Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from *Mycobacterium tuberculosis*

Shan Shan,[*,†,1] Xuehui Chen,[‡,1] Ting Liu,[§] Hanchao Zhao,[*,†] Zihe Rao,[*,†,‡,§] and Zhiyong Lou[*,†,§,2]

*Structural Biology Laboratory and †MOE Laboratory of Protein Science, Tsinghua University, Beijing, China; ‡National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing, China; and §High-Throughput Molecular Drug Discovery Center, Tianjin Joint Academy of Biotechnology and Medicine, Tianjin, China

ABSTRACT **Isoprenoid precursors, which are a large group of natural products and play key roles in many biological pathways, can only be biosynthesized by the 2-C-methyl-D-erythritol 4-phosphate pathway in *Mycobacterium tuberculosis*. The 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE), which is an essential enzyme in the isoprenoid precursor biosynthesis pathway, catalyzes ATP-dependent phosphorylation of 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) to 4-diphosphocytidyl-2C-methyl-D-erythritol-2-phosphate and plays a crucial role in *M. tuberculosis* survival. Therefore, IspE is characterized as an attractive and potential target for antimicrobial drug discovery. However, no experimental structure of *M. tuberculosis* IspE has been reported, which has hindered our understanding of its structural details and mechanism of action. Here, we report the expression and purification of fully active full-length *M. tuberculosis* IspE and solve the high-resolution crystal structures of IspE alone and in complex with either the substrate CDP-ME or nonhydrolyzable ATP analog or ADP. The structures present a characteristic galactose/homoserine/mevalonate/phosphomevalonate kinase superfamily α/β-fold with a catalytic center located in a cleft between 2 domains and display clear substrate and ATP binding pockets. Our results also indicate distinct differences in ligand binding of *M. tuberculosis* IspE with other reported IspEs. Combined with the results of mutagenesis and enzymatic studies, our results provide useful information on the structural basis of IspE for future anti-*M. tuberculosis* drug discovery targeting this kinase.—Shan, S. Chen, X., Liu, T., Zhao, H., Rao, Z., Lou, Z. Crystal structure of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE) from *Mycobacterium tuberculosis*. *FASEB J.* 25, 1577–1584 (2011). www.fasebj.org**

Key Words: mechanism · drug discovery

Mycobacterium tuberculosis is currently one of the most widespread human pathogens; it is responsible for millions of deaths annually and nearly 10 million new infections every year. Unfortunately, this situation has recently been exacerbated due to poor living conditions, coinfection with human immunodeficiency virus (HIV), and the emergence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) *M. tuberculosis* strains (1, 2). Currently approved anti-*M. tuberculosis* drugs may have limited effectiveness in the event of the emergence and reemergence of MDR and XDR *M. tuberculosis* pandemics; therefore, identifying new drug targets and understanding the underlying mechanisms of these targets are crucial for the development of new anti-*M. tuberculosis* therapies (3).

Isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) are the precursors of isoprenoids, an extensive group of natural products consisting of essential primary and secondary metabolites (*e.g.*, dolichol, ubiquinones, carotenoids, and sterols) that contribute to many important biological functions, including electron transport, photosynthesis, apoptosis, meiosis, protein cleavage and degradation, and hormone-based signaling (4–6). In the life cycle of *M. tuberculosis* and several other bacterial pathogens, isoprenoids play crucial roles in the biosynthesis of structural components of the cell wall, and thus are characterized as attractive and potential targets for antimicrobial drug discovery (7, 8).

Currently, 2 different routes have been identified to biosynthesize IPP and DMAPP: the well-known mevalonate pathway in mammals, higher plants, archaea, and most eukaryotes (9, 10); and the 2-C-methyl-D-erythritol 4 phosphate (MEP) pathway, which is also known as the 1-deoxy-D-xylulose-5-phosphate (DOXP) or nonmevalonate pathway, in bacteria, plant plastids, and especially in several pathogenic microorganisms (*e.g.*, *M. tuberculosis* and *Plasmodium falciparum*; ref. 11). The MEP pathway consists of 8 steps and 7 enzymes (12): DOXP synthase;

[1] These authors contributed equally to this work.

[2] Correspondence: Laboratory of Structural Biology, New Life Science Bldg., Tsinghua University, Beijing 100084, China. E-mail: louzy@xtal.tsinghua.edu.cn

doi: 10.1096/fj.10-175786

DOXP reductase, 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase (IspD); 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (IspE); 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase (IspF); (*E*)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate (HMB-PP) synthase (IspG); and HMB-PP reductase (LytB or IspH) (13). Because the MEP pathway is vital for *M. tuberculosis* survival and absent in mammalian cells, it is an attractive potential antimicrobial drug target (14–18). Indeed, fosmidomycin (3-[formyl(hydroxy)amino] propoylphosphonic acid), an antibiotic used in clinical trials to treat malaria (19), targets on inhibiting its second step of the MEP pathway (20). This result suggests that compounds that can inhibit the MEP pathway may be potential anti-*M. tuberculosis* drug candidates.

Our research concentrates on the fourth step of the MEP pathway, in which the IspE converts 4-diphosphocytidyl-2-C-methyl-D-erythritol (CDP-ME) to 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate (CDP-ME2P) in an ATP-dependent manner (**Fig. 1**). IspE has been identified as one of the genes that is crucial for pathogenic bacteria survival and was recently shown to be essential in *Mycobacterium smegmatis* (21). Due to its importance in bacteria and absence in mammalian cells, IspE has long been considering an ideal antibacterial drug target (22) and the deeper research on *M. tuberculosis* IspE will help us find the better way to treat *M. tuberculosis*.

