



NIH Public Access

Author Manuscript

Biochemistry. Author manuscript; available in PMC 2014 September 24.

Published in final edited form as:

Biochemistry. 2013 September 24; 52(38): . doi:10.1021/bi400845b.

Structure of a dinuclear iron cluster-containing beta hydroxylase active in antibiotic biosynthesis

Thomas M. Makris^{†,§}, Cory J. Knoot[†], Carrie M. Wilmot^{*}, and John D. Lipscomb^{*}

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

Abstract

A family of dinuclear iron cluster-containing oxygenases was recently described that catalyze -hydroxylation tailoring reactions in natural product biosynthesis by nonribosomal peptide synthetase (NRPS) systems (Makris, T. M., Chakrabarti, M., Münck, E., and Lipscomb, J. D. (2010) *Proc. Natl. Acad. Sci. U.S.A.* 107, 15391–15396). Here, the 2.17 Å X-ray crystal structure of the archetypal enzyme from the family, CmlA, is reported. CmlA catalyzes -hydroxylation of L-p-aminophenylalanine during chloramphenicol biosynthesis. The fold of the N-terminal domain of CmlA is unlike any previously reported, but the C-terminal domain has the -fold of the metallo- -lactamase (MBL) superfamily. The diiron cluster bound in the C-terminal domain is coordinated by an acetate, three His, two Asp, one Glu and a bridging oxo moiety. One of the Asp ligands forms an unusual monodentate bridge. No other oxygen-activating diiron enzyme utilizes this ligation or the MBL protein fold. The N-terminal domain facilitates dimerization, but using computational docking and a sequence-based structural comparison to homologs, we hypothesize that it likely serves additional roles in NRPS recognition and the regulation of O₂ activation.

Biosynthesis of many complex secondary metabolites by bacteria and fungi is effected by nonribosomal peptide synthetases (NRPS). These large, modular enzymes act as ‘assembly lines’ in which each module plays a role in extending or modifying the growing natural product.¹ Generally, a parent amino acid or adduct is covalently tethered to the 4-phosphopantetheine (PPant) moiety of a modified serine residue on a peptidyl-carrier protein (PCP) domain of the first module of the NRPS by the action of a preceding adenylation (A) domain. The resulting S-acyl amino acid intermediates are either chemically modified by following NRPS domains or transferred to the S-acyl-linked amino acid adduct on the PCP domain of the next module by the action of a condensation (C) (or similar) domain, forming a new peptide bond. When the synthesis is complete, the natural product is released by the action of a thioesterase (TE) or a reductase (R) domain.

*Corresponding authors: C.M.M.: Tel: (612) 624-2406; Fax, (612) 624-5121; Wilmo004@umn.edu. J.D.L.: Tel: (612) 625-6454; Fax, (612) 624-5121; Lipsc001@umn.edu.

[†]Author contributions

The contributions of these authors to the study should be considered to be equal.

[§]Present Address

Department of Chemistry and Biochemistry, University of South Carolina, 631 Sumter Street, Columbia, SC 29208

The authors declare no competing financial interests.

ACCESSION NUMBERS

The crystallography statistics, atomic coordinates, and structure factors coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 4JO0)

SUPPLEMENTAL INFORMATION

The supplemental information for this article includes one table and seven figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

In some cases, the chemical modifications of the developing natural product are catalyzed by “tailoring” enzymes that are not modules of the NRPS but are rather encoded elsewhere, often in the same genetic operon.¹ A common modification for tethered amino acids is -hydroxylation, which is crucial for the function of many natural products including the vancomycin and coumarin antibiotic families as well as the antitumor compounds of the bleomycin family.^{2,3} The newly introduced hydroxyl function can serve as a site for glycosylation or act as a nucleophile during release of the natural product from the NRPS to form large macrocyclic molecules. Study of the structural and regulatory mechanisms that control the specificity and spatiotemporal activity of these accessory enzymes is necessary in order to understand the biosynthetic origins of complex natural products. Learning how to harness these capabilities may enable the production of new antibiotics and pharmaceuticals.^{4,5}

Sequencing of the chloramphenicol biosynthetic operon⁶ led to the discovery of a tailoring enzyme (CmlA) that catalyzes -hydroxylation of the precursor molecule L-*p*-aminophenylalanine (L-PAPA) to form L-*p*-aminophenylserine.³ This is a key step in chloramphenicol synthesis as illustrated in Figure 1.⁷ The amino acid sequence of CmlA suggests that it has two domains: The N-terminal domain has no known homologs, but the C-terminal domain contains the signature sequence of the -type metallo-lactamase (MBL) superfamily.^{3,8} The vast majority of MBLs catalyze hydrolysis reactions utilizing a dizinc cluster coordinated by the signature His-X-His-X-Asp-His scaffold, although MBLs with both mono- and di-metal clusters containing other metals are known.⁹ In contrast, spectroscopic characterization of CmlA showed that it contains a diiron cluster.^{3,10}

Prior to the characterization of CmlA, no MBL-like enzymes catalyzing O₂ activation and hydroxylation reactions were known to exist, although a branch of the MBL family that utilizes a diiron cluster to reduce O₂ or NO in obligate anaerobes has been described.^{9,11} Typically, -hydroxylation reactions in NRPS pathways are catalyzed by -ketoglutarate-dependent oxygenases or cytochrome P450s.² CmlA represents a third class, and it is the only known diiron enzyme to catalyze -hydroxylation of an amino acid-like substrate. Oxygen activation and hydroxylation reactions involving diiron clusters are generally carried out by members of the bacterial multicomponent monooxygenase (BMM) family,¹² which includes enzymes such as the soluble methane monooxygenase (sMMO).¹³ In contrast to CmlA, BMMs coordinate the diiron cluster in a 4-helix bundle. The typical cluster coordination consists of four carboxylate (Asp/Glu) and two histidine ligands.¹⁴ Numerous kinetic, spectroscopic, and structural studies have shown that both the diiron cluster and the protein scaffold play important roles in orchestrating and regulating the oxygenation chemistry.^{13,15,16}

Genome database searching based on the CmlA gene sequence has now shown that there are many enzymes serving similar roles in natural product biosynthesis that are homologous to both the N- and C-terminal regions of CmlA.³ To date, no enzyme from this group of homologous enzymes has been structurally characterized, and no role for the N-terminal domain has been forthcoming.

Here, we present the crystal structure of CmlA at 2.17 Å resolution. The structure confirms the predicted MBL C-terminal fold and reveals the structure of the N-terminal domain. The detailed structure of the dinuclear iron cluster differs from those of previously characterized oxidoreductases from the MBL family, as well as those from the O₂-activating BMM family. The overall organization of the enzyme suggests different roles for the N- and C-terminal regions as well as the basis for regulation of O₂ activation through interaction with the NRPS from the chloramphenicol biosynthetic pathway, CmlP.

EXPERIMENTAL PROCEDURES

Crystallization of CmlA

CmlA and CmlP were expressed, purified and quantified as previously described.³ Crystals of resting CmlA were acquired by the hanging drop vapor diffusion method at 277 K by addition of 3 μ l of a 15 mg/ml solution of CmlA in 50 mM HEPES pH 7.5 to an equal volume of mother liquor containing 100 mM HEPES pH 7.5, 10–15 % polyethylene glycol (F.W. 20,000), 100 mM potassium acetate, and 10 % glycerol. Square-bipyramidal CmlA crystals typically grew to 0.2 – 0.6 mm over the course of several days. Selenomethionine-labeled CmlA was generated by feedback inhibition during expression.¹⁷ Crystals were soaked in mother liquor supplemented with 25 % glycerol prior to looping and flash-freezing in liquid nitrogen. For the amino acid co-crystallization, crystallization and freezing conditions were identical to the non-co-crystallized samples except the mother liquor and cryoprotectant contained L-phenylalanine, L-tyrosine or L-PAPA at final concentrations of 10–20 mM.

Data Collection and Modeling

Data were collected at the Structural Biology Center at the Advanced Photon Source, Argonne National Labs on beamline 19-ID at 100 K. Diffraction data were indexed, integrated and scaled using the HKL2000 software package¹⁸ and phased using single-wavelength anomalous dispersion of SeMet labeled enzyme with Phenix.¹⁹ The initial model was prepared with Phenix followed by manual rebuilding. The preliminary model was then used in direct Fourier phasing of a higher-resolution native data set using Refmac5²⁰ from the CCP4 suite.²¹ The final model was prepared using iterative cycles of modeling using Coot²² and refinement with Refmac5. Iron-ligand bond distances were not restrained during any of the refinement cycles. Model geometries were evaluated using SFCHECK and PROCHECK.²¹ X-ray data processing and refinement statistics are summarized in Table 1. All structure figures were produced using PyMOL Molecular Graphics System, Version 1.5, Schrödinger, LLC.

