

# 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

Jing Huang,<sup>†</sup> Chunsen Li,<sup>\*,‡</sup> Binju Wang,<sup>§</sup> Dina A. Sharon,<sup>§</sup> Wei Wu,<sup>\*,†</sup> and Sason Shaik<sup>\*,§</sup>

<sup>†</sup>The State Key Laboratory of Physical Chemistry of Solid Surfaces, iChEM, Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry and College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, Fujian 361005, China

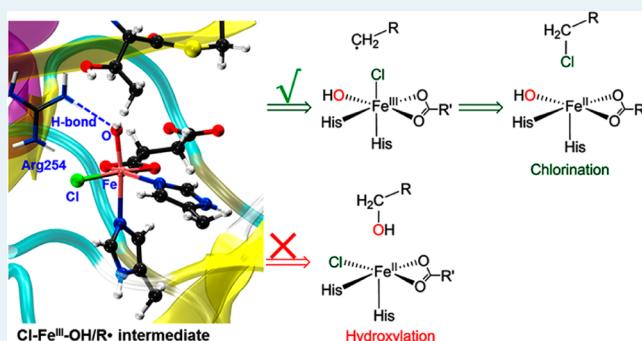
<sup>‡</sup>State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, Fujian 350002, China

<sup>§</sup>Institute of Chemistry and the Lise Meitner-Minerva Center for Computational Quantum Chemistry, The Hebrew University of Jerusalem, 91904, Israel

## Supporting Information

**ABSTRACT:** The enzyme SyrB2 employs an  $\text{Fe}^{\text{IV}}\text{-oxo}$  species to achieve selective C–H halogenation of L-threonine. Herein, we use combined quantum mechanical/molecular mechanical (QM/MM) calculations and molecular dynamics (MD) simulations to decipher the mechanism of selective halogenation by SyrB2. Our QM/MM calculations show the presence of three  $\text{Cl}\text{-Fe}^{\text{IV}}\text{-oxo}$  isomers which interconvert, and only the one having its oxo ligand pointing toward the target C–H bond is active during the hydrogen atom abstraction (H-abstraction) process. The fate of the formed  $\text{Cl}\text{-Fe}^{\text{III}}\text{-OH/R}^{\bullet}$  intermediate is determined by a hydrogen-bonding interaction between the Arg254 residue and the OH ligand of  $\text{Cl}\text{-Fe}^{\text{III}}\text{-OH}$ . The hydrogen bond not only prevents the OH group from participating in the followup rebound step to form a hydroxylated product but also facilitates the isomerization of the  $\text{Cl}\text{-Fe}^{\text{III}}\text{-OH/R}^{\bullet}$  intermediate so that the Cl is directed toward the alkyl radical. The role of Arg254 in regulating the selectivity of chlorination is further discussed and connected to the experimentally observed effect of mutations of Arg247 (Arg247Glu and Arg247Ala) in the related CurA halogenase. The Ala118Asp and Ala118Glu mutants of SyrB2 were investigated by MD simulations, and they were found to suppress the H-bonding interaction of Arg254 with  $\text{Cl}\text{-Fe}^{\text{III}}\text{-OH}$ : this result is in accord with their experimentally observed suppressed chlorination activity. This novel mechanism highlights the role of the H-bonding interaction between the protein and a reaction intermediate.

**KEYWORDS:** selective chlorination, SyrB2 enzyme, QM/MM calculation, MD simulation, hydrogen bond, protein environment, reaction mechanism, nonheme enzymes



## 1. INTRODUCTION

Mononuclear nonheme iron enzymes catalyze a myriad of chemically challenging transformations, including C–H bond hydroxylation, heteroatom oxidation, desaturation, epoxidation, and halogenation, which are important for the biosynthesis and degradation of compounds in metabolic processes.<sup>1–7</sup> Among these versatile enzymes, the specific family that requires  $\text{Fe}^{\text{II}}$  as a cofactor and  $\alpha$ -ketoglutarate ( $\alpha\text{KG}$ ) as a cosubstrate to activate aliphatic compounds to their halogenated products is called  $\alpha\text{KG}$ -dependent halogenases.<sup>8–13</sup> As halogenated compounds are the precursors of many organometallic species in the synthesis of pharmaceutical compounds,<sup>14–16</sup> unraveling the mechanisms of carbon–halogen bond formation is extremely important.

Syringomycin biosynthesis enzyme 2 (SyrB2), found in the bacterium *Pseudomonas syringae* pv *syringae* B301D,<sup>17</sup> chlori-

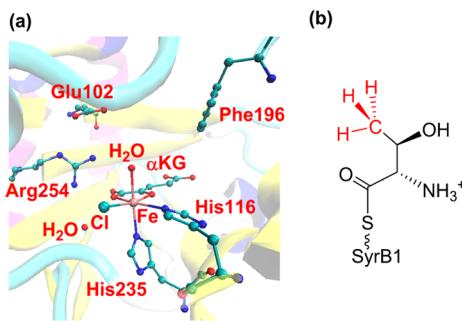
nates the  $\gamma$ -carbon of L-threonine in the biosynthesis of a phytotoxin called syringomycin E.<sup>9</sup> SyrB2 is a member of the  $\alpha\text{KG}$ -dependent halogenase family. Recently, wild-type SyrB2 and the Ala118Gly mutant were also found to be able to perform C–N bond formation reactions by substituting the native halide ion with the nitrogenous anions  $\text{N}_3^-$  and  $\text{NO}_2^-$ .<sup>18,19</sup> These results hint that the Cl is not by itself the cause of the selectivity of halogenation over hydroxylation, and they thereby underscore the need to understand why, unlike other similar enzymes, C–H activation does not end with hydroxylation.

As shown in Figure 1a, the active site of substrate-free SyrB2 taken from the X-ray structure (PDB code 2FCT)<sup>20</sup> contains a

Received: December 12, 2015

Revised: March 3, 2016

Published: March 11, 2016

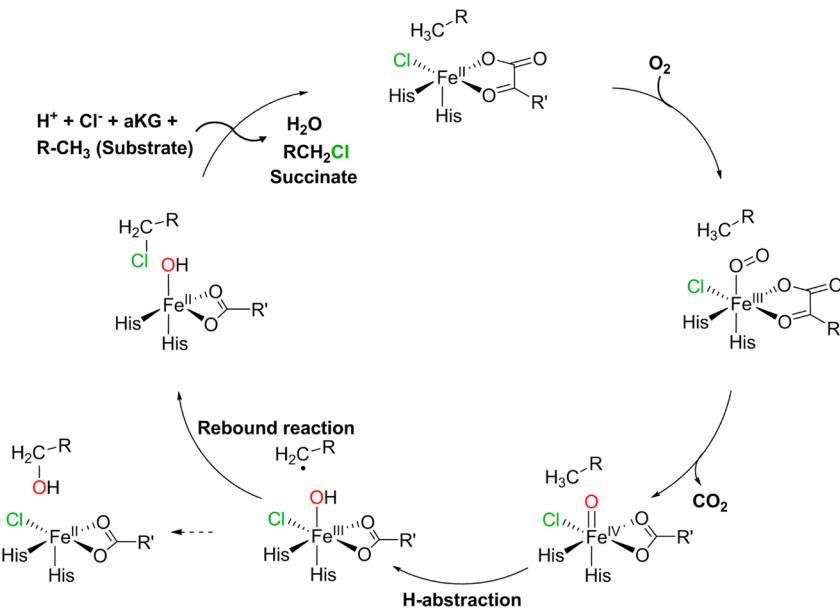


**Figure 1.** (a) Active site of substrate-free halogenase SyrB2 (PDB 2FCT), with some key amino acid residues and other ligands labeled. (b) Native substrate L-Thr-S-SyrB1 with its  $\gamma$ -C–H bonds in red.

hexacoordinated ferrous complex where the Fe<sup>II</sup> center is bound to  $\alpha$ KG, a chloro ligand, one crystal water, and two histidine residues (His116 and His235). The native substrate L-Thr-S-SyrB1 is shown in Figure 1b, and the  $\gamma$ -methyl group of the L-Thr moiety that is chlorinated in the catalytic cycle is highlighted in red.

Scheme 1 displays the catalytic cycle of the SyrB2 halogenase which was proposed in previous experimental studies.<sup>9,19–29</sup> The catalytic cycle is initiated by the binding of the O<sub>2</sub> molecule to the Fe<sup>II</sup> center to form a ferric–superoxide species. Subsequently, the distal oxygen atom nucleophilically attacks the  $\alpha$ KG ligand to form a peroxy-bridged Fe<sup>IV</sup> species. The subsequent cleavage of the bridged O–O bond and the decarboxylation of  $\alpha$ KG lead to the formation of a Cl–Fe<sup>IV</sup>–oxo complex. Thereafter, a hydrogen atom is abstracted from the substrate L-Thr by the Cl–Fe<sup>IV</sup>–oxo species to form an alkyl radical and an Fe<sup>III</sup>–OH complex. This is followed by a rebound reaction of the radical with the chlorine atom to generate the chlorinated product. A competitive hydroxylation reaction may occur via the rebound of the alkyl radical to the Fe<sup>III</sup>–OH group. The catalytic cycle finishes with the binding of another molecule of  $\alpha$ KG and another molecule of the substrate to reproduce the resting ferrous reactant complex.

### Scheme 1. Catalytic Cycle of the $\alpha$ KG-Dependent Halogenase SyrB2

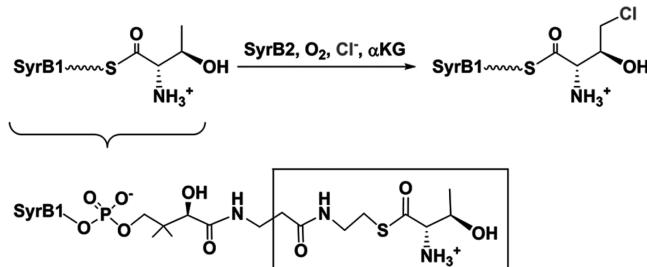


Experimental and quantum mechanical (QM) investigations have been performed to gain insight into the selective halogenation by SyrB2, and various mechanistic explanations have been proposed on the basis of such studies.<sup>29–34</sup> De Visser and Latifi have shown that the byproduct of CO<sub>2</sub> from O<sub>2</sub> activation shuttles the hydroxyl group of the Fe<sup>III</sup>OH moiety away to prevent the occurrence of hydroxylation by forming bicarbonate.<sup>30</sup> Kulik et al. proposed that chlorination following hydrogen atom abstraction (H-abstraction) is barrierless; therefore, the halogenases do not hydroxylate their substrates.<sup>31</sup> Hillier et al. suggested that the predominant chlorination should be caused by the protonation of the hydroxo intermediate, which inhibits the hydroxylation reaction.<sup>32</sup> Noack and Siegbahn proposed that the barriers for hydroxylation and halogenation by the biomimetic equivalent of SyrB2 are comparable.<sup>33</sup> On the basis of a model study considering first-shell and some of the second-shell residues of the iron center of SyrB2, Borowski et al. found a carbon-radical intermediate that favors chlorination is generated by H-abstraction via a second oxoferryl species isomer, which has a lower barrier in H-abstraction than the original oxoferryl species. Furthermore, they proposed that the interchange of oxo and chloride ligands relative to their positions in the crystal structure could control the selectivity.<sup>33</sup> Using nuclear resonance vibrational spectroscopy (NRVS) experiments and density functional theory (DFT) calculations, Wong et al. found that the oxoferryl species which has its Fe<sup>IV</sup>–oxo bond perpendicular to the C–H bond of the native substrate L-Thr favors chlorination.<sup>34</sup> Moreover, experimental findings led to the proposal that the different positions of the methylene radical vis-à-vis the chlorine and hydroxyl ligands can lead to the different halogenation and hydroxylation reactivities not only for the native substrate but also for non-native substrates.<sup>19,25,36,37</sup> Using DFT methods augmented with a Hubbard U correction, Kulik and Drennan found that the propensity of the native substrate toward halogenation is strongly dependent on deeper substrate delivery into the active site following H-abstraction.<sup>38</sup>

The above mechanistic proposals seem to be quite diverse. Furthermore, it was found that the halogenation/hydroxylation selectivity strongly depends on the protein environment. For

example, as found by Blasiak et al.,<sup>20</sup> Ala118Asp and Ala118Glu mutations completely suppress chlorination activity of the enzyme. As previous QM studies did not consider the complete protein environment of this enzymatic reaction, we deemed it necessary to carry out combined quantum mechanical/molecular mechanical (QM/MM)<sup>39–43</sup> calculations and molecular dynamics (MD) studies to thoroughly decipher the mechanism of halogenation/hydroxylation selectivity by SyrB2. Indeed, QM/MM studies of other nonheme enzymes, such as TauD,<sup>44</sup> AlkB,<sup>45,46</sup> and HEPD,<sup>47</sup> have clearly and effectively elucidated the crucial influence of protein environments on complex enzymatic reactions. Herein, we started from the active species of Fe<sup>IV</sup>–oxo, which has been experimentally characterized in many halogenases,<sup>11,25,34,36,48</sup> to explore the selective mechanism of halogenation vs hydroxylation in SyrB2. As the crystal structure for the substrate's carrier protein SyrB1 is absent,<sup>20</sup> we truncated the L-Thr substrate according to the strategy used in a previous study,<sup>33</sup> as specified in **Scheme 2**. As will be demonstrated below,

**Scheme 2.** Truncation of the Substrate Component Involving L-Thr from L-Thr-S-SyrB1<sup>a</sup>



<sup>a</sup>The fragment in the inner frame is the substrate used in this work.

selectivity arises from the H-bonding interaction between the Arg254 residue and the OH ligand of the Cl–Fe<sup>III</sup>–OH species after H-abstraction. This H-bonding interaction prevents the OH group from participating in the followup rebound step and in a nutshell steers the reaction's outcome toward selective chlorination.

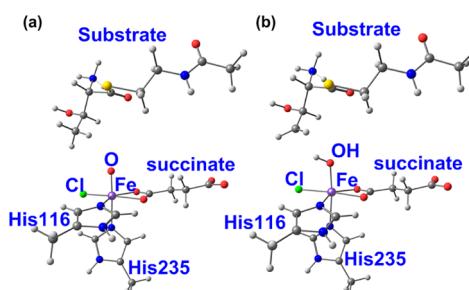
## 2. METHODS

**2.1. Setup of the System.** The system was first constructed using the oxoferryl species to model the H-abstraction process. The initial coordinates were taken from the X-ray crystal structure (PDB code 2FCT). Only chain A was kept in our study.<sup>20</sup> The missing residues, Ile57–Ser58–Gly59–Gly60, in the inner protein were added by superimposing 2FCT on another crystal structure<sup>20</sup> (PDB code 2FCU, chain A) to determine the missing coordinates. The missing residues on the protein surface, His0–Met1–Ser2, were ignored. The detergent molecules Dsu896 and Cl903 were removed, as they are irrelevant for the catalytic reaction. To generate the channel for the whole substrate L-Thr with a length of about 17 Å, we rotated the side chain of Phe196 around the C<sub>α</sub>–C<sub>β</sub> bond with an angle of 120°, as suggested by Borowski et al.<sup>33</sup> and previous experimental work.<sup>20</sup> Subsequently, the substrate was docked into the channel above the active site by using the AutoDock 4.2 program.<sup>49</sup> We assigned the protonation states of acidic and basic residues, using the pK<sub>a</sub> values given by the empirical PROPKA<sup>50</sup> procedure, and we verified the results through visual inspection. The detailed protonation scheme is included in the Supporting Information. The missing hydrogen atoms were added by the HBUILD

module<sup>51</sup> embedded in CHARMM.<sup>52</sup> The positions of the hydrogen atoms were optimized with 100 steps of adopted basis Newton–Raphson (ABNR) minimization. A water layer of 16 Å thickness was constructed around the enzyme using the VMD software package.<sup>53</sup> The inner 8 Å of the solvent was minimized by 2000 steps of ABNR followed by 500 steps of steepest descent (SD) and then equilibrated by MD for 50 ps at 300 K while the remainder of the system was kept fixed.

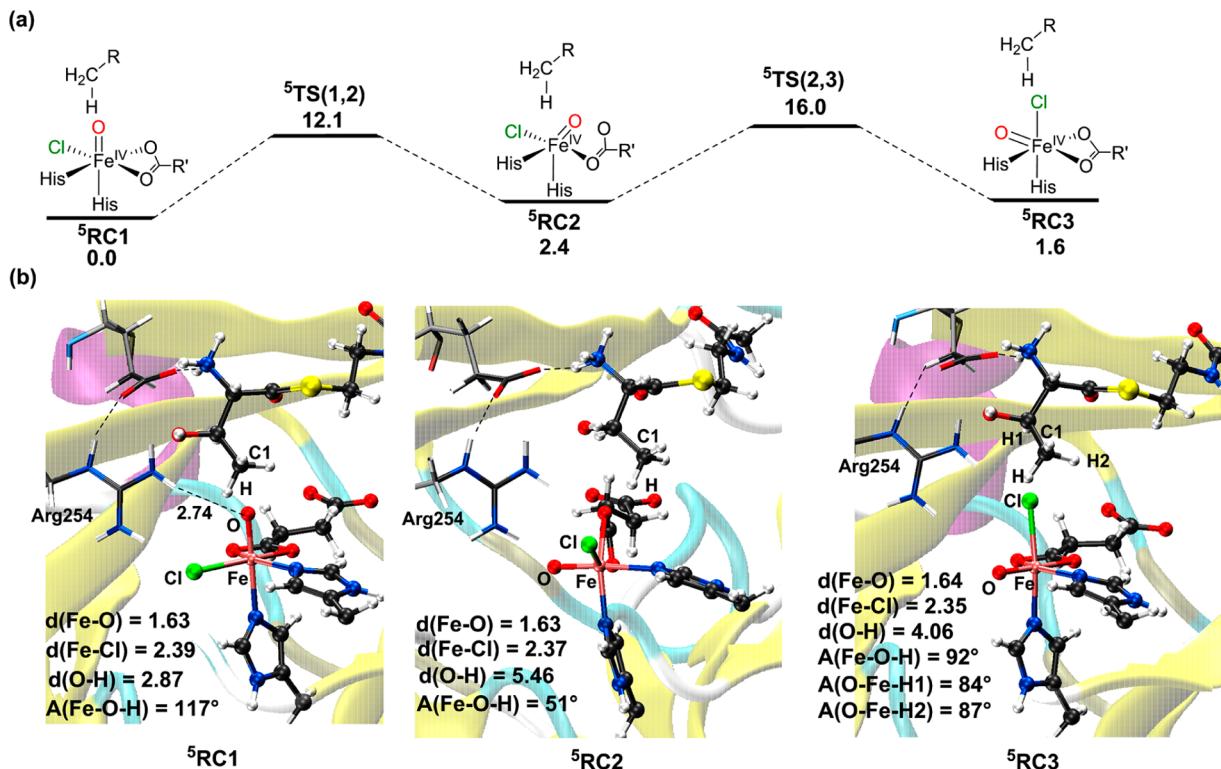
For MD runs, to investigate the effect of surrounding residue fluctuations on the halogenation/hydroxylation selectivity,<sup>19,30,32,36,38</sup> the entire system was also rebuilt as the ferric hydroxyl radical intermediate. The atomic electrostatic potential (ESP) charges were taken from the QM/MM calculations, while all other features were kept the same as for the oxoferryl species. Both the oxoferryl and ferric hydroxyl radical species contain 26971 atoms, including 22104 atoms of solvent for the wild-type SyrB2 (see Figure S1 in the Supporting Information). After the full solvation procedure, a productive MD simulation of 20 ns was run for both the oxoferryl species and the ferric hydroxyl radical intermediate using the CHARMM22 force fields implemented in the CHARMM program. The coordinates of the entire Cl–Fe<sup>IV</sup>=O or Cl–Fe<sup>III</sup>–OH unit and the metalligating residues as well as the outer 8 Å of the solvent layer were kept fixed during all MD simulations. Two representative snapshots from the MD trajectory of the oxoferryl species were selected for studying the H-abstraction reaction (snapshots at 2 and 10 ns), while four representative snapshots at 1, 1.5, 2, and 5 ns from the MD trajectory of the ferric hydroxyl radical intermediates were selected to investigate the competition of the halogenation and hydroxylation reactions. To understand the root cause of the loss of reactivity in the Ala118Asp and Ala118Glu mutants, we performed 20 ns MD simulations on the ferric hydroxyl radical intermediate of Ala118Asp and Ala118Glu variants using the same setup as for the wild-type SyrB2 enzyme except for the specific mutated residues.

**2.2. QM/MM Methodology.** The QM region in our QM/MM calculations contained 70 atoms for the oxoferryl species (or ferric hydroxyl/carbon-radical intermediate), including Fe<sup>IV</sup>=O (or Fe<sup>III</sup>–OH), the chloro ligand, succinate, two 4-methylimidazole groups, and the substrate, as shown in Figure 2.<sup>31–34</sup>

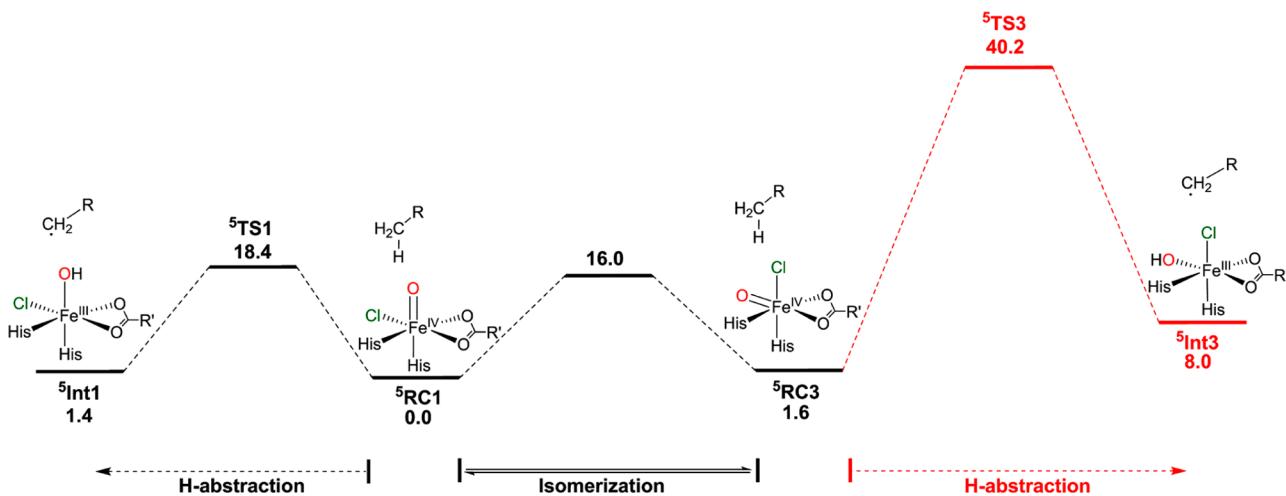


**Figure 2.** QM regions for (a) the oxoferryl species and (b) the ferric hydroxyl/carbon-radical intermediate in the QM/MM calculations.

All calculations were carried out within ChemShell,<sup>54</sup> combining Turbomole<sup>55</sup> as the QM code and DL\_POLY<sup>56</sup> as the MM code. The electronic embedding scheme<sup>57</sup> was applied to include the polarizing effect of the enzyme environment on the QM region. The QM/MM boundary was treated by hydrogen link atoms<sup>58</sup> with the charge shift model.<sup>59</sup> For the QM region, the unrestricted B3LYP (UB3LYP)<sup>60–62</sup> functional was employed with two basis sets. The choice of UB3LYP relied



**Figure 3.** (a) QM/MM (UB3LYP/B2) relative energies (in kcal/mol) for the isomerizations of three  $\text{Fe}^{\text{IV}}=\text{O}$  isomers,  ${}^5\text{RC1}$ ,  ${}^5\text{RC2}$ , and  ${}^5\text{RC3}$ , in the snapshot at 10 ns from the MD trajectory of the oxoferryl species. The ZPE and dispersion corrections are included in the relative energies. (b) QM/MM-optimized geometries of  ${}^5\text{RC1}$ ,  ${}^5\text{RC2}$ , and  ${}^5\text{RC3}$ , along with the key structural parameters (in Å) at UB3LYP/B1.

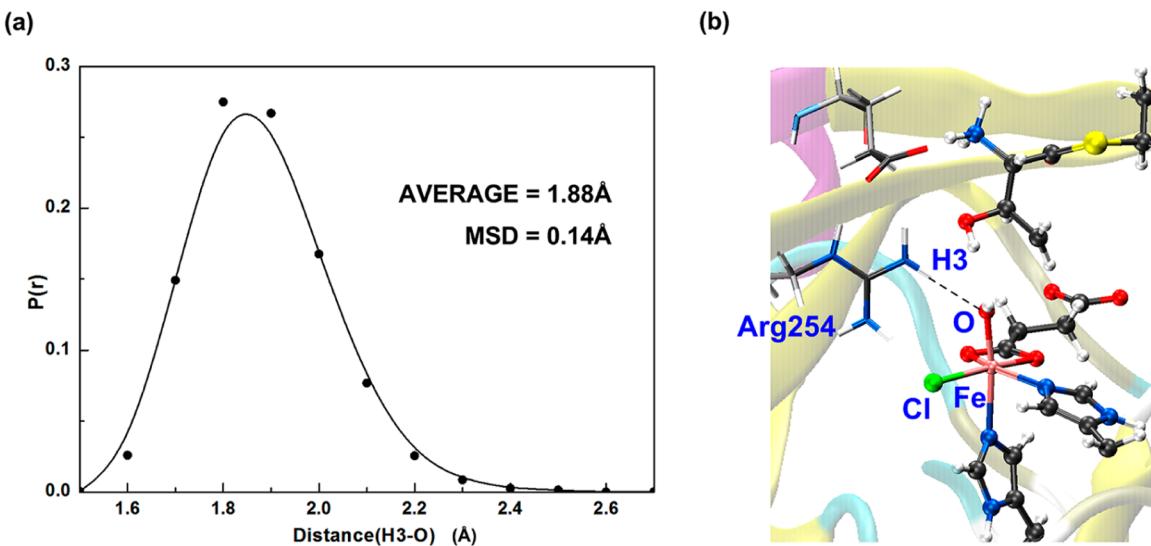


**Figure 4.** QM/MM (UB3LYP/B2) relative energies (in kcal/mol) for H-abstraction from the native substrate by  ${}^5\text{RC1}$  and  ${}^5\text{RC3}$  and isomerization between two isomers in the snapshot at 10 ns from the MD trajectory of the oxoferryl species. The ZPE and dispersion corrections are included in the relative energies. The pathway from  ${}^5\text{RC3}$  to yield  ${}^5\text{Int3}$  is not taken and is drawn as a red line.

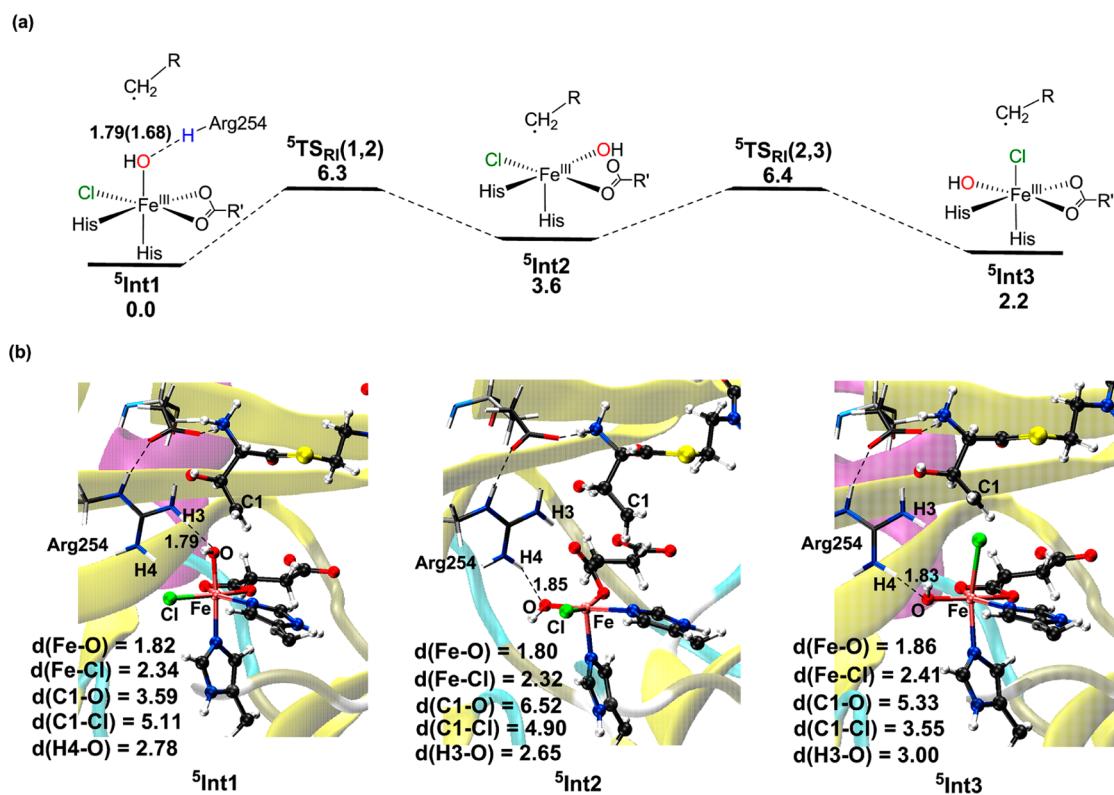
on past findings of a good match between UB3LYP and CCSD(T)<sup>63</sup> and between UB3LYP/MM and CASSCF/MM<sup>64</sup> for high-valent  $\text{Fe}^{\text{IV}}$ -oxo oxidants. A recent careful benchmark study by Thiel and Neese showed that B3LYP is the most suitable functional for  $\text{Fe}^{\text{IV}}$ -oxo systems.<sup>65</sup>

A split valence polarized basis set (B1), LACVP for iron<sup>66</sup> and 6-31G\*\* for other atoms, was used for the geometry optimization and zero point energy (ZPE) calculations, while an extended all-electron basis set (B2), def2-TZVP<sup>67</sup> for all atoms, was used to correct the relative energies of the species. The Grimme DFT-D2 correction<sup>68</sup> was added to the relative

energies of the species obtained from UB3LYP for the QM atoms. The MM region was described by the CHARMM22 force field. The HDLC optimizer<sup>69</sup> was employed in the geometry optimization, and the core regions of optimization were within 8 Å of the QM regions (see section 1 in the Supporting Information). All transition states (TSs) were located by relaxed potential energy surface (PES) scans followed by full TS optimizations using the P-RFO optimizer implemented in the HDLC code, and all TSs and minima were confirmed by frequency calculations to have only 1 and 0 imaginary vibrations, respectively.



**Figure 5.** Statistical analyses for the H-bond between Arg254 and the OH group of the active site iron in  ${}^5\text{Int1}$ . (a) Probability distribution  $P(r)$  for H-bond lengths between the H3 atom of the Arg254 residue and the O atom of the OH group of the active site. The H-bond distance is  $r$ . The average H3–O distance and the mean square deviation (MSD) are also shown. (b) Model of  ${}^5\text{Int1}$ .



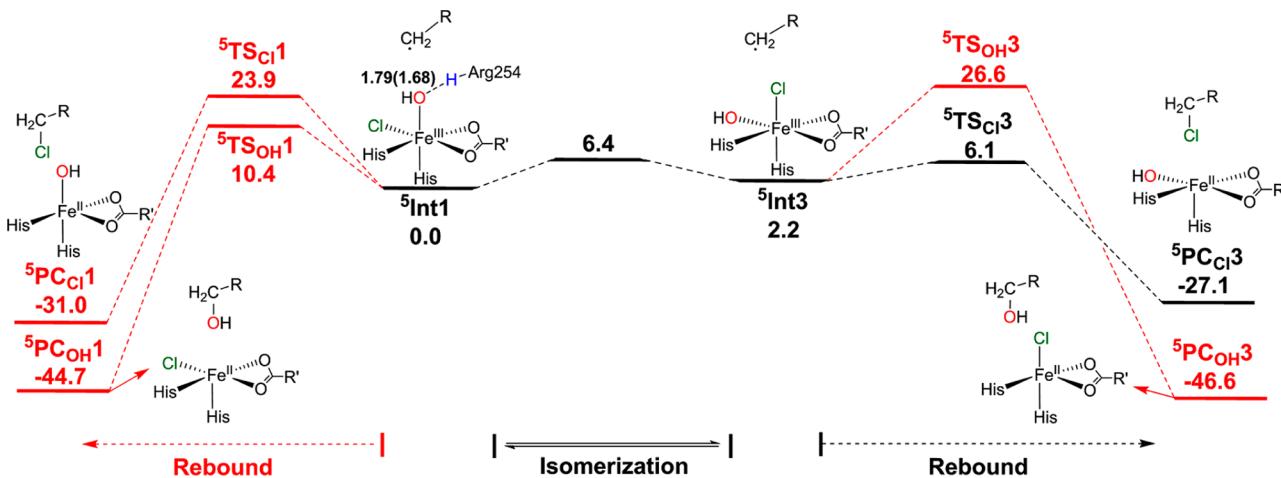
**Figure 6.** (a) QM/MM (UB3LYP/B2) relative energies (in kcal/mol) for the isomerization of the three  $\text{Fe}^{\text{III}}-\text{OH}$  isomers  ${}^5\text{Int1}$ ,  ${}^5\text{Int2}$ , and  ${}^5\text{Int3}$  in the snapshot at 2 ns from the MD trajectory of the ferric hydroxyl radical intermediate with an H-bond between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$ . The ZPE and dispersion corrections are included in the relative energies. The H-bond between the Arg254 residue and the OH ligand of  $\text{Fe}^{\text{III}}-\text{OH}$  is highlighted in  ${}^5\text{Int1}$ . The distance ( $\text{\AA}$ ) outside of the parentheses is the QM/MM optimized distance between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$ , while that in parentheses is taken from the MD trajectory. (b) QM/MM-optimized geometries of  ${}^5\text{Int1}$ ,  ${}^5\text{Int2}$ , and  ${}^5\text{Int3}$ , along with the key structural parameters (in  $\text{\AA}$ ) at UB3LYP/B1.

### 3. RESULTS AND DISCUSSION

**3.1. H-Abstraction.** QM-only calculations<sup>31–34</sup> have previously showed that only the quintet state is relevant to the reactivity of the  $\text{Fe}^{\text{IV}}=\text{O}$  species in nonheme enzymes. To test this conclusion with QM/MM, we investigated the H-abstraction processes for the triplet, quintet, and septet states and reaffirmed

(Figure S2 in the Supporting Information) the quintet state as the lowest in energy among the three spin states for all species. Therefore, all the following results and discussion are based on the calculations done for the quintet state.

The computed electronic energy barrier and free energy barrier are generally close for H-abstraction in P450 enzymes, as Thiel and Senn discovered by sampling over the protein's



**Figure 7.** QM/MM (UB3LYP/B2) relative energies (in kcal/mol) for the rebound reactions of Fe<sup>III</sup>–OH isomers of <sup>5</sup>Int1 and <sup>5</sup>Int3 in the snapshot at 2 ns from the MD trajectory of the ferric hydroxyl radical intermediate, where the H-bond between Arg254 and Fe<sup>III</sup>–OH is well maintained. The ZPE and dispersion corrections are included in the relative energies. The H-bond between the Arg254 residue and the OH ligand of Fe<sup>III</sup>–OH is highlighted in <sup>5</sup>Int1. The distance (in Å) outside of the parentheses is the QM/MM optimized distance between Arg254 and Fe<sup>III</sup>–OH, while that in parentheses is taken from the MD trajectory. The paths not taken are indicated in red.

fluctuations.<sup>70,71</sup> Sampling procedures (e.g., using umbrella sampling or free energy perturbation theory) are highly computationally demanding with ab initio or DFT QM components, and this is why calculations of QM/MM free energy for iron-containing enzymes are still rare. Herein we utilized thermal and entropic corrections just for the QM system to produce a similar result. H-abstraction by <sup>5</sup>RC1 has a barrier of 18.4 kcal/mol with ZPE and the dispersion correction incorporated. The QM entropic correction ( $-T\Delta S$ ) increases the barrier by 1.5 kcal/mol, and the QM thermal contributions to the internal energy ( $\Delta U^{\text{th}}$ ) decreases the barrier by 1.1 kcal/mol. The net effect is a close match between the computed barrier of 18.4 kcal/mol (with ZPE and dispersion correction) and the computed free energy barrier of 18.8 kcal/mol. However, this practice neglects protein fluctuations; therefore, the computed barriers may not be accurate.<sup>70,71</sup> Thus, given the prohibitively high computational cost of QM/MM MD-based free energy calculations, we treated the barriers with ZPE and dispersion corrections as estimates of the free energy barriers in the enzyme.

Mössbauer spectroscopic experiments demonstrated that there are two rapidly interconverting isomers of the Fe<sup>IV</sup>–oxo species, and at least one of these Cl–Fe<sup>IV</sup>–oxo species abstracts hydrogen from the substrate.<sup>11,25</sup> In this study, we successfully located three isomers of the Fe<sup>IV</sup>–oxo species, labeled as <sup>5</sup>RC1, <sup>5</sup>RC2, and <sup>5</sup>RC3, which were previously reported by QM studies.<sup>33,34</sup> As such, we first investigated the isomerization of the three Cl–Fe<sup>IV</sup>–oxo species, and the results are summarized in Figure 3. The calculated Fe<sup>IV</sup>=O and Fe–Cl bond distances in all of the isomers are in the ranges of 1.63–1.64 and 2.35–2.39 Å, respectively (Figure 3b), which are in good agreement with the extended X-ray absorption fine structure (EXAFS) results of 1.66 and 2.31 Å.<sup>25</sup> Other features in Figure 3b (such as the O–Fe–H1 angle of 84° and the O–Fe–H2 angle of 87° in <sup>5</sup>RC3) match those in previous reports.<sup>37</sup>

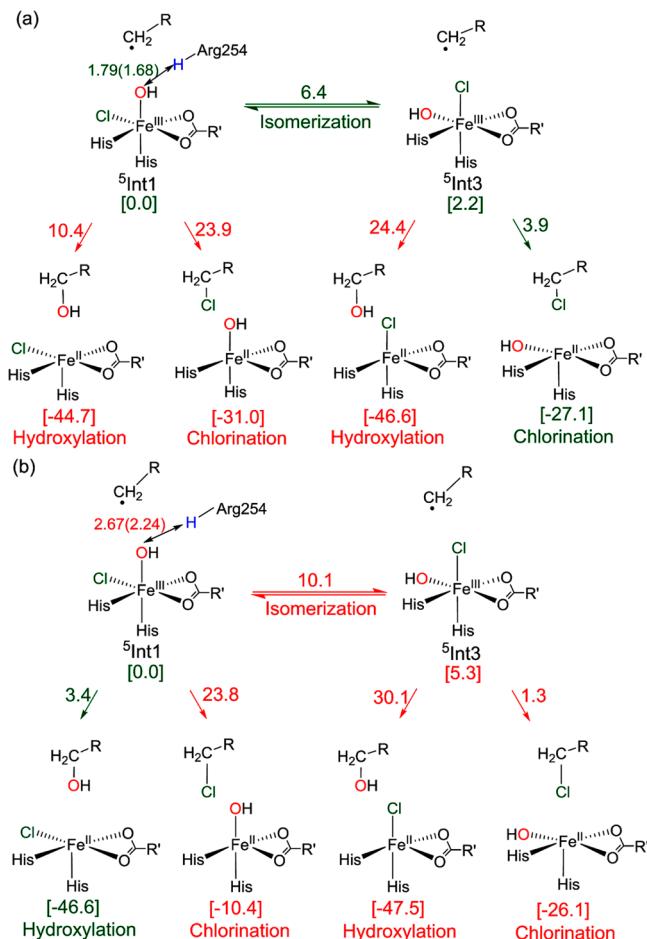
As shown in Figure 3, <sup>5</sup>RC1 can isomerize to <sup>5</sup>RC3 via <sup>5</sup>RC2 by exchanging the positions of the oxo and chloro ligands. In this isomerization process, the oxo ligand first rotates down via <sup>5</sup>TS(1,2) to form the intermediate <sup>5</sup>RC2, and then the chloro ligand rotates up via <sup>5</sup>TS(2,3) to form <sup>5</sup>RC3. The whole process has an overall barrier of 16.0 kcal/mol, in close agreement with

previous QM calculations.<sup>33</sup> However, all attempts to locate another plausible pathway for the isomerization of <sup>5</sup>RC1 to <sup>5</sup>RC3, as suggested by a previous QM study, where the chloro ligand is transferred to the site trans to His116,<sup>33</sup> failed as the optimization automatically led back to <sup>5</sup>RC1. As shown in Figure 3b, the chloro ligand is pointing toward the targeted C–H bond in <sup>5</sup>RC3, while the oxo ligand is 4.06 Å away from the targeted C–H bond.

Subsequently, we turned to investigate the corresponding H-abstraction reactions from <sup>5</sup>RC1 and <sup>5</sup>RC3, in Figure 4. In the QM/MM optimized structure of <sup>5</sup>RC1 (Figure 3b), the O atom of Fe<sup>IV</sup>=O is located 2.87 Å from the H atom of the H–C moiety of the substrate. Thus, starting from <sup>5</sup>RC1 and moving along the left energy profile leads to H-abstraction from the substrate via <sup>5</sup>TS1, yielding the Cl–Fe<sup>III</sup>–OH/carbon radical intermediate, <sup>5</sup>Int1, with a barrier of 18.4 kcal/mol. Moving to the right-hand energy profile (in red), <sup>5</sup>RC1 is converted first to <sup>5</sup>RC3. The subsequent H-abstraction in <sup>5</sup>RC3 requires a barrier of 38.6 kcal/mol, which is practically a forbidden reaction.

We also considered the H-abstraction reaction for the unstable intermediate of <sup>5</sup>RC2; however, all our attempts to abstract the hydrogen atom directly from the substrate by the Cl–Fe<sup>IV</sup>–oxo species led back to <sup>5</sup>RC1 (see Figure S3 in the Supporting Information). We therefore conclude that our QM/MM calculations single out the Cl–Fe<sup>IV</sup>–oxo species <sup>5</sup>RC1, whose oxo ligand is pointing toward the target C–H bond, as the sole reactive species in the C–H activation. The other isomer, <sup>5</sup>RC3, in which the oxo and chloro ligands exchange positions in comparison to <sup>5</sup>RC1, does not participate in C–H activation.

Comparison of our results with those generated by QM-only model calculations shows that the protein plays a major role. Thus, Borowski et al. proposed that <sup>5</sup>RC1 can isomerize into <sup>5</sup>RC3 and the latter species is the more efficient H-abstractor.<sup>33</sup> However, our QM/MM study shows that the barrier for H-abstraction by <sup>5</sup>RC3 is much larger than that for H-abstraction by <sup>5</sup>RC1 in all of the investigated snapshots (see Figure 4 and Tables S1 and S2 in the Supporting Information). Inspection of the structure of QM/MM-optimized <sup>5</sup>TS3 (Figure S4 in the Supporting Information) shows that the protein environment, such as Arg254 and Glu102 around the substrate, exhibits a large



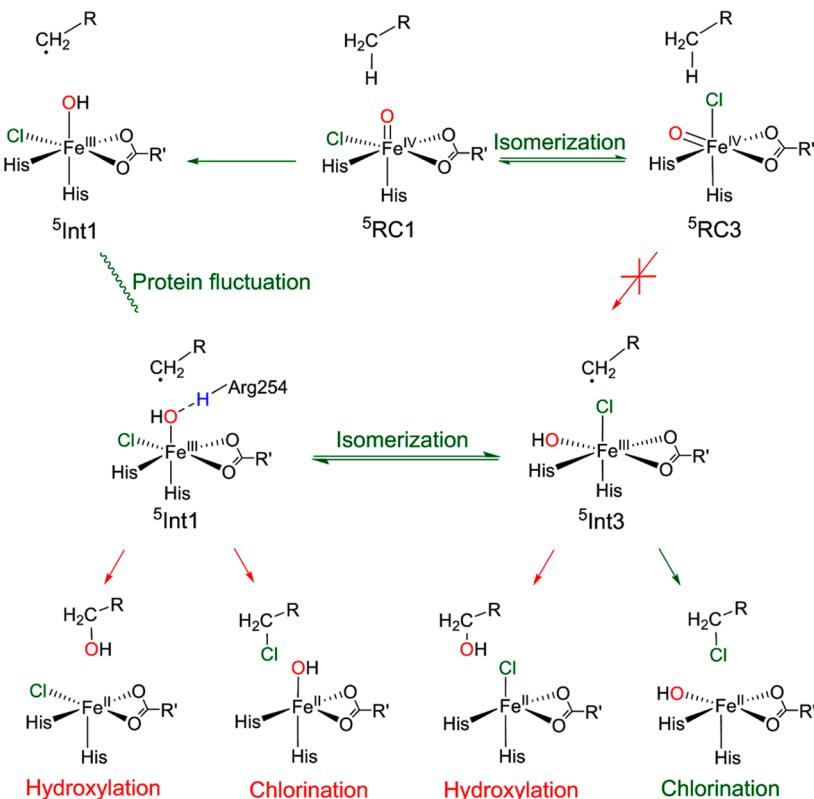
**Figure 8.** Comparison of the reaction network for  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  species  $\text{^5Int1}$  as a function of the H-bonding strength between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$ . The red pathways are those not taken, while the green pathways are allowed: (a) The H-bond between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$  is present in the  $\text{^5Int1}$  species in the snapshot at 2 ns from the MD trajectory of the ferric hydroxyl radical intermediates. (b) The H-bond between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$  is absent in the  $\text{^5Int1}$  species in the snapshot at 1 ns from the MD trajectory of the ferric hydroxyl radical intermediates. The distances (in Å) outside of the parentheses are the QM/MM-optimized distances between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$ , while those in parentheses are taken from the MD trajectories. The barriers near the respective arrows are in kcal/mol, while in brackets we show the energies of the species relative to  $\text{^5Int1}$ . The H-bond interaction between the Arg254 residue and the OH ligand of  $\text{Fe}^{\text{III}}-\text{OH}$  is highlighted in  $\text{^5Int1}$ .

displacement relative to  $\text{^5RC3}$ , which may cause the much higher energy barrier in the QM/MM-calculated H-abstraction by  $\text{^5RC3}$  via  $\text{^5TS3}$ . The conformational dynamics of  $\text{^5RC3}$  leads to a 4.2 kcal/mol difference in the H-abstraction barriers. However, the barrier for H-abstraction by  $\text{^5RC3}$  is still much larger (more than 20 kcal/mol) than that for H-abstraction by  $\text{^5RC1}$  (see Table S2). The QM model study may yield a biased conformation of the substrate and the protein environment in the gas-phase study and thus make  $\text{^5RC3}$  appear to be an efficient H-abstractor, in contrast to the QM/MM findings herein. Wong et al. proposed that  $\text{^5RC2}$ , with its  $\text{Fe}^{\text{IV}}-\text{oxo}$  vector perpendicular to the substrate C–H bond, corresponds to the conformation in which C–H activation occurs, as the position of the substrate radical is favorable for the subsequent Cl rebound.<sup>34</sup> However, our QM/MM calculations do not support this proposal, as  $\text{^5RC2}$

is not stable, and it converts back to  $\text{^5RC1}$  during the H-abstraction. The QM model study by Wong et al.<sup>34</sup> predicted a very short oxo–H distance of 2.75 Å in  $\text{^5RC2}$  (note this study did not consider the protein environment), while the present QM/MM calculations show that in  $\text{^5RC2}$  there is a long distance of 5.46 Å due to the constraints of the protein environment.

**3.2. Rebound Reaction.** In the last section, we demonstrated that the H-abstraction reaction will be predominantly carried out by the  $\text{Cl}-\text{Fe}^{\text{IV}}=\text{O}$  species  $\text{^5RC1}$ , leading to the  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  radical intermediate  $\text{^5Int1}$ . In this section, we proceed to discuss the transformations of  $\text{^5Int1}$ . In order to investigate the effect of surrounding residues on the halogenation/hydroxylation selectivity in  $\text{^5Int1}$ , we performed 20 ns MD simulations on  $\text{^5Int1}$ . Figure 5a shows the probability distribution for the hydrogen bond (H-bond) distance between the H3 atom of the Arg254 residue and the O atom of the OH group in the 20 ns trajectory of  $\text{^5Int1}$ . The average value of the H-bond distance is 1.88 Å, with a mean square deviation of 0.14 Å. As shown in Figure 5a, the H-bond between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$  is maintained in a majority of the frames. The H-bonding interaction between Arg254 and the O atom of  $\text{Fe}^{\text{III}}-\text{OH}$  plays a key role in halogenation/hydroxylation selectivity, as discussed below. As such, we first considered a representative snapshot from the MD trajectory of  $\text{^5Int1}$ , in which the H-bond between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$  is well maintained.

Similar to the case for the  $\text{Fe}^{\text{IV}}=\text{O}$  species, the  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  species also has three isomers,  $\text{^5Int1}$ ,  $\text{^5Int2}$ , and  $\text{^5Int3}$ , which interconvert readily under the reaction conditions as shown in the calculated potential energy surface in Figure 6. The barrier for the isomerization of  $\text{^5Int1}$  into  $\text{^5Int3}$  is 6.4 kcal/mol, which is close to the value of 4.9 kcal/mol found in a previous QM calculation.<sup>33</sup> The previously found alternative pathway where the chloro ligand is on the site trans to His116<sup>33</sup> could not be located in the current QM/MM calculations. Figure 7 shows the competitions of Cl/OH rebound processes from the  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  isomers  $\text{^5Int1}$  and  $\text{^5Int3}$ . Going to the left side on the red profile,  $\text{^5Int1}$  either undergoes Cl-rebound to form the chlorinated product or OH-rebound to form the hydroxylated product. The calculated barriers are 10.4 kcal/mol for hydroxylation and 23.9 kcal/mol for chlorination, respectively. Clearly, chlorination is not favored over hydroxylation in  $\text{^5Int1}$ , as the carbon radical center has a much shorter distance to the OH ligand (3.59 Å), in comparison with its distance to the Cl ligand (5.11 Å) (see Figure 6b). In the middle part of the energy profile in Figure 7,  $\text{^5Int1}$  can isomerize into  $\text{^5Int3}$ , and thereafter,  $\text{^5Int3}$  participates in Cl/OH rebound by going to the right side along the black energy profile. It is apparent that now the chlorination (with a barrier of 3.9 kcal/mol relative to  $\text{^5Int3}$ ) is much more favorable than hydroxylation (with a barrier of 24.4 kcal/mol relative to  $\text{^5Int3}$ ) by  $\text{^5Int3}$ . Clearly,  $\text{^5Int3}$  is the right conformation for the selective chlorination, in which the C1-radical center is much closer to Cl (with a distance of 3.55 Å) than to the OH ligand (5.34 Å). We also considered the OH-rebound reactions of  $\text{^5Int2}$ ; however, as shown in Figure S5 in the Supporting Information, all calculations inevitably led back to  $\text{^5Int1}$ . Meanwhile, halogenation from  $\text{^5Int2}$  was found to lead to the formation of its halogenated product with an energetic barrier of 13.2 kcal/mol, which is significantly higher than the chlorination barrier through  $\text{^5Int3}$ . Further inspection of the red profile in Figure 7 reveals that the OH rebound barrier of 10.4 kcal/mol is quite significant. This is caused by the H-bonding interaction between  $\text{Fe}^{\text{III}}-\text{OH}$  and Arg254. It is this effect that



**Figure 9.** Reaction network of the selective halogenation by SyrB2 starting from the  $\text{Fe}^{\text{IV}}=\text{O}$  species of <sup>5</sup>RC1. The green pathway is the sole pathway for halogenation of the L-Thr substrate, while the red paths are not taken. The H-bonding interaction between the Arg254 residue and the OH ligand of  $\text{Fe}^{\text{III}}-\text{OH}$  is highlighted in <sup>5</sup>Int1.

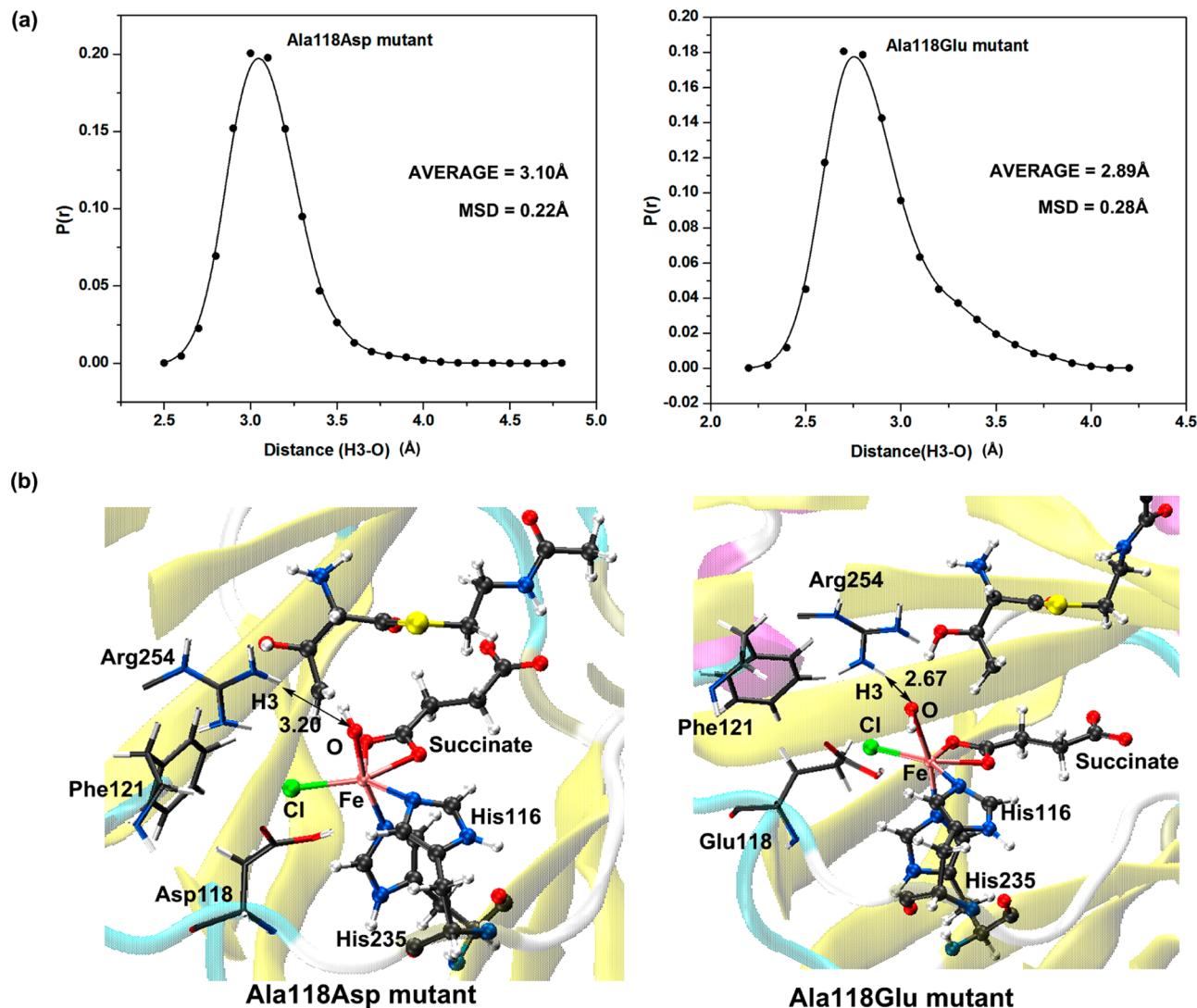
ends up favoring isomerization and then Cl rebound, with an overall barrier of 6.4 kcal/mol, hence resulting in exclusive Cl rebound.

To probe the H-bonding effect further, we also considered a snapshot wherein the H-bond between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$  is practically absent. Figure 8 summarizes and compares the reactivity network of both snapshots with or without H-bonding interactions between Arg254 and  $\text{Fe}^{\text{III}}-\text{OH}$ . If this H-bond is maintained, the OH rebound will experience a higher barrier (10.4 kcal/mol in Figure 8a) relative to the isomerization of <sup>5</sup>Int1 to <sup>5</sup>Int3 (6.4 kcal/mol in Figure 8a), as the H-bond will hinder the rebound of the OH ligand to the carbon radical center and facilitate the isomerization. In <sup>5</sup>Int1 (Figure 6b), the guanidinium group of the Arg254 residue maintains a hydrogen-bonding interaction with the OH group of  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  (the O–H3 distance is 1.79 Å), while the O–H4 distance is 2.78 Å. During the isomerization of  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$ , the O–H4 distance decreases to 1.83 Å (see <sup>5</sup>TS<sub>RI(1,2)</sub> in Figure S6 in the Supporting Information). Clearly, the Arg254 residue facilitates the isomerization by stabilizing the transition state through additional H-bonding interactions. As such, the isomerization from <sup>5</sup>Int1 to <sup>5</sup>Int3 is favored in comparison with the Cl/OH rebound reactions via <sup>5</sup>Int1. Once <sup>5</sup>Int3 is formed, the Cl-rebound reaction will occur readily to yield the chlorinated product (as the green pathway in Figure 8a shows). In contrast, if Arg254 is not H-bonded or weakly H-bonded to the OH group (see the QM/MM optimized H-bond length of 2.67 Å in Figure 8b), the OH group will easily rebound to the carbon radical, as the OH is adjacent to it (see the barrier of 3.4 kcal/mol in Figure 8b). Inspection of the results in Figure 8 reveals the dual roles of H-bonding interactions between the Arg254 residue and the OH

ligand of the  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  in controlling the reactivity and selectivity of the  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  species. This H-bond not only prevents the rebound of OH to form a hydroxylation product but also facilitates the isomerization of the  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  intermediate.

**3.3. Articulating the New Mechanism of Selective Halogenation by SyrB2.** The above QM/MM calculations in Figure 8 suggest a new mechanism of selective halogenation by SyrB2, which is illustrated in Figure 9, where the green arrows depict the most energetically favorable path. Thus, the reaction starts from the active  $\text{Cl}-\text{Fe}^{\text{IV}}-\text{oxo}$  species <sup>5</sup>RC1, in which the oxo group is pointing toward the target C–H bond. H-abstraction by this  $\text{Cl}-\text{Fe}^{\text{IV}}-\text{oxo}$  complex leads to the  $\text{Fe}^{\text{III}}-\text{OH}/\text{carbon-radical}$  intermediate <sup>5</sup>Int1, in which the OH ligand is pointing toward the carbon radical of the substrate. Protein fluctuations around the formed <sup>5</sup>Int1 establish an H-bond between Arg254 and  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$ . This H-bond increases the hydroxylation barrier for <sup>5</sup>Int1 and decreases the isomerization barrier of <sup>5</sup>Int1 to <sup>5</sup>Int3 via exchange of the positions of the Cl and OH ligands. As such, in <sup>5</sup>Int3 the Cl ligand points toward the carbon radical of the substrate and halogenation takes place favorably over hydroxylation. Our QM/MM-derived mechanism shows therefore that selective halogenation is achieved via the isomerization of the  $\text{Cl}-\text{Fe}^{\text{III}}-\text{OH}$  intermediate, rather than by the isomerization of the  $\text{Cl}-\text{Fe}^{\text{IV}}=\text{O}$  active species as previously proposed.<sup>33</sup>

**3.4. Halogenation vs Hydroxylation: Comparison with Experiment.** Mass spectrometric (MS) studies identified a partitioning of 87% halogenation versus 13% hydroxylation for the native L-Thr substrate,<sup>36</sup> indicating the minor OH-rebound reaction still transpires. In the subsequent section, we show that



**Figure 10.** Statistical analyses for the H-bond between Arg254 and the OH group of the active site in <sup>5</sup>Int1 species for the Ala118Asp and Ala118Glu mutants. (a) Probability distribution  $P(r)$  for H-bond lengths between the H3 atom of Arg254 residue and the O atom of OH group of the active site.  $r$  is the distance of the H-bond for the Ala118Asp and Ala118Glu mutants. The average H3–O distance and the mean square deviation (MSD) are also shown. (b) Active structures of <sup>5</sup>Int1 species for Ala118Asp and Ala118Glu mutants taken from their representative snapshots with some key amino acid residues labeled. The distances (in Å) between the H3 atom of Arg254 and the O atom of the active site are taken from the MD trajectories.

Arg254 plays key roles in determining halogenation/hydroxylation selectivities. As discussed above, the H-bond interaction of Arg254 plays dual roles in the selective halogenation by SyrB2 and thus not only prevents the OH-rebound reaction but also facilitates the isomerization of <sup>5</sup>Int1 to <sup>5</sup>Int3 leading to chlorination. The MD simulation in Figure 5 shows that this H-bond is well preserved in the majority of frames along the trajectory (71% and 88% conformers have H3–O distances of less than 1.9 and 2.0 Å, respectively), so that the majority of reactions will lead to the chlorination reactions. In the minority of frames of the trajectory, where the OH is released from the H-bond's grip, hydroxylation will occur (Figure 8b). The present findings are in good accord with experiment.<sup>36</sup> More evidence was found in another αKG-dependent CurA halogenase, whose Arg247 occupies the same site as Arg254 in SyrB2, as its halogenation reactivity was completely suppressed in the Arg247Glu and Arg247Ala mutants.<sup>12</sup>

To further support our conclusions, we performed 20 ns MD simulations on the Ala118Asp and Ala118Glu mutants of the SyrB2 enzyme. It was observed that chlorination does not

transpire in either mutant<sup>20</sup> and the hydroxylation is enhanced in the Ala118Glu mutant.<sup>36</sup> As can be seen from the statistical analyses in Figure 10, unlike the wild-type enzyme, the Ala118Asp and Ala118Glu mutants show no H-bonding interaction between Arg254 and Fe<sup>III</sup>–OH during the simulations. The average H3–O distance is 3.10 (2.89) Å with a mean square deviation of 0.22 (0.28) Å in the Ala118Asp (Ala118Glu) mutant, respectively. Ala118 is a second-shell residue located near the His116 residue and the chlorine atom. As such, the Ala118Asp and Ala118Glu mutations lead to the rearrangement of second-shell residues, and the side chain of Phe121 moves to the position between Arg254 and Fe<sup>III</sup>–OH. In this manner, the steric effect and hydrophobic nature of the benzyl side chain Phe121 causes the breakage of the H-bonding interaction between Arg254 and Fe<sup>III</sup>–OH. According to the mechanism proposed above, these two mutants should not lead to any chlorination products, which is in full accord with experimental findings.<sup>20</sup>



- (40) Senn, H. M.; Thiel, W. In *Atomistic Approaches in Modern Biology*; Reiher, M., Ed.; Springer: Berlin, Heidelberg, New York, 2007; pp 173–290.
- (41) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (42) Lin, H.; Truhlar, D. *Theor. Chem. Acc.* **2007**, *117*, 185–199.
- (43) Warshel, A. *Angew. Chem., Int. Ed.* **2014**, *53*, 10020–10031.
- (44) Godfrey, E.; Porro, C. S.; de Visser, S. P. *J. Phys. Chem. A* **2008**, *112*, 2464–2468.
- (45) Quesne, M. G.; Latifi, R.; Gonzalez-Ovalle, L. E.; Kumar, D.; de Visser, S. P. *Chem. - Eur. J.* **2014**, *20*, 435–446.
- (46) Wang, B.; Usharani, D.; Li, C.; Shaik, S. *J. Am. Chem. Soc.* **2014**, *136*, 13895–13901.
- (47) Du, L.; Gao, J.; Liu, Y.; Zhang, D.; Liu, C. *Org. Biomol. Chem.* **2012**, *10*, 1014–1024.
- (48) Galonić Fujimori, D.; Barr, E. W.; Matthews, M. L.; Koch, G. M.; Yonce, J. R.; Walsh, C. T.; Bollinger, J. M.; Krebs, C.; Riggs-Gelasco, P. J. *J. Am. Chem. Soc.* **2007**, *129*, 13408–13409.
- (49) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- (50) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Genet.* **2005**, *61*, 704–721.
- (51) Brünger, A. T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 148–156.
- (52) MacKerell, A. D.; Bashford, D.; Bellott, Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (53) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (54) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1–28.
- (55) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (56) Smith, W.; Yong, C.; Rodger, P. *Mol. Simul.* **2002**, *28*, 385–471.
- (57) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.
- (58) de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. *J. Phys. Chem. B* **1999**, *103*, 6133–6141.
- (59) Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss.* **1997**, *106*, 79–92.
- (60) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
- (61) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (62) Becke, A. D. *J. Chem. Phys.* **1992**, *97*, 9173–9177.
- (63) Chen, H.; Lai, W.; Shaik, S. *J. Phys. Chem. Lett.* **2010**, *1*, 1533–1540.
- (64) Chen, H.; Song, J.; Lai, W.; Wu, W.; Shaik, S. *J. Chem. Theory Comput.* **2010**, *6*, 940–953.
- (65) Altun, A.; Breidung, J.; Neese, F.; Thiel, W. *J. Chem. Theory Comput.* **2014**, *10*, 3807–3820.
- (66) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299–310.
- (67) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (68) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (69) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177–2186.
- (70) Senn, H. M.; Thiel, S.; Thiel, W. *J. Chem. Theory Comput.* **2005**, *1*, 494–505.
- (71) Senn, H. M.; Kästner, J.; Breidung, J.; Thiel, W. *Can. J. Chem.* **2009**, *87*, 1322–1337.