MOLECULAR DYNAMICS SIMULATIONS
OF THE ADENYLATE KINASE ENZYME
AND ITS MUTANT STATES

by

BEATA DULKO-SMITH

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2022

Acknowledgements

I would like to thank my advisor, Dr Kwangho Nam, for his unyielding patience which I have challenged on many an occasion, for generosity in sharing his scientific expertise, insight and wisdom, and for his immense compassion. Most of all, though, I am grateful to him for pushing me head-first into FORTRAN programming which, quite apart from being a powerful mediator between human and the machine, on a more mundane level, liberated me from the shackles of Excel.

My thanks and gratitude for their time and insight go to the members of my doctoral committee: Dr Peter Kroll, whose disarming sense of humor, honesty, and perspective never failed to lighten up my mood; Dr Jongyun Heo whose formative influence on my character and challenging me to go beyond my insecurities and stand my ground is priceless and appreciated; Dr Robin Macaluso for her openness, support, and volunteering her time to fill a vacancy on my committee, and finally, Dr Brad Pierce, who helped me jump the initial hurdle of the proposal defense. I would also like to extend thanks to our collaborators at Umeå University in Sweden, Dr Magnus Wolf-Watz for his comments, mentoring, and contagious enthusiasm, Dr Pedro Ojeda-May for his friendly presence on the other side of email, and Dr Jörgen Ådén for his meticulous proofreading of our manuscript.

As is often the case, no undertaking can ever be successful without the support of the administrative staff who *de facto* preside over the fate of many a student. Stephanie Henry, Debbie Cooke, and now retired Jill Howard have been the people to turn to in times of bureaucratic distress, low blood sugar/caffeine levels, or when in need of a friendly chat. Most importantly, I am indebted to Stephanie, an insightful cat behaviorist, for always generously sharing her expertise and widening my horizons in areas where I never thought I would venture. I have been very fortunate to cross paths at UTA with some of the most wonderful people whom I feel privileged to call friends: Thanh Thuy Vuong, Ilia Ponomarev, Poroshat Taheri, Muhammed Shameem Kakkattuparambil, and Abdul Raafik Arattu Thodika, as well as my colleagues Dr James Mao and Dr Yunwen Tao.

Words quite literally cannot express my complete gratitude and appreciation for the generosity I have received from people who welcomed me into their family and furnished me with a home, car and a cat asking for nothing in return: Lorette Whitebread, Sidnee and Bernie Silva, Linda Green, and Jane Munyaneza are the people who showed me the meaning of friendship, generosity, kindness, reliability, and unabated loyalty. I would have achieved nothing without them. Equally, I am forever grateful, honored, and humbled to have met my dearest Zen master, Thich Minh Thien, who would take my mind on a rollercoaster ride with one piece of advice for all that troubled me: sit. Thanks to the wonders of technology, I could draw on the generous support from my friends on the other side of the Atlantic who stood by me in times of need and whom I treasure for their wisdom, encouragement, and comradeship: Hanh Thao Ho, Olga Krajewska-Kajzer, Magda Pietrzyk, Anna Wyzykowska, and Piotr Mielcarek.

Last but not least, I am deeply indebted for the continuous encouragement, care, and unconditional love that I received from my beloved family, Danuta Dulko and Marcus Smith, who endured the years of separation that I inflicted upon them in order to pursue this degree. It is only through their unwavering support that I got thus far. Most importantly though, I must acknowledge the significance of someone who certainly cares the least for such gestures but has nevertheless been my most faithful companion throughout those years: Molly McGee-Kocinski, my most cherished feline friend.

In loving memory of

my father, Tadeusz Dulko

and Jeffri Tori, my friend

Table of Contents

List of figures and tables

CHAPTER 1

CHAPTER 2

CHAPTER 3

Table of abbreviations

ADP	Adenosine 5’-diphosphate
AK / AdK	Adenylate Kinase [enzyme protein]
AMP	Adenosine 5’-monohosphate
Ap5A / AP5	P1,P5-Di(adenosine-5’)pentaphosphate
ATP	Adenosine 5’-triphosphate
CHARMM	Chemistry at HARvard Macromolecular Mechanics
COM	Center Of Mass
GPU	Graphics processing unit
GTP	Guanosine 5’-triphosphate
MD	Molecular Dynamics
ns	Nanosecond
PDB	Protein Data Bank
PME	Particle Mesh Ewald [summation]
PS	Product state
RMSD	Root-mean-square deviation [of atomic positions]
RMSF	Root-mean-square fluctuations [of atomic positions]
RS	Reactant state
QM/MM	[hybrid] Quantum Mechanical / Molecular Mechanical [calculations]
TIP3P	Transferable Intermolecular Potential with 3 Points
WT	Wild type
µs	Microsecond

Abstract

Adenylate kinase (AK) is a small, monomeric enzyme present across all domains of life. It is responsible for maintaining homeostasis of an intracellular adenylate pool and catalyzes a reversible transfer of the phosphoryl group from NTP (nucleoside triphosphate) to NMP producing two NDP molecules. While most known isoforms take ATP as phosphate donor, AK3 present in mitochondrial matrix employs GTP. Crystallographic investigations revealed that the basis for its selectivity for GTP over ATP came from an additional hydrogen bond between guanosine's O6 and backbone nitrogen of threonine is responsible for recognition of GTP. Molecular dynamics (MD) simulations have demonstrated that stabilization and activation of the substrate are facilitated through a double stacking interaction of guanine's pyrimidine ring with arginine and phenylalanine, while the AK isoform operating on ATP substrate stabilizes the pyrimidine ring with one π-stacking to an arginine.

AK is not compartmentalized into functionally distinct modules, which would separately control the catalytic step and product release. Therefore, the two events might be connected and mediated by catalytic residues. Enzyme kinetics measurements indicate that the rate of enzymatic turnover and the rate of enzyme opening are of the same magnitude (~ 330 s^1), which throws into question the accepted notion that the chemical step is rate limiting. Kinetic data of AK variants, where single-point mutations of the catalytic residues to alanine showed that catalytic activity is quenched upon mutation, did not differentiate if the mutations impacted the chemical step or enzyme dynamics. Using computational methods, the impact on enzyme dynamics is studied with MD simulations and the impact on the energy barrier for the phosphoryl transfer is calculated with hybrid quantum mechanical / molecular mechanical methods (QM/MM). The results show that while QM/MM calculations produced a trend of energy barrier values among the mutant case which linearly correlates to that from the experimental data, AK variants transition towards open conformation quicker than the wild type and suggest that disruptions to the interaction network formed by the catalytic residues trigger untimely opening event. It has been revealed that while the catalytic residues individually control

the progression of the chemical step, their organization into a network maintained by a mesh of hydrogen bonds is sensitive to reaction status and responsible for product release.

Introduction

Computer-based experiments are *in silico* experimental environments where the instruments and chemical compounds are replaced by pieces of software and atomic coordinates in the Cartesian system, and natural processes and phenomena under investigation are recreated by mathematical representations of the laws of nature. Just as in a real-world experiment, the choice of a question will dictate the choice of a method, which one will employ in search for the answers. If one is interested in the dynamic behavior of a biochemical system, *e.g.* a protein, nucleic acid, or even a membrane mosaic, one may take a molecular mechanics approach to model one's system. To investigate chemical reactions or phenomena which involve electronic rearrangements, it is necessary to turn to quantum mechanical calculations. The reasons for this are the level at which those two approaches can accurately represent reality. While approximations are inherent in quantum mechanical modelling, *e.g.* evaluation of certain integrals is replaced with empirically-derived parameters (semi-empirical methods) or electrons are treated with functionals describing them as spatially dependent density of electron gas (density-functional theory), molecular mechanics yanks the atoms out of their natural environments and casts into the realm of the classical mechanics where they exist as hard spheres connected by springs under the thumb of Sir Isaak Newton. As such, they move in accordance with Newton's second law of motion:

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} = m_i \frac{d^2 r_i}{dt^2} \qquad (1)$$

where *F* is force, *m* is particle's mass, and *a* acceleration, of which the first and second derivates are the velocity and particle's position, respectively. Potential energy of the system is expressed as an additive formulation of bonded and non-bonded terms[1]:

$$
\begin{aligned}
U = & \sum_{bonds} K_b(b-b_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \sum_{\text{Urey-Bradley}} K_{\mathrm{UB}}(\mathrm{S}-\mathrm{S}_0)^2 \\
& + \sum_{dihedrals} K_\chi(1+\cos(n\chi-\delta)) + \sum_{impropers} K_{imp}(\varphi-\varphi_0)^2 \\
& + \sum_{non-bonded} \left(\varepsilon_{\mathrm{ij}}^{\mathrm{min}} \left[\left(\frac{\mathrm{R}_{\mathrm{ij}}^{\mathrm{min}}}{r_{ij}} \right)^{12} - 2 \left(\frac{\mathrm{R}_{\mathrm{ij}}^{\mathrm{min}}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon r_{ij}} \right) \\
& + \sum_{\text{residues}} \mathrm{U}_{\mathrm{CMAP}}(\varphi, \psi)
\end{aligned}
\tag{2}
$$

, where the harmonic terms describe the bond, angle, dihedral, and improper angle terms, and the non-bonded interactions are the summations over the Lennard-Jones and Coulomb potentials, for dispersion and electrostatic interactions, respectively.

Realistic conditions of temperature and pressure for the function of biological systems are maintained in the simulation by algorithms with thermostat and barostat functionalities, *i.e.* the simulations are carried out in the isothermal-isobaric ensemble. One of the most used thermostats is that proposed by Shuichi Nosé and further developed by William Hoover, the Nosé-Hoover thermostat[2-4]. The idea is to imagine a “heat bath” at our target temperature, to which our system is coupled, and can refer to, to ensure that its temperature is consistent with the reference bath. Both, our simulation and the heat bath are part of the same system, *i.e.* form an extended system. Since phenomenological temperature of a macroscopic system reflects its kinetic energy, the coupling to the heat bath can be mediated by an artificial variable / particle (*i.e.* an external degree of freedom) with associated mass and velocity. The magnitude of its mass determines the magnitude of the coupling and the degree of fluctuations in the temperature.

The most approachable way of visualizing pressure control would be to see it as rescaling of atomic positions within the system following the adjustment of the size of the unit cell. The popular choice of the method is the Langevin piston method[5], philosophically based on the idea of an extended system, where the introduced degree of freedom would correspond to the volume of the cubic cell in the simulation, which adjusts itself such that the pressure between our simulation and the reference value equalizes. In the piston method, the volume degree of freedom is replaced by a piston with a mass value. The magnitude of the mass will influence a time interval in which volume fluctuations are to decay. Andersen[6], who originally developed the extended system method, ingeniously proposed the choice of mass for the piston based on the length of time that would be required for a sound wave to travel through our simulation cell. The Langevin piston method is a further variation on the theme in that here, the piston degree of freedom is coupled to a heat bath and incorporated into the Langevin equations of motion.

This work is a result of the application of computational methods and algorithms in order to understand the behavior of adenylate kinase enzymes on the molecular level. Adenylate kinases are evolutionarily ancient enzymes present in all domains of life[7] and have been divided into three classes: (i) monomeric with a long nucleosidetriphosphate(NTP)-binding lid[8], which are most associated with or derived from prokaryotic organisms; (ii) monomeric with a short NTP-binding lid[7], occurring within eukaryotes; (iii) and trimeric isoforms with a short NTP-binding lid[9] found in archaeal extremophiles. They function within a crowded cellular environment and reversibly shuttle the terminal phosphoryl group from the ATP donor onto an AMP substrate to yield two ADP products. There also exists an isoform, AK3, which selectively uptake GTP as the phosphoryl donor – the mechanistic understanding of which forms the discourse presented in chapter 1. AK3 is associated with the mitochondrial matrix and is belongs to the first group (i). Through the application of molecular dynamics simulations my collaborators and I were able to show that the backbone recognition via the nucleobase carbonyl oxygen and amide proton of Thr201, and subsequent stabilization via a stacking of the guanosine moiety were responsible for the preferential uptake and detention of GTP within AK3's active site.

