Materials:

Unless otherwise noted, all reagents were obtained from commercial suppliers and used without further purification. Deuterated solvents were obtained from Cambridge Isotope labs. Silicycle silica gel plates (250 mm, 60 F254) were used for analytical TLC, and preparative chromatography was performed using SiliCycle SiliaFlash silica gel (230-400 mesh). Oligonucleotides were purchased from Integrated DNA Technologies (San Diego, CA). Plasmids pET-28a/RebF and pET-28a/RebH in BL-21 DE3 *E. coli* were provided by the Walsh group of Harvard Medical School, Boston, MA.[1] The pLIC-MBP plasmid was provided by the Bottomley group of Monash University, Clayton, Australia.[2] The pGro7 plasmid encoding the groES and groEL chaperone set was purchased from Takara (Otsu, Shiga, Japan). BL21(DE3) *E. coli* cells were purchased from Invitrogen (Carlsbad, CA). T4 DNA polymerase and Phusion HF polymerase were purchased from New England Biolabs (Ipswitch, MA). Luria broth (LB) and Terrific broth (TB) media were purchased from Research Products International (Mt. Prospect, IL). Library colonies were picked using an automated colony picker (Norgren Systems). Qiagen Miniprep Kits were purchased from QIAGEN Inc. (Valencia, CA) and used according to the manufacturer's instructions. All genes were confirmed by sequencing at the University of Chicago Comprehensive Cancer Center DNA Sequencing & Genotyping Facility (900 E. 57th Street, Room 1230H, Chicago, IL 60637). Electroporation was carried out on a Bio-Rad MicroPulser using method Ec3. Ni-nitrilotriacetic acid (Ni-NTA) resin and Pierce® BCA Protein Assay Kits were purchased from Fisher Scientific International, Inc. (Hampton, NH), and the manufacturer's instructions were following when using both products (for Ni-NTA resin, 5 mL resin were used, with buffers delivered by a peristaltic pump at a rate of 1 mL/min, in a 4 °C cold cabinet). Amicon® 30 kD spin filters for centrifugal concentration were purchased from EMD Millipore (Billerica, MA) and used at 4,000 g at 4 °C. The glucose dehydrogenase GDH, FAD, and NAD were purchased from Codexis (Redwood City, CA). NADH was purchased from Chem-Impex International (Wood Dale, IL). DOWEX™ 50WX8 strong cation exchange resin was purchased from Sigma-Aldrich. Biotage reverse phase columns (SNAP-KP-C18-HS) were purchased from Biotage.

General Procedures:

Standard molecular cloning procedures were followed.[3] Reactions were monitored using HPLC (Agilent 1200 UHPLC with an Agilent Eclipse Plus C18 4.6 x 150 mm column, 3.5 μM particle size; C18 4.6 x 50 mm column, 3.5 μM particle size; and C18 2.1 x 50 mm column, 1.8 μM particle size; solvent A = H_2O/0.1% TFA, solvent B = CH_3CN). Gel filtration was performed using a HiLoad 16/600 Superdex 200 column (GE Healthcare Life Sciences). Reverse phase preparative chromatography was carried out using a Biotage Isolera One. ^1H spectra were recorded at 500 MHz on a Bruker DMX-500 or DRX-500 spectrometer, and chemical shifts are reported relative to residual solvent peaks.[4] The T_m's of enzymes were found using a AVIV 202 CD Spectrometer with Peltier temperature controller and SigmaPlot (Systat Software, San Jose, CA). Crystallization data were collected at NE-CAT beamline 24-ID-E at the Advanced Photon Source at Argonne National Laboratory.

Experimental Procedures:

Library construction, expression, and screening: All genes encoding RebH were cloned into pET-28a between the NdeI and HindIII digestion sites. Mutant libraries were constructed by error-prone PCR, using *Taq* polymerase with 150 μM $MnCl_2$ (round 1) or 100 μM $MnCl_2$ (rounds 2 and 3). PCR was performed in a volume of 50 μL with conditions of 95 °C 30 s, (95 °C 30 s, 55 °C 30 s, 72 °C 90 s) for 20 cycles, 72 °C 10 min. Beneficial mutations were recombined via overlap extension [5] with PCR conditions of 98 °C 30 s, (98 °C 10 s, 72 °C 50 s) for 35 cycles, 72 °C 10 min. Plasmids were transformed by electroporation into *E. coli* containing the chaperone pGro7. Library colonies were picked using an automated colony picker (Norgren Systems) and arrayed in 1-ml 96-well plates containing 300 μL LB with 50 μg/mL kanamycin and 20 μg/mL chloramphenicol. Cells were grown overnight at 37 °C, 250 rpm, and 50-100 μL of overnight culture was used to inoculate 1 mL TB (with 50 μg/mL kanamycin and 20 μg/mL chloramphenicol) in 2-mL 96-well plates. Following growth at 37 °C, 250 rpm, to an OD_{600} = 0.9-1, enzyme expression was induced with IPTG and arabinose to final concentrations of 10 μM and 0.2 mg/mL, respectively. Protein expression continued for ~20 h at 30 °C, 250 rpm, after which cultures were harvested by centrifugation and stored at -80 °C until use.

