Revealing Substrate Positioning Dynamics in Non-heme Fe(II)/αKG-dependent Halogenases Through Spectroscopically Guided Simulation

Rimsha Mehmood[1,2], Vyshnavi Vennelakanti[1,2] and Heather J. Kulik[1,*]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139*

[2]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139*

ABTRACT: Non-heme iron halogenases, such as SyrB2, WelO5, and BesD, halogenate unactivated carbon atoms of diverse substrates at ambient conditions with exquisite selectivity seldom matched by non-biological catalysts. Although crystallography, spectroscopy, and kinetic measurements provide foundational knowledge of enzyme structure and function, critical gaps remain in our understanding of how the protein environment dynamically reorganizes to a catalytically active state capable of halogenation. Using experimentally-guided molecular dynamics (MD) simulations augmented with multi-scale (i.e., QM/MM) simulations of substrate-bound complexes of BesD and WelO5, we investigate substrate/active-site dynamics that enable selective halogenation. Our simulations reveal that active-site configurational isomerization is necessary in WelO5 to attain substrate/active-site geometry consistent with its observed chemo- and regioselectivity. Conversely, a slight reorientation of the substrate from its crystal-structure position is sufficient to enable regioselective chlorination in BesD without the need to invoke active-site isomerization. We observe how substrate-protein interactions evolve during experimentally-motivated MD of halogenases. We relate the nature of these interactions to the distinct substrates. For BesD, we resolve the uncertainty around the mechanistic relevance of Asn219 based on a prior mutagenesis study. By quantifying the presence of thermodynamically competitive C-H bonds on substrates of SyrB2, BesD, and WelO5, we confirm the need for the protein environment to strategically position the substrate to impart regioselectivity to halogenases. Our simulations reveal that the optimum substrate/active-site geometry also outweighs interactions between active site ligands and the protein environment in facilitating the required chemoselectivity in halogenases.

Keywords: halogenation, C-H activation, QM/MM, enzyme catalysis, non-heme iron

1. Introduction.

Selective halogenation of aliphatic C-H bonds is essential for the synthesis of bioactive molecules[1-3]. Nevertheless, halogen functionalization via synthetic routes requires harsh chemical conditions[4] and is often marked by poor selectivity[5]. In contrast, enzymes can readily halogenate natural products in biosynthetic pathways, as exemplified by non-heme iron and alpha-ketoglutarate (αKG) dependent halogenases[6-13]. These enzymes functionalize unactivated carbon atoms of diverse substrates (e.g. polar[14] vs hydrophobic[11]) with halogens at ambient temperature and pressure, with exquisite chemo-, regio- and stereoselectivity[6-14]. Non-heme iron halogenases are closely related to their well-studied, non-heme iron hydroxylase[8, 15] counterparts that catalyze challenging hydroxylation reactions on unactivated carbon centers by invoking radical chemistry via an iron-oxo reactive intermediate[16-19]. The similar structure but distinct reactivity motivates a mechanistic understanding of the unique roles of metal cofactors and protein scaffolds in imparting superior halogenation selectivity to non-heme iron halogenases. Nevertheless, the fleeting nature of the highly reactive iron-oxo intermediate[20-21] has challenged efforts to crystallize halogenases in catalytically active states and elucidate the mechanism behind selective halogenation over hydroxylation. Additionally, most non-heme iron halogenases require carrier proteins for substrate delivery[22-23], which has hindered structural characterization of their enzyme-substrate complexes.

The limitations in crystallographic characterization have sparked vigorous computational studies[24-29] alongside spectroscopic[30-34] and kinetic[31, 35] studies to understand how non-heme iron halogenases engage their substrates for preferential halogenation in the presence of an iron-oxo moiety that is also capable of substrate hydroxylation. Precise substrate/active-site positioning has been proposed[25, 34-35] as a critical mechanistic strategy employed by non-heme iron

halogenases to suppress hydroxylation. Kinetic[35] and spectroscopic[34] studies have suggested that proximity of substrate to Cl instead of the oxo moiety in the halogenase active site compromises hydrogen atom abstraction (HAA) efficiency in order to ensure chlorination reactivity. Some prior studies[28-29, 31, 33-34] have shown that the proximity to the Cl vs oxo moiety can also be adjusted via configurational isomerization of the active site. Reaction selectivity in halogenases has been correlated to the angle between the HAA target C-H bond on the substrate and the iron-oxo moiety[33-34], with a more obtuse angle activating the π-pathway to favor halogenation while an acute angle favors the σ-pathway for hydroxylation.[36-38] Although these studies have provided foundational knowledge of substrate positioning in controlling reaction outcome in halogenases, the role of protein dynamics and greater enzyme environment remains unclear. Computational modeling of the dynamics of substrate-enzyme complexes can reveal the possible role of protein-substrate interactions in governing selective halogenation via strategic positioning. Additionally, simulations with otherwise experimentally elusive iron-oxo intermediates provide insight on catalytically operative active-site configurational isomers that are most aligned with experimentally-motivated (i.e., as guided by spectroscopic evidence) substrate positioning for halogenases[34-35, 39].

We recently used spectroscopically-derived[34] distance and angle metrics for substrate/active-site geometry to guide molecular dynamics (MD) simulations of halogenase SyrB2 with its substrates tethered to carrier protein.[39] This approach was necessary[39] because free molecular dynamics with approximate force fields can provide inaccurate descriptions of hydrogen bonding interactions and otherwise fail to adequately sample short, non-covalent distances.[40-43] Our simulations provided insight on previously unknown protein-substrate interactions that give rise to experimentally-measured substrate positions, which in turn impact

reaction selectivity in SyrB2. In this work, we apply the same computational protocol to recently discovered carrier-protein independent non-heme iron halogenases WelO5[11, 44-45] and BesD[14, 46] for which substrate-bound crystal structures have been obtained. These simulations provide the first guidance on the evolution of substrate-protein interactions from the resting state captured in crystal structures to a catalytically active moiety capable of halogenation in non-heme halogenases. This work also presents the first large-scale (i.e., > 200 atom) multi-scale, quantum mechanics/molecular mechanics (QM/MM) treatment of WelO5 and BesD that incorporate not just the minimal active site in the QM region but also the substrate and critical protein residues with potential mechanistic relevance. We expect insights from this work to guide current biomimetic catalyst design efforts inspired by non-heme iron enzymes.[47-51]

2. Overview of Experimental Guidance on Halogenases.

Although a wide number of non-heme Fe(II)/αKG-dependent enzymes are known, we focus here on three representative halogenases, SyrB2[6-7, 28-29, 31, 33-36, 39], BesD[14, 46], and WelO5[11, 44-45], with diverse substrates and review commonalities and differences established from crystal structures and biochemistry of their active sites (Figure 1 and Table 1). The crystal structures of each of the three halogenases were solved with an iron center coordinated to two His residues, bidentate αKG, and a Cl^- ligand[7, 11, 14] (Figure 1). The remaining sixth coordinate site, expected to be occupied by the oxo moiety upon formation of the iron-oxo intermediate, is at the axial position for all three halogenases. This site is vacant in the crystal structure of BesD[14] but is occupied by water or an NO ligand in the active sites of SyrB2[7] and WelO5[11].

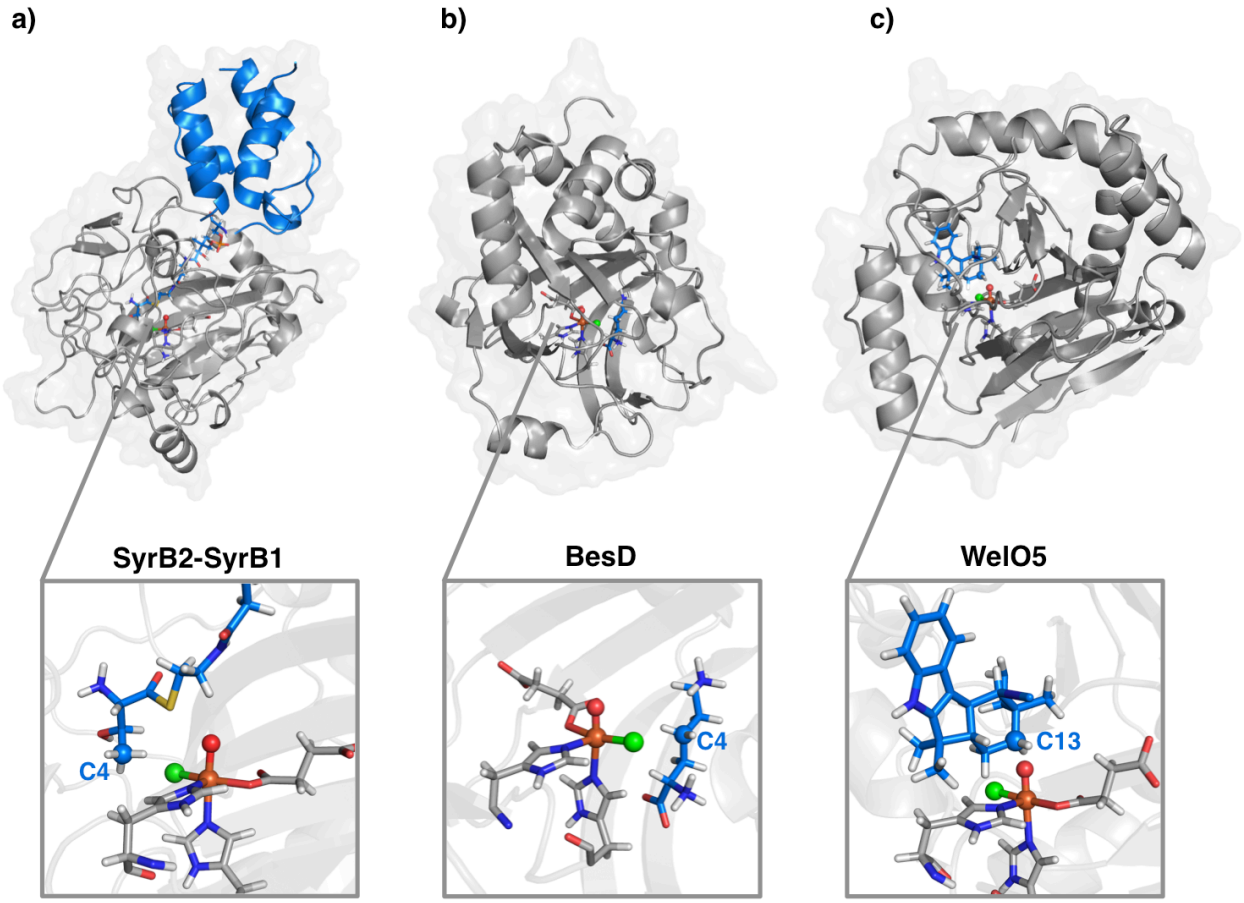

Figure 1. Crystal structures in cartoon representation and modeled iron-oxo intermediates in active sites for a) SyrB2-SyrB1 in complex with substrate Thr on the PPant arm[39], b) BesD in complex with substrate Lys[14], and c) WelO5 in complex with substrate 12-*epi*-fischerindole U[11]. The protein is shown in gray cartoon, and the substrate is shown in blue sticks or cartoon (i.e., for the SyrB1 carrier protein). Atoms in the inset active site structures are colored as: H in white, O in red, Cl in green, Fe in orange, S in yellow, and N in blue.

