Published in final edited form as:

J Biol Inorg Chem. 2016 September ; 21(5-6): 589–603. doi:10.1007/s00775-016-1363-x.

Crystal structure of CmlI, the arylamine oxygenase from the chloramphenicol biosynthetic pathway

Cory J. Knoot¹, Elena G. Kovaleva², and John D. Lipscomb¹

¹Department of Biochemistry Molecular Biology and Biophysics and the Center for Metals in Biocatalysis, University of Minnesota, Minneapolis, MN 55455, USA

²Stanford Synchrotron Radiation Lightsource, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

Abstract

The diiron cluster-containing oxygenase CmlI catalyzes the conversion of the aromatic amine precursor of chloramphenicol to the nitroaromatic moiety of the active antibiotic. The X-ray crystal structures of the fully active, N-terminally truncated CmlI 33 in the chemically reduced Fe²⁺/Fe²⁺ state and a *cis* μ -1,2(η^1 : η^1)-peroxo complex are presented. These structures allow comparison with the homologous arylamine oxygenase AurF as well as other types of diiron cluster-containing oxygenases. The structural model of CmlI 33 crystallized at pH 6.8 lacks the oxo-bridge apparent from the enzyme optical spectrum in solution at higher pH. In its place, residue E236 forms a μ -1,3(η^1 : η^2) bridge between the irons in both models. This orientation of E236 stabilizes a helical region near the cluster which closes the active site to substrate binding in contrast to the open site found for AurF. A very similar closed structure was observed for the inactive dimanganese form of AurF. The observation of this same structure in different arylamine oxygenases may indicate that there are two structural states that are involved in regulation of the catalytic cycle. Both the structural studies and single crystal optical spectra indicate that the observed *cis* μ -1,2(η^1 : η^1)-peroxo complex differs from the μ -(η^1 : η^1)-peroxo proposed from spectroscopic studies of a reactive intermediate formed in solution by addition of O₂ to diferrous CmlI. It is proposed that the structural changes required to open the active site also drive conversion of the μ -1,2-peroxo species to the reactive form.

Keywords

Oxygen activation; Non-heme iron; Diiron cluster; Arylamine oxygenase; Antibiotic biosynthesis; Peroxo intermediate

Introduction

Nitro-substituted aromatic compounds (nitroaromatics) are used as synthetic building blocks for various chemicals including explosives, pesticides, and pharmaceuticals. Although

Correspondence to: John D. Lipscomb.

John D. Lipscomb lipsc001@umn.edu.

Electronic supplementary material The online version of this article (doi:10.1007/s00775-016-1363-x) contains supplementary material, which is available to authorized users.

relatively rare in nature, several natural pathways for the biosynthesis of nitroaromatics have been discovered and characterized [1, 2]. Many of these compounds are produced by soil bacteria of the *Pseudomonas* and *Streptomyces* genera and have antibiotic and medicinal properties [3–5]. The primary route to the generation of biological nitroaromatics is by direct enzymatic nitration of substrates [1]. An alternative pathway involves oxygenase enzymes, which catalyze the direct incorporation of the O atoms of molecular dioxygen into the substrate to form the nitro group [6–8]. One of the most well-known and widely used natural nitroaromatic compounds is the antibiotic chloramphenicol (CAM), which is synthesized by *Streptomyces venezuelae* [4, 8–10]. The biosynthesis of CAM employs a non-ribosomal peptide synthetase (NRPS) system and several tailoring enzymes to chemically modify a para-amino phenylalanine precursor to form the final product [8, 10–14]. The final step in CAM biosynthesis is oxidation of the arylamine function of the immediate CAM precursor D-threo-1-(4-aminophenyl)-2-dichloroacetyl-amino-1,3-propanediol (NH₂-CAM) by the N-oxygenase CmlI (Scheme 1) [8, 15].

CmlI is part of a large and diverse family of enzymes that use a carboxylate-bridged diiron cluster cofactor to activate O₂ for oxidation reactions or the hydroxylation of organic substrates [16]. Some of the enzymes found in this broad family include the hydroxylase component of the soluble methane monooxygenase (MMOH) [17, 18], the R2 subunit of class I ribonucleotide reductase (RNR-R2) [19], plant fatty acid desaturases [20–22], aldehyde-deformylating oxygenase [23, 24], aromatic hydroxylases [25, 26] and many more. All of the aforementioned enzymes coordinate the diiron cluster in a highly conserved 4-helix bundle fold and are part of the “ferritin-like” superfamily [27]. The as-isolated CmlI cluster is diferric and only reacts with O₂ upon reduction to the diferrous state [15]. Exposure of the diferrous cluster to O₂ generates a diferric peroxo species (**P**), which reacts directly with NH₂-CAM. Most of the peroxo species that have been characterized in other diiron cluster-containing enzymes have a *cis* μ -1,2(η^1 : η^1) geometry in which each O atom of the peroxo species coordinates a different iron [28–33]. In contrast, the CmlI **P** species is proposed to have a μ - η^1 : η^2 geometry in which one O atom coordinates to both metals but the other to only one [15]. At pH 9.0, the CmlI peroxo intermediate was found to be highly stable, having a $t_{1/2}$ of nearly 3 h at 4 °C. The exceptional longevity of the species makes it amenable to interrogation with various structural and spectroscopic techniques, possibly including X-ray crystallography. CmlI differs from most other members of the diiron cluster oxygenase family in that the reactive intermediate appears to be the peroxo species rather than a subsequent high valent species [15]. Also, CmlI will not catalyze the hydrocarbon hydroxylation reactions commonly observed for other members of the superfamily. It is important to determine how the active site structure of CmlI differs from those of the other diiron cluster oxygenases in order to understand how the unique geometry and chemistry of **P** are realized.

A homolog of CmlI called AurF has been characterized biochemically and structurally [7, 34–37]. AurF is isolated from *Streptomyces thioluteus* and catalyzes the conversion of para-aminobenzoate to para-nitrobenzoate in the biosynthesis of a precursor of the antibiotic aureothin [5]. Crystal structures of this enzyme have been reported in the diiron (Fe-AurF) [37] and dimanganese (Mn-AurF) forms [36]. The dimanganese form is now believed to be

catalytically inactive [37, 38]. AurF has a core 4-helix bundle structural motif, which is used to house a μ -oxo- and carboxylate-bridged diiron cluster [36, 37]. The structure of Fe-AurF with the reaction product bound showed that the product, and most likely substrate, binds in a cavity above the cluster but not directly to the metals [37]. Like CmlI, AurF forms a reactive peroxo intermediate [34]. This species is significantly less stable than that formed by CmlI, but it exhibits similar optical spectral features and reactivity. A structural comparison between AurF and CmlI might be useful in understanding the basis for stabilization of this new type of reactive peroxo intermediate as well as the mechanism of N-oxidation, but no structure of CmlI has yet been reported.

Here we report the crystal structure of an N-terminally truncated variant of CmlI in two states that have not been previously reported for this enzyme family. The truncated form of CmlI exhibits the same reactivity as the full-length enzyme and generates CAM product. Although the crystal structures reveal that CmlI 33 has an overall fold similar to that of AurF, there are marked differences in some of the secondary structure elements, the diiron cluster architecture, and substrate channel. The differences in the clusters provide insights into the geometric variability of diiron cofactors and permits structural comparisons to other family members including MMOH, RNR-R2 and the fatty acid desaturases. The structural malleability of the arylamine oxygenase clusters may represent a new type of catalytic regulation relevant to their ability to effect biological N-oxygenation, control the progression of catalysis, and determine the type of oxidizing species that is generated.

Materials and methods

Mutagenesis, overexpression, and purification

Polymerase chain reaction (PCR) was used to generate the 33 amino acid N-terminally truncated version of CmlI (CmlI 33) in a pVP91A expression plasmid (Arizona State University plasmid repository). The primers used to generate the truncated coding sequence from the full-length pVP template were as follows (gene coding sequence underlined and restriction sites bolded):

5' AAAAGCGATCGCGGAGAACGCGGTCATC 3' (forward).

5' AAAAGTTTAAACTCATCGGGTCACCGTC 3' (reverse).

The amplicon generated from PCR was cloned into a pCR4-TOPO vector (Thermo Fisher Scientific) and transformed into XL1 Blue cells. The CmlI 33-TOPO vector was isolated and digested with *SgfI* and *PmeI* restriction enzymes (New England Biolabs) and ligated into a similarly digested pVP91A vector before transforming into XL1 Blue cells for plasmid propagation. Overexpression and purification of the truncated enzyme were accomplished using the same methods as previously described for the full-length enzyme [15].

Kinetic characterization of CmlI 33 and enzyme reaction product analysis

NH₂-CAM was purchased from Toronto Research Chemicals and stored as a concentrated stock in methanol at -20 °C prior to use. All assays were performed in 50 mM Bicine buffer at pH 9.0 in order to facilitate comparisons to the full-length enzyme. The **P** intermediate

was generated using the methods described previously [15]. Transient kinetic experiments were performed using an Applied Photophysics SX.18MV stopped-flow spectrophotometer at 4.5 °C. For these experiments, the **P** was generated beforehand, loaded into the instrument and reacted with NH₂-CAM in aerobic Bicine buffer. These reactions were monitored at 500 nm to follow the decay of **P**. Product analysis was performed by incubating preformed **P** (150–200 µM) with 0.5 equivalents of NH₂-CAM substrate on ice for 30 min. The reaction product was separated from the enzyme using an Amicon centrifugal filter (3 kDa MWCO). The flow-through was acidified by addition of formic acid to 1 % (vol/vol). High-performance liquid chromatography (HPLC) was performed using a Waters 1525 model binary HPLC pump and elution monitored using a Waters 2487 detector. The acidified samples were injected into a Phenomenex Luna hexyl-phenyl column (5 µm particle size, 150 × 4.6 mm) and eluted using a gradient from 5 % to 95 % acetonitrile in 0.1 % formic acid at a flow-rate of 2 ml/min. Under these conditions NH₂-CAM had a retention time of 1.1 min and CAM 5.8 min. The elution was monitored at 280 nm. Liquid chromatography/mass spectrometry (LC/MS) samples were prepared by quenching the enzyme reaction with 2.5 % (vol/vol) trifluoroacetic acid and centrifuging to collect the precipitated protein pellet. The samples were separated by reverse-phase chromatography on a Waters Acquity UPLC system and analyzed using a Waters Synapt G2 QTOF mass spectrometer at the University of Minnesota Chemistry Department Mass Spectrometry Laboratory. The product had a *m/z* of 321.01 ([M-H][−]), the same as observed for authentic CAM standard.

Crystallization of CmlI 33 and structure determination

Initial crystallization conditions of CmlI 33 were determined from screens at the University of Minnesota Nanoliter Crystallization Facility. 5–10 µm crystals were obtained from solutions containing 10–20 % polyethylene glycol (PEG) 3350 and 0.2 M ammonium L-(+)-tartrate. The conditions were optimized for hanging-drop crystallization and modified to contain 10–20 % PEG 1500, 0.15 M ammonium v-(+)-tartrate, and 0–2 % glycerol. Typical drop ratios contained 1–1.5 µl mother liquor and 1 µl CmlI 33 at 15–20 mg/ml concentration with a well reservoir volume of 0.5–1.0 ml. The pH of these drops was near 6.8. Crystals were obtained at 20–22 °C and at 4 °C. Crystals of CmlI 33 were chemically reduced by placing single crystals in a drop containing Ar-sparged anaerobic mother liquor with 25 % glycerol, 40 mM sodium dithionite and a catalytic amount of methyl viologen. The crystals were incubated for 10 min in an anaerobic chamber before flash-freezing in liquid N₂.

Initial diffraction data of the crystals were collected in-house at the University of Minnesota Kahlert Structural Biology Laboratory at 100 K. Crystals were soaked in mother liquor supplemented with 25 % glycerol as cryoprotectant and flash-frozen in liquid N₂. Data reduction of the initial data sets indicated that the crystals belonged to a trigonal P321 space group. This is an enantiomorphic space group, which could possess either a P3₂ or P3₁ screw-axis symmetry. As such, the diffraction data were processed separately in both space groups. The P3₂21 enantiomorph was determined to be the correct space group during the molecular replacement process. The structure was solved by molecular replacement with the Phaser MR software [39] using a trimmed model of an AurF monomer [37]. Following the initial solution, CmlI was iteratively modeled using Coot [40] and refined using Refmac5

[41]. Synchrotron-derived diffraction data for the reported structures were collected at the Structural Biology Center at Argonne National Laboratories, Beamline 19-ID at 100 K. All structure figures were produced using PyMOL Molecular Graphics System, Version 1.5, Schrödinger, LLC.

Single crystal micro-spectrophotometry

Optical absorbance spectra of CmlI and CmlA crystals at 100 K were collected using an on-line micro-spectrophotometer system (Hamamatsu D4H light source, Ocean Optics QE65000 spectrograph, Newport reflective Schwarzschild objectives) at Stanford Synchrotron Radiation Lightsource, Beamline 11-1.

Detection and quantitation of hydrogen peroxide in polyethylene glycol stocks

Crystalline bovine liver catalase was purchased from Sigma. PEG 1500 stocks (50 % w/vol) were purchased from Hampton Research and stored in the dark at 4 °C until use; these same stocks were also used for the crystallization of CmlI [33]. Hydrogen peroxide (H₂O₂, 29–32 %, ACS grade) was purchased from Millipore. Assays to detect H₂O₂ were performed using a Hansatech oxygen polarimeter (<http://www.hansatech-instruments.com>) to observe the formation of O₂ produced from the disproportionation of H₂O₂ catalyzed by catalase at 23 °C. Assays contained 0.2 mg catalase, 30 % PEG 1500 and 20 mM HEPES buffer at pH 7.0. The method was calibrated using assays with a known amount of H₂O₂. Assays containing 30 % glycerol instead of PEG were used as a negative control. Three different PEG 1500 stocks were tested for the presence of H₂O₂ and these experiments showed that the total concentration of H₂O₂ ranged between 20 and 80 μM.

Reaction of CmlI with hydrogen peroxide and product analysis

To monitor the reaction between CmlI and hydrogen peroxide in solution, 100 μM enzyme (post mix) was reacted with a 40-fold excess of hydrogen peroxide using a stopped-flow spectrophotometer. These reactions were performed in 0.1 M Bicine buffered at pH 7.8 at 4.5 °C. In order to avoid photoreduction of CmlI from the intense white light during the long experiments, the lamp shutter of the instrument was manually opened immediately before each spectrum was collected and otherwise kept closed between scans. To assay for product in the H₂O₂ reactions, 100 μM CmlI was incubated with 1 equivalent of NH₂-CAM and varying concentrations of H₂O₂ (0.6–40 mM) for 3 h at 23 °C. The reactions were stopped by precipitating the protein with 2.5 % TFA and centrifuging to collect the pellet. The supernatant was assayed for product using HPLC as described above.

Results

Rationale for generation of the N-terminal CmlI truncation variant

Attempts to crystallize the full-length wild-type (WT) CmlI with or without the N-terminal His purification tag were unsuccessful. This was unexpected because AurF, a homolog with 37 % sequence identity with CmlI (based on a full-length sequence alignment), has been crystallized in two different space groups and structurally characterized [36, 37]. The two enzymes are of similar overall length (Fig. S1), but a sequence alignment showed that CmlI contains a stretch of ten amino acids near the N-terminus that are not present in AurF.

Analysis of the available crystal structures of AurF shows that the N-terminus is positioned near a protomer packing interface in the unit cell (PDB structures 3CHU and 3CHI) [37]. It was theorized that the additional ten amino acid stretch of residues in CmlI could be interfering with crystal nucleation and growth. To test this possibility, a truncated variant of the enzyme was generated where the first 33 amino acids were removed (CmlI 33, start site marked with a black star in Fig. S1). This site was chosen based on the observation that in two of the three AurF crystal structures the first 23 amino acids are not visible presumably because of conformational flexibility [37]. Either due to deletion of the differential sequence or the decrease in conformational disorder, CmlI 33 could be crystallized and its structure solved.

Biochemical characterization of CmlI 33

The reaction of CmlI 33 was studied with the native substrate NH₂-CAM and the results compared to those reported for the full-length enzyme [15]. CmlI is able to generate the long-lived diferric-peroxo species **P** ($t_{1/2}$ ~3 h, 4 °C) at pH 9.0 after exposure of the chemically reduced enzyme to O₂. **P** is characterized by a weak optical band centered at 500 nm (ϵ ~0.5 mM⁻¹ cm⁻¹). This species was also generated in CmlI 33 using the same methods and conditions as for full-length enzyme (Fig. 1a, red trace). Titration of **P** with NH₂-CAM indicated that roughly 0.35 equivalents of substrate were sufficient to fully consume the species under the conditions employed (Fig. 1b). A similar 3:1 stoichiometry was observed for full-length CmlI under the same conditions [15]. CmlI 33-**P** was also able to generate the CAM product in high yield upon reaction with NH₂-CAM as determined using high-performance liquid chromatography and liquid chromatography/mass spectrometry (LC/MS) (see Fig. 1c; “Materials and methods”). Substrates have been shown to react with the **P** in what appears kinetically to be a second order reaction [15]. Accordingly, reaction of pre-formed CmlI 33-**P** with different concentrations of NH₂-CAM substrate in buffer at 4.5 °C suggested that the truncated variant also reacts directly with substrate, as evidenced by the linearity of the peroxo decay rate constant vs. NH₂-CAM concentration plot (Fig. 1d). The zero-intercept and linearity of the plot means that the reaction is effectively second order and irreversible. The slope of the fit to the plot gives the apparent bimolecular rate constant for the reaction of CmlI 33-**P** with substrate. The value for CmlI 33 was comparable to that obtained for the full-length enzyme, 12 vs. 7.1 mM⁻¹ s⁻¹, respectively. The experiments performed here demonstrate that CmlI 33 is functionally the same as the full-length enzyme. Thus, the findings derived from a structure of the truncated mutant are likely to be applicable to the natural form of CmlI. The lifetime of **P** is much shorter at pH 6.8, obviating the detailed analysis possible at pH 9.0. However, addition of NH₂-CAM and O₂ to diferrous CmlI 33 at pH 6.8 rapidly yields CAM in the expected yields.

Crystallization of CmlI 33 and its overall structure

Crystals of CmlI 33 were obtained in conditions containing 10–20 % PEG 1500, 0.15 M ammonium L-(+)-tartrate and 0–2 % glycerol. The enzyme only crystallized near pH 6.8; attempts to generate crystals at higher pH values 7.5–9.0 were unsuccessful. Data collection and model refinement statistics for the enzyme crystal structures before (PDB ID: 5HYG) and after (PDB ID: 5HYH) chemical reduction are given in Table 1. The enzyme crystallized

in a trigonal P₃2₁ space group with one CmlI 33 monomer in the asymmetric unit (Fig. 2a). Density was missing for the N-terminal purification tag, a portion of the C-terminal section of the protein and a disordered stretch of residues between K52 and F62 (dashed line in Fig. 2a). CmlI contains ten α -helices (α 1– α 10) which are labeled in Fig. 2a. Like many other diiron cluster-containing enzymes, CmlI coordinates the diiron cluster in a 4-helix bundle. His and Glu residues in helices α 2, α 3, α 5, and α 6 are ligands for the cluster. A structural search of the CmlI protein using the Dali server (http://ekhidna.biocenter.helsinki.fi/dali_server/) shows that the closest structural hit is Mn-AurF (RMSD 1.5 Å, main chain atoms) [36] followed by Fe-AurF (RMSD 2.1 Å, main chain atoms) [37]. The next closest hits are the R2-subunits of class I ribonucleotide reductases and the α -subunit of the hydroxylase component of soluble methane monooxygenase (9–12 % sequence identity based on a full-length alignment, RMSDs 3.0–3.4 Å) [18, 19]. The biological unit of CmlI is a dimer and its interface lies along one of the crystallographic symmetry axes in the unit cell (Fig. 2b). The CmlI dimer interface buries an area of 1100 Å² on the surface of helices α 2 and α 3. This results in a 22 Å separation of the two diiron clusters, the same as in AurF.

In the proposed catalytic model for CmlI, the arylamine substrate binds near the diiron cluster for reaction with activated O₂ [15]. The crystal structures of Fe-AurF revealed a clear channel for substrate ingress to the interior of the protein and to the buried cluster [37]. In contrast, in Mn-AurF this channel is collapsed with no clear open path for substrate to access the active site [36]. A putative substrate channel was also identified in CmlI 33 (Fig. S2a). The entrance is located between helices α 2, α 5, α 7, and α 9 (Fig. S2b and orange arrow in Fig. S2a). The channel is lined by the following amino acids: F69, M73, P75, Y101, N104, I108, L171, S174, I176, I208, Y211, F273, D277, F278, W281 and L309. Some of these residues also encompass a relatively large cavity that is just inside the entrance to the substrate channel (starred in Fig. S2a). Above the diiron cluster, this channel contracts to a diameter that is too small to accommodate NH₂-CAM.

Comparison of the overall structure of CmlI 33 to those of Fe-AurF and Mn-AurF

A structural alignment of CmlI 33 with Fe-AurF shows that there are differences in the arrangements of some of the α -helices. The most notable of these is a break in the secondary structure of one of the four core helices (equivalent to α 5 in CmlI 33). In Fe-AurF, there is a stretch of residues numbered 197–201 that are not α -helical and are pushed out away from the substrate channel and the cluster (Fig. 3a, purple cartoon). In contrast, in CmlI 33, this region remains α -helical (Fig. 3a, gray cartoon). This has the effect of collapsing the channel above the cluster in CmlI 33, whereas it remains open in Fe-AurF (Fig. 3b, c). In particular, the side-chain of residue I208 (I199 in AurF) occludes the space above the CmlI 33 cluster, in contrast to Fe-AurF where this side-chain is pointing away from the channel. Interestingly, the structures of CmlI 33 and Mn-AurF are nearly identical in this respect. In both structures, the region shown in Fig. 3a remains α -helical and I208/I199 is pushed toward the channel and blocks access to the active site. A structural alignment of the CmlI 33 structure with the Fe-AurF para-nitrobenzoate product complex (PDB ID: 3CHT) is presented in Fig. 3d. The side chain of I208 in CmlI 33 occludes the binding site for the

product and, likely, substrate. It will be shown in the next section that the position of I208 also has an important effect on one of the diiron cluster ligands.

In an attempt to crystallize an alternate form of CmlI where the channel was open as in Fe-AurF, we tried to co-crystallize CmlI 33 with various substrates including NH₂-CAM. Unfortunately, this did not yield any alternate crystal forms, nor was there any evidence for substrate binding to CmlI 33 in crystals prepared under these conditions. Chemical reduction of the cluster had no significant effect on the size of the substrate channel or the orientation of the residues that line the channel. Due to the similarity of the CmlI 33 structure to Mn-AurF, the metal content of the preparation was determined. Inductively coupled plasma mass spectrometry (ICP-MS) of the acid-digested enzyme used for the crystallization revealed that that iron was present in roughly stoichiometric ratio (1.9:1) relative to the amount of protein. Iron was also present at nearly 20-fold higher concentration than any of the other metals, making it highly unlikely that these other metals were present in the active site above a trace level. No detectable amount of Mn was present in these samples. It should also be noted that if CmlI 33 were loaded with non-cognate metals, it would not be expected to exhibit the reactivity or be able to generate the product at the yield that we observe.

Structure of the CmlI 33 diiron cluster prior to chemical reduction

The diiron cluster is coordinated by four Glu and three His ligands (Fig. 4a, b). All cluster ligands are derived from the core 4-helix bundle: E109 (from helix α 2), E144 and H147 (α 3), E205 (α 5), and H232, E236, and H239 (α 6). Both metals are in a six-coordinate distorted octahedral environment with five protein-derived ligands each. Fe1 is coordinated by E109, H232 (N^e mode), H147 (N^δ mode) and E144 and E236 (monodentate mode). Fe2 is bound by the two bridging carboxylates E144 and E236 (bidentate mode), H239 (N^δ mode) and E205. E144 bridges the irons in a μ -1,3 mode and E236 in a μ -1,3(η^1 : η^2). Both irons are bound to a bridging exogenous ligand (see below). In this and the chemically reduced structure, both Fe atoms were modeled at full occupancy based on the atomic B-factors after refinement and the ICP-MS analysis. The Fe-Fe distance in the cluster is 3.3 Å, which is similar to the distance in several other diiron clusters containing enzymes [19, 22, 23, 42, 43], but it is shorter than in the diferric Fe-AurF homolog (3.45 Å) [37].