Previous results indicate that the IspE belongs to the galactose/homoserine/mevalonate/phosphomevalonate (GHMP) kinase superfamily, which is named after several different proteins in mammals (*e.g.*, galactose kinase, homoserine kinase, alonate kinase, and phosphomevalonate kinase; refs. 23, 24). Currently, 3 crystal structures of *M. tuberculosis* IspE orthologues have been reported: the *Escherichia coli* orthologue (*E. coli* IspE) in complex with the ATP analog adenosine 5′-(β,γ-imino)triphosphate (AMP-PNP) and the substrate CDP-ME in 2003 (13), the *Aquifex aeolicus* IspE (25), and the *Thermus thermophilus* IspE, and they all tallied with great GHMP superfamily character (13). The amino acid sequence similarities between these 3 enzymes and the *M. tuberculosis* IspE suggest similar overall structures and enzymatic mechanisms. However, no experimental structure of *M. tuberculosis* IspE has been reported, which has hindered our understanding of its structural details and mechanism of action (13, 22).

Here, we report the construction, expression, and purification of full active full-length recombinant *M. tuberculosis* IspE. We also determined the 3-D crystal structure of apo *M. tuberculosis* IspE, as well as in the complex containing either substrate or AMP-PNP or ADP, providing structural details to complement our enzyme study and future drug discovery efforts targeting the *M. tuberculosis* IspE to treat MDR and XDR strains.

MATERIALS AND METHODS

Protein expression and purification

We used primers 5′-TGCAACT**CATATG**GTCAGTGAAGCGTCCG-3′ (F) and 5′-CAT**CTCGAG**TTCAGCCACGTACTACACCA-3′ (R) to amplify the gene encoding the 306-aa full-length IspE from *M. tuberculosis* H37Rv. The target gene was inserted into the pET-28a vector (Novagen, San Diego, CA,

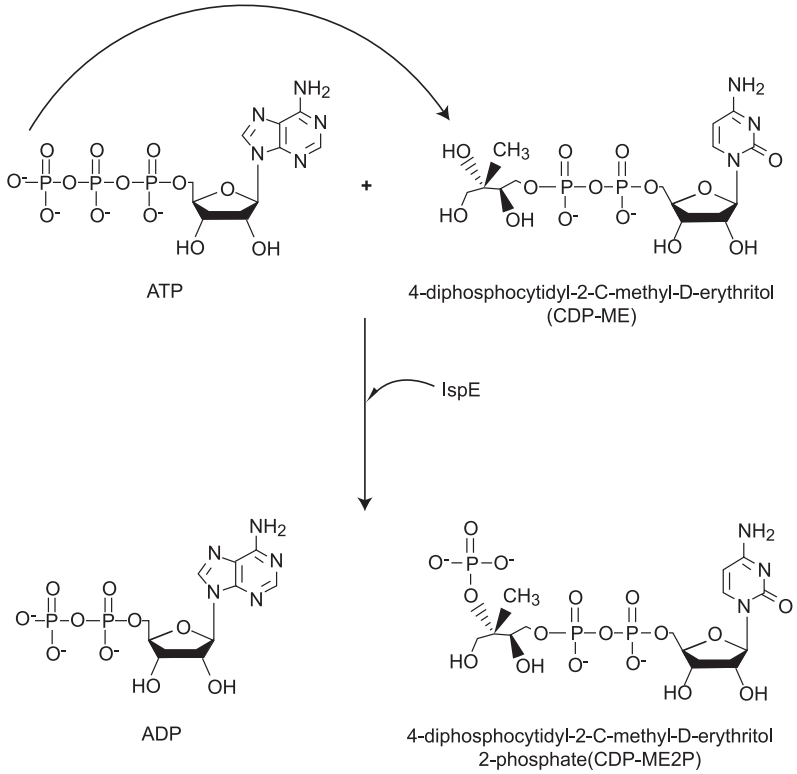

Figure 1. The fourth-step reaction of the MEP pathway catalyzed by IspE.

USA) using the *Nde*I and *Xho*I (bold and underscored in the primers) restriction sites.

The recombinant plasmid was transformed into *E. coli* strain BL21 (DE3) and overexpressed as a 6xHis-tag fusion protein. The cells were grown with shaking for 4–5 h at 37°C in 1 L Luria-Bertani medium containing 50 μg/ml kanamycin until the OD_{600} reached 0.4–0.6. Then, the cells were transferred to 16°C, and IspE expression was induced by incubation for 16–20 h in the presence of 0.3 mM IPTG. The cells were harvested; resuspended in 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10 mM imidazole, and 5 mM β-mercaptoethanol (β-ME); and then lysed by sonication. The lysate was centrifuged at 16,000 *g* for 35 min to remove cell debris.

The supernatant was loaded twice onto a Ni-NTA column (Qiagen, Valencia, CA, USA) equilibrated with lysis buffer. Nonspecifically bound proteins were washed off with 5 column volumes of 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 30 mM imidazole, and 5 mM β-ME; the target protein was eluted with 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 100 mM imidazole, and 5 mM β-ME. Eluted protein was dialyzed against 20 mM Tris-HCl (pH 7.0), 50 mM NaCl, 1 mM DTT, and 20 U thrombin overnight at 4°C to remove the N-terminal 6xHis tag and then reloaded onto the Ni-NTA column to remove the cleaved tag. Relevant mutants were generated by site-directed mutagenesis and purified *via* the same procedure.

The recombinant IspE was further purified using a Resource Q anion exchange column (GE Healthcare, Little Chalfont, UK) and a Superdex-200 gel filtration column (GE Healthcare) according to the manufacturer's recommendations. Fractions were analyzed with SDS-PAGE, and the purity was >95%. The purified protein was finally concentrated to 15 mg/ml in 20 mM Tris-HCl (pH 7.0), 50 mM NaCl, and 5 mM DTT for crystallization.

Crystallization

Crystallization of the *M. tuberculosis* IspE was performed at 18°C using the hanging-drop vapor-diffusion method. The crystals grew in drops containing 1 μl IspE and 1 μl reservoir solution against 200 μl reservoir solutions. Small and twin crystals appeared after 1 mo of growth in 0.1 M bis-tris propane (pH 7.0) and 3.2 M NaCl, with poor diffraction quality after initial screening. Microseeding was used to optimize the crystals to a final size of 80 × 80 × 100 μM, which diffracted to 2.1 Å. IspE in complex with different small molecules was prepared from crystals grown in 0.1 M bis-tris propane (pH 7.5) and 1.5 M Li_2SO_4 by soaking them for 16 h at 18°C in a mixture consisting of reservoir solution and small molecules (60 mM CDP-ME and/or saturated AMP-PNP or saturated ADP). The crystals were then soaked in a cryoprotectant solution consisting of the reservoir solution and 20% (v/v) glycol and then flash-frozen in liquid nitrogen for X-ray diffraction data collection.