Chapters 2 and 3 turn their focus to another isoform of adenylate kinase found in *Escherichia coli* (*E. coli*), AdK, also a member of group (i), but associated with the cytosolic environment. Kinetic experiments established that the catalytic turnover (k_{cat}) of 263 ± 30 s^{-1} and the rate of opening of the ATP-binding lid (k_{open}) 286 ± 85 s^{-1} were coincidental[11]. However, the rate for the lid closing (k_{close}) was measured to be 1,374 ± 110 s^{-1} (11). This inspired the probing of a potential connection between the catalytic step and the rate at which the enzyme undergoes a conformational shift to the open state and revisit the established idea of the enzyme opening being the rate limiting step. A series of molecular dynamics simulations were performed to establish a causal connection between the motions of the catalytic residues and the overall protein dynamics. To this end, our experimental collaborators measured the k_{cat} of six variants whose catalytic residues have been replaced with a different side chain. The goal was to create, set up and simulate those same variants and analyze their dynamic behavior, quantifying large domain motions and qualitatively relate them to the orientation of the preserved residues in the active site. The simulations showed that those same residues whose individual role was to facilitate the chemical step by coordinating to a specific part of the substrates, also organized themselves into a unit whose internal network of interactions controls lid motions.

My collaborators on the computational part, Drs Kwangho Nam and Pedro Pjeda-May, are credited with the QM/MM calculations of the energetic barriers for the phosphoryl transfer in the wild-type AdK and its mutant variants. The names and affiliations of all persons who contributed to the published (Chapters 1 and 2) and unpublished (Chapter 3) work are listed at the beginning of each chapter.

References

[1] Mackerell, A. D., Jr. (2004). Empirical force fields for biological macromolecules: overview and issues. *Journal of Computational Chemistry*, *25*(13), 1584–1604.

[2] Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, *81*(1), 511–519.

[3] Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A: General Physics*, *31*(3), 1695–1697.

[4] Frenkel, D., & Smit, B. (2001). *Understanding molecular simulation: From algorithms to applications* (2nd ed.). Academic Press.

[5] Feller, S. E., Zhang, Y., Pastor, R. W., & Brooks, B. R. (1995). Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics*, *103*(11), 4613–4621

[6] Andersen, H. C. (1980). Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, *72*(4), 2384–2393.

[7] Moon, S., Kim, J., & Bae, E. (2017). Structural analyses of adenylate kinases from Antarctic and tropical fishes for understanding cold adaptation of enzymes. *Scientific Reports*, *7*(1), 1–12

[8] Berry, M. B.; Meador, B.; Bilderback, T.; Liang, P.; Glaser, M.; Phillips, G. N., (1994) The closed conformation of a highly flexible protein - The structure of Escherichia-coli adenylate kinase with bound AMP and AMPPNP. Proteins 19, 183-198.

[9] Vonrhein, C.; Bonisch, H.; Shafer, G.; Schulz, G. E., (1998) The structure of a trimeric archaeal adenylate kinase. J. Mol. Biol. 282, 167-179.

[10] McGuffee, S. R.; Elcock, A. H., (2010) Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. PLoS Comput. Biol. 6, 18.

[11] Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., & Kern, D. (2004). Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nature Structural & Molecular Biology*, *11*(10), 945–949.

Structural basis for GTP versus ATP selectivity in the NMP kinase AK3

Per Rogne[1], Beata Dulko-Smith[2], Jack Goodman[1], Marie Rosselin[1], Christin Grundström[1], Christian Hedberg[1], Kwangho Nam[2], A. Elisabeth Sauer-Eriksson[1], and Magnus Wolf-Watz[1]

[1]Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden, [2]Department of Chemistry and Biochemistry, University of Texas at Alington, Arlington, TX 76019-0065, USA

Biochemistry, **2020**, *59*(38), 3570-3581.

Abstract

ATP and GTP are exceptionally important molecules in biology with multiple, and often discrete, functions. Therefore, enzymes that bind to either of them must develop robust mechanisms to selectively utilize one or the other. Here, this specific problem is addressed by molecular studies of the human NMP kinase AK3 which uses GTP to phosphorylate AMP. AK3 plays an important role in the citric acid cycle where it is responsible for GTP/GDP recycling. By combining a structural biology approach with functional experiments, we present a comprehensive structural and mechanistic understanding of the enzyme. We discovered that AK3 functions by recruitment of GTP to the active site, while ATP is rejected and non-productively bound to the AMP binding site. Consequently, ATP acts as an inhibitor with respect to GTP and AMP. The overall features with specific recognition of the correct substrate and non-productive binding by the incorrect substrate bears strong similarity to previous findings for the ATP specific NMP kinase adenylate kinase. Taken together we are now able to provide the fundamental principles for GTP and ATP selectivity in the large NMP kinase family. As a side-result originating from non-linearity of chemical shifts in GTP and ATP titrations, we find that protein surfaces offer a general and weak binding affinity for both GTP and ATP. These non-specific interactions likely act to lower the available intra-cellular GTP and ATP concentrations and may have driven evolution to adapt the Michaelis constants of NMP kinases accordingly.

Introduction

Enzymes perform the formidable task of recognizing their endogenous substrates within the complex[1] and crowded[2] cellular milieu. The complexity rises even further when substrates with similar chemical structures needing to be accepted or rejected at the active site are considered. To overcome these difficulties, enzymes can employ positive and negative selection mechanisms to recruit and reject substrates respectively, both of which require evolution of distinct molecular mechanisms[3]. A particular example is the ability of enzymes to discriminate between adenosine triphosphate (ATP) and guanosine triphosphate (GTP) (Figure 1). ATP is utilized as a general cellular energy carrier[4] and a phosphoryl donor in phosphorylation[5]. For this broad role of ATP, various enzymes bind ATP for their functions in the cell. GTP, on the other hand, is found to have a narrower range of cellular functions in, for instance, protein synthesis[6] and signaling[7]. Therefore, the selectivity of GTP from ATP is necessary for functional aspects and also to maintain the cellular GTP pool.

Given the importance of ATP and GTP for cellular function we undertook a systematic investigation to decode molecular principles of enzymatic ATP and GTP recognition. To this end, we have focused on members of the nucleotide monophosphate (NMP) kinase family[8] as model systems. Members of this family are selective either for ATP[9] or GTP[10] and use these substrates as phosphoryl donors. Based on their structures, NMP kinases are divided into three major classes: (1) monomeric and long[9], (2) monomeric and short[11] and (3) trimeric and short[12]. Long and short refer to the presence or absence of an INSERT segment in the nucleotide triphosphate binding domain, respectively (Figure S1). Specifically, we have studied two monomeric and long NMP kinases: adenylate kinase isolated from *E. coli*, $\mathrm{AK_{eco}}$, and human adenylate kinase isoform 3, AK3. The two enzymes phosphorylate adenosine monophosphate (AMP) by making use of GTP (in AK3) or ATP (in $\mathrm{AK_{eco}}$), respectively.

We have previously discovered principles for; (1) the molecular mechanisms of $\mathrm{AK_{eco}}$ that governs selective recognition of ATP[13], (2) nucleation of an induced fit transition by ATP[14] and (3) negative selection of GTP[13]. In short, the key interaction for selective recognition of ATP is a hydrogen bond formed between N_6 of the adenine base and the backbone carbonyl oxygen of

Lys200[13]. This backbone mediated ATP recognition is conserved not only among NMP kinases which use ATP but also in eukaryotic protein kinases[13, 15]. Activation of AK_{eco} by a large-scale conformational change of the ATP binding domain (ATPlid) is nucleated by formation of a cation-π (or stacking) interaction between the sidechain of Arg119 and the aromatic system of the adenine base[14]. This cationic "sensing" of the aromatic system of the substrate is conserved for NMP kinases[14]. Finally, the 60-fold difference of the catalytic activity of AK_{eco} between ATP and GTP is attributed to a mechanism where GTP binds to and arrests the enzyme in an open and catalytically incompetent state[13]. Here, we focus on the molecular understanding of GTP recognition and GTP versus ATP selectivity of AK3. This enzyme is expressed in the human mitochondrial matrix[10] and its role is to shuttle GTP into GDP as GDP is used by succinyl-CoA synthetase in the citric acid (TCA) cycle[10]. Through an integrative structural biology approach combining X-ray crystallography, NMR spectroscopy and molecular dynamics simulations, we reveal the molecular mechanisms that underlie the GTP selectivity of AK3. In addition, and by examining observations off non-linearity of chemical shifts in GTP and ATP titrations, we find that protein surfaces offer a general and weak affinity for both GTP and ATP.

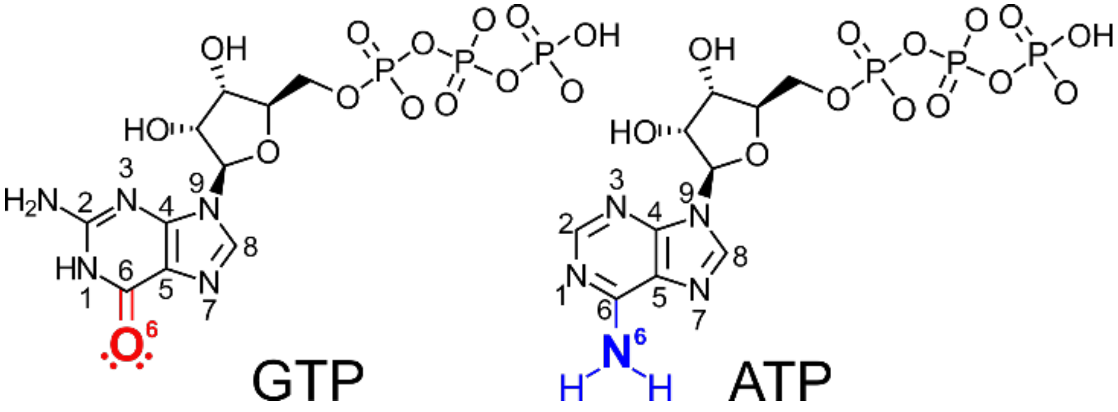


Figure 1. Chemical structures of GTP and ATP
The nomenclature of the nucleobases are indicated together with the substituents at the C6 position that are key for selective binding of GTP and ATP. In GTP the free electron pairs at the O_6 oxygen serves as hydrogen bond acceptors, while the hydrogens on the N_6 atom of ATP serves as hydrogen bond donors.