Cell pellets were thawed and suspended in 100 µL HEPES buffer (25 mM, pH 7.4) containing 0.75 mg/mL lysozyme. After incubation at 37 °C, 250 rpm, cells were flash frozen in liquid nitrogen and thawed in a 37 °C water bath. Ten microliters of DNaseI at 1 mg/mL were added and the cells incubated at 37 °C, 250 rpm, for 15 min. After centrifugation, 50 µL of supernatant were transferred to a microtiter plate for screening.

Libraries were sealed (AeraSeal, Research Products International), incubated at 42 °C for 2 h (round 1), 51 °C for 2 h (round 2), or 54 °C for 3 h (round 3) and then immediately cooled in an ice water bath. Similar to what has been described previously for halogenation reactions,[6] tryptophan (0.5 mM final concentration) was added to 50 µL lysate as a solution of 25 mM HEPES (pH 7.4). NaCl (10 mM final concentration), FAD (100 µM final concentration), NAD (100 µM final concentration), MBP-RebF (2.5 µM final concentration) and glucose dehydrogenase (50 U/mL final concentration) were also added to reactions as solutions of 25 µM HEPES buffer. A solution of glucose (20 mM final concentration) was added to the reaction mixtures to initiate. Reactions were mixed, the plates sealed, and left overnight on the benchtop. Reactions were quenched with an equal volume of methanol and centrifuged, and the supernatant was filtered and analyzed for 7-chlorotryptophan production via HPLC (Agilent 1200 UHPLC with an Agilent Eclipse Plus C18 2.1 x 50 mm column, 1.8 µM particle size; solvent A = H_2O/0.1 % TFA, solvent B = CH_3CN; 0-0.5 min, B = 16 %; 0.5-1.5 min, B = 16-80 %).

Enzyme purification and residual activity determination: The MBP-RebF and RebH used for 1 and 10 mg bioconversions was grown, expressed, lysed and purified according to a previous report.[6] Enzyme expression and purification procedures for analytical bioconversions were adapted from this report.[6] An overnight starter culture was used to inoculate 50 mL TB (with 50 µg/mL kanamycin and 20 µg/mL chloramphenicol). Following growth at 37 °C, 250 rpm, until OD_{600} = 0.6-0.8, enzyme expression was induced with IPTG and arabinose to final concentrations of 100 µM and 2 mg/mL, respectively. Protein expression continued for ~20 h at 30 °C, 250 rpm, after which cultures were harvested by centrifugation and stored at -80 °C until use. Cell pellets were thawed, suspended in cold 15 mL 20 mM HEPES (pH 7.4), 150 mM NaCl, and lysed by sonication while kept on ice (Qsonica S-4000 with a 0.5" horn; 8 x 30 s with 45 s rests, 20 % duty cycle delivering 40-50 W). After clarification by centrifugation, halogenases were purified by Ni-NTA affinity chromatography and exchanged into a buffer of 20 mM HEPES (pH 7.4), 150 mM NaCl, and

10 % glycerol. For crystallography, mutant RebH was further purified by gel filtration chromatography using a HiLoad 16/600 Superdex 200 column (GE Healthcare Life Sciences) into a buffer of 20 mM HEPES (pH 7.4). Protein concentration was determined using A_{280} and extinction coefficients calculated based on amino acid composition (Protein Calculator v3.3, http://www.scripps.edu/~cdputnam/protcalc.html).

The residual activity was determined following incubation of 50 µL of pure protein at 49 °C for 2 h in 1.5-mL microcentrifuge tubes. Tryptophan halogenation reactions consisted of the same reagents used during library screening with the following exceptions: pure protein was substituted for lysate, and the buffer was 20 mM HEPES (pH 7.4), 6.7% glycerol, and 100 mM NaCl. Reactions were conducted overnight on the benchtop, and quenched and analyzed the following day like the library reactions.

