J|A|C|S ARTICLES

Published on Web 03/12/2009

Structural Analysis of an Open Active Site Conformation of Nonheme Iron Halogenase CytC3

Cintyu Wong,[†] Danica Galonić Fujimori,[II,⊥] Christopher T. Walsh,[II] and Catherine L. Drennan*[,†,‡,§]

Departments of Chemistry, Biology and the Howard Hughes Medical Institute, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, Massachusetts 02139 and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave., Boston, Massachusetts 02115

Received December 13, 2008

Abstract: CytC3, a member of the recently discovered class of nonheme Fe(II) and α-ketoglutarate (αKG)-dependent halogenases, catalyzes the double chlorination of L-2-aminobutyric acid (Aba) to produce a known *Streptomyces* antibiotic, γ,γ-dichloroaminobutyrate. Unlike the majority of the Fe(II)-αKG-dependent enzymes that catalyze hydroxylation reactions, halogenases catalyze a transfer of halides. To examine the important enzymatic features that discriminate between chlorination and hydroxylation, the crystal structures of CytC3 both with and without αKG/Fe(II) have been solved to 2.2 Å resolution. These structures capture CytC3 in an open active site conformation, in which no chloride is bound to iron. Comparison of the open conformation of CytC3 with the closed conformation of another nonheme iron halogenase, SyrB2, suggests two important criteria for creating an enzyme-bound Fe—Cl catalyst: (1) the presence of a hydrogen-bonding network between the chloride and surrounding residues, and (2) the presence of a hydrophobic pocket in which the chloride resides.

Introduction

Many halogenated natural products have important therapeutic properties, such as anticancer and antifungal activities.[1] Examples of chlorinated natural products include salinisporamide A, a potent anticancer agent,[2] and syringomycin, an antifungal compound.[3] The incorporation of chlorine into natural products can be important for tailoring a compound's activity. For example, the addition of a single chlorine atom to the syringomycin scaffold increases its antifungal properties 4-fold.[4] The biosynthesis of natural products with cyclopropane ring structures can involve so-called "cryptic halogenations", where a halide is used to functionalize a nonreactive group and is subsequently displaced by a substrate-derived enolate to generate the three member ring. Since chlorination by chemical synthesis can be environmentally unfriendly and technically challenging,[5] there is interest in enzymatic halogenation chemistry for the production of various natural products.

The biosynthesis of γ,γ-dichloroaminobutyrate in soil *Streptomyces* is carried out by a four component nonribosomal peptide synthetase (NRPS) assembly, CytC1-CytC4.[6] CytC1 is a free-standing adenylation domain that loads the nonproteinogenic amino acid L-2-aminobutyric acid (Aba) onto the phosphopantetheine arm of a thiolation domain, CytC2 (Figure 1). While the amino acid is tethered to CytC2, CytC3 then catalyzes tandem chlorinations of the γ-methyl of Aba. Lastly, CytC4 is the thioesterase that hydrolytically releases chlorinated aminobutyrate from the phosphopantetheine arm of CytC2.

CytC3 is part of a newly discovered class of bacterial halogenase enzymes that use α-ketoglutarate (αKG) and nonheme Fe(II) for catalysis.[6] The only nonheme iron halogenase that has been structurally characterized thus far is the syringomycin biosynthetic enzyme II (SyrB2), which is involved in the biosynthesis of the antifungal agent syringomycin.[7−9] CytC3 is highly similar to SyrB2, with 58% identity and 71% homology.[10] All αKG/Fe(II)-dependent enzymes share a common fold, referred to as a cupin fold, comprising antiparallel β-strands in a jelly roll motif.[11] The iron cofactor is located in the center of the jelly roll and is typically coordinated by a

[†] Department of Chemistry, Massachusetts Institute of Technology.
[‡] Department of Biology, Massachusetts Institute of Technology.
[§] Howard Hughes Medical Institute, Massachusetts Institute of Technology.
[II] Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School.
[⊥] Current address: Departments of Cellular and Molecular Pharmacology and Pharmaceutical Chemistry, University of California San Francisco, 600 16th St. MC 2280, San Francisco, CA 94158-2280.

(1) Walsh, C. *Nature* **2000**, *406*, 775–781.
(2) Fenical, W.; Jensen, P. R. *Nat. Chem. Biol.* **2006**, *2*, 666–73.
(3) Bender, C. L.; Alarcon-Chaidez, F.; Gross, D. C. *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 266–292.
(4) Grgurina, I.; Barca, A.; Cervigni, S.; Gallo, M.; Scaloni, A.; Pucci, P. *Experientia* **1994**, *50*, 130–133.
(5) Schnarr, N. A.; Khosla, C. *Nature* **2005**, *436*, 1094–1095.
(6) Ueki, M.; Galonic, D. P.; Vaillancourt, F. H.; Garneau-Tsodikova, S.; Yeh, E.; Vosburg, D. A.; Schroeder, F. C.; Osada, H.; Walsh, C. T. *Chem. Biol.* **2006**, *13*, 1183–1191.
(7) Vaillancourt, F. H.; Yin, J.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10111–10116.
(8) Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. *Nature* **2006**, *440*, 368–71.
(9) Vaillancourt, F. H.; Vosburg, D. A.; Walsh, C. T. *Chembiochem* **2006**, *7*, 748–752.
(10) Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T. J.; Higgins, D. G.; Thompson, J. D. *Nucleic Acids Res.* **2003**, *31*, 3497–3500.
(11) Dunwell, J. M.; Purvis, A.; Khuri, S. *Phytochemistry* **2004**, *65*, 7–17.

10.1021/ja8097355 CCC: $40.75

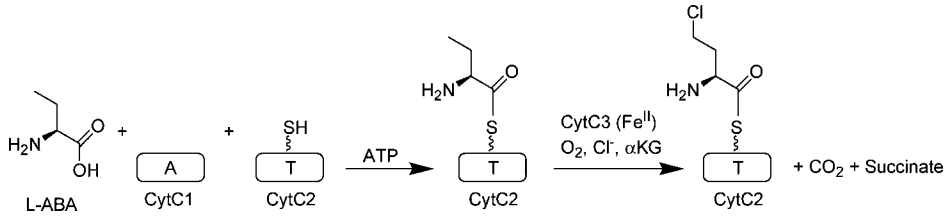

Figure 1. Production of 4-Cl-L-Aba by CytC1−3. CytC1 loads L-Aba-AMP onto the phosphopantetheine arm (wavy line) on CytC2. CytC3 chlorinates the tethered substrate to form 4-Cl-L-Aba. CytC1 is an adenylation domain (A) and CytC2 is a thiolation domain (T).

