Supporting Information

Khare et al. 10.1073/pnas.1006738107

SI Text

Materials and Methods. ***Cloning, expression, and protein purification.*** DNA encoding the Hal domain (CurA residues 1600–1919, 320 amino acids, 37.1 kDa) was amplified from plasmid pET28b::Cur Hal encoding CurA residues 1600–1930 (1), using primers for ligation-independent cloning (Table S2). The amplified product was cloned into the pMCSG7 vector (2) to create pMCSG7::Hal. To prepare wild-type Hal, *Escherichia coli* BL21(DE3) cells were transformed with pMCSG7::Hal, grown at 37 °C in LB medium to an OD_{600} of 0.8, transferred to 18 °C, induced with 0.4 mM IPTG, grown for 18 h, harvested, and lysed by ultrasonication on ice. Soluble protein was bound to a HisTrap Ni^{2+} column (GE Healthcare), washed with Buffer A (100 mM Tris pH 8.0, 500 mM NaCl) containing 5 mM imidazole, and eluted by a linear gradient of 5 mM–500 mM imidazole in Buffer A. Purified Hal was gel filtered on Superdex 200 column (GE healthcare) equilibrated with 10 mM Tris pH 8.0, 25 mM NaCl, 3 mM αKG, and reconstituted by addition of 0.25 mM $(NH_4)_2Fe(SO_4)_2$. Hal eluted with an apparent molecular mass of 80 kDa, corresponding to a dimer, and was concentrated to 10 mg/mL and stored at −80 °C. Selenomethionyl (SeMet) Hal was produced and purified as for the wild type, except that cells of *E. coli* strain B834(DE3), a methionine auxotroph, were transformed with pMCSG7::Hal and grown in SeMet minimal media (Molecular Dimensions). SeMet Hal protein was concentrated to 7 mg/mL and stored at −80 °C. Mutant proteins were generated by site directed mutagenesis of pMCSG7::Hal using primers listed in Table S2 and produced as for the wild type through the nickel affinity chromatography step (Fig. S5). Fractions containing Hal were pooled, concentrated, and dialyzed against 10 mM Tris pH 7.8, 50 mM NaCl, 10% glycerol. The mutant proteins were reconstituted with 3 mM αKG and freshly prepared 0.25 mM $(NH_4)_2Fe(SO_4)_2$.

Crystallization and data collection. Crystals of four forms were obtained at 20 °C by the sitting drop vapor diffusion method from the protein stock solutions described above. Crystal form I (apo Hal) grew from a 1:1 mixture of protein stock and reservoir solution (200 mM sodium citrate, 20% PEG 3350). Crystals were cryoprotected by quick passage through a solution containing reservoir solution with 10% (v/v) glycerol. Crystals of form II (holo Hal) grew from a 1:1 mixture of protein and reservoir solution (4.0 M sodium formate) and were cryoprotected in the mother liquor supplemented with 15% glycerol (v/v). Crystal form III was obtained out of a mixture of equal volume of Hal and the precipitant (100 mM sodium fluoride and 18% PEG 3350). Crystal form IV was obtained by replacing αKG with 3 mM sodium succinate during protein purification and crystals grew under conditions similar to form II. Wild-type and SeMet Hal crystallized in form I under identical conditions. The unbuffered crystallization mixtures had measured pH in the range of 8–9. All crystals were flash frozen in liquid nitrogen. All diffraction data were collected at 100 K on beamline 23ID-D (GM/CA CAT) at the Advanced Photon Source (APS). For structure determination, a 2.6-Å single-wavelength anomalous diffraction (SAD) dataset was collected from a crystal of SeMet Hal at the energy of peak absorption above the Se K edge ($\lambda = 0.9794$ Å). A 2.2-Å dataset from a crystal of wild-type Hal (form II) was collected at a wavelength of 0.9793 Å. The presence of Fe in form II crystals was confirmed by recording fluorescence spectra of crystals using an excitation energy of 12.66 keV. Data were integrated and scaled with program HKL2000 (3). A summary of the data statistics for crystal form I (apo Hal, *I*23, 1 dimer per asymmetric unit), crystal form II (holo Hal, $P4_32_12$, one polypeptide per asymmetric unit, dimer formed by crystallographic symmetry), crystal form III (holo Hal, $P2_1$, 1 dimer per asymmetric unit) and crystal form IV ($P2_12_12_1$, 1 dimer per asymmetric unit) is given in Table S1.

