Mutation of Archaeal Isopentenyl Phosphate Kinase Highlights Mechanism and Guides Phosphorylation of Additional Isoprenoid Monophosphates

Nikki Dellas[†,‡] and Joseph P. Noel[†,*]

[†]Howard Hughes Medical Institute, Salk Institute for Biological Studies, Jack H. Skirball Center for Chemical Biology and Proteomics, 10010 Torrey Pines Road, La Jolla, California 92037 and [‡]Department of Chemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093

Isopentenyl diphosphate (IPP) and its isomeric partner dimethylallyl diphosphate (DMAPP) are precursors for a diverse collection of primary and secondary isoprenoid metabolites in all organisms. Following its formation, successive units of IPP are used together either with DMAPP, formed by the action of types I or II IPP isomerases, or with the IPP extended isoprenoid diphosphate chain, to biosynthesize C10, C15, or C20 oligoprenyl diphosphates known as geranyl diphosphate (GPP), farnesyl diphosphate (FPP), and geranylgeranyl diphosphate (GGPP), respectively, as well as larger isoprenoid diphosphates. In plants and some microorganisms, GPP, FPP, and GGPP also serve as starting materials for the biosynthesis of a large class of specialized and often cyclic terpene hydrocarbons (1). FPP is the most ubiquitous of the three isoprenoid diphosphate building blocks, as it resides at the juncture of bifurcating branches of the general isoprenoid biosynthetic pathway leading to both primary and secondary metabolites. Squalene, hopanoids, and steroids, serve as critical components of cellular membranes and, in the case of steroids, also serve as transcription modulators through nuclear hormone receptor engagement (2, 3). Moreover, dolichols play essential roles in N-glycosylation and membrane anchorage of sugars in eukaryotes and archaea (4). The 20-carbon GGPP molecule functions as the precursor to all carotenoids, the latter of which provides photoprotection in plants, fungi, algae, bacteria, and some archaea (5, 6). Interestingly, GGPP also is a precursor to the isoprenoid-derived hydrocarbon moiety of lipids that is present exclusively in archaea (see ref (7) for a review on archaeal lipids).

Over the last two decades, two distinct pathways have been characterized that biosynthesize IPP and

ABSTRACT The biosynthesis of isopentenyl diphosphate (IPP) from either the mevalonate (MVA) or the 1-deoxy-D-xylulose 5-phosphate (DXP) pathway provides the key metabolite for primary and secondary isoprenoid biosynthesis. Isoprenoid metabolism plays crucial roles in membrane stability, steroid biosynthesis, vitamin production, protein localization, defense and communication, photoprotection, sugar transport, and glycoprotein biosynthesis. Recently, an alternative branch of the MVA pathway was discovered in the archaeon *Methanocaldococcus jannaschii* involving a small molecule kinase, isopentenyl phosphate kinase (IPK). IPK belongs to the amino acid kinase (AAK) superfamily. *In vitro*, IPK phosphorylates isopentenyl monophosphate (IP) in an ATP and Mg^{2+}-dependent reaction producing IPP. Here, we describe crystal structures of IPK from *M. jannaschii* refined to nominal resolutions of 2.0−2.8 Å. Notably, an active site histidine residue (His60) forms a hydrogen bond with the terminal phosphate of both substrate and product. This His residue serves as a marker for a subset of the AAK family that catalyzes phosphorylation of phosphate or phosphonate functional groups; the larger family includes carboxyl-directed kinases, which lack this active site residue. Using steady-state kinetic analysis of H60A, H60N, and H60Q mutants, the protonated form of the N_{ε_2} nitrogen of His60 was shown to be essential for catalysis, most likely through hydrogen bond stabilization of the transition state accompanying transphosphorylation. Moreover, the structures served as the starting point for the engineering of IPK mutants capable of the chemoenzymatic synthesis of longer chain isoprenoid diphosphates from monophosphate precursors.

*Corresponding author, noel@salk.edu.

DMAPP, namely the mevalonate (MVA) pathway and the more recently discovered 1-deoxy-D-xylulose 5-phosphate (DXP) pathway (8). The MVA pathway is utilized by animals, plants (cytosol), fungi, and certain bacteria, while the DXP pathway resides in plant plastids, a number of eubacteria, cyanobacteria, and certain parasitic organisms (9). In archaea, orthologs for almost all of the genes encoding the MVA pathway are present except for two; interestingly, the last two genes
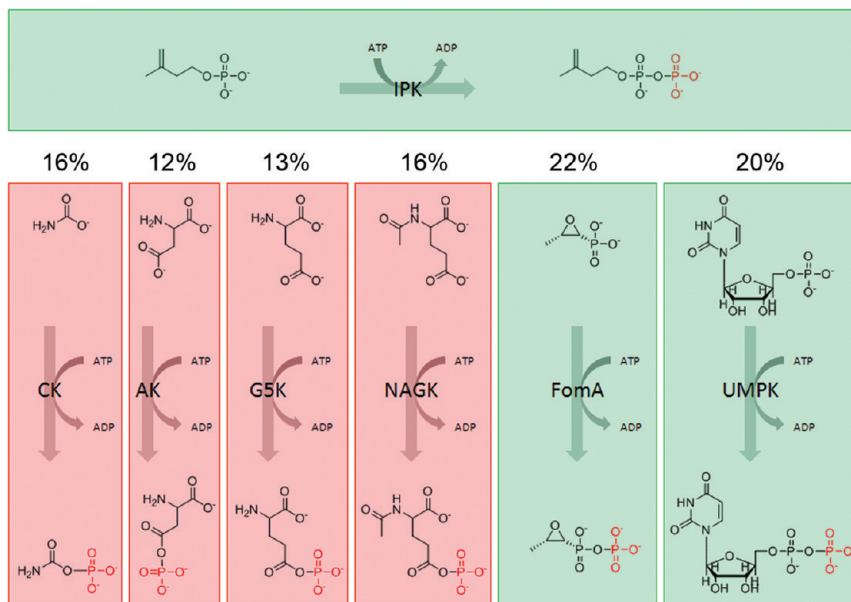

Figure 1. The amino acid kinase (AAK) family members. Isopentenyl phosphate kinase (IPK) reaction depicted across the top. Representative family members displayed from left to right: carbamate kinase (CK), aspartokinase (AK), glutamate-5-kinase (G5K), *N*-acetyl-L-glutamate kinase (NAGK), fosfomycin resistance kinase (FomA), and uridine monophosphate kinase (UMPK). The percent sequence identities relative to IPK are listed above each enzyme. Reactions shaded green utilize a phosphate or phosphonate phosphoryl acceptor, while the reactions shaded red utilize carbamate or carboxylate groups as phosphate acceptors.

encoding phosphomevalonate kinase and diphosphomevalonate decarboxylase appear to be missing from the genomes of almost all archaea. For this reason, the isoprenoid pathway in archaea is referred to as "The Lost Pathway" (10). In 2006, Grochowski *et al.* discovered an enzyme and its associated gene in the archaeon *Methanocaldococcus jannaschii* that belongs to the larger family of amino acid kinases (AAK) but catalyzes the ATP-dependent phosphorylation of IP, thereby producing IPP (11). This enzyme, named isopentenyl phosphate kinase (IPK), appears to be a starting point for the functional reconstruction of The Lost Pathway, representative of a completely unexpected biosynthetic variation of the MVA pathway.

IPK shares significant sequence homology with proteins in the AAK superfamily (Pf000696, Figure 1). Members of this family employ $Mg^{2+}-ATP$ to catalyze phosphorylation of carboxylate, carbamate, phosphonate, or phosphate functional groups. Here, crystal structures of IPK from *M. jannaschii* are presented in the 'apo' form and in complex with substrate (IP) and product (IPP). These structures allow for rational mutagenesis and bio-

chemical analyses of residue(s) near the phosphate moiety and the isopentenyl tail of IP, respectively. Mutation of a residue near the phosphate of IP demonstrates a key role for His-directed hydrogen bonding in the phosphorylation of phosphate or phosphonate groups. Mutation of residues near the isoprenyl moiety of IP establishes IPK as a starting point for engineering the phosphorylation of alternative phosphate/phosphonate bearing small molecules, including geranyl monophosphate (GP) and farnesyl monophosphate (FP). This sets the stage for a multitude of applications in chemoenzymatic syntheses including diphosphate analog synthesis, low cost radio-labeling of isoprenoid diphosphates with ^{32}P or ^{35}S containing β-phosphates, and the possible *in vivo* recycling of isoprenoid monophosphates formed upon FPP up-regulation and degradation in heterologous hosts.

RESULTS AND DISCUSSION

Three-Dimensional Architecture. IPK represents the newest member of the AAK superfamily to be structurally determined. The overall fold, commonly referred to

as the open αβα sandwich, was first discovered in carbamate kinase from *E. faecalis* (*12*). IPK is architecturally most similar to fosfomycin resistance kinase (FomA) from *S. wedmorensis* with a root-mean-square deviation (rmsd) of 2.0 Å for superimposed backbone atoms (NH−Cα−C) and a sequence identity of 22%. However, it shares the highest sequence identity, 25%, with uridine monophosphate kinase (UMPK) from *A. fulgidus*. Two subdivisions of the AAK superfamily exist, referred to here as the phosphate and carboxylate subdivisions, respectively. Enzymes in the phosphate subdivision, including IPK, FomA, and UMPK, catalyze phosphorylation of a phosphate or phosphonate moiety. Enzymes in the carboxylate subdivision, including carbamate kinase, *N*-acetylglutamate kinase (NAGK), aspartokinase, and glutamate-5-kinase, catalyze phosphorylation of a carbamate or carboxylate group (Figure 1).

Like all other AAKs, IPK adopts a dimeric quaternary structure, and each monomer folds into structurally distinct N- and C-terminal domains (Figure 2). The N-terminal domain, spanning residues 1−171, binds the nucleophilic phosphate group (IP in IPK). The C-terminal domain, spanning residues 171−260, coordinates a Mg^{2+} ion and binds the phosphate donor ATP. Although all attempts to crystallize *M. jannaschii* IPK with ATP, ADP, AMPPNP, and a variety of other analogs have thus far been unsuccessful, the location of the nucleotide-binding site is structurally conserved among all family members, affording a reasonable model for ATP binding. Each monomer of IPK contains 16 β-strands, 8 α-helices, and 1 3_{10} helix. The core of the open αβα sandwich, represented by 8 β-strands, β14, β16, β15, β11, β1, β2, β8, and β5, resides between 4 α-helices on one side, αF, αH, αE, and αD, and 3 on the other, αG, αA, and αC (Figure 2). Four β-hairpins, one α-helix, and one 3_{10} helix (η1) decorate the periphery of the central β-sandwich. Three of the hairpins, β3−β4, β6−β7, and β9−β10, lie in the N-terminal domain and surround the back wall and one side of the IP binding pocket; the αB helix covers the remaining side of the isopentenyl-binding surface. The fourth β-hairpin, β12−β13, located within the C-terminal domain, resides in close proximity to the expected location for the adenine ring of ATP. Finally, the 3_{10} helix links one end of the central β5 strand and the β6−β7 hairpin. Although the β1−αA junction is depicted as a loop in Figure 2, it also adopts a helical structure in some cases.

The IPK crystalline dimer is consistent with its oligomeric state deduced by gel filtration chromatography (Supplementary Figure 1). The dimer orients around a noncrystallographic two-fold axis. This dyadic axis sits perpendicular to the extended β-sheet spanning the length of the dimer (16 β-strands with 8 β-strands per monomer). Although every AAK family member utilizes a similar dimerization interface, each dimer is unique in that its monomers orient differently with respect to one another (*13*). The IPK dimer closely resembles that of UMPK, with the αC helices from each monomer crossing one another at an angle of 190° (*13*) (Figure 2, panel b). In IPK, this interface is comprised of 8 charged hydrogen bonds and 29 noncharged hydrogen bonds with 14 residues participating in van der Waals interactions (*14*). The majority of the hydrogen-bonding interactions stitch together three structural motifs: (i) the αC helices of each monomer; (ii) the αD helix of one monomer and the β9−β10 hairpin of the dyad related monomer; and (iii) the 3_{10} helix of one monomer and β5 of its dyadic partner. Hydrophobic interactions between the two monomers include residues from the αC and αD helices, the 3_{10} helix, and the β4, β5, β6, β8, β9, and β10 strands. These residues form an intimate hydrophobic interface further cementing the monomers together and burying a considerable amount of accessible surface area (1869 $Å^2$ of buried surface area per monomer).

Active Site Architecture. The refined 'apo' structure contains two active site sulfate molecules bound per monomer. One sulfate superimposes onto the position of the monophosphate of IP in the IP-bound structure. The second sulfate is present in all structures determined thus far and lies in the approximate location of the β-phosphate of ATP observed in other structures of AAK family members (*15−17*). This second sulfate ion is in close proximity to Gly9, Lys6, Lys221, and Thr179 (Figure 3). The equivalent residues in other AAK family members stabilize the β-phosphate of ADP or ATP analogues (PDBIDs: 2hmf (*18*), 1ohb (*19*), 2j0w (*20*), 2bri (*16*), 3c1m (*21*), 3d41 (*15*), 1gs5 (*17*)). Therefore, the sulfate ion appears to serve as a reasonable spatial mimic of the β-phosphate of ATP and is referred to as the β-sulfate ion.

The structures of IPK in complex with IP and IPP define the secondary structural elements comprising the IP-binding pocket and include the β2−αB glycine-rich loop, the αB helix, the β3−β4 hairpin, the β4−αC loop, the N-terminal section of the αC helix, and the β9−β10 hair-
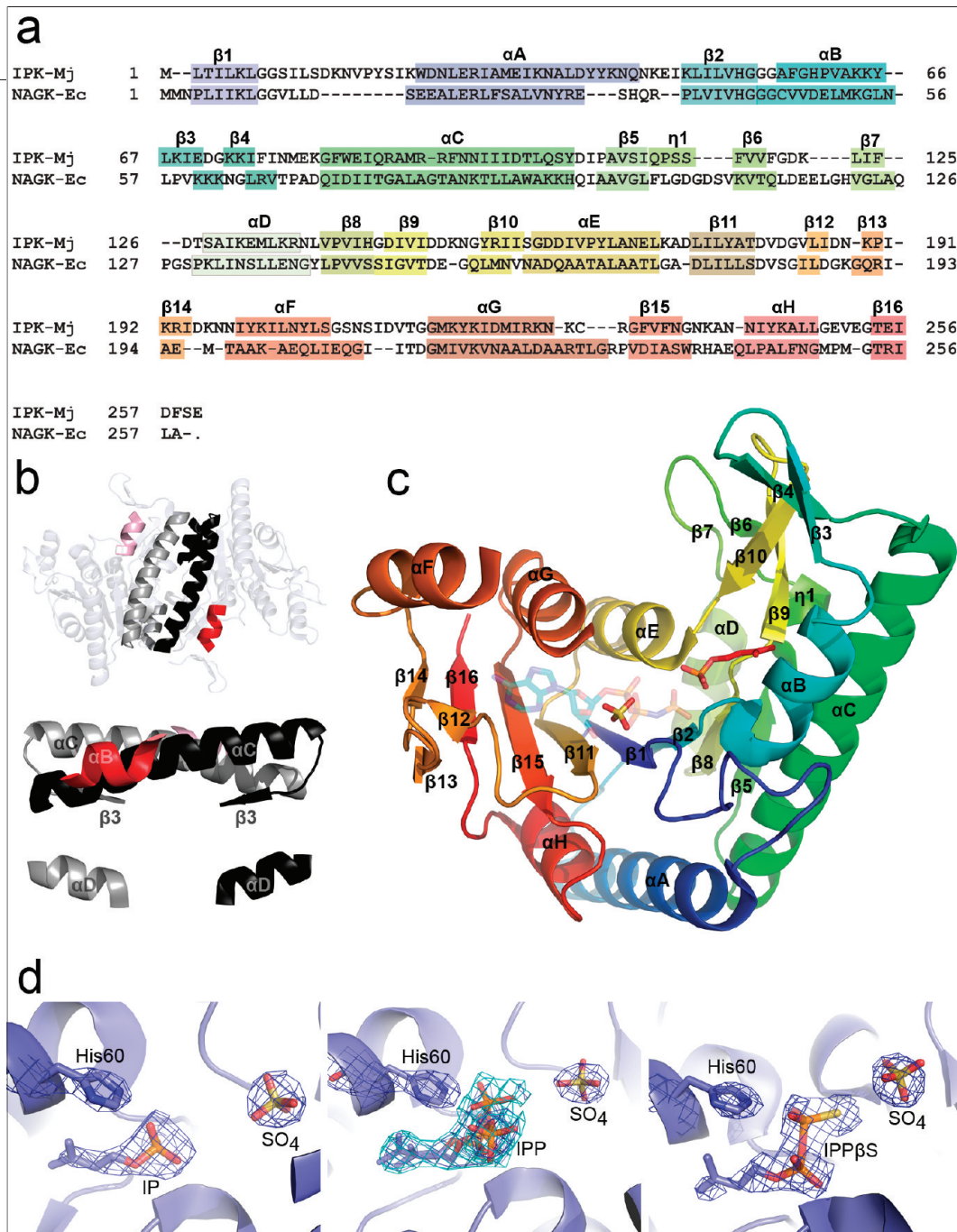
Figure 2. Primary sequence, tertiary architecture, and active site snapshots of IPK. a) Primary sequence of IPK from *M. jannaschii* aligned with *E. coli* NAGK. The color coding of each motif correlates with its color shown on the three-dimensional model. **b)** Global view of the IPK dimer (top) and a close-up view of the dimerization interface (bottom). Motifs positioned near the dimerization interface are gray (or pink) for one monomer and black (or red) for the other. **c)** Ribbon diagram of the IPK monomer. The structure is colored using a blue to red gradient from the N- to C-terminus. The C-terminal ATP-binding domain contains a β-sulfate residing in a location coinciding with the β-phosphate of ATP. The ATP analog AMPPNP is faintly colored and blended into the background (modeled from PDBID: 1gs5) and serves as a reference for the putative location of ATP in IPK. The crystallographically observed isopentenyl phosphate (IP) substrate is shown bound within the N-terminal domain. **d)** The active sites of IPK complexed with IP (left), IPP (middle), and IPPβS (right). Electron density surrounding each ligand (dark and light blue are contoured to 1σ and 0.6σ, respectively) shown as 2Fo−Fc omit electron density maps, where the ligands were removed before a round of refinement and subsequent phase and map calculations.
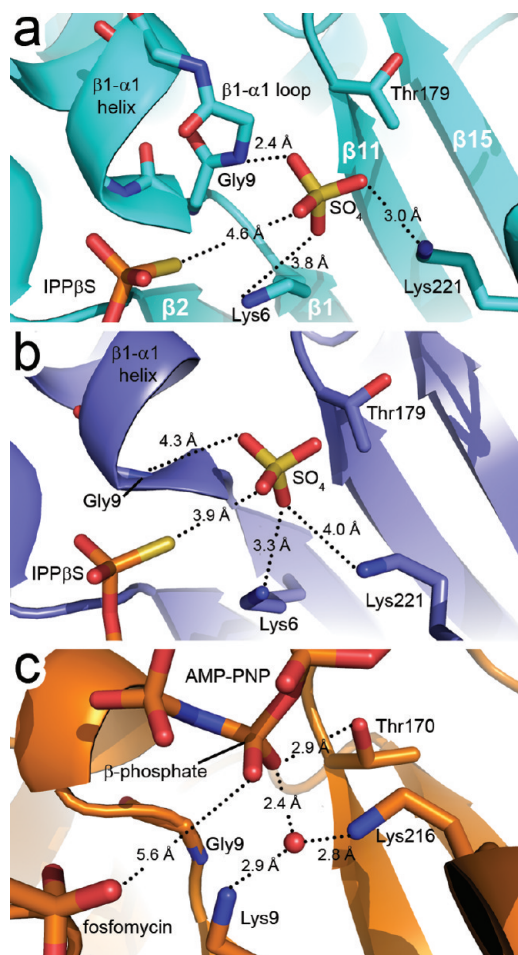
Figure 3. Comparative close-up views of the nucleotide phosphate-binding region of the IPK and FomA active sites. a) Monomer A of the IPK–IPPβS complex depicting the β-sulfate ion and the surrounding residues. b) Monomer B of the IPK–IPPβS complex oriented as in panel a. c) FomA complexed with the ATP analog AMPPNP and fosfomycin (PDB ID: 3d41) (15). As depicted here, the β-sulfate ion in both IPK monomers shares a similar position and interacts with the same residues as does the β-phosphate group on AMPPNP in FomA.

Figure 4. a) Close-up view of the N-terminal domain depicting the isopentenyl tail and the surrounding hydrophobic residues. The motifs surrounding the active site are colored as follows: β2−αB glycine-rich loop (red), αB helix (magenta), β3−β4 hairpin (yellow), β4−αC loop (green), N-terminal portion of the αC helix (cyan), and the β9−β10 hairpin (blue). Residues within van der Waals contact of the isopentenyl chain include Ile86, Met90, and Ile156. b) Dual conformation of the β1−αA loop in monomer A of the IPK−IP complex. One conformation places the loop close to the β2−αB loop and the IP substrate, while the other conformation places the loop in close proximity to the β-sulfate ion.

pin (Figure 4, panel a). The β2−αB loop is one of two conserved glycine-rich loops present throughout the AAK family. It is thought to contribute to charge neutralization in the transition state during phosphoryl transfer (15, 19). Notably, the orientation of the αB helix is conserved only in the phosphate division of the AAK superfamily (including IPK, FomA, and UMPK). In FomA, the αB helix orders when substrate is present but is other-
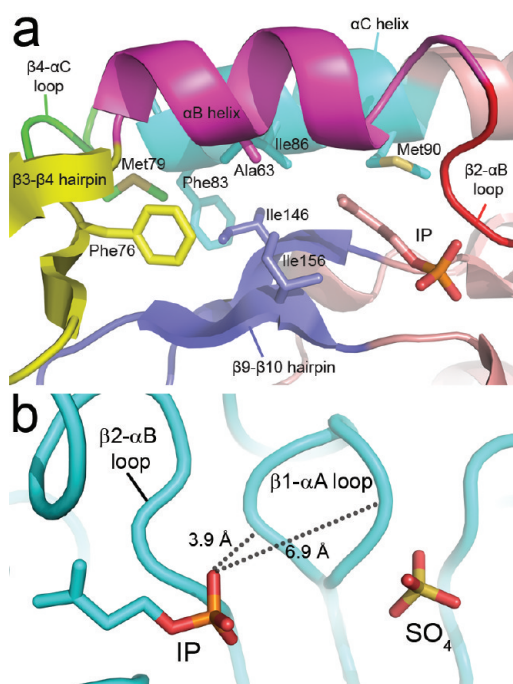
wise disordered (15). In contrast, the αB helix in IPK is ordered in both 'apo'- and IP-bound structures. Of even more limited familial distribution, the β3−β4 hairpin is present only in IPK and NAGK. In NAG-bound NAGK structures, the hairpin often exists in a closed conformation (17, 22); in contrast, all structures of IPK reveal the motif in an open conformation. Regardless, the hairpin may play a role in shielding the substrate-binding pocket from the surrounding solvent in both enzymes.

The branched C5 tail of the substrate resides in a pocket surrounded principally by hydrophobic residues, including Ala63, Phe76, Met79, Phe83, Ile86, Met90, Ile146, and Ile156, (Figure 4, panel a). The arrangement of residues within the cavity suggests that transmuta-
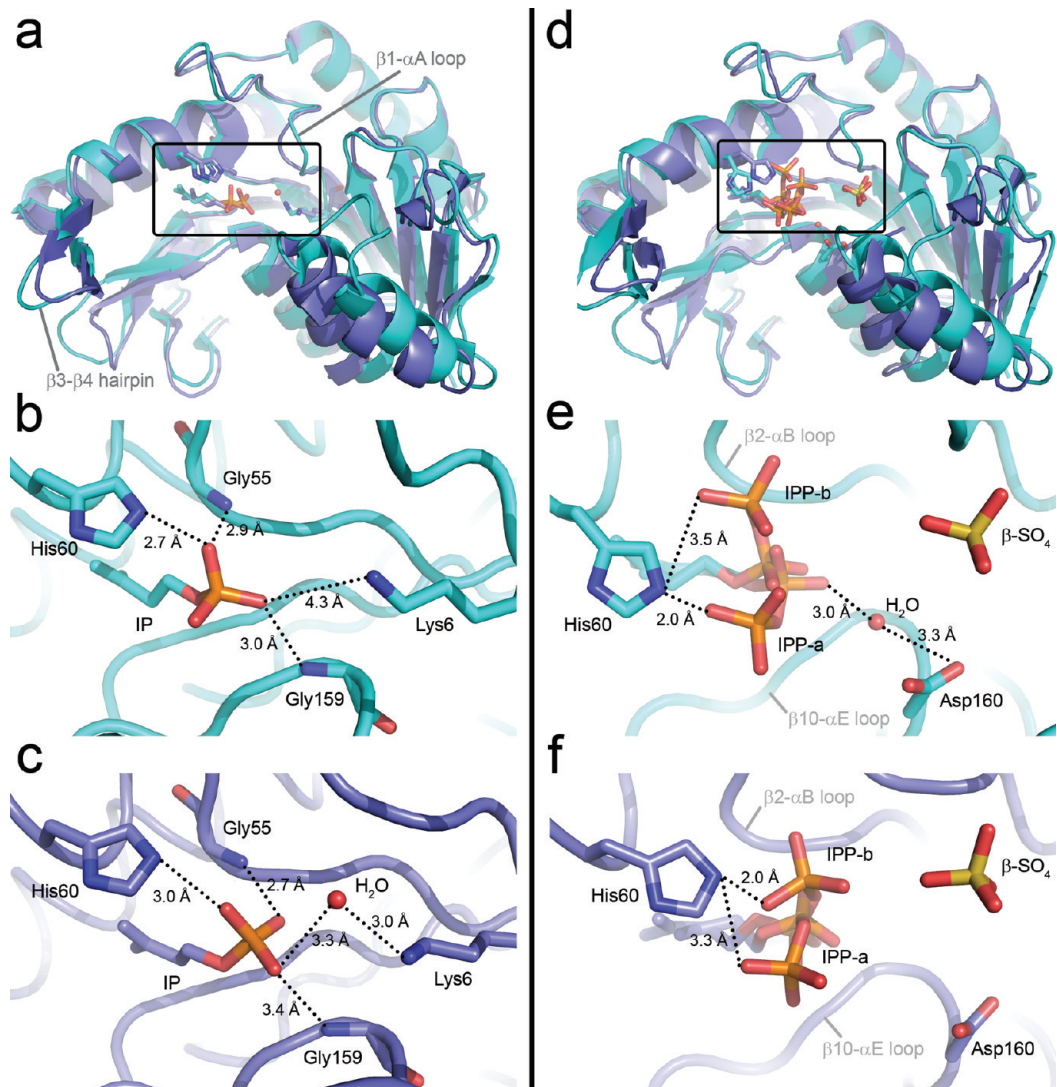
Figure 5. a) Tertiary structure superposition of monomers A and B of the IPK−IP complex. The rmsd between the two monomers is 1.31 Å. b−c) Close-up views of residues proximal to and hydrogen bonding with the α-phosphate of IP in monomers A (panel b) and B (panel c). In monomer B, a water molecule bridges the side-chain amino group of Lys6 and a nonbridging oxygen atom of the IP phosphate. d) Tertiary structure superposition of monomers A and B of the IPK−IPP complex. The rmsd between the two monomers is 1.39 Å. e−f) Views of the multiple conformers of IPP (labeled as IPP-a and IPP-b) in both monomers A (panel e) and B (panel f).

tion of the isopentenyl binding pocket to accommodate longer hydrocarbon chains may be relatively facile. The phosphate moiety of IP occupies the active site region between three short motifs including His60 and the β2−αB (residues 54−56) and β10−αE (residues 157−159) loops (Figure 5, panels a−c). In both monomers, the $N_{\varepsilon 2}$ atom of His60 and the N atoms of Gly55

and Gly159 when protonated could stabilize the three nonbridging O atoms on the phosphate of IP through hydrogen bonds (Figure 5, panels b−c). In monomer B, Lys6 also interacts with one of these nonbridging O-atoms indirectly through hydrogen-bonding interactions with an intervening water molecule; in monomer A, there is no water molecule to facilitate this interaction.

In monomer A, a loop at the β1—αA junction (Gly8—Leu12), residing near the active site, can adopt two distinct substrate-binding conformations based upon refinement of alternative conformations and the observed electron density. In one conformation, the loop lies near the active site β-sulfate ion, while in the other, the loop lies closer to the β2—αB loop (Figure 4, panel b). None of the residues in this loop participate in hydrogen-bonding interactions with IP; however, the dual binding conformations are not observed for the 'apo' structure, suggesting that loop movement is partially dependent on the presence of substrate. In monomer B, the loop adopts one binding conformation roughly equidistant between the two modes present in monomer A.

Multiple Conformations of IPP in a Single Active Site. The crystal structure of IPK with its product bound reveals that IPP adopts two distinct conformers designated conformers A and B (Figure 5, panels d—f). These conformers are similar except for the orientation of the β-phosphate group and the adjacent bridging O atom. In conformer A, these two moieties sit closer to the β10—αE loop, while in conformer B, they reside closer to the β2—αB loop (Figure 5, panels e—f). In both conformers, a nonbridging O atom of the β-phosphate group hydrogen bonds to the protonated $N\varepsilon_2$ atom of His60. A superposition of the two monomers (Figure 5, panel d), reveals that His60 sits in a different location in each monomer, which may reflect conformations of this residue that are dynamically accessible as the phosphorylation reaction proceeds.

In monomer A only, a water molecule rests between a nonbridging O atom from the α-phosphate of IPP and the carboxylate moiety of Asp160 (Figure 5, panel e). This water molecule is also found in substrate-bound structures of FomA kinase (PDBID 3d41) (15), E. coli NAGK (pdbID 1gs5) (17), P. furiosus UMPK (PDBID 2bmu) (16), and E. coli UMPK (pdbID 2bne) (23), and it is stabilized in a similar fashion in each of these structures. Asp160 of IPK is highly conserved among the AAK family and has been suggested to function as an active site base and a central organizing residue (15, 24).