Data collection, processing, and structure determination

The apo-IspE crystal diffracted to 2.1-Å resolution at 100 K using a Rigaku FR-E rotating anode generator with an R-Axis IV++ image plate (Rigaku, Tokyo, Japan) at 45 kV and 45 mA, while the complex crystal data were collected with a Rigaku MM-007 rotating anode generator with an R-Axis IV++ image plate at 40 kV and 20 mA. Data were processed, integrated, and scaled using the HKL2000 package (26). The crystals belong to the $P2_12_12_1$ space group with unit cell parameters of a = 52.5 Å, b = 72.3 Å, c = 107.3 Å, and α = β = γ = 90°. There is only 1 molecule/asymmetric unit, with a Matthews coefficient of 3.4 $Å^3$/Da, corresponding to 63% solvent content (27).

The Phaser program (28) was used to find the correct solution *via* the molecular replacement method, using the *T. thermophilus* IspE crystal structure [Protein Data Bank (PDB) code 1UEK; shares 38% sequence identity with the *M. tuberculosis* IspE] as the initial searching model. Subsequently, the *M. tuberculosis* apo-IspE crystal structure was used as a starting model to solve the structure of the *M. tuberculosis* IspE in complex with ligands.

Manual model building and refinement were performed with Coot (29) and Phenix (30). During the later stages of positional refinement, restraints were relaxed, and a bulk solvent correction was applied under the guidance of R_{free}. Model geometry was verified using the program ProCheck (31). Solvent molecules were located from stereochemically reasonable peaks in the σA-weighted F_o-F_c difference electron density map. Final refinement statistics are summarized in **Table 1**. Structural figures were drawn with the program PyMOL (32).

M. tuberculosis IspE enzymatic assay

M. tuberculosis IspE activity was determined by a spectrophotometric assay that uses pyruvate kinase and L-lactate dehydrogenase as auxiliary enzymes to transfer the action of the IspE activity to an observable decrease in the absorption of NADH at 340 nm (33). To measure the kinetic properties of wild-type IspE and relevant mutants, enzymatic assays were conducted based on the activity of IspE observed above. Briefly, we used 100 mM Tris-HCl (pH 7.5), 1 mM DTT, 20 mM KCl, 10 mM $MgCl_2$, 0.45 mM NADH, 4 mM phosphoenolpyruvate, 8 U pyruvate kinase, 9 U lactate dehydrogenase, 5 μg *M. tuberculosis* IspE, and 1.5 mM ATP, which was added last, in a total volume of 50 μl. The reaction mixture was incubated at 37°C for 5 min, and the A_{340} value was measured using a Thromo spectrophotometer (Thermo Scientific, Waltham, MA, USA) immediately after adding an additional 1.5 mM ATP.

RESULTS AND DISCUSSION

Overall *M. tuberculosis* IspE structure

The crystal structure of the 31-kDa, full-length apo-IspE was solved by the molecular replacement method and refined to 2.1-Å resolution, resulting in a final R_{work} value of 19% (R_{free}=24%). The crystal belongs to the $P2_12_12_1$ space group, and there is 1 IspE monomer/asymmetric unit, with a Matthews coefficient of 3.4 $Å^3$/Da (corresponding to 63% solvent content; ref. 34). The first 4 N-terminal residues are missing due to a lack of density, which did not affect the correct molecular folding of the entire structure. Gel filtration and analytical ultracentrifugation (data not shown) show that IspE exists as a monomer in solution, which is consistent with the crystallographic result and suggests that the monomer is the biological unit *in vivo*.

The *M. tuberculosis* IspE displays a compact, cylindrical, 60- × 50- × 35-Å structure consisting of 2 relatively separated domains: the ATP binding domain (N-terminal domain) and substrate binding domain (C-terminal domain), which is the canonical GHMP kinase fold (35). The ATP binding domain (residues Met1-Gly148)

TABLE 1. *X-ray data collection and refinement statistics*

Parameter	CDP-ME kinase	CDP-ME kinase/CDP-ME	CDP-ME kinase/AMP-PNP	CDP-ME kinase/ADP
Data collection				
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell parameters (Å, deg)	a = 52.5, b = 72.3, c = 107.3 Å; α = β = γ = 90°	a = 52.7, b = 73.6, c = 106.2 Å; α = β = γ = 90°	a = 52.9, b = 73.4, c = 106.3 Å; α = β = γ = 90°	a = 51.2, b = 73.1, c = 106.9 Å; α = β = γ = 90°
Wavelength (Å)	1.0000	1.5418	1.0000	1.0000
Resolution range (Å)[a]	50.00–2.10 (2.18–2.10)	40.00–2.00 (2.07–2.00)	50.00–1.70 (1.76–1.70)	50.00–2.00 (2.07–2.00)
R_{merge} (%)[a]	9.2 (61.5)	16.4 (55.2)	4.9 (48.5)	6.8 (55.8)
Observed reflections	153,128 (12,911)	367,673 (20,451)	672,696 (66,532)	126,041 (12,631)
Unique reflections	24,569 (2391)	28,610 (2622)	46,412 (4557)	27,581 (2746)
Completeness (%)	99.8 (99.0)	99.3 (92.7)	100.0 (100.0)	97.3 (98.5)
$I/\sigma(I)$	13.4 (2.4)	32.8 (3.4)	59.3 (6.7)	16.9 (2.3)
Refinement				
Resolution range (Å)	37.57–2.10	60.52–2.00	33.41–1.70	30.18–1.99
Reflections used	23,346	27,105	45,280	26,350
R_{work} (%)	19.4	21.5	21.6	22.3
R_{free} (%)[b]	24.4	25.0	23.6	24.7
Protein atoms	2343	2570	2556	2366
Overall average B factor (Å^2)	35.08	33.49	30.88	38.72
RMSD				
RMSD bond lengths (Å)	0.008	0.024	0.005	0.025
RMSD bond angles (deg)	1.177	2.183	1.075	1.657

