Structure of 4-diphosphocytidyl-2-C-methylerythritol synthetase involved in mevalonate-independent isoprenoid biosynthesis

Stéphane B. Richard[1], Marianne E. Bowman[1], Witek Kwiatkowski[1], Ilgu Kang[2], Cathy Chow[2], Antonietta M. Lillo[2], David E. Cane[2] and Joseph P. Noel[1]

The YgbP protein of *Escherichia coli* encodes the enzyme 4-diphosphocytidyl-2-C-methylerythritol (CDP-ME) synthetase, a member of the cytidyltransferase family of enzymes. CDP-ME is an intermediate in the mevalonate-independent pathway for isoprenoid biosynthesis in a number of prokaryotic organisms, algae, the plant plastids and the malaria parasite. Because vertebrates synthesize isoprenoid precursors using a mevalonate pathway, CDP-ME synthetase and other enzymes of the mevalonate-independent pathway for isoprenoid production represent attractive targets for the structure-based design of selective antibacterial, herbicidal and antimalarial drugs. The high-resolution structures of *E. coli* CDP-ME synthetase in the apo form and complexed with both CTP–Mg^{2+} and CDP-ME–Mg^{2+} reveal the stereochemical principles underlying both substrate and product recognition as well as catalysis in CDP-ME synthetase. Moreover, these complexes represent the first experimental structures for any cytidyltransferase with both substrates and products bound.

Isopentenyl diphosphate (IPP) and the isomeric compound, dimethylallyl diphosphate (DMAPP), are the fundamental building blocks of isoprenoids in all organisms. The isoprenoids include >23,000 naturally occurring molecules of both primary and secondary metabolism. Isoprenoids include hopane triterpenes, ubiquinones and menaquinones in bacteria; carotenoids, plastoquinones, mono-, sesqui-, di- and triterpenes and the prenyl side chains of chlorophylls in plants; and quinones, dolichols, steroids and retinoids in mammals[1]. IPP was generally assumed to be derived solely from mevalonate synthesized from the condensation of three molecules of acetyl-CoA. However, independent studies demonstrated the existence of a novel, mevalonate-independent pathway for IPP synthesis known as the 1-deoxy-D-xylulose 5-phosphate / 2-C-methyl-D-erythritol 4-phosphate (DXP/MEP) pathway. This mevalonate-independent pathway utilizes pyruvate and glyceraldehyde 3-phosphate as starting materials for production of IPP[2] (Fig. 1).

The DXP/MEP pathway occurs in a variety of eubacteria, including several pathogenic species such as *Mycobacterium tuberculosis*, in algae, in the plastids of plant cells[2] and in the apicoplast of *Plasmodium falciparum*[3], the parasite that causes malaria. Given the essential nature of the DXP/MEP pathway in these organisms and the absence of this pathway in mammals, the enzymes comprising the DXP/MEP pathway represent potential targets for the generation of selective antibacterial[4], antimalarial[3] and herbicidal[5] molecules. 4-diphosphocytidyl-2-C-methylerythritol (CDP-ME) synthetase[6,7] (alternatively known as MEP cytidyltransferase) encoded by the *ygbP* open reading frame of *Escherichia coli* catalyzes the formation of CDP-ME from MEP and CTP. This reaction constitutes the third step of the DXP/MEP pathway of isoprenoid biosynthesis. We now report the three-dimensional structures of CDP-ME synthetase both in the apo form and complexed to CTP–Mg^{2+} and CDP-ME–Mg^{2+}, respectively. These structures reveal active site features responsible for the stereochemical control of the cytidyltransferase reaction and serve as three-dimensional templates for future efforts directed at structure-based inhibitor design and evaluation.

Overall three-dimensional architecture

CDP-ME synthetase organizes as a homodimer with each subunit related by a crystallographic two-fold axis. We have refined the 'apo' form of CDP-ME synthetase to 1.55 Å resolution, the complex with CTP-Mg^{2+} to 1.5 Å resolution (Fig. 2*a*) and the complex with CDP-ME–Mg^{2+} to 1.8 Å resolution (Fig. 2*b*). The 'apo' form of CDP-ME synthetase was crystallized in the presence of 10 mM MEP but showed no extra electron density in the active site after structural elucidation. Each subunit of the homodimer is comprised of two structurally distinct domains. The larger core domain (residues 1–136 and 160–236) is globular in shape and maintains an α/β structure that resembles a Rossmann fold[8] but which displays a distinct α/β connectivity pattern with an insertion of two β-strands, β7 and β10, into a parallel 3-2-1-4-5 β-sheet. The second, much smaller lobe or subdomain (residues 137–159) resembles a curved arm that interlocks in *trans* with its symmetry related arm to mediate dimer formation. Moreover, the interlocking arms form part of the MEP binding site and organize portions of the catalytic surface responsible for cytidyltransferase activity (Fig. 2*b*).

Relationship to other cytidyltransferases

A search for related three-dimensional structures in the Protein Data Bank[9] using the DALI[10] server retrieved a number of

[1]Structural Biology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. [2]Department of Chemistry, Brown University, Providence, Rhode Island 02912, USA.

Correspondence and requests for materials should be addressed to J.P.N. *email: noel@sbl.salk.edu*

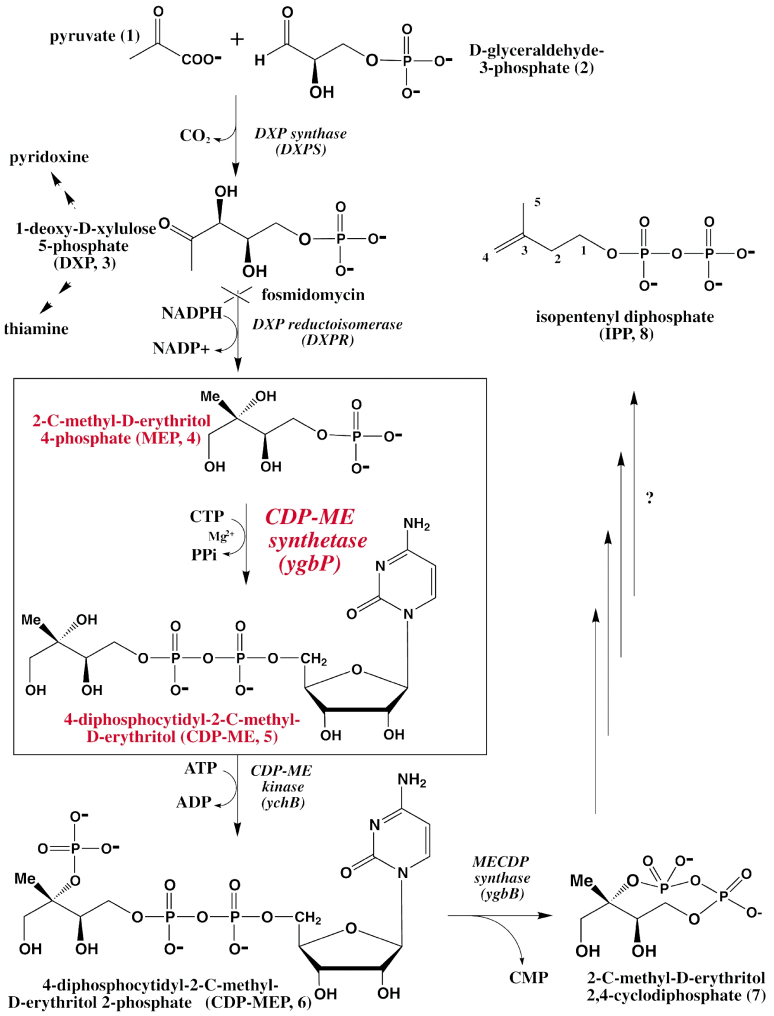


Fig. 1 Biosynthesis of the isoprenoid precursor isopentenyl diphosphate (IPP, **8**) *via* the alternative, mevalonate-independent DXP/MEP pathway. All reactions are depicted in the forward direction. The synthesis of the C5 IPP skeleton begins with the condensation of a C2 moiety from the decarboxylation of pyruvate (**1**) and a C3 moiety from glyceraldehyde 3-phosphate (**2**) to form 1-deoxy-D-xylulose 5-phosphate (DXP, **3**) through the action of DXP synthase[32,33] (DXPS or DXS). Next, DXP is converted into 2-C-methyl-D-erythritol 4-phosphate (MEP, **4**) by DXP reductoisomerase[34,35] (DXPR or DXR) and subsequently transformed into 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME, **5**) by CDP-ME synthetase[6,7] (YgbP protein). Fosmidomycin acts as an effective inhibitor of DXPR[3]. CDP-ME is phosphorylated on the 2-hydroxy group to form 4-diphosphocytidyl-2-C-methyl-D-erythritol 2-phosphate (CDP-MEP; **6**) in an ATP-dependent reaction by the enzyme CDP-ME kinase encoded by the *ychB* gene of *E. coli*[36,37]. Subsequent formation of 2-C-methyl-D-erythritol 2,4-cyclodiphosphate (**7**) is catalyzed by the enzyme MECDP synthase encoded by the gene *ygbB*[17,38]. Additional steps, which remain to be elucidated, ultimately form IPP (**8**).

enzyme cores containing a nucleotide binding fold. Currently, three cytidyltransferase structures, which contain a core nucleotide binding domain, have been described and include capsule specific CMP:2-keto-3-deoxy-*manno*-octonic acid synthetase[11] (K-CKS), CTP:glycerol-3-phosphate cytidyltransferase[12] (GCT) and, most recently, CMP acylneuraminate synthetase[13] (CMP-NeuAc synthetase).

CDP-ME synthetase, CMP-NeuAc synthetase and GCT share similarly folded cores. However, the five parallel β-strands of GCT, which maintain a 3-2-1-4-5 topology, are interrupted in CDP-ME synthetase and CMP-NeuAc synthetase by the insertion of two antiparallel β-strands (β7 and β10) between β-strands 4 and 5 (β6 and β11) of the core β-sheet (Fig. 2*c*). This insertion extends the central β-sheet and leads to a structural alteration of the nucleotide binding region and subsequent formation of a distinct and spatially nonoverlapping CTP binding motif in CDP-ME synthetase and CMP-NeuAc synthetase. The coordinates for K-CKS have yet to be deposited in the PDB, making a detailed comparison of K-CKS with either CMP-NeuAc synthetase or CDP-ME synthetase impossible at this stage.

By superimposing GCT and CDP-ME synthetase to constrain the connectivity pattern of the central β-sheets and the surrounding α-helices, the GCT and CDP-ME synthetase active sites were shown to reside on the same C-terminal side of the central β-sheets, but the orientations of the bound CTP molecules in each enzyme are arranged in opposite fashions. In contrast, alignments of the two structures using the bound CTP molecules as a guide reveals an internal pseudosymmetry in the core α/β fold of both GCT and CDP-ME synthetase (Fig. 2*c*).

Although the cytidine base and ribose exhibit nearly identical orientations in GCT and CDP-ME synthetase, the curved triphosphate ends of the bound CTP molecules pucker in opposite directions. These alternative orientations of the triphosphate tail of CTP position each of the α-phosphates of the CTP substrates for nucleophilic attack by MEP in CDP-ME synthetase or glycerol 3-phosphate in GCT. The structure of CMP-NeuAc synthetase was solved in the apo form and complexed to a partially disordered CDP molecule, which serves as a mimic of the true CTP substrate. Structural alignments of *E. coli* CDP-ME synthetase or *Neisseria meningitidis* CMP-NeuAc synthetase[13] (data not shown) with GCT reveal that the catalytic machinery of CDP-ME synthetase and CMP-NeuAc synthetase are spatially conserved with GCT but not shared at the primary amino acid level with GCT[14]. The spatial correspondence includes the putative catalytic residues His 14 and His 17 of GCT[15] with Lys 27 and Lys 213 of CDP-ME synthetase and Lys 21 and Asp 209 of CMP-NeuAc synthetase[13].

Active site architecture

CTP and CDP-ME are sequestered in the active site of CDP-ME synthetase by a glycine-rich loop spanning Pro 13–Arg 20. Selectivity for the pyrimidine base is achieved through hydrogen bonding interactions and steric constrictions in the base-binding pocket that do not allow for the sequestration of larger purine bases. This selectivity is specifically achieved through hydrogen bonds formed between the backbone amides of Ala 14 and Ala 15, the carbonyl oxygens of Gly 82 and Asp 83, and the hydroxyl group of Ser 88. The cytosine base is stacked between the flexible loop spanning β1 and β2, and the methylene portion of the Arg 85 side chain projecting outward from the β5–αE catalytic loop. The 2′ and 3′ hydroxyl groups of the ribose moiety participate in hydrogen bonds with the backbone carbonyl oxygen of Pro 13 and the backbone amides of Gly 16 and Ala 107 (Fig. 3*a*).

