THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 27, pp. 18723–18733, July 7, 2006

Crystal Structure of Mammalian Cysteine Dioxygenase

*A NOVEL MONONUCLEAR IRON CENTER FOR CYSTEINE THIOL OXIDATION**

Received for publication, February 17, 2006, and in revised form, March 30, 2006 Published, JBC Papers in Press, April 11, 2006, DOI 10.1074/jbc.M601555200

Chad R. Simmons[‡], Qun Liu[§], Qingqiu Huang[§], Quan Hao[§], Tadhg P. Begley[¶], P. Andrew Karplus[‖1], and Martha H. Stipanuk[‡2]

From the [‡]Division of Nutritional Sciences, [§]Macromolecular Diffraction Facility at Cornell High Energy Synchrotron Source (MacCHESS), and the [¶]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-8001 and the [‖]Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331-4501

Cysteine dioxygenase is a mononuclear iron-dependent enzyme responsible for the oxidation of cysteine with molecular oxygen to form cysteine sulfinate. This reaction commits cysteine to either catabolism to sulfate and pyruvate or the taurine biosynthetic pathway. Cysteine dioxygenase is a member of the cupin superfamily of proteins. The crystal structure of recombinant rat cysteine dioxygenase has been determined to 1.5-Å resolution, and these results confirm the canonical cupin β-sandwich fold and the rare cysteinyl-tyrosine intramolecular cross-link (between Cys^{93} and Tyr^{157}) seen in the recently reported murine cysteine dioxygenase structure. In contrast to the catalytically inactive mononuclear Ni(II) metallocenter present in the murine structure, crystallization of a catalytically competent preparation of rat cysteine dioxygenase revealed a novel tetrahedrally coordinated mononuclear iron center involving three histidines (His^{86}, His^{88}, and His^{140}) and a water molecule. Attempts to acquire a structure with bound ligand using either co-crystallization or soaking crystals with cysteine revealed the formation of a mixed disulfide involving Cys^{164} near the active site, which may explain previously observed substrate inhibition. This work provides a framework for understanding the molecular mechanisms involved in thiol dioxygenation and sets the stage for exploration of the chemistry of both the novel mononuclear iron center and the catalytic role of the cysteinyl-tyrosine linkage.

The cytosolic enzyme cysteine dioxygenase (CDO)[3] (EC 1.13.11.20) catalyzes the irreversible oxidation of cysteine to cysteine sulfinate (Reaction 1). This reaction is required for a variety of critical metabolic pathways (1). CDO initiates the catabolism of cysteine to pyruvate and sulfate, which is essential for the provision of adequate inorganic sulfate and allows pyruvate to enter central pathways of metabolism. Also, the oxidation and excretion of the sulfur of methionine depends on CDO, because the sulfur atoms of methionine and homocysteine are only oxidized after their transfer, via the transsulfuration pathway, to serine to yield cysteine. In addition, CDO activity is essential for the biosynthesis of taurine, which is formed by the decarboxylation of cysteine sulfinate to hypotaurine and further oxidation of hypotaurine to taurine.

Clinical evidence indicates that a block in cysteine catabolism, thought to be at CDO, leads to an altered cysteine to sulfate ratio that is associated with sulfate depletion and other adverse effects (1). The prevalence of impaired cysteine catabolism has been reported to be increased in patient populations afflicted with rheumatoid arthritis, liver diseases, Parkinson disease, Alzheimer disease, motor neuron disease, and systemic lupus erythematosus (2–5). These patients frequently exhibit low levels of sulfate in plasma (and in synovial fluid), elevated fasting plasma cysteine concentrations, elevated plasma cysteine to sulfate ratios, and an impaired capacity for sulfation reactions *in vivo*. Reduced cysteine catabolism would cause both depletion of the products sulfate and taurine and an accumulation of the substrate cysteine, either of which would lead to adverse effects. Large doses of cysteine or cystine have been shown to be toxic in several species (6–8). Cysteine is thought to be neuroexcitotoxic, acting via effects on glutamate transport by systems X_C^- and X_{AG}^- and on the *N*-methyl-D-aspartate subtype of the glutamate receptor (9, 10), and cysteine can form toxins by reacting with other compounds (11). Taurine status is associated with sulfur amino acid intake and thus with its synthesis from cysteine, and a lack of adequate taurine has been associated with a number of abnormalities, most commonly with dilated cardiomyopathy, impaired neurological development, and retinal photoreceptor cell abnormalities and photoreceptor cell death (12).

CDO was first described by Ewetz and Sorbo (13), who postulated that it might be a mixed function oxidase. Subsequently, Lombardini *et al.* (14) demonstrated that the enzyme was a dioxygenase and did not require NAD(P)H as an electron donor. CDO was purified from rat liver by Yamaguchi *et al.* (15), who showed it to have a high specificity for cysteine as compared with various cysteine analogs. Little additional work had been done to further characterize the structure or catalytic mechanism of CDO until our recent purification of catalytically active recombinant CDO with kinetic properties that match those observed for CDO in rat liver homogenates: a K_m for cysteine of 0.45 mM, a requirement for ferrous ions, and a pH optimum of 6.1 (16). This study also demonstrated that recombinant CDO is expressed as both active and inactive isoforms, indicating that significant attention to isolation of the active species would be necessary for structural studies.

The function of CDO has been studied most thoroughly in mammals, where it is expressed primarily in liver hepatocytes (17-23). In the rat and mouse, CDO is expressed in a highly tissue-specific manner, but CDO abundance in tissues where it is expressed is regulated largely, if not entirely, by cysteine-mediated regulation of CDO degradation (20–21).

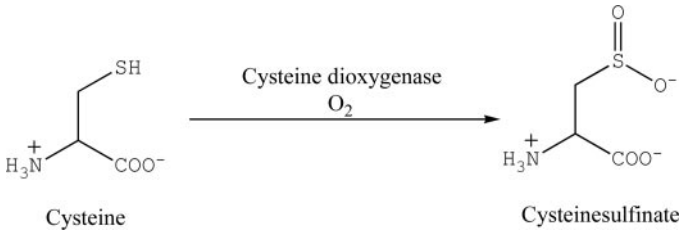
* This work was supported by National Institutes of Health Grants PHS DK056649 (to M. H. S.), PHS DK044083 (to T. P. B.) and RR-01646 (to Q. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2B5H and 2GH2) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence may be addressed: Dept. of Biochemistry and Biophysics, Oregon State University, 2011 Ag. & Life Sciences Bldg., Corvallis, OR 97331. Tel.: 541-737-3200; Fax: 541-737-0481; E-mail: karplusp@science.oregonstate.edu.

[2] To whom correspondence may be addressed: Div. of Nutritional Sciences, Cornell University, 227 Savage Hall, Ithaca, NY 14853. Tel.: 607-255-2683; Fax: 607-255-1033; E-mail: mhs6@cornell.edu.

[3] The abbreviations used are: CDO, cysteine dioxygenase; SAD, single-wavelength anomalous diffraction; PDB, Protein Data Bank; QDO, quercetin dioxygenase.



REACTION 1

The reaction catalyzed by CDO is notably different from those catalyzed by other classes of dioxygenases that have been studied. First, cysteine dioxygenation involves the oxidation of a sulfhydryl group rather than cleavage of a C–C bond or hydroxylation of a carbon atom, and second, both oxygen atoms from the oxygen molecule are transferred to a single sulfur atom rather than distributed between two carbon atoms. These other classes of dioxygenases are: (i) the Fe(II)-containing vicinal oxygen chelate or type I extradiol dioxygenases that catalyze aromatic ring cleavage of catechols at a C–C bond adjacent to an ortho-hydroxyl substituent; (ii) the Fe(II)/Fe-S-center-containing Rieske dioxygenases that catalyze the *cis*-hydroxylation of an arene double bond; (iii) the Fe(III)-containing intradiol dioxygenases that cleave aromatic rings between two carbons that each bear a hydroxyl group; (iv) several transition metal-dependent dioxygenases that belong to the cupin superfamily and cleave C–C bonds; and (v) the α-ketoglutarate-dependent Fe(II) dioxygenases (hydroxylases), many of which have also been described as cupins, that couple the oxidative decomposition of α-ketoglutarate to the hydroxylation of a cosubstrate.

Mammalian CDO was assigned to the cupin superfamily (24) by the presence of two short but partially conserved sequence motifs, $GX_5HXHX_{3-6}EX_6G$ and $GX_{5-7}PXGX_2HX_3N$, that are separated by 29 residues. Proteins in the cupin family have a wide range of enzymatic and biological functions and often show very low overall sequence similarity but share a canonical cupin "jelly roll" β-barrel (25). Determined as part of a structural genomics effort, the recently reported structure of recombinant *Mus musculus* CDO-1 (26) confirmed the cupin fold, revealed the geometry of the active site when it contains a catalytically incompetent nickel ion, and revealed the presence of a rare cysteinyl-tyrosine cross-link.

To give mechanistic work on CDO a firm foundation, we independently initiated crystallographic studies of recombinant *Rattus norvegicus* CDO (identical in sequence to mouse CDO). Here, we describe structures at 1.5-Å resolution of both the native iron-containing CDO and a substrate-inhibited complex. The observed iron metallocenter geometry is distinct from that of the nickel center reported for the mouse CDO structure, and this has major ramifications for mechanistic proposals. The active site geometry reported here provides a framework for understanding the molecular mechanisms involved in thiol dioxygenation and sets the stage for exploring the chemistry of this new type of mononuclear iron center.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Crystallization of CDO—Native *R. norvegicus* cysteine dioxygenase (SwissProt/TrEMBL P21816) was prepared as described previously (27). The purified protein used for crystallization had a k_{cat} of ~43 min^{-1} and a K_m of 0.45 mM for L-cysteine when assayed in the presence of ferrous ions (16). Expression and purification of selenomethionine (Se-Met)-substituted CDO followed a protocol adapted from Doublié (28). BL21(DE3) cells (Novagen) transformed with the pET32a expression vector as described previously (16) were grown in M9 salts supplemented with 2 mM $MgSO_4$, 0.4% glucose, 0.002% thiamine (vitamin B_1), 0.1 mM $CaCl_2$, and 100 μg/ml carbenicillin at 37 °C to A_{600} ~ 0.6. Inhibition of bacterial methionine biosynthesis was targeted by the addition of lysine, phenylalanine, and threonine at 100 mg/liter each, isoleucine, leucine, and valine at 50 mg/liter, and Se-Met at 60 mg/liter. After 15 min, the expression of CDO was induced with 1 mM isopropyl β-D-thiogalactoside, and the cells were incubated at 25 °C overnight. The cells were harvested via centrifugation at 6000 × *g* for 10 min. Cell lysis and protein purification were performed as described for native CDO (16, 27), except all buffers were supplemented with 5 mM dithiothreitol to prevent oxidation of the Se-Met. The final protein concentrations of the purified native and Se-Met CDO that were used for crystallization were 7.5 and 6 mg/ml, respectively.

