Supplementary Information

Structural basis for halogenation by iron- and 2-oxo-glutarate-dependent enzyme WelO5

Andrew J. Mitchell,[1] Qin Zhu,[3] Ailiena O. Maggiolo,[1] Nikhil Ananth,[1] Matthew L. Hillwig,[3]

Xinyu Liu[3*] and Amie K. Boal[1, 2*]

[1]Departments of Biochemistry and Molecular Biology and of [2]Chemistry, The Pennsylvania State University, University Park, PA 16802, [3]Department of Chemistry, University of Pittsburgh, 219 Parkman Avenue, Pittsburgh, PA 15260

SUPPLEMENTARY RESULTS

Supplementary Table 1 Data collection and refinement statistics for the x-ray structure of WelO5[1] in the absence of substrate.

	WelO5 (Fe peak)	**WelO5** (native)
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	70.93, 85.39, 141.06	70.76, 86.35, 141.31
Resolution (Å)	50.00-2.65 (2.70-2.65)*	50.00-2.00 (2.03-2.00)
R_{merge}	0.19 (0.99)	0.12 (0.61)
$I/\sigma I$	20.3 (3.85)	20.8 (3.2)
Completeness (%)	98.6 (96.8)	99.3 (96.1)
Redundancy	14.3 (13.9)	13.3 (12.4)
Refinement		
Resolution (Å)		50.00-2.00
No. reflections		55713
R_{work}/ R_{free}		0.19/0.21
No. atoms		
Protein		6661
Ligand/ion		36
Water		456
B-factors		
Protein		15.25
Ligand/ion		18.06
Water		29.18
r.m.s deviations		
Bond lengths (Å)		0.004
Bond angles (°)		0.953

[1]The above datasets correspond to two different crystals that were used to solve the structure.
*Highest resolution shell shown in parenthesis.

Supplementary Table 2 Data collection and refinement statistics for the x-ray structures of wt and G166D WelO5[1] solved in the presence of substrate **1** (+/- NO).

	WelO5 + 1	**G166D WelO5 + 1**	**WelO5 + 1 + NO**
Data collection			
Space group	$P2_12_12_1$	$C2$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c* (Å)	71.05, 83.07, 141.94	74.10, 54.13, 74.64	71.30, 85.18, 143.89
Resolution (Å)	50.00-2.40 (2.44-2.40)	50.00-2.50 (2.54-2.50)	50.00-2.40 (2.44-2.40)
R_{merge}	0.13 (0.72)	0.06 (0.29)	0.12 (0.97)
$I/\sigma I$	16.7 (1.9)	21.18 (3.06)	17.3 (2.06)
Completeness (%)	99.9 (99.2)	99.0 (92.7)	100 (100)
Redundancy	7.1 (6.4)	3.6 (3.2)	7.4 (7.4)
Refinement			
Resolution (Å)	50.00-2.40	50.00-2.50	50.00-2.37
No. reflections	32172	7919	33111
R_{work}/ R_{free}	0.19/0.23	0.22/0.26	0.20/0.24
No. atoms			
Protein	6641	2116	6523
Ligand/ion	82	34	109
Water	251	35	157
B-factors			
Protein	33.52	51.40	40.72
Ligand/ion	46.72	51.71	47.79
Water	43.17	43.82	37.90
r.m.s deviations			
Bond lengths (Å)	0.005	0.0054	0.005
Bond angles (°)	0.970	1.039	1.003
Molprobity clashscore	1.44	2.38	1.15
Rotamer outliers	0	1	4
Ramachandran			
Preferred	801	250	798
Allowed	21	9	16
Outliers	3	2	1

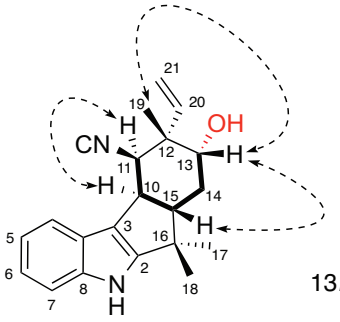
[1]A single crystal was used to solve each structure.
*Highest resolution shell is shown in parenthesis.

Supplementary Table 3 Data collection statistics for Br-substituted wt WelO5 (soaked in substrate **1**) before and after O_2 exposure.

	Br-WelO5 + 1	**Br-WelO5 + 1 + O_2**
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	70.83, 85.77, 142.44	71.04, 86.62, 142.69
Resolution (Å)	50.00-2.35 (2.39-2.35)	50.00-2.60 (2.64-2.60)
R_{merge}	0.12 (0.94)	0.18 (---)
$I/\sigma I$	15.81 (1.66)	12.55 (1.25)
Completeness (%)	99.7 (98.0)	93.5 (76.9)
Redundancy	5.9 (4.1)	10.3 (6.9)

[1]Each dataset corresponds to a single crystal.
*Highest resolution shell is shown in parenthesis.

Supplementary Table 4 Tabulated 1-D and 2-D NMR spectral data of **3**



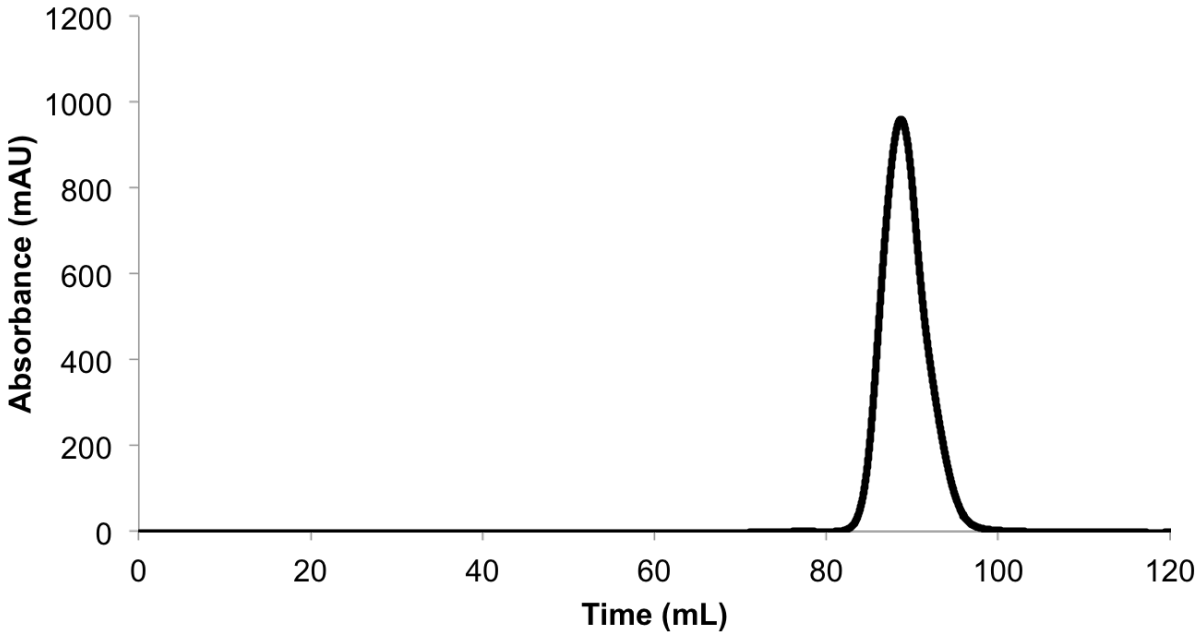
COSY —

NOESY ◄- - -►

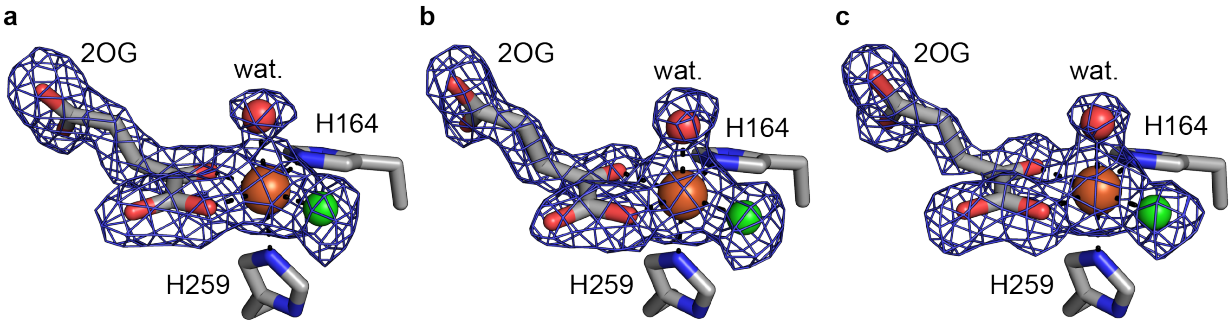
13*R*-hydroxyl-12-*epi*-fisherindole U (**3**)