T_m and T_{opt} analyses: Melting temperature measurements were conducted in 20 mM HEPES (pH 7.4), 150 mM NaCl, and 10 % glycerol, with a protein concentration of 20 µM. Thermal denaturation was irreversible and monitored by circular dichroism spectroscopy using an AVIV 202 CD Spectrometer with Peltier temperature controller. Denaturation was monitored at 222 nm in 2 °C increments from 20-90 °C with 2 min equilibration at each temperature. The midpoint of the denaturation curve was determined with SigmaPlot (Systat Software, San Jose, CA) after fitting to a 4-parameter sigmoid.

Activity-temperature profiles were constructed using purified enzyme with 75 µL reactions in 1.5-mL microcentrifuge tubes. Reactions were conducted in a buffer of 20 mM HEPES (pH 7.4), 6.7% glycerol, and 100 mM NaCl, with 0.5 mM L-tryptophan, 20 mM DTT, and 100 µM FAD. Reactions were conducted overnight at temperatures ranging from 21-45 °C. These were quenched and analyzed by the same methods as the library reactions.

General Procedure for 1 and 10 mg Bioconversions[6]: Substrate (1 or 10 mg) was added to a beaker (100 mL) or crystallization dish (100 x 50 mm) as a solution in HEPES buffer (25 mM HEPES, pH = 7.4) or isopropanol (5% v/v). Solutions of NAD (0.2 equiv., 100 µM final concentration), FAD (0.2 equiv., 100 µM final concentration), NaCl (200 equiv., 100 mM final concentration), and a glucose dehydrogenase (50 U/mL final concentration GDH) were added to the reaction. This was diluted to the appropriate volume with HEPES buffer, and RebH (0.002-0.1 equiv., 1-50 µM final concentration) and MBP-RebF (0.005 equiv., 2.5 µM final concentration) were added as solutions of HEPES/glycerol buffer (25 mM HEPES, pH

7.5, 10% glycerol v/v). The reaction was initiated with a solution of 1 M glucose (40 equiv., 20 mM final concentration), sealed with an AeraSeal film, and placed in incubator at the desired temperature. During initial screens of substrates at different temperatures, it was found that wild-type RebH halogenates in higher yield when agitated, whereas 3-LSR has optimal activity in the absence of shaking. Because of this, bioconversions with wild-type enzyme were agitated with 110 rpm; those with 3-LR or 3-LSR were not agitated.

To the 1 mg bioconversions, an internal standard (phenol, 0.5 mM final concentration) was added as a 10 mM solution in DMSO after quenching with one volume of methanol. These reactions were analyzed using the HPLC (Agilent 1200 UHPLC with an Agilent Eclipse Plus C18 4.6 x 150 mm column, 3.5 μM particle size; solvent A = H_2O/0.1% TFA, solvent B = CH_3CN). The following method was used for all substrates: 0-10 min, B = 15%; 10-20 min, B = 15-100%; 20-24 min, B = 100%. The area of product to internal standard was used to compare the activity of the mutant enzymes to wild-type.

The 10 mg bioconversions were quenched with HCl (1 M, until pH<2) and saturated with NaCl. Precipitated protein was filtered out through a pad of Celite. The filtrate was extracted into CH_2Cl_2 or submitted to strong cation exchange chromatography. The crude material was purified by normal or reverse phase chromatography.

Procedure for Strong Cation Exchange Chromatography[6]: After protein removal from bioconversions, aqueous filtrate was submitted to strong cation exchange to remove salts if extraction into organic solvent was not possible. DOWEX™ 50WX8 resin was slurry-packed with methanol in a 250 mL chromatography column. The resin was washed with ~300 mL of methanol and ~300 mL of deionized water. The resin was acidified with HCl (1 M) until the pH of flow through was less than 2. The resin was washed with deionized water until neutral. Acidic filtrate (pH < 2) was added to resin. The resin was washed with ~500 mL of deionized water. The product was eluted with NH_4OH (1 M). Product-containing fractions were concentrated to dryness using a rotary evaporator under high vacuum.

Detailed Isolation and Characterization:

7-chloro-L-tryptophan (1): The 10 mg bioconversion was conducted according to the general procedure. Substrate was added in a solution of HEPES (25 mM HEPES, pH = 7.4). RebH was added to a final concentration of 2 μM (0.004 equiv.) and the reaction was incubated at

40°C. After reaction completion, the reaction mixture was filtered through Celite, submitted to strong cation exchange chromatography, and dry-loaded onto Celite. This was packed into a Biotage samplet, loaded into a reverse phase column (Biotage SNAP-KP- C18-HS) and purified by reverse phase chromatography (gradient from water 0.1% TFA to 15% CH_3CN/water 0.1%TFA) to afford **1** in 69% yield (11.9 mg of **1**·TFA, 0.034 mmol). ^{1}HNMR spectrum was consistent with previous reports of this compound.[6] ^{1}H NMR (500 MHz, Deuterium Oxide with one drop TFA). δ 7.35 (d, *J* = 5.7 Hz, 1H), 7.13 (s, 1H), 7.04 (d, *J* = 7.6 Hz, 1H), 6.88 (t, *J* = 7.8 Hz, 1H), 4.15 (t, *J* = 4.8 Hz, 1H), 3.23 (m, 2H).

