X-ray Crystal Structure of *Escherichia coli* Taurine/α-Ketoglutarate Dioxygenase Complexed to Ferrous Iron and Substrates[†,‡]

Jonathan M. Elkins,[§] Matthew J. Ryle,[||] Ian J. Clifton,[§] Julie C. Dunning Hotopp,[||] John S. Lloyd,[§] Nicolai I. Burzlaff,[§] Jack E. Baldwin,[§] Robert P. Hausinger,*[,||] and Peter L. Roach*[,§,⊥]

Dyson Perrins Laboratory, Oxford University, South Parks Road, Oxford, England, OX1 3QY, and Departments of Microbiology & Molecular Genetics and Biochemistry & Molecular Biology, Michigan State University, East Lansing, Michigan 48824-1011

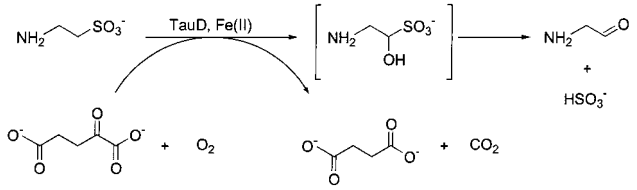
Received December 6, 2001; Revised Manuscript Received February 14, 2002

ABSTRACT: Taurine/α-ketoglutarate dioxygenase (TauD), a non-heme Fe(II) oxygenase, catalyses the conversion of taurine (2-aminoethanesulfonate) to sulfite and aminoacetaldehyde concurrent with the conversion of α-ketoglutarate (αKG) to succinate and CO_2. The enzyme allows *Escherichia coli* to use taurine, widely available in the environment, as an alternative sulfur source. Here we describe the X-ray crystal structure of TauD complexed to Fe(II) and both substrates, αKG and taurine. The tertiary structure and fold of TauD are similar to those observed in other enzymes from the broad family of Fe(II)/αKG-dependent oxygenases, with closest structural similarity to clavaminate synthase. Using the TauD coordinates, a model was determined for the closely related enzyme 2,4-dichlorophenoxyacetate/αKG dioxygenase (TfdA), supporting predictions derived from site-directed mutagenesis and other studies of that biodegradative protein. The TauD structure and TfdA model define the metal ligands and the positions of nearby aromatic residues that undergo post-translational modifications involving self-hydroxylation reactions. The substrate binding residues of TauD were identified and those of TfdA predicted. These results, along with sequence alignment information, reveal how TauD selects a tetrahedral substrate anion in preference to the planar carboxylate selected by TfdA, providing insight into the mechanism of enzyme catalysis.

Sulfonates and sulfates are widely available in the environment where they are utilized as alternatives to inorganic sulfur sources (*1*). For example, *Escherichia coli* possesses *tauABCD* and *ssuEADCB* operons for utilizing taurine (2-aminoethanesulfonate) and aliphatic sulfonates, respectively (*2*, *3*). These sulfur sources are taken up by ATP-binding cassette-type transporters, TauABC and SsuABC (*4*), and subsequently oxidized by TauD and SsuDE to release sulfite for cell growth (*5*, *6*). Similarly, aromatic sulfonates and alkyl sulfate esters are metabolized by many microorganisms (*1*) such as *Pseudomonas putida* S-313, which possesses SsuED and AtsK for oxidation of these substrates (*7*, *8*). The work described here features TauD, an oxygenase from *E. coli* required for the decomposition of taurine and selected other sulfonates.

TauD is an Fe(II)- and α-ketoglutarate (αKG)[1]-dependent dioxygenase that catalyzes the conversion of taurine to sulfite and aminoacetaldehyde (Scheme 1) (*5*). Like many members

Scheme 1: Reaction Catalyzed by TauD



of the αKG dioxygenase family (*9*), TauD couples the oxidative decarboxylation of αKG to the hydroxylation of a second substrate. In this case, the hydroxylated taurine intermediate is unstable and decomposes to yield the final products. Sequence analyses indicate that TauD is closely related to several members (referred to as group II) of the αKG dioxygenase family (*10*). For example, a HX(D/E)-X_{23-26}(T/S)$X_{114-183}$HX_{10-13}R motif is found in TauD, YSD (a broad specificity yeast sulfonate/αKG dioxygenase; (*11*)), TfdA (2,4-dichlorophenoxyacetate (2,4-D)/αKG dioxygenase, a herbicide-degrading enzyme; (*12*)), AtsK (an alkyl sulfate/αKG dioxygenase; (*8*)), clavaminate synthase (CAS; (*13*)), and γ-butyrobetaine hydroxylase (GBBH; (*14*)). Other variations of this motif with a significantly shortened

[†] This work was supported by the BBSRC, MRC, EPSRC, the Royal Society through a fellowship to P.L.R., and the National Institutes of Health through Postdoctoral Fellowship GM20196 to M.J.R.

[‡] Atomic coordinates have been deposited in the Protein Data Bank under codes 1gqw (TauD) and 1gqx (TfdA model).

* To whom correspondence should be addressed. E-mail: hausinge@msu.edu or P.L.Roach@soton.ac.uk.

[§] Dyson Perrins Laboratory.

[||] Michigan State University.

[⊥] Current address: Department of Chemistry, University of Southampton, Highfield, Southampton, SO17 1BJ.

[1] Abbreviations: αKG, α-ketoglutarate; TauD, taurine/αKG dioxygenase; Tris, tris(hydroxymethyl)aminomethane; CAS, clavaminic acid synthase; 2,4-D, 2,4-dichlorophenoxyacetic acid; TfdA, 2,4-D/αKG dioxygenase; AtsK, alkyl sulfate/αKG dioxygenase; YSD, yeast sulfonate/αKG dioxygenase; GBBH, γ-butyrobetaine hydroxylase; DAOCS, deacetoxycephalosporin C synthase.

10.1021/bi016014e CCC: $22.00
Published on Web 03/27/2002

Table 1: Data Collection and Phasing Statistics for TauD and Its Derivatives

	native	Xe	$Me_3Pb(OAc)_2$	$(AuCl_4)^-$
wavelength (Å)	1.54179	1.54179	1.54179	1.54179
resolution range[a] (Å)	40.49−3.00 (3.16−3.00)	45.18−3.50 (3.69−3.50)	19.17−3.00 (3.17−3.00)	38.35−3.50 (3.69−3.50)
no. of unique observations	16727	10885	16933	10880
no. of total observations	98529	138591	157963	194517
completeness[a] (%)	99.3 (99.8)	99.9 (99.9)	99.2 (97.0)	99.9 (100.0)
multiplicity[a]	5.9 (6.0)	5.0 (5.0)	5.8 (5.6)	6.0 (6.3)
R_{merge}[a,b] (%)	6.5 (27.0)	5.3 (13.2)	7.0 (30.5)	10.3 (19.2)
$\langle I/\sigma(I)\rangle$[a]	10.5 (2.8)	12.6 (5.6)	10.5 (2.5)	6.2 (3.6)
R_{cullis} (cen/acen/anom)		0.80/0.77/0.88	0.85/0.85/0.88	0.85/0.82/0.98
phasing power (cen/acen/anom)		1.23/1.61/1.42	1.18/1.34/1.38	0.98/1.27/0.74
isomorphous difference (%)		23.4	25.5	13.8
no. of sites		2	4	2

[a] The numbers in parentheses indicate values for the highest resolution bin. [b] $R_{merge} = \sum_j\sum_h|I_{h,j} - \langle I_h\rangle|/\sum_j\sum_h\langle I_h\rangle \times 100$.

sequence separating the HX(D/E) motif from the second His residue are found in the extended family of αKG dioxygenases. Structurally characterized representatives of the αKG dioxygenase superfamily include CAS (*13*), deacetoxycephalosporin C synthase (DAOCS) (*15*), proline 3-hydroxylase (*16*), anthocyanidin synthase (*17*), and isopenicillin N synthase (*18*, *19*), a related enzyme that does not utilize αKG. Each of these family members possesses a core motif of eight beta-sheets arranged in a "jellyroll" manner, with the order of strands being identical in each structure and with similarly arranged Fe(II) ligands (the Asp/Glu and two His residues shown in the above motif). All of the family members are likely to function by related mechanisms, and observations made on one representative may allow accurate predictions for the others.

Here, we structurally characterize TauD, use it to model the structure of TfdA, and interpret the structures in terms of earlier mechanistic studies. In particular, the structures reveal the proximity of the iron center to residues known to undergo post-translational processing by self-hydroxylation reactions in both TauD[2] and TfdA (*20*). Additionally, the structures show the unique features involved in distinguishing tetrahedral anionic substrates from planar carboxylate substrates within this enzyme family.

EXPERIMENTAL PROCEDURES

Enzyme Purification and Assay. TauD apoprotein was purified and assayed as previously described (*21*). The reconstituted enzyme exhibited specific activities ranging from 4.2 to 7.8 μmol min^{-1} (mg of $protein^{-1}$).

Crystallization. Crystals were grown by the hanging drop method using Linbro multiwell plates under an anaerobic atmosphere of argon (0.1−0.2 ppm O_2) at 17 °C. All materials except the protein solution were placed in the anaerobic environment at least 24 h prior to crystallization experiments. The protein solution was prepared by adding 1.4 molar equivalents of $FeSO_4$ from a 15 mM solution to 75 μL of protein (18 mg/mL in 20 mM Tris, pH 7.0). After 1 min, 5 μL of 200 mM αKG, 5 μL of 200 mM taurine (pH 7.0), and 5 μL of 100 mM DTT were added. Crystals grew from a variety of poly(ethylene glycol) (PEG) precipitants and buffers. Optimization of crystal size led to a well solution (500 μL) containing 20−28% PEG 1000, 75 mM imidazole (pH 7.0), and 20% ethylene glycol. Drops were prepared by mixing equal volumes of protein and well solutions with a total drop size of 4−8 μL. Hexagonal diamond-shaped crystals were obtained after 2 days, reaching a general size of 0.2 × 0.1 × 0.1 mm after 6 days. Careful control of the crystallization conditions, in particular, the Fe(II) concentration and the time spent deoxygenating the protein, allowed less nucleation, resulting in crystals of a maximum size of 1.0 × 0.5 × 0.5 mm. The crystals were shown to belong to space group $P6_222$ ($a = b = 117$ Å, $c = 202$ Å), and the largest crystals were found to diffract X-rays to 2.8 Å resolution. The solvent content of the crystals was 60%, giving a Matthews coefficient of 3.0. The crystals were very sensitive to changes in their surrounding solution, which initially caused difficulties in finding a suitable cryoprotectant, but this problem was resolved by modification of the crystallization conditions such that no subsequent cryoprotection step was necessary.

Data Collection and Structure Determination. An interpretable electron density map was obtained using the multiple isomorphous replacement method. Crystals were cryocooled in a stream of nitrogen (Oxford Cryosystems) at 100 K as soon as possible after removal from the anaerobic environment. No additional cryoprotectant was used after removal of the crystal from the hanging drop. Diffraction data were collected using a Rikagu RU-200 rotating anode generator with Osmic focusing mirrors and a 345 mm Mar Research image plate. The xenon derivative was prepared by exposing a crystal mounted in a nylon loop to xenon gas (10 bar) for 10 min. The other heavy atom derivatives were prepared by addition of a concentrated solution of the heavy atom (0.5 μL) to the hanging drop. The crystals were highly sensitive to changes in their environment, and extensive difficulties were encountered due to lack of isomorphism between different crystals. It was necessary to screen a large number of crystals before suitable derivatives were obtained. Data statistics are shown in Table 1. Data were processed with MOSFLM (*22*) and the CCP4 suite of programs (*23*). Heavy atom sites for the xenon derivative were determined using CNS (*24*), and sites for the other derivatives were determined by difference Patterson methods. Initial phases were calculated using the program SHARP (*25*) and subsequently improved by solvent flattening, histogram matching, and 2-fold noncrystallographic symmetry averaging using the program DM (*26*). The overall figure of merit increased from 0.42 to 0.68 following density modification. Skeletonized electron density was calculated using MAPMAN (*27*), and

[2] Unpublished observations by M. J. Ryle, A. Liu, R. B. Muthukumaran, B. S. Phinney, R. Y. N. Ho, J. McCracken, L. Que, Jr., and R. P. Hausinger.

Table 2: TauD Refinement Statistics

resolution limits (Å)	40.44−3.00
R_{cryst} (%)	28.1
R_{free} (%)	32.0
rmsd bonds (Å)	0.009
rmsd angles (°)	1.4
rmsd dihedrals (°)	24.6
rmsd impropers (°)	0.89
average *B*-factor (Å^2)	63.15

a protein model was built using the program O (*28*). Refinement of the model was carried out by simulated annealing with torsion angle dynamics using CNS. A maximum temperature of 5000 K and a cool rate of 50 K were used. Grouped *B*-factor refinement with two groups per residue was performed after each simulated annealing step. The initial model was built using strict noncrystallographic symmetry, changing to NCS restraints when the majority of the model was built, and a significant drop in R_{free} resulted. Clear peaks in the electron density difference maps identified the positions of the iron and the two substrates. Restraints on the binding distances and angles of the Fe(II) ligands were used due to the low resolution of the data. The values were taken from the published CAS structure (PDB ID, 1DRY). Residues not visible in the density were omitted from the model, while those whose side-chains were not visible were modeled as alanine. The final model contains 4348 protein atoms, out of a theoretical maximum of 4582. Refinement statistics are shown in Table 2.

Modeling Methods. The PHD predicted TfdA secondary structure elements (*29*) were aligned with the crystallographically determined secondary structures of TauD (determined here) and CAS1 (PDB ID: 1DS1). The TfdA sequence was modeled onto the TauD chain A and CAS1 crystal structures ignoring iron, αKG, and primary substrates using Modeller4 (*30*). TfdA residues for which there were not coordinates of an equivalent residue in TauD were masked. From the initial model, the loop refinement method was used to correct an α-helix distortion near residues 34−36 and to remove six disallowed phi/psi angles. This method was also used in efforts to remodel the loop formed by residues 86−110, which contained a knot. Numerous possible models were generated for this region, even including the additional criterion that residues 96−103 form an α-helix as suggested by PHD secondary structure prediction. No model was convincingly superior to the others, and since this region of TfdA contains a 16 amino acid insertion when compared to TauD, residues 86−111 were omitted from the model. αKG and iron were inserted into the TfdA model based on alignment of its 2 His − 1 carboxylate facial triad with that of TauD, with direct substitution of the αKG and iron positions. The residues interacting with the C-5 carboxylate of αKG (Thr141 and Arg274) were adjusted to optimize hydrogen bonding and ionic interactions, based on the CAS1 and TauD coordinates. 2,4-D was manually positioned into the modeled TfdA active site using InsightII (Molecular Simulations Inc.). As an initial constraint, the sites of hydroxylation (i.e., the α carbon of the 2,4-D side chain and the carbon adjacent to the sulfur of taurine) were superimposed. Since TfdA selectively oxidizes the *pro-R* hydrogen of 2,4-D (*31*), this hydrogen was positioned to face the iron active site. The substrate carboxylate was rotated to allow favorable interactions with Arg278, Lys71, His214, and the backbone amide of Ser117.

RESULTS AND DISCUSSION

Crystal Growth. TauD crystals were obtained from anaerobic protein solutions containing Fe(II), αKG, taurine, buffer, and precipitants. The inclusion of both taurine and αKG was absolutely required for crystal growth, suggesting that substrate binding confers a more rigid structure to the protein. Such results are consistent with the reported increased affinity for αKG in the presence of taurine (*21*). Furthermore, spectroscopic studies carried out on copper-substituted TfdA, a protein that is ∼30% identical in sequence to TauD, showed that binding of the two substrates resulted in enhanced definition of the metal site environment (*10*). The solution used for crystal growth was pink due to low-lying metal-to-ligand charge-transfer transitions arising from chelation of Fe(II) by the C-1 carboxylate and C-2 keto groups of αKG (λ_{max} = 520 nm and ϵ_{520} = 180 M^{-1} cm^{-1}, (*21*)). Although this chelate structure is present within the crystal (see below), the crystals appeared colorless.

Upon exposure to oxygen, the crystals developed a strong yellow color over a period of less than 1 h. It is possible that this chromophore derives either from histidyl-Fe(III) charge-transfer transitions or, based on separate studies,[2] from a tyrosyl radical (λ_{max} = 408 nm) generated from a protein side chain. When the oxygen level was allowed to exceed 1 ppm during setup of the crystallization, a green color was observed and subsequent crystal growth was poor. Again, based on separate studies,[2] it is likely that the green chromophore derives from a catecholate-Fe(III) species generated by self-hydroxylation reactions of the enzyme.

Structure Refinement. The structure of TauD complexed with the essential cofactor, ferrous iron, and two substrates, αKG and taurine, has been solved at 3.0 Å resolution (Figure 1). The initial electron density map following density modification was of good quality for residues[3] 94−150 and 180−280, and for the two additional alpha helices present in residues 1−93. A model was constructed of those regions that were well defined in the electron density map and with polyalanine traces through the other visible secondary structure elements. This model produced stable refinement by simulated annealing, and the resulting improved phases allowed the majority of the protein chain to be modeled, accounting for 277 of the 283 residues. There are two molecules in the asymmetric unit, with few significant differences between them. Chain A has an average *B*-factor of 49.7, while chain B has an average *B*-factor of 77.0. Residues 164−168, present in chain A, were not well defined in chain B and were omitted from the model for chain B. The side chains of 13 residues in chain A (including residues 165−169) and 14 in chain B were not well defined, and these residues were modeled as alanine. All of these residues are in loops on the surface of the protein, and none are part of the structural jellyroll motif or active site environment. According to Procheck, there are no Ramachandran outliers.

[3] The numbering schemes used for identifying amino acid residues of TauD and TfdA are based on the gene sequences, rather than the positions in the isolated proteins. The amino terminal methionine residues of both proteins are processed by methionine aminopeptidase. The TauD and TfdA numbering schemes used here differ by the addition of one residue compared to those we have used previously.

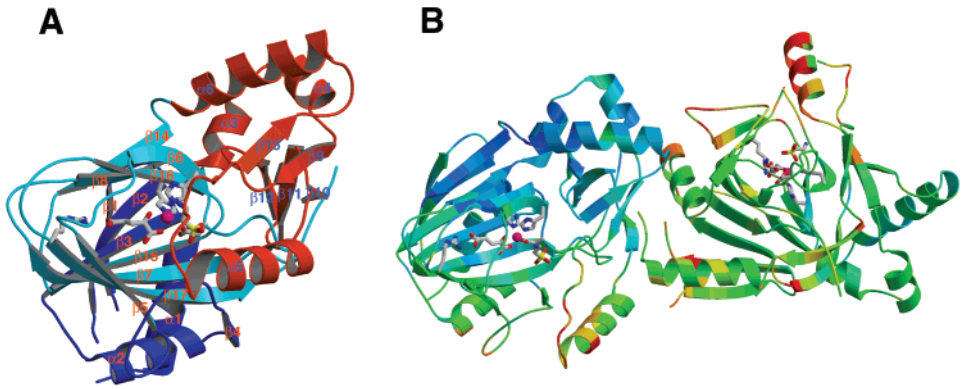


FIGURE 1: Structure of TauD. (A) Ribbon diagram of the TauD subunit with labeled secondary structure. The N-terminal region is in dark-blue, the jellyroll motif is in light-blue, and the extended insert into the jellyroll motif is in red. (B) The two molecules of TauD per asymmetric unit colored by *B*-factor. The Fe(II)- and αKG- binding ligands and substrates are shown. Figures 1−3 were produced using Molscript (*43*), Figures 4 and 6 were produced with Bobscript (*44*), and all pictures were rendered with Raster3D (*45*).

Dimerization Interface. The two TauD molecules per asymmetric unit were not related by simple 2-fold rotation and are unlikely to account for the observed dimeric nature of this enzyme (*5*). Rather, we propose that a dimer formed by a crystallographic 2-fold rotation represents the biological dimer (Figure 2). By application of crystallographic symmetry, the same dimeric arrangement can be generated for both molecules in the asymmetric unit. A significant number of hydrogen bonds are formed between the monomers including Thr122 OG1 − Pro12 N, Arg237 NH2 − Gln235 O, Glu233 OE1 − Tyr256 OH, Arg239 NE − Leu10 O, Asn258 OD1 − Glu233 N, Asn258 OD1 − Glu233 OE1, Arg237 NH2 − Pro232 O, Glu233 OE2 − Ala259. The dimer has a buried surface area of 1770 Å^2, which is comparable to an average value of 2000 Å^2 for similar sized proteins (*32*). Furthermore, the number of hydrogen bonds formed between the monomers is comparable to that found in other dimers with a similar buried surface area (*32*). The residues at the contact area are in regions of lower average *B*-factor, as expected for residues in the interior of a multimer. The dimer contact area is on the side of the protein opposite to the active site, and comprises mainly residues of α6 and β14 and the connecting loop between them, with additional interactions involving β1 and the loop between β7 and β8.

Topology. The main chain of the TauD subunit contains 17 strands of β-sheet, eight of which form a jellyroll motif (β5−8 and β14−17) (Figure 1). The β-sheets form the core of the structure, around which are arranged six α-helices. Flexible regions, indicated by higher *B*-factors, are found in the external loops 74−80 (β4-β5), 18−26 (β2-α1), and the extended loop 156−183. The TauD structure shows a high degree of similarity to that of CAS (*13*) (Figure 3), despite the low extent of sequence identity between these proteins (7.7% when compared over the length of alignment). The similarity is most clear in the β-sheet core and helices 3, 4, and 6, which are adjacent to the jellyroll motif β-strands in the primary sequence. Overall, 48% of the protein backbones of TauD and CAS superimpose with a rms deviation of 1.64 Å. The most significant difference in the overall structures is in the extended loop between strands β9 and β10 (Arg156-Pro183), which extends over the principle substrate-binding pocket. Substantial differences are observed in the loop Glu66−Val79 (between helix 2 and strand 5) which defines part of the taurine-binding pocket. Other residues bordering the substrate-binding pocket also show considerable differences, including the loop between Thr87 and Asn97.

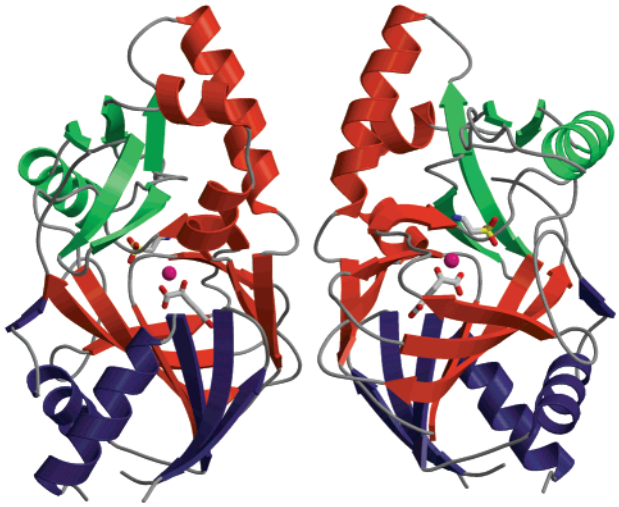


FIGURE 2: Diagram of the proposed biological dimer of TauD.

Iron and αKG binding. The positions of the ferrous iron and two substrates could clearly be seen in the difference electron density map (Figure 4). Although the resolution of the electron density was not sufficient to allow determination of the exact orientations of the substrates, this information could be inferred from the hydrogen bonding possibilities with protein side chains. The pentacoordinate ferrous iron is bound to the protein through a 2 His − 1 carboxylate facial triad (*33*) made up of residues His99, Asp101, and His255. The αKG is bound in a bidentate manner with the C-1 carboxylate and the 2-oxo group coordinated to Fe, while the C-5 carboxylate forms a salt bridge with Arg266 and a hydrogen bond with Thr126. Both of these residues, along with the three metal-binding ligands, are conserved across this subfamily (Figure 5) and are certain to perform the same function in each member.

Taurine Binding. The structure implies potential hydrogen bonds between three amino acid side chains and the taurine amine group. These H-bonds form a distorted tetrahedral environment around the amine and involve the phenolic side chain of Tyr73, the amide oxygen of Asn95, and the hydroxyl of Ser158. The taurine sulfonate interacts with Arg270, His70, and the backbone N−H of Val102. Neutralization of the sulfonate charge is achieved by an ionic bond to Arg270.

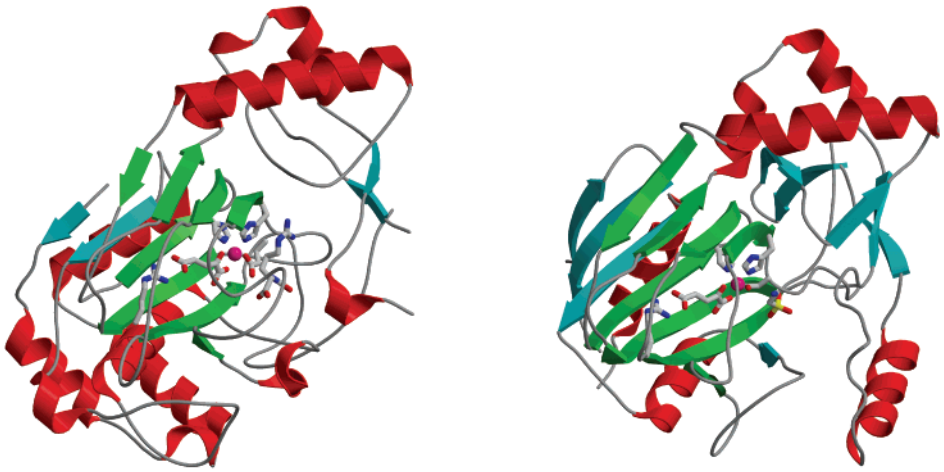


FIGURE 3: Comparison of the structures of TauD (on the right) and CAS1 (PDB ID 1DRY, on the left). The beta-strands of the conserved jellyroll motif are colored separately in green. Similarity of the core structures is apparent, while the main differences involve the extended loops over the active sites.

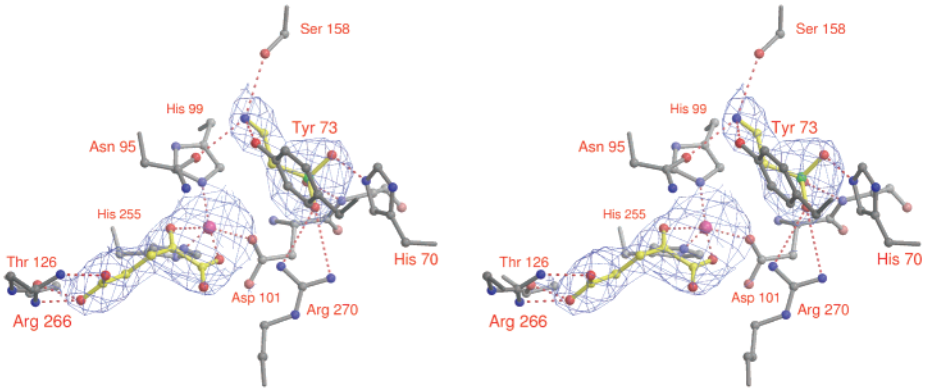


FIGURE 4: Stereodiagram of the active site of TauD. Shown are the protein ligands to the iron, the two substrates and other important protein side-chains. A $2mF_o-DF_c$ electron density map is shown contoured at 1.0 sigma around the substrate molecules.

His70 both interacts with substrate and forms a hydrogen bond with the backbone N−H of Val72; thus, His70 is probably uncharged and is likely to form a hydrogen bond from atom NE2 of the imidazole to the sulfonate group. While Arg270 is partially conserved across a range of the group II dioxygenases (Figure 5), presumably to act as a counterion to the substrate anion, the hydrogen bonds from Val102 and from the side chain of His70 permit TauD to select a tetrahedral sulfonate anion. Consistent with this hypothesis, His70 is conserved within those dioxygenases that metabolize a sulfonate or sulfate substrate (e.g., AtsK (*8*) and YSD (*11*) in Figure 5), but is not found in those αKG dependent dioxygenases that bind substrates containing a carboxylate (for example, CAS, TfdA, or CarC in this figure). These three interactions explain how TauD can select a tetrahedral sulfonate anion in preference to, for example, a carboxylate group. In contrast, the residues interacting with the taurine NH_2 are less well conserved. Tyr73 does not appear to be conserved with other enzymes. Asn95 and Ser158 are conserved with TfdA, but the TfdA equivalents are predicted to occupy alternative positions (see below). In addition to the lack of conservation of the taurine amine ligands, the taurine amine group is relatively solvent exposed, and there is a sizable cavity adjacent to the amine. TauD is known to convert a number of alternative substrates (*5*), including certain much larger sulfonates such as 3-(*N*-morpholino)propanesulfonic acid (MOPS), which can be understood in relation to the available substrate-binding cavity and presumably the higher importance of the sulfonate-binding ligands over those binding the taurine amine group.

Modeling of the TfdA Structure. TfdA is particularly well characterized by biochemical and biophysical methods (*10*, *12*, *31*, *34*−*39*) and therefore is worth modeling to extend our knowledge of the αKG-dioxygenase superfamily. The predicted overall fold of TfdA is very similar to that of TauD, as expected for such closely related proteins (∼30% sequence identity). Analysis of the TfdA model by Procheck gave an acceptable overall average G factor of −0.22. Interactions with the 2,4-D carboxylic acid are proposed to involve Arg278, Lys71, His214, and the backbone amide of Ser117 (Figure 6). Two of these interactions correspond to the taurine sulfonate-binding residues, Arg270 and the N−H of Val102 of TauD. Gly67 in TfdA replaces the third sulfonate-binding residue, His70, whereas TfdA residues His214 and Lys71 do not have similar counterparts in TauD. In addition to the potential interaction with the carboxylate, TfdA Lys71 (conserved with the taurine amine ligand Tyr73 in TauD) could reasonably interact with the ether oxygen of 2,4-D. Additional interactions with the substrate are certain to involve residues 86−111 which could not be convincingly modeled. However, some potential models suggested that Lys95 could orient to hydrogen bond with the ether oxygen of 2,4-D. The remaining residues analogous to the hydrogen

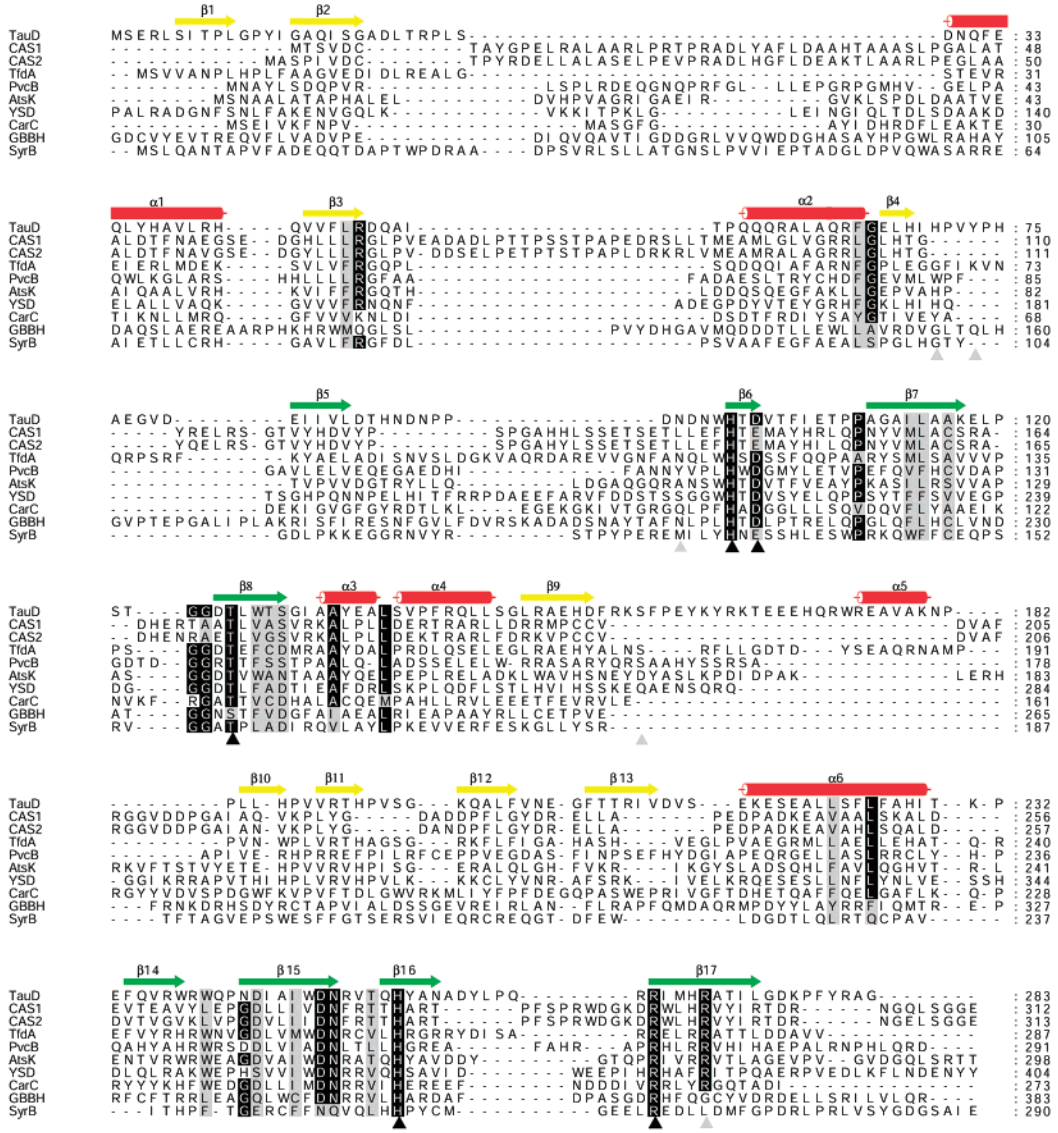


FIGURE 5: Secondary structure based sequence alignment of the group II (*10*) αKG-dependent dioxygenases. An initial alignment was produced based on a 3D alignment of CAS1 (PDB ID: 1DRY) and TauD. The other sequences were then aligned to the existing alignment. The sequences (with Swiss-Prot accession numbers) are TauD, *E. coli* taurine/αKG dioxygenase (P37610); CAS1, CAS2, *Streptomyces clavuligerus* clavaminate synthase 1 and 2 (Q05581 and Q05582); TfdA, *Ralstonia eutropha* 2,4-D/αKG dioxygenase (P10088); PvcB, *Pseudomonas aeruginosa* pyoverdine biosynthesis enzyme (O30371); AtsK, *Pseudomonas putida* alkyl sulfate/αKG dioxygenase (Q9WWU5); YSD, *Saccharomyces cerevisiae* sulfonate/αKG dioxygenase (Q12358); CarC, *Erwinia carotovora* carbapenam biosynthesis enzyme (Q9XB59); GBBH, *Pseudomonas sp.* γ-butyrobetaine hydroxylase (P80193); SyrB, *Pseudomonas syringae* syringomycin biosynthesis enzyme I (P97053). The TauD secondary structure is colored red for helix, green for β-strands of the jellyroll motif, and yellow for other β-strands. Conserved Fe(II)- and αKG-binding ligands are marked with black triangles. Taurine-binding ligands within TauD are marked below the alignments with gray triangles. Produced using Alscript (*46*).

bond partners of the taurine amine in TauD (e.g., Ser173) are predicted not to interact with the 2,4-D ether.

The TfdA model is consistent with previous observations and predictions made about TfdA. For example, Arg78 is surface exposed in the model and accounts for the reported protease sensitivity of this site (*12*). In addition, the model supports previous predictions based on site-directed mutagenesis and spectroscopic analyses that His114, Asp116, and His263 bind iron (*10*). The model vindicates a prior suggestion that Arg274 forms a salt bridge with the C-5 carboxylate of αKG (*10*). Thr141 is positioned to hydrogen bond this moiety of αKG, consistent with conservation of a hydroxylated amino acid at that position in group II dioxygenases. The model is in agreement with previous results suggesting that His214 functions in 2,4-D binding or catalysis

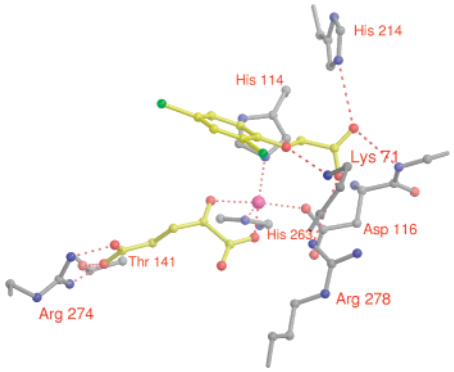
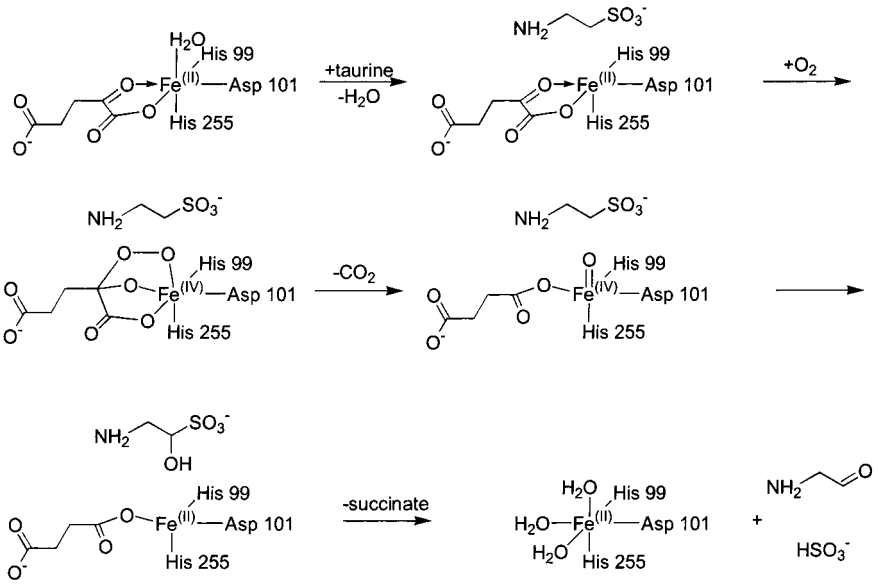


FIGURE 6: Model of the TfdA active site. The predicted active site of TfdA is shown with protein ligands to the iron and residues that stabilize binding of the two substrates, αKG and 2,4-D.

Scheme 2: Proposed Mechanism of TauD



(*10*). Specifically, His214 interaction with a carboxylate oxygen of 2,4-D accounts for the 10-fold increase in K_m of 2,4-D in a H214A mutant. Although substitution of His217 by Ala was also shown to lead to a 2.5-fold increase in K_m of 2,4-D (*10*), this residue is more distant from the active site and is predicted neither to bind substrate directly nor participate in catalysis. The observed effects may alternatively arise from reduced substrate access to the active site or to altered positioning of a 2,4-D ligand. Finally, evidence for an interaction with the substrate ether atom, presumed to involve Lys71, but possibly also Lys95, is provided by studies involving the substrate 2,4-dichlorocinnamic acid (*38*). This compound closely resembles 2,4-D in structure, but lacks the ether bond leading to a much higher K_m.

Sites of Post-translational Modifications in TauD and TfdA. The availability of the TauD structure and TfdA model allowed us to examine the locations of post-translational modifications that are known to occur within the enzymes. Mass spectrometric methods have revealed that upon exposure of the αKG-bound enzymes to oxygen Trp128, Trp240, and Trp248 in TauD[2] and Trp113 in TfdA (*20*) become hydroxylated. These modifications would not be expected to be present within the anaerobically prepared TauD crystals, but the low resolution precludes the use of structural data to confirm the absence of hydroxylation of these residues. The mechanism by which these modifications are introduced is unclear. Although the three TauD Trp residues all lie within 10.5 Å of the metallocenter (Trp128 ∼ 7.5 Å, Trp240 ∼ 10.5 Å, Trp248 ∼ 6.0 Å), these residues are positioned close to His255, trans to the proposed oxygen binding site and the taurine binding cavity. These residues comprise a tight hydrophobic cluster in TauD that is conservatively substituted with alternative hydrophobic residues in other family members (Figure 5). The hydroxylated Trp residue in TfdA occurs adjacent to the His114 ligand, approximately 10.5 Å from the metal center. TauD possesses Trp at the corresponding position (Trp98), yet hydroxylation of this residue has not been observed in the taurine-degrading enzyme.

In addition to the hydroxy-Trp modifications, a catecholate species was detected in oxygen-exposed solutions containing TauD, Fe(II), and αKG.[2] Mass spectrometric analysis has tentatively identified this unique modification as arising from Tyr256. The aromatic ring of Tyr256 lies 9 Å from the metal center, on the opposite side of β-strand 15. The modified Tyr256 is suspected to form a chromophore involving the oxidized metal ion,[2] implying that the position of the metal also moves in the modified form of the enzyme. Efforts to obtain crystals of cross-linked protein containing this green chromophore were unsuccessful. A transient tyrosyl radical also has been detected in TauD.[2] The location of this radical has not been determined, but it is reasonable to suggest Tyr256 as the site for radical formation during synthesis of the modification described above. Alternatively, the taurine-binding ligand Tyr73 is located only 7 Å from the metallocenter and is another potential site for radical formation. Structural studies are unable to address this point, and alternative approaches will be needed to identify the location of the radical species.

Reaction Mechanism. In Scheme 2 we propose a mechanism for the reaction catalyzed by TauD that is also applicable to TfdA. The orientation of αKG binding relative to the 2 His − 1 carboxylate facial triad is analogous to that found in CAS (*13*), and suggests that oxygen binds opposite the second His in the motif. This arrangement differs from that in DAOCS where oxygen is thought to bind opposite the first His (*15*). Spectroscopic measurements are consistent with a shift from 6-coordinate to 5-coordinate geometry around the TauD and TfdA metal sites promoted by the binding of taurine or 2,4-D, respectively (*21*, *37*, *40*). Substrate binding thus creates an oxygen-binding site, with bound oxygen ideally positioned for insertion into the appropriate C−H bonds of the respective substrates. This situation is clearly analogous to the hydroxylation reaction performed by CAS (*13*). Such reactions of non-heme ferrous iron enzymes are generally suggested to proceed through an Fe(IV)═O intermediate (*13*, *41*, *42*), as shown in Scheme 2. In the absence of substrate, this intermediate or some other activated oxygen species decays to generate hydroxylated aromatic residues within the TfdA (*20*) or TauD[2] proteins. The TauD structure and TfdA model provide important information about the roles of residues surrounding the metal center that function in substrate binding and catalysis. These

studies set the stage for more detailed analysis of the TauD and TfdA enzyme mechanisms using other mutagenic, biochemical, and biophysical techniques.

ACKNOWLEDGMENT

We thank Karl Harlos for much advice on the structure solution, Minakshi Ghosh for assistance with the X-ray machine, Raghavakaimal Padmakumar for providing initial protein samples, and Leslie Kuhn for help with TfdA modeling.

REFERENCES

1. Kertesz, M. A. (2000) *FEMS Microbiol. Rev. 24*, 135−75.
2. van der Ploeg, J. R., Weiss, M. A., Saller, E., Nashimoto, H., Saito, N., Kertesz, M. A., and Leisinger, T. (1996) *J. Bacteriol. 178*, 5438−46.
3. van der Ploeg, J. R., Iwanicka-Nowicka, R., Bykowsky, T., Hryniewicz, M. M., and Leisinger, T. (1999) *J. Biol. Chem. 274*, 29358−29365.
4. Eichhorn, E., van der Ploeg, J. R., and Leisinger, T. (2000) *J. Bacteriol. 182*, 2687−95.
5. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) *J. Biol. Chem. 272*, 23031−6.
6. Eichhorn, E., van der Ploeg, J. R., and Leisinger, T. (1999) *J. Biol. Chem. 274*, 26639−46.
7. Kahnert, A., Vermeij, P., Wietek, C., James, P., Leisinger, T., and Kertesz, M. A. (2000) *J. Bacteriol. 182*, 2869−78.
8. Kahnert, A., and Kertesz, M. A. (2000) *J. Biol. Chem. 275*, 31661−7.
9. Prescott, A. G., and Lloyd, M. D. (2000) *Nat. Prod. Rep. 17*, 367−83.
10. Hogan, D. A., Smith, S. R., Saari, E. A., McCracken, J., and Hausinger, R. P. (2000) *J. Biol. Chem. 275*, 12400−9.
11. Hogan, D. A., Auchtung, T. A., and Hausinger, R. P. (1999) *J. Bacteriol. 181*, 5876−9.
12. Fukumori, F., and Hausinger, R. P. (1993) *J. Biol. Chem. 268*, 24311−7.
13. Zhang, Z., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2000) *Nat. Struct. Biol. 7*, 127−33.
14. Ruetschi, U., Nordin, I., Odelhog, B., Jornvall, H., and Lindstedt, S. (1993) *Eur. J. Biochem. 213*, 1075−80.
15. Valegård, K., van Scheltinga, A. C., Lloyd, M. D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H. J., Baldwin, J. E., Schofield, C. J., Hajdu, J., and Andersson, I. (1998) *Nature 394*, 805−9.
16. Clifton, I. J., Hsueh, L.-C., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2001) *Eur. J. Biochem. 268*, 6625−6636.
17. Wilmouth, R. C., Turnbull, J. J., Welford, R. W. D., Clifton, I. J., Prescott, A., and Schofield, C. J. (2002) *Structure 10*, 93−103.
18. Roach, P. L., Clifton, I. J., Fülöp, V., Harlos, K., Barton, G. J., Hajdu, J., Andersson, I., Schofield, C. J., and Baldwin, J. E. (1995) *Nature 375*, 700−4.
19. Roach, P. L., Clifton, I. J., Hensgens, C. M., Shibata, N., Schofield, C. J., Hajdu, J., and Baldwin, J. E. (1997) *Nature 387*, 827−30.
20. Liu, A., Ho, R. Y. N., Que, L., Jr., Ryle, M. J., Phinney, B. S., and Hausinger, R. P. (2001) *J. Am. Chem. Soc. 123*, 5126−5127.
21. Ryle, M. J., Padmakumar, R., and Hausinger, R. P. (1999) *Biochemistry 38*, 15278−86.
22. Leslie, A. G. W. (1999) *Acta Crystallogr. D55*, 1696−1702.
23. Collaborative Computational Project Number 4. (1994) *Acta Crystallogr. D50*, 760−763.
24. 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., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
25. De La Fortelle, E., and Bricogne, G. (1997) in *Methods in Enzymology, Macromolecular Crystallography* (Sweet, R. M., Carter, C. W., Jr., Eds.) pp 472−494, Academic Press, New York.
26. Cowtan, K. (1994) in *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, pp 34−38.
27. Kleywegt, G. J., and Jones, T. A. (1996) *Acta Crystallogr. D52*, 826−828.
28. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
29. Rost, B., and Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.
30. Sali, A., and Blundell, T. L. (1993) *J. Mol. Biol. 234*, 779−815.
31. Saari, R. E., Hogan, D. A., and Hausinger, R. P. (1999) *J. Mol. Catal. B-Enzym. 6*, 421−428.
32. Jones, S., and Thornton, J. M. (1995) *Prog. Biophys. Mol. Biol. 63*, 31-&.
33. Que, L., Jr. (2000) *Nat. Struct. Biol. 7*, 182−184.
34. Fukumori, F., and Hausinger, R. P. (1993) *J. Bacteriol. 175*, 2083−6.
35. Saari, R. E., and Hausinger, R. P. (1998) *Biochemistry 37*, 3035−42.
36. Cosper, N. J., Stalhandske, C. M., Saari, R. E., Hausinger, R. P., and Scott, R. A. (1999) *J. Biol. Inorg. Chem. 4*, 122−9.
37. Hegg, E. L., Whiting, A. K., Saari, R. E., McCracken, J., Hausinger, R. P., and Que, L., Jr. (1999) *Biochemistry 38*, 16714−26.
38. Dunning Hotopp, J. C., and Hausinger, R. P. (2001) *J. Mol. Catal. B-Enzym. 15*, 155−162.
39. Whiting, A. K., Que, L., Jr., Saari, R. E., Hausinger, R. P., Fredrick, M. A., and McCracken, J. (1997) *J. Am. Chem. Soc. 119*, 3413−3414.
40. Ho, R. Y. N., Mehn, M. P., Hegg, E. L., Liu, A., Ryle, M. J., Hausinger, R. P., and Que, L., Jr. (2001) *J. Am. Chem. Soc. 123*, 5022−5029.
41. Burzlaff, N. I., Rutledge, P. J., Clifton, I. J., Hensgens, C. M., Pickford, M., Adlington, R. M., Roach, P. L., and Baldwin, J. E. (1999) *Nature 401*, 721−4.
42. 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) *Chem. Rev. 100*, 235−349.
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
44. Esnouf, R. M. (1997) *J. Mol. Graph. Model. 15*, 132−134.
45. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.
46. Barton, G. J. (1993) *Protein Eng. 6*, 37−40.

BI016014E