A region of strong residual electron density is observed between the two irons which is significantly larger than can be accounted for by a single water. A bridging *cis* μ -1,2-peroxo ligand modeled into this density accounts well for the electron density, as shown by the ligand-omit map of Fig. 5. This peroxo moiety was modeled at an occupancy of 0.7 as this resulted in comparable B-factors with the nearby metals and residue atoms. The refined bond distances for the peroxo are 1.8 Å (Fe1-O1), 2.0 Å (Fe2-O2), and 1.5 Å (O1-O2). The model angles are 126° ($\angle$ Fe1-O1-O2), 107° ($\angle$ Fe2-O2-O1) and dihedral -38° ($\angle$ Fe1-O1-O2-Fe2). These values compare well with other structurally characterized diferric *cis* μ -1,2-peroxo species in diiron proteins or model compounds (Table 2) [44–49]. Overall, the CmlI 33 iron-peroxo bond distances (1.8, 2.0 Å) are more similar to those observed in inorganic model complexes than in other diiron protein crystals, which exhibit longer bonds. In CmlI, the carboxylate functions of E109 and E205 are positioned within bonding distance

of both peroxo O atoms (2.6–2.9 Å). The orientation of the CmlI peroxo unit is distinct from that in the toluene 4-monooxygenase hydroxylase (T4moH) *cis* μ -1,2-peroxo complex (PDB ID: 3I63) [47] (Fig. S3), likely due to the different arrangement of the terminal carboxylate ligands.

The source of the oxygen for the putative peroxo moiety is unclear. O₂ might bind to the diiron cluster in the crystal if the irons become reduced. However, the enzyme is crystallized and frozen at 100 K before being exposed to X-ray radiation (a possible source of reducing species) and binding of O₂ at cryogenic temperatures is unlikely. We have observed that the enzyme in solution is readily photoreduced by intense white light, so it is possible that long-term aerobic incubation in ambient light conditions during crystallization might allow formation of the peroxo adduct. However, a more likely route leading to formation of observed peroxo complex in crystallo is provided by the components in the crystallization solution. As shown in Fig. S4, PEG used for the crystallization contains a low level of H₂O₂, presumably generated during its slow decomposition. Exposure of aqueous solutions of PEGs to higher temperatures in the presence of O₂ and light is known to accelerate its decomposition, generating H₂O₂ [50]. Thus, during crystallization, long-term incubation of the diferric CmlI in solution containing PEG (and residual H₂O₂) might directly produce a peroxo adduct. In order to test this possibility, as-isolated diferric CmlI was mixed with varying concentrations of H₂O₂ in solution. With a single equivalent of H₂O₂, no changes were observed in the optical spectrum even after several hours, but exposure to an excess of H₂O₂ (40 equivalents) yielded changes in the optical spectrum within several minutes. A new species with a broad chromophore with substantial absorbance at 600 nm and above formed over the course of an hour with an isosbestic point observable near 400 nm (Fig. 6a, inset). The spectrum of this species (Fig. 6a, black trace) is distinct from that of the reactive **P** intermediate (red trace). This new species was found to be quite stable, and no decay was observed even after several hours of incubation at 4.5 °C. However, no CAM product was detected when CmlI was incubated with substrate and H₂O₂ for extended periods.

No evidence for a μ -oxo ligand is observable in the CmlI crystal structure, whereas the optical characteristics of the as-isolated enzyme in solution at pH 9.0 suggested the presence of such a moiety (Fig. 1a inset, blue trace) [15, 51]. The optical spectrum of the CmlI 33 in solution at pH 6.8 exhibits only minor changes throughout the visible range, but retains the intense charge-transfer band near 350 nm that is characteristic of the oxo-bridge (Fig. 1a inset, magenta trace). It is possible that the continued presence of the strong oxo-bridge in solution accounts for the failure of the peroxo adduct formed by addition of H₂O₂ to the diferric enzyme to convert to **P**.

In order to investigate the discrepancy between the absence of a μ -oxo-bridge in the diffraction data and the presence of its optical signature in solution, single crystal optical absorbance spectra were collected of CmlI 33 crystals. As shown in Fig. 6b, the optical spectra of CmlI 33 crystals before and after exposure to X-rays at the synchrotron (magenta and purple traces, respectively) were largely bleached in the visible, consistent with the lack of μ -oxo-bridge. As a control, the spectrum of the enzyme CmlA in a crystal form was also recorded and showed the intense spectral band of the μ -oxo-bridge, which is known to be present in that diiron enzyme (Fig. 6b, orange trace). In solution, the oxo-iron charge-

transfer bands of diferric CmlI and CmlA are of similar intensity (Fig. 6b, inset). The spectrum of the CmlI 33 in crystalline form is also distinct from that of **P** because the latter intermediate exhibits absorption similar to that of the diferric CmlI 33 in the 350 nm region (Fig. S5). In the absence of a crystal containing **P**, it is difficult to determine whether the crystal of CmlI 33 peroxo species also lacks the characteristic chromophore in the 500–550 nm region. Attempts to obtain an optical spectrum of chemically reduced CmlI 33 crystals were unsuccessful due to the presence of sodium dithionite in the surrounding mother liquor, which has a significant absorbance in the near-UV region.

The structure of the CmlI 33 cluster overlays well with that of Fe-AurF with a few exceptions. The most dramatic differences are the presence of the peroxo-bridge in CmlI and the rotation of the carboxylate of the CmlI 33 E236 to occupy the site where the μ -oxo-bridge is found in the Fe-AurF structure (Fig. 7a) [37]. As noted, E236 bridges between the metals in a μ -1,3($\eta^1:\eta^2$) mode, and the carboxylate function lies in the same plane as the Fe–Fe vector. This same coordination mode of the Glu residue is observed in Mn-AurF crystal structure, which also lacked a μ -oxo-bridge (Fig. 7b) [36]. In Fe-AurF, the equivalent residue to E236 of CmlI is E227, which is instead found in a monodentate mode, coordinated only to Fe2. The difference in orientation of this residue is due to a projection of the side-chain of CmlI residue I208 towards the cluster (Fig. 3a, gray sticks and spheres), which was discussed in a previous section. This positioning of I208 forces the carboxylate and C γ of E236 into the equatorial coordination plane of the cluster. In AurF, E227 is presumably able to assume the monodentate orientation because the side-chain of the equivalent I199 residue is pointing away from the cluster (Fig. 3a, purple sticks and spheres). In MMOH and the D84E variant of RNR-R2, steric effects from a second sphere residue are proposed to change the coordination mode of a bridging carboxylate ligand [52]. It is interesting to note that both the CmlI μ -1,2-peroxo complex and AurF with an oxo-bridge have overall neutral charge at the diiron center. This control over buried charge is commonly observed for O₂ activating diiron cluster-containing enzymes, and it is maintained at each stage of their reaction cycles [26, 53]. Accordingly, it is unlikely that the structurally characterized state of diferric AurF would have a high affinity for peroxide, because this would dramatically change the net charge.

Structure of the chemically reduced diiron cluster

Crystals of CmlI 33 were chemically reduced under anaerobic conditions using methyl viologen as mediator dye and sodium dithionite as a source of reducing equivalents. No large changes in the crystal unit cell parameters were noted compared to the crystals prior to reduction. The largest change upon reduction is the expected loss of the putative peroxo moiety. Otherwise, only relatively subtle changes were found to occur (Figs. 4c, d, 7c). Both of the metals are roughly five-coordinate in a nearly square pyramidal geometry. There is an open site on each iron where the peroxo ligand was bound. The coordination mode of all of the protein-derived ligands remains essentially the same, but the overall Fe–ligand bond distances are increased relative to the CmlI cluster described above (Fig. 4a, c). The position of the I201 side-chain is unaltered, locking the side-chain of E236 into the same plane as in the peroxo-bound state [54, 55]. Reduction and loss of the peroxo ligand cause an increase in the Fe–Fe separation from 3.3 to 3.6 Å. The increase in Fe–Fe separation is

accommodated by a mild deformation of the two terminal cluster ligands E109 (Fe1 ligand) and E205 (Fe2).

Discussion

The structure of CmlI 33 reported here in two states presents the opportunity to make comparisons with the homologous arylamine oxygenase AurF as well as other diiron cluster-containing oxygenases that catalyze different types of reactions. In particular, the availability of the structure of a *cis* μ -1,2-peroxo complex allows comparison to the structures of other diiron peroxo complexes as well as the significantly different proposed structure of the reactive **P** intermediate of CmlI. The **P** intermediate is formed by adding O₂ to the diferrous enzyme in solution, and one form of the fully reduced enzyme is now structurally characterized for the first time. It is significant that both the diferrous and *cis* μ -1,2-peroxo-bound states of CmlI have active sites that are too constricted to accommodate the substrate. A similar restricted active site also has been reported for Mn-AurF, suggesting that this alternative structure may have an unanticipated importance to catalysis. It is possible that the structural changes required to allow substrate binding by CmlI also engender conversion of the *cis* μ -1,2-peroxo intermediate observed in the crystal to the active **P** state previously characterized in solution. This possibility and its relationship to regulation are discussed here vis-à-vis the unique reactivity of the aromatic amine oxygenases. Indeed, CmlI is unable to catalyze hydrocarbon oxidation despite the presence of a diiron cluster that can form *cis* μ -1,2-peroxo intermediate similar to those found in other oxygenases such as MMOH and T4moH [47, 53, 56]. The basis for this alternative reactivity as well as the greatly increased stability of the reactive **P** intermediate is currently unknown. These aspects of the reactions of diiron arylamine oxygenases are further evaluated here.