Values in parentheses refer to the highest resolution shell. [a]$R_{merge} = \Sigma_h \Sigma_l |I_{ih} - \langle I_h \rangle| / \Sigma_h \Sigma_I \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean of multiple observations I_{ih} of a given reflection h. [b]$R_{work} = \Sigma \| F_{p(obs)}| - |F_{p(calc)} \| / \Sigma |F_{p(obs)}|$; R_{free} = R factor for a selected subset (5%) of reflections that was not included in prior refinement calculations.

consists of 4 β strands (β1, β4, β5, and β6), 4 α helices (α1 to α4), and 1 3_{10} helix (η1), while the substrate binding domain is composed of 10 β strands (β2, β3, and β5 to β13) and 6 α helices (α5 to α10) (**Fig. 2*A***). β1, β4, β5, and β6 form a 4-strand β sheet; β4 is antiparallel to the other 3. The concave side of this 4-strand β sheet is linked to 4 α helices to form the N-terminal domain. This 4-stranded β sheet and another (consisting of β6 and β9, β11, β12, and β13) are wrapped in the middle of 6 α helices to form the main structure of the C-terminal domain.

The 3 known conserved motifs in the GHMP kinase superfamily that are involved in creating the catalytic center are also found in the *M. tuberculosis* IspE (21). Motif A (Lys13 to Leu18), which is located in the β2 strand, is involved in forming the substrate binding site. Motif B (Gly102 to Ser107) is a glycine-rich phosphate binding loop, which is located in the loop between β6

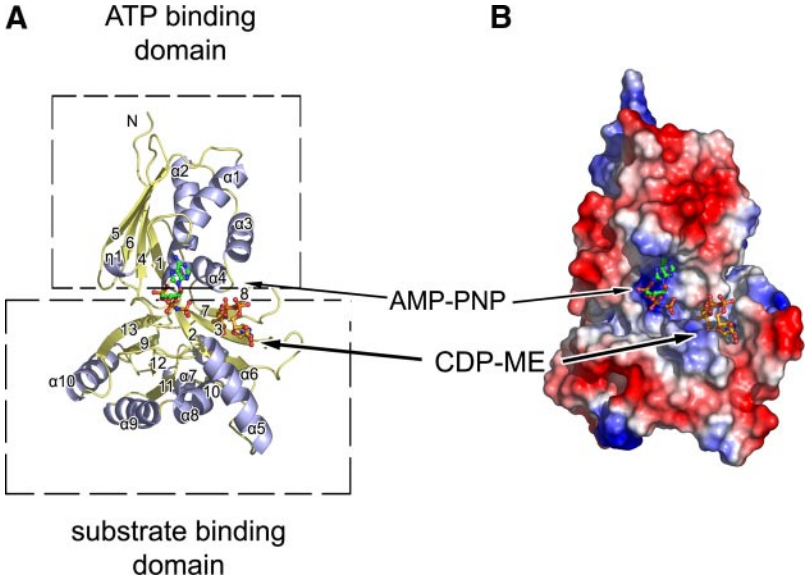

Figure 2. Overall structure of *M. tuberculosis* IspE. *A*) Overview of crystal structure, shown in cartoon view. The α and 3_{10} helices are shown in light blue; β strands and the loop regions are in yellow. All secondary structure elements are labeled; N terminus and the 2 functional domains are marked. CDP-ME atoms are shown as sticks: C, gold; O, red; N, blue; P, orange. AMP-PNP atoms are shown as sticks: C, green; O, red; N, blue; P, orange. *B*) View of the molecular surface of *M. tuberculosis* IspE. Surface coloring indicates electrostatic potential (red, negative; blue, positive). CDP-ME and AMP-PNP atoms are shown as sticks and colored as in panel A.

and α2 and interacts with the triphosphate moiety of ATP. Moreover, motif B is also described as the P loop, which is found in the septin family and some NTP hydrolases (36). Motif C (Val254 to Gly258), which is located in the loop between β11 and β12, helps to stabilize the conformation of motifs A and B rather than interacting with ligands directly. Thus, motif C is likely a linker to sustain the stability of motifs A and B, akin to a corner of a triangle that stabilizes the other 2 corners. All these motifs together create the catalytic center (Figs. 2*A* and **3*B***).

The *M. tuberculosis* IspE shares high sequence identity with IspE orthologues from the other bacteria, with the highest sequence identities of 38% (*T. thermophilus*), 37% (*E. coli*), and 30% (*A. aeolicus*). Unsurprisingly, *M. tuberculosis* IspE also shares high structural similarity with its homologues in these species: *T. thermophilus* [PDB code 1UEK, *Z* score 30.0, root mean square deviation (RMSD) of Ca 2.343], *E. coli* (PDB code 1OJ4, *Z* score 28.8, RMSD of Ca 10.151), and *A. aeolicus* (PDB code 2V2Z, *Z* score 27.7, RMSD of Ca 3.122) (37). Overlaying *M. tuberculosis* IspE and *T. thermophilus* IspE demonstrates that the overall folding and structural conformation are highly conserved. However, there is an extra 3_{10} helix (η1) between β3 and α1 in the *M. tuberculosis* structure. Further, the C terminus of *M. tuberculosis* IspE (which is longer than in the *T. thermophilus* IspE) forms a β strand that is missing in the *T. thermophilus* IspE structure, and the loop between α1 and β4 is slightly different (Fig. 3*A*).

Substrate and ATP binding sites

Because the substrate and ATP binding sites are potential sites for drug design targeting in protein kinases, which has been well demonstrated with cyclin-dependent kinases (CDKs), mitogen-activated kinases (MAPKs), glycogen synthase kinase 3 (GSK-3), and CDK-like kinases (CLKs) (38), it is crucial for us to know the precise substrate and ATP binding sites in the *M. tuberculosis* IspE. Therefore, we determined the high-resolution crystal structures of *M. tuberculosis* IspE in complex with its substrate CDP-ME, the stable ATP analog AMP-PNP, and the reaction product ADP, respectively. Although our efforts to obtain the ternary structure of IspE in complex with its substrate and ATP ultimately failed, our results still present a trusted model of the *M. tuberculosis* IspE in complex with substrate and ATP at the atomic level and shed light on its enzymatic mechanism. The K_m and V_{max} of CDP-ME and ATP were determined to be 666.7 μm and 3.0 μm/min and 285.7 μm and 3.0 μm/min, respectively. Our CDP-ME data are consistent recently reported *M. tuberculosis* IspE kinetic parameters (21), but our ATP values are higher, which may be due to the different *in vitro* conditions used.

The CDP-ME binding site is a deep cavity between the substrate binding domain and the ATP binding domain, which is formed by β3, β4, α4, and the loop regions linking β3 to β4 and β7 to β8 (Fig. 2*A*).The cytosine moiety of CDP-ME is deeply buried in the cavity, sandwiched by the aromatic side chains of Tyr28

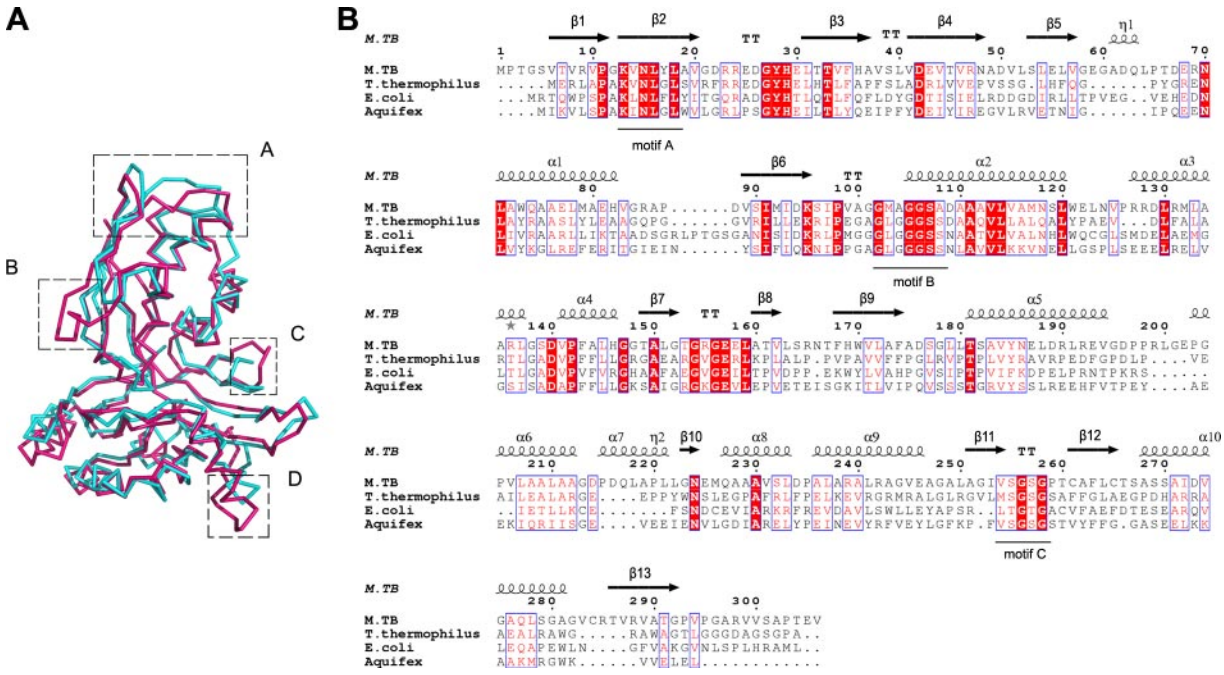

Figure 3. Comparison of *M. tuberculosis* IspE and orthologues. *A*) Comparison of *M. tuberculosis* IspE and *T. thermophilus* IspE. *M. tuberculosis* IspE is indicated by hot pink ribbons; *T. thermophilus* IspE is indicated as cyan ribbons. Positions of difference are marked with dashed boxes (labeled A–D). *B)* Sequence alignment of the IspE from *M. tuberculosis T. thermophilus, E. coli*, and *A. aeolicus*. Three conserved motifs are underscored and labeled. Secondary structure elements of *M. tuberculosis* IspE are indicated at top of the alignment; α helices and β strands are presented as curves and arrows, respectively. Number of residues at top refers to *M. tuberculosis* IspE sequence.

and Tyr185. These 2 residues stabilize the cytosine moiety with π-π stacking interactions to fix the position of CDP-ME. Tyr185 is conservatively replaced by phenylalanine in *E. coli* IspE but not in the *A. aeolicus* IspE. Tyr28 plays an key role in cytosine head binding due to its nearly 90° conformational change after CDP-ME binding, *i.e.*, from parallel to the β2–β3 loop to vertical relative to the β2–β3 loop (**Fig. 4*A*, *B***). These aromatic residues in the active site may act as a doorkeeper to close the site during substrate binding and open it for product release. His29 also participates in stabilizing the cytosine head by forming 3 hydrogen bonds with the pyrimidine. The His29 imidazole ND1 and amide in the main chain donate hydrogen bonds to the cytosine O2 and N3, respectively, of CDP-ME, and the His29 main-chain carboxyl accepts another hydrogen bond from cytosine N4. This interaction is conserved in other IspE substrate binding pockets (Fig. 4*A*).

The CDP-ME ribose position is directed out of the cleft formed by Tyr28 and Tyr185 and has access to the solvent. Based on the chemical bonds with the cytosine head, the ribose and phosphates are clearly in the density map, though they lack direct interactions with the other residues. The α-phosphate accepts a hydrogen bond from Tyr28 the hydroxyl. The CDP-ME tail inserts into the catalytic center, where the catalytic key residues Lys13 and Asp140 are found. There is a network of hydrogen bonds with and around the catalytic residues and ATP to achieve the catalytic reaction. However, without the ATP cofactor present, the density of the erythritol moiety is not clear, making the CDP-ME tail flexible. This flexibility of CDP-ME may regulate the insertion depth of the erythritol moiety by changing its conformation to control different reaction steps. The loop linking β7 and β8 is also altered in the structure containing CDP-ME. This loop is flexible in apo-IspE, but in the IspE-CDP-ME structure, a network of water-mediated interactions stabilizes the loop.

The ATP binding site is in a cavity in the N-terminal domain that is surrounded by the loop linking β6 to α2 and the loop linking β5 to α1. This cavity also includes a 3_{10} helix and is positioned adjacent to the CDP-ME binding site. In our structures, the 2 different ligands (ADP and AMP-PNP) display different binding conformations. In the structure containing AMP-PNP, only 3 phosphates can be unambiguously determined. This lack of clear density may be due to the flexibility of adenine in the binding pocket when no CDP-ME is present. The 3 phosphates can be clearly observed near the glycine-rich loop, where 7 hydrogen bonds are involved in binding. The O2G atom of the γ-phosphate is stabilized by Gly102 *via* a 3.1-Å hydrogen bond with the carboxyl O atom, and the O3G and O1G atoms of the γ-phosphate share a hydrogen bond with Wat169 and Wat16, respectively. Similarly, Wat16 also forms a 2.7-Å hydrogen bond to Ser107. The linker NB3 atom and the O2B atom of the β-phosphate accept hydrogen bonds donated from the main-chain amides of Met103 and Gly102, respectively. The α-phosphate forms 2 hydrogen bonds: between the O3A atom and Wat46 (2.8 Å) and the O2A atom and main-chain amide of

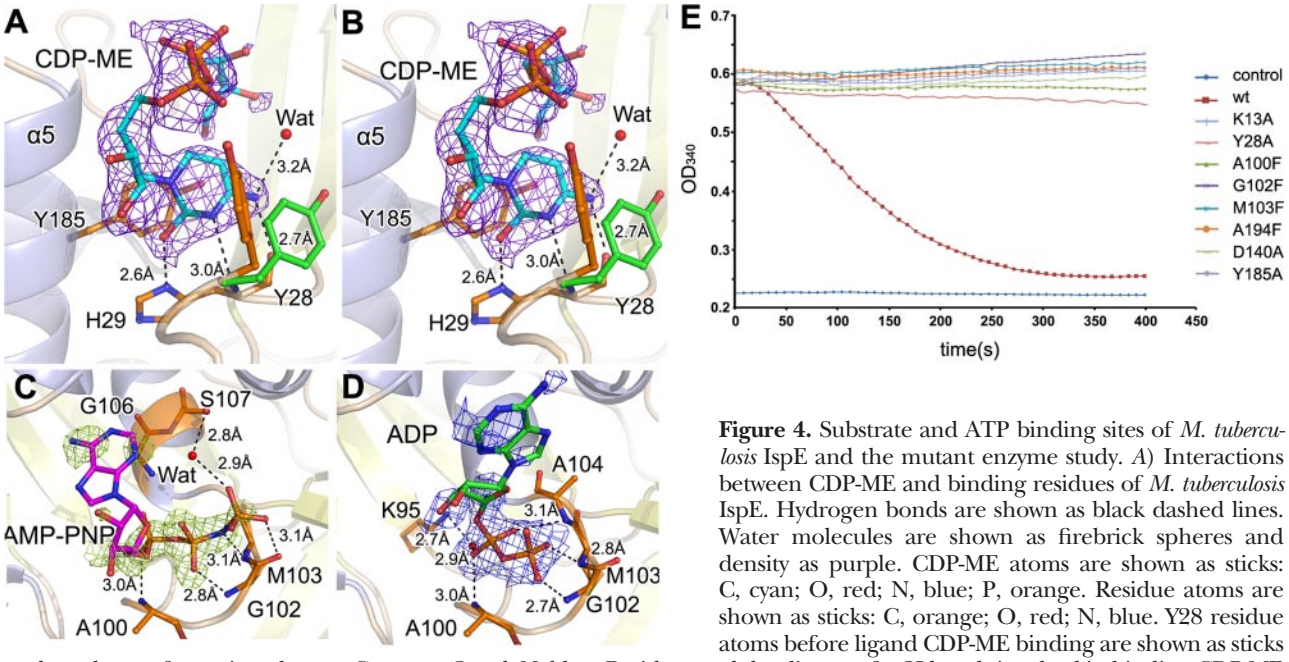

Figure 4. Substrate and ATP binding sites of *M. tuberculosis* IspE and the mutant enzyme study. *A*) Interactions between CDP-ME and binding residues of *M. tuberculosis* IspE. Hydrogen bonds are shown as black dashed lines. Water molecules are shown as firebrick spheres and density as purple. CDP-ME atoms are shown as sticks: C, cyan; O, red; N, blue; P, orange. Residue atoms are shown as sticks: C, orange; O, red; N, blue. Y28 residue atoms before ligand CDP-ME binding are shown as sticks to show the conformation change: C, green; O red; N, blue. Residues and the distance for H bonds involved in binding CDP-ME are labeled. *B*) Stereo view of panel *A*. *C*) Interactions between AMP-PNP and binding residues. Hydrogen bonds are shown as black dashed lines. Water molecules are shown as firebrick spheres and density as pea green. AMP-PNP atoms are shown as sticks: C, magenta; O, red; N, blue; P, orange. Residue atoms are shown as sticks: C, orange; O, red; N, blue; S, yellow. Residues and distance for H bonds involved in binding AMP-PNP are labeled. *D*) Interactions between ADP and binding residues. Hydrogen bonds are shown as black dashed lines. Water molecules are shown as firebrick spheres and density as blue. ADP atoms are shown as sticks: C, green; O, red; N, blue; P, orange. Residue atoms are shown as sticks: C, orange; O, red; N, blue; S, yellow. Residues and distance for H bonds involved in binding ADP are labeled. *E*) Mutant enzyme activity assays.

Ala100 (2.9 Å). These residues and water form a hydrogen bond network to firmly fix the phosphate portion of AMP-PNP in place (Fig. 4*C*).

The ADP binding conformation is different from that in the AMP-PNP-bound structure, which likely reflects the ADP-release step of the catalytic cycle. Such a conformation was not previously observed in the *E. coli* or *A. aeolicus* IspE structures. The adenine and ribose are solvent accessible, and a lack of interactions with the IspE residues results in vague density. The phosphates are still bound near the P loop, but their positions are slightly different to those in AMP-PNP. The O1B and O2A atoms of the α-phosphate accept hydrogen bonds donated by the main-chain amide of Ala104 and Ala100, respectively, whereas the O2A atom forms a water-mediated interaction with Lys95. The O1B and O3B atoms of the β-phosphate contact the main-chain amide of Gly102 and Met103 at distances of 2.7 and 2.9 Å. These residues may participate in the process of releasing the ADP product from the channel (Fig. 4*D*).

Although our attempts to generate the ternary complex of both factors failed, the amino acid conservation in active sites between the IspE orthologues in *E. coli*, *T. thermophilus*, *A. aeolicus*, and other members of the GHMP kinase superfamily indicates that Lys13 and Asp140 should also participate in the enzymatic reaction. Mutagenesis and enzymatic assessments of the K13A and D140A mutants demonstrated that both mutations perturb catalysis, confirming our hypothesis of the crucial role of Lys13 and Asp140 (Fig. 4*E*).

Moreover, although divalent cations are reported to play an essential role in the catalytic reaction (21), we did not observe any ordered density from metal ions in any of the 4 structures, despite the facts that relevant cations were added to the crystal growth solutions and the enzyme displayed normal catalytic activity. Indeed, metal ions are also absent in the *E. coli* (39) and *A. aeolicus* (25) IspE structures.

Implications for drug design targeting IspE

The structures of *M. tuberculosis* IspE alone and in complex with its substrate and ADP/AMP-PNP shed light on the detailed interactions of these molecules and the precise mechanism of catalysis. Combined with our mutagenesis study, these structural data provide valuable information for the development of new drugs targeting *M. tuberculosis* IspE that may ultimately be used to treat MDR and XDR infections.

While the observed interaction details between the substrate and IspE focus on the cytosine moiety, the interaction mainly involves the phosphates in the ATP binding site. These results suggest that new inhibitors designed for competitive inhibition should focus on altering the cytosine head to a high affinity state in order to decrease the activity of the kinase or to take the ATP binding site with a higher affinity than ATP. In addition, the new inhibitors can be designed by focusing on the CDP-ME tail, which may inhibit IspE function by restraining tail flexibility. We designed 6 mutants (*e.g.*, Y28A, A100F, G102F, M103F, A104F, and Y185A; Fig. 4*E*) for structure/function and enzymatic studies to determine the residues involved in binding. Our enzyme activity assays showed that phosphate transfer by all these mutants was completely inhibited, indicating that the residues responsible for the binding interactions with CDP-ME and the cofactor AMP-PNP are crucial for catalysis.

On the basis of the crystal structures, we may be able to design compounds with high binding affinity to block the favorable interactions between *M. tuberculosis* IspE and CDP-ME or ATP. For example, the side chain of the G106F mutant blocks the binding pocket from the ribose moiety of ATP, and loss of the benzene ring in the Y28A and Y185A mutant decreases the interaction with the CDP-ME cytosine head, thus increasing the flexibility of the substrate. In addition, the change of side chain of the A100F, G102F, M103F, and A104F mutant blocks the ATP binding site, decreasing the size of the binding pocket. In summation, this mutant work suggests targets in the *M. tuberculosis* IspE structure for new drug design is an operable plan.

CONCLUSIONS

The IspE is involved in the fourth step in the bacterial MEP pathway, which is crucial for the survival of many pathogens and is absent in mammalian cells. Thus, it has been characterized as an attractive and potential drug target to treat *M. tuberculosis* infection. Here, we report the high-resolution crystal structures of the *M. tuberculosis* IspE alone and in complex with substrate CDP-ME, the nonhydrolyzable ATP analog AMP-PNP, or ADP, and we provide an efficient method to produce soluble and stable full-length recombinant protein for research on crystal structure and high-throughput screening of small molecular inhibitors. The structures show clear details of interaction among the substrate, ATP, and residues and make it feasible to compare with other reported IspEs in ligand binding. The mutant enzyme activity results indicate that the key residues participated in CDP-ME and ATP binding are crucial for the enzyme function and destruction of the interaction can decrease the enzyme activity.These *M. tuberculosis* IspE crystal structures and the mechanistic insight gleaned from them lay the foundation for future structure-based drug design and high-throughput screening of small molecular inhibitors to treat *M. tuberculosis* infection worldwide. FJ

This work was supported by the National Natural Science Foundation of China (grants 30870486, 30770516, and 30730022), the 973 Project (grants 2007CB914304 and 2010CB530100), and the National Major Projects (grants 2009ZX09311-001 and 2009ZX10004-304).

REFERENCES

1. Comas, I., and Gagneux, S. (2009) The past and future of tuberculosis research. *PLoS Pathog.* **5,** e1000600
2. Ma, W. C. Y., Zhao, Y., Zheng. W., Zhang. W., Jiang, X., and Zhang, W. (2010) Progress of antibacterial mechanisms study on nanoparticles. *Acta Biophys. Sin.* **26,** 638–648
3. Li, A. T. T., Zhang, H., Wang, Q., Tian, J., and Li, Y. (2010) Modification of bacterial magnetosomes and application of magnetosome-antibody complex in pathogen detection. *Acta Biophys. Sin.* **26,** 680–690
4. Gershenzon, J., and Dudareva, N. (2007) The function of terpene natural products in the natural world. *Nat. Chem. Biol.* **3,** 408–414
5. Edwards, P. A., and Ericsson, J. (1999) Sterols and isoprenoids: signaling molecules derived from the cholesterol biosynthetic pathway. *Annu. Rev. Biochem.* **68,** 157–185
6. Zhang, S., Du, Y., Wang, Y., and Liu, D. (2010) Lipid droplet—a cellular organelle for lipid metabolism. *Acta Biophys. Sin.* **26,** 97–105
7. Mikusova, K., Mikus, M., Besra, G. S., Hancock, I., and Brennan, P. J. (1996) Biosynthesis of the linkage region of the mycobacterial cell wall. *J. Biol. Chem.* **271,** 7820–7828
8. Brennan, P. J., and Crick, D. C. (2007) The cell-wall core of *Mycobacterium tuberculosis* in the context of drug discovery. *Curr. Top. Med. Chem.* **7,** 475–488
9. Fritz, G. (2009) Targeting the mevalonate pathway for improved anticancer therapy. *Curr. Cancer Drug Targets* **9,** 626–638
10. Lou, Z.Y., and Zhang, X. X. (2010) Protein targets for structure-based anti-Mycobacterium tuberculosis drug discovery. *Protein Cell* **1,** 435–442
11. Eisenreich, W., Bacher, A., Arigoni, D., and Rohdich, F. (2004) Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell. Mol. Life Sci.* **61,** 1401–1426
12. Rohdich, F., Kis, K., Bacher, A., and Eisenreich, W. (2001) The non-mevalonate pathway of isoprenoids: genes, enzymes and intermediates. *Curr. Opin. Chem. Biol.* **5,** 535–540
13. Wada, T., Kuzuyama, T., Satoh, S., Kuramitsu, S., Yokoyama, S., Unzai, S., Tame, J. R., and Park, S. Y. (2003) Crystal structure of 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol kinase, an enzyme in the non-mevalonate pathway of isoprenoid synthesis. *J. Biol. Chem.* **278,** 30022–30027
14. Testa, C.A., and Brown, M. J. (2003) The methylerythritol phosphate pathway and its significance as a novel drug target. *Curr. Pharm. Biotechnol.* **4,** 248–259
15. Rohmer, M. (1998) Isoprenoid biosynthesis via the mevalonate-independent route, a novel target for antibacterial drugs? *Prog. Drug Res.* **50,** 135–154
16. Hasan, S., Daugelat, S., Rao, P. S., and Schreiber, M. (2006) Prioritizing genomic drug targets in pathogens: application to Mycobacterium tuberculosis. *PLoS Comput. Biol.* **2,** e61
17. de Ruyck, J., and Wouters, J. (2008) Structure-based drug design targeting biosynthesis of isoprenoids: a crystallographic state of the art of the involved enzymes. *Curr. Protein Pept. Sci.* **9,** 117–137
18. Eoh, H., Brennan, P. J., and Crick, D. C. (2009) *The Mycobacterium tuberculosis* MEP (2C-methyl-d-erythritol 4-phosphate) pathway as a new drug target. *Tuberculosis (Edinb.)* **89,** 1–11
19. Missinou, M. A., Borrmann, S., Schindler, A., Issifou, S., Adegnika, A. A., Matsiegui, P. B., Binder, R., Lell, B., Wiesner, J., Baranek, T., Jomaa, H., and Kremsner, P. G. (2002) Fosmidomycin for malaria. *Lancet* **360,** 1941–1942
20. Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Türbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H. K., Soldati, D., and Beck, E. (1999) Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285,** 1573–1576
21. Eoh, H., Narayanasamy, P., Brown, A. C., Parish, T., Brennan, P. J., and Crick, D. C. (2009) Expression and characterization of soluble 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase from bacterial pathogens. *Chem. Biol.* **16,** 1230–1239
22. Rodriguez-Concepcion, M. (2004) The MEP pathway: a new target for the development of herbicides, antibiotics and antimalarial drugs. *Curr. Pharm. Des.* **10,** 2391–2400
23. Zhou, T., Daugherty, M., Grishin, N. V., Osterman, A. L., and Zhang, H. (2000) Structure and mechanism of homoserine kinase: prototype for the GHMP kinase superfamily. *Structure* **8,** 1247–1257
24. Cheek, S., Zhang, H., and Grishin, N. V., (2002) Sequence and structure classification of kinases. *J. Mol. Biol.* **320,** 855–881
25. Sgraja, T., Alphey, M. S., Ghilagaber, S., Marquez, R., Robertson, M. N., Hemmings, J. L., Lauw, S., Rohdich, F., Bacher, A., Eisenreich, W., Illarionova, V., and Hunter, W. N. (2008) Characterization of Aquifex aeolicus 4-diphosphocytidyl-2C-methyl-d-erythritol kinase–ligand recognition in a template for antimicrobial drug discovery. *FEBS J.* **275,** 2779–2794
26. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326
27. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
28. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
29. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. Allogr. D Biol. Crystallogr.* **60,** 2126–2132
30. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger,T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **58,** 1948–1954
31. Laskowski, R., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26,** 283–291
32. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA, USA
33. Bernal, C., Mendez, E., Terencio, J., Boronat, A., and Imperial, S. (2005) A spectrophotometric assay for the determination of 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase activity. *Anal. Biochem.* **340,** 245–251
34. Matthews, B. W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33,** 491–497
35. Bonanno, J. B., Edo, C., Eswar, N., Pieper, U., Romanowski, M. J., Ilyin, V., Gerchman, S. E., Kycia, H., Studier, F. W., Sali, A., and Burley, S. K. (2001) Structural genomics of enzymes involved in sterol/isoprenoid biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 12896–12901
36. Tuteja, N., and Tuteja, R. (2004) Unraveling DNA helicases. Motif, structure, mechanism, and function. *Eur. J. Biochem.* **271,** 1849–1863
37. Holm, L., and Sander, C. (1996) Alignment of three-dimensional protein structures: network server for database searching. *Methods Enzymol.* **266,** 653–662
38. Jirage, D., Keenan, S. M., and Waters, N. C., Exploring novel targets for antimalarial drug discovery: plasmodial protein kinases. *Infect. Disord. Drug Targets* **10,** 134–146
39. Miallau, L., Alphey, M. S., Kemp, L. E., Leonard, G. A., McSweeney, S. M., Hecht, S., Bacher, A., Eisenreich, W., Rohdich, F., and Hunter, W. N. (2003) Biosynthesis of isoprenoids: crystal structure of 4-diphosphocytidyl-2C-methyl-D-erythritol kinase. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 9173–9178

Received for publication October 30, 2010.
Accepted for publication January 13, 2011.