Materials and methods

Expression and purification of human AK3

The gene for AK3 (UniProt Accession ID: Q9UIJ7) was cloned into a pET24 plasmid and the protein was expressed in BL21 (DE3) *E. coli* cells. Two labeling schemes for NMR experiments were used: ^{2}H, ^{13}C and ^{15}N triple labeling for samples used in sequential assignment and ^{15}N labeling for chemical shift perturbation and binding experiments. Both types of labeling were performed by growing the expression cells in the M9 minimal media where suitable labeled nutrients were added. The triple labeling media was based on 100% D_2O and used ^{13}C glucose and $^{15}NH_4Cl$ as the sole sources for carbon and nitrogen respectively. For ^{15}N-labeling, the medium was based on H_2O with $^{15}NH_4Cl$ used as the sole source of nitrogen. The expressed protein was purified using a two stage purification using first a Blue Sepharose (GE healthcare) column and in the second step a HiPrep 26/60 Sephacryl S-100 HR (GE healthcare) gel filtration column, the same procedure as for *E. coli* adenylate kinase published earlier[16]

Activity of AK3

The activity of AK3 was determined using a coupled enzymatic assay where the production of ADP and GDP was coupled to the oxidation of NADH to NAD^+ using pyruvate kinase and lactate dehydrogenase. The oxidation of NADH was monitored by absorption spectroscopy at 340 nm17. Reaction rates, V (unit M s^{-1}) were determined for substrate concentrations ranging between 0.6 and 220 µM. In all cases, 300 µM of the complementary substrate was used. 2 mM of $MgCl_2$ was added to provide the essential Mg^{2+} ions to the enzyme. The reaction rates were normalized by the total enzyme concentration, $[E]_0$ prior to analysis and K_M and k_{cat} where obtained by fitting of normalized reaction rates, $\frac{V}{[E]_0}$, to the Michaelis-Menten equation (equation 1).

$$\frac{V}{[E]_o} = \frac{k_{cat}[S]}{K_M + [S]} \qquad \text{(eq 1)}$$

In inhibition studies, 1.7 mM of ATP was added to the reaction mixture for the enzymatic assay, and the data were analyzed by non-linear fits of the data to different kinetic models. The modified Michaelis-Menten equation for the case of mixed inhibition is:

$$\frac{V}{[\mathrm{E}]_o} = \frac{k_{cat}^{obs}[S]}{K_M^{obs}+[S]} \qquad \text{(eq 2)}$$

where;

$$k_{cat}^{obs} = \frac{k_{cat}}{(1+\frac{[I]}{K_I})}$$

and

$$K_M^{obs} = K_M(\frac{K_I K_I' + K_I'[I] + K_I[I]}{K_I K_I' + K_I'[I]})$$

K_I and K_I' corresponds to inhibition constants for non-competitive and competitive inhibition, respectively, and [I] corresponds to the concentration of the inhibitor ATP.

ITC

Isothermal calorimetry titrations (ITC) were conducted using a MicroCal Auto-iTC200 at 25 $^\mathrm{o}$C. A solution of either 1.6 mM of ATP or 1.5 mM of GTP, was titrated into a solution of 140 µM AK3 in 20 steps up to 2:1 molar ratio of ligand to protein. Data was analyzed to determine dissociation constant (K_d), binding enthalpy (ΔH^0), the binding entropy ($-\mathrm{T}\Delta S^0$), and binding ratio (n) using the analysis program provided by the ITC manufacturer. All experiments were conducted in 30 mM MOPS buffer at pH 7.0 supplemented with 50 mM NaCl. Three technical replicates were run to determine the accuracy of the measurements.

NMR

All NMR experiments were recorded in a 30 mM MOPS buffer with 50 mM of NaCl at pH 7.0 with 7% D_2O used in all samples to provide the lock signal. The temperature was 298 K or 25 $^\mathrm{o}$C for all experiments. Experiments for backbone assignments were recorded using 180 µl 700 µM ^{2}H, ^{13}C, ^{15}N labeled AK3 in a 3 mm NMR tube. The experiments were acquired using targeted acquisition[18] at a Bruker Avance III spectrometer operating at 800 MHz with a 3 mm cryoprobe at the Swedish NMR center, Gothenburg. The acquired data was processed using compressed sensing IRLS algorithm in the mddnmr software[19, 20]. AK3 sequential assignments were obtained by combining the targeted acquisition approach[19, 21, 22] and automated assignment using the FLYA module of CYANA[23]. The assignments were manually checked and extended using CARA[24]. The assignments were deposited in Biological Magnetic Resonance Data Bank (BMRB ID: 50369).

To determine binding properties and chemical shift perturbations ^{15}N-HSQC spectra were recorded using samples of 200 µM ^{15}N-labeled AK3. Ligands were titrated into the sample over nine steps across a range of 50 µM to 20 mM. The spectra were recorded using either an 850 MHz Bruker Avance III spectrometer equipped with a 5mm HCN cryoprobe or a 600 MHz Bruker Avance III equipped with a BBO broadband cryoprobe. The chemical shift perturbations induced by each ligand were determined by comparing spectra with saturating concentration to spectra recorded in the absence of ligand (apo-state).

Binding data were analyzed by fitting of either a one-site model (equation 3) or a two independent sites model (equation 4) to the chemical shift induced by ligand binding as a function of the ligand concentration.

$$\Delta\omega([L]) = \Delta\omega_{sat} * \frac{(K_a*([E]+[L])+1)}{2*K_a*[E]} - \sqrt{\left(\frac{(K_a*([E]+[L])+1)}{2*K_a*[E]}\right)^2 - \frac{[L]}{[E]}} \qquad \text{(eq. 3)}$$

$$\Delta\omega([L]) = \Delta\omega_{sat1} * \frac{(K_{a1}*([E]+[L])+1)}{2*K_{a1}*[E]} - \sqrt{\left(\frac{(K_{a1}*([E]+[L])+1)}{2*K_{a1}*[E]}\right)^2 - \frac{[L]}{[E]}} + \Delta\omega_{sat2} *$$
$$\frac{(K_{a2}*([E]+[L])+1)}{2*K_{a2}*[E]} - \sqrt{\left(\frac{(K_{a2}*([E]+[L])+1)}{2*K_{a2}*[E]}\right)^2 - \frac{[L]}{[E]}} \qquad \text{(eq 4)}$$

Where Δω([L]) is the chemical shift change as a function of ligand concentration in PPM, $\Delta\omega_{sat}$ is the chemical shift change at saturating ligand concentration in PPM, [E] is the enzyme concentration in M, [L] is the ligand concentration in M, and K_a is the association constant (K_d^{-1}) in M^{-1}. In the two site model the 1 and 2 in K_a and $\Delta\omega_{sat}$ indicate the different sites.

The ^{1}H and ^{15}N chemical shifts of each amino acid residue with its peak assigned in the ^{15}N-HSQC spectra were fitted against one or two, depending on the model, global K_as using a MatLab script written by the authors.

X-ray crystallography

Purified AK3 was dialyzed against 50 mM NaCl and 30 mM 3-(N-morpholino)propanesulfonic acid (MOPS) buffer (pH 7.0) and concentrated to 15-18 mg·ml^{-1} using an Amicon ultracentrifugal filter device (Millipore; 3 kDa molecular weight cutoff) and co-crystallized at 18°C with a threefold excess of: P1 -(5’-guanosyl)-P5 - (5’-adenosyl) pentaphosphate (Gp5A), P1,P5-di(adenosine-5′)pentaphosphate (Ap5A), or ATP, using the vapor-diffusion hanging drop method. In general, crystallization drops of all three AK3-complexes contained between 0.1-1 µL of protein solution mixed with 0.1-1 µL of precipitant and equilibrated against 1 mL of precipitant solution containing for AK3-Gp5A: 20% isopropanol, 0.1 M NaCitrate pH 5.6 and 20% PEG 4000; for AK3-Ap5A: 0.2 M $CaCl_2$, 0.1 M Tris-HCl pH 8.0 and 20% PEG 6000 (AK3-Ap5A); and for AK3-ATP: 0.2 M $MgCl_2$, 0.1 M Tris-HCl pH 5.8, and 20% PEG 6000. Crystals of hAK3-Ap5A and hAK3-ATP were cryo-protected in 30% PEG 6000 before vitrification in a N_2 gas stream maintained at 100 K using a Cryostream Cooler (Oxford Cryosystems). Crystals of AK3-Gp5A were taken directly from the growing drop and vitrified. High-resolution synchrotron diffraction data at 100 K were collected at the ESRF laboratory (Grenoble, France), and using a Bruker Microstar-H rotating anode equipped with a Platinum 135 CCD Detector. Data were processed and scaled using XDS[25], Pointless, and Aimless from the CCP4 suite[26]. Data- collection statistics are listed in Table 1. The phases for structure determination were obtained by molecular replacement using the program PHASER from the PHENIX suite[27] and the AK3 structure (pdb code 1zd8, unpublished) as the search model. Noteworthy, the 1zd8 structure includes an ADP molecule that is not modelled in the coordinates, but clearly visible in the electron density. Analysis of the AK3-Gp5A diffraction data showed that molecules packed in the crystal were related with translational non-crystallographic symmetry (TNCS). PHASER automatically consider TNCS when detected in the diffraction data. However, in this case only 4 of the 6 protein molecules in the asymmetric unit were related with TNCS. Therefore, the structure was solved in two steps: First, the 4 structures related by TNCS were found, and secondly the remaining 2 protein molecules were found by i) including and fix the correctly placed 4 structures, and ii) inactivate the option to account for translational NCS in the PHASER program. All AK3-complex structures were built and refined using the programs COOT[28] and PHENIX REFINE. The atomic coordinates and structure factors (pdb codes 6ZJB (hAK3-Gp5A), 6ZJD (hAK3-APT), and 6ZJE (hAK3-Ap5A)) were deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (Hypertext).

Table 1. Data collection and refinement statistics for three hAK3 complexes.

Data-collection statistics	**hAK3-Gp5A**	**hAK3-ATP**	**hAK3-Ap5A**
Synchrotron /In-house	ESRF-ID29	In-house	ESRF-ID23-2
Wavelength (Å)	0.976	1.5418	0.873
Space group	P1	$P2_1$	$P2_1$
Unit-cell parameters (Å,°)	45.47, 66.76, 105.91 93.79, 93. 96, 89.71	42.52, 62.52, 49.80 90.0, 108.21, 90.0	42.72, 62,64, 50.15 90, 108.01, 90
Resolution limits[a] (Å)	45.35-1.88 (1.95-1.88)	26.86-1.75 (1.81-1.75)	42.72-1.48 (1.53-1.48)
No. of unique reflections	103037 (9068)	45601 (2335)	39892 (3174)
Multiplicity	1.9 (1.8)	7.0 (1.9)	6.0 (4.3)
Completeness (%)	93.1 (80.9)	96.2 (79.9)	95.1 (76.1)
R_{merge}	0.064 (0.644)	0.058 (0.420)	0.050 (0.852)
R_{PIM}	0.093 (0.600)	0.023 (0.203)	0.033 (0.707)
<I/σ (I)>	7.3 (1.1)	22.3 (4.4)	14.9 (1.3)
CC1/2	0.995 (0.632)	0.999 (0.887)	0.999 (0.521)
R factor (%)	0.200 (0.361)	0.169 (0.276)	0.161 (0.215)
R free (%)	0.255 (0.433)	0.222 (0.326)	0.198 (0.283)
No. of protein atoms	10429	1772	1867
No. of water molecules	773	432	308
No. of ligand atoms	354	69	61
Bond length (Å)	0.014	0.010	0.009
Bond angles (°)	1.59	1.24	1.05
Clash score	5.27	2.17	3.3
Residues in most favored regions (%)	99.1	100.0	100.0
Residues in disallowed regions (%)	0.0	0	0.0
Average B-factor protein (Å^2)	38.4	17.4	27.4
Average B-factor ligands (Å^2)	29.7	24.4	60.5
Average B-factor solvent (Å^2)	36.2	29.5	42.0
PDB code	6ZJB	6ZJD	6ZJE

[a] Values in parenthesis are for the high-resolution shell. Resolution limits were determined by applying a cut-off based on the mean intensity correlation coefficient of half-datasets, CC1/2.

Molecular dynamics (MD) simulations

Atomic coordinates for the Gp5A-bound AK3 were based on the X-ray structure determined in the present study (PDB code: 6ZJB), and the model for the Ap5A-bound Adk was based on the PDB X-ray structure (PDB code: 1AKE)[29]. Protonation states of the residues with titratable groups were determined from their pKa values in water while the protonation of the histidine imidazole side chain was assigned to maintain the optimal hydrogen bond network within the protein. The ATP, AMP, and GTP ligands were re-constructed from the inhibitor coordinates by the removal of one phosphate group. The systems were solvated in a 69 Å (AK3) and 72 Å (AK_{eco}) cubic box with explicit water model TIP3P[30], followed by removal of any water molecule that overlapped with the protein atoms. 0.15 M of NaCl buffer was added to neutralize each protein, and the salt ions were distributed throughout the solvent phase randomly. The systems' total number of atoms were 30,372 (AK3) and 36,245 (AK_{eco}), respectively. All atoms were described with the CHARMM27 [31, 32] force field parameters, and the CMAP[33, 34] correction was applied to the backbone dihedrals to maintain structural stability of the protein. MD simulations were performed for 500 ns using the leap-frog Verlet scheme with a 2 fs integration timestep following application of the SHAKE[35] algorithm to constrain all bonds to hydrogen atoms. Throughout the simulation, the temperature and pressure of the systems were maintained at 300 K and 1 bar with the Langevin thermostat and piston methods[36]. Electrostatic interactions were handled by the particle mesh Ewald summation (PME) method[37] with the real-space cutoff distance of 12 Å. The force-based switching function was used for the evaluation of the van der Waals interactions between 10 Å and 12 Å. The energy minimization and equilibration were performed using the CHARMM c42a2[38] while the production MD was handled by the OpenMM2[39] with GPU acceleration.

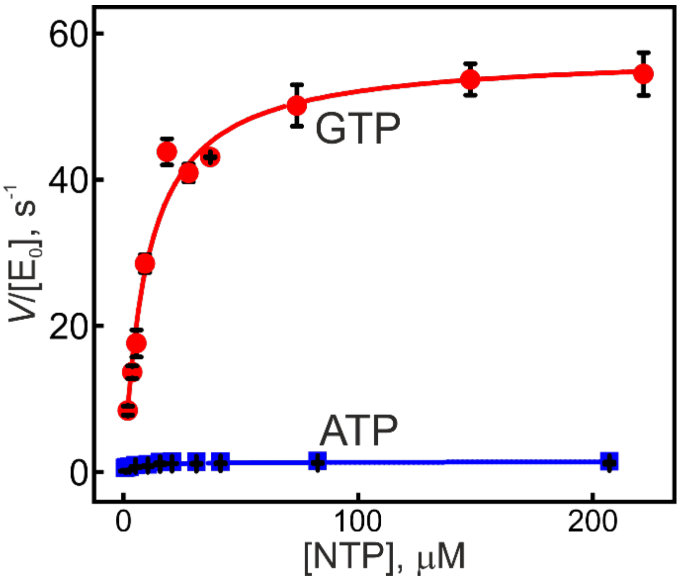


Figure 2. AK3 is selective for GTP over ATP
The phosphoryl donor capacity of GTP (red) and ATP (blue) were quantified at constant AMP concentration by using the coupled ATPase assay. The velocity of the reaction, V, as a function of substrate concentration is scaled with the total enzyme concentration ($[E_0]$). Catalysis follows the Michaelis-Menten kinetics and the best-fitted curves (equation 1) are shown as solid lines, which asymptotically converge towards $k_{\text{cat}}=V_{\text{max}}/[E_0]$. The errors are estimated as the standard deviations from three technical replicates.

Results and discussion

Catalytic GTP versus ATP selectivity of AK3

AK3 displays a distinct preference for GTP over ATP as a substrate in the phosphorylation of AMP (Figure 2), revealed by their quantification with a coupled ATPase assay[17]. Substrate turnover by AK3 follows Michaelis-Menten kinetics, and k_{cat}^{GTP} (57 ± 2 s^{-1}) and K_M (10 ± 1.5 μM) are on the same order of magnitude as previously reported for AK_{eco} [40](k_{cat}^{ATP} of 360 ± 10 s^{-1} and K_M 63 ± 7 μM, respectively). In terms of the specificity constant, k_{cat}/K_M, the values are close to identical for the two enzymes with 5.7 ± 0.9 (s^{-1} μM^{-1}) and 5.7 ± 0.7 (s^{-1} μM^{-1}) for AK3 and AK_{eco}, respectively. Nevertheless, the k_{cat}^{GTP} for AK3 is smaller by around 6-fold than k_{cat}^{ATP} of AK_{eco}. As we show in the following sections, this is related to the dependence of k_{cat} on the free energy of substrate binding. The ratio of the observed k_{cat} values for AK3 defined as, $k_{cat}^{GTP}/k_{cat}^{ATP}$ equals to 41 ± 3 and again, this is similar to the ratio of k_{cat} for ATP versus GTP turnover, $k_{cat}^{ATP}/k_{cat}^{GTP}$ of 59 ± 17 reported for AK_{eco}[13]. For AK_{eco} we have demonstrated that the catalytic ATP versus GTP selectivity is predominantly dependent on the coordination of GTP to the enzyme in an open and catalytically inactive conformation[13]. Here, as we present below, we find that AK3 adopts a related mode of substrate selection mechanism.

Structural basis for GTP recognition by AK3

In order to investigate substrate binding and GTP selectivity in AK3 we solved a 1.9 Å co-crystal structure of AK3 in complex with the bi-substrate mimic P^1-(guanosine-5´)-P^5-(adenosine-5´) pentaphosphate (Gp5A) (Figure 3A and Table 1). Gp5A is essentially a GTP molecule fused with an AMP molecule with an additional phospho-diester bond (Figure 3B). Thus, in contrast to the scenario with the endogenous substrates (GTP and AMP) containing four phosphorus atoms, the Gp5A molecule contains five phosphorus atoms and is therefore not a transition state mimic but rather a tight binding inhibitor. The co-crystal structure establishes that AK3 has the conserved NMP-kinase fold with two nucleotide binding domains (GTPlid and AMPbd) and a central core domain[41]. An overlay of the structure with that of closed AK_{eco} in complex with P1,P5-di(adenosine-5')pentaphosphate (Ap5A) demonstrates that AK3 is in a closed conformation in the Gp5A complex (Figure S2). Overall, the recognition of both the tri- and mono- nucleotide phosphates follows the patterns expected for an NMP kinase. The GTP moiety of Gp5A is sandwiched between residues in the GTPlid and core sub-domains and the AMP moiety is engulfed by residues in the AMPbd (Figure 3A).

The central element in AK3 catalysis is the specific recognition of the guanosine base. Inspection of the structure reveals two key interactions both of which follow the pattern identified previously for ATP recognition by AK_{eco}[13, 14]. The first interaction is a hydrogen bond formed between the backbone amine of Thr201 and the O_6 carbonyl oxygen in GTP (Figure 3C). The corresponding hydrogen bonding interaction in AK_{eco} occurs between the backbone carbonyl oxygen of Lys200 and the N_6 amine group of ATP (Figure 3C). The second interaction is the cation-π (or stacking) interaction between the aromatic ring of the guanosine base and Arg124 in the GTPlid domain (Figure 3A). The corresponding interaction in AK_{eco} occurs between Arg119 in the ATPlid and the adenosine base[14], and we have shown that this interaction nucleates the ATP dependent closure of the ATPlid in AK_{eco}[14]. On basis of the critical role of backbone recognition of GTP by Thr201 in AK3 and ATP by Lys200 in AK_{eco}, we denote the loop where they are located as the "selectivity loop", which connects the C-terminal beta strand with the C-terminal alpha helix (Figure 3C). Taken together, the specific recognition of GTP and ATP has evolved by precise three-dimensional positioning of a hydrogen bond donor (the GTP case) or a hydrogen bond acceptor (the ATP case). Unlike AK_{eco}, the selectivity loop of AK3 forms an additional hydrogen bonding interaction between the sidechain of Thr201 and the N_7 nitrogen of GTP. This dual involvement of backbone and side-chain interactions from one amino acid residue is related to the recognition of GTP by H-Ras where the backbone and sidechain of Ser145, both hydrogen bond to the C_6 oxygen[42]. In addition to these interactions the guanosine base is stacked in between Arg124 and Phe142 and this interaction is discussed below in conjunction with quantification of binding affinities.

A conserved feature of long NMP-kinases is that they exist in open substrate-free states and closed and active substrate bound states, where the closed states can be stabilized for instance with an inhibitor like Ap5A (Figure S3). It has been shown that substrate turn-over by AK_{eco} is rate-limited by the transition from the closed to the open state in complex with bound substrate/product[43, 44]. Chemical shifts are sensitive markers for conformational states of proteins[45-47] and large chemical shift perturbations are observed following saturation of AK3 with either GTP or AMP (Figures 3D, E and Figures S4-S6). To determine the spatial location of the shifted residues we assigned 90% of the backbone resonances of AK3 using a combination of targeted acquisition[18] and both automated[22] and manual assignment procedures. The accuracy of the assignments was validated by a comparison of the secondary structure elements predicted from chemical shifts with those from the co-crystal structure with Gp5A (Figure S4). This is indicative of substantial conformational changes for both nucleotide binding domains following substrate binding, comparable to the open to closed transition observed by AK_{eco}[48]. Thus, although we do not have the atomic structure of substrate free AK3, the NMR observation indicates that AK3 has an open substrate-free structure as expected for an NMP kinase. On the other hand and consistent with the NMR observations, direct experimental observation of an open state of the GTPlid in AK3 is available in crystallographic co-crystal structures of AK3 with both Ap5A and ATP as described in the next paragraph and in the section "Catalysis: Mixed inhibition of AK3 by ATP.

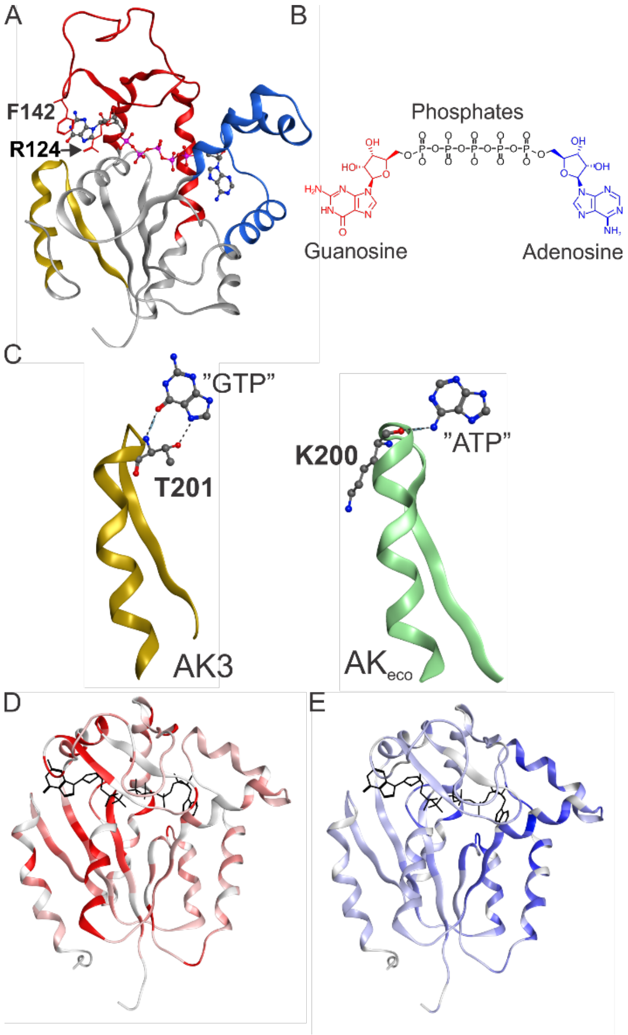


Figure 3. Structural basis for GTP recognition by AK3

(A) 1.75 Å crystallographic structure of closed AK3 in complex with the inhibitor Gp5A. The GTP and AMP binding domains are colored red and blue, respectively, and the selectivity motif that specifically recognizes the O_6 atom of guanosine is colored in gold. Phe140 and Arg124 that stack to the guanosine base are indicated. (**B**). Chemical structure of the inhibitor Gp5A. (**C**) Comparison of the GTP specific selectivity loop in AK3 with the corresponding ATP specific selectivity loop in AK_{eco} (PDB ID: 1AKE). In AK3 the O_6 atom of GTP is recognized by donation of a hydrogen bond from the backbone amide proton of Thr201. Furthermore, N_7 is hydrogen bonded by the sidechain of Thr201. In AK_{eco} the N_6 atom of ATP is recognized by a hydrogen bond formed with the backbone carbonyl oxygen of Lys200. (**D**) Chemical shift perturbations for AK3 in response to GTP binding displayed on the closed AK3 structure. The red color intensity is proportional to the magnitude of the chemical shift perturbation. (**E**) Chemical shift perturbations for AK3 in response to AMP binding displayed on the closed AK3 structure. The blue color intensity is proportional to the magnitude of the chemical shift perturbation. See Figure S5 and S6 primary data.

Even though GTP is the endogenous substrate for phosphoryl transfer by AK3, the enzyme has residual activity also with ATP (Figure 2). In an attempt to understand the structural basis for this activity we determined a 1.5 Å crystal structure with the inhibitor Ap5A which is similar to Gp5A (Figure 4) but is fully symmetric with two adenosine bases. Ap5A has been an exceptionally useful compound in structural biology studies of ATP dependent NMP kinases in that it stabilizes the fully closed conformation$^{29, 48}$. The most prevalent feature of the complex is that although the adenine moiety is correctly coordinated to the AMP binding domain the GTPlid fails to collapse over the adenosine moiety corresponding to ATP (Figure 4). Thus, the Ap5A complex provides the structure of the GTPlid in an open conformation. Inspection of three key residues important for coordination of GTP namely Thr201, Arg124 and Phe142 reveals that the adenosine base of ATP is in a position that is incompatible with recruitment of their respective and specific interactions. Specifically, the distance between the backbone N^H atom of Thr201 and the N_6 atom of the adenosine base is 5 Å which can be compared to the corresponding distance of 2.8 Å in the Gp5A complex. The open orientation of the GTPlid in the AP5A complex does not explain the residual ATP dependent activity of AK3, hence it is possible that the activity is explained by a small fraction of closed enzyme in complex with ATP in solution. On the other hand, it is possible that the AP5A co-crystal structure is resembling an encounter complex where a search for a specific hydrogen bond to the C_6 substituent of the NTP has been interrupted due to the chemical difference in this position between ATP and GTP.

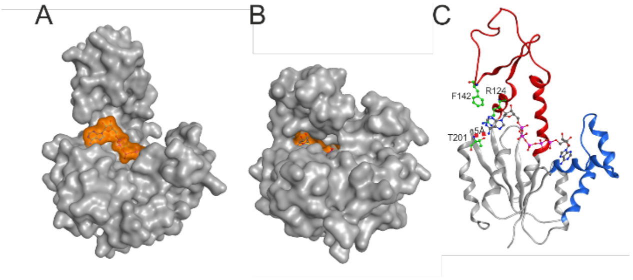


Figure 4. 1.5 Å co-crystal structure of AK3 with Ap5A

(**A**) Surface representation of AK3 in complex with Ap5A. AK3 is colored in gray and Ap5A is colored in transparent orange. (**B**) Surface representation of AK3 in complex with Gp5A. AK3 is colored in gray and Gp5A is colored in transparent orange. (**C**) Ribbon representation of AK3 in complex with Ap5A (ball & sticks). The GTPlid and AMPbd are colored in red and blue respectively. Key residues for the specific interactions with GTP (Arg124, Phe142 and Thr201) are highlighted together with the distance between the backbone N^H atom of Thr201 and the N_6 atom of the adenosine base corresponding to ATP.

Table 2. Thermodynamics for binding of GTP and ATP to AK3

	K_d[a)] (μM)	$\Delta H^0_{\mathrm{ITC}}$ (kcal mol^{-1})	$-\mathrm{T}\Delta S^0_{\mathrm{ITC}}$ (kcal mol^{-1})	$\Delta G^0_{\mathrm{ITC}}$ (kcal mol^{-1})	n
GTP	2.1 ± 0.5	-11.3 ± 0.3	3.5 ± 0.4	-7.8 ± 0.1	1.07 ± 0.025
ATP	13 ± 2	-11.1 ± 0.5	4.4 ± 0.6	-6.7 ± 0.1	0.94± 0.015

[a]Reported values and standard deviations are averages based on three technical replicates.

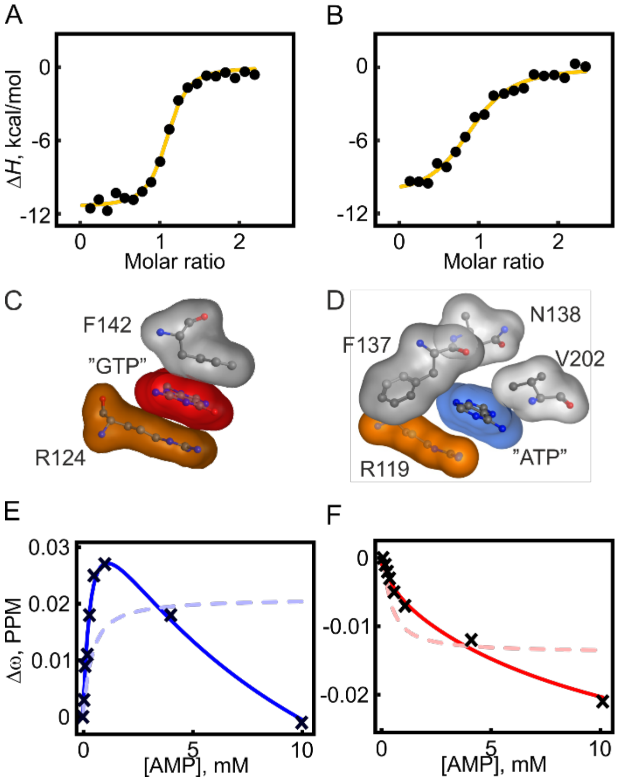


Figure 5. Quantification of GTP, ATP and AMP binding affinities to AK3

Dissociation constants (Kd) were quantified with ITC for GTP and ATP and with NMR spectroscopy for AMP. The choice of technique is dependent on the magnitude of the Kd values, i.e. ITC for stronger binding and NMR for weaker binding. (A) ITC quantification of GTP binding, shown are integrated heats of injections together with the best fitted one-site binding model (yellow line). (B) ITC quantification of ATP binding (details as in A). (C) Structural details of the stacking of the guanosine base (red transparent surface) between Phe142 and Arg124 in AK3. (D) The corresponding interaction network for the adenosine base (blue transparent surface) in AKeco (1AKE 29) consists of Arg119, Phe137, Asn138 and Val202. (E, F) NMR detected binding of AMP at residue Asp94. One-site (equation 2, dashed line) and two-sites (equation 3, solid lines) binding models were fitted to both 15N chemical shifts (E) and 1H chemical shifts (F).

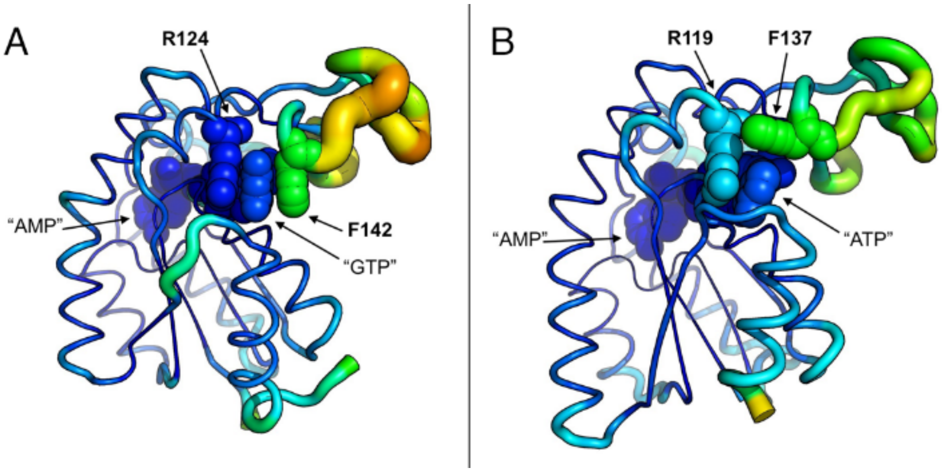


Figure 6. Variation of atomic displacement from the average position in AK3 and AK$_{eco}$
RMSF values for the backbone atoms are used to quantify AK3 and AK$_{eco}$ flexibility over the course of 500 ns MD simulation and projected on the respective average structures. The color scale ranging from blue to orange represents a transition from the most rigid to the most flexible end of the spectrum. (A) The structure of AK3 in complex with the AMP and GTP ligands. The stacking interaction of Arg124 and Phe142 effectively arrests the GTP base in place despite high mobility of the enclosing GTPlid domain. (B) The structure of AK$_{eco}$ in complex with the AMP and ATP ligands. Atomic fluctuations of the ATPlid are weaker than those in AK3, but penetrate deeper into the ATP binding region, and include Arg119 which fails to maintain a stable stacking with the ATP base, thus making it discernibly more flexible than the GTP base.

Binding: GTP and ATP specific and non-specific effects

Enzymatic catalysis following the Michaelis-Menten mechanism can broadly be divided into two consecutive steps corresponding to substrate binding and catalysis. These two fundaments of AK3 catalysis is explored in this and the following section. Equilibrium affinity measurements with isothermal titration calorimetry (ITC) reveal that GTP and ATP bind to AK3 with K_d values of ~2 and ~13 µM, respectively (Figure 5A, B & Table 2). Binding either of the two nucleotide triphosphates is enthalpically driven with a one to one stoichiometry and the overall thermodynamic signatures are very similar to each other. The binding affinity of AK3 for the "endogenous" substrate GTP is relatively high compared to the binding affinity of ATP to AK$_{eco}$ (K_d = 50 µM[48]. The tight binding of GTP to AK3 can be attributed to the ordered stacking formed between the guanosine base, Phe142, and Arg124 (Figure 5C). The equivalent stacking in AK$_{eco}$ is less well ordered and built up from, in addition to Arg119, several amino acid residues (Figure 5D). This is consistent with the molecular dynamics (MD) simulation results showing the tight stacking of GTP between Arg124 and Phe142 during the 500 ns MD simulations, whereas the corresponding MD simulations with ATP for AK_{eco} show larger fluctuation of the ATP base and Arg119 (Figure 6 & S7). In addition, this difference results in a larger overall RMSD than the corresponding value of AK3 (Figure S8).

For AMP binding, due to the relatively low binding affinity, this event was quantified with NMR spectroscopy. A number of residues displayed curved chemical shift perturbation patterns indicating the existence of two binding events with distinguishable binding affinities (Figure 5E, 5F). Fits to these data were improved from a statistical standpoint (F-test) by invoking two independent binding sites and the resulting K_d vales were 210 ± 7 µM and 4.8 ± 0.2 mM, respectively. The stronger binding affinity is close to the K_d for AMP binding to AK_{eco} (210 ± 66 µM[48]).

In light of the weak secondary binding affinity of AK3 to AMP we investigated if similar effects were also present for ATP and GTP. For both ATP and GTP, curved chemical shift perturbation patterns were observed indicating at least two binding events with distinguishable affinities. For the first binding event, their binding affinities (K_d = 2 µM and 13 µM from ITC above) were too strong to be accurately quantified with NMR. On the other hand, the weak biding affinities could be quantified to be 1.7 ± 1 mM and 9.0 ± 3.8 mM for GTP and ATP respectively (Figure 7A, B**)**. Since attractive interactions between any pair of molecules arise from the correlation between the motions of their electrons by London dispersion interactions[49], we assumed that the weak binding events are due to non-specific interactions with the protein surface. To test this hypothesis, we quantified the ATP binding affinity of AK_{eco} with its adenosine binding sites blocked with the tight binding inhibitor AP5A (Müller et. al.[29] and Figure 7C). In this experiment the strong ATP binding event was absent while the weak binding event was preserved (12 ± 2.1 mM). We further tested the hypothesis by quantifying the ATP binding affinity of two proteins void of known ATP binding sites, human α-synuclein[50], and LcrG[51] from *Yersinia pseudotuberculosis*. Both proteins displayed chemical shift perturbation patterns indicative of weak ATP binding events (Figure 7D, E). The resulting affinities were 13 ± 2.4 mM and 12 ± 0.8 mM for α-synuclein and LcrG, respectively. Together, we conclude that the weak binding affinity of ATP and GTP observed by NMR is due to non-specific interactions with protein surfaces.

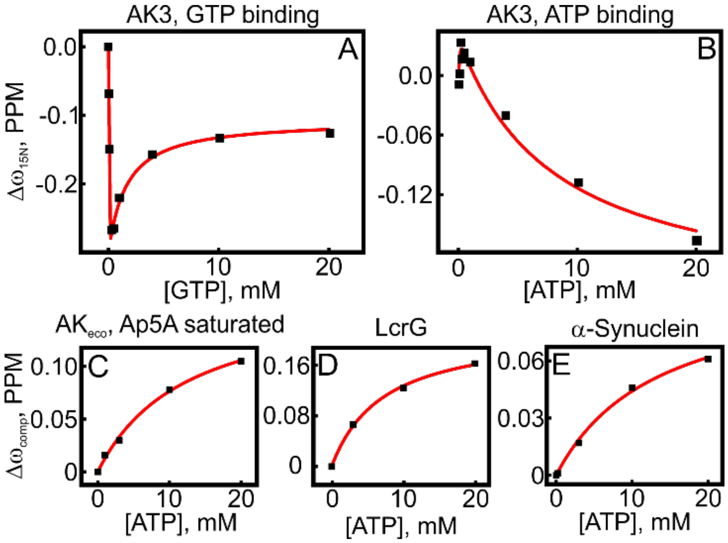


Figure 7. Non-specific binding of GTP and ATP to protein surfaces

Weak binding affinities of GTP to AK3 and ATP to three different proteins, in addition to AK3, were quantified with NMR spectroscopy. In all panels the GTP or ATP induced chemical shift perturbations for one example amino acid residue are shown against the ATP/GTP concentration. (**A**) and (**B**) GTP and ATP binding curves to AK3 fitted to a two-site model (equation 4) are shown by the red line. (**C**) ATP binding to AK_{eco} where the endogenous ATP and AMP binding sites have been occluded by presence of excess of the tight binding inhibitor Ap5A. **(D)** ATP binding to the *Yersinia pseudotuberculosis* protein LcrG. (**E**) ATP binding to the human protein α-synuclein. In C-E the best fitted one-site model (equation 3) is shown as a solid red line.

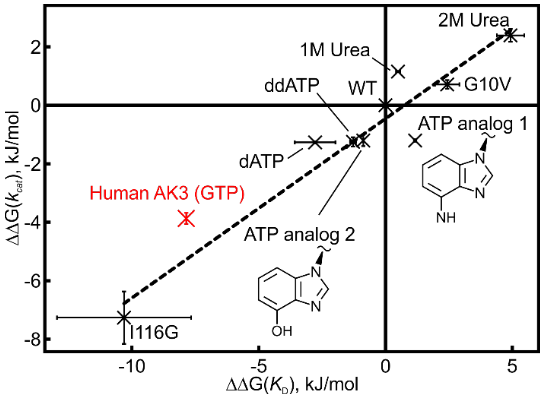


Figure 8. Linear free energy relationship between substrate turnover and binding affinities

A linear free energy relationship between substrate turnover and binding affinity has previously been established for AKeco53. The ΔΔG values are calculated relative to wild-type AKeco. In the original representation of correlation we compared different AKeco variants and also wild-type AKeco performed in different urea concentrations. Here the corresponding values for AK3 were added to the correlation and it is evident that AK3 follows the general energetic relationship as observed for AKeco. In addition, values for AKeco catalysis with modified nucleotide tri-phosphates have been added. These variants include dATP, ddATP and two designed ATP analogs (analogs 1 & 2) where analog 1 has the hydrogen bond donor at C6 replaced from "NH2" to "OH" and analog 2 has the N1 and N3 nitrogens replaced with carbons. Analog 2 was used previously in our study of the inhibited AKeco:GTP complex13. The errors, in both dimensions, are estimated as the standard deviations based on three technical replicates.

The core of our catalytic model for the NMP kinase AK_{eco} is that the turnover of substrates is rate-limited by product release, caused by the slow opening of substrate binding sub-domains[43, 44]. It should also be noted that there exists a contradicting model where ultra-fast opening and closing dynamics are utilized by AK_{eco} to stochastically populate a catalytically competent state[52]. In line with our model, we have previously established a thermodynamic signature for this rate limitation with a linear relation between free energy values of catalysis versus binding[53]. When displayed together with various AK_{eco} variants, different substrates, and AK_{eco} under different conditions, AK3 thermodynamics are consistent with the established linearity (Figure 8). Thus, we argue that AK3 substrate turn-over is rate-limited by product release and that the k_{cat} of around 60 s^{-1} (compare to 360 s^{-1} for AK_{eco}) is a consequence of the strong GTP binding affinity.

Catalysis: Mixed inhibition of AK3 by ATP

Although GTP is the endogenous nucleoside triphosphate used for phosphorylation by AK3, the similarity of the chemical structures between GTP and ATP (Figure 1) prompted us to explore the potential inhibitory power of ATP.

To this end, we employed the standard ATPase activity[17] assay with a constant concentration of ATP (1.7 mM) and varying concentrations of either GTP or AMP, with a subsequent fit of the data to different kinetic models. These models included: (1) no changes to either k_{cat} or K_M, i.e. no inhibitory effect of ATP, (2) competitive inhibition, i.e. an increased observed K_M (3) noncompetitive inhibition, i.e. a decreased observed k_{cat}, and (4) mixed inhibition, i.e. an increased observed K_M together with a decreased observed.

k_{cat}. Fits to the four kinetic models for both variation of GTP and AMP are shown in Figure 9. It is evident that mixed inhibition models (eq.2 in materials and methods) provides the best explanation of the experimental data for both GTP and AMP and that in both cases the second best model was non-competitive inhibition. The statistical relevance for the improved fits in the mixed models were accessed with F-tests. In the case of GTP the F-value for mixed versus noncompetitive inhibition was 9.25 which makes the mixed inhibition model statistically relevant at the 90% confidence level (critical F-value equals 5.46). Likewise, in the AMP case the F-value for mixed versus noncompetitive inhibition was 50 which makes the mixed inhibition model statistically relevant at the 99% confidence level (critical F-value equals 31). For the scenario of mixed inhibition, the inhibition constants are defined as K_I and K_I' for noncompetitive and competitive inhibition, respectively. For GTP variation, K_I and K_I' were found to be 850 and 655 µM, respectively, whereas the corresponding values for AMP variation were 1700 and 200 µM. Among these inhibition constants, the strongest effect was observed for competitive inhibition of AMP with an inhibition constant of 200 µM.

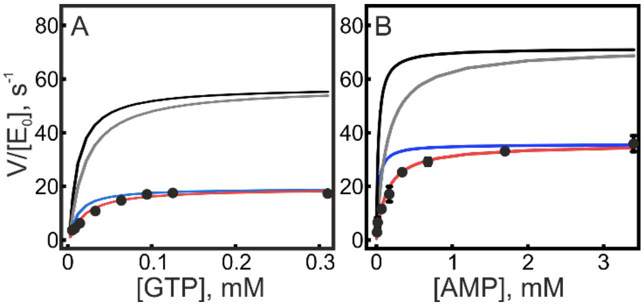


Figure 9. ATP inhibition of AK3

The effect on ATP (at 1.7 mM) on AK3 catalysis was quantified by fitting kinetic data from coupled ATPase assays to different models. Kinetic rates equal to V/E_0 are plotted against (**A**) GTP concentration or (**B**) AMP concentration. The solid lines corresponds to fits to kinetic models as follows; black: no inhibitory effect of ATP, gray: competitive inhibition, blue: noncompetitive inhibition and red: mixed inhibition.

These ATP inhibition patterns implies that the binding of ATP likely interferes with the AMP binding site. To verify this model, we used solution state NMR chemical shift perturbation analysis. Structural changes were monitored by quantifying the ligand-induced changes in chemical shifts in ^{1}H-^{15}N HSQC experiments which provided a unique fingerprint of the structural state occupied by the protein under defined conditions. By comparing spectra of apo and ATP saturated AK3 (Figure S9), it is evident that the chemical shifts of a set of amino acid residues were significantly perturbed as expected for a binding process with K_d of 13 µM.

Quantitative analysis was performed by comparing the AK3 spectra in three defined states saturated respectively with AMP, GTP and ATP. Here the spectra of AK3 in complex with AMP and GTP served as reference states. The ATP saturated spectra showed a striking similarity with the AMP reference spectra while the resemblance of the GTP state was much weaker. This difference was quantified by linear regression of the chemical shifts of the ATP state with those of the reference states (Figure 10A, B**)**. The analysis reveals a robust correlation (R^2=0.93 and a slope of 1.06) between ATP and AMP chemical shifts whereas the ATP and GTP chemical shifts are essentially uncorrelated. Collectively, these NMR results demonstrate that the primary site of ATP binding is the AMP site. This binding mode is consistent with the competition pattern observed for ATP against GTP and AMP (Figure 9) and also with the inability of AK3 to close its GTPlid over Ap5A (Figure 4). In order to determine the molecular details of ATP binding to the AMP site, we solved a 1.75 Å co-crystal structure of AK3 in complex with ATP. As predicted from the NMR analysis the structure has ATP coordinated to the AMP between AK3 and the adenosine group in the ATP inhibited structure is very similar to the pattern observed in the closed structure in complex with Gp5A. Taken together, we have discovered the molecular underpinning of the inhibitory effect of ATP on AK3 catalysis from a combination of enzymatic assays, NMR and x-ray crystallography. The mechanism of ATP inhibition by binding to the AMP site is dependent on the bi-substrate nature of AK3 catalysis and the chemical similarity between ATP and AMP.

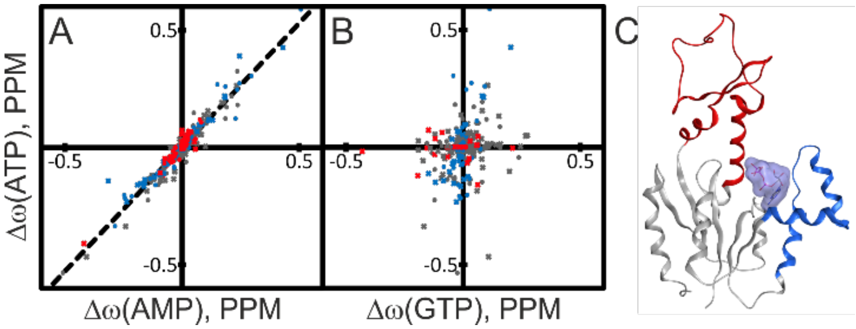
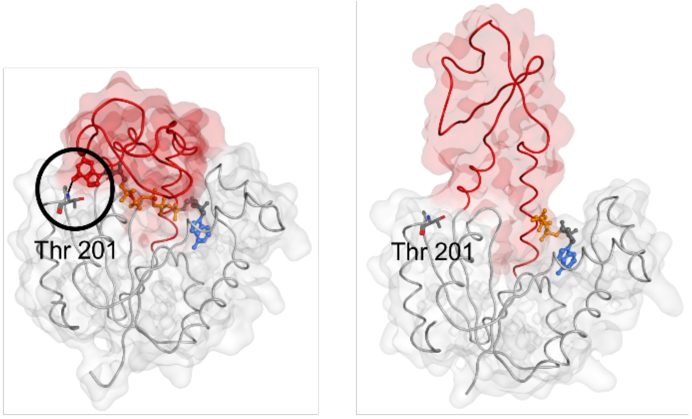
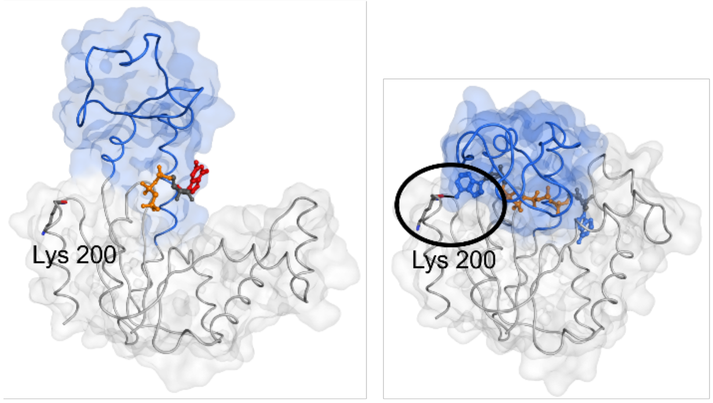


Figure 10. Structural basis for ATP inhibition of AK3
The spatial location of ATP binding to AK3 was identified from comparisons of chemical shift perturbations induced by ATP with those induced by AMP and GTP. (**A**) Correlation between chemical shift perturbations induced by ATP and AMP. The identity of residues in the AMP binding (red), GTP binding (blue) and core sub-domains (gray) are indicated in the plot. The best fitted linear correlation is indicated with the straight line (Δω(ATP)=1.06* Δω(AMP), R^2: 0.93). The strong linear correlation indicates that ATP binds to the AMP binding site of AK3. (**B**) Comparison of ATP and GTP induced chemical shift perturbations with color-coding as in (A). The lack of correlation excludes the possibility that ATP binds to the GTP binding site. (**C**) 1.75 Å co-crystal structure of AK3 in complex with ATP. ATP binds to the AMP biding site as predicted from the NMR results in (A) and (B). The GTP and AMP binding sites are colored in blue and red, respectively.



AK3+GTP*
Catalytically active

AK3+ATP
Catalytically inactive



AK_{eco}+GTP
Catalytically inactive

AK_{eco}+ATP*
Catalytically active

Figure 11. Principles of ATP and GTP selectivity in NMP kinases
The principles are based on the findings for AK3 (GTP specific NMP kinase) presented here and for AK_{eco} (ATP specific NMP kinase) described previously[13]. The function of the enzymes is dependent on positive selection via formation of a catalytically active, closed, state with the "correct" substrate, and negative selection via formation of a catalytically inactive, open, complex with the "incorrect" substrate. **Top:** For AK3 the catalytically active complex (left panel, PDB entry 6ZJB) with the endogenous substrate GTP is, in part, dependent on a hydrogen bond between the backbone amide proton of Thr201 and the O_6 atom of the guanosine base (encircled). A catalytically incompetent state of AK3 in complex with ATP (right panel, PDB entry 6ZJD) is formed by coordination of ATP to the AMP binding domain. **Bottom:** In the case of the ATP specific NMP kinase AK_{eco} a catalytically incompetent state with GTP (left panel) is obtained through arrest of an open conformation mediated through contacts with the guanosine base by residues in the ATP binding domain (PDB entry 6F7U[13]). The catalytically competent state of AK_{eco} with ATP (right panel) is, in part, stabilized with a backbone hydrogen bond between Lys200 and the N_6 hydrogen of the adenosine base (encircled) (PDB entry 1AKE[29]).
*Figures of the active complexes are shown with the inhibitors Gp5A and Ap5A for AK3 and AK_{eco} respectively, showing the closed state of the NTP binding lid.

Conclusions

The production of the universal energy carrier, ATP, from the combined action of the citric acid (TCA) cycle and oxidative phosphorylation is a key event in biology. One of the steps in the citric acid cycle is the conversion of succinyl-CoA to succinate by the action of succinyl-CoA synthase[10]. During the succinyl-CoA conversion, GDP is phosphorylated to GTP, which implies that GTP must be recycled into GDP in the mitochondrial matrix. Here, we provide molecular insights into the function of AK3 that is responsible for GTP to GDP recirculation, and a detailed insight in its selectivity for GTP over the similar and abundant ATP. The key aspect for specific recognition of the guanosine moiety of GTP by AK3 is the formation of a backbone-mediated hydrogen bond between the amide proton of Thr201 and the O_6 carbonyl oxygen of the nucleobase. Backbone-mediated hydrogen bonding for specific recognition of GTP has also been found in other enzymes. For example, in H-Ras, GTP is, in part, coordinated by a hydrogen bond donated from the amide proton of Ser145[42], and in EF-Tu it is hydrogen bonded in a similar manner by residue Leu176 (1EFT.pdb)[54]. The involvement of backbone atoms in the mechanism underlying binding or catalysis is robust from an evolutionary perspective because the backbone is insensitive to mutational perturbation. A well-established example of this is the proteolytic mechanism of serine proteases, which depends on an oxyanion hole formed by backbone amide protons to stabilize the enzyme-acyl intermediate[55].

Our structural data reveal that the almost 40-fold catalytic GTP versus ATP specificity of AK3 is primarily dependent on ATP being coordinated by the endogenous AMP binding site. As such, ATP acts as a non-competitive inhibitor with respect to GTP and this mechanism provides a robust way to ensure that GTP is the primary phosphoryl donor used by AK3 inside the mitochondrial matrix. This mechanism is dependent on the bi-substrate nature of AK3 and bears similarities to how AK_{eco} selects ATP over GTP. In AK_{eco}, GTP binding arrests the enzyme in an open, catalytically incompetent state, shown by the X-ray crystallographic structure of AK_{eco} in complex with a non-hydrolysable GTP analogue determined to 1.4 Å resolution[13]. The findings for AK3 and AK_{eco} summarize how members of the NMP-kinase family have evolved to utilize positive and negative selections for GTP and ATP (Figure 11). During the course of the discovery that ATP binds to the AMP binding site in AK3, we have also found that GTP and ATP have a tendency to bind non-specifically to proteins. This effect complicates the analysis of NMR titrations by causing a curvature in the ligand concentration dependent chemical shift perturbations. From a functional standpoint the dissociation constant (K_d) for non-specific interactions with proteins is in the range of 1.3 – 13 mM which is on par with the intracellular ATP concentrations that are in the milli-molar range[56]. Thus, the intracellular activity of GTP and ATP, is expected to be lowered by non-specific protein interactions, and this concept has been discussed previously in different context where ATP has been put forward as a hydrotrope[57]. In contrast, the K_M values for GTP for AK3 (10 µM) and for ATP for AK_{eco} (63 µM[40]) are unexpectedly low compared to the millimolar intracellular concentrations of ATP and GTP[56]. This apparent discrepancy may have evolved to compensate the limited availability of free ATP as a result of its non-specific interactions with protein surfaces.

References

(1) Wennerström, H.; Vallina Estrada, E.; Danielsson, J.; Oliveberg, M., (2020) Colloidal stability of the living cell. Proc. Natl. Acad. Sci. U. S. A. 117, 10113-10121.

(2) McGuffee, S. R.; Elcock, A. H., (2010) Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. PLoS Comput. Biol. 6, 18.

(3) Wong, K. B.; Clarke, J.; Bond, C. J.; Neira, J. L.; Freund, S. M. V.; Fersht, A. R.; Daggett, V., (2000) Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. J. Mol. Biol. 296, 1257-1282.

(4) Berg, J. M.; Tymoczko, J. L.; Stryer, L., Biochemistry. Fifth ed.; W.H. Freeman and Company: New York, USA, 2002.

(5) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P., Molecular biology of the cell. 4th ed.; Garland Science: New York, NY, USA, 2002.

(6) Grela, P.; Helgstrand, M.; Krokowski, D.; Boguszewska, A.; Svergun, D.; Liljas, A.; Bernado, P.; Grankowski, N.; Akke, M.; Tchorzewski, M., (2007) Structural characterization of the ribosomal P1A-P2B protein dimer by small-angle X-ray scattering and NMR spectroscopy. Biochemistry 46, 1988-1998.

(7) Corrons, J. L. V.; Garcia, E.; Tusell, J. J.; Varughese, K. I.; West, C.; Beutler, E., (2003) Red cell adenylate kinase deficiency: molecular study of 3 new mutations (118G > A, 190G > A, and GAC deletion) associated with hereditary nonspherocytic hemolytic anemia. Blood 102, 353-356.

(8) van Rompay, A. R.; Johansson, M.; Karlsson, A., (2000) Phosphorylation of nucleosides and nucleoside analogs by mammalian nucleoside monophosphate kinases. Pharmacol. Therapeut. 87, 189-198.

(9) Berry, M. B.; Meador, B.; Bilderback, T.; Liang, P.; Glaser, M.; Phillips, G. N., (1994) The closed conformation of a highly flexible protein - The structure of Escherichia-coli adenylate kinase with bound AMP and AMPPNP. Proteins 19, 183-198.

(10) Noma, T.; Fujisawa, K.; Yamashiro, Y.; Shinohara, M.; Nakazawa, A.; Gondo, T.; Ishihara, T.; Yoshinobu, K., (2001) Structure and expression of human mitochondrial adenylate kinase targeted to the mitochondrial matrix. Biochem. J. 358, 225-232.

(11) Moon, S.; Kim, J.; Bae, E., (2017) Structural analyses of adenylate kinases from Antarctic and tropical fishes for understanding cold adaptation of enzymes. Sci. Rep. 7, 12.

(12) Vonrhein, C.; Bonisch, H.; Shafer, G.; Schulz, G. E., (1998) The structure of a trimeric archaeal adenylate kinase. J. Mol. Biol. 282, 167-179.

(13) Rogne, P.; Rosselin, M.; Grundström, C.; Hedberg, C.; Sauer, U. H.; Wolf-Watz, M., (2018) Molecular mechanism of ATP versus GTP selectivity of adenylate kinase. Proc. Natl. Acad. Sci. U. S. A. 115, 3012-3017.

(14) Rogne, P.; Andersson, D.; Grundström, C.; Sauer-Eriksson, E.; Linusson, A.; Wolf-Watz, M., (2019) Nucleation of an Activating Conformational Change by a Cation−π Interaction. Biochemistry 58, 3408-3412.

(15) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S., (2002) The protein kinase complement of the human genome. Science 298, 1912-1934.

(16) Reinstein, J.; Brune, M.; Wittinghofer, A., (1988) Mutations in the nucleotide binding loop of adenylate kinase of Escherichia coli. Biochemistry 27, 4712-4720.

(17) Rhoads, D. G.; Lowenstein, J. M., (1968) Initial velocity and equilibrium kinetics of myokinase. J. Biol. Chem. 243, 3963-3972.

(18) Jaravine, V. A.; Orekhov, V. Y., (2006) Targeted acquisition for real-time NMR spectroscopy. J. Am. Chem. Soc. 128, 13421-13426.

(19) Kazimierczuk, K.; Orekhov, V. Y., (2011) Accelerated NMR Spectroscopy by Using Compressed Sensing. Angew. Chem. 50, 5556-5559.

(20) Mayzel, M.; Kazimierczuk, K.; Orekhov, V. Y., (2014) The causality principle in the reconstruction of sparse NMR spectra. ChemComm. 50, 8947-8950.

(21) Jaravine, V. A.; Zhuravleva, A. V.; Permi, P.; Ibraghimov, I.; Orekhov, V. Y., (2008) Hyperdimensional NMR spectroscopy with nonlinear sampling. J. Am. Chem. Soc. 130, 3927-3936.

(22) Isaksson, L.; Mayzel, M.; Saline, M.; Pedersen, A.; Rosenlow, J.; Brutscher, B.; Karlsson, B. G.; Orekhov, V. Y., (2013) Highly efficient NMR assignment of intrinsically disordered proteins: application to B- and T cell receptor domains. PLoS One 8, 8.

(23) Schmidt, E.; Guntert, P., (2012) A new algorithm for reliable and general NMR resonance assignment. J. Am. Chem. Soc. 134, 12817-12829.

(24) Keller, R. J. L., The computer aided resonance assignment tutorial. Cantina verlag: 2004.

(25) Kabsch, W., (2010) XDS. Acta Crystallogr. D 66, 125-132.

(26) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D 67, 235-242.

(27) Adams, P. D.; Afonine, P. V.; Bunkoczi, 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.; Zwart, P. H., (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D 66, 213-221.

(28) Emsley, P.; Cowtan, K., (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D 60, 2126-2132.

(29) Müller, C. W.; Schulz, G. E., (1992) Structure of the complex between adenylate kinase from Escherichia coli and the inhibitor AP5A refined at 1.9 A resolution - A model for a catalytic transition-State. J. Mol. Biol. 224, 159-177.

(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., (1983) Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926-935.

(31) Foloppe, N.; MacKerell, A. D., (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. J. Comput. Chem. 21, 86-104.

(32) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B 102, 3586-3616.

(33) Mackerell, A. D.; Feig, M.; Brooks, C. L., (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J. Comput. Chem. 25, 1400-1415.

(34) MacKerell, A. D.; Feig, M.; Brooks, C. L., (2004) Improved treatment of the protein backbone in empirical force fields. J. Am. Chem. Soc. 126, 698-699.

(35) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., (1977) Numerical-integration of cartesian equations of motions of a system with constraints - molecular-dynamics of N-alkanes. J. Comput. Phys. 23, 327-341.

(36) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R., (1995) Constant-pressure molecular-dynamics simulation - the langevin piston method. J. Chem. Phys. 103, 4613-4621.

(37) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., (1995) A smooth particle mesh ewald method. J. Chem. Phys. 103, 8577-8593.

(38) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., (2009) CHARMM: The Biomolecular Simulation Program. J. Comput. Chem. 30, 1545-1614.

(39) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y. T.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S., (2017) OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS Computat. Biol. 13.

(40) Åden, J.; Verma, A.; Schug, A.; Wolf-Watz, M., (2012) Modulation of a pre-existing conformational equilibrium tunes adenylate kinase activity. J. Am. Chem. Soc. 134, 16562-16570.

(41) Rundqvist, L.; Åden, J.; Sparrman, T.; Wallgren, M.; Olsson, U.; Wolf-Watz, M., (2009) Noncooperative folding of subdomains in adenylate kinase. Biochemistry 48, 1911-1927.

(42) Pai, E. F.; Krengel, U.; Petsko, G. A.; Goody, R. S.; Kabsch, W.; Wittinghofer, A., (1990) Refined crystal-structure of the triphosphate conformation of H-RAS P21 at 1.35 A resolution - implications for the mechanism of GTP hydrolysis. EMBO J. 9, 2351-2359.

(43) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D., (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. Nat. Struct. Mol. Biol. 11, 945-949.

(44) Kovermann, M.; Åden, J.; Grundström, C.; Sauer-Eriksson, A. E.; Sauer, U. H.; Wolf-Watz, M., (2015) Structural basis for catalytically restrictive dynamics of a high-energy enzyme state. Nat. Commun. 6.

(45) Kovermann, M.; Rogne, P.; Wolf-Watz, M., (2016) Protein dynamics and function from solution state NMR spectroscopy. Q. Rev. Biophys. 49.

(46) Selvaratnam, R.; VanSchouwen, B.; Fogolari, F.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G., (2012) The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. Biophys. J. 102, 630-639.

(47) Masterson, L. R.; Yu, T.; Shi, L.; Wang, Y.; Gustavsson, M.; Mueller, M. M.; Veglia, G., (2011) cAMP-Dependent Protein Kinase A Selects the Excited State of the Membrane Substrate Phospholamban. J. Mol. Biol. 412, 155-164.

(48) Åden, J.; Wolf-Watz, M., (2007) NMR identification of transient complexes critical to adenylate kinase catalysis. J. Am. Chem. Soc. 129, 14003-14012.

(49) Evans, D. F.; Wennerström, H., The colloidal domain: Where physics, chemistry, biology, and technology meet. 2nd ed.; Wileey-VCH: New Jersey, USA, 1999.

(50) Eliezer, D.; Kutluay, E.; Bussell, R.; Browne, G., (2001) Conformational properties of alpha-synuclein in its free and lipid-associated states. J. Mol. Biol. 307, 1061-1073.

(51) Bergman, T.; Håkansson, S.; Forsberg, A.; Norlander, L.; Macellaro, A.; Backman, A.; Bolin, I.; Wolfwatz, H., (1991) Analysis of the V-antigen lcrGVH-yopBD operon of Yersinia pseudotuberculosis - evidence for a regulatory role of LcrH and LcrV. J. Bacteriol. 173, 1607-1616.

(52) Aviram, H. Y.; Pirchi, M.; Mazal, H.; Barak, Y.; Riven, I.; Haran, G., (2018) Direct observation of ultrafast large-scale dynamics of an enzyme under turnover conditions. Proc. Natl. Acad. Sci. U. S. A.115, 3243-3248.

(53) Rogne, P.; Wolf-Watz, M., (2016) Urea-dependent adenylate kinase activation following redistribution of structural states. Biophys. J. 111, 1385-1395.

(54) Fersht, A. R., Structure and mechanism in protein science. W.H.Freeman Co Ltd: New York, USA, 1999.

(54) Kjeldgaard, M.; Nissen, P.; Thirup, S.; Nyborg, J., (1993) The crystal-structureof elongation-factor EF-TU from Thermus aquaticus in the GTP conformation. Structure 1, 35-50.

(55) Hedstrom, L., (2002) Serine protease mechanism and specificity. Chem. Rev. 102, 4501-4523.

(56) Traut, T. W., (1994) Physiological concentrations of purines and pyrimidines. Mol. Cell. Biochem. 140, 1-22.

(57) Patel, A.; Malinovska, L.; Saha, S.; Wang, J.; Alberti, S.; Krishnan, Y.; Hyman, A. A., (2017) ATP as a biological hydrotrope. Science 356, 753-756.

Appendix I

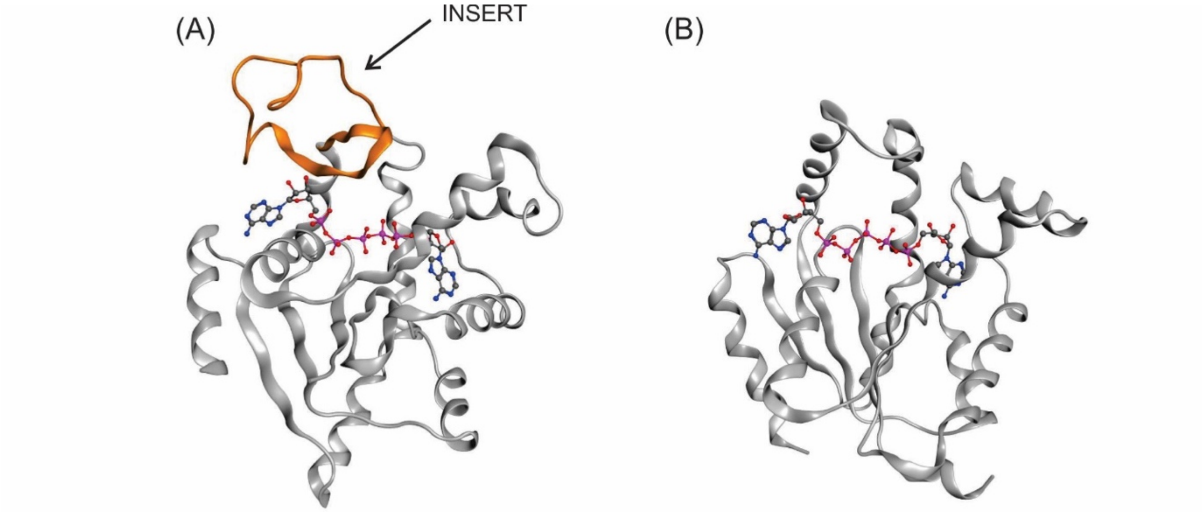


Figure S1. Structures of long and short monomeric NMP kinases

(**A**) Example of a long monomeric NMP kinase with the INSERT segment indicated and colored in orange. Adenylate kinase from *Escherichia Coli* in complex with the inhibitor Ap5A[1] (1AKE). (**B**) Example of a short monomeric NMP kinase. Adenylate kinase from the Antarctic fish *Notothenia coriiceps* in complex with Ap5A[2] (5XZ2).

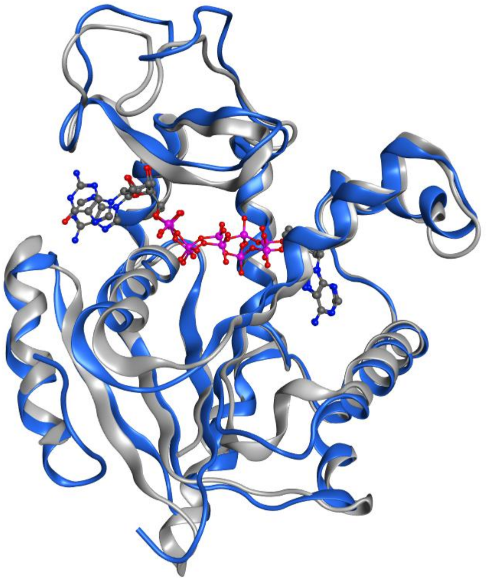


Figure S2. AK3 in a closed conformation in complex with the inhibitor
Shown is a superimposition of AK3 (blue ribbon) in complex with Gp5A with AK_{eco} (gray ribbon) in complex with Ap5A[1] (1AKE).