7-chloro-2-methyltryptamine (2): The 10 mg bioconversion was conducted according to the general procedure. Substrate was added in a solution of isopropanol (5% v/v). RebH was added to a final concentration of 5 μM (0.01 equiv.) and the reaction was incubated at 40°C. After reaction completion, the reaction mixture was filtered through Celite, extracted into CH_2Cl_2, and dry-loaded onto Celite. This was packed into a Biotage samplet, loaded into a reverse phase column (Biotage SNAP-KP- C18-HS) and purified by reverse phase chromatography (gradient from water 0.1% TFA to 15% CH_3CN/water 0.1%TFA) to afford **2** in 56% yield (10.4 mg of **2**·TFA, 0.032 mmol). ^{1}HNMR spectrum was consistent with previous reports of this compound.[6] ^{1}H NMR (500 MHz, Methanol-d_4). δ 7.40 (d, *J* = 7.8 Hz, 1H), 7.06 (d, *J* = 7.6 Hz, 1H), 6.98 (t, *J* = 7.7 Hz, 1H), 3.16-3.10 (m, 2H), 3.07-3.02 (m, 2H), 2.43 (s, 3H).

1-chloro-2-aminonaphthalene (3): The 10 mg bioconversion was conducted according to the general procedure. Substrate was added in a solution of isopropanol (5% v/v). RebH was added to a final concentration of 4 μM (0.008 equiv.) and the reaction was incubated at 21°C. After reaction completion, the reaction mixture was filtered through Celite and extracted into CH_2Cl_2. The crude material was purified by flash column chromatography (SiO_2, 1% ethyl acetate/hexanes) to afford **3** in 62% yield (7.7 mg of **3**, 0.043 mmol). ^{1}HNMR spectrum was consistent with previous reports of this compound.[6] ^{1}H NMR (500 MHz, Methylene Chloride-d_2). δ 8.00 (d, *J* = 8.5 Hz, 1H), 7.70 (d, *J* = 8.1 Hz, 1H), 7.59 (d, *J* = 8.7 Hz, 1H), 7.49 (t, *J* = 7.7 Hz, 1H), 7.27 (t, *J* = 8.0 Hz, 1H), 7.02 (d, *J* = 8.7 Hz, 1H), 3.58 (s, 3H).

5-chloro-tryptoline and 6-chloro-tryptoline (4): The 10 mg bioconversion was conducted according to the general procedure. Substrate was added in a solution of isopropanol (5% v/v). RebH was added to a final concentration of 50 μM (0.1 equiv.) and the reaction was incubated at 21°C. After reaction completion, the reaction mixture was filtered through Celite, extracted into CH_2Cl_2, and dry-loaded onto Celite. This was packed into a Biotage samplet, loaded into a reverse phase column (Biotage SNAP-KP- C18-HS) and purified by reverse phase chromatography (gradient from water 0.1% TFA to 30% CH_3CN/water 0.1%TFA) to afford **4** in 67% yield (12.5 mg of **4**·TFA, 0.039 mmol). ^{1}HNMR spectrum was consistent with previous reports of this compound.[6] ^{1}H NMR (500 MHz, Methanol-d_4). δ 7.47 (s, 0.80H), 7.44 (d, J = 8.4 Hz, 0.12H), 7.36 (s, 0.10H), 7.32 (d, J = 8.6 Hz, 0.81H), 7.11 (dd, J = 8.6, 2.0 Hz, 0.80H), 7.04 (dd, J = 8.4, 1.8 Hz, 0.10H), 4.44 (s, 2H), 3.58 (t, J = 6.1 Hz, 2H), 3.05 (t, J = 6.0 Hz, 2H).

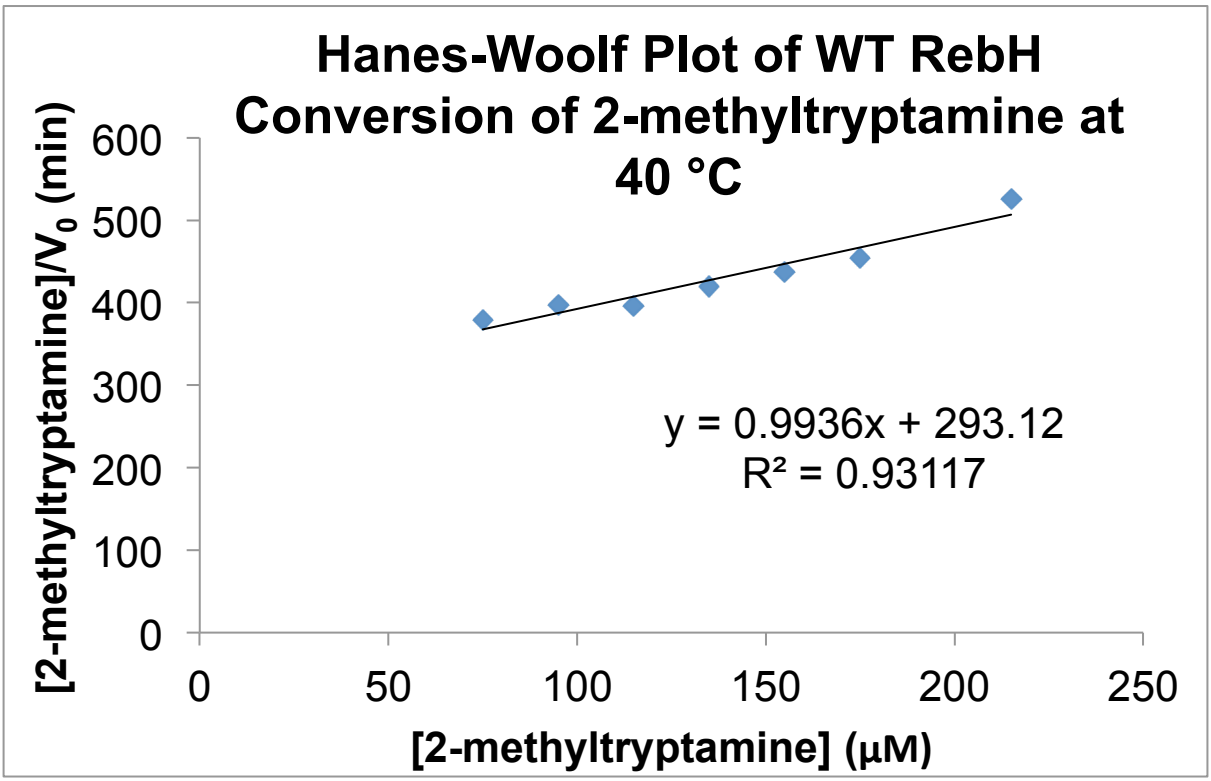
Determination of Kinetic Parameters for WT RebH and 3-LSR: Kinetic parameters were determined for WT RebH and 3-LSR at 21 °C and 40 °C with the unnatural substrate 2-methyltryptamine. Rates were determined by monitoring the conversion of 75-215 μM substrate in the presence of NAD (100 μM final concentration), FAD (100 μM final concentration), NaCl (100 mM final concentration), MBP-RebF (2.5 μM final concentration), glucose dehydrogenase (50 U/mL final concentration GDH), glucose (20 mM final concentration), and phenol as an internal standard (0.5 mM final concentration) at a final volume of 75 μL in a microtiter plate. WT RebH or 3-LSR was added at a final concentration of 4-25 μM (20 μM WT RebH at 21 °C; 25 μM 3-LSR at 21 °C; 4 μM WT RebH at 40 °C; 8 μM 3-LSR at 40 °C). The reactions were left shaking at 600 rpm in an IKA® microtiter shaker at room temperature, then quenched at 30-105 minutes (all time points were collected in triplicate) by addition of 75 μL of MeOH. The precipitated protein was then removed by centrifugation and the reactions were filtered and analyzed by HPLC (Agilent 1200 UHPLC with an Agilent Eclipse Plus C18 4.6 x 50 mm column, 3.5 μM particle size; solvent A = H_2O/0.1% TFA, solvent B = CH_3CN) using the following method: 0-2 min, B = 15%; 2-2.5 min, B = 15-20%; 2.5-4.5 min, B = 20%; 4.5-5.5 min, B = 20-30%; 5.5-6 min, B = 30-70%; 6-7 min, B = 70%. Product formation was quantitated by calculating the ratio of product to internal standard and fitting that value to a calibration curve prepared from known concentrations of each material. The kinetic parameters (K_m and k_{cat}) for each substrate were

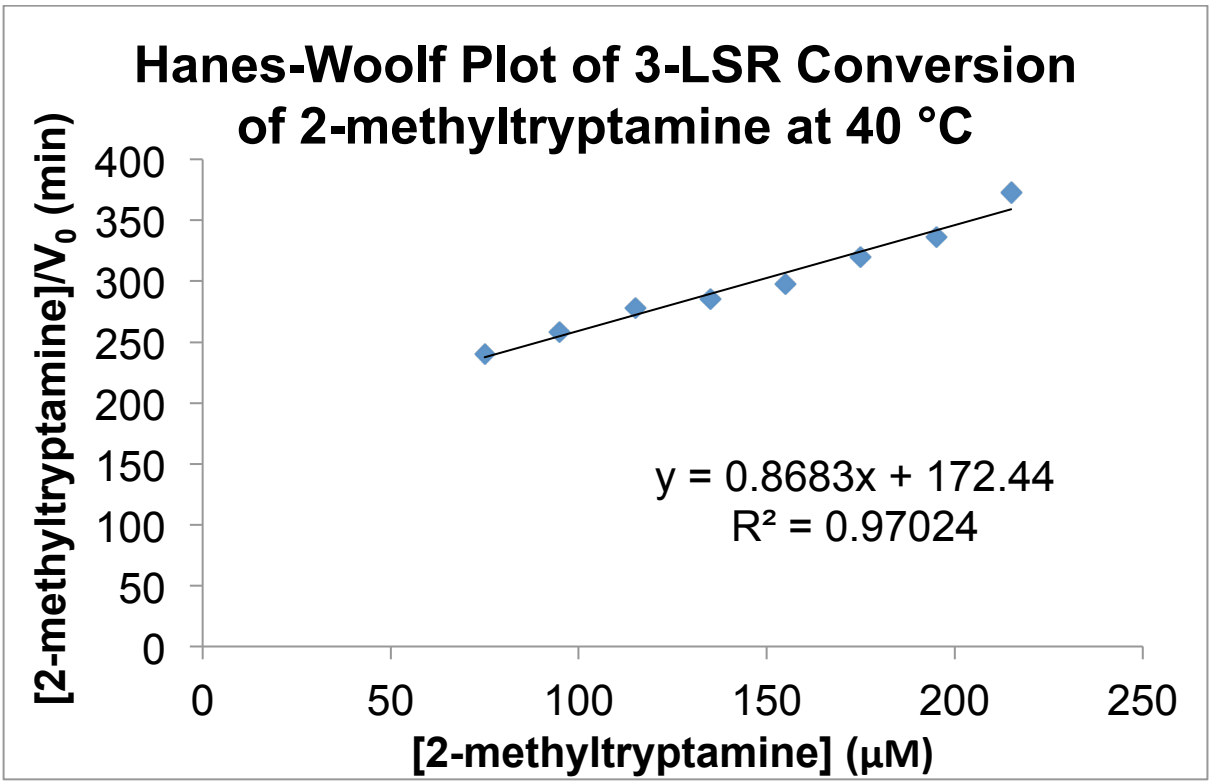
determined using the Hanes-Woolf plots (see p. 10-11) constructed from the substrate concentrations and the observed initial rates.

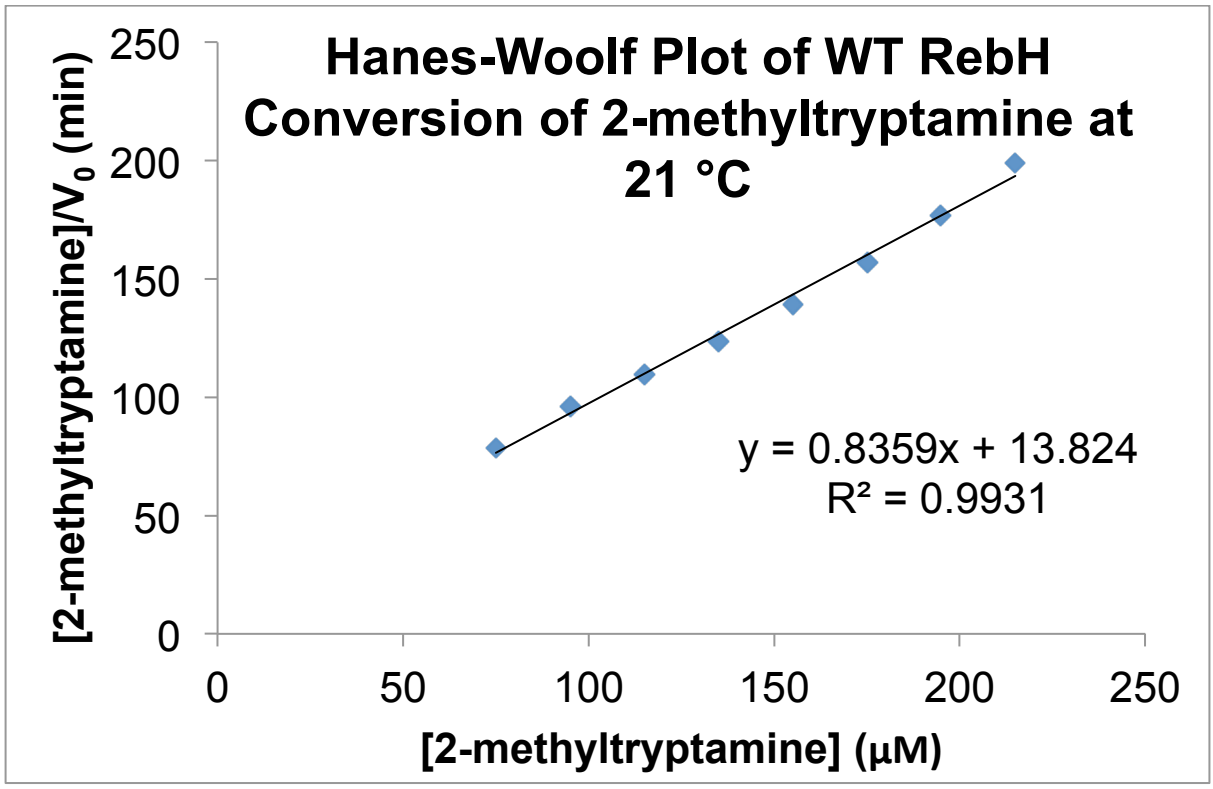
Time Courses: Conversion vs. time courses for halogenation by WT RebH and 3-LSR were plotted for the substrate 2-methyltryptamine. 10 mg bioconversions were conducted according to the general procedures for 1 and 10 mg bioconversions. WT RebH and 3-LSR bioconversions were each conducted at a final concentration of 5 μM RebH. The WT bioconversion was mixed at 110 rpm. 3-LSR was not agitated. Aliquots (75 μL) were pulled from 10 mg bioconversions at various times, were quenched with MeOH (75 μL) and were analyzed using the HPLC method described in the general procedures for 1 and 10 mg bioconversions.

Crystallization and structure determination: Purified protein was concentrated to 11 mg/mL, and crystals were grown at 20 °C using the hanging drop vapor diffusion method with a reservoir solution of 1.4 M Na/K phosphate buffer (pH 6.8). Rod-like crystals grew in 2-3 weeks and were flash frozen in liquid nitrogen following cryoprotection with the reservoir solution supplemented with 16 % glycerol. Data were collected at NE-CAT beamline 24-ID-E at the Advanced Photon Source at Argonne National Laboratory, and processed using HKL2000.[7] Phases were determined via molecular replacement using Phaser [8] and wild-type RebH (PDB ID 2OAM) as the search model. Manual model building was performed in Coot,[9] and the structure was refined with PHENIX.[10] Figures were prepared with PyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC). Structural data have been deposited in the RCSB Protein Data Bank with accession code 4LU6.

Hanes-Woolf Plots:






Hanes-Woolf Plot of WT RebH Conversion of 2-methyltryptamine at 21 °C
[2-methyltryptamine]/V0 (min)
250
200
150
100
50
0
0
50
100
150
200
250
[2-methyltryptamine] (µM)
y = 0.8359x + 13.824
R² = 0.9931

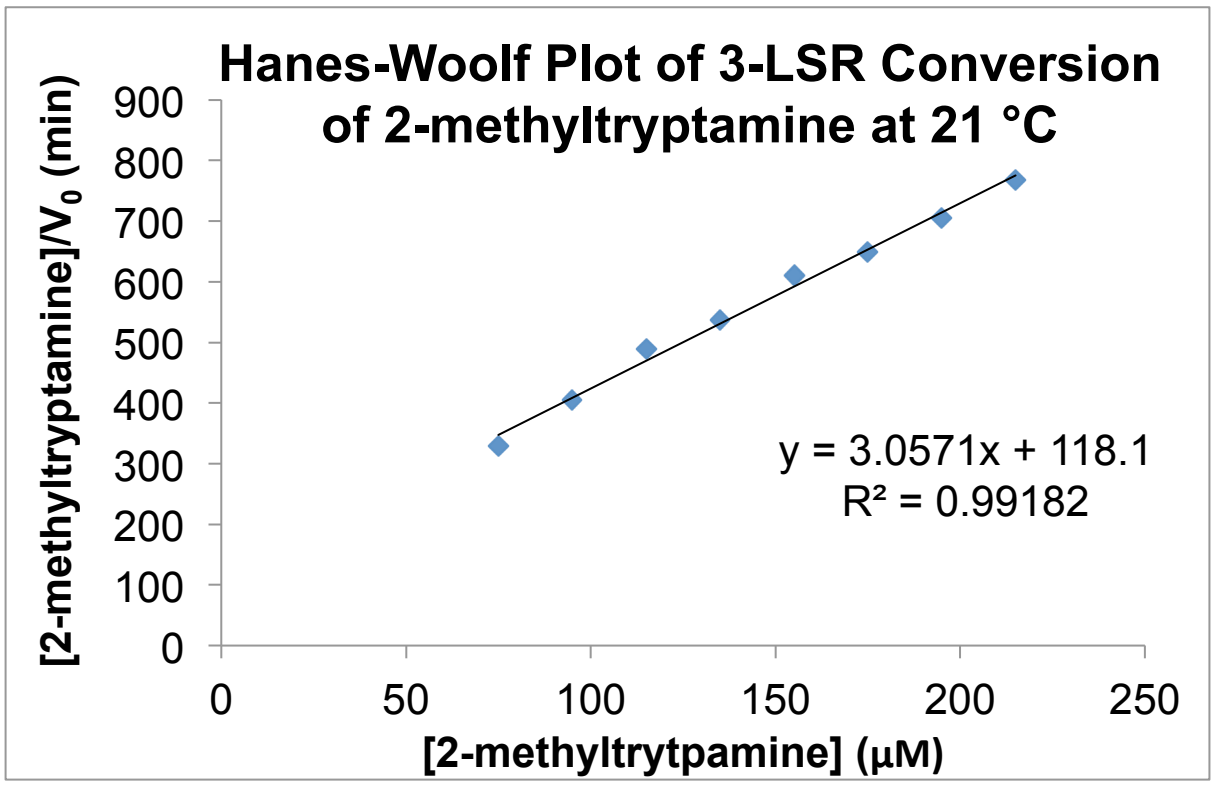

Hanes-Woolf Plot of 3-LSR Conversion of 2-methyltryptamine at 21 °C
[2-methyltryptamine]/V0 (min)
900
800
700
600
500
400
300
200
100
0
0
50
100
150
200
250
[2-methyltrytpamine] (µM)
y = 3.0571x + 118.1
R² = 0.99182

Table S1. RebH 3-LSR Data Collection and Refinement Statistics

Data Collection and Crystal Parameters	
Space group	$P6_2$
Cell dimensions	
a, b, c (Å)	115, 115, 230
α, β, γ (°)	90, 90, 120
Wavelength (Å)	0.97919
Resolution (Å)	80-3.05 (3.10-3.05)[a]
R_{merge} (%)	25.0 (89.6)
I/σI	7.7 (2.0)
Completeness (%)	99.3 (99.8)
Multiplicity	4.0 (4.0)
Refinement	
Resolution (Å)	75.2-3.05 (3.16-3.05)
R_{work}/R_{free} (%)	17.8/23.8 (26.6/34.4)
Rmsd bond lengths (Å)	0.01
Rmsd bond angles (°)	1.3
Ramachandran plot statistics (%/#)	
Preferred regions	94.4/991
Allowed regions	5.6/59
Outliers	0/0
PDB ID	4LU6

[a]Values for the highest resolution shell are shown in parentheses.

References

[1] E. Yeh, S. Garneau, C.T. Walsh, *Proc. Nat. Acad. Sci.* **2005**, *102*, 3960-3965.

[2] L.D. Cabrita, W. Dai, S.P. Bottomley, *BMC Biotech.* **2006**, *6*, 12.

[3] J. Sambrook, E.F. Frisch, T. Maniatis (**1989**) *Molecular cloning: a laboratory manual*. 2, Cold Spring Harbor Laboratory Press, New York.

[4] H.E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem*. **1997**, *62*, 7512-7515.

[5] K. L. Heckman, L. R. Pease, *Nat. Protoc*. **2007**, *2*, 924-932.

[6] J. T. Payne, M. C. Andorfer, J. C. Lewis, *Angew. Chem. Int. Ed*. **2013**, *52*, 5271-5274.

[7] Z. Otwinowski, W. Minor, *Methods Enzymol.* **1997**, *276*, 307-326.

[8] A. J. McCoy, R. W. Grosse-Kunstleve, L. C. Storoni, R. J. Read, *Acta Crystallogr. D Biol. Crystallogr.* **2005**, *61*, 458-464.

[9] J. Fauvarque, *Pure Appl. Chem.* **1996**, *68*, 1713-1727.

[10] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213-221.