Crystallization of native and Se-Met CDO was performed as described (27) in sitting drops at 25 °C using a reservoir of 0.1–0.25 M ammonium acetate, 0.1 M tri-sodium citrate, pH 5.6, with 22–26% (w/v) polyethylene glycol 4000 (28). Equivalent crystals could also be grown using a reservoir of 0.15 M ammonium sulfate, 0.2 M sodium cacodylate, pH 6.5, with 26% (w/v) polyethylene glycol 8000, and the co-crystal with 5 mM cysteine were grown under these conditions. In all crystallization setups 1.5 μl of concentrated protein solution was mixed with an equal volume of reservoir solution. Cryomounting of CDO crystals was done as described previously (27).

Crystallographic Data Collection—Crystallographic data collection of native CDO crystals was performed at the National Synchrotron Light Source X12b beamline on an ADSC Q4 CCD detector as reported elsewhere (27). Single wavelength anomalous diffraction (SAD) data on cryo-cooled Se-Met CDO crystals were collected at the selenium K-edge (peak) at the Cornell High Energy Synchrotron Source F2 station using an ADSC Q210 CCD detector. The x-ray wavelength was set at 0.9790 Å based on a fluorescence scan. A total of 360 1° frames (180 + 180 by inverse beam geometry with 5° wedges) were recorded from one crystal. Diffraction data from frozen CDO-cysteine co-crystals were also collected at the Cornell High Energy Synchrotron Source F2 station. All data were reduced using the HKL package (29); the data quality statistics are summarized in Table 1. All CDO crystals used in this study were isomorphous, belonging to space group $P4_32_12$.

Structure Determination and Refinement—The CDO structures were solved by SAD phasing. Four Se sites were located using the SAPI program (30). The correct space group, $P4_32_12$, was selected by using the program ABS (31) based on the four Se sites and SAD data to 3.0-Å resolution. The Se substructure was then fed into the program SOLVE (32) for refinement and phase calculation, resulting in an average figure of merit of 0.43 for all reflections between 20 and 2.3 Å. Combined with the native data set, the 2.3-Å SAD phases were gradually extended to 1.5 Å by solvent flipping implemented in the program SOLOMON (33), and an initial model accounting for 97% of the structure (r = 19.2 R_{free} = 28.0) was automatically built with ArpWarp (34).

Manual model building was performed in O (35), alternating with crystallographic refinement using CNS (Crystallography and NMR System) software (36) until final completion of the model. R_{free} calculations were based on 10% of the reflections. During further refinement, water molecules were added in places with difference density >3.5 r_{rms}, $2F_o - F_c$ density >1.0 r_{rms} and having a reasonable environment. A close approach of Cys^{93}-Sγ to Tyr^{157}-Cϵ2 indicated the presence of a covalent link, and this led us to loosely restrain it to a bond length of 1.93 Å (based on an Sδ–Cϵ bond distance of methionine). Refinement was terminated when the remaining significant difference peaks were associated with alternate conformations of some water sites and of a few

disordered side chains that were not near the active site (His[20] and Val[36]). When refinement was complete, water molecules were numbered based on electron density strength, with Wat1 having the strongest density and Wat339 the weakest.

The structure of the CDO-cysteine co-crystal was solved by difference Fourier, using as the initial model the final native CDO structure with active site waters removed. Clear movements indicated for the side chains of Arg[60], Cys[164], and Met[179] and the backbone near Cys[164] were accounted for manually, but solvent structure in the active site pocket and the density for a molecule apparently covalently attached to Cys[164] were initially left uninterpreted. As refinement progressed, it became apparent that the active site was a mixture of a minor component indistinguishable from the native structure and a major component having Cys[164] in an apparent disulfide link with an unknown ligand. Given this mixture, we decided to model the residual active site density (including the disulfide-linked sulfur site) as a series of water sites at the significant density maxima even though these sites were in some cases too close to each other and to protein atoms. These waters are numbered 401–428. Some residual density for the original native positions of Met[179] and Arg[60] remained. Final statistics for both refined models are given in Table 1.

Coordinates—The atomic coordinates and structure factors of native CDO and the CDO-cysteine co-crystal have been deposited in the Protein Data Bank (PDB) with accession codes 2B5H and 2GH2, respectively.

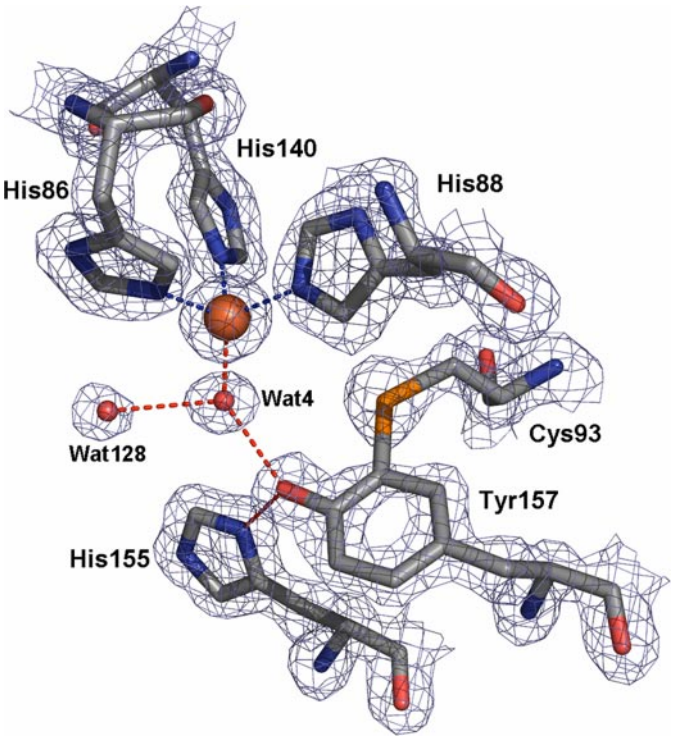


FIGURE 1. **Electron density evidence for key features of the CDO active site.** $2F_o - F_c$ electron density is shown contoured at 1.6 σ. Stick representations of select protein residues, including the Cys-Tyr linkage, are shown with iron (*orange sphere*) and active site waters (*red spheres*). All structural figures within this report were prepared using PyMOL (62).

RESULTS

Overall Structure of CDO—The crystal structure of *R. norvegicus* CDO, solved by SAD phasing using Se-Met-substituted protein, yielded a final refined model with $r = 18.0$ and $R_{free} = 20.8$ at 1.5-Å resolution. A total of 186 of the 200 residues in the protein (residues 5–190) were well defined in the electron density map (Fig. 1) and are included in the final model. No non-Gly residues have outlier ϕ,ψ angles, and there is one *cis*-peptide preceding Pro[159]. The structure as a whole is highly similar to that of murine CDO (PDB accession code 2ATF; 100% sequence identity; root-mean-square deviation = 0.2 Å), including the cysteinyl-tyrosine linkage (Fig. 1). The only salient difference involves the metallocenter, as further discussed below.

Briefly, the overall structure of CDO consists of a small α-helical domain containing three α-helices near the N terminus followed by 13 β-strands subdivided into a main β-sandwich domain and two β-hairpins at the C terminus (Fig. 2). A short 3_{10} helix is observed between β1 and β2. The entire β-sandwich is composed of seven anti-parallel β-strands (β1, β2, β4, β7, β9, β12, and β13) on the lower side and six anti-parallel β-strands (β3, β5, β6, β8, β10, and β11) on the upper side. N-terminal helices pack against the outside of the lower face of the sandwich to build a second non-polar core.

Alignment of CDO sequences across multiple species reveals that the elements of secondary structure seen in the core β-sandwich of rat CDO are conserved in other CDOs, with all insertions and deletions occurring between the secondary structural elements (Fig. 3). In contrast, the C-terminal β-hairpins may be dispensable. The

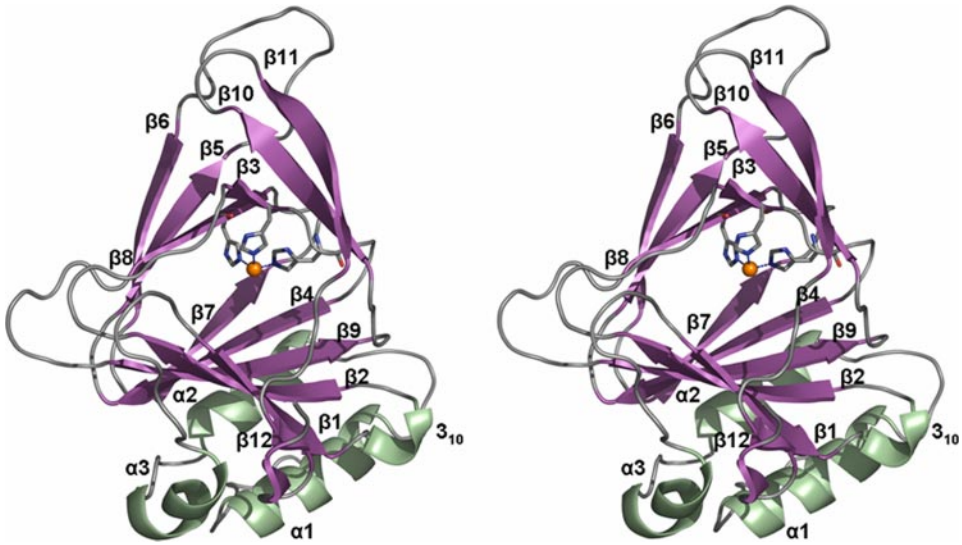


FIGURE 2. **Overall structure of CDO.** Stereoview ribbon illustration of the CDO structure with the central active site iron and its ligands shown. β-Strands (*purple*) and α-helices (*green*) are labeled, with the exception of β13 because of space limitations.

