Biochemistry

 Volume 42, Number 19 May 20, 2003

Accelerated Publications

Substrate-Induced Conformational Changes in *Escherichia coli* Taurine/α-Ketoglutarate Dioxygenase and Insight into the Oligomeric Structure‡

Jessica R. O'Brien,§ David J. Schuller,|| Victoria S. Yang,§ Bret D. Dillard,§ and William N. Lanzilotta*,§

Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853, and Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia 30602

Received January 20, 2003; Revised Manuscript Received March 25, 2003

ABSTRACT: The enzymes in the α-ketoglutarate (αKG) dependent dioxygenase superfamily represent the largest class of non-heme iron oxidases and have important medical, ecological, and biotechnological roles. One such enzyme, taurine/α-ketoglutarate dioxygenase (TauD), catalyzes the conversion of 2-aminoethanesulfonate (taurine) to sulfite and aminoacetaldehyde while decomposing αKG to succinate and CO_2. This αKG dependent dioxygenase is expressed in *Escherichia coli* under sulfur starvation conditions and allows the cell to utilize taurine, and other similar sulfonates in the environment, as an alternative sulfur source. In this work, we report the structures of the apo and holo forms of TauD to 1.9 Å resolution (R_{cryst} = 21.2%, R_{free} = 24.9%) and 2.5 Å resolution (R_{cryst} = 22.5%, R_{free} = 27.8%), respectively. The models reported herein provide significant new insight into the substrate orientations at the active site and the conformational changes that are induced upon taurine binding. Furthermore, analysis of our crystallographic data coupled with reanalysis of the crystallographic model (resolution = 3.0 Å, R_{cryst} = 28.1, R_{free} = 32.0) presented by Elkins et al. (*Biochemistry* (2002) *41*, 5185−5192) reveals an alternative oligomeric arrangement for the enzyme that is consistent with the conserved primary and secondary structure elements of other αKG dependent dioxygenases.

Oxygenases and oxidases that contain iron as the sole cofactor constitute a large family of redox enzymes that utilize either di-iron centers or mono-iron centers for catalysis. Mono-nuclear iron enzymes can be iron(III)-dependent, such as lipoxygenase or the intradiol cleaving catechol dioxygenase, and iron(II)-dependent, such as the extradiol cleaving catechol dioxygenases and the αKG[1] dependent oxygenases. The members of the latter family, the αKG dependent oxygenases, represent the largest class of non-heme iron oxidases. These enzymes require ferrous iron and αKG as a cosubstrate to catalyze biological oxidations of otherwise inactivated carbon−hydrogen bonds (*1*, *2*). Like their heme counterparts, these oxygenases perform a variety of essential reactions, including hydroxylations, desaturations, and oxidative ring closures in the biosynthesis of antibiotics, modified amino acids/peptides, carnitine, and signaling molecules. Notably, in higher organisms these enzymes catalyze the posttranslational hydroxylation of

* Address correspondence to the following author. Phone: (706) 542-1324. Fax: (706) 542-1738. E-mail: wlanzilo@bmb.uga.edu.

‡ Atomic coordinates have been deposited in the Protein Data Bank under codes 1OS7 (Apo model) and 1OTJ (Holo model).

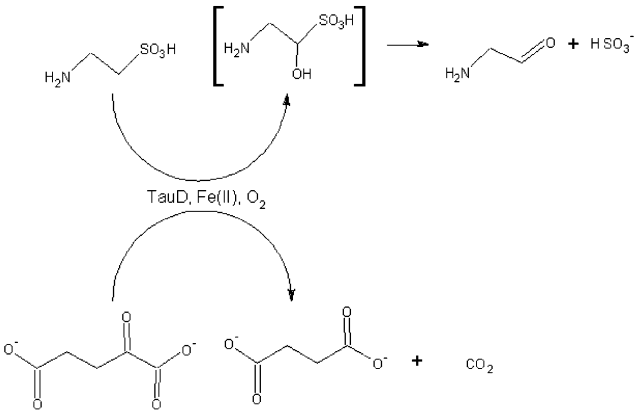
§ University of Georgia.

|| Cornell University.

[1] Abbreviations: αKG, α-ketoglutarate; TauD, taurine/αKG dioxygenase; TRIS, tris(hydroxymethyl)aminomethane; HEPES, (*N*-[2-hydroxyethyl]piperazine-*N*′-[4-butanesulfonic acid]); MAD, multiwavelength with anomalous dispersion; CAS, clavaminic acid synthase.

10.1021/bi0341096 CCC: $25.00
Published on Web 04/19/2003

Scheme 1



proline residues during collagen biosynthesis and in the hypoxia inducible factor (HIF)-mediated pathway for dioxygen sensing (*3*−*5*).

Sequence analysis of many members within αKG dependent superfamily has revealed a conserved His-X-Asp-$X_{(\sim 55)}$-His sequence (*6*) that forms a 2-His-carboxylate facial triad at the active site of the enzyme (*7*−*9*). The three amino acids that form the 2-His-carboxylate facial triad represent the only protein ligands to a bound iron atom. Previous studies have shown that in the absence of αKG and substrate, solvent molecules occupy the remaining three coordination sites of the 6-coordinate iron (*7*, *10*). Two of these solvent ligands are displaced when αKG binds to the ferrous center through its C-1 carboxylate oxygen and C-2 carbonyl oxygen (*7*, *10*).

The overall reaction catalyzed by *Escherichia coli* taurine/α-ketoglutarate dioxygenase (TauD) is summarized in Scheme 1. TauD is a member of the group II αKG dependent dioxygenases, and a crystal structure of TauD with Fe(III), taurine, and αKG bound at the active site has been reported to 3.0 Å resolution (R_{cryst} = 28.1%, R_{free} = 32.0%) (*11*). Although the exact orientations of the substrates could not be determined, it was clear from this work that the 2-His-carboxylate facial triad architecture was preserved in TauD. Similar to what has been proposed for other αKG dependent dioxygenases, this structural work combined with spectroscopic measurements suggested that the binding of taurine displaces a solvent molecule resulting in a shift from 6-coordinate to 5-coordinate geometry at the active site and creation of an oxygen binding site (*11*−*15*). In the absence of taurine, oxygen can still react with the αKG- and Fe(II)-bound form of TauD and leads to hydroxylation at tryptophan residues 128, 240, and 248 (*16*). At the present time, the mechanism by which these modifications are made and their physiological role remains unclear.

Biochemical analysis suggests that TauD functions as a dimer (*17*), and a model for the physiological dimer was proposed based on the crystallographic data (*11*). Interestingly, the proposed oligomeric structure was inconsistent with some of the conserved primary and secondary structural elements found throughout the group II αKG-dependent dioxygenase family of iron dioxygenases. Specifically, all group II αKG-dependent dioxygenase, which have been structurally characterized thus far, contain a jelly roll motif around the active site as well as a pair of structurally conserved α-helices (*11*). For TauD and clavaminate synthase (CAS), these are helices 4 and 6. The amino acids that make up these helices are composed of primarily conserved hydrophobic residues, and when the jelly roll motifs are aligned, these two helices also align. A similar pair of structurally conserved helices has also been observed in the structure of proline 3-hydroxylase and has been proposed to be the site of dimer formation (*18*).

To provide further insight into the substrate interactions and the oligomeric nature of TauD, we have determined the crystal structure of apo and holo TauD to 1.9 and 2.5 Å resolution, respectively. In both cases, the substrate taurine was an absolute requirement for crystallization. The apo structure therefore represents crystals prepared in the presence of taurine, while the holo structure contains taurine, αKG, and ferrous iron bound at the active site. Because of the crystallographic symmetry, four copies of the TauD monomer are present in the asymmetric unit. In both the holo and the apo structures, three of the monomers contain taurine bound at the active site, while the fourth does not. Significant conformational changes are observed in the monomer without taurine bound. In both the apo and the holo data, the exact orientations of the substrate molecule can be observed in the monomers with taurine bound. Taken together, the data presented here provides an improved understanding of the substrate interactions within the active site and insight into the conformational changes associated with substrate binding. In addition, evidence for an alternative dimeric arrangement for TauD is presented that is consistent with the conserved sequence and structural elements found in other αKG-dependent enzymes.

EXPERIMENTAL PROCEDURES

Enzyme Expression, Purification, and Assay. The *tauD* gene was cloned from the *E. coli* genome by using traditional PCR methods. The oligos used for amplification of the *tauD* gene were designed such that the correct flanking sequences required for subcloning into the pET-TOPO vectors (Invitrogen Corporation, Carlsbad, CA) were incorporated in the PCR product. Using these vectors, gene expression is under the control of the T7 RNA polymerase/promoter system (*19*). One advantage to using this approach is that a 6x His tail can be attached at either the N- or C-terminus for rapid isolation. For optimal expression levels using this system, 5 mL of cells (0.6−1.0 $OD_{600\ nm}$) harboring the expression plasmid were inoculated into 1 L of TB media and grown for 12 h at 34 °C. The cells were then induced by addition of IPTG (1 mM final) and grown an additional 4 h before being harvested. Purification of the tagged TauD was performed in a single step using chelating sepharose (Amersham Pharmacia, Piscataway, NJ) as previously described (*20*), except that salt was omitted from all elution buffers. Typical yields were between 40 and 60 mg of pure protein from 1 L of cell material. To ensure that no cationic metals were contaminating the purified protein, purified enzyme was exchanged from 10 mM TRIS buffer pH 8.0 into 100 mM EDTA pH 8.0 and back into the original buffer. The reconstituted enzyme had a specific activity of 14.0 ± 1.5 μmol min^{-1} (mg TauD)$^{-1}$ at pH 7.0 when assayed by the method described by Eichhorn et al. (*17*).

Crystallization. Crystals were grown by batch method in an anaerobic environment of 95% nitrogen and 5% hydrogen gas. The temperature of the glovebox was 22 °C, and the

oxygen level was maintained below 1 ppm at all times. All solutions for crystallization were made anaerobic by several rounds of vacuum degassing and flushing with argon using a vacuum manifold. The protein-to-precipitant ratio was 2:3 in the batch experiment, and the initial protein concentration was 30 mg/mL. Because of protein requirement for crystallization and the ease of purification, we decided to use TauD that was tagged at the N-terminus. Diffraction quality crystals typically appeared within two weeks after the conditions were set up. Crystals were grown using the TauD protein as isolated and a precipitating solution consisting of 0.2 M di-sodium tartrate dihydrate, 5 mM taurine, and 20% PEG 3350. These crystals represent the apo crystals used in this study. Interestingly, once these crystals were formed, they could be transferred to mother liquor containing 0.1 M TRIS-HCl pH 7.5 in place of di-sodium tartrate dehydrate. This could only be accomplished if iron ($FeSO_4$), ascorbate, and αKG were also present at 10 mM. These crystals represent the holo crystals used in this study. To prepare both the apo and the holo crystals for freezing, the PEG 3350 was replaced with PEG 4000, and the concentration was increased to 30% in 2.5% increments.

Data Collection and Structure Determination. Data for the apo-crystals was collected at the Cornell High Energy Sychrotron Source (CHESS) on beam line A1. Data were collected for the iron-treated crystals at the University of Georgia on a Rigaku RU-200 rotating anode equipped with Osmic focusing mirrors and a R-axisIIc image plate detector. Data were processed using MOSFLM (*21*) and the CCP4 suite of programs (*22*). Our initial approach was to treat data collected from the apo crystal as native data and data from the iron-treated crystal as a derivative. Anomalous difference Patterson methods indicated four weak signals in the asymmetric unit. Prior to applying for beam time to collect MAD data, a 3.0 Å structure (R_{cryst} = 28.1%, R_{free} = 32.0%) of TauD was reported by Elkins et al. (*11*). Using a poly-alanine version of the momomer reported by Elkins et al., without any substrate or cofactors, molecular replacement was performed using CNS (*23*) and gave a good solution that was consistent with four monomers in the asymmetric unit. Subsequent rounds of model building and refinement were performed using the programs O (*24*) and CNS (*23*). Protein atoms were placed prior to modeling the water molecules, and no atoms or molecules were treated as rigid groups during refinement. All figures were generated using the programs Molscript (*25*), Xtalview (*26*), and Raster3D (*27*).

RESULTS AND DISCUSSION

Crystal Growth and Structure Refinement. Consistent with Elkins et al. (*11*), we observed that the exposure of iron-treated crystals to oxygen had a negative impact on diffraction (exposure during crystallization or post-crystallization). Therefore, prior to freezing the crystals for data collection, sodium dithionite was added to a final concentration of 2 mM. A crystal, in mother liquor containing dithionite, was then removed from the glovebox and immediately frozen. We found that taurine, αKG, and molecules with structure similar to αKG (e.g., tartrate or citrate) could also be used in crystal growth. In addition, once crystals were formed, they could be transferred to mother liquor in which the original buffer had been replaced with a variety of buffers including 50 mM imidazole (pH 6.5−7.0), 50 mM HEPES pH 7.5, or 50 mM TRIS (pH 7.5−8.5). The only essential requirement for preserving the crystal quality appeared to be taurine and the precipitant (Either PEG 3350 or PEG 4000). Taken together, these data indicate that conformational changes are occurring in TauD upon substrate binding or during turnover.

The structure of TauD with 5 mM taurine bound (apo TauD) is described to 1.9 Å resolution. In addition, the structure of TauD with αKG, Fe(II), and taurine bound (holo TauD) is described to 2.5 Å resolution. In both cases, the crystals belong to the space group $P2_12_12_1$ with similar unit cell dimensions (See Table 1). There are four monomers in the asymmetric unit. In all models, the main chain atoms were clearly visible in their respective $2F_o - F_c$ composite omit maps for residues 3−282. The 6x poly-histidine tail was not observed in any of the models. All side chain atoms could be placed with the exception of residues 168−170 in the D monomer. In this case, the residues were modeled as alanine.

The overall fold of the TauD monomer forms an $\alpha-\beta$ structure with a jelly roll motif at the core of the structure. Despite the low sequence identity (7.7%) the TauD monomer shows a high degree of structural similarity to clavaminic synthase (CAS) (*28*). Forty-eight percent of the backbone atoms align with a rms deviation of 1.64 Å (*11*). Not too surprisingly, the region of greatest alignment is centered on the jelly roll motif, where the active site typically resides in this superfamily. Interestingly, the other region of structural alignment is found around helices 3, 4, and 6.

In both the apo and the holo models, substrate is bound in only three of the four monomers. If taurine was modeled in the active site of the fourth monomer, both R_{cryst} and R_{free} would increase by 0.5%. Because of this observation and the high B-factors of residues covering the active site, we suspected that the fourth monomer might be in multiple conformations because of partial occupancy of taurine in the active site. To test this, we modeled monomer D with taurine having an occupancy of 0.5 and the residues that cover the active site in both an open and a closed conformation. Like taurine, each of the conformations was assigned an occupancy of 0.5. This approach resulted in an increase in both R_{cryst} and R_{free} by ∼2.5%. These results indicate that the structure that best fits the data is a crystallographic tetramer with three monomers in a closed conformation, while the fourth monomer is in an open conformation. The nature of the open and closed conformations is discussed in the sections that follow.

Interestingly, if all four monomers were in the closed conformation, then the crystallographic tetramer would exhibit 222 symmetry. Considering that two of the unit cell edges are very similar (Table 1) and the observation that all angles are 90°, this explains why our crystals were initially indexed in a tetragonal space group. However, integration and scaling of the data using *P*4 resulted in unreasonable values for R_{sym}. Subsequently, all data reduction and refinement was performed with the orthorhombic space group *P*222. Further analysis of the systematic absences suggested that the correct space group was $P2_12_12_1$. This space group was used for all subsequent refinement for both the apo and the holo data, and the same set of reflections were used in both cases for calculating R_{cryst} and R_{free}. The structure reported by Elkins et al. described crystals belonging to the hexagonal space group $P6_222$ with unit cell dimensions a =

Table 1: Data and Refinement Statistics

Data Statistics				
data set	λ (Å)	resolution range	completeness (%)	R_{sym}[a]
apo	0.97	50.0−1.90 (1.97−1.90)[b]	96.9 (78.9)	0.049 (0.334)
holo	1.54	38.0−2.50 (2.59−2.50)	99.9 (98.4)	0.059 (0.431)

Refinement Statistics		
model	apo	holo
unit cell[c]	95.8 117.8 118.2	92.5 118.8 118.8
protein/cofactor non-hydrogen atoms	6810	6850
solvent molecules	706	353
resolution range (Å)	50.0−1.9	36.5−2.5
total reflections	99 459	46 047
total reflections used in R_{free}	5034	2321
R_{cryst}	0.212	0.225
R_{free}	0.254	0.278
rms deviations		
bond distances (Å)	0.005	0.006
bond angles (deg)	1.34	1.35
B-factors (Å^2)		
average	34.5	47.9
minimum	13.3	10.8
maximum	76.9	98.0

[a] $R_{sym} = \sum_{hkl}[\sum_I(|I_{hkl,I} - \langle I_{hkl}\rangle|)]/\sum_{hkl,I}\langle I_{hkl}\rangle$, where I_{hkl} is the intensity of an individual measurement of the reflection with indices *hkl*, and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection. [b] The numbers in parentheses refer to the outer resolution bin used in data processing. [c] The numbers refer to cell edges *a*, *b*, and *c*, respectively. The space group is $P2_12_12_1$ so all angles are 90°.

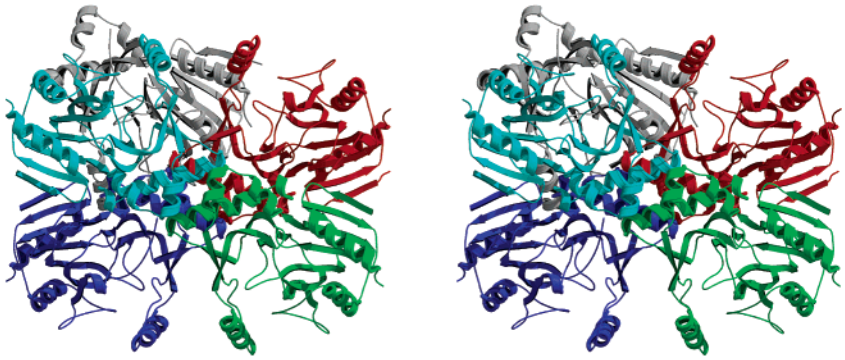


FIGURE 1: Cross-eyed stereoview showing the identification of the asymmetric unit for space group $P2_12_12_1$ relative to the asymmetric unit (space group $P6_222$) presented by Elkins et al. (*11*). The monomers colored in gray and red represent the asymmetric unit for the space group $P6_222$, and the monomers colored in red, green, cyan, and blue represent the asymmetric unit found in the space group $P2_12_12_1$. The gray and red monomers are taken from the coordinates submitted by Elkins et al. (PDB 1GQW), and the green, cyan, and blue monomers were generated using the appropriate symmetry operators for the space group $P6_222$. The biological dimer proposed by Elkins et al. is represented by the red and green monomers.

$b = 117$ and $c = 202$ Å. This was interesting because of the similar cell dimensions observed for our crystals (Table 1) and would suggest a small movement of the unique monomer as well as the tetramer itself could generate two 6-fold symmetry axes. Consistent with this proposal is the observation of a tetramer in the data presented by Elkins et al. that is similar to the tetramer found in our asymmetric unit (RMSD for α-carbon atoms <2.0 Å). This is shown in Figure 1. Figure 1 was generated using the coordinates deposited by Elkins et al. (PDB 1GQW) and shows the position of this tetramer relative to their asymmetric unit. The observation of this tetramer in the hexagonal unit cell as well as in our asymmetric unit is significant because it suggests two possible arrangements for the biological dimer.

Dimerization Interface. The proposed oligomeric arrangement that we feel best explains the dimeric nature of the enzyme (*17*) is shown in Figure 2 and is in contrast to what was previously proposed (*11*). The reason we favor this dimer arrangement stems from several lines of evidence. First, the two helices that form the dimer interface are conserved elements of secondary structure in other αKG-dependent iron

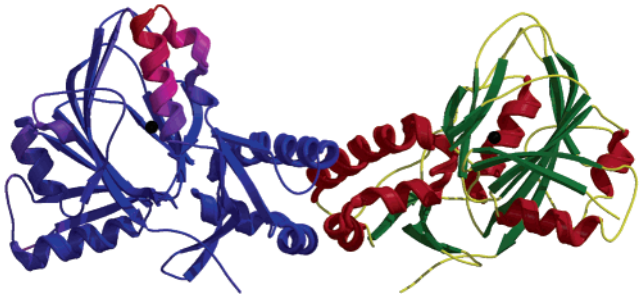


FIGURE 2: Diagram of the proposed biological dimer for TauD. The monomer on the left is colored by B-factor with the color ranging from blue to red and corresponds to a B-factor range from 10 to 95 Å^2. The monomer on the right is colored by secondary structure with helices shown in red, strands in green, and coils in yellow. The figure was generated using the coordinates from monomer A (right side) and monomer D (left side) of the Holo model.

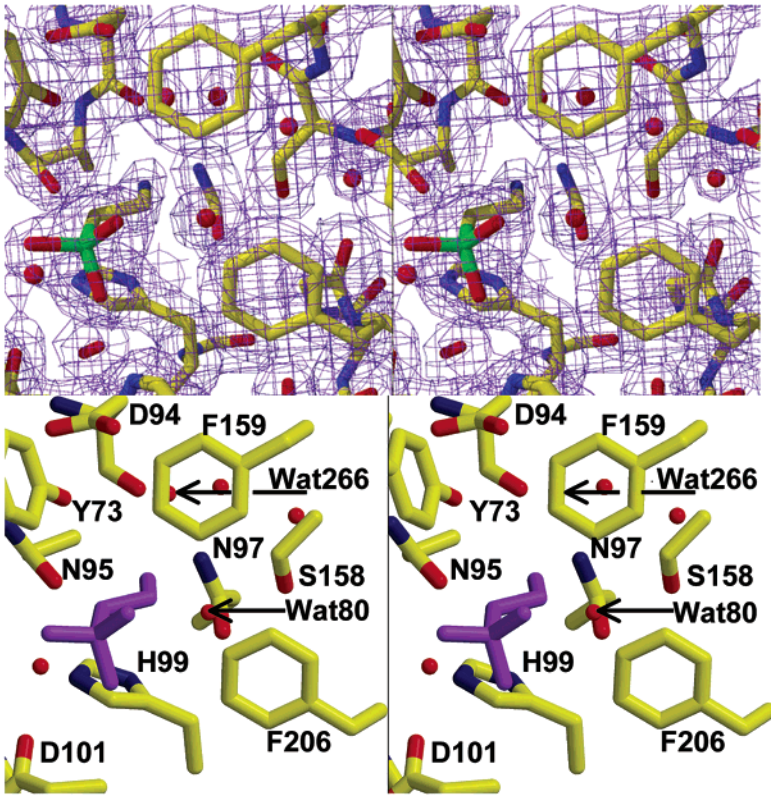


FIGURE 3: Cross-eyed stereoview of the $2F_o - F_c$ composite omit map contoured at 1σ and the corresponding model of the active site region for apo TauD with taurine bound. Top panel: stereoview of the substrate taurine and key residues in the active site of the apo model and the corresponding $2F_o - F_c$ composite omit map (purple cage). The composite omit map was generated using the simulated annealing protocol with 5% of the asymmetric unit being omitted at a time. Bottom panel: stereoview taken from the same perspective as the top panel with taurine colored purple. For additional clarity, the amino acid D94 and the side chain atoms of F159, Y73, N97, N95, S158, H99, F206, as well as D101 are labeled. Five water atoms atoms are also shown, and unless stated otherwise, all atoms are shown in stick format using cpk colors.

dioxygenases (*18*, *29*). Moreover, many of the residues that are found in the two helices at the dimer interface are highly conserved and hydrophobic in nature (*11*). Specifically, four highly amphipathic helices (two from each subunit) form a hydrophobic core that extends back into the hydrophobic center of each monomer. The tertiary structure that comprises our proposed dimer interface can therefore be described as an antiparallel four helical bundle. At the center of the dimer interface are residues L145 and L221 (two from each subunit). An axis of symmetry can be drawn between the two equivalent pairs, and a 2-fold rotation of one monomer around this axis will generate the other. Using the CCP4 program AREAMOL, the calculated buried surface for this dimer is ~400 Å^2, as compared with ~550 Å^2 for the alternative dimer arrangement (*11*). The hydrophobic core at the dimer interface is then covered by the helical structure itself and by a number of hydrogen bonds between the monomers at either end of the helices. It is very clear, especially in the 1.9 Å data, that water is completely excluded from this interface. Extensive work by Fernandez et al. (*30*) indicates that there is significant hydrogen bond desolvation observed in the interior of proteins. These authors have extended their analysis to show that insufficiently dehydrated hydrogen bonds are good determinants of protein interactions (*31*). Consistent with their findings is the observation that formation of our dimer interface allows these hydrophobic helices to dehydrate the hydrogen bonds that are central to the stability of the secondary structure. There are no such interactions at the dimer interface proposed by Elkins et al. (*11*). Additional support for this dimer arrangement comes from the observed B-factors for atoms at our proposed dimer interface. The B-factor can be described as a measurement of thermal motion. In Figure 2 it is clear that the region of the model with the lowest thermal motion is centered on our proposed dimer interface. Finally, a dimeric arrangement that is similar to our proposed dimer has been observed in the high-resolution structures of other αKG-dependent dioxygenases (*18*).

Taurine Binding in the Apo Enzyme. The binding mode of taurine, prior to the addition of σKG and iron, is presented in Figure 3. The focus of Figure 3 is on the amine group of taurine. This is primarily because the taurine sulfonate interactions are similar to what was reported by Elkins et al. Briefly, this includes residues R270, H70, and the backbone N-H of V102. Interactions of the taurine sulfonate clearly support the hypothesis presented by these authors to explain the selection of substrates with tetrahedral sulfonate moieties by TauD.

The taurine amine group is bound in a distorted tetrahedral environment with hydrogen bonds to two water molecules and the amide oxygen of N95. The phenolic oxygen of Y73 extends across the substrate from a loop region to come within 3.0 Å of the carboxylate oxygens of D94. While the phenolic oxygen of Y73 is 3.7 Å away from amine group of the taurine, the interaction of Y73 with D94 positions the phenolic ring within 3.5 Å of the C2 carbon of taurine. This interaction essentially locks the substrate in place, with two additional phenyl side chains (F159 and F206) approaching the C1 and C2 carbon atoms of the taurine molecule from the other side.

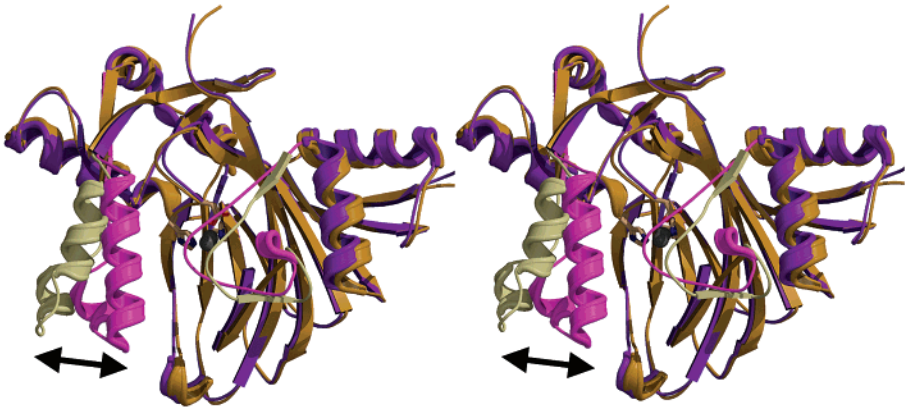


FIGURE 4: Cross-eyed stereoview depicting the difference between the open and the closed conformation of TauD. Overlay of the C and D monomers from the apo model. Monomer C is shown in purple, while monomer D is shown in brown. Regions of movement are highlighted by a lighter shade. The difference in the relative position of helix 5 is further highlighted by the black arrow.

Taurine-Induced Conformational Changes. For both the apo and the holo data sets, the crystallographic tetramer was best modeled with three monomers in a closed conformation with substrate bound. As previously mentioned, the position of taurine was not well-defined in the electron density of the fourth monomer. Given that the position of taurine was not well-defined in the fourth monomer (monomer D) for both the apo and the holo data sets, as well as the lack of any strong interactions between this region of monomer D and other molecules in the unit cell, some conclusions can be made regarding conformational changes that may be due to taurine binding alone. The overall secondary structure for the open conformation is compared with that of the closed conformation in Figure 4. Two regions of secondary structure are seen to undergo dramatic movement. One possible explanation for the conformational change from the closed to the open state is the loss of several key interactions with the substrate taurine. In the region of random coil (Figure 4), both H70 and Y73 have key interactions with taurine. On the opposite side of taurine is a hydrophobic cluster of side chains including F159, Y164, and W174. In the absence of taurine, the hydrogen bonding network in the active site is lost. This action releases the residues 60−80 (random coil region) and exposes a significant portion of these hydrophobic side chains in the region of the helix (see highlighted in helix in Figure 4). Conformational adjustments must be made to shelter these hydrophobic side chains from solvent. This is accomplished in part by a shift from α-helical structure to that of a $^{3-10}$helix in the region from T167 to W174.

Iron and αKG Binding. The binding of iron and αKG has minimal impact on the orientation of the substrate taurine (Figure 5). However, the hydrogen bonding network around the taurine amine group is in contrast to what was reported by Elkins et al. (*11*). In our model, the hydroxyl oxygen of S158 is positioned 2.6 Å from the backbone carbonyl oxygen of G205 and 2.5 Å from a bound water molecule that is interacting with the amine group of taurine. This water molecule is 2.5 Å away from the amide oxygen of N97. It is clear in both apo and holo data sets that water bound near the amino group of taurine provides a short solvent channel to the bulk solvent. Residues in the active site that have hydrogen bond interactions with the ordered water include D94, N97, and S158. Consistent with what was proposed by Elkins et al., we clearly see a pentacoordinate ferrous iron bound to the protein through a 2-His-carboxylate facial triad consisting of residues H99, D101, and H255. The αKG molecule is bound to the iron in a bidentate manner with the opposite carboxylate end being locked into place by the strictly conserved residues R266 (salt bridge) and T126 (hydrogen bond). In the holo structures, all four monomers in the asymmetric unit show good density for the iron and αKG molecules. As was observed in the apo data, the density in the active site of the fourth monomer (D monomer) was consistent with no taurine being bound. In addition, the B-factors for atoms in the active site region of the D monomer were also higher than average suggesting greater thermal motion in the absence of substrate.

Implications for the Mechanism of TauD. All αKG dependent dioxygenase family members demonstrate a common jelly roll fold with three iron ligands found in a common 2-His-carboxylate facial triad motif. While the source of the carboxylate ligand can be either an aspartate or glutamate residue, the relative positions of the histidine and carboxylate side chain further divide these enzymes into three distinct groups. Both TauD and CAS belong to the group II αKG-dioxygenase subfamily because of the additional 138−207 residues that separate the carboxylate ligand from the second histidate ligand (*32*). While TauD and CAS demonstrate low sequence identity (<10%), a remarkable 48% of the protein backbone atoms align with an RMS deviation of 1.64 Å (*11*). In fact, if the residues that form the 2-His-carboxylate facial triad in both enzymes are explicitly aligned (residues 144−146 in the CAS model, and residues 99−101 in the TauD model), then two regions of high homology are immediately clear. The first region of homology is the *β*-strand section, or jelly roll motif, that is important for the active site structure. The second region of alignment is centered on the two helices that we propose are involved in dimer formation.

Given the complex regulation of metal ions in the cytoplasm (*33*), the relative binding constants for substrates (*12*), and the absolute requirement for taurine and molecules with structures similar to αKG for crystal growth, a reasonable hypothesis is that in vivo activity of TauD may ultimately be limited by the availability of iron. Furthermore, given the similar positions of taurine, αKG, and the 2-His-carboxylate facial triad of TauD, as compared to the equivalent positions in CAS, it is reasonable to predict that the oxygen binding mode between these two group II αKG-dependent dioxygenases is similar (*28*). In this mode, oxygen binds opposite the second histidine residue and is positioned

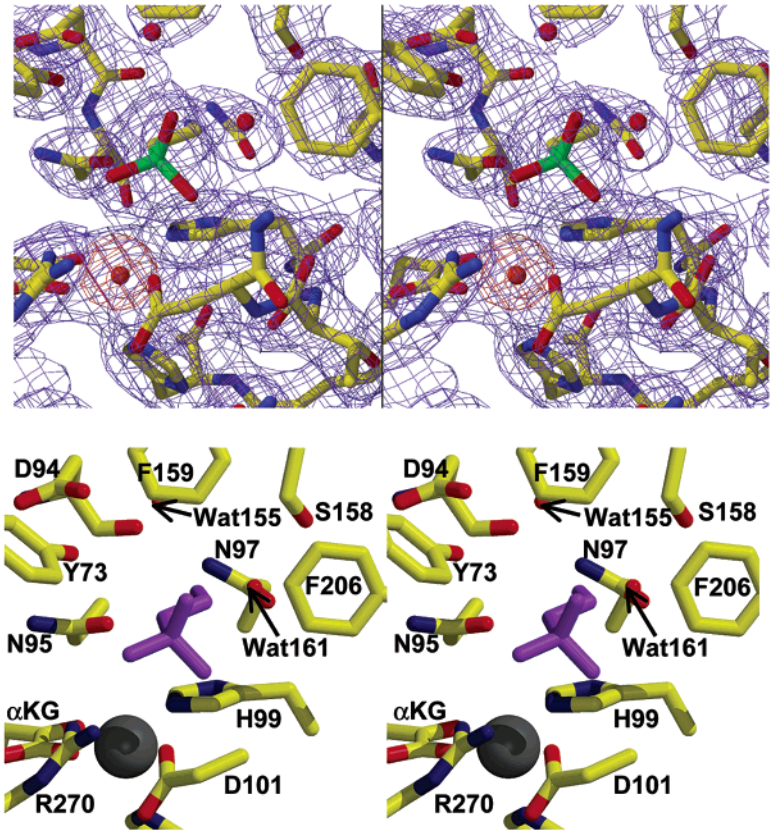


FIGURE 5: Cross-eyed stereoview of the $2F_o - F_c$ composite omit map contour at 1σ and corresponding model of the active site region for holo TauD with taurine bound. Top panel: stereoview of the substrate taurine and key residues in the active site of the holo model and the corresponding $2F_o - F_c$ composite omit map (purple cage). The $2F_o - F_c$ composite omit map was generated using the simulated annealing protocol with 5% of the asymmetric unit being omitted at a time. The anomalous difference map is also shown contoured at 12σ around the iron position (red cage). Bottom panel: stereoview taken from a similar perspective with taurine colored purple. For additional clarity, the amino acid D94 and the side chain atoms of F159, S158, Y73, N97, F206, N95, H99, R270, as well as D101 are labeled. The iron atom is colored black, αKG is labeled, and two water atoms are also shown. Unless stated otherwise, all atoms are shown in stick format using cpk colors.

for insertion into the appropriate C−H bond of αKG. However, recent work using nitric oxide (NO) as an oxygen analogue has shown the potential for Fe−O rearrangements (*29*) at the active site in CAS. At the present time this has not been tested in TauD. In any case, decomposition of αKG has been proposed to occur prior to hydroxylation of taurine (*11*, *12*). It has further been suggested that the hydroxylation reaction will proceed through an Fe(IV)═O intermediate in this family of enzymes (*28*, *34*, *35*). At the present time, there is no crystallographic information for these intermediate states or the product-bound form of TauD.

One final point regarding the structural information presented here must be incorporated into the current model for the mechanism of TauD and the observed properties of the enzyme. In addition to the conformational changes that are revealed in this work, we have also observed significant differences in the positions of side chains near the amino group of the substrate taurine. In contrast to what was reported previously, the hydroxyl oxygen of S158 and the phenolic oxygen of Y73 are not interacting directly with taurine. Alternatively, the hydroxyl side chain of S158 is hydrogen bonded to the backbone carboxylate oxygen of residue 205, and the phenolic oxygen of Y73 is hydrogen bonding with the carboxylate side chain of D94. It is interesting to note that, despite poor sequence and structural alignment in the variable loop regions covering the active site, a similar interaction is observed in CAS between Y149 and D202. While the functional role of these interactions remains to be investigated, the observation of a short solvent channel to the bulk solvent is significant. Specifically, expulsion of the water atoms allows additional room to accommodate larger sulfonates, consistent with the activity of TauD (*17*).

REFERENCES

1. Prescott, A. G., and Philip, J. (1996) *Annu. Rev. Plant Physiol Plant Mol. Biol. 47*, 245−71.
2. Que, L., Jr., and Ho, R. Y. (1996) *Chem. Rev. 96*, 2607−24.
3. Myllyharju, J., and Kivirikko, K. I. (1997) *EMBO J. 16*, 1173−80.
4. Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., Kriegsheim, A., Hebestreit, H. F., Mukherji, M., Schofield, C. J., Maxwell, P. H., Pugh, C. W., and Ratcliffe, P. J. (2001) *Science 292*, 468−72.
5. Epstein, A. C., Gleadle, J. M., McNeill, L. A., Hewitson, K. S., O'Rourke, J., Mole, D. R., Mukherji, M., Metzen, E., Wilson, M. I., Dhanda, A., Tian, Y. M., Masson, N., Hamilton, D. L., Jaakkola, P., Barstead, R., Hodgkin, J., Maxwell, P. H., Pugh, C. W., Schofield, C. J., and Ratcliffe, P. J. (2001) *Cell 107*, 43−54.
6. Borovok, I., Landman, O., Kreisberg-Zakarin, R., Aharonowitz, Y., and Cohen, G. (1996) *Biochemistry 35*, 1981−7.
7. Valegard, 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.
8. 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.
9. Hegg, E. L., and Que, L., Jr. (1997) *Eur. J. Biochem. 250*, 625−9.

10. Pavel, E. G., Zhou, J., Busby, R. W., Gunsior, M., Townsend, C. A., and Solomon, E. I. (1998) *J. Am. Chem. Soc. 120*, 743−753.
11. Elkins, J. M., Ryle, M. J., Clifton, I. J., Dunning Hotopp, J. C., Lloyd, J. S., Burzlaff, N. I., Baldwin, J. E., Hausinger, R. P., and Roach, P. L. (2002) *Biochemistry 41*, 5185−92.
12. Ryle, M. J., Padmakumar, R., and Hausinger, R. P. (1999) *Biochemistry 38*, 15278−86.
13. Hegg, E. L., Whiting, A. K., Saari, R. E., McCracken, J., Hausinger, R. P., and Que, L., Jr. (1999) *Biochemistry 38*, 16714−26.
14. Barton, G. J. (1993) *Protein Eng. 6*, 37−40.
15. Zhou, J., Gunisor, M., Bachmann, B. O., Townsend, C. A., and Solomon, E. I. (1998) *J. Am. Chem. Soc. 120*, 13539−40.
16. Liu, A., Ho, R. Y., Que, L., Jr., Ryle, M. J., Phinney, B. S., and Hausinger, R. P. (2001) *J. Am. Chem. Soc. 123*, 5126−7.
17. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) *J. Biol. Chem. 272*, 23031−6.
18. Clifton, I. J., Hsueh, L. C., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2001) *Eur. J. Biochem. 268*, 6625−36.
19. Tabor, S. (1990) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K., Eds.) p 16.2.1−16.2.11, Green Publishing Associates and Wiley-Interscience, New York.
20. Chan, J. M., Wu, W., Dean, D. R., and Seefeldt, L. C. (2000) *Biochemistry 39*, 7221−8.
21. Powell, H. R. (1999) *Acta Crystallogr. D 55 (Pt 10)*, 1690−5.
22. Winn, M. D. (2003) *J. Synchrotron. Radiat. 10*, 23−5.
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., and Warren, G. L. (1998) *Acta Crystallogr. D 54 (Pt 5)*, 905−21.
24. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr. A 47 (Pt 2)*, 110−9.
25. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 945−949.
26. McRee, D. E. (1999) *J. Struct. Biol. 125*, 156−65.
27. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−24.
28. Zhang, Z., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2000) *Nat. Struct. Biol. 7*, 127−33.
29. Zhang, Z., Ren, J., Harlos, K., McKinnon, C. H., Clifton, I. J., and Schofield, C. J. (2002) *FEBS Lett. 517*, 7−12.
30. Fernandez, A., Sosnick, T. R., and Colubri, A. (2002) *J. Mol. Biol. 321*, 659−75.
31. Fernandez, A., and Scheraga, H. A. (2003) *Proc. Natl. Acad. Sci. U.S.A. 100*, 113−8.
32. Hogan, D. A., Smith, S. R., Saari, E. A., McCracken, J., and Hausinger, R. P. (2000) *J. Biol. Chem. 275*, 1200−12409.
33. Arnesano, F., Banci, L., Bertini, I., Ciofi-Baffoni, S., Molteni, E., Huffman, D. L., and O'Halloran, T. V. (2002) *Genome. Res. 12*, 255−71.
34. 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.
35. 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*, 235−350.

BI0341096