**Associative Mechanism for Phosphoryl Transfer: A Molecular Dynamics Simulation of *Escherichia coli* Adenylate Kinase Complexed With Its Substrates**

Harini Krishnamurthy,1,2,4 Hongfeng Lou,3,4 Adam Kimple,3,4 Claire Vieille,1 and Robert I. Cukier3,4*

1Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan
2Graduate Program in Cell and Molecular Biology, Michigan State University, East Lansing, Michigan
3Department of Chemistry, Michigan State University, East Lansing, Michigan
4MSU Center for Biological Modeling, Michigan State University, East Lansing, Michigan

**ABSTRACT**  The ternary complex of *Escherichia coli* adenylate kinase (ECAK) with its substrates adenosine monophosphate (AMP) and Mg-ATP, which catalyzes the reversible transfer of a phosphoryl group between adenosine triphosphate (ATP) and AMP, was studied using molecular dynamics. The starting structure for the simulation was assembled from the crystal structures of ECAK complexed with the bisubstrate analog diadenosine pentaphosphate (AP5A) and of *Bacillus stearothermophilus* adenylate kinase complexed with AP5A, Mg2+, and 4 coordinated water molecules, and by deleting 1 phosphate group from AP5A. The interactions of ECAK residues with the various moieties of ATP and AMP were compared to those inferred from NMR, X-ray crystallography, site-directed mutagenesis, and enzyme kinetic studies. The simulation supports the hypothesis that hydrogen bonds between AMP’s adenine and the protein are at the origin of the high nucleoside monophosphate (NMP) specificity of AK. The ATP adenine and ribose moieties are only loosely bound to the protein, while the ATP phosphates are strongly bound to surrounding residues. The coordination sphere of Mg2+, consisting of 4 waters and oxygens of the ATP β- and γ-phosphates, stays approximately octahedral during the simulation. The important role of the conserved Lys13 in the P loop in stabilizing the active site by bridging the ATP and AMP phosphates is evident. The influence of Mg2+, of its coordination waters, and of surrounding charged residues in maintaining the geometry and distances of the AMP α-phosphate and ATP β- and γ-phosphates is sufficient to support an associative reaction mechanism for phosphoryl transfer.

**INTRODUCTION**

Adenylate kinases (AKs) are small, monomeric phosphor yl transferases that catalyze the reversible transfer of a phosphoryl group from adenosine triphosphate (ATP) to adenosine monophosphate (AMP) by nucleophilic attack on the γ-phosphate of ATP.1 The net reaction is schematized as

\[
\text{Mg}^{2+}\text{-ATP} + \text{AMP} \rightleftharpoons \text{Mg}^{2+}\text{-ADP} + \text{ADP} \quad (\text{Scheme 1})
\]

AKs are composed of 3 domains: the core, LID, and AMP-binding (AMP-bd) domains (Fig. 1). In the absence of substrates, AKs are in an open form. In the presence of ATP, AMP, and Mg2+, the LID and AMP-bd undergo major conformational rearrangements, resulting in the enzyme closing to form the ternary complex and expelling waters to prevent ATP and AMP hydrolysis.2 The best-known AK families are the short AKs represented by the mammalian cytosolic enzymes, and the long AKs represented by bacterial and mitochondrial enzymes. The long AKs contain a 20- to 30-residue insertion inside the LID domain.3,4 Among the short AKs, only the porcine cytosolic AK has been crystallized in the absence of substrate (Protein Data Bank (PDB) code: 3ADK).5 On the other hand, the long AKs from *Escherichia coli* (ECAK), *Bacillus stearothermophilus*, *Saccharomyces cerevisiae*, and bovine mitochondria have been crystallized with various substrate–inhibitor combinations or with the 5-phosphate bisubstrate analog, diadenosine pentaphosphate (i.e., AP5A).6–15 The crystal structures of the *E. coli* and bovine mitochondrial AKs in the absence of any ligands are also known.10,16 ATP binds to the catalytic site as the Mg2+-ATP complex, in which the catalytically essential Mg2+ is coordinated to ATP’s β- and γ-phosphates and to 4 water molecules. Thus, the functional catalytic assembly is a ternary complex. In this work, we focus on ECAK, since a wealth of kinetic,17 NMR,18–22 and mutagenesis23–25 data are available for this enzyme.

Crystallography can provide structures of enzymes complexed with products, with inhibitors that resemble sub-

---

1 Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan
2 Graduate Program in Cell and Molecular Biology, Michigan State University, East Lansing, Michigan
3 Department of Chemistry, Michigan State University, East Lansing, Michigan
4 MSU Center for Biological Modeling, Michigan State University, East Lansing, Michigan

**Key words:** adenylate kinase; ATP and AMP; molecular dynamics; phosphoryl transfer

© 2004 Wiley-Liss, Inc.

**INTRODUCTION**

Adenylate kinases (AKs) are small, monomeric phosphoryl transferases that catalyze the reversible transfer of a phosphoryl group from adenosine triphosphate (ATP) to adenosine monophosphate (AMP) by nucleophilic attack on the γ-phosphate of ATP. The net reaction is schematized as

\[
\text{Mg}^{2+}\text{-ATP} + \text{AMP} \rightleftharpoons \text{Mg}^{2+}\text{-ADP} + \text{ADP} \quad (\text{Scheme 1})
\]

AKs are composed of 3 domains: the core, LID, and AMP-binding (AMP-bd) domains (Fig. 1). In the absence of substrates, AKs are in an open form. In the presence of ATP, AMP, and Mg2+, the LID and AMP-bd undergo major conformational rearrangements, resulting in the enzyme closing to form the ternary complex and expelling waters to prevent ATP and AMP hydrolysis. The best-known AK families are the short AKs represented by the mammalian cytosolic enzymes, and the long AKs represented by bacterial and mitochondrial enzymes. The long AKs contain a 20- to 30-residue insertion inside the LID domain. Among the short AKs, only the porcine cytosolic AK has been crystallized in the absence of substrate (Protein Data Bank (PDB) code: 3ADK). On the other hand, the long AKs from *Escherichia coli* (ECAK), *Bacillus stearothermophilus*, *Saccharomyces cerevisiae*, and bovine mitochondria have been crystallized with various substrate–inhibitor combinations or with the 5-phosphate bisubstrate analog, diadenosine pentaphosphate (i.e., AP5A). The crystal structures of the *E. coli* and bovine mitochondrial AKs in the absence of any ligands are also known. ATP binds to the catalytic site as the Mg2+-ATP complex, in which the catalytically essential Mg2+ is coordinated to ATP’s β- and γ-phosphates and to 4 water molecules. Thus, the functional catalytic assembly is a ternary complex. In this work, we focus on ECAK, since a wealth of kinetic, NMR, and mutagenesis data are available for this enzyme.

Crystallography can provide structures of enzymes complexed with products, with inhibitors that resemble sub-

---

**Key words:** adenylate kinase; ATP and AMP; molecular dynamics; phosphoryl transfer

© 2004 Wiley-Liss, Inc.
strates, and even with inhibitors that model transition states. In this work, we create a model of the ECAK-AMP ternary reaction complex based on the ECAK/AP5A structure (PDB code: 1AKE) and on the B. stearothermophilus AK (BSAK) structure crystallized with AP5A and Mg²⁺ (PDB code: 1ZIO). Constructing a reaction complex model allows one to study the stabilization of a possible catalytically competent complex that is more realistic than a model inferred from a crystal structure containing substrate mimics. Thus, the ternary complex model will be used as the starting point of a molecular dynamics (MD) simulation that can assess how ATP and AMP are maintained in a catalytically competent configuration with the aid of the surrounding protein, the Mg²⁺ ion, and its 4 coordinating waters. The highly charged complex suggests a focus on strong hydrogen bonds (H-bonds) and salt bridges that can serve the purpose of maintaining a catalytically relevant geometry.

Two simulations of AKE have appeared previously. One, a 300-ps study of ECAK complexed with AP5A, was carried out in the gas and solvated states. A number of residues that would be candidates for strong interaction with a Mg²⁺ ion were suggested in this study. The other simulation used a weighted masses MD method to explore the nature of the open conformation of apo-ECAK that could be reached starting from a number of closed form (with different bound substrates) crystallographic structures.

The similarities and differences between various experimental conclusions based mainly on studies using AP₅A, and the simulation of the ECAK-Mg-ATP-AMP structure are analyzed. While there are some differences in details, the suggested distinction between the high specificity of AKs for the AMP substrate and the lower specificity for ATP is supported by the simulation. Particular attention is devoted to investigating the interactions of the conserved Lys13 (the numbering refers to ECAK) located at the end of the P loop, with the AMP and ATP phosphate groups. The simulation shows that Lys13 does play an important role in substrate stabilization by bridging the AMP and ATP phosphates.

There is great interest, in general, in the mechanism of phosphoryl transfers, whether it is an associative or a dissociative mechanism, these being extremes of a continuum of mechanisms qualitatively defined by the phosphorous–phosphorous distance, which increases in going from an associative to a dissociative mechanism. There is evidence that a dissociative mechanism is favored in solution, while an associative mechanism is favored in enzymes. The nature of the transition state (or reaction intermediate) also is an issue, with the most frequently suggested geometry invoking a pentavalent phosphorous. As we discuss, the simulation carried out herein maintains the AMP and ATP terminal phosphates at a distance that supports an associative phosphoryl transfer mechanism.

**METHODS**

**Starting Structure**

The 1.9 Å resolution crystal structure of the ECAK-AP₅A complex (PDB code: 1AKE) was used as the starting point for constructing an ECAK-AMP-Mg²⁺-ATP ternary complex, which will be denoted as 1AKE*. The 1AKE PDB file contains the coordinates of the 2 enzyme–inhibitor complexes (A and B) present in the crystal asymmetric unit. We used complex B to build 1AKE*. AP₅A (Fig. 1) is considered to be a bisubstrate mimic of the reactant state, since ECAK catalyzes the reaction of ATP and AMP to produce two ADPs (Scheme 1). ATP and AMP coordinates
were obtained by eliminating AP$_5$A’s δ-P0$_4$ group (see Fig. 1).

While the 1AKE structure provides the AP$_5$A coordinates needed to create well-positioned ATP and AMP molecules within the protein, it does not contain the catalytic Mg$^{2+}$ and its 4 waters of coordination. To create the starting structure for the MD simulation, we imported the coordinates of Mg$^{2+}$ and its coordinated waters from the crystal structure of the BSAK · Mg$^{2+}$-AP$_5$A complex (PDB code: 1ZIO). The 1ZIO and 1AKE structures were superimposed by fitting 1ZIO’s AP$_5$A molecule onto that of 1AKE. 1ZIO’s Mg$^{2+}$ and its 4 coordination waters were then merged with the ECAK · AMP · ATP structure to create 1AKE*. This structure contains Mg$^{2+}$ positioned between nonbridging oxygens of ATP’s β- and γ-phosphates and the 4 Mg$^{2+}$ coordinating waters, as displayed in Figure 1.

**Force Field**

The GROMOS96 force-field was used for the MD simulations. This force field is a united atom force field with explicit polar hydrogens, and hydrogens on the phenylalanine, tryptophan, and tyrosine rings. The GROMOS96 force field contains parameters for all protein residues, as well as for ATP, but it does not include parameters for AMP. Parameters for AMP were created based on the ATP parameters.

The ionization states of all residues were set to reflect a neutral pH. Histidines were assumed to be neutral, with the Nδ atom protonated, and the lysine, arginine, aspartate, and glutamate residues were assumed to be charged. To assess whether this charge assignment is reasonable in the binding pocket, where numerous positively charged residues bind the substrate phosphates, we adopted the following procedure. The accessible surface area (ASA) of each residue was evaluated using the Surface Racer program, for the residues in the 1AKE* complex and in the 1AKE* structure in the absence of the substrates. For each residue, the ASA in the presence of substrate was subtracted from the ASA in the absence of substrate. The nonzero ASA difference values then single out those residues that are buried at the protein–substrate interface. The charged residues buried at the protein–substrate interface, Lys13, Arg36, Asp84, Arg88, Arg119, Arg123, Arg156, and Asp158, have a net charge of +4. These charges are almost compensated, through various interactions discussed below, by the net charge of −2 from ATP (−3, with 1 terminal oxygen protonated), AMP (−1), and Mg$^{2+}$ (+2). A look at the charged residues buried at the protein–substrate interface shows that, with the exception of Arg123, all the ionizable groups on positively charged residues are engaged in charge–charge interactions with the substrate phosphates. All these charge–charge interactions are stable throughout the 3-ns simulation period (see below). The side-chain of Asp84 is engaged in stable H-bonds with 2 of the waters coordinated to Mg$^{2+}$, namely, 219 and 221. Arg123 forms a salt bridge with Asp159, which is also stable during the simulation. Note that Asp159, along with its neighboring Asp158 at the C-terminal end of the LID domain, are conserved in all AKs implicating a crucial role for them in AKs. These 2 aspartates form salt bridges with Arg123 and Arg156 side-chains only in the substrate bound form. Müller and Schulz suggested that salt bridges Arg123–Asp159 and Arg156–Asp158 (formed with the aid of ATP and AMP, which help in positioning the side-chains) are instrumental in the conformational change leading to LID closure. Thus, the conventional ionization states assumed in this study are appropriate, since almost all of the charged groups buried in the 1AKE* complex have their charges satisfied either through protein–substrate, or protein–protein interactions.

**Molecular Dynamics**

The MD simulation was carried out using CUKMODY, a code designed for the efficient simulation of proteins and other large solutes. A combination of a cell index method with linked lists and a Verlet neighbor list is used to provide linear scaling with the number of atoms in the pair list routine, essential for the large systems considered here. For the Verlet neighbor list, the outer distance was $r_l = 11.84$ Å and the inner distance was $r_c = 10.42$ Å. The pair list was updated whenever any atom moved a distance greater than 0.5($r_l - r_c$), leading to updates roughly every 25 steps. Electrostatic interactions are evaluated using the charge-group method, to be consistent with the parametrization of the GROMOS force field. The SHAKE algorithm was used to constrain bond lengths permitting a 2-fs timestep. Periodic boundary conditions were used. The simulation was carried out at constant number, volume, and temperature (NVT), with velocity scaling to control the temperature at 303 K. The runs were all performed on a PC with dual 1.6GHz AMD Athlon processors. A 1-ns simulation takes about 2 weeks on a single processor.

**Equilibration**

The 1AKE* system was equilibrated according to the following protocol. A simulation box of side 59.19 Å was filled with an equilibrated sample of 6912 waters, and the 1AKE* structure, totaling 2162 atoms, was centered in the simulation cell. The waters that overlapped any 1AKE* atom were discarded if $r_{OJ} < \sigma_{OJ}$, where $r_{OJ}$ is the distance between water oxygen (atom O) and protein atom (atom j), and $\sigma_{OJ}$ is the van der Waals distance parameter. This procedure eliminated 1470 waters. One-body forces, with force constant $k = 30$ kcal/molÅ$^2$, were used on all protein and ligand atoms for the initial 50,000 steps. This procedure allowed the waters to equilibrate with the protein and with each other. The van der Waals length ($\sigma$) and well depth ($\epsilon$) values for the water oxygen–Mg$^{2+}$ interactions were designed to maintain coordination throughout the simulations. To enforce the initial water coordination of Mg$^{2+}$, the parameters listed in Table I were used to equilibrate the system and then modified as indicated for the duration of the simulation.
TABLE I. Water-Mg\textsuperscript{2+} Van der Waals Parameters for the Mg\textsuperscript{2+} Coordination Waters During the MD Equilibration and Simulation Periods

<table>
<thead>
<tr>
<th>Equilibrate with</th>
<th>Run with</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$ (Å)</td>
<td>$\sigma$ (Å)</td>
</tr>
<tr>
<td>$\epsilon$ (K)</td>
<td>$\epsilon$ (K)</td>
</tr>
<tr>
<td>Water 218</td>
<td>2.16137</td>
</tr>
<tr>
<td>Water 219</td>
<td>2.07814</td>
</tr>
<tr>
<td>Water 220</td>
<td>1.88086</td>
</tr>
<tr>
<td>Water 221</td>
<td>2.19517</td>
</tr>
</tbody>
</table>

\(a\)The van der Waals interaction between atoms \(i\) and \(j\) is
\[ V = 4\epsilon_i (r_{ij} / \sigma_i)^{12} - (r_{ij} / \sigma_i)^6, \]
where \(\sigma_{ij} = (\sigma_i + \sigma_j) / 2\) and \(\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}\).

After the initial 50,000 steps with one-body forces on, the one-body force constant was linearly reduced to 0 over 30,000 steps. The system was then run for another 20,000 steps to make sure that the system’s strain was sufficiently reduced to initiate unconstrained dynamics. A 0.5-ns trajectory was run at 273 K, after which the temperature was linearly brought to 303 K over 10,000 steps. After the temperature had reached 303 K, the MD production trajectory proceeded for 3 ns.

H-bond–Salt Bridge Analysis

Interactions of AMP and ATP with 1AKE\textsuperscript{*} residues were analyzed through H-bond analysis. Donor and acceptor H-bond–Salt Bridge Analysis steps to make sure that the system’s strain was sufficient to induce differences. After the initial 50,000 steps with one-body forces on, the one-body force constant was linearly reduced to 0 over 30,000 steps. The system was then run for another 20,000 steps to make sure that the system’s strain was sufficiently reduced to initiate unconstrained dynamics. A 0.5-ns trajectory was run at 273 K, after which the temperature was linearly brought to 303 K over 10,000 steps. After the temperature had reached 303 K, the MD production trajectory proceeded for 3 ns.

H-bond–Salt Bridge Analysis

Interactions of AMP and ATP with 1AKE\textsuperscript{*} residues were analyzed through H-bond analysis. Donor and acceptor H-bond–Salt Bridge Analysis steps to make sure that the system’s strain was sufficiently reduced to initiate unconstrained dynamics. A 0.5-ns trajectory was run at 273 K, after which the temperature was linearly brought to 303 K over 10,000 steps. After the temperature had reached 303 K, the MD production trajectory proceeded for 3 ns.

H-bond–Salt Bridge Analysis

Interactions of AMP and ATP with 1AKE\textsuperscript{*} residues were analyzed through H-bond analysis. Donor and acceptor H-bond–Salt Bridge Analysis steps to make sure that the system’s strain was sufficiently reduced to initiate unconstrained dynamics. A 0.5-ns trajectory was run at 273 K, after which the temperature was linearly brought to 303 K over 10,000 steps. After the temperature had reached 303 K, the MD production trajectory proceeded for 3 ns.

RESULTS AND DISCUSSION

Validation of Simulation

The radius of gyration of the protein increases very slowly during the first 740 ps, with an average of 17.18 ± 0.09 Å during that period and, subsequently, suddenly increases and stabilizes to an average of 17.60 ± 0.11 Å during the remainder of the simulation, indicating that the overall dimension of the protein is stable. (The radius of gyration of 1AKE in the crystal structure is 16.635 Å.)

The RMSD (i.e., root-mean-square deviation from the starting 1AKE\textsuperscript{*} structure) for all Cα atoms changes slowly from 1.5 to 3.1 Å during the first 800 ps of the production trajectory; then, between 1100 and 1300 ps, it increases more sharply to finally plateau between 4.0 and 5.0 Å during the rest of the simulation. The RMSD for the Cα atoms of the core domain (residues 1–30, 60–120, and 160–214) increases very slowly to 2.3 Å during the first 1 ns of the production trajectory, then plateaus between 2.1 and 2.75 Å before finally increasing after 2800 ps. The RMSD for the Cα atoms of the core plus the AMP-bd domains (residues 1–120 and 160–214) increases during the first 800 ps at the same rate as the RMSD for all Cα atoms, then oscillates between 2.5 and 3.4 Å during the next 2000 ps, before finally increasing after 2800 ps. These results indicate that the deviation from the crystal structure is not distributed uniformly along the sequence, and that most of the large deviations (in particular after the first 1200 ps) occur in the LID domain.

Coordination Waters During the MD Equilibration

The RMSD (i.e., root-mean-square deviation from the average structure) were calculated for all Cα carbons as
\[
\text{RMSF}_\alpha = \sqrt{\frac{1}{n} \sum (r_{n} - \langle r_n \rangle)^2},
\]
with \(r_n\) being the average position of the Cα carbon of residue \(n\) over time, and \(\langle r_n \rangle\) being the position at every step. Simulation RMSF values are compared with 1AKE crystallographic RMSF values in Figure 2. Crystallographic RMSF values were calculated from the B values using the equation
\[
\text{RMSF} = \sqrt{\frac{3\text{B}}{8\pi^2}}.
\]
Regions with low RMSF values (e.g., helices \(\alpha_4\) and \(\alpha_9\), and strand \(\beta_3\)) in the simulation coincide well with the regions with low crystallographic RMSF values. In contrast, the regions showing high simulation RMSF values and high crystallographic RMSF values do not always match each other. Residues 41–48 (the loop between helices \(\alpha_2\) and \(\alpha_3\), and 132–140 (the loop between strands \(\beta_2\) and \(\beta_3\)) show high simulation RMSF values that were not expected from their average crystallographic RMSF values. Residues 74–79 (the loop between helix \(\alpha_4\) and strand \(\beta_3\)) show the opposite trend: Although they have the highest crystallographic RMSF values in 1AKE (1.72 ± 0.17 Å for 74–79 Cα’s against an average B factor of 1.001 ± 0.19 Å for all Cα’s), these residues show simulation RMSF values (2.16 ± 0.49 Å) only slightly above the average for the protein (1.93 ± 0.94 Å).

The low crystallographic RMSF values associated with residues 42–48 in 1AKE might be due to the fact that residues 44, 45, 47, and 48 are involved in crystal contacts. In the simulation, the absence of crystal contacts results in a larger conformational freedom for this region of the protein. In this respect, it is interesting to note that residues 46–48 adopt different secondary structures in 1AKE complexes I and II. Müller et al. later noticed that residues 44–48 are part of helix \(\alpha_3\) in the apo-enzyme (PDB code: 4AKE). These crystallographic observations suggest that this region is highly susceptible to conformational changes, and they agree with the large RMSF values observed for residues 41–48 in our simulation.

Because residues 74–79 have the highest crystallographic RMSF values in 1AKE, and because these residues have lower crystallographic RMSF values in the apo-enzyme, it was suggested that this loop is part of an energy counterweight in ECAK. The low RMSF values observed for residues 74–79 in our simulation do not support this energy counterweight hypothesis. Another possibility is that the high crystallographic RMSF values of residues 74–79 are an artifact of crystallization. In 1AKE, this loop is involved in crystal contacts: Glu75 forms a H-bond and salt bridge with Asp76 and Arg78 of another molecule, respectively. These interactions might force this loop into a conformation that is not 100% occupied or that would not be favored in solution. Some of the different conditions may be attributable not only to the different conditions involved in crystalline versus liquid
states but, of course, also to the 1AKE* model, where $\text{AP}_5\text{A}$ is replaced by the ATP-AMP-Mg$^{2+}$-4(H$_2$O) complex. The highly charged ATP and AMP phosphate chains and Mg$^{2+}$ can exert significant forces that extend far into the protein. Thus, differences can arise from this replacement.

The integrity of the ATP-AMP-Mg$^{2+}$-4(H$_2$O) complex as the simulation proceeds is well maintained, as displayed in Figure 3. The Mg$^{2+}$ keeps its ligation to the ATP $\beta$- and $\gamma$-PO$_4$ oxygens, and the 4 waters are coordinated to the Mg$^{2+}$, forming an octahedral coordination sphere. The distance between the ATP $\gamma$-PO$_4$ and AMP $\alpha$-PO$_4$ is reasonable for a catalytically competent configuration, and residues thought to be key for catalysis are positioned close to the reactant complex, as we detail below.

The stereochemical quality of 5 snapshots of 1AKE$^\kappa$ taken between 900 and 1500 ps was checked using the software PROCHECK v 3.5.4. The results show that 74–79% of the residues are in the most favored regions of the Ramachandran plot, and 17–23% are in the additional allowed regions (around 97% total). These numbers are comparable to values obtained from solution structures of proteins from NMR. These observations suggest that the protein does not go through disallowed dihedral conformations during the simulation.

**Substrate Specificity**

**AMP binding site**

AKs, in general, are known to have higher specificity for their nucleoside monophosphate (NMP) substrate than for their nucleoside triphosphate (NTP) substrate. Extensive mutagenesis studies and structural data have shown that AKs control AMP specificity through several interactions with highly conserved residues. The H-bonding pattern between AMP and ECAK was analyzed...
over the complete 3-ns simulation (Fig. 4). Figure 5 shows all the interactions of AMP and ATP with ECAK residues that are present during the 3-ns simulation.

In the 1AKE crystal structure of the ECAK - AP5A complex, AMP's adenine interacts with ECAK through 5 strong H-bonds (2 H-bonds to backbone atoms and 3
H-bonds to side-chains) and several other less polar interactions. A strictly conserved glutamine residue, Gln92, plays a critical role both in catalysis and AMP specificity through H-bonds with AMP's adenine N6. The orientation of Gln92's carboxamide is fixed by additional H-bonds to a backbone amide (from Arg88) and a conserved carboxylate (Asp61). In the simulation, plots (not shown) of the H-bonds between Asp61, Arg88, and Gln92 versus time show that Gln92's side-chain is indeed fixed by very stable H-bonds. Specifically, Gln92:NE2 H-bonds with Asp61 side-chain oxygens OD1 and OD2, and Gln92:OE1 H-bonds with Arg88:N. Two other H-bonds between Thr31:OG1, Gly85:O, and the adenine N6 that are stable over the trajectory also help in discriminating against guanosine monophosphate (GMP), which lacks the NH2 group. Additionally, Thr31:OG1 forms a stable H-bond with AMP’s adenine N7. Other H-bonds like adenine N6 to Phe86:N and adenine N7 to Gly32:N are on and off during the simulation (Fig. 4). Another residue, Arg167, forms transient H-bonds with the AMP adenine N6 and N7, as well as with AMP’s αPO4, and with the AMP ribose O3 (Fig. 4). These bonds are possible because Arg88’s side-chain is positioned parallel to the π rings of AMP’s adenine. In 1AKE, Arg88 occupies 2 different conformations in the 2 molecules of the asymmetric unit. In the BSAK·AP5A complex, Arg88 adopts yet another configuration, suggesting that specific interactions between Arg88 and AMP are not what makes this residue essential in AMP binding. Arg88 is highly conserved in AKs, and its function in substrate binding and/or catalysis is the subject of ongoing discussions.

Only 1 good (2.68 Å) H-bond between AMP ribose and Lys57:O is seen in the 1AKE crystal structure. Interactions between Lys57:O and Gly56:O are present only during the first 500 ps of the simulation (Fig. 4). However, a H-bond between the side-chain of Arg167 (involving NH1, NH2, or NE) and the ribose is present during the entire simulation (Fig. 4). In 1AKE, Arg167 is within H-bonding distance of the fourth (β) phosphate in AP5A. Examination of snapshots from the simulation show that the side-chain of Arg167 moves toward the ribose moiety early in the simulation. The difference in Arg167 orientation between the 1AKE crystal structure and our simulation may be due to the presence in 1AKE of the extra phosphate in AP5A. In 1AKE, Glu170’s side-chain points away from the ribose. Starting after 700 ps of our simulation, Glu170 (OE1 and OE2) forms a stable, double H-bond with AMP ribose O2 and O3 (Fig. 4). Whether this double H-bond is significant for AMP is open to question, because Glu170 is replaced by an alanine or a valine in the B. stearothermophilus, Thermotoga neapolitana, yeast, and bovine heart mitochondrial AKs.

Several arginines (Arg36, Arg88, Arg119, Arg123, Arg131, and Arg167) are situated in the active site cleft surrounding AP5A’s phosphate groups in 1AKE, presumably to help stabilize the negatively charged phosphates. Some of these arginines have been shown to participate in phosphate binding and catalysis. In 1AKE, AP5A’s e-PO4 (that corresponds to AMP’s α-PO4) is well fixed through interactions with R36:NH2, R88:NH2, and R156:NH2, the latter interaction being weaker than the first two. AP5A’s e-PO4 sits between the head groups of Arg88 and Arg36. Chemical modification of ECAK with phenylglyoxal completely inactivates the enzyme. Reinstein et al showed that this complete inactivation is obtained by modification of a single arginine: Arg88. Mutant R88G is
1000 times less active than the wild-type ECAK, the Michaelis constant (Km) values for ATP and AMP increasing up to 5-fold and 85-fold, respectively.

The simulation results show that R88:NH2 and R88:NE have very stable interactions with AMP's α-PO₄. These interactions are salt bridge–like considering the charges of the groups involved; hence, they are much stronger than H-bonds. In a previously published 300-ps simulation of the ECAK·AP₅A complex, the interactions between Arg88 and AP₅A's ε-PO₄, analyzed as H-bonds, were not preserved. In the 1AKE* simulation, H-bonds involving Arg36:NH₁ and NH₂, and Gln170:OE₁ and OE₂ appear after 500 ps. The Arg36 H-bonds are to the ribose, whereas in the 1AKE crystal structure, Arg36 H-bonds an ε-PO₄ oxygen. It seems that the removal of AP₅A's ε-PO₄ to create ATP and AMP (with the resulting loss of 2 negative charges) weakens some of the charged and polar interactions between phosphates and arginines and rearranges the H-bond and salt bridge–like interactions. The presence of Mg²⁺ can also influence the distribution of charged interactions between the protein Arg residues and the phosphate groups. Permanent H-bonds are formed between AMP's α-PO₄ and the Lys13:NZ. Lys13, situated in the P loop of all AKs, has been described as the “invariant lysine” due to dramatic loss of activity by chemical modification. In summary, interactions of Thr31, Arg88, Gly85, and Gln92 with the adenine N6 are very persistent providing the possibility of discrimination. Arg88 and Lys13 are crucial for binding the α-PO₄ and positioning it suitably to receive a phosphate group from ATP.

**ATP binding site**

The ATP-binding site is formed by residues in the P-loop (residues 7–15), part of the LID domain, and residues in the C-terminal helix with its connecting loop region (residues 198–214). The ATP H-bonds found in the simulation are shown in Figure 4, and the stable interactions are indicated in Figure 5. Similar to other AKs, ECAK is less specific for ATP than for AMP, and most of the tight interactions between the protein and Mg·ATP involve the ATP phosphates. In the ECAK·AP₅A complex, ATP's adenine moiety is only loosely bound to the protein through a single H-bond between Lys200:O and ATP's adenine N6. This H-bond leads to a preference for ATP over guanosine triphosphate (GTP). The simulation shows (Fig. 4) that this H-bond breaks and then reforms during the 3 ns, due to fluctuations in the position of ATP's adenine. The position of Lys200 itself does not deviate much from the crystal structure, and the H-bond between Lys200:O and adenine N6 is formed whenever the adenine moiety of ATP moves to the appropriate position.

In the crystal structure, the ATP adenine is parallel to the Arg119 guanidinium group, suggesting a cation–π interaction. The backbone of Pro201 and Val202 are in van der Waals contact distance of the adenine atoms. The Arg119–ATP interaction is not stable during the simulation, owing again to the large movements of ATP's adenine. N7 of the adenine ring forms an H-bond with Arg119:NH1 and NH2. Although this bond is relatively infrequent (occurring only about 8% of the time), it occurs on and off during the simulation (Fig. 4), concurrent with the movement of the adenine ring. A similar unstable H-bond is also seen between Val202:O and ATP:AN6. Two new H-bonds between ATP:AN6 and either Tyr199:O or Gly198:O also form more or less in synchronization with the movement of the adenine ring. Taken together, none of the H-bonds are strong enough to hold ATP's adenine in place. Moreover, these interactions (with the exception of the H-bond involving Lys200) are independent of the base (it can be either adenine or guanine). While Lys200 is not conserved among the AKs, Arg119 is conserved in all known NMP kinases. These observations concur with experimental data confirming that ATP adenine is not tightly bound to the enzyme, and that it can be substituted with other NTPs such as GTP. In ECAK, the ATP adenine is quite solvent accessible, and so is Arg119; these observations probably lead to the extensive sampling of conformation space observed for ATP adenine in our MD simulation. They might also explain why the R128A mutation (equivalent to R119A in ECAK) only moderately perturbed the kinetic parameters of the chicken muscle AK1 enzyme.

Only one strong H-bond between the ATP ribose O3 and Tyr133:O exists in the crystal structures of ECAK·AP₅A and ECAK·AMP·AMPPNP. It is thought that this H-bond plays a significant role in the closing of the LID over ATP, since it is one of the few interactions between the LID and AP₅A. The H-bond between Tyr133 and ribose is present only during the first 750 ps of our simulation. Even during this time, the distance between Tyr133:O and ribose:O3 is close to 3.5 Å, making it a very weak interaction and hence unlikely to be a driving force in LID closure. A weak H-bond involving His134:O and ATP's ribose O2 (at 3.48 Å in 1AKE) is present until 1000 ps. Since His134 lies in the LID domain that shows considerable backbone flexibility during the simulation (Fig. 2), this H-bond is present only during the first nanosecond. An H-bond between ATP ribose O3 and Arg124:N appears after ~1290 ps. This H-bond is fairly stable, occurring 78% of the time. Taken together, the ATP ribose, like the adenine, is also not very tightly bound to the enzyme.

In contrast to ATP ribose and adenine, the ATP phosphates have extensive contacts with the protein. The triphosphate-binding site is a highly charged pocket formed by the LID and the core. It is lined with several arginines and the P loop (glycine-rich), and it is often described as a giant anion hole. Phosphate binding, orientation, and transfer are mediated by Arg36, Arg88, Arg123, Arg156, Arg167, and Lys13, all interacting with the P-loop backbone. As in the case of AMP, interactions between the phosphate oxygens and the positively charged side-chains of arginines and lysines are salt bridge–like in character and hence much stronger than H-bonds.

In 1AKE, α-PO₄ shows 3 good H-bonds. It is connected in the P loop to Thr15:N (at 2.85 Å), Thr15:OG1 (at 2.71 Å), and to Gly12:N (at 2.98 Å). The α-PO₄:O2 is also within H-bonding distance from Arg123:NH2 (at 2.91 Å). Thr23 (Thr15 in ECAK) in chicken muscle AK1 displays strong
nuclear Overhauser effects (NOEs) with the adenosine moiety of Mg-ATP, although it is within H-bonding distance from the ATP -PO4 in the AK1 crystal structure. Phosphate stereoechemistry experiments in AK1 also suggest that Thr23 directly interacts with -PO4. This result agrees with the 1AKE structure, in which Thr15:OG1 forms 2 stable H-bonds with -PO4. Based on these observations, it was suggested that in AK1, Thr23 plays a role in catalysis through direct interaction with ATP. Although Thr23 is conserved in all known AKs, the mild perturbations of mutant T23A’s kinetic parameters (kcat and Km) are insufficient to confirm a functional role for Thr23. In our simulation, Thr15:OG1 and Thr15:N form permanent H-bonds with ATP adenine N6, N7, and N9, in keeping with its proximity to ATP adenine, as seen in the solution structure of AK1 and in the crystal structure of the yeast AK·Mg-AP5A complex. These H-bonds are weaker, lasting for at most 500 ps.

The H-bond between Gly12:N and ATP’s -PO4, seen in the 1AKE structure, is present only during the simulation. The on–off nature of this H-bond in the simulation originates from fluctuations in the position of the phosphate relative to Gly12. Instead, we observe a stable H-bond between ATP -PO4 and Gly14:N throughout the simulation. The salt bridge–like interaction with Arg123:NH1 is present only for 100 ps, with the side-chain of Arg123 moving away from ATP early in the simulation. The ATP phosphate chain adopts a slightly different position in the equilibrated 1AKE compared to that in 1AKE, -PO4 is further away from Arg123 in 1AKE than in 1AKE, leading to a weaker salt bridge–like interaction. The side-chain of Arg123 is thus able to fluctuate and move away from -PO4 during the simulation. This side-chain, however, moves within 4.5 Å of ATP -PO4, thus creating another potential salt bridge interaction. Note that in the 1AKE crystal structure, Arg123:NH1 is 3.94 Å from a -PO4. In our simulation, Arg123:NH1 is within H-bonding distance of Lys13, and within 4.5 Å of the side-chains of Arg123 (NH1 and NH2) and Arg156. These H-bonds have donor–acceptor distances of 3.3 Å or more, indicating weaker interactions with the protein as compared to ATP’s -PO4 and -PO4. In our simulation, the salt-bridges between and Arg156:NH1 and NH2 are present as in the crystal structure. The interactions with Lys13 and Arg123 are not stable in the simulation. They are replaced by other interactions arising from the presence of Mg2+, and from the enhanced conformational range available to -PO4, which is no longer covalently attached to AP5A’s δ-PO4. -PO4 is now free to rotate and, with its negative charges, it strongly interacts with Mg2+. An additional H-bond not seen in the crystal structures is seen in our simulation between Gly10:N and -PO4. -PO4 is stable throughout the simulation and it is the only link between -PO4 and the P loop.

In summary, we see about the same interactions between the ATP phosphate chain and the P loop (main-chain and side-chains) as in the crystal structure, but we see more and stronger interactions with Arg156 than with Arg123, for both and -PO4. Without exception, all permanent H-bonds between ATP and the protein in 1AKE in the simulation involve the 3 ATP phosphate groups (Figs. 4 and 5), the ATP adenine and ribose moieties being only loosely bound to the protein.

Role and conformation of active site residues around Mg2+

Kinases require a divalent cation, usually Mg2+, for catalysis. The ATP and -phosphates are coordinated to Mg2+, and the true substrate for AKs is Mg2+-ATP. Mg2+ serves to neutralize the highly negative charge on the ATP triphosphate, thereby reducing electrostatic repulsion of the transferred phosphate and increasing the efficiency of nucleophilic attack. In AKs, conserved serine, threonine, or aspartate residues participate in binding the Mg2+ cation, coordinating it either directly in the first coordination sphere or indirectly through water molecules.

In porcine muscle AK·CoATP, AK·CoGTP, and AK·CoGDP, 31P NMR relaxation rates indicate that Co2+ is directly coordinated to only -PO4 (GDP) or -PO4 and -PO4 (ATP and GTP). In addition, -PO4 is too far from Co2+ to be in the first coordination sphere, but also too close to be in the second coordination sphere (not enough
room for a water molecule). These results suggest that ATP–Mg$^{2+}$ binds AK as a βγ bidentate complex rather than an αβγ tridentate complex. In the structure of the BSAK · Mg$^{2+}$–AP$_5$A complex, Mg$^{2+}$ is coordinated to the phosphates corresponding to ATP’s β- and γ-phosphates, and to 4 water molecules numbered 300–303.\textsuperscript{11} Mg$^{2+}$–AP$_5$A is anchored to BSAK through several H-bond networks mediated by the 4 Mg$^{2+}$-coordinating water molecules. Waters 300 and 303 help orient the α-PO$_4$ of AMP. Water 300 is also positioned to interact with Arg36 and Asp33, thus linking Mg$^{2+}$–AP$_5$A to the protein. Water 302 is within H-bond distance of ATP’s 3 phosphate groups, reinforcing the linkage between the phosphate chain and Mg$^{2+}$. Water 301 H-bonds to Asp84:OD2 and Gly14:N, and it is the only water that does not interact with the phosphates. It is evident that this H-bond network is important for orienting the donor and acceptor phosphates and aiding the ternary complex to proceed to the transition state. Based on their 300-ps simulation results of ECAK · AP$_5$A, Kern et al. predicted that 3 aspartate residues, Asp84, Asp33, and Asp110, are possible candidates for binding the Mg$^{2+}$ complex.\textsuperscript{29} Ser30, Thr31, and Thr89 were also suitably placed and were available to interact with Mg$^{2+}$. Orientation of the side-chains of these residues and the position of the phosphate chain left just enough room for the Mg$^{2+}$ ion.

In the simulation, Mg$^{2+}$ remains coordinated to 4 waters (218–221) as dictated by the force field (see the Methods section). The water–Mg$^{2+}$ distances are short in the crystal structure, indicating that these waters are ligated to the ion. Then, one can choose to modify the force field by inserting specific “covalent” bonds, or by using noncovalent but relatively strong interactions,\textsuperscript{55} as we have done. Even though no such interactions were introduced between Mg$^{2+}$ and ATP β- and γ-phosphates, we find that throughout the simulation Mg$^{2+}$ is also coordinated to ATP’s β-PO$_4$ through its O2 oxygen, and to ATP’s γ-PO$_4$ through its O1 oxygen. As shown in Figure 6, the coordination is close to octahedral. Furthermore, the additional flexibility introduced by having ATP/AMP versus AP$_5$A in the structure, as in the 1ZIO crystal structure, still preserves the octahedral geometry.

A second coordination sphere for Mg$^{2+}$ includes Asp84: OD1 and OD2 that have waters 219 and 221 (equivalent to waters 301 and 303 in BSAK) screening this direct salt bridge–like interaction. The H-bonds between water molecules 219 and 221 and the side-chain of Asp84 are very stable over the 3-ns trajectory (not shown), indicating the importance of Asp84 in binding the Mg$^{2+}$–water complex. In AK1, mutation of the corresponding Asp93 to alanine decreased the $k_{cat}$ 650-fold.\textsuperscript{56} The D93A mutant showed no structural perturbation as evidenced by NMR analysis, indicating a local effect. Moreover, the AK1 D93A mutant had a markedly lower affinity for Mg$^{2+}$.\textsuperscript{56} Similar effects were also observed for the ECAK D84A mutant.\textsuperscript{26} Considering the fact that Mg$^{2+}$ plays a dual structural and chemical role in catalysis, these results suggest that Asp84 is important in the structural role of Mg$^{2+}$. Specifically, in the AK1 D93A mutant, Mg$^{2+}$ may be unable to orient the phosphate chain for the transfer.\textsuperscript{56}

Throughout the simulation, Asp84’s main- and side-chains form a stable network of H-bonds. Besides the above-mentioned H-bonds with waters 219 and 221, Asp84 forms stable H-bonds with Ser30 and Thr31. The H-bond between Asp84 and Thr31 backbone atoms is particularly significant, since Thr31 is also stably linked to AMP’s adenine. Thus, these interactions anchor AMP to the Mg$^{2+}$ complex involving the ATP phosphate chain. Asp84 interacts with Ser30 and Ile29 through strong H-bonds and also forms a persistent salt bridge with Lys13:NZ in the P-loop. Additionally, Ser30 is by itself coordinated to the Mg$^{2+}$ ion through water molecule 219.

As in BSAK, water molecule 220 (water 302 in 1ZIO), is within H-bonding distance of ATP α-, β-, and γ-phosphates, strengthening the Mg$^{2+}$–phosphate complex and properly orienting the ATP phosphate chain. Interactions not seen in BSAK but transiently present in the 1AKE\textsuperscript{a} simulation include H-bonds between Arg156 and waters 220 and 218. Arg156 has been implicated in the stabilization of the transition state. Another rather weak H-bond is also seen between Gly14:N and water molecule 219. A corresponding H-bond is present in 1ZIO involving Gly14 and water molecule 301.

Lysine, with its backbone amide, long and flexible side-chain, and charged head group, is well-suited for a multifunctional role. These features are apparent in Figure 6, where Lys13 stabilizes the active site by forming H-bonds and salt bridge–like interactions with both ATP and AMP using its flexibility and multifunctionality. The
other interactions of Lys13 are discussed in previous sections.

**Catalytic Mechanism**

The 1AKE* simulation can be used to suggest whether the mechanism of phosphoryl transfer is associative or dissociative, at least from the perspective of geometrical requirements. The consensus is that in enzymes in general, and in AKs in particular, an associative mechanism is operative. The geometric criterion for an associative mechanism is that the transition state has sufficiently short entering and leaving group distances to support reasonably large fractional bond numbers, \( n \). For example, if an SN2 mechanism were operative (\( n = \frac{1}{2} \)), the \( P-O \) distance would be 1.91 Å, while if a fully associative mechanism were operative (\( n = 1 \)), the \( P-O \) distance would be 1.73 Å, the \( P-O \) single-bond distance.

In Figure 7, the ATP \( \gamma-P \) to AMP \( \alpha-P \) distance is plotted during the 3-ns simulation time. For most of the time, this distance ranges between 4.5 Å and 5.0 Å, about 1 Å greater than the 3.8-Å van der Waals \( P-P \) contact distance. In the absence of the \( Mg^{2+} \) cation and Lys13, with its ability to span ATP and AMP, it would be difficult to support the close approach of these (negatively charged) phosphate groups. The \( P-P \) distance suggests that an associative mechanism could be operative in AKE, as illustrated in Figure 8. The atoms ATP \( \beta-P \), Lys13:NZ and AMP \( \alpha-P \) form an essentially equilateral triangle with apex Lys13:NZ and equal legs ATP \( \beta-P \)-Lys13:NZ and Lys13:NZ-AMP \( \alpha-P \). The atoms Lys13:NZ, ATP \( \beta-P \), ATP \( \gamma-P \), and AMP \( \alpha-P \) form a plane, where the ATP \( \beta-P \)-AMP \( \alpha-P \) distance is approximately twice that of the ATP \( \beta-P \)-ATP \( \gamma-P \) distance. Thus, there would be room to rotate the ATP \( \gamma-P_4 \) into a line formed by ATP \( \beta-P \), ATP \( \gamma-P \), and AMP \( \alpha-P \) and produce a trigonal bipyrimidal phosphorane transition state, as suggested by Reinstein et al. In this configuration, Lys13:NZ would be positioned above the 3 apical oxygens, those in the plane perpendicular to the ATP \( \beta-P \), ATP \( \gamma-P \), and AMP \( \alpha-P \) line. The \( Mg^{2+} (4H_2O) \) complex is below the above-mentioned plane, with \( Mg^{2+} \) equidistant from the oxygens (shown as spheres) associated with ATP \( \beta-P \) and ATP \( \gamma-P \). From these geometrical considerations, the role played by \( Mg^{2+} (4H_2O) \) and Lys13 in maintaining the proximity and orientation of the ATP and AMP phosphates appropriate for phosphoryl transfer is clarified.

**CONCLUSIONS**

MD simulation of the ECAK \( \cdot Mg^{2+} \)-ATP \( \cdot AMP \) ternary reactant complex that we constructed shows general agreement with conclusions drawn from the analysis of the crystallographic and other studies of ECAK complexed with AP5A. The strong repulsive interaction between the terminal, negative phosphate groups that results from the split of AP5A into ATP and AMP is counteracted by the presence of \( Mg^{2+} \) and, presumably, some of the positively charged surrounding residues. The AMP adenine N6 strongly interacts with the protein through H-bonds, which
forms the basis for NMP specificity. In contrast, and also in agreement with conclusions drawn from the crystallographic data, the ATP adenine and ribose moieties are only loosely bound to the protein, while the phosphates interact strongly with nearby residues.

Octahedral coordination of the Mg$^{2+}$ by 4 waters and by oxygens of ATP ß- and ß-phosphates is maintained throughout the simulation. The conserved Lys13 in the P loop bridges the ATP and AMP phosphates, a phenomenon that relies on the lysine side-chain’s unique properties of flexibility and H-bond–salt bridge capability. In addition to Lys13, the Mg$^{2+}$’s coordination waters, and some surrounding charged residues maintain the AMP ß-phosphate and ATP ß- and ß-phosphates in a configuration suggesting that phosphoryl transfer occurs by an associative mechanism in AK.

The simulation results could provide a starting point for a combined quantum mechanical–molecular mechanical simulation of the reaction mechanism for phosphoryl transfer.

REFERENCES