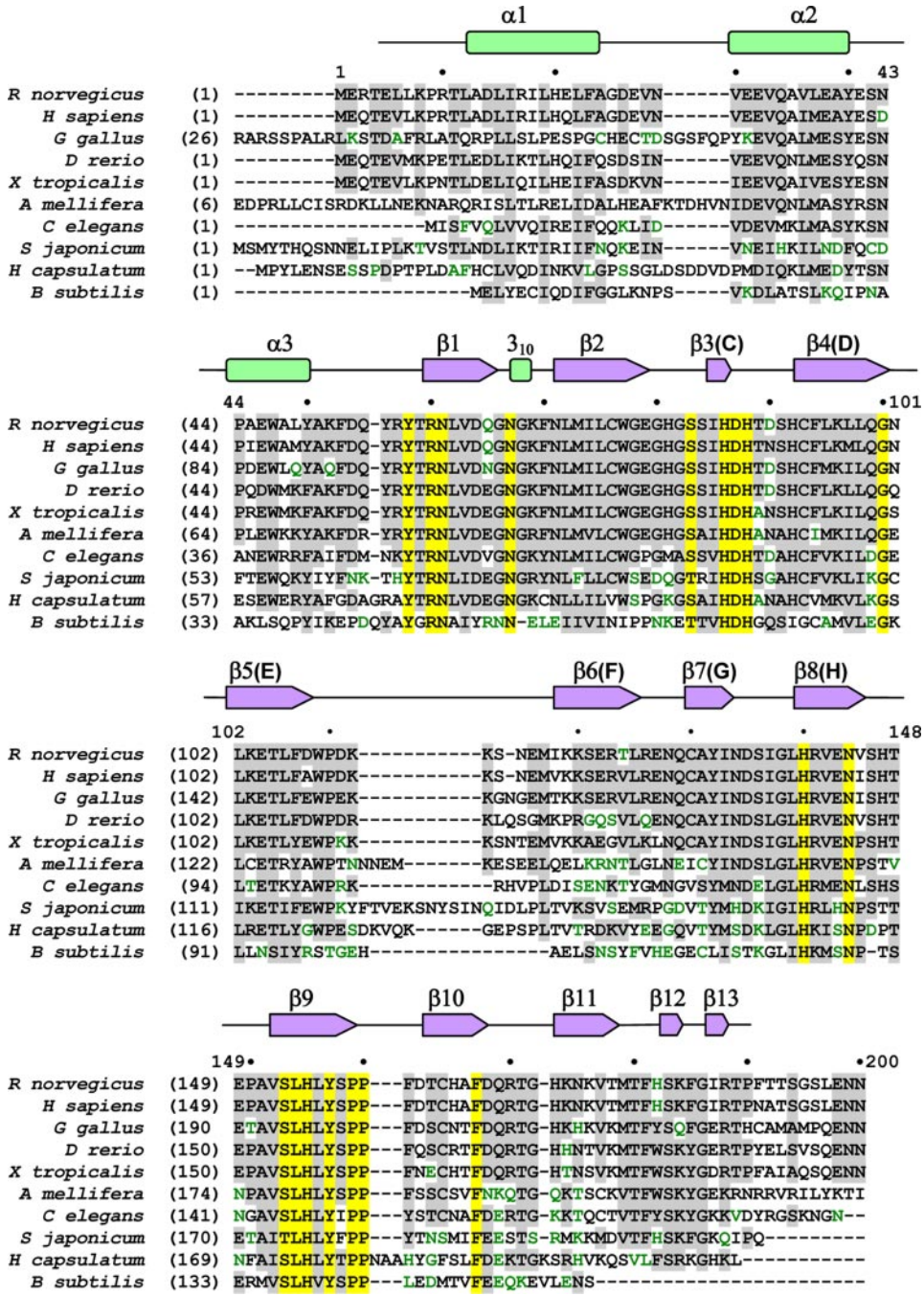


FIGURE 3. **Sequence alignment of select CDOs.** The secondary structure for rat CDO is indicated. From *top* to *bottom*, the sequences are from *Rattus norvegicus* (GenBank™ accession number AAA40904), *Homo sapiens* (2024212A), *Gallus gallus* (XP424964), *Danio rerio* (AAH67344), *Xenopus tropicalis* (AAH61333), *Apis mellifera* (XP396935), *Caenorhabditis elegans* (Q20893), *Shistosoma japonicum* (AAD52701), *Histoplasma capsulatum* (Q5RLY7), and *Bacillus subtilis* (CAB15092). Residues conserved (allowing for Ser/Thr replacements) in all 10 sequences are highlighted in *yellow*, and those identical in 5 or more sequences are shown in *gray*. Conservative replacements are also indicated by *gray highlighting* and weakly similar replacements by *green letters*. The sequence alignment was prepared using AlignX software (Vector NTI, Invitrogen) and the Blosum62mt scoring matrix.

sequence alignment also reveals notably strong conservation of 18 residues (*yellow* in Fig. 3). Among these, 14 are present in the active site area, and their roles will be discussed below. The remaining four residues are likely important for structural reasons: Gly^{100} at the end of cupin motif 1 at a corner of the β-sandwich has a Gly-specific conformation of $\phi,\psi = (+85, -157)$; Asn^{144}, the terminal residue of cupin motif 2, is fully buried, stabilizing a loop spanning residues 144–150 by hydrogen bonding to three backbone groups (Ser^{146}-N, Gly^{78}-O, and Glu^{149}-O) and a water molecule; Asn^{61} is fully buried and hydrogen bonds to Ser^{183}-Oγ, Thr^{59}-Oγ and the Ile^{74}-O, apparently significant for formation of the zigzag chain path of segment 180–188 that contains two of the metal ligands; and Asn^{67}, with $\phi,\psi = (+55, +28)$, is in a G-N-G tripeptide with each residue in the α-L conformation, creating a short left-handed 3_{10} helix.

CDO Active Site—The CDO active site is identified by the mononuclear iron center that is fully occupied and is coordinated via a roughly tetrahedral geometry by the conserved residues His^{86}, His^{88}, His^{140} and the water molecule Wat4 (Figs. 1 and 4*A*). The B-factors for the iron (12 Å^2) and Wat4 (10 Å^2) are similar to those of the coordinating His residues (11–14 Å^2), indicating full occupancy of the metal and no heterogeneity in the coordination geometry. The metallocenter is located in the central portion of the cupin β-sandwich (Fig. 2), consistent with what has been observed in other cupin structures. This metallocenter geometry appears different from the hexacoordinated Ni(II) center seen in the *M. musculus* CDO structure (26), but the difference can be seen simply to involve the additional presence in *M. musculus* CDO of two waters with long coordination distances (Fig. 4*B*). However, although geometrically small,

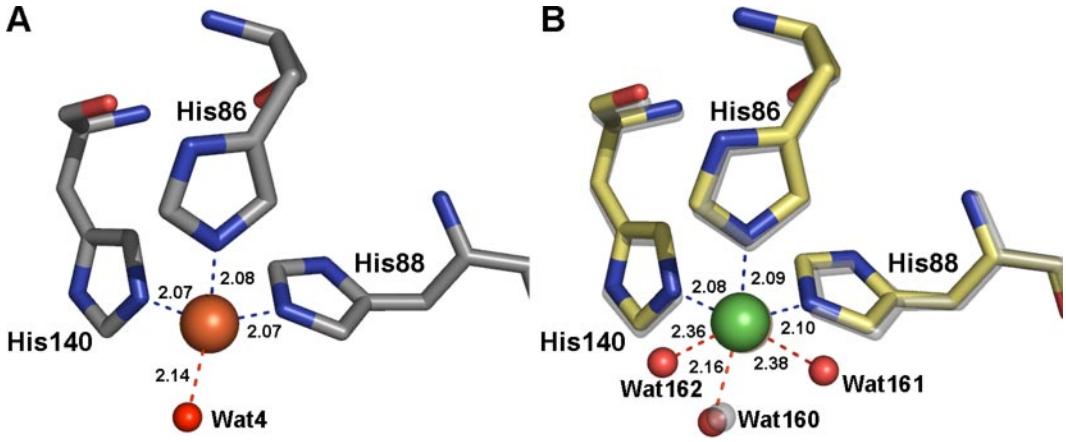


FIGURE 4. **Metal coordination comparison between the *R. norvegicus* and *M. musculus* CDO models.** *A*, tetrahedrally coordinated iron center in native *R. norvegicus* CDO. The CDO iron (*orange sphere*) is coordinated by the Nϵ2 atoms of three histidine ligands (His^{86}, His^{88}, and His^{140}) and water molecule Wat4 (*red sphere*) bound to the catalytic iron. The angles between coordinating ligands are as follows: Wat4-Fe-His^{140} = 110.1°, His^{140}-Fe-His^{86} = 100.6°, His^{86}-Fe-His^{88} = 100.4°, His^{86}-Fe-Wat4 = 118.8°, and His^{88}-Fe-Wat4 = 122.6°. The refined B-factors for the iron coordinating ligands are His^{86}-Nϵ2 = 13.7 Å^2, His^{88}-Nϵ2 = 11.2 Å^2, His^{140}-Nϵ2 = 11.0 Å^2, and Wat4 = 9.9 Å^2. The refined iron B-factor of 11.7 Å^2 suggests that it is fully occupied. Chelation distances are indicated and are consistent with the distances typically found in iron metalloenzymes. *B*, superposition of the hexacoordinated nickel center in the *M. musculus* CDO model with the iron center of *R. norvegicus* CDO (*translucent gray*) Residues (*pale yellow*), nickel (*green sphere*), and waters (*red spheres*) are indicated. Nickel chelation distances are also indicated. The views depicted in each panel are similar to the orientation shown in Fig. 2.

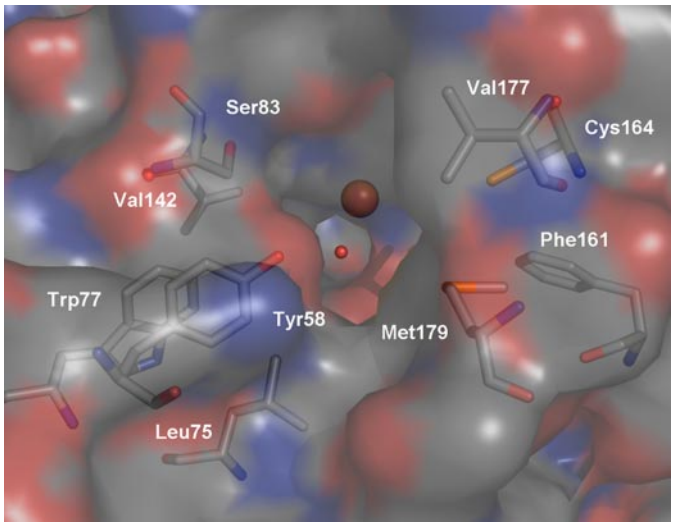


FIGURE 5. **Accessibility of the substrate binding pocket viewed from the enzyme surface.** The molecular surface of the enzyme is 60% transparent to allow visualization of the iron (*orange sphere*), the coordinated water molecule (*red sphere*), and the residues surrounding the opening and lining the binding pocket (*stick* representations).

this difference has major ramifications for mechanistic proposals (see "Discussion").

The iron is roughly 8 Å away from the protein surface and is surrounded by a solvent-filled pocket that connects to bulk solvent (Fig. 5). The remaining 14 conserved residues not mentioned above are all associated with the active site pocket, either lining it or adjacent to it (Fig. 5).

Conspicuous among the conserved residues within the substrate binding pocket is Tyr^{157}. Tyr^{157}-OH forms a short (2.6 Å) hydrogen bond with Wat4, and as observed previously by McCoy *et al.* (26), very clear electron density shows that Cys^{93}-Sγ is covalently bonded to Tyr^{157}-Cϵ2 forming a cysteinyl-tyrosine linkage (Fig. 1). That Cys^{93}-Sγ and Cys^{93}-Cβ are roughly coplanar with the aromatic ring of Tyr^{157} indicates that the Cys^{93}-Sγ to Tyr^{157}-Cϵ2 bond has partial double bond character. This geometry was first observed by Ito *et al.* (37) in galactose oxidase and more recently by Schnell *et al.* (38) in NirA, a sulfite reductase. These are the only other structurally known examples of proteins containing a cysteinyl-tyrosine linkage.

Additional highly conserved residues directly lining the active site include Tyr^{58}, Arg^{60}, Ser^{153}, and His^{155} (Fig. 6). Tyr^{58} and Arg^{60} H-bond to waters in the active site and are thus well positioned to be directly involved in substrate coordination/catalysis. Ser^{153}-Oγ H-bonds to His^{155}-Nϵ2 (2.68 Å), and His^{155}-Nδ2 H-bonds to Tyr^{157}-OH (2.69 Å), forming a Ser^{153}·His^{155}·Tyr^{157} triad reminiscent of the Asp·His·Ser catalytic triad in chymotrypsin-like serine proteases (39, 40). The conservation of Leu^{154}, buried in a neighboring aliphatic pocket, and *cis*-Pro^{159}-Pro^{160}, located in a loop between β9 and β10, would seem related to ensuring accurate positioning of this triad of residues. Similarly, the conservation of Ser^{83}, which H-bonds to the backbone NH of residue 142 (very close to the metal ligand His^{140}), and Phe^{167}, packed behind the main chain containing the metal ligands His^{86} and His^{88}, may play a role in maintaining the integrity of the metallocenter. In addition, Asp^{87} is positioned between the iron-coordinating residues His^{86} and His^{88}, and the Asp^{87} carboxylate interacts electrostatically with Asp^{87}-N and Thr^{89}-N, maintaining the position of the iron ligands, as well as with the backbone and side chain of His^{165} in the neighboring β-strand that contributes to the active site.

Other residues that provide a non-polar lining to the active site pocket are Leu^{75}, Trp^{77}, Val^{142}, Phe^{161}, Cys^{164}, Val^{177}, and Met^{179}. Overall, the pocket corresponds reasonably well to the space that would be required for a single molecule of cysteine. In addition to this main pocket, there is a smaller pocket located behind Wat4 immediately adjacent to Cys^{93}. This small pocket contains a single bound water (Wat129), is lined with the hydrophobic residues Leu^{95} and Ile^{133}, and is barely separated from bulk solvent by a loop comprising residues 134–138. Sequence alignment of CDO across multiple species reveals strong conservation of the sequence of this loop.

Mixed Disulfide Structure—Several attempts were made to acquire a CDO-substrate ligand-bound complex structure. Data were collected on crystals soaked with substrate (cysteine) or 2-aminoethanethiol (cysteamine), for times ranging from 5 min to 24 h, as well as crystals co-crystallized with substrate or the cysteine analog selenocysteine. All crystals were isomorphous with native CDO, but none yielded convincing evidence for ligand binding at the active site. The tetrahedral geometry of the iron center and the Cys^{93}-Tyr^{157} thioether observed in the native structure were also observed in all of these structures (data not shown).

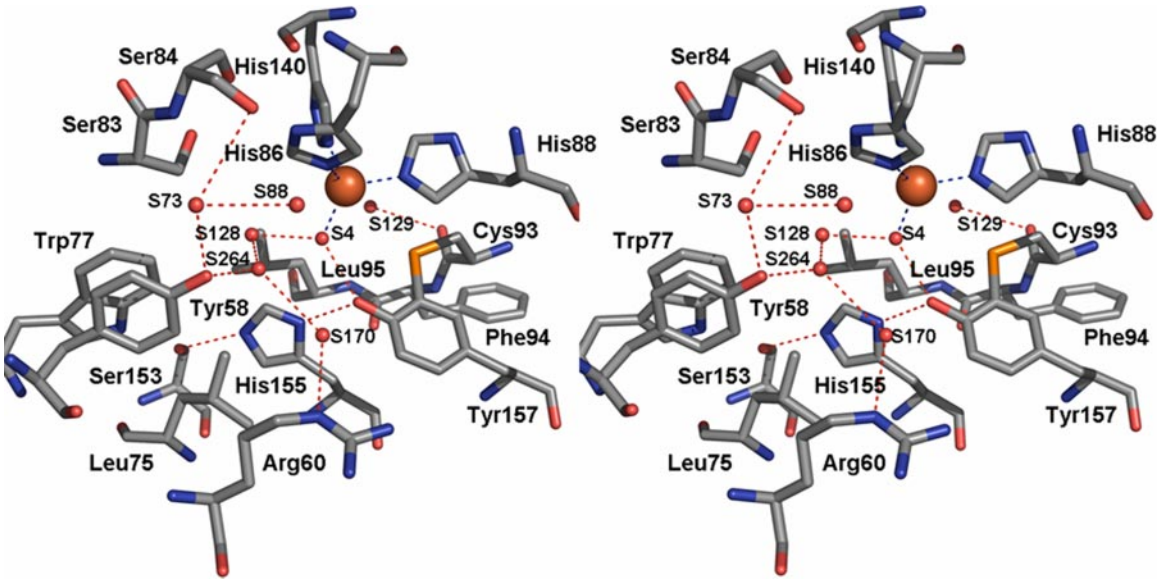


FIGURE 6. **Stereoview of the residues comprising the native CDO substrate binding pocket.** Residues (atom coloring), iron (*orange sphere*), active site waters (*red spheres*), and H-bonds (*dashed lines*) are shown. Active site waters S151 and S156 have been removed for clarity. Water S151 H-bonds with waters S156, S88, and S264 (where S stands for solvent or water). This representation is similar to the orientation shown in Fig. 1. A plausible substrate binding mode would place an end-on iron-bound dioxygen oriented along the S4–S128 axis and the sulfur of Cys near the position of S264.

TABLE 1

Crystallographic data and refinement statistics

NSLS, National Synchrotron Light Source; CHESS, Cornell High Energy Synchrotron Source; NA, not applicable; r.m.s.d., root mean square deviation.

	Native	Se-Met	Cys co-crystal
Data collection			
Experimental station	NSLS X12b	CHESS F2	CHESS F2
X-ray wavelength (Å)	1.0	0.9790	0.9790
Exposure time (s)	30	15	20
Oscillation range (°)	1.0	1.0	1.0
Cell dimensions (Å)	$a = b = 57.55$, $c = 123.06$	$a = b = 57.49$, $c = 122.27$	$a = b = 57.48$, $c = 122.80$
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$
Resolution (Å)[a]	30–1.50 (1.55-1.50)	30–1.80 (1.86-1.80)	30–1.50 (1.53-1.5)
Unique reflections	33453	19865	32390
Multiplicity	13.1	17.8	13.8
I/σ	52.4 (7.0)	31.8 (8.6)	25.1 (1.4)
R_{sym} (%)	4.6 (31.8)	8.7 (46.9)	8.4 (71.3)
Completeness (%)	98.2 (95.8)	100.0 (100.0)	100.0 (100.0)
Refinement			
R (%)/R_{free} (%)	18.0/20.8	NA	19.6/22.3
r.m.s.d. bonds (Å)	0.02	NA	0.01
r.m.s.d. angles (°)	2.06	NA	1.93

[a] Values in parentheses are from the highest resolution shell.

In all of the Cys soaks or co-crystals, however, a set of structural changes was consistently observed in the binding pocket that indicated some change had taken place. Using the best of these data sets, acquired from a co-crystallization experiment with 5 mM cysteine, these structural changes were determined to 1.5 Å resolution (final $r = 19.8$ and $R_{free} = 22.4$; Table 1). The analysis revealed very strong (~10 σ) difference density ~2 Å from the Sγ of Cys^{164} (~8 Å from the metallocenter), as well as clear evidence for conformational changes of Met^{179} and Arg^{60}. Although there was no clear density beyond the strong peak adjacent to Cys^{164}, the observed structural changes appear to all be related to what has been interpreted as the formation of a mixed disulfide between Cys^{164} and a substrate molecule (Fig. 7*A*). The lack of definitive density attributable to the remainder of the cysteine involved in the disulfide is mysterious, but as no other potential disulfide-forming molecules are known to be present, we attribute the lack of density to disorder and for the purposes of illustration have modeled it as a methyl sulfide.

The presence of the disulfide leaves a significant "footprint" upon the surrounding residues (Fig. 7), as the new sulfur atom is incompatible with the native position of Met^{179}-Sδ (2.6 Å), and this causes Met^{179} χ_1 to rotate ~120°. This new position of Met^{179} overlaps with the native position of the guanidinium of Arg^{60}, causing it to rotate around χ_2 ~80° to a point deeper into the substrate-binding cavity (Fig. 7*B*). The movement of Arg^{60} into the active site cavity disrupts the remaining native solvent structure (not including Wat4), but the observed solvent structure cannot be interpreted in terms of a single well defined constellation of non-overlapping water sites. This complexity could be due to partial occupancy of the modification. Indeed, some residual electron density is visible for the native conformations of Met^{179} and Arg^{60} (Fig. 7*A*), leading us to estimate that the mixed disulfide is present at an occupancy of about 75%.

Comparison with Structurally Similar Proteins—To assess the similarity of CDO with other known protein structures, a Dali search (41) of the whole Protein Data Bank was performed. A Z-score = 6.0 is recommended by Dali as the cutoff for consideration of structural homologs. As many cupin structures are available, a cutoff of Z-score = 8.0 has been employed here to select the top 12 structural homologs (Table 2). Although all of these homologs are classified as

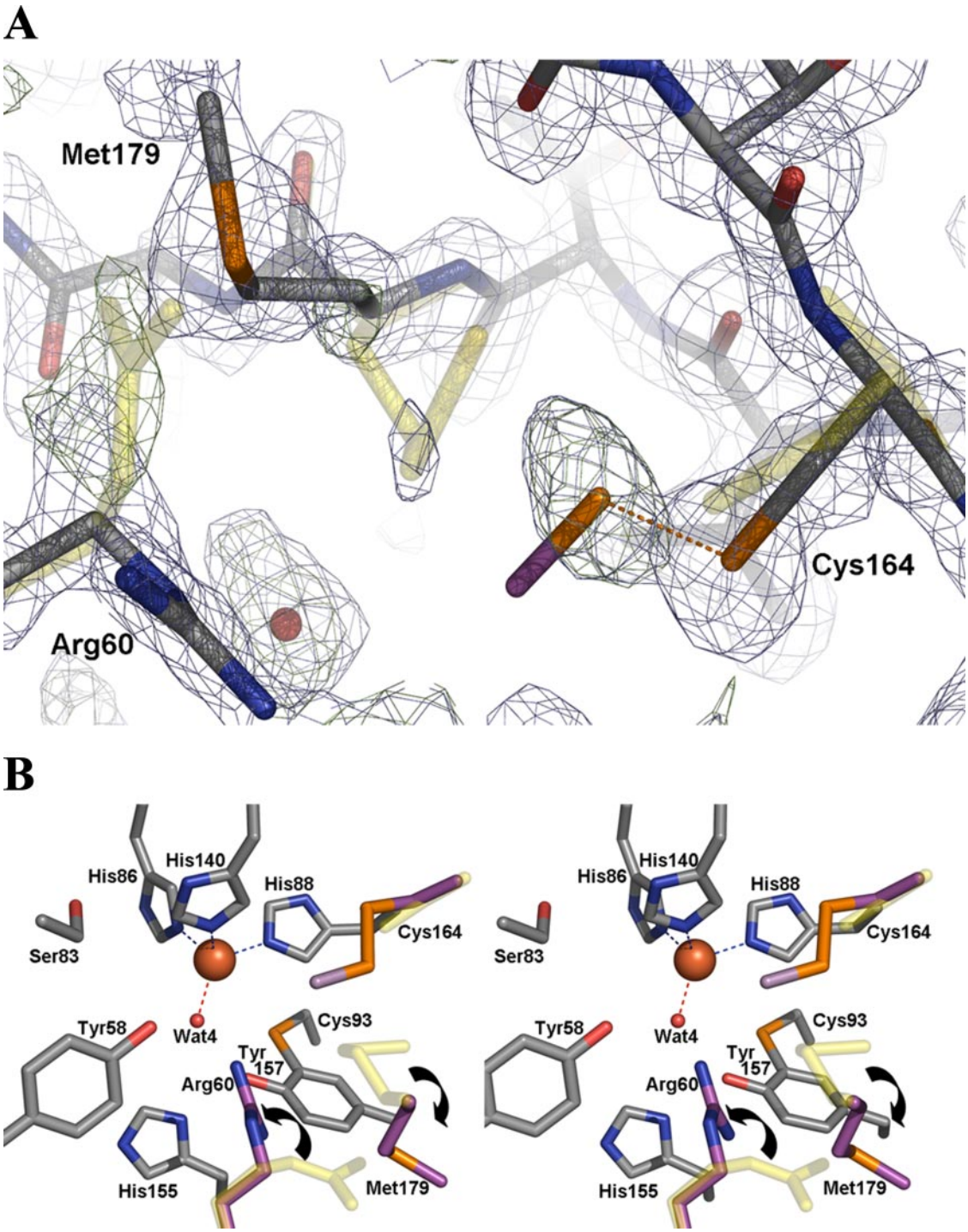


FIGURE 7. **The mixed disulfide observed in the co-crystal with cysteine.** *A*, evidence for the mixed disulfide is shown as $2F_o - F_c$ electron density contoured at 1.0 σ (*blue*), and omit $F_o - F_c$ electron density contoured at 3.0 σ (*green*). The mixed disulfide model (*gray*) is superimposed with the native model (*translucent yellow*). Sγ and Cβ of the putative bound cysteine molecule (*i.e.* just the methyl sulfide portion) are modeled with a Cys^{93}-Sγ-substrate-Sγ-substrate-Cβ angle of 104°. The orange *dashed line* indicates the putative disulfide bond at a distance of 2.06 Å. *Black arrows* indicate movement of residues. Note the weak residual electron density for the native positions of Met^{179}-Sδ and the Arg^{60} guanidino. In the final co-crystal structure, waters have been modeled at the position of the substrate sulfur and in the native position of Met^{179}-Sδ (water molecules are labeled S401 and S362, respectively, in the Protein Data Bank). *B*, stereoview of the mixed disulfide structure substrate binding site. Arg^{60}, Cys^{164}, and Met^{179} from the native model (*translucent yellow*) and in the mixed disulfide complex (*magenta* carbons) are shown. Also shown are the mixed disulfide bond (*orange*), additional pertinent active site residues (*gray* carbons), and the metallocenter (colored as described for Fig. 4*A*).

cupin superfamily members, none of them is very similar to CDO (42–51). The highest sequence identity is 18% and the closest structural similarity is 2.2 Å over 105 residues with quercetin dioxygenase (QDO) from *Bacillus subtilis* (42). In comparison with the structurally similar proteins, CDO has the longest intermotif domain, and, along with YML079w (43), lacks the consensus Gly at the beginning of motif 2. Six of the structural neighbors bind metals at the active site. Of these six, three (TM1459, *B. subtilis* QDO, and *Aspergillus japonicus* QDO) contain all four of the canonical cupin family metal-binding ligands (*i.e.* the His, His, Glu, and His residues of the cupin motifs), but only *B. subtilis* QDO consistently uses all four as metal binding ligands (42, 44, 45). Hydroxyanthranilate dioxygenase and

TABLE 2

Closest structural neighbors of CDO

Structural neighbors of CDO were determined using the Dali network service (41). For each protein identified, the structure given in this table represents that with the highest similarity. All of the closest neighbors had identifiable conserved cupin-like motifs. Six of these have coordinated metals and six do not. The metal ion in the cited structure and its coordination number are listed in Part A. Alignment of sequences for the conserved cupin motifs is shown in Part B. Conserved metal binding residues are highlighted in yellow, and those that have been shown to coordinate with a transition metal ion in crystal structures are underscored. (For *Aj*QDO, E73 was coordinated with copper in a minor observed geometry but not in the major tetrahedral coordination. In OxdC, E162 (outside of the cupin motifs) also coordinates with the metal.) Other highly conserved cupin motif residues are highlighted in green. The Cys^{93} residue in CDO is shown in red. The six conserved β-strands that give rise to the cupin (jelly roll) fold are shown by arrows. Proteins (PDB no. in parentheses) are: 1) TM1459 (1VJ2) from *Thermotoga maritima* (44); 2) Q8U9W0 (1ZNP), a hypothetical protein from *Agrobacterium tumefaciens* (A. P. Kuzin, Y. Chen, F. Forouhar, S. M. Vorobiev, R. Xiao, L.-C. Ma, T. Acton, G. T. Montelione, J. F. Hunt, and L. Tong, unpublished); 3) germin (1FI2), oxalate oxidase from *Hordenum vulgare* (47); 4) proglycinin (1FXZ) from *Glycine max* (49); 5) canavalin (1DGW) from *Canavalia ensiformis* (50); 6) HAD (1YFU), 3-hydroxyanthranilate 3,4-dioxygenase from *Ralstonia metallidurans* (46); 7) YML079w (1XE7) from *Saccharomyces cerevisiae* (43); 8) *Bs*QDO (1Y3T), quercetin dioxygenase from *Bacillus subtilis* (42); 9) glycinin (1OD5) from *G. max* (51); 10) OxdC (1UW8), oxalate decarboxylase from *B. subtilis* (48); 11) YhhW (1TQ5) from *Escherichia coli* (M. A. Adams and Z. Jia, unpublished); and 12) *Aj*QDO (1JUH), quercetin dioxygenase from *Aspergillus japonicus* (45).