Docking of the CmlP PCP Domain and PPant Amino Acid Substrate

Homology modeling of the CmlP PCP domain was performed using Swiss-Model²³ and docked to CmlA using the HADDOCK online server.²⁴ In order to guide the docking of the PCP domain, we defined several residues to be part of the interface based on studies of our CmlA model and prior sequence information. For CmlA we input W438, D496, and H501. These surface residues were chosen based on their proximity to the active site channel. For the CmlP T domain homology model, we input only the serine identified as the putative PPant attachment from sequence comparisons to other structurally characterized PCP domains, with the rationale that this residue must be poised above the channel upon binding of the NRPS to CmlA.

Docking of L-PAPA-PPant to CmlA was performed using the Sybyl-X 2.0 Surflex-Dock Suite [Sybyl-X 2.0 Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA]. The final CmlA model was prepared by deleting all water molecules and non-protein ligands, except the Fe atoms, followed by addition of hydrogen to residues. The structure was then minimized using the Amber7 FF99 package in the docking suite, but Fe atoms and ligand positions were unaltered. The substrate was modeled manually in Sybyl-X and atomic partial charges for both the protein and substrate were automatically calculated and assigned, with the exception of Fe atoms which were manually assigned as +3 and +2 for Fe1 and Fe2, respectively. The parameters for defining the active site were as follows: threshold 0.65, bloat 0 and the active site was automatically identified by the program. The protein was not allowed to flex during docking.

Cloning and Mutagenesis

The expression construct for the E430A mutant of CmlA was produced using an Agilent Quickchange II Site Directed Mutagenesis kit. The following primer pair was used to introduce the mutation into a pet28a expression plasmid for wild-type (WT) CmlA: CGTCTTCATCGGCATGGAGCTCATGGCGCGG (forward) and CCGGCCGATGAGCTCCATGCCATGAAGACG (reverse). Mutated plasmid was transformed into chemically competent XL1-Blue *E. coli* and plated onto media containing 50 µg/ml kanamycin. Individual colonies were screened *via* colony PCR and sequenced to confirm the presence of the mutation. E430A enzyme was expressed and purified as described above. Mass spectrometry of CmlA was done with a QSTAR XL Quadrupole TOF MS from AB Sciex at the University of Minnesota Center for Mass Spectrometry and Proteomics and used to confirm the presence of the desired mutation in purified enzyme. The deconvoluted mass spectrum showed a major peak at 62,204 Da and 62,146 Da for the WT and E430A, respectively, the difference corresponding to the expected loss of a carboxylate and CH₂ group in the mutant.

Stopped Flow Kinetics

All stopped-flow experiments were performed with an Applied Photophysics model SX. 18MV stopped flow device at 4 °C. Stopped-flow experiments and CmlP amino acid loading were performed as previously described,³ with the exception that reduced CmlA was pre-mixed with L-PAPA loaded CmlP_{AT} in an anaerobic chamber and rapidly mixed with O₂ saturated buffer on the instrument. Excess CmlA reducing reagents and CmlP loading reagents were removed prior to the experiment using a PD10 desalting column.

RESULTS AND DISCUSSION

Overall Structure of CmlA

The structure of CmlA was determined by single wavelength anomalous dispersion of selenomethionine-enriched enzyme. The enzyme crystallized in spacegroup P4₃2₁2 with a single monomer in the asymmetric unit. The CmlA model was refined to 2.17 Å resolution with an R_{work} of 0.203 and R_{free} of 0.230 with 99.6% of residues in the allowed regions of the Ramachandran plot (Table 1). Overall, 522 residues and 151 solvent molecules were modeled into the electron density map. The CmlA monomer (~ 62 kD) is globular with several protrusions that extend from the N-terminal domain (Figures 2 and S1). This domain (residues 1–236, green in Figure 2A) has a mixed α -topology, whilst the C-terminal domain (residues 249–532, grey with secondary structure indicated in Figure 2A) has the topology of the MBL superfamily.⁸ The metal cluster and the active site of the enzyme are located in the C-terminal domain, thereby identifying it as the catalytic domain. A stretch of disordered residues separates the domains (residues 237–248, dashed lines Figure 2). Residues from both domains form the boundaries of a 30 Å long groove that runs along the surface and above the active site channel. These include a kinked β -helical region (termed Helix N1) composed of residues 26–67 in the N-terminal domain (magenta in Figure 2A) and residues 429–460 (cyan in Figure 2A) and 493–505 from the catalytic domain. A channel located near one end of the groove extends 10 Å from the surface to the active site diiron cluster (Figure 2A, right side). The cluster appears to be easily accessible. This is radically different than observed in the BMMs, where either the active site is buried deep within the enzyme^{25,26} or solvent access is severely restricted.^{14,27}

The N-terminal domain is a mixture of β -helices and three sets of antiparallel β -strands. Structural alignments of this domain made using the Dali Server²⁸ found no significant matches, with the highest scoring hit having a calculated RMSD of 6.6 Å for the C atoms. This domain is elongated, measuring roughly 70 Å by 34 Å at its longest and widest points,

respectively. It packs against the catalytic domain through multiple β -helices and antiparallel β -strands which run parallel to the long axis. A 50 Å long, bent amphipathic helix termed Helix N2 (Figure 2) formed by residues 73–107 spans nearly half the length of the domain and perpendicularly packs against Helix N1 (27–47). Helix N1 forms an L-shaped arm (“Arm 1” in Figure 2A and 2B) which projects over the catalytic domain forming one side of the groove that runs between the domains.

The overall structure of the catalytic domain is similar to that of other MBLs, having a core of two antiparallel β -sheets arranged in a β -sandwich flanked by several α -helices. On the side that is packed against the N-terminal domain of CmlA, the α -helices are notably truncated relative to other members of the MBL superfamily, where these α -helices are typically solvent exposed. Based on a Dali structural search, the catalytic domain of CmlA is structurally most similar to the MBL L-ascorbate-6-phosphate lactonase UlaG²⁹ (15% sequence identity, with an RMSD of 2.9 Å for the C atoms). An elongated arm (“Arm 2” in Figure 2A and 2B) is formed by residues 429–460 and extends upward from the β -core (cyan in Figure 2A). Similar extensions have been observed in other members of the MBL superfamily and are relevant to binding⁸ and recognition³⁰ of substrate in these enzymes. However, the arm in CmlA is longer and has a distinctive orientation. The main portion of Arm 2 forms one side of the surface groove and active site pocket and, on the opposite face, interacts with another CmlA monomer in the crystal.

Analysis of the crystal packing identified a buried interface of roughly 2700 Å² between two CmlA monomers in the unit cell, suggesting the enzyme may function as a dimer in solution. In order to investigate this possibility, we performed Native-PAGE, which confirmed the dimeric quaternary structure observed in the crystal (Figure S2). A projecting “dimerization arm” (residues 108–146) from the N-terminal domain of CmlA mediates the interaction between the monomers (Figure 3) by packing against the dimerization arm and Arm 2 from the symmetry-related protomer. This arrangement places one dimerization arm within 17 Å of the active site of the other protomer.

Structure of the Diiron Cluster

The active site of CmlA is situated above the β -sheets of the β -core of the catalytic domain. Structural alignments of this domain show that the active sites of other members of the MBL-superfamily are found in similar spatial positions relative to the rest of the fold. The iron atoms are ligated in part by four ligands derived from a loop containing the conserved metal binding motif, which extends upward between the fourth β -strand and second α -helix of the catalytic domain (Figures 4A and S1). The only apparent access to the active site is from the channel above, although some solvent-occupied caverns can be seen extending from the substrate binding pocket.

The two iron atoms in the enzyme active site are in distorted octahedral environments with six ligands each (Figures 4A and S3). Herein, iron atoms are referred to as Fe1 and Fe2 for Fe600 and Fe601 in the model, respectively. Iron atom occupancies refined to 1.0 for Fe1 and 0.75 for Fe2 (see Supplemental Information). The iron–iron distance was determined to be 3.39 Å, in close agreement with EXAFS studies.¹⁰ Strong density for a bridging atom is also apparent in the difference map following refinement of the cluster (Figure 4B) and, based on previous spectroscopic studies of the diferric state of CmlA,^{3,10} we have modeled this density with an oxo atom (O602). The iron–oxo bond distances refined to 1.77 Å and 2.15 Å for Fe1 and Fe2, respectively. EXAFS studies¹⁰ suggested the average iron–oxo distance is ~1.8 Å, raising the possibility that Fe2 may have been reduced to the ferrous state during data collection, forming a mixed-valent (MV) Fe(III)Fe(II) cluster. Consistent with this, the overall average ligand distances refined to 2.0 Å for Fe1 and 2.3 Å for Fe2. We investigated the possibility of reduction during data collection by several standard

approaches (Supporting Information). However, none of these methods provided definitive evidence for a mixed-valence (MV) cluster.

The cluster in CmlA is coordinated by one less endogenous amino acid carboxylate and one more histidine ligand than found for most other diiron hydroxylases. All nitrogen ligands derive from the conserved MBL metal-binding scaffold. The N-terminal domain contributes no cluster ligands. Fe2 is coordinated by D309, H310, a monatomic bridging D403, an acetate molecule (ACT603) derived from the crystallization buffer, and the bridging O atom. Fe1 is coordinated by a chelating E377, bridging D403, H305, H307 and the bridging O atom. It is likely that in solution the coordinating acetate is replaced by solvent because no acetate is added to the purification buffer. Exposure of CmlA to 100 mM potassium acetate in buffered solution results in no detectable changes in the chromophore.

