

# X-ray Structures of Isopentenyl Phosphate Kinase

Mark F. Mabanglo<sup>†</sup>, Heidi L. Schubert<sup>‡</sup>, Mo Chen<sup>†,§</sup>, Christopher P. Hill<sup>‡</sup>, and C. Dale Poulter<sup>†,\*</sup>

<sup>†</sup>Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112 and <sup>‡</sup>Department of Biochemistry, University of Utah School of Medicine, 15 North Medical Drive, Salt Lake City, Utah 84112. <sup>§</sup>Present address: Department of Pharmacology, University of Pennsylvania School of Medicine, 135 John Morgan Building, 3620 Hamilton Walk, Philadelphia, PA 19104-6084.

Isoprenoid compounds constitute the largest and most structurally diverse family of natural products (1). The isoprenoid pathway is found in all organisms, where the molecules perform numerous functions, including roles as hormones (gibberellins, sterols), structural components in cell membranes (cholesterol, hopanoids), photosynthetic pigments (chlorophyll), visual pigments (retinal), insect pheromones (*cis*-verbenol), defensive agents (phytoalexins), electron carriers (ubiquinone), and protein modifiers for membrane targeting (farnesyl and geranylgeranyl diphosphate) (2–4). Many isoprenoid compounds are currently employed or under investigation as antibacterial, antifungal, and anticancer drugs. The carbon skeletons of isoprenoid molecules are constructed from two 5-carbon building blocks, isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP). Addition of IPP to DMAPP and to a succession of allylic diphosphates of increasing chain length produces geranyl diphosphate (C<sub>10</sub>), farnesyl diphosphate (C<sub>15</sub>), geranylgeranyl diphosphate (C<sub>20</sub>), and longer polyprenyl diphosphates (5, 6). These compounds lie at branch points in the isoprenoid pathway that lead to the huge variety of isoprenoid compounds found in nature (2, 7).

Two orthogonal routes are known for the biosynthesis of IPP and DMAPP. Acetyl CoA is the source of the carbon atoms in isoprene units synthesized in the mevalonate (MVA) pathway, whereas glyceraldehyde phosphate and pyruvate are used in the methylerythritol phosphate (MEP) pathway. Animals, fungi, and *Archaea* utilize the MVA pathway exclusively. Plants utilize both pathways in distinct cellular compartments, the MVA pathway in the cytoplasm and the MEP pathway in chloroplasts. Eubacteria employ either the MVA or MEP pathway, and in a few cases, both (8).

**ABSTRACT** Isoprenoid compounds are ubiquitous in nature, participating in important biological phenomena such as signal transduction, aerobic cellular respiration, photosynthesis, insect communication, and many others. They are derived from the 5-carbon isoprenoid substrates isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP). In *Archaea* and *Eukarya*, these building blocks are synthesized *via* the mevalonate pathway. However, the genes required to convert mevalonate phosphate (MP) to IPP are missing in several species of *Archaea*. An enzyme with isopentenyl phosphate kinase (IPK) activity was recently discovered in *Methanocaldococcus jannaschii* (MJ), suggesting a departure from the classical sequence of converting MP to IPP. We have determined the high-resolution crystal structures of isopentenyl phosphate kinases in complex with both substrates and products from *Thermoplasma acidophilum* (THA), as well as the IPK from *Methanothermobacter thermautotrophicus* (MTH), by means of single-wavelength anomalous diffraction (SAD) and molecular replacement. A histidine residue (His50) in THA IPK makes a hydrogen bond with the terminal phosphates of IP and IPP, poising these molecules for phosphoryl transfer through an in-line geometry. Moreover, a lysine residue (Lys14) makes hydrogen bonds with non-bridging oxygen atoms at P<sub>α</sub> and P<sub>γ</sub> and with the P<sub>β</sub>-P<sub>γ</sub> bridging oxygen atom in ATP. These interactions suggest a transition-state-stabilizing role for this residue. Lys14 is a part of a newly discovered “lysine triangle” catalytic motif in IPKs that also includes Lys5 and Lys205. Moreover, His50, Lys5, Lys14, and Lys205 are conserved in all IPKs and can therefore serve as fingerprints for identifying new homologues.

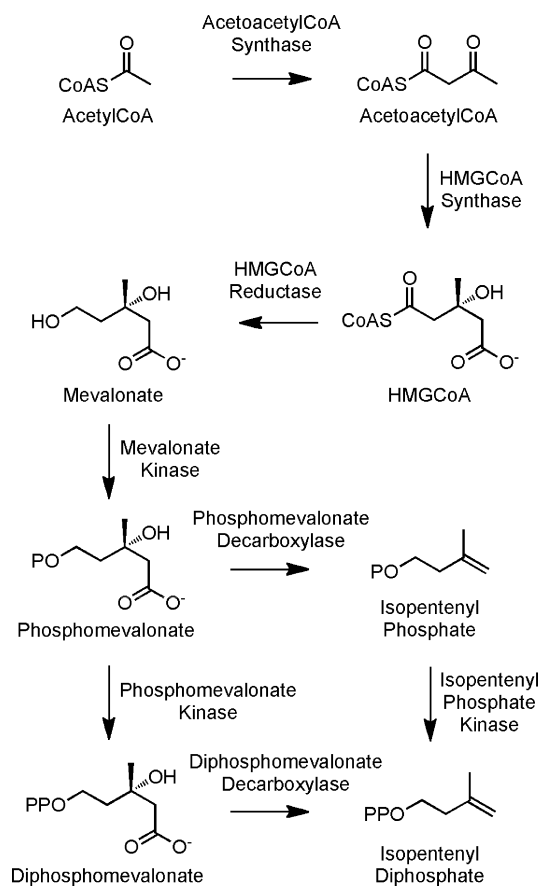
\*Corresponding author:  
poulter@chemistry.utah.edu.

Received for review February 4, 2010  
and accepted April 19, 2010.

Published online April 20, 2010

10.1021/cb100032g

© 2010 American Chemical Society



**Figure 1. Alternate route in the MVA pathway. In *Archaea*, the last two steps in the formation of IPP are reversed. Phosphomevalonate is decarboxylated to form IP, which is then phosphorylated to give IPP.**

The MVA pathway and the enzymes that catalyze the steps in the pathway are shown in Figure 1 (9). Four of the corresponding enzymes—acetoacetyl CoA synthase, HMG CoA synthase, HMG CoA reductase, and mevalonate kinase have been identified in *Archaea* (10). Furthermore, *Archaea* have a form of IPP isomerase, the type 2 enzyme, that is genetically unrelated to the type 1 IPP isomerase found in eukaryotes (10). Homologues for the enzymes that catalyze the last two steps in the MVA pathway, phosphomevalonate kinase (PMK) and diphosphomevalonate decarboxylase (PPMD), are generally missing in *Archaea*. There is evidence for homologues of PMK and PPMD in species of *Sulfolobus*, while homologues for PPMD, but not PMK, have been reported in *Halobacteriales* and *Thermoplasmatales* (10).

However, these PPMD homologues do not align with *Sulfolobus* PPMD but rather cluster with bacterial PPMDs. Smit and Mushiegan have proposed candidate PMK and PPMD genes in *Archaea* (11), although these assignments have not been tested experimentally.

Recently, Grochowski *et al.* reported that the translated protein from MJ0044 in *Methanocaldococcus jannaschii* (MJ) catalyzes the ATP-dependent phosphorylation of isopentenyl phosphate (IP) rather than mevalonate or mevalonate phosphate, and the protein was assigned as an IP kinase (IPK) (12). This *M. jannaschii* protein and homologous proteins from *Thermoplasma acidophilum* (THA) and *Methanothermobacter thermautotrophicus* (MTH) have high catalytic efficiencies for phosphorylation of IP ( $k_{\text{cat}}/K_m \approx 10^6 \text{ M}^{-1} \text{ s}^{-1}$ ) (13) that are consistent with those of established enzymes in the isoprenoid pathway. Moreover, MJ0044 does not cluster with other MJ genes involved in the MVA pathway, although homologues of MJ0044 are found in the genomes of representatives from all of the families in *Archaea* (13). Grochowski *et al.* suggested a variation of the “classic” MVA pathway in *Archaea* where mevalonate phosphate is decarboxylated to give IP, which is then converted to IPP by IPK (12).

In this paper, we present crystal structures of IPKs from THA and MTH. Structures were obtained for THA IPK in ternary complexes with the substrates, IP and ATP, and with the products, IPP and ADP. These structures reveal active site residues that participate in the phosphoryl transfer reaction.

## RESULTS AND DISCUSSION

We have determined structures of IP kinases from two archaeal species, THA and MTH, with crystals obtained under a variety of conditions. All crystals were grown in the presence of IP and ATP. The THA IPK structure was determined by single wavelength anomalous diffraction (SAD) using crystals of selenomethionine-substituted protein grown in 0.1 M citrate buffer, pH 5.0. The asymmetric unit contained two molecules of IPK, each of which contained IP and ADP in the active site (IPK-IP-ADP). The apparent presence of an unphosphorylated IP molecule (substrate) and ADP (product) did not illuminate the phosphoryl transfer reaction and will not be included in the discussion. This combination of ligands presumably resulted from hydrolysis of ATP that occurred during sample preparation and crystallization. This refined

**TABLE 1. Data collection and refinement statistics<sup>a</sup>**

	IPK-IP-ATP	IPK-IPP-ADP	IPK-GOL-ADP
Data collection			
Space group	C2	C2	C2
Cell dimensions	<i>a</i> = 124.31 <i>b</i> = 44.28, <i>c</i> = 91.76, $\beta$ = 109.45	<i>a</i> = 187.10 <i>b</i> = 42.79 <i>c</i> = 134.90 $\beta$ = 113.19	<i>a</i> = 119.67 <i>b</i> = 96.36 <i>c</i> = 72.43 $\beta$ = 120.59
Molecules/asu	2	4	2
Wavelength (Å)	1.54	0.97887	1.54
Resolution range (Å) (outer shell)	29.30–2.00 (2.07–2.00)	50.00–1.99 (2.06–1.99)	30–2.15 (2.18–2.15)
Unique reflections	31,495	65,827	38,177
$R_{\text{sym}}^b$ (%)	7.5 (47.5)	6.20 (23.4)	9.1 (45.4)
Completeness (%)	98.0 (95.2)	94.6 (80.9)	92.3 (66.2)
Refinement			
Resolution range	29.30–2.00	38.29–1.99	30–2.15
$R_{\text{factor}}^c$ (work/test)(%)	17.6/23.0	19.2/23.4	20.7/26.0
rmsd bonds (Å)	0.018	0.019	0.022
rmsd angles (deg)	1.7	1.9	2.0
Ramachandran plot <sup>d</sup>			
Most favored (%)	98.70	97.81	94.02
Allowed (%)	1.30	2.19	5.92
Disallowed (%)	0.00	0.00	0.06

<sup>a</sup>Statistics for highest resolution shell are in parentheses. <sup>b</sup> $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$ , where *I* is the observed intensity, and  $\langle I \rangle$  is the average intensity of multiple observations of symmetry-related reflections. <sup>c</sup> $R_{\text{factor}} = \sum |F_{\text{obs}}| - |F_{\text{calc}}| / \sum |F_{\text{obs}}|$ , where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factors, respectively. <sup>d</sup>Determined by Molprobity (31).

structure was used as a search model in subsequent molecular replacement solutions of other IPK structures. THA IPK crystals grown in 0.1 M malonate/imidazole/borate (MIB) buffer, pH 7.0, also contained two molecules of the enzyme in the asymmetric unit, but each contained the substrates IP and ATP in the active site (IPK-IP-ATP). In addition, THA IPK crystals grown in propionate/cacodylate/bis-Tris (PCB) buffer, pH 6.0 contained four molecules in the asymmetric unit where two molecules contained IP and ATP (IPK-IP-ATP) while the other two contained the products, IPP and ADP (IPK-IPP-ATP), formed during crystallization. The presence of substrate and product complexes in the asymmetric unit is reasonable given an equilibrium constant for the reaction,  $K_{\text{eq}} = 6.3$ , for uncomplexed substrates and products (13). Finally, a structure for MTH IPK in complex with glycerol and ADP (IPK-GOL-ADP) was determined. All of these structures

were refined to  $R_{\text{free}}$  values of 23–26% against data to 2.0–2.15 Å resolution. All four structures have excellent geometry, with more than 95% of the residues in favored Ramachandran regions. Crystallographic statistics are given in Table 1.

All of the crystallized proteins of THA IPK included N-terminal Phe(−1) and Thr(0) residues that were the remnants of a TEV protease cleavage site. In some cases, one or both of these residues were visible in the electron density. Moreover, electron density for residues Asp(−3) and Pro(−2), also parts of the TEV protease cleavage site, were found in the asymmetric unit of the crystal grown in PCB buffer. This asymmetric unit was interesting because of the combination of substrates and products in the molecules mentioned earlier, as well as the significant movement of loops in two molecules. Briefly, molecules A and B contained bound products, while molecules C and D contained bound substrates

Molecules A and C and molecules B and D formed dimers in the asymmetric unit, where each dimer contained the enzyme-substrate complex in one subunit and the enzyme-product complex in the other structure. In molecules C and D, the  $\eta$ 1- $\alpha$ A loop was in a conformation different than those in molecules A and B, which allowed these molecules to make crystal lattice contacts through Tyr17 (Supplementary Figure 1, panel a). This loop movement resulted in a significant shift in the side chain position of Lys14, an invariant residue in IPKs with direct involvement in catalysis that will be discussed later. Moreover, in molecule D, alternate conformations were found for another catalytic residue, His50 (Supplementary Figure 1, panel b).

An active site water molecule found in enzymes structurally related to IPK was found only in molecules A, C, and D. The unique characteristics of these asymmetric unit molecules allowed us to choose molecule A as the best structure that represents the THA IPK product complex (IPK-IPP-ADP).

The two asymmetric unit molecules of IPK-IP-ATP were almost identical, and both contained IP and ATP in the active site. The same is true for the MTH structure, except that the asymmetric unit molecules contained ADP and either glycerol or water molecules in the IP binding site. This glycerol molecule presumably displaced a bound IP molecule when the crystal was being prepared for cryogenic data collection.

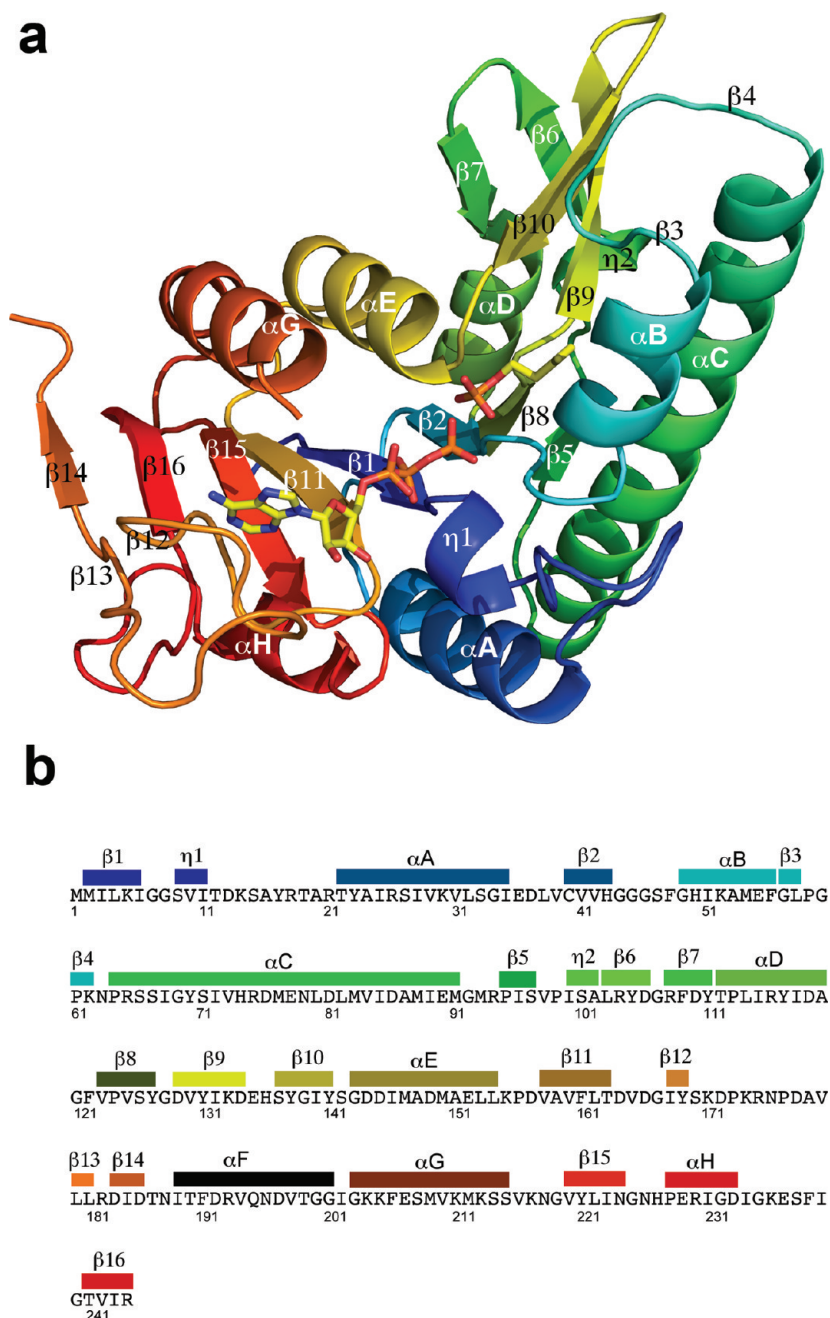
**Fold and Quaternary Structure of IPK.** IPK belongs to the amino acid kinase (AAK) family of enzymes that also includes *N*-acetyl-L-glutamate kinase (NAGK) (14), the aspartokinase (AK) N-terminal domain (15), glutamate-5-kinase (G5K) (16), carbamate kinase (CK) (17), uridine monophosphate kinase (UMPK) (18), and fosfomycin kinase (FomA) (19). This family is characterized by the open  $\alpha\beta\alpha$  sandwich fold first observed in *E. faecalis* CK (17) and catalyzes the transfer of a phosphoryl group from a nucleotide to a small molecule acceptor. Depending on the functional group that is phosphorylated, this family can be divided into two subdivisions. Enzymes in the “carboxylate” subdivision catalyze the transfer of a phosphoryl group to a carboxylate or carbamate group (NAGK, CK, G5K and AK), while those in the “phosphate” subdivision transfer a phosphoryl group to a phosphate or phosphonate (UMPK, FomA, IPK). THA IPK is structurally most similar to FomA, where the backbone rmsd is 2.6 Å over 224 residues. Overlap of FomA and NAGK with THA IPK also aligns their respective sub-

strates and most of their active site residues, suggesting similar catalytic mechanisms. As expected, MTH IPK shares high structural homology with THA IPK with rmsd of 1.9 Å over 223 C $\alpha$  atoms, although sequence identity is only 32%. These structural comparisons confirm the correct assignment of THA and MTH IPKs to the AAK family of enzymes.

The THA IPK structure contains an N-terminal domain (residues 1–155) that binds IP and a C-terminal domain (residues 156–245) that binds ATP. Each molecule of THA IPK consists of 16  $\beta$ -strands, eight  $\alpha$ -helices, and two  $3_{10}$  helices interconnected by loops and turns (Figure 2, panel a). The open  $\alpha\beta\alpha$  sandwich fold formed by these secondary structures is highly similar to that of NAGK (14): the same structural elements form the central  $\beta$ -sheet core ( $\beta$ 5,  $\beta$ 8,  $\beta$ 2,  $\beta$ 1,  $\beta$ 11,  $\beta$ 15,  $\beta$ 16,  $\beta$ 14) and the two flanking layers of  $\alpha$ -helices ( $\alpha$ C,  $\alpha$ A,  $\alpha$ H on one side and  $\alpha$ D,  $\alpha$ E,  $\alpha$ G,  $\alpha$ F on the other). Three  $\beta$ -hairpin turns reside in the N-terminal domain, two of which form the surrounding walls of the aliphatic tail of IP ( $\beta$ 3- $\beta$ 4,  $\beta$ 9- $\beta$ 10) and another forming one side of the IP binding pocket ( $\beta$ 6- $\beta$ 7). The  $\beta$ 3- $\beta$ 4 hairpin turn of THA IPK is shorter than that of NAGK but is in the same closed conformation that allows it to cap the IP binding site and shield its hydrophobic surroundings from solvent molecules (14). The remaining  $\beta$ 12- $\beta$ 13 hairpin turn resides in the C-terminal domain and hovers above the purine ring of ATP. This  $\beta$ -hairpin is found within the segment that connects the core  $\beta$ -sheet strands  $\beta$ 11 and  $\beta$ 14 and is longer than the corresponding segment in NAGK. The longest helix in THA IPK,  $\alpha$ C, aligns well with  $\alpha$ C in NAGK but is one turn longer.

The segment between  $\beta$ 14 and  $\alpha$ G (residues 189–201) in the THA IPK structure is disordered in all molecules in the asymmetric unit, as evidenced by an uninterpretable electron density. This segment contains the helix,  $\alpha$ F, and the  $\alpha$ F- $\alpha$ G junction loop that form one side of the ATP binding site. The corresponding segment in MTH IPK is missing (204–218), as is also seen in one molecule of G5K (2J5T) (16), suggesting that the flexibility of these secondary structures might be important for the binding and release of ATP.

The THA and MTH IPKs are dimeric in the crystal, consistent with the observed oligomerization state during gel filtration chromatography (Figure 3, panels a and b). The dyadic axis is perpendicular to the core  $\beta$ -sheet that runs across the dimer (16  $\beta$ -sheets, 8 per subunit). As in all enzymes of this family that catalyze phos-



**Figure 2.** Tertiary structure of THA IPK. **a)** Ribbon diagram of the THA IPK monomer with the helices and strands labeled. The region that connects  $\beta 14$  and  $\alpha G$  is disordered in the crystal and not modeled in the structure. **b)** THA IPK primary sequence and secondary structures in the same color scheme as in panel **a**. The residues are numbered below the primary sequence. The segment between  $\beta 14$  and  $\alpha F$  (which contains  $\alpha F$ ) is not modeled.

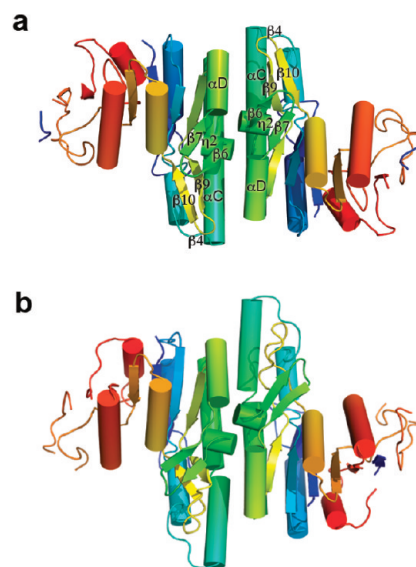
phoryl transfer, each active center is confined to a single subunit. However, dimer formation may be structurally

quaternary architectures are generated by changing the angle between the long interfacial helices  $\alpha C$  of the

significant since N-terminal domain loops surrounding the active site are anchored to the other subunit of the dimer (14).

Analysis of the dimerization surface of THA IPK using the EBI PISA (20) server resulted in a complexation significance score of 0.797, corresponding to a tight dimer that buries a total of 3275 Å<sup>2</sup> of solvent accessible area between the two monomers. The compactness arises from a total of nine hydrogen bonding interactions between residues in the long interfacial  $\alpha C$  helices of both subunits and between the  $\alpha D$  helix of one subunit and the  $\beta 4$ ,  $\beta 6$ , and  $\beta 9$  strands and the  $\eta 2$  helix of the other. Eight electrostatic interactions between the  $\alpha C$  helices of the two subunits and between the  $\beta 6$  strand of one subunit and the  $\beta 7$  strand of the other further stabilize the dimer. Finally, hydrophobic interactions mediated by residues in the  $\alpha C$ ,  $\alpha D$ , and  $\eta 2$  helices and the  $\beta 4$ ,  $\beta 5$ ,  $\beta 6$ ,  $\beta 9$ , and  $\beta 10$  strands further bind the two molecules.

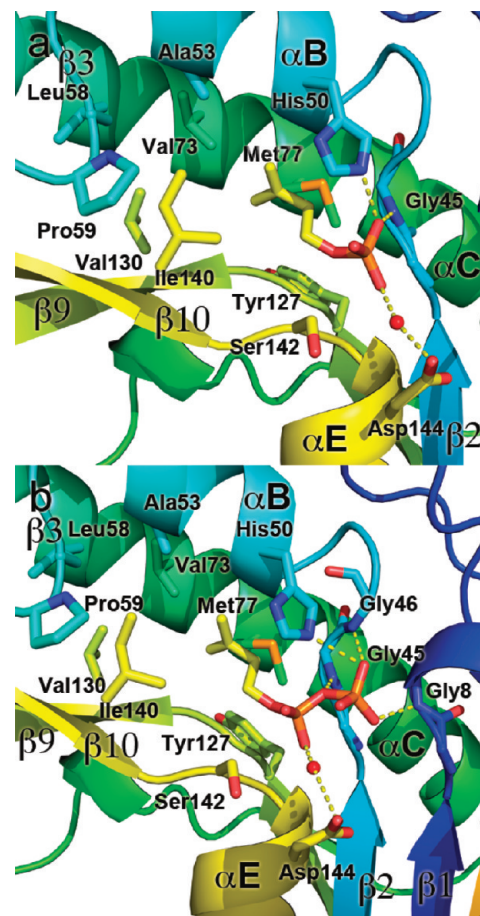
Although essentially the same secondary structures are used by the AAK enzymes in dimerization, different



**Figure 3.** Homodimers of THA and MTH IPKs. **a)** THA IPK homodimer, with secondary structures involved in the dimer interface labeled. **b)** MTH IPK homodimer. The dyadic axis is perpendicular to the core  $\beta$ -sheet that runs across the dimer and projects toward the viewer.

dimer subunits (16). In the case of THA IPK, the crossover point of these helices is found on the fifth turn from the N-terminus, and the rotation angle around an axis that penetrates the intersubunit interface perpendicularly at this point is  $\sim 220^\circ$ . In MTH IPK, the crossover point shifts to the sixth turn from the N-terminus while the rotation angle remains at  $\sim 220^\circ$ . In NAGK and CK the crossover points in  $\alpha C$  occur in the third turn and shifts to the fifth turn in G5K, while the rotation angles are  $110^\circ$  in both NAGK and CK and  $260^\circ$  in G5K (16).

**The Active Site.** The (IPK-IP-ATP) structures presented here are the first to show the molecular details of binding an IP molecule by an enzyme (Figure 4, panel a). Many characteristics of AAK enzymes have been described previously in the discussion of the NAGK structure, and several of these traits are conserved in both THA and MTH IPKs. Briefly, the IP (IPP) binding pocket is located in the N-terminal domain formed by the  $\beta 2$ - $\alpha B$  loop, the  $\alpha B$  helix, the  $\beta 3$ - $\beta 4$  hairpin, the N-terminus of the  $\alpha C$  helix, the  $\beta 8$ - $\beta 9$  loop, the  $\beta 9$ - $\beta 10$  hairpin turn, and the  $\beta 10$ - $\alpha E$  loop. The  $\alpha B$  helix contributes to binding IP (IPP) through the orientation of the positive end of the helix dipole and the hydrogen bonding and electrostatic interactions made by the charged His50 residue. In the FomA structure, the  $\alpha B$  helix is ordered in the pres-



**Figure 4.** IP and IPP binding sites. **a)** IP binding site showing the amino acid residues that form the hydrophobic interior of the pocket and limit the chain length specificity of the enzyme. Hydrogen bonds between active site residues and the phosphate moiety of IP are shown as broken yellow lines. A water molecule in the active site is modeled as a red sphere, acting as a hydrogen bonding bridge between IP and Asp144. **b)** IPP binding site.

ence of fosfomycin and disordered in its absence (19). Comparison of AAK enzyme structures revealed that only the enzymes in the “phosphate” subdivision (UMPK, FomA and IPK) align their respective  $\alpha B$  helices. Moreover, an aligning residue at the His50 position is found exclusively in these enzymes.

The His50 residue in IPK-IP-ATP poises the bound IP for nucleophilic attack on  $P_\gamma$  of ATP by hydrogen bonding with a nonbridging oxygen atom in the terminal phosphate of IP. In the IPK-IPP-ADP structure, this residue moves slightly to make a hydrogen bond with a non-

bridging oxygen atom in the IPP terminal phosphate (Figure 4, panel b), facilitating the reverse reaction. In fact, in comparing IPK·IP·ATP and IPK·IPP·ADP (rmsd 0.4 Å), the slight change in the position of His50 to hydrogen bond with the terminal phosphate of the product is one of the very few coordinate shifts that were observed. In FomA, His58 positions fosfomycin for nucleophilic attack on the phosphate donor through hydrogen bonding with an intervening water molecule (19). The corresponding residue in *E. coli* UMPK, Arg62, hydrogen bonds with the terminal phosphates of both substrate and product and stabilizes P<sub>γ</sub> of ATP through electrostatic interactions (21). These observations together with the invariance of this histidine residue in putative IPK homologues (13) (Supplementary Table S2), the apparent pK<sub>a</sub> of 6.5 for THA IPK (6.8 for MTH IPK) from pH dependence studies (13) suggest that this His50 may perform direct catalytic roles similar to those proposed for Arg62 in UMPK, a signature trait of IPKs.

The specificity of IPK arises from hydrophobic residues that cradle the aliphatic portion of IP and a constellation of conserved residues that hydrogen bond with its phosphate moiety (Figure 4, panel a). The residues Met77 and Val73 (αC N-terminus), Val130 and Ile140 (β9-β10 hairpin), and Ala53 (αB C-terminus) and Tyr127 (β8-β9 loop) all form the hydrophobic pocket for the apolar tail of IP. In addition, Leu58 and Pro59 of the β3-β4 hairpin enhance the hydrophobic surroundings by capping the IP binding pocket. Notably, the β3-β4 hairpin is found only in the structures of NAGK and IPK. This motif is in the closed conformation in substrate-bound NAGK structures (14) as in the IPK·IP·ATP and IPK·IPP·ADP structures. The side chains of some, if not all, of these hydrophobic residues may give rise to the chain length specificity of IPK. Geranyl phosphate, which is one isoprene unit longer than IP, is a poor substrate for IPK, indicating that the IP binding site cannot easily accommodate a C<sub>10</sub> isoprenoid chain (13).

The phosphate moiety of IP is recognized in the active site by structural motifs that are well-conserved in the AAK family. In IPK·IP·ATP, the three nonbridging oxygen atoms of IP make hydrogen bonding interactions with the N<sub>ε</sub> of His50, the N atom of Gly45 in the β2-αB loop, and Asp144 through a bound water molecule (Figure 4, panel a). This water molecule is found in the holoenzyme structures of *E. coli* NAGK (pdb ID 1GS5), *P. furiosus* UMPK (pdb ID 2BMU), *E. coli* UMPK (pdb ID 2BNE), and FomA (pdb ID 3D41). A water molecule in a

similar position is present in IPK·IPP·ADP (Figure 4, panel b). The β2-αB loop, which contains Gly45, is one of the conserved glycine-rich loops in the AAK family and is thought to stabilize the transition state intermediate and the phosphorylated product of these enzymes (19, 22). In IPK·IPP·ADP, Gly45 of this loop makes a hydrogen bond with the P<sub>α</sub>-P<sub>β</sub> bridging oxygen atom of IPP. The other glycine-rich loop, β1-η1, contains Gly8 that stabilizes IPP by making a hydrogen bond with a nonbridging oxygen atom of its terminal phosphate.

The β10-αE loop completes the structural elements involved in binding the polar end of IP. This loop contains Ser142, which is conserved in the “phosphate” subdivision (Thr120 in *P. furiosus* UMPK, Ser149 in FomA) (18, 19) and makes a hydrogen bond with the terminal phosphate of the substrates. In IPK·IP·ATP and IPK·IPP·ADP, the orientation of the side chain hydroxyl group of Ser142 does not allow a hydrogen bond with substrates or products (Figure 4, panel a and b). However, an alternate rotamer could provide hydrogen bond stabilization as the substrates moved to the transition state. The positive end of the helix dipole of αE, initiated by the β10-αE loop, may further stabilize the polar head of the IP molecule as well as the negatively charged transition state intermediate, as in other structures of AAK enzymes.

Thus, the specificity of IPK toward IP arises from a combination of hydrophobic interactions with the hydrocarbon tail and hydrogen bonding and electrostatic interactions with the phosphate headgroup to match the amphiphatic character of IP. This differentiates the IPK active site from other enzymes in the family whose substrates do not have this intrinsic amphiphaticity, such as that of NAGK whose active site is laden with both polar and charged residues that interact with the acetyl and carboxyl groups of NAG.

The location of the ATP binding site in both THA and MTH IPKs superimpose well with those of NAGK and other AAK enzymes. Conserved amino acid residues in these enzymes that have previously been identified to form hydrogen bonding, electrostatic, and hydrophobic interactions with the purine ring, sugar, and polyphosphate moieties of ATP are also found in IPK. For instance, the NAGK residues Met214 (purine stacking), Asp181 (ribose 3'-OH binding), Asp162 (polyphosphate positioning), and Lys8 (phosphate hydrogen bonding) (14) are also conserved in THA IPK as Ile202, Asp164, Asp144, and Lys5, respectively. In MTH IPK, the corre-



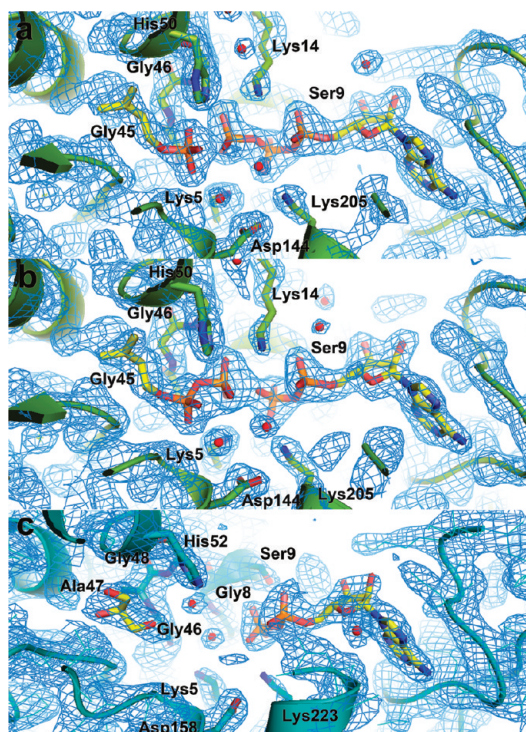


tive in the presence of  $Mn^{2+}$ ,  $Zn^{2+}$ ,  $Ni^{2+}$ ,  $Cd^{2+}$ , and  $Co^{2+}$  (Supplementary Figure 3, panel b).

In IPK-IP-ATP, Lys14 is located within hydrogen bonding distance of the nonbridging oxygen atoms of  $P_{\alpha}$  (3.0 Å) and  $P_{\gamma}$  (2.8 Å) and the  $P_{\beta}$ - $P_{\gamma}$  bridging oxygen atom (3.1 Å) in ATP. In IPK-IPP-ADP, Lys14 makes a hydrogen bond with a nonbridging oxygen atom of  $P_{\beta}$  (2.7 Å) in IPP and the nonbridging oxygen atoms of  $P_{\alpha}$  (2.7 Å) and  $P_{\beta}$  (3.0 Å) of ADP. Interestingly, a corresponding residue, Lys18, is found in the (FomA-fofosfomycin-AMPPNP) complex (19). However, in this structure, Lys18 only forms a hydrogen bond with a  $P_{\gamma}$  oxygen atom in AMPPNP, which in turn is not properly oriented for nucleophilic attack by fosfomycin. The relative positions of IP and ATP in THA IPK are consistent with phosphoryl transfer. Thus, in THA IPK, Lys5, Lys14, and Lys205 form a “lysine triangle” surrounding the phosphate residues in the substrates of IPKs.

Along with His50, Lys14 is a conserved feature of the IPKs. These residues may have evolved to preferentially stabilize transition states that contain more negative charge than those formed during phosphoryl transfers involving carboxylated and carbamated substrates, thus giving rise to IP kinases. The conservation of these traits in FomA along with other key catalytic residues seen in IPK is also consistent with the acquisition of antibiotic resistance toward fosfomycin in *Streptomyces* by a lateral gene transfer of an ancestral IPK gene from *Archaea*.

**Conclusion.** The crystal structures of IPK from two species of *Archaea* are highly similar despite their low amino acid sequence identities, and key catalytic residues are conserved in both proteins. The catalytic machinery of *Archaeal* IPKs is related to that found in other AAK enzymes except for the incorporation of two additional residues, His50 and Lys14, in the active site. These residues appear to orient the phosphate residues in the bound substrates and are positioned to stabilize the transition state for phosphoryl transfer. Moreover, a catalytic lysine triangle formed by Lys5, Lys14, and Lys205 constitutes a new catalytic motif in IPKs. The



**Figure 6.**  $2F_o - F_c$  electron density maps of the active sites of THA and MTH IPKs contoured at 1 rmsd. a) THA IPK active site with bound substrates, IP and ATP. b) THA IPK active site with bound products, IPP and ADP. The Gly8 residue is obscured by  $P_{\gamma}$  of ATP. c) MTH IPK active site with bound glycerol and ADP, showing the putative catalytic residues, Gly8 and Ser9. Putative catalytic residues are labeled, and water molecules are shown as red spheres. Ligand omit maps for IP and ATP, and IPP and ADP, can be found in Supplementary Figure 2, panels a and b.

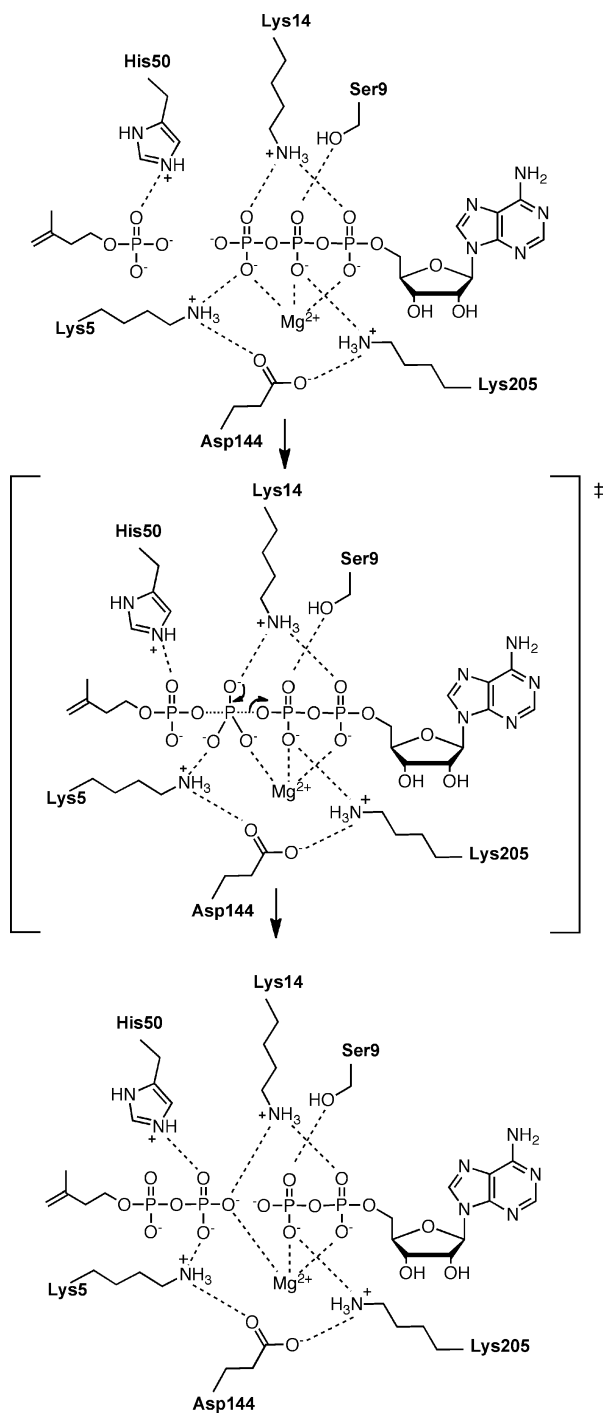
presence of these key IPK residues in bacterial FomA suggests that fosfomycin resistance in *Streptomyces* was acquired by a horizontal transfer of an *Archaeal* gene for IPK genes. Finally, homologues of IPK are found in the genomes of representatives from all of the families of *Archaea*. Although the existence of a PMD is yet to be established, it is likely that isoprenoid compounds are synthesized by the alternate MVA pathway in these organisms.

## METHODS

**Crystallization.** All crystals of THA IPK were grown by the sitting drop vapor diffusion method at 21 °C using a protein mixture comprising 0.30 mM protein, 1.5 mM  $MgCl_2$ , 1.5 mM IP, and 1.5 mM ATP that was prepared and incubated for at least 2 h before crystallization. In all cases the crystals grew to maximum di-

mensions in about 1 week in drops comprising a 1:1 mixture of protein and reservoir solutions.

Selenomethionine-substituted THA IPK crystals complexed with IP and ADP were grown using a solution of 0.1 M sodium citrate, pH 5.0, and 15% PEG 6000. Crystals grew to maximum dimensions in about 1 week. The IPK-IP-ATP crystal was grown



**Figure 7.** Proposed catalytic mechanism for IPK. Catalytic residues and selected hydrogen bonds are shown. The “missing”  $Mg^{2+}$  ion is positioned on the basis of overlap with the structure of Foma.

using native protein and a reservoir solution of 0.1 M MIB buffer (2:3:3 molar ratio of sodium malonate, imidazole, and boric acid), pH 7.0, and 25% PEG 1500. The IPK-IPP-ADP crystal was obtained using selenomethionine-substituted protein and a reservoir solution of 0.1 M PCB buffer (2:1:2 molar ratio of sodium propionate, sodium cacodylate, and bis-Tris propane), pH 6.0, containing 25% PEG 1500. In this case the substrate included with the crystallized protein was converted and preferentially bound as product under the crystallization conditions.

Crystals of MTH IPK grew by mixing 2  $\mu\text{L}$  of protein (5  $\text{mg mL}^{-1}$  MTH IPK, 5  $\text{mM MgCl}_2$ , 1  $\text{mM ATP}$ ) with a precipitant solution containing 12% PEG 6000 and 2 M NaCl. The active site was found to contain glycerol and ADP.

**Crystal Structure Determination.** Prior to data collection, all crystals were cryo-protected by immersion for 30 s in a modified reservoir solution that contained 30% glycerol and cooled by plunging into liquid nitrogen. Diffraction data were collected at 100 K on a rotating anode source or at beamline 11-1 of the SSRL synchrotron and processed using HKL2000 (24). The THA IPK structure in complex with IP and ADP was determined by single-wavelength anomalous diffraction (SAD) using PHENIX (25). A total of 45 selenium sites were located and used to calculate initial phases to 1.90  $\text{\AA}$ , and the phase estimates were improved by density modification to generate a readily interpretable electron density map. Model building was done using COOT (26). Refinement used REFMAC5 (27) with the final refinement cycles performed using PHENIX (25). This refined model was used to determine all of the other structures by molecular replacement using PHASER (28) in the CCP4 suite (29).

The DaliLite server (30) was used to superimpose the  $C\alpha$  atoms of two protein structures and measure the rmsd. Molprobity (31) was used for structure validation. PyMol (32) and the command “Ray” was used to create images of the protein structures.

**Accession Codes:** The coordinates of the THA IPK-IP-ATP, THA IPK-IPP-ADP and MTH IPK-GOL-ADP complex structures have been deposited in the Protein Data Bank and are listed as 3LKK, 3LL5, and 3LL9, respectively.

**Acknowledgment:** We thank J. Muller of the Department of Chemistry of the University of Utah for assistance in mass spectrometry. This research was supported by NIH grants GM 25521 (C.D.P.) and GM 56775 (H.L.S., C.P.H.). Part of the work was carried out at the Stanford Synchrotron Radiation Light Source (SSRL), a national research facility operated by Stanford University on behalf of the Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and by the U.S. National Institutes of Health (NIH), National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences.

**Supporting Information Available:** This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

1. Reiling, K. K., Yoshikuni, Y., Martin, V. J. J., Newman, J., Bohlmann, J., and Keasling, J. D. (2004) Mono and diterpene production in *Escherichia coli*, *Biotechnol. Bioeng.* 87, 200–212.
2. Davis, E. M., and Croteau, R. (2000) Cyclization enzyme in the biosynthesis of monoterpenes, sesquiterpenes, and di-terpenes, *Top. Curr. Chem.* 209, 53–95.
3. Liang, P., Ko, T., and Wang, A. H. (2002) Structure, mechanism and function of prenyltransferases, *Eur. J. Biochem.* 269, 3339–3354.
4. Sacchetini, J. C., and Poulter, C. D. (1997) Creating isoprenoid diversity, *Science* 277, 1788–1789.

- Ogura, K., and Koyama, T. (1998) Enzymatic aspects of isoprenoid chain elongation, *Chem. Rev.* **98**, 1263–1276.
- Poulter, C. D. (2006) Farnesyl diphosphate synthase. A paradigm for understanding structure and function relationships in E-polyprenyl diphosphate synthases, *Phytochem. Rev.* **5**, 17–26.
- Song, L., and Poulter, C. D. (1994) Yeast farnesyl diphosphate synthase: site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3044–3048.
- Rohdich, F., Bacher, A., and Eisenreich, W. (2004) Perspectives in anti-infective drug design. The late steps in the biosynthesis of the universal terpenoid precursors, isopentenyl diphosphate and dimethylallyldiphosphate, *Bioorg. Chem.* **32**, 292–308.
- Kuzuyama, T. (2002) Mevalonate and nonmevalonate pathways for the biosynthesis of isoprene units, *Biosci. Biotechnol. Biochem.* **66**, 1619–1627.
- Boucher, Y., Kamekura, M., and Doolittle, W. F. (2004) Origins and evolution of isoprenoid biosynthesis in *Archaea*, *Mol. Microbiol.* **52**, 515–527.
- Smit, A., and Mushieng, A. (2000) Biosynthesis of isoprenoids via mevalonate in *Archaea*: the lost pathway, *Genome Res.* **10**, 1468–1484.
- Grochowski, L. L., Xu, H. M., and White, R. H. (2006) *Methanocaldococcus jannaschii* uses a modified mevalonate pathway for biosynthesis of isopentenyl diphosphate, *J. Bacteriol.* **188**, 3192–3198.
- Chen, M., and Poulter, C. D. (2010) Characterization of thermophilic and archaeal isopentenyl phosphate kinases, *Biochemistry* **49**, 207–217.
- 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 family, during catalysis, *Structure* **10**, 329–342.
- 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.
- 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.
- 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 precursors of pyrimidines and arginines, *Protein Sci.* **8**, 934–940.
- 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.
- Pakhomova, S., Bartlett, S., Augustus, A., Kuzuyama, T., and Newcomer, M. (2008) Crystal structure of fosfomycin resistance kinase FomA from *Streptomyces wedmorensis*, *J. Biol. Chem.* **283**, 28518–28526.
- Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J. Mol. Biol.* **372**, 774–797.
- 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.
- Gil-Ortiz, F., Ramon-Maiques, S., Fita, I., and Rubio, V. (2003) The course of phosphorous in the reaction of *N*-acetyl-L-glutamate kinase, determined from the structures of crystalline complexes, including a complex with AlF<sub>4</sub>(<sup>-</sup>) transition state mimic, *J. Mol. Biol.* **331**, 231–244.
- Tian, G., Kane, L. S., Holmes, W. D., and Davis, S.T. (2002) Modulation of cyclin-dependent kinase 4 by binding of magnesium(II) and manganese(II), *Biophys. Chem.* **95**, 79–90.
- Otwinowski, Z., Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods in Enzymology, Macromolecular Crystallography* (Carter, C. W., Jr. and Sweet, R. M., Eds.), Vol. 276, part A, pp 307–326, Academic Press, New York.
- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 213–221.
- Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr.* **60**, 2126–2132.
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
- McCoy, A. J., Grosse-Kunstleve, Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J. Appl. Crystallogr.* **40**, 658–674.
- (1994) Collaborative Computational Project, Number 4. The CCP4 Suite: programs for protein crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.
- Holm, L., and Park, J. (2000) DALI: workbench for protein structure comparison, *Bioinformatics* **16**, (6), 566–567.
- Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J., Kapral, G. J., Wang, X., Murray, L., Arendall, B., III, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) Molprobity: all-atom contacts and structure validation for proteins and nucleic acids, *Nucleic Acids Res.* **35**, W375–W383, Web Server issue.
- DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA, USA.