Published in final edited form as:

J Mol Biol. 2008 December 31; 384(5): 1353–1368. doi:10.1016/j.jmb.2008.10.048.

Specificity and Reactivity in Menaquinone Biosynthesis: The Structure of *Escherichia coli* MenD (2-Succinyl-5-Enolpyruvyl-6-Hydroxy-3-Cyclohexadiene-1-Carboxylate Synthase)

Alice Dawson, Paul K. Fyfe, and William N. Hunter*

Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK

Abstract

The thiamine diphosphate (ThDP) and metal-ion-dependent enzyme 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate synthase, or MenD, catalyze the Stetter-like conjugate addition of α -ketoglutarate with isochorismate to release 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate and carbon dioxide. This reaction represents the first committed step for biosynthesis of menaquinone, or vitamin K₂, a key cofactor for electron transport in bacteria and a metabolite for posttranslational modification of proteins in mammals. The medium-resolution structure of MenD from *Escherichia coli* (EcMenD) in complex with its cofactor and Mn²⁺ has been determined in two related hexagonal crystal forms. The subunit displays the typical three-domain structure observed for ThDP-dependent enzymes in which two of the domains bind and force the cofactor into a configuration that supports formation of a reactive ylide. The structures reveal a stable dimer-of-dimers association in agreement with gel filtration and analytical ultracentrifugation studies and confirm the classification of MenD in the pyruvate oxidase family of ThDP-dependent enzymes. The active site, created by contributions from a pair of subunits, is highly basic with a pronounced hydrophobic patch. These features, formed by highly conserved amino acids, match well to the chemical properties of the substrates. A model of the covalent intermediate formed after reaction with the first substrate α -ketoglutarate and with the second substrate isochorismate positioned to accept nucleophilic attack has been prepared. This, in addition to structural and sequence comparisons with putative MenD orthologues, provides insight into the specificity and reactivity of MenD and allows a two-stage reaction mechanism to be proposed.

Keywords

crystal structure; enzyme mechanism; menaquinone biosynthesis; thiamine diphosphate cofactor

Introduction

Menaquinone, or vitamin K₂, is a lipid-soluble molecule with distinct biological functions according to the type of organism in which it is found. Its chemical properties are exploited in the electron transport chain in Gram-positive aerobes and facultative Gram-negative bacteria under anaerobic conditions.^{1,2} In an aerobic environment, Gram-negative organisms utilize ubiquinone.^{3,4} In mammals, which acquire menaquinone from intestinal microflora and diet, menaquinone is an important enzyme cofactor involved in glutamate γ -

carboxylation.⁵ This is an important posttranslational modification of certain proteins involved in blood coagulation, bone metabolism, and vascular biology.

The biosynthesis of menaquinone has been extensively studied, mainly in *Escherichia coli*, and shown to involve a pathway of eight or nine proteins exploiting a range of cofactors.^{1,2} The starting point is chorismate, the branch point intermediate of the shikimate pathway.⁶ Chorismate is isomerized to isochorismate and then converted, in two stages, to o-succinyl-1-benzoate. Two further enzyme-catalyzed reactions, involving ATP and CoA (coenzyme A), led to a CoA-naphthalene derivative and a thioesterase removes the CoA moiety to release 1,4-dihydroxy-2-naphthanoate. Prenylation and methylation then produce the reduced form of menaquinone. The structure–function–mechanism relationships for several enzymes involved in menaquinone biosynthesis have been elucidated. These include MenB,^{7–9} MenF,^{10,11} and MenH.¹² A notable exception is the enigmatic enzyme that catalyzes the first committed step in menaquinone biosynthesis, MenD (2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate synthase).

Initially, MenD was assigned as the bifunctional enzyme 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) synthase^{1,13} and reactions catalyzed were considered to be decarboxylation of α -ketoglutarate and ligation of a second substrate, isochorismate, to produce 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate (SEPHCHC). Cleavage of the pyruvate moiety then produced SHCHC (Fig. 1a). Recently, however, more rigorous analyses provided evidence that the enzyme is actually SEPHCHC synthase and a complicating factor in the characterization of MenD was the slow, spontaneous release of pyruvate from SEPHCHC (Fig. 1a).^{14,15} The cleavage of pyruvate from SEPHCHC is catalyzed by MenH (also known as YfbB, Fig. 1a).¹² In *E. coli*, MenH is encoded after MenD on the same operon as the majority of the other menaquinone biosynthetic genes.

The observation that MenD is dependent on the cofactor thiamine diphosphate (ThDP) and a divalent metal cation, with Mg^{2+} and Mn^{2+} providing optimal activity, has structural and mechanistic implications.^{13,14} ThDP, the active form of vitamin B₁, consists of a diphosphate moiety, aromatic thiazolium, and aminopyrimidine rings. Tautomerization of the aminopyrimidine is an important chemical property of the cofactor (Fig. 1b). It has been demonstrated by spectroscopic methods that the N4'-aminopyrimidine and N1',N4'-iminopyrimidine tautomers are present in a number of pyruvate decarboxylases with the latter form in particular present in the predecarboxylation intermediate.^{16,17} Within the confines of an enzyme active site, ThDP adopts a bent conformation that supports deprotonation of C2 by the basic imine N4' to produce the activated form of the cofactor, a nucleophilic ylide (Fig. 1b). The interaction of a glutamate, highly conserved in ThDP-dependent decarboxylases, with the aminopyrimidine N1' supports the generation of the N4'-iminopyrimidine form that can acquire a proton from the thiazolium C2. It is this almost universal triumvirate of glutamate–iminopyrimidine–thiazolium in ThDP-dependent enzymes that determines reactivity at C2 and the generation of the reactive ylide.^{18–21} MenD is the only enzyme known to catalyze addition of a ThDP intermediate to the β -carbon of a second substrate, a reaction that is similar to the Stetter reaction, which is a 1,4-addition, or conjugate addition, of an aldehyde to a β -unsaturated compound.²² The Stetter reaction is catalyzed by cyanide or thiazolium salts.

A complication in identifying potential active-site residues is the low level of sequence conservation of ThDP-dependent enzymes.²¹ Even within MenD sequences from different organisms, the identity is typically 20–30% (data not shown), and structural data are critical to delineate the important constituents of the enzyme active site. We set out therefore to characterize the structure of MenD and improve our understanding of menaquinone

biosynthesis. Crystallization of the *E. coli* enzyme (*EcMenD*) has previously been reported²³ and we now present the medium-resolution structure of the *EcMenD*:ThDP:Mn²⁺ complex. The structure has been exploited to construct a model of the first covalent intermediate formed following reaction with α -ketoglutarate, together with the second substrate isochorismate. This, in combination with the analysis of orthologous *MenD* sequences and comparisons with related enzymes, provides a framework for discussion of substrate recognition, the chemical determinants of catalysis, and the enzyme mechanism in the first committed step of menaquinone biosynthesis.

Results and Discussion

Quaternary structure

It was important to characterize the quaternary structure of *EcMenD* (556 residues; molecular mass, ~61.4 kDa) since all ThDP-dependent enzymes form dimeric or tetrameric structures, with the cofactor binding and active sites formed by contributions from at least two subunits. An efficient recombinant protein expression system for an N-terminal hexahistidine-tagged *EcMenD* was prepared and a highly purified enzyme was obtained with a yield of approximately 100 mg L⁻¹ of culture. The final purification step employed analytical gel filtration and indicated that the bulk of the apoprotein, approximately 75%, eluted with a mass of about 150 kDa and that 25% eluted with a mass of approximately 300 kDa (data not shown) irrespective of whether the histidine tag was present (native *EcMenD*) or had been removed by proteolysis as in the case of a selenomethionine derivative (SeMet *EcMenD*). Although these data indicate that a dimer is the predominant form of *EcMenD* in solution with a smaller quantity of tetramer present, we note that the observed mass differs significantly from the theoretical mass of the dimer (~123 kDa). The peak corresponding to the dimer was isolated and that protein was used for crystallization and analytical ultracentrifugation (AU). AU was performed on untagged apo-*EcMenD* and in the presence of ThDP and Mn²⁺ (data not shown). The apo-*EcMenD* sample displayed a single species with a mass of ~117 kDa, which more closely agrees with the theoretical mass of the dimer than observed by gel filtration. In the presence of ThDP and a divalent cation, a dimer (~130 kDa) was still predominant but there was also evidence for monomer (~60.8 kDa), trimer (~190 kDa), and tetramer (~252 kDa). The data indicate that the most abundant species of *EcMenD* is a dimer but that higher-order assembly into a tetramer is also observed. Under the conditions used for AU, the holoenzyme quaternary structure is not particularly stable, hence the observation of a monomer, the presence of which allows for interaction with the dimer to provide a trimeric species. The differences observed between apo- and holoenzyme suggest the possibility that ThDP and metal ion binding may induce gross conformational changes.

Structure determination

The apoenzyme did not crystallize but holo-*EcMenD* provided two similar hexagonal forms, both of which diffract to medium resolution. The native protein, with the N-terminal His-tag, crystallized readily in space group *P*6₅22 with approximate unit cell lengths of $a=b=93$ Å and $c=465$ Å. A SeMet derivative was prepared but failed to provide useable crystals under conditions used for the native enzyme. Removal of the histidine tag by proteolytic cleavage and further screening identified different conditions for crystal growth. These SeMet *EcMenD* crystals display space group *P*6₅ with approximate unit cell lengths of $a=b=96$ Å and $c=463$ Å. Despite similar unit cell dimensions, the merging of diffraction data for SeMet *EcMenD* in point group 622 gave an R_{merge} of 32% while native *EcMenD* data gave an R_{merge} of 8.5% in this high-symmetry system (Table 1). A self-rotation function for the SeMet *EcMenD* data in *P*6 indicated that a noncrystallographic symmetry (NCS) 2-fold axis was displaced by 3.5° from the position of a crystallographic 2-fold in *P*622. The

Matthews coefficient²⁵ is estimated as 2.4 Å³/Da, and the bulk solvent content is at approximately 50% for each crystal form assuming that the asymmetric units of native and SeMet *EcMenD* contain two and four subunits, respectively.

A single-wavelength anomalous diffraction (SAD) phasing experiment at the Se K-edge provided phases to 2.8 Å resolution for SeMet *EcMenD* and an interpretable electron density map. Once two subunits had been sufficiently modeled, a switch to the more ordered native structure data at 2.5 Å resolution was made using molecular replacement and the model was refined. Once completed, the native model was used to determine the SeMet *EcMenD* structure to investigate any structural differences between dimer and tetramer forms.

Analysis of crystallographic statistics and model geometry using MolProbity²⁶ indicates that acceptable medium-resolution models have been produced (Table 1). Continuous electron density is observed for almost the entire polypeptide chains in the SeMet *EcMenD* structure, missing only the additional N-terminal residues (Gly–His) that remain after proteolytic cleavage of the His-tag. The native *EcMenD* still carries the His-tag but there is no electron density for those residues or for Gly176 and Glu177 of subunit A of native *EcMenD*. These are in the loop linking β6 and α6, a region of polypeptide that displays high thermal parameters in all subunits. An example of electron density is shown when the cofactor is described below.

Six independent *EcMenD* subunit structures have been determined from the two crystal forms. These are highly conserved within each asymmetric unit and between native and SeMet structures. For example, the 554 C^α positions of subunits A and B in native *EcMenD* overlay with a root-mean-square deviation (r.m.s.d.) of 0.36 Å, while subunit A in the native structure overlays with an r.m.s.d. of 0.32 Å onto subunit A of SeMet *EcMenD*. Intriguingly, the similarity extends to a dimer-of-dimers association. In native *EcMenD*, a crystallographic 2-fold axis, one-third along c bisecting the *a*–*b* plane, generates a tetramer. In SeMet *EcMenD*, this crystallographic axis is converted to an NCS 2-fold axis and the asymmetric unit is a tetramer. The asymmetric unit tetramer of SeMet *EcMenD* superposes on the asymmetric dimer plus symmetry-related dimer of native *EcMenD* with an r.m.s.d. of 0.53 Å over 2220 C^α positions. Given such a high degree of similarity, it is only necessary to detail the more ordered native structure. We concentrate on subunit A, and where contributions from the partner subunit are important, the subunit is assigned in residue naming.

The *EcMenD* subunit and dimer formation

The *EcMenD* subunit has approximately 65% of residues in elements of secondary structure and displays the three-domain architecture typical of ThDP-dependent enzymes (Figs. 2 and 3a).²⁸ The N-terminal domain (I) covers approximately 200 residues. An ordered segment of extended structure links domain I to domain II, which spans residues 210–340. A slightly shorter linker region then leads to the C-terminal domain III. Domains I and III display similar topology, each consisting of a central six-stranded parallel β-sheet sandwiched between several α-helices (Figs. 2 and 3a). The principal difference in the two domains resides in the loop linking β4 and β5 and the equivalent in domain III, the segment linking β17 to β18. The polypeptide adopts a different orientation, displays additional elements of secondary structure, and is longer in domain III compared to domain I. The smaller central domain (II) has closely related topology, with a five-stranded parallel β-sheet and six helices. The sheet gains an extra strand by extension across a dimerization interface (discussed below).

Structural comparisons between *EcMenD* and other ThDP-dependent enzymes were explored with DALI²⁹ and the Secondary Structure Matching server³⁰ and now clearly

indicate that, as predicted, MenD belongs in the pyruvate oxidase (POX) family of ThDP-dependent enzymes, the largest of the five ThDP-dependent enzyme families.^{28,30} Monomer A of native *EcMenD* overlays with a monomer of benzoylformate decarboxylase [Protein Data Bank (PDB) code: 1BFD³¹] with an r.m.s.d of 2.54 Å over 426 aligned C α positions, despite a sequence identity of only 16%. Similarly, a monomer of oxalyl-CoA decarboxylase (OXC; PDB code: 2C31³²), which shares 14% sequence identity, overlays with *EcMenD* with an r.m.s.d. of 2.56 Å over 424 aligned C α positions. The least-squares fit of individual domains gives r.m.s.d. values that are less than the overall alignments. For example, overlay of the three domains of *EcMenD* with the individual domains of OXC results in r.m.s.d. values of 2.1 Å (156 C α positions), 2.5 Å (94 C α positions), and 1.9 Å (150 C α positions), respectively. This suggests that domains I and III, which bind the pyrimidine and diphosphate components of the cofactor, respectively, are structurally more conserved between *EcMenD* and POX family members than the central domain II. Several ThDP-dependent enzymes, including POX and acetohydroxy acid synthase,³³ bind a nucleotide in domain II. This nucleotide can fulfill a purely structural role, like the flavin adenine dinucleotide in acetohydroxy acid synthase, or alternatively participate in catalysis as exemplified by flavin adenine dinucleotide in POX.³⁴ MenD is not dependent on binding a nucleotide in domain II, and the overlay of *EcMenD* with POX shows that the nucleotide-binding region of POX is replaced by the loop linking β 11 with α 11 in *EcMenD* (data not shown). This loop appears therefore to substitute for the stabilizing effect a nucleotide has on the structure of a subset of ThDP-dependent enzymes.

A comparison of 78 MenD sequences indicates a low level of sequence conservation for this enzyme with only 46 residues out of 556 being highly conserved (Fig. 2). The MenD sequence list was derived by omitting those that displayed a sequence identity of $\geq 90\%$ since they were all from closely related species. The lowest level of conservation occurs in the central domain II (2 residues), followed by domain III (with 16), and the most conserved domain is domain I with 28 positions conserved.

The *EcMenD* dimer is formed by interactions involving domains I and III (Fig. 3b). We estimate that 42 hydrogen bonds and six salt bridge interactions are involved in stabilizing the dimer together with the additional contribution from two ThDP molecules binding to each of the subunits. Approximately 50% of atoms at the interface are nonpolar, indicating a significant contribution of van der Waals forces to oligomerization. Analysis using the PISA server³⁵ indicates that 15% (3340 Å²) of the surface area of a monomer (23,200 Å²) is occluded on formation of the dimer.

Both crystal structures contain a dimer-of-dimers, constituting the asymmetric unit of the SeMet protein (subunits A, B, C, and D) and generated by a crystallographic 2-fold axis applied to subunits A and B in the native structure as described earlier. Subunits C and D of the SeMet structure correspond to crystallographic equivalents of subunits A and B in the native structure, respectively. Residues on domain II make the most significant contributions to the formation of the tetramer, which involves $\sim 8\%$ (1860 Å²) of the surface area of a subunit. This is primarily due to interaction with one partner subunit A with D or B with C (1500 Å²) with a smaller area of interaction (360 Å²) formed by subunit A with C or B with D (Fig. 4). Two secondary-structure elements contribute to the interface. An interaction between β 8 strands forms a two-stranded antiparallel β -sheet, while β 7 of subunit A interacts with β 13 of subunit D to continue the β -sheet of domain II across the tetramer interface, so that the sheet in subunit A extends into subunit D (data not shown). This crystallographic dimer-of-dimers is markedly similar to the structure of benzoylformate decarboxylase, which has the same secondary-structure elements present at the dimer-of-dimers interface and is tetrameric in both solution and crystalline states.³¹

Cofactor binding

The ThDP-binding site in *EcMenD*, in common with all other ThDP-dependent enzymes, is formed by contributions from the two subunits that comprise the dimer, and each dimer contains two active sites approximately 20 Å apart (Fig. 3b). The cofactor is positioned by hydrogen bond and van der Waals interactions with the protein in addition to metal ion complexation. The residues important for cofactor binding are contributed from domains I and III, the most highly conserved domains across MenD sequences (Fig. 2).

The diphosphate and associated metal ion are firmly anchored by a number of interactions, and a conserved motif (Gly441-Asp442-X₂₅-Asn468-Asn469 in *EcMenD*)²⁰ generally allows identification of ThDP binding proteins in the absence of any other data. Although the Asn-Asn combination is not strictly conserved across MenD sequences, an asparagine always appears in at least one position (data not shown). Three segments of the polypeptide (the residues linking α15-β15, β17-α18, and the N-terminal section of α17) create a polar pocket in which Mn²⁺ is bound with octahedral coordination. The ligands are two oxygen atoms from the diphosphate, the main-chain carbonyl of Gly471, the side-chain carbonyl of Asn469, the side chain of Asp442, and a water molecule (Fig. 5). The metal ion interactions with the cofactor diphosphate are supplemented by nine direct cofactor-protein hydrogen bonds formed with main-chain amides (Leu392, Asp442, Leu443, Ser444, Gln473, and Ile474), a side-chain amide (Asn469), and side-chain hydroxyls (Ser391 and Ser444). Two ordered water molecules, one of which is a ligand for the metal ion, also provide bridging interactions between the diphosphate and the protein (data not shown). Two of these residues, Ser444 and Asn469, are not strictly conserved and interact using the functional groups on their side chains. Ser444 is present in 37 MenD sequences and replaced by threonine in 12 sequences, suggesting that the side-chain interaction with the cofactor is often conserved. In the remaining 29 sequences, an alanine occupies this position, perhaps allowing enough room for a water molecule to satisfy the hydrogen-bonding capacity of the phosphate group. The Asn469 interaction with Mn²⁺ is also likely to be conserved. Asn469 is conserved in 65 sequences, and an aspartate is observed in the other 13 MenD sequences. The substitution of asparagine would result in the loss of the side-chain amide interaction with the phosphate group.

The thiazolium forms few direct interactions with the protein but rather is anchored in place by the interactions of the pyrimidine and diphosphate groups. This may allow a small degree of flexibility of this part of the cofactor.³⁶ Ser391 OG is 2.8 Å distant from the thiazolium sulfur, suggesting the formation of a hydrogen bond. This serine side chain also interacts with the diphosphate, separated by 2.8 Å, and it is possible that the diphosphate O is protonated to support a hydrogen-bonding interaction. There are van der Waals interactions formed by the thiazolium with Ile418 and Phe475 (not shown) located on either side of the inverted V formed between the pyrimidine and thiazolium rings and, at the side, Ile474. Hydrophobic residues are always found in these positions in the cofactor-binding pocket of ThDP-dependent enzymes. Ile418 forms van der Waals interactions also with the pyrimidine component and is strictly conserved in MenD sequences. In addition, the thiazolium methyl group interacts with Pro30 of the partner subunit (data not shown).

The pyrimidine moiety is tucked into a cleft formed by the turn between β15-α16 of subunit A and the N-terminal sections of α3 and α4 of subunit B. The three functional groups participate in hydrogen-bonding interactions with subunit B. N1' appears to form a bifurcated hydrogen bond with the carboxylate of Glu55. N3' accepts a hydrogen bond donated from the main-chain amide of Ile418, and N4' donates a hydrogen bond to the main-chain carbonyl of Ser416. Ile418 of subunit A is on one side of the pyrimidine, as described above, and on the other side is a polar segment where an ordered water molecule fills the gap between this part of the cofactor and the protein (data not shown).

Glu55 is strictly conserved in MenD sequences, and mutagenesis of this residue in *EcMenD* abolishes enzyme activity.¹³ Glu55, together with Asp54 and Arg56, forms a conserved motif in MenD. A network of hydrogen bonds and salt bridges formed by Asp54 and Arg56 with the equivalent residues in the partner subunit and the presence of Tyr447 (conserved as Tyr or His in MenD) form a tight turn at the start of $\alpha 3$ and, together with the bulk of the aromatic side chain, make contributions to the assembly of the MenD dimer and serve to position Glu55 to bind the cofactor and thereby influence the tautomeric state of the pyrimidine (Figs. 5 and 6). This group of residues is worthy of note with respect to comparisons of ThDP-dependent enzymes.

ThDP-dependent enzymes are oligomeric, and investigations into aspects of communication between active sites have been reported.^{21,37} In the *Bacillus stearothermophilus* (now *Geobacillus stearothermophilus*) E1 subunit of pyruvate dehydrogenase, a heavily hydrated cavity spanning some 20 Å between the two ThDP molecules in the dimer was observed. This channel is lined with acidic residues (four aspartates and six glutamates), and it has been postulated that this channel and the associated water molecules may assist the shuttling of protons back and forth through this tunnel, allowing the cofactors to act in reciprocal fashion as general acids/bases in catalysis, and that this explained the requirement for oligomers, conformational asymmetry, alternating site reactivity, and the kinetic properties of E1 and other thiamine-dependent enzymes.³⁷ While such a theory may apply to some, it is unlikely to be a universal feature of all ThDP-dependent enzymes since the channel is absent in transketolase,¹⁸ OXC,³² and *N*²-(2-carboxyethyl)arginine synthase (CEAS)³⁸ and not all members of the family display alternating site reactivity. There is also no channel linking the active sites in *EcMenD* where the Asp54–Glu55–Arg56 motif and Tyr447 block access between the two active sites (Fig. 6).

A key feature of ThDP-dependent enzymes is the V-shaped conformation forced upon the cofactor by its binding site, with either end tethered to different domains (Figs. 5 and 7). The net result is the close proximity of the pyrimidine N4' and thiazolium C2. In crystal structures of ThDP itself, the distance is 3.9 Å.^{40,41} In *EcMenD*, the average distance between N4' and C2, from the six independent subunit structures, is 3.1 Å. This distance agrees well with other ThDP-dependent enzymes with the exception of CEAS where the separation is much shorter at around 2.3 Å. The proximity of these two atoms, in conjunction with the presence of a glutamic acid–N1' interaction, pyrimidine tautomerization, and the positively charged N3 of the thiazolium, allows abstraction of the acidic C2 hydrogen by N4' to produce the activated ylide ready to react with substrate.

With one notable exception, ThDP-dependent enzymes characterized thus far have the equivalent of Glu55 interacting with N1', an interaction that is vital to activate the cofactor and support formation of the covalent intermediate with the first substrate. The exception is *E. coli* glyoxylate carboligase where a valine occupies the position near to N1'.⁴² This enzyme appears to compensate for the lack of a polar interaction with the pyrimidine moiety with an aliphatic, nonpolarizable environment for the thiazol group. Such an environment stabilizes the neutral ylide 4'-aminopyrimidine form of the cofactor but reduces the reactivity at C2. The reduction in reactivity is likely balanced by the enhanced reactivity of the substrate in this case, an aldehyde as opposed to the ketone substrates processed by the majority of ThDP-dependent enzymes. The presence of a carboxylate to interact with N1' accelerates deprotonation at C2 and in so doing prevents the enzyme getting stuck in a low-energy intermediate.⁴²

A model for substrate (α -ketoglutarate and isochorismate) binding

Only a limited number of ThDP-dependent enzyme substrate complexes have been experimentally determined due to the transient nature of the intermediates formed by the

cofactor-substrate complexes.^{32,43} An elegant study of OXC, using modified ThDP derivatives, provided snapshots of several stages in the decarboxylation cycle (OXC has no ligase activity) and revealed that the covalent ThDP-substrate intermediate has an internal hydrogen bond between the hydroxyl group, remaining after decarboxylation, and the pyrimidine N4'.³² This study provided a template to assist our modeling of the first covalent intermediate in *EcMenD*. The second substrate, isochorismate, was positioned such that the reactive carbon was 1.5 Å distant from the nucleophilic carbon of the activated intermediate and oriented to minimize steric clash with the protein (Fig. 7). It was clear that conformational changes involving residues in the $\beta 1$ – $\alpha 2$ turn and several side chains were likely to be required for binding of the second substrate in a productive fashion. However, no attempts were made to generate an energy-minimized structure of the complex since we preferred to use the model to simply infer regions of the active site that might be relevant to specificity and reactivity.

The *EcMenD* active site is, as described, formed by contributions from two subunits. It consists of a narrow cavity having a width and length of approximately 8 and 15 Å, respectively, and the distance from the surface of the protein to ThDP C2 is nearly 20 Å (Fig. 7). The wall of the cavity is primarily basic due to the presence of five arginine residues (Arg33A, Arg107A, Arg395A, Arg293B, and Arg413B) together with a lysine (Lys292B). A hydrophobic patch on one side of the active site is created by the side chains of residues Ile474B, Phe475B, and Leu478B. These residue types match well to the acidic substrates and, in the case of isochorismate, a significant hydrophobic character. Seven of these nine amino acids are strictly conserved in MenD sequences (Fig. 2). The two exceptions, Lys292 and Arg293, are more distant from the catalytic center than the other residues discussed (Fig. 7) and are located in domain II, which contains only two highly conserved residues. Lys292 is conserved or replaced by arginine in more than 90% of MenD sequences, and the side chain of this residue could reach over to interact with the second substrate depending upon which rotamer is present. Arg293 is not well conserved and unlikely to bind the substrates directly, but in *EcMenD*, it may simply contribute to the basic character of the active site.

The model predicts that the nonpolar component of isochorismate is placed to interact with the hydrophobic patch formed by the strictly conserved residues, Ile474B, Phe475B, and Leu478B, on one side of the active site. This isoleucine/phenylalanine pair is important for binding the cofactor as described earlier. The four conserved arginine residues could potentially interact with three carboxylate and one hydroxyl group of the substrates helping to align the reagents for catalysis. These arginine side chains all participate in hydrogen-bonding interactions that serve to position them in the active site (data not shown). Arg413 lies in the loop linking $\beta 15$ and $\alpha 16$ and is held in place by forming a hydrogen bond to Asn390, a residue moderately conserved as asparagine or serine in MenD sequences. Arg395 is located towards the end of $\alpha 15$ and interacts with the main-chain carbonyl of Ser392. We also note that an ordered chloride was observed in the SeMet *EcMenD* active site bound to the side chain of Arg395 (data not shown). In our model, the carboxylate group of the covalent intermediate occupies this anion recognition site. Arg107 is in the loop section following $\beta 4$ and interacts with the main-chain carbonyl of Asn117 while Arg33 associates with the main-chain carbonyl of Phe170 and forms a salt bridge with Glu172.

The model suggests that there are no obvious residues that could be involved in cleaving the pyruvate side chain of the isochorismate derivative, an observation that is consistent with the assignment of MenD as SEPHCHC and not as SHCHC synthase.¹⁴

Mechanistic considerations

Our current knowledge of ThDP-dependent enzymes in conjunction with the structural data now presented supports the mechanism for *EcMenD* depicted in Fig. 8. The mechanism is described in two stages corresponding to the reactions with each of the two substrates. In similar fashion to CEAS,³⁸ we note the absence of a suitable residue to function as a general acid/base and the cofactor N4' is therefore a critical component of the mechanism.

At the onset of stage I, the Glu55 ThDP N1' interaction supports tautomerization of 1',4'-iminopyrimidine to the 4'-aminopyrimidine form and promotes increased basic character of N4' (Figs. 1a and 8). This aids acquisition of the acidic proton donated from the thiazolium C2 to generate a carbanion ylide. The cationic thiazolium N3 helps to stabilize this nucleophile. Attack on α -ketoglutarate C2 occurs to generate a covalent ThDP intermediate and N4' reverts to the imino-form following donation of a proton that converts the α -ketoglutarate O5 carbonyl to a hydroxyl group. Decarboxylation occurs to produce an enamine and resonance leads to formation of an activated carbanion intermediate.

In stage II, the second substrate binds and the intermediate carbanion attacks isochorismate C2. The ThDP 4'-imino group can abstract a proton from the intermediate hydroxyl and revert to the aminoform and in so doing a carbonyl group is formed. C-C bond cleavage and the release of SEPHCHC complete the reaction. In common with the observation above concerning activation of the cofactor, we do not identify a likely proton-donating residue in the vicinity of isochorismate C1–C2. While we invoke the cofactor itself in the activation stage, for stage II of the reaction, it is possible that a water molecule participates here in the Stetter-like addition. Further data would be required to address this hypothesis.

Materials and Methods

Reagents, preparation of recombinant expression systems, and protein purification

Chemicals, of the highest quality available, were sourced from Sigma-Aldrich and VWR International except where stated otherwise. The *menD* gene was amplified by PCR from *E. coli* (K12) genomic DNA using 5'-**CATATGTCAGTCAGCGCATTTAACC**-3' and 5'-**GGATCCTCATAAATGGCTTACCTGC**-3' as the forward and reverse primers, respectively (Thermo Scientific). These oligonucleotides include 5' NdeI and 3' BamHI restriction sites respectively, indicated in bold. The PCR product was ligated into TOPO-BLUNT-II (Invitrogen) and then sub-cloned into a modified pET15b vector (Novagen), which produces a histidine-tagged protein with a tobacco etch virus protease site. The native and SeMet-labeled proteins were obtained using the methionine auxotrophic strain *E. coli* B834 (DE3) (Stratagene). Cells were grown in 1 L of Luria-Bertani media supplemented with 50 $\mu\text{g mL}^{-1}$ of carbenicillin for production of native protein, while for the SeMet protein, bacteria were cultured in minimal media (Molecular Dimensions) supplemented with SeMet following an established protocol.⁴⁴ Gene expression was induced at 20 °C using 0.5 mM isopropyl- β -D-thiogalactopyranoside, and growth continued for 16 h at room temperature.

Cells were harvested by centrifugation for 25 min at 40,000g at 4 °C, resuspended in lysis buffer (50 mM Tris-HCl, pH7.5, 250 mM NaCl, and 20 mM imidazole) containing DNase I (0.1 mg) and a single tablet of a cocktail of ethylenediaminetetraacetic-acid-free protease inhibitors (Roche), and lysed using a French press at 1000 psi. Insoluble debris was separated by centrifugation at 39,000g for 25 min at 4 °C, and the soluble fraction was loaded onto a 5-mL HiTrap Chelating HP column (GE Healthcare) precharged with Ni^{2+} . A linear concentration gradient was applied to the column and *EcMenD* was eluted at 170 mM imidazole. Fractions were analyzed using sodium dodecyl sulfate polyacrylamide gel

electrophoresis and those containing *EcMenD* were pooled. The native protein, still carrying the affinity tag, was further purified using a Superdex 200 26/60 size-exclusion column (GE Healthcare) equilibrated with 50 mM Tris-HCl and 250 mM NaCl, pH7.5. Selected fractions were pooled and dialyzed into 20 mM Tris-HCl, 200 mM NaCl, and 10% glycerol, pH8, and concentrated using Amicon Ultra devices (Millipore) to 5 mg mL⁻¹. The protein concentration was determined spectrophotometrically using a theoretical extinction coefficient of 108,775 M⁻¹ cm⁻¹ at 280 nm calculated using ProtParam.⁴⁵ Purification of the SeMet protein involved two additional steps; after elution from the HiTrap column, the hexahistidine tag was cleaved by incubation with a histidine-tagged tobacco etch virus protease for 3 h at 30 °C followed by dialysis overnight at 4 °C to remove imidazole. The sample was subjected to a second affinity chromatography step, and the flow through, containing untagged SeMet *EcMenD*, was collected and applied to the Superdex column as described for the native protein. The full incorporation of SeMet and the high level of sample purity for both samples were verified by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and sodium dodecyl sulfate polyacrylamide gel electrophoresis (data not shown). The yield of SeMet *EcMenD* (~35 mg L⁻¹ of culture) was less than that obtained for the native protein, ~100 mg L⁻¹ of culture. This is likely due to the reduction in bacterial cell mass observed under conditions required for SeMet incorporation.

Analytical gel filtration and ultracentrifugation

Analytical gel filtration experiments were conducted on a Superdex 200 26/60 column calibrated with molecular mass standards: blue dextran (>2000 kDa), thyroglobulin (669 kDa), ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29.5 kDa), ribonuclease A (13.7 kDa), and aprotinin (6.5 kDa) (GE Healthcare; data not shown). Sedimentation velocity experiments were performed (wavelength of 280 nm, rotor AN50-TI at 45,000 rpm and 20 °C), using a Beckman Coulter XL-1 analytical ultracentrifuge. Both native (tagged) and SeMet (untagged) *EcMenD* were run in 50 mM Tris-HCl, pH7.5, and 250 mM NaCl at concentrations of 0.25, 1.0, and 3.0 mg mL⁻¹ in the presence of ThDP and MnCl₂ at concentrations of 0.5 and 2 mM, respectively. Samples were centrifuged simultaneously, and *A*₂₈₀ measurements were taken at 5-min intervals for 16 h. The resultant data were analyzed using the programs SEDFIT and SEDNTERP.^{46,47}

Crystal growth, structure determination, and refinement

Crystallization—Sitting and hanging drop vapor diffusion methods were used for screening and optimization, respectively. The apoenzyme did not crystallize. Holoenzyme was prepared by incubation with 2 mM MnCl₂ and 0.5 mM ThDP, and crystals were obtained using a reservoir (500 µL) comprising 100 mM sodium citrate, pH5.6, 8% isopropanol, and 14% polyethylene glycol 6K for the native protein and 100 mM Tris-HCl, pH8.5, 22% polyethylene glycol 8K, 15% glycerol, and 200 mM NaCl for the SeMet derivative. In both cases, the best crystals were obtained from drops consisting of 1 µL protein and 2 µL of reservoir. Crystals with hexagonal rod morphology appeared after 1-2 days at room temperature. The maximum dimensions of native crystals were 0.25 mm×0.1 mm×0.1 mm, and the SeMet derivative crystals were much smaller at 0.05 mm×0.02 mm×0.02 mm.

X-ray data collection and processing—Crystals were flash cooled by plunging them into liquid nitrogen, mounted on a goniostat, and maintained at -173 °C in a flow of cooled nitrogen. Diffraction properties were characterized with a Rigaku Micromax 007 rotating anode R-Axis IV⁺ image plate system. The best samples were stored for use at the European Synchrotron Radiation Facility (Grenoble, France). Native data were collected on

beamline ID29, and SAD data were measured on beamline BM14, both with ADSC Q315 CCD detectors. The SAD data were collected at the Se absorption edge determined using X-ray Absorption Near Edge Structure spectroscopy at a wavelength of 0.97872 Å. Data processing and scaling were performed using XDS⁴⁸ and SCALA,⁴⁹ and statistics are presented in Table 1. Initial attempts to record native data at a synchrotron source were compromised by severe radiation damage; in addition, both crystal forms displayed a large unit cell c dimension (>460 Å) with mosaic spread estimated at 0.6°. We had to balance the resolution to which we could observe diffracted intensities with being able to accurately record complete and highly redundant data sets with minimum loss of data due to reflection overlaps or radiation damage. The resulting data sets are therefore artificially truncated, and this is evident from the high $\langle I/\sigma(I) \rangle$ and low R_{merge} values in the outer shells (Table 1). The high level of redundancy will, however, have contributed in a positive way to the $\langle I/\sigma(I) \rangle$ values.

Structure solution and refinement—Initial phases were obtained by SAD phasing using the positions of 23 selenium atoms (out of a possible 32) that were identified using the program SOLVE.⁵⁰ The initial figure of merit was 0.34 for data to 2.8 Å and the Z-score was 56.4. An interpretable electron density map was obtained only for space group P6₅. Density modification using RESOLVE,⁵¹ with an estimated solvent content of 50%, improved the figure of merit to 0.77. These phases produced an electron density map (calculated using the CCP4 suite of programs⁵²) that was readily interpreted (Fig. 4). Automated interpretation of the map with Buccaneer⁵³ identified a few segments of secondary structure and provided a starting point for model building in Coot.⁵⁴ Two subunits, of the four that comprise the asymmetric unit, were modeled at which point the higher-resolution native data set, with only a dimer in the asymmetric unit, became available. A monomer was used for molecular replacement into the 2.5-Å data using MOLREP.⁵⁵ Two monomers were positioned giving an *R*-factor of 43% and a score of 0.58, with the second highest solution giving a score of 0.21. This model was subsequently refined using REFMAC5⁵⁶ with NCS restraints applied during the early stages of the refinement, subsequently relaxed as the refinement progressed and then removed altogether. Refinement, including isotropic *B*-factors, was combined with manual rebuilding in Coot and the addition of water molecules and ligands. The identification of Mn²⁺ positions was confirmed by comparison of peak heights in omit difference density maps, with those observed on S of ordered methionine residues.

For the sake of completeness and to investigate possible differences in quaternary structure, we then used the refined native model to solve the SeMet *EcMenD* structure. Molecular replacement calculations positioned four molecules, with four clear peaks in the rotation function, and the refinement was completed with similar protocols to those adopted for the native structure. Model validation was carried out using MolProbity²⁶ and Coot.

The refinement statistics for both of these models are summarized in Table 1. It has been our general experience that for medium- to high-resolution macromolecular structures, the Wilson plot *B*-value is similar to the overall isotropic *B*-factors displayed by the model. In the structures reported here, we judge that the artificial truncation of data, described above, is a likely explanation for the differences (about 20 and 15 Å² for the SeMet and native forms, respectively) in these two terms.

PRODRG⁵⁷ was used to construct models and libraries for both the ThDP–oxoglutarate intermediate and isochorismate, and modeling of the *EcMenD* substrate complexes was carried out in Coot. Structural superpositions were carried out with LSQKAB.⁵⁸ Figures were prepared with Chemdraw (Adept Scientific), PyMOL[†], and ALINE (C. S. Bond, personal communication).

PDB accession numbers

Coordinates and structure factors have been deposited with accession codes 2jla and 2jlc.

Acknowledgments

This study was supported by awards from the Biotechnology and Biological Sciences Research Council (Structural Proteomics of Rational Targets, BBS/B/14434) and The Wellcome Trust (grant numbers 082596 and 083481). We thank Mark Agacan for AU measurements and the European Synchrotron Radiation Facility for beam time and excellent staff support.

Abbreviations used

AU	analytical ultracentrifugation
CEAS	<i>N</i> ² -(2-carboxyethyl)arginine synthase
CoA	coenzyme A
NCS	noncrystallographic symmetry
OXC	oxalyl-CoA decarboxylase
PDB	Protein Data Bank
POX	pyruvate oxidase
SAD	single-wavelength anomalous diffraction
SEPHCHC	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexadiene-1-carboxylate
SeMet	L-selenomethionine
SHCHC	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate
ThDP	thiamine diphosphate

References

1. Meganathan, R. Biosynthesis of the isoprenoid quinones menaquinone (vitamin K2) and ubiquinone (coenzyme Q). In: Neidhardt, FC.; Curtiss, R.; Ingraham, JL.; Lin, ECC.; Low, KB.; Magasanik, B.; Reznikoff, WS.; Riley, M.; Schaechter, M.; Umberger, HE., editors. *Escherichia coli and Salmonella: Cellular and Molecular Biology*. American Society for Microbiology; Washington, DC: 1996. p. 642-656.
2. Meganathan R. Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q): a perspective on enzymatic mechanisms. *Vitam. Horm.* 2001; 61:173–218. [PubMed: 11153266]
3. Soballe B, Poole RK. Microbial ubiquinones: multiple roles in respiration, gene regulation and oxidative stress management. *Microbiology*. 1999; 145:1817–1830. [PubMed: 10463148]
4. Bekker M, Kramer G, Hartog AF, Wagner MJ, de Koster CG, Hellingwerf KJ, de Mattos MJ. Changes in the redox state and composition of the quinone pool of *Escherichia coli* during aerobic batch-culture growth. *Microbiology*. 2007; 153:1974–1980. [PubMed: 17526854]
5. Bandyopadhyay PK. Vitamin K-dependent gamma-glutamylcarboxylation: an ancient posttranslational modification. *Vitam. Horm.* 2008; 78:157–184. [PubMed: 18374194]
6. Knaggs AR. The biosynthesis of shikimate metabolites. *Nat. Prod. Rep.* 2003; 20:119–136. [PubMed: 12636087]
7. Johnston JM, Arcus VL, Baker EN. Structure of naphthoate synthase (MenB) from *Mycobacterium tuberculosis* in both native and product-bound forms. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2005; 61:1199–1206. [PubMed: 16131752]

[†]<http://pymol.sourceforge.net>

8. Truglio JJ, Theis K, Feng Y, Gajda R, Machutta C, Tonge PJ, Kisker C. Crystal structure of *Mycobacterium tuberculosis* MenB, a key enzyme in vitamin K2 biosynthesis. *J. Biol. Chem.* 2003; 278:42352–42360. [PubMed: 12909628]
9. Ulaganathan V, Agacan MF, Buetow L, Tulloch LB, Hunter WN. Structure of *Staphylococcus aureus* 1,4-dihydroxy-2-naphthoyl-CoA synthase (MenB) in complex with acetoacetyl-CoA. *Acta Crystallogr., Sect. F: Struct. Biol. Crystallogr. Commun.* 2007; 63:908–913.
10. Parsons JF, Shi KM, Ladner JE. Structure of isochorismate synthase in complex with magnesium. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2008; 64:607–610. [PubMed: 18453696]
11. Kolappan S, Zwahlen J, Zhou R, Truglio JJ, Tonge PJ, Kisker C. Lysine 190 is the catalytic base in MenF, the menaquinone-specific isochorismate synthase from *Escherichia coli*: implications for an enzyme family. *Biochemistry.* 2007; 46:946–953. [PubMed: 17240978]
12. Jiang M, Chen X, Guo ZF, Cao Y, Chen M, Guo Z. Identification and characterization of (1R, 6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase in the menaquinone biosynthesis of *Escherichia coli*. *Biochemistry.* 2008; 47:3426–3434. [PubMed: 18284213]
13. Bhasin M, Billinsky JL, Palmer DR. Steady-state kinetics and molecular evolution of *Escherichia coli* MenD [(1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase], an anomalous thiamin diphosphate-dependent decarboxylase-carboligase. *Biochemistry.* 2003; 42:13496–13504. [PubMed: 14621995]
14. Jiang M, Cao Y, Guo ZF, Chen M, Chen X, Guo Z. Menaquinone biosynthesis in *Escherichia coli*: identification of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate as a novel intermediate and re-evaluation of MenD activity. *Biochemistry.* 2007; 46:10979–10989. [PubMed: 17760421]
15. Jiang M, Chen M, Cao Y, Yang Y, Sze KH, Chen X, Guo Z. Determination of the stereochemistry of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate, a key intermediate in menaquinone biosynthesis. *Org. Lett.* 2007; 9:4765–4767. [PubMed: 17956107]
16. Nemeria N, Chakraborty S, Baykal A, Korotchkina LG, Patel MS, Jordan F. The 1',4'-iminopyrimidine tautomer of thiamin diphosphate is poised for catalysis in asymmetric active centers on enzymes. *Proc. Natl Acad. Sci. USA.* 2007; 104:78–82. [PubMed: 17182735]
17. Nemeria N, Baykal A, Joseph E, Zhang S, Yan Y, Furey W, Jordan F. Tetrahedral intermediates in thiamine diphosphate-dependent decarboxylations exist as a 1',4'-imino tautomeric form of the coenzyme, unlike the Michaelis complex or the free coenzyme. *Biochemistry.* 2004; 43:6565–6575. [PubMed: 15157089]
18. Lindqvist Y, Schneider G, Ermler U, Sundström M. Three-dimensional structure of transketolase, a thiamine diphosphate dependent enzyme, at 2.5 Å resolution. *EMBO J.* 1992; 11:2373–2379. [PubMed: 1628611]
19. Kern D, Kern G, Neef H, Tittmann K, Killenberg-Jabs M, Wikner C, et al. How thiamine diphosphate is activated in enzymes. *Science.* 1997; 275:67–70. [PubMed: 8974393]
20. Jordan F. Current mechanistic understanding of thiamin diphosphate-dependent enzymatic reactions. *Nat. Prod. Rep.* 2003; 20:184–201. [PubMed: 12735696]
21. Frank RAW, Leeper FJ, Luisi BF. Structure, mechanism and catalytic duality of thiamine-dependent enzymes. *Cell. Mol. Life Sci.* 2007; 64:892–905. [PubMed: 17429582]
22. Stetter H. Catalyzed addition of aldehydes to activated double bonds—a new synthetic approach. *Angew. Chem., Int. Ed.* 1976; 15:639–647.
23. Sieminska EA, Macova A, Palmer DR, Sanders DA. Crystallization and preliminary X-ray analysis of (1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) synthase (MenD) from *Escherichia coli*. *Acta Crystallogr., Sect. F: Struct. Biol. Crystallogr. Commun.* 2005; 61:489–492.
24. Cruikshank DWJ. Remarks about protein structure precision. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 1999; 55:583–601. [PubMed: 10089455]
25. Matthews BW. Solvent content of protein crystals. *J. Mol. Biol.* 1968; 33:491–497. [PubMed: 5700707]
26. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 2007; 35:W375–W383. [PubMed: 17452350]

27. The UniProt Consortium. The Universal Protein Resource (UniProt). *Nucleic Acids Res.* 2008; 36:D190–D195. [PubMed: 18045787]
28. Duggleby RG. Domain relationships in thiamine diphosphate-dependent enzymes. *Acc. Chem. Res.* 2006; 39:550–557. [PubMed: 16906751]
29. Holm L, Sander C. DALI: a network tool for protein structure comparison. *Trends Biochem. Sci.* 1995; 20:478–480. [PubMed: 8578593]
30. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2004; 60:2256–2268. [PubMed: 15572779]
31. Hasson MS, Muscate A, McLeish MJ, Polovnikova LS, Gerlt JA, Kenyon GL, et al. The crystal structure of benzoylformate decarboxylase at 1.6 Å resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes. *Biochemistry.* 1998; 37:9918–9930. [PubMed: 9665697]
32. Berthold CL, Toyota CG, Moussatche P, Wood MD, Leeper F, Richards NG, Lindqvist Y. Crystallographic snapshots of oxalyl-CoA decarboxylase give insights into catalysis by nonoxidative ThDP-dependent decarboxylases. *Structure.* 2007; 15:853–861. [PubMed: 17637344]
33. Pang SS, Duggleby RG, Guddat LW. Crystal structure of yeast acetohydroxyacid synthase: a target for herbicidal inhibitors. *J. Mol. Biol.* 2002; 317:249–262. [PubMed: 11902841]
34. Muller YA, Lindqvist Y, Furey W, Schulz GE, Jordan F, Schneider G. A thiamin diphosphate binding fold revealed by comparison of the crystal structures of transketolase, pyruvate oxidase and pyruvate decarboxylase. *Structure.* 1993; 1:95–103. [PubMed: 8069629]
35. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 2007; 372:774–797. [PubMed: 17681537]
36. Mosbacher TG, Mueller M, Schulz GE. Structure and mechanism of the ThDP-dependent benzaldehyde lyase from *Pseudomonas fluorescens*. *FEBS J.* 2005; 272:6067–6076. [PubMed: 16302970]
37. Frank RAW, Titman CM, Pratap JV, Luisi BF, Perham RN. A molecular switch and proton wire synchronize the active sites in thiamine enzymes. *Science.* 2004; 306:872–876. [PubMed: 15514159]
38. Caines ME, Elkins JM, Hewitson KS, Schofield CJ. Crystal structure and mechanistic implications of N²-(2-carboxyethyl)arginine synthase, the first enzyme in the clavulanic acid biosynthesis pathway. *J. Biol. Chem.* 2004; 279:5685–5692. [PubMed: 14623876]
39. 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]
40. Pletcher J, Sax M. Crystal and molecular structure of thiamine pyrophosphate hydrochloride. *J. Am. Chem. Soc.* 1972; 94:3998–4005. [PubMed: 5037985]
41. Carlisle CH, Cook DS. The crystal structure of thiamine pyrophosphate. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* 1969; 25:1359–1367.
42. Kaplun A, Binshtein E, Vyazmensky M, Steinmetz A, Barak Z, Chipman DM, et al. Glyoxylate carboligase lacks the canonical active site glutamate of thiamine-dependent enzymes. *Nat. Chem. Biol.* 2008; 4:113–118. [PubMed: 18176558]
43. Wille G, Meyer D, Steinmetz A, Hinze E, Golbik R, Tittmann K. The catalytic cycle of a thiamine diphosphate enzyme examined by cryocrystallography. *Nat. Chem. Biol.* 2006; 2:324–328. [PubMed: 16680160]
44. Hall DR, Gourley DG, Leonard GA, Duke EBH, Anderson LA, Boxer DH, Hunter WN. The high-resolution crystal structure of the molybdate-dependent transcriptional regulator (ModE) from *Escherichia coli*; a novel combination of domain folds. *EMBO J.* 1999; 18:1435–1446. [PubMed: 10075916]
45. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, MR.; Appel, RD.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In: Walker, JM., editor. *The Proteomics Protocols Handbook*. Humana Press Inc.; Totowa, NJ: 2005. p. 571–607.

46. Schuck P. Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* 2000; 78:1606–1619. [PubMed: 10692345]
47. Lebowitz J, Lewis MS, Schuck P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci.* 2002; 11:2067–2079. [PubMed: 12192063]
48. Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* 1993; 26:795–800.
49. Evans P. Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2006; 62:72–82. [PubMed: 16369096]
50. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 1999; 55:849–861. [PubMed: 10089316]
51. Terwilliger TC. Maximum-likelihood density modification. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2000; 56:965–972. [PubMed: 10944333]
52. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 1994; 50:760–763. [PubMed: 15299374]
53. Cowtan K. The Buccaneer software for automated model building. 1 Tracing protein chains. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2006; 62:1002–1011. [PubMed: 16929101]
54. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
55. Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* 1997; 30:1022–1025.
56. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 1997; 53:240–255. [PubMed: 15299926]
57. Schüttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 2004; 60:1355–1363. [PubMed: 15272157]
58. Kabsch W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* 1976; 32:922–923.

Fig. 1.

The reactions catalyzed by MenD and MenH and tautomeric forms of the pyrimidine moiety of the cofactor. (a) MenD catalyzes the reaction of isochorismate (5S,6S)-5-[(1-carboxylatoethenyl)oxy]-6-hydroxycyclohexa-1,3-diene-1-carboxylate with α -ketoglutarate to produce SEPHCHC and carbon dioxide. MenH catalyzes the cleavage of pyruvate from SEPHCHC to release SHCHC. This reaction can also occur spontaneously. M²⁺ represents Mg²⁺ or Mn²⁺. (b) Chemical structure and numbering scheme for ThDP. Three tautomeric forms of the pyrimidine moiety are implicated in ylide generation. PP represents the cofactor diphosphate.

Fig. 2.

The primary and assigned secondary structure of *EcMenD*. α -Helices are shown as cylinders and β -strands are shown as arrows, colored according to the domain to which they have been assigned. Residues with a black background are conserved in 90% of MenD sequences listed in UniProt,²⁷ which have been filtered to remove those with 90% sequence identity level (78 sequences in total). The ThDP interacting and metal binding residues are indicated by black stars and potential substrate binding residues are denoted by yellow circles.

Fig. 3.
The secondary, tertiary, and dimer structure of *EcMenD*. (a) A ribbon diagram of the *EcMenD* subunit showing the elements of secondary structure. ThDP is depicted in stick representation colored according to atom type: C, black; N, blue; O, red; S, yellow; and P, orange. The N- and C-terminal positions are labeled, as are the elements of secondary structure using the scheme employed in Fig. 2. A black star marks the position of ThDP C2. (b) Ribbon diagram of the dimer. Subunit A is red and in the same orientation as in (a); subunit B is blue and the position of the dimer twofold (NCS) is marked. The cofactor is depicted as in (a).

Fig. 4.
The *EcMenD* tetramer. The tetramer is shown as a van der Waals surface with subunits labeled and colored differently. The dimer of subunits A and B is in the same orientation as in Fig. 3b.

Fig. 5.

The cofactor-binding site. The cofactor is shown as in Fig. 3, and C atoms of the protein residues are colored green. Mn^{2+} and a water molecule are shown as purple and red spheres, respectively. Hydrogen bonds are depicted as black broken lines, and interactions between Mn^{2+} and O ligands are shown as purple lines. In the view selected, the interaction between the cation and Gly471 is obscured. The difference density omit map (black chicken wire) for ThDP calculated with $(F_o - F_c)$ and a_c coefficients and contoured at the 3σ level. F_o represents the observed structure factors, F_c denotes the calculated structure factors, and a_c indicates the phases calculated without any contributions from ThDP atoms.

Fig. 6.

A network of interactions at the dimer interface between two cofactor-binding sites. A similar color scheme to Fig. 5 is used with the addition that C atoms of residues in subunit B are gray.

Fig. 7.

A stereoview of the model for substrates binding to *EcMenD*. The post-decarboxylation covalent adduct of ThDP and α -ketoglutarate (ThDP*) is shown, and the reactive C atoms of this intermediate and isochorismate are 1.6 Å apart. This separation is indicated by a magenta broken line. The polypeptide main chain is shown as a gray ribbon and atomic positions are colored as in the previous figures except that the C atoms of the second substrate, isochorismate, are yellow. A calculated electrostatic surface potential (blue, positive; red, negative; white, neutral) showing the active-site cleft. The figure was produced using PyMOL¹ with Adaptive Poisson–Boltzmann Solver³⁹ with electrostatic potential isocontours set at +5 kT/e (blue) and –5 kT/e (red).

Fig. 8.

A two-stage mechanism for catalysis by *EcMenD*. An asterisk marks the isochorismate C2, which is attacked by the carbanion intermediate.

Table 1Crystallographic statistics for *EcMenD*

	SeMet derivative	Native
Resolution range (Å)	83.0–2.8	92.8–2.5
Unit cell dimensions: <i>a</i> , <i>c</i> (Å)	95.83, 463.17	93.43, 465.47
Space group	<i>P</i> 6 ₅	<i>P</i> 6 ₅ 22
No. of reflections	598,083	286,462
No. of unique reflections	57,944	42,932
Redundancy	10.3	6.7
Wilson <i>B</i> (Å ²)	40.8	44.2
Completeness ^a	99.5 (96.8)	99.3 (95.4)
$\langle I/\sigma(I) \rangle$	20.6 (6.4)	18.2 (4.5)
<i>R</i> _{merge}	9.9 (26.2)	8.5 (28.4)
Protein residues (A/B/C/D)	556/556/556/556	554/556
Water molecules/Cl [−]	395/14	246/0
<i>R</i> _{work} / <i>R</i> _{free}	16.9/23.3	21.1/29.3
Average <i>B</i> -factors per subunit [overall/side chain/main chain] (Å ²)		
A	20.1/20.5/19.7	30.5/30.9/30.1
B	20.0/20.5/19.6	29.7/30.2/29.2
C	19.9/20.3/19.5	—
D	19.9/20.3/19.5	—
Waters/ThDP/Mn ²⁺ /Cl [−]	14.6/17.5/17.8/29.6	26.6/22.3/25.1/—
Ramachandran plot analysis		
Favorable	96.16	96.2
Additionally allowed	3.47	3.44
Generously allowed	0.36	0.36
DPI ^b (Å)	0.35	0.34

^aValues in parentheses refer to the highest-resolution shell; 2.88–2.8 Å (SeMet), 2.56–2.5 Å (native).

^bThe diffraction-component precision index as defined by Cruickshank.²⁴