Table 1. Comparison of key features across the three halogeanses SyrB2, WelO5 and BesD.

feature	SyrB2	WelO5	BesD
function			
native function	chlorination	chlorination	chlorination
chlorination regioselectivity	C4	C13	C6
HAA stereoselectivity	-	*pro-R* H	*pro-R* H
substrate delivery	tethered	free	free
carrier protein	SyrB1	none	none
substrate			
native substrate	threonine	12-epi-fischerindole U	lysine
nature of substrate	amino acid	indole	amino acid
target C-H type	aliphatic	aliphatic	aliphatic
target C position	chain terminus	ring	mid-chain
target C chirality	achiral	prochiral	prochiral

crystal structure			
crystal structure with substrate	none	solved	solved
predicted oxo position	axial	axial	axial
target C-Fe distance	-	4.5 Å	4.8 Å
target C-Cl distance	-	5.7 Å	4.0 Å
experimental distances and angles			
Fe-H distance	4.2 ± 0.3 Å	3.7 Å	3.9 Å
H-Fe-O angle (axial-oxo)	85 ± 10°	50°	91°
H-Fe-O angle (equatorial-oxo)	85 ± 10°	120°	78°

Even though the crystal structures obtained thus far suggest the formation of the axial-oxo (ax-O) configurational isomer during catalysis for SyrB2[7], WelO5[11] and BesD[14], isomerization of the reactive metal-oxo species has been proposed for all three halogenases. Mössbauer[31] and nuclear resonance vibrational spectroscopy[33] identified and characterized the ax-O and equatorial-oxo (eq-O) configurational isomers of the iron-oxo intermediate in SyrB2. Additionally, DFT studies[28-29, 33] evaluated the role of active-site configurational isomers in controlling halogenation vs. hydroxylation reactivity observed in SyrB2[6, 35]. A spectroscopic study of SyrB2[34] invoked active-site isomerization to rationalize observations on substrate/active-site positioning, which we recently explored further using computational modeling[39]. Such studies are lacking for both BesD and WelO5 beyond the initial proposals made alongside published crystal structures.[11, 14] For WelO5, the eq-O configurational isomer was suggested[11] to be necessary to disfavor rebound hydroxylation. The reduced halogenation selectivity of WelO5 upon mutation of Ser189[11], a residue predicted to form a hydrogen bond (HB) with the oxo moiety if it occupies an equatorial site, provided support for this proposal. For BesD, the eq-O configurational isomer was proposed as a means to aid HAA based solely on the substrate/active-site geometry observed in the crystal structure[14] but has not been characterized experimentally or computationally.

Despite sharing common active sites, SyrB2, BesD, and WelO5 have diverse substrates

and modes of substrate delivery. SyrB2 and BesD both react on polar amino acid substrates, whereas WelO5 reacts on a bulky and hydrophobic indole substrate. As would be expected due to differences in their substrates, crystal structures show that the second-sphere residues of the active site are mostly polar or charged in BesD[14] (e.g., Arg67, Asp140, Asn219) and SyrB2[7] (e.g., Glu102, Asn123, Arg254) but hydrophobic (e.g., Ala82, Ile84, Val81) in WelO5[11] (Supporting Information Figure S1). Corresponding to the expected efficiency afforded by smaller, free-standing substrates, as in BesD and WelO5, compared to substrate complexes with carrier proteins, as in SyrB2, kinetic analysis reveals that SyrB2 only catalyzes 7 ± 2 turnovers before inactivation[6], whereas WelO5 can catalyze ~ 75 turnovers[45]. The total turnover number for BesD is not known.

The chemo- and regioselectivity of the three halogenases has been determined experimentally. Liquid chromatography with mass spectrometry analysis has confirmed that the primary products of SyrB2[6], WelO5[45] and BesD[14, 46] are their respective mono-chlorinated substrates, even though trace amounts of hydroxylated products were detected for these halogenases. Mutation of Ala for SyrB2[7] and Gly for BesD[46] and WelO5[11] to carboxylate-containing residues displaces the active site Cl and recover both the facial triad characteristic of hydroxylases[52-53] and the expected hydroxylation reactivity. Additionally, nuclear magnetic resonance analysis has been used to determine the regioselectivity of SyrB2[6-7], BesD[14, 46] and WelO5[44-45]. Both SyrB2[6-7] and BesD[14, 46] regioselectively chlorinate the γ carbon atom (i.e., C4) on the aliphatic side chains of their respective amino acid substrates (Figure 1). Although WelO5 performs regioselective halogenation[44-45], its reactivity is distinct, since it halogenates the aliphatic carbon atom in the cyclohexane ring (i.e., C13) of its substrate (Figure 1). In contrast to SyrB2, which targets the achiral carbon atom of its substrate during HAA and radical rebound,

both WelO5[45] and BesD[14] react on pro-chiral carbon atoms of their substrates, thus enabling stereoselectivity in addition to regioselectivity (Table 1).

While both BesD and WelO5 active sites share many structural and catalytic commonalties with the SyrB2 active site, the crystal structure of WelO5 does not have the expected substrate/active-site positioning that has been observed via spectroscopy in SyrB2 to coincide with halogenase activity (Figure 1). Spectroscopic studies[34] on SyrB2 have revealed that placement of the HAA target on substrate further from the iron moiety (4.2 ± 0.3 Å) and at a significant angle relative to the putative iron-oxo bond (85 ± 10°) ensures chlorination selectivity, while shorter distances (3.4 ± 0.2 Å) and more acute angles (64 ± 7°) are characteristic of hydroxylation selectivity. The substrate-isomer combination observed in the crystal structure of WelO5[11] is positioned to be more similar to a hydroxylase than a halogenase because it is more proximal (by ca. 0.6 Å) to the iron-oxo moiety than typically observed for halogenases (e.g., SyrB2[34]) and more distant (by ca. 1.7 Å) to Cl than would be expected[39] of a halogenase (Table 1). Moreover, the angle between the target substrate C-H bond and the iron-oxo moiety is either too small (ca. 40°) or too large (ca. 120°) in the crystal-structure-derived configurational isomers of WelO5 compared to the optimum angle (85 ± 10°) expected for chlorination[33] (Table 1).

In contrast to WelO5, the substrate-isomer combination observed in the crystal structure of BesD[14] is favorable for halogenation, in terms of both proximity to Cl and the angle between the target C-H bond and iron-oxo moiety[14] (Table 1). The crystal structure distance between the ax-O moiety and the C4 chlorination target of the substrate Lys, however, is significantly larger[14] (ca. 5.2 Å) than values typical (ca. 3.8 Å) for HAA. If instead we consider the eq-O configurational isomer of BesD, proposed[14] to aid HAA, the oxo moiety comes too close to C4 to

avoid hydroxylation. Thus, active site isomerization alone is not sufficient to explain the observed reactivity of halogenases. These observations thus motivate computational modeling of enzyme dynamics (i.e., with MD) with restraints guided by experimentally-motivated distances to investigate the interplay of substrate positioning, active-site isomerization, and substrate-protein interactions in WelO5 and BesD.

3. Results and Discussion.

3a. WelO5: The Role of Bulky Substrates in Dynamics.

To understand possible substrate-active-site dynamics with a bulky substrate, we enforce known, experimentally-motivated[34] distances and angles for halogenases in WelO5 (see Sec. 5). Although the distance restraints are derived from other halogenases, this approach allows us to observe how substrate-active site orientation can be expected to evolve during dynamics even for cases (e.g., in SyrB2) where we have observed[39] force fields to be insufficiently accurate to predict these quantities *a priori*. Regardless of the configurational isomer chosen, we are able to sample the experimentally-motivated substrate position with respect to the iron-oxo moiety in WelO5 (Figure 2 and Supporting Information Figure S2 and Table S1). In the ax-O isomer, an increase in the Fe···H distance and H···Fe-O angle is accompanied by a concomitant shift in the 12-*epi*-fischerindole U substrate position further away from the equatorial plane of iron in the active site (Figure 2). Conversely, when the oxo is in the equatorial plane, as is the case for the other two isomers, the substrate moves closer to the active site (Figure 2). In the eq-O isomer, the substrate orients directly over the equatorial plane of the active site due to the open proximal axial site, whereas for the axial-Cl (ax-Cl) isomer, the substrate is oriented away from the axial site occupied by Cl (Figure 2). While the substrate reorients in different ways across the three configurational isomers, a single pose in each case is observed to satisfy the experimentally-

motivated positioning constraints (Supporting Information Tables S2–S4).

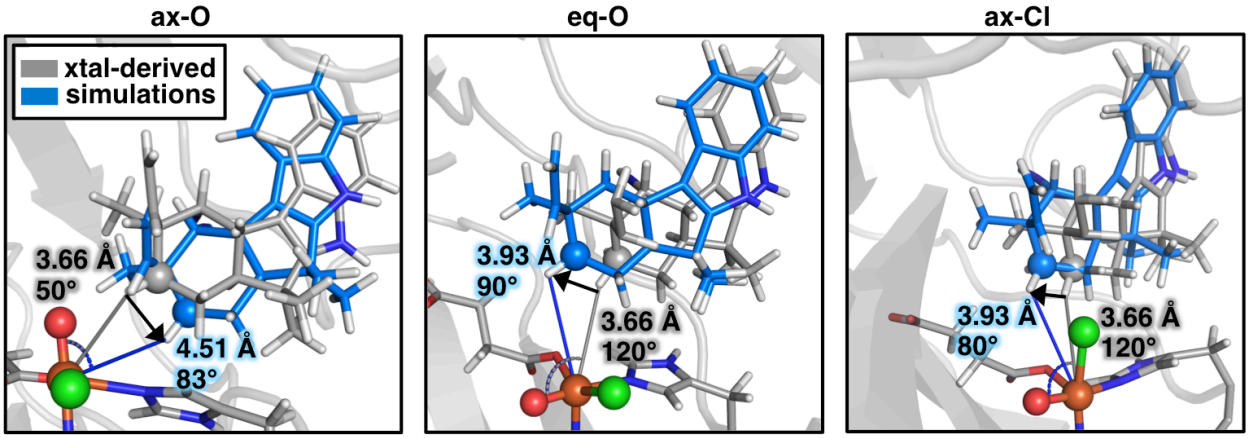

Figure 2. Comparison of Fe···H distance (in Å) and H···Fe-O angle (in °) for the ax-O (left), eq-oxo (middle), and ax-Cl (right) configurational isomers of WelO5. The substrate 12-*epi*-fischerindole U is shown in the starting position derived from crystal structure (gray sticks) and compared to a representative structure from experimentally-guided MD (blue sticks), with corresponding distances and angles labeled with a gray line and black text or blue line and text, respectively. The black arrow indicates the shift in position of C13 relative to iron-oxo after application of experimentally-motivated restraints. The chlorination target carbon C13 on substrate is shown as a ball, as are iron in orange, oxo in red, and Cl in green. The protein environment is shown as translucent gray cartoon.

Based on prior work[39], we anticipate satisfying experimentally-motivated positions, i.e., by increasing the distance of the reacting C13 to the iron-oxo moiety, should also correspond to shortening of the C13 distance to Cl in accordance with observed trends in other halogenases[34-35]. Over the three configurational isomers, the expected behavior is recovered in the ax-Cl isomer, with C13 remaining much closer to Cl (ca. 3.6 Å vs. 5.4 Å) compared to the oxo moiety (Figure 3 and Supporting Information Tables S4–S5). Conversely, in the ax-O configurational isomer, C13 is closer to the oxo moiety (ca. 5.2 Å) than Cl (ca. 6.7 Å) despite satisfying the experimentally-motivated distance to the iron-oxo moiety (Figure 3 and Supporting Information Tables S4–S5). Prior studies[11, 54] have suggested that the isomerization of the oxo moiety to the equatorial plane in the active site would result in the expected proximity of C13 to Cl. During

experimentally-guided MD of the eq-O isomer, C13 does sample distances somewhat closer to Cl (ca. 5.4 Å) than observed for the ax-O isomer (ca. 6.7 Å), but only when the C13 distance to the oxo (ca. 5.1 Å) is more proximal (Figure 3 and Supporting Information Tables S4–S5).

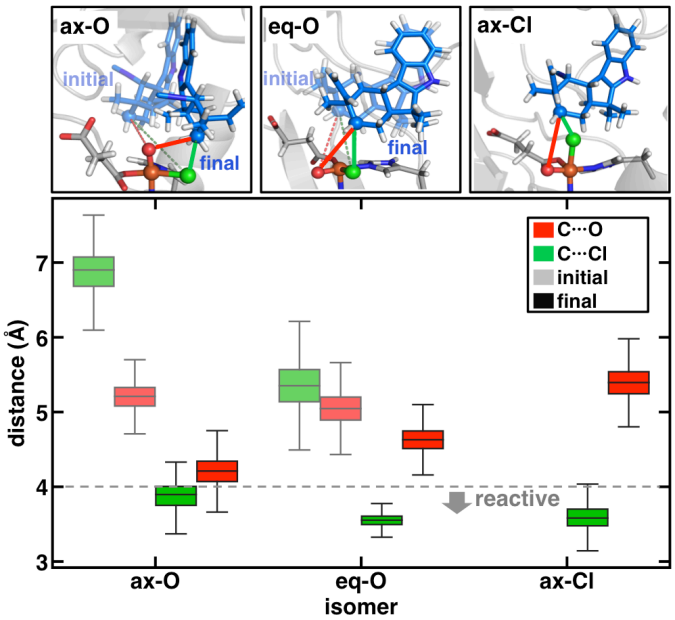

Figure 3. (top) The distance definitions for chlorination target carbon C13 on substrate 12-*epi*-fischerindole U to Cl (C···Cl, labeled in green) and oxo (C···O, labeled in red) ligands in the active site for the ax-O (left) eq-oxo (middle), and ax-Cl (right) configurational isomers of WelO5. The modified crystal structure is shown in translucent blue sticks (initial) along with the substrate after achieving the expected proximity to Cl vs. oxo (final) is shown in solid blue sticks, with the substrate C13 shown as a ball in both cases. (bottom) Target C13 C···Cl (green) and C···O (red) distance distributions (in Å) for ax-O, eq-O, and ax-Cl configurational isomers of WelO5 from MD in initial and final configurations are shown as box plots with the whiskers representing the range of the data and the box from the lower to upper quartile.

We next investigated whether we could enforce positioning of the substrate C13 closer to Cl than the oxo moiety while maintaining the experimentally-motivated distance and angle to the Fe(IV)=O. To do so, we added additional constraints during dynamics for both the eq-O and ax-O configuration isomers (Figure 3 and Supporting Information Tables S6–S8). In the eq-O isomer, rearrangement to satisfy this additional constraint is modest, with the substrate moving over the active site equatorial plane to shorten its distance from Cl (ca. 3.6 Å) while maintaining

the expected longer distance (ca. 4.6 Å) to the oxo moiety (Figure 3 and Supporting Information Tables S6–S8). For the ax-O isomer, the substrate must instead undergo a more substantial reorientation around the oxo to approach Cl (Figure 3 and Supporting Information Tables S6–S8). Although this reorientation positions C13 closer to Cl (ca. 3.9 Å vs 4.2 Å) than to the iron-oxo moiety, the reorientation disrupts substrate-protein interactions that could be expected to influence regioselectivity.

Based on the experimentally observed regioselectivity of WelO5[11, 44-45], we expect the mechanistically relevant substrate position to preferentially orient the target *pro-R* hydrogen atom on C13 close to the iron-oxo moiety. In the ax-O isomer with the additional restraint, however, a hydrogen atom of a non-target, methyl substituent on the substrate approaches close to the oxo moiety at a much shorter distance, d(O···H) = 2.8 Å, relative to the target hydrogen atom on C13 with d(O···H) = 4.4 Å (Figure 4). Thus, dynamics reveal that even if the substrate can be positioned in ax-O isomer with the C13 target carbon oriented closer to Cl, active-site positioning needed for regioselective chlorination cannot be simultaneously attained. Conversely, in both ax-Cl and eq-O isomers, only the target *pro-R* hydrogen of the chlorination target C13 remains closest to the oxo moiety with both the initial and additional restraints (see Sec. 5, Figure 4 and Supporting Information Figure S3). Prior work on WelO5[11, 54] had only considered the isomerization of the oxo moiety to the equatorial plane to attain expected substrate positioning in the active-site to favor regioselective chlorination. However, our dynamics simulations strongly suggest that a subsequent isomerization of Cl to the axial plane merits consideration to explain the observed reactivity.

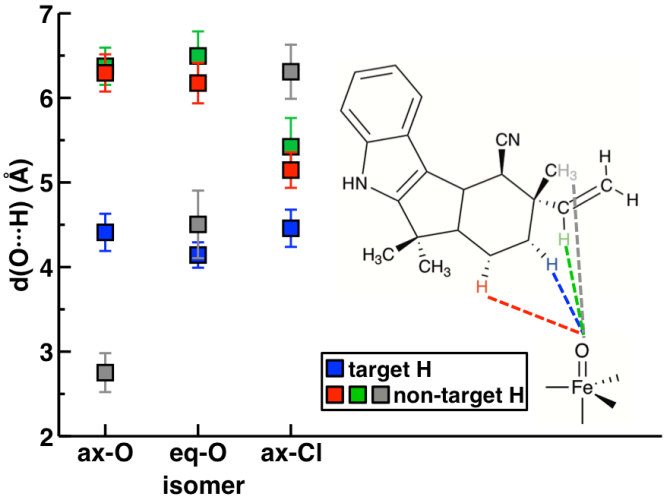

Figure 4. The average (square symbols) and standard deviation (error bars) of the distance of active-site-facing hydrogen atoms for the substrate 12-*epi*-fischerindole U to the oxo moiety, *d*(O···H) in Å, from experimentally-motivated MD of ax-O, eq-O, and ax-Cl configurational isomers of WelO5. The distance between the oxo moiety and the HAA target on the substrate is shown in blue, and the distances between the oxo moiety and non-target hydrogen atoms are shown in red, green and gray, as indicated in the inset at right. The iron-oxo bond vector is shown for the active site to define the relevant distances to the hydrogen atoms on the substrate.

Next, we investigated the effect of reorienting the substrate into an experimentally-motivated position on the interactions of the substrate with the greater protein environment. Classical energy decomposition analysis (see Sec. 5) reveals that WelO5's bulky 12-*epi*-fischerindole U substrate primarily interacts with the protein environment via dispersive interactions (Supporting Information Table S9). Given the large size and hydrophobic nature of the substrate, the observation of primarily dispersive interactions is expected. The substrate favorably interacts with several nonpolar residues (i.e., Val81, Ile84, Val90, Ile161, Phe169) and also exhibits van der Waals stabilization from polar Met221 and Met225 residues in all configurational isomers (Figure 5 and Supporting Information Table S9). Unsurprisingly due to the hydrophobic and bulky nature of the substrate, classical hydrogen bonding analysis indicates that the substrate forms few hydrogen bonds (HBs) with nearby protein residues in the experimentally-motivated positions for any isomer (Supporting Information Table S10). The

only HB observed, i.e., between the isocyanide on 12-*epi*-fischerindole U and the backbone N–H of Ala82, is preserved in most isomers and was previously observed in the crystal structure[11] (Figure 5 and Supporting Information Table S10). This HB is observed throughout all simulations except in the case where we simultaneously enforce a high substrate-oxo distance and a relatively shortened distance to Cl in the ax-O isomer with the additional restraint (see Sec. 5 and Supporting Information Table S10).

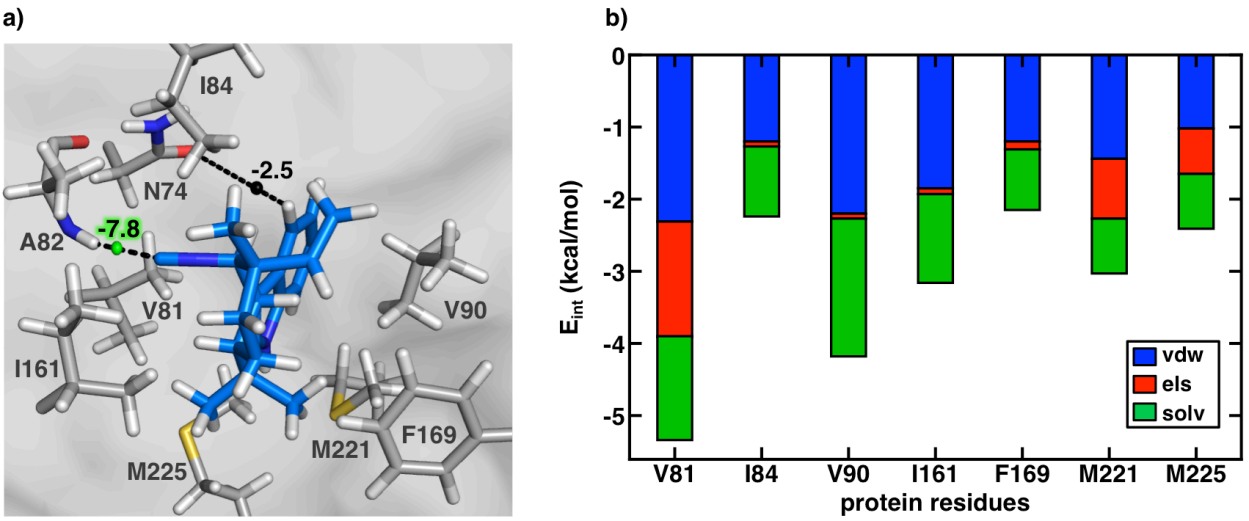

Figure 5. The interactions between substrate 12-*epi*-fischerindole U and the WelO5 protein environment. a) 12-*epi*-fischerindole U is shown in blue sticks with interacting protein residues shown as gray sticks. The HBs are indicated by black dashed lines, and the corresponding BCP is shown as a sphere along with its HB energy (in kcal/mol), with the stronger HB shown as a green sphere and a weaker (i.e., < 4.0 kcal/mol)[40-41] shown as a black sphere. b) The classical (GBSA) noncovalent interaction energies (in kcal/mol) between 12-*epi*-fischerindole U and WelO5 residues shown as a stacked bar chart of the van der Waals (vdw in blue), electrostatic (els in red), and sum of polar and non-polar solvation interaction energies (solv).

To confirm the observations from classical interaction analysis, we also carried out multi-scale QM/MM modeling of the WelO5 active site. We optimized representative configurations from dynamics with QM/MM using large QM regions (i.e., over 240 atoms) that incorporate essential protein residues (e.g., HB-forming Ala82) not included in prior computational studies[6]

of WelO5 (see Sec. 5 and Supporting Information Tables S11–S12). In contrast with prior studies, we include additional protein residues within 4.0 Å of 12-*epi*-fischerindole U, such as Asn74 and Val81, in our QM regions to detect any HB interactions that could have otherwise gone undetected[39-40, 42-43] with geometric criteria or classical interaction analysis from MD. Electronic structure analysis of these QM/MM-optimized snapshots confirms that 12-*epi*-fischerindole U only forms a strong HB with Ala82 in eq-O and ax-Cl isomers (Figure 5 and Supporting Information Table S13). During QM/MM optimizations, the HB donor-acceptor N-H···C distance shortens by 0.6 Å compared to the crystal structure-predicted distance (to ca. 2.6 Å), resulting in an even stronger HB than observed with MM. Bader analysis (i.e., QTAIM)[55-56] to detect bond critical points (BCPs) indicates that this HB is stabilizing by up to 7.8 kcal/mol (Figure 5 and Supporting Information Table S13). We only identify one additional BCP between the Asn74 sidechain carboxamide and a C-H on the benzene ring of the substrate. However, the QTAIM stabilization energy for this BCP is too weak (ca. 2.5 kcal/mol) for it to be classified as a strong HB, based on previously developed criteria[40] (Figure 5 and Supporting Information Table S13). This absence of a strong HB network between 12-*epi*-fischerindole U and the protein environment is in accordance with the bulky nature of the substrate and the presence of only a single polar group.

3b. BesD: Hydrogen Bonding to a Charged Substrate.

To compare the effect of free substrate dynamics for a case where the substrate is more highly charged and polar, we next performed constrained, experimentally-motivated MD on the carrier-independent BesD halogenase in its configurational isomers. We again aimed to identify the most probable adjustment in substrate positioning or active site isomerization needed to facilitate HAA and favor chlorination over hydroxylation (Figure 6 and Supporting Information

Figure S4 and Tables S14–S18). If we do not modify the BesD crystal structure configurational isomer from its ax-O orientation, the hydrogen atom on the chlorination target C4 of the substrate is already distant from the iron center (Fe···H distance 3.9 Å), as is expected for halogenases (Figure 6). This isomer also forms the expected optimum angle (H···Fe-O angle 91°) for activating the π-pathway for HAA resulting in subsequent chlorination[33] (Figure 6). Thus, experimentally-guided MD of this ax-O isomer exhibit very limited rearrangement from the initial configuration (Figure 6 and Supporting Information Figure S4). In this isomer, the chlorination target C4 also remains closer to Cl (ca. 3.7 Å) compared to the oxo moiety (ca. 5.4 Å) throughout dynamics (Figure 7 and Supporting Information Tables S17–S18). During MD, the hydrogen atom on C4 occasionally approaches even closer to the oxo moiety (d(O···H) = 3.5 Å) than what was expected from the crystal structure[14] (d(O···H) = 4.2 Å) for the ax-O orientation (Figure 7). When this occurs, the C4 still remains closer to Cl (d(C···Cl) = 3.7 Å) than the oxo moiety (d(C···O) = 4.6 Å), consistent with hypotheses[35] of sluggish HAA and preferential chlorination (Figure 7).

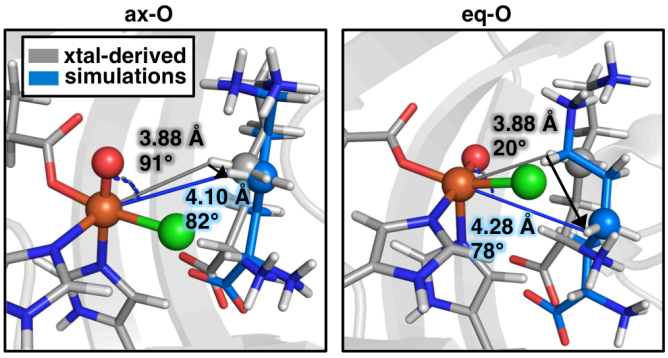

Figure 6. Comparison of the Fe···H distance (in Å) and H···Fe-O angle (in °) for the ax-O (ax-O, left) and eq-O (eq-O, right) configurational isomers of BesD. The substrate Lys in the starting position derived from the crystal structure for dynamics is shown as gray sticks and the substrate in experimentally-guided MD simulations is shown as blue sticks, with the corresponding distance and angle labeled in black and blue text or lines, respectively. The chlorination target carbon on substrate is highlighted as a ball as are iron in orange, oxo in red, and Cl in green. The black arrow indicates the shift in target carbon position from initial to final orientations.

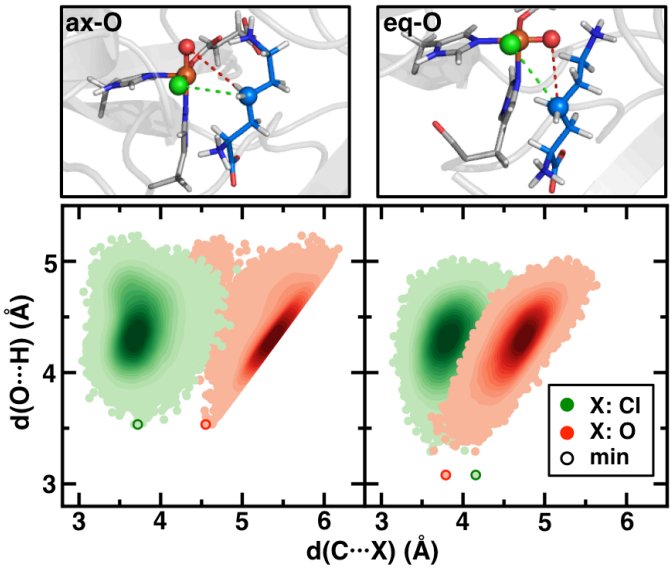

Figure 7. (top) The distance definitions for the target carbon on substrate Lys to Cl (C···Cl, labeled in green) and oxo (C···O, labeled in red) in the active site for the ax-O (left) and eq-O (right) configurational isomers of BesD. (bottom) Two-dimensional histograms with Gaussian kernel density estimation of the 2D distributions for C···Cl vs O···H (green) and C···O vs O···H (red) distances (in Å) for each isomer from experimentally-motivated MD. The point with the minimum *d*(O···H) of each distribution is shown by an outlined circle.

Starting from the modified eq-O configurational isomer that was suggested[14] to aid HAA, considerable substrate reorientation instead occurs during MD in order to satisfy the optimal substrate angle to the iron-oxo moiety (Figure 6 and Supporting Information Figure S4). The ca. 60° increase required for the H···Fe-O angle to ca. 78° falls within the experimentally-motivated range but reorients the substrate below the equatorial plane of the iron active site. As a result of this rearrangement, the C–H bond that is the target for HAA becomes oriented away from the iron-oxo moiety in a way that is suboptimal for activating the π-pathway for HAA that is proposed to lead to favored chlorination[33] (Figure 6). For this starting, modified eq-O configurational isomer, the substrate C4 is also closer to the oxo (ca. 3.4 Å) than Cl (ca. 4.0 Å). The restrained, experimentally-motivated dynamics do, however, shift relative distances and causes the target C4 to reside closer to Cl (ca. 3.8 Å) and more distant (ca. 4.7 Å) from the oxo

on average (Figure 7 and Supporting Information Tables S17–S18). In contrast to the ax-O isomer, the shortest distance of the C4 hydrogen atom to the oxo moiety sampled by the eq-O isomer (d(O···H) = 3.1 Å) also positions the C4 carbon closer to the oxo moiety (d(C···O) = 3.8 Å) than the Cl (d(C···Cl) = 4.2 Å, Figure 7). Thus, these dynamics indicate that the previously proposed[14] isomerization of the oxo moiety to the equatorial plane to aid HAA would disfavor typical distances expected to be key for favoring chlorination.

The eq-O isomer can be further ruled out based on analysis of the experimentally observed regioselectivity of BesD[46]. The experimentally-motivated substrate position in dynamics should exclusively orient the hydrogen atom on the C4 carbon toward the iron-oxo moiety. While this expectation is satisfied during the dynamics in the ax-O isomer, the eq-O exhibits distinct behavior (Figure 8). For the eq-O isomer, the hydrogen atom of the C6 carbon is unexpectedly more proximal to the oxo moiety (ca. 3.6 Å on average) relative to the C4 hydrogen atom (ca. 4.3 Å on average, Figure 8). This orientation is inconsistent with experimentally-observed regioselectivity of BesD (see Sec. 2).

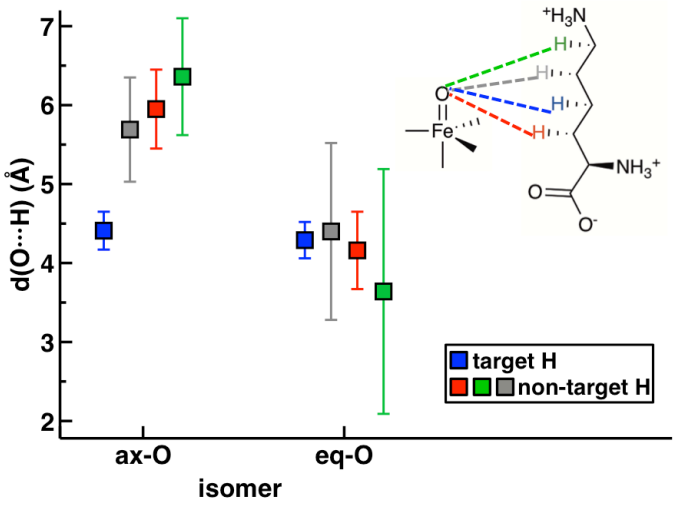

Figure 8. The average distance of active-site-facing hydrogen atoms on the substrate Lys to the oxo ligand in experimentally motivated dynamics of ax-O and eq-O configurational isomers of BesD. The average (square) and standard deviation (error bars) of the distance between oxo and

the target hydrogen atom is shown as blue, and the corresponding distances between oxo and non-target hydrogen atoms are shown in red, green, and gray as indicated in the inset at right. The iron-oxo bond vector is shown in the active site to define the relevant distances to the hydrogen atoms on substrate.

While the experimentally-motivated MD suggests the most likely substrate placement, we again employ for BesD classical and QM/MM interaction analysis, as we did in WelO5, to identify the mediating interactions in the greater protein environment that favor this placement. Classical analysis (see Sec. 5) indicates that the substrate Lys consistently forms multiple HBs with the protein environment during dynamics with both configurational isomers but especially in the ax-O case (Figure 9 and Supporting Information Table S19). As would be expected for a polar substrate forming several HBs with the protein environment, classical energy decomposition analysis suggests that the substrate-protein interactions are strongly electrostatic (i.e., rather than dispersive) in nature (Supporting Information Table S20). The HBs observed in the ax-O isomer[14] (i.e., that in the solved crystal structure) between the Lys substrate carboxylate and polar or charged protein residues (i.e., Arg74, His134, and Trp238) are maintained throughout the experimentally-motivated, restrained MD of both configurational isomers (Figure 9 and Supporting Information Table S19). The HBs to the two Lys substrate amine functional groups that were observed in the crystal structure[14] (i.e., with Trp138, Asp140, and Asn219) are also maintained throughout dynamics of ax-O isomer (Supporting Information Table S19). However, these HBs are not observed during dynamics of the eq-O isomer (Supporting Information Table S19).

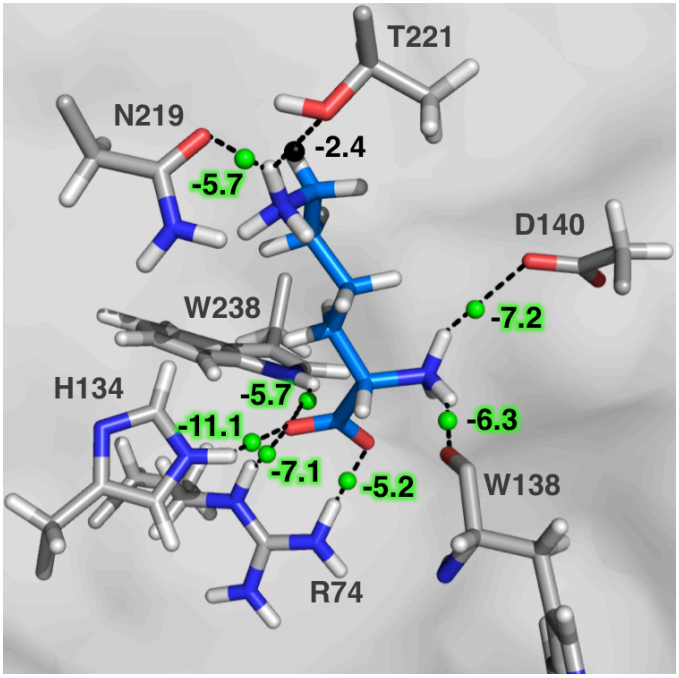

Figure 9. The hydrogen bonding interactions of the substrate Lys with the greater protein environment of BesD for the ax-O isomer (N219, D140, and W138 HBs are absent in the eq-O isomer). The substrate is shown in blue sticks, and protein residues are shown in gray sticks. The HBs are indicated by black dashed lines, and the corresponding BCPs are shown as spheres (green for strong HBs, black for weaker ones) along with their HB energies (in kcal/mol).

To strengthen the observations from classical modeling, we optimized snapshots from MD with QM/MM using large QM regions (see Sec. 5 and Supporting Information Tables S21–S22). The large QM region (over 240 atoms, see Sec. 5) employed for BesD incorporates all protein residues identified via classical HB analysis to interact with the Lys substrate, including His134, Asn219, and Trp238. Electronic structure analysis of these QM/MM-optimized snapshots indicates the presence of a strong hydrogen bonding network around the carboxylate of the Lys substrate in both configurational isomers, with significant stabilization energy from BCP analysis (Figure 9 and Supporting Information Table S23). The strength of this HB network appears to drive the observed downward shift of the substrate relative to the eq-O configurational isomer's iron-oxo plane to satisfy the experimentally-motivated position. This reorientation in the eq-O isomer disrupts the HBs mediated via Lys amine groups to strengthen the HBs with the

carboxylate. Conversely in the ax-O isomer, QM/MM analysis suggests that Lys attains the experimentally-motivated position while maintaining HBs mediated via both its carboxylate and amine groups.

The HBs between Lys substrate amine functional groups and Trp138, Asp140, and Asn219 in the ax-O isomer are confirmed with BCP analysis, with significant (i.e., 6 to 7 kcal/mol) stabilization energies for each interaction (Figure 9 and Supporting Information Table S23). The distances for these charge-mediated HBs are relatively unchanged after QM/MM geometry optimization, shortening by at most 0.1 Å for the case of the His134 N–H···O HB to the Lys carboxylate (i.e., from 2.06 to 1.96 Å). Overall, the protein environment provides significantly greater cumulative stabilization to the Lys substrate via HBs in ax-O isomer (ca. 51 kcal/mol) compared to the eq-O isomer (ca. 32 kcal/mol). The trends we observe for the role of the protein environment in positioning the Lys substrate are consistent with mutagenesis experiments[14] on BesD that indicated loss of halogenation selectivity when His134 was mutated to Ala. The His134 HB to the Lys carboxylate contributes 11 kcal/mol stabilization in our QM BCP analysis (Figure 9 and Supporting Information Table S23). The role of Asn219 in positioning the substrate revealed in our simulations and its correspondence to the reactivity of N219A mutant will be revisited next in relation to its proposed[14] interaction with the oxo moiety.

3c. Comparison of Diverse Halogenases.

Our simulations reveal differences in substrate-protein interactions across the three halogenases that correspond to the unique nature of their respective substrates, i.e., the bulky substrate in WelO5, the free, polar substrate in BesD, and the carrier-tethered substrate in SyrB2. In WelO5, the bulkier and neutral 12-*epi*-fischerindole U substrate forms only one HB with Ala82 via its isocyanide moiety. This single HB stabilizes 12-*epi*-fischerindole U by ca. 7

kcal/mol (Figure 5). The remaining favorable interactions between 12-*epi*-fischerindole U and the protein environment are all dispersive in nature (Supporting Information Table S9). Since WelO5 contains only one dominant, directional substrate-protein interaction, the 12-*epi*-fischerindole U can more flexibly orient in the active site. This flexibility affords the substrate the capability to reorient from the crystal structure position to the experimentally motivated position in both the eq-O and ax-Cl configurational isomers. In contrast for BesD, the small Lys substrate forms numerous charge-assisted HBs with the protein environment, acting both as an HB donor and acceptor via its carboxylate and amine functional groups respectively (Figure 9 and Supporting Information Table S19). This QM analysis, which represents the first QM modeling of the BesD active site and essential protein residues, indicates that these HBs cumulatively stabilize the Lys substrate by 32-51 kcal/mol depending on the isomer (Figure 9). Finally in analysis of the SyrB2-SyrB1 complex from our previous study[39], the carboxylate group of the native Thr substrate is covalently bonded to the PPant arm and does not form any HBs with the protein environment[39]. Instead, the backbone amine and side chain hydroxyl group on Thr act as an HB donor and acceptor, respectively. This ambifunctional HB configuration formed with Asn123, stabilizes Thr by 21 kcal/mol[39] (i.e., intermediate stabilization between WelO5 and BesD).

To understand the interplay between substrate/active-site dynamics and the inherent reactivity of the C–H bonds on the unique substrates of the three halogenases, we computed C-H bond dissociation energies (BDEs) using accurate domain-localized pair natural orbital-coupled cluster theory calculations (DLPNO-CCSD(T), see Sec. 5). The BDE of the target C-H bond on the SyrB2 substrate Thr is the highest, with the WelO5 12-*epi*-fischerindole U substrate and BesD Lys substrates both having lower BDEs, by 8.4 and 6.9 kcal/mol respectively, in

comparison (Figure10 and Supporting Information Table S24). Thus, the carrier protein-dependent halogenase SyrB2 must catalyze a more thermodynamically unfavorable reaction compared to carrier protein-independent halogenses BesD and WelO5. Additionally, we find that there are similar or even lower energy C–H BDEs for neighboring, non-target C–H bonds on the substrates, including the β carbon of Thr by 5.2 kcal/mol and target-carbon-neighboring atoms in the cyclohexane ring of 12-*epi*-fischerindole U by 0.2 kcal/mol. While in BesD the target C–H bond on Lys is indeed the weakest, adjacent, non-target C-H bonds are also of similar (i.e., within 2.7 kcal/mol) strength (Figure 10 and Supporting Information Table S24). The presence of these thermodynamically competitive C–H bonds on substrates adjacent to the HAA target C–H bond underscores the critical role of substrate positioning in overriding the influence of BDEs, and thus controlling regioselectivity in halogenases (Figure 10).

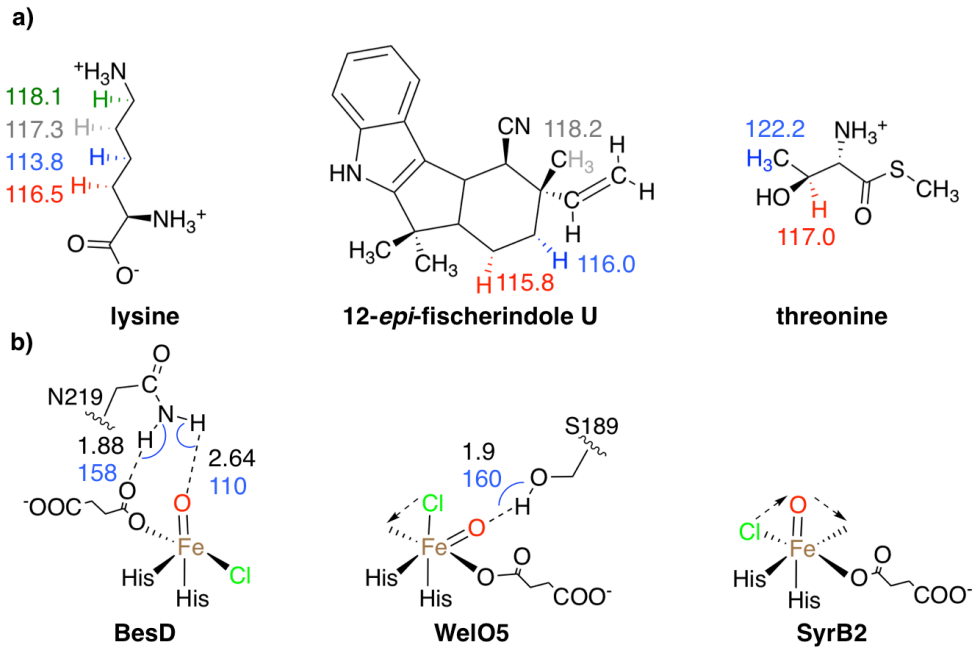

Figure10. a) The C–H BDEs (in kcal/mol) of substrates Lys (from BesD), 12-*epi*-fischerindole U (from WelO5) and Thr (from SyrB2, with truncated PPant arm). The HAA target hydrogen atom on each substrate is colored in blue, and the non-target hydrogen atoms are colored in red, green and gray where applicable. b) The QM/MM geometry-optimized structures for the most mechanistically relevant configurational isomers for BesD (ax-O, left), WelO5 (eq-O and ax-Cl, middle) and SyrB2-SyrB1 halogenases (ax-O and eq-O, right). The arrows with dashed lines indicate the multiple configurational isomers possible for WelO5 and SyrB2. The HB distances

(black dashed line and black text, in Å) and angles (blue arcs and blue text, in °) between succinate/oxo moiety and relevant protein residues are shown for BesD and WelO5.

Having confirmed the importance of substrate-positioning especially in BesD and SyrB2 to ensure preferential HAA for relatively high-energy C–H bonds, we return to the question of the role of hydrogen bonding interactions between the oxo moiety and nearby protein residues in suppressing hydroxylation reactivity.[11, 14, 54, 57] Previous studies have attributed the suppression of hydroxylation to hydrogen bonding interactions of the oxo/hydroxyl moiety to Arg254[57] in SyrB2, Ser189[11, 54] in WelO5, and Asn219[14] in BesD. In the SyrB2-SyrB1 complex, MD with the experimentally-motivated substrate position indicates that Arg254 is too far away, with a minimum N-H···O distance ca. 3.8 Å, to form a HB with the oxo moiety for any configurational isomer.[39] Instead, we observed an essential Asn123 interaction with the substrate to be critical for substrate positioning.[39] For WelO5, we identify a strong HB between the oxo moiety and the hydroxyl group on the sidechain of Ser189 for both the eq-O and ax-Cl isomers, which is stabilizing by around 8 kcal/mol in both cases (Figure 10 and Supporting Information Tables S25–S26). This HB promotes the mechanistically relevant active site isomerization of the oxo moiety from the axial plane to the equatorial plane prior to the radical rebound step.[54] While the Ser189 HB to the oxo moiety aids active site isomerization, our simulations indicate that this HB is not essential for suppressing hydroxylation reactivity. Experimentally-motivated dynamics reveals that the substrate itself can reorient to attain closer proximity to Cl compared to the oxo moiety to favor chlorination over hydroxylation reactivity in the two configurational isomers of WelO5.

Next, we characterized the possible HB between Asn219 and the oxo moiety in BesD since this HB has not been studied beyond predictions based on analysis of the crystal

structure[14]. In our simulations of BesD, the side chain amide group of Asn219 rarely approaches the oxo moiety in the ax-O isomer and remains distal and unable to form a HB in the eq-O isomer (Figure 10 and Supporting Information Table S25). In instances when Asn219 does approach the oxo moiety in the ax-O isomer, the distance is still long (N–H···O distance ca. 2.80 Å) and the angle relatively small (N–H···O angle ca. 109°) to form a strong HB interaction. The HB donor-acceptor distance shortens slightly (by 0.21 Å) during QM/MM geometry optimization, but the angle (ca. 110°) remains small (Figure 10 and Supporting Information Table S26). Using QM/MM BCP analysis (see Sec. 5), we confirm this interaction to be minimal, with a stabilization of less than 1.6 kcal/mol (Supporting Information Table S26).

The Asn219 residue instead forms strong HBs with the succinate cofactor in the active site and the Lys substrate (Figures 9–10). The QM/MM optimized structures reveal much shorter distances of N-H···O (ca. 1.88 Å) and N-H···N (ca. 2.05 Å) for the Asn219-succinate and Asn219-Lys HBs respectively, compared to the crystal-structure-derived distances[14] of 2.62 Å and 3.09 Å. These Asn219-Lys/succinate HBs are also more linear (ca. 141° and 158°) and thus stabilizing (by ca. 7.4 and 5.7 kcal/mol) compared to the Asn219-oxo HB (Figure 10 and Supporting Information Tables S19 and S22). Therefore, Asn219 interacts favorably with succinate and Lys, which compromises its ability to form a strong HB with the oxo moiety, due to restriction in motion imposed by the planarity of the Asn219 sidechain. Our simulations indicate that the loss of halogenation selectivity in the N219A[14] mutant is likely due to the disruption in Asn219-Lys and Asn219-succinate interactions, rather than the previously proposed Asn219-oxo HB.

4. Conclusions.

Using experimentally-motivated, restrained MD and large-scale QM/MM calculations,

we studied the roles of substrate positioning, active-site isomerization, and the greater protein environment in facilitating selective halogenation in the representative carrier-independent halogenases BesD and WelO5 in comparison to the carrier-dependent SyrB2. In BesD, the crystal-structure-like ax-O isomer is most aligned with experimentally-motivated substrate positioning and reactivity observations. We instead found that active-site isomerization is indeed necessary in WelO5 to achieve a geometry that favors regioselective chlorination. Besides the previously investigated eq-O isomer of WelO5, we found that the ax-Cl isomer was equally probable, thus motivating further studies that are underway in our lab. Our simulations confirmed expected differences in substrate-protein interactions between the small, charged Lys substrate of BesD and the bulky, neutral 12-*epi*-fischerindole U substrate of WelO5 during dynamics.

As judged through QM/MM analysis with large (> 200-atom) QM regions, the BesD protein environment stabilizes Lys the most (ca. 51 kcal/mol) through a number of HBs, followed by moderate stabilization (ca. 21 kcal/mol) of Thr by SyrB2 and limited stabilization (7 kcal/mol) of 12-*epi*-fischerindole U by WelO5. This quantitative comparison indicated that the stabilization provided by the protein environment in halogenases is correlated to the charge and polarity of their distinct substrates more than whether the enzyme requires a carrier protein.

Further QM analysis revealed some key similarities across SyrB2, WelO5 and BesD. We observed thermodynamically competitive C–H bonds on the substrates of all the three halogenases, which emphasized the critical role of strategic substrate positioning in ensuring desired regioselectivity. Additionally, we concluded that adjustment in substrate/active-site geometry via substrate reorientation, as in BesD and SyrB2, or active site isomerization, as in WelO5, is sufficient to suppress hydroxylation in all halogenases, instead of the previously

proposed oxo-anchoring HB in active site. For BesD, we also resolved the uncertainty around the mechanistic relevance of Asn219, which we showed stabilizes the substrate and succinate instead of anchoring the oxo moiety.

This work sets the stage for using experimentally-guided simulations for other non-heme iron enzymes, especially hydroxylases such as TauD[21, 58-59] and VioC[60] for which spectroscopically-determined distance and angle measurements needed to guide dynamics are available in the literature. Comparison of hydroxylases that are expected to exhibit distinct substrate poses for hydroxylation[21, 58-60] are expected to provide insights that further explain the divergent reaction outcomes in non-heme iron enzymes.

5. Computational Details.

Protein structure and preparation. The protein-substrate complexes of BesD with Lys and WelO5 with 12-epi-fischerindole were both prepared following the same protocol. The trajectories for analysis on SyrB2/SyrB1 with tethered L-Thr substratewere obtained from Ref. 39. In both BesD and WelO5, crystal structures were prepared by removing all crystallizing agents and omitting N-terminal unresolved residues. For BesD, the tetramer crystal structure (PDB ID: 6NIE[14]) with bound Lys was prepared. For WelO5 the trimer crystal structure (PDB ID: 5IQV[11]) with substrate 12-*epi*-fischerindole was used for simulation. The WelO5 chain A structure was also missing three mid-chain residues (i.e., A215, S216, and K217), which we added back using Modeller[61] loop refinement. In both cases, the active site in the crystal structure was modified using PyMOL[62] and Avogadro[63]. To model a ferryl-oxo intermediate with succinate, we manually removed a CO_2 group from αKG present in the crystal structures, and we inserted an oxo (i.e., for BesD) or modified the bound NO to an oxo (i.e., for WelO5).

To generate candidate active site configurational isomers, the oxo was placed: i) in an

axial position (ax-O), ii) in the equatorial plane with equatorial Cl^- (eq-O), and iii) in the equatorial plane with Cl^- in an axial position (i.e., ax-Cl, Supporting Information Figure S5). The charge states of amino acids were assigned using the H++ web server[64-66] assuming a pH of 7.0 with all other defaults applied. After manual charge assignment of residues adjacent to cofactors/substrates, the apo BesD structure has a net charge of -4 (Supporting Information Table S27). The charges from the substrate Lys (+1) and active site (i.e., Fe(IV)=O, succinate, and Cl^-, -1) contributes an overall net charge of 0, meaning the holoenzyme also has a net charge of -4 (Supporting Information Table S28). The WelO5 apoenzyme has a net charge of -6, and addition of the neutral substrate and active site (i.e., -1) gives the holoenzyme a net charge of -7 (Supporting Information Tables S29–S30). The protein-substrate complexes were solvated in a periodic rectangular prism box with at least a 10-Å buffer of TIP3P[67] water and neutralized with Na^+ counterions for a total simulation of over 35k atoms (WelO5: 36,160 and BesD: 35,127 atoms). Starting topology and coordinate files are provided in the Supporting Information.

The AMBER ff14SB force field[68] was used for all standard protein residues. The generalized AMBER force field (GAFF)[69] with restrained electrostatic potential (RESP)[70] charges obtained with Hartree-Fock/6-31G*[71] using GAUSSIAN16[72] were employed for the non-standard residues (i.e., succinate and substrates). The AMBER metal center parameter builder[73] (MCPB.py) was used to obtain force field parameters for the iron active site for both BesD and WelO5 in all isomers (Supporting Information Figures S6–S7 and Tables S31–S32). MCPB.py employed GAUSSIAN16[72] with the UB3LYP[74-76] functional and the LANL2DZ effective core potential[77] on Fe and 6-31G*[71] basis set for the remaining atoms to complete the geometry optimization, force constant calculation, and RESP charge calculations.

MM Equilibration and Dynamics. All MM MD used the GPU-accelerated PMEMD code

in AMBER18[78] and followed the same equilibration protocol: i) restrained (1000 steps) and unrestrained (2000 steps) minimization, ii) 10-ps NVT heating to 300 K with a Langevin thermostat with collision frequency of 5.0 ps^{-1} and a random seed, and iii) 1-ns NpT equilibration using the Berendsen barostat with a pressure relaxation time of 2 ps. The SHAKE algorithm[79] was applied in conjunction with a 2-fs timestep for all steps. The particle mesh Ewald method was used with a 10-Å real space electrostatic cutoff. Starting from step ii) of the equilibration procedure, harmonic restraints were added to enforce target (i.e., experimentally-motivated) distances and angles. The restraint values and force constants were adjusted iteratively to enforce sampling of target distances and angles (Supporting Information Table S33). Production dynamics (100 ns) were carried out during restraint adjustment, and 250 ns was obtained for each structure with the final restraint values.

Analysis of MD Trajectories. Snapshots from trajectories spaced 16 ps apart were clustered by the root mean square deviation (RMSD) of substrate heavy atoms and active site iron-oxo using the cpptraj utility of AMBER. The bottom-up, average-linkage algorithm with minimum distance between clusters of 4 Å was employed for a target of five clusters, based on guidelines in Ref. [80]. We employed AMBER MMPBSA.py[81] with the Generalized Born (GB)[82] approximation using the "OBC1" model[83], as suggested by benchmarks[84] for classical interaction analysis based on Ref. [81] for each of the clusters. Specifically, up to 625 snapshots spaced 1000 frames apart were obtained from each cluster for this analysis to compute pairwise residue electrostatic and van der Waals' interactions. Geometric hydrogen bond analysis and computed distances and angles in the active site were obtained with the cpptraj utility in AMBER18[78].

QM/MM or QM-only Simulation and Analysis. Snapshots from MD production were extracted for QM/MM geometry optimizations for all isomers of BesD and WelO5. The periodic

box was post-processed using PyMOL[62] to generate a 35-Å radius spherical droplet centered around the center of mass of each protein and further prepared with tleap. All QM/MM simulations were carried out using a developer version of TeraChem v1.9[85-86] for the QM portion and AMBER18[78] for the MM portion. The QM modeling employed unrestricted density functional theory (DFT) with the range-separated hybrid ωPBEh[87] (ω=0.2 $bohr^{-1}$) and a basis consisting of the LANL2DZ effective core potential[77] on Fe and 6-31G*[71] for the other atoms. Both enzymes were modeled in the high-spin, sextet state with near neutral net charge (BesD: -1, WelO5: +1, Supporting Information Tables S11 and S21). The QM region contained around 240 atoms BesD: 242 atoms, WelO5: 241 atoms) including link atoms (BesD: 16 atoms, WelO5: 18 atoms, Supporting Information Tables S11 and S21). Quantum theory of atoms in molecules (QTAIM) bond critical points (BCPs)[55] were obtained with Multiwfn[88] on QM/MM snapshots, and HB energies were estimated from the potential energy density of the closest BCP.[56]

Bond dissociation energies (BDEs) were calculated using domain-localized pair natural orbital coupled cluster with singles, doubles, and perturbative triples (DLPNO-CCSD(T))[89-90] as implemented in Orca v4.2.1. These energies were computed as the difference in energy between substrate and sum of the energy of an isolated hydrogen atom and substrate radical. Dunning-style correlation consistent double-ζ and triple-ζ (i.e., aug-cc-pVDZ and aug-cc-pVTZ) basis sets were employed for two-point[91-93] extrapolation to the complete basis set (CBS) limit (Supporting Information Table S24).

ASSOCIATED CONTENT

Supporting Information. Additional details and parameters of WelO5 and BesD QM, QM/MM, and MM MD simulations, including restraint adjustment, GBSA, and HB analysis from geometry as well as QTAIM. (PDF)

Starting topology and coordinate files for MD simulations (ZIP)

This material is available free of charge.

AUTHOR INFORMATION

Corresponding Author

*email: hjkulik@mit.edu phone: 617-253-4584

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The authors acknowledge primary support by the National Science Foundation under grant number CBET-1704266 (to R.M.) and CBET-1846426 (to R.M. and V.V.). H.J.K. holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund and an AAAS Marion Milligan Mason Award, which supported this work. This work was carried out in part using computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562. The authors acknowledge Adam H. Steeves for providing a critical reading of the manuscript.

References

1. Blunt, J. W.; Copp, B. R.; Hu, W.-P.; Munro, M. H.; Northcote, P. T.; Prinsep, M. R. Marine Natural Products. *Nat. Prod. Rep.* **2009,** *26*, 170-244.

2. Nakama, Y.; Yoshida, O.; Yoda, M.; Araki, K.; Sawada, Y.; Nakamura, J.; Xu, S.; Miura, K.; Maki, H.; Arimoto, H. Discovery of a Novel Series of Semisynthetic Vancomycin Derivatives Effective against Vancomycin-Resistant Bacteria. *J. Med. Chem.* **2010,** *53*, 2528-2533.
3. Latham, J.; Brandenburger, E.; Shepherd, S. A.; Menon, B. R.; Micklefield, J. Development of Halogenase Enzymes for Use in Synthesis. *Chem. Rev.* **2018,** *118*, 232-269.
4. Prakash, G. S.; Mathew, T.; Hoole, D.; Esteves, P. M.; Wang, Q.; Rasul, G.; Olah, G. A. N-Halosuccinimide/Bf3− H2O, Efficient Electrophilic Halogenating Systems for Aromatics. *J. Am. Chem. Soc.* **2004,** *126*, 15770-15776.
5. Alonso, F.; Beletskaya, I. P.; Yus, M. Metal-Mediated Reductive Hydrodehalogenation of Organic Halides. *Chem. Rev.* **2002,** *102*, 4009-4092.
6. Vaillancourt, F. H.; Yin, J.; Walsh, C. T. SyrB2 in Syringomycin E Biosynthesis Is a Nonheme FeII A-Ketoglutarate- and O2-Dependent Halogenase. *Proc. Natl. Acad. Sci.* **2005,** *102*, 10111-10116.
7. Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. Crystal Structure of the Non-Haem Iron Halogenase SyrB2 in Syringomycin Biosynthesis. *Nature* **2006,** *440*, 368-371.
8. Galonić, D. P.; Barr, E. W.; Walsh, C. T.; Bollinger Jr, J. M.; Krebs, C. Two Interconverting Fe (IV) Intermediates in Aliphatic Chlorination by the Halogenase Cytc3. *Nat. Chem. Biol.* **2007,** *3*, 113.
9. Wong, C.; Fujimori, D. G.; Walsh, C. T.; Drennan, C. L. Structural Analysis of an Open Active Site Conformation of Nonheme Iron Halogenase Cytc3. *J. Am. Chem. Soc.* **2009,** *131*, 4872-4879.
10. Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'connor, S. E.; Walsh, C. T. Cryptic Chlorination by a Non-Haem Iron Enzyme During Cyclopropyl Amino Acid Biosynthesis. *Nature* **2005,** *436*, 1191.
11. Mitchell, A. J.; Zhu, Q.; Maggiolo, A. O.; Ananth, N. R.; Hillwig, M. L.; Liu, X.; Boal, A. K. Structural Basis for Halogenation by Iron-and 2-Oxo-Glutarate-Dependent Enzyme Welo5. *Nat. Chem. Biol.* **2016,** *12*, 636-640.
12. Khare, D.; Wang, B.; Gu, L.; Razelun, J.; Sherman, D. H.; Gerwick, W. H.; Håkansson, K.; Smith, J. L. Conformational Switch Triggered by A-Ketoglutarate in a Halogenase of Curacin a Biosynthesis. *Proc. Natl. Acad. Sci.* **2010,** *107*, 14099-14104.
13. Pratter, S. M.; Light, K. M.; Solomon, E. I.; Straganz, G. D. The Role of Chloride in the Mechanism of O2 Activation at the Mononuclear Nonheme Fe (II) Center of the Halogenase Hctb. *J. Am. Chem. Soc.* **2014,** *136*, 9385-9395.
14. Neugebauer, M. E.; Sumida, K. H.; Pelton, J. G.; McMurry, J. L.; Marchand, J. A.; Chang, M. C. A Family of Radical Halogenases for the Engineering of Amino-Acid-Based Products. *Nat. Chem. Biol.* **2019,** *15*, 1009-1016.
15. Mitchell, A. J.; Dunham, N. P.; Bergman, J. A.; Wang, B.; Zhu, Q.; Chang, W.-c.; Liu, X.; Boal, A. K. Structure-Guided Reprogramming of a Hydroxylase to Halogenate Its Small Molecule Substrate. *Biochemistry* **2017,** *56*, 441-444.
16. Groves, J. T. Key Elements of the Chemistry of Cytochrome P-450: The Oxygen Rebound Mechanism. *J. Chem. Educ.* **1985,** *62*, 928.
17. Price, J. C.; Barr, E. W.; Glass, T. E.; Krebs, C.; Bollinger, J. M. Evidence for Hydrogen Abstraction from C1 of Taurine by the High-Spin Fe (IV) Intermediate Detected During Oxygen Activation by Taurine: A-Ketoglutarate Dioxygenase (Taud). *J. Am. Chem. Soc.* **2003,** *125*, 13008-13009.

18. Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M.; Krebs, C. The First Direct Characterization of a High-Valent Iron Intermediate in the Reaction of an A-Ketoglutarate-Dependent Dioxygenase: A High-Spin Fe (IV) Complex in Taurine/A-Ketoglutarate Dioxygenase (Taud) from Escherichia Coli. *Biochemistry* **2003,** *42*, 7497-7508.
19. Krebs, C.; Galonic Fujimori, D.; Walsh, C. T.; Bollinger Jr, J. M. Non-Heme Fe (IV)–Oxo Intermediates. *Acc. Chem. Res.* **2007,** *40*, 484-492.
20. Bollinger Jr, J. M.; Krebs, C. Stalking Intermediates in Oxygen Activation by Iron Enzymes: Motivation and Method. *J. Inorg. Biochem.* **2006,** *100*, 586-605.
21. Martinie, R. J.; Pollock, C. J.; Matthews, M. L.; Bollinger Jr, J. M.; Krebs, C.; Silakov, A. Vanadyl as a Stable Structural Mimic of Reactive Ferryl Intermediates in Mononuclear Nonheme-Iron Enzymes. *Inorg. Chem.* **2017,** *56*, 13382-13389.
22. Agarwal, V.; Miles, Z. D.; Winter, J. M.; Eustáquio, A. S.; El Gamal, A. A.; Moore, B. S. Enzymatic Halogenation and Dehalogenation Reactions: Pervasive and Mechanistically Diverse. *Chem. Rev.* **2017,** *117*, 5619-5674.
23. Smith, D. R.; Grüschow, S.; Goss, R. J. Scope and Potential of Halogenases in Biosynthetic Applications. *Curr. Opin. Chem. Biol.* **2013,** *17*, 276-283.
24. de Visser, S. P.; Latifi, R. Carbon Dioxide: A Waste Product in the Catalytic Cycle of A-Ketoglutarate Dependent Halogenases Prevents the Formation of Hydroxylated by-Products. *J. Phys. Chem. B* **2009,** *113*, 12-14.
25. Kulik, H. J.; Drennan, C. L. Substrate Placement Influences Reactivity in Non-Heme Fe (II) Halogenases and Hydroxylases. *J. Biol. Chem.* **2013,** *288*, 11233-11241.
26. Kulik, H. J.; Blasiak, L. C.; Marzari, N.; Drennan, C. L. First-Principles Study of Non-Heme Fe (II) Halogenase SyrB2 Reactivity. *J. Am. Chem. Soc.* **2009,** *131*, 14426-14433.
27. Rugg, G.; Senn, H. M. Formation and Structure of the Ferryl [Fe=O] Intermediate in the Non-Haem Iron Halogenase SyrB2: Classical and QM/MM Modelling Agree. *Phys. Chem. Chem. Phys.* **2017,** *19*, 30107-30119.
28. Borowski, T.; Noack, H.; Radon, M.; Zych, K.; Siegbahn, P. E. Mechanism of Selective Halogenation by SyrB2: A Computational Study. *J. Am. Chem. Soc.* **2010,** *132*, 12887-12898.
29. Huang, J.; Li, C.; Wang, B.; Sharon, D. A.; Wu, W.; Shaik, S. Selective Chlorination of Substrates by the Halogenase SyrB2 Is Controlled by the Protein According to a Combined Quantum Mechanics/Molecular Mechanics and Molecular Dynamics Study. *ACS Catal.* **2016,** *6*, 2694-2704.
30. Galonić, D. P.; Barr, E. W.; Walsh, C. T.; Bollinger, J. M.; Krebs, C. Two Interconverting Fe (IV) Intermediates in Aliphatic Chlorination by the Halogenase Cytc3. *Nat. Chem. Biol.* **2007,** *3*, 113-116.
31. Matthews, M. L.; Krest, C. M.; Barr, E. W.; Vaillancourt, F. H.; Walsh, C. T.; Green, M. T.; Krebs, C.; Bollinger Jr, J. M. Substrate-Triggered Formation and Remarkable Stability of the C− H Bond-Cleaving Chloroferryl Intermediate in the Aliphatic Halogenase, SyrB2. *Biochemistry* **2009,** *48*, 4331-4343.
32. Solomon, E. I.; Light, K. M.; Liu, L. V.; Srnec, M.; Wong, S. D. Geometric and Electronic Structure Contributions to Function in Non-Heme Iron Enzymes. *Acc. Chem. Res.* **2013,** *46*, 2725-2739.
33. Wong, S. D.; Srnec, M.; Matthews, M. L.; Liu, L. V.; Kwak, Y.; Park, K.; Bell III, C. B.; Alp, E. E.; Zhao, J.; Yoda, Y., et al. Elucidation of the Fe (IV)= O Intermediate in the Catalytic Cycle of the Halogenase SyrB2. *Nature* **2013,** *499*, 320.

34. Martinie, R. J.; Livada, J.; Chang, W.-c.; Green, M. T.; Krebs, C.; Bollinger Jr, J. M.; Silakov, A. Experimental Correlation of Substrate Position with Reaction Outcome in the Aliphatic Halogenase, SyrB2. *J. Am. Chem. Soc.* **2015,** *137*, 6912-6919.
35. Matthews, M. L.; Neumann, C. S.; Miles, L. A.; Grove, T. L.; Booker, S. J.; Krebs, C.; Walsh, C. T.; Bollinger, M. J. Substrate Positioning Controls the Partition between Halogenation and Hydroxylation in the Aliphatic Halogenase, SyrB2. *Proc. Natl. Acad. Sci.* **2009,** *106*, 17723-17728.
36. Srnec, M.; Wong, S. D.; Matthews, M. L.; Krebs, C.; Bollinger Jr, J. M.; Solomon, E. I. Electronic Structure of the Ferryl Intermediate in the A-Ketoglutarate Dependent Non-Heme Iron Halogenase SyrB2: Contributions to H Atom Abstraction Reactivity. *J. Am. Chem. Soc.* **2016,** *138*, 5110-5122.
37. Srnec, M.; Wong, S. D.; England, J.; Que, L.; Solomon, E. I. π-Frontier Molecular Orbitals in S= 2 Ferryl Species and Elucidation of Their Contributions to Reactivity. *Proc. Natl. Acad. Sci.* **2012,** *109*, 14326-14331.
38. Srnec, M.; Solomon, E. I. Frontier Molecular Orbital Contributions to Chlorination Versus Hydroxylation Selectivity in the Non-Heme Iron Halogenase SyrB2. *J. Am. Chem. Soc.* **2017,** *139*, 2396-2407.
39. Mehmood, R.; Qi, H. W.; Steeves, A. H.; Kulik, H. J. The Protein's Role in Substrate Positioning and Reactivity for Biosynthetic Enzyme Complexes: The Case of SyrB2/SyrB1. *ACS Catal.* **2019,** *9*, 4930-4943.
40. Vennelakanti, V.; Qi, H. W.; Mehmood, R.; Kulik, H. J. When Are Two Hydrogen Bonds Better Than One? Accurate First-Principles Models Explain the Balance of Hydrogen Bond Donors and Acceptors Found in Proteins. *Chem. Sci.* **2021,** *12*, 1147-1162.
41. Yang, Z.; Liu, F.; Steeves, A. H.; Kulik, H. J. Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action across Methyltransferases. *J. Phys. Chem. Lett.* **2019,** *10*, 3779-3787.
42. Yang, Z.; Mehmood, R.; Wang, M.; Qi, H. W.; Steeves, A. H.; Kulik, H. J. Revealing Quantum Mechanical Effects in Enzyme Catalysis with Large-Scale Electronic Structure Simulation. *React. Chem. Eng.* **2019,** *4*, 298-315.
43. Qi, H. W.; Kulik, H. J. Evaluating Unexpectedly Short Non-Covalent Distances in X-Ray Crystal Structures of Proteins with Electronic Structure Analysis. *J. Chem. Inf. Model.* **2019,** *59*, 2199-2211.
44. Hillwig, M. L.; Fuhrman, H. A.; Ittiamornkul, K.; Sevco, T. J.; Kwak, D. H.; Liu, X. Identification and Characterization of a Welwitindolinone Alkaloid Biosynthetic Gene Cluster in the Stigonematalean Cyanobacterium Hapalosiphon Welwitschii. *ChemBioChem* **2014,** *15*, 665-669.
45. Hillwig, M. L.; Liu, X. A New Family of Iron-Dependent Halogenases Acts on Freestanding Substrates. *Nat. Chem. Biol.* **2014,** *10*, 921-923.
46. Marchand, J.; Neugebauer, M.; Ing, M.; Lin, C.-I.; Pelton, J.; Chang, M. Discovery of a Pathway for Terminal-Alkyne Amino Acid Biosynthesis. *Nature* **2019,** *567*, 420-424.
47. Oloo, W. N.; Que Jr, L. Bioinspired Nonheme Iron Catalysts for C–H and C-C Bond Oxidation: Insights into the Nature of the Metal-Based Oxidants. *Acc. Chem. Res.* **2015,** *48*, 2612-2621.
48. Nam, W. High-Valent Iron (IV)–Oxo Complexes of Heme and Non-Heme Ligands in Oxygenation Reactions. *Acc. Chem. Res.* **2007,** *40*, 522-531.

49. Pangia, T. M.; Davies, C. G.; Prendergast, J. R.; Gordon, J. B.; Siegler, M. A.; Jameson, G. N.; Goldberg, D. P. Observation of Radical Rebound in a Mononuclear Nonheme Iron Model Complex. *J. Am. Chem. Soc.* **2018,** *140*, 4191-4194.
50. Hill, E. A.; Weitz, A. C.; Onderko, E.; Romero-Rivera, A.; Guo, Y.; Swart, M.; Bominaar, E. L.; Green, M. T.; Hendrich, M. P.; Lacy, D. C., et al. Reactivity of an Feiv-Oxo Complex with Protons and Oxidants. *J. Am. Chem. Soc.* **2016,** *138*, 13143-13146.
51. Rana, S.; Biswas, J. P.; Sen, A.; Clémancey, M.; Blondin, G.; Latour, J.-M.; Rajaraman, G.; Maiti, D. Selective C–H Halogenation over Hydroxylation by Non-Heme Iron (Iv)-Oxo. *Chem. Sci.* **2018,** *9*, 7843-7858.
52. Koehntop, K. D.; Emerson, J. P.; Que, L. The 2-His-1-Carboxylate Facial Triad: A Versatile Platform for Dioxygen Activation by Mononuclear Non-Heme Iron (II) Enzymes. *JBIC, J. Biol. Inorg. Chem.* **2005,** *10*, 87-93.
53. Que Jr, L. One Motif—Many Different Reactions. *Nat. Struct. Mol. Biol.* **2000,** *7*, 182.
54. Zhang, X.; Wang, Z.; Gao, J.; Liu, W. Chlorination Versus Hydroxylation Selectivity Mediated by the Non-Heme Iron Halogenase Welo5. *Phys. Chem. Chem. Phys.* **2020,** *22*, 8699-8712.
55. Bader, R. F. A Quantum Theory of Molecular Structure and Its Applications. *Chem. Rev.* **1991,** *91*, 893-928.
56. Espinosa, E.; Molins, E.; Lecomte, C. Hydrogen Bond Strengths Revealed by Topological Analyses of Experimentally Observed Electron Densities. *Chem. Phys. Lett.* **1998,** *285*, 170-173.
57. Pandian, S.; Vincent, M. A.; Hillier, I. H.; Burton, N. A. Why Does the Enzyme SyrB2 Chlorinate, but Does Not Hydroxylate, Saturated Hydrocarbons? A Density Functional Theory (DFT) Study. *Dalton Trans.* **2009**, 6201-6207.
58. Casey, T. M.; Grzyska, P. K.; Hausinger, R. P.; McCracken, J. Measuring the Orientation of Taurine in the Active Site of the Non-Heme Fe (II)/A-Ketoglutarate-Dependent Taurine Hydroxylase (Taud) Using Electron Spin Echo Envelope Modulation (Eseem) Spectroscopy. *J. Phys. Chem. B* **2013,** *117*, 10384-10394.
59. Muthukumaran, R. B.; Grzyska, P. K.; Hausinger, R. P.; McCracken, J. Probing the Iron− Substrate Orientation for Taurine/A-Ketoglutarate Dioxygenase Using Deuterium Electron Spin Echo Envelope Modulation Spectroscopy. *Biochemistry* **2007,** *46*, 5951-5959.
60. Mitchell, A. J.; Dunham, N. P.; Martinie, R. J.; Bergman, J. A.; Pollock, C. J.; Hu, K.; Allen, B. D.; Chang, W.-c.; Silakov, A.; Bollinger Jr, J. M. Visualizing the Reaction Cycle in an Iron (II)-and 2-(Oxo)-Glutarate-Dependent Hydroxylase. *J. Am. Chem. Soc.* **2017,** *139*, 13830-13836.
61. Šali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993,** *234*, 779-815.
62. Schrodinger, L. L. C., The PyMOL Molecular Graphics System, Version 1.7.4.3. 2010.
63. Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf.* **2012,** *4*, 17.
64. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating pK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res.* **2012,** *40*, W537-W541.

65. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating pKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005,** *33*, W368-W371.
66. Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. A Simple Clustering Algorithm Can Be Accurate Enough for Use in Calculations of pKs in Macromolecules. *Proteins: Struct., Funct., Bioinf.* **2006,** *63*, 928-938.
67. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983,** *79*, 926-935.
68. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015,** *11*, 3696-3713.
69. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004,** *25*, 1157-1174.
70. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993,** *97*, 10269-10280.
71. Harihara, P. C.; Pople, J. A. Influence of Polarization Functions on Molecular-Orbital Hydrogenation Energies. *Theor Chim Acta* **1973,** *28*, 213-222.
72. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H., et al. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.
73. Li, P.; Merz Jr, K. M. Mcpb. Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016,** *56*, 599-604.
74. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994,** *98*, 11623-11627.
75. Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988,** *37*, 785--789.
76. Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993,** *98*, 5648-5652.
77. Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for the Transition Metal Atoms Sc to Hg. *J. Chem. Phys.* **1985,** *82*, 270-283.
78. Case, D.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham III, T.; Cruzeiro, V.; Darden, T.; Duke, R.; Ghoreishi, D.; Gilson, M., et al. Amber 2018. *University of California, San Francisco* **2018**.
79. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977,** *23*, 327-341.
80. Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theory Comput.* **2007,** *3*, 2312-2334.
81. Miller III, B. R.; McGee Jr, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. Mmpbsa. Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012,** *8*, 3314-3321.

82. Tsui, V.; Case, D. A. Theory and Applications of the Generalized Born Solvation Model in Macromolecular Simulations. *Biopolymers* **2000,** *56*, 275-291.
83. Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large‐Scale Conformational Changes with a Modified Generalized Born Model. *Proteins: Struct., Funct., Bioinf.* **2004,** *55*, 383-394.
84. Hou, T.; Wang, J.; Li, Y.; Wang, W. Assessing the Performance of the MM/Pbsa and MM/Gbsa Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2011,** *51*, 69-82.
85. Ufimtsev, I. S.; Martínez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009,** *5*, 2619-2628.
86. Petachem. http://www.petachem.com. (accessed 10/15/18).
87. Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. A Long-Range-Corrected Density Functional That Performs Well for Both Ground-State Properties and Time-Dependent Density Functional Theory Excitation Energies, Including Charge-Transfer Excited States. *J. Chem. Phys.* **2009,** *130*, 054112.
88. Lu, T.; Chen, F. Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* **2012,** *33*, 580-592.
89. Riplinger, C.; Neese, F. An Efficient and near Linear Scaling Pair Natural Orbital Based Local Coupled Cluster Method. *J. Chem. Phys.* **2013,** *138*, 034106.
90. Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural Triple Excitations in Local Coupled Cluster Calculations with Pair Natural Orbitals. *J. Chem. Phys.* **2013,** *139*, 134101.
91. Zhong, S. J.; Barnes, E. C.; Petersson, G. A. Uniformly Convergent N-Tuple-Zeta Augmented Polarized (nZaP) Basis Sets for Complete Basis Set Extrapolations. I. Self-Consistent Field Energies. *J. Chem. Phys.* **2008,** *129*, 184116.
92. Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated Ab Initio Methods? *J. Chem. Theory Comput.* **2011,** *7*, 33-43.
93. Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997,** *106*, 9639-9646.