As discussed previously, the β1—αA loop occupies two conformations in monomer A of the IPK—IP complex: one that places it in close proximity to the β2—αB loop, and another that interacts with the β-sulfate ion (Figure 4, panel b). In monomer A of the IPK—IPP complex, the latter conformation is the major binding mode

observed. The former conformation can also be seen in the electron density, although this binding mode is so subtle that it did not refine well and was not built into the final structure. This minor binding mode may, however, hold some significance, as it places Gly9 of the β1—αA loop in close proximity to the β-phosphate group in conformer B of IPP. The β1—αA loop is often reported to interact with the β- and γ-phosphate groups from ATP analogs; however, there are also examples of this loop interacting with the β-phosphate of the product (in UMPK from E. coli) (19, 23).

The catalytically relevant conformer for IPP is most likely conformer B. This hypothesis is supported by three pieces of information: (i) The β1—αA loop, which is thought to play a key role during phosphoryl transfer, accesses a minor binding mode that is in close proximity to the β-phosphate of conformer B (19); (ii) A superposition of UDP-bound UMPK from E. coli and IPP-bound IPK demonstrates that the phosphate moiety of UDP superimposes with conformer B of IPP (23); and (iii) The ATPγS/IP/Mg^{2+} complex structure (discussed below) exhibits clear electron density for a IPPβS molecule bound in a single conformation that superimposes with conformer B of IPP from the IPP-bound structure. Conformer A of IPP may, therefore, represent a post-reaction enzyme—product (EP) complex.

Product-Bound Active Site Containing IPPβS. When a crystal of IPK was soaked in a stabilizing solution containing IP, Mg^{2+}, and ATPγS, the product IPPβS was observed bound in the active site. This product resembles IPP except one of the nonbridging O atoms on the β-phosphate is clearly replaced with an S atom, as evidenced by the additional electron density associated with the β-thiophosphate. Notably, no electron density for the second product ADP is seen. This is the only structure determined where both substrates, IP and ATPγS, were soaked into the crystal resulting in a catalyzed reaction in the crystal lattice. Interestingly, this structure reveals only one binding mode for IPPβS consistent with the orientation of conformer B in the IPP-bound structure. In both monomers, Gly159 from the β10—αE loop stabilizes the α-phosphate group of IPPβS; the β-thiophosphate group remains in close proximity to His60. However, in monomer B (compared to monomer A), the substrate migrates to a position that shifts away from this residue and resides closer to the β-sulfate ion (Figure 3). The intermediate location of the β-phosphate group in monomer B coupled with the in-

TABLE 1. Kinetic Data for IPK-Mj Wild-Type and H60Q at 25°C

protein name	$K_{m, ATP}$ (μM)	$K_{m, IP}$ (μM)	k_{cat} (s^{-1})	$k_{cat}/K_{m,IP}$ (s^{-1}μM^{-1})
IPK—*M. jannaschii*	198.2 ± 32.7	4.30 ± 0.58	1.46 ± 0.03	0.34
IPK—*M. jannaschii* H60Q	559.3 ± 116.9	34.5 ± 7.2	0.040 ± 0.002	0.001

ferred heightened dynamics of certain loops within this monomer suggests that the monomer B structure depicts an earlier phase in the transphosphorylation reaction compared to monomer A.

His60 Plays a Key Role in Binding and Catalysis. From the results discussed above, it is evident that His60 plays an important role in both substrate and product sequestration. This binding role is accomplished through a hydrogen-bonding interaction between the protonated Nε_2 atom of His60 and a non-bridging O atom from the terminal phosphate group on either the substrate (IP) or the product (IPP). His60 was mutated to Asn, Gln, and Ala; the Asn and Gln mutations are isosteric with the protonated Nε_1 and Nε_2 groups on His, respectively. The three mutants were assayed at 25 °C using the pyruvate kinase/lactate dehydrogenase coupled reaction to detect kinase activity. Turnover for the H60A and H60N mutants was not detected. In contrast, the H60Q mutant, whose −NH$_2$ side-chain moiety mimics the protonated Nε_2 nitrogen of His60, catalyzed a measurable transphosphorylation of IP. Notably, the apparent $K_{m,IP}$ values for H60Q and wild-type IPK are 34.5 and 4.3 μM, respectively, while the apparent k_{cat} values are 0.04 s^{-1} and 1.46 s^{-1}, respectively. These experimental values yield an apparent catalytic efficiency, k_{cat}/K_m, 340 times lower for the H60Q IPK mutant (Table 1).

These steady-state kinetic results suggest several interpretations for the unique role of His60 in catalysis: (i) Since both H60A and H60N exhibit no measurable activity, while H60Q remains active (albeit at a fraction of wild-type activity), binding and/or catalysis appears to be dependent on the presence of a hydrogen bond donor that is isosteric with the protonated Nε_2 nitrogen of His60; 2) Given that the H60Q mutant possesses a higher apparent K_m than wild-type, His60 also appears to be important for ground-state binding. Additional flexibility in the Gln side chain relative to the imidazole group of His60 may hinder its ability to bind substrate as effectively as wild-type IPK; 3) The k_{cat}/K_m value is

more than 300 times higher for wild-type compared to that of the H60Q mutant, which again suggests that His60, through its added charge and lowered conformational flexibility, plays a role in stabilization of the more negatively charged transition state accompanying phosphoryl transfer.

Through comparison of the solved IPK structures, it is evident that His60 shifts from stabilizing the α-phosphate on the substrate IP to stabilizing the β-phosphate on the product IPP. In FomA, His58 (the equivalent residue to His60 of IPK) indirectly stabilizes the substrate through an intervening water molecule that is within hydrogen-bonding distance of both His58 and fosfomycin (*15*). In UMPKs, an arginine residue that aligns with His60 appears to serve two roles: (i) in bacterial UMPKs, this Arg interacts with GTP, which is an allosteric activator for all bacterial UMPKs (*25*); and (ii) in both bacterial and archaeal UMPKs, this Arg stabilizes the phosphate intermediate throughout the course of the reaction (*23, 26*). In comparisons of crystal structures of *E. coli* UMPK, Arg62 points in opposite directions to fulfill these two roles (*25*), indicating that the length and the conformational freedom of this residue are important for its functional versatility. It is therefore not surprising that a mutation to histidine at this equivalent position results in loss of GTP activation, thereby reducing the enzyme's catalytic activity (*27*).

The phosphate division of the AAK family encompasses the three enzymes discussed above, which are the only currently known family members that contain a residue aligning with His60 of IPK; the four family members in the carboxylate division do not possess a residue or a motif that structurally aligns with this region of IPK. Regardless of other roles that this residue may play, His60 in IPK, His58 in FomA, and the aligning arginine in all UMPKs most likely play a role in substrate/product binding and/or transition-state stabilization.

IPK Mutants Can Phosphorylate Oligoprenyl Monophosphates. As mentioned previously, the tail of the IP substrate is sequestered within a hydrophobic binding pocket (Figure 4, panel a). At the back of the
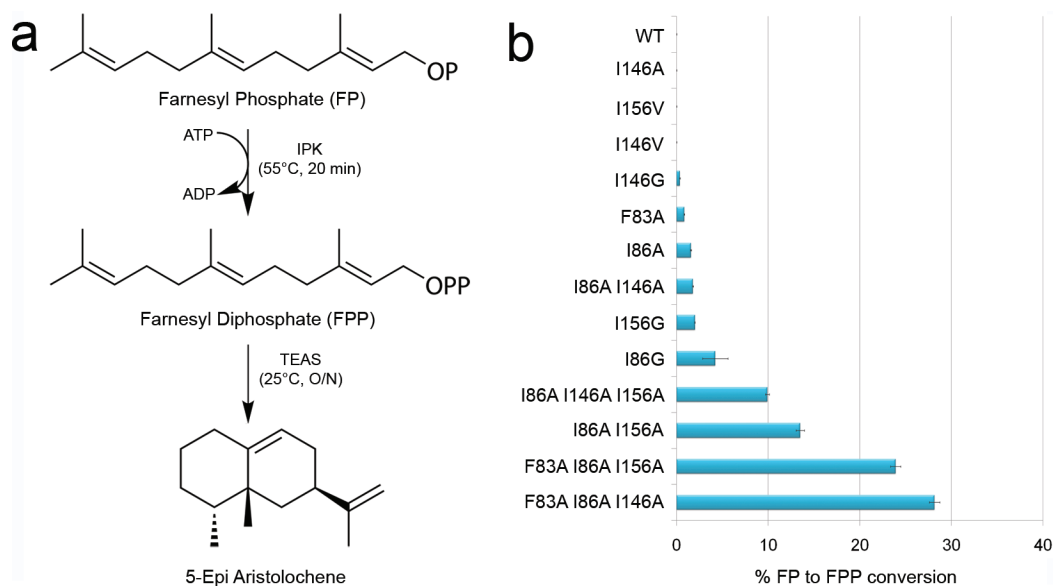
Figure 6. Farnesyl phosphate (FP) phosphorylation by IPK chain length mutants. a) The coupled IPK—sesquiterpene synthase reaction used to test for FP transphosphorylation. b) Comparative bar graph depicting several IPK tunnel mutants qualitatively tested for their ability to convert FP to FPP (expressed as a percentage of maximal production of 5-epi-aristolocene produced from IPK-generated FPP using wild-type IPK and identical concentrations of wild-type tobacco 5-epi-aristolochene synthase incubated for equivalent lengths of time).

pocket, several residues, Ile86, Ile146, and Ile156, can be mutated to smaller amino acids to accommodate the binding of ligands with extended carbon chains (such as GP and FP). With this idea in mind, several point mutants were generated including I86A, I86G, I146A, I146G, and I156A, and most of these were able to convert FP to FPP while the wild-type enzyme lacked an equivalent activity (Figure 6). These observations support the idea that mutation to smaller residues widens the cavity to allow for the binding of an extended isoprenoid tail, while mutation to bulkier residues hinders this ability. It was found that several double and triple mutant combinations displayed improved FP to FPP conversion by an order of magnitude compared to the single mutants, providing evidence that a deeper or larger cavity is more effective for FP binding and catalysis (Figure 6). It is also reassuring that these mutations are contextually dependent, meaning that mutations at the very back of the active site (at position 83) are not effective unless they are present with mutations closer to the front of the active site (at positions 86, 146, or 156).

Modeling of an FP molecule within the active site of IPK suggests that a C15 isoprenyl tail can orient in several different directions without introducing a large num-ber of steric clashes. These various orientations could be explored by mutating the appropriate amino acid side chains (to relieve putative steric clashes) and by utilizing a high-throughput coupled assay to test each mutant for its ability to convert FP to FPP. The coupled kinase/terpene synthase assay (Figure 6) is ideal for rapid and qualitative analysis, which is necessary here since a large number of mutants must be screened to explore all possible FP orientations and mutant combinations. Thus far, we have obtained active mutants that were designed to accommodate one putative FP orientation. Crystal structures of FP-bound IPK mutants will assist in the future design of more robust mutants. Finally, exploration of the above mutations in the context of a mesophilic ortholog should afford improvement in kinetic activity for new phosphate-bearing substrates; a more accurate picture of dynamic behavior can be painted for a mesophilic IPK being studied at ambient temperature rather than thermophilic IPK (used here) being studied at ambient temperature.

CONCLUSIONS

Isopentenyl phosphate kinase (IPK) from *M. jannaschii* is the newest member of the large amino acid kinase

(AAK) family to be structurally characterized. The phosphate division of the AAK family is comprised of three proteins [IPK, fosfomycin resistance kinase (FomA), and uridine monophosphate kinase (UMPK)] that exclusively align with one another along their αB helices. More importantly, they all contain a superimposable residue at position 60 (in IPK) that indirectly or directly stabilizes the terminal phosphate group of the substrate or product. Using the His60 marker, we have been able to identify putative IPK homologues from a number of phylogentically diverse eukarya. Work to be presented elsewhere is focused on *in vitro* and *in vivo* analyses of these latter proteins, given that, in most cases, the organisms in question also contain the full complement of predicted mevalonate (MVA) pathway enzymes.

Finally, we have shown that our initial goal to rationally engineer IPK to accept longer chain substrates is possible. These mutants along with an expanded mutational analysis in IPKs from mesophilic hosts can serve a variety of applications. For example, they can be used to recycle isoprenyl monophosphates, which are thought to be one possible *in vivo* byproduct of farnesyl diphosphate (FPP) degradation (through the action of an alkaline phosphatase) (28, 29). This recycling mechanism would be useful in an *in vivo* system designed to overproduce isoprenoids (such as terpenes) by means of the MVA pathway in a fungal or bacterial host. The IPK chain-length mutants would also be useful in the chemo-enzymatic synthesis of radio-labeled geranyl diphosphate (GPP) or FPP as well as a variety of other analogs including fluorescently tagged isoprenyl tails (30, 31). We have demonstrated that IPK can be rationally engineered to accept and phosphorylate oligoprenyl monophosphate substrates, such as FP. Future work with IPK from *M. jannaschii* (and with orthologous IPKs) will focus on redesigning the enzyme(s) such that they can bind and more efficiently turn over a variety of bulky GP and FP analogs.

METHODS

Activity Assays and Steady-State Kinetic Analyses. All specific activity and kinetic measurements were performed using the pyruvate kinase−lactate dehydrogenase coupled assay (32). The reaction in 200 μL includes 7 U of pyruvate kinase, 10 U of lactate dehydrogenase, 2 mM of phosphoenolpyruvate, 0.16 mM of NADH, 50 mM of Tris−HCl, pH of 8.0, 100 mM of KCl, 8 mM of $MgCl_2$, and varying concentrations of ATP or IP (purchased from Larodan Fine Chemicals and Isoprenoids, LLC). When IP was varied, the concentration of ATP was fixed at 4 mM. When ATP was varied, the concentration of IP was fixed at 100 μM for wild-type IPK and 500 μM for the H60Q mutant. The reaction was initiated by the addition of IPK (0.15 μg mL^{-1} final concentration) and followed by observing the depletion of NADH at 340 nm, expressed as $\Delta(AU_{340})/\Delta t$ and converted to $\Delta(ADP)/\Delta t$. These values were plotted against substrate concentration in GraphPad Prism (Version 5.01 for Windows) to compute the kinetic parameters k_{cat} and K_m, using the "nonlinear regression enzyme kinetic analysis" option.

Kinase/Terpene Synthase Coupled Assay for Chain-Length Mutants. The coupled assay consists of two steps: the kinase reaction followed by the terpene synthase reaction (Figure 6). The 50 μL of kinase reaction includes 4 mM of ATP, 8 mM of $MgCl_2$, 1 mM of FP (or GP), and 50 mM of Tris−HCl, pH of 8.0. The reaction was initiated by the injection of IPK (1 μM of the final concentration), incubated at 55 °C for 20 min, and then cooled on ice for 10 min. The subsequent 500 μL of terpene cyclase reaction includes 10 μL of the aforementioned IPK reaction, 8 mM of $MgCl_2$, 30 μg of tobacco 5-epi aristolochene synthase (TEAS), and 50 mM of Tris−HCl, pH 8.0. Each sample was overlaid with an equal volume of ethyl acetate, incubated at RT overnight, and vortexed for terpene extraction and subsequent injection onto the GC-MS (Hewlett-Packard, 6890/5973 system) equipped with a HP-5MS column (0.25 mm × 30 m × 0.25 μm). The method employed is similar to that reported in O'Maille *et al.* (33) with an injection temperature reduced from 250 to 200 °C.

Structure Solution and Refinement. Data were processed and scaled with XDS (34). The reduced single anomalous diffraction (SAD) data from the IPK 'apo' crystal treated with ethyl mercuric phosphate was used in SOLVE (35) to locate and refine the positions of two Hg atoms per asymmetric unit, followed by phasing (mean figure of merit: 0.33). The program RESOLVE (35) was then employed to build 424 out of 520 residues into the SAD-derived model. Model building and phase improvement were accomplished using ARP/wARP (36, 37). The refined model was used as the starting model for the structure determination of IPK in complex with IP, IPP, and IPPβS. Simulated annealing and rigid-body refinements were performed in CNS version 1.2 for each structure before additional rounds of refinement in CNS version 1.2 (38) and CCP4 (36). COOT was used for map-model visualization and manual model building (39). Areaimol was used for a calculation of the buried surface area per monomer (36). PROCHECK was used to assess the quality of all models (40). The data and refinement statistics are listed in Table 2.

Additional programs used to view, analyze, and manipulate structure information include SSM Superpose, a program within COOT that superimposes the Cα atoms of one structure onto another generating an rmsd value (41), PyMOL, a molecular graphics program used to create images of the protein structure (42), and Adobe Photoshop CS4, used to label and manipulate images created with PyMOL.

Accession Codes 3K4O: IPK 'apo'; 3K52: IPK−IP complex; 3K4Y: IPK−IPP complex; and 3K56: IPK−IPPβS complex.

Gene cloning, protein expression, purification, crystallization and data collection are detailed in the Supporting Information.

TABLE 2. X-ray Diffraction Data Processing and Refinement Statistics

	IPK 'apo'	IPK−IP complex	IPK−IPP complex	IPK−IPPβS complex
PDB ID	3K4O	3K52	3K4Y	3K56
ligand (crystal drop)	none	IP + MgCl$_2$	IPP + MgCl$_2$	IP + ATPγS + MgCl$_2$
ligand (observed)	none	IP	IPP	IPPβS
Data Collection and Processing				
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
resolution (Å)	2.0	2.7	2.55	2.35
cell dimensions:				
a (Å)	76.05	77.86	78.09	77.68
b (Å)	99.61	100.80	99.23	100.24
c (Å)	87.60	87.32	87.41	87.79
$\alpha = \beta = \gamma$(°)	90	90	90	90
molecules in asymmetric unit	2	2	2	2
no. measured reflections[a]	280 086 (41 535)	109 680 (15 190)	128 795 (17 570)	195 455 (25 991)
no. unique reflections[a]	43 888 (6469)	18 614 (2598)	21 910 (3104)	27 728 (3786)
redundancy	6.38 (6.42)	5.89 (5.85)	5.87 (5.66)	7.05 (6.87)
R_{merge} (%)[ab]	6.9 (32.9)	8.4 (38.5)	7.9 (33.99)	7.5 (39.8)
completeness (%)[a]	95.6 (88.6)	95.5 (84.5)	95.5 (85.2)	93.9 (80.9)
$I/\sigma(I)$[a]	16.23 (4.58)	16.79 (3.70)	16.52 (4.34)	15.54 (3.84)
Refinement				
resolution range (Å)	50.0−2.0	50.0−2.7	50.0−2.55	50.0−2.35
no. reflections:				
working set	40 472	18 323	21 505	27 724
test set	2105	983	1141	1462
R_{work}/R_{free}[c]	0.223/0.241	0.224/0.286	0.224/0.287	0.220/0.259
no. atoms:				
protein	4079	4079	4074	4074
ligand	0	20	28	28
water	153	61	96	68
rmsd bond lengths (Å)[d]	0.008	0.022	0.022	0.022
rmsd bond angles (°)	1.3	1.98	1.99	1.99
mean B-factor (Å^2)	41.29	48.71	42.06	51.76
Ramachandran plot:				
% core	88.3	83.0	82.7	84.3
% allowed	11.1	15.7	16	14
% generously allowed	0.7	1.3	1.3	1.8
% disallowed	0.0	0.0	0.0	0.0
refinement program	CNS	CNS, Refmac	CNS, Refmac	CNS, Refmac

[a]Values in parentheses represent data from the highest resolution shell. [b]$R_{merge} = \sum_{hkl} \sum_i I_i(hkl) - <I(hkl)>|/\sum_{hkl} \sum_i I_i(hkl)$. [c]$R_{factor} = \sum ||F_{obs}| - |F_{calc}||/\sum |F_{obs}|$. R_{work} is the R_{factor} calculated using all diffraction data included in the refinement. R_{free} is the R_{factor} calculated using the randomly chosen 5% of diffraction data that were not included in the refinement. [d]Rmsd is the root-mean-square deviation.

REFERENCES

1. Gershenzon, J., and Dudareva, N. (2007) The Function of Terpene Natural Products in the Natural World, *Nat. Chem. Biol. 3*, 408–414.

2. Novakova, Z., Surin, S., Blasko, J., Majernik, A., and Smigan, P. (2008) Membrane Proteins and Squalene-Hydrosqualene Profile in Methanoarchaeon Methanothermobacter Thermautotrophicus Resistant to N, N′-Dicyclohexylcarbodiimide, *Folia Microbiol. (Prague, Czech Repub.) 53*, 237–240.

3. Ourisson, G., Rohmer, M., and Poralla, K. (1987) Prokaryotic Hopanoids and Other Polyterpenoid Sterol Surrogates, *Annu. Rev. Microbiol. 41*, 301–333.

4. Eichler, J., and Adams, M. W. (2005) Posttranslational Protein Modification in Archaea, *Microbiol. Mol. Biol. Rev. 69*, 393–425.

5. Sieiro, C., Poza, M., de Miguel, T., and Villa, T. G. (2003) Genetic Basis of Microbial Carotenogenesis, *Int. Microbiol. 6*, 11–16.

6. Hemmi, H., Ikejiri, S., Nakayama, T., and Nishino, T. (2003) Fusion-Type Lycopene Beta-Cyclase from a Thermoacidophilic Archaeon Sulfolobus Solfataricus, *Biochem. Biophys. Res. Commun. 305*, 586–591.

7. Koga, Y., and Morii, H. (2007) Biosynthesis of Ether-Type Polar Lipids in Archaea and Evolutionary Considerations, *Microbiol. Mol. Biol. Rev. 71*, 97–120.

8. Rohmer, M. (1999) The Discovery of a Mevalonate-Independent Pathway for Isoprenoid Biosynthesis in Bacteria, Algae and Higher Plants, *Nat. Prod. Rep. 16*, 565–574.

9. Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000) Isoprenoid Biosynthesis: The Evolution of Two Ancient and Distinct Pathways Across Genomes, *Proc. Natl. Acad. Sci. U.S.A. 97*, 13172–13177.

10. Smit, A., and Mushegian, A. (2000) Biosynthesis of Isoprenoids Via Mevalonate in Archaea: The Lost Pathway, *Genome Res. 10*, 1468–1484.

11. Grochowski, L. L., Xu, H., and White, R. H. (2006) Methanocaldococcus Jannaschii Uses a Modified Mevalonate Pathway for Biosynthesis of Isopentenyl Diphosphate, *J. Bacteriol. 188*, 3192–3198.

12. Marina, A., Alzari, P. M., Bravo, J., Uriarte, M., Barcelona, B., Fita, I., and Rubio, V. (1999) Carbamate Kinase: New Structural Machinery for Making Carbamoyl Phosphate, the Common Precursor of Pyrimidines and Arginine, *Protein Sci. 8*, 934–940.

13. Marco-Marin, C., Gil-Ortiz, F., Perez-Arellano, I., Cervera, J., Fita, I., and Rubio, V. (2007) A Novel Two-Domain Architecture within the Amino Acid Kinase Enzyme Family Revealed by the Crystal Structure of Escherichia Coli Glutamate 5-Kinase, *J. Mol. Biol. 367*, 1431–1446.

14. Krissinel, E., and Henrick, K. (2007) Inference of Macromolecular Assemblies from Crystalline State, *J. Mol. Biol. 372*, 774–797.

15. Pakhomova, S., Bartlett, S. G., Augustus, A., Kuzuyama, T., and Newcomer, M. E. (2008) Crystal Structure of Fosfomycin Resistance Kinase FomA from Streptomyces Wedmorensis, *J. Biol. Chem. 283*, 28518–28526.

16. Marco-Marin, C., Gil-Ortiz, F., and Rubio, V. (2005) The Crystal Structure of Pyrococcus Furiosus UMP Kinase Provides Insight into Catalysis and Regulation in Microbial Pyrimidine Nucleotide Biosynthesis, *J. Mol. Biol. 352*, 438–454.

17. Ramon-Maiques, S., Marina, A., Gil-Ortiz, F., Fita, I., and Rubio, V. (2002) Structure of Acetylglutamate Kinase, a Key Enzyme for Arginine Biosynthesis and a Prototype for the Amino Acid Kinase Enzyme Family, during Catalysis, *Structure 10*, 329–342.

18. Faehnle, C. R., Liu, X., Pavlovsky, A., and Viola, R. E. (2006) The Initial Step in the Archaeal Aspartate Biosynthetic Pathway Catalyzed by a Monofunctional Aspartokinase, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun. 62*, 962–966.

19. Gil-Ortiz, F., Ramon-Maiques, S., Fita, I., and Rubio, V. (2003) The Course of Phosphorus in the Reaction of N-Acetyl-L-Glutamate Kinase, Determined from the Structures of Crystalline Complexes, Including a Complex with an AlF(4)(·) Transition State Mimic, *J. Mol. Biol. 331*, 231–244.

20. Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006) Structures of R- and T-State Escherichia Coli Aspartokinase III. Mechanisms of the Allosteric Transition and Inhibition by Lysine, *J. Biol. Chem. 281*, 31544–31552.

21. Liu, X., Pavlovsky, A. G., and Viola, R. E. (2008) The Structural Basis for Allosteric Inhibition of a Threonine-Sensitive Aspartokinase, *J. Biol. Chem. 283*, 16216–16225.

22. Ramon-Maiques, S., Fernandez-Murga, M. L., Gil-Ortiz, F., Vagin, A., Fita, I., and Rubio, V. (2006) Structural Bases of Feed-Back Control of Arginine Biosynthesis, Revealed by the Structures of Two Hexameric N-Acetylglutamate Kinases, from Thermotoga Maritima and Pseudomonas Aeruginosa, *J. Mol. Biol. 356*, 695–713.

23. Briozzo, P., Evrin, C., Meyer, P., Assairi, L., Joly, N., Barzu, O., and Gilles, A. M. (2005) Structure of Escherichia Coli UMP Kinase Differs from that of Other Nucleoside Monophosphate Kinases and Sheds New Light on Enzyme Regulation, *J. Biol. Chem. 280*, 25533–25540.

24. Marco-Marin, C., Ramon-Maiques, S., Tavarez, S., and Rubio, V. (2003) Site-Directed Mutagenesis of Escherichia Coli Acetylglutamate Kinase and Aspartokinase III Probes the Catalytic and Substrate-Binding Mechanisms of these Amino Acid Kinase Family Enzymes and Allows Three-Dimensional Modelling of Aspartokinase, *J. Mol. Biol. 334*, 459–476.

25. Meyer, P., Evrin, C., Briozzo, P., Joly, N., Barzu, O., and Gilles, A. M. (2008) Structural and Functional Characterization of Escherichia Coli UMP Kinase in Complex with its Allosteric Regulator GTP, *J. Biol. Chem. 283*, 36011–36018.

26. Jensen, K. S., Johansson, E., and Jensen, K. F. (2007) Structural and Enzymatic Investigation of the Sulfolobus Solfataricus Uridylate Kinase shows Competitive UTP Inhibition and the Lack of GTP Stimulation, *Biochemistry 46*, 2745–2757.

27. Bucurenci, N., Serina, L., Zaharia, C., Landais, S., Danchin, A., and Barzu, O. (1998) Mutational Analysis of UMP Kinase from Escherichia Coli, *J. Bacteriol. 180*, 473–477.

28. Song, L. (2006) A Soluble Form of Phosphatase in Saccharomyces Cerevisiae Capable of Converting Farnesyl Diphosphate into E, E-Farnesol, *Appl. Biochem. Biotechnol. 128*, 149–158.

29. Coleman, J. E. (1992) Structure and Mechanism of Alkaline Phosphatase, *Annu. Rev. Biophys. Biomol. Struct. 21*, 441–483.

30. Rose, M. W., Rose, N. D., Boggs, J., Lenevich, S., Xu, J., Barany, G., and Distefano, M. D. (2005) Evaluation of Geranylazide and Farnesylazide Diphosphate for Incorporation of Prenylazides into a CAAX Box-Containing Peptide using Protein Farnesyltransferase, *J. Pept. Res. 65*, 529–537.

31. Hovlid, M. L., Edelstein, R. L., Henry, O., Ochocki, J., DeGraw, A., Lenevich, S., Talbot, T., Young, V. G., Hruza, A. W., Lopez-Gallego, F., Labello, N. P., Strickland, C. L., Schmidt-Dannert, C., and Distefano, M. D. (2010) Synthesis, Properties, and Applications of Diazotrifluropropanoyl-Containing Photoactive Analogs of Farnesyl Diphosphate Containing Modified Linkages for Enhanced Stability, *Chem. Biol. Drug Des. 75*, 51–67.

32. Lindsley, J. E. (2001) Use of a Real-Time, Coupled Assay to Measure the ATPase Activity of DNA Topoisomerase II, *Methods Mol. Biol. 95*, 57–64.

33. O'Maille, P. E., Chappell, J., and Noel, J. P. (2004) A Single-Vial Analytical and Quantitative Gas Chromatography-Mass Spectrometry Assay for Terpene Synthases, *Anal. Biochem. 335*, 210–217.

34. Kabsch, W. (1993) Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants, *J. Appl. Crystallogr. 26*, 795–800.

35. Terwilliger, T. (2004) SOLVE and RESOLVE: Automated Structure Solution, Density Modification and Model Building, *J. Synchrotron Radiat. 11*, 49–52.

36. (1994) The CCP4 Suite: Programs for Protein. *Acta Crystallogr. D* 50, 760–763.

37. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated Protein Model Building Combined with Iterative Structure Refinement, *Nat. Struct. Biol. 6*, 458–463.

38. Brunger, A. T. (2007) Version 1.2 of the Crystallography and NMR System, *Nat. Protoc. 2*, 2728–2733.

39. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905–921.

40. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures, *J. App. Cryst. 26*, 283–291.

41. Krissinel, E., and Henrick, K. (2004) Secondary-Structure Matching (SSM), a New Tool for Fast Protein Structure Alignment in Three Dimensions, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2256–2268.

42. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA.