% yield in 30 min at 45 °C.[41] We were able to achieve 100% conversion of Trp to 6-Cl-Trp in a 25 min reaction at 45 °C by increasing the enzyme concentrations to 1.2 mol% BorH and 2 mol% holo-BorF.

Steady-state kinetic analysis of Trp chlorination by BorH was carried out by measuring initial velocities using 3–35 μM Trp at 45 °C. Nonlinear fitting to the Michaelis–Menten equation was used to determine the kinetic parameters $k_{cat}$ (turnover number) = 4.42 $min^{-1}$, and $K_M$ (Michaelis constant for Trp) = 9.78 μM, yielding a specificity constant ($k_{cat}/K_M$) of 0.45 $min^{-1}$ $\mu M^{-1}$ (Figure 4). This specificity constant is 1.8 times higher than that reported for the thermostable halogenase Th-Hal at 45 °C, 1.2 times higher than Th-Hal at 30 °C, and 2.8–45 times higher than that reported for PyrH, SttH, PrnA, RebH, and KtzR at 30 °C (Table S1).[41] Compared with BorH's $k_{cat}$ at 45 °C, Th-Hal's $k_{cat}$ is higher at 45 °C and approximately equal at 30 °C. Th-Hal's $K_M$ for Trp is higher at both temperatures than BorH's $K_M$ at 45 °C.

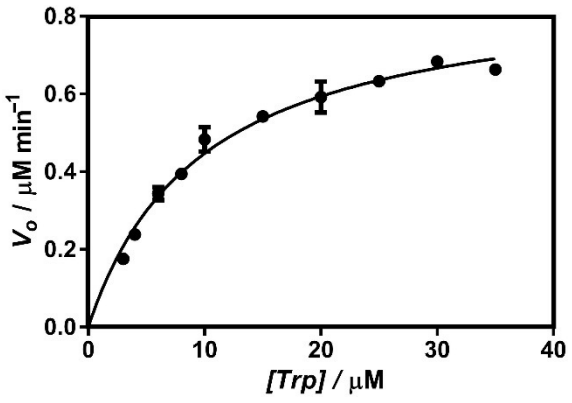



**Figure 4.** Steady-state kinetic analysis of conversion of Trp to 6-Cl-Trp by BorH/BorF at 45 °C and pH 8.0. Reactions (final volume of 1 mL) containing 3–35 μM Trp, 0.2 μM BorH, 0.35 μM holo-BorF, 0.5 mM NADH, and 10 mM NaCl in the presence of NADH (0.5 mM) and NaCl (10 mM) were carried out at 45 °C. Aliquots of 65 μL were taken at 1 min intervals for 5 min, quenched and analyzed by RP-HPLC to determine product formation. Initial velocities determined from progress curves were fit to the Michaelis–Menten equation to determine kinetic parameters.

### Preparative scale chlorination of Trp by BorH/BorF

To test the performance of BorH/BorF under conditions of higher substrate loading, 100 mg of Trp was reacted with chloride using 0.2 mol% of BorH (4.5 μM BorH/2 mM Trp). We adapted our reaction conditions to include glucose dehydrogenase (GDH) as a cofactor (NADH) regeneration system.[27,47,48,59] A crude yield of 52% 6-Cl-Trp was observed after 19 h, corresponding to a total turnover number (TTN; moles product per mole enzyme) of ≈225. The crude yield of halogenated product produced by BorH in this pilot study was similar to previously reported yields at this scale for other halogenases despite the use of significantly lower catalyst loading. Lysate containing 3 mol% RebH has been used to chlorinate 100 mg Trp with 74% crude yield,[47] and SttH has been shown to chlorinate 100 mg anthranilamide with 25% isolated yield[27] and 2.5 mol% at a 25 mg scale, Th-Hal was able to chlorinate tryptophan, *N*-methyltryptophan, kynurenine, anthranilic acid and anthranilamide with isolated yields of 62.5, 12, 21, 8, and 19% respectively.[41] At a 10 mg scale, thermostabilized RebH halogenates tryptophan, 2-aminonapthalene, 2-methyltryptamine, and tryptoline with 69, 62, 56, and 67% yields, respectively.[59] Gram scale halogenation of Trp has also been reported using RebH cross-linked with flavin reductase and alcohol dehydrogenase.[62] We anticipate that optimization of BorH, BorF and dehydrogenase concentrations, as well as temperature and volume; will enable this yield to be further improved.

### BorH halogenates a variety of aromatic compounds with and without indole groups

To investigate the substrate scope of BorH, we first tested its ability to halogenate the closely related metabolites 5-hydroxytryptophan and serotonin; 1.2 mol% BorH and 2 mol% BorF could chlorinate 0.5 mM serotonin or 5-hydroxytryptophan with 41% and 35% crude yield, respectively in a 1 h reaction (Figures S15 and S16), indicating that BorH is tolerant of substitutions to the indole group. We then expanded our approach to examine both chlorination and bromination of indole-containing aromatic substrates (3-indolepropionic acid, tryptamine) and non-indole containing aromatic substrates (L-kynurenine, anthranilamide, 6-aminoquinoline). These experiments used 0.5 mM substrate and 1.2 mol% BorH and 2 mol% BorF with 35 U $mL^{-1}$ GDH as a cofactor regeneration system in 12 h reactions (Table 1, Figures 5 and S9–S14). Under these conditions, we saw quantitative chlorination of tryptophan and 83% bromination of Trp, and between 10–62% chlorination and 4–53% bromination of the other substrates based on crude HPLC yields (Figure 5 and Table S2).

BorH halogenates serotonin, 5-hydroxytryptophan, and tryptamine much more efficiently than 3-indolepropionic acid, suggesting that the amino group of Trp is more important than the carboxylate in substrate recognition. Surprisingly, the non-indole containing substrates (kynurenine, anthranilamide, and 6-aminoquinolone) were halogenated at similar or higher levels than the non-Trp indole-containing substrates. Relative yields for chlorination and bromination varied by substrate as

**Table 1.** BorH-Trp data collection and refinement statistics.

| Data collection | | Refinement | |
|---|---|---|---|
| wavelength [Å] | 1.0332 | $R_{work}$ | 0.1929 (0.2541) |
| resolution | 31.89–1.98 | $R_{free}$ | 0.2288 (0.2862) |
| range [Å] | (2.05–1.98) | CC (work) | 0.941 (0.831) |
| space group | $P2_1$ | CC (free) | 0.933 (0.761) |
| unit cell | | no. of non- | 18164 |
| *a*, *b*, *c* [Å] | 73.48, 157.92, | hydrogen atoms: | |
| | 113.25 | protein | 16753 |
| $\alpha$, $\beta$, $\gamma$ [°] | 90, 104.06, 90 | ligands | 135 |
| total | 338736 (34139) | water | 1276 |
| reflections | | protein residues: | 2081 |
| unique | 170593 (17172) | RMS (bonds) | 0.006 |
| reflections | | RMS (angles) | 0.88 |
| multiplicity | 2.0 (2.0) | Ramachandran | 97 |
| completeness | 98.17 (99.16) | favored [%] | |
| mean $I/\sigma$ ($I$) | 6.04 (2.30) | Ramachandran | 3 |
| Wilson *B*-factor | 24.61 | allowed [%] | |
| $R_{merge}$ | 0.077 (0.348) | Ramachandran | 0 |
| $R_{meas}$ | 0.108 (0.493) | outliers [%] | |
| $R_{pim}$ | 0.077 (0.348) | Clashscore | 3.28 |
| $CC_{1/2}$ | 0.986 (0.690) | average *B* factor | 35.40 |
| $CC^*$ | 0.996 (0.904) | protein | 35.20 |
| | | ligands | 59.70 |
| | | water | 34.60 |

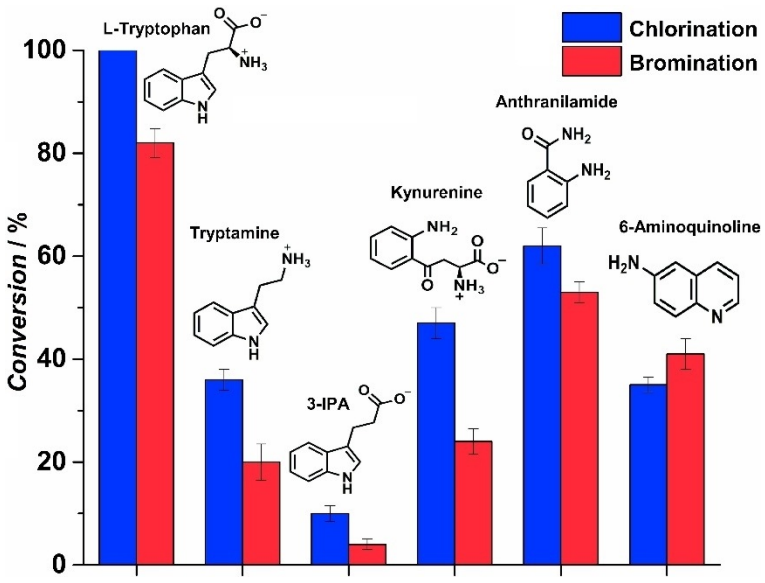



**Figure 5.** BorH halogenates aromatic substrates with indole, benzene, and quinoline groups. Substrates shown (0.5 mM) were reacted for 12 h with 6 μM BorH, 10 μM holo-BorF, 35 U mL$^{-1}$ glucose dehydrogenase, 20 mM D-glucose, and 200 μM NADH in 10 mM HEPES pH 8.0. Chlorination reactions contained 50 mM NaCl and bromination reactions contained 250 mM NaBr. Reactions were analyzed by RP-HPLC and integrated peak areas were used to determine product yields (Figures S9–S14).

well, ranging from ≈2:1 for L-kynurenine and tryptamine, ≈1.2:1 for Trp and anthranilamide, and ≈1:1.2 for 6-aminoquinoline. Based on this initial data, we can conclude that BorH is capable of halogenating indole, benzene or quinoline derivatives with exocyclic amino groups, but further work is necessary to determine the extent of the substrate scope and the regioselectivity of the non-Trp halogenations.

When comparing the substrate scope of BorH to that reported for other tryptophan-6-halogenases, we observed enhanced activity for some substrates and decreased activity for others. Relative to Th-Hal, BorH has higher activity against anthranilamide, similar chlorination activity against 5-hydroxytryptophan, and decreased activity against L-kynurenine.[41] Thal was more effective at halogenating 6-aminoquinoline than BorH (88% versus 34% yield), though this may be due to higher catalyst loading and longer reaction time in the Thal study.[57]

## Crystal structure of BorH-Trp

Crystals of BorH were grown, soaked in Trp, cryoprotected, and frozen in liquid $N_2$. X-ray diffraction data were collected at the Life Sciences Collaborative Access Team beamline 21-ID-D at the Advanced Photon Source, Argonne National Laboratory to 1.98 Å, and the structure was determined by molecular replacement using RebH (65% sequence identity, PDB ID: 2OAM[33]) as a search model (Table 1). BorH crystallized in space group $P2_1$ with four molecules in the asymmetric unit. Each chain BorH is a single domain with 26 β-strands, 19 α-helices, six $3_{10}$ helices, and one π-helix (Figures 6A and S1). As originally observed in PrnA[31] and seen in all subsequent FDH structures,[27,28,33,41,43,44,59,63] the fold comprises two subdomains: a rectangular subdomain containing a Rossmann-like fold on one end and a six-stranded mixed β-sheet on the other end spanned by several long helices, and a smaller helical pyramid-shaped subdomain composed of α5–α7 and α14–α18. The Trp binding site is located at the interface between the box and the pyramid, and the FAD binding site is located in the box-shaped subdomain. The BorH-Trp structure has been deposited in the Protein Data Bank under the ID: PDB ID: 6UL2.

The final refined model ($R_{work}$ = 19.3%, $R_{free}$ = 22.9%) contains all amino acids except for 1–2 residues at the N terminus and two residues at the C terminus of each protomer, and loop β2-β3 (residues 39–45) in chains B, C, and D. In previously determined halogenase structures, this loop is involved in FAD binding and is flexible in the absence of FAD. The four BorH protomers have Cα root mean square deviations (RMSD) of between 0.19 and 0.36 Å with one another (Table S3).

BorH is the fifth Trp-6-halogenase to have its structure determined,[27,28,41,42] but only the second (after Thal) to be solved with bound Trp. Clear electron density was visible in initial maps in all four chains for the substrate Trp. The orientation of the Trp and the interactions with BorH are identical to those observed in the Trp-bound structure of Thal (Figures 6B, C and S20A).

The Trp α-amino group forms a salt bridge with the side chain carboxylate of Glu460 and donates two H-bonds, one to the hydroxy group of Tyr453 and the other to the backbone carbonyl oxygen of Phe464. The Trp carboxylate accepts H-bonds from the hydroxy group of Tyr454 and two ordered

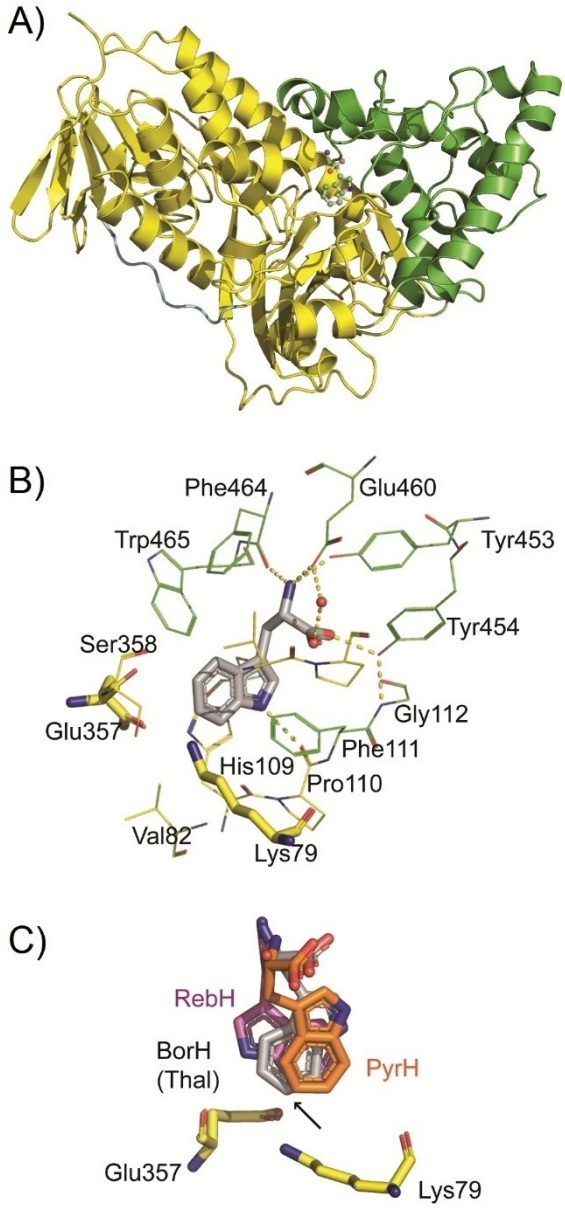



**Figure 6.** Crystal structure of BorH-Trp. A) Ribbon diagram of chain A, with the FAD-binding rectangular box subdomain in yellow and the pyramidal subdomain in green. The loop between β2 and β3, which is disordered in chains B, C, and D, and has been shown to participate in FAD binding in other halogenases, is in cyan. Trp is shown as a ball and stick model. B) The Trp binding site in BorH. The orientation of Trp and contacts between Trp and BorH are identical to those observed in Thal (PDB ID: 6H44) The catalytic residues (Lys79 and Glu357) and the Trp are shown as stick models, and the other residues lining the binding pocket are shown as thin lines. Hydrogen bonds are indicated by yellow dashed lines. There are two ordered water molecules (red spheres) hydrogen bonded to the Trp carboxylate. Both are conserved in the structures of Thal and RebH. C) Orientation of Trp in Trp-5-halogenase PyrH (PDB ID: 2WEU, orange), Trp-6-halogenase BorH (gray), and Trp-7-halogenase RebH (PDB ID: 2E4G, magenta). The arrow indicates the position of halogenation.

water molecules, which are conserved in the structures of Thal and RebH. The more extensive interactions between BorH and the amino group relative to the carboxylate is consistent with the observed preference for halogenating tryptamine over 3-IPA. The indole nitrogen donates a H-bond to the backbone oxygen of Pro111. The binding pocket for the Trp is bounded by Val52 and Pro53 (from loop α2-β1), His109, Pro110, and Phe111 (from β8), Glu357 (on loop β26–α11), Tyr453, Tyr454 (on α15), Glu460 (on α16), Phe464 and Trp465 (from loop α16–α17), and Ser469 (on α17), all of which are conserved between BorH and Thal.

The putative catalytic residues Lys79 and Glu357 perfectly align with the corresponding residues in the other halogenase crystal structures. The common orientation of the Trp in BorH and Thal positions C6 of Trp closest to Lys79 (3.7 Å away from Nε in BorH). In RebH C7 is closest to Lys79, and in PyrH C5 is closest to Lys79 (Figure 6C).

As expected from the sequence identity (65–70%), the overall tertiary structure of BorH is nearly identical that of Thal and RebH, with no significant insertions or deletions or substantial backbone differences. Chain A of BorH has a Cα root mean square deviation (RMSD) of 0.9 Å (526 residues aligned) with Thal-Trp (PDB ID: 6H44, 70% identity to BorH) and 1.0 Å (527 residues aligned) with RebH-Trp (PDB ID: 2E4G, 65% identity).

A second cluster of structural homologues includes the Trp-5-halogenase structure PyrH-Trp (PDB ID: 2WEU), and the Trp-6-halogenase structures Th-Hal (PDB ID: 5LV9), SttH-FAD (PDB ID: 5HY5), and Tar14-FAD (PDB ID: 6NSD), which have 37–39% sequence identity and Cα RMSD values of 1.4–1.8 Å over 470–484 residues aligned. The areas that do not align well between BorH/Thal/RebH and PyrH/Th-Hal/SttH/Tar14 comprise two insertions in the BorH family and one insertion in the Th-Hal family, and result in a dramatic reorganization of one face of the pyramidal subdomain and a complete difference in the topology of the Trp binding site, as originally noted when the structure of PyrH was solved and compared with the structure of RebH.[44]

The backbone-binding portion of the Trp binding site is made up of a topologically different insertion in the BorH family than it is in the Th-Hal family (Figure S19). The BorH insertion is between α14 and α17, which contains 28 residues in BorH but only 17 residues in Th-Hal/SttH/Tar14/PyrH. In BorH, this region forms two short roughly perpendicular α-helices (α15 and α16) that contact the Trp amino and carboxylate group (Tyr453, Tyr454 on α15, Glu460 on α16, and Phe464 and Trp465 on loop α16–α17). In Th-Hal and related structures, the space that is occupied by this insertion in BorH is occupied by a different insertion in the Th-Hal family, which is between α7 and β11 of BorH. There are no structures of Trp bound to the Trp-6-halogenases from this family (Th-Hal, SttH, Tar14), but the Trp-bound structure of PyrH indicates that residues in this region (Gln160, Gln163) donate hydrogen bonds to the carboxylate of Trp.

The second BorH insertion is between β5 and β8, which are the second and first strands in the seven-strand mixed β-sheet in the box-shaped subdomain. In Th-Hal, these strands are connected by an eight residue loop, but in BorH, there are a total of 21 residues between these two strands, including a short β-hairpin (β6-β7) that forms a flap covering one corner of the pyramidal subdomain.

PrnA (59% identity) is much closer in sequence to BorH than to the Th-Hal structural family, and contains the α14–α17 inser-

tion but has a shorter β5–β8 insertion lacking the β-hairpin. The PrnA-FAD-Trp structure (PDB ID: 2AQJ) has a Cα RMSD value of 1.5 Å for 510 residues aligned with BorH.

In summary, we have shown that BorH, which was predicted to be a tryptophan-6-halogenase based on sequence homology, not only regioselectively halogenates Trp, but is also able to chlorinate and brominate structurally distinct aromatic substrates. BorH and BorF (its companion flavin reductase subunit) have higher thermal stability than most halogenases, reflecting their origins in a desert soil bacterial gene cluster, and their thermal stability and relatively high optimal reaction temperature are attractive characteristics in a green biocatalyst for synthetic applications. Production of $>50$ mg of 6-Cl-Trp in a pilot study using 0.2 mol% BorH indicates that BorH has sufficient stability and catalytic efficiency to be useful for preparative scale halogenation.

## Acknowledgements

*The authors thank Dr. Yong Wah Kim and Istiak Hossain for assistance with HPLC and NMR experiments, and Dr. Dragan Isailovic and Krishani K. Rajanayake for assistance with mass spectrometry experiments, LS-CAT beamline staff for assistance with X-ray data collection, and the University of Toledo and US National Institutes of Health (NIH) grant R15GM110679 (to J.B.) for funding. Dr. Zheng Ma originally cloned the BorF construct, and the* bor *gene cluster was generously provided by Dr. Sean Brady at Rockefeller University. This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).*

## Conflict of Interest

*The authors declare no conflict of interest.*

**Keywords:** biocatalysis · flavin · halogenation · thermophilic · tryptophan

[1] C. M. Harris, R. Kannan, H. Kopecka, T. M. Harris, *J. Am. Chem. Soc.* **1985**, *107*, 6652 – 6658.
[2] G. W. Gribble, *Chemosphere* **2003**, *52*, 289 – 297.
[3] R. Wilcken, M. O. Zimmermann, A. Lange, A. C. Joerger, F. M. Boeckler, *J. Med. Chem.* **2013**, *56*, 1363 – 1388.
[4] Z. Xu, Z. Yang, Y. Liu, Y. Lu, K. Chen, W. Zhu, *J. Chem. Inf. Model.* **2014**, *54*, 69 – 78.
[5] P. Jeschke, *Pest. Manag. Sci.* **2017**, *73*, 1053 – 1066.
[6] V. F. Slagt, A. H. M. de Vries, J. G. de Vries, R. M. Kellogg, *Org. Process Res. Dev.* **2010**, *14*, 30 – 47.
[7] P. Ruiz-Castillo, S. L. Buchwald, *Chem. Rev.* **2016**, *116*, 12564 – 12649.
[8] A. Podgorsek, M. Zupan, J. Iskra, *Angew. Chem. Int. Ed.* **2009**, *48*, 8424 – 8450; *Angew. Chem.* **2009**, *121*, 8576 – 8603.
[9] D. A. Petrone, J. Ye, M. Lautens, *Chem. Rev.* **2016**, *116*, 8003 – 8104.
[10] L. C. Blasiak, C. L. Drennan, *Acc. Chem. Res.* **2009**, *42*, 147 – 155.
[11] D. R. M. Smith, S. Grüschow, R. J. M. Goss, *Curr. Opin. Chem. Biol.* **2013**, *17*, 276 – 283.
[12] V. Agarwal, Z. D. Miles, J. M. Winter, A. S. Eustaquio, A. A. El Gamal, B. S. Moore, *Chem. Rev.* **2017**, *117*, 5619 – 5674.
[13] C. Schnepel, N. Sewald, *Chemistry* **2017**, *23*, 12064 – 12086.
[14] A. E. Fraley, D. H. Sherman, *Bioorg. Med. Chem. Lett.* **2018**, *28*, 1992 – 1999.
[15] J. Latham, E. Brandenburger, S. A. Shepherd, B. R. K. Menon, J. Micklefield, *Chem. Rev.* **2018**, *118*, 232 – 269.
[16] M. M. Huijbers, S. Montersino, A. H. Westphal, D. Tischler, W. J. van Berkel, *Arch. Biochem. Biophys.* **2014**, *544*, 2 – 17.
[17] M. L. Mascotti, M. Juri Ayub, N. Furnham, J. M. Thornton, R. A. Laskowski, *J. Mol. Biol.* **2016**, *428*, 3131 – 3146.
[18] T. Heine, W. J. H. van Berkel, G. Gassner, K. H. van Pée, D. Tischler, *Biology* **2018**, *7*, 42.
[19] E. Romero, J. R. Gomez Castellanos, G. Gadda, M. W. Fraaije, A. Mattevi, *Chem. Rev.* **2018**, *118*, 1742 – 1769.
[20] K. H. van Pée, E. P. Patallo, *Appl. Microbiol. Biotechnol.* **2006**, *70*, 631 – 641.
[21] V. N. Burd, K.-H. van Pée, *Biochemistry* **2004**, *69*, 674 – 677.
[22] E. Yeh, S. Garneau, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3960 – 3965.
[23] S. Zehner, A. Kotzsch, B. Bister, R. D. Süssmuth, C. Méndez, J. A. Salas, K.-H. van Pée, *Chem. Biol.* **2005**, *12*, 445 – 452.
[24] C. Seibold, H. Schnerr, J. Rumpf, A. Kunzendorf, C. Hatscher, T. Wage, A. J. Ernyei, C. Dong, J. H. Naismith, K.-H. van Pée, *Biocatal. Biotransform.* **2006**, *24*, 401 – 408.
[25] J. R. Heemstra, C. T. Walsh, *J. Am. Chem. Soc.* **2008**, *130*, 14024 – 14025.
[26] J. Zeng, J. Zhan, *Biotechnol. Lett.* **2011**, *33*, 1607 – 1613.
[27] S. A. Shepherd, B. R. K. Menon, H. Fisk, A.-W. Struck, C. Levy, D. Leys, J. Micklefield, *ChemBioChem* **2016**, *17*, 821 – 824.
[28] A. C. Moritzer, H. Minges, T. Prior, M. Frese, N. Sewald, H. H. Niemann, *J. Biol. Chem.* **2019**, *294*, 2529 – 2542.
[29] D. S. Gkotsi, H. Ludewig, S. V. Sharma, J. A. Connolly, J. Dhaliwal, Y. Wang, W. P. Unsworth, R. J. K. Taylor, M. M. W. McLachlan, S. Shanahan, J. H. Naismith, R. J. M. Goss, *Nat. Chem.* **2019**, *11*, 1091 – 1097.
[30] F.-Y. Chang, S. F. Brady, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2478 – 2483.
[31] C. Dong, S. Flecks, S. Unversucht, C. Haupt, K. H. van Pée, J. H. Naismith, *Science* **2005**, *309*, 2216 – 2219.
[32] S. Flecks, E. P. Patallo, X. Zhu, A. J. Ernyei, G. Seifert, A. Schneider, C. Dong, J. H. Naismith, K. H. van Pée, *Angew. Chem. Int. Ed.* **2008**, *47*, 9533 – 9536; *Angew. Chem.* **2008**, *120*, 9676 – 9679.
[33] E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan, C. T. Walsh, *Biochemistry* **2007**, *46*, 1284 – 1292.
[34] K. Abdurachim, H. R. Ellis, *J. Bacteriol.* **2006**, *188*, 8153 – 8159.
[35] J. Sucharitakul, T. Phongsak, B. Entsch, J. Svasti, P. Chaiyen, D. P. Ballou, *Biochemistry* **2007**, *46*, 8611 – 8623.
[36] J. Valton, C. Mathevon, M. Fontecave, V. Niviere, D. P. Ballou, *J. Biol. Chem.* **2008**, *283*, 10287 – 10296.
[37] R. Tinikul, W. Pitsawong, J. Sucharitakul, S. Nijvipakul, D. P. Ballou, P. Chaiyen, *Biochemistry* **2013**, *52*, 6834 – 6843.
[38] E. Morrison, A. Kantz, G. T. Gassner, M. H. Sazinsky, *Biochemistry* **2013**, *52*, 6063 – 6075.
[39] J. Sucharitakul, R. Tinikul, P. Chaiyen, *Arch. Biochem. Biophys.* **2014**, *555 – 556*, 33 – 46.
[40] D. Milbredt, E. P. Patallo, K.-H. van Pée, *ChemBioChem* **2014**, *15*, 1011 – 1020.
[41] 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.
[42] H. Luhavaya, R. Sigrist, J. R. Chekan, S. M. K. McKinnie, B. S. Moore, *Angew. Chem. Int. Ed.* **2019**, *58*, 8394 – 8399; *Angew. Chem.* **2019**, *131*, 8482 – 8487.
[43] E. Bitto, Y. Huang, C. A. Bingman, S. Singh, J. S. Thorson, G. N. Phillips, Jr., *Proteins* **2008**, *70*, 289 – 293.
[44] X. Zhu, W. De Laurentis, K. Leang, J. Herrmann, K. Ihlefeld, K. H. van Pée, J. H. Naismith, *J. Mol. Biol.* **2009**, *391*, 74 – 85.
[45] W. S. Glenn, E. Nims, S. E. O'Connor, *J. Am. Chem. Soc.* **2011**, *133*, 19346 – 19349.
[46] A. Lang, S. Polnick, T. Nicke, P. William, E. P. Patallo, J. H. Naismith, K.-H. van Pée, *Angew. Chem. Int. Ed.* **2011**, *50*, 2951 – 2953; *Angew. Chem.* **2011**, *123*, 3007 – 3010.

[47] J. T. Payne, M. C. Andorfer, J. C. Lewis, *Angew. Chem. Int. Ed.* **2013**, *52*, 5271 – 5274; *Angew. Chem.* **2013**, *125*, 5379 – 5382.
[48] J. T. Payne, C. B. Poor, J. C. Lewis, *Angew. Chem. Int. Ed.* **2015**, *54*, 4226 – 4230; *Angew. Chem.* **2015**, *127*, 4300 – 4304.
[49] S. Brown, S. E. O'Connor, *ChemBioChem* **2015**, *16*, 2129 – 2135.
[50] J. T. Payne, M. C. Andorfer, J. C. Lewis, *Methods Enzymol.* **2016**, *575*, 93 – 126.
[51] M. C. Andorfer, H. J. Park, J. Vergara-Coll, J. C. Lewis, *Chem. Sci.* **2016**, *7*, 3720 – 3729.
[52] K. H. van Pée, D. Milbredt, E. P. Patallo, V. Weichold, M. Gajewi, *Methods Enzymol.* **2016**, *575*, 65 – 92.
[53] P. Chankhamjon, Y. Tsunematsu, M. Ishida-Ito, Y. Sasa, F. Meyer, D. Boettger-Schmidt, B. Urbansky, K. D. Menzel, K. Scherlach, K. Watanabe, C. Hertweck, *Angew. Chem. Int. Ed.* **2016**, *55*, 11955 – 11959; *Angew. Chem.* **2016**, *128*, 12134 – 12138.
[54] M. Frese, C. Schnepel, H. Minges, H. Voß, R. Feiner, N. Sewald, *ChemCatChem* **2016**, *8*, 1799 – 1803.
[55] B. R. K. Menon, E. Brandenburger, H. H. Sharif, U. Klemstein, S. A. Shepherd, M. F. Greaney, J. Micklefield, *Angew. Chem. Int. Ed.* **2017**, *56*, 11841 – 11845; *Angew. Chem.* **2017**, *129*, 12003 – 12007.
[56] M. C. Andorfer, J. E. Grob, C. E. Hajdin, J. R. Chael, P. Siuti, J. Lilly, K. L. Tan, J. C. Lewis, *ACS Catal.* **2017**, *7*, 1897 – 1904.
[57] M. C. Andorfer, J. C. Lewis, *Annu. Rev. Biochem.* **2018**, *87*, 159 – 185.
[58] J. T. Payne, P. H. Butkovich, Y. Gu, K. N. Kunze, H. J. Park, D.-S. Wang, J. C. Lewis, *J. Am. Chem. Soc.* **2018**, *140*, 546 – 549.
[59] C. B. Poor, M. C. Andorfer, J. C. Lewis, *ChemBioChem* **2014**, *15*, 1286 – 1289.
[60] M. H. Hefti, F. J. Milder, S. Boeren, J. Vervoort, W. J. van Berkel, *Biochim. Biophys. Acta Gen. Subj.* **2003**, *1619*, 139 – 143.
[61] S. A. Seabrook, J. Newman, *ACS Comb. Sci.* **2013**, *15*, 387 – 392.
[62] M. Frese, N. Sewald, *Angew. Chem. Int. Ed.* **2015**, *54*, 298 – 301; *Angew. Chem.* **2015**, *127*, 302 – 305.
[63] M. A. Ortega, D. P. Cogan, S. Mukherjee, N. Garg, B. Li, G. N. Thibodeaux, S. I. Maffioli, S. Donadio, M. Sosio, J. Escano, L. Smith, S. K. Nair, W. A. van der Donk, *ACS Chem. Biol.* **2017**, *12*, 548 – 557.

Manuscript received: November 1, 2019
Accepted manuscript online: November 6, 2019
Version of record online: December 16, 2019