Crystal Structure of the Alkylsulfatase AtsK: Insights into the Catalytic Mechanism of the Fe(II) α-Ketoglutarate-Dependent Dioxygenase Superfamily[†,‡]

Ilka Müller,[||] Antje Kahnert,[⊥,§] Thomas Pape,[||] George M. Sheldrick,[||] Wolfram Meyer-Klaucke,[#] Thomas Dierks,[▽] Michael Kertesz,[⊗] and Isabel Usón*[,||,$]

Lehrstuhl für Strukturchemie, Universität Göttingen, Tammannstrasse 4, D-37077 Göttingen, Germany, Institute of Microbiology, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zürich, Switzerland, EMBL Outstation Hamburg, Notkestrasse 85, D-22603 Hamburg, Germany, Institut für Biochemie und Molekulare Zellbiologie, Abteilung Biochemie II, Universität Göttingen, Heinrich-Düker-Weg 12, D-37073 Göttingen, Germany, and School of Biological Sciences, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, U.K.

Received September 29, 2003; Revised Manuscript Received January 15, 2004

ABSTRACT: The alkylsulfatase AtsK from *Pseudomonas putida* S-313 belongs to the widespread and versatile non-heme iron(II) α-ketoglutarate-dependent dioxygenase superfamily and catalyzes the oxygenolytic cleavage of a variety of different alkyl sulfate esters to the corresponding aldehyde and sulfate. The enzyme is only expressed under sulfur starvation conditions, providing a selective advantage for bacterial growth in soils and rhizosphere. Here we describe the crystal structure of AtsK in the apo form and in three complexes: with the cosubstrate α-ketoglutarate, with α-ketoglutarate and iron, and finally with α-ketoglutarate, iron, and an alkyl sulfate ester used as substrate in catalytic studies. The overall fold of the enzyme is closely related to that of the taurine/α-ketoglutarate dioxygenase TauD and is similar to the fold observed for other members of the enzyme superfamily. From comparison of these structures with the crystal structure of AtsK and its complexes, we propose a general mechanism for the catalytic cycle of the α-ketoglutarate-dependent dioxygenase superfamily.

The alkylsulfatase, AtsK,[1] from *Pseudomonas putida* S-313 catalyzes the oxygenolytic release of inorganic sulfate from a range of different aliphatic sulfate esters to provide the strain with the sulfur that it requires for growth in sulfate-limited conditions (*1*). The enzyme decarboxylates one molecule of α-ketoglutarate (αKG) into succinate per molecule of sulfate ester that is cleaved (*1*) and is hence a member of the family of non-heme iron(II), αKG-dependent dioxygenases (*2*). Members of this remarkable enzyme family bind iron using a two histidine, one carboxylate facial triad, a feature that is one of nature's recurring multifunctional bioinorganic motifs (like the heme group or iron sulfur clusters). The family is widespread in both prokaryotic and eukaryotic organisms and includes a diverse variety of enzymes, which catalyze a range of energetically demanding biosynthetic and degradative reactions such as the stereoselective desaturation of unactivated carbon−carbon single bonds, oxidative ring closure, and hydroxylation processes. AtsK cleaves alkyl sulfate esters to yield the corresponding aldehyde and sulfate, and the initial hydroxylation in the cleavage reaction occurs at the nonactivated carbon atom adjacent to the sulfate ester group (*1*).

AtsK is a very flexible enzyme because it will desulfate a broad spectrum of linear and branched chain sulfate esters (K_m values 40−400 μM) using several different cosubstrates, as well as αKG. The AtsK protein has been reported in *P. putida* S-313 (*1*) and in *P. aeruginosa* (*3*), and homologues of *atsK* are also present in the genomes of several other pseudomonads. Its expression is regulated by the availability of inorganic sulfate and requires the LysR-family regulator SftR (*4*). AtsK is an intracellular enzyme and is therefore coexpressed with a sulfate ester transporter (AtsRBC). Interestingly, the sequenced strain of *P. putida*, strain KT2440, does not contain AtsK or AtsRBC and is unable to desulfurize sulfate esters as a result.

The first crystal structure of a member of the non-heme iron(II) αKG-dependent dioxygenase family was reported for deacetoxycephalosporin C synthase (DAOCS) (*5*), the enzyme that catalyzes the ring expansion of penicillin N in the biosynthesis of deacetoxycephalosporin C. Since then, crystal structures of several other αKG-dependent dioxygenases have been solved. These include the following:

[†] This work was supported by the Deutsche Forschungsgemeinschaft, the BMBF, and the Fonds der Chemischen Industrie.

[‡] Atomic coordinates have been deposited in the Protein Data Bank under codes 1oih (apo enzyme), 1oij (αKG−AtsK complex), 1oii (Fe−αKG−AtsK complex), and 1oik (Fe−αKG−2-ethylhexyl-1-sulfate−AtsK complex).

* Corresponding author. E-mail: iufcri@ibmb.csic.es. Fax: +34 93 2045904. Tel: +34 93 4006147.

[||] Lehrstuhl für Strukturchemie, Universität Göttingen.

[§] Current address: Max Planck Institute for Infection Biology, Department of Immunology, Schumannstrasse 21/22, 10117 Berlin, Germany.

[⊥] Swiss Federal Institute of Technology.

[#] EMBL Outstation Hamburg.

[▽] Institut für Biochemie und Molekulare Zellbiologie, Abteilung Biochemie II, Universität Göttingen.

[⊗] University of Manchester.

[$] Current address: ICREA at Instituto de Biología Molecular de Barcelona (IBMB-CSIC). C Jordi Girona 18-26, 08034 Barcelona, Spain.

[1] Abbreviations: AtsK, alkylsulfatase; αKG, α-ketoglutarate; DAOCS, deacetoxycephalosporin C synthase; CAS, clavaminate synthase; P3H, proline 3-hydroxylase; ANS, anthocyanidin synthase; TauD, taurine dioxygenase; XAS, X-ray absorption spectroscopy; XANES, X-ray absorption near-edge structure; EPR, electron paramagnetic resonance.

10.1021/bi035752v CCC: $27.50
Published on Web 02/27/2004

clavaminate synthase (CAS), which catalyzes three different reactions during clavaminate synthesis, hydroxylation, oxidative ring closure, and desaturation (*6*); proline 3-hydroxylase (P3H), which hydroxylates proline to *cis*-3-hydroxyproline (*7*) and is the prokaryotic homologue of the human P4H the malfunction of which under lack of vitamin C underlies scurvy; anthocyanidin synthase (ANS), which catalyzes a desaturation reaction in the synthesis of a plant pigment (*8*); taurine dioxygenase (TauD), which carries out the oxidative cleavage of the C−S bond in the sulfonate taurine (*9*). The hypothetical oxygenase product of the *Escherichia coli gab* gene has been determined within a structural genomics project, but its function is not yet known (*10*). These proteins display very low levels of sequence identity. Still, they all share the structural fold known as "jelly roll" and accommodate strikingly similar active sites in equivalent positions within this fold. The reactions catalyzed by members of this family are highly varied and very selective and of considerable biotechnological interest. This has prompted us to undertake a systematic structural study on the catalytic mechanism of AtsK to gain further insights into both the general mechanism and the structural differences underlying specificity within this family.

In this paper, we report the crystal structure of the AtsK enzyme in the apo form, and in three complexes: with the cosubstrate αKG, with αKG and iron, and finally with αKG, iron, and the substrate 2-ethylhexyl-1-sulfate used in catalytic studies. In addition, the nature of the iron present in the complexes was established by XAS.

EXPERIMENTAL PROCEDURES

Crystallization. The apo enzyme was concentrated to 18 mg/mL in 20 mM Tris/HCl at pH 7.5. The hanging drop vapor diffusion method was used at 20 °C under aerobic conditions. Crystals grew out of a mixture of equal volumes of protein and well solution (25 mM ammonium acetate or 25−50 mM ammonium sulfate, 25 mM sodium citrate, pH 5.6, 6−7% PEG 4000, 30% glycerol or MPD) to a typical size of $0.05 \times 0.2 \times 0.5$ mm^3 over 1−4 weeks and could be flash frozen in liquid nitrogen without further cryoprotectant. The crystals belong to the orthorhombic space group $P2_12_12_1$ (a = 72.4 Å, b = 145.5 Å, and c = 159.2 Å). The asymmetric unit contains four protein molecules and 60% solvent.

Crystals were derivatized by soaking in the precipitant solution containing 3 mM $HgCl_2$ for 2 days.

To obtain crystals of the cosubstrate complex, equal volumes of protein solution and a solution containing the cosubstrate αKG as a buffer (50 mM ammonium sulfate, 50 mM sodium αKG, pH 5.5, 8−10% PEG 4000, 30% glycerol) were mixed under aerobic conditions. Crystals appeared after 3 weeks.

To introduce iron into the cosubstrate complex, the crystals were transferred to a solution containing 50 mM ammonium sulfate, 50 mM sodium αKG, pH 5.5, 10% PEG 4000, 30% glycerol, and 2 mM $Fe(NH_4)_2(SO_4)_2$ under anaerobic conditions for 1−3 days.

Formation of the substrate complex was achieved by anaerobic soaking of the crystals in the above iron-containing mother liquor for several hours and addition of 2-ethylhexyl-1-sulfate under anaerobic conditions to a final concentration of 2 mM. Because the substrate range of the enzyme covers variable linear and branched aliphatic sulfate esters, we soaked with butyl-1-sulfate, hexyl-1-sulfate, and 2-ethylhexyl-1-sulfate. Only in the case of soaks with 2-ethylhexyl-1-sulfate did we succeed in getting data of suitable quality and resolution. In all three cases, the space group was shown to change from $P2_12_12_1$ to $I2_12_12_1$ (data not shown).

Data Collection and Processing. X-ray diffraction data of the apo form of AtsK were collected at 100 K on the beamline BW7B at the EMBL outstation c/o DESY in Hamburg to a resolution of 1.9 Å (λ = 0.8439 Å) using a MAR345 image plate detector. A mercury derivative dataset was collected in-house using a Bruker Cu rotating anode with Osmic mirrors (λ = 1.54178 Å) and a MAR345 image plate detector to a resolution of 3.30 Å. Lack of isomorphism hindered SIRAS solution of the structure with these two datasets. Thus, a further in-house native dataset was collected to a resolution of 3.8 Å.

Data collection of the αKG, the αKG−iron, and the αKG−iron−2-ethylhexyl-1-sulfate complexes of AtsK was performed at the beamline X11 at the EMBL outstation c/o DESY in Hamburg, using a MARCCD detector at wavelengths of 0.8126, 0.8098, and 0.8111 Å, respectively. The space group of the apo form, the αKG, and the αKG−iron complex was $P2_12_12_1$, whereas the space group of the αKG−iron complex crystals changed upon soaking with substrate to $I2_12_12_1$. Data were processed with the programs DENZO and SCALEPACK (*11*) and evaluated and prepared with Bruker program XPREP. Details are summarized in Table 1.

XAS data were gathered in fluorescence mode at the EMBL outstation c/o DESY in Hamburg, Germany, using a 13 element Ge-detector. The incoming X-ray beam was monochromated by a Si(111) double crystal. Harmonics were rejected by detuning of the monochromator to 50% of its peak intensity and by a focusing mirror with an energy cutoff of about 20.5 keV. The sample was kept at 30 K in a modified Oxford instruments closed-cycle cryostat during the data collection. Energy calibration and data analysis were performed as described in ref *12*. The X-ray absorption near-edge structure (XANES) resembles the one reported for Fe^{3+} human tyrosine hydroxylase (*12*), an enzyme with the two histidine, one carboxylate facial triad active site (*13*). The preedge peak intensity, characteristic for the local symmetry, suggests an octahedral iron ion, which is consistent with the crystal structure.

Structure Solution and Refinement. For SIRAS, the native and derivative in-house datasets of the apo enzyme crystals were used at a resolution of 3.8 Å. Ten mercury positions per asymmetric unit could be located by Patterson-aided dual-space methods as implemented in SHELXD (*14*).

The mercury sites were analyzed using the Bruker program XP to find four pairs of symmetry-related atoms consistent with the self-rotation function at κ = 180°, which suggested 222 noncrystallographic symmetry. Indeed, three orthogonal axes could be located among eight of the mercury atoms. Two mercury atoms were found to lie on one of the noncrystallographic 2-fold axes. Averaging matrices were calculated by the program LSQKAB (*15*) as implemented in the CCP4 package (*16*) and initial phases were calculated using a Harker construction (mean phase error = 53.6°).

Table 1: Summary of Crystallographic Data on AtsK and AtsK Complexes

	native 1 (*apo*-form) (synchrotron)	native 2 (*apo*-form) (in-house)	Hg derivative	α-KG complex	α-KG−Fe complex	substrate complex
cell [Å]	$a = 72.24$	$a = 71.74$	$a = 72.21$	$a = 72.45$	$a = 72.42$	$a = 72.03$
	$b = 145.34$	$b = 145.49$	$b = 145.31$	$b = 144.88$	$b = 147.33$	$b = 141.89$
	$c = 159.46$	$c = 158.91$	$c = 159.37$	$c = 160.69$	$c = 158.57$	$c = 161.72$
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$I2_12_12_1$
resolution limits [Å]	20−1.89	12−3.80	15−3.30	38.93−2.10	20−2.19	25−2.05
total no. of reflns	719 684	109 036	151 676	472 571	564 785	270 566
no. of unique reflns	134 420	29 248	45 894	102 501	86 268	51 154
completeness						
overall	99.9	94.5	98.0	98.8	97.5	99.1
outer shell	99.6 (1.9−2.0)	88.4 (3.8−3.9)	97.0 (3.3−3.4)	94.6 (2.2−2.07)	83.1 (2.3−2.19)	97.0 (2.15−2.06)
R_{merge}[a]						
overall	0.048	0.077	0.159	0.104	0.064	0.057
outer shell	0.357	0.128	0.344	0.410	0.452	0.321
mean $I/\sigma(I)$						
overall	19.21	11.26	10.76	13.51	19.96	13.53
outer shell	4.32	5.97	6.51	3.06	5.10	4.34

[a] $R_{merge} = \sum|I_i - \langle I\rangle|/\sum I_i$ over all reflections.

NCS averaging, solvent flattening, histogram matching, and phase extension to the resolution of 1.9 Å were carried out by the program dm (*17*) (mean phase error = 22.6°). After autobuilding of the main chain by ARP/wARP (*18*), the asymmetric unit consisted of 934 residues and 16 chains (connectivity index 0.96, *R*-factor 18.1%).

This main chain model and the sequence were used in ARP/wARP to build the side chain atoms. Additional manual rebuilding was performed with XtalView (*19*) and the experimental density from dm.

The resulting model was refined isotropically with REFMAC5 (*20*), interspersed with manual rebuilding of the model in a $2mF_o-DF_c$ and a F_o-F_c map using XtalView. Water molecules were included in the model by XtalView and evaluated manually with respect to *B*-factors and suitable stereochemistry. Within all four monomers of the asymmetric unit, residual electron density was detected in the active site that could not be interpreted sufficiently. The residual electron density was 0.5 e/Å^3.

Phases for the structures of the protein complexes were obtained by a rigid body refinement followed by simulated annealing with the program CNS (*21*). All four monomers in the asymmetric unit of the apo enzyme structure were used as individual rigid groups for refinement of the αKG and the αKG−iron complexes. Because the asymmetric unit of substrate complex crystal contains two independent molecules, monomer A of the apo form with truncated side chains in the region of the active site was used for the molecular replacement search by the program MOLREP (*22*), followed by simulated annealing by CNS (the chain ID of the second chain was chosen to be D to assign the chain ID according to the crystal contacts). Initial rigid body refinement and further refinement of the complex structures including ncs-restraints was performed with REFMAC5. Water molecules were built in by XtalView, and their *B*-factors and stereochemistry were evaluated.

In all structures, hydrogen atoms were included in geometrically calculated positions whenever R_{free} dropped upon refinement.

PROCHECK (*23*) and WHATIF (*24*) were used to check stereochemical and geometrical outliers in the final structures. Ramachandran plot outliers are L212 (generously allowed region) and K53 (disallowed region) in all of the structures. The electron density of these amino acids was carefully inspected, and no inconsistency between the geometry as described by the model and the electron density was detected.

Table 2: Data on Refinement of the Different AtsK Structures

dataset	apo form	α-KG complex	α-KG−Fe complex	substrate complex
overall *R* factor (%)	17.0	19.1	19.5	19.9
free *R* factor (%)	19.1	21.5	22.0	22.1
free reflns (%)	5.1	4.8	4.9	5.0
RMS bond length (Å)	0.02	0.02	0.02	0.02
RMS bond angles (deg)	1.70	1.60	1.65	1.68
no. of residues per monomer				
monomer A	254	254	253	240
monomer B	255	255	253	−
monomer C	242	238	237	−
monomer D	244	236	245	254
no. of water molecules	760	694	574	246

Figures were created using the programs BobScript (*25*) and MolScript (*26*).

Table 2 shows details of the refinement.

RESULTS AND DISCUSSION

To gain further insight into the catalytic mechanism of AtsK action, we solved a number of different complexes of the protein intended to mimic the various stages of the reaction cycle. We investigated the enzyme itself (apo form) and the structures of AtsK complexes with the cosubstrate αKG (αKG complex), with both cosubstrate and iron (αKG−Fe complex), and as a cosubstrate−iron−substrate complex (αKG−Fe−$ROSO_3$ complex). The overall structure did not change upon complex formation, and the differences between the structures are in the same range as those among the four monomers in the asymmetric unit of the apo enzyme structure where noncrystallographic symmetry restraints were not used during refinement, as shown by the program ESCET (*27*). Despite this, soaking of the apo enzyme with substrate invariably led to a change in the space group from $P2_12_12_1$ to $I2_12_12_1$, which can be associated with the fixation of one part of a flexible loop coordinating the substrate, because this loop was not visible in any of the $P2_12_12_1$ structures.

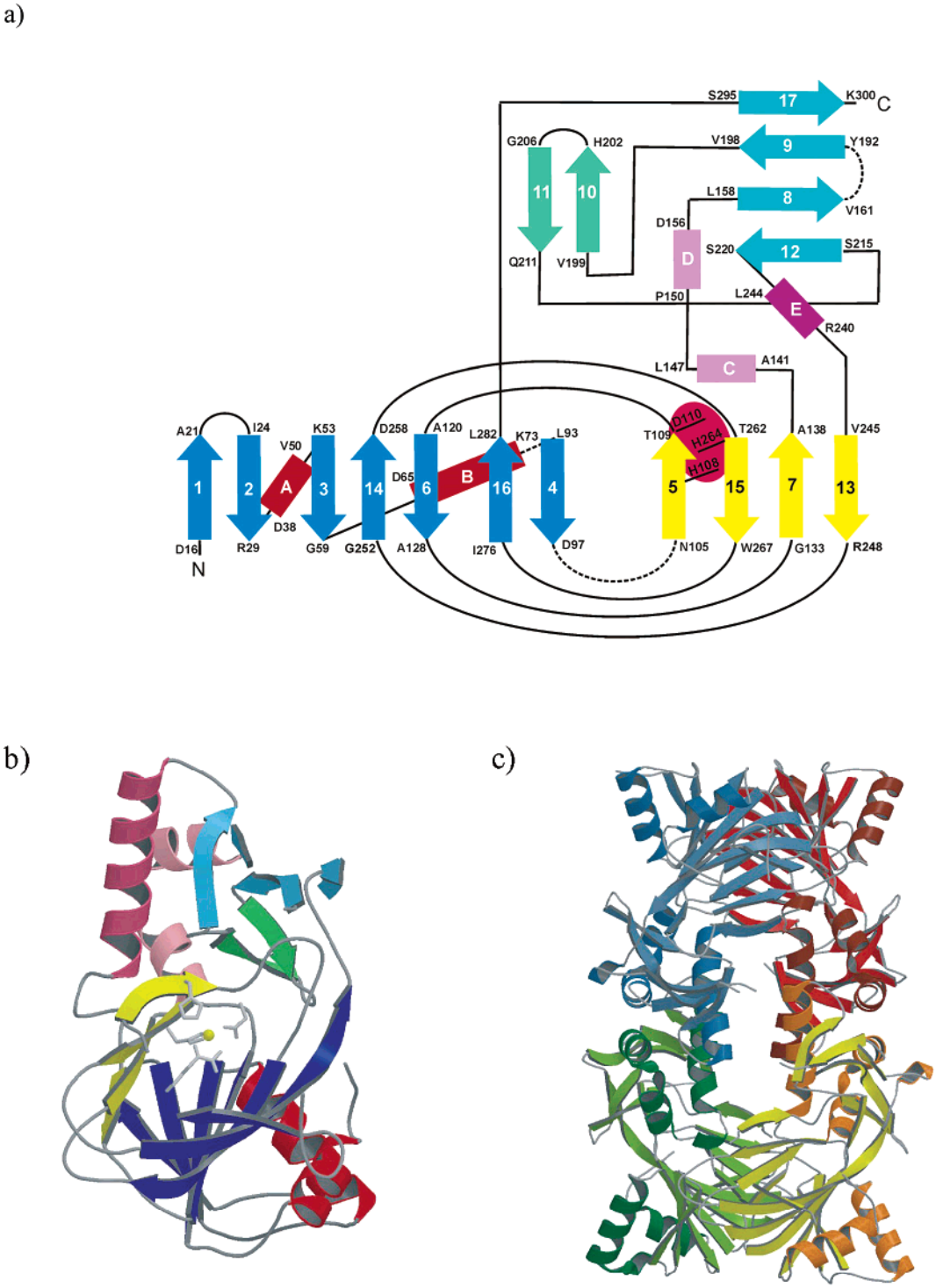

FIGURE 1: The overall fold of AtsK. Panel a presents the topology diagram of the AtsK structure. β strands are represented by arrows, and α helices are represented by cylinders. The first and the last residue number in each secondary structure element is given. The colors used match those in Panel b. Panel b shows a ribbon presentation of the AtsK structure highlighting the secondary structure elements and the iron cation (yellow sphere), the cosubstrate αKG, and relevant side chains in the residues constituting the iron-binding site (pale gray). Panel c shows a ribbon presentation of the putative biologically active tetramer of AtsK.

The quality of the structures of the monomers in the asymmetric unit, as represented by the *B*-values and the number of disordered amino acids, varies in the four structures. The best defined monomer carries the chain ID A in all structures except for the substrate complex, for which monomer D was best defined (the same chain ID identifies those monomers in the asymmetric unit forming the same crystal contacts). The apo structure yielded the highest resolution (1.9 Å), and in the following, we describe the structure of the best defined monomer in the apo form, up to the point where the different active sites are presented. Table 1 summarizes the X-ray data statistics. Table 2 shows details of the refinement.

Description of the Fold: AtsK Shows the Jelly Roll Motif Present in the Other Known Fe−*αKG-dependent dioxygenases.* Figure 1 shows a topology diagram and a schematic view of the structure of the protein. The overall fold displays striking similarities with the one of TauD (*9*), although the

sequence identity is only 38%. The core α/β-domain containing the active site folds into a jelly roll motif. The larger seven-stranded β-sheet (β1, β2, β3, β14, β6, β16, β4) in it is twisted and flanked by the α-helices αA and αB. The smaller four-stranded β-sheet of the jelly roll motif consists of four antiparallel strands (β5, β15, β7, β13). Two of the surface loops of the jelly roll motif that are located close to the active site are disordered and cannot be determined from the experimental electron density map. The active site (see below) is located between strands β5 and β15 inside the wider end of a polar funnel-like gap in the interior of the jelly roll. An extended insert (A141−N243) consisting of three α-helices (αC, αD, αE) and two antiparallel β-sheets of two (β10, β11) and three strands (β9, β8, β12) protrudes from the jelly roll core. The latter β-sheet is completed by a C-terminal β-strand (β17). Several residues of the inserted domain interact with the main core of the protein via salt bridges (E207 to K53, E146 to R23, E285 to R200, and D258 to H202) and hydrogen bonds (R23 to E146, R23 to N243, S106 to H214, W107 to H214, E211 to R260, T112 to H196, T112 to H162, F113 to T297, E115 to T297, R200 to E285, and H202 to D258). No significant hydrophobic contacts are present between the two domains. A large loop, from Y166 to T190, in the extended insert close to the active site cannot be modeled into the electron density. Additionally, the C-terminal residue and 12 N-terminal amino acids are missing in the electron density map of all the structures (Table 2).

The protein does not contain any cysteine residues and thus has no disulfide bridges. It also lacks methionine, consistent with its expression by *P. putida* under sulfur starvation conditions.

AtsK Forms Tetramers under Physiological Conditions. The molecular weight of the active species, determined by gel filtration chromatography, is 121 kDa (*1*), which corresponds well with a tetrameric structure, and a dimer of dimers with 222 symmetry was indeed present in the crystal. The formation of the dimer between the monomers A and B and monomers C and D is mediated through their α-helices αD and αE. The helices αD(A) and αD(B) and αE(A) and αE(B) are antiparallel to each other.

These four helices together build a hydrophobic cavity. On the outside of this cavity two salt bridges between helices αD(A) and αD(B) (E153 to K157) and hydrogen bonds between helices αE(A) and αE(B) (D226 to H237 and H229 to H229) assist the hydrophobic interactions.

The contact between the two dimers AB and CD to form a dimer of dimers is mediated through hydrophobic interactions between the four-stranded β-sheet in the jelly roll motif and the loop between strand β13 and helix αE. Additionally, direct hydrogen bonds (P19 to R248, G22 to D269, N105 to E242, E242 to Y265, E242 to D268, and T244 to R246) and water-mediated hydrogen bonds stabilize the dimer of dimers.

An analogous tetramer can be found in the crystal packing of TauD (PDB ID 1gpw). Nevertheless, TauD is reported to be active as a dimer (*34*).

The formation of this tetramer leaves a cavity surrounded by protein but not entirely isolated from the external solvent. The polar surface of this hole is covered by water molecules conserved among the structures. The active site in all four enzyme molecules within this dimer of dimers is accessible both from the inner cavity and from the surface of the enzyme. The tetramer hole is large enough to accommodate the missing amino acids between R165 and Y191. This loop could participate in substrate recognition and protection from undesired side reactions by covering the active site as a lid during the catalytic cycle, or it could even provide a side chain to quench activated oxygen species that might be formed in the absence of substrate by a nonconcerted mechanism (*2*, *28*). In the mercury derivative structure, the two heavy atoms lying on one of the noncrystallographic 2-fold axes were found at the interface of the four monomers inside the tetramer hole coordinated by H229 and Q228 and their symmetry equivalents.

Description of the Active Site. Table 3 summarizes the relevant interatomic distances found within the active site in the different structures of AtsK and its complexes. The only conserved sequence pattern within this protein family is H-X-D/E-X_n-H, where the last histidine is found 39 (P3H) to 154 (AtsK, TauD) residues away from the H-X-D/E sequence. This motif constitutes the iron binding site, as was visualized in the first crystal structure of an αKG-dependent oxygenase, the structure of the cephalosporin synthase DAOCS (*5*). All other iron-containing structures of enzymes belonging to this family so far determined have shown an equivalent coordination site.

The Apo Form Binds a Sodium Cation. In the case of AtsK, the metal binding amino acids are H108, D110, and H264. In the apo enzyme, high residual electron density appears in the iron binding site at too short a distance to the protein side chains to correspond to a coordinated water molecule. Because sodium was the only metal ion present in the crystallization liquor, this site has to be filled by this cation. The occupancy of the sodium had to be set at values between 0.5 and 0.7 for the different monomers to match the *B* values of other atoms in the same area. This and the ellipsoidal shape of the electron density indicates a substitutional disorder between the nearer sodium cation and a water molecule, somewhat further, in agreement with the distances of a water molecule to the coordinating amino acids as found in the apo structure of DAOCS (H183−H_2O 4.4 Å, D185−H_2O 2.3 Å, and H243−H_2O 3.3 Å). The resolution of the structure does not allow these peaks to be resolved. After refinement of this region with a partially occupied sodium cation, as shown in Figure 2, difference electron density remains between the residues constituting the cation binding site (H108, D110, H264, and Na) and R275 and T135 at the bottom of the active site cavity. This electron density is most likely the result of a loose coordination to the sodium ion of carboxylate compounds present in the crystallization liquor, which contained a mixture of citrate and acetate. The best defined feature is the carboxylate group salt-bridged to R275, whereas the rest represents the presence of a mixture.[2]

αKG Is Held by a Salt Bridge to a Conserved Arginine and a Hydrogen Bond to a Hydroxyl Side Chain. The structure of the αKG complex shows the region just described to be the cofactor binding site (Figure 2). To obtain the complex, crystals were grown with αKG in the crystallization buffer, avoiding the presence of citrate and acetate

[2] The structure refined including a citrate molecule modeled into this density can be obtained from the authors on request.

Table 3: Summary of Selected Distances in the Active Site of the Structure of AtsK and Its Complexes

AtsK complex	His108(NE2) −Fe/Na [Å]	Asp110(OD1) −Fe/Na [Å]	His264(NE2) −Fe/Na [Å]	H_2O/sub(C1) −Fe/Na [Å]	αKG(O2) −Fe [Å]	αKG(O5) −Fe [Å]	Arg279(NH1) −αKG(O1) [Å]	Arg275(NH2) −αKG(O3) [Å]	Arg275(NH1) −αKG(O4) [Å]	Thr135(OG1) −αKG(O3) [Å]
AtsK + Na[a] (apo form)	2.60	2.24	2.27							
	2.84	2.35	2.19							
	2.63	2.76	3.11							
	2.37	2.27	2.65							
AtsK + Na + αKG	2.40	2.06	2.26		2.33	2.19	4.01/3.92 [a]	3.39	2.87	2.84
	2.51	2.18	2.40		2.51	2.20	4.57/4.43 [a]	3.17	2.74	2.70
	2.50	2.23	2.16		2.38	2.14	4.37/3.90 [a]	3.12	2.74	2.69
	2.45	2.24	2.19		2.31	2.18	*c*	3.21	2.64	2.58
AtsK + Fe + αKG	2.21	1.95	2.13	1.88	2.11	1.92	2.90	2.77	3.02	2.63
	2.23	2.00	2.14	1.88	2.04	2.04	2.98	2.94	3.02	2.69
	2.22	2.05	2.10	1.93	2.11	1.92	3.12	2.84	2.81	2.56
	2.17	2.07	2.12	2.11	1.97	2.03	3.21	2.84	3.00	2.60
AtsK + Fe + αKG + $ROSO_3$	2.19	1.94	2.14	4.86	2.19	2.02	2.51[a]	2.77	3.01	2.71
	2.20	1.93	2.02	3.40	2.00	1.96	2.66[a]	2.91	3.01	2.54
AtsK−sodium citrate	2.50	2.23	2.21	1.98	2.12 (O4)	3.12 (O5)	3.06 (O6)	3.16	2.93	2.63
	2.53	2.36	2.13	1.99	2.42 (O4)	2.96 (O5)	2.99 (O6)	3.13	2.94	2.48
	2.61	2.76	3.23	3.00	3.30 (O4)	2.47 (O5)	2.86 (O3)	3.36	2.99	2.56
	2.36	2.27	2.76	2.31	2.41 (O4)	2.85 (O5)	2.94 (O4)	3.33	2.95	3. 06

[a] Possible positional disorder of sodium and a water molecule and therefore bond length not reliable. [b] αKG(O2)−Arg279(NH2). [c] No Arg279 side chain modeled.

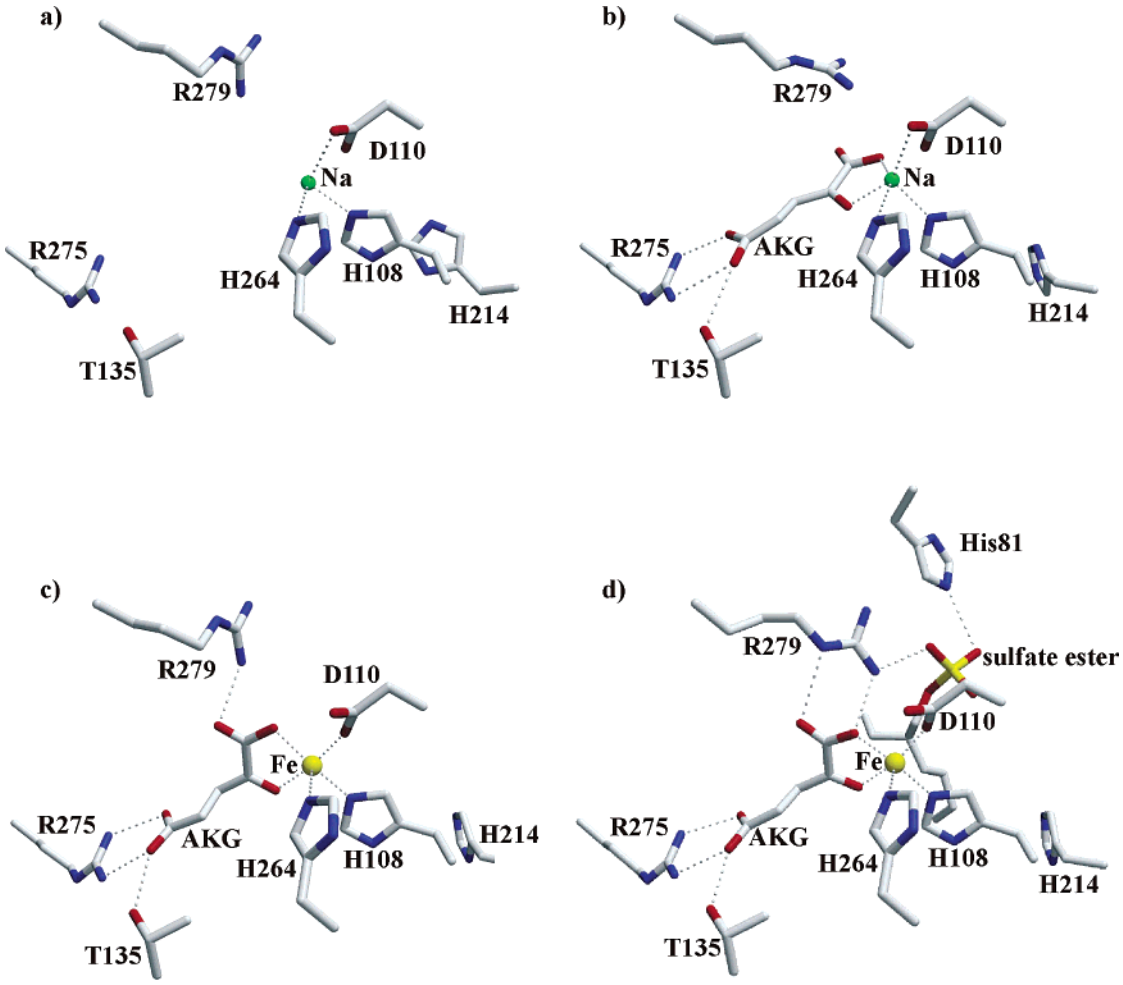

FIGURE 2: Active site region of AtsK: (a) apo enzyme; (b) αKG complex; (c) αKG−Fe complex; (d) αKG−Fe−substrate complex. The iron is coordinated by the enzyme through H108, H264, and D110. The αKG coordinates to the iron atom by the carboxyl group at C1 and the oxo group at C2. The water molecule occupying the sixth coordination site in the αKG iron complex is not present in the substrate complex. The carbon atom of the sulfate ester molecule to be oxidized is close to the metal atom (see text for details).

to prevent competition for the cofactor binding site by other carboxylated species. In the structure of the αKG complex, the iron binding site is occupied predominantly by a sodium cation (occupancies vary between 0.7 and 1) with similar bond lengths to those seen in the apo structure. The occupancies of the αKG molecules were made similar to those of the sodium atoms. The sodium cation is coordinated by the carboxylate group of the cosubstrate at αKG(C1) through αKG(O2) and by the oxo-atom αKG(O5). The carboxylate group of the αKG therefore lies opposite to H264, and the carbonyl oxygen of αKG is trans to D110. The sixth coordination site at the sodium ion is opposite to H108 and close to R279. Three of the four monomers in the asymmetric unit showed no peak of high electron density in this region. The α-carboxyl group of the αKG that coordinates the cation displays higher *B*-values than the other protein atoms, indicating a disorder of this carboxyl group. Indeed, in the empty sixth coordination site in monomer B, residual electron density could be found the shape of which was less typical of a water than of a carboxyl group. This accounts for the distortion in the electron density in the coordination sphere around the sodium atom seen in the remaining monomers. However, a disorder of the cosubstrate could not be modeled.

As well as being coordinated to the sodium cation, αKG is anchored to the protein by a salt bridge to R275 through its γ-carboxyl group at αKG(C5). This carboxyl group forms a hydrogen bond with T135(Oγ1) through αKG(O3). This results in an overall arrangement for the major conformation of the αKG in which all atoms lay approximately in one plane.

Despite not being directly involved in either metal or substrate binding, H214, at the outer end of the active site cavity, changes its orientation upon αKG binding (Figure 3). This movement enlarges the cleft leading to the iron center, which could either provide a cavity suited for longer aliphatic substrate side chains or generate more room to accommodate the flexible loop R165 to V191. However there is no unique hydrophobic pocket in which to bury the aliphatic chain of the substrate, in agreement with the broad substrate range observed for the enzyme.

Iron Coordination Causes a Rotation of the αKG(C1) Carboxylate Group. The iron(II) cation is coordinated by the protein through H108(Nϵ2), D110(Oδ1), and H264(Nϵ2). Whether iron or sodium is present in the active site or not does not lead to a significant change in the side chain positions of these amino acids. Comparison of the metal bond lengths in the iron−αKG and the αKG complex shows a decrease for the iron complex of approximately 0.2 Å for each bond as compared to sodium. The octahedral coordination sphere of the iron cation in the αKG−Fe complex is completed by the carboxyl group αKG(O2) and the oxo

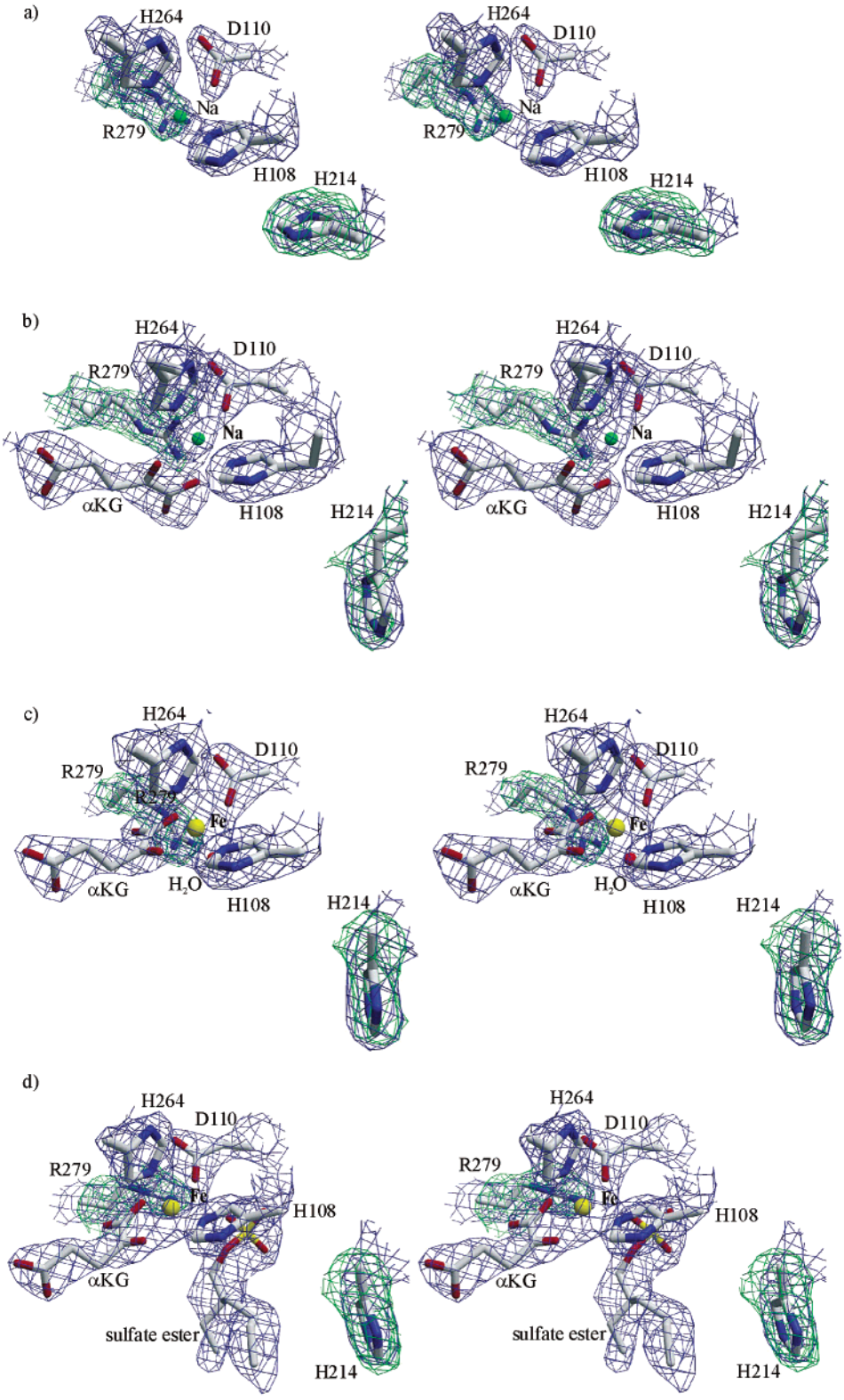

FIGURE 3: Stereoview of the active site region of (a) the apo form, (b) the Na−αKG−AtsK complex, (c) the Fe−αKG−AtsK complex, and (d) the Fe−αKG−substrate−AtsK complex. A $2mF_o-DF_c$ map contoured at 1σ is shown in blue; a σA weighted (*40*) difference electron density map for the side chains of H214 and R279 is shown in green. The conformational change of the side chain of H214 upon αKG binding and reorientation of R279 upon substrate binding is clearly visible.

group αKG(O5) of the cosubstrate, as well as by a water molecule. The Fe−O/N bond lengths range from 1.9 to 2.5 Å.

The conformation of the αKG differs from the major conformation seen in the sodium−αKG complex by a 90° rotation of the carboxyl group at αKG(C1), taking it out of the plane of the other atoms in the molecule (Figure 3). It is comparable to the minor conformation suggested by the residual electron density in the sodium−αKG complex, despite the difference in coordination number between both structures. In the octahedral geometry around the iron atom, the carboxyl group of the αKG(O2) therefore occupies the

position opposite to H108 and the water is in trans position to H264. The αKG is coordinated by the same amino acids as in the sodium-containing complex except that the carboxyl group at αKG(C1) now forms a hydrogen bond between αKG(O1) and R279(Nχ2). The side chain of the latter changed its conformation relative to that observed in the sodium−αKG complex (Figure 3). Spectroscopic measurements on enzymes of this family have already shown the octahedral coordination of the iron(II) center in the cosubstrate complex (*29*). In that case, it was proposed that coordinative saturation could protect the enzyme from oxygen binding and formation of radical species in the absence of a proper substrate, thus hampering the occurrence of the nonconcerted mechanism. Nevertheless, in the case of AtsK further reaction was only prevented under anaerobic conditions, as was shown in spectroscopic analysis.

The metal oxidation state and the local symmetry could be determined by X-ray absorption near-edge structure (XANES) on the AtsK−Fe−αKG sample in solution under aerobic conditions. The XANES pattern is indicative for Fe^{3+} bound to histidine and carboxyl ligands

Cocrystallization or soaking with Fe(II) under aerobic conditions never yielded crystals in which iron was bound to the active site (over 10 complete data sets collected on such crystals, data not shown), suggesting that once iron oxidation takes place, it will not bind to the enzyme. Therefore, the Fe(II) solution added to the aerobic protein/αKG sample was prepared anaerobically to avoid oxidation of the metal prior to protein binding. On the other hand, oxidation of the coordinated iron, when the complex is kept under aerobic conditions, does not lead to dissociation of the metal, as the XANES data clearly demonstrate protein binding of iron. In conclusion, we showed that the oxidation of the iron inside the protein is not coupled to substrate turnover and, as a side reaction, can take place in the absence of the latter.

Substrate Binding Fixes a Flexible Loop, Which Closes on the Cavity and Induces a Change in Space Group. In the αKG−Fe−$ROSO_3$ complex, the iron center is pentacoordinated, as the site previously occupied by a water molecule is now held empty. Substrate coordination does not leave sufficient space between the hydrophobic chain of the alkyl sulfate and the iron center to accommodate a water ligand (3.4 Å in diameter). This pentacoordination in the substrate/cosubstrate/iron-bound form of the enzyme has also been shown spectroscopically (*29*) and invoked to explain the role of substrate binding in enzyme activation. The αKG shows the same orientation as that in the αKG−Fe complex, presenting the position opposed to H264 to the substrate. The three terminal sulfate oxygen atoms of the 2-ethylhexyl-1-sulfate are coordinated by the protein through the nitrogen atoms V111(N), H81(Nϵ2), and R279(Nχ2) (see Figure 2 and Figure 5). In the substrate-free structures, H81 is part of a flexible disordered loop (H81−V84), and only in this structure is there a clear electron density for this amino acid. This is the only structural change that can be correlated with the change in space group from $P2_12_12_1$ to $I2_12_12_1$ that is observed upon soaking with substrate. In TauD, taurine binding induces dramatic conformational changes in the equivalent random coil region and a helix that close over the substrate-filled active site (*30*).

The side chain of R279 alters its orientation to coordinate the substrate (see Figure 3), thereby changing the hydrogen bonding to αKG so that R279(Nχ1) now interacts with αKG-(O2) and R279(Nϵ) with αKG(O1). The tight binding to this carboxyl group could be important in initiating decarboxylation of the cosubstrate during the catalytic cycle. The bridging oxygen atom of the substrate is weakly coordinated by the flexible loop (E165−V191), but the electron density map does not allow precise characterization of this feature. Except for this last putative residue, the substrate binding side chains are the same as those for TauD sulfonate binding.

The rather weak coordination of the fourth sulfate oxygen by the flexible loop could prevent tight coordination of the inorganic sulfate produced in the reaction. Indeed sulfate was not found to bind to the protein in any of the structures, despite being present in concentrations of up to 75 mM during crystallization. The aliphatic chain of the substrate is oriented so that the carbon atom C1 of the sulfate ester which is oxidized during catalysis is in close proximity to the iron atom. The electron density for the aliphatic side chain of the substrate is not very clear (Figure 3) because the substrate was used in racemic form and the occupancy of the substrate molecule was modeled as only 0.6. Besides, AtsK is active on a broad range of aliphatic sulfates, so binding of the aliphatic chain must not be very specific.

Comparison of Different αKG-Dependent Dioxygenases. Table 4 summarizes the available crystal structures of other αKG-dependent dioxygenases and their complexes. The overall fold of all enzymes includes the jelly roll motif in the core domain, hosting a conserved active site in the equivalent position in each protein. The main differences among the structures of the enzymes of this family occur in the region of the helices that flank the jelly roll β-sheets and in the extended regions. AtsK(141−243) and TauD-(131−234) (*9*) show an extended insert region, which is also present in CAS(177−258) (*6*, *31*) and GAB(192−271) (*10*). A three-dimensional superposition search with the program DALI (*32*) of this insert from AtsK confirmed the structural relationship among those inserts but no other protein with a similar structural element could be identified. Instead of an extended insert, DAOCS (*5*, *33*), ANS (*8*), and P3H (*7*) display a C-terminal extension in the same part of the protein structure.

The active site is located in all these proteins in the funnel-like cavity within the jelly roll. Two areas can be differentiated within the active site: the part of the structure responsible for metal and cofactor binding is highly conserved, whereas the substrate binding environment differs among the various enzymes. The first part includes the conserved motif H-X-D(E)-X_n-H, which coordinates the iron cation, and one arginine that forms a salt bridge with αKG γ-carboxylate, which does not bind the iron cation. Comparison of all crystal structures permits the assignment of further conserved residues of the active site, even though no metal or cosubstrate complex structures are available for some of the proteins (Table 4). Thus, the cosubstrate coordination site includes either a threonine (AtsK, CAS, and TauD) or a serine (DAOCS, ANS, and GAB) to form a hydrogen bond with the γ-carboxylate of the αKG. In the case of P3H, three different residues could fill this role: H135 (the side chain of which corresponds to the position occupied by T135 in AtsK), H43 (corresponding to S260 in

Table 4: Summary of Available Crystal Structures of αKG-Dependent Dioxygenases

protein	fold[a]	complexes[b]	OH[c]	αKG[d]	anchoring amino acids for αKG to: γ-COO[e]	γ-COO	α-COO[f]	PDB-ID (*41*)
AtsK	insert	apo enzyme (1.9 Å)	+	−	T135	R275	R279	1oih
		+ Na, αKG (2.07 Å)		p				1oij
		+ Fe, αKG (2.2 Å)		np				1oii
		+ Fe, αKG, 2-ethylhexyl-1-sulfate (2.05 Å)		np				1oik
DAOCS	C-ext	apo enzyme (1.3 Å)	−	−	S260	R258	R162	1dcs
		+ Fe (1.5 Å)		−				1rxf
		+ Fe, αKG (1.5 Å)		p				1rxg
		+ Fe, succinate, CO_2 (1.96 Å)		−				1e5h
		+ Fe, α-ketohexanoic acid (1.6 Å)		−				1hjf
		+ Fe, α-ketopentanoic acid (1.5 Å)						1hjg
CAS	insert	apo enzyme (1.63 Å)	+	−	T172	R293	R297	1ds0
		+ Fe, αKG (1.08 Å)	−	np				1ds1
		+ Fe, αKG, NAA[i] (1.40 Å)		np				1dry
		+ Fe, αKG, proclavaminic acid A5[j] (2.10 Å)		np				1drt
		+ Fe, αKG, DGP,[g] NO[k] (1.54 Å)		p				1gvg
P3H	C-ext	apo enzyme (2.3 Å)	+	−	H135	R168	R95, R97	1e5r
		+ Fe (2.4 Å)					or R122	1e5s
ANS	C-ext	apo enzyme + αKG (2.1 Å)	+	p	S300	R298	N215	1gp4
		+ Fe, αKG, DHQ (substrate) (2.2 Å)[h]		p				1gp5
		+ Fe, succinate, product (1.75 Å)		−				1gp6
Gab	insert	+ Fe (2.0 Å)	?	−	S188	R305	R309 or R311	1jr7
TauD	insert	apo enzyme (1.9Å)	+	−	T126	R266	R270	1os7
		+ Fe, αKG (2.50 Å)		np				1otj
		+ Fe, αKG, taurine (2.50 Å)		np				1gy9

[a] The proteins fold into a major and a minor domain; the latter can be an insert or a C-terminal extension. [b] Complex of the protein for which a crystal structure is available (resolution given in parentheses). [c] Hydroxylation of the substrate during catalysis: + denotes yes; − denotes no. [d] Conformation of the αKG in the structure: p denotes planar; np denotes not planar; − denotes no data available (see text for details). [e] Hydrogen bond donor for γ-carboxyl group of αKG that binds to the arginine given in the next column via salt bridge. [f] Hydrogen bond donor for carboxyl group of αKG that coordinates to Fe. [g] Deoxyguanidinoproclavaminate (substrate for hydroxylation). [h] Published as a substrate complex, but comparison of this structure and the published product complex structure reveals discrepancies. These include high *B*-values for the carboxyl group of the αKG and hydrogen bonding between the carboxyl group of the highly acidic αKG and a carbonyl group of the alleged substrate (a water molecule instead of this carboxyl group in the iron coordination sphere could explain both the hydrogen bond and the poor electron density for the modeled carboxyl group); the distance of the substrate alcohol carbon atom to be oxidized during catalysis (6.26 Å) is identical to that of the corresponding carbonyl-carbon atom in the product (6.22 Å) and too high to allow a direct oxygen transfer from the iron to the carbon atom. Finally, it is not likely that there are no significant changes in substrate or protein conformation from the substrate to the product complex. Therefore, the structure refined as containing the substrate can be suspected to be rather the product containing structure. Accordingly, it has not been taken into account in the discussion. [i] Substrate analogue for hydroxylation. [j] Substrate for ring closure. [k] CO_2 analogue.

DAOCS), and S170, which has been proposed as a putative cosubstrate binding site (*7*). Superposition of the various active site models determined in the presence of cosubstrate suggests that the first of these residues, H135, is best suited because the others lie too far away to form a hydrogen bond (over 5 Å in the superposition). Each structure possesses a group capable of hydrogen bonding to this end of the cosubstrate, but interestingly, the relevant amino acid may be located on either side of the γ-carboxyl group, thus binding O3 or O4. For DAOCS and ANS, the serine side chain binds one end of the carboxyl group (corresponding to αKG(O4) in AtsK structure), whereas the respective amino acids in the other protein structures bind to the other side of the cosubstrate γ-carboxylate (αKG(O3) in AtsK).

Second, an arginine (R279 in AtsK) close to the iron-coordinating carboxyl group at αKG(C1) forms a hydrogen bond with the cosubstrate. In ANS, the arginine is replaced by an asparagine (N215). The arginine and asparagine are important for the orientation of the C1 carboxyl group of αKG in the iron coordination sphere. For AtsK, TauD, CAS, P3H, and GAB, it is in trans position to the *H*-X-D(E)-X_n-H, liberating the position opposite the second histidine to a water molecule (Figure 4a). The conformation of the αKG can be described as a rotation of 50°−90° by this iron-coordinating carboxyl group relative to the other atoms of the molecule, which lie approximately in one plane. For ANS, DAOCS, and the NO complex of CAS, the cosubstrate conformation is the same as the one found in the AtsK−αKG complex, containing pentacoordinated sodium rather than hexacoordinated iron, where all atoms of the cosubstrate lie roughly in one plane. This leads in the case of ANS and DAOCS to an iron coordination sphere where the carboxyl group of αKG and the water molecule change places, so that the water molecule is in trans position to the *H*-X-D(E)-X_n-H (Figure 4b). It has been discussed that the coordination sphere of the iron rearranges during catalysis (*31*). The variation observed in these different structures could therefore represent the coordination sphere of iron in different stages of catalysis.

The described arginine/asparagine represents a linker between the iron/cosubstrate and substrate binding sites. As found in the substrate structures of AtsK, TauD, and CAS, the side chain of the arginine/asparagine binds both to αKG-(O2) and to either a carboxyl or the sulfate or sulfonate group of the substrate.

The substrate is placed opposite to H-X-D(E)-X_n-*H* in all of these complexes. For AtsK, TauD, and CAS, the water in this site of the metal coordination sphere could be replaced by O_2 during catalysis, allowing the substrate to react directly with a ferryl intermediate (Figure 6). A change in the coordination sphere and hence a different mechanism could also be possible admitting the CAS−NO structure as a O_2 complex analogue.

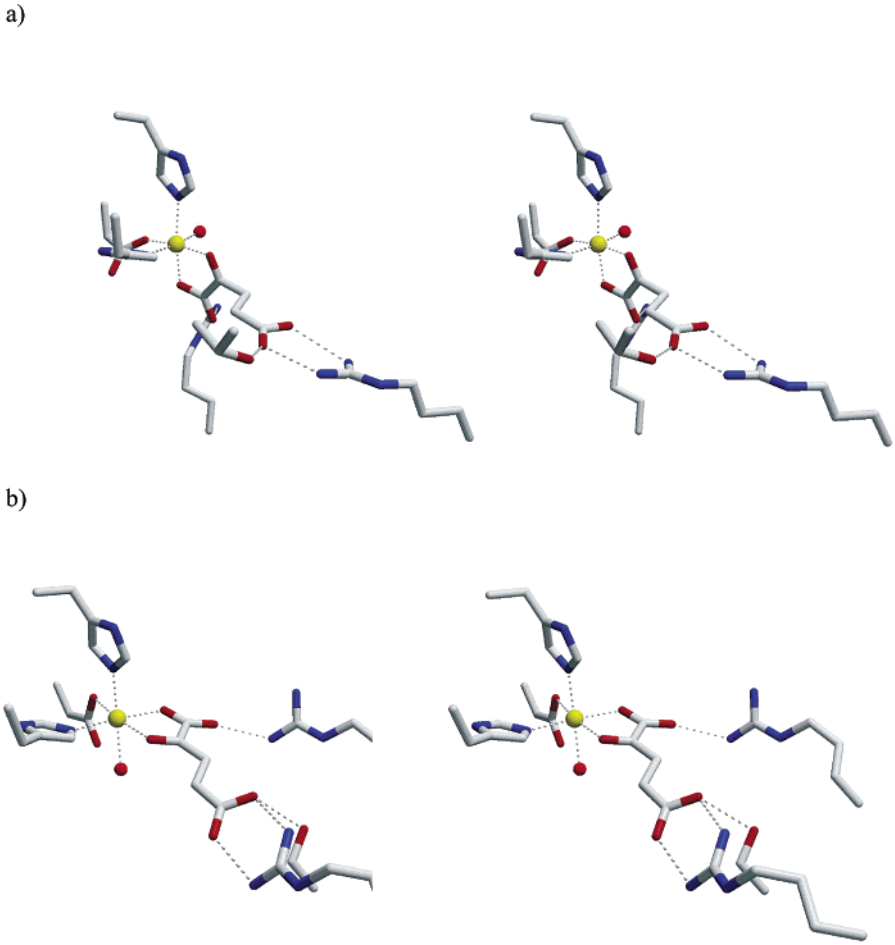

FIGURE 4: Stereoview of the active sites of the iron−cosubstrate complexes of (a) AtsK (substrate and cosubstrate complex) and (b) DAOCS. Comparison shows the different arrangement of the cosubstrate in the iron coordination sphere.

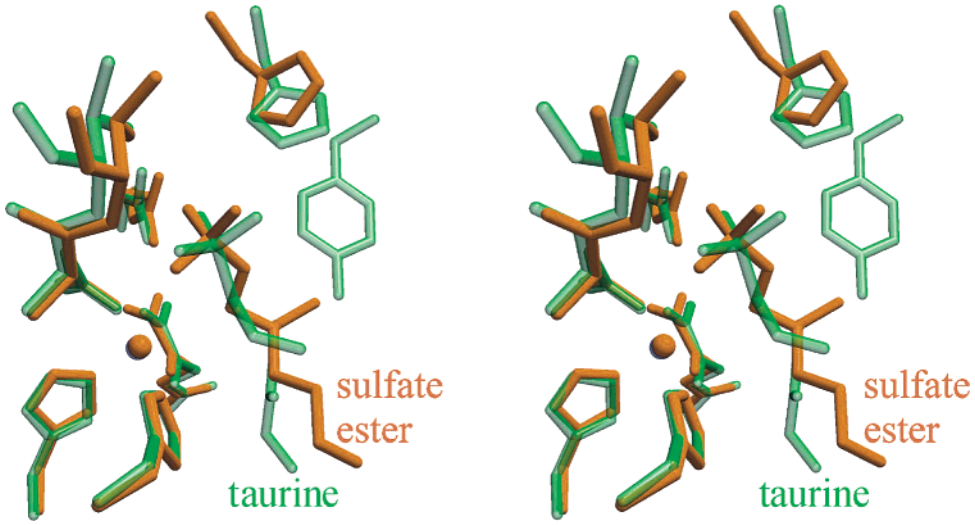

FIGURE 5: Stereoview of the superposition of AtsK (orange) and TauD (green). The sulfate and the sulfonate group, respectively, are coordinated in the same way but shifted relative to each other to position the substrate atom to be oxidized in proximity to the iron. In case of taurine, the substrate molecule is anchored via its amino group to an asparagine and tyrosine side chain.

The Fe−αKG−substrate−AtsK structure shows that the sulfate ester is coordinated by the equivalent arginine, histidine, and valine that bind the sulfonate group of taurine in the Fe−αKG−taurine−TauD complex (Figure 5). The coordination fixes the SO_3 moiety in sites that are shifted from each other by roughly 0.5 Å and rotated by approximately 25° and the substrate adopts a different conformation. As a consequence, the atom to be oxidized, namely, the nearest substrate atom to the iron center, is in either case the first carbon atom in the aliphatic chain of the substrate, bonded to the sulfur in the sulfonate and to the ester oxygen in the sulfate ester. The weak binding of the ester oxygen by the flexible loop could play a role in this altered orientation of the SO_4 moiety, thus contributing to its correct

positioning for catalysis. This poses the question whether the sulfonate and the sulfate ester are arranged in this way because of differences in the active site of the proteins or because of the chemical structure of the substrates. For TauD, only activity assays including sulfonate compounds are reported (*34*). In case of AtsK, no release of sulfite from taurine was detected. This implies that either AtsK is not able to cleave sulfonates or the polar amine group of taurine prevents the molecule from entering the active site. Further investigations on the ability of AtsK to bind and cleave sulfonates are underway.

For the other enzymes, no substrate or product complexes are available, or the substrates are even unknown. At best, this allows a tentative assignment of the linking amino acids on the basis of the structural superposition.

Nevertheless, for DAOCS, it is most probable that R162, which coordinates the cosubstrate in the Fe−αKG complex, also binds a carboxyl group of the substrate. The CC-bond at which the oxidative C-insertion occurs to form the six-membered ring in the product would be located close to the iron center by coordination of R162 to this carboxyl group of the product. It is worth mentioning that DAOCS crystallizes as a trimer and that the C-terminus of a symmetry-related protein molecule extends into this putative substrate binding pocket in the crystal structure (*33*).

In the case of P3H, several arginine side chains (R95, R97, and R122) are located in a position suitable for coordination to the carboxyl group of αKG(C1). These three arginines communicate via hydrogen bonds. R122 is most likely to coordinate the αKG and R95 and R97 could fix the substrate proline at its carboxylic end to bring its C3 close to the iron.

The Catalytic Mechanism. Figure 6 shows a scheme of the proposed general catalytic mechanism for enzymes of the αKG-dependent dioxygenase family, generalizing the individual reaction mechanisms. The resting state of the active enzyme corresponds to the structure of the complex with iron(II) and αKG. Kinetic studies have shown that catalysis is triggered upon substrate binding (*35*). In the case of AtsK, the sulfate group of the substrate is held fixed by three hydrogen bridges to the main chain atom N of V111 and side chain atoms Nχ2 of R279 and Nϵ2 of H81. This last residue belongs to a flexible loop that is missing from the electron density maps of the remaining three structures, so substrate binding determines its partial immobilization. No interaction is seen for the bridging oxygen of the sulfate group, but the coordination to a second flexible loop (R165−Y191) indicated by residual electron density close to the oxygen atom may be more specific for the optimal reaction substrate, which is so far undetermined. Concomitantly to the sulfate ester binding, the water molecule previously coordinated to the iron center is displaced, which would contribute to activate the enzyme for catalysis upon substrate binding. The water molecule should be the most labile ligand in the iron coordination sphere, and indeed, the AtsK enzyme structure complexed with αKG and the lower charged sodium cation lacks a water molecule in this position. According to kinetic studies (*35*), the next step is the coordination of O_2 to the iron center, which implies the displacement of this water molecule from the sixth coordination position. Studies on a CAS−Fe(II)−αKG−NO complex have shown this enzyme to bind NO at the bottom of the active site, at a position opposed to the substrate. A hydrophobic tunnel is proposed to lead from the bottom of the jelly roll to the iron center. The amino acid side chains inside the jelly roll core in all structures of AtsK are packed more tightly than the ones in the CAS structure. Alternatively to the diffusion of oxygen through the jelly roll core, the oxygen molecule could approach the iron center from the substrate binding site where the water coordination site is already held empty. The next step of the catalytic cycle is considered to be oxidation of the cosubstrate, which is thereby decarboxylated, leading to the formation of carbon dioxide and succinate. At the same time, the remaining oxygen atom forms a ferryl intermediate (Fe(IV)═O) with the iron atom (characterized by Mössbauer and EPR for TauD (*36*)). The side chain of the conserved R279, placed to interact both with the carboxyl αKG (C1) group and with the substrate, may assist this reaction by hydrogen bonding the resulting carbon dioxide and removing it from the iron coordination sphere. In that case, a water molecule may fill the sixth coordination site again, or succinate may coordinate in a bidentate manner. R279 not only directs the cosubstrate but is also involved in substrate binding, as we deduce from the substrate complex structures of AtsK, TauD, and CAS. In DAOCS, the corresponding arginine side chains are located at the open face of the active center.

The highly reactive ferryl intermediate (Fe(IV)═O) now abstracts a hydrogen atom from the substrate atom held in closest proximity by the substrate binding area of the active site, thus generating a radical species on the substrate and a reduction of the metal center (Fe(III)−OH). After radical formation, the radical center should be the most proximate atom of the substrate to the iron. Consequently, abstraction of the hydroxyl group from the iron center, leading to hydroxylation of the substrate should be the most immediate reaction and therefore the fate of most substrates in this enzyme family including AtsK. The oxygen atom in the hydroxylated substrate has been shown to derive both from molecular oxygen and from water in experiments using isotopically labeled reagents on deacetoxy/deacetylcephalosporin C synthase (*37*). This may be explained by equilibrium exchange of a proton between hydroxyl and water ligands in the iron coordination sphere, rather than by the reaction of a free water molecule with the substrate radical, since the latter mechanism would give rise to undesired and uncontrolled side reactions. In some cases, a reaction alternative to the hydroxylation is observed. In these cases, another hydrogen atom is subtracted from the substrate radical by the Fe(III)−OH species. CAS catalyzes three different types of oxidations, namely, hydroxylation, ring formation, and desaturation (*6*), and the reaction fate should not only be directed by the features of the active site. More likely, the faster the reaction is (in case of CAS, the intramolecular ring closure reaction) or the more stable the product is (in case of CAS, formation of the desaturated product rather than a semiacetal), the more likely the reaction will be.

Oxidation of the substrate radical regenerates the iron(II) center, which is ready for the next turnover once the resulting succinate has left the active site and has been replaced by new αKG.

The role of ascorbate is still unclear. It is required stoichiometrically by some of the enzymes in this family (e.g., human proline 4-hydroxylase) and has been shown to

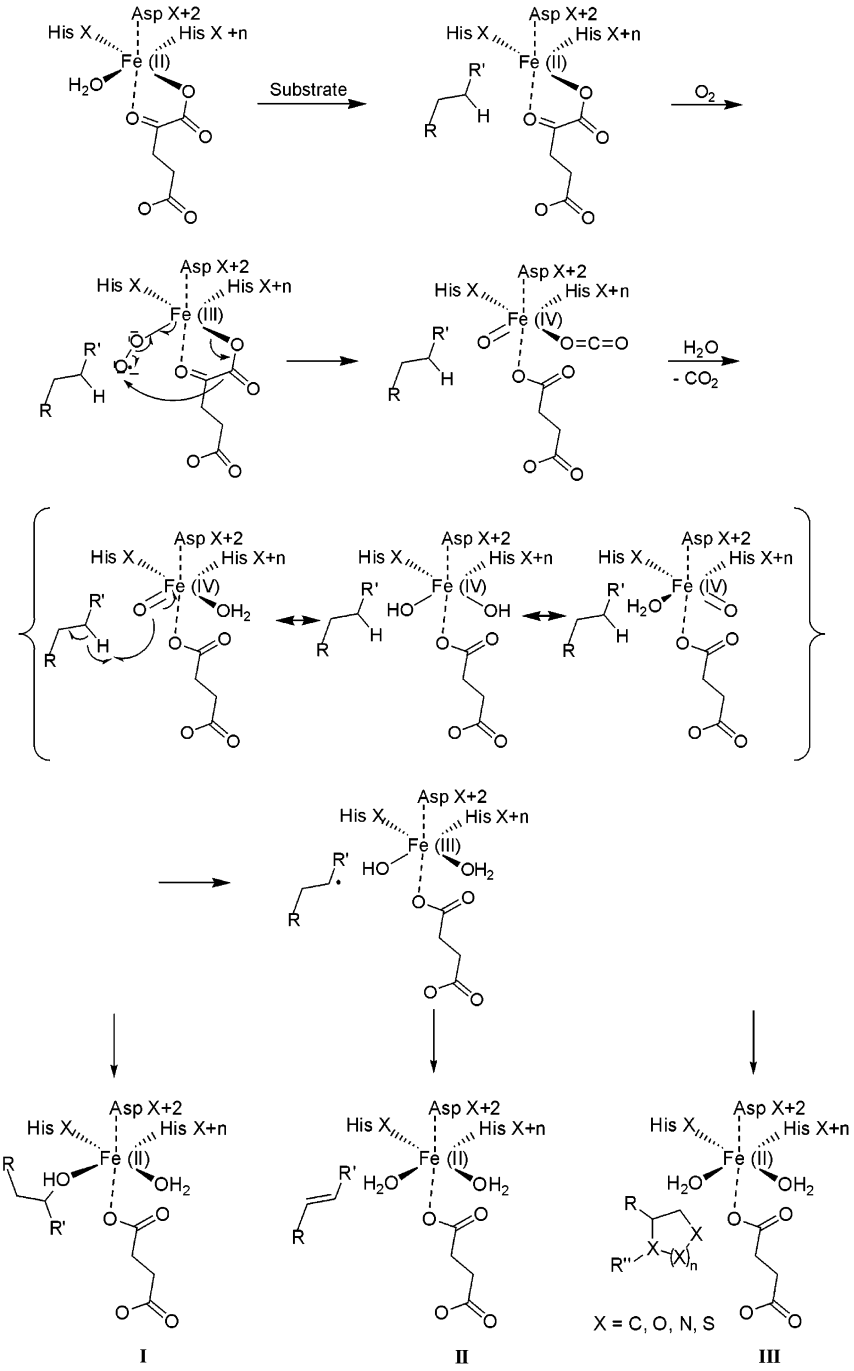

FIGURE 6: Proposed mechanism for enzymatic reactions catalyzed by αKG-dependent dioxygenases. In the first step, the substrate coordinates to the iron replacing a water molecule in the iron coordination sphere. Second, a dioxygen molecule coordinates to the iron. Oxidation and decarboxylation of the cosubstrate occurs, and the ferryl Fe(IV)═O species is formed. Dissociation of CO_2 from the iron and addition of a water molecule to the highly positively charged ferryl species possibly followed by hydrogen exchange to an oxo or hydroxyl species then allow a rearrangement of the iron ligands. This is necessary to bring one oxygen atom coordinated to the iron close to the substrate to abstract a hydrogen atom. Iron is thereby reduced to Fe(III). The radical intermediate reacts further with the iron species either through hydroxyl transfer or abstraction of a second hydrogen atom from the substrate to form either the hydroxylated (I) or the desaturated species (II), or it leads to a cyclization or ring expansion (III). This is accompanied by the reduction of the iron to Fe(II). As a side reaction, the hydroxyl radical could be transferred from a water molecule to the radical intermediate, leaving an inactivated Fe(III)−protein complex that requires a reducing agent to start a new catalytic cycle. Ascorbate is known to play this role for a number of αKG-dependent dioxygenases (*39*).

accelerate the reaction in all cases where the kinetic effect has been characterized (*38*). AtsK activity does not require ascorbate but is strongly stimulated by it. It has been proposed to play a part in the prevention of oxidative self-inactivation (*39*), but the mechanism has yet to be elucidated.

ACKNOWLEDGMENT

Data collection at the synchrotron beamlines BW7b and X11 in Hamburg at the EMBL outstation c/o DESY and assistance from Christofer Enroth and Ehmke Pohl are gratefully acknowledged.

SUPPORTING INFORMATION AVAILABLE

Figure of merit-weighted mean phase error (wMPE) and XAS spectrum for the Fe(III)−αKG−AtsK complex in comparison with the Fe^{2+}- and Fe^{3+}-human tyrosine hydroxylase. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Kahnert, A., and Kertesz, M. A. (2000) Characterization of a sulfur-regulated oxygenative alkylsulfatase from *Pseudomonas putida* S-313, *J. Biol. Chem. 275*, 31661−31667.
2. Ryle, M. J., Hausinger, R. P. (2002) Non-heme iron oxygenases, *Curr. Opin. Chem. Biol. 6*, 193−201.
3. Quadroni, M., James, P., Dainese-Hatt, P., and Kertesz, M. A. (1999) Proteome mapping, mass spectrometric sequencing and reverse transcriptase-PCR for characterisation of the sulfate starvation-induced response in *Pseudomonas aeruginosa* PAO1. *Eur. J. Biochem. 266*, 986−996.
4. Kahnert, A., Mirleau, P., Wait, R., and Kertesz, M. A. (2002) The LysR-type regulator SftR is involved in soil survival and sulfate ester metabolism in *Pseudomonas putida*, *Environ. Microbiol. 4*, 225−237.
5. Valegård, K., Terwisscha 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. (1997) Structure of a cephalosporin synthase, *Nature 394*, 805−809.
6. Zhang, Z.-H., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2000) Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase, *Nat. Struct. Biol. 7*, 127−133.
7. Clifton, I. J., Hsueh, L.-C., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2001) Structure of proline 3-hydroxylase, *Eur. J. Biochem. 268*, 6625−6636.
8. Wilmouth, R. C., Turnbull, J. J., Welford, R. W. D., Clifton, I. J., Prescott, A. G., and Schofield, C. J. (2002) Structure and mechanism of anthocyanidin synthase from *Arabidopsis thaliana*, *Structure 10*, 93−103.
9. 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) X-ray crystal structure of *Escherichia coli* taurine/α-ketoglutarate dioxygenase complexed to ferrous iron and substrates, *Biochemistry 41*, 5185−5192.
10. Chance, M. R., Bresnick, A. R., Burley, S. K., Jiang, J.-S., Lima, C. D., Sali, A., Almo, S. C., Bonanno, J. B., Buglino, J. A., Boulton, S., Chen, H., Eswar, N. Y., He, G., Huang, R., Ilyin, V., McMaham, L., Pieper, U., Ray, S., Vidal, M., and Wang, L. K. (2002) Structural genomics: A pipeline for providing structures for biologists. *Protein Sci. 11*, 723−738.
11. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
12. Meyer-Klaucke, W., Winkler, H., Schünemann, V., Trautwein, A. X., Nolting, H.−F., and Haavik, J. (1996) Mossbauer, electron-paramagnetic-resonance and X-ray-absorption fine-structure studies of the iron environment in recombinant human tyrosine hydroxylase, *Eur. J. Biochem. 241*, 432−439.
13. Que, L., Jr. (2000) One motif-many different reactions. *Nat. Struct. Biol. 7* (3), 182−184.
14. Usón, I., and Sheldrick, G. M. (1999) Advances in direct methods for protein crystallography, *Curr. Opin. Struct. Biol. 9*, 643−648.
15. Kabsch, W. (1976) A solution for the best rotation to relate two sets of vectors, *Acta Crystallogr. A 32*, 922−923.
16. Collaborative Computational Project, Number 4 (CCP4) (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr. D 50*, 760−763.
17. Cowtan, K. (1994) in *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, (Bailey, S., and Wilson, K., Eds.) pp 34−38 Daresbury, U.K.
18. Perrakis, A., Morris, R. J. H., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol. 6*, 458−463.
19. McRee, D. E. (1999) XtalView/Xfit − A versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol. 125*, 156−165.
20. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method, *Acta Crystallogr. D 53*, 240−255.
21. Brünger, 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) Crystallography and NMR systems: a new software suite for macromolecular structure determination, *Acta Crystallogr. D 54*, 905−921.
22. Vagin, A., and Teplyakow, A. (1997) MOLREP: an automated program for molecular replacement, *J. Appl. Crystallogr. 30*, 1022−1025.
23. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
24. Vriend, G. (1990) WHAT IF: A molecular modeling and drug design program, *J. Mol. Graphics 8*, 52−56.
25. Esnouf, R. M. (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities, *J. Mol. Graphics 15*, 132−134.
26. Kraulis, P. J. (1991) Molscript: A program to produce both detailed and schematic plots of protein structure, *J. Appl. Crystallogr. 24*, 946−950.
27. Schneider, T. R. (2002) A genetic algorithm for identification of conformationally invariant regions in protein molecules, *Acta Crystallogr. D 58*, 195−208.
28. Ryle, M. J., Liu, A., Muthukumaran, R. B., Ho, R. Y. N., Koehntop, K. D., McCracken, J., Que, L., Jr., and Hausinger, R. P. (2003) O_2- and α-ketoglutarate-dependent tyrosyl radical formation in TauD, an α-ketoacid-dependent non-heme iron dioxygenase, *Biochemistry 42*, 1854−1862.
29. Solomon, E. I., Brunold, T. C., Davis, M. I., Kemsley, J. N., Lee, S.-K., Lehnert, N., Neese, F., Skulan, A. J., Yang, Y.-S., and Zhou, J. (2000) Geometric and electronic structure/function correlations in non-heme iron enzymes, *Chem. Rev. 100*, 235−349.
30. O'Brien, J. R., Schuller, D. J., Yang, V. S., Dillard, B. D., and Lanzilotta, W. N. (2003) Substrate-induced conformational changes in *Escherichia coli* taurine/α-ketoglutarate dioxygenase and insight into the oligomeric structure, *Biochemistry 42*, 5547−5554.
31. Zhang, Z.-H., Ren, J., Harlos, K., McKinnon, C. H., Clifton, I. J., and Schofield, C. J. (2002) Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO-complex: evidence for metal centred rearrangements, *FEBS Lett. 517*, 7−12.
32. Holm, L., and Sander, C. (1996). Mapping the protein universe, *Science 273*, 595−602.
33. Lee, H.-J., Lloyd, M. D., Harlos, K., Clifton, I. J., Baldwin, J. E., and Schofield, C. J. (2001) Kinetic and Crystallographic studies on deacetoxycephalosporin C synthase (DAOCS), *J. Mol. Biol. 308*, 937−948.
34. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*, *J. Biol. Chem. 272*, 23031−23036.
35. Holme, E. (1975) A kinetic study of thymine 7-hydroxylase from *Neurospora crassa*, *Biochemistry 14*, 4999−5003.
36. Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr., and Krebs, C. (2003) The first direct characterization of a high-valent iron intermediate in the reaction of an α-ketoglutarate-dependent dioxygenase: a high-spin Fe(IV) complex in taurine/α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*, *Biochemistry 42*, 7497−7508.
37. Baldwin, J. E., Adlington, R. M., Crouch, N. P., and Pereira, I. A. C. (1993) Incorporation of ^{18}O-labeled water into oxygenated products produced by the enzyme deactetoxy/deacetylcephalosporin C synthase, *Tetrahedron 49*, 7499−7518.
38. Myllylä, R., und Kivirikko, K. I. (1997) Characterization of the iron- and 2-oxoglutarate binding sites of human prolyl 4-hydroxylase, *EMBO J. 16*, 1173−1180.
39. Que, L., and Ho, R. Y. N. (1996) Dioxygen activation by enzymes with mononuclear non-heme iron active sites, *Chem. Rev. 96*, 2607−2624.
40. Read, R. J. (1986) Improved coefficients for maps using phases from partial structures with phases, *Acta Crystallogr. A 42*, 120−129.
41. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Brice, M. D. Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) Protein Data Bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol. 112*, 535−542.

BI035752V