

Site-directed Mutagenesis of 2,4-Dichlorophenoxyacetic Acid/ α -Ketoglutarate Dioxygenase

IDENTIFICATION OF RESIDUES INVOLVED IN METALLOCENTER FORMATION AND SUBSTRATE BINDING*

(Received for publication, January 18, 2000)

Deborah A. Hogan^{‡§¶}, Sheila R. Smith^{||}, Eric A. Saari^{||}, John McCracken^{||}, and Robert P. Hausinger^{‡§***‡‡}

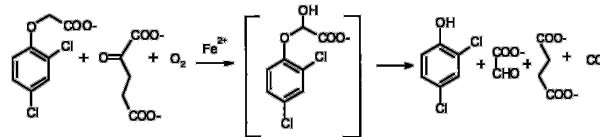
From the [‡]Center for Microbial Ecology and Departments of [§]Microbiology, ^{||}Chemistry, and ^{**}Biochemistry, Michigan State University, East Lansing, Michigan 48824

2,4-Dichlorophenoxyacetic acid (2,4-D)/ α -ketoglutarate (α -KG) dioxygenase (TfdA) is an Fe(II)-dependent enzyme that catalyzes the first step in degradation of the herbicide 2,4-D. The active site structures of a small number of enzymes within the α -KG-dependent dioxygenase superfamily have been characterized and shown to have a similar HXDX_{50–70}HX₁₀RXS arrangement of residues that make up the binding sites for Fe(II) and α -KG. TfdA does not have obvious homology to the dioxygenases containing the above motif but is related in sequence to eight other enzymes in the superfamily that form a distinct consensus sequence (HX(D/E)X_{138–207}HX₁₀R/K). Variants of TfdA were created to examine the roles of putative metal-binding residues and the functions of the other seven histidines in this protein. The H167A, H200A, H213A, H245A, and H262A forms of TfdA formed inclusion bodies when overproduced in *Escherichia coli* DH5 α ; however, these proteins were soluble when fused to the maltose-binding protein (MBP). MBP-TfdA exhibited kinetic parameters similar to the native enzyme. The H8A and H235A variants were catalytically similar to wild-type TfdA. MBP-H213A and H216A TfdA have elevated K_m values for 2,4-D, and the former showed a decreased k_{cat} , suggesting these residues may affect substrate binding or catalysis. The H113A, D115A, MBP-H167A, MBP-H200A, MBP-H245A and MBP-H262A variants of TfdA were inactive. Gel filtration analysis revealed that the latter two proteins were highly aggregated. The remaining four inactive variants were examined in their Cu(II)-substituted forms by EPR and electron spin-echo envelope modulation (ESEEM) spectroscopic methods. Changes in EPR spectra upon addition of substrates indicated that copper was present at the active site in the H113A and D115A variants. ESEEM analysis revealed that two histidines are bound equatorially to the copper in the D115A and MBP-H167A TfdA variants. The experimental data and sequence analysis lead us to conclude that His-113, Asp-115, and His-262

are likely metal ligands in TfdA and that His-213 may aid in catalysis or binding of 2,4-D.

2,4-Dichlorophenoxyacetic acid (2,4-D)¹/ α -ketoglutarate (α -KG) dioxygenase (TfdA) is an Fe(II)- and α -KG-dependent enzyme that catalyzes the first step in degradation of the herbicide 2,4-D. This enzyme couples the oxidative decarboxylation of α -KG to the hydroxylation of a side chain carbon atom. The resultant hemiacetal spontaneously decomposes to form 2,4-dichlorophenol and glyoxalate (1). Mechanistically, TfdA resembles numerous other α -KG-dependent dioxygenases from plants, animals, fungi, and bacteria that catalyze similar hydroxylation reactions at unactivated carbon centers (2, 3).

Members of the α -KG-dependent dioxygenase superfamily are not closely related by their sequences but rather appear to fall into one of three groups of related enzymes or fall into a fourth group of unrelated sequences (4). The best studied α -KG-dependent hydroxylases, including prolyl and lysyl hydroxylase (5) and flavanone hydroxylase (3, 6), have an HXDX₅₅HX₁₀RXS motif in common (7, 8), and this motif is present in the more than 20 enzymes, defined here as Group I, within the α -KG dependent dioxygenase superfamily (8, 9). Site-directed mutagenesis studies have confirmed the importance of these residues for activity (8, 10–13), and the crystal structures of two Group I enzymes, isopenicillin N synthase (IPNS) and deacetoxycephalosporin C synthase (DAOCS), indicate that these residues comprise the metalcenter and α -KG-binding site (14–17). TfdA is not closely related in sequence to the Group I α -KG-dependent dioxygenases described



SCHEME 1

* This research was supported in part by National Science Foundation Grants DEB9120006 and MCB 9603520 (to R. P. H.), National Institutes of Health Grant GM 54065 (to J. M.), and the Michigan State University Agricultural Experiment Station. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[¶] Supported in part by National Institutes of Health Biotechnology Training Grant T32-GM08350.

^{‡‡} To whom correspondence should be addressed: 160 Giltner Hall, Dept. of Microbiology, Michigan State University, East Lansing, MI 48824. Tel.: 517-353-9675; Fax.: 517-353-8957; E-mail: Hausinger@pilot.msu.edu.

above but is clearly homologous (25–30% identity) to *Escherichia coli* taurine/ α -KG dioxygenase (TauD) (18) and sulfonate/ α -KG dioxygenase from *Saccharomyces cerevisiae* (19). Furthermore, PSI-BLAST analyses (20) find additional rela-

¹ The abbreviations used are: 2,4-D, 2,4-dichlorophenoxyacetic acid; α -KG, α -ketoglutarate; TfdA, 2,4-D/ α -KG dioxygenase; TauD, taurine/ α -KG dioxygenase; cw-EPR, continuous wave electron paramagnetic resonance; ESEEM, electron spin echo envelope modulation; IPNS, isopenicillin N synthase; DAOCS, deacetoxycephalosporin synthase; MBP, maltose-binding protein; FT, Fourier transform; MOPS, 4-morpholinepropanesulfonic acid.

TABLE I
Consensus motifs for subgroups within the α -KG-dependent dioxygenase superfamily

Group I (IPNS, DAOCS, and related enzymes)	HXD	X ₅₀₋₇₀	HX ₁₀ (R/K)XS
Group II (TfdA, TauD, clavamate synthase and related enzymes)	HX(D/E)	X ₁₃₈₋₂₀₇	HX ₁₀₋₁₃ R
Group III (Phytanoyl-CoA hydroxylase, proline hydroxylase and related enzymes)	HXD	X ₇₂₋₁₀₁	HX ₁₀ (R/K)XS

tionships to γ -butyrobetaine hydroxylase and clavamate synthase. Alignment of these Group II sequences indicates the conservation of two histidines and one aspartate (His-113, His-262, and Asp 115 in TfdA) as well as an invariant arginine that may be analogous to the α -KG-binding arginine in DAOCS and related enzymes (Table I). A third set of enzyme sequences (Group III) from members of the α -KG-dependent dioxygenase superfamily, including phytanoyl-CoA hydroxylase and proline hydroxylase, exhibit the presence of a third related motif despite the lack of overall sequence similarity to Group I or Group II enzymes.

In this study, we used site-directed mutagenesis methods to examine the roles of potential metal-binding residues in the above motif (His-113, His-262, and Asp-115) and the remaining seven histidines in TfdA. Previously published work showed that TfdA was inactivated by diethylpyrocarbonate, a histidine-selective reagent, and provided evidence consistent with the presence of multiple histidines in the active site (21). Spectroscopic studies of TfdA showed the presence of two equatorially bound imidazole nitrogens as ligands to the active site metal and indicated that one imidazole ligand may be displaced or shifted to an axial position upon substrate binding (22–24). Based on analyses of different TfdA variants, we identify several likely metal ligands and provide evidence that another one or two histidines may aid in substrate binding.

EXPERIMENTAL PROCEDURES

Recombinant Plasmids—All plasmids were constructed from pUS311 (21), a pUC19 derivative that contains the *Ralstonia eutropha* JMP134 *tfdA* gene (Fig. 1). The H8A, D115A, H213A, H216A, H235A, H245A, and H262A TfdA variants were created by direct mutation of *tfdA* in pUS311 by the Stratagene Quickchange System (Stratagene, La Jolla, CA). All mutagenic primers are listed in Table II. Two alternative approaches were used to construct the three remaining variants. Plasmids encoding H113A and H167A TfdA variants were created by CLONTECH mutagenesis of pXHTfdA, a pUC19 plasmid containing the 5'-*XbaI*-*HindII* fragment of the *tfdA* gene (Fig. 1). To create the complete gene containing the indicated mutations, the *XbaI*-*HindII* fragment was cloned into pHKtfdA, which contains the 3' end of the *tfdA* gene. pHKtfdA was constructed in two steps. First, the 1.4-kilobase pair *XbaI*-*SalI* fragment from pUS311, containing the complete *tfdA* gene, was cloned into pBC KS⁺ (Stratagene) cut with *XbaI* and *XhoI* to create pBCtfdA. This step had the benefit of eliminating a *HindII* site that interfered with further cloning steps. The 727-base pair *HindII*-*KpnI* fragment of pBCtfdA was subcloned into pBC KS⁺ cut with the same enzymes to give pHKtfdA. Similarly, the gene encoding H200A TfdA was made by mutagenesis of pXHTfdA, a pUC19 plasmid containing the 5'-*XbaI*-*XhoI* fragment of *tfdA*. The altered *XbaI*-*XhoI* fragment was then inserted into pHKtfdA cut with *XbaI* and *XhoI* to give the complete H200A *tfdA* gene. The identity of all final constructs was confirmed by sequence analysis. To insert the genes encoding H167A and H262A variants of TfdA into a plasmid that would allow for isopropyl-1-thio- β -D-galactopyranoside-controlled expression, the *XbaI*-*SalI* fragments from the corresponding plasmids described above were cloned into pET23a (Novagen) prepared with the same enzymes.