H-position	observed (700 MHz, $CDCl_3$)					
	δ (ppm)	splitting	*J* (Hz)	^{1}H-^{1}H COSY	NOESY	# of H
1	7.92	br s		H-7	H-17, H-18	1H
4	7.43	d	7.7		H-11	1H*
5	7.08	ddd	7.7, 7.7, 1.1			1H
6	7.12	ddd	7.7, 7.7, 1.1			1H
7	7.34	d	7.7			1H*
10	3.29	dtd	10.9, 3.9, 2.3	H-15	H-20, H-11, H-15, H-14ax, H-17	1H
11	4.43	s			H-4, H-11, H-19	1H
13eq						
13ax	3.88	dd	11.0, 4.7	H-14ax, H-14eq	H-19, H-15, H-14ax, H-14eq	1H
14eq	1.95	ddd	11.9, 3.8, 2.3	H-13	H-13ax, H-15, H-14eq	1H
14ax	1.70	dd	12.2, 11.7	H-13	H-20, H-13ax, H-10, H-15, H-14eq	1H
15	2.48	ddd	13.1, 10.9, 2.3	H-10, H-14	H-10, H-13ax,H-14ax, H-14eq H-18	1H
17	1.08	s			H-1, H-10, H-14eq	3H
18	1.42	s			H-1, H-15, H-14eq	3H
19	1.50	s			H-11, H-13ax,	3H
20	6.25	dd	17.5. 11.0	H-21	H-10, H-14ax, H-19	1H
21*E*	5.41	d	11.0	H-20		1H
21*Z*	5.40	d	17.6	H-20		1H

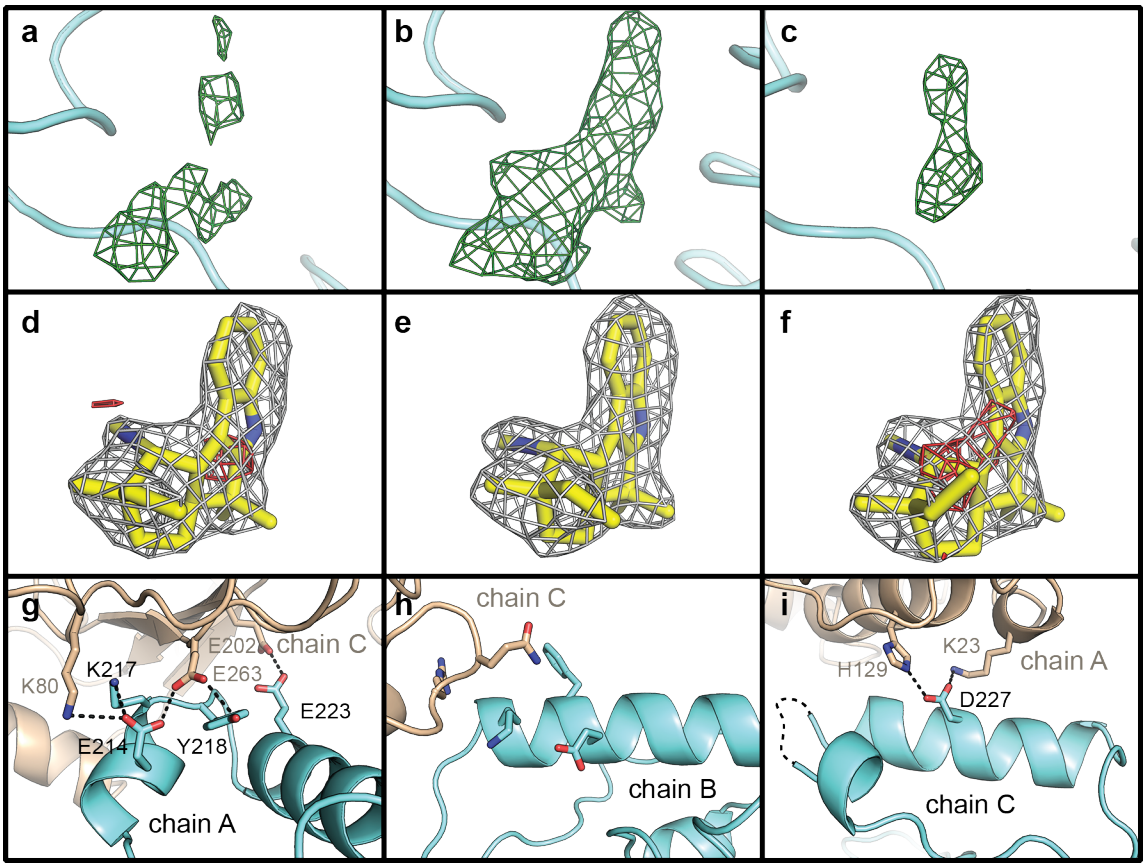
* Further splitting was not distinguishable in the spectrometer used for this measurement.



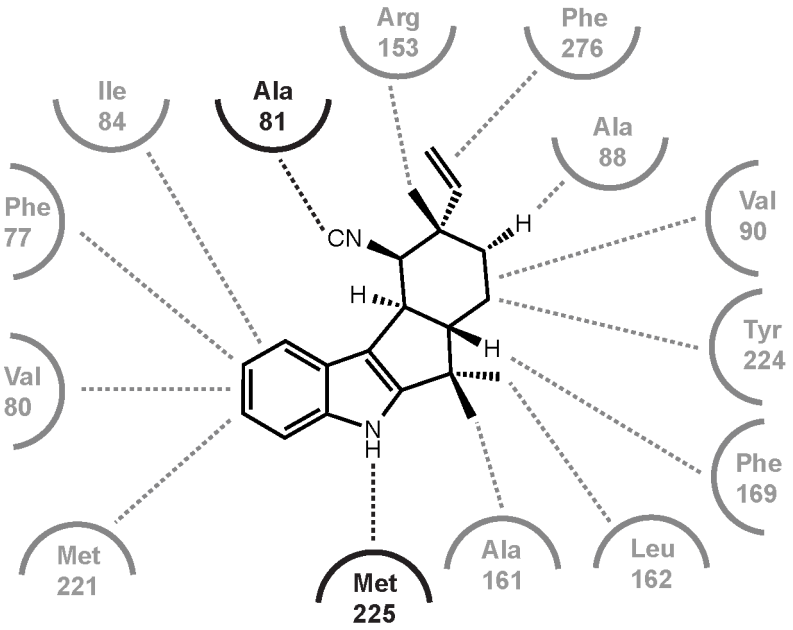
Supplementary Figure 1. Size exclusion chromatography of WelO5 indicates the protein is a monomer in solution (20 mM HEPES pH 7.5, 200 mM NaCl). WelO5 elutes at 88.65 mL on a GE HiLoad 16/600 Superdex 200 pg column (20 mM HEPES pH 7.5, 200 mM NaCl), corresponding to an apparent molecular weight of 30.4 kDa. The value is consistent with the calculated MW (30.2 kDa) of a WelO5 monomer.