Monodentate Diiron Cluster Bridging Structure

The monodentate μ -² binding mode of D403 in CmlA was unexpected, because in most oxygen activating diiron enzymes and model complexes the carboxylate forms a μ -1,3 bridge. The monodentate bridging mode has been observed in purple acid phosphatase, the reduced form of the D84E mutant of *E. coli* ribonucleotide reductase R2 subunit, and reduced sMMO hydroxylase component.^{31–35} However, in the latter two cases, the bridging mode is μ -(¹, ²) where the second oxygen of the carboxylate coordinates to one of the irons. Also, in both cases, a traditional μ -1,3 bridging carboxylate is also present. It is unlikely that D403 in CmlA can assume a μ -1,3 orientation without significant rearrangements of the backbone due to several constraints: (i) D403 is sterically restrained and essentially limited to rotation around the C-C bond, (ii) D403 needs to rotate towards Fe1 to bring the second oxygen close enough to coordinate Fe2; not only would this introduce severe strain on the sp^3 hybridization of the D403 C, but it would result in steric clashes with E377 and a too-short iron-oxygen bond to Fe1. Finally, (iii) if the backbone rearranges, it would be required to undergo a nearly 90° rotation to bring the carboxylate into an allowed conformation close enough to coordinate both metals. Indeed, rubredoxin:oxygen oxidoreductase has a bidentate Asp bridge and the backbone runs perpendicular to the iron-iron axis, whereas in CmlA and purple acid phosphatase the backbone runs more parallel. Thus, the bridge in CmlA is essentially “locked” into a single orientation during the catalytic cycle in the absence of significant structural rearrangement.

Structure of the Substrate Binding Pocket

The substrate binding pocket is elongated, roughly 8 Å by 13 Å (Figure 5). The pocket is bordered on one side by L458 and I439, on the opposite side by M493 and V492 and on top by the partially disordered surface-exposed W438, which sits at the base of the surface groove between domains. Q308 is positioned 8 Å from the cluster at the opposite end of the active site pocket, as measured from the Q308 carboxamide to the O602 in the cluster (“conformation 1”). A second conformation (“conformation 2”) of Q308 observed in the electron density map positions the side chain further from the cluster, expanding the long dimension of the binding pocket (Fig. S4). Y440 is at the back of the pocket and appears to engage in a hydrogen bond with the -nitrogen of Fe1 ligand H307 (Figure 4A).

In all models, additional areas of residual positive F_o-F_c density were observed in the part of the pocket above the cluster. We attempted to model this density with solvent molecules, but this resulted in significant residual positive density after refinement. Co-crystals of CmlA with 10–20 mM amino acids L-tyrosine, L-phenylalanine or L-PAPA resulted in isomorphous crystals, but insufficient density to model any amino acid substrates in the binding pocket. This density appeared nearly identical to that observed in crystals that were not co-crystallized with amino acids (Figure S5) and does not correspond well with

components of the crystal mother liquor, thus its source remains unknown. The absence of significant electron density for any amino acids in these co-crystallized samples demonstrates that free amino acids have little affinity for binding in the CmlA active site. This was also observed in a cyt. P450 tailoring system where no amino acid binding was detected even with a 500-fold excess over protein. In contrast, the NRPS-tethered substrate had a binding affinity in the low micromolar range.³⁶ The possible binding mode of the NRPS-bound substrate in CmlA is discussed and modeled in context below.

The monooxygenase chemistry catalyzed by CmlA requires delivery of two protons and two electrons to the reactive center during turnover, leading to release of a water molecule. We searched for possible endogenous proton donors in CmlA. One candidate is H258 which is positioned only 3 Å from Fe2 in the same plane as the cluster (Figure 4A). Ramachandran analysis shows H258 to be in a strained orientation, likely because the residue engages in a hydrogen bond with D309. On the opposite side, another His residue H378 is positioned 5 Å from Fe1 but points away from the cluster. Alternatively, the surface-exposed E430 could act as a proton shuttle from bulk solvent.

Conserved Structural Features of CmlA and Homologs

The structure of CmlA permits the sequence-based structural comparison of family members. In order to identify residues that may be functionally relevant, we generated an alignment of CmlA with 13 homologs from diverse species, having between 32–41 % sequence identity (Table S1). In the N-terminal domain, the conserved residues appear to primarily serve structural roles by engaging in interactions with nearby residues or facilitating turns on the protruding arms (Figure 6, G110). A cluster of conserved residues also appears to stabilize the dimerization arm in the N-terminal domain (Figure S6). S44 and H52 engage in a hydrogen bonding interaction that maintains a kink in Helix N1. The similar length and high sequence identity of the homologs (Table S1), along with the nature of the noted conservations, indicates that the overall shape of the N-terminal domain is conserved. However, the intervening regions are highly variable in their chemical character, including in Helix N2 (residues 73–107) and in the tips of Arms 1 and 2 (Figure S6). In the active site of the catalytic domain, all of the iron ligands, W441, Y443 and the potential proton donors H258, H378 and E430, are conserved in all homologs. None of the residues that line the binding pocket (I439, L458, V492, M493, Q308) are conserved.

Putative electron delivery routes have been identified in other diiron enzymes,³⁷ and we attempted to find such a route in CmlA, as well as identify a possible ferredoxin (Fd) or flavo-iron-sulfur reductase interaction site. From the sequence alignments, we noted a partially conserved buried aromatic network starting at the solvent-exposed Y479 that runs to underneath the cluster through the interior of the catalytic domain terminating at Y255, which is hydrogen bonded to the carbonyl group of H258 (Figure 7A). In more distant homologs, some positions are substituted by other aromatic residues. CmlA surface electrostatics reveal a strongly acidic patch (residues D423, E474, D476 and E477) at the base of the catalytic domain (Figure 7B) that might form a binding site for an electron transfer partner. Y479 is next to this patch and engages in a hydrogen bond with the carboxylate of E477. The acidic nature of this region would be more compatible with direct electron transfer from a flavo-iron-sulfur reductase³⁸ than a Fd,³⁹ which is likely to carry a strong negative charge. Unfortunately, a specific reductase protein for CmlA has not been identified, and the Cml operon appears not to encode such a protein.⁶ Thus, a definitive test of this hypothesis through mutagenesis is not currently possible.

Interaction of CmlA with the NRPS

CmlA hydroxylates only L-PAPA tethered to the carrier PCP domain of CmlP via a thioester bond to a PPant linker, minimally necessitating a specific interaction between the holo-PCP domain and CmlA. In order to identify the binding site of the PCP domain of CmlP and elucidate possible features conferring NRPS specificity, we docked a homology model of the highly conserved CmlP PCP domain^{40,41} to the crystal structure of CmlA (Figure 8 and blue residues in Figure S6). In this model, the carrier protein binds directly above the channel to the active site, as observed in similar systems⁴² and the serine residue where the PPant cofactor attaches is poised above the channel to the active site. The PCP domain interacts primarily with Arm 2 that extends from the catalytic domain of CmlA and the residues around the channel to the active site. The binding site is minimally conserved (Figure S5) consistent with the high specificity of the hydroxylation reaction.

Sequence-based structural comparisons of CmlA and homologs show that the surface composition of the N-terminal domain is highly variable. Nevertheless, there are 12 conserved residues, and these appear to be required to maintain the overall folded structure of each of the protruding arms on CmlA. This suggests that each arm is functionally relevant, but the small size of the CmlP carrier domain would limit its interaction to a portion of the arms identified in the docked model (Figure S6). Analysis of CmlA surface electrostatics reveal that the tips of the arms, which form the groove between domains, have a net positive charge relative to the rest of the surface (Figure 2C), suggesting that it may be used as a mechanism to electrostatically target the correct NRPS. One interpretation is that NRPS specificity of the *trans*-acting tailoring enzymes from the CmlA family is conferred by interactions with the entire NRPS module and not just the CmlP carrier domain. Resolution of this possibility awaits further study.

