Published in final edited form as:

Metallomics. 2013 April ; 5(4): 287–301. doi:10.1039/c3mt20153h.

Imposing function down a (cupin)-barrel: secondary structure and metal stereochemistry in the α KG-dependent oxygenases

John A. Hangasky, Cornelius Y. Taabazuing, Meaghan A. Valliere, and Michael J. Knapp*

Department of Chemistry, University of Massachusetts, Amherst, MA 01003, USA

Abstract

The Fe(II)/ α ketoglutarate (α KG) dependent oxygenases catalyze a diverse range of reactions significant in biological processes such as antibiotic biosynthesis, lipid metabolism, oxygen sensing, and DNA and RNA repair. Although functionally diverse, the eight-stranded β -barrel (cupin) and HX(D/E)XnH facial triad motifs are conserved in this super-family of enzymes. Crystal structure analysis of 25 α KG oxygenases reveals two stereoisomers of the Fe cofactor, Anti and Clock, which differ in the relative position of the exchangeable ligand position and the primary substrate. Herein, we discuss the relationship between the chemical mechanism and the secondary coordination sphere of the α KG oxygenases, within the constraints of the stereochemistry of the Fe cofactor. Sequence analysis of the cupin barrel indicates that a small subset of positions constitute the second coordination sphere, which has significant ramifications for the structure of the ferryl intermediate. The competence of both Anti and Clock stereoisomers of Fe points to a ferryl intermediate that is 5 coordinate. The small number of conserved close contacts within the active sites of α KG oxygenases can be extended to chemically related enzymes, such as the α KG-dependent halogenases SyrB2 and CytC3, and the non- α KG dependent dioxygenases isopenicillin N synthase (IPNS) and cysteine dioxygenase (CDO).

Introduction

Fe(II)/ α ketoglutarate (α KG) dependent oxygenases are a large class of enzymes that catalyze a wide array of biomedically relevant reactions. The α KG oxygenases are typically grouped according to the types of reactions they catalyze. These include protein side chain modifications, repair of alkylated DNA and RNA, biosynthesis of antibiotics, biosynthesis of plant products, lipid metabolism, and other assorted functions.¹ Protein sequence analysis suggested that there were as many as ten different sub-families defined by their various loops or domains inserted between the strands of the cupin barrel.² All of the α KG oxygenases contain an eight-stranded beta-barrel (the cupin barrel) as their core structural element, with a number of variable inserted loops and elements such as a lid on the active site. For example, our research on hypoxia sensing specifically centers on the chemical mechanism of FIH and PHD2, enzymes with very different active site geometry.^{3,4} Despite this variety, the Fe cofactor in all α KG oxygenases contains a His₂(Asp/Glu) facial triad of ligands near the top of the cupin barrel, with one exchangeable ligand site that is frequently

occupied by H₂O. When αKG binds as a bidentate ligand, two stereoisomers of the Fe, Anti and Clock, are observed (Fig. 1). The explosion of crystallographically determined structures over the past five years provides a basis for comparing Fe stereochemistry among enzymes of high structural homology.

The purpose of this review is to define how Fe stereochemistry relates to enzyme function. While such chirality in αKG oxygenases has been noted previously,^{5–7} the relationship between Fe stereochemistry and the enzyme structure and function has not been discussed. This metal stereochemistry is crucial to understanding the chemistry of αKG oxygenases, as it dictates the proximity between the site of O₂ binding and the primary substrate, as well as which active site residues may form polar or hydrophobic contacts with the reactants. The facial triad and any other ligands directly bonded to Fe form the primary coordination sphere, whereas a small number of residues in the cupin barrel form close contacts to these ligands, and constitute a secondary coordination sphere. As substrate binds at the top of the cupin barrel, the relative position of substrate and the exchangeable ligand site differs between the Anti and Clock stereoisomer (Fig. 1), raising the potential for different chemical reactions for the two stereoisomers.

Research on these enzymes is motivated by two overarching goals. The first goal is to predict enzyme function from primary sequence and/or structure, encompassing both identification of substrates as well as the chemistry performed by the active site Fe. Although αKG oxygenases can typically be identified from the conserved HX(D/E)X_nH motif, aligning the remainder of the sequence is a real challenge. Additionally, the structural factors that ensure high fidelity between O₂ activation to form the ferryl intermediate, and hydroxylation of the substrate by the ferryl intermediate remain poorly understood. We propose that conserved second coordination sphere interactions within the cupin barrel may aid in future sequence alignments, as well as provide insight into the mechanism of O₂ activation. The second goal is to develop inhibitors or activators for these enzymes for therapeutic applications. While certain classes of compounds have shown promise as inhibitors of the αKG oxygenases,^{8–12} improved understanding of the active site structure, including the chirality of the Fe center, will lead to selective enzyme effectors.

This review summarizes our analysis of Fe stereochemistry and second coordination sphere positions in those αKG oxygenases that have a high degree of structural similarity to FIH and PHD2 in their cupin fold. We first describe the common features of these αKG oxygenases, including chemical mechanism, key chemical intermediates, and protein topology. Next, our analysis of Fe stereochemistry is described, with the conclusion that changes in the Fe stereochemistry alters which cupin barrel positions constitute the second coordination sphere.

As the mechanism for O₂ activation in the αKG oxygenases is not known, these second coordination sphere interactions may be crucial for reactivity. The extension of these sequence correlations to function in structurally related enzymes, such as the αKG dependent halogenases and non-αKG dependent oxygenases, suggests that our analysis of cupin barrel positions may be broadly applicable.

The cupin fold and the consensus mechanism

Members of this enzyme superfamily all possess a cupin (“little barrel”) fold, characterized by two conserved β -sheets consisting of four β -strands, with several variable loops. The core of all α KG oxygenases exhibit this cupin fold, in which substrate binds between the top of the barrel and a variable loop which forms an active site lid.¹⁵ Additional β -strands, α -helices, and loops surround the core, and likely confer auxiliary stability to the enzymes. At the heart of the β -strand core, lies the active site, which contains metal binding residues arranged in a HX(D/E)X_nH motif occupying one face of the octahedral Fe(II) center. This is referred to as a 2-His 1-carboxylate facial triad.¹⁶ For all α KG oxygenases, the facial triad is always localized on the first sheet and second strands of the front β -sheets (F1 and F2 in Fig. 2). Residues that comprise the second coordination sphere are all located on the back β -sheet. In effect, all residues that have the potential to significantly influence the active site chemistry generally come from the top two strands of the front and back β -sheets. A similar number of interactions in the metal-binding vicinal oxygen chelate proteins were described in a recent review.¹⁷

These enzymes utilize three substrates (O_2 , α KG, and an oxidizable primary substrate), forming three products (CO_2 , succinate, and an oxidized product); typically we will simply refer to the oxidizable primary substrate as ‘substrate’ or ‘primary substrate’ for brevity. The Fe^{2+} center is coordinated by a bidentate α KG using its C-1 carboxylate and C-2 ketone. Additional stability is achieved via interaction of the α KG C-5 carboxylate with other protein side chains. The C-1 carboxylate binds opposite to either of the two His residues while the C-2 ketone binds opposite to the acidic (Asp/Glu) residue, leading to a chiral coordination geometry for the Fe (Fig. 1). These two modes of binding, which dictate the stereochemistry at the active site, can have important ramifications for substrate binding, positioning, and the subsequent chemistry that ensues.

Although the consensus mechanism was initially proposed in the early 1980s¹⁸ and has been subsequently refined by a number of researchers, molecular detail is lacking for a number of proposed intermediates and for the role of the protein active site in catalysis. In broad terms, these enzymes use molecular oxygen and α KG to form a ferryl intermediate that performs a two-electron oxidation to the primary substrate. This imposes certain requirements on ligands binding to the Fe during catalysis, which has been addressed by a combination of spectroscopy on resting enzyme and rapidly frozen intermediates.^{19–27}

Crystal structure and magnetic circular dichroism (MCD) data reveal that the resting form of the Fe-bound form of the enzyme consists of a six coordinate octahedral Fe^{2+} center coordinated to the facial triad residues and three water molecules.^{20,28} According to the consensus mechanism, the first step is binding of α KG, forming the (Fe + α KG) form of enzyme (Scheme 1). Next, primary substrate binds, however it does not coordinate directly to the metal center; rather it induces aquo release, thereby opening a coordination site for O_2 binding. The crystal structures of (Fe + α KG + Substr) bound forms of FIH and TauD, as examples, show a 5-coordinate Fe^{2+} center, which is consistent with this view.⁴ In addition, MCD and CD spectra of a number of enzymes loaded with (Fe + α KG + Substr) corroborate the formation of a 5-coordinate Fe(II) center upon substrate binding in solution.^{20,23,29}

The next steps in the consensus mechanism are incompletely understood, but are proposed using the hydrogen atom abstraction (HAT)/rebound paradigm of the P450 oxygenases. Substrate binding significantly increases reactivity towards O₂, as loss of the aquo ligand opens a coordination site that allows O₂ to bind.³⁰ It's thought that O₂ binds to form a Fe(III)-O₂-adduct, which attacks the keto group of αKG to form an intermediate that can be formulated as a bicyclic peroxybridged Fe⁴⁺. Decarboxylation and heterolytic cleavage of the O-O bond yields an Fe(IV)-oxo intermediate which has been spectroscopically observed in TauD and, and identified as the species responsible for HAT from primary substrate.^{19,31} Hydroxyl rebound to the substrate radical yields the hydroxylated product.

Although hydroxylation is the typical reaction for the αKG oxygenases, a number of different products can be found for this class of enzyme. In several cases the final product is a result of hydrolysis of the initially formed hydroxylated product. In certain other cases, the HAT step is followed by steps other than hydroxyl rebound, leading to halogenation or desaturation.

Demethylation

Fe(II)/αKG dependent enzymes are also utilized by cells to catalyze dealkylation of alkylated DNA and RNA. *E. coli* AlkB and its human homologues ABH2 and ABH3 oxidatively demethylate DNA bases in which the Fe(IV)-oxo intermediate hydroxylates the lesion, followed by spontaneous hydrolysis to generate the canonical purine or pyrimidine.³² XAS studies on ABH2 demonstrate that (Fe + αKG + Substr)ABH2 contains a 5-coordinate Fe(II) center, as suggested by the consensus mechanism.³³

Autohydroxylation

In the absence of prime substrate, many of the αKG hydroxylases undergo autohydroxylation (for review see ref. 34).³⁵⁻⁴⁰ In these self-hydroxylation reactions, O₂ activation leads to oxidative decarboxylation of αKG and the ferryl intermediate. As substrate is absent, the ferryl intermediate attacks a nearby amino acid group with subsequent rebound forming a hydroxylated residue. Autohydroxylation chemistry indicates that the oxidative and reductive half reactions are separable, as proposed by the steps of HAT followed by rebound in the consensus mechanism.

Chlorination

The Fe(II)/αKG dependent halogenases are related family of enzymes, distinguished by the absence of a protein derived carboxylate ligand. These enzymes likely evolved from the αKG hydroxylases as a halide atom coordinates the iron cofactor in place of the carboxylate ligand typically found in the facial triad. The chemical mechanism of these halogenases is thought to be very similar to that of the αKG hydroxylases, in which oxidative decarboxylation of αKG leads to the ferryl intermediate reacting *via* HAT with the substrate. However, the reductive half reaction takes the unique direction of chlorination due to rebound between Cl[•] and the substrate radical; such chlorination chemistry has precedent from small molecule models.⁴¹ Transient and spectroscopic studies of the nonheme iron halogenase from *Streptomyces* sp., CytC3, have characterized two ferryl species that decay

slowly in the presence of deuterated substrate.⁴² Additional studies on Syringomycin halogenase (SyrB2) showed the primary substrate is positioned to allow HAT and favor the subsequent Cl• radical-rebound over •OH rebound.⁴³ Although SyrB2 is capable of •OH rebound, Cl• rebound is highly favored due to proximity between the Fe³⁺-Cl bond and the substrate radical.

Desaturation

Some Fe(II)/ α KG oxygenases such as CAS, CarC and DAOCS are multifunctional, using a common ferryl intermediate to oxidize substrate *via* hydroxylation, ring-closure, and desaturation. While the chemical steps of the desaturation remain to be fully elucidated, the ferryl intermediate has been proposed to undergo HAT to produce an initial substrate radical and Fe³⁺-OH. Subsequently, a second HAT from the substrate has been proposed to lead to formation of the double bond and Fe²⁺-OH₂.⁴⁴

Spatial relationship between Fe(IV)=O and substrate. Ramifications to mechanism

Hydroxylation requires that the oxo group and the C-H bond be in close proximity, both for the HAT step as well as for the rebound step. Consequently the stereochemistry of the ferryl intermediate is the key to reactivity within α KG hydroxylases. Based on DFT calculations of TauD,⁴⁵ both 6-coordinate and 5-coordinate geometries about the Fe(IV)-oxo are of similar stability, the difference between the geometries being whether the succinate binds as a bidentate or monodentate ligand. Two limiting models have been presented to explain the oxo reactivity from the Clock and Anti stereoisomers.^{1,5} In the first model, ligand rearrangement leads to the oxo group appearing proximate to the substrate, which is the position *trans* to F2–3(His). One limit for this model is for an early rearrangement preceding O₂ binding, so that O₂ could bind at a site other than that occupied by the aquo ligand crystallographically. Examples supporting this include the observation of rearranged ligands in both CAS and AlkB under different conditions. Another limit for this first model is for ligand rearrangement after ferryl formation, the ‘ferryl flip’, which was proposed as an explanation for isomerism in CAS - essentially the oxo ligand migrates to the position *trans* to position F2–3(His).⁵ We prefer the second model in which O₂ binding at either position, *trans* to F1–7(His) or F2–3(His), leads to a 5-coordinate ferryl, which overcomes such proximity limitations by allowing the oxo group to form near substrate from both the Anti and the Clock isomer. This second model is supported by DFT calculations⁴⁵ and by the prevalence of both stereoisomers.

Non-heme ferryl centers have a short Fe-O bond length (~ 1.65 Å), as shown by EXAFS analysis of α KG oxygenases and synthetic models.^{46,47} A similar short equilibrium distance of 1.66 Å was reported for the (Fe⁴⁺-O) bond in a variety of heme and non-heme systems by DFT calculations,^{45,48} suggesting that the (Fe⁴⁺-O) bond length is not sensitive to the coordination geometry. Due the short Fe⁴⁺-O bond length, the oxo group must be located close to the substrate for any reaction to occur.

Electronic structure calculations indicate that HAT for a high-spin ferryl occurs through a linear Fe-O-HR geometry,⁴⁸ imposing significant restrictions on ferryl reactivity. This could be met either by a 6-coordinate Fe with the oxo group located directly below the substrate, or by a 5-coordinate Fe in which the oxo group is near to the substrate. Calculations predict that the transition state for oxidative decarboxylation on TauD retains an octahedral Fe,⁴⁸ and the ferryl adopts inequivalent bonding interactions with the two O-atoms of succinate, leading to a geometry that is intermediate between 5 and 6-coordinate.⁴⁵ As a result, the oxo group was calculated to occupy a coordination site near that used to initially bind O₂. In the limit of a 6-coordinate ferryl the Anti stereoisomer would be ideally suited for HAT chemistry while the Clock stereo-isomer is poorly suited for HAT, as substrate always binds near the top of the Fe (Fig. 3). This raises the question of how an enzyme containing the Clock stereoisomer can perform HAT, barring a 5-coordinate ferryl.

Ligand rearrangement could permit a Clock isomer to convert to an Anti isomer at some point during turnover, as both stereoisomers have been observed in CAS and AlkB under different ligand/substrate bound states.^{1,5,7,50} One limit would be a ligand rearrangement prior to O₂ binding, which would lead to O₂ binding, and oxo formation, in the coordination site appropriate for HAT chemistry – that corresponding to the Anti isomer. The existing crystallographic data slightly disfavors this notion, as some (Fe + α KG + Substr) forms of CAS and AlkB possess the Clock stereochemistry. Nevertheless, it appears that the Anti and Clock stereoisomers are of similar enough energy in these enzymes, and by extension others, that the stereoisomers could interconvert during turnover. Another limit would be a ligand rearrangement after oxo formation, the ‘ferryl flip’.⁵ While possible, this limiting model is unlikely as it would require changes in bonding to the double-bonded oxo concomitant with a rearrangement of the succinate ligand.

The second model to explain ferryl reactivity is that the ferryl intermediate is 5-coordinate, which would remove any challenge imposed by the Clock/Anti isomerism, and appears to be well supported by crystallographic and kinetic data. Support for this model comes from DFT calculations and ¹⁸OH₂ exchange into the hydroxylated product. DFT calculations indicate that the 5-coordinate geometry is of similar stability as the 6-coordinate geometry – in fact, the bidentate succinate found in the 6-coordinate ferryl contains Fe-O(succinate) bonds that are very unequal, suggesting that succinate is essentially monodentate.⁴⁵ A number of reports indicate that some fraction of ¹⁸O coming from ¹⁸OH₂ can be incorporated into the product of α KG oxygenases, which has been explained as occurring through exchange of ¹⁸OH₂ with an Fe³⁺-OH intermediate.^{51–53} Although occurring through the Fe³⁺-OH intermediate, this ¹⁸O exchange supports the 5-coordinate ferryl, as such exchange likely requires that ¹⁸OH₂ and OH be simultaneously bound by the Fe³⁺, thereby requiring a monodentate succinate.

The extent of ¹⁸OH₂ exchange is highly variable among the tested enzymes, ranging from 0% to 100% ¹⁸O incorporation into the hydroxylation product. The extreme case is the autohydroxylation of TauD (¹⁸OH₂ incorporated into the catecholate formed at Tyr⁷³), in which 100% of O-atoms incorporated into the hydroxylated product came from water.⁵⁴ At the other extreme is PHD2, for which no ¹⁸OH₂ exchange into product was observed.⁵⁵ The variable ¹⁸O incorporation can be rationalized if one views the lifetime of the Fe³⁺-OH/R•

intermediate as varying with the specific enzyme and reactant - a long-lived intermediate will undergo extensive exchange prior to the rebound step, leading to a high level of ^{18}O incorporation from $^{18}\text{OH}_2$.

Stereochemistry of αKG oxygenase active sites

Although the ligand-set found within the active sites of αKG hydroxylases can lead to a large number possible isomers in the absence of protein linkage, the only isomers found in nature are the chiral pair noted above. This means that the Fe of every αKG oxygenase adopts the same geometrical arrangement of ligands. In order to understand the reasons for this restriction on Fe geometry, we considered each potential isomer using an idealized ligand set with Bailar's method⁵⁶ and then considered the merits of each isomer in the face of protein structure and chemical imperatives.

All of the isomers considered were octahedral Fe complexes coordinated by six ligands; two equivalent histidines, a protein derived carboxylate, one aquo ligand, and a bidentate αKG . Each ligand was treated as a single point, and protein connectivity was ignored. A convenient frame of reference is the keto O-atom of αKG , as this allows us to consider the isomers according to which ligand is *trans* to this O-atom, and the keto O is the highest priority ligand (Fig. 4). There are seven possible geometrical isomers, five of which are chiral.

Each isomer was named using the IUPAC naming convention for metal complexes.⁵⁷ This notation indicates the idealized site symmetry of the metal (O = octahedral), the number of ligands (6) and then two numbers indicating the priority of the ligands located *trans* to the highest priority ligands. The ligands were ranked according to the following priority: αKG keto > αKG carboxylate > Asp/Glu > H_2O > His. When the crystal structure contained a monodentate ligand other than water, this ligand was given the same priority as H_2O , for consistent identification of the exchangeable ligand site.

The geometrical constraint on O_2 reacting with the keto group excludes certain isomers, as the three protein derived ligands can be arranged in a facial (*fac*) or meridional (*mer*) orientation. As O_2 must bind at one of the positions occupied by non-protein derived ligands, only the *fac* isomers enable O_2 to bind in a fashion permitting nucleophilic attack on the keto position. The *mer* isomers cannot support O_2 activation, and therefore cannot be found in a functional oxygenase. This is sufficient to explain why a facial triad is a defining feature of the αKG oxygenases. As the three *fac* isomers are chiral, there are six total isomers which could support O_2 activation.

Of the three *fac* isomers, only one has the keto group of αKG *trans* to the facial triad carboxylate. This is the best arrangement for O_2 activation from an electronics perspective, as it supports good electron flow between the Fe- O_2 p-bonding orbitals and keto group of αKG . The *trans* arrangement also ensures that the keto group remains bound in the resting enzyme, as the potential for Fe-keto – back-bonding is enhanced when the keto group is located *trans* to the anionic ligand of the facial triad – typically Asp. Consequently, the only Fe isomer that is viable for oxygenation is the naturally observed OC-6–35(A/C) isomer.

Solomon *et al.* used calibrated electronic structure calculations of a facial triad coordinating ($\text{Fe}^{2+} + \text{NO} + \alpha\text{KG}$), an $[\text{FeNO}]^7$ system, to gain insight into electronic effects on O_2 -binding in this class of enzyme.⁵⁸ This work showed a strong interaction between Fe^{2+} and the keto group, due to the frontier orbitals of αKG . The αKG -based HOMO was a lone pair delocalized across the C-1 carboxylate and the keto group, which had π in-plane symmetry toward the Fe. The αKG LUMO was a π^* symmetry orbital on the keto group of αKG , which happens to also have π^* symmetry (out of plane) with respect to the Fe–keto bond. As the αKG HOMO pushed significant electron density onto the coordinated NO through an intervening Fe-based d-orbital, it was concluded that αKG plays a dominant role in activating coordinated gas (NO or O_2).

Calculation of the $[\text{FeO}_2]^8$ system, a facial triad coordinating ($\text{Fe}^{2+} + \text{O}_2 + \alpha\text{KG}$), further bolstered the leading role of the keto group of αKG in O_2 activation.⁵⁸ The occupied π^* (in-plane) orbital of O_2 mixes with an empty Fe-based d-orbital, to create a bonding interaction. This same π^* (in-plane) orbital is positioned to donate electron density into the αKG LUMO – leading directly to the nucleophilic attack on the keto group of αKG . A more recent computational study by Neese *et al.* similarly supports a leading role for the αKG HOMO and LUMO in directing O_2 activation.⁵⁹

Our analysis of protein linkage and chemical reactivity indicates that the only geometric isomer that could support O_2 activation is the enantiomeric pair found in nature, which is described by the IUPAC convention as OC-6–35-(A/C). As αKG oxygenases with OC-6–35-C or OC-6–35-A stereochemistry are of similar prevalence, both isomers must be competent for O_2 activation (Table 1).

VAST analysis of αKG oxygenase structures

Tertiary structure alignment of αKG oxygenases was performed using the online Vector Alignment Search Tool (VAST) developed by the National Center for Biotechnology Information,⁶⁰ using the molecular modeling database (MMDB).⁶¹ VAST identifies proteins according to their 3-D structures, and provides great insight into proteins that bear strong structural similarity despite low sequence homology.⁶² As we sought to understand the relationship between the Fe stereochemistry and the secondary coordination sphere within the diverse αKG oxygenase family, we chose several structures as initial VAST queries: FIH (1H2L), CAS (1DS1), PAHX (2A1X), and BBOX1 (3O2G); subsequent iterations included proteins from the jumonji family in order to obtain as many inter-related structures of high structural similarity as possible (Table 1). Structures selected for further evaluation of the second coordination sphere were those with reasonable similarity scores (VAST score > 10), known function, and for which the enzyme was constituted with (Fe + αKG) or (Metal + NOG). PyMOL was used to view these structures.⁶³

The protein structures consisted of a cupin barrel core with variable loops connecting the beta strands. Substrate typically binds at the top of the cupin barrel, with one of the variable loops forming a lid to close the active site. The VAST aligned structures were then truncated to the two 4-stranded b-sheets (front and back) comprising the cupin barrel, as this contained the primary coordination sphere of the Fe and would account for conserved secondary-

sphere interactions with the Fe. The HX(D/E)X_nH facial triad motif was formed by residues at the end of the F1 and F2 strands, providing a common location for the Fe at the tops of the cupin barrel in each protein (Fig. 1 and Fig. 5).

Protein phylogenetic analysis of the cupin-barrel sequences using BLOSUM scoring within the Jalview⁸⁴ alignment software sorted the 25 sequences into 4 families. The jumonji family (FIH-1) and the PHD family (PHD2) contain HIF-hydroxylases, whereas the AlkB family (AlkB), and the small-molecule activating family (CAS) contained additional members of biomedical significance. What is immediately striking is the low level of sequence conservation across these sequences and even within each family, which likely reflects the multitude of substrates recognized by these proteins and the evolutionary divergence of this family.

Classifications of the α KG oxygenases are typically based on the reaction catalyzed,^{6,85,86} however, such divisions can be arbitrary. A sequence profile analysis in 2001 showed ten sequence-based families within the α KG oxygenase super-family;² three of our α KG oxygenase families are similar to those previously reported families. In agreement with Aravind and Koonin, we identified distinct AlkB, PHD, and small-molecule activating families; our alignment differs in that the cupin sequence of DAOCS correlates better with the AlkB family than with the small-molecule activating family. Many of the structures in our PHD family and the jumonji family were not known or had unknown function in 2001, making these distinct families differ from any of those listed in the earlier publication.

Second coordination sphere

The α KG and the Asp/Glu residue of the facial triad form specific close contacts with several residues on the back and front sheets of the cupin barrel. As these residues have the potential to interact with the ligands of the Fe center, they are referred to as the secondary coordination sphere.⁷⁸ Key positions are those that line the cupin interior near the carboxylates of the α KG and the Asp/Glu of the facial triad. We identified 4 residues from the back sheet of the barrel as forming a ‘diamond’ of close contacts with the C1-carboxylate of α KG in a number of structures (Fig. 1 and Fig. 6). These residues come from the upper strands, underscoring the importance of the residues at the top of the barrel in defining the chemical environment of the active site.

B2–6 and B2–8 formed close contacts with the C-1 carboxylate of α KG or the facial triad in every structure. These positions were frequently hydrogen bond donors, with neutral polar residues such as Ser/Thr/Asn highly prevalent, suggesting a functional role for residues on this strand for all α KG oxygenases. Data from the literature suggest that these residues are extremely critical for binding α KG and substrate. A kinetics analysis indicated that $k_{\text{cat}}/K_{\text{M}}$ for α KG decreased by 25-fold upon mutation of the residue at position B2–6(Asn²⁹⁴ → Ala) in FIH-1 (1H2L), indicating that this polar residue was crucial for turnover.⁸⁷

Crystallographic analysis of AlkB (3I2O) indicated that the primary substrate formed close contacts with the side chains of residues at position B2–8(Arg²¹⁰) and B3–3(Leu¹¹⁸), supporting a functional role in substrate binding.⁸⁸ Similarly, for TauD mutating the residue

B2-6(Arg²⁷⁰ → Lys) eliminates all activity for the enzyme,⁸⁹ underscoring the critical requirement of precise polar contacts from the B2 strand.

Two additional positions, B1-3 and B3-3, were in close contact with the C1-carboxylate of αKG in many structures. Due to the antiparallel β-sheet structure, these positions were in close registry with the B2-8 position. A variety of residues were found at B1-3, ranging from Lys to Trp, however the B1 strand is at the top of the barrel, so it is likely that these residues are also involved in substrate recognition, as seen for AlkB.⁸⁸ The residues found at B3-3 were either bulky hydrophobics (Val/Ile), or small polar residues (Ser/Asn). As this beta strand is buried inside the barrel, the presence of polar residues in some structures suggested a functional significance. One study of FIH found a moderate decrease in activity upon mutating the residue at position B3-3(Asn²⁰⁵ → Ala), suggesting that this position was not as important to catalysis as B2-6.⁸⁷

Four other positions exhibited variable contacts with αKG, with reasonable conservation within three of the families. Two positions on the B1 sheet formed hydrogen bonds to αKG in two of the families, whereas two positions were alternated to form a salt-bridge to the C5-carboxylate. Both the AlkB and CAS-like families exhibited close contacts between position B1-5 and the C1-carboxylate of αKG. In the case of the AlkB family, the consensus B1-5(Asn) residue can serve as a hydrogen-bond donor to the C1-carboxylate, which is likely an important contributor to the arrangement of this ligand around the Fe. In contrast, the consensus B1-5(Val/Leu) for the CAS-like family is hydrophobic, simply providing sterics to define the active site shape. A conserved B1-7(Tyr) in the AlkB family formed a hydrogen-bond with the C5-carboxylate of αKG, which serves to anchor the αKG within the active site.

The cationic residue that formed a salt bridge with the C5 carboxylate of αKG, either F3-1(Lys), or B2-2(Arg), was highly conserved within each family. The jumonji-family used F3-1(Lys) exclusively for this buried salt-bridge, whereas the other three families used B2-2(Arg). This buried salt bridge effectively locks the αKG into the active site.

Fe stereochemistry is correlated with second coordination sphere

The nearly equal distribution of Fe stereochemistry suggests that both Clock and Anti isomers are competent for oxygenase chemistry, which is consistent with a 5-coordinate ferryl intermediate as discussed above. Understanding how the active site contributes to O₂ activation requires that key functional groups be identified, and these are likely to reflect the Fe stereochemistry. The structural alignment provides us with insight into key residues that make contacts with αKG, which have further been implicated as having functional significance in a handful of mutagenesis studies.^{7,87,89}

The αKG oxygenases were separately analyzed as Anti and Clock groupings (Fig. 7) in order to identify second coordination sphere interactions that were specific to Fe stereochemistry. Three positions were highly correlated with stereochemistry due to their contacts with αKG. The first two positions are B3-3 and B2-8, which either contact the C1-carboxylate of αKG for Anti isomers, or they form close contacts with the Asp/Glu of the facial triad for Clock isomers. The Anti stereochemistry is typically found when residues at

these positions are capable of hydrogen bond donation (such as Asn or Trp), as the C1 carboxylate of α KG is positioned downward in between the B2 and B3 strands. In contrast, the Clock stereochemistry is typically found when these positions are nonpolar residues. A notable exception is found for B2–8(Arg), which typically leads to a Clock stereochemistry as the C1-carboxylate of α KG must rotate upward to form the salt-bridge with the long side chain of arginine.

The stereochemistry is also highly correlated with the position that forms a salt-bridge to the C5 carboxylate of α KG. Despite the fact that this interaction is remote from those atoms bonded to the Fe, the Clock stereochemistry is correlated with the salt-bridge from B2–2(Arg) whereas the Anti stereochemistry is correlated with a salt-bridge from F3–1(Lys). This may reflect a bias in our VAST alignment, but more likely it reflects the overall positioning of the α KG within the active site.

Overall, the picture which emerges is that the secondary coordination sphere subtly differs between the Anti and Clock stereoisomers, with significant bearing on the structural basis for coupling aquo release to substrate binding (*cf.* Scheme 1).^{91,92} The typical Anti isomer is arranged such that positions B3–3 and B2–6 contact the C1 carboxylate of α KG, whereas position B2–8 is in a close contact with the Asp/Glu of the facial triad (Fig. 8). Two examples of this pattern can be found in FIH (1H2L)⁴ and Rph1 (3OPT).⁷³ In FIH, B2–6(Asn²⁹⁴) and B3–3(Asn²⁰⁵) hydrogen bond to α KG, and B2–8(Trp²⁹⁶) is in close contact (within 4 Å) to both α KG and F1–7(Asp²⁰¹) of the facial triad. Nearly equivalent contacts are found in (Ni + hydroxyglutarate)-Rph1, as B2–6(Ser335) is 4.3 Å away from α KG, B3–3(Ser²⁴³) hydrogen bonds to the hydroxyglutarate, and B2–8(Asn³³⁷) forms hydrogen bonds to both the coordinated aquo ligand as well as F1–7(Glu²³⁷) of the facial triad. This suggests that changes in the position of the B2–8 sidechain induced by substrate binding could be coupled to aquo release.

The second coordination sphere about a typical Clock isomer is distinct from that of an Anti isomer, despite the common close contact between B2–6 and the C1 of α KG. Clock isomers differ from Anti isomers by the structurally favored close contact between the C1-carboxylate of α KG and position B1–3, and the close contact between B3–3 and F1–7(Asp/Glu) (Fig. 9). These contacts are exemplified by (Fe + α KG + Substr)ANS, (PDBID 2BRT)⁷⁴ where B3–3(Thr²³⁹) forms a 2.6 Å hydrogen bond with F1–7(Asp²³⁴) of the facial triad, and B1–3(Lys²¹³) forms an indirect hydrogen bond with α KG; B2–6(Ala³⁰²) and B2–8(Phe³⁰⁴) are nonpolar side chains within 4 Å of α KG. The B1–7–B3–3 hydrogen bond lengthens (2.7 Å) upon product formation in the structure of (Fe + succ + Product)ANS (PDBID 1GP6),⁹⁰ suggesting that strengthening this hydrogen bond upon substrate binding may induce aquo release.

The (Mn + α KG)AlkB structure (3O1M)⁵⁰ provides a notable example of the second coordination sphere in the distinct sub-group of Clock isomers with the B2–8(Arg) residue. This arginine forms hydrogen bonds to both the C1 of α KG as well as to the facial triad F1–7(Asp¹³³), and B2–6(Thr²⁰⁸) forms an indirect hydrogen bond to the coordinated aquo ligand. Comparison to the (Fe + α KG + Substr)AlkB structure (PDBID 3I2O)⁵⁰ indicates that changes in these hydrogen bonds may transduce substrate binding to release of the aquo

ligand. When substrate binds, B2–6(Thr) and B2–8(Arg) reorient to form a new hydrogen bond array, removing polar interactions with the coordinated aquo ligand, thereby inducing aquo release.

Stereochemical flexibility within the active site

The subtle correlations between sequence and stereochemistry suggest that both the Anti and the Clock stereochemistry may be of similar stability within many active sites. Two prominent examples are AlkB and CAS, which have been crystallized as both Anti and Clock stereoisomers under varied substrate/product-bound conditions. These ‘switchable’ active sites strongly suggest that ligands can exchange during turnover, and that the secondary coordination sphere can be fluid.

Multiple substrate/product bound structures have been reported for *E. coli* AlkB, which oxidatively repairs alkylation damage to the purine and pyrimidine bases of DNA. One crystallographic study of the substrate adducts of AlkB proposed a key role for protein flexibility in hydroxylation, and further implied that the Fe stereochemistry changed from Clock to Anti during catalysis.⁷ The Fe center in (Fe + α KG + Substr)AlkB exhibited Clock stereochemistry (PDBID 2FD8), however upon exposure to air these crystals underwent oxidative decarboxylation to form succinate within the active site. This (Fe + Succ + Substr)AlkB structure exhibited bidentate succinate coordination with two H₂O ligands and an achiral Fe center (PDBID 2FDK). Interestingly, the Fe in crystals separately prepared containing (Fe + Succ + Substr)AlkB exhibited a bidentate succinate ligand and adopted the Anti stereochemistry (PDBID 2FDG). These coordination changes at Fe, particularly the differences observed for succinate binding, indicates that ligand reorientation is relatively facile and supports the intermediacy of a monodentate succinate ligand along with a 5-coordinate geometry for the ferryl. Notably, succinate formation was more rapid than substrate hydroxylation for crystals of (Fe + α KG + Substr)AlkB upon air exposure, which the authors attributed to restricted protein flexibility in the crystal.⁷

A second study of AlkB with varied substrates further supports a 5-coordinate ferryl intermediate. Crystals of (Fe + α KG + Substr)AlkB in which a varied substrate was cross-linked to enzyme were grown,⁵⁰ then exposed to air for crystallographic analysis of structures formed upon in situ oxidation. The substrate-bound forms of (Mn + α KG + Substr)AlkB were also prepared, in order to prevent spurious oxidation. The metal (Mn or Fe) in each of four independent (Mn + α KG + Substr)AlkB structures adopted Clock stereochemistry; in contrast the metal exhibited variable stereochemistry in the (Fe + Succ + Prod)AlkB structures.

A variety of substrate-bound structures reveal that the Fe of CAS can also adopt Anti and Clock stereochemistry depending on the specific substrate bound, illustrating the similar stability of these isomers within this active site. CAS is an unusual enzyme in that it catalyzes three sequential steps in clavaminic acid synthesis, each of which requires an initial HAT from closely related substrates. Although the sequence of CAS is more consistent with a secondary coordination sphere expected for the Clock stereoisomer, CAS adopts the Anti stereochemistry as (Fe + α KG)CAS, or when bound to either of two

different substrates as the (Fe + α KG + Substr)CAS form, in which the diffusible ligand is bound *trans* to F2–3(His²⁷⁹) (Table 2).⁶⁵ In contrast, exposing (Fe + α KG + A1)CAS to NO leads to the (Fe + α KG + A1 + NO)CAS which adopts the Clock stereochemistry, in which NO is bound *trans* to F1–5(His¹⁴⁴).⁵ This isomerization was proposed to be a necessary part of the catalytic cycle by the original authors of this report,⁵ who proposed that O₂ would bind at the site occupied by NO, then flip to the site closer to substrate prior to HAT. We favor an alternate explanation in which ferryl formation from either the Clock or Anti isomer leads to a 5-coordinate ferryl, in which the oxo group is positioned in a site between the F1–7(Asp) and succinate, placing it near to the substrate for subsequent HAT chemistry.

Inhibition of Fe(II)/ α KG dependent enzymes

Rational inhibitor design depends on inhibitor structure with the binding properties of the active site residues and steric constraints. The Fe is a focus for inhibitor design due to the potential for tight binding via metal coordination, however selectivity remains a challenge. As second coordination sphere residues serve to orient substrate and cofactor, the second coordination sphere molds the topology of the active site. In fact, it was shown for FIH and TauD that the second sphere residues significantly affect the rate of hydroxylation.^{87,93} Virtual screening of FIH inhibitors revealed that expanding the bulk of the inhibitor match contacts with nearby residues increased potency and selectivity.⁹⁴ The concurrent establishment of multiple hydrogen bonds and hydrophobic interactions are critical binding interactions that also stabilize α KG as well as inhibitors at the active site.⁹⁵

Our VAST analysis supports the notion that second-sphere residues from the B1 and B2 strands define key interactions with α KG and/or inhibitors, influencing the affinity and specificity of inhibitors. In a comparison of inhibitor selectivity toward FIH and PHD2, it was demonstrated that certain functional groups distinguished between the two active sites.^{11,14} It was shown that bulky and chiral inhibitors can confer selectivity between FIH and PHD2, as derivatization of the N-oxalylglycine (NOG) framework with chiral amino acid side chains showed that all D-isomers were stronger inhibitors of FIH over PHD2.¹¹ In related studies demonstrating the potential for further inhibitor selectivity, stereospecific derivatization at the C1 position of NOG did not increase potency when tested with the JMJD2E enzyme;¹² however, the hydrophobic interactions gained from derivatization improved selectivity between α KG dependent hydroxylases.¹¹ A related study showed that FIH was more indulgent of sterically bulky inhibitors than PHD2, with EPR data suggesting that inhibitors were more conformationally constrained in the PHD2 active site.¹⁴

Relationship to the oxidative halogenases and non- α KG dependent oxygenases

The cupin barrel sequence analysis permits us to extend our predictions of key second sphere residues to functionally diverse enzymes related to the α KG oxygenases, such as the α KG halogenases CytC3 and SyrB2 and the non- α KG dependent oxygenases cysteine dioxygenase (CDO) and isopenicillin dioxygenase (IPNS). These enzymes were aligned with PHD2 (3OUJ) using VAST, then the cupin barrel sequences were analyzed using

BLOSUM to identify nearest neighbors from the α KG oxygenases. A number of correlations between sequence and chemistry were evident (Fig. 10).

Oxidative halogenases SyrB2 and CytC3

The α KG-dependent halogenases contain an active site Fe^{2+} that resembles those found in the α KG oxygenases, with the replacement of the B1–7(Asp/Glu) residue by an Ala residue, with a chloride ligand to the Fe.^{96,97} The chemical mechanisms are also similar, with oxidative decarboxylation leading to a ferryl intermediate that performs HAT from the primary substrate, but the product is formed by $\text{Cl}^\bullet$ rebound.⁹⁸ The structural factors that make oxidative decarboxylation lead to chlorination rather than hydroxylation are not clear, but the phylogenetic analysis of the second coordination sphere indicates that there is a unique combination of second coordination sphere residues in the halogenases.

A number of studies using alternative substrates or point mutants to identify the structural determinants of halogenation-hydroxylation in related active sites have been reported, establishing that the difference between a hydroxylase and a halogenase is more involved than the presence or absence of the Asp/Glu ligand from the facial triad. It was shown that SyrB2 will hydroxylate alternate substrates, suggesting that the partitioning between chlorination and hydroxylation is due to the proximity of substrate to the ferryl oxo ligand, which affects the relative efficiency of rebound from the $^\bullet\text{OH}$ and $^\bullet\text{Cl}$ groups.⁴³ Hausinger *et al.*, prepared the Asp¹⁰¹ $\rightarrow$ Ala mutant of TauD, however that enzyme was inactive for both hydroxylation and halogenation;⁹⁹ similar attempts by Schofield *et al.*, to convert FIH into a halogenase by preparing the Asp²⁰¹ $\rightarrow$ Ala mutant were also unsuccessful.¹⁰⁰ Preparation of the Ala $\rightarrow$ Asp and Ala $\rightarrow$ Glu mutants of SyrB2 by Drennan *et al.*, led to enzymes that were similarly inactive for hydroxylation and halogenation.⁹⁷

Comparison of the cupin sequences of the halogenases to the oxygenases, and consideration of Fe stereochemistry, provides insight into residues that may be crucial for halogenase activity. The orientation of the α KG within the active site of SyrB2 leads to an Anti Fe stereochemistry, if we treat the Cl ligand as having the same priority of the Asp/Glu of the facial triad. Key second coordination sphere interactions for Anti stereoisomers are provided by positions B3–3 and B2–8. Anti isomers typically have a small polar residue at position B2–8 forming a hydrogen bond to the C1 of α KG - but SyrB2 and CytC3 both contain B2–8(Arg), which forms a polar interaction with the chloride ligand in SyrB2. We think that B2–8(Arg) is important for Cl binding, however this has been tested only once, resulting in an inactive enzyme. Schofield *et al.* recognized the potential importance of B2–8(Arg), preparing the double mutant of FIH F1–7(Asp $\rightarrow$ Ala), B2–8(Trp $\rightarrow$ Arg), but found that this mutant lost all hydroxylation activity yet made no comment on its competence for halogenation.¹⁰⁰ Nevertheless, the primary substrate for FIH is a long peptide, making this enzyme less than ideal for engineering. Better test cases in our opinion are enzymes with simpler substrates, such as the TauD, in which the second coordination sphere can be mutated to mimic that found in SyrB2 and CytC3.

Cysteine dioxygenase (CDO) and isopenicillin N synthase (IPNS)

CDO is a non-heme Fe(II), non- α KG dependent dioxygenase which incorporates both atoms of O₂ into cysteine to form cysteine sulfinic acid.¹⁰¹ The Fe cofactor of CDO is bound by a 3 His facial triad, deviating from that found in the α KG oxygenases. The crystal structure of (Fe + Substr)CDO (2IC1) reveals bidentate coordination by the substrate, cysteine, with an open coordination site *trans* to position F1-5(His⁸⁶) as the putative site for O₂ binding.¹⁰² As the thiolate group of substrate is bound *trans* to position F3-3(His¹⁴⁰), the putative O₂ coordination site is directly analogous to that found in the Clock isomers of the α KG oxygenases. Alignment of the cupin barrel sequence of CDO with the α KG oxygenases revealed that two key residues at positions typically found to interact with α KG in the α KG oxygenases, B2-8(Tyr¹⁵⁷) and B3-3(Cys⁹³), form a cysteinyltyrosine linkage thought to be important to CDO function.¹⁰³ CDO's chemical mechanism is subject to a number of competing mechanistic proposals, and the significance of the cysteinyltyrosine is not fully known,¹⁰¹ nevertheless, this cysteinyltyrosine linkage conserves the position of second-coordination sphere interactions found in the α KG oxygenases.

IPNS is responsible for the bicyclization of (L- α -aminoadipoyl)-L-cysteinyl-D-valine (ACV) forming bicyclic isopenicillin N. IPNS structure and mechanism is closely related to those of the α KG hydroxylases. IPNS coordinates the active site iron with a 2His-1Asp facial triad. However, unlike the α KG hydroxylases which use α KG as the source of two electrons for O₂ activation, IPNS uses the redox active substrate (ACV) to complete O₂ activation.¹⁰⁴ Although the active site stereochemistry cannot be assigned as the (Fe + Substr)IPNS form of the enzyme contains two aquo ligands, the second coordination sphere bears significant similarity to that of DAOCS. Alignment of the IPNS (PDBID 2WO7) beta barrel to the α KG hydroxylases reveals B2-6(Pro) and B2-8(Phe) residues that forms a part of a substrate-binding pocket, with a B3-3(Thr) residue that hydrogen bonds with the facial triad.

Although there has been some debate over the coordination site in IPNS for O₂, the crystal structure of (Fe + Substr + NO)-IPNS (PDBID 1WO4)¹⁰⁵ shows the thiolate of substrate bound *trans* to F2-3(His²⁷⁰), and NO bound *trans* to F1-7(Asp²¹⁶) of the facial triad, implying that this is the site for O₂ binding under normal turnover conditions.¹⁰⁴ Such a binding site for O₂ is unlike that found for any of the α KG oxygenases discussed herein, as the site *trans* to the (Asp/Glu) residue of the facial triad is always occupied by the keto group of α KG. However the stereochemistry of O₂ binding in the α KG-dependent oxygenases is thought to involve significant orbital interaction with the cis-keto group of the α KG (see above), whereas it is likely that the ACV thiolate dictates the site of O₂ binding in IPNS. Recent crystallography and product analysis using an alternate substrate suggests that O₂ binds *trans* to F1-5(His²¹⁴),¹⁰⁶ in a site expected for a Clock isomer of the α KG oxygenases. This flexibility in O₂ binding by IPNS may result from concerted binding of O₂ to iron and thiolate, or may reflect the dominant role of thiolate in directing O₂ addition.

Conclusions

Structural alignment of the α KG dependent oxygenases using VAST revealed a limited number of interactions with the iron cofactor that constitute a conserved second coordination

sphere. These close contacts with the Fe ligands vary depending on the stereochemistry of the metallocofactor, which has significant ramifications for the structure of the ferryl center, the reactive oxidant, as well as predicting key active site positions impacting the chemistry of closely related enzymes. The competence of both Anti and Clock stereoisomers of Fe points to a ferryl intermediate that is 5 coordinate, and therefore able to react with primary substrate with greater flexibility. The small number of conserved close contacts within the active sites of α KG oxygenases suggests that the α KG-dependent halogenases maintain similar second-sphere interactions for function.

Notes and references

1. Hausinger RP. Crit. Rev. Biochem. Mol. Biol. 2004; 39:21–68. [PubMed: 15121720]
2. Aravind L, Koonin EV. Genome Biol. 2001; 2:1–8.
3. Rosen MD, Venkatesan H, Peltier HM, Bembenek SD, Kanelakis KC, Zhao LX, Leonard BE, Hocutt FM, Wu X, Palomino HL, Brondstetter TI, Haugh PV, Cagnon L, Yan W, Liotta LA, Young A, Mirzadegan T, Shankley NP, Barrett TD, Rabinowitz MH. ACS Med. Chem. Lett. 2010; 1:526–529. [PubMed: 24900242]
4. Elkins JM, Hewitson KS, McNeill LA, Seibel JF, Schlemminger I, Pugh CW, Ratcliffe PJ, Schofield CJ. J. Biol. Chem. 2003; 278:1802–1806. [PubMed: 12446723]
5. Zhang Z, Ren J-S, Harlos K, McKinnon CH, Clifton IJ, Schofield CJ. FEBS Lett. 2002; 517:7–12. [PubMed: 12062399]
6. Hausinger RP. Crit. Rev. Biochem. Mol. Biol. 2004; 39:21–68. [PubMed: 15121720]
7. Yu B, Edstrom WC, Benach J, Hamuro Y, Weber PC, Gibney BR, Hunt JF. Nature. 2006; 439:879–884. [PubMed: 16482161]
8. Müller I, Kahnert A, Pape T, Sheldrick GM, Meyer-Klaucke W, Dierks T, Kertesz M, Usón I. Biochemistry. 2004; 43:3075–3088. [PubMed: 15023059]
9. Clifton IJ, Doan LX, Sleeman MC, Topf M, Suzuki H, Wilmouth RC, Schofield CJ. J. Biol. Chem. 2003; 278:20843–20850. [PubMed: 12611886]
10. Banerji B, Garcia AC, McNeill LA, McDonough MA, Buck MRG, Hewitson KS, Oldham NJ, Schofield CJ. Chem. Commun. 2005:5438–5440.
11. McDonough MA, McNeill LA, Tilliet M, Papamicael CA, Chen QY, Banerji B, Hewitson KS, Schofield CJ. J. Am. Chem. Soc. 2005; 127:7680–7681. [PubMed: 15913349]
12. Rose NR, Ng SS, Mecinovi J, Liénard BTMR, Bello SH, Sun Z, McDonough MA, Oppermann U, Schofield CJ. J. Med. Chem. 2008; 51:7053–7056. [PubMed: 18942826]
13. Rose NR, McDonough MA, King ONF, Kawamura A, Schofield CJ. Chem. Soc. Rev. 2011; 40:4364–4397. [PubMed: 21390379]
14. Flagg SC, Martin CB, Taabazuing CY, Holmes BE, Knapp MJ. J. Inorg. Biochem. 2011; 113:25–30. [PubMed: 22687491]
15. Dunwell JM, Purvis A, Khuri S. Phytochemistry. 2004; 65:7–17. [PubMed: 14697267]
16. Hegg EL, Que L Jr. Eur. J. Biochem. 1997; 250:625–629. [PubMed: 9461283]
17. He PQ, Moran GR. J. Inorg. Biochem. 2011; 105:1259–1272. [PubMed: 21820381]
18. Hanauske-Abel HM, Gunzler V. J. Theor. Biol. 1982; 94:421–455. [PubMed: 6281585]
19. Hoffart LM, Barr EW, Guyer RB, Bollinger JM, Krebs C. Proc. Natl. Acad. Sci. U. S. A. 2006; 103:14738–14743. [PubMed: 17003127]
20. Pavel EG, Zhou J, Busby RW, Gunsior M, Townsend CA, Solomon EI. J. Am. Chem. Soc. 1998; 120:743–753.
21. Bollinger JM, Price JC, Hoffart LM, Barr EW, Krebs C. Eur. J. Inorg. Chem. 2005:4245–4254.
22. Brown CD, Neidig ML, Neibergall MB, Lipscomb JD, Solomon EI. J. Am. Chem. Soc. 2007; 129:7427–7438. [PubMed: 17506560]

23. Neidig ML, Brown CD, Light KM, Fujimori DG, Nolan EM, Price JC, Barr EW, Bollinger JM, Krebs C, Walsh CT, Solomon EI. *J. Am. Chem. Soc.* 2007; 129:14224–14231. [PubMed: 17967013]
24. Price JC, Barr EW, Tirupati B, Bollinger JM Jr, Krebs C. *Biochemistry.* 2003; 42:7497–7508. [PubMed: 12809506]
25. Solomon EI, Brunold TC, Davis MI, Kemsley JN, Lee SK, Lehnert N, Neese F, Skulan AJ, Yang YS, Zhou J. *Chem. Rev.* 2000; 100:235–350. [PubMed: 11749238]
26. Zhou J, Gunsior M, Bachmann BO, Townsend CA, Solomon EI. *J. Am. Chem. Soc.* 1998; 120:13539–13540.
27. Zhou J, Kelly WL, Bachmann BO, Gunsior M, Townsend CA, Solomon EI. *J. Am. Chem. Soc.* 2001; 123:7388–7398. [PubMed: 11472170]
28. Valegard K, van Scheltinga ACT, Lloyd MD, Hara T, Ramaswamy S, Perrakis A, Thompson A, Lee H-J, Baldwin JE, Schofield CJ, Hajdu J, Andersson I. *Nature.* 1998; 394:805–809. [PubMed: 9723623]
29. Neidig ML, Brown CD, Kavana M, Choroba OW, Spencer JB, Moran GR, Solomon EI. *J. Inorg. Biochem.* 2006; 100:2108–2116. [PubMed: 17070917]
30. Price JC, Barr EW, Hoffart LM, Krebs C, Bollinger JM. *Biochemistry.* 2005; 44:8138–8147. [PubMed: 15924433]
31. Price JC, Barr EW, Tirupati B, Bollinger JM, Krebs C. *Biochemistry.* 2003; 42:7497–7508. [PubMed: 12809506]
32. Mishina Y, Duguid EM, He C. *Chem. Rev.* 2006; 106:215–232. [PubMed: 16464003]
33. Giri NC, Sun H, Chen H, Costa M, Maroney MJ. *Biochemistry.* 2011; 50:5067–5076. [PubMed: 21510633]
34. Mantri M, Zhang Z, McDonough MA, Schofield CJ. *FEBS J.* 2012; 279:1563–1575. [PubMed: 22251775]
35. Chen YH, Comeaux LM, Herbst RW, Saban E, Kennedy DC, Maroney MJ, Knapp MJ. *J. Inorg. Biochem.* 2008; 102:2120–2129. [PubMed: 18805587]
36. Liu A, Ho RYN, Que L, Ryle MJ, Phinney BS, Hausinger RP. *J. Am. Chem. Soc.* 2001; 123:5126–5127. [PubMed: 11457355]
37. Koehntop KD, Marimanikkuppam S, Ryle MJ, Hausinger RP. *JBICJ. Biol. Inorg. Chem.* 2006; 63–72.
38. Ge Y, Lawhorn BG, ElNaggar M, Sze SK, Begley TP, McLafferty FW. *Protein Sci.* 2003; 12:2320–2326. [PubMed: 14500890]
39. Sundheim O, Vagbo CB, Bjoras M, Sousa MML, Talstad V, Aas PA, Drablos F, Krokan HE, Tainer JA, Slupphaug G. *EMBO J.* 2006; 25:3389–3397. [PubMed: 16858410]
40. Mantri M, Webby CJ, Loik ND, Hamed RB, Nielsen ML, McDonough MA, McCullagh JSO, Bottger A, Schofield CJ, Wolf A. *Med. Chem. Commun.* 2012; 3:80–85.
41. Kojima T, Leising RA, Yan SP, Que L. *J. Am. Chem. Soc.* 1993; 115:11328–11335.
42. Galonic DP, Barr EW, Walsh CT, Bollinger JM, Krebs C. *Nat. Chem. Biol.* 2007; 3:113–116. [PubMed: 17220900]
43. Matthews ML, Neumann CS, Miles LA, Grove TL, Booker SJ, Krebs C, Walsh CT, Bollinger JM. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:17723–17728. [PubMed: 19815524]
44. Kershaw NJ, Caines MEC, Sleeman MC, Schofield CJ. *Chem. Commun.* 2005:4251–4263.
45. Sinnecker S, Svensen N, Barr EW, Ye S, Bollinger JM, Neese F, Krebs C. *J. Am. Chem. Soc.* 2007; 129:6168–6179. [PubMed: 17451240]
46. Riggs-Gelasco PJ, Price JC, Guyer RB, Brehm JH, Barr EW, Bollinger JM, Krebs C. *J. Am. Chem. Soc.* 2004; 126:8108–8109. [PubMed: 15225039]
47. Rohde J-U, Torelli Sp, Shan X, Lim MH, Klinker EJ, Kaizer Jz, Chen K, Nam W, Que L. *J. Am. Chem. Soc.* 2004; 126:16750–16761. [PubMed: 15612713]
48. Ye S, Neese F. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108:1228–1233. [PubMed: 21220293]
49. Elkins JM, Hewitson KS, McNeill LA, Seibel JF, Schlemminger I, Pugh CW, Ratcliffe PJ, Schofield CJ. *J. Biol. Chem.* 2003; 278:1802–1806. [PubMed: 12446723]

50. Yi CQ, Jia GF, Hou GH, Dai Q, Zhang W, Zheng GQ, Jian X, Yang CG, Cui QA, He CA. *Nature*. 2010; 468:330–333. [PubMed: 21068844]
51. Baldwin JE, Adlington RM, Crouch NP, Pereira IAC, Aplin RT, Robinson C, Soc JChem. *Chem. Commun.* 1993:105–108.
52. Lindblad B, Lindsted G, Lindsted S. *J. Am. Chem. Soc.* 1970; 92:7446–7449. [PubMed: 5487549]
53. Sabourin PJ, Bieber LL. *J. Biol. Chem.* 1982; 257:7468–7471. [PubMed: 7085633]
54. Ryle MJ, Liu A, Muthukumaran RB, Ho RY, Koehntop KD, McCracken J, Que L Jr, Hausinger RP. *Biochemistry*. 2003; 42:1854–1862. [PubMed: 12590572]
55. Wu M, Begley TP, Myllyharju J, Kivirikko KI. *Bioorg. Chem.* 2000; 28:261–265. [PubMed: 11133144]
56. Bailar JC Jr. *J. Chem. Educ.* 1957; 34:334–338.
57. Block, BP.; Powell, WH.; Ferneliuss, WC. *Inorganic Chemical Nomenclature: Principles and Practice*. Washington, D.C: The American Chemical Society; 1990.
58. Diebold AR, Brown-Marshall CD, Neidig ML, Brownlee JM, Moran GR, Solomon EI. *J. Am. Chem. Soc.* 133:18148–18160. [PubMed: 21981763]
59. Ye SF, Riplinger C, Hansen A, Krebs C, Bollinger JM, Neese F. *Chem.-Eur. J.* 2012; 18:6555–6567. [PubMed: 22511515]
60. NCBI – National Center for Biotechnology Information, U.S. National Library of Medicine, 8600 Rockville Pike 20894 USA Bethesda MD.
61. Madej T, Address KJ, Fong JH, Geer LY, Geer RC, Lanczycki CJ, Liu CL, Lu SN, Marchler-Bauer A, Panchenko AR, Chen J, Thiessen PA, Wang YL, Zhang DC, Bryant SH. *Nucleic Acids Res.* 2012; 40:D461–D464. [PubMed: 22135289]
62. Gibrat J-F, Madej T, Bryant SH. *Curr. Opin. Struct. Biol.* 1996; 6:377–385. [PubMed: 8804824]
63. DeLano, WL. *PyMOL™ Incentive Product – Copyright© 2008. DeLano Scientific LLC; 2008.*
64. Leung IKH, Krojer TJ, Kochan GT, Henry L, von Delft F, Claridge TDW, Oppermann U, McDonough MA, Schofield CJ. *Chem. Biol.* 2010; 17:1316–1324. [PubMed: 21168767]
65. Zhang Z, Ren J, Stammers DK, Baldwin JE, Harlos K, Schofield CJ. *Nat. Struct. Mol. Biol.* 2000; 7:127–133.
66. O'Brien JR, Schuller DJ, Yang VS, Dillard BD, Lanzilotta WN. *Biochemistry*. 2003; 42:5547–5554. [PubMed: 12741810]
67. Yang Y, Hu L, Wang P, Hou H, Lin Y, Liu Y, Li Z, Gong R, Feng X, Zhou L, Zhang W, Dong Y, Yang H, Lin H, Wang Y, Chen CD, Xu Y. *Cell Res.* 2010; 20:886–898. [PubMed: 20567261]
68. Han Z, Liu P, Gu L, Zhang Y, Li H, Chen S, Chai J. *Front. Sci.* 2007; 1:52–67.
69. Ng SS, Kavanagh KL, McDonough MA, Butler D, Pilka ES, Lienard BMR, Bray JE, Savitsky P, Gileadi O, von Delft F, Rose NR, Offer J, Scheinost JC, Borowski T, Sundstrom M, Schofield CJ, Oppermann U. *Nature*. 2007; 448:87–91. [PubMed: 17589501]
70. Chen Z, Zhang G. Coordinates deposited in PDB on 07/25/2008.
71. Hong X, Zang J, White J, Wang C, Pan C-H, Zhao R, Murphy RC, Dai S, Henson P, Kappler JW, Hagman J, Zhang G. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:14568–14572. [PubMed: 20679243]
72. Horton JR, Upadhyay AK, Qi HH, Zhang X, Shi Y, Cheng X. *Nat. Struct. Mol. Biol.* 2010; 17:38–43. [PubMed: 20023638]
73. Chang Y, Wu J, Tong XJ, Zhou JQ, Ding J. *Biochem. J.* 2011; 433:295–302. [PubMed: 21067515]
74. Welford RWD, Clifton IJ, Turnbull JJ, Wilson SC, Schofield CJ. *Org. Biomol. Chem.* 2005; 3:3117–3126. [PubMed: 16106293]
75. McDonough MA, Kavanagh KL, Butler D, Searls T, Oppermann U, Schofield CJ. *J. Biol. Chem.* 2005; 280:41101–41110. [PubMed: 16186124]
76. Zhang Z, Butler D, McDonough M, Kavanagh K, Bray J, Ng S, Von Delft F, Arrowsmith C, Weigelt J, Edwards A, Sundstrom M, Schofield C, Oppermann U. deposited in PDB. 2008
77. You Z, Omura S, Ikeda H, Cane DE, Jogl G. *J. Biol. Chem.* 2007; 282:36552–36560. [PubMed: 17942405]
78. Shook RL, Borovik AS. *Inorg. Chem.* 2010; 49:3646–3660. [PubMed: 20380466]

79. Yang C-G, Yi C, Duguid EM, Sullivan CT, Jian X, Rice PA, He C. *Nature*. 2008; 452:961–965. [PubMed: 18432238]
80. Pastore C, Topalidou I, Forouhar F, Yan AC, Levy M, Hunt JF. *J. Biol. Chem.* 2012; 287:2130–2143. [PubMed: 22065580]
81. Valegard K, van Scheltinga ACT, Lloyd MD, Hara T, Ramaswamy S, Perrakis A, Thompson A, Lee H-J, Baldwin JE, Schofield CJ, Hajdu J, Andersson I. *Nature*. 1998; 394:805–809. [PubMed: 9723623]
82. Kato M, Araiso Y, Noma A, Nagao A, Suzuki T, Ishitani R, Nureki O. *Nucleic Acids Res.* 2011; 39:1576–1585. [PubMed: 20972222]
83. Sengoku T, Yokoyama S. *Genes Dev.* 2011; 25:2266–2277. [PubMed: 22002947]
84. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. *Bioinformatics.* 2009; 25:1189–1191. [PubMed: 19151095]
85. McDonough MA, Li V, Flashman E, Chowdhury R, Mohr C, Lienard BMR, Zondlo J, Oldham NJ, Clifton IJ, Lewis J, McNeill LA, Kurzeja RJM, Hewitson KS, Yang E, Jordan S, Syed RS, Schofield CJ. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:9814–9819. [PubMed: 16782814]
86. Loenarz C, Schofield CJ. *Nat. Chem. Biol.* 2008; 4:152–156. [PubMed: 18277970]
87. Saban E, Chen Y-H, Hangasky JA, Taabazuing CY, Holmes BE, Knapp MJ. *Biochemistry.* 2011; 50:4733–4740. [PubMed: 21456582]
88. Yu B, Hunt JF. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:14315–14320. [PubMed: 19706517]
89. Grzyska PK, Ryle MJ, Monterosso GR, Liu J, Ballou DP, Hausinger RP. *Biochemistry.* 2005; 44:3845–3855. [PubMed: 15751960]
90. Wilmouth RC, Turnbull JJ, Welford RWD, Clifton IJ, Prescott AG, Schofield CJ. *Structure.* 2002; 10:93–103. [PubMed: 11796114]
91. Hangasky JA, Saban E, Knapp MJ. *Biochemistry.* 2013
92. Flagg SC, Giri N, Pektas S, Maroney MJ, Knapp MJ. *Biochemistry.* 2012; 51:6654–6666. [PubMed: 22747465]
93. McCusker KP, Klinman JP. *Proc. Natl. Acad. Sci. U. S. A.* 2009; 106:19791–19795. [PubMed: 19892731]
94. Kang J, Park K-S, Chong Y, Choe J. *Bull. Korean Chem. Soc.* 2010; 31:1407–1410.
95. Park H, Ko S, Jeon YH. *J. Mol. Graphics Modell.* 2010; 29:221–228.
96. Kulik HJ, Blasiak LC, Marzari N, Drennan CL. *J. Am. Chem. Soc.* 2009; 131:14426–14433. [PubMed: 19807187]
97. Blasiak LC, Vaillancourt FH, Walsh CT, Drennan CL. *Nature.* 2006; 440:368–371. [PubMed: 16541079]
98. Vaillancourt FH, Yin J, Walsh CT. *Proc. Natl. Acad. Sci. U. S. A.* 2005; 102:10111–10116. [PubMed: 16002467]
99. Grzyska PK, Muller TA, Campbell MG, Hausinger RP. *J. Inorg. Biochem.* 2007; 101:797–808. [PubMed: 17350690]
100. Hewitson KS, Holmes SL, Ehrismann D, Hardy AP, Chowdhury R, Schofield CJ, McDonough MA. *J. Biol. Chem.* 2008; 283:25971–25978. [PubMed: 18611856]
101. Joseph CA, Maroney MJ. *Chem. Commun.* 2007:3338–3349.
102. Ye S, Wu Xa, Wei L, Tang D, Sun P, Bartlam M, Rao Z. *J. Biol. Chem.* 2007; 282:3391–3402. [PubMed: 17135237]
103. McCoy JG, Bailey LJ, Bitto E, Bingman CA, Aceti DJ, Fox BG, Phillips GN. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:3084–3089. [PubMed: 16492780]
104. Roach PL, Clifton IJ, Hensgens CMH, Shibata N, Schofield CJ, Hajdu J, Baldwin JE. *Nature.* 1997; 387:827–830. [PubMed: 9194566]
105. Long AJ, Clifton IJ, Roach PL, Baldwin JE, Rutledge PJ, Schofield CJ. *Biochemistry.* 2005; 44:6619–6628. [PubMed: 15850395]
106. Ge W, Clifton IJ, Stok JE, Adlington RM, Baldwin JE, Rutledge PJ. *J. Am. Chem. Soc.* 2008; 130:10096–10102. [PubMed: 18620394]

Fig. 1. Active sites with Anti (left) and Clock (right) stereochemistry

The beta-strand position for each ligand is indicated (*e.g.*: F1-3 is the third residue of the first strand in the front β sheet) (a) Anti: AtsK (1oii);¹³ (b) Clock: CarC (1nx8).¹⁴

Fig. 2. Topology diagram of the α KG oxygenases

Positions of interest are noted: the second coordination sphere (blue rectangles), the facial triad (red ovals) and residues contacting the C5-carboxylate of α KG (yellow hexagons).

Fig. 3. Schematic of active site stereochemistry within cupin barrels, showing the relative positions of primary substrate (R-H) and the exchangeable ligand site (OH₂).

Fig. 4.
Possible isomers for an iron center coordinated by a bidentate αKG, an aquo ligand, two equivalent histidine residues (H) and a single carboxylate ligand (D).

Fig. 5.
Sequence alignment of the cupin barrel sequences using BLOSUM scoring.

Fig. 6.
Orientation of key residues from the back sheet for Anti and Clock stereoisomers (H = histidiny; D = aspartyl/glutamyl residue).

Fig. 7.
Sequence alignment of the cupin barrel within the Anti (top) and the Clock (bottom) stereoisomers.

Fig. 8. Second coordination sphere interactions in Anti isomers: (a) (Fe + αKG + Substr)FIH (PDBID 1H2L)⁴; (b) (Ni + hydroxyglutarate)Rph1 (3OPT).⁷³

Fig. 9. Second coordination sphere interactions in Clock isomers: (a) (Fe + α KG + Subst)ANS (2BRT);⁷⁴ (b) (Fe + α KG + Subst)AlkB (3I2O).⁸⁸

Fig. 10.
Sequence alignment of the cupin-barrel sequences of selected α KG oxygenases with SyrB2, CytC3, CDO, and IPNS

Scheme 1.
Consensus chemical mechanism for α KG oxygenases.

Table 1

Structures used for the VAST alignment and sequence analysis

PDB ID	Stereo-isomer	Protein name	Bound in active site	Resolution (Å)
1oi1 ⁸	A	Alkylsulfatase	Fe + aKG	2.19
3o2g ⁶⁴	C	Human γ -butyrobetaine dioxygenase	Zn + NOG + substr	1.8
1nx8 ⁹	C	Carbapenem synthase	Fe + aKG + substr	2.3
1ds1 ⁶⁵	A	Clavaminase synthase	Fe + aKG	1.08
1OS7 ⁶⁶	A	Taurine/aketoglutarate dioxygenase	Fe + aKG + substr	2.50
3n9n ⁶⁷	A	Jumonji C-domain-containing histone demethylase	Fe + NOG + substr	2.30
1h2l ⁴⁹	A	Factor inhibiting HIF-1a	Fe + aKG	2.25
2yu1 ⁶⁸	—	Jumonji C-domain-containing histone demethylase 1A	Fe + aKG	2.7
2oq7 ⁶⁹	A	Jmjd2a	Ni + NOG	2.15
3dxu ⁷⁰	—	Jmjd2d	Fe + NOG	2.2
3Ldb ⁷¹	C	Jmjd6	Fe + aKG	2.70
3kvb ⁷²	C	KIAA1718	Ni + NOG	2.7
3opt ⁷³		RPH1	Ni + aKG	2.2
2brt ⁷⁴	C	ANS	Fe + aKG + substr	2.20
2a1x ⁷⁵	C	PAHX	Fe + aKG	2.5
3ouj ³	C	PHD2	Fe + aKG	2.30
3obz ⁷⁶	A	phyhd1	Fe + aKG	2.15
2rds ⁷⁷	A	PuH	Fe + NOG + substr	1.65
3f2O ⁸⁸	C	AlkB	Fe + aKG + substr	1.70
3buc ⁷⁹	C	ABH2	Mn + aKG + substr	2.59
2iuw ³⁹	C	ABH3	Fe + aKG	1.5
3THT ⁸⁰	C	AlkB homolog 8	Mn + aKG	3.01
1RXG ⁸¹	C	DAOCs	Fe + aKG	1.5
3AL6 ⁸²	A	TyW5	Ni + aKG	2.8
3AVS ⁸³	C	Utx/KDM6A	Ni + NOG	1.85

Table 2

Metal stereochemistry in AlkB and CAS structures

	PDBID	Metal-bound ligands	Substrate/product ^a	Stereochemistry
AlkB	3O1M	(Mn + αKG + H ₂ O)	3MeT ^b	Clock ⁵⁰
	3O1S	(Fe + succ + 2H ₂ O)	HO-3MeC	Achiral ⁵⁰
	3O1T	(Fe + succ + H ₂ O)	HO-3MeT	5-Coord. ⁵⁰
	3O1U	(Fe + succ + H ₂ O + ROH)	HO-εA	Pseudo-Anti ⁵⁰
	3O1V	(Fe + succ + H ₂ O)	HO-3MeC	Anti ⁵⁰
AlkB	2FD8	(Fe + αKG + H ₂ O)	TmAT	Clock ⁷
	2FDK	(Fe + succ + 2 H ₂ O)	TmAT	Achiral ⁷
	2FDG	(Fe + succ + H ₂ O)	TmAT	Anti ⁷
	3I2O	(Fe + αKG + H ₂ O)	TmAT	Clock ⁸⁸
CAS	1DS1	(Fe + αKG + H ₂ O)	No substrate	Anti ⁶⁵
	1DRY	(Fe + αKG + H ₂ O)	N-α-L-Acetyl arginine	Anti ⁶⁵
	1DRT	(Fe + αKG)	Proclavaminic acid	Anti ⁶⁵
	1GVG	(Fe + αKG + NO)	Deoxyguanidino-proclavaminic acid (A1)	Clock ⁵

^a εA, 1,N(6)-etheno adenine; 3MeT, N(3)-methyl thymine; 3MeC, N(3)-methyl cytosine in a 13 bp fragment.

^b Clock stereochemistry was also observed for (Mn + αKG + Substr)AlkB: εA (3O1P), 3MeC (3O1R).