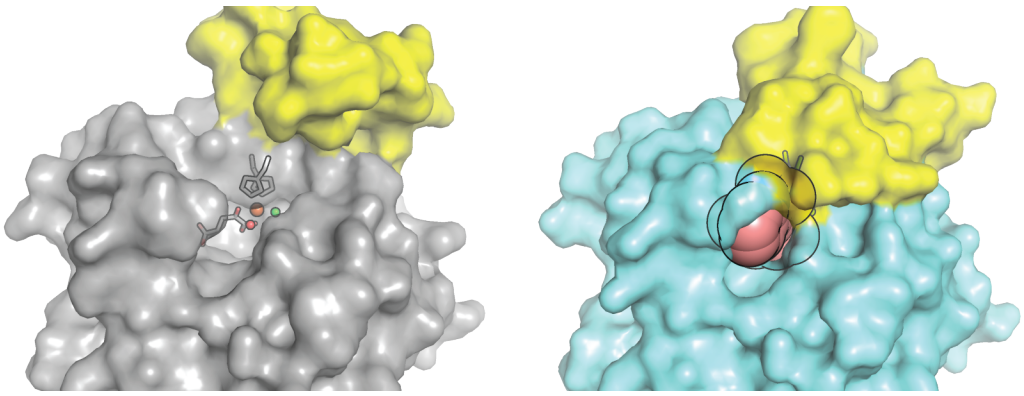
Supplementary Figure 2. An F_o-F_c omit map (blue mesh, contoured at 3.0 σ) for the Fe^{II} ion and exogenous ligands in the three equivalent views of the active site in the substrate-free (apo) WelO5 structure. Chains A, B, and C are shown in the corresponding panels above. The Fe^{II} ion is shown in orange, chloride in green, and water in red. The views reveal identical coordination environments for the Fe^{II} cofactor in all three monomers.



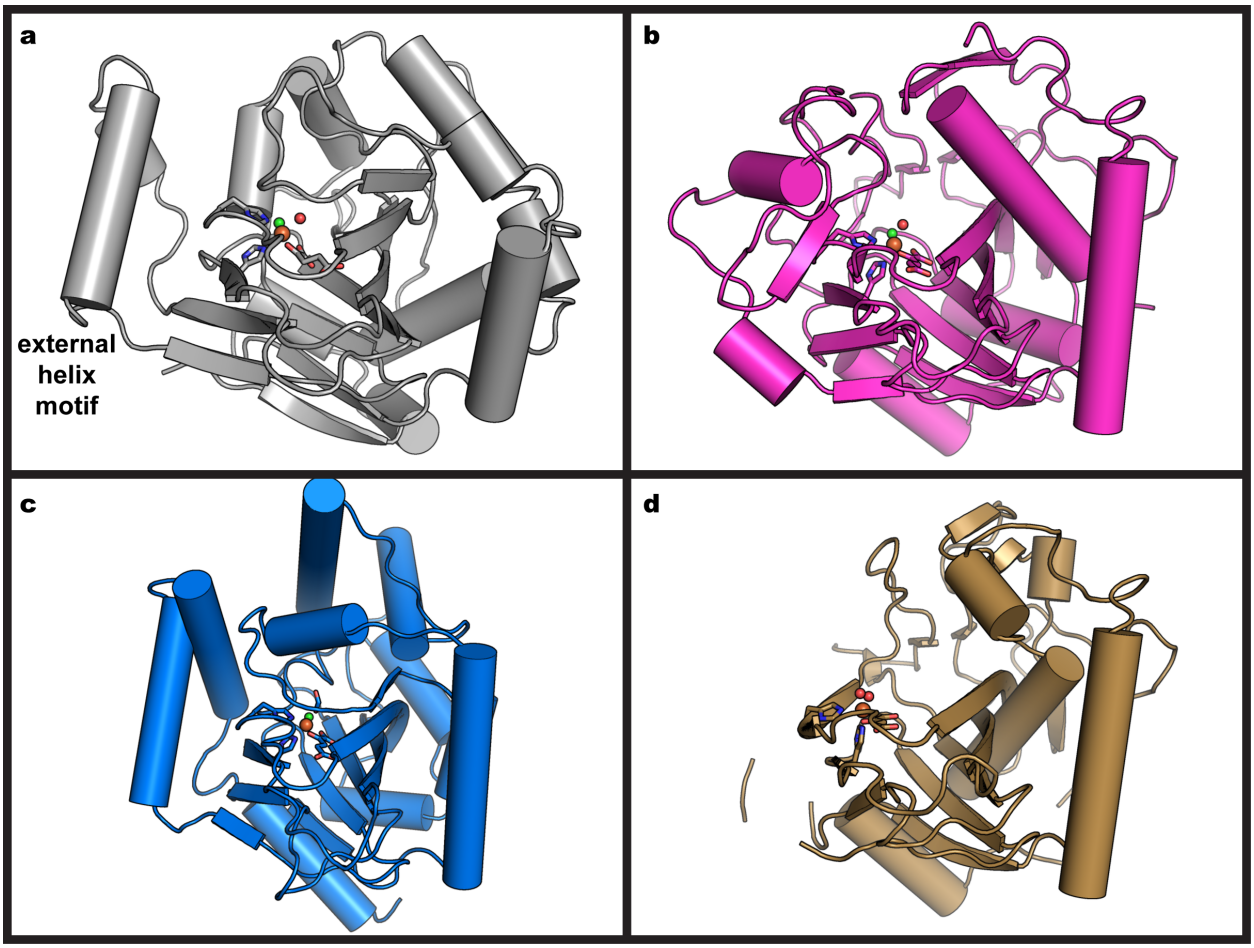
Supplementary Figure 3 (**a-c**) F_o-F_c omit map (green mesh, contoured at 3.0σ) for WelO5 substrate **1**, generated after ligand deletion and refinement, in chains A-C. (**d-f**) The resulting 2F_o-F_c electron density map (gray mesh, contoured at 1.0σ) after modeling **1** into each chain. Negative difference electron density (F_o-F_c map, red mesh, contoured at -3.0σ) in chains A and C is likely a consequence of diminished substrate occupancy and we omitted the molecule from the model in chain C. (**g-i**) Lattice contacts between the external helix and adjacent molecules in the ASU in chains A (**g**, blue cartoon representation) and C (**i**, blue cartoon representation) indicate a correlation between fewer contacts in this region, as in chain B (**h**, blue cartoon representation), and increased occupancy of substrate when incorporated via soak methods.



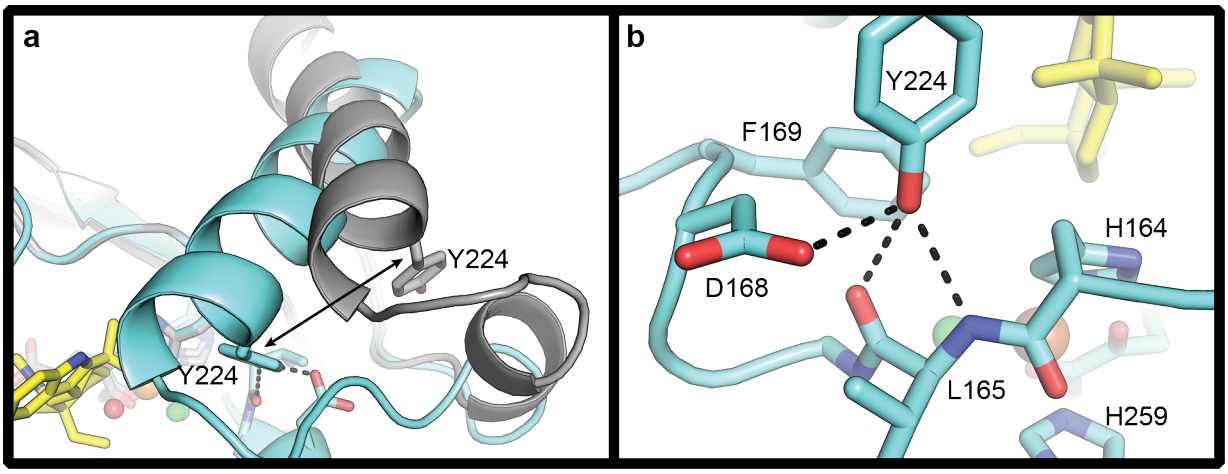
Supplementary Figure 4. The binding pocket for **1**, shown in a 2D cartoon representation. Hydrogen-bonding interactions are depicted in bold and van der Waals contacts are shown in gray (restricted to those within 4.0 Å). The substrate-binding site is predominantly composed of nonpolar amino acid side chains, consistent with the hydrophobic nature of **1**.



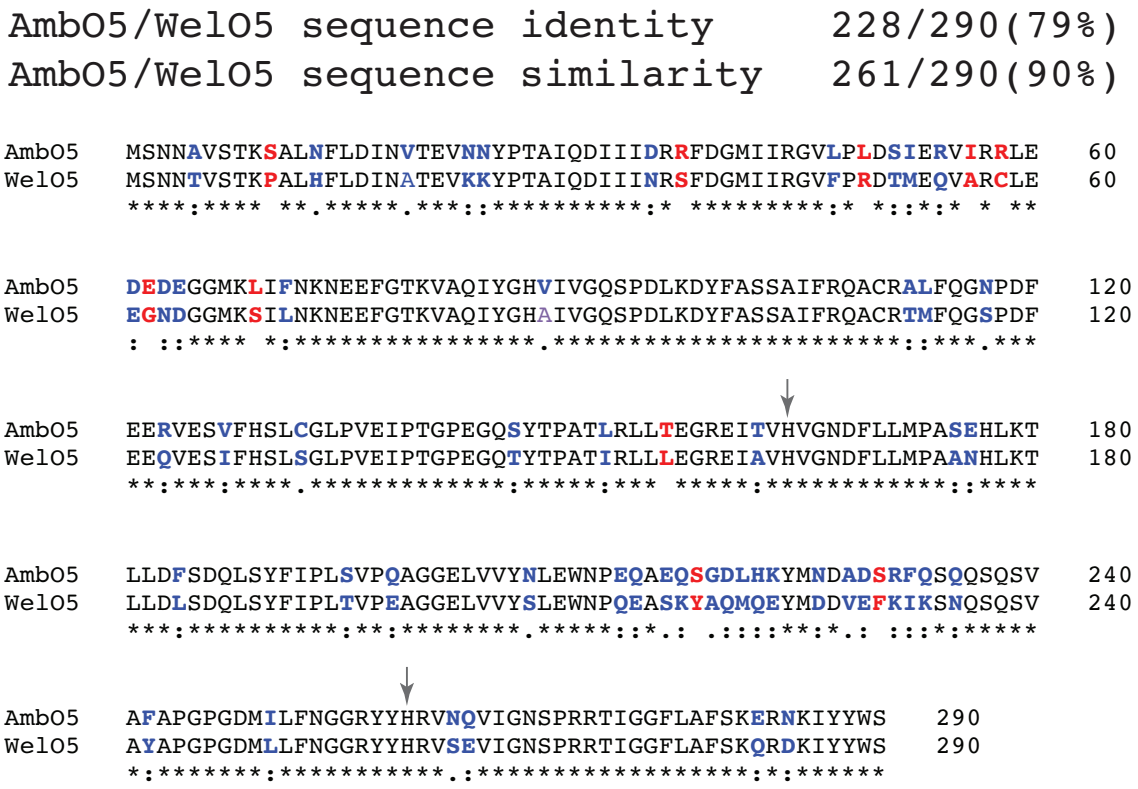
Supplementary Figure 5. A surface representation of the WelO5 active site (left) illustrating solvent exposure of the active site and the change in accessibility upon interaction with **1** (right). Closure of the active site upon substrate binding arises from the intrinsic bulk of **1** and a conformational change in the external helix motif (yellow).



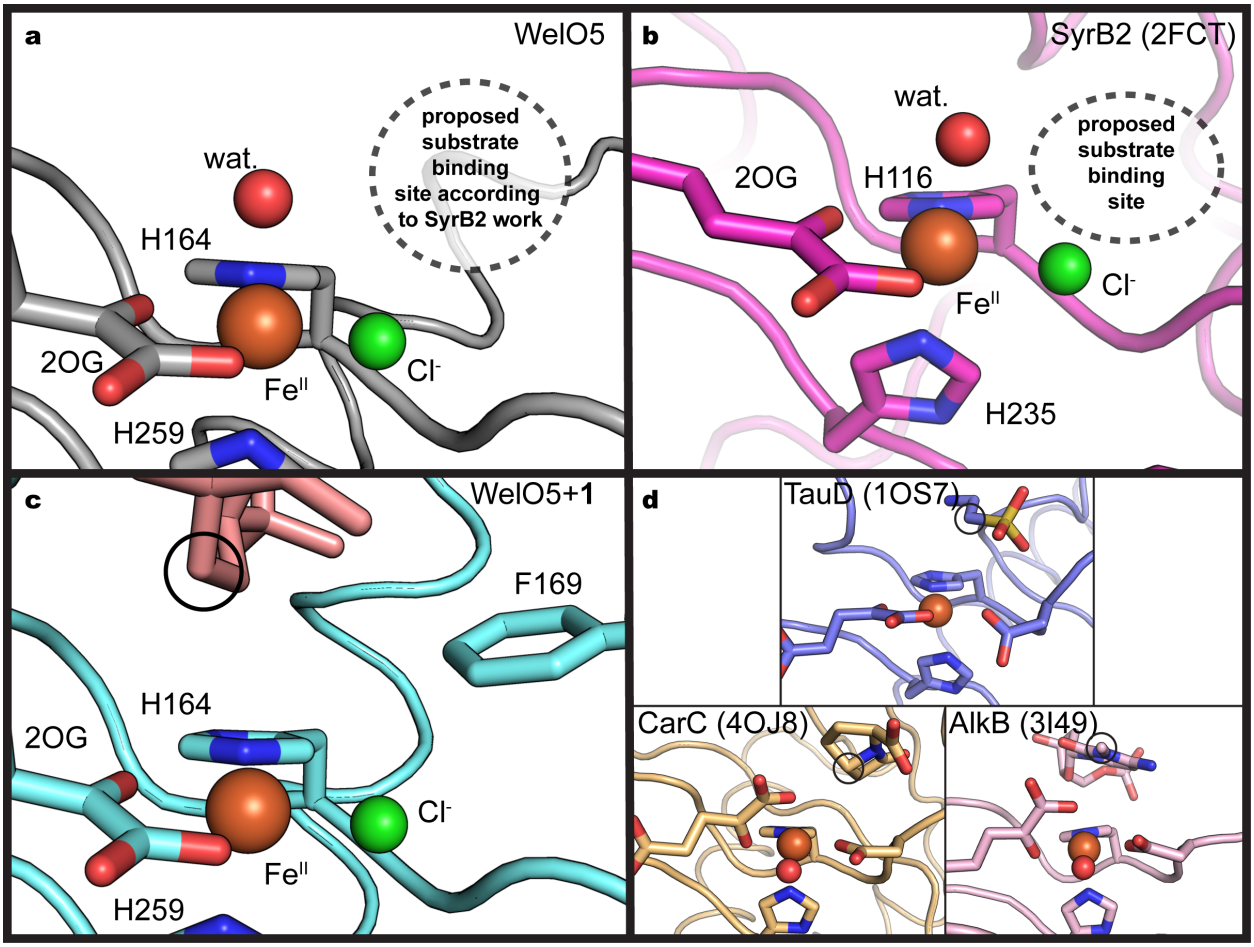
Supplementary Figure 6. Comparative views of the overall fold in WelO5, SyrB2 (PDB accession code 2FCT), CurA (PDB accession code 3NNF), and CytC3 (PDB accession code 3GJB).



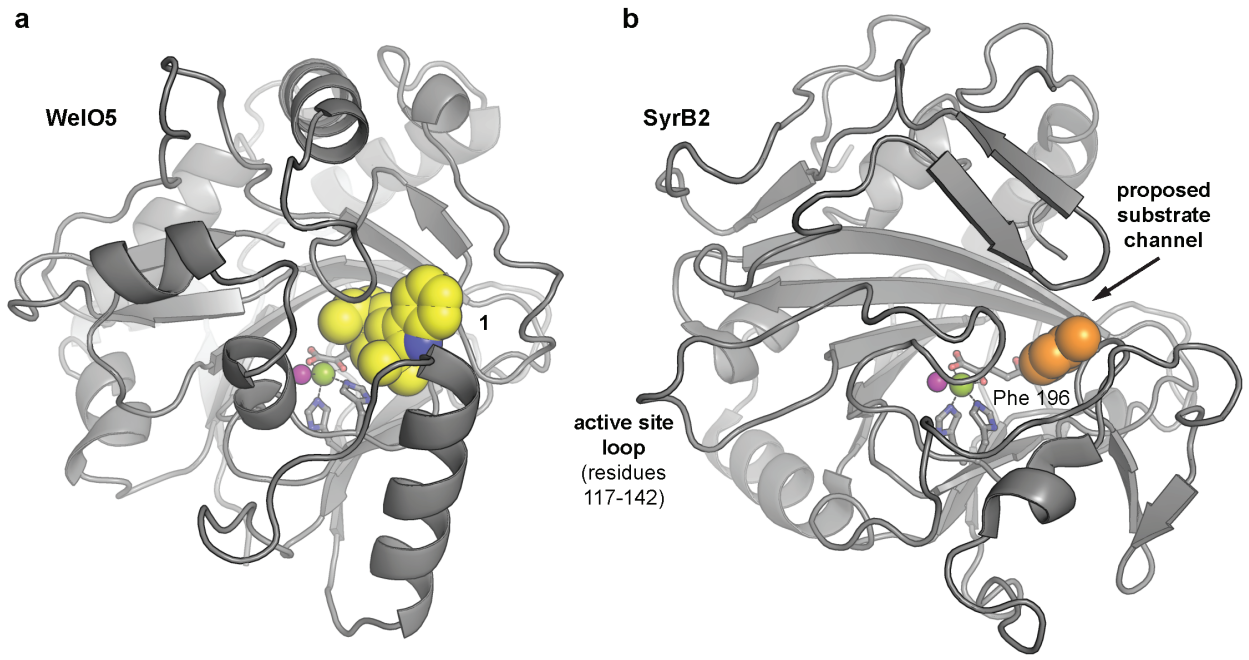
Supplementary Figure 7. (**a**) A zoomed-in view of the shift in the external helix structure upon substrate binding. (**b**) The interactions between conserved residue Tyr224 and the region immediately following Fe^{II} ligand His164 that promote closure of the external helix upon substrate binding and conformational change in residue Phe169.



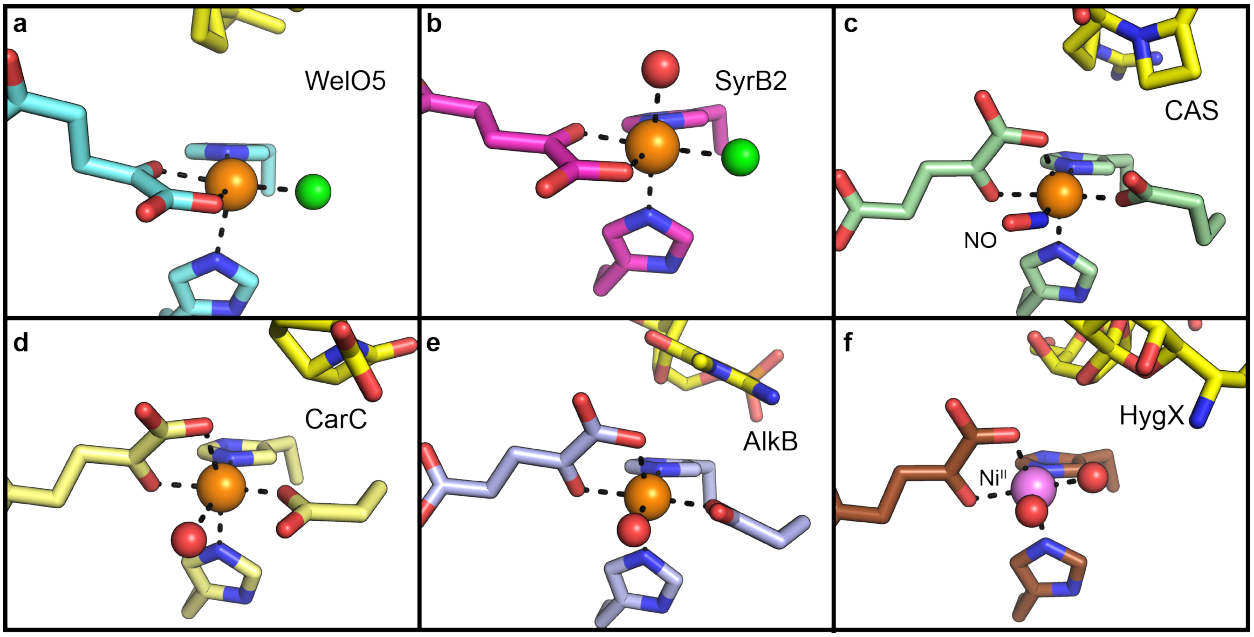
Supplementary Figure 8. Sequence alignment of WelO5 and AmbO5. Identical amino acid residues are shown in black, similar sites are shown in blue, and residues that are distinct are highlighted in red. The largest number of substitutions are observed in the external helix motif spanning residues 220-235.



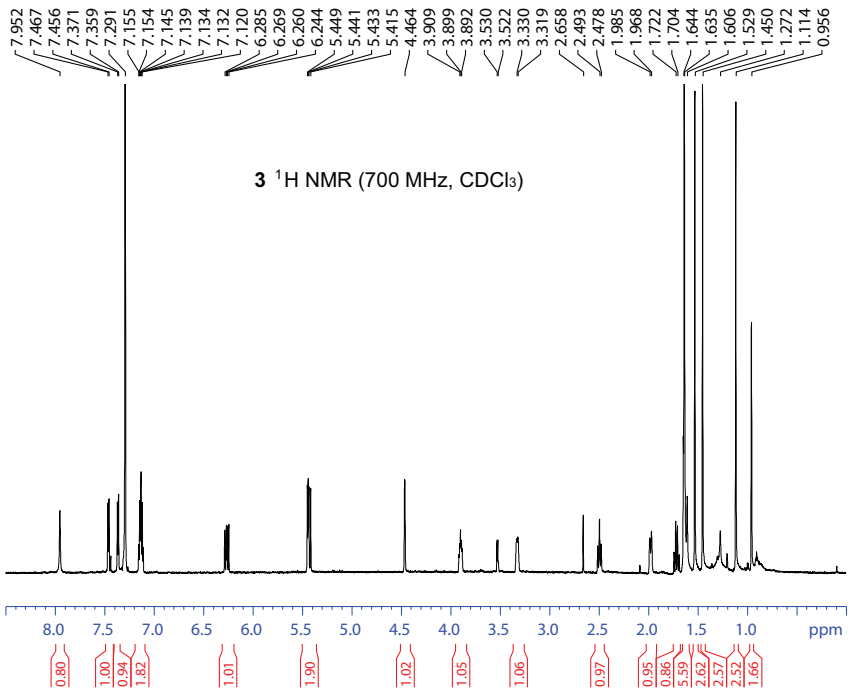
Supplementary Figure 9. WelO5 uses a distinct substrate-binding site compared to other structurally characterized Fe/2OG enzymes. The high degree of solvent exposure in the active site in WelO5 in the absence of substrate (**a**) theoretically allows for the possibility of substrate interaction close to the –Cl ligand, as anticipated in carrier protein dependent halogenases (**b**). Instead, the substrate is localized to the opposite quadrant near His164 (**c**). Interestingly, the originally anticipated location for the hydrogen atom abstraction target in SyrB2 is similar to that of other Fe/2OG enzymes (**d**) that catalyze hydroxylation/stereoinversion reactions.



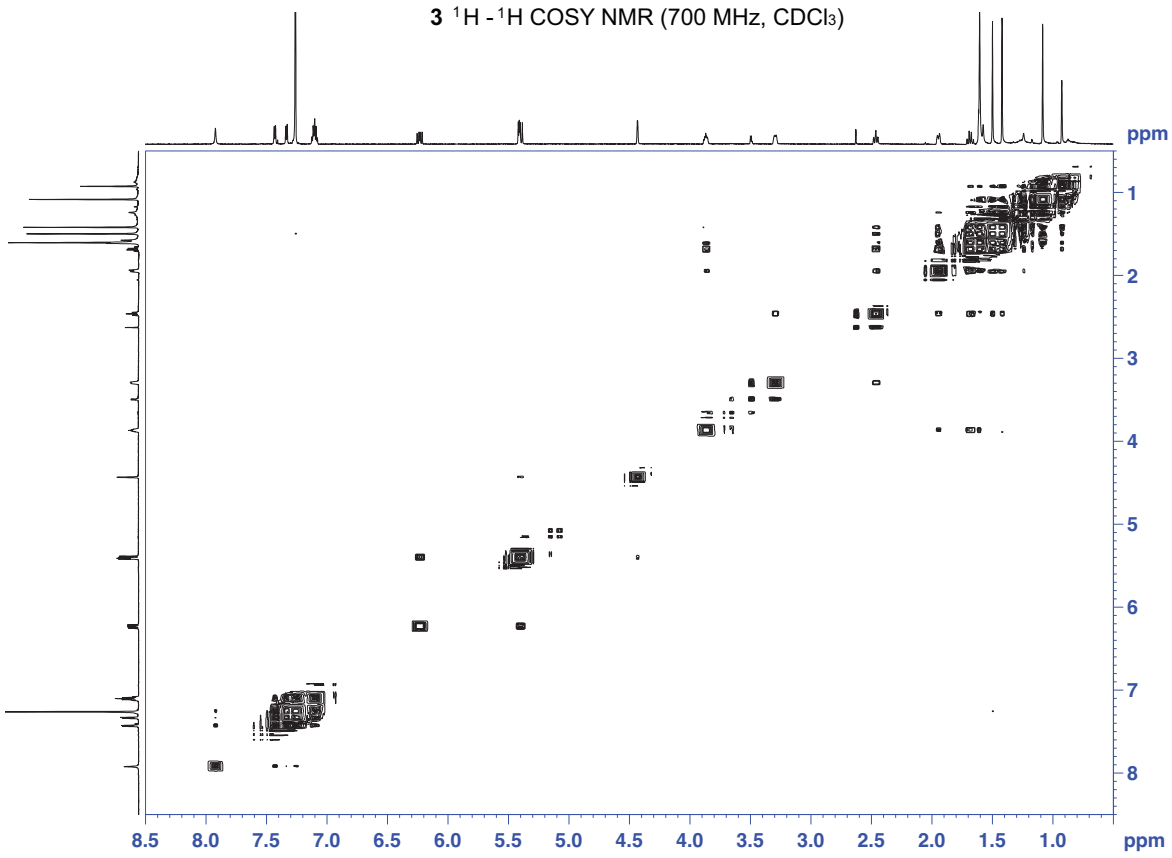
Supplementary Figure 10. A comparison of the substrate binding site in **1**-bound WelO5 (**a**) to the proposed Thr-phosphopantetheine channel in apo SyrB2 (PDB accession code 2FCT) (**b**), as demarcated by Phe 196. The putative channel in SyrB2 would represent the same approach to the active site as that observed in the experimentally determined substrate-bound structure WelO5. Both sites terminate near His 164 (WelO5 numbering) in a location distinct from the other proposed binding site near the halogen ligand, occluded by an active site loop (residues 117-142) in SyrB2. The substrate-binding site is challenging to predict in the carrier-protein dependent halogenases because significant conformational change in peripheral loop structures that flank the active site in SyrB2 could occur upon complex formation with SyrB1.



Supplementary Figure 11. If the oxo moiety of the Fe^{IV}=O intermediate in WelO5 occupies the open coordination site in the structure of the reactant state (**a**), a hydroxylation outcome would seem most likely to occur. The open site in WelO5 is equivalent to the site occupied by the water molecule in apo SyrB2 (PDB accession code 2FCT) (**b**). To achieve selective halogenation in SyrB2 and WelO5, we propose the Fe^{IV}=O must occupy the position held by the C1 carboxylate of 2OG in the reactant complex structures. 2OG conformational change has been observed upon binding of NO (an O_2 surrogate) in clavaminate synthase (PDB accession code 1GVG) (**c**) and in the resting Fe^{II} state in CarC (**d**) (PDB accession code 4OJ8), AlkB (**e**) (PDB accession code 3I49), and HygX (**f**) (PDB accession code 4XCB). Alternative positions for the oxo moiety may be a universal characteristic of other non-canonical Fe/2OG oxygenases.

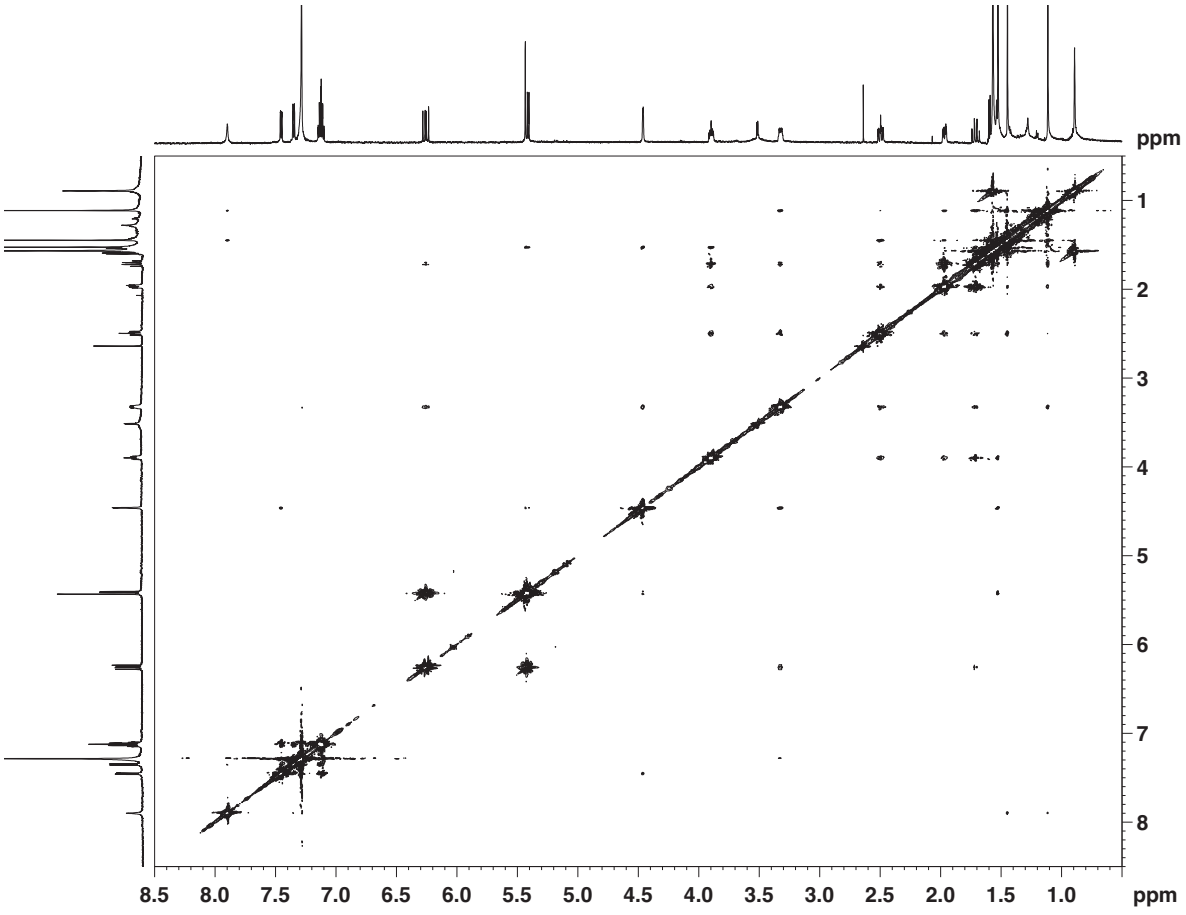


Supplementary Figure 12. ^{1}H NMR spectrum of **3.**

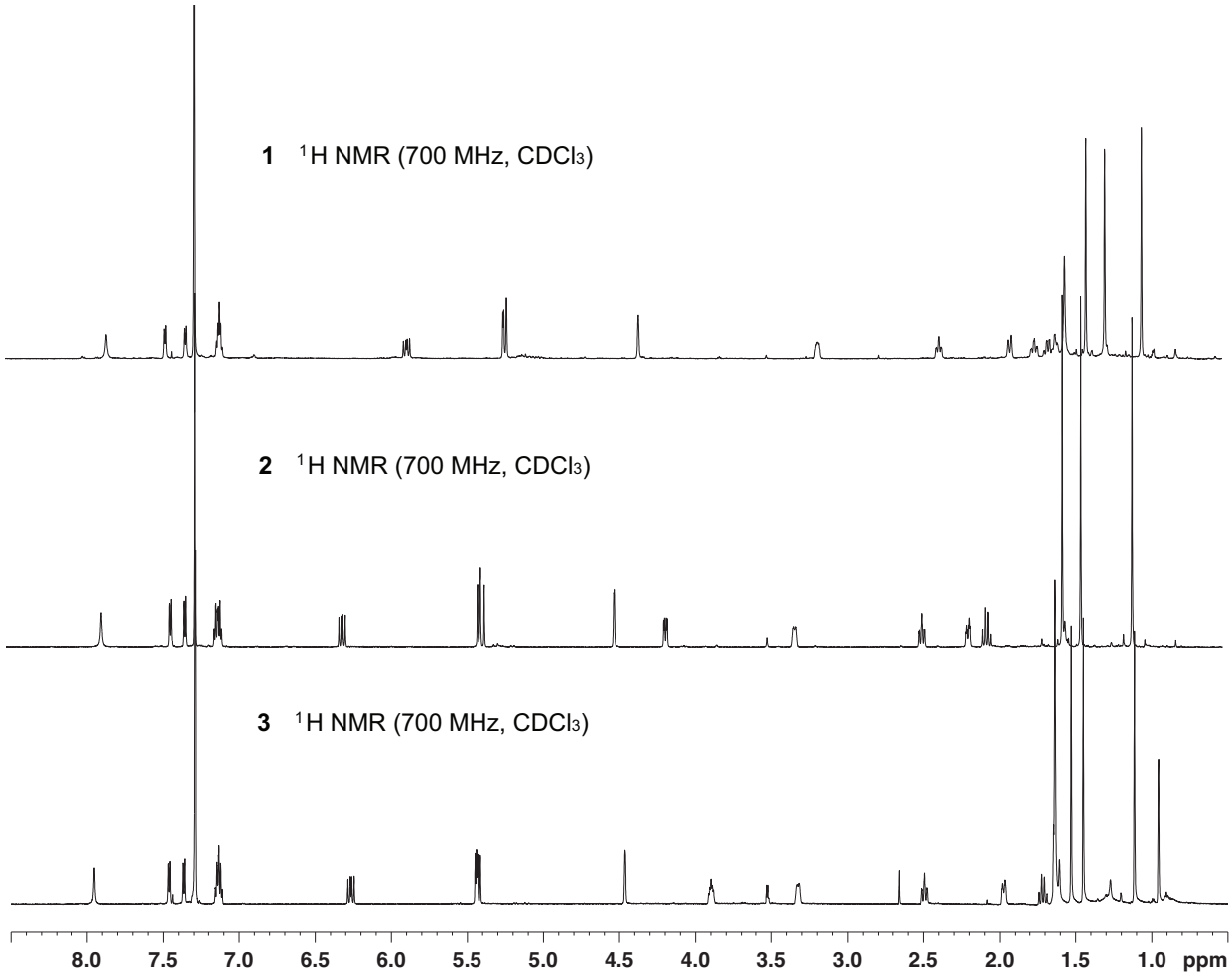


Supplementary Figure 13. ^{1}H–^{1}H COSY spectrum of **3**.

3 ^{1}H -^{1}H NOESY NMR (600 MHz, $CDCl_3$)



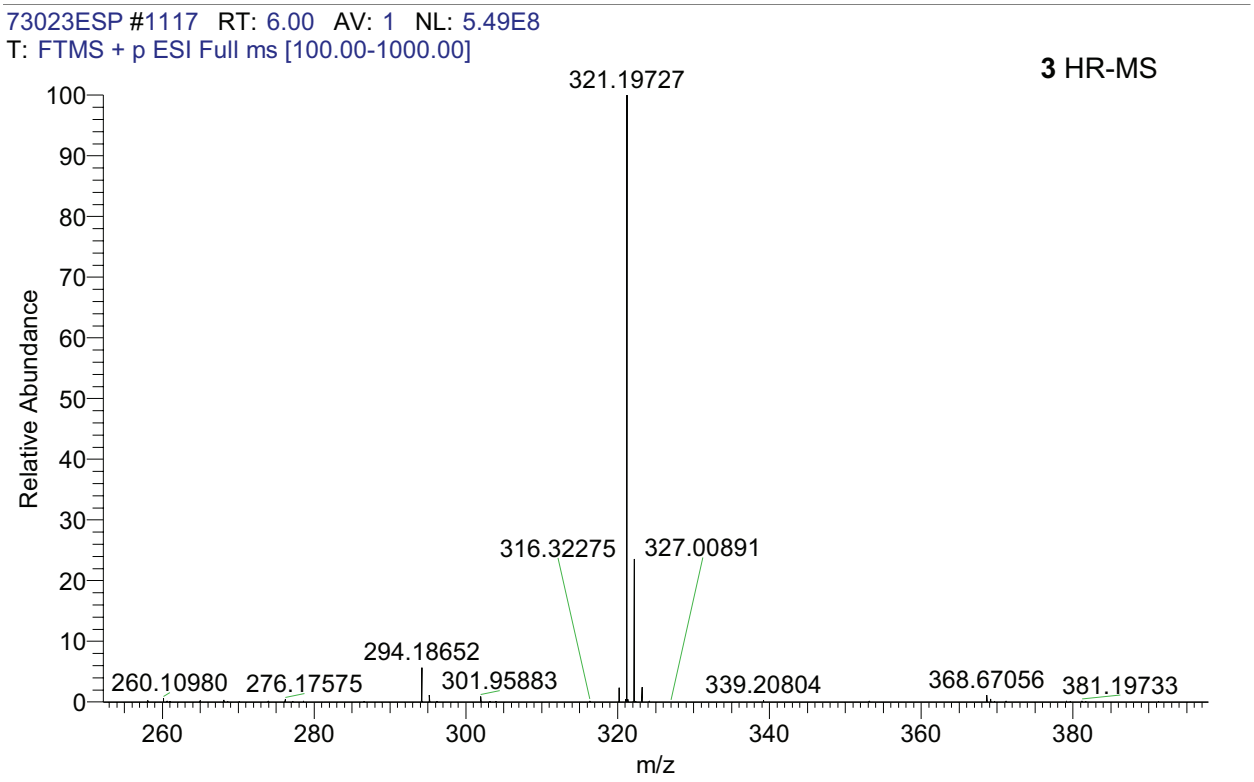
Supplementary Figure 14. ^{1}H–^{1}H NOESY spectrum of **3**.



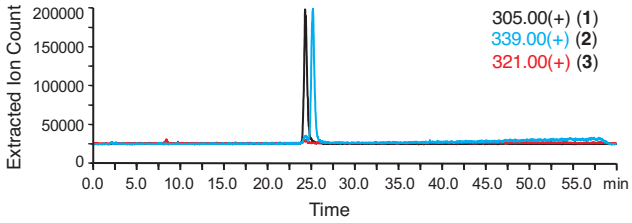
Supplementary Figure 15. ^{1}H NMR spectral overlay of **1**, **2**, **3**.

73023ESP#1117 RT: 6.00
T: FTMS + p ESI Full ms [100.00-1000.00]
m/z= 244.95192-373.31731

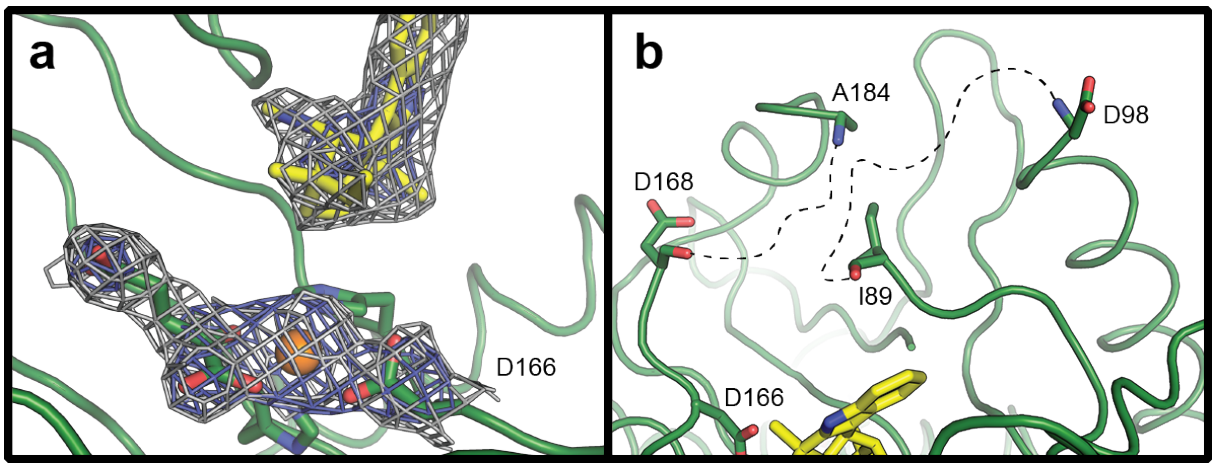
m/z	Intensity	Relative	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
321.19727	554111104.0	100.00	321.19669	1.80	10.5	$C_{21}H_{25}ON_2$



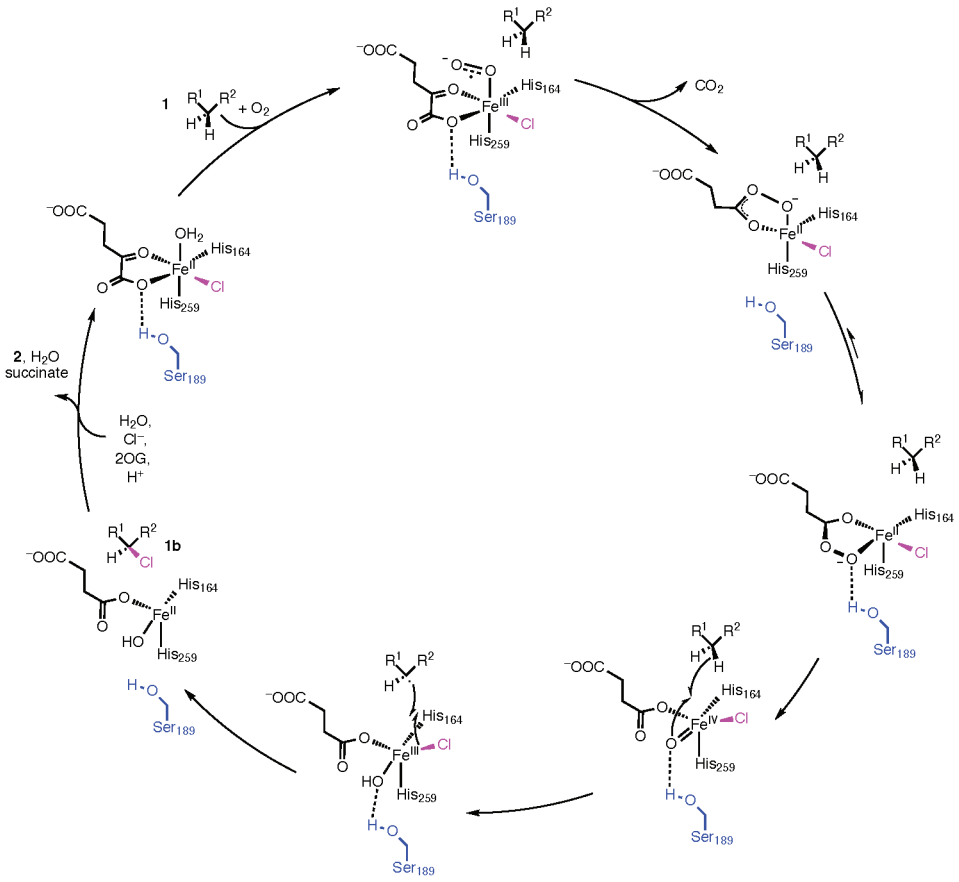
Supplementary Figure 16. High-resolution mass spectrum of **3**.



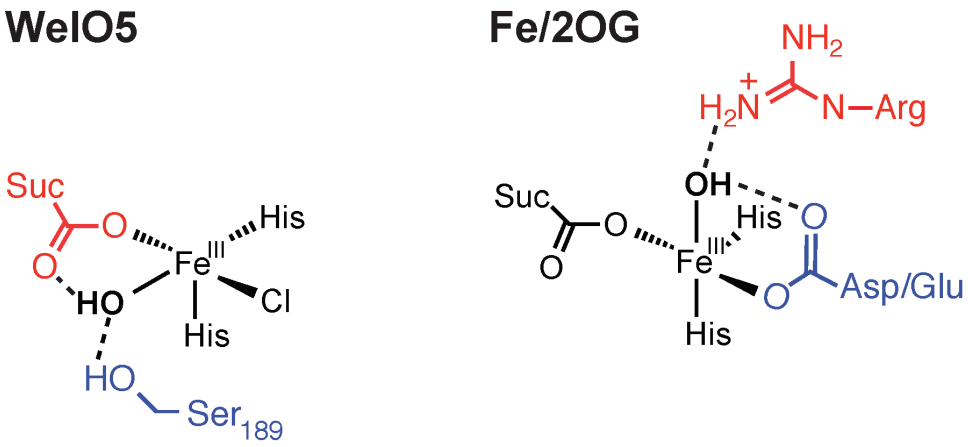
Supplementary Figure 17. LC-MS analysis of the *wt*-WelO5 enzymatic assays with **1** as the substrate indicates the formation of trace amounts of hydroxylated product **3** in addition to the major chlorinated product **2**.



Supplementary Figure 18. The x-ray structure of G166D WelO5 (**a**) illustrates that the variant uses the same substrate binding site as wt WelO5. A F_o-F_c omit map contoured at 3.0σ is shown in blue mesh and a $2F_o$-F_c map contoured at 1.0σ is shown in grey mesh. Substitution of Gly166 for an Asp residue disorders two key regions of the protein (**b**) which results in failure to close the external helix in the crystals after soak of **1** into the substrate binding site.



Supplementary Figure 19. Proposed mechanism for selective halogenation in Fe/2OG halogenases.



Supplementary Figure 20. Outer-sphere interactions in WelO5 (left) distinct from those in prototypical Fe/2OG hydroxylases (right) that would constrain the position of the oxo/hydroxo unit in WelO5 to the equatorial plane defined by succinate, Cl, and the proximal His ligand (top right in left panel).