NRPS-Loaded Substrate Binding in the Active Site of CmlA

In the P450_{biol}-ACP crystal complex, PPant-tethered substrate is funneled into a long U-shaped cavity that positions the hydroxylation site above the heme for oxidation.⁴² We propose that a similar mechanism functions in CmlA, and that substrate positioning is primarily directed by the shape of the pocket and the length restraints of the linker, resulting in the positioning of the α -carbon of L-PAPA above the cluster. The shape of the pocket suggests an orientation where the substrate-tethered PPant linker bends roughly 45° over the cluster, towards Q308 and between the side chains of M493 and I439 (Figure 5). The length of the linker requires that the substrate be pushed farther down towards Q308, with the aromatic portion extending beyond the cluster. The Q308 side-chain may then adopt conformation 2 (Figure S4) farther back in the pocket, accommodating the substrate and stabilizing it in a specific orientation through hydrogen bonding to the *p*-NH₂ group of L-PAPA. In CmlA homologs, Q308 is substituted with other residues capable of H-bonding to substrate. In Tyr-accepting homologs (from *A. teichomyceticus*, *Nonomurea sp.*, and *S. toyocaensis*) involved in synthesis of teicoplanins and other glycopeptides, Q308 is substituted with Ser. In a putative His-accepting homolog from *S. verticillus* involved in bleomycin synthesis, Q308 is replaced by His (Table S1). The α -carbon of L-PAPA is the site of hydroxylation, and must be positioned near the cluster after binding of substrate. In our docked CmlA-PCP domain model, the distance between the PPant attachment site and the cluster is 13 Å, which matches well with observations from other systems.^{42,43} Upon binding, this would position the L-PAPA amide somewhere above the cluster and α -carbon slightly beyond it, near the site occupied by the oxo bridge in the resting state structure.

In order to test this possibility, we docked a model of the PPant-L-PAPA substrate to our structure of CmlA (Figure 8B). We separately docked the substrate to CmlA with Q308 in each conformation (Figure S4), but only conformation 2 shown in Figure 8B gave a result

consistent with the observed chemistry. In this model, the aromatic group of L-PAPA is positioned in the pocket identified in Figure 5, and *p*-NH₂ group interacts with nearby carbonyl groups from V492 and the side-chain of Q308. The *α*-carbon is positioned above the cluster, with the correct pro-chiral hydrogen placed within 3.4 Å of both Fe atoms. These results support our proposed binding orientation and are consistent with a role for Q308 in orienting substrate in the pocket.

Regulation of Oxygen Activation

Intermediates formed during O₂ activation are potentially dangerous oxidants to living systems.^{44,45} BMMs have evolved sophisticated regulatory mechanisms to only generate these oxidizing species in the presence of substrate and/or regulatory protein binding partners in order to mitigate undesired side-reactions.^{13,25,27,46} The triggering of O₂ activation by CmlA only upon binding to substrate-loaded CmlP is a salient example, and it implies a regulatory role for the NRPS.³ The mechanism of regulation of O₂ activation in the BMM family often involves structural changes. For example, the binding of the regulatory protein MMOB to the hydroxylase component of sMMO causes structural changes near the diiron cluster,^{27,47} and similar perturbations may also occur in CmlA. Comparison of the predicted and observed cluster ligand structures of CmlA suggests a possible regulatory mechanism.

In our earlier study, the cluster ligands reported here were correctly predicted based on sequence comparisons to other members of the MBL superfamily and alignments to CmlA homologs.³ However, E430 was also identified as a likely fourth carboxylate ligand to the diiron cluster. Analysis of the electron density maps shows that this residue is positioned nearby on a loop (430–437) which runs along the base of the groove between the domains above the acetate binding site of Fe2. It does not coordinate the cluster and the side chain of E430 appears disordered. One possible rotameric conformation of E430 places the carboxylate group within ~3 Å of Fe2 (green in Figure 4C), an orientation that partially overlaps with the acetate ligand (and the solvent(s) that presumably replaces the acetate when the enzyme is in solution). The residue is far enough away that it cannot coordinate to the cluster without a concomitant shift in the backbone. Coordination of E430 to Fe2 would build the carboxylate rich diiron cluster that differentiates O₂-carrier proteins from O₂-activating enzymes.^{3,10} Consequently, a conformational change that allows E430 to coordinate might serve as the switch to promote O₂ activation. E430 is located on a loop directly below the tentative binding site of the CmlP PCP domain identified in the docked structure (blue residue in Figure 8), so that the docking of CmlP may shift E430 close enough to coordinate to Fe2. Conservation of the nearby G433 residue could indicate it is required in order to maintain the backbone flexibility required for this change.

An initial examination of the hypothesized role of E430 was made using the E430A variant. A single turnover of fully reduced E430A in complex with L-PAPA-loaded CmlP_{AT} resulted in a 25-fold decrease in the maximal rate constant for reaction with O₂ as shown in Figure 9. In contrast, the autooxidation rate in the absence of L-PAPA-loaded CmlP was unchanged from that observed for wild type CmlA. Total iron loading the active site of the E430A variant was comparable to that of wild type CmlA, but exhibited a blue-shifted optical absorption spectrum (Figure S7). This demonstrates that E430 interacts strongly with the diiron cluster despite its remote position from Fe2. As expected, the optical spectra of diferrrous CmlA and E430, with or without L-PAPA loaded CmlP_{AT} bound, were bleached and featureless in the visible.³ These results support the hypothesis that E430 has a role in establishing a diiron cluster structure that is capable of rapidly activating O₂. However, the fact that E430A still accelerates the rate of O₂ reaction over autooxidation (Figure 9) suggests that the complex with the loaded NRPS has additional structural and/or electronic

effects on this reaction. Initial attempts to crystallize E430A were successful, but the resulting crystal forms did not diffract to high resolution.

CONCLUSION

Tailoring of natural products during biosynthesis in NRPS-based systems is highly specific. The X-ray crystal structure of the archetypal diiron cluster-containing -hydroxylase from an NRPS system described here provides insight into the basis for this specificity. It is likely to involve interactions directly with the substrate, L-PAPA, in the CmlA active site and also with the NRPS to which L-PAPA is covalently attached. The two structural domains identified in CmlA may serve different roles in establishing specificity. The structure demonstrates directly that the C-terminal domain houses the active site in which the diiron cluster responsible for oxygen activation is bound. The full role of the N-terminal domain with its previously uncharacterized structural motif remains open to speculation. While the structure shows that this domain facilitates dimerization in this family of enzymes, its proximity to the channel into the active site and the structural conservation of CmlA homologs also imply a role in establishing the NRPS complex. Finally, the NRPS may also act as a regulator of O₂ activation, delivering substrate and engendering conformational changes in CmlA which ensure that the oxidative chemistry only occurs in the presence of the tethered amino acid.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This work was supported by NIH grants GM40466 (to J.D.L.), GM40466-S1 (to J.D.L.), GM100943 (to J.D.L.), GM66569 (to C.M.W.) and graduate traineeship GM08700 (to C.J.K.).

We thank Drs. Lawrence Que, Jr. and Erik Yukl for valuable discussions. Diffraction data were collected at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We are grateful for resources from the Supercomputing Institute and the Kahlert Center for Structural Biology at the University of Minnesota.

ABBREVIATIONS USED

| | |
|---------------|--|
| NRPS | nonribosomal peptide synthetase |
| CmlA | -hydroxylase of the chloramphenicol biosynthetic pathway |
| CmlP | NRPS of the chloramphenicol biosynthetic pathway |
| sMMO | soluble form of methane monooxygenase |
| MMOB | regulatory component B of the sMMO system |
| L-PAPA | L-p-aminophenylalanine |
| PPant | 4-phosphopantetheine cofactor |
| PCP | peptidyl-carrier protein. This CmlP domain contains the serine attachment site for PPant and the amino-acid substrates that are subsequently covalently linked to PPant. PCP is alternatively termed the thiolation (T) domain |

| | |
|--------------------|---|
| CmlP _{AT} | truncated version of full-length CmlP lacking the C-terminal reductase domain of the NRPS |
|--------------------|---|