Trapping of the *cis* μ -1,2-peroxo complex

The ability to trap and structurally characterize the effectively stable *cis* μ -1,2-peroxo complex in a crystal of CmlI is unexpected because in previous studies of CmlI, a structurally different peroxo species that slowly decays was observed [15]. However, in several recent studies, we have observed that intermediates with limited lifetime in solution are greatly stabilized when the reaction is carried out *in crystallo* [57–59]. It is also possible that significant structural rearrangements in the cluster structure are required to convert the observed peroxo species to the reactive **P** intermediate, and this may not be possible in the crystal lattice as discussed below.

Significance of the single crystal optical spectra

The lack of intense near-UV optical bands in the CmlI 33 crystals (Fig. 6b) is entirely consistent with the observed lack of a μ -oxo-bridge in the crystal structure (Fig. 4a, b). Furthermore, the very similar optical spectra of the CmlI 33 crystals pre- and post-X-ray exposure also strongly argues against the possibility that photoreduction during X-ray data collection is responsible for the altered diiron cluster structure and channel relative to those in Fe-AurF. It is also inconsistent with the generation of the observed *cis* μ -1,2-peroxo adduct by reaction of O₂ with the photoreduced diiron cluster under cryogenic conditions. This means that the CmlI 33 cluster is already in the observed orientation before X-ray

exposure and that the peroxo adduct is generated prior to freezing of the crystals. Since CmlI in solution is clearly able to react with free H_2O_2 (Fig. 6a), the most likely source of the peroxo adduct in the CmlI 33 crystals is the residual H_2O_2 present in the crystallization solution.

Several lines of evidence suggest that the peroxo species characterized structurally and in the single crystal optical spectra is not **P**. First, the spectroscopic properties of **P** are distinct from the peroxo intermediate described here as well as those of previously characterized *cis* μ -1,2-peroxo intermediates. In particular, the intense band at ~350 nm of **P** is missing in the single crystal optical spectra of CmlI 33 (compare Figs. 6b; S5). Second, the conversion of the observed *cis* μ -1,2-peroxo adduct to the μ - η^1 : η^2 peroxo structure proposed for **P** would require a rearrangement of the E236 cluster ligand to the monodentate mode observed in Fe-AurF. This is unlikely to occur in crystals of the enzyme as the transition requires a rearrangement of the protein backbone near the active site which, based on the protomer interactions in the unit cell, would destabilize the crystal lattice.

Significance of the collapsed substrate channel to the CmlI catalytic cycle

The absence of the μ -oxo-bridge in the crystal structure of CmlI 33 likely to be of little mechanistic consequence in-and-of itself, as this moiety will probably be lost upon reduction of the cluster in solution. Unfortunately, no structure has been reported of diferrous Fe-AurF, which would allow evaluation of the continued presence of the bridge (or a protonated form) in that enzyme. The more mechanistically significant differences between the structures of CmlI 33 and Fe-AurF are the position of CmlI E236 versus AurF E227 and the constricted substrate-binding pocket of CmlI 33. As described above, these differences are related to each other by the effects of the E236/E227 residue on the stability of a critical helical region near the cluster. The shift from helix to coil in this region determines whether substrate can bind effectively. This large structural change opens the question of whether the closed active site structure has a role in the catalytic cycle. The finding of the nearly identical closed structure in the Mn-AurF supports the possibility that the closed structure is a real alternative form of the amine oxygenase that has a role in regulation of catalysis. It is clearly possible to react the reduced forms of CmlI, CmlI 33, and Fe-AurF with O_2 to form the peroxo-intermediates in the absence of substrate. However, to date, there is no experimental evidence for the formation of a substrate complex prior to O_2 binding. Indeed, the kinetics of substrate reaction with the peroxo-intermediates are apparently second order to the limit of substrate solubility, consistent with a reaction that occurs effectively by collision of the substrate with the pre-formed peroxo-species [15]. One possibility is that the observed structures for CmlI 33 (and Mn-AurF) and Fe-AurF represent extremes of the conformational range. O_2 binding to the reduced states followed by reorganization to the μ - η^1 : η^2 peroxo intermediate **P** locks CmlI E236 or AurF E227 in a rotated orientation bound only to Fe2 that, in turn, opens the channel to allow substrate to bind and the reaction to proceed. This scenario, explored further in the context of MMOH below, would represent a novel strategy to achieve the common goal of diiron oxygenases to coordinate formation of the reactive species with its exposure to substrate [43, 60–63].

Dynamic cluster ligands potentially related to activity

In MMOH, the terminal Fe2 ligand E209 has been implicated in gating O₂ binding to the diferrous cluster in response to the binding of the regulatory protein MMOB [61]. Computations based on VTVH MCD measurements indicate that this residue shifts several degrees relative to the Fe–Fe vector as MMOB binds to the MMOH surface to open access to the available iron ligands sites from bulk solvent during formation of the *cis* μ -1,2-peroxo intermediate. The equivalent residue E205 in diferrous CmlI 33 is approximately in the putative shifted position of the MMOH E209 (Fig. 8, starred). This may, in part, explain why diferrous CmlI 33 readily binds O₂ to ultimately form intermediate **P** while the reaction rate constant of diferrous MMOH with O₂ is decreased 1000-fold in the absence of MMOB [64].

Another dynamic residue found in several diiron cluster-containing enzymes is the iron ligand equivalent to CmlI E236. As described above, this residue forms a μ -1,3(η^1 : η^2) carboxylate bridge in CmlI 33. In diferric RNR-R2 and MMOH, the equivalent residue is a terminal monodentate Fe2 ligand very similar to that seen in Fe-AurF. In these other oxygen activating enzymes, the carboxylate bridge of CmlI 33 E236 is replaced by either a μ -oxo or μ -hydroxo-bridge. Reduction of the RNR-R2 cluster results in loss of the μ -oxo-bridge and a shift in the monodentate Fe2 Glu ligand to a bridging μ -1,3 mode [65]. In MMOH upon cluster reduction, the Glu residue in this position (E243) undergoes a significant shift to μ -1,3(η^1 : η^2) bridging mode that is superimposable with that of E236 in CmlI 33 [54] (Fig. 8). MMOH E243 has a major regulatory role in the catalytic cycle of the enzyme in that it is part of the process that converts the stable, coordinately saturated ferric diiron cluster [18, 66] to a cluster with vacant sites or readily displaced solvents in the diiron cluster coordination sphere, allowing it to bind O₂. A comparison of the structures of reduced CmlI 33 and reduced MMOH shows that the equivalent sites for O₂ binding are available in each case.

Parallels in the interconversion of oxy-intermediates in CmlI and MMOH

Although CmlI 33 and MMOH/MMOB can both form *cis* μ -1,2-peroxo species, the fates of these species in the catalytic cycle and the reactions catalyzed by the enzymes are quite different. In CmlI the *cis* μ -1,2-peroxo species would be a logical precursor to intermediate **P** currently assigned as a μ - η^1 : η^2 peroxo species [15]. If **P** does, in fact, evolve from the *cis* μ -1,2-peroxo complex, at least one peroxo-oxygen must reorient significantly and a second ligand site on one of the irons for peroxo-oxygen binding must become available. In doing so, the most likely new ligand site to occupy is that of the μ -oxo moiety (seen in solution, open conformation) or E236 (seen in the crystal, closed conformation). Recently, we have studied the structure of the reactive intermediate of the MMOH reaction cycle, compound **Q** [67]. Continuous flow resonance Raman spectra show that **Q** has a diamond core structure with two μ -oxo-bridges where both oxygens were derived from the same O₂ molecule. One way to form **Q** with this structure from the *cis* μ -1,2-peroxo precursor is to effect a *cis*-to-*trans* μ -peroxo reorganization prior to O–O bond cleavage. The resulting intermediate would be structurally reminiscent of the μ - η^1 : η^2 species proposed for CmlI-**P** (Scheme 2). Indeed, evaluation of the potential ligand sites suggests that the MMOH μ -1,3(η^1 : η^2) bridging E243 would be displaced in the reorganization. This residue is the equivalent to μ -1,3(η^1 : η^2) bridging E236 carboxylate that we propose must be displaced in CmlI in order to convert the

observed *cis* μ -1,2 peroxo to CmlI-**P**. The formation reaction of **Q** is pH dependent and exhibits a linear proton inventory plot, suggesting that a proton from a cluster ligand or an immediately adjacent residue plays a key role in breaking the O–O bond [53, 56, 67, 68]. One difference between the diiron cluster structures of MMOH and CmlI is the replacement of Fe1 water in the former by a H232 in the latter (Fig. 8). If this water is the ultimate source of the proton involved in **Q** formation, the inability of a ligand His to provide an equivalent proton may account for the stability of the **P** intermediate of CmlI. It is noteworthy, that this scenario would place the reactive oxygen of **Q** and the peroxo-oxygen facing the substrate-binding pocket of CmlI in approximately the same position (Scheme 2). Thus, CmlI and MMOH may follow very similar paths to the reactive species both mechanistically and structurally up to the point where the O–O bond breaks. This final step would determine the type of reactivity exhibited by each enzyme.

Conclusions

The overall structure of CmlI 33 is similar to that reported for Fe-AurF. However, the detailed structure of the active site diiron cluster has distinct differences. These differences recapitulate the structural differences reported for Fe-AurF and inactive Mn-AurF. In contrast to Mn-AurF, CmlI 33 is fully active. These finding suggests that the observed CmlI 33 structure is relevant to the functional enzyme, despite the lack of the μ -oxo-bridge detected optically in solution. The lack of large structural changes between reduced CmlI 33 and the *cis* μ -1,2-peroxo complex suggests that, in the normal order of catalysis, the peroxo complex could form by O₂ binding to the closed form of the enzyme. However, it is evident that the most dynamic residue, E236, must shift at some point to open the active site for substrate binding. The trigger for this shift may be a progression in the O₂ binding process to yield intermediate **P**. This complex process may be part of a mechanism to limit adventitious chemistry of the activated O₂ species, a common feature of most oxygenase reaction cycles.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health Grants GM 100943 and GM 118030 (to J.D.L.) and NIH graduate traineeship GM 08700 (to C. J. K.). We would like to thank Carrie Wilmot and her research group for many helpful discussions. We also thank Ed Hoeffner at the University of Minnesota Kahlert Structural Biology Laboratory for suggestions in indexing the CmlI 33 data and collecting the data sets. We also thank Klaus Lovendahl for generating the expression construct for the truncated enzyme and Anna Komor and Brent Rivard for assistance with biochemical characterization. Diffraction data were collected at Argonne National Laboratory, Structural Biology Center Beamline 19-ID at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the US Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We are thankful for computational resources from the Supercomputing Institute and the facilities at the Kahlert Structural Biology Laboratory at the University of Minnesota.

Abbreviations

CmlI Full length wild-type CmlI enzyme

CmlI 33	33 amino acid N-terminally truncated variant of CmlI used for crystallization
CAM	Chloramphenicol
NH₂-CAM	D-threo-1-(4-aminophenyl)-2-dichloroacetyl-amino-1,3-propanediol, the arylamine precursor of CAM and natural CmlI substrate
Fe-AurF	The functional diiron variant of the aureothin arylamine oxygenase, AurF
Mn-AurF	The catalytically inactive dimanganese variant of AurF
P	Diferriic peroxo intermediate of CmlI or CmlI 33
RNR-R2	The R2 subunit of ribonucleotide reductase which houses the diiron cluster
MMOH	Hydroxylase component of the soluble form of methane monooxygenase, MMO
T4moH	Toluene 4-monooxygenase hydroxylase

References

- Winkler R, Hertweck C. *ChemBioChem*. 2007; 8:973–977. [PubMed: 17477464]
- Ju K-S, Parales RE. *Microbiol Mol Biol Rev*. 2010; 74:250–272. [PubMed: 20508249]
- Kirner S, Hammer PE, Hill DS, Altmann A, Fischer I, Weislo LJ, Lanahan M, van Pée K-H, Ligon JM. *J Bacteriol*. 1998; 180:1939–1943. [PubMed: 9537395]
- He J, Magarvey N, Pirae M, Vining LC. *Microbiology*. 2001; 147:2817–2829. [PubMed: 11577160]
- He J, Hertweck C. *Chem Biol*. 2003; 10:1225–1232. [PubMed: 14700630]
- Lee J, Simurdiak M, Zhao H. *J Biol Chem*. 2005; 280:36719–36727. [PubMed: 16150698]
- Simurdiak M, Lee J, Zhao H. *ChemBioChem*. 2006; 7:1169–1172. [PubMed: 16927313]
- Lu HG, Chanco E, Zhao HM. *Tetrahedron*. 2012; 68:7651–7654.
- Doull J, Ahmed Z, Stuttard C, Vining LC. *J Gen Microbiol*. 1985; 131:97–104. [PubMed: 3989509]
- Makris TM, Chakrabarti M, Münck E, Lipscomb JD. *Proc Natl Acad Sci USA*. 2010; 107:15391–15396. [PubMed: 20713732]
- Fischbach MA, Walsh CT. *Chem Rev*. 2006; 106:3468–3496. [PubMed: 16895337]
- Walsh CT, Chen H, Keating TA, Hubbard BK, Losey HC, Luo L, Marshall CG, Miller DA, Patel HM. *Curr Opin Chem Biol*. 2001; 5:525–534. [PubMed: 11578925]
- Pacholec M, Sello JK, Walsh CT, Thomas MG. *Org Biomol Chem*. 2007; 5:1692–1694. [PubMed: 17520135]
- Pirae M, White RL, Vining LC. *Microbiology*. 2004; 150:85–94. [PubMed: 14702400]
- Makris TM, Vu VV, Meier KK, Komor AJ, Rivard BS, Münck E, Que L Jr, Lipscomb JD. *J Am Chem Soc*. 2015; 137:1608–1617. [PubMed: 25564306]
- Wallar BJ, Lipscomb JD. *Chem Rev*. 1996; 96:2625–2657. [PubMed: 11848839]
- Fox BG, Froland WA, Dege JE, Lipscomb JD. *J Biol Chem*. 1989; 264:10023–10033. [PubMed: 2542319]
- Rosenzweig AC, Frederick CA, Lippard SJ, Nordlund P. *Nature*. 1993; 366:537–543. [PubMed: 8255292]
- Nordlund P, Sjöberg BM, Eklund H. *Nature*. 1990; 345:593–598. [PubMed: 2190093]
- Fox BG, Shanklin J, Somerville C, Münck E. *Proc Natl Acad Sci USA*. 1993; 90:2486–2490. [PubMed: 8460163]
- Lindqvist Y, Huang WJ, Schneider G, Shanklin J. *EMBO*. 1996; 15:4081–4092.

22. Guy JE, Whittle E, Kumaran D, Lindqvist Y, Shanklin J. *J Biol Chem*. 2007; 282:19863–19871. [PubMed: 17463003]
23. Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB. *Science*. 2010; 329:559–562. [PubMed: 20671186]
24. Li N, Chang W-C, Warui DM, Booker SJ, Krebs C, Bollinger JM Jr. *Biochemistry*. 2012; 51:7908–7916. [PubMed: 22947199]
25. Pikus JD, Studts JM, Achim C, Kauffmann KE, Münck E, Steffan RJ, McClay K, Fox BG. *Biochemistry*. 1996; 35:9106–9119. [PubMed: 8703915]
26. Sazinsky MH, Bard J, Di Donato A, Lippard SJ. *J Biol Chem*. 2004; 279:30600–30610. [PubMed: 15096510]
27. Andrews SC. *Biochim Biophys Acta*. 2010; 1800:691–705. [PubMed: 20553812]
28. Skulan AJ, Brunold TC, Baldwin J, Saleh L, Bollinger JM Jr, Solomon EI. *J Am Chem Soc*. 2004; 126:8842–8855. [PubMed: 15250738]
29. Srncic M, Rokob TA, Schwartz JK, Kwak Y, Rulisek L, Solomon EI. *Inorg Chem*. 2012; 51:2806–2820. [PubMed: 22332845]
30. Liu KE, Valentine AM, Qiu D, Edmondson DE, Appelman EH, Spiro TG, Lippard SJ. *J Am Chem Soc*. 1995; 117:4997–4998.
31. Broadwater JA, Ai J, Loehr TM, Sanders-Loehr J, Fox BG. *Biochemistry*. 1998; 37:14664–14671. [PubMed: 9778341]
32. Yun D, Garcia-Serres R, Chicalet BM, An YH, Huynh BH, Bollinger JM Jr. *Biochemistry*. 2007; 46:1925–1932. [PubMed: 17256972]
33. Vu VV, Emerson JP, Martinho M, Kim YS, Münck E, Park MH, Que L Jr. *Proc Natl Acad Sci USA*. 2009; 106:14814–14819. [PubMed: 19706422]
34. Korboukh VK, Li N, Barr EW, Bollinger JM Jr, Krebs C. *J Am Chem Soc*. 2009; 131:13608–13609. [PubMed: 19731912]
35. Li N, Korboukh VK, Krebs C, Bollinger JM Jr. *Proc Natl Acad Sci USA*. 2010; 107:15722–15727. [PubMed: 20798054]
36. Zocher G, Winkler R, Hertweck C, Schulz GE. *J Mol Biol*. 2007; 373:65–74. [PubMed: 17765264]
37. Choi YS, Zhang H, Brunzelle JS, Nair SK, Zhao H. *Proc Natl Acad Sci USA*. 2008; 105:6858–6863. [PubMed: 18458342]
38. Winkler R, Zocher G, Richter I, Friedrich T, Schulz GE, Hertweck C. *Angew Chem Int Ed*. 2007; 46:8605–8608.
39. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. *J Appl Crystallogr*. 2007; 40:658–674. [PubMed: 19461840]
40. Emsley P, Cowtan K. *Acta Crystallogr*. 2004; E60:2126–2132.
41. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Longa F, Vagin AA. *Acta Cryst*. 2011; D67:355–367.
42. Holmes MA, Le Trong I, Turley S, Sieker LC, Stenkamp RE. *J Mol Biol*. 1991; 218:583–593. [PubMed: 2016748]
43. Makris TM, Knoot CJ, Wilmot CM, Lipscomb JD. *Biochemistry*. 2013; 52:6662–6671. [PubMed: 23980641]
44. Ookubo T, Sugimoto H, Nagayama T, Masuda H, Sato T, Tanaka K, Maeda Y, Okawa H, Hayashi Y, Uehara A, Suzuki M. *J Am Chem Soc*. 1996; 118:701–702.
45. Kim K, Lippard SJ. *J Am Chem Soc*. 1996; 118:4914–4915.
46. Zhang X, Furutachi H, Fujinami S, Nagatomo S, Maeda Y, Watanabe Y, Kitagawa T, Suzuki M. *J Am Chem Soc*. 2005; 127:826–827. [PubMed: 15656607]
47. Bailey LJ, Fox BG. *Biochemistry*. 2009; 48:8932–8939. [PubMed: 19705873]
48. Han Z, Sakai N, Boettger LH, Klinke S, Hauber J, Trautwein AX, Hilgenfeld R. *Structure*. 2015; 23:882–892. [PubMed: 25865244]
49. Fiedler AT, Shan X, Mehn MP, Kaizer J, Torelli S, Frisch JR, Kodera M, Que L Jr. *J Phys Chem A*. 2008; 112:13037–13044. [PubMed: 18811130]
50. Ray WJ, Puvathingal JM. *Anal Biochem*. 1985; 146:307–312. [PubMed: 4025798]

51. Xiong J, Kurtz DM Jr, Ai J, Sanders-Loehr J. *Biochemistry*. 2000; 39:5117–5125. [PubMed: 10819979]
52. Wei P-P, Skulan AJ, Miti N, Yang Y-S, Saleh L, Bollinger JM Jr, Solomon EI. *J Am Chem Soc*. 2004; 126:3777–3788. [PubMed: 15038731]
53. Banerjee R, Meier KK, Münck E, Lipscomb JD. *Biochemistry*. 2013; 52:4331–4342. [PubMed: 23718184]
54. Rosenzweig AC, Nordlund P, Takahara PM, Frederick CA, Lippard SJ. *Chem Biol*. 1995; 2:409–418.
55. Voegtli WC, Khidekel N, Baldwin J, Ley BA, Bollinger JM Jr, Rosenzweig AC. *J Am Chem Soc*. 2000; 122:3255–3261.
56. Tinberg CE, Lippard SJ. *Acc Chem Res*. 2011; 44:280–288. [PubMed: 21391602]
57. Kovaleva EG, Lipscomb JD. *Science*. 2007; 316:453–457. [PubMed: 17446402]
58. Kovaleva EG, Lipscomb JD. *Biochemistry*. 2008; 47:11168–11170. [PubMed: 18826259]
59. Knot CJ, Purpero VM, Lipscomb JD. *Proc Natl Acad Sci USA*. 2015; 112:388–393. [PubMed: 25548185]
60. Wallar BJ, Lipscomb JD. *Biochemistry*. 2001; 40:2220–2233. [PubMed: 11329291]
61. Miti N, Schwartz JK, Brazeau BJ, Lipscomb JD, Solomon EI. *Biochemistry*. 2008; 47:8386–8397. [PubMed: 18627173]
62. Yang Y-S, Broadwater JA, Pulver SC, Fox BG, Solomon EI. *J Am Chem Soc*. 1999; 121:2770–2783.
63. Schwartz JK, Wei P-P, Mitchell KH, Fox BG, Solomon EI. *J Am Chem Soc*. 2008; 130:7098–7109. [PubMed: 18479085]
64. Liu Y, Nesheim JC, Lee S-K, Lipscomb JD. *J Biol Chem*. 1995; 270:24662–24665. [PubMed: 7559577]
65. Logan DT, Su XD, Aberg A, Regnstrom K, Hajdu J, Eklund H, Nordlund P. *Structure*. 1996; 4:1053–1064. [PubMed: 8805591]
66. Elango N, Radhakrishnan R, Froland WA, Wallar BJ, Earhart CA, Lipscomb JD, Ohlendorf DH. *Protein Sci*. 1997; 6:556–568. [PubMed: 9070438]
67. Banerjee R, Proshlyakov Y, Lipscomb JD, Proshlyakov DA. *Nature*. 2015; 518:431–434. [PubMed: 25607364]
68. Lee SK, Lipscomb JD. *Biochemistry*. 1999; 38:4423–4432. [PubMed: 10194363]

Fig. 1.

Results of the biochemical characterization of CmlI 33 at pH 9.0 in 50 mM Bicine buffer. **a** Optical spectra of 200 μM CmlI 33 in the chemically reduced state (*green*), the diferric peroxo state **P** (*red*) and its decay upon titration with of $\text{NH}_2\text{-CAM}$ substrate (*purple traces*). The *blue trace* is the final spectrum of diferric CmlI after **P** is fully consumed. *Inset* optical spectra of resting diferric CmlI 33 at pH 9.0 (*blue*) and pH 6.8 (*magenta*). **b** Consumption of the CmlI 33 **P** when titrated with $\text{NH}_2\text{-CAM}$ based on the observed intensity at 500 nm. **c** HPLC analysis of the CmlI 33 reaction products. Extracted CmlI 33 reaction product from reaction of 200 μM **P** with 100 μM $\text{NH}_2\text{-CAM}$ (*black*), buffer-only negative control (*red*), 60 μM $\text{NH}_2\text{-CAM}$ substrate standard (*green*), 60 μM CAM standard (*blue*). The data for the CmlI 33 reaction product and CAM were multiplied by 0.5 in order to fit on the figure. **d** Stopped-flow transient kinetic analysis of the dependence of **P** decay on $\text{NH}_2\text{-CAM}$ concentration as monitored at 500 nm and 4.5 $^\circ\text{C}$. The *red line* shows a linear fit to the data

Fig. 2.

Overall structure of CmlI 33. **a** The CmlI 33 monomer in the crystal asymmetric unit with secondary structure shown. The helices are labeled and referred to throughout the text. The *dotted line* represents a disordered stretch of residues that was not modeled. **b** Two views of the CmlI 33 dimer. One monomer is in *gray* and the other in *blue*. In both *panels* the iron ions are shown as *brown spheres*

Fig. 3.

Comparison of the substrate channels in CmlI and Fe-AurF and the structural difference observed in helix $\alpha 5$. **a** Alignment of the structure of CmlI 33 (*gray*) and Fe-AurF (*purple*, PDB ID: 3CHH) showing the region of helix $\alpha 5$ which differs between the structures. In CmlI this region remains helical, but in Fe-AurF there is a break in the loop as some of the residues lose their secondary structure (*yellow star*). The side-chains of residue I208 (CmlI 33, *gray*) and I199 (Fe-AurF, *purple*) are shown as sticks and the atomic radii as transparent spheres. I199 points away from the channel and cluster in Fe-AurF. **b** The

collapsed substrate channel in CmlI 33. Important residues are shown as *sticks* and labeled. Channels in the interior of the protein are shown as a *grey* surface. **c** The open substrate channel in Fe-AurF with important residues shown as *sticks* and labeled. Iron ions are shown as *brown spheres*. **d** Alignment of the structure of CmlI 33 (*gray*) and the para-nitrobenzoate (pNB) product complex of Fe-AurF (*purple*, PDB ID: 3CHT). The pNB molecule carbon atoms are colored *yellow*

Fig. 4. Structure of the CmlI 33 diiron cluster in two states. **a** Peroxo-bound cluster structure with bond distances in Å. **b** Peroxo-bound cluster $2[F_o]-|F_c|$ electron density map contoured at 1.6σ . **c** Chemically reduced cluster structure with bond distances in Å. **d** Chemically reduced cluster $2[F_o]-|F_c|$ electron density map contoured at 1.3σ . Oxygen atoms are red and nitrogen atoms blue. Iron atoms are colored brown

Fig. 5. Evidence for the peroxo adduct in the CmlI 33 crystal structure. Two views of the $|F_o|-|F_c|$ ligand-omit difference map (*green mesh*) contoured at $+4\sigma$ after several refinement cycles of the model without the peroxo ligand. Coloring is as in Fig. 4

Fig. 6.

Optical spectra of as-isolated CmlI reacted with H₂O₂ in solution and of single crystals. **a** Reaction of WT CmlI with H₂O₂ in solution: resting CmlI (*blue*), the final spectrum of the enzyme after addition of 40 equivalents of H₂O₂ (*black*), chemically reduced enzyme (*green*), and intermediate **P** generated by exposing chemically reduced CmlI to O₂ (*red*). The H₂O₂ reactions were performed at pH 7.8 in 0.1 M Bicine buffer at 4.5 °C using 100 μM CmlI. *Inset* optical changes during the reaction of CmlI with H₂O₂ monitored over the course of an hour. **b** Optical spectra of single crystals of CmlI 33 at 100 K collected before

(*magenta*) and after (*purple*) exposure to synchrotron X-ray radiation. For comparison, a single crystal spectrum of a CmlA, a diiron enzyme that retains an oxo-bridge in the crystal form, is shown in orange. The spectrum of CmlA crystal was scaled by 0.75 to approximately account for differences in crystal sizes between CmlA and CmlI. The crystallization solution alone (*brown*) makes no contribution in the near-UV or visible range. *Inset* comparison of the optical spectra of as-isolated CmlI (*blue*, pH 9.0) and CmlA (*orange*, pH 7.5) in solution (100 μ M enzyme; 23 °C)

Fig. 7.

Overlay of the CmlI and AurF clusters. **a** Peroxo-bound CmlI 33 (*gray* carbon atoms and *black labels*) and Fe-AurF (*green* atoms and labels, PDB ID: 3CHH, 2.0 Å). The *green spheres* are the iron ions in the Fe-AurF structure and the *red sphere* is the μ -oxo ligand. A water ligand bound to AurF Fe1 has been omitted for clarity. **b** Peroxo-bound CmlI 33 (*gray*) and Mn-AurF (*purple*, PDB ID: 2JCD, 2.1 Å). The *purple spheres* are the manganese ions. Other coloring is the same as in Fig. 4. **c** Overlay of the peroxo-bound (*gray*) and reduced (*slate*) CmlI 33 clusters

Fig. 8. Overlay of the reduced CmlI 33 and MMOH diiron clusters. **a** Overlay of chemically reduced CmlI 33 and MMOH (PDB ID: 1FYZ). CmlI 33 and MMOH residues are colored *gray* and *blue*, respectively. Labels in *black* and *blue* are for the equivalent residues in CmlI 33 and sMMOH, respectively. In the sMMOH model, *blue spheres* are the iron ions and the smaller *blue sphere* is a water ligand bound to Fe1. E209 and E243 have been implicated as key residues that regulate the catalytic cycle of MMOH

Scheme 1.
Overall reaction catalyzed by the arylamine oxygenase CmlI

Scheme 2.

Hypothetical parallels in the formation of MMOH-Q and CmlI-P. The * marks the substrate-binding cavity and the oxygen that will be transferred to form product

Table 1

X-ray diffraction data collection and model statistics

	CmII 33 peroxo	CmII 33 chemically reduced
PDB ID	5HYG	5HYH
Data collection		
Space group	P3 ₂ 2 1	P3 ₂ 2 1
Cell dimensions		
a, b, c (Å)	56.54, 56.54, 150.47	56.96, 56.96, 151.10
α, β, γ (°)	90, 90, 120	90, 90, 120
Wavelength (Å)	0.97926	0.97933
Resolution (Å)	50–2.03	50–2.03
Total/unique reflections	84,870/15,532	131,570/19,182
<i>R</i> _{merge} (%) ^{a,b}	8.1 (40.8)	9.1 (37.4)
<i>I</i> /σ ^a	29 (6.9)	38 (5.4)
Completeness (%) ^a	82.8 (88.4)	99.7 (99.9)
Redundancy ^a	5.5 (5.1)	6.9 (6.9)
Model refinement		
Resolution (Å)	22.0–2.03	30.0–2.03
<i>R</i> _{free} test set size (%) ^c	8.0	7.8
<i>R</i> _{work} / <i>R</i> _{free} (%) ^c	18.5/25.1	18.9/23.9
ESU (Å) ^d	0.13	0.12
Residues modeled	285	287
Average B-factors		
Protein atoms (Å ²)	37	34
Overall (Å ²)	39	36
RMSDs		
Bond lengths (Å)	0.0089	0.0089
Bond angles (°)	1.23	1.23
Planes (Å)	0.006	0.006
Ramachandran analysis		
Favored (%)	99	96
Allowed (%)	1	4
Outlier (%)	0	0

RMSD root mean square deviation from ideal geometry

All data collected on synchrotron beamline APS SBC-CAT 19ID-D

^aHighest resolution shell is shown in parentheses^b $R_{\text{sym}} = \frac{\sum_i \sum_j |F_{\text{hkl}}| |F_{\text{hkl},i}| - \langle F_{\text{hkl}} \rangle \sum_i |F_{\text{hkl}}| |F_{\text{hkl},i}|}{\sum_i |F_{\text{hkl}}| |F_{\text{hkl},i}|}$, where F_{hkl} is the intensity of a reflection and $\langle F_{\text{hkl}} \rangle$ is the average of all observations of the reflection

$c_{R_{\text{free}}}$, R-factor calculated from a portion of data set excluded from refinement. The same reflection list was used for refinement of both models

d Estimated overall coordinate error (ESU) based on maximum likelihood

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Comparison of structurally characterized diferric *cis-μ*-1,2-peroxo complexes in proteins and inorganic model compounds

Protein/complex	Method	Fe1-O1	Fe2-O2	Fe1-Fe2	∠Fe1-O1-O2	∠Fe2-O2-O1	∠Fe1-O1-O2-Fe2	References
CmlI 33	Crystal	1.8	2.0	3.3	126	107	-38	This work
T4moH	Crystal	2.2	2.4	3.3	116	107	29	[47]
hDOHH ^a	Crystal	2.2 (2.2)	2.2 (2.2)	3.8 (3.7)	133 (151)	95 (89)	-58 (-32) 9.89	[48]
[Fe ₂ (Ph-bimp)(C ₆ H ₅ COO)(O ₂)] ²⁺	Crystal	1.94	1.86	3.33	118.1	121.4		[44]
[Fe ₂ (O ₂)(O ₂ CCH ₂ Ph) ₂ ·(HB[pz'] ₃) ₂]	Crystal	1.88	1.88	4.00	128.9	129.7	53.5	[45]
[Fe ₂ (6-Me ₂ -BPP) ₂ (OH)(O ₂)]	Crystal	1.87	1.89	3.40	123.1	120.4	-14.5	[46]
[Fe ₂ (O)(O ₂)(6-Me ₃ -TPA) ₂] ²⁺	XAS	1.78	1.78	3.17	118	118		[49]

All bond distances are given in Å and angles in degrees

^aMeasurements given for hDOHH chain A and chain B in parentheses