A. Structural similarity statistics and metal coordination

Protein	PDB	Z-score	RMSD (Å)	Length of Alignment (# residues)	Length of Sequence (# residues)	Sequence Identity (%)	Metal	Metal Coordination Number
CDO	2B5H				186		Fe	4
TM1459	1VJ2	10.2	3.0	106	114	13	Mn	6
Q8U9W0	1ZNP	9.7	2.4	110	138	14	--	--
Germin	1FI2	9.7	3.5	122	201	12	Mn	6
Proglycinin	1FXZ	9.4	3.5	116	371	5	--	--
Canavalin	1DGW	9.4	3.7	113	178	10	--	--
HAD	1YFU	9.0	3.4	122	174	15	Fe	6
YML079w	1XE7	8.8	3.2	115	186	8	--	--
*Bs*QDO	1Y3T	8.7	2.2	105	330	18	Fe	5
Glycinin	1OD5	8.6	3.3	114	382	6	--	--
OxdC	1UW8	8.5	2.9	108	377	12	Mn	6
YhhW	1TQ5	8.4	2.2	86	227	15	--	--
*Aj*QDO	1JUH	8.3	2.7	112	334	11	Cu	4/5

B. Cupin motif alignment

Protein	Cupin Motif 1	Intermotif	Cupin Motif 2
	PGXXXXXHXHXXXX--E-XXXXXXG	(~10-50)	GXXXXX----PXGXXHXXXN
CDO	$_{79}$EGHGSSIHDHSDSH--C-FLKLLQG$_{100}$	(29)	$_{130}$CAYIND----SIG-LHRVEN$_{144}$
TM1459	$_{57}$PGGLIDRHSHPWEH--E--IFVLKG$_{77}$	(15)	$_{93}$GFYIFVE---PNEI-HGFRN$_{108}$
Q8U9W0	$_{48}$LEKGVRSHWHRVTDAVE-VWHYYAG$_{70}$	(25)	$_{96}$GERPQVIV--PANCWQSAES$_{112}$
Germin	$_{44}$PNHRTDYHDDPLE---E-FFYQLRG$_{64}$	(19)	$_{84}$GDIFLL----PPHVRHSPQR$_{99}$
Proglycinin	$_{298}$KNAMFVPHYNLNANS---IIYALNG$_{368}$	(21)	$_{390}$GRVLIV----PQNFVVAARS$_{405}$
Canavalin	$_{52}$PNTLLLPH-HSDS---DLLVLVLEG$_{72}$	(19)	$_{92}$GDAIKI----QAGTPFYLIN$_{107}$
HAD	$_{44}$PNHRTDYHDDPLE---E-FFYQLRG$_{64}$	(19)	$_{84}$GDIFLL----PPHVRHSPQR$_{99}$
YML079w	$_{35}$NDWQLIKHREGGYFK-E-TDRSPYT$_{57}$	(28)	$_{85}$LLTPDS----PIGKFHKNIN$_{100}$
*Bs*QDO	$_{55}$KGDAFPLHVHKDTH--E-GILVLDG$_{76}$	(15)	$_{92}$GDYANI----PAGTPHSYRM$_{107}$
Glycinin	$_{376}$NGIYS-PHWNLNANS---YVTRGKG$_{398}$	(8)	$_{407}$GNAVFDGELRRGQLLVVPQN$_{426}$
OxdC	$_{88}$PGAIRELHWHKEA---E--AYMIYG$_{108}$	(18)	$_{127}$GEGDLWYF--PSGLPHSIQA$_{144}$
YhhW	$_{50}$AGQGFGTHPHKDX---EILTYVLEG$_{71}$	(16)	$_{88}$GEFQIXSA--GTGIRHSEYN$_{105}$
*Aj*QDO	$_{59}$DALGVLPHIHQKHY--E-NFYCNKG$_{80}$	(20)	$_{101}$GDYGSV----PRNVTHTFQI$_{116}$

germin each lack one of the conserved His residues in cupin motif 1 (46, 47); oxalate decarboxylase (OxdC), which lacks both a His and Glu in motif 1, is unusual in that it uses a downstream Glu for metal coordination (48). CDO uniquely contains a conserved Cys at the consensus motif 1 Glu position. Further, only *B. subtilis* QDO and *A. japonicus* QDO display metal centers with fewer than six ligands (42, 45). *B. subtilis* QDO has a pentacoordinate iron center, whereas *A. japonicus* QDO has been observed with both a tetra- and penta-coordinate Cu center. Among all known cupins, CDO is the only metal-binding protein that contains a four-coordinate iron center.

DISCUSSION

The mononuclear iron center observed in the crystal structure of rat CDO, along with the differences in sequence and structure compared with all other structurally known cupins, makes it clear that CDO represents a new cupin subfamily. One point, however, that requires discussion before making interpretations about structure-function relations for CDO is the distinct coordination geometry of the nickel-bound and iron-bound metallocenters seen in the mouse and rat CDO structures, respectively (Fig. 4). Previous studies have shown that CDO

cannot be activated by nickel (15, 52), implying that the *M. musculus* structure is an inactive form of the enzyme. In contrast, characterization of CDO purified directly from rat liver determined that CDO activity is iron-dependent and contains one atom of iron per molecule (15). The recombinant CDO used in this study showed ferrous ion-dependent activity that is identical to that of the enzyme in rat liver homogenates (16). The presence of iron in recombinant *R. norvegicus* CDO has been verified more recently by Chai *et al.* (52) as well as by transmission emission microscopy studies performed on the recombinant CDO from our laboratory (data not shown). Fluorescence scans at the nickel absorption edge of purified CDO in solution and the crystals used here demonstrated no observable presence of nickel in either case. It is important to note that the protein prepared for the structure reported here was purified based upon activity, with the enzyme being separated into active and inactive isoforms. Only the highly active population of CDO was used for crystallization; purified rat CDO had a k_{cat} of 43 min^{-1} and a K_m for cysteine of 0.45 mM, compared with $k_{cat} = 1.8\ min^{-1}$ and $K_m = 3.4$ mM for the enzyme used in determining the murine CDO structure (26). Thus, we conclude that between these two structures, the iron-bound structure reported here is the more physiologically relevant structure to consider for deriving mechanistic insights into CDO. Although Fe(II) is required for activity, we are not able to assert whether the iron center in the crystal is ferrous or ferric, although it is probably ferric due to weeks of aerobic storage. Given the active site rigidity, we suspect the coordination will be unchanged in the ferrous *versus* ferric states, but it is conceivable that the coordination sphere of an oxidized Fe(III) center expands upon reduction to Fe(II).

The Iron Metallocenter Is Novel—The most surprising feature of the metallocenter is its tetrahedral coordination, because metal-binding proteins in the cupin family, as well as mononuclear iron enzymes in other protein families, typically have penta- or hexacoordinated metal centers. In fact, to our knowledge, such pure tetrahedral coordination of mononuclear iron has not been seen apart from that found in the $Fe(Cys)_4$ iron centers of rubredoxin-like electron carriers. Interestingly, the most similar centers are the non-heme iron metallocenters in the Rieske family of dioxygenases and the copper center of the cupin QDO.

For the Rieske dioxygenases, representative structures have been determined for biphenyl dioxygenase (BphA1A2; PDB entry 1UL1), naphthalene dioxygenase (PDB entry 1NDO) and cumene dioxygenase (PDB entry 1WQL) (53–55). In terms of protein ligands, the biphenyl dioxygenase and cumene dioxygenase metal centers are coordinated by two histidines and one Asp carboxylate atom, whereas in naphthalene dioxygenase the second Asp carboxylate oxygen appears to be a distant (2.3 Å) fourth protein ligand. Each of these enzymes is, however, distinct from CDO in that the coordination sphere appears to be completed not just by a single well ordered water atom but by extended electron density that may represent a pair of disordered water atoms or perhaps a dioxygen-like molecule (53–55). This unusual ligation effectively makes these centers pentacoordinate, supporting proposals that these dioxygenases bind oxygen in a "side-on" fashion to facilitate the chemistry necessary for the *cis*-specific hydroxylation of the hydrocarbon ring of the substrate (54).

In the case of the cupin superfamily member QDO, Fusetti *et al.* (44) reported that the copper center of QDO from *A. japonicus* displayed a mixture of a tetrahedral form (~70%), chelated by the three conserved cupin His residues (His^{66}, His^{68}, and His^{112}) and a water molecule, and a minor (~30%) pentacoordinate form in which the copper was additionally coordinated by the conserved cupin Glu residue (Glu^{73}). EPR analysis of Cu-QDO was consistent with the absence of a carboxylate ligand in the *A. japonicus* Cu-QDO, further suggesting that the absence of a carboxylate ligand may facilitate active site chemistry in QDO (56). CDO, however, is distinct from both QDO and the Rieske dioxygenases in that it does not have any carboxylate available near the iron center to serve as an additional ligand.

An Explanation for Substrate Inhibition in Vitro—The metal center of CDO contains only one exchangeable ligand (Wat4) and therefore would appear to only provide access to either a molecule of substrate or an oxygen molecule. In the substrate soaks and co-crystals, the crystal structures revealed no apparent electron density close to the metal ion that corresponds to a molecule of ordered oxygen, nor did they reveal any convincing density within the active site pocket for a molecule of cysteine that was either bound to the metal center directly or coordinated by nearby residues.