To create the maltose-binding protein (MBP)-TfdA fusion proteins, the wild-type *tfdA* gene was amplified from pUS311 with TfdA-MBPF and TfdA-MBPR primers (Table II) to create an *XbaI* site directly upstream of the GTG start codon of the *tfdA* gene and a *HindIII* restriction site 54 base pairs downstream of the stop codon. The polymerase chain reaction product was cloned directly into the pGEM-T vector according to the manufacturer's instructions (Promega, Madison,

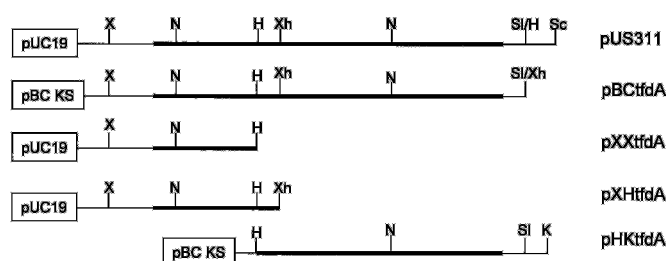


FIG. 1. Restriction maps of the plasmids used in construction of *tfdA* mutants. The abbreviations for the restriction enzyme sites are as follows: *H*, *HindII*; *X*, *XbaI*; *Xh*, *XhoI*; *SI*, *SalI*; *K*, *KpnI*; *Sc*, *SacI*; and *N*, *NruI*. The vector backbone is indicated in the box on the left. The *tfdA* open reading frame is shown as a wide, solid line.

TABLE II
Sequences of mutagenic primers used to create altered *tfdA* genes

Mutant	Primer sequence ^a
H8A	5'-CGC AAA TCC CCT TGC TCC TCT TTT CGC C-3'
H113A	5'-GTC GCT GGC CCA AGC TGG-3'
D115A	5'-GCA CAG CGC CAG CTC CTT TCA-3'
H167A	5'-GCG TGC CGA GCA GTA CGC ACT G-3'
H200A	5'-GGT TCG AAC CGC CGC CGG CTC-3'
H213A	5'-GCT CGC GGC CGC GCC GAT-3'
H216A	5'-CCT TCG ACG GCG CTC GCG-3'
H235A	5'-GCT TCT CGA GGC GAC ACA G-3'
H245A	5'-GTG TAC CGG GCT CGC TGG AAC-3'
H262A	5'-CGT TCT TGC ACG CGG ACG CAG-3'
TfdA-MBPF	5'-TCT CTA GAG TGA GCG TCG TCG CAA ATC C-3'
TfdA-MBPR	5'-GTC AAG CTT GGT TGC GTA CAT CTT GTG G-3'

^a The reverse primer used for the creation of all mutants was the complement of the forward primer.

WI). The *XbaI*-*HindIII* fragment was isolated from the resulting plasmid and cloned into the pMAL-c2 vector that had been digested with the same enzymes. The identity of the newly created *malE*-*tfdA* gene fusion was confirmed by sequencing. Substitution of the mutation-containing internal *NruI* fragment for the same fragment of the wild-type *malE*-*tfdA* gene created MBP-fusion forms of altered TfdAs. First, pMAL-*tfdA* was digested with *NruI*, and the vector fragment was purified and religated to create pMAL-*tfdA*Δ*NruI*. The resultant plasmid was linearized with *NruI* and dephosphorylated with calf intestine alkaline phosphatase prior to ligation with the *NruI* fragments isolated from the previously described mutant genes. Constructs were confirmed by restriction analysis.

Protein Purification—H8A, H113A, D115A, H216A, H235A, and wild-type TfdA proteins were purified from *E. coli* DH5α cells carrying pUS311 and its mutated derivatives according to a previously described protocol (21). In addition, the non-mutated enzyme and the TfdA variants H167A, H200A, H213A, H245A, and H262A were purified as MBP-TfdA fusion proteins from *E. coli* DH5α by the protocol described in the pMAL Protein Fusion and Purification System Manual (New England Biolabs, Beverly, MA).

Analysis of Kinetic Parameters—Specific activities of the wild-type and variant TfdA proteins were determined by a previously described spectrophotometric assay (21). The typical assay mixture contained 1 mM 2,4-D, 1 mM α -KG, 100 μ M (NH₄)₂Fe(SO₄)₂, and 100 μ M ascorbic acid in 10 mM MOPS buffer (pH 6.75) at 30 °C. The reactions were quenched by the addition of EDTA to a concentration of 5 mM. 2,4-Dichlorophenol was quantified by reaction with 4-aminoantipyrene followed by measurement of the absorbance at 510 nm. One unit of activity was defined as the amount of enzyme required to produce 1 μ mol of dichlorophenol·min⁻¹. Protein concentrations were determined using

the Bio-Rad Protein Assay with bovine serum albumin as a standard. For calculation of the k_{cat} values, the TfdA variants were assumed to have $M_r = 31,600$ and the MBP-TfdA variants were assigned $M_r = 74,500$.

The low K_m values for α -KG (~ 2 – $5 \mu\text{M}$ for the wild-type enzyme) precluded use of the 4-aminoantipyrine assay for accurate determination of this value. The alternative method used to measure the K_m values for α -KG quantified the amount of $^{14}\text{CO}_2$ liberated from α -[1- ^{14}C]KG during the course of the reaction (21).

Native Protein Analysis by Gel Filtration—Size exclusion chromatography was used to estimate the native molecular weights of TfdA, MBP-TfdA, and mutant proteins. The proteins were chromatographed on a Superose 6 gel filtration column ($1.0 \times 30 \text{ cm}$, Amersham Pharmacia Biotech) in 20 mM Tris buffer (pH 7.5), 1 mM EDTA, and 200 mM NaCl at a flow rate of $0.2 \text{ ml} \cdot \text{min}^{-1}$. The elution volumes were compared with those for gel filtration standards (Bio-Rad) including thyroglobulin, 670 kDa; bovine gamma globulin, 158 kDa; chicken ovalbumin, 44 kDa; myoglobin, 17 kDa; and vitamin B₁₂, 1350 Da.

Spectroscopic Analysis—Proteins for electron paramagnetic resonance (EPR) and electron spin-echo envelope modulation (ESEEM) spectroscopic analyses were exchanged into 25 mM MOPS (pH 6.75) by repeated concentration and dilution in Centricon 30 (Amicon) centrifugal concentrators. The final subunit concentration was 0.5 mM for the non-fusion forms of TfdA and 0.4 mM for the MBP-TfdA proteins. CuCl_2 was added to a concentration of 450 and 350 μM , respectively. Buffered solutions of α -KG and 2,4-D were added to final concentrations of 5 mM. Glycerol was present at 40% in all samples.

X-band EPR spectra were obtained at 77 K on a Bruker ESP-300E spectrometer. ESEEM data were collected on a home-built spectrometer; the microwave bridge of this instrument has been previously described in detail (25). Data collection and analyses were controlled by a Power Computing model 200 Power PC using software written with LabView version 5.01 (National Instruments). Electron spin echoes were digitized, averaged, and integrated by a Tektronix model 620B digital oscilloscope interfaced to the spectrometer computer via an IEEE-488 bus. Two four-channel delay and gate generators (Stanford Research Systems model DG535), a Bruker BH-15 magnetic field controller, and a Hewlett-Packard model 8656B radiofrequency synthesizer were also interfaced using IEEE-488 protocol. Data were collected using a reflection cavity that employed a folded microstrip resonator (26). A three-pulse stimulated echo sequence (90° - τ - 90° - T - 90°) was used. ESEEM spectra were generated by Fourier transformation of the time domain data using dead time reconstruction (27). Simulations of the experimental data were performed on a Sun SparcII work station. Simulation programs were written in FORTRAN and based on the density matrix formalism developed by Mims (28). Software for the frequency analysis of the experimental and simulated data was written in Matlab (Mathworks, Natick, MA).

Sequence Comparisons—Related sequences were initially detected by BLAST (29) and PSI-BLAST (20) analyses. Alignments were generated with the CLUSTAL algorithm (30), and the figure was prepared using Genedoc (31).

RESULTS

Production of the Mutant TfdAs—Initially, all of the mutant genes were expressed from their pUC19-based plasmids except for those encoding H113A, H167A, and H200A TfdA, which were in pBC KS⁻-derived plasmids. By using the standard protocol to produce soluble, wild-type TfdA (growth at 30 °C to early stationary phase), only the H8A, H113A, D115A, H216A, and H235A variants existed as soluble proteins. All of the other TfdA variants were present as inclusion bodies even when grown at lower temperatures (22 °C), in M9 minimal medium, or in LB broth containing 660 mM sorbitol and 2.5 mM betaine (32). In addition, isopropyl-1-thio- β -D-galactopyranoside-controlled production of H167A and H262A proteins from mutant genes cloned into pET23a did not yield soluble samples even when the harvested cell pellets were suspended in buffer containing 20% glycerol to limit protein aggregation.

To overcome the solubility problems for the five TfdA variants, MBP-TfdA fusion proteins were created. Wild-type TfdA and the MBP-TfdA fusion protein had essentially identical k_{cat} and very similar K_m values for α -KG and 2,4-D (Table III). A slight increase in the apparent K_D for Fe(II) may reflect some

TABLE III
Summary of the kinetic parameters for active TfdA variants

TfdA sample	k_{cat}	K_m α -KG	K_m 2,4-D	K_D Fe(II)
	min^{-1}	μM	μM	μM
Wild type	442 ± 33	4.9 ± 0.73	19.3 ± 3.7	2.0 ± 0.5
MBP fusion	411 ± 23	9.9 ± 1.8	33.7 ± 2.5	15.2 ± 5.4
H8A ^a	284 ± 6.9	6.5 ± 1.8	17.7 ± 3.8	1.9 ± 0.7
MBP-H213A	22.1 ± 2.9	8.3 ± 2.9	318 ± 44	27.7 ± 9.8
H216A	474 ± 47	6.05 ± 2.6	52 ± 5	2.4 ± 1.7
H235A	284 ± 25	9.6 ± 1.6	34 ± 8	1.4 ± 1.7

^a More than 75% of the protein was present as the degradation product. The k_{cat} was estimated from the active, full-length fraction.

metal binding capacity of MBP. Since the presence of the fusion protein did not appear to greatly affect the kinetic parameters of wild-type enzyme, similar fusion proteins were created for the H167A, H200A, H213A, H245A, and H262A TfdA variants.

Kinetic Analyses of Altered TfdAs—Results from kinetic analyses of the four active mutant proteins are summarized in Table III. H8A TfdA was soluble and active but was rapidly proteolyzed to an inactive form. By electrophoretic comparisons, the cleavage site appeared to be the same as in wild-type TfdA (between Arg-77 and Phe-78) (21). The rate of proteolysis of H8A TfdA was enhanced compared with that seen for the wild-type enzyme despite the presence of EDTA and protease inhibitors in the purification buffer. Because purified H8A TfdA was more than 75% degraded, the catalytic rate constant was calculated with the estimated amount of intact enzyme. These calculations indicate rates and K_m values similar to those for the wild-type enzyme. Similarly, the kinetic parameters for H235A TfdA were comparable to the native enzyme. In contrast, two variants exhibited differences from wild-type enzyme in their kinetic parameters. The H213A MBP-TfdA variant exhibited a 20-fold reduction in k_{cat} and a 10-fold increase in K_m for 2,4-D. In addition, H216A TfdA had a modest (2.5-fold) increase in the K_m for 2,4-D and no change in catalytic rate. The other kinetic parameters for H213A MBP-TfdA and H216A TfdA (K_m for α -KG and K_D for ferrous ion) did not differ significantly from the wild-type values. Six soluble TfdA variants (H113A, D115A, MBP-H167A, MBP-H200A, MBP-H245A, and MBP-H262A) exhibited no activity even when assayed with elevated substrate and cofactor concentrations (10 mM α -KG, 5 mM 2,4-D, and 250 μM Fe(II)).

Evaluation of the Structural Consequences of the Mutations—To assess whether the inactive mutant proteins assumed conformations similar to the wild-type enzyme, their apparent molecular weights were estimated by gel filtration analysis. The observed size of wild-type TfdA was found to be 51 kDa by comparison to protein standards, suggesting that TfdA forms a compact dimer or an elongated monomer. The elution volume for both H113A and D115A corresponded exactly to wild-type TfdA indicating that these proteins are not significantly altered in their quaternary structure. MBP-TfdA eluted both in the void volume (approximately 25% of the protein) and at a position corresponding to 216 kDa (roughly 75% of the protein), suggesting that MBP-TfdA forms at least a dimer. Because each MBP-TfdA subunit is comprised of two domains separated by a 13-amino acid linker, the resultant protein may migrate with a larger apparent molecular weight. MBP-H167A and MBP-H200A samples demonstrated the same two-peak profile as MBP-TfdA but with larger proportions eluting in the void volume. MBP-H245A and MBP-H262A proteins were soluble; however, gel filtration analysis indicated the presence of only highly aggregated material eluting in the void volume. Because these mutant proteins exhibited aberrations in their folding properties, the catalytic role of His-245 and His-262, if any, could not be assessed.

EPR Spectroscopic Characterization of Variants with Altered Metal Sites—The metalcenter properties for selected TfdA variants were probed by EPR spectroscopy. To circumvent the problems that arise in EPR measurements of integer spin paramagnetic centers, Fe(II) was substituted with cupric ion. Although the Cu(II) form of TfdA is inactive, Cu(II) binds competitively with respect to Fe(II) ($K_i = 1\text{--}3\ \mu\text{M}$), and copper-substituted TfdA has been used previously to study the metal coordination environment of this enzyme in the presence and absence of substrates (22–24). Spectral parameters of wild-type Cu(II)-TfdA, Cu(II)-TfdA + α -KG, and Cu(II)-TfdA + α -KG + 2,4-D (Fig. 2A and Table IV), agreed well with those reported previously (23, 24). Earlier studies of copper-substituted wild-type TfdA indicated that the metal is bound in a type 2 environment with a mixture of O and N ligands in the equatorial plane. Upon addition of α -KG and 2,4-D to the enzyme, the spectral parameters are altered to a more rhombic signal with accompanying resolution of ligand hyperfine coupling (Fig. 2A). These results suggest that binding of the co-substrates to the enzyme leads to a better defined copper site with α -KG binding directly to the metalcenter (23, 24). The small $A_{\parallel}$ (less than 14 mT) for the α -KG- and 2,4-D-bound sample indicates a significant distortion from planarity.

The four inactive mutant forms of TfdA with quaternary structures similar to the corresponding wild-type protein (H113A, D115A, MBP-H167A, and MBP-H200A) were analyzed by EPR spectroscopy to assess the metal coordination environments. No significant differences between the spectra of MBP-TfdA and the non-fusion wild-type TfdA were observed (data not shown). EPR spectra of the copper-substituted samples, in all cases, showed contributions from multiple copper sites indicating a mixture of copper centers, most likely resulting from copper binding in multiple conformations or at sites other than the active site. The presence of alternative copper-binding sites is not surprising in an enzyme with nine histidines.

The EPR spectra for Cu-H113A TfdA alone and in the presence of α -KG and 2,4-D (Fig. 2B and Table IV) differ significantly from spectra of wild-type enzyme and show modest changes in Cu(II) g values and hyperfine tensor principal values upon substrate additions. The broadening of the EPR signal in the $g_{\parallel}$ region upon addition of α -KG again suggests a mixture of copper site conformations. Thus, it appears that alteration of His-113 significantly affects the metal binding properties for TfdA such that copper is no longer constrained to a single active site configuration.

The EPR spectrum of D115A-TfdA (Fig. 2C) has parameters similar to those observed in the copper-substituted wild-type protein (Table IV), although with less resolution of $A_{\parallel}$. Addition of α -KG (in the presence or absence of 2,4-D) has a dramatic effect on the appearance of the D115A data, enhancing resolution of the Cu(II) hyperfine peaks at $g_{\parallel}$ and the ligand hyperfine structure in the $g_{\perp}$ region (Fig. 2C). The appearance of these superhyperfine interactions may result from a subtle shift in the orientation of the principal g tensor such that the imidazole ligands occupy an increasingly equatorial position (33). These results are consistent with the formation of a tighter or more regular copper-binding site upon addition of the co-substrate.

The EPR spectrum for MBP-H167A TfdA (Fig. 2D), like that for the H113A variant, is poorly resolved, probably due to binding of copper in multiple configurations instead of formation of one major conformation. Additionally, few significant changes are seen upon addition of either α -KG or 2,4-D. The MBP-H200A EPR spectrum (Fig. 2E) is better resolved than that of H113A or MBP-H167A and clearly shows a second set of resonances of lower amplitude with parameters identical to those observed for the copper-substituted wild-type enzyme

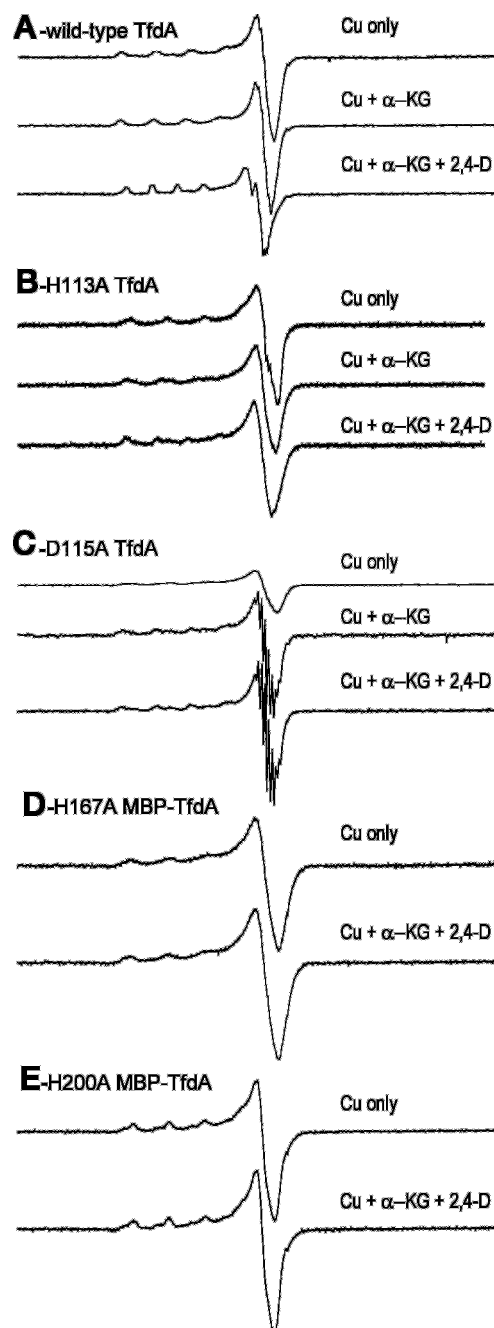


FIG. 2. X-band CW-EPR spectra of Cu(II)-substituted TfdA variants. EPR spectra were obtained at 77 K for the following samples (substrate concentrations were 5 mM, when present). *Top to bottom*, wild-type Cu-TfdA in the absence of substrates, with α -KG, and with α -KG + 2,4-D (A). H113A Cu-TfdA in the absence of substrates, with α -KG, and with α -KG + 2,4-D (B). D115A Cu-TfdA in the absence of substrates, with α -KG, and with α -KG + 2,4-D (C). MBP-H167A Cu-TfdA in the absence of substrates, with α -KG, and with α -KG + 2,4-D (D). MBP-H200A Cu-TfdA in the absence of substrates and in the presence of α -KG + 2,4-D (E). Spectral parameters are listed in Table IV.

(Table IV). The presence of this wild-type signal suggests that His-200 is most likely not a copper-binding ligand in TfdA. As found for H113A and MBP-H167A, addition of α -KG and/or 2,4-D has little effect on the spectrum.

ESEEM Spectroscopic Characterization of Variants with Altered Metal Sites—Pulsed EPR (ESEEM) spectroscopy has proven useful for determining the number of histidines bound to copper in proteins. This approach has been previously used to determine that copper-substituted wild-type TfdA binds cop-

TABLE IV
Summary of the EPR spectral parameters for
Cu(II)-substituted TfdA variants

	$g_{\parallel}$	$A_{\parallel}$	$g_{\perp}$
		mT	
Wild-type TfdA	2.34	16.3	2.07
Wild-type TfdA/ α -KG	2.36	15.1	2.07
Wild-type TfdA/ α -KG/2,4-D	2.38	12.0	2.09
H113A	2.30	16.6	2.07
H113A/ α -KG	2.33	15.8	2.07
H113A/ α -KG/2,4-D	2.35	14.0	2.08
D115A	2.30	17.0	2.06
D115A/ α -KG	2.30	16.2	2.06
	2.34	16.8	2.06
D115A/ α -KG/2,4-D	2.30	16.2	2.06
	2.34	16.8	2.06
MBP-H167A	2.30	17.1	2.07
MBP-H167A/ α -KG/2,4-D	2.30	17.1	2.07
MBP-H200A	2.30	17.1	2.07
	2.34	16.1	2.07
MBP-H200A/ α -KG/2,4-D	2.30	17.1	2.07
	2.34	16.1	2.07

per in a site with two histidyl residues directly coordinated to the metal. As previously reported (23), three-pulse ESEEM spectra collected in the $g_{\parallel}$ region for wild-type TfdA exhibit sharp peaks at 0.6, 0.9, and 1.5 MHz and a broad feature at 3.5 MHz (Fig. 3B). This signature is typical of imidazole bound in an equatorial position to copper. The additional appearance of narrow combination bands at 2.1, 2.5, and 3.1 MHz in these spectra indicated the presence of at least two such imidazole ligands bound to the copper (34–36). Spectral simulations using the density matrix approach of Mims (28) were used to analyze these data further including determination of the number of histidyl ligands coordinated to copper. The $g_{\parallel}$ ESEEM of Cu(II)-substituted wild-type TfdA was simulated with magnetic parameters for copper equatorially coordinated to two identical histidines with a simulation program using the angle selection scheme developed for ENDOR analysis (37, 38). For each simulation, a background decay function was applied to allow the amplitudes of the initial 1.5 μ s of the simulations to match the data. The resulting simulated data sets (*dashed curves*) are superimposed on the experimental data in Fig. 3. Further analysis of these data suggested that formation of the ternary complex with α -KG and 2,4-D results in the probable loss or reorientation of one of these imidazole ligands, as evidenced by a substantial decrease in the modulation intensity and disappearance of the combination bands (23).

Three-pulse ESEEM patterns for the H113A (*trace A*), MBP-H167A (*trace B*), and D115A (*trace C*) variants of TfdA are shown in Fig. 4. Each data set was collected under identical conditions and normalized so that the integrated echo amplitudes range from zero, determined by shifting the integration window off of the signal at the end of a scan, to one which marks the largest measured echo amplitude. Low frequency modulations indicative of equatorially bound histidyl ligand(s) are observed for all four mutant TfdAs. The ESEEM from H113A (*trace A*) and MBP-H200A (not shown) variants are nearly identical and show modulations that are considerably weaker than those measured for MBP-H167A (*trace B*) and D115A (*trace C*). Each of these data contains a large unmodulated or DC component (well over 80% of the total signal intensity in the case of H113A), contrasting sharply with the ESEEM for copper-substituted wild-type TfdA, where the DC component constituted less than 40% of the total signal intensity (Fig. 3A).

The frequency spectra, obtained by Fourier transformation (FT) of the data shown in Fig. 4, are characterized by major

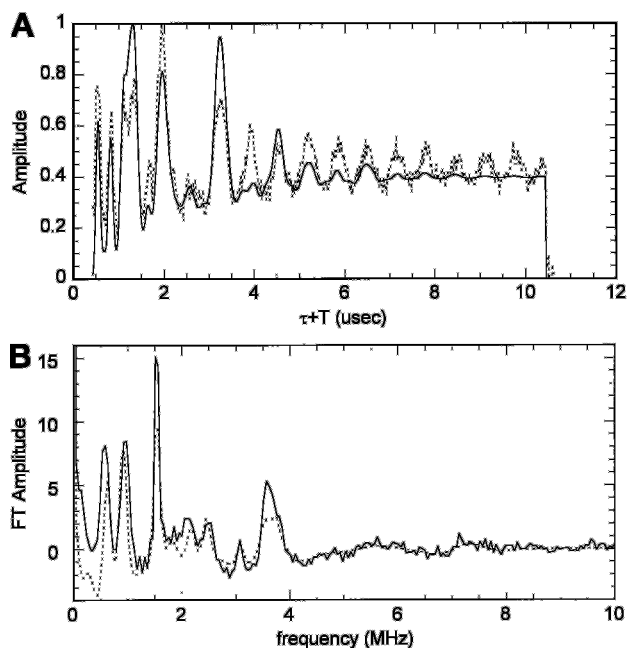


FIG. 3. **Three-pulse ESEEM spectrum and Fourier transformation of Cu(II)-substituted wild-type TfdA with 2-histidine simulation data.** ESEEM spectrum of wild-type Cu(II)-TfdA + α -KG (*solid*) and computer-simulated data for Cu(II) bound by two identical histidines (*dashed*) (A) and the corresponding frequency domain spectrum (B). Parameters for data collections were magnetic field = 2800 G, τ = 375 ns, T = 50 ns, ν = 8.98 GHz; 4.2 K. Cu(II) hyperfine parameters: g_{xy} = 2.05, g_z = 2.34, $A_{xx} = A_{yy} = 20$ MHz, $A_{zz} = 487$ MHz; Superhyperfine parameters: $g_N = 0.40347$, $A_{iso} = 1.7$ MHz, radius = 2.75 Å [θ, ϕ] = [2.37, 2.07] radians, $e^2qQ = 1.60$ MHz, $\eta = 0.75$, NQI [α, β, γ] = [0.50, 1.1, 0.1] radians. The simulated data were treated with an exponential decay function to give $y_{ij} = (y_{ij} - 0.17) \cdot \exp((\tau + T)/3500)^{1.5} + 0.17$. For each simulation, a background decay function was applied to allow the amplitudes of the initial 1.5 μ s of the simulations to match the data.

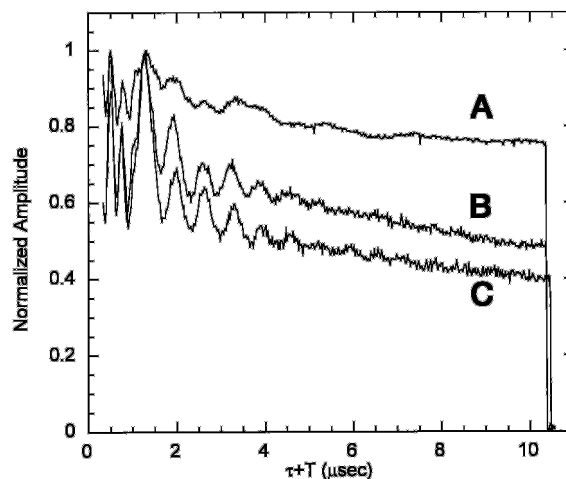


FIG. 4. **Three-pulse ESEEM time domain spectra of TfdA variants.** H113A Cu-TfdA (A), magnetic field = 3050 G, τ = 300 ns, T = 40 ns. MBP-H167A Cu-TfdA (B), magnetic field = 3100 G, τ = 300 ns, T = 40 ns. D115A Cu-TfdA (C), magnetic field = 3050 G, τ = 385 ns, T = 55 ns, ν = 8.8 GHz; 4.2 K.

features at 0.7, 1.5, and 4.0 MHz that are indicative of histidyl imidazole equatorially coordinated to Cu(II) (36). The spectra obtained for D115A and MBP-H167A variants are shown in Figs. 5B and 6B (*solid lines*), respectively. The resolution in these spectra is poor when compared with similar data obtained for rigid Cu(II) proteins or model complexes (39, 40). In Fig. 5A, the normalized three-pulse ESEEM data for D115A TfdA (*solid line*) are shown with the computer-simulated data

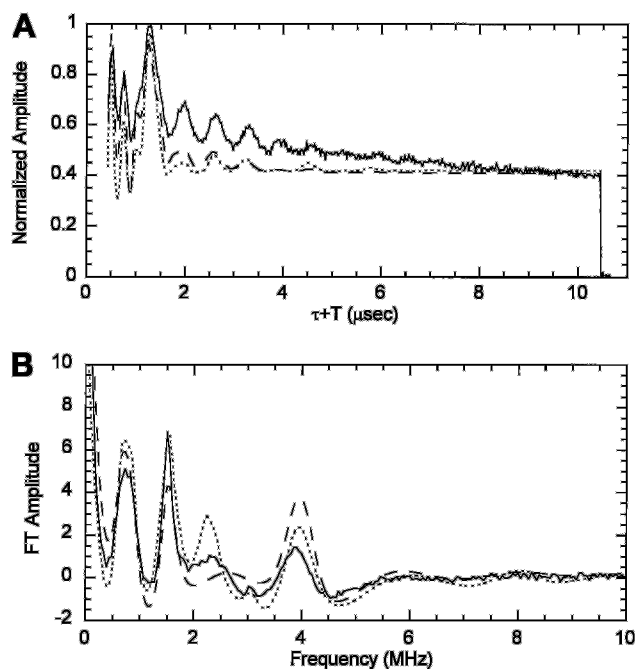


FIG. 5. **Spectral simulation of ESEEM data for D115A-Cu-TfdA.** Experimental (solid line), computer-simulated 1-His (dashed line), and simulated 2-His (dotted line) time domain spectra (A) and the corresponding Fourier-transformed data (B) for copper-substituted D115A TfdA are shown. Magnetic field = 3050 G, $\tau = 230$ ns, $T = 50$ ns, $\nu = 8.8$ GHz, 4.2 K, $g_N = 0.40347$, $e^2qQ = 1.55$ MHz, $\eta = 0.87$, $A_{xx} = 1.6$ MHz, $A_{yy} = 1.9$ MHz, $A_{zz} = 2.2$ MHz. The simulated data were treated with an exponential decay function to give $y_{ij} = (y_{ij} - 0.05) \cdot \exp((\tau + T)/600)^{0.8} - 0.05$.

for one (dashed line) and two (dotted line) coordinated histidyl ligands obtained using spin Hamiltonian parameters typical for the remote nitrogen of imidazole equatorially bound to Cu(II). To compare the amplitudes of the ESEEM simulations with those obtained experimentally, the simulated ESEEM patterns of Fig. 5A were multiplied by background decay functions characterized by an e^{-1} time of 600 ns. This decay time constant is a factor of 6 faster than that used to make the comparison between simulation and experiment for wild-type Cu-TfdA in Fig. 3. Unfortunately, the severity of this decay impairs our ability to distinguish whether 1 or 2 nitrogen nuclei are giving rise to the ESEEM of D115A based on modulation depths.

The Fourier transforms of the experimental and simulated ESEEM functions of Fig. 5A are shown in Fig. 5B. Although the combination lines are weaker in the data for D115A (solid line) than for wild-type TfdA, the combination line at ~ 3.0 MHz is observed, and its frequency is predicted nicely by the 2-His simulation. The broader line widths of the FT spectra found for D115A TfdA, and the subsequent difficulty in observing the combination lines most likely reflects a distribution of ^{14}N superhyperfine couplings for the Cu(II) sites of the variant. To account for the observed line widths and relative peak amplitudes in our simulations, it was necessary to average several calculated data sets obtained with ^{14}N isotropic hyperfine coupling constants that ranged from 1.6 to 1.8 MHz. This assumption of a distribution of hyperfine couplings results in better agreement between experiment and theory for our ESEEM simulations and is in accord with the poor resolution observed in the cw-EPR spectrum of Cu(II)-D115A TfdA in the absence of substrates.

A parallel analysis was undertaken for the MBP-H167A data, and the results are provided in Fig. 6. The dash and dot patterns of Fig. 6A are simulation results for one and two

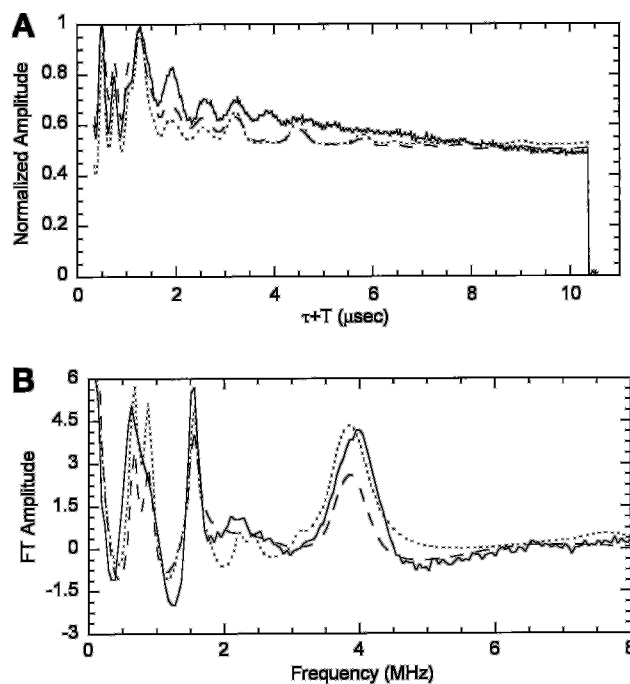


FIG. 6. **Spectral simulation of ESEEM data for MBP-H167A-Cu-TfdA.** Experimental (solid line), computer-simulated 1-His (dashed line), and 2-His (dotted line) time domain spectra (A) and the corresponding Fourier-transformed data (B) are shown. Magnetic field = 3100 G, $\tau = 300$ ns, $T = 40$ ns, $\nu = 8.8$ GHz, 4.2 K, $g_N = 0.40347$, $e^2qQ = 1.60$ MHz, $\eta = 0.82$, $A_{xx} = 1.55$ MHz, $A_{yy} = 1.9$ MHz, $A_{zz} = 2.0$ MHz. The simulated data were treated with an exponential decay function to give $y_{ij} = (y_{ij} - 0.03) \cdot \exp((\tau + T)/350)^{0.5} - 0.03$.

equatorially bound histidyl ligands, respectively. The broadness of the lines at 0.6 and 0.9 MHz results in these frequencies being unresolved in the experimental data. This effect could be simulated using a higher asymmetry parameter, η , for the nuclear quadrupole interaction but not without significant deficiencies in the prediction of the double quantum peak at 4 MHz. A more satisfactory method of accounting for this coalescence in the simulations was through the averaging of calculation results obtained using a distribution of isotropic hyperfine coupling constants, as in the simulations for D115A. The simulations in Fig. 6 are the result of averaging five such data sets where the ^{14}N isotropic hyperfine coupling ranged from 1.6 to 1.9 MHz. This method resulted in the broadening of the low frequency peaks to the point where they began to overlap, without reducing the magnitude of the double quantum feature at 4 MHz. As in the analysis of D115A, different background decay functions were applied to these simulations to provide the best match to the modulation intensities for the initial 1.5 μs of the ESEEM data (solid line). The background decay required to reduce the initial amplitudes of both one and two imidazole simulations was a factor of 10 faster than that used to simulate wild-type Cu(II)-TfdA ESEEM. This problem, combined with the lack of resolved combination peaks in the data, makes it difficult to determine the number of bound histidines. Although the modulation intensities are best accounted for by the one-histidine model, the features of the ESEEM spectrum are better simulated assuming two coupled nitrogens. The ESEEM patterns associated with the H113A and MBP-H200A mutants were too shallow to be consistent with single histidyl coordination, and it is likely that these mutant proteins have multiple Cu(II)-binding sites, some with equatorially coordinated histidine ligand(s) and some without.

The Cu(II)-substituted forms of H113A, D115A, MBP-H167A, and MBP-H200A were also analyzed by ESEEM spec-

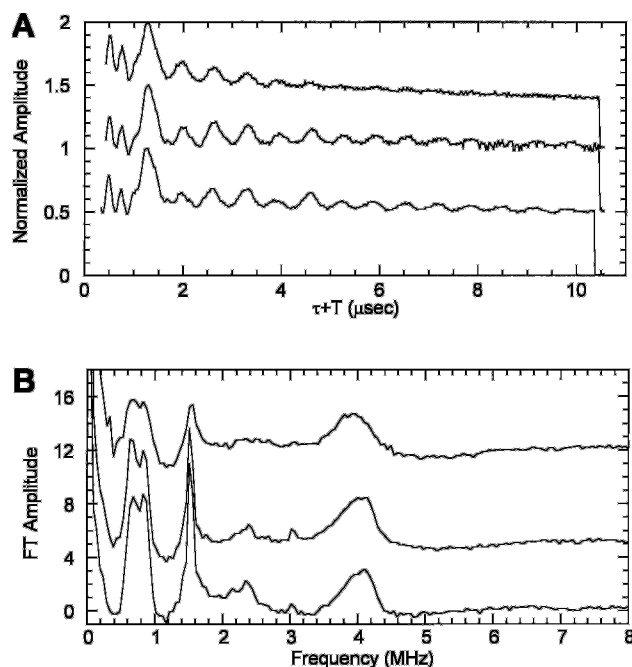


FIG. 7. **Three-pulse ESEEM of D115A TfdA (4.2K).** Top to bottom, D115A Cu-TfdA in the absence of substrate, with 5 mM α -KG, and with 5 mM α -KG + 5 mM 2,4-D (A) and the corresponding frequency domain spectra (B). Conditions for these measurements were identical to those of Fig. 3.

troscopy in the presence of α -KG and 2,4-D. Only slight variations in the line shape of the higher frequency peak at ~ 4.0 MHz were observed for the three histidine mutants. This result contrasts sharply with those for copper-substituted wild-type TfdA where addition of the co-substrates resulted in data that suggested a shift from two equatorially bound histidyl ligands to only one. The lack of similar changes in the ESEEM of these mutants upon addition of the co-substrates supports the conclusion that alteration of these key histidyl residues significantly effects the metal binding properties of the TfdA enzyme. Three-pulse ESEEM spectra for D115A TfdA in the presence of α -KG and α -KG + 2,4-D (Fig. 7) show that addition of co-substrates leads to improved resolution in the ESEEM spectra without significant changes in the amount of unmodulated intensity. These results mirror the improved resolution observed in the cw-EPR spectra. For the ternary complex with both co-substrates bound to the D115A variant, all three combination lines from the electron spin manifold that gives rise to the intense low frequency peaks are resolved (Fig. 7B). Furthermore, the cw-EPR spectra of D115A treated with co-substrates show well resolved ligand hyperfine structure at $g_{\perp}$. Taken together, these data indicate that two histidyl ligands remain equatorially bound during addition of the co-substrates and support our proposal that the differences between theory and experiment found for the two-histidine simulation of Fig. 5 most likely stem from a distribution of structures present prior to substrate addition.

DISCUSSION

EPR, ESEEM, and extended x-ray absorption fine structure spectroscopies of both the Fe(II)- and Cu(II)-forms of wild-type TfdA had previously indicated that the resting enzyme has two imidazoles and a mixture of bound nitrogens and oxygens as ligands to the metal (22–24). ESEEM and x-ray absorption data indicated that both imidazole ligands are retained upon addition of α KG, whereas addition of 2,4-D leads to the displacement or g tensor reorientation of a histidine out of the equatorial plane (22–24). Our present results confirm the im-

portance of multiple histidines and one aspartate for TfdA activity and allow us to assign functions to selected residues.

Mutation of His-113 or Asp-115 to alanine both eliminates activity and alters the structure of Cu(II) metallocenter site. Both variants show EPR spectra consistent with at least two distinct Cu(II)-binding sites. Substrate binding perturbs these spectra giving rise to better resolution of $g_{\parallel}$ features associated with each metal site suggesting that the copper is likely present at the same binding site as in the wild-type enzyme. His-113 and Asp-115 comprise the HX(D/E) motif that is strictly conserved in enzymes related to TfdA (Fig. 8) and more distantly related α -KG-dependent dioxygenases (Table I) (7). Site-directed mutagenesis studies analogous to those described here demonstrated the importance of this motif in a number of Group I enzymes including aspartyl, lysyl, prolyl, and flavanone hydroxylases (8, 10–13). Furthermore, crystal structures of IPNS and DAOCS established these residues as ligands to the Group I metal centers. Our studies support the proposal that residues in the HX(D/E) motif are ligands to the metal in TfdA and, most likely, other Group II enzymes.

Cu(II)-D115A TfdA ESEEM and EPR spectra (particularly for the case where substrates are bound) indicate the presence of two imidazole ligands at one of the Cu(II) sites. The intensity of the non-modulated echo signal was greater for this variant, as compared with that found for the wild-type enzyme, providing evidence for another copper binding environment that lacks equatorially bound histidines. The absence of the aspartate ligand alters the chemistry that occurs at the active site upon addition of the co-substrates. The geometry of the Cu(II) site shows less dispersion when α -KG and 2,4-D are bound as evidenced by the increased resolution in the EPR spectrum and ESEEM pattern. These observations are consistent with those made for the wild-type Cu(II) TfdA; however, the structural rearrangement triggered by formation of the Cu(II)- α KG-2,4-D ternary complex in wild-type TfdA is not observed for D115A TfdA. Together, these data indicate that Asp-115 is very likely a metal-binding residue and may also be involved in formation of the proper catalytic conformation. Because these observations pertain to the Cu(II) form of TfdA, it is difficult to draw conclusions about the specific electronic role of Asp-115 in the native Fe(II) TfdA based on these observations.

The identity of the second metal-binding histidine, predicted by various spectroscopic studies, is less clear. Spectral simulations of H167A are consistent with the presence of two imidazole nitrogens per copper in this protein. Furthermore, His-167 is replaced with an arginine in the TfdA sequence encoded *tfdA* on pIJB in *Burkholderia cepacia* (GenBankTM accession number AAB47567). Thus, even though this residue most closely matches the spacing observed in the Group I sequences, it is unlikely that His-167 is the third metal ligand to the TfdA iron metallocenter. EPR and ESEEM analyses of H200A TfdA suggest that His-200 is important for creation of the metallocenter. Although His-200 is conserved among the enzymes with high identities to TfdA, such as TauD from *E. coli* and sulfonate/ α -KG dioxygenase from yeast, a corresponding histidine is not found in more distantly related Group II enzymes such as γ -butyrobetaine hydroxylase or clavamate synthase (Fig. 8). Mutational analysis of prolyl 4-hydroxylase (10) provided evidence for a non-ligand histidine at the active site that is important for α -KG binding and for controlling the decarboxylation of α -KG in the absence of substrate. Perhaps His-200 serves a similar function in TfdA, providing an explanation for the changes in active site structure upon its alteration. Sequence comparisons of all Group II enzymes, including several putative dioxygenases with overall similarity to TfdA, reveal that His-262 in TfdA is the only other invariant histidine (Fig.

		8			
TfdA	:	-----MSVVANPLHP--LF-AAGMED---IDLREALGSTEVREIERLEDEKSVLVFRGQP-I---SQDQOI	:	56	
TauD	:	-----MSERLSITPLGP--YI-GAQSISG---ADLTRPLSDNQFEQLYHAILRHQVVFRLRDOA-I---TPQQQR	:	58	
AtsK	:	MSNAALATAPHALELDVHPVAG--RI-GAEIRG---VKLSPLDAAATVEAIQAAIVRHKVIFFRGQTHL---DDQSQE	:	69	
YSD	:	GNFSNLFAKENVGQLKVKKITP--KL-GLEENG---IQLT-DLSDAAKDELALLAQKGVVVFNRQ---NFADGE	:	65	
CarC	:	-----MSEIVKFNPMAS--GF-GAYIDH---RD---FLEAKTETIKNLMRQGFVVVKV--L---DIDSD	:	52	
PvcB	:	-----MAYLS--DQ---PRL---SPLRDEQGNQPRFGLLEPGRPG-MH-VGELP---AQWLK	:	47	
GBBH	:	NQVIIDEATNSLQIVWIDGHQS--KFKIGNIR---EGKVEKNVSNDNKIYELWNSKSL--KD-VPRI---SKSTL	:	65	
GBBH_CEL	:	AVTIGDDG--RLVVQWDDGHAS--AYHPG-WIRAHAYDAQSLAEREAPRPHK-HRWMOGLSLPVD-HGAY---MQDDD	:	69	
SyrB	:	-----MSLQANTAPVFADEQQT--DAPTWPDR--ADPSVRLSLLATGNS-LPVMIEPTADGLD-PVQWASARREAE	:	67	
CS2	:	-----MASPIVDCTPYRDEL---LALASELPEVPRADL-HGFEDAEKTLAARLPEGIAAA-LDTFN	:	56	
CS1	:	-----MTS--VDCTAYGPDL---RALAARLPRTPRADL-YAFIDAAHTAAASLPGALATA-LDTFN	:	54	
TfdA	:	A-----FARNFGP-----LEGGFIKVNQRPS-----RFKYA---ELADISNVSHDGKVAQRDARE-----V--	:	104	
TauD	:	A-----LAQRFG-----ELHIHP-----VYPHA---E-GVDEIIVDTH--NDNPPD-----	:	94	
AtsK	:	G-----FAKLLG-----EPVAHP-----TPPVV---D-GTRYLLQLDG---AQGQR-----	:	103	
YSD	:	PD-YVTEYGRHFCK---LHIHQTSQHPQNNPE---LHITPRP---D-AEEFARVFD---DSTSS---G--	:	117	
CarC	:	T---FRDIYSAYG-----TIVEYAD---EKIGV---GFGYRDTK---LEGEK---GKI	:	91	
PvcB	:	GLARSHHLLLLRGF-----AAFADAESLTRYCHDFGEVML-----WPFQAVLEVEQGAEDHIFA-----NN	:	105	
GBBH	:	SLQFSKNLVKYGV-----IIVDGVETSEATEKLCQSLVPVHD-----TFFGQFWVFSNSATNDEPAYE-----DTA	:	128	
GBBH_CEL	:	TLLEWLLAVRDVGL-----TQLHGVPTEPGALPLAKRISFIRE-----SNFGVLFDRSKADADSNAI-----TA	:	130	
SyrB	:	TLLCRHGAVLFRGF-----DLPSVAAFEGFA-----EAGSP-----GLHGTYGDPKKEGGGRNVYRS-----TP	:	121	
CS2	:	AVGSEGGYLLLRGLPVD-DSELPETPTSTPAPLDRKRLVMEARALAGRRLGLHTGYQELRSGTVYHDVYPSGAHYLSS	:	135	
CS1	:	AEGSEDDGHLRLGLPVEADADLPTTPSSTPAPEDRSLLTMEANLGLVGRRLGLHTGYRELRSGTVYHDVYPSGAHYLSS	:	134	
		113 115		167	
TfdA	:	VGNFANQLPHSSSQPPAARYSML--SAVVVPPSSEDWE--FCD-MRAAYDALP---RDLQSEIE-GLRAEH--YALNS	:	173	
TauD	:	-----NDNWHEDVTIETPAGAIL--AAKELPSTGGDIL--WTS-GIAAYEALS---VPFRQLS-GLRAEHDFRKSFP	:	160	
AtsK	:	-----ANSWHDVTVEAYKASIL--RSVVAPASGGDTV--WAN-TAAAYQELPEPLREKADK-WAVHSNEYDYASLKP	:	173	
YSD	:	-----GNHTDVSRELQPPSYTFF--SVVEGPDGGGDIL--FAD-TIEAFDRLS---KPLQDFLSTL---HVIHS	:	176	
CarC	:	VTGRGQLPFRHADGGLLSQVDQVFLYAAEIKNVKFRSAT--VCD-HALACQEMPA--HLLRVLE--EETFE--VRVLE	:	161	
PvcB	:	Y-----VPLHWDGMLETVEFQVFHCVDPAGDTDGGRTFSST--PAALQLAD--SSELELWR---RASAR	:	165	
GBBH	:	YG-SDEIGPHDGTEDQTEGIQVFHCL--TPAKTGGDTVLVDSFYCAEKLRLNES--PEDFEIMC-----NTKIS	:	193	
GBBH_CEL	:	F-----NLPHHTDLPTRELQGLQLFHLCL--VNDATGDNSTFVDGFAIAEALRIEA--PAAYRLLC-----ETPVE	:	192	
SyrB	:	YPEREMILPHNESSHLESWERKQWFFCE--QPSRVGSAEPL--A-----DIRQVL--AYLPKEV---ERFES	:	180	
CS2	:	ETSETLLEPHTEMAHILQPNYVMLACS-RADHENRATILV--G-----SVRKAL--PLLDKTR---ARLFD	:	195	
CS1	:	ETSETLLEPHTEMAHRLQPNYVMLACS-RADHERTAATLV--A-----SVRKAL--PLLDERTR---ARLLD	:	194	
		200		213 216	
TfdA	:	RFLGDDTD-YSEAQRNAMP--PVNWPLVR---THAGSG-RKFLFIGA-HAS-HV--EGLPV-----AEGRMLLAE	:	232	
TauD	:	EYKYRKTEEEHQWRWEAVAKNPPLLPVVR---THPVSG-QALFVNEGFTT-RI--VDVSE-----KESEALLSF	:	224	
AtsK	:	DIDPAKLERHRKVFSTSTVYET--EHPVVR---VHPISG-ERALQLGH-FVK-RI--KGYSL-----ADSQHLFAV	:	233	
YSD	:	KEQAENSQRQGGIKRRAPV--THIHLVR---VHPVLK-KCLYVNRASF-RK--VELKR-----QESSELLNF	:	237	
CarC	:	RGYYVDVS-PDGWFKVPVF--TDLGWR---KMLIYF-PFDEGQPASWEP-RI--VGFTD-----HETQAFFQE	:	220	
PvcB	:	Y-QRSAAH-----YSSRS---AAPIVE-RHPRREFPILRFCEPPVEGDASF--INPSEFH--YDGAIEPQRGELLAS	:	228	
GBBH	:	H-HYLEGS-PPGSSIHVSLS--EKPIVE-RNSFGNITQIRF--NPYDRAPFSCNLSSEAS--A--AETIKFYEAYEK	:	259	
GBBH_CEL	:	F-RNKDRH-S-----DYRC---TAVVIA-LDSSGVEIREIL--ANFLRAPFQ-MDAQRMPP-----DYLLAYRR	:	246	
SyrB	:	K--GLLYS-----RT---FTAGVE--PSWESF--FGTSERSV--I--EQRC--REQGT--DFEWLDG	:	225	
CS2	:	RKVPCCVD-----VA---FRGGVDDPGAIAIVKPL---YGDANDPF--LGYDRELLAPEDPA--DKEAVAH	:	251	
CS1	:	RRMPCCVD-----VA---FRGGVDDPGAIAQVKPL---YGDADDPF--LGYDRELLAPEDPA--DKEAVAA	:	250	
		235		245	
TfdA	:	L-LEHATQREFVYRHRWNV--DLVWMDNRCVLE--RG-RRYDISARRELRRATTLDDAVV-----	:	287	
TauD	:	L-FAHITKPEFQVRWRWQPNDAIWDNRVTOH--YA-NADYLPQRIMHRATILGDKPFYRAG-----	:	283	
AtsK	:	L-QGHVTRLENTVRWRWEAGDVAIWDNRATQH--YA-VDDYGTQPRIVRRVTLAGEVPVGVGDQLSRTRKG	:	301	
YSD	:	L-YNLVESSHDLQLRAKWEPSVVIWDMRRVQH--SAVIDWEEPIERHAFRITPQAERPVEDLKFLNDENYPPSSLTLDI	:	314	
CarC	:	L-GAFLKQPRYYYKHFWEDDELIMDMRRVIE--EREENDDDIIVRRLYRGQTADI-----	:	273	
PvcB	:	L-RRCLYHPQAHYHRWRSDDLVIAONLTLHGREAF--AHRAPRHLRRVH-IHAEPALRN--PHL-QRD-----	:	291	
GBBH	:	L-SKICHNPDSNIESLRPSWIFIDNFRILSRITSF--QGYRQMCGCY-LSRDNFMKAIPFL--SGKSQTSL----	:	327	
GBBH_CEL	:	F-IQMTREPRFCFTRRLEAGQLWCFNRRVLRHARDAFD-PASGDRHFQGCY-VDRDELLSR-ILVL-QR-----	:	310	
SyrB	:	D-TLQLRTQCPAVITHPFTGERCFVQVQLH---PYCMGEELREDLLD--MFGPDRL-----PRLVSYGDSIAEDPV	:	293	
CS2	:	L-SQALDDVTGVKLVLP--GDLIATDNFRTTI--ARTPFSRWDGKDRWLHRVYIIRTDNRN---GEL-SGGERAGDTISF	:	321	
CS1	:	L-SKALDEVTEAVYLEP--GDLIATDNFRTTI--ARTPFSRWDGKDRWLHRVYIIRTDNRN---GQL-SGGERAGDVVAF	:	320	

FIG. 8. Sequence comparison of TfdA and other Group II α -KG-dependent dioxygenases. The accession numbers and sequence identifiers are shown in parentheses for the following Group II sequences: *R. eutropha* 2,4-D/ α -KG dioxygenase (TfdA, M16730); *E. coli* taurine/ α -KG dioxygenase (TauD, P37610); putative α -KG dioxygenase from sulfate-ester transport gene cluster in *Pseudomonas putida* (AtsK, AF126201); *S. cerevisiae* sulfonate/ α -KG dioxygenase (YSD, Q12358); *Erwinia carotovora* carbapenem biosynthesis enzyme C (CarC, AAD38231); *Pseudomonas aeruginosa* pyoverdine biosynthesis enzyme (PvcB, AF002222); *Pseudomonas syringae* SyrI, syringomycin biosynthesis enzyme I (SyrB, AAB63253); *Pseudomonas* sp. γ -butyrobetaine hydroxylase (GBBH, P80193); *Caenorhabditis elegans* γ -butyrobetaine hydroxylase (GBBH_CEL, Z66523); *Streptomyces clavuligerus* clavamate synthase 1 and 2 (CS1, A44241; CS2, L06214). Highly conserved residues (black boxes) and residues conserved in more than 75% of the sequences (dark gray boxes) are highlighted. Amino acid numbers for each sequence are shown at the right of each line; mutated TfdA residues are numbered above.

8); thus, we suggest that this is most likely the third metal-binding residue. In support of our assignment of His-262 as a metal ligand, this residue is 10 amino acids away from Arg-273 creating an HX₁₀R motif similar to that in DAOCS and related enzymes (8). This arginine is conserved in the other Group II enzymes (Fig. 8). Together, these data suggest a more general motif for α -KG-dependent dioxygenases, HX(D/E)X_nHX₁₀₋₁₃R where X_n varies with the different subgroups within the superfamily (Group I, $n = 44-65$; Group II $n = 123-153$; and Group III, $n = 74-104$) (4).

His-213 and His-216 may also be present at or near the TfdA active site. MBP-H213A and H216A TfdA variants exhibited increased 2,4-D K_m values, and the former enzyme also demonstrated a decreased catalytic rate. These data may indicate that His-213 (perhaps also assisted by His-216) facilitates 2,4-D binding by electrostatic interactions between the imidazole nitrogen and the substrate carboxylate group; however, additional studies are required to test this hypothesis. The importance of the 2,4-D carboxylate for binding to TfdA is indicated by the inability of phenoxyethanol to serve as a substrate (21). Neither His-213 nor His-216 are conserved among other related α -KG-dependent dioxygenases. The divergence of sequences in this region may reflect differences in the residues required to bind substrates unique to each enzyme.

Our results provide no evidence to indicate that His-8, His-235, or His-245 are critical for catalysis. The H8A variant exhibits enhanced protease sensitivity, but the full-length protein possesses near wild-type activity. Similarly, the H235A protein is nearly unaffected in its activity. Neither of these residues is conserved in homologous sequences. As described for the H262A TfdA variant, His-245 appears to be structurally important. Because H262A and H245A derivatives of TfdA exist in non-native, highly aggregated conformations, we were unable to assess directly whether these residues may also play a catalytic role. Although His-245 is not conserved in these sequences, it is in the C-terminal half of the protein which in other α -KG-dependent dioxygenases show a higher level of similarity than the overall sequence comparisons (41). This region might be more constrained for structural reasons.

The HX(D/E)X_nH motif, which is found in all known α -KG dependent dioxygenases, is an example of the more general structural arrangement (2-His-1-carboxylate) found in non-heme Fe(II) enzymes including extradiol dioxygenases and pterindependent hydroxylases (42). It is proposed that these three ligands anchor the iron in the active site of the enzyme while leaving three open coordination sites available for the binding of substrate, cofactor, and oxygen. Several studies suggest that O₂ reacts much more readily with such iron centers after the displacement of solvent molecules by more electronegative substrate molecules (42). In addition, it is the diversity of exogenous ligands that enable non-heme Fe(II)-dependent oxygenases to catalyze such a wide range of reactions.

Our proposed model of the active site of TfdA is shown in Fig. 9. His-113, Asp-115, and a third residue tentatively identified as His-262 are shown as ligands to the metal. As in the crystal structures of both DAOCS and IPNS, water molecules are thought to occupy the three remaining coordination sites in substrate-free enzyme. The model includes a likely binding mode for α -KG based on the DAOCS crystal structure (17). Recent spectroscopic studies of clavaminic synthase (43), TauD (44), and TfdA (24) are consistent with a similar mode of α -KG binding in these enzymes. We have no evidence to suggest direct binding of 2,4-D to the metal center and indicate the presence of a separate 2,4-D-binding site adjacent to the metal in the model. Mutation of His-213 affects both substrate binding and the rate of catalysis and therefore is shown to be

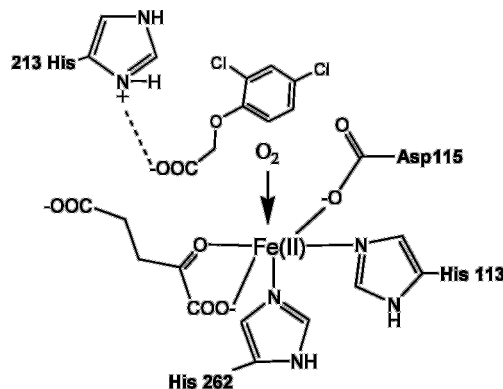


FIG. 9. Model for the active site of TfdA. The three residues proposed to ligate the ferrous ion of TfdA are shown along with a possible coordination mode for the substrates, α -KG and 2,4-D. His-213, implicated in catalysis and/or binding of 2,4-D, is also illustrated.

present near the active site to indicate potential substrate interactions, but the precise role of the residue is still to be defined. In the absence of a crystal structure, the site-directed mutagenesis, metalcenter spectroscopy, and kinetic analysis studies described here provide the most complete working model of the TfdA active site. This model begins to define the necessary elements for catalysis in Group II α -KG-dependent dioxygenases and finds similarities between these enzymes and other well studied representatives of this superfamily.

Note Added in Proof—The recently reported structure of clavaminic synthase (45) confirms our proposed roles for residues in the Hx(D/E)x₁₃₈₋₂₀₇-Hx₁₀ (R/K) motif for the Group II α -KG-dependent dioxygenases.

REFERENCES

1. Fukumori, F., and Hausinger, R. P. (1993) *J. Bacteriol.* **175**, 2083-2086
2. De Carolis, E., and De Luca, V. (1994) *Phytochemistry* **36**, 1093-1107
3. Prescott, A. G., and John, P. (1996) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**, 245-271
4. Hogan, D. A. (1999) *Characterization and Function of 2,4-Dichlorophenoxyacetic Acid/ α -Ketoglutarate Dioxygenase and Related Enzymes*, Ph.D. thesis, Michigan State University, East Lansing, MI
5. Kivirikko, K. I., and Pihlajaniemi, T. (1998) *Adv. Enzymol. Relat. Areas Mol. Biol.* **72**, 328-398
6. Hedden, P. (1992) *Biochem. Soc. Trans.* **20**, 373-377
7. Borovok, I., Landman, O., Kreisberg-Zakarin, R., Aharonowitz, Y., and Cohen, G. (1996) *Biochemistry* **35**, 1981-1987
8. Lukacin, R., and Britsch, L. (1997) *Eur. J. Biochem.* **249**, 748-757
9. Kreisberg-Zakarin, R., Borovok, I., Yanko, M., Aharonowitz, Y., and Cohen, G. (1999) *Antonie Van Leeuwenhoek* **75**, 33-39
10. Myllyharju, J., and Kivirikko, K. I. (1997) *EMBO J.* **16**, 1173-1180
11. Lamberg, A., Pihlajaniemi, T., and Kivirikko, K. I. (1995) *J. Biol. Chem.* **270**, 9926-9931
12. Pirskanen, A., Kaimio, A.-M., Myllylä, R., and Kivirikko, K. I. (1996) *J. Biol. Chem.* **271**, 9398-9402
13. McGinnis, K., Ku, G. M., VanDusen, W. J., Fu, J., Garsky, V., Stern, A. M., and Friedman, P. A. (1996) *Biochemistry* **35**, 3957-3962
14. Lloyd, M. D., Lee, H. J., Harlos, K., Zhang, Z. H., Baldwin, J. E., Schofield, C. J., Charnock, J. M., Garner, C. D., Hara, T., Terwisscha van Scheltinga, A. C., Valegård, K., Viklund, J. A., Hajdu, J., Andersson, I., Danielsson, A., and Bhikhabhai, R. (1999) *J. Mol. Biol.* **287**, 943-960
15. Roach, P. L., Clifton, I. J., Fülöp, V., Harlos, K., Barton, G. J., Hajdu, J., Andersson, I., Schofield, C. J., and Baldwin, J. E. (1995) *Nature* **375**, 700-704
16. Roach, P. L., Clifton, I. J., Hensgens, C. M. H., Shibata, N., Schofield, C. J., Hajdu, J., and Baldwin, J. E. (1997) *Nature* **387**, 827-830
17. Valegård, K., van Scheltinga, A. C. T., Lloyd, M. D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H.-J., Baldwin, J. E., Schofield, C. J., Hajdu, J., and Andersson, I. (1998) *Nature* **394**, 805-809
18. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A., and Leisinger, T. (1997) *J. Biol. Chem.* **272**, 23031-23036
19. Hogan, D. A., Auchtung, T. A., and Hausinger, R. P. (1999) *J. Bacteriol.* **181**, 5876-5879
20. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389-3402
21. Fukumori, F., and Hausinger, R. P. (1993) *J. Biol. Chem.* **268**, 24311-24317
22. Cosper, N. J., Stålhandske, C. M. V., Saari, R. E., Hausinger, R. P., and Scott, R. A. (1999) *J. Biol. Inorg. Chem.* **4**, 122-129
23. Whiting, A. K., Que, L., Jr., Saari, R. E., Hausinger, R. P., Fredrick, M. A., and McCracken, J. (1997) *J. Am. Chem. Soc.* **119**, 3413-3414
24. Hegg, E. L., Whiting, A. K., McCracken, J., Hausinger, R. P., and Que, J. L. (1999) *Biochemistry* **38**, 16714-16726

25. McCracken, J., Shin, D.-H., and Dye, J. L. (1992) *Appl. Mag. Res.* **3**, 305–316
26. Lin, C. P., Bowman, M. K., and Norris, J. R. (1985) *J. Magn. Reson.* **65**, 369–374
27. Mims, W. B. (1984) *J. Magn. Reson.* **59**, 291–306
28. Mims, W. B. (1972) *Phys. Rev. B: Solid State* **5**, 2409–2419
29. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–10
30. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680
31. Nicholas, K. B., Jr., Nicholas, H. B., Jr., and Deerfield, D. W., II (1997) *EMB News* **4**, 14
32. Blackwell, J. R., and Horgan, R. (1991) *FEBS Lett.* **295**, 1–12
33. Rotilio, G., Agrò, A. F., Calabrese, L., Bossa, F., Guerrieri, P., and Mondovì, B. (1971) *Biochemistry* **10**, 616–621
34. McCracken, J., Cornelius, J. B., and Peisach, J. (1989) in *Pulsed EPR: A New Field of Applications* (Keijzers, C. P., Riejerse, E. J., and Schmidt, J., eds) pp. 155–161, Elsevier/North-Holland Biomedical Press, Amsterdam
35. McCracken, J., Pember, S., Benkovic, S. J., Villafranca, J. J., Miller, R. J., and Peisach, J. (1988) *J. Am. Chem. Soc.* **110**, 1069–1074
36. Mims, W. B., and Peisach, J. (1978) *J. Chem. Phys.* **69**, 4921–4930
37. Hoffman, B. M., Martinsen, J., and Venters, R. A. (1984) *J. Magn. Reson.* **59**, 110–123
38. Hurst, G. C., Henderson, T. A., and Kreilick, R. W. (1985) *J. Am. Chem. Soc.* **107**, 7294–7299
39. McCracken, J., Peisach, J., and Dooley, D. M. (1987) *J. Am. Chem. Soc.* **109**, 4064–4072
40. Jiang, F., Peisach, J., Ming, L.-J., Que, L., Jr., and Chen, V. J. (1991) *Biochemistry* **30**, 11437–11445
41. Jia, S., McGinnis, K., VanDusen, W. J., Burke, C. J., Kuo, A., Griffin, P. R., Sardana, M. K., Elliston, K. O., Stern, A. M., and Friedman, P. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 7227–7231
42. Hegg, E. L., and Que, L., Jr. (1997) *Eur. J. Biochem.* **250**, 625–629
43. Pavel, E. G., Zhou, J., Busby, R. W., Gunsior, M., Townsend, C. A., and Solomon, E. I. (1998) *J. Am. Chem. Soc.* **120**, 743–754
44. Ryle, M. J., Padmakumar, R., and Hausinger, R. P. (1999) *Biochemistry* **38**, 15278–15286
45. Zhang, Z., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K., and Schofield, C. J. (2000) *Nat. Struct. Biol.* **7**, 127–133