2-His-1-carboxylate (from Asp or Glu) triad.[12] Reactions catalyzed by this enzyme superfamily include epimerization, cyclization, epoxidation, and hydroxylation among others, where the hydroxylation reaction is the most common.[13] A major difference between hydroxylases and halogenases is that the hydroxylases coordinate the iron with the typical 2-His-1-carboxylate motif, whereas the halogenases coordinate the iron with only two histidines. Alanine is in place of the aspartate/glutamate of the 2-His-1-carboxylate motif in the halogenases, the small size of the alanine substitution and the noncoordinating nature of the side chain allow room for chloride to bind directly to iron in the position that is typically occupied by the side chain carboxylate.[8] In addition to the protein ligands, the remaining coordination sites of iron are taken by chloride, αKG, and molecular oxygen.[13] Upon decarboxylation of αKG, an Fe(IV)-oxo species is generated, which is proposed to be the active intermediate responsible for hydrogen abstraction in both hydroxylases and halogenases.[14−17] The step after hydrogen abstraction is where the two mechanisms diverge; in hydroxylases, it is hypothesized that Fe(III)-OH undergoes a hydroxyl radical rebound to form the hydroxylated product.[14] In halogenases, the substrate radical is postulated to abstract a chlorine atom to form the chlorinated product.[17] It is as yet unclear how the proteins selectively promote hydroxyl or chloride transfer to the substrate.

Efforts have been made to convert the halogenase to hydroxylase and vice versa. Substitution of alanine with aspartate or glutamate in halogenase SyrB2, and substitution of aspartate for alanine in hydroxylase TauD abolishes all activity without the gain of new function in both cases.[8,18] However, it is unclear if there is chloride binding to the active site iron in the hydroxylase mutant. From a protein engineering perspective, it is important to understand the features responsible for chloride binding in these halogenases. However, with only one structurally characterized nonheme iron halogenase, SyrB2, it is difficult to pinpoint the essential attributes for chloride binding. In this study, the structure of a second halogenase, CytC3, was solved showing an open conformation of the active site. The comparison of the open conformation of CytC3 with the closed conformation of SyrB2 provides some insights into the features important for chloride binding.

Materials and Methods

Overexpression, Purification and in Vitro Reconstitution of CytC3. The halogenase CytC3 was overproduced and purified as described[6] with the exception that the iron-reconstitution step was performed after the gel filtration. Following reconstitution, the protein was desalted on a Bio-Gel P6 DG (Bio-Rad, Hercules, CA) column equilibrated in 20 mM Hepes, pH 7.5, 80 mM NaCl. Fractions containing proteins were pooled, concentrated in Amicon stirred ultrafiltration cells (Millipore, Billerica, MA) under N_2 pressure, flash-frozen in liquid N_2 in 25 μL aliquots, and stored at −80 °C.

Crystallization, Data Collection and Processing. CytC3 crystals were obtained using the hanging drop method by mixing 1 μL of protein (38 mg/mL CytC3 in 20 mM HEPES, pH 7.5) with 1 μL of 2.8 M sodium acetate trihydrate, pH 7.0 precipitant solution (Hampton Research, Aliso Viejo, CA), and placing the drop over 0.5 mL of precipitant solution at room temperature. αKG (Sigma, St. Louis, MO) was added to the protein stock solution to a final concentration of 5 mM prior to crystallization setup. Square bipyramidal crystals appeared after 2−3 weeks. For crystals containing iron, the crystals grown aerobically without iron were brought into an anaerobic glovebox (Coy Laboratory, Grass Lake, MI) and soaked in degassed mother liquor with $Fe(NH_4)_2(SO_4)_2$ (10 mM), αKG (5 mM), and chloride (80 mM). Prior to cryo-cooling in liquid nitrogen, crystals were soaked in mother liquor with 15% trehalose overnight. $Fe(NH_4)_2(SO_4)_2$, αKG, and chloride were also present in the mother liquor with the cryo-protectant when crystals were soaked overnight. Data were collected to 2.2 Å resolution for both the apo and iron structures at Stanford Synchrotron Radiation Laboratory on beamline 11−1. Data were processed and scaled using Denzo and Scalepack (Table 1).[19] To locate the iron sites, a data set was collected at the iron peak wavelength using crystals prepared as described above. An iron anomalous data set was collected to 2.75 Å resolution at the Argonne Advanced Photon Source on beamline NE-CAT-24-ID-C.

Structure Determination and Refinement. The structure of CytC3 was solved by molecular replacement using PHASER[20] with the apo CytC3 data set to 3 Å resolution. The structure of SyrB2 (2FCU), with 58% identity and 71% homology to CytC3, without any ligands or waters was used as a search model.[8] The best rotational and translational solution has a correlation coefficient of 16 with two CytC3 molecules per asymmetric unit. The resulting electron density map was solvent flattened in SHARP.[21] Side chains were added to the model at 3 Å resolution, followed by iterative rounds of model building and refinement using XtalView and CNS, respectively.[22,23] Phases were extended to 2.2 Å resolution, waters

(12) Hegg, E. L.; Que, L., Jr. *Eur. J. Biochem.* **1997**, *250*, 625–629.
(13) Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J. *J. Inorg. Biochem.* **2006**, *100*, 644–669.
(14) Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M.; Krebs, C. *Biochemistry* **2004**, *43*, 1134–1134.
(15) Price, J. C.; Barr, E. W.; Glass, T. E.; Krebs, C.; Bollinger, J. M. *J. Am. Chem. Soc.* **2003**, *125*, 13008–13009.
(16) Bollinger, J. M.; Price, J. C.; Barr, E. W.; Tirupati, B.; Krebs, C. *J. Inorg. Biochem.* **2003**, *96*, 63–63.
(17) Galonic, D. P.; Barr, E. W.; Walsh, C. T.; Bollinger, J. M., Jr.; Krebs, C. *Nat. Chem. Biol.* **2007**, *3*, 113–116.
(18) Grzyska, P. K.; Muller, T. A.; Campbell, M. G.; Hausinger, R. P. *J. Inorg. Biochem.* **2007**, *101*, 797–808.
(19) Otwinowski, Z.; Minor, W. *Macromolecular Crystallography, Pt A* **1997**, *276*, 307–326.
(20) Read, R. J. *Acta Crystallographica Section D-Biological Crystallography* **2001**, *57*, 1373–1382.
(21) delaFortelle, E.; Bricogne, G. *Macromol. Crystallogr., A* **1997**, *276*, 472–494.