Although all attempts to acquire a CDO-cysteine complex structure were unsuccessful, we were able to gain insights into why higher concentrations of cysteine in *in vitro* activity assays inhibit CDO activity (16). The cysteine adduct involved in the mixed disulfide with Cys^{164}-Sγ is thought to have a profound influence upon catalysis by inhibiting access of substrate to the active site. In particular, the displacement of Arg^{60} from the native position causes it to protrude ~4 Å into the pocket (Fig. 7), and given the small size of the active site pocket, this position of Arg^{60} would block cysteine binding.

Cys^{164} is conserved from humans to *Caenorhabditis elegans* but is not found in CDO of lower eukaryotes or in any putative prokaryotic CDO.[4] Although Cys^{164} is positioned in a seemingly inefficient place in the protein, where it lines the substrate binding pocket, it is likely that this effect would not occur in an *in vivo* reducing environment, allowing for efficient catalysis. This observation does suggest, however, that mutation of Cys^{164} could improve the overall efficiency of CDO *in vitro*, especially in the presence of multimillimolar cysteine concentrations.

Mechanistic Proposals—In the absence of a CDO-substrate bound complex, the native unliganded structure nevertheless provides sufficient context to allow us to propose a catalytic mechanism for the enzyme. The formal addition of nucleophiles to molecular oxygen is a commonly occurring reaction in dioxygenase chemistry. This reaction does not occur by a simple nucleophilic addition, because molecular oxygen has a triplet ground state. Typically, this addition is catalyzed by a transition metal ion, which can stabilize the superoxide radical by coordinating to it (58, 59). Based on this precedent, we propose for CDO the mechanistic scheme outlined in Fig. 8*A*. Upon end-on binding to the Fe(II) center, O_2 accepts an electron from the iron and an H-bond from the hydroxyl of Tyr^{157} to yield complex **2**. Meanwhile, a putative active site base, which we propose is likely to be Tyr^{58}, deprotonates the thiol of the substrate and the resulting thiolate transfers an electron to the Fe(III) to yield **3**. Radical coupling then gives the peroxysulfenate **4**. Cleavage of the oxygen-oxygen bond, facilitated by protonation by Tyr^{157}, would give **5**. Rotation about the carbon–sulfur bond gives **6,** which followed by addition of the iron-bound oxygen to the sulfur (and proton transfer back to Tyr^{157}) would give **7**, and product dissociation would complete the reaction. This proposal is consistent with the observed incorporation of both atoms of molecular oxygen into the sulfinate product and with known iron/oxygen chemistry (14, 58, 59). The sulfur and the two oxygens of the peroxysulfenate could occupy the positions occupied by waters 264, 128, and 4, respectively, in the native structure, and the bond rotation converting **5** to **6** would move the oxygen from the site occupied by water 128 to the site occupied by water 151 (Fig. 5). Although peroxysulfenate chemistry is proposed, there has

[4] Dominy, J. E., Jr., Simmons, C. R., Karplus, P. A., Gehring, A., and Stipanuk, M. H. (2006) *J. Bacteriol.*, in press.

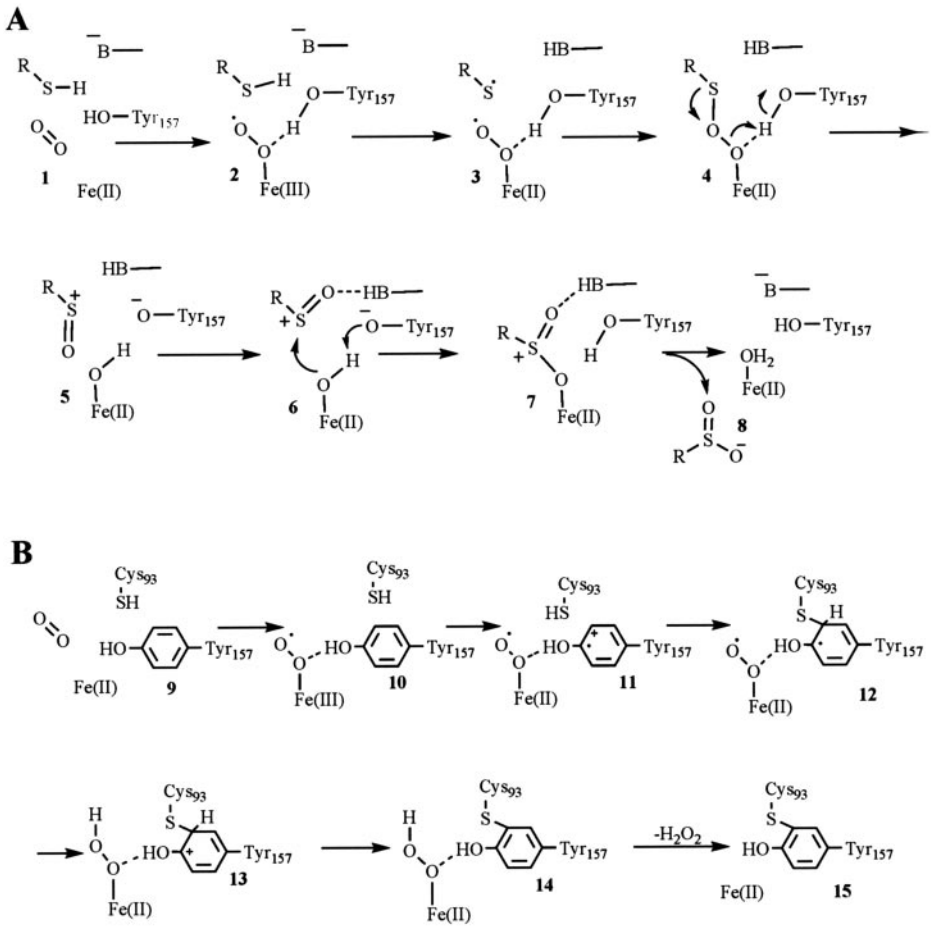


FIGURE 8. **Mechanistic proposals for CDO.** *Scheme A*, mechanistic proposal for the catalytic cycle of cysteine oxidation. The *letter B* in this *Scheme A* indicates a putative active site base. *Scheme B*, mechanistic proposal for the single turnover event generating the cysteinyl-tyrosine thioether cross-link. Each mechanism is discussed in the text (see "Discussion").

been no experimental characterization of such an intermediate and there is little literature in this area (60). Therefore, CDO is potentially a valuable model for gaining a better understanding of this chemistry, which is likely to be of general applicability to thiol oxidation to sulfenic, sulfinic, and sulfonic acids.

Several examples of posttranslational redox modification of enzymes have been described, but apart from CDO, galactose oxidase and NirA are the only other proteins known to contain a cysteinyl-tyrosine covalent modification (37, 38). The mechanisms of these modifications are often related to the catalytic mechanism of the enzyme (61). Such a proposal for the modification of Tyr^{157} in CDO is outlined in Fig. 8*B*. This mechanism is similar to the catalytic mechanism proposed in Fig. 8*A* in so far as it requires electron transfer from the active site Fe(II) to molecular oxygen followed by the formation of a superoxide complex **10**. Electron transfer from Tyr^{157} to the Fe(III) would then give **11**. Addition of the thiol to the resulting electron-deficient benzene ring to give **12** followed by an electron transfer would give **13**. Aromatization followed by dissociation of hydrogen peroxide would complete the posttranslational modification reaction.

Although we have proposed an important mechanistic role for Tyr^{157}, the lack of conservation of Cys^{93} in many bacterial CDOs suggests the thioether linkage is not crucial for activity.[4] Thus, our working model is that for mammalian enzymes the cross-linking to Cys^{93} aids the positioning of Tyr^{157}, but it is the hydrogen bonding to His^{155} that primarily increases its acidity.

In addition to Tyr^{157} (activated by the supporting residues Ser^{153} and His^{155}), Cys^{93}, and Tyr^{58}, for which specific roles have already been proposed above, the active site of CDO contains several conserved residues that are strong candidates for substrate coordination due to both proximity and sequence conservation. In particular, Arg^{60} and Ser^{83} (Fig. 6) appear well suited to directly coordinate the α-carboxylate and α-amino groups of cysteine.

Outlook—Thiol oxidation and reduction are important reactions in biology, yet the enzymes responsible are largely unidentified. The absence of close homologs of CDO may be because of the limited requirement for enzymes that catalyze oxidation of sulfhydryl groups. Within mammals, the only two reactions in which a thiol is oxidized to a sulfinate are those catalyzed by cysteine dioxygenase and the putative cysteamine dioxygenase. *Acidithiobacillus* and *Acidiphilium* spp., which can thrive chemolithotrophically by oxidation of sulfur compounds, appear to use a glutathione-dependent sulfur dioxygenase (EC 1.13.11.18) to oxidize elemental sulfur via a glutathione persulfide substrate (GS_nH) that is formed nonenzymatically from glutathione and elemental sulfur (S_8) (57). Given the limited requirement in biology for reactions that catalyze sulfur dioxygenation, it is plausible that the subfamily of cupins represented by CDO has very few members.

The CDO structure, as the representative member of a new class of cupin family proteins, raises many questions. The mechanism we propose here provides a starting point for further experimental testing of the mechanistic possibilities for CDO-catalyzed oxidation of cysteine to cysteine sulfinate. Future efforts will focus on experiments to determine substrate- and product-bound complexes as well as mutagenesis experiments aimed at understanding how the unique iron center coordination and conserved active site residues contribute to the sulfur chemistry catalyzed by CDO.

Acknowledgments—We thank Peter Meyers, David Schuller, and Ganapathy Sarma for help with crystallographic refinement and modeling and John E. Dominy, Jr., Stephen Ealick, and Linda Nicholson for useful scientific and technical discussions. We also acknowledge Dieter Schneider and Alexei Soares at the National Synchrotron Light Source (NSLS) for strong technical support. NSLS financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the United States Dept. of Energy and from the National Center for Research Resources of the National Institutes of Health. The Cornell High Energy Synchrotron Source is supported by the National Science Foundation and National Institutes of Health under Awards DMR-0225180 and RR-01646.