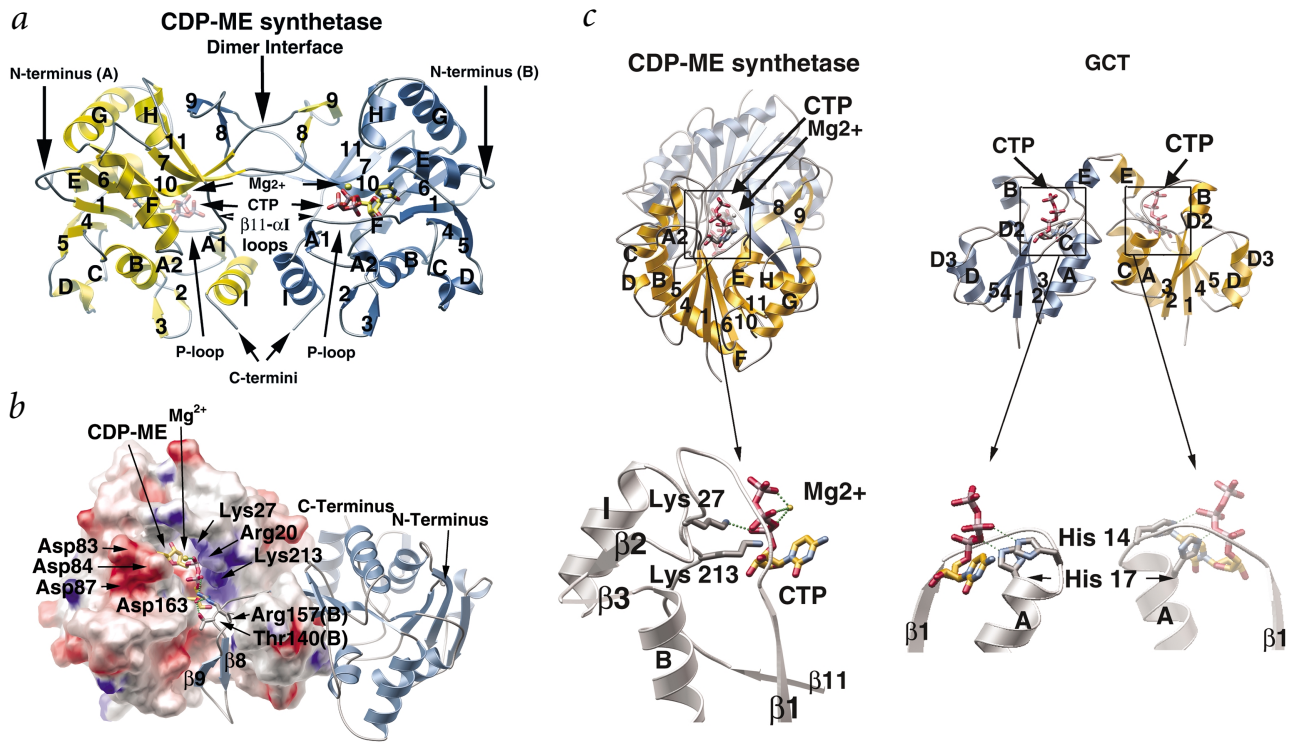


Fig. 2 Overall architecture of *E. coli* CDP-ME synthetase. ***a***, Ribbon representation of the CDP-ME synthetase homodimer complexed with CTP–Mg^{2+}. Subunit A is shown in gold and subunit B in blue. The CTP molecule and Mg^{2+} ion are represented as rendered stick and ball models, respectively. The catalytic loop surrounding the CTP binding site is indicated. The secondary structure is annotated according to the cytidyltransferases nomenclature based on the CMP-NeuAc synthetase structure[13] aligned with CDP-ME synthetase. ***b***, A combined surface and ribbon view of the CDP-ME–Mg^{2+} complex highlighting the extended product-binding pocket. The molecular surface representation of subunit A (left) was calculated with GRASP[31], with acidic and basic residues colored red and blue, respectively. Subunit B (right) is colored blue and represented as a ribbon, with the side chains of Thr 140 and Arg 157 shown as rendered sticks color coded by atom type. Dotted green lines depict hydrogen bonds. This orientation is derived from the view depicted in (*a*) after a 180° rotation around the horizontal axis. ***c***, Contrasting CTP binding modes in GCT[12] and CDP-ME synthetase. The gold colored subunits of each structure maintain the same orientation of secondary structural elements and opposing orientations for the CTP molecules. In contrast, the gold-colored subunit of CDP-ME synthetase and the blue-colored subunit of GCT exhibit identical orientations for their respective CTP molecules. The overall orientation of CDP-ME synthetase depicted is derived from the view shown in (*a*) following a 90° clockwise rotation in the plane of the figure and a 45° rotation around the horizontal axis.

The main chain atoms of each of the CDP-ME synthetase subunits superpose with a root mean square (r.m.s) deviation of 0.264 Å, 0.773 Å, and 0.754 Å between the apo and CTP–Mg^{2+}, the apo and CDP-ME–Mg^{2+}, and CTP–Mg^{2+} and CDP-ME–Mg^{2+} bound forms, respectively. The largest backbone differences occur in the loop linking β1 and β2, which links the 3_{10}-helices, A1 and A2. In CDP-ME synthetase, the P-loop comprising residues 17–25 is not defined in the apo form, is poorly defined in the CDP-ME–Mg^{2+} complex and is ordered in the CTP–Mg^{2+} complex.

The side chain amino group of Lys 27 is in direct contact with the α-phosphates of both CTP and CDP-ME. Arg 20, through both its backbone amide and its side chain δ-guanido moiety, forms a series of hydrogen bonds with the α- and γ-phosphate oxygens of CTP (Fig. 3*a*,*b*). Arg 20 and Lys 27 may play important roles in transition state stabilization during CDP-ME formation because of their positions near the α-phosphate of CTP, their absolute conservation in both functionally-characterized and putative CDP-ME synthetases from bacteria and plants (data not shown), and the potential accumulation of negative charges in the pentacoordinate transition state during CDP-ME formation.

Mg^{2+}, which is essential for cytidyltransferase activity[6], forms coordination bonds with the α-, β- and γ-phosphate oxygens of both CTP (Fig. 3*a*,*d*) and the α-phosphate oxygen of CDP-ME (Fig. 3*c*,*e*). There are no coordination bonds that occur between CDP-ME synthetase and Mg^{2+} in any of the complexes examined to date. Lys 213 and Arg 157 from the dyad-related subunit do not participate in electrostatic hydrogen bonds with CTP but both are in direct contact with the MEP derived portions of CDP-ME (Fig. 3*a*–*c*).

Catalytic mechanism

Theoretically, two catalytic mechanisms are possible in cytidyltransferases. The first, a dissociative mechanism, includes the transient formation of a reactive metaphosphate intermediate at the α-phosphate position of CTP upon loss of pyrophosphate. Subsequent capture of the reactive metaphosphate intermediate by the 4-phosphate group of MEP would form CDP-ME. The second, associative mechanism uses a negatively charged pentacoordinate transition state upon nucleophilic attack on the α-phosphate of CTP by the 4-phosphate of MEP. Collapse of this charged state would lead to pyrophosphate release and CDP-ME formation. Although the determination of the contribution of each pathway to the CDP-ME synthetase reaction mechanism is beyond the scope of the present study, the substrate and product bound structures that include Mg^{2+} ions provide useful models to assess the roles of specific residues in the catalytic mechanism of CDP-ME synthetase.

The catalytic roles of Lys 27 and Lys 213 have been examined. Lys 27 has been mutated to both Ala and Ser, and Lys 213 has

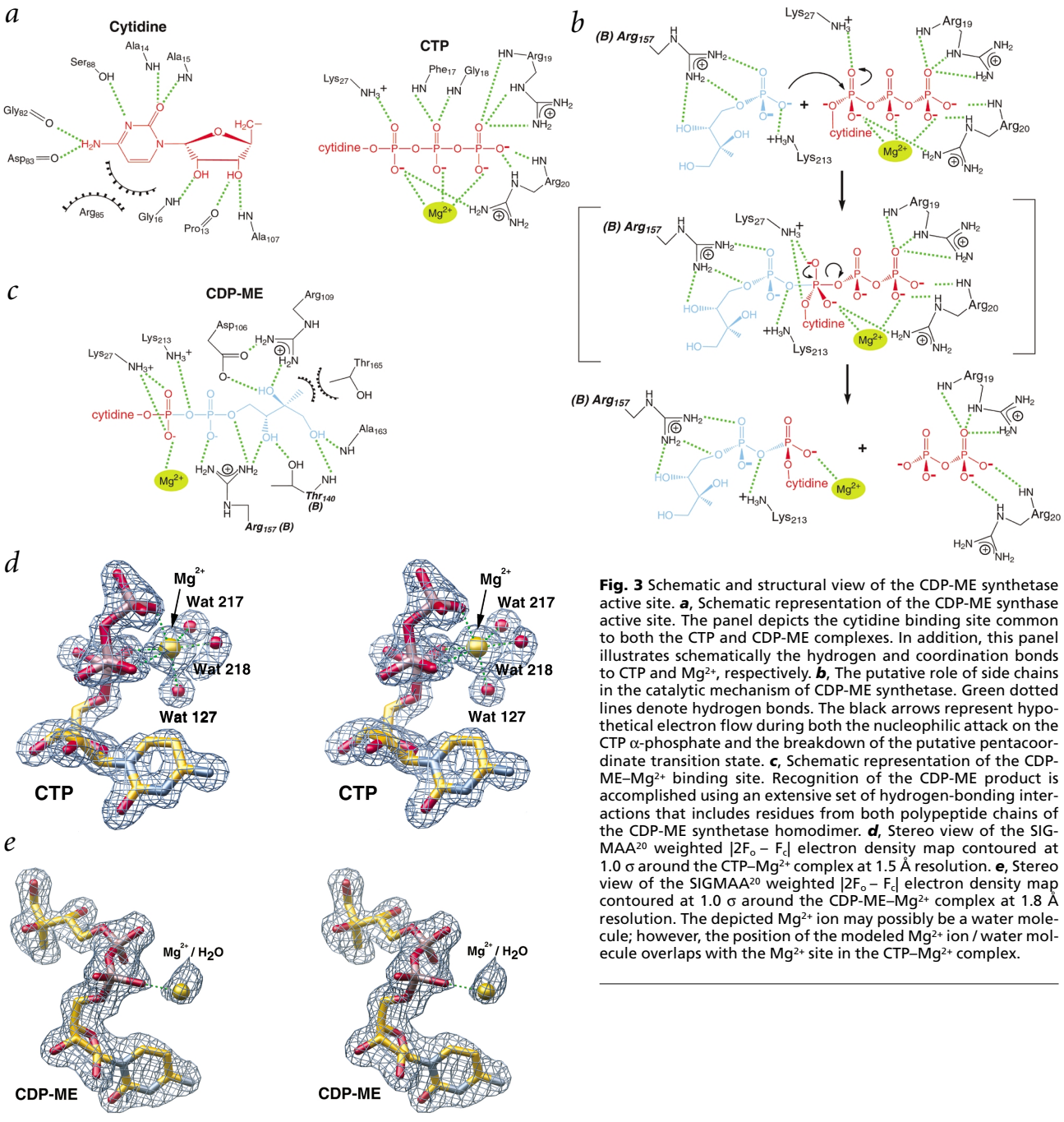


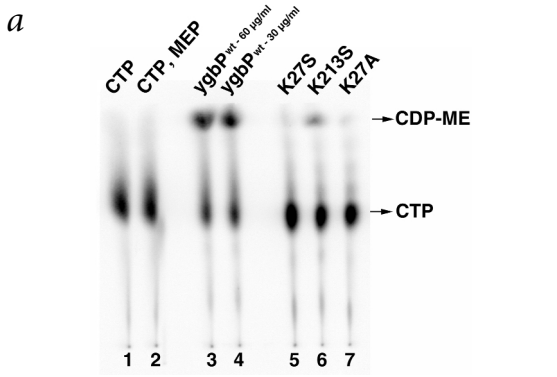
Fig. 3 Schematic and structural view of the CDP-ME synthetase active site. ***a***, Schematic representation of the CDP-ME synthase active site. The panel depicts the cytidine binding site common to both the CTP and CDP-ME complexes. In addition, this panel illustrates schematically the hydrogen and coordination bonds to CTP and Mg^{2+}, respectively. ***b***, The putative role of side chains in the catalytic mechanism of CDP-ME synthetase. Green dotted lines denote hydrogen bonds. The black arrows represent hypothetical electron flow during both the nucleophilic attack on the CTP α-phosphate and the breakdown of the putative pentacoordinate transition state. ***c***, Schematic representation of the CDP-ME–Mg^{2+} binding site. Recognition of the CDP-ME product is accomplished using an extensive set of hydrogen-bonding interactions that includes residues from both polypeptide chains of the CDP-ME synthetase homodimer. ***d***, Stereo view of the SIGMAA[20] weighted $|2F_o - F_c|$ electron density map contoured at 1.0 σ around the CTP–Mg^{2+} complex at 1.5 Å resolution. ***e***, Stereo view of the SIGMAA[20] weighted $|2F_o - F_c|$ electron density map contoured at 1.0 σ around the CDP-ME–Mg^{2+} complex at 1.8 Å resolution. The depicted Mg^{2+} ion may possibly be a water molecule; however, the position of the modeled Mg^{2+} ion / water molecule overlaps with the Mg^{2+} site in the CTP–Mg^{2+} complex.

been mutated to Ser. Assays conducted on these mutants and compared quantatively to the wild type protein indicate that Lys 27 plays an essential role in catalysis because the effective turnover rate is ~0.025% of the rate of the wild type enzyme. The activity of the K213S mutant is compromised resulting in an effective turnover rate that is ~0.4% of the rate of wild type enzyme. Nevertheless, the K213S mutant retains the capacity to form the CDP-ME product, although with significantly reduced efficiency (Fig. 4*a*). A more quantitative kinetic and structural analysis of these and other active site mutants is currently in progress.

In summary, Arg 20 interacts with the α- and γ-phosphates of CTP, possibly positioning the α-phosphate for nucleophilic attack. Arg 20, Lys 27, and Mg^{2+} may serve as complementary charges for the negatively charged pentacoordinate transition state during CMP transfer. Arg 157 positioned on the dyad-related subunit most likely positions the attacking nucleophile near the α-phosphate of CTP in CDP-ME synthetase. Finally, Lys 213 may act as an electrostatic guide for the MEP phosphate prior to nucleophilic attack on CTP (Figs 3*b*, 4*a*,*b*).

Conclusion

The three-dimensional structures of *E. coli* CDP-ME synthetase complexed with both substrate and product provide the first three-dimensional view of intermediate formation in the DXP/MEP pathway of isoprenoid biosynthesis. Additional functional studies can be undertaken using the three-dimensional structures presented in this study as a guide to better understand



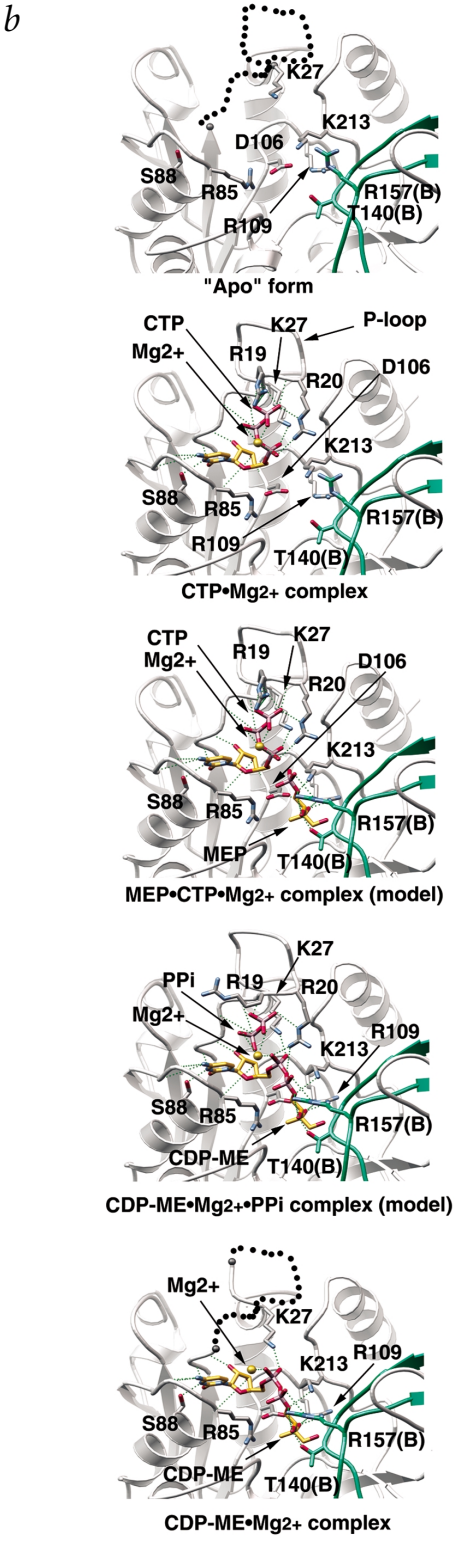


Fig. 4 Catalytic mechanism of CDP-ME synthetase. ***a***, Thin-layer chromatography assays of CDP-ME synthetase wild type and the K27A, K27S and K213S mutants. Lanes 1 and 2 are controls with no enzyme added and either CTP added or both CTP and MEP added, respectively. Lanes 3 and 4 contain, in addition to CTP–Mg^{2+} and MEP, CDP-ME synthetase at a concentration of 60 μg ml^{-1} (lane 3) or 30 μg ml^{-1} (lane 4). Lanes 5, 6, and 7 included 30 μg ml^{-1} of either the K27S mutant (lane 5), the K213S mutant (lane 6), or the K27A mutant (lane 7), respectively. The quantity of CDP-ME produced by the K27S, K213S and K27A mutants are 1.5%, 25% and 1.5%, respectively, normalized to 100% for wild type CDP-ME synthetase formed over 2 hr. The reaction with wild type protein is complete within 1–2 min (not shown). Extrapolating the 2 hr time period back to 2 min indicates a decrease in the rate of turnover for the K27S, K213S, and K27A mutants of factors of 0.00025, 0.004, and 0.00025, respectively. ***b***, Rendered view of the ground state complexes of CDP-ME synthetase. The model for the MEP–CTP–Mg^{2+} complex is based on the observed positions of CTP and the MEP-derived portion of CDP-ME. The model for the product complex that includes diphosphate (PP_i) is based on the observed position of the β- and γ-phosphates of CTP. The backbone of the dyad related monomer is colored green, and hydrogen bonds are shown as green dashed lines. These closeup views are shown in an orientation identical to that depicted in Fig. 2*b*.

the role of the active site residues in CDP-ME synthetase. Finally, mutants that critically impair catalytic activity can be used to structurally characterize additional steps in the catalytic mechanism of CDP-ME synthetase. All of the enzymes comprising this pathway are potentially targets for the design of novel drugs directed against pathogenic bacteria and the malaria parasite[5]. The structures of several CDP-ME synthetase complexes are the first steps in efforts to use structural templates for the synthesis of effective inhibitors of the DXP/MEP pathway for isoprenoid biosynthesis.

Methods

Expression, mutagenesis and purification. The *E. coli ygbP* gene (GenBank accession number AP002562) encoding CDP-ME synthetase was PCR amplified from *E. coli* K12 and inserted into the pHIS8 expression vector. The CDP-ME synthetase K27S, K27A and K213S mutants were generated with the QuickChange (Stratagene) PCR method. N-terminal His_8-tagged protein was expressed in BL21(DE3) cells. Tagged CDP-ME synthetase was purified from sonicates using a Ni^{2+}-NTA (Qiagen) column. The histidine tag was removed using thrombin, and the cleaved synthetase was purified to greater then 99% homogeneity by gel filtration on a Superdex-S200 (Pharmacia) FPLC column. CDP-ME synthetase containing fractions were combined and concentrated to 35 mg ml^{-1}.

Synthesis of methyl-D-erythritol-4-phosphate (MEP). MEP was prepared on a 2 mmol scale using a one-pot, coupled enzymatic synthesis with malate as source of pyruvate and fructose-1,6-diphosphate as the source of glyceraldehyde-3-phosphate, and a NADPH recycling system using malic enzyme and deoxyxylulose-5-phosphate reductoisomerase (DXPR). The reaction mixture contained: fructose 1,6-diphosphate (10.0 mmol), malate (20.0 mmol), pyruvate (2 mmol), $NADP^+$ (1.24 mmol), $MgCl_2$ (0.5 mmol), Tris-Cl at pH 7.5 (40 mmol), DTT (0.1 mmol), TPP (0.06 mmol), aldolase (570 units), isomerase (9700 units), malic enzyme (50 units), deoxyxylulose phosphate synthetase (DXPS2, ~50 units) and DXPR (~50 units), both from *Streptomyces coelicolor*[16], in total volume of 100 ml. This solution was incubated for 3 d at 30 °C.

Table 1 Crystallographic data used for phasing

	Native	SeMet,λ1	SeMet,λ2	SeMet,λ3	DMA	$KAu(CN)_4$
Resolution (Å)	90–1.2	90–1.35	90–1.35	90–1.35	90–1.65	90–1.75
Unique reflections[1,2]	59,518 (1,539)	94,279 (4,556)	93,977 (4,277)	94,976 (4,529)	51,709 (1,714)	41,116 (1,502)
Redundancy[1,2]	2 (2.2)	2.4 (2.4)	2.5 (2.4)	2.5 (2.4)	3.3 (2.1)	2.2 (1.7)
Completeness[1,2] (%)	86.4 (44.9)	94.5 (91.4)	94.2 (85.8)	95.2 (90.3)	94.6 (64.3)	91.7 (66.3)
$I / \sigma I$[1]	22.5 (2.3)	20.1 (1.8)	20.8 (1.8)	19.8 (1.5)	33.5 (2.9)	27.4 (2.5)
R_{sym}[1,3] (%)	3.5 (37.9)	2.7 (35.6)	2.6 (35.8)	2.9 (43.2)	3.4 (26.2)	3.1 (22.5)
Number of sites					4	5
Phasing power[4]						
Centric iso					0.626	1.466
Acentric iso					1.384	1.066
Acentric ano					1.184	0.806
R_{cullis}[5]						
Centric iso					0.846	0.911
Acentric iso					0.815	0.917
Acentric ano					0.715	0.831

[1]Number in parentheses is for highest resolution shell.
[2]For the SeMet, dimercurial acetate (DMA) and $KAu(CN)_4$ data sets, F^+ and F^- were considered nonequivalent when calculating the number of unique reflections and completeness.
[3]$R_{sym} = \Sigma_h |I_h - \langle I_h \rangle| / \Sigma_h (I_h)$, where $\langle I_h \rangle$ is the average intensity over symmetry equivalent reflections.
[4]Phasing power = $\langle |F_{H(calc)}| / |E| \rangle$, where $F_{H(calc)}$ is the calculated heavy-atom structure factor and E is the estimated lack-of-closure error. Iso is isomorphous and ano is anomalous.
[5]$R_{cullis} = \Sigma |E| / \Sigma |F_{PH} - F_P|$.

The crude reaction mixture was cleared with activated charcoal, depleted of Tris buffer using a cation exchange column (DOWEX 50W-X8, 2.5 × 50 cm^2) in the H^+ form, and further fractionated on a cellulose column (2.5 × 50 cm^2) equilibrated and eluted with an acetonitrile:water:TFA mixture (90:10:1). Final purification of the MEP containing fractions was achieved by loading onto an anion exchange column (DOWEX 1X8-100, formate form) and then washing the column with water. Finally, MEP was eluted with a 1:1 mixture of ammonium formate (1 M) and formic acid. The MEP-containing fractions were combined and lyophilized. Analysis by ^{1}H-NMR showed the yield of MEP to be 11.0 mmol (55% yield based on the fructose-1,6-diphosphate starting material), with a purity of >85 %.

Synthesis of 4-diphosphocytidyl-2-C-methylerythritol (CDP-ME). CDP-ME was prepared on a 0.1 mmol scale. The reaction mixture contained 104 μmol MEP (purified as described above), CTP (104 μmol), $MgCl_2$ (16.8 μmol), Tris-HCl at pH 7.8 (168 μmol), NaOH (120 μmol), 1.3 mg of *E. coli* CDP-ME synthetase and inorganic phosphatase (20 units) in a total volume of 1.68 ml. This solution was incubated for 3 h at 37 °C. The crude reaction mixture was deproteinized (500 NMWL filter, Millipore), lyophilized and dissolved in an aqueous solution containing 40% (v/v) methanol and 0.1 M ammonium formate. Aliquots were fractionated on an ion exchange analytical HPLC column (Nucleosil SB100-10 4.6 × 250 mm^2, M. Nagel) and eluted with the same solution[17]. The peak fractions corresponding to CDP-ME were concentrated and freed of formate by repeated lyophilizations. CDP-ME was obtained in 30% overall yield and >85 % purity as determined by ^{1}H NMR analysis.

Enzyme assays. Assay mixtures at 25 °C contained 30 μg ml^{-1} CDP-ME synthetase, 600 μM CTP, 600 μM MEP, 5 mM $MgCl_2$ and 0.2 μCi ml^{-1} [α-^{32}P] CTP (400 Ci $mmol^{-1}$) in 0.1 M Tris-HCl, pH 8.0, and 2 mM DTT in a final volume of 25 μl. Reactions were initiated by adding CDP-ME synthetase at a stock concentration of 2.25 mg ml^{-1}. After incubation for 2 h, the samples were boiled for 5 min and aliquots were spotted on polyethyleneimine (PEI)-cellulose TLC plates. Ascending chromatography was accomplished at 4 °C in rectangular glass tanks containing 100 ml of 0.8 M $(NH_4)_2SO_4$ to within 1 cm of the top of the TLC plate. The plate was dried, exposed for 8 h with an imaging plate and scanned with a PhosphorImager (Molecular Dynamics).

Crystallization. Crystals of CDP-ME synthetase (800 μm × 600 μm × 200 μm) were obtained by the vapor diffusion method at 4 °C in 2 μl hanging drops containing a 1:1 mixture of a 35 mg ml^{-1} protein solution with crystallization buffer (10% (w/v) PEG 8000, 0.2 M calcium acetate, 2 mM DTT and 0.1 M PIPES, pH 6.5) containing up to 19% (v/v) ethylene glycol. CDP-ME synthetase crystals belong to space group C2 with unit cell dimensions of a = 130.6 Å, b = 47.1 Å,

Table 2 Crystallographic data and refinement statistics

	CTP–Mg^{2+}	Apo	CDP-ME–Mg^{2+}
Resolution (Å)	90–1.5	90–1.55	90–1.8
Unique reflections	34,216 (1,353)	30,652 (880)	20,268 (641)
Redundancy	4.7 (3.9)	3.2 (2.1)	2.8 (2.9)
Completeness (%)	92.2 (72.6)	90.3 (51.8)	98.6 (98.6)
I / σ	25.6 (1.9)	32.7 (6)	26.4 (2.7)
R_{sym}[1] (%)	4.5 (34.5)	2.6 (13.4)	2.6 (30.8)
R_{cryst}[2] / R_{free}[3] (%)	22.7 / 24.9	24.6 / 26.8	22.1 / 27.4
Missing residues	1–4, 230–236	1–4, 16–26, 229–236	1–3, 229–236
Protein atoms	1,713	1,625	1,717
Water molecules	328	370	103
Ions bound	1 Mg^{2+}, 1 Ca^{2+}	1 Ca^{2+}	1 Mg^{2+}
Ligand atoms[4]	29	0	33
R.m.s. deviation			
Bonds (Å)	0.0099	0.0051	0.0089
Angles (°)	1.592	1.23	1.35
Average B-factor (Å^2)			
Protein	24.8	25.5	35.4
Water	38.2	40.3	45.7
Ligand	31.1		30.5

[1]$R_{sym} = \Sigma_h |I_h - \langle I_h \rangle| / \Sigma_h I_h$, where $\langle I_h \rangle$ is the average intensity over symmetry equivalent reflections.
[2]$R_{cryst} = \Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$, where summation is over the data used for refinement.
[3]R_{free} factor is R_{cryst} calculated using 5% of data (test set) excluded from refinement.
[4]Ligand atoms refer to a CTP molecule in the CTP–Mg^{2+} complex and to a CDP-ME molecule in the CDP-ME–Mg^{2+} complex.

c = 38.1 Å, β = 94°, with one monomer per asymmetric unit, and a solvent content of 42%. Complexes of CDP-ME synthetase with CTP or CDP-ME were obtained as above by crystallization in the presence of 10 mM CTP and 10 mM $MgCl_2$, and 10 mM CDP-ME and 10 mM $MgCl_2$, respectively.

Data collection and processing. A data set was collected on beamline 9.2 (λ = 0.9848 Å) of the Stanford Synchrotron Radiation Laboratory (SSRL) equipped with an ADSC Quantum 4 CCD detector on a single crystal cocrystallized with 1 mM CTP, to a resolution of 1.2 Å. A MAD data set was collected on beamline 9-2 (selenium edge with λ_1 = 0.9797 Å, λ_2 = 0.9795 Å, λ_3 = 0.8952 Å) at SSRL on a single crystal grown from Se-met containing CDP-ME synthetase cocrystallized with 1 mM CTP, to a resolution of 1.35 Å (Table 1). Diffraction data were collected on beamline 7-1 (λ = 1.08 Å) at SSRL on a single crystal cocrystallized with 10 mM CTP to a resolution of 1.5 Å, on a crystal cocrystallized with 10 mM MEP (referred to as the 'apo' form) to a resolution of 1.55 Å, and on beamline 9-1 (λ = 0.773 Å) on a crystal cocrystallized with 10 mM CDP-ME to a resolution of 1.8 Å; all data were collected on a 180/345 mm MAR imaging plate system detector (Table 2). Heavy atom derivatives were obtained by soaking native crystals for 12–16 h in the presence of either 13 mM $KAu(CN)_4$ or a saturated solution of dimercurial acetate (DMA). MIR data sets were collected in-house at 100 K using a DIP 2030 imaging plate system (Mac Science Corporation). CuKα radiation was produced by a rotating anode operated at 45 kV and 100 mA that was equipped with double-focusing Pt/Ni coated mirrors (Table 1). All data were indexed and integrated using DENZO[18] and scaled with the program SCALEPACK[18]. Intensities were transformed into amplitudes using TRUNCATE[19]. The heavy atom derivative data sets were scaled against the 1.2 Å resolution native data set with the program SCALEIT[20].

Structure determination. MAD experiments attempting to solve the structure of Se-met containing CDP-ME synthetase yielded two potential Se sites and, after MAD phasing, low quality experimental maps that did not permit chain tracing. However, the phases obtained from SHARP[21] were sufficient to identify initial metal-binding sites for the Au and Hg derivatives by difference Fourier analysis. These initial sites were verified by inspection of difference Patterson maps using XTALVIEW[22] and initially refined using MLPHARE[23]. Final refinement of heavy atom parameters, identification of minor heavy atom binding sites and phase-angle calculations were performed with the program SHARP[21] using the 1.2 Å resolution native data set. Solvent flipping using the CCP4 program SOLOMON[20] significantly improved and extended phases to 1.2 Å.

Model building and refinement. The initial atomic model was generated using wARP[24]. Subsequent model building was carried out with the program O[25], and refinement steps consisting of bulk-solvent correction, positional and torsion angle-simulated annealing and B-factor refinement were carried out with CNS[26]. However, due to the low occupancy of the CTP binding site and the corresponding partial disorder surrounding this site, further refinement using the 1.2 Å data set was halted. Subsequent building and refinement steps were carried out using data obtained from a crystal cocrystallized with 10 mM CTP and 10 mM $MgCl_2$. This data set yielded a well-ordered CTP–Mg^{2+} binding site with full occupancy. This refined model served as the starting model for construction and refinement of the apo form and CDP-ME–Mg^{2+} complex. PROCHECK[27] analysis of all models shows 92% of the main chain torsion angles in the most favored regions and no residues in the disallowed regions. Figures were generated using MOLSCRIPT[28] and BOBSCRIPT[29] and rendered with POV-RAY[30]. The electrostatic surface potential was generated using GRASP[31].

Coordinates. Coordinates and structure factors for the CDP-ME synthetase apo (accession code 1INJ), the CTP-Mg^{2+} complex (accession code 1I52) and the CDP-ME-Mg^{2+} complex (accession code 1INI) have been deposited in the Protein Data Bank[9].

Acknowledgements

This work is based in part on research conducted at the Stanford Synchrotron Radiation Laboratory (SSRL), which is funded by the Department of Energy (BES, BER) and the National Institutes of Health (NCRR, NIGMS). We thank the staffs of the SSRL crystallographic beam lines (7-1, 7-2 and 9-2) and members of the Noel laboratory for technical assistance. This work was supported by grants from the Philippe Foundation Inc. to S.B.R. and the NIH (NIGMS) to J.P.N. and to D.E.C.

Received 26 March, 2001; accepted 9 May, 2001.

1. Sacchettini, J.C., & Poulter, C.D. Creating isoprenoid diversity. *Science* **277** 1788–1789 (1997).
2. Rohmer, M. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat. Prod. Rep.* **16**, 565–574 (1999).
3. Jomaa, H. *et al.* Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285**, 1573–1576 (1999).
4. Kuzuyama, T., Shimizu, T., Takahashi, S. & Seto, H. Fosmidomycin, a specific inhibitor of 1-deoxy-D-xylulose 5-phosphate reductoisomerase in the non-mevalonate pathway for terpenoid biosynthesis. *Tetrahedron Lett.* **39**, 7913–7916 (1998).
5. Lichtenthaler, H.K., Zeidler, J., Schwender, J. & Muller, C. The non-mevalonate isoprenoid biosynthesis of plants as a test system for new herbicides and drugs against pathogenic bacteria and the malaria parasite. *Z. Naturforsch.* **55**, 305–313 (2000).
6. Rohdich, F. *et al.* Cytidine 5′-triphosphate-dependent biosynthesis of isoprenoids: YgbP protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-C-methylerythritol. *Proc. Natl. Acad. Sci. USA* **96**, 11758–11763 (1999).
7. Kuzuyama, T., Takagi, M., Kaneda, K., Dairi, T. & Seto, H. Formation of 4-(cytidine 5′-diphospho)-2-C-methyl-erythritol from 2-C-methyl-erythritol 4-phosphate by 2-C-methyl-erythritol 4-phosphate cytidylyltransferase, a new enzyme in the nonmevalonate pathway. *Tetrahedron Lett.* **41**, 703–706 (2000).
8. Rossmann, M.G., Liljas, A., Branden, C.-I. & Banaszak, L.J. *The enzymes* (Academic Press, New York; 1975).
9. Berman, H.M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
10. Holm, L. & Sander, C. Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138 (1993).
11. Jelakovic, S., Jann, K. & Schulz G.E. The three-dimensional structure of capsule-specific CMP:2-keto-3-deoxy-*manno*-octonic acid synthetase from *Escherichia coli*. *FEBS Lett.* **391**, 157–161 (1996).
12. Weber, C.H., Park, Y.S., Sanker, S., Kent, C. & Ludwig, M.L. A prototypical cytidylyltransferase: CTP:glycerol-3-phosphate cytidylyltransferase from *Bacillus subtilis*. *Structure Fold. Des.* **7**, 1113–1124 (1999).
13. Mosimann, S.C. *et al.* Structure of a sialic acid activating synthetase, CMP acylneuraminate synthetase in the presence and absence of CDP. *J. Biol. Chem.* **276**, 8190–8196 (2001).
14. Bork, P., Holm, L., Koonin, E.V. & Sander, C. The cytidylyltransferase superfamily: identification of the nucleotide-binding site and fold prediction. *Proteins* **22**, 259–266 (1995).
15. Veitch, D.P., Gilham, D. & Cornell, R.B. The role of histidine residues in the HXGH site of CTP:phosphocholine cytidylyltransferase in CTP binding and catalysis. *Eur. J. Biochem.* **255**, 227–234 (1998).
16. Cane, D.E., Chow, C., Lillo, A. & Kang, I. Molecular cloning, expression and characterisation of the first three genes in the mevalonate independant isoprenoid pathway in *Streptomyces coelicolor. Bioorg. Med. Chem.* **9**, 1467–1477 (2001).
17. Herz, S. *et al.* Biosynthesis of terpenoids: YgbB protein converts 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate to 2C-methyl-D-erythritol 2,4-cyclodiphosphate. *Proc. Natl. Acad. Sci. USA* **97**, 2486–2490 (1999).
18. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. **276**, 307–326 (1997).
19. French, G.S. & Wilson, K.S. On the treatment of negative intensity observations. *Acta Crystallog. A* **34**, 517–525 (1978).
20. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallog. D* **50**, 760–763 (1994).
21. La Fortelle E. de & Bricogne G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).
22. McRee, D.E. A visual protein crystallographic software system for X11/Xview. *J. Mol. Graph.* **10**, 44–46 (1992).
23. Otwinowski, Z. in *Proceedings of the CCP4 study weekend* 307–326 (Daresbury Laboratories, Warrington, UK; 1991).
24. Perrakis, A., Sixma, T.K., Wilson K.S. & Lamzin, V.S. wARP: improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy atomic models. *Acta Crystallogr. D* **53**, 448–455 (1997).
25. Jones, T.A. & Kjeldgaard, M.O. Electron-density map interpretation. *Methods Enzymol.* **277**, 173–208 (1997).
26. Brunger, A.T. *et al.* Crystallography and NMR System (CNS): a new software system for macromolecular structure determination. *Acta Crystallog. D* **54**, 905–921 (1998).
27. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291 (1993).
28. Kraulis, P.J. MOLSCRIPT: a program to produce both detailled and schematic plots of structures. *J. Appl. Crystallog.* **24**, 946–950 (1991).
29. Esnouf, R. An extensively modified version of Molscript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 133–138 (1997).
30. Amundsen, S. X-POV-Team POV-Ray: persistence of vision ray-tracer. *http://www.povray.org* (1997).
31. Nicholls, A. & Honig, B. GRASP. *J. Comput. Chem.* **12**, 435–445 (1991).
32. Sprenger G.A. *et al.* Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proc. Natl. Acad. Sci. USA* **94**, 12857–12862 (1997).
33. Lois, L.M. *et al.* Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. *Proc. Natl. Acad. Sci. USA* **95**, 2105–2110 (1997).
34. Takahashi, S., Kuzuyama, T., Watanabe, H. & Seto, H. A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc. Natl. Acad. Sci. USA* **95**, 9879–9884 (1998).
35. Kuzuyama, T., Takahashi, S., Takagi, M. & Seto, H. Characterization of 1-deoxy-D-xylulose 5-phosphate reductoisomerase, an enzyme involved in isopentenyl diphosphate biosynthesis, and identification of its catalytic amino acid residues. *J. Biol. Chem.* **275**, 19928–19932 (2000).
36. Luttgen, H. *et al.* Biosynthesis of terpenoids: YchB protein of *Escherichia coli* phosphorylates the 2-hydroxy group of 4-diphosphocytidyl-2C-methyl-D-erythritol. *Proc. Natl. Acad. Sci. USA* **97**, 1062–1067 (2000).
37. Kuzuyama, T. *et al.* Studies on the nonmevalonate pathway: conversion of 4-(cytidine 5′-diphospho)-2-C-methyl-erythritol to its 2-phospho derivative by 4-(cytidine 5′-diphospho)-2-C-methyl-erythritol kinase. *Tetrahedron Lett.* **41**, 2925–2928 (2000).
38. Takagi, M. *et al.* Studies on the nonmevalonate pathway: formation of 2-C-methyl-D-erythritol 2,4-cyclodiphosphate from 2-phospho-4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol. *Tetrahedron Lett.* **41**, 3395–3398 (2000).