Structure determination and refinement. The structure was solved by SAD using SeMet Hal in crystal form I. 12 of 14 expected Se positions were located based on the Bijvoet differences in the SAD dataset using SOLVE (4) and the resulting experimental map was improved by density modification using RESOLVE (5), both programs running within the PHENIX suite (6). A partial model was built with the program RESOLVE. Manual adjustment and rebuilding of the model was done using COOT (7). All excepting the two C-terminal residues were visible in two or more crystal structures (Fig. S1). Crystallographic refinement was carried out using both PHENIX and Refmac5 (8). Although Hal was reconstituted with $(NH_4)_2Fe(SO_4)_2$ and αKG prior to crystallization, citrate in the crystallization solution chelated Fe^{3+} ($K_{stability} = 11.85$ (9)), accounting for the lack of iron in the form I structure.

The structure in crystal form II was solved by molecular replacement using the model from crystal form I after deletion of 11 residues encoded by the vector preceding the natural Hal sequence. In the active site of the form II structure, the position of iron and αKG could be seen clearly in different electron density maps (Fig 2*A* and Fig. S2*A*). Fluorescence scans at the iron *K* edge confirmed the presence of iron. Although Hal was reconstituted with ferrous iron, Fe^{3+} is assumed to be the bound species because Hal was crystallized under aerobic conditions. During refinement additional positive density appeared in the F_o-F_c electron density map at the position of a modeled water molecule near the Fe^{3+}. The large peak suggested the presence of a more electron dense species than water (Fig S2*A*). A chloride ion coordinating Fe^{3+} was modeled in the electron density. Another positive density within coordination distance of the Fe^{3+} was assigned to formate, which was present in the crystallization solution.

The structure in crystal form III (two polypeptides per asymmetric unit) was solved readily by molecular replacement using one subunit from the form I structure and the other from the form II structure. Crystal form IV, related to form II, was solved from the form II structure. Iron was missing from the form IV structure, apparently due to chelation by succinate ($K_{stability} = 7.49$ (9)) in the crystallization solution. Statistics for all four crystal structures are summarized in Table S1. All crystallographic models were validated with MolProbity (10), structure similarity searches were done with DALI (11), sequences were aligned with Clustal (12), subunit interfaces were quantitated with PISA (13), and figures were prepared using Pymol (http://www.pymol.org).

Halogenation assay. The Hal reaction, in a 110-μL mixture containing 5 μM Hal, 10 mM Tris pH 7.8, 100 μM fresh $Fe(NH_4)_2(SO_4)_2$, 25 μM (S)-HMG ACP and 1 mM αKG, was initiated by addition of enzyme, incubated at 24 °C for 10 min, and quenched by addition of 20 μL formic acid. Aliquots were analyzed by reverse phase HPLC using a Source 15RPC column (GE Healthcare) with a linear gradient from 30%–70% CH_3CN. Fractions containing substrate (S)-HMG ACP and product Cl-(S)-HMG ACP, which eluted together, were pooled and analyzed by FTICR-MS and infrared multiphoton dissociation (IRMPD).

Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) analysis. Samples were prepared for FTICR-MS analysis as

previously described (1) All samples were analyzed with an actively shielded 7 Tesla quadrupole-FTICR-MS (APEX-Q, Bruker Daltonics, Billerica, MA). Target analytes were directly infused into an electrospray ionization (ESI) source (Apollo II, Bruker Daltonics) operating in positive-ion mode at a flow rate of 70 μL/h. Multiply protonated ions generated by ESI were externally accumulated in a hexapole and transferred via high voltage ion optics to the ICR cell for analysis. For infrared multiphoton dissociation (IRMPD), ions were irradiated for 100 ms by 10.6 μm photons at 10 W laser power (25 W CO_2 laser, Synrad, Mukilteo, WA) without precursor ion selection. All data were acquired with XMASS software (version 6.1, Bruker Daltonics) in broadband mode from $m/z = 200$–2,000 with 512 K data points and summed over 10–20 scans. Mass spectra were analyzed with the MIDAS analysis software (14). External frequency-to-m/z calibration was performed if needed with a two-term calibration equation (15) based on the 10^+ and 11^+ charge states of ubiquitin (Sigma).

IRMPD-based quantification. Recently we showed that wild-type Hal specifically converts (*S*)-HMG-ACP to Cl-(*S*)-HMG-ACP in presence of iron and αKG (1). In that work, we also showed that IRMPD-based phosphopantetheine (PPant) ejection assay (16) releases PPant-bound Cl-HMG from Cl-(*S*)-HMG-ACP, resulting in a product ion at m/z 439.12, whereas the nonchlorinated PPant-HMG fragment ejected from (*S*)-HMG-ACP yields a product ion at m/z 405.16. The relative ratio of these PPant ejection products were used to estimate the conversion efficiency of the (*S*)-HMG-ACP substrate to the Cl-(S)-HMG-ACP product. Hal-catalyzed addition of a single chlorine atom to the ACP-bound substrate is unlikely to influence ESI efficiency. Therefore, the ratio of the (*S*)-HMG-ACP and Cl-(*S*)-HMG-ACP PPant ejection products represents the Hal reaction efficiency. A time course experiment was done with the wild-type Hal (Fig. S3). After 10 min, wild-type Hal converted most of the substrate to the chlorinated form, and this reaction time was chosen to monitor the IRMPD product ions at m/z 405 or 439 on IRMPD formed by Hal wild type and mutant forms (Figs. S4 and S5).

1. Gu L, et al. (2009) Metamorphic enzyme assembly in polyketide diversification. *Nature* 459:731–735.
2. Stols L, et al. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* 25:8–15.
3. Otwinowski Z, Minor W, (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology* 276:307–326.
4. Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D* 55:849–861.
5. Terwilliger TC (2000) Maximum-likelihood density modification. *Acta Crystallogr D* 56:965–972.
6. Adams PD, et al. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D* 58:1948–1954.
7. Emsley P, Cowtan K (2004) COOT: model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
8. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
9. Sillén LG, Martell AE (1964) *Stability constants of metal-ion complexes* (The Chemical Society, Burlington House, London) 2nd Ed.
10. Davis IW, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–383.
11. Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. *Bioinformatics* 24:2780–2781.
12. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
13. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372:774–797.
14. Senko MW, Canterbury JD, Guan S, Marshall AG (1996) A high-performance modular data system for Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Commun Mass Spectrom* 10:1839–1844.
15. Ledford EB, Jr., Rempel DL, Gross ML (1984) Space charge effects in Fourier transform mass spectrometry. Mass calibration. *Anal Chem* 56:2744–2748.
16. Dorrestein PC, et al. (2006) Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry. *Biochemistry* 45:12756–12766.

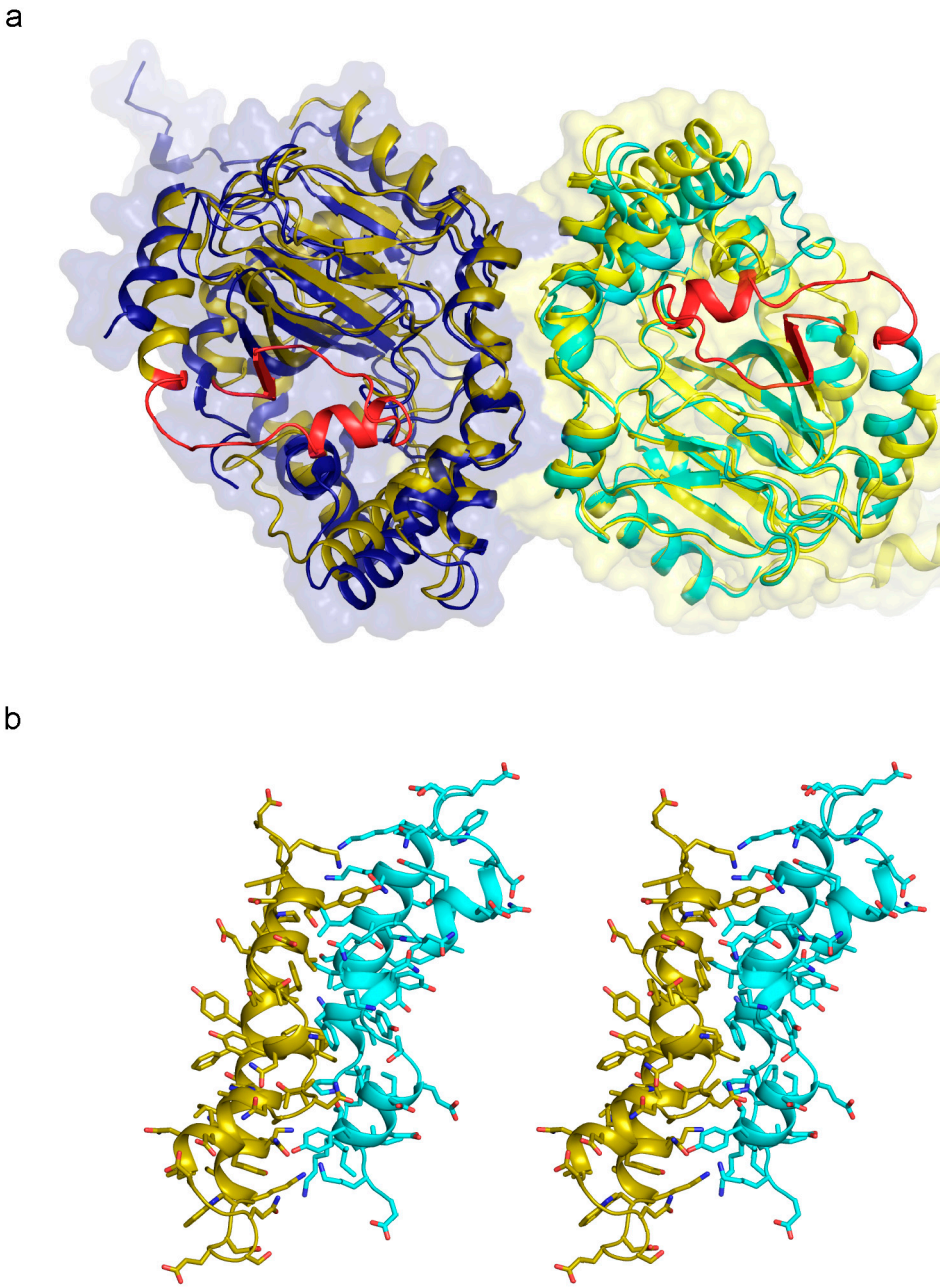

Fig. S1. Cur Hal dimer interface. (*A*) Superposition of open-form and closed-form Hal dimer. Open-form Hal from crystal form I (blue monomer at left and yellow monomer at right) is shown with closed-form Hal from crystal form II (yellow monomer at left and cyan monomer at right with active-site lid in red). The transparent surface for the open form shows shape complementarity of monomer surfaces in the interface. The superposition is based on the core of the subunits on the right. (*B*) Side chain packing in the dimer interface. Interacting regions for the closed-form dimer (subunits in yellow and cyan) are shown in stereo with side chains in stick form (blue N, red O, C colored as the helices). The view in both *A* and *B* is along the dimer axis.

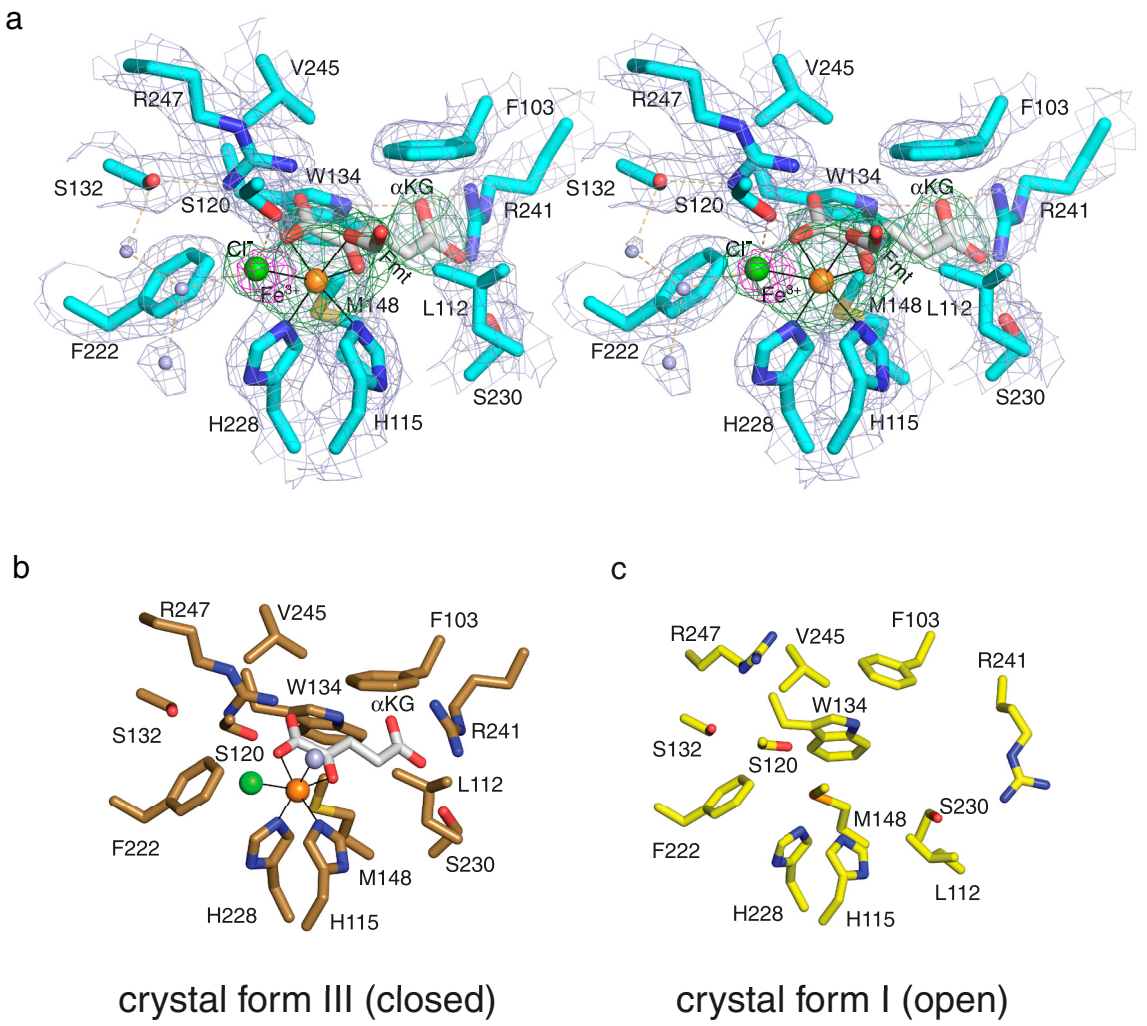

Fig. S2. Cur Hal active site. (*A*) Electron density at the Fe center of closed-form Cur Hal. The F_o-F_c omit map for the active-site ligands is colored in green and contoured at 5σ. The F_o-F_c omit map for Cl^- is colored in magenta and contoured at 6σ. The 2F_o-F_c map is colored in blue and contoured at 1.5σ. Protein side chains are shown with cyan C, and αKG and formate (Fmt) are shown with white C; other atoms are blue N, red O and yellow S. Single atoms are represented as spheres with Fe^{3+} in orange, Cl^- in green and water O in light blue. (*B*) Closed-form Hal with Fe^{3+}/aKG/Cl^-. Iron has a complete ligand sphere in Subunit A (brown C) of the mixed-form dimer in crystal form III. (*C*) Open-form Hal with no ligands. This is the active site structure in crystal form I (yellow C).

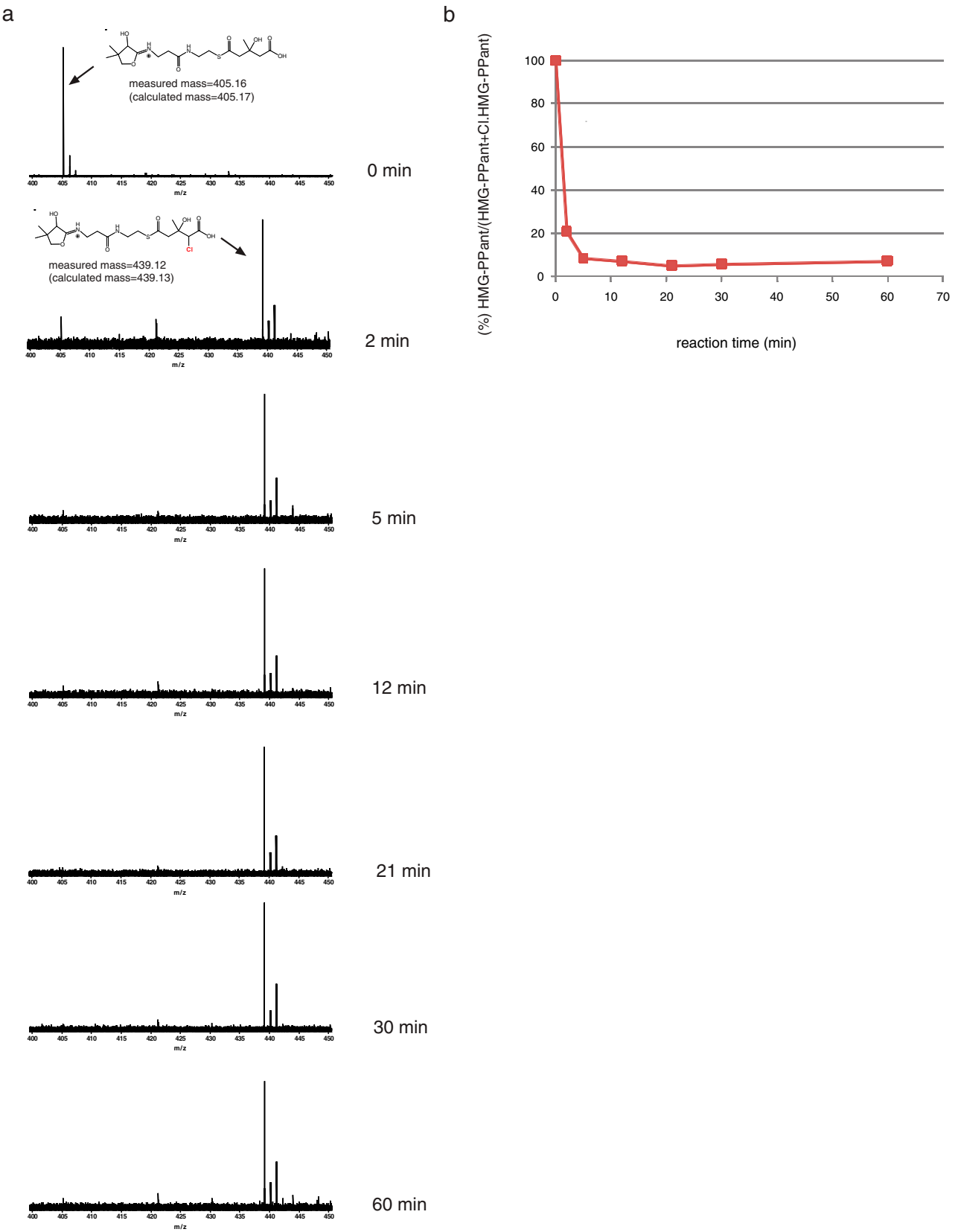

Fig. S3. Time course of chlorinated product formation by wild-type Cur Hal. (*A*) Time course experiment for wild-type Hal. FTICR-MS analysis using IRMPD to release PPant-bound Cl-HMG from Cl-(*S*)-HMG-ACP, resulting in $m/z = 439.12$. The (*S*)-HMG-ACP starting material releases PPant-HMG, resulting in $m/z = 405.16$. (*B*) Time course plot. Peak abundances from PPant ejection products were used to estimate the extent of reaction. The level of released Cl-(*S*)-HMG-PPant as a percent of total released PPant is plotted vs. reaction time.

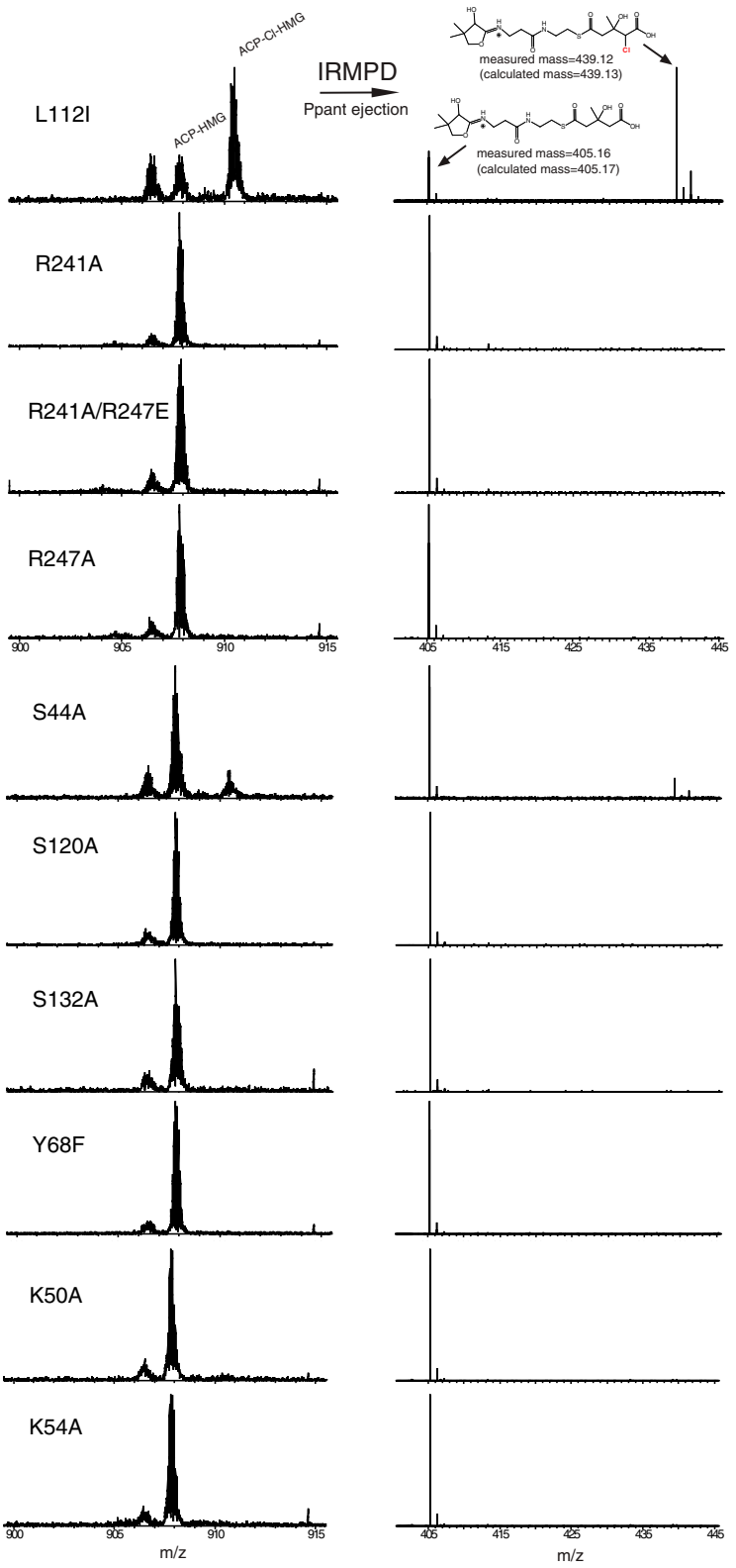

Fig. S4. Catalytic efficiency of Hal wild-type and site-directed mutants. The FTICR traces of ACP-bound species after the 10-min reaction period are shown on the left. Mass traces of PPant ejection products of IRMPD are shown on the right.

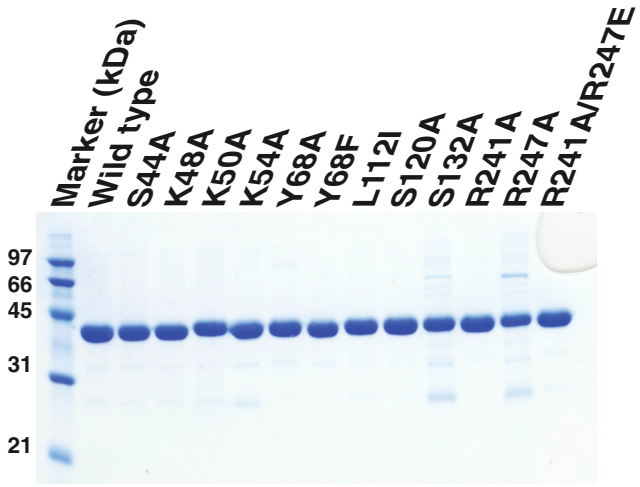

Fig. S5. Purity of wild-type and mutant Hal proteins. For activity assays, Hal proteins were purified by nickel affinity chromatography. Fractions containing Hal were pooled, concentrated, and dialyzed against 10 mM Tris pH 7.8, 50 mM NaCl, 10% glycerol. The figure shows SDS PAGE analysis of purified proteins, including molecular weight markers in the left-most lane.

Table S1. Summary of Crystal Structures

Crystal form	Form I (SeMet)	Form II	Form III	Form IV
Diffraction data				
X ray source	APS 23ID-D	APS 23ID-D	APS 23ID-D	APS 23ID-D
Wavelength (Å)	0.9794	0.9793	1.0332	1.0333
Space group	*I*23	$P4_32_12$	$P2_1$	$P2_12_12_1$
Cell dimensions				
a, *b*, *c* (Å)	$a = b = c = 182.8$	$a = b = 87.6$, $c = 157.7$	50.0, 79.4, 85.2	87.7, 87.9, 157.7
α, β, γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = \beta = \gamma = 90$	$\alpha = 90$, $\beta = 103.0$, $\gamma = 90$	$\alpha = \beta = \gamma = 90$
d_{min} (Å)	2.60 (2.69–2.60)*	2.20 (2.28–2.20)	2.90 (3.00–2.90)	2.70 (2.80–2.70)
R_{merge} (%)	11.1 (73.0)	6.7 (74.8)	6.5 (68.0)	5.6 (61.7)
Avg I/σ_I	31.4 (2.2)	26.3 (1.6)	21.9 (2.5)	34.9 (3.8)
Completeness (%)	99.0 (99.6)	96.8 (86.8)	99.7 (100)	99.4 (100)
Avg redundancy	7.0 (6.1)	7.0 (4.7)	3.8 (3.8)	6.2 (6.2)
Unique reflections (#)	31,330	31,942	14,714	34,399
Refinement				
Data range (Å)	48.8–2.6	33.6–2.2	40.4–2.9	39.3–2.7
Reflections (#)	29,815	30,842	14,644	32,955
R_{work}/R_{free} (%)	18.9/23.8	19.5/23.1	25.8/28.7	18.5/22.9
Number of atoms				
Protein	4,999	2,614	4,930	5,228
Water	152	119	5	108
Ligands	0	30	13	13
RMS deviations				
Bond lengths (Å)	0.009	0.009	0.013	0.008
Bond angles (°)	1.183	1.156	1.440	1.120
B-factors				
Protein	64	63	102	75
Ligand/ion	-	61	81	74
Water	56	62	90	66
Ramachandran plot				
Allowed (%)	99.7	99.7	99.5	99.7
Outliers (%)	0.3	0.3	0.5	0.3
Molecular state				
Asymmetric unit	1 dimer	½ dimer	1 dimer	1 dimer
Chain conformation	A: open B: open	A: closed	A: closed B: open	A: closed B: closed
Active site ligands	none		chain A chain B	
Fe^{2+} site		Fe^{3+}	Fe^{3+} Fe^{3+}	
αKG site		αKG	αKG	formate
Cl^- site		Cl^-	Cl^-	
O_2 site		formate		
Missing residues	A: 40–65; 319–320 B: 40–66; 319–320	A: 319–320	A: 48–59; 319–320 B: 40–65; 319–320	A: 319–320 B: 319–320
PDB code	**3NNJ**	**3NNF**	**3NNL**	**3NNM**

*Values in parentheses are for the highest-resolution shell.

Table S2. Primers for cloning of constructs encoding wild-type and mutant Hal

Construct	Primer sequence (5′–3′)
Wild-type Hal	TACTTCCAATCCAATGCCATGAACCGGGAACAAGTTGAAC
	TTATCCACTTCCAATGCTATTAAGCTCCAGCTGCCATCC
S68A	CTTGGATCAGGCTAAACAGGAAGCCCAGATATCTAAGGAAAAAG
	CTTTTTCCTTAGATATCTGGGCTTCCTGTTTAGCCTGATCCAAG
K72A	TCAGGCTAAACAGGAAAGCCAGATATCTGCGGAAAAAGTAACTTTGAAATTG
	CAATTTCAAAGTTACTTTTTCCGCAGATATCTGGCTTTCCTGTTTAGCCTGA
K74A	AACAGGAAAGCCAGATATCTAAGGAAGCAGTAACTTTGAAATTGGGAGG
	CCTCCCAATTTCAAAGTTACTGCTTCCTTAGATATCTGGCTTTCCTGTT
K78A	CAGATATCTAAGGAAAAAGTAACTTTGGCATTGGGAGGGATTGATGAT
	ATCATCAATCCCTCCCAATGCCAAAGTTACTTTTTCCTTAGATATCTG
Y92A	CGATACCAATGATCATGCCGCTGACTTGGTGAAATACG
	CGTATTTCACCAAGTCAGCGGCATGATCATTGGTATCG
Y92F	CGATACCAATGATCATGCCTTTGACTTGGTGAAATACG
	CGTATTTCACCAAGTCAAAGGCATGATCATTGGTATCG
L136I	GGAACCAAACCATAAAGGCATACCTTGGCATGTAGGAGTCG
	CGACTCCTACATGCCAAGGTATGCCTTTATGGTTTGGTTCC
S144A	TTGGCATGTAGGAGTCGGTGCGTTTTCCTTTACC
	GGTAAAGGAAAACGCACCGACTCCTACATGCCAA
S156A	CCGAAGACTTTGGTGCTGCCATCTGGATCCCC
	GGGGATCCAGATGGCAGCACCAAAGTCTTCGG
R265A	CCCTTAAAACCTGGCTTACATAAGTTAGCTCGTGCTTTTGTGA
	TCACAAAAGCACGAGCTAACTTATGTAAGCCAGGTTTTAAGGG
R271A	CATAAGTTACGTCGTGCTTTTGTGATTGCACTGGTAGATTATG
	CATAATCTACCAGTGCAATCACAAAAGCACGACGTAACTTATG
R271E	TACGTCGTGCTTTTGTGATTGAGCTGGTAGATTATGACACG
	CGTGTCATAATCTACCAGCTCAATCACAAAAGCACGACGTA

For each construct, forward primers are on the top line and reverse primers on the bottom line.