REFERENCES

1. Stipanuk, M. H. (2004) *Annu. Rev. Nutr.* **24,** 539–577
2. Heafield, M. T., Fearn, S., Steventon, G. B., Waring, R. H., Williams, A. C., and Sturman, S. G. (1990) *Neurosci. Lett.* **110,** 216–220
3. Heafield, M. T., and Williams, A. C. (1992) *Curr. Opin. Neurol. Neurosurg.* **5,** 288–294
4. Bradley, H., Gough, A., Sokhi, R. S., Hassell, A., Waring, R., and Emery, P. (1994) *J. Rheumatol.* **21,** 1192–1196
5. Davies, M. H., Ngong, J. M., Pean, A., Vickers, C. R., Waring, R. H., and Elias, E. (1995) *J. Hepatol.* **22,** 551–560
6. Pedersen, O. O., and Karlsen, R. L. (1980) *Investig. Ophthalmol. Vis. Sci.* **19,** 886–892
7. Karlsen, R. L., Grofova, I., Malthe-Sorenssen, D., and Fonnum, F. (1981) *Brain Res.* **208,** 167–180
8. Mathisen, G. A., Fonnum, F., and Paulsen, R. E. (1996) *Neurochem. Res.* **21,** 293–298
9. McBean, G. J., and Flynn, J. (2001) *Biochem. Soc. Trans.* **29,** 717–722
10. Schubert, D., and Piasecki, D. (2001) *J. Neurosci.* **21,** 7455–7462
11. Li, H., and Dryhurst, G. (1997) *J. Neurochem.* **69,** 1530–1541
12. Huxtable, R. J. (1992) *Physiol. Rev.* **72,** 101–163
13. Ewetz, L., and Sorbo, B. (1966) *Biochim. Biophys. Acta* **128,** 296–305
14. Lombardini, J. B., Singer, T. P., and Boyer, P. D. (1969) *J. Biol. Chem.* **244,** 1172–1175
15. Yamaguchi, K., Hosokawa, Y., Kohashi, N., Kori, Y., Sakakibara, S., and Ueda, I. (1978) *J. Biochem.* (*Tokyo*) **83,** 479–491
16. Simmons, C. R., Hirschberger, L. L., Machi, M. S., and Stipanuk, M. H. (2005) *Protein Expression Purif.* **47,** 74–81
17. Hirschberger, L. L., Daval, S., Stover, P. J., and Stipanuk, M. H. (2001) *Gene* (*Amst.*) **277,** 153–161
18. Stipanuk, M. H., Londono, M., Hirschberger, L. L., Hickey, C., Thiel, D. J., and Wang, L. (2004) *Amino Acids* **26,** 99–106
19. Stipanuk, M. H., Hirschberger, L. L., Londono, M. P., Cresenzi, C. L., and Yu, A. F. (2004) *Am. J. Physiol.* **286,** E439–E448
20. Dominy, J. E., Jr., Hirschberger, L. L., Coloso, R. M., and Stipanuk, M. H. (2006) *Biochem. J.* **394,** 267–273
21. Stipanuk, M. H., Londono, M., Lee, J. I., Hu, M., and Yu, A. F. (2002) *J. Nutr.* **132,** 3369–3378
22. Hwang, L., Hocking-Murray, D., Bahrami, A. K., Andersson, M., Rine, J., and Sil, A. (2003) *Mol. Biol. Cell* **14,** 2314–2326
23. Sacco, M., Maresca, B., Kumar, B. V., Kobayashi, G. S., and Medoff, G. (1981) *J. Bacteriol.* **146,** 117–120
24. Dunwell, J. M., Khuri, Sawsan, and Gane, P. J. (2000) *Microbiol. Mol. Biol. Rev.* **64,** 153–179
25. Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R., and Goodenough, P. W. (2001) *Trends Biochem. Sci.* **26,** 740–746
26. McCoy, J. G., Bailey, L. J., Bitto, E., Bingman, C. A., Aceti, D. J., Fox, B. G., and Phillips, G. N., Jr. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 3084–3089
27. Simmons, C. R., Hao, Q., and Stipanuk, M. H. (2005) *Acta Crystallogr. Sect. F* **61,** 1013–1016
28. Doublié, S. (1997) *Methods Enzymol.* **276,** 523–530
29. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
30. Hao, Q., Gu, Y. X., Yao, J. Y., Zheng, C. D., and Fan, H. F. (2003) *J. Appl. Crystallogr.* **36,** 1274–1276
31. Hao, Q. (2004) *J. Appl. Crystallogr.* **37,** 498–499
32. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D* **55,** 849–861
33. Abrahams, J. P., and Leslie, A. G. (1996) *Acta Crystallogr. Sect. D* **52,** 30–42
34. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
35. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
36. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* **54,** 905–921
37. Ito, N., Phillips, S. E., Stevens, C., Ogel, Z. B., McPherson, M. J., Keen, J. N., Yadav, K. D., and Knowles, P. F. (1991) *Nature* **350,** 87–90
38. Schnell, R., Sandalova, T., Hellman, U., Lindqvist, Y., and Schneider, G. (2005) *J. Biol. Chem.* **280,** 27319–27328
39. Kraut, J. (1977) *Annu. Rev. Biochem.* **46,** 331–358
40. Dodson, G., and Wlodawer, A. (1998) *Trends Biochem. Sci.* **23,** 347–352
41. Holm, L., and Sander, C. (1998) *Nucleic Acids Res.* **26,** 316–319
42. Gopal, B., Madan, L. L., Betz, S. F., and Kossiakoff, A. A. (2005) *Biochemistry* **44,** 193–201
43. Zhou, C.-Z., Meyer, P., Quevillon-Cheruel, S., De La Sierra-Gallay, I. L., Collinet, B., Graille, M., Blondeau, K., Leulliot, N., Sorel, I., Poupon, A., Janin, J., and Van Tibeurgh, H. (2005) *Protein Sci.* **14,** 209–215
44. Jaroszewski, L., Schwarzenbacher, R., von Delft, F., McMullan, D., Brinen, L. S., Canaves, J. M., Dai, X., Deacon, A. M., DiDonato, M., Elsliger, M. A., Eshagi, S., Floyd, R., Godzik, A., Grittini, C., Grzechnik, S. K., Hampton, E., Levin, I., Karlak, C., Klock, H. E., Koesema, E., Kovarik, J. S., Kreusch, A., Kuhn, P., Lesley, S. A., McPhillips, T. M., Miller, M. D., Morse, A., Moy, K., Ouyang, J., Page, R., Quijano, K., Reyes, R., Rezezadeh, F., Robb, A., Sims, E., Spraggon, G., Stevens, R. C., van den Bedem, H., Velasquez, J., Vincent, J., Wang, X., West, B., Wolf, G., Xu, Q., Hodgson, K. O., Wooley, J., and Wilson, I. A. (2004) *Proteins* **56,** 611–614
45. Fusetti, F., Schroter, K. H., Steiner, R. A., van Noort, P. I., Pijning, T., Rozeboom, H. J., Kalk, K. H., Egmond, M. R., and Dijkstra, B. W. (2002) *Structure* (*Lond.*) **10,** 259–268
46. Zhang, Y., Colabroy, K. L., Begley, T. P., and Ealick, S. E. (2005) *Biochemistry* **44,** 7632–7643
47. Woo, E. J., Dunwell, J. M., Goodenough, P. W., Marvier, A. C., and Pickersgill, R. W. (2000) *Nat. Struct. Biol.* **7,** 1036–1040
48. Just, V. J., Stevenson, C. E. M., Bowater, L., Tanner, A., Lawson, D. M., and Bornemann, S. (2004) *J. Biol. Chem.* **279,** 19867–19874
49. Adachi, M., Takenaka, Y., Gidamis, A. D., Mikami, B., and Utsumi, S. (2001) *J. Mol. Biol.* **305,** 291–305
50. Ko, T. P., Ng, J. D., and McPherson, A. (1993) *Plant Physiol.* **101,** 729–744
51. Adachi, M., Kanamori, J., Masuda, T., Yagasaki, K., Kitamura, K., Mikami, B., and Utsumi, S. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 7395–7400
52. Chai, S. C., Jerkins, A. A., Banik, J. J., Shalev, I., Pinkham, J. L., Uden, P. C., and Maroney, M. J. (2005) *J. Biol. Chem.* **280,** 9865–9869
53. Furusawa, Y., Nagaarajan, V., Tanokura, M., Masai, I., Fukuda, M., and Senda, T. (2004) *J. Mol. Biol.* **342,** 1041–1052
54. Kauppi, B., Lee, K., Carredano, E., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (1998) *Structure* (*Lond.*) **6,** 571–586
55. Dong, X., Fushinobu, S., Fukuda, E., Terada, T., Nakamura, S., Shimizu, K., Nojiri, H., Omori, T., Shoun, H., and Wakagi, T. (2005) *J. Bacteriol.* **187,** 2483–2490
56. Kooter, I. M., Steiner, R. A., Dijkstra, B. W., van Noort, P. I., Egmond, M. R., and Huber, M. (2002) *Eur. J. Biochem.* **269,** 2971–2979
57. Rohwerder, T., and Sand, W. (2003) *Microbiology* **149,** 1699–1709
58. Que, L., Jr., and Ho, R. Y. (1996) *Chem. Rev.* **96,** 2607–2624
59. Bugg, T. D. H. (2003) *Tetrahedron* **59,** 7075–7101
60. Bugg, T. D. H., and Winfield, C. J. (1998) *Nat. Prod. Rep.* **15,** 513–530
61. Rogers, M. S., and Dooley, D. M. (2001) *Adv. Protein Chem.* **58,** 387–436
62. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA