

Substrate Placement Influences Reactivity in Non-heme Fe(II) Halogenases and Hydroxylases*

Received for publication, September 29, 2012, and in revised form, February 18, 2013. Published, JBC Papers in Press, February 28, 2013, DOI 10.1074/jbc.M112.415570

Heather J. Kulik^{†1} and Catherine L. Drennan^{§2}

From the [†]Department of Chemistry, Stanford University, Stanford, California 94305 and the [§]Department of Chemistry and Biology and the Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Background: SyrB2 is a non-heme Fe(II) halogenase that reacts on tethered amino acid substrates.

Results: Hydroxylation and hydrogen abstraction are less sensitive to substrate positioning at the active site than halogenation.

Conclusion: Observed halogenation of native L-Thr substrate by SyrB2 can be explained by positioning effects.

Significance: Our work could inform how to redesign tether-dependent enzymes toward alternative products.

We employ error-corrected density functional theory methods to map out the dependence of reactivity on substrate position for SyrB2, a member of a family of non-heme iron halogenases and hydroxylases that are only reactive toward amino acid substrates delivered via prosthetic phosphopantetheine arms. For the initial hydrogen abstraction step, the inherent flexibility of the phosphopantetheine molecule weakens the position dependence for both the native substrate (threonine for SyrB2) and alternative substrates. Over a 5 Å window of substrate positions, the tethered hydrogen abstraction step proceeds with nearly identical activation energies and donor-acceptor distances in the transition state. The propensity of a particular substrate toward halogenation or hydroxylation is found to depend strongly on the substrate placement following hydrogen abstraction, with deeper substrate delivery into the active (for native substrates) site favoring halogenation and shallower substrate delivery favoring hydroxylation.

Enzymes efficiently activate relatively unreactive species, such as alkanes, at ambient temperature and in conditions much less harsh than are typically required for analogous industrial processes. A straightforward route to functionalize small molecules, *e.g.*, by halogenating or hydroxylating, is key for applications ranging from therapeutic drug development (1) to biorenewable product functionalization (2). Nonribosomal peptide and polyketide synthases (3, 4), which synthesize clinically important natural products (5, 6), can associate with tailoring enzymes, which act on small molecule substrates that are delivered by a phosphopantetheine (PPant)³ tether bound to a carrier protein (Fig. 1) (4, 7–9). The PPant-dependent halogenase, SyrB2 (10, 11), chlorinates the methyl group of tethered L-Thr as part of the biosynthetic pathway for syringomycin E,

an antifungal antibiotic that acts by forming channels in lipid bilayers (10, 12). Although chlorination of synthesized drugs often increases their efficacy (1), chlorination at unactivated alkanes is very difficult for synthetic chemists (5, 12, 13). SyrB2 efficiently halogenates the aliphatic carbon of the small L-Thr substrate but only when the substrate is delivered via PPant tether on the upstream enzyme SyrB1, and the substrate-tether binding affinity controls specificity toward L-Thr (10). It is crucial to understand the role that PPant plays in SyrB2 activity to determine whether the enzyme can be engineered to react on small, untethered substrates.

SyrB2 is a mononuclear, non-heme Fe(II)-, α -ketoglutarate-, chlorine-, and O₂-dependent enzyme that shares similarities in structure and mechanism to more well studied hydroxylases (10, 11, 14). This class of enzymes activates O₂ at a non-heme Fe(II) center to cleave an α -ketoglutarate molecule to CO₂ and succinate (15–18). The subsequently formed high spin, Fe(IV)=O intermediate (19, 20) abstracts a hydrogen from the substrate. Theoretical studies of TauD, a non-heme Fe(II) hydroxylase, showed that 5-fold coordination stabilizes the high spin Fe(IV)=O intermediate, and strong favorability of the forming O–H bond enhances enzyme activity (21). A number of model complexes have been developed and studied experimentally to further elucidate the role that a high spin Fe(IV)=O intermediate plays in hydrogen abstraction (22–26). Unlike the enzyme, most model complexes have triplet ground states with a low lying high spin quintet state (26). Nevertheless, these model studies have shown that controlled spin inversion will likely give rise to preference for the quintet surface in hydrogen abstraction (25).

Following hydrogen abstraction, the mechanisms of halogenases and hydroxylases diverge (Scheme 1): in halogenases, the radical reacts with an equatorial chlorine, whereas in hydroxylases there is a rebound abstraction of the recently formed axial hydroxyl. Comparison of the structure of SyrB2 to homologous hydroxylases first suggested (11) that the substitution of a single residue proximal to the active site iron (Ala in halogenases and Asp in hydroxylases) was crucial for explaining the difference in function between these homologous enzymes. The smaller Ala in halogenases provides room for chlorine to bind to iron, whereas Asp would prevent binding of chlorine to the iron center. However, mutation of Ala \rightarrow Asp in SyrB2 (11) and Asp \rightarrow

* This work was supported in part by computational resources from the Lawrence Livermore National Laboratory (to H. J. K.).

¹ Recipient of a Career Award at the Scientific Interface from the Burroughs Wellcome Fund. To whom correspondence should be addressed: Dept. of Chemistry, Stanford University, 333 Campus Dr., Mudd Bldg., Rm. 121, Stanford, CA, 94305. E-mail: hkulik@stanford.edu.

² Howard Hughes Medical Institute Investigator.

³ The abbreviations used are: PPant, phosphopantetheine; DFT, density functional theory; DFT+U, density functional theory augmented with a Hubbard U correction; L-Aba, L-ababutyric acid; L-Nva, L-norvaline.

Placement and Reactivity in Non-heme Fe(II) Enzymes

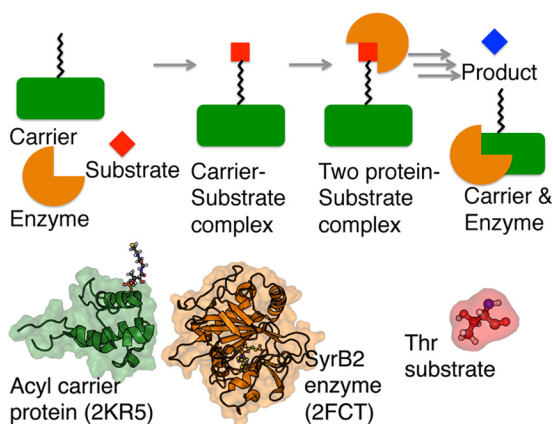
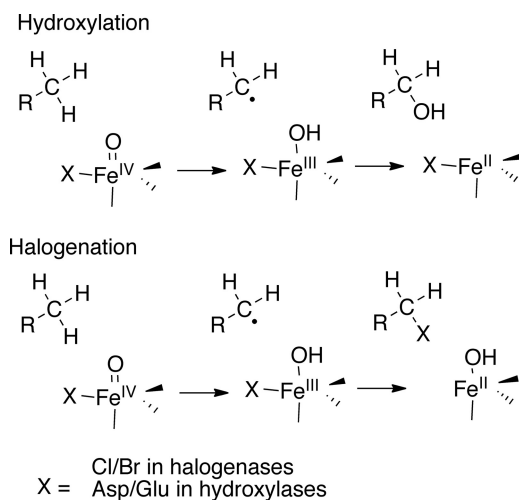


FIGURE 1. **Depiction of tethered substrate delivery common in biosynthetic pathways.** A cartoon of key binding events (top) is compared with structures of the model acyl carrier protein, SyrB2 enzyme, and Thr substrate used in this study (bottom).



SCHEME 1. **Proposed mechanisms for hydroxylation and halogenation in non-heme Fe(II), α -ketoglutarate-dependent enzymes.**

Ala in prototypical hydroxylase TauD (27) failed to convert enzyme activities. Structural studies of the halogenase CytC3 (28), which is homologous to SyrB2, provide insight into why this mutation fails to modify the enzymatic activity. These studies highlighted the fact that numerous additional interactions with the protein environment are critical in ensuring chlorine can bind to the iron center (28). Thus, altering a hydroxylase to catalyze halogenation would require formation of a chlorine-binding site. Previous density functional theory studies augmented with a Hubbard U term (DFT+U) of models of SyrB2 and related hydroxylases (29) showed no barrier to halogenation if chlorine was bound, suggesting again that the lack of chlorination in mutant hydroxylases is due to poor uptake and binding of chlorine in the active site. In an effort to explain the inability of A118D SyrB2 mutants to hydroxylate, the previous DFT+U study identified that substrate hydrogen bonding effects and binding energy differences were responsible (29). Biochemical experiments later led to similar observations (30).

Biochemical studies have further shown that SyrB2 can hydroxylate the non-native substrates (structures shown in Fig. 2) L-ababutyric acid (L-Aba) and L-norvaline (L-Nva) and that substrate identity and placement strongly control product type

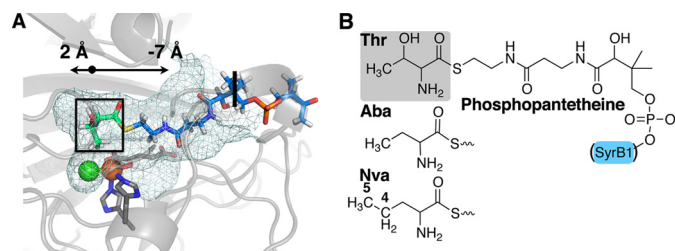


FIGURE 2. **Model for *ab initio* simulations from PPant/SyrB2 docking and two-dimensional representation of structures.** A, stick representation of full docked model (PPant truncation indicated as a vertical black line) with Thr substrate in green and outlined. A cavity in the protein structure is shown as cyan wireframe mesh. The active site contains iron (rust sphere), chlorine (green sphere), α -ketoglutarate/succinate (with gray carbons), and two histidine ligands (with gray carbons) to iron. The range of displacements in the simulations is depicted by a double-headed arrow with respect to the origin in the figure. B, depiction of PPant chemical structure covalently bonded to Thr (highlighted in gray) alongside the structures of Aba and Nva. Connection to SyrB1 is indicated in parentheses for reference, but the SyrB1 structure is not shown in A.

and reaction rates (30). For this reason, the interactions of a tethered substrate in the active site would appear to take on a principal role in determining the function of these halogenases and hydroxylases. Nevertheless, structural models of the enzyme and tethered substrate are unavailable because of challenges associated with co-crystallizing SyrB2 with the carrier protein SyrB1 when either substrates or analogs are loaded on PPant via thioester linkage. The established accordance between DFT+U theory and experimental biochemistry motivates further study into identifying substrate-dependent and position-dependent reactivity in larger models of both SyrB2 and the PPant tether. Importantly, DFT+U preserves the favorable scaling of standard DFT approaches used in previous SyrB2 studies (31, 32) while providing even more accurate descriptions of transition-metal chemistry without any of the model size limitations of more expensive approaches (33). These first principle DFT+U simulations of SyrB2 with PPant tethers will elucidate tether active site interactions and identify the role of tethered substrate delivery in determining the type of product formed in this class of alkane-functionalizing enzymes.

Here, we present mechanistic results on a model of the interactions of the SyrB2 active site with the SyrB1-associated PPant arm. We identify the position dependence of reactivity for the hydrogen abstraction, chlorination, and alternative hydroxylation steps. We identify the most probable substrate positioning in the active site by comparing the results of our position-dependent reaction coordinate to known results from experimental biochemistry. We also analyze the structure and flexibility of phosphopantetheine during the reaction to understand the role that arm flexibility plays in reactivity.

THEORY

PPant structures selected from a NMR ensemble of an acyl carrier protein (Protein Data Bank code 2KR5) (7) were docked inside a model of SyrB2 (Protein Data Bank code 2FCT) (11) using the Glide software package, and the complex was solvated with water (34). The pK_a values of ionizable side chains for the protein were calculated in the environment of each surrounding residue and assigned for a pH of 7.4 (35–37). The starting PPant structure was obtained from solution NMR results

because they provided more information about structural heterogeneity in the PPant structure, but we note that comparable x-ray structures of acyl carrier proteins are available. A roughly linear PPant structure was used, and rotation about PPant bonds was sampled during the docking optimization procedure. We do not assume that free phosphopantetheine would successfully deliver the substrate to the core active site but rather aim to model what would happen upon substrate delivery to the active site via the full acyl carrier protein. Following the suggestion of Blasiak *et al.* (11), we modified the anticipated channel of entry for PPant by rotating a blocking phenylalanine residue (Phe-196) to an alternate rotamer. This rotation permitted the docking of PPant inside SyrB2 for delivery of the substrate to the active site. Other channels of delivery were considered by searching for cavities in the holoenzyme with the 3V (38) and CASTp (39) packages, but no alternative cavities that were of sufficient size to both accommodate PPant and reach the active site were found. It is possible that allosteric effects cause significant rearrangements that open new channels to the active site, but further investigation is outside the scope of this work.

The five lowest energy structures from our docking results were evaluated by single point energy and forces from DFT inside an extended 8 Å radius (541 atoms) model of the active site and surrounding SyrB2 protein. We confirmed that incorporation of remote residues does not greatly perturb direct ligand orientation by relaxing the reactant complex in spheres of remote residues of increasing size from 4 to 10 Å (in 2 Å increments) and aligning the relaxed structures (data not shown). As an additional test, we calculated the hydrogen abstraction minimum energy pathway for L-Thr with second sphere residues surrounding both PPant and the active site in a 676-atom model and found the calculated barrier height to be within 1 kcal/mol of the minimal model.

After aligning docked PPant structures, we found these structures to be relatively consistent, where deviations between structures were localized toward the phosphate end of the arm. We therefore truncated the PPant model by removing the terminal $\text{C}(\text{CH}_3)_2\text{CH}_2\text{PO}_4$, which does not substantially affect the electronic structure of the tether close to the substrate (data not shown). The energetically most favorable structure, as characterized by DFT, was used as a starting point for more extensive DFT and DFT+U calculations.

We studied the reactivity of a SyrB2 model complex with L-Thr, L-Aba, and L-Nva substrates that were bound via thioester linkage to a truncated PPant model (*e.g.*, L-Thr-S-PPant), following a recent comparison of the substrate-dependent product formation rates (30). In each case, we started from the results of the docked L-Thr-S-PPant structure in SyrB2, and we modified the substrate when studying L-Aba and L-Nva. For each substrate, we carried out three key steps: substrate radical formation, halogenation, and hydroxylation (Scheme 1). We began our study at a midpoint in the full reaction using a SyrB2 model complex obtained after the initial oxygen activation step with an Fe(IV)=O intermediate coordinated to succinate with a quintet spin multiplicity (40). Triplet and septet reaction coordinates were also calculated for varying PPant position, but dependence of spin energetics on arm positioning was found to

be weak and not investigated further. The stabilization of the quintet surface is also consistent with large scale computational studies on AlkB, a DNA repair dioxygenase (41).

The PPant position-dependent studies are carried out by fixing the end of the truncated PPant that is distant from the substrate by constraining a single carbon atom in the arm to one position in space, whereas the rest of the arm is free to move. Additionally, the iron center of the SyrB2 model complex is fixed in space to prevent additional relative displacements between the SyrB2 model and PPant. The reaction steps are driven forward through constrained displacements along each reaction coordinate, and a minimum energy path is obtained with a nudged elastic band method using climbing image and variable springs extensions (42) to ensure population of images at the barrier. Transition states are only approximately identified as the highest energy structure here. Therefore, the geometries and energetics are only reasonable approximations of the true transition state geometry and barrier height energy (43).

The starting positions of the PPant arm are incremented by ~ 1 Å at a time over a 9 Å window along the channel observed in the protein structure (Fig. 2). Because docked structures were qualitatively observed to position the substrate past the oxo-intermediate and well into the calculated cavity (Fig. 2), the majority of structures considered were displaced further out of the active site, whereas only one set of calculations was considered through further displacement into the active site. We note that the range of distances considered likely exceeds that sampled *in vivo*, but this enables us to identify the subset of positions that is most probable under physiological conditions. We do not model in the full protein channel around the PPant. Channel lining residues could hinder the flexibility of the PPant arm slightly through hydrogen bonding interactions that are not incorporated in our simulations. However, as we will later show, much of the key flexibility in PPant is observed in the portion closest to the substrate and thus unlikely to be affected by channel-lining residue interactions.

DFT calculations were carried out using the Quantum-ESPRESSO package (44). An ultrasoft plane wave pseudopotential approach was used with a cutoff of 30 Rydbergs for the wave function and 300 Rydbergs for the density. The pseudopotential for iron included both semi-core (3s, 3p) and valence (3d, 4s) states in the valence. Following previous results (29), DFT+U is employed to obtain improved descriptions of electronic and geometric structure. We employ the Perdew-Burke-Ernzerhof generalized gradient approximation for the exchange correlation functional (45). We apply a U of 5.4 eV, as obtained from averaging linear response values on key intermediates. We note that the Hubbard U value is slightly reduced from the value of 5.6 eV published in previous work (29) because of slight differences in the intermediates considered, but variations of Hubbard U under 1 eV are not likely to have a significant effect on the quantitative or qualitative descriptions of reaction coordinates.

RESULTS

We have studied the position-dependent reactivity of a SyrB2 model on native substrate and unnatural amino acid substrates tethered to a truncated form of phosphopantetheine. To com-

TABLE 1

Arm position dependence of key measurements for approximate transition states in the hydrogen abstraction step for L-Thr

	Relative arm position										
	2.0	0.0	-0.5	-1.0	-1.5	-2.0	-3.0	-4.0	-5.0	-6.0	-7.0
C-H (Å)	1.19	1.18	1.17	1.17	1.17	1.18	1.18	1.19	1.20	1.21	1.21
O-H (Å)	1.39	1.39	1.40	1.39	1.39	1.39	1.39	1.39	1.39	1.39	1.39
C-O (Å)	2.58	2.56	2.56	2.55	2.56	2.57	2.57	2.58	2.58	2.59	2.59
C-H-O (°)	174	171	170	172	174	175	178	175	172	174	170

putationally probe placement-dependent reactivity, we fixed the remote end of our truncated PPant and substrate assembly to various incremental positions and studied the energetic and structural properties of each reaction step (hydrogen abstraction, halogenation, and hydroxylation) under this constraint starting from the initial position obtained with docking (Fig. 2). Properties, such as the flexibility of phosphopantetheine, are straightforwardly probed by identifying the rearrangement that occurs in the tether as the reaction proceeds and the difference in tether length when PPant is positioned near or far to the SyrB2 active site. We focus on the relative barrier heights of reaction steps for differing substrates and compare the results with experimental measurements, enabling an approximate mapping of the substrate placement for the full *in vivo* protein-protein complex. We first compare the native L-Thr substrate to L-Aba because they differ only by a single hydroxyl and then compare effects of adding bulk with L-Nva.

During hydrogen abstraction, we observe the key distances in the hydrogen transfer around the saddle point to be very similar, regardless of initial PPant complex positioning (Table 1). Other atoms in the complex, including the L-Thr hydroxyl side chain, show some variation in the transition state (Fig. 3). The correspondence of donor-acceptor atom distances in the transition states suggests that arms that are displaced from sub-ideal positioning must undergo rearrangement to achieve optimal donor-acceptor distances for hydrogen transfer, in accordance with experimental observations (46, 47) for hydrogen transfer.

The flexibility of the PPant complex facilitates the formation of nearly identical transferring distances regardless of PPant positioning, but for extrema of the positions considered, we observe an increase in barrier heights. By calculating only the energy and forces of the isolated PPant-substrate complex from each of these transition states, we observe that this contribution comes from an incurred energetic penalty associated with stretching the PPant arm too far (Fig. 4). In total, there is a roughly 5 Å arm displacement window over which L-Thr hydrogen abstraction barrier heights are essentially equivalent (Fig. 4). These activation energies are consistent with our previous findings for reactions between model SyrB2 complexes and free L-Thr (29).

We now compare the hydrogen abstraction results for L-Thr with those for L-Aba, which differs from L-Thr only by the absence of a hydroxyl group. The arm displacement simulations repeated with L-Aba yield roughly the same barrier heights. Although L-Aba has been observed to react 13 times faster than the native L-Thr substrate (30), our simulations suggest that if the two substrates can be positioned identically, the reactivity should be indistinguishable. Instead, we anticipate that residues not included in our model distinguish L-Thr and

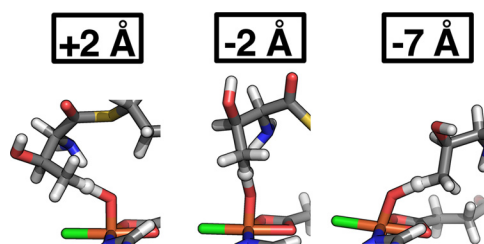


FIGURE 3. Comparison of transition states for L-Thr hydrogen abstraction. Selected approximate transition states with the PPant displacements 2 Å, -2 Å, and -7 Å. The leaving C-H and forming H-O bonds are of comparable length between different transition states.

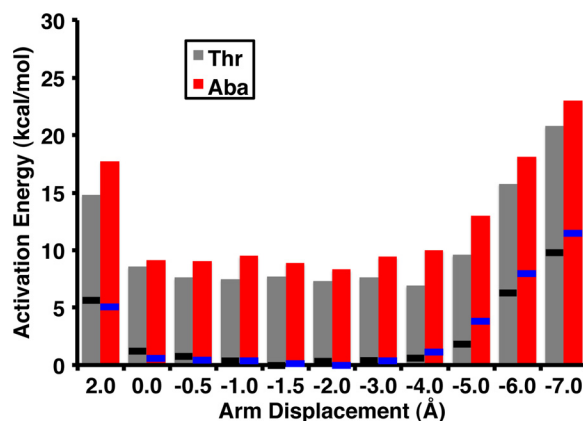


FIGURE 4. Activation energy for hydrogen abstraction from tethered L-Thr (gray) and L-Aba (red) over a range of PPant model arm displacements with relative energy contribution from strain in PPant arm indicated as dashes (black for L-Thr and blue for L-Aba).

L-Aba substrate positioning, thereby informing the relative reactivity of the two substrates. If L-Thr is positioned over chlorine (denoted between +2 and 0 Å in our simulations) and L-Aba is positioned over iron (approximately -0.5 Å), our simulations predict a lower barrier height and faster kinetics for L-Aba. This prediction is consistent both with experimental observations of a 13-fold higher rate for L-Aba over L-Thr and with the previous hypothesis that positioning L-Thr toward chlorine lowers the rate of hydrogen abstraction (30).

One aspect of phosphopantetheine that makes it an effective delivery mechanism for small substrates to enzyme active sites is its flexibility (48, 49). Following the assessment of hydrogen abstraction barriers, we considered metrics of the flexibility of the arm in more detail from our simulation. As we displace the model arm with respect to the active site, minimum energy path structures must include elongation of PPant for the substrate to reach the SyrB2 model complex. Alternatively, if the arm is unable to elongate sufficiently, a minimum energy pathway, if found, would involve the uncatalyzed release of H⁺ from the substrate. The energetically costly, uncatalyzed hydrogen abstraction (known to be ~100 kcal/mol

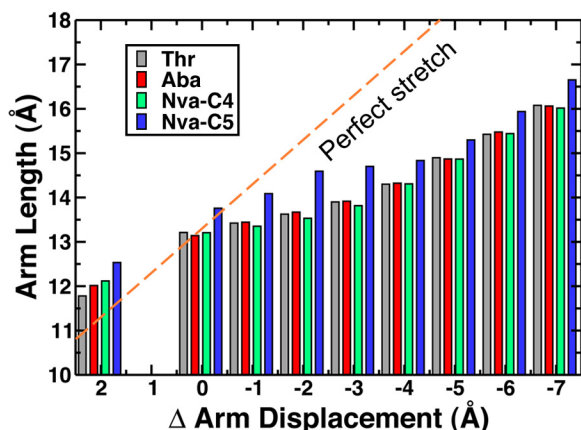


FIGURE 5. Total length of truncated PPant model and linked substrate (in Å) in the hydrogen abstraction transition state over a range of arm displacements for the L-Thr, L-Aba, and L-Nva (C4 or C5) substrates. The length of the PPant model from simulations is compared with a linear stretch with an orange dashed line.

(29, 50)) was not observed over the displacement range considered (a total of 9 Å).

By examining the length from end to end of PPant based on the distance separating the terminal carbon of PPant and the reacting carbon of the substrate, we can approximate the length of the PPant complex and compare against the starting position of PPant in each set of reactions. We observe (Fig. 5) that PPant arms do not stretch the full rigid “perfect” stretch associated with taking the length of the arm at the initial position from docking and then adding or subtracting the distance change of the arm to the active site from alternative starting positions. However, the total arm length from the compressive starting position at 2 Å to the extended starting position at −7 is nearly 5 Å. This total extension is nearly 50% of the length of the unextended PPant model arm and substrate complex. We therefore are interested in understanding the bond lengths and bond angles that contribute most strongly to elongation.

A detailed, first principles analysis of the different contributions to PPant flexibility and elongation can uncover avenues for design modifications to PPant for biomimetic catalysis. In addition to previous experimental work in mapping out properties of acyl carrier protein through site-directed mutagenesis (51, 52), computationally redesigned and synthesized PPant-like probes could provide new insight into these multienzyme complexes. Perhaps unsurprisingly, the greatest degree of flexibility is observed at or around the sulfur linker between PPant and the substrate molecule. The C–S–C angle is observed to increase by up to 9°, and the C–S bonds stretch by nearly 0.04 Å (Fig. 6). In isolation, the bond elongation of 0.04 Å over 1.84 Å for the average S–C bond adds ~0.06 Å to the C–C distance, which we use here as a metric for total PPant complex length. The angle changes in isolation on the other hand of ~9° over a starting value of 96° adds nearly 0.2 Å to the total length of the PPant complex. The soft, angular degrees of freedom in the PPant arm are thus most responsible for the significant elongation observed in accommodation of transition states when the arm and substrate complex is displaced away from the active site.

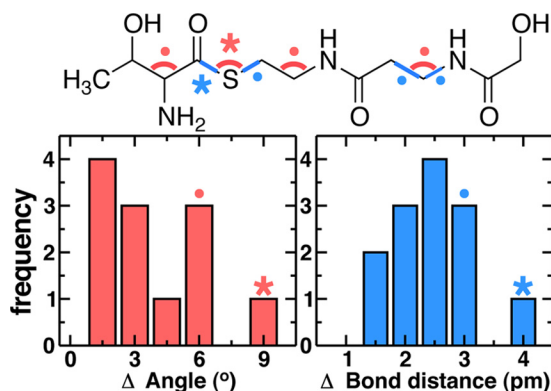


FIGURE 6. Identification of most flexible bonds and angles in truncated L-Thr-S-PPant. Top panel, overlay of structure with key bonds and angles highlighted (most flexible bond or angle indicated by an asterisk, and second through fourth most flexible indicated by filled circles). Bottom panel, histogram of changes in bond angles (left) and bond distances (right) in the model system under stretching conditions.

SyrB2 has been shown to react on L-Nva substrates as well as L-Aba and L-Thr (30). In the case of L-Nva, there is an additional methylene group, and therefore there are two sites from which hydrogen can be easily abstracted: the C4 or fourth carbon or C5 or fifth, terminal carbon. For free (*i.e.*, untethered) L-Nva, the calculated hydrogen abstraction at C4 is less energetically costly by ~2 kcal/mol with respect to C5 (Table 2), in agreement with experimental observations about radical stability (50). Nevertheless, recent biochemical studies have observed that SyrB2 reacts on L-Nva at C5 at a faster rate than at C4, suggesting that substrate positioning plays a significant role in these reaction rates (30). In our calculation of hydrogen abstraction from C4 or C5 in tethered L-Nva (Fig. 7), we observe the relative barrier heights to depend on arm positioning. For substrates positioned further into the active site, hydrogen abstraction at C4 is more favorable than at C5. When substrates are displaced away from the iron-oxo intermediate, the C5 carbon instead is more accessible and requires less rearrangement of PPant and of the substrate for the hydrogen to be abstracted. In both cases, hydrogen abstraction energies are relatively flat over a large, 5 Å window of arm displacements, consistent with the previously observed stretching of PPant to accommodate low energy hydrogen abstraction for L-Thr. When our results are taken in combination with experimental observations that L-Nva hydrogen abstraction is faster at C5, it leads to the conclusion that the larger L-Nva substrate is likely prevented from fully entering the same active site pocket occupied by the native L-Thr substrate. In cases where C5 is deuterated to encourage reactions at C4 in experiments, we would expect that hydrogen abstraction would be limited by the ability of PPant to position C4 of the L-Nva substrate further into the active site and proximal to the Fe(IV)=O species.

For all of the substrates considered, hydrogen abstraction is found to be equally possible over a large range of arm displacements. However, given the inequivalence of an equatorial chlorine and an axial hydroxyl (*e.g.*, differences in their calculated binding energies (29)), we would expect to observe greater position dependence in the partitioning of subsequent halogenation or hydroxylation steps. In fact, as shown in Fig. 8, we observe that halogenation is favorable for a relatively small window of

TABLE 2

The activation energies (E_a) and exothermicity (ΔE_r) in kcal/mol of key reaction steps (hydrogen abstraction or H_{abs} , chlorination or +Cl, and hydroxylation or +OH) of free amino acids (Thr, Aba, and Nva) at SyrB2

Step	Thr		Aba		Nva-C5		Nva-C4	
	E_a	ΔE_r	E_a	ΔE_r	E_a	ΔE_r	E_a	ΔE_r
	kcal/mol	kcal/mol	kcal/mol	kcal/mol	kcal/mol	kcal/mol	kcal/mol	kcal/mol
H_{abs}	7	-4	8	-4	9	-3	7	-5
+Cl	0	-20	0	-17	0	-19	0	-15
+OH	4	-29	1	-25	1	-26	3	-30

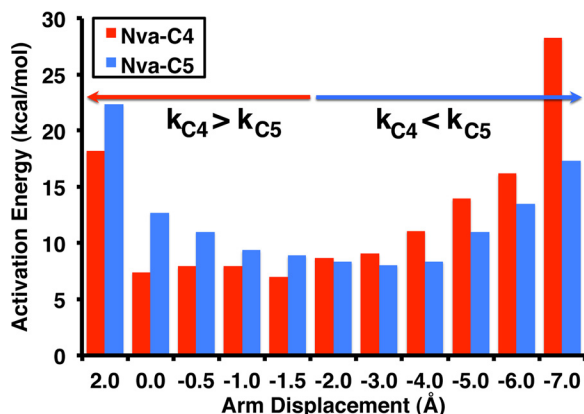


FIGURE 7. Activation energy for hydrogen abstraction from tethered L-Nva at the fourth carbon (C4, in red) or fifth and terminal carbon (C5, in blue) over a range of PPant model arm displacements. The arrows indicate regions in which the rate of hydrogen abstraction is faster at C4 or C5.

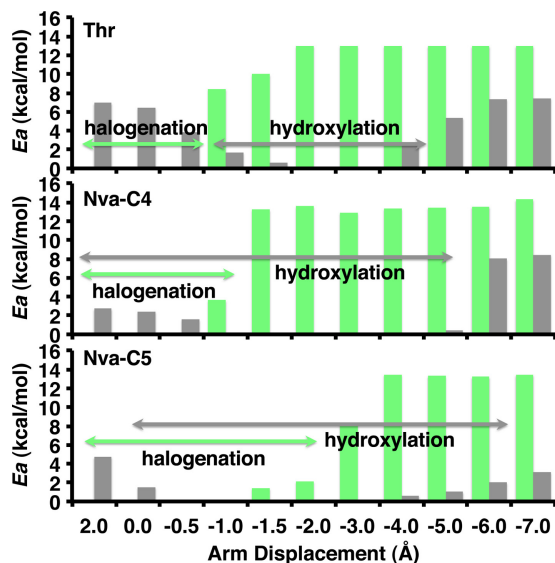


FIGURE 8. Activation energy for competing halogenation (green) and hydroxylation (gray) steps on L-Nva (at C4 and C5) and L-Thr over a range of PPant model arm displacements. The arrows indicate regions in which halogenation or hydroxylation is feasible because of a low energetic barrier. Overlapping arrows indicate that either reaction step could occur and would be in competition with one another.

L-Thr-S-PPant arm positions corresponding to the substrate being displaced well into the active site. Beyond this narrow window, the substrate and PPant assembly cannot stretch far enough to access the equatorial chlorine. The resulting barrier heights are obtained from uncoupling the cleavage of the iron–chlorine bond (approximately costing 13 kcal/mol (29)) and formation of a new L-Thr–chlorine bond.

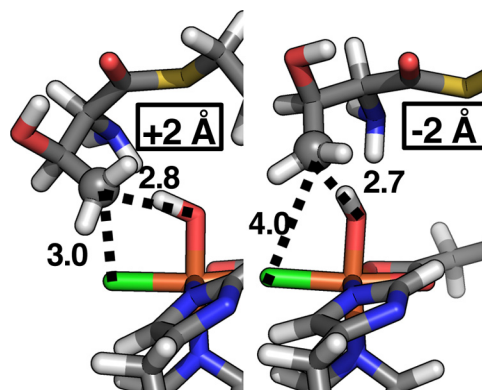


FIGURE 9. Depiction of L-Thr radical carbon distances to either chlorine or hydroxyl groups in +2 Å (left panel) and -2 Å (right panel) starting positions.

Hydroxylation, on the other hand, is most favorable when the substrate is not fully delivered into the central active site. We can rationalize the difference in activation energies by observing an L-Thr radical intermediate tethered in the +2 Å displacement position and comparing it with that obtained for a -2 Å displacement (Fig. 9). In the +2 Å case, the L-Thr radical is equivalently positioned between chlorine and OH, whereas for shallower arm displacements, the L-Thr radical easily accesses the hydroxyl but becomes more and more distant from the chlorine. Biochemical experiments have identified a partitioning of 87% halogenation *versus* 13% hydroxylation for L-Thr and 31% halogenation *versus* 69% hydroxylation for L-Aba (30). These observations are consistent with the hypothesis that the hydroxyl side group on L-Thr positions the natural substrate more fully into the active site (positions equivalent to 0 to +2 Å). A slight displacement from the L-Thr positioning of approximately -0.5 Å to -1 Å for L-Aba because of the absence of this hydrogen bond would be consistent with the experimentally observed partitioning ratios. The extent to which this differential binding is due to interactions between the hydroxyl on L-Thr and the second sphere residues of the SyrB2 active site will be investigated in future work.

We briefly considered an alternate form of the reactive complex, which has been shown to form through a rearrangement involving a high energy transition state (33), with an equatorial Fe(IV)=O and axial chlorine to identify whether reactivity differences could be a result of changes in the position of the Fe(IV)=O. By studying the reactivity of L-Thr with this alternate complex, we found hydrogen abstraction to be less favorable and hydroxylation to be preferred. These results combined with a high barrier for interconversion to this intermediate suggest that it is not a likely explanation for reactivity patterns in SyrB2, but further study with accelerated quantum chemical methods (53) to observe whether rearrangement occurs dynamically may be useful.

Although the importance of substrate positioning to the partitioning between hydroxylation and halogenation has previously been reported (30), now we reveal a structural basis for this partitioning that is consistent with the most likely channel of entry for PPant into SyrB2. The positioning of the hydroxyl for tethered L-Thr is likely different from the previously reported observations of free L-Thr. Positioning of L-Thr

toward chlorine to enhance the preference for halogenation like that observed for free L-Thr (29) is still a viable mechanism (Fig. 9). However, if substrate placement is shallow (rather than deep) in the active site, hydroxylation becomes competitive with or even overtakes halogenation.

Because substrate placement is key for explaining why L-Thr substrate is preferentially halogenated, the source of differences in product branching ratios for radicals formed at the C4 and C5 sites of L-Nva is a compelling question for computation to answer. Experimentally, L-Nva reacts 130 times faster than L-Thr and preferentially at the C5 site (30). The rate of hydrogen abstraction of L-Nva at C4 is comparable to that of L-Aba (30). In our simulations, we observed that L-Nva hydrogen abstraction at C5 is most favorable when L-Nva is positioned away from chlorine and closer to the entrance of the SyrB2 active site, corresponding to shallower tethered-substrate placement (approximately -2 Å). In these cases, C4 becomes inaccessible. Experimentally, hydrogen abstraction at C4 may be observed by slowing down C5 abstraction through deuteration (30).

We consider halogenation and hydroxylation as a function of all arm displacement positions over the full 9 Å range. The C4–C5 carbon bond is ~ 1.5 Å, meaning that the terminal carbon can extend effectively up to 1.5 Å further into the active site than C4. The C4 is positionally equivalent to the radical formation sites in L-Thr and L-Aba, whereas C5 is more chemically equivalent. For halogenation (Fig. 8), we observe that chlorination of L-Nva at C4 roughly follows the same pattern as L-Thr. A narrow window corresponding to the C4 of L-Nva being positioned over chlorine yields barrierless chlorination *versus* small barriers for hydroxylation. For the C5 of L-Nva, we observe nearly barrierless hydroxylation over a wide range of substrate placements. Halogenation becomes competitive if the substrate is inserted more deeply into the active site, but shallow insertion is more consistent for L-Nva C5 based on observations for hydrogen abstraction. In this shallow regime (approximately -2 Å), L-Nva C5 would be preferentially hydroxylated, consistent with experimental partitioning ratios (93% hydroxylation *versus* 7% halogenation) (30). In the case of C5, halogenation becomes feasible for a broader range of arm positions, roughly equivalent to the length added by the extra C–C bond. However, hydrogen abstraction patterns suggest that poor fit of the substrate in the active site makes it difficult for the substrate to access positions at which halogenation is favorable. The extra C–C bond also results in a greater range of motion for the substrate overall with respect to L-Thr or L-Aba.

In light of the observed reactivity patterns for hydrogen abstraction on L-Nva, a consistent prediction can be made from computational results with respect to product formation. If a hydrogen is abstracted from C4 of L-Nva, as is the case when the C5 site is blocked via deuteration, it is likely because the substrate is displaced further into the active site. At substrate positions approximately 0.0 to -0.5 Å, the hydrogen abstraction barrier for C4 is lowered and becomes equivalent to that of C5 when L-Nva is positioned at approximately -2 Å. For a deeper substrate positioning required to facilitate C4 hydrogen abstraction, halogenation and hydroxylation become competitive in our simulations. These results are consistent with exper-

imental observations that 65% of C5-deuterated L-Nva are halogenated at C4 (30). These computational results suggest that modification of PPant or SyrB1 to shorten the length of the linker would prevent L-Thr from fully entering the active site and would result in a greater degree of hydroxylated product.

DISCUSSION

We have presented a model system for studying the position dependence of substrate reactivity in the active site of the non-heme Fe(II) halogenase SyrB2. Because SyrB2 acts only on the native substrate when delivered on a long, flexible PPant arm, we have considered in detail the specific role that PPant plays in reactivity. For hydrogen abstraction, we observe that donor-acceptor distances were largely independent of substrate placement. The associated barrier heights are dependent upon the strain induced in the phosphopantetheine-substrate complex. For a reacting L-Thr substrate, hydrogen abstraction barrier heights were determined to be flat over roughly a 5 Å window. This range of low energy abstractions had a similar width for L-Aba. When the end points where hydrogen abstraction barrier heights began to rise were probed, we observed that the energy increase comes from increased strain in the phosphopantetheine arm. The PPant arm elongated by ~ 5 Å or $\sim 50\%$ over the full 9 Å range of arm positions considered in our simulations. The majority of elongation came from soft, angular degrees of freedom in PPant, suggesting that in the full protein-protein complex, there should be variability in the positioning of the substrate. Chemical modification of PPant to reduce flexibility or alter length is of future interest—such modifications could be used as probes of the enzyme active site or to test hypotheses about directing the partitioning of product formation.

For the two sites of hydrogen abstraction in the unnatural amino acid L-Nva, we observe differences in reactivity depending upon substrate positioning. If the substrate is positioned further away from the reactive SyrB2 active site complex, it is more favorable to abstract hydrogen from C5. If the substrate is instead proximal, especially to the chlorine side of the model complex, then hydrogen abstraction from C4 is preferred. Based upon experimental radical stability trends (50), C4 should be the more stable radical. However, previous experimental results (30) indicated that the reaction proceeded faster at the C5 site of L-Nva. Taken together, these computational and experimental observations suggest that a poor fit of the L-Nva substrate into the SyrB2 active site means that the substrate is positioned in a way that favors hydrogen abstraction at L-Nva C5. Positioning of the substrate during hydrogen abstraction influences product formation. When substrate radicals are inserted further into the active site toward chlorine, these reactive intermediates undergo barrierless chlorination. Analysis of the geometry of tethered substrates in this positioning regime suggests that it is difficult for non-native substrates to achieve placement directly over chlorine.

Hydroxylation, on the other hand, is favorable with low energy barriers ~ 2 –4 kcal/mol over a larger range of substrate positions. The side chain hydroxyl group present on L-Thr and absent from L-Aba and L-Nva therefore likely positions the substrate past the reactive intermediate and over chlorine. This

preferential positioning for L-Thr should occur via interactions with the second sphere residues in the active site. In the case of L-Nva, the most probable substrate position following hydrogen abstraction influences the partitioning of product formation. That is, both the C4 and C5 site favor chlorination when positioned over the chlorine ligand, whereas they both have lower barriers for hydroxylation when positioned further out of the active site. However, C4 hydrogen abstraction is most probable when the substrate is positioned toward chlorine, whereas C5 is most probable when the substrate is further out of the active site. Therefore, we anticipate a hydroxylated C5 product if the L-Nva residue does not fit well in the active site, whereas we anticipate primarily halogenated C4 product if the L-Nva residue is positioned further into the active site. This computational observation is consistent with experimental measurements that showed deuterating the C5 position results in preferentially halogenating at C4 (30). By preventing hydrogen abstraction at C5, hydrogen abstraction may occur at C4, but only when C4 is proximal to the Fe(IV)=O species (and also proximal to chlorine).

This analysis of our computed position-dependent reactivity provides a consistent explanation of previously observed experimental reactivity patterns (30). In particular, we are able to understand why L-Nva reacts more rapidly at C5, despite the fact that formation of a radical at C4 is more energetically favorable. We also now can identify the geometric basis for the initially counterintuitive preference for chlorination of L-Nva at C4 and hydroxylation of L-Nva at C5. This computational study has relied on the accuracy (29) of Hubbard-augmented density functional theory for describing transition metal chemistry while maintaining affordable scaling that permits us to study systems several hundreds of atoms in size. It is the combined accuracy and efficiency of this approach that allows us to obtain a position energy mapping to elucidate the source of reactivity patterns in phosphopantetheine-dependent enzymes. Future work will focus on recently developed accelerated quantum chemistry methods (53) to study even larger models of a SyrB1-SyrB2 complex to observe dynamic rearrangement of the active site and flexibility of the substrate/tether in real time. Importantly, our mapping scheme coordinates with experimental biochemistry to predict physiologically relevant positioning effects where experimental structural data are unavailable. The development of this approach now paves the way to redesigning tethers as reactivity probes, to developing biomimetics, and to understanding reactivity in other enzymes that rely on phosphopantetheine for substrate delivery.

Acknowledgments—We thank Sarah E. J. Bowman, Michael A. Funk, Kateryna Kozyrytska, and Adam H. Steeves for helpful suggestions.

REFERENCES

- Gribble, G. W. (2004) Natural organohalogens. A new frontier for medicinal agents? *J. Chem. Educ.* **81**, 1441
- Lowe, J. R., Martello, M. T., Tolman, W. B., and Hillmyer, M. A. (2011) Functional biorenewable polyesters from carvone-derived lactones. *Polym. Chem.* **2**, 702–708
- Tran, L., Broadhurst, R. W., Tosin, M., Cavalli, A., and Weissman, K. J. (2010) Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases. *Chem. Biol.* **17**, 705–716
- Leibundgut, M., Jenni, S., Frick, C., and Ban, N. (2007) Structural basis for substrate delivery by acyl carrier protein in the yeast fatty acid synthase. *Science* **316**, 288–290
- Neumann, C. S., Fujimori, D. G., and Walsh, C. T. (2008) Halogenation strategies in natural product biosynthesis. *Chem. Biol.* **15**, 99–109
- Wagner, C., El Omari, M., and König, G. M. (2009) Biohalogenation. Nature's way to synthesize halogenated metabolites. *J. Nat. Prod.* **72**, 540–553
- Wattana-amorn, P., Williams, C., Ploskoń, E., Cox, R. J., Simpson, T. J., Crosby, J., and Crump, M. P. (2010) Solution structure of an acyl carrier protein domain from a fungal type I polyketide synthase. *Biochemistry* **49**, 2186–2193
- Liu, Y., Zheng, T., and Bruner, S. D. (2011) Structural basis for phosphopantetheinyl carrier domain interactions in the terminal module of non-ribosomal peptide synthetases. *Chem. Biol.* **18**, 1482–1488
- Chan, D. I., and Vogel, H. J. (2010) Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem. J.* **430**, 1–19
- Vaillancourt, F. H., Yin, J., and Walsh, C. T. (2005) SyrB2 in syringomycin E biosynthesis is a nonheme Fe-II α -ketoglutarate- and O-2-dependent halogenase. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10111–10116
- Blasiak, L. C., Vaillancourt, F. H., Walsh, C. T., and Drennan, C. L. (2006) Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440**, 368–371
- Grgurina, I., Barca, A., Cervigni, S., Gallo, M., Scaloni, A., and Pucci, P. (1994) Relevance of chlorine-substituent for the antifungal activity of syringomycin and syringotoxin, metabolites of the phytopathogenic bacterium *Pseudomonas-syringae* pv. *Syringae*. *Experientia* **50**, 130–133
- Vaillancourt, F. H., Yeh, E., Vosburg, D. A., Garneau-Tsodikova, S., and Walsh, C. T. (2006) Nature's inventory of halogenation catalysts. Oxidative strategies predominate. *Chem. Rev.* **106**, 3364–3378
- Yi, C., Jia, G., Hou, G., Dai, Q., Zhang, W., Zheng, G., Jian, X., Yang, C. G., Cui, Q., and He, C. (2010) Iron-catalysed oxidation intermediates captured in a DNA repair dioxygenase. *Nature* **468**, 330–333
- Solomon, E. I., Brunold, T. C., Davis, M. I., Kemsley, J. N., Lee, S. K., Lehnert, N., Neese, F., Skulan, A. J., Yang, Y. S., and Zhou, J. (2000) Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* **100**, 235–350
- Costas, M., Mehn, M. P., Jensen, M. P., and Que, L. (2004) Dioxygen activation at mononuclear nonheme iron active sites. Enzymes, models, and intermediates. *Chem. Rev.* **104**, 939–986
- Hausinger, R. P. (2004) Fe(II)/ α -ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol.* **39**, 21–68
- Krebs, C., Fujimori, D. G., Walsh, C. T., and Bollinger, J. M., Jr. (2007) Non-heme Fe(IV)-oxo intermediates. *Acc. Chem. Res.* **40**, 484–492
- Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr., and Krebs, C. (2003) The first direct characterization of a high-valent iron intermediate in the reaction of an α -ketoglutarate-dependent dioxygenase. A high-spin Fe(IV) complex in taurine/ α -ketoglutarate dioxygenase (TauD) from *Escherichia coli*. *Biochemistry* **42**, 7497–7508
- Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M., Jr., and Krebs, C. (2006) Direct spectroscopic detection of a C–H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14738–14743
- de Visser, S. P. (2006) Propene activation by the oxo-iron active species of taurine/ α -ketoglutarate dioxygenase (TauD) enzyme. How does the catalysis compare to heme-enzymes? *J. Am. Chem. Soc.* **128**, 9813–9824
- Cho, K.-B., Chen, H., Janardanan, D., de Visser, S. P., Shaik, S., and Nam, W. (2012) Nonheme iron-oxo and -superoxo reactivities. O₂ binding and spin inversion probability matter. *Chem. Commun. (Camb.)* **48**, 2189–2191
- Janardanan, D., Wang, Y., Schyman, P., Que, L., Jr., and Shaik, S. (2010) The fundamental role of exchange-enhanced reactivity in C–H activation by S=2 oxo iron(IV) complexes. *Angew. Chem. Int. Ed. Engl.* **49**, 3342–3345
- Klinker, E. J., Shaik, S., Hirao, H., and Que, L. (2009) A two-state reactivity model explains unusual kinetic isotope effect patterns in C–H bond cleavage by nonheme oxoiron(IV) complexes. *Angew. Chem. Int. Ed. Engl.* **48**,

- 1291–1295
25. Seo, M. S., Kim, N. H., Cho, K.-B., So, J. E., Park, S. K., Clemaney, M., Garcia-Serres, R., Latour, J.-M., Shaik, S., and Nam, W. (2011) A mononuclear nonheme iron(IV)-oxo complex which is more reactive than cytochrome P450 model compound I. *Chem. Sci.* **2**, 1039–1045
26. Shaik, S. (2010) Biomimetic chemistry. Iron opens up to high activity. *Nat. Chem.* **2**, 347–349
27. Grzyska, P. K., Müller, T. A., Campbell, M. G., and Hausinger, R. P. (2007) Metal ligand substitution and evidence for quinone formation in taurine/ α -ketoglutarate dioxygenase. *J. Inorg. Biochem.* **101**, 797–808
28. Wong, C., Fujimori, D. G., Walsh, C. T., and Drennan, C. L. (2009) Structural analysis of an open active site conformation of nonheme iron halogenase CytC3. *J. Am. Chem. Soc.* **131**, 4872–4879
29. Kulik, H. J., Blasiak, L. C., Marzari, N., and Drennan, C. L. (2009) First-principles study of non-heme Fe(II) halogenase SyrB2 reactivity. *J. Am. Chem. Soc.* **131**, 14426–14433
30. Matthews, M. L., Neumann, C. S., Miles, L. A., Grove, T. L., Booker, S. J., Krebs, C., Walsh, C. T., and Bollinger, J. M. (2009) Substrate positioning controls the partition between halogenation and hydroxylation in the aliphatic halogenase, SyrB2. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17723–17728
31. Pandian, S., Vincent, M. A., Hillier, I. H., and Burton, N. A. (2009) Why does the enzyme SyrB2 chlorinate, but does not hydroxylate, saturated hydrocarbons? A density functional theory (DFT) study. *Dalton Trans.* 6201–6207
32. Comba, P., and Wunderlich, S. (2010) Iron-catalyzed halogenation of alkanes. Modeling of nonheme halogenases by experiment and DFT calculations. *Chem. Eur. J.* **16**, 7293–7299
33. Borowski, T., Noack, H., Radon, M., Zych, K., and Siegbahn, P. E. (2010) Mechanism of selective halogenation by SyrB2. A computational study. *J. Am. Chem. Soc.* **132**, 12887–12898
34. Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., and Banks, J. L. (2004) Glide. A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **47**, 1750–1759
35. Gordon, J. C., Myers, J. B., Foltz, T., Shojia, V., Heath, L. S., and Onufriev, A. (2005) H+++. A server for estimating pK_a s and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **33**, W368–W371
36. Myers, J., Grothaus, G., Narayanan, S., and Onufriev, A. (2006) A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins* **63**, 928–938
37. Olsson, M. H., Søndergaard, C. R., Rostkowski, M., and Jensen, J. H. (2011) Propka3. Consistent treatment of internal and surface residues in empirical pK(a) predictions. *J. Chem. Theory Comput.* **7**, 525–537
38. Voss, N. R., and Gerstein, M. (2010) 3V. Cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res.* **38**, W555–W562
39. Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp. Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **34**, W116–W118
40. Galonic, D. P., Barr, E. W., Walsh, C. T., Bollinger, J. M., Jr., and Krebs, C. (2007) Two interconverting Fe(IV) intermediates in aliphatic chlorination by the halogenase CytC3. *Nat. Chem. Biol.* **3**, 113–116
41. Liu, H., Llano, J., and Gauld, J. W. (2009) A DFT study of nucleobase dealkylation by the DNA repair enzyme AlkB. *J. Phys. Chem. B* **113**, 4887–4898
42. Henkelman, G., Uberuaga, B. P., and Jonsson, H. (2000) A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901
43. Sheppard, D., Terrell, R., and Henkelman, G. (2008) Optimization methods for finding minimum energy paths. *J. Chem. Phys.* **128**, 134106
44. Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G. L., Cococcioni, M., Dabo, I., Dal Corso, A., de Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougousis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A. P., Smogunov, A., Umari, P., and Wentzcovitch, R. M. Quantum ESPRESSO. A modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**:395502, 2009
45. Perdew, J. P., Burke, K., and Ernzerhof, M. (1996) Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868
46. Benkovic, S. J., and Hammes-Schiffer, S. (2003) A perspective on enzyme catalysis. *Science* **301**, 1196–1202
47. Nagel, Z. D., and Klinman, J. P. (2006) Tunneling and dynamics in enzymatic hydride transfer. *Chem. Rev.* **106**, 3095–3118
48. Mootz, H. D., Finking, R., and Marahiel, M. A. (2001) 4-Phosphopantetheine transfer in primary and secondary metabolism of *Bacillus subtilis*. *J. Biol. Chem.* **276**, 37289–37298
49. Joshi, A. K., Witkowski, A., Berman, H. A., Zhang, L., and Smith, S. (2005) Effect of modification of the length and flexibility of the acyl carrier protein-thioesterase interdomain linker on functionality of the animal fatty acid synthase. *Biochemistry* **44**, 4100–4107
50. McMillen, D. F., and Golden, D. M. (1982) Hydrocarbon bond-dissociation energies. *Annu. Rev. Phys. Chem.* **33**, 493–532
51. Lai, J. R., Fischbach, M. A., Liu, D. R., and Walsh, C. T. (2006) A protein interaction surface in nonribosomal peptide synthesis mapped by combinatorial mutagenesis and selection. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5314–5319
52. Flaman, A. S., Chen, J. M., Van Iderstine, S. C., and Byers, D. M. (2001) Site-directed mutagenesis of acyl carrier protein (ACP) reveals amino acid residues involved in ACP structure and acyl-ACP synthetase activity. *J. Biol. Chem.* **276**, 35934–35939
53. Kulik, H. J., Luehr, N., Ufimtsev, I. S., and Martinez, T. J. (2012) *Ab initio* quantum chemistry for protein structures. *J. Phys. Chem. B* **116**, 12501–12509