References

1. Fischbach MA, Walsh CT. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem Rev.* 2006; 106:3468–3496. [PubMed: 16895337]
2. Chen H, Thomas MG, O'Connor SE, Hubbard BK, Burkart MD, Walsh CT. Aminoacyl-S-enzyme intermediates in -hydroxylations and -desaturations of amino acids in heptide antibiotics. *Biochemistry.* 2001; 40:11651–11659. [PubMed: 11570865]
3. Makris TM, Chakrabarti M, Münck E, Lipscomb JD. A family of diiron monooxygenases catalyzing amino acid beta-hydroxylation in antibiotic biosynthesis. *Proc Natl Acad Sci USA.* 2010; 107:15391–15396. [PubMed: 20713732]
4. Nguyen KT, Ritz D, Gu JQ, Alexander D, Chu M, Miao V, Brian P, Baltz RH. Combinatorial biosynthesis of novel antibiotics related to daptomycin. *Proc Natl Acad Sci USA.* 2006; 103:17462–17467. [PubMed: 17090667]
5. Zhang C, Griffith BR, Fu Q, Albermann C, Fu X, Lee IK, Li L, Thorson JS. Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. *Science.* 2006; 313:1291–1294. [PubMed: 16946071]
6. He J, Magarvey N, Pirae M, Vining LC. The gene cluster for chloramphenicol biosynthesis in *Streptomyces venezuelae* ISP5230 includes novel shikimate pathway homologues and a monomodular non-ribosomal peptide synthetase gene. *Microbiology.* 2001; 147:2817–2829. [PubMed: 11577160]
7. Doull J, Ahmed Z, Stuttard C, Vining LC. Isolation and characterization of *Streptomyces venezuelae* mutants blocked in chloramphenicol biosynthesis. *J Gen Microbiol.* 1985; 131:97–104. [PubMed: 3989509]
8. Bebrone C. Metallo-beta-lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily. *Biochem Pharmacol.* 2007; 74:1686–1701. [PubMed: 17597585]
9. Frazão C, Silva G, Gomes CM, Matias P, Coelho R, Sieker L, Macedo S, Liu MY, Oliveira S, Teixeira M, Xavier AV, Rodrigues-Pousada C, Carrondo MA, Le Gall J. Structure of a dioxygen reduction enzyme from *Desulfovibrio gigas*. *Nat Struct Mol Biol.* 2000; 7:1041–1045.
10. Vu VV, Makris TM, Lipscomb JD, Que L. Active-site structure of a beta -hydroxylase in antibiotic biosynthesis. *J Am Chem Soc.* 2011; 133:6938–6941. [PubMed: 21506543]
11. Silaghi-Dumitrescu R, Kurtz DM Jr, Ljungdahl LG, Lanzilotta WN. X-ray crystal structures of *Moorella thermoacetica* FprA. Novel diiron site structure and mechanistic insights into a scavenging nitric oxide reductase. *Biochemistry.* 2005; 44:6492–6501. [PubMed: 15850383]
12. Notomista E, Lahm A, Di Donato A, Tramontano A. Evolution of bacterial and archaeal multicomponent monooxygenases. *J Mol Evol.* 2003; 56:435–445. [PubMed: 12664163]
13. Wallar BJ, Lipscomb JD. Dioxygen activation by enzymes containing binuclear non-heme iron clusters. *Chem Rev.* 1996; 96:2625–2657. [PubMed: 11848839]
14. Rosenzweig AC, Frederick CA, Lippard SJ, Nordlund P. Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane. *Nature.* 1993; 366:537–543. [PubMed: 8255292]
15. Kovaleva EG, Neibergall MB, Chakrabarty S, Lipscomb JD. Finding intermediates in the O₂ activation pathways of non-heme iron oxygenases. *Acc Chem Res.* 2007; 40:475–483. [PubMed: 17567087]
16. Tinberg CE, Lippard SJ. Dioxygen activation in soluble methane monooxygenase. *Acc Chem Res.* 2011; 44:280–288. [PubMed: 21391602]
17. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol.* 1993; 229:105–124. [PubMed: 7678431]

18. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 1997; 276:307–326.
19. Adams PD, Afonine PV, Bunkóczki G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr.* 2010; D66:213–221.
20. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.* 1997; D53:240–255.
21. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR. Overview of the CCP4 suite and current developments. *Acta Crystallogr D.* 2011; 67:235–242. [PubMed: 21460441]
22. Emsley P, Cowtan K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr, Sect D: Biol Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
23. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* 2009; 37:D387–D392. [PubMed: 18931379]
24. de Vries SJ, van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols.* 2010; 5:883–897.
25. Bailey LJ, McCoy JG, Phillips GN Jr, Fox BG. Structural consequences of effector protein complex formation in a diiron hydroxylase. *Proc Natl Acad Sci USA.* 2008; 105:19194–19198. [PubMed: 19033467]
26. Sazinsky MH, Bard J, Di Donato A, Lippard SJ. Crystal structure of the toluene/o-xylene monooxygenase hydroxylase from *Pseudomonas stutzeri* OX1. Insight into the substrate specificity, substrate channeling, and active site tuning of multicomponent monooxygenases. *J Biol Chem.* 2004; 279:30600–30610. [PubMed: 15096510]
27. Lee SJ, McCormick MS, Lippard SJ, Cho US. Control of substrate access to the active site in methane monooxygenase. *Nature.* 2013; 494:380–384. [PubMed: 23395959]
28. Holm L, Rosenström P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 2010; 38:W545–W549. [PubMed: 20457744]
29. Garces F, Fernández FJ, Montellà C, Penya-Soler E, Prohens R, Aguilar J, Baldomà L, Coll M, Badia J, Vega MC. Molecular architecture of the Mn²⁺-dependent lactonase UlaG reveals an RNase-like metallo- -lactamase fold and a novel quaternary structure. *J Mol Biol.* 2010; 398:715–729. [PubMed: 20359483]
30. de la Sierra-Gallay IL, Pellegrini O, Condon C. Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z. *Nature.* 2005; 433:657–661. [PubMed: 15654328]
31. Strater N, Klabunde T, Tucker P, Witzel H, Krebs B. Crystal structure of a purple acid phosphatase containing a dinuclear Fe(III)-Zn(II) active site. *Science.* 1995; 268:1489–1492. [PubMed: 7770774]
32. Voegtli WC, Khidekel N, Baldwin J, Ley BA, Bollinger JJM, Rosenzweig AC. Crystal structure of the ribonucleotide reductase R2 mutant that accumulates a μ-1,2-peroxodiiron(III) intermediate during oxygen activation. *J Am Chem Soc.* 2000; 122:3255–3261.
33. Rosenzweig AC, Nordlund P, Takahara PM, Frederick CA, Lippard SJ. Geometry of the soluble methane monooxygenase catalytic diiron center in two oxidation states. *Chem Biol.* 1995; 2:409–418.
34. Whittington DA, Lippard SJ. Crystal structures of the soluble methane monooxygenase hydroxylase from *Methylococcus capsulatus* (Bath) demonstrating geometrical variability at the dinuclear iron active site. *J Am Chem Soc.* 2001; 123:827–838. [PubMed: 11456616]
35. Wei, P-p; Skulan, AJ.; Miti , N.; Yang, YS.; Saleh, L.; Bollinger, JM., Jr; Solomon, EI. Electronic and spectroscopic studies of the non-heme reduced binuclear iron sites of two ribonucleotide reductase variants: comparison to reduced methane monooxygenase and contributions to O₂ reactivity. *J Am Chem Soc.* 2004; 126:3777–3788. [PubMed: 15038731]
36. Cryle MJ, Meinhart A, Schlichting I. Structural characterization of OxyD, a cytochrome P450 involved in -hydroxytyrosine formation in vancomycin biosynthesis. *J Biol Chem.* 2010; 285:24562–24574. [PubMed: 20519494]

37. Stubbe J, Nocera DG, Yee CS, Chang MCY. Radical initiation in the class I ribonucleotide reductase: Long-range proton-coupled electron transfer? *Chem Rev.* 2003; 103:2167–2201. [PubMed: 12797828]
38. Muller J, Lugovskoy AA, Wagner G, Lippard SJ. NMR structure of the [2Fe-2S] ferredoxin domain from soluble methane monooxygenase reductase and interaction with its hydroxylase. *Biochemistry.* 2002; 41:42–51. [PubMed: 11772001]
39. Sobrado P, Lyle KS, Kaul SP, Turco MM, Arabshahi I, Marwah A, Fox BG. Identification of the binding region of the [2Fe-2S] ferredoxin in stearoyl-acyl carrier protein desaturase: Insight into the catalytic complex and mechanism of action. *Biochemistry.* 2006; 45:4848–4858. [PubMed: 16605252]
40. Koglin A, Mofid MR, Löhr F, Schäfer B, Rogov VV, Blum MM, Mittag T, Marahiel MA, Bernhard F, Dötsch V. Conformational switches modulate protein interactions in peptide antibiotic synthetases. *Science.* 2006; 312:273–276. [PubMed: 16614225]
41. Lai JR, Koglin A, Walsh CT. Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. *Biochemistry.* 2006; 45:14869–14879. [PubMed: 17154525]
42. Cryle MJ, Schlichting I. Structural insights from a P450 carrier protein complex reveal how specificity is achieved in the P450BioI ACP complex. *Proc Natl Acad Sci USA.* 2008; 105:15696–15701. [PubMed: 18838690]
43. Tanovic A, Samel SA, Essen LO, Marahiel MA. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science.* 2008; 321:659–663. [PubMed: 18583577]
44. Farquhar ER, Koehntop KD, Emerson JP, Que L Jr. Post-translational self-hydroxylation: a probe for oxygen activation mechanisms in non-heme iron enzymes. *Biochem Biophys Res Commun.* 2005; 338:230–239. [PubMed: 16165090]
45. Kurtz DM Jr. Avoiding high-valent iron intermediates: superoxide reductase and rubrerythrin. *J Inorg Biochem.* 2006; 100:679–693. [PubMed: 16504301]
46. Froiland WA, Andersson KK, Lee SK, Liu Y, Lipscomb JD. Methane monooxygenase component B and reductase alter the regioselectivity of the hydroxylase component-catalyzed reactions. A novel role for protein-protein interactions in an oxygenase mechanism. *J Biol Chem.* 1992; 267:17588–17597. [PubMed: 1325441]
47. Miti N, Schwartz JK, Brazeau BJ, Lipscomb JD, Solomon EI. CD and MCD studies of the effects of component B variant binding on the biferrous active site of methane monooxygenase. *Biochemistry.* 2008; 47:8386–8397. [PubMed: 18627173]
48. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA.* 2001; 98:10037–10041. [PubMed: 11517324]

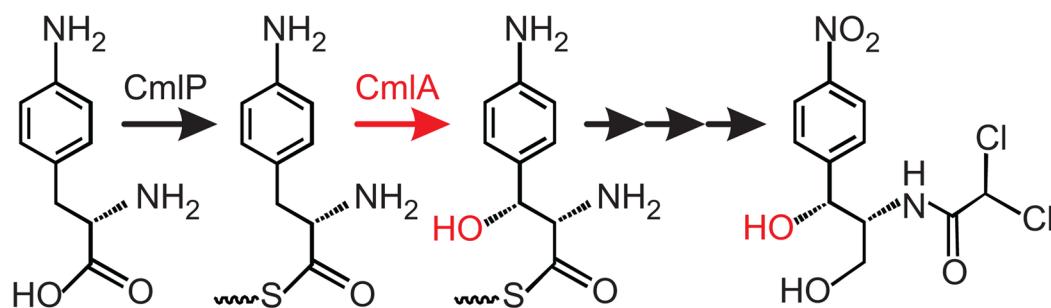


Figure 1.

Steps in the biosynthesis of chloramphenicol. The -hydroxyl shown in red is derived from O_2 .

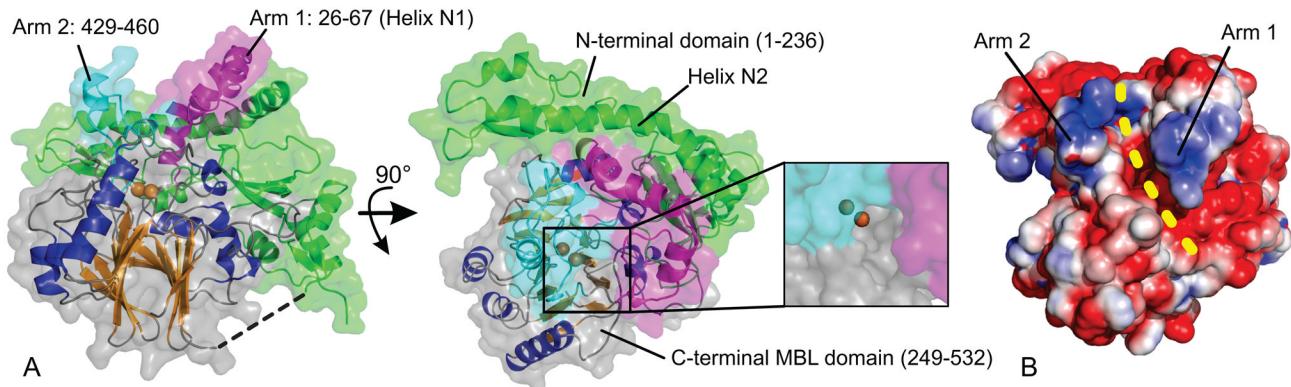


Figure 2.

Overall structure of CmlA showing important structural features and electrostatic properties. (A) The N-terminal domain is shown in green and the MBL-like catalytic domain in grey with α -helices blue and β -sheets orange. Iron atoms are displayed as orange spheres. A stretch of disordered residues separates the domains, indicated by the dashed line. Arms extending from CmlA and bordering the surface groove are denoted in magenta and cyan and the residue range is shown. The channel to the active site is positioned (see zoomed panel on right) at the base of a groove on the surface of CmlA, which runs between Arms 1 and 2. (B) Surface electrostatics map of CmlA. The tips of Arms 1 and 2 have notable basic character. The surface groove of CmlA lies between these arms and is denoted by the yellow dashed line. The map was generated using APBS⁴⁸ and contoured at 2 kT/e.

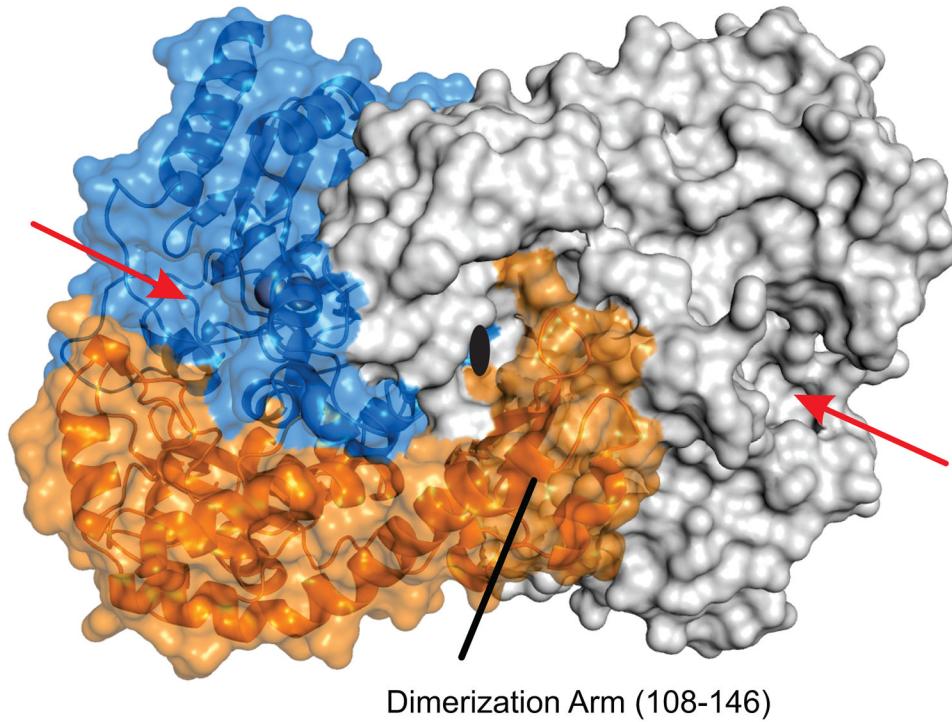
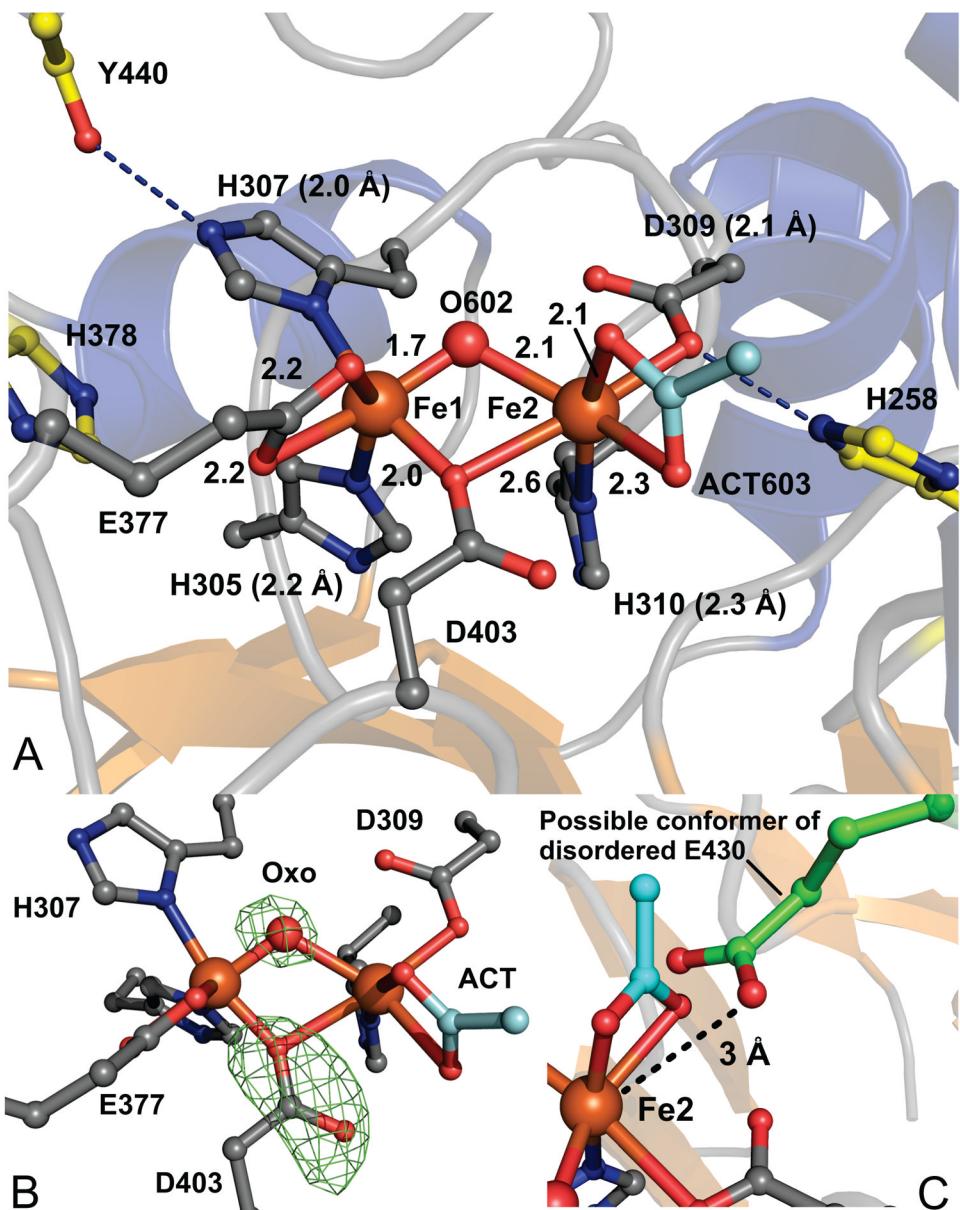


Figure 3.

CmlA is a dimer and the interaction between protomers is mediated by an extension from the N-terminal domain. Shown is a surface view of the CmlA dimer with the dimerization arm and corresponding residue range indicated. One protomer surface is transparent, with secondary structure shown in cartoon. In this protomer, the N-terminal domain is colored orange and the catalytic domain blue. The other protomer is shown in white. Red arrows indicate the locations of the active site channels. The black oval indicates the location of the 2-fold crystallographic symmetry axis that runs through the dimer.

**Figure 4.**

Structure of the CmlA active site and nearby residues. (A) Cluster geometry and bond distances in Ångströms, with iron ligands in grey, second-sphere conserved residues in yellow and the bridging oxo atom as a red sphere. Blue dashes indicate hydrogen bonds. The exogenous acetate ligand (Act) is shown in cyan. When not directly indicated, distances are given in parenthesis. (B) Positive $F_o - F_c$ difference electron density (green) following removal of the oxo bridge atom and trimming of D403 back to Ala, followed by 5 cycles of refinement. $F_o - F_c$ difference electron density contoured at $+4\sigma$ for the oxo and at $+7\sigma$ for D403. The map was calculated using a resolution range of 38.7–2.17 Å. (C) One possible rotameric conformation of the disordered E430 side chain (green) places the carboxylate within 3 Å of Fe2, overlapping partially with the acetate (cyan) binding site.

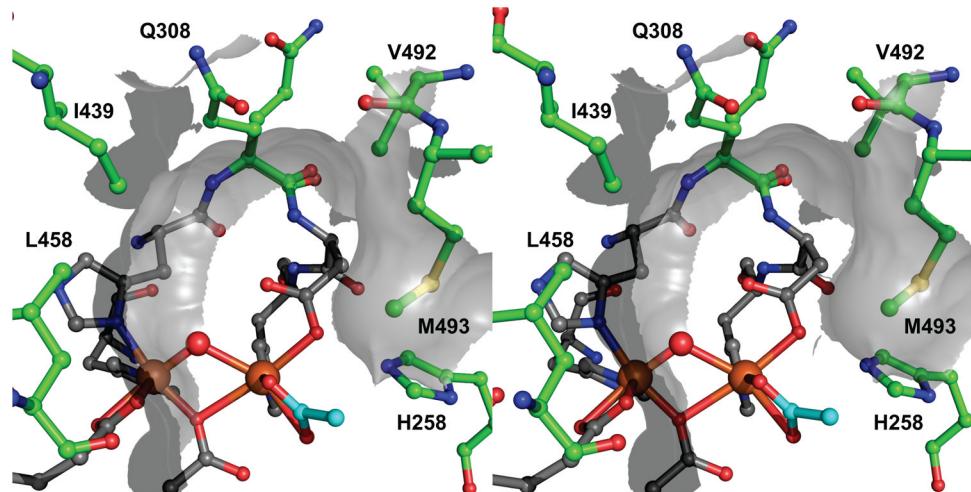
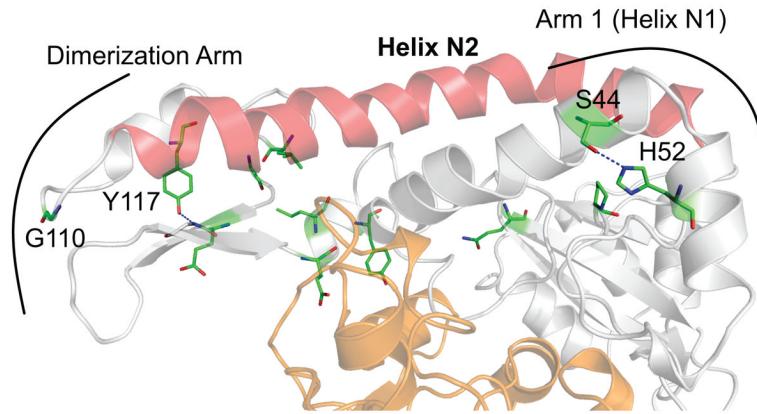


Figure 5.

Stereo-view diagram of the CmlA active site pocket as viewed from the channel. Residues that form the pocket are represented in green and labeled. The cluster is represented as in Figure 4. Grey surface contours illustrate the shape of the pocket and define a likely binding orientation for the aromatic substrate. W438 is positioned above the pocket and has been omitted for clarity.

**Figure 6.**

Conserved residues of the N-terminal domain identified from sequence alignments with CmlA homologs and mapped onto the crystal structure. Conserved residues are shown in green, with the catalytic domain in orange and N-terminal domain in white. Residues primarily serve structural roles by interacting with nearby residues through hydrogen bonds and hydrophobic packing or facilitating turns on the arms (G110). Two hydrogen bonded pairs are denoted by blue dashed lines: S44 and H52 and Y117 and the amide of a nearby residue. Helix N2 in the N-terminal domain is colored red.

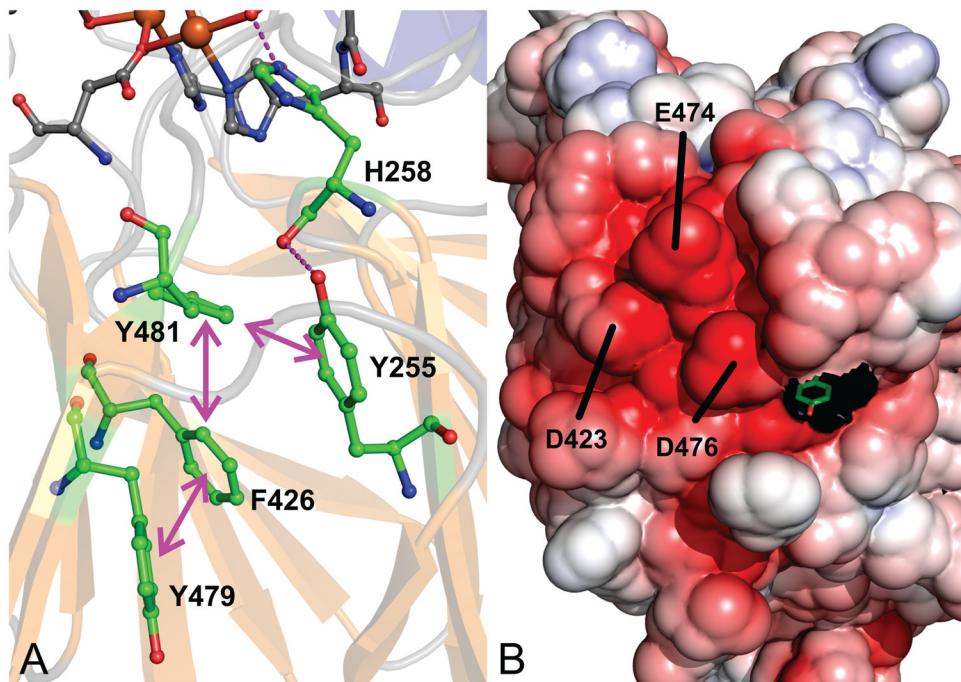
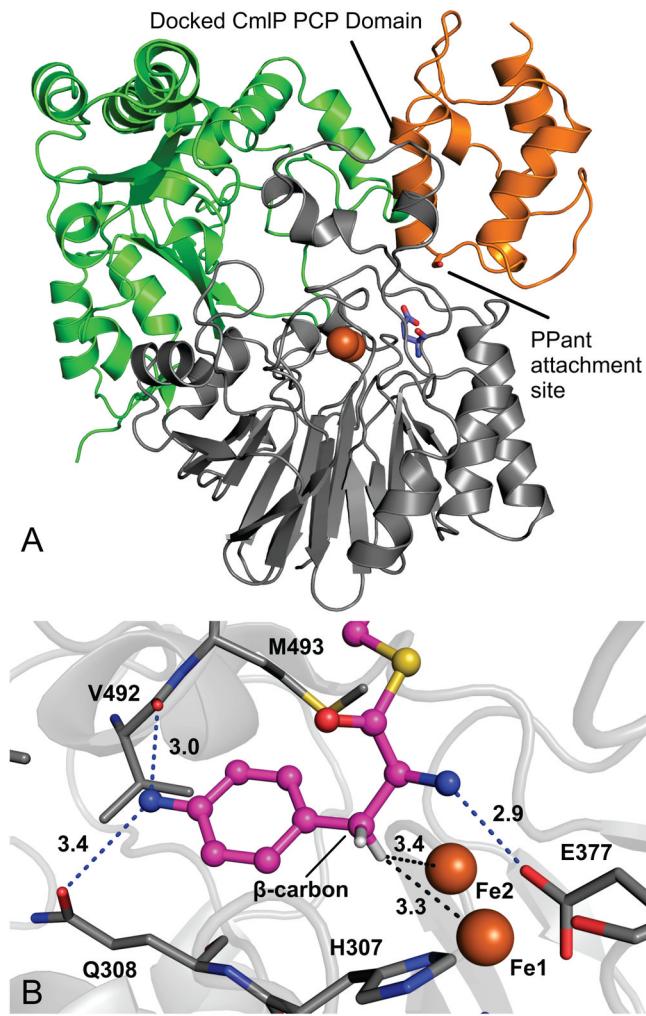
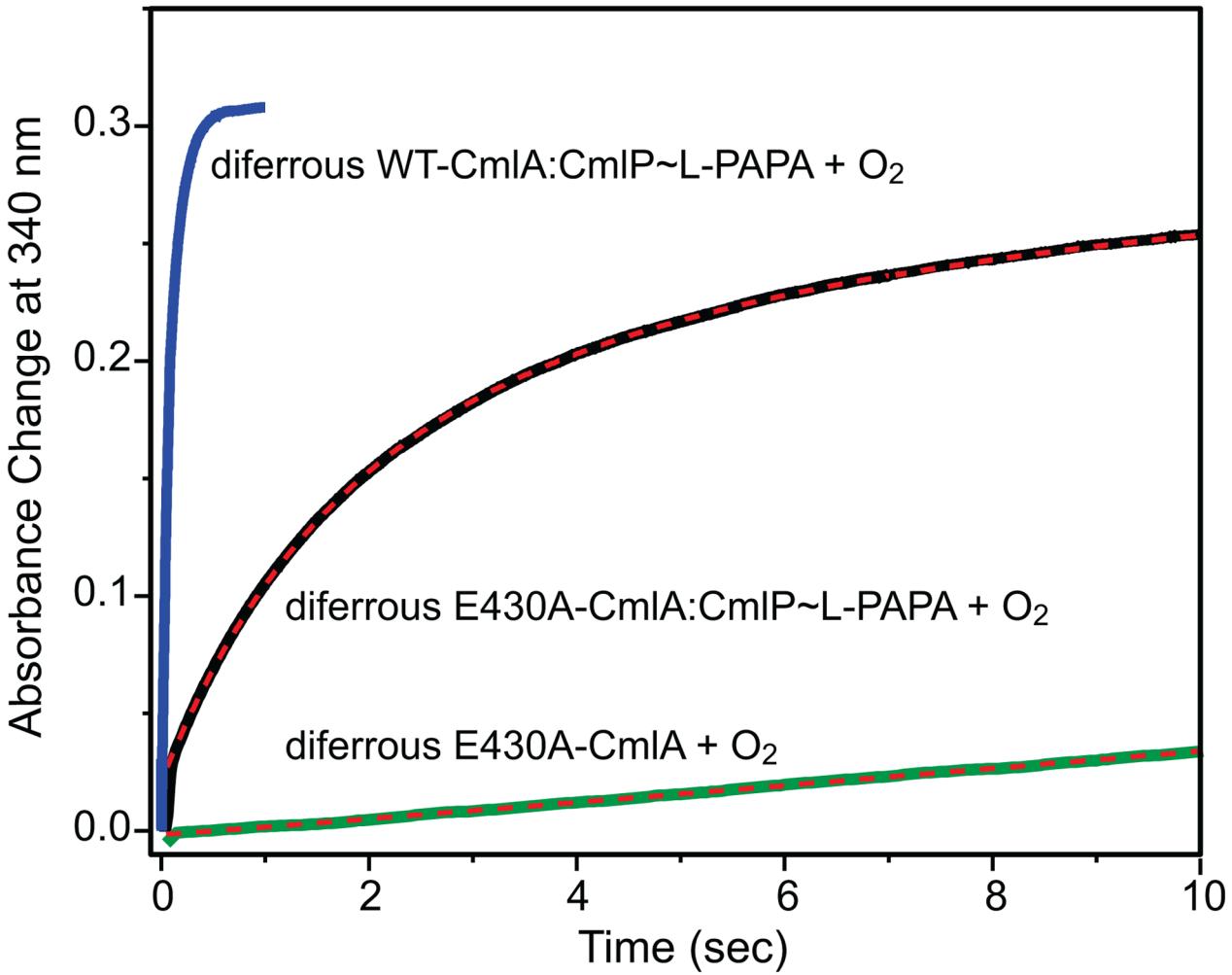


Figure 7.

Identification of a possible electron delivery network and reductase binding site in CmlA. (A) Conserved aromatic network which extends from solvent-exposed Y479 to the diiron cluster. (B) Surface electrostatics map of CmlA showing a possible acidic binding site at the base of catalytic domain. Y479 is drawn in stick colored by atom (carbon, green). The map was generated using APBS⁴⁸ and contoured at 5 kT/e.

**Figure 8.**

Computationally docked structures. (A) The homology-modeled CmlP PCP domain and CmlA. The PCP domain likely binds above the channel to the active site but interacts minimally with the N-terminal domain. The PCP domain is shown in orange with the putative PPant attachment site labeled. The CmlA N-terminal and catalytic domains are colored in green and grey, respectively. E430 is positioned on a loop below the putative PCP domain binding site and is shown in blue. (B) Model of the PPant-L-PAPA substrate docked in the active site of CmlA. PPant-L-PAPA is shown in magenta, protein residues in grey and hydrogens on the β -carbon are shown in white. All distances are in Å. Hydrogen bonds are indicated by blue dashes. Docking was performed using Sybyl-X 2.0.

**Figure 9.**

Stopped flow reaction between 100 μM chemically reduced wild type or E430A CmlA in complex with 2.5-fold excess L-PAPA-loaded CmlP_{AT} rapidly mixed with O₂ saturated buffer at 4 °C at pH 7.5. The reaction of wild type CmlA (blue trace) is complete within 1 s ($1/\tau_1 = 17 \text{ s}^{-1}$ and $1/\tau_2 = 5.2 \text{ s}^{-1}$). The E430A reaction (black trace) can be fit by a sum of three exponentials (red dashed line) with $1/\tau_1 = 0.68 \pm 0.09 \text{ s}^{-1}$, $1/\tau_2 = 0.18 \pm 0.04 \text{ s}^{-1}$ and $1/\tau_3 = 0.020 \pm 0.006 \text{ s}^{-1}$. The basis for the multiple exponential time course is unknown. Autooxidation of the cluster upon exposure of reduced E430A to O₂ in the absence of CmlP (green trace) can be fit (red dashed line) to a single exponential with $1/\tau_1 = 0.014 \pm 0.007 \text{ s}^{-1}$, the same as that observed for wild type CmlA autooxidation. The reaction of wild type CmlA (blue trace) is from ref.³

Table 1

Data Collection and Refinement Statistics

| | Native | Se |
|---|----------------------------------|----------------------------------|
| Data collection | | |
| Space group | P4 ₃ 2 ₁ 2 | P4 ₃ 2 ₁ 2 |
| Cell dimensions | | |
| a, b, c (Å) | 153.2, 153.2, 93.4 | 153.3, 153.3, 92.7 |
| , , (°) | 90, 90, 90 | 90, 90, 90 |
| Wavelength (Å) | 0.97857 | 0.97926 |
| Resolution (Å) [*] | 38.7-2.17 (2.21-2.17) | 20.9-2.60 (2.64-2.60) |
| Unique reflections | 56972 | 35553 |
| R _{merge} (%) ^{*,†} | 6.3 (43.5) | 10.8 (45.0) |
| I/I [*] | 28 (3.1) | 27 (4.5) |
| Completeness (%) [*] | 96.4 (81.6) | 99.9 (100) |
| Redundancy [*] | 7.3 (7.1) | 7.2 (7.4) |
| Overall figure of merit | | 0.405 |
| Cruickshank's Diffraction Precision Index (DPI) | 0.154 | |
| Refinement | | |
| Resolution, Å | 38.7-2.17 | |
| Number of reflections | 54097 | |
| R _{work} /R _{free} (%) [‡] | 20.3/23.0 | |
| Averaged B factor (Å ²) | 35.1 | |
| RMSDs | | |
| Bond lengths (Å) | 0.0151 | |
| Bond angles (°) | 1.78 | |
| Ramachandran analysis | | |
| Favored regions (%) | 95.6 | |
| Allowed regions (%) | 4.0 | |
| Disallowed regions (%) | 0.4 | |

RMS, root mean square

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

^{*}Highest resolution shell is shown in parentheses[†]R_{sym} = $\frac{\sum_{hkl} \sum_i |I_{hkl,i} - \langle I \rangle_{hkl}|}{\sum_{hkl} \sum_i |I_{hkl,i}|}$, where I_{hkl} is the intensity of a reflection and $\langle I \rangle_{hkl}$ is the average of all observations of the reflection[‡]R_{free}, R-factor calculated from 5% of the data excluded from refinement.