Table 1. Data and Refinement Statistics

structure	apo CytC3	iron CytC3	iron anomalous CytC3[d]
crystal parameters			
space group	$P4_32_12$	$P4_32_12$	$P4_32_12$
unit cell dimension (Å)	90.0, 90.0, 249.2	89.0, 89.0, 248.4	89.8, 89.8, 249.1
no. of molecules per ASU	2	2	2
data collection			
wavelength (Å)	1.0	1.0	1.73989
resolution limit (Å)[b]	50 − 2.2 (2.28 − 2.20)	50 − 2.2 (2.28 − 2.20)	50 − 2.75 (2.85 − 2.75)
R_{sym}[a,b]	5.6 (20.0)	8.2 (48.0)	9.8 (35.9)
no. of unique reflections	50903	50583	50102
redundancy[b]	13.7 (2.6)	10.9 (7.1)	11.0 (2.1)
completeness[b]	99.1 (81.5)	97.6 (81.3)	98.3 (91.2)
I/σ[b]	33.8 (3.7)	27.8 (2.5)	11.3 (2.3)
refinement			
R_{cryst}/R_{free}[c]	21.8/26.0	23.8/27.8	
no. of protein atoms	4313	4321	
no. of metal ion	0	2	
no. of ligands	5	13	
no. of water molecules	249	247	
disordered residues	170−172, 178−220 (chain A) 179−215 (chain B)	178−220 (chain A) 179−216 (chain B)	
rmsd bond lengths (Å)	0.0065	0.0077	
rmsd bond angles (°)	1.33	1.43	
Ramachandran analysis			
most favored (%)	89.2	86.4	
allowed (%)	9.7	12.1	
generously allowed (%)	0.9	0.7	
disallowed (%)	0.2	0.9	
B-factors (Å^2)			
protein	41	46	
iron		49	
αKG		67	
water	45	50	

[a] $R_{sym} = (\Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|)/\Sigma_{hkl}\Sigma_i I_i(hkl)$ for n independent reflections and i observations of a given reflection. $\langle I(hkl)\rangle$ = average intensity of the ith observation. [b] Numbers for the highest resolution shell are shown in parentheses. [c] $R_{cryst} = \Sigma_{hl}||F_o(hkl)| - |F_c(hkl)||/\Sigma_{hkl}|F_o(hkl)|$, where F_o and F_c are the observed and calculated structure factors, respectively. R_{free} is calculated the same way with a test set of reflections (10%) that are not used during refinement. [d] An iron anomalous CytC3 data set was collected at iron peak wavelength and scaled anomalously.

were added to the model after few rounds of refinement, and further model building and refinements were done in COOT[24] and CNS, respectively. No noncrystallographic symmetry (NCS) or sigma cutoff was used during the refinement. The final apo structure was refined to 2.2 Å resolution with $R_{cryst} = 21.8\%$ and $R_{free} = 26.0\%$. Since the iron CytC3 crystals were isomorphous with the apo CytC3 crystals, a rigid body refinement followed by simulated annealing refinement was used. The iron structure was refined using a patch that restrains the bond lengths and angles of the ligands to the metal using higher weights for protein ligands (500 for both bond length and angle weights) and αKG (500 bond length weight and 100 angle weight), and lower weights for the water molecules (50 bond length weight and no angle weight). The final iron structure was refined to 2.2 Å resolution with $R_{cryst} = 23.8\%$ and $R_{free} = 27.8\%$ with the same reflection test set as for the apo structure. Composite omit maps calculated using CNS simulated annealing were used to validate the models, and an anomalous difference map was used to verify the positions of the irons. For both the apo and iron structures, residues 8−317 were observed out of a total of 319 residues. There is one major chain break in both structures (residues 178−220 in molecule A, and 179−216 in molecule B), and a minor chain break only in molecule A of the apo structure (residues 170−172). Figures 2, 4−6 were generated in pymol. Figure 3 was generated in ESPript server v.2.1 (Laboratorie des Interactions Plantes-Microorganismes, France).

The CytC3 coordinates have been deposited to the Protein Databank: apo CytC3 (PDB ID 3GJA) and iron CytC3 (PBD ID 3GJB).

Results

Structural Overview. We have determined the structures of apo and iron/αKG-bound CytC3, both at 2.2 Å resolution. Both structures belong to the $P4_32_12$ space group, and the asymmetric unit contains a dimer of two CytC3 molecules related by 2-fold noncrystallographic symmetry. Although CytC3 crystallizes as a dimer, gel filtration chromatography indicates that CytC3 functions as a monomer.[6] The core structure of CytC3 is a β-sandwich "jelly roll" motif, common to the αKG nonheme iron enzyme family, and it comprises β-strands β2, β3, β5, β6, and β9−12 (Figure 2a). Dimerization occurs between β-strand 3 of each monomer, each strand hydrogen bonding to form a continuous eight-stranded β-sheet (consists of β3, β6, β9, and β11 from each monomer). In addition to the β-strands, there are five α-helices near the N-terminus packed on the outside of the dimer (Figure 2a). The apo and iron/αKG-bound CytC3 structures are very similar, with an rmsd of 0.461 Å over 263 C_α atoms.

Each monomer in the CytC3 dimer has a large disordered region encompassing residues 178 to 219, which we will refer to as the "missing loop" (Figure 2a). There is electron density around the vicinity of the missing loop indicating the presence of mobile amino acids; however, the quality of the density is too poor to model. Residues near the ends of the missing loop have elevated B-factors compared to the rest of the protein, 75

(22) McRee, D. E. *J. Struct. Biol.* **1999**, *125*, 156–165.
(23) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905–921.
(24) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.

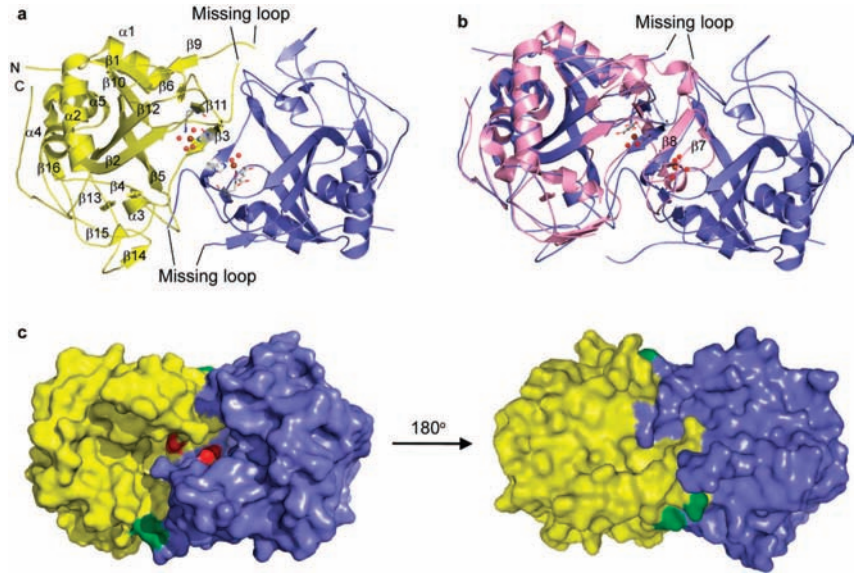

Figure 2. Overall structure of CytC3 dimer. (a) Crystallographic dimer is colored by molecule. The iron ligands: two protein residues (His118 and His240) and αKG are shown in stick representation, and waters in the active site are shown in a spherical representation. The ends of the disordered region are indicated as "missing loop", which consists of residues 178−219. (b) Structural alignment of SyrB2 monomer (pink) and CytC3 dimer (blue). (c) Surface representation of CytC3 dimer showing access to active site on one face of the crystallographic dimer. Active site waters are colored in red; each monomer of CytC3 is colored yellow and blue. The ends of the disordered region are colored in green.

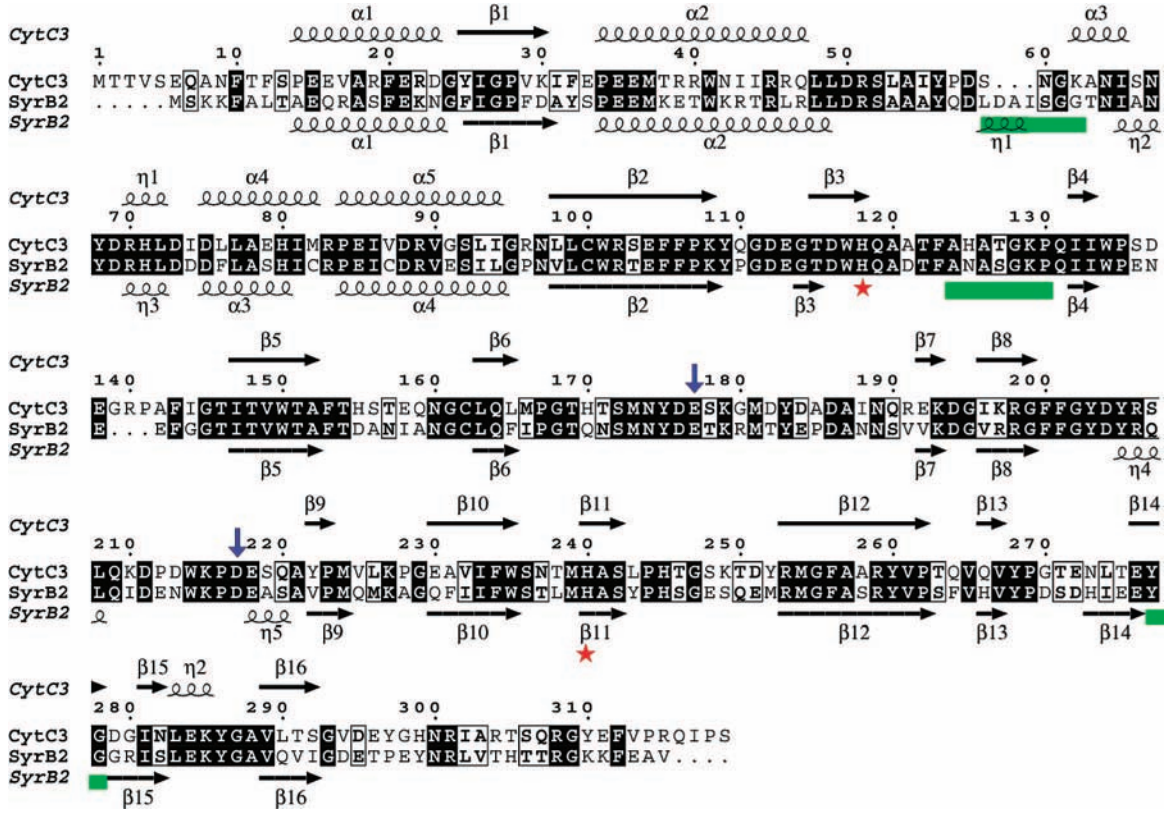

Figure 3. Sequence alignment of CytC3 and SyrB2 with secondary structures defined above the sequence. Alpha helices are indicated by helical lines with α labels, beta strands are indicated by arrows with β labels, and 3_{10} helices are indicated by helical lines with η labels. Red stars indicate the active site residues, blue arrows indicate the ends of the missing loop, and green bars indicate the chloride binding loops in SyrB2.

Å^2 compared to 45 Å^2. Structural alignment between CytC3 and SyrB2, with an rmsd of 1.16 Å over 221 C_α atoms, predicts that two β-strands (β7 and β8) should be located in the missing loop (Figures 2b and 3). These two β-strands and the residues of the corresponding missing loop are very important in SyrB2 as they cover the active site.[8]

The structural alignment between CytC3 and SyrB2 strongly indicates that the interactions between the two molecules of CytC3 in the asymmetric unit are a crystallographic artifact. The core secondary structural elements in both CytC3 and SyrB2 align well, with the exception of the two β-strands and the two 3_{10} helices that cover the active site in SyrB2 that are missing

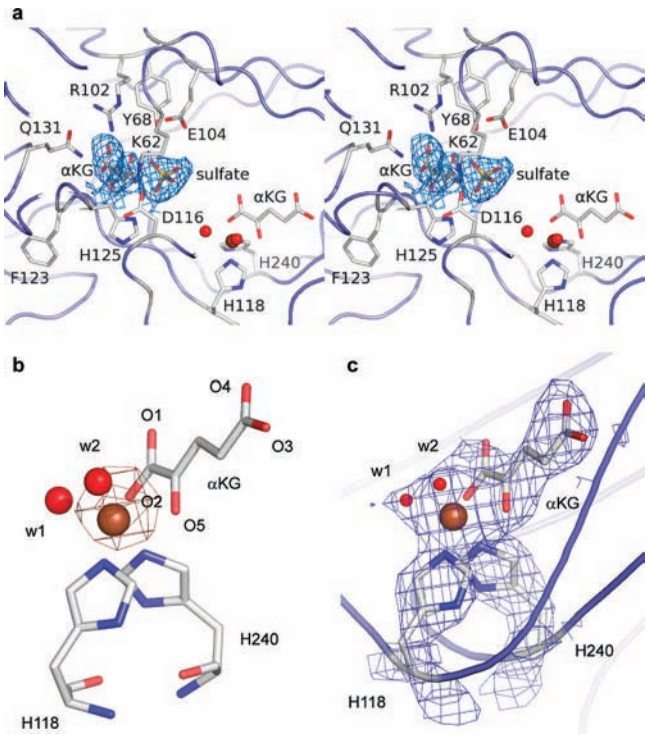

Figure 4. CytC3 active site region. (a) Open active site in CytC3 structure allows for molecules from crystallization solution to bind, shown in stereo view. A $2F_o - F_c$ composite omit map contoured at 1σ is shown in blue mesh around excess ligands (αKG and sulfate). Iron is shown in brown, and waters are shown in red. (b) An iron anomalous difference map contoured at 12σ is shown in brown mesh around the iron. The αKG is labeled with the numbering nomenclature for each oxygen atoms. (c) A $2F_o - F_c$ composite omit map contoured at 2σ is shown in blue mesh around the active site ligands.

in CytC3. This part of the structure should occupy the space that is occupied by the second molecule in the CytC3 noncrystallographic dimer (Figure 2b). Therefore, crystal packing has allowed us to capture an open active site conformation of the CytC3 protein. On the basis of sequence alignment, it is predicted that in solution the intermolecular hydrogen bonds across the two CytC3 molecules are not present, and the missing loop in CytC3 folds over the active site as it does in SyrB2 (Figure 3). Given that dimerization of CytC3 requires shifting the missing loop (residues 178−220) out of the way, one can imagine that this region of the structure has some mobility, and it may be important in interacting with the substrate, which is the amino-acylated thiolation domain (CytC2).

Active Site Features. The active sites of the two CytC3 monomers are located at the dimer interface positioning them very close to each other; the irons are within 10 Å. Surface representations reveal that the active sites are highly solvent-exposed from one side of the dimer (Figure 2c). Unlike most other enzymes in the cupin superfamily, CytC3 and SyrB2 only contain two histidine residues as protein ligands to the iron.[8] Other iron coordination sites are available to bind chloride, waters and/or αKG. Since cocrystallization with iron, chloride, and αKG yielded slow growing apo crystals, we obtained a structure with iron and αKG bound by soaking the apo CytC3 crystals in high concentrations of $Fe(NH_4)_2(SO_4)_2$ (10 mM), αKG (5 mM), and chloride (80 mM). In the soaked structure, both CytC3 molecules have one iron bound; however, even with such high concentrations of external ligands, only one molecule in the dimer has αKG in the active site, and no chloride is observed in either molecule.

In addition to the iron and the iron-bound αKG, the open and solvent accessible active sites of CytC3 can accommodate other molecules from the crystallization conditions such as excess αKGs, acetate, and sulfate. There are many charged residues near the active site that interact with these molecules. For example, the side chains of R102 and D116 and the backbone of H125 and F123 interact with an extra molecule of αKG, and a sulfate ion hydrogen bonds with D116 and K62 (Figure 4a). It is unlikely that these excess binding events are physiologically relevant.

To verify the presence of iron in the active site, a data set was collected at the iron peak wavelength, 1.73989 Å. One iron per monomer is observed at the active site as judge by the 12σ peak in the anomalous difference map (Figure 4b). Additionally, if water molecules are placed in the iron sites and refined after fixing the B-factor to the average B-factors of the protein at 45 Å^2, the resultant occupancies for waters are 3.19 and 2.91, suggesting the presence of a heavier atom at the site. The occupancies for iron atoms refined in the iron sites are 0.93 and 0.84, which are consistent with the conclusion based on the anomalous difference map.

Chloride is observed bound to iron in the SyrB2 structure, an interaction that is expected to be important for the chlorination activity of the protein.[8] Since CytC3 and SyrB2 are structurally similar to each other and perform the same chemistry, it is predicted that the mechanism by which CytC3 chlorinates is the same as that for SyrB2. Therefore, a chloride ion is expected to be one of the ligands for the iron. To examine whether water or chloride is bound to iron, B-factors were refined for modeled water and modeled chloride after fixing occupancies to one. The B-factors for chloride ions modeled as ligands are 185.3 Å^2 and 178.3 Å^2, which are much higher than the average B-factors for the protein, suggesting that an atom lighter than chloride is bound. The B-factors for modeled waters are 43.5 Å^2 and 32.2 Å^2, which are consistent with the average B-factors for the rest of the protein, indicating the presence of water molecules in the other open iron coordination sites (Figure 4c). It is surprising that no chloride ion is bound to the iron given the concentration of chloride used in the soaks was 80 mM. In contrast, iron-containing SyrB2 was unable to be crystallized without a chloride bound.[8]

The composite omit map of the active site only shows electron density for αKG in one of the CytC3 monomers (Figure 4c, data for the monomer without αKG is not shown), suggesting that the open active site conformation of CytC3 is not an optimal environment for αKG binding. The B-factor for the active site αKG is also higher than the average B-factor for the protein, 67 Å^2 versus 45 Å^2, indicating that the αKG is not as well ordered. The αKG binds to the iron through its O2 and O5 oxygens (Figure 4b for numbering nomenclature) in a bidentate and planar fashion. Bidentate binding of αKG to iron is strictly conserved in the αKG/Fe(II)-dependent enzyme family, although the geometry of αKG varies.[25]

CytC3 and SyrB2 Active Site Comparison. Our results suggest that residues that do not directly contact the iron are also important in binding αKG and chloride in these halogenating enzymes. These surrounding residues are highly conserved between CytC3 and SyrB2; however, due to the difference in crystal lattice contacts, the interactions of these CytC3 residues

(25) Hausinger, R. P. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21–68.

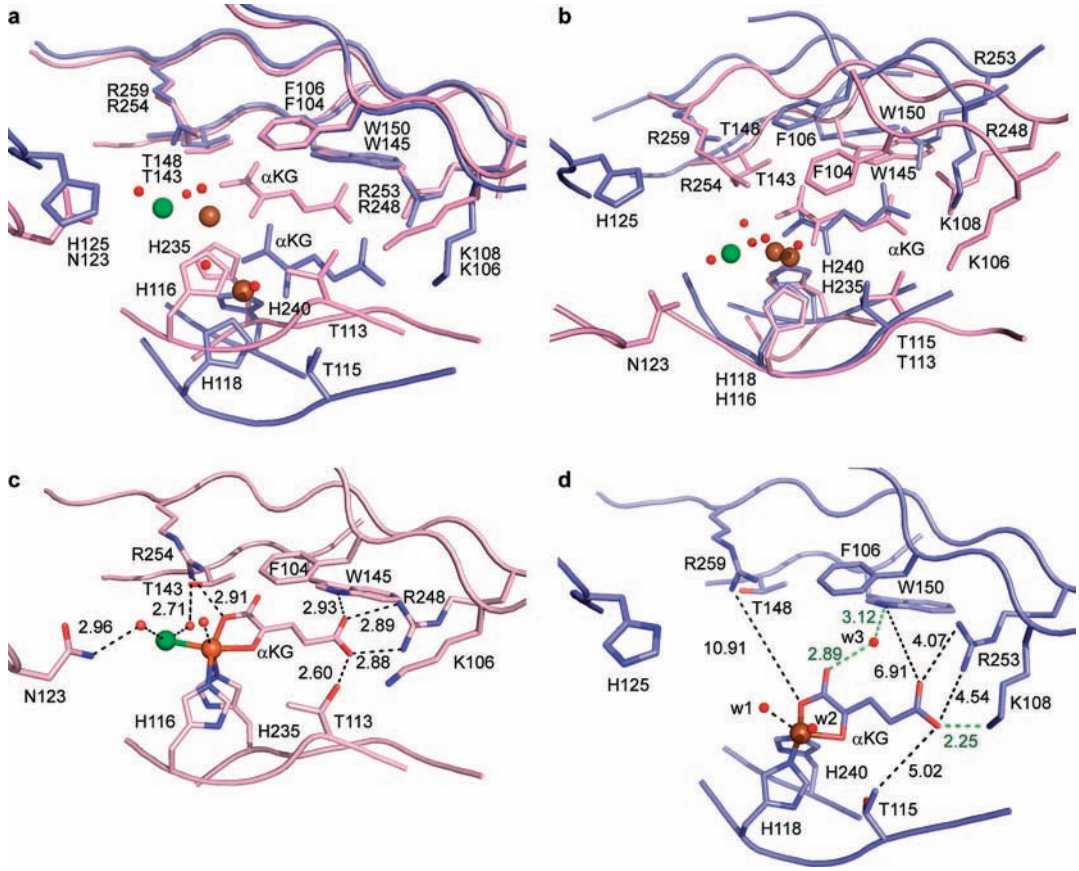

Figure 5. Active site comparison for CytC3 and SyrB2. Iron is shown in brown, chloride ion is shown in green, and waters are shown in red. (a) Comparison between CytC3′s "open" active site (blue) and SyrB2's "closed" active site (pink), where structures are aligned by F106, K108, H125, W150, T148, R253, and R259 of CytC3 with the corresponding residues in SyrB2. (b) Comparison of CytC3 (blue) and SyrB2 (pink) where structures are aligned by the two active site histidines. (c) Active site residues of SyrB2 and relevant distances. Dash lines indicate hydrogen bonding with labeled distances. (d) Active site residues of CytC3 and relevant distances. The corresponding hydrogen bond interactions between the protein and αKG as observed in SyrB2 are indicated by black dashes. New interactions observed between CytC3 and αKG are indicated by green dashes with labeled distances.

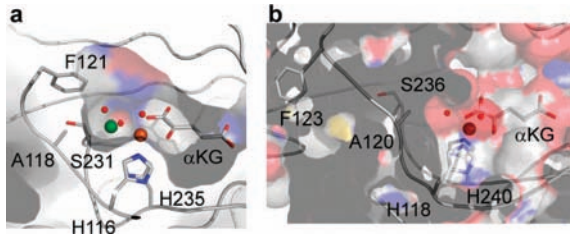

Figure 6. Chloride binding pocket of SyrB2 and CytC3. Iron is shown in brown, chloride ion is shown in green, and waters are shown in red. (a) Residues A118, F121, and the β-carbon of S231 form a nice hydrophobic pocket for chloride in the closed state of SyrB2 active site. (b) The corresponding residues, A120, F123, and S236, in the open state of the CytC3 active site are too far from the chloride binding site to form a hydrophobic pocket environment.

with αKG are not the same as those observed in SyrB2. If the structures of CytC3 and SyrB2 are aligned by all the residues on one side of the active site (H125, R259, T148, F106, W150, R253, and K108 in CytC3, and N123, R254, T143, F104, W145, R248, and K106 in SyrB2), then the agreement of the aligned residues is good. However, the agreement between the unaligned residues (H118, H240, and T115 for CytC3, and H116, H235, and T113 for SyrB2) is not (Figure 5a). Likewise, when residues on the other side of the active site (T115 in CytC3, T113 in SyrB2) and (H118 and H240 in CytC3, H116 and H235 in SyrB2) are aligned, the upper half of the active site in CytC3 is no longer superimposed with that in SyrB2 (Figure 5b). Collectively, these differences dramatically affect the contacts between protein, αKG and chloride.

There are a number of interactions that anchor the αKG in the SyrB2 active site, such as hydrophobic stacking against F104 and many hydrogen binding interactions (Figure 5c). Hydrophobic stacking between αKG and the protein is not observed in CytC3 due to the slightly altered conformation of αKG, and the greater distance between the αKG and the corresponding phenylalanine (F106) (Figure 5d). There are also multiple hydrogen bonds between surrounding residues and αKG in SyrB2: R254 interacts with O2 of αKG (Figure 4b for αKG numbering nomenclature), W145 and R248 interact with O3 of αKG, R248 and T115 interact with O4 of αKG (Figure 5c). However, none of these interactions are preserved in the "open" CytC3 active site; in each case, the distances are too long to represent hydrogen bonds (Figure 5d). Instead, new interactions are observed in this conformation of the CytC3 active site to stabilize the αKG: K108 interacts with O4 of αKG, and W150 interacts with O1 of αKG through a water molecule (w3) (Figure 5d). Overall, the "open" active site conformation of CytC3 results in αKG having fewer total stabilizing interactions compared to SyrB2, which is consistent with the higher degree of disorder observed for this αKG molecule. We predict that these weak interactions between αKG and CytC3 are only present when the active site is in a relatively open state as observed in this structure, and that αKG binding will mimic

that in SyrB2 in a closed structure of the active site as would occur during catalysis.

The comparison of CytC3 and SyrB2 active sites provides insight into why no chloride is bound to iron in CytC3. In SyrB2, a hydrogen-bonding network between residues N123, T143, R254, and chloride connected by water molecules appears to stabilize the chloride in the active site (Figure 5c). However, no such hydrogen-bonding network is observed in CytC3, as the corresponding residues H125, T148, and R259 are too far from the chloride binding site for a single water molecule to form a hydrogen-bonding network. Additionally, the chloride in SyrB2 is situated in a large hydrophobic pocket, consisting of A118, F121, and the β-carbon of S231 (Figure 6a), that might be important for its binding.[8] This hydrophobic pocket is not present in this open structure of CytC3; while the alanine is in the same relative place, both the corresponding phenylalanine (F123) and serine (S236) residues are more than 12 Å and 8 Å away, respectively (Figure 6b).

There are three chloride-binding loops in SyrB2 (55−61, 122−128, and 271−275) that undergo conformational changes to close the active site upon chloride binding.[8] On the basis of the structural alignment between CytC3 and SyrB2, the predicted chloride-binding loops in CytC3 consist of residues 60−63, 124−130, and 278−280 (Figure 3). The first predicted chloride-binding loop in CytC3 is shortened by a three-residue gap compared to that of SyrB2, and residue 63 is part of α-helix 3 instead of being part of a loop. The second predicted chloride-binding loop in CytC3 precedes β-strand 4 in the structure, as is the case for SyrB2. The third predicted chloride-binding loop in CytC3 is located at a hairpin between β-strands 14 and 15. No movement in these predicted chloride-binding loops is observed when the apo and the Fe(II)/αKG/Cl^- soaked CytC3 structures are compared, which is not surprising since no chloride is bound to the iron center in CytC3. It is difficult to confirm the locations, the lengths, or the roles of these chloride-binding loops in CytC3 without a chloride bound structure for comparison.

Discussion

The metal centers in CytC3 and SyrB2 vary from the common iron centers found in other members of the αKG-mononuclear iron enzyme family. These halogenases harvest the oxidizing power of the simple metal center of other αKG/Fe(II)-dependent enzymes and adapt it to bind chloride for chlorination of unactivated carbons. The metal centers in hydroxylases with three-protein ligands (two His and one carboxylate from Asp or Glu) are transformed to a center with two-protein ligands (only two His) in halogenases; in the place of the carboxylate, an alanine residue is present in the halogenases, leaving room for chloride to bind to the iron.[8] There has been much interest in understanding the features that are important for chloride binding to an iron center, converting hydroxylation chemistry to that of halogenation.[26] Is it a simple steric issue or are there other factors involved, such as hydrogen-bonding or hydrophobic interactions through the secondary coordination sphere of the metal? The structural comparison of a halogenase containing no bound chloride with one containing bound chloride has allowed us to investigate these questions. Although the CytC3 used in these studies is fully active, a crystalline state of this enzyme has been captured to which chloride binding is impaired. While in solution, this "open" chloride-free state would be in equilibrium with a "closed" chloride-bound state, in the crystal, the transformation between states is prevented by lattice contacts, allowing us to observe this otherwise transient enzyme conformation.

With only two protein ligands in iron halogenases, iron does not bind tightly in the absence of αKG. When apo CytC3 crystals are soaked in 10 mM $Fe(NH_4)_2(SO_4)_2$, or soaked in 10 mM $Fe(NH_4)_2(SO_4)_2$ first, followed by 5 mM αKG, or vice versa, little density is observed in the active site for αKG or iron, indicating very low occupancies of both. Iron is only bound in high occupancy to the crystalline CytC3 active site when the crystals are soaked in both 10 mM $Fe(NH_4)_2(SO_4)_2$ and 5 mM αKG together, which suggests the binding of αKG helps to keep iron in the active site. Poor iron binding is also observed in the D101A mutant of TauD, where the iron content is low (0.50 ± 0.49 mol Fe/mol subunit) when it is measured without the presence of αKG.[18] Like CytC3, SyrB2 can only be reconstituted with 0.3−0.4 mol Fe/mol subunit in the absence of αKG, but the iron content is much higher (0.85−0.95 mol Fe/mol subunit) in the presence of αKG.[7]

The hydrogen-bonding network between residues surrounding the active site and the chloride appears to be an important feature for chloride binding in CytC3 and SyrB2. The "closed" active site in SyrB2 allows residues near the metal center (R254, T143, and N123) to hydrogen bond with chloride through two water molecules. In the light-driven chloride pump halorhodopsin, the chloride is also stabilized in a hydrogen-bonding network with protein residues via water molecules.[27] However, due to the "open" active site in this CytC3 structure, the corresponding residues are too far from the chloride binding site for hydrogen bonding, which may contribute to the lack of observed chloride binding. Secondary coordination sphere hydrogen bonds have been shown to be important for modulating ligand affinity to metal centers in both metalloproteins as well as synthetic model compounds that mimic enzymatic active sites. For example, the removal of a hydrogen bond distal to the active site in horseradish peroxidase changes the metal center from five coordinate to six coordinate in the resting state.[28] Likewise, in a series of model compounds with varying numbers of secondary coordination sphere hydrogen bonds, a range of affinities for oxygen binding to cobalt is observed.[29]

Another feature that may be important for chloride binding to iron is the presence of a suitable hydrophobic pocket around the chloride. In the closed SyrB2 active site, hydrophobic residues (Ala, Phe, and the β-carbon of Ser) are within 4 Å of chloride, providing a hydrophobic pocket for chloride binding. Corresponding residues in the open CytC3 active site are more than 8−12 Å from the chloride, and are thereby unable to provide a similar environment. Hydrophobic chloride binding sites have been observed in the light driven chloride pump halorhodopsin and ClC chloride binding channel as well. In both of these proteins, the chloride binding environment is made up of hydrophobic residues such as Gly, Ile, Phe, Trp, the β-carbon of Ser, and the γ-carbon of Thr.[27,30] Along the lines of a binding pocket for chloride, Tolman, Que, and co-workers have recently

(26) Straganz, G. D.; Nidetzky, B. *Chembiochem* **2006**, *7*, 1536–1548.
(27) Kolbe, M.; Besir, H.; Essen, L. O.; Oesterhelt, D. *Science* **2000**, *288*, 1390–1396.
(28) Mukai, M.; Nagano, S.; Tanaka, M.; Ishimori, K.; Morishima, I.; Ogura, T.; Watanabe, Y.; Kitagawa, T. *J. Am. Chem. Soc.* **1997**, *119*, 1758–1766.
(29) Lucas, R. L.; Zart, M. K.; Murkerjee, J.; Sorrell, T. N.; Powell, D. R.; Borovik, A. S. *J. Am. Chem. Soc.* **2006**, *128*, 15476–15489.
(30) Dutzler, R.; Campbell, E. B.; Cadene, M.; Chait, B. T.; MacKinnon, R. *Nature* **2002**, *415*, 287–294.

explored the use of a sterically bulky α-ketocarboxylate ligands to prepare structural models of iron-halogenase active sites.[31] They found that the use of bulky α-ketocarboxylate ligands leads to inactive iron complexes that structurally mimic halogenase active sites, complete with an Fe−Cl bond, but are unable to catalyze oxidative decarboxylation of α-keto ligand. However, the use of less bulky α-ketocarboxylate ligands leads to the formation of iron complexes that lack an Fe−Cl bond, but are able to catalyze oxidative decarboxylation of the α-keto ligand. Thus, while steric hindrance is useful for the creation or stability of the Fe−Cl bond in these models, it impedes decarboxylation activity, presumably by preventing the attack of the bound O_2 on the α-keto group. The challenge will be to create the appropriate binding pocket for chloride without interfering with the oxidative decarboxylation reaction.

Since the discovery of αKG/Fe(II) dependent halogenases, several groups have tried to convert hydroxylases to halogenases and vice versa. In an attempt to convert SyrB2 to do hydroxylation chemistry, A118D and A118E mutant proteins were prepared. These mutations abolished chlorination activity, but no hydroxylation activity was observed.[8] Hausinger et al. tried to convert hydroxylase TauD to do a chlorination reaction by mutating the active site carboxylate to an alanine, but the D101A variant showed no chlorination activity.[18] However, in the latter case, it is unclear if the lack of chlorination was due to a lack of chloride and/or iron binding to the active site or if the mutation introduced a perturbation to the protein structure. The data presented here suggest that the selective activity toward hydroxylation or chlorination depends on more than just the steric issue of the number of protein ligands to iron. On the basis of our observations in the open CytC3 active site structure, chloride binding to iron depends on a combination of a hydrogen-bonding network and/or hydrophobic environment, in addition to an open coordination at the iron site.

(31) Friese, S. J.; Kucera, B. E.; Young, V. G., Jr.; Que, L., Jr.; Tolman, W. B. *Inorg. Chem.* **2008**, *47*, 1324–1331.

Conclusions

The fortuitous "open" active site structure of CytC3 has allowed us to investigate the protein features that allow chloride to bind directly to the iron center, a necessary step for chlorination and not hydroxylation in this family of enzymes. The structure of SyrB2 revealed that the substitution of Ala for Asp/Glu creates space in the protein structure for chloride to bind directly to iron, however, as the biochemistry shows, this substitution is insufficient to switch a hydroxylase to a halogenase. Therefore, the SyrB2 structure alone did not indicate which other protein features might be important for chloride binding. Here, we find that chloride binding is impaired when a hydrogen-bonding network to the surrounding residues is interrupted and when a hydrophobic pocket is not properly formed. These findings will assist in the protein design of halogenating enzymes and in the creation of model complexes.

Acknowledgment. We thank K.S. Ryan for help with data collection. This work was supported in part by grants from the NIH (GM65337 to C.L.D., and GM49338 to C.T.W.), from the MIT Center for Environmental Health Sciences NIEHS P30 ES002109, and from the Damon Runyon Cancer Research Foundation Postdoctoral Fellowship DRG-1893-05 (to D.G.F.). The Argonne Photon Source and Stanford Synchrotron Radiation Laboratory are supported by the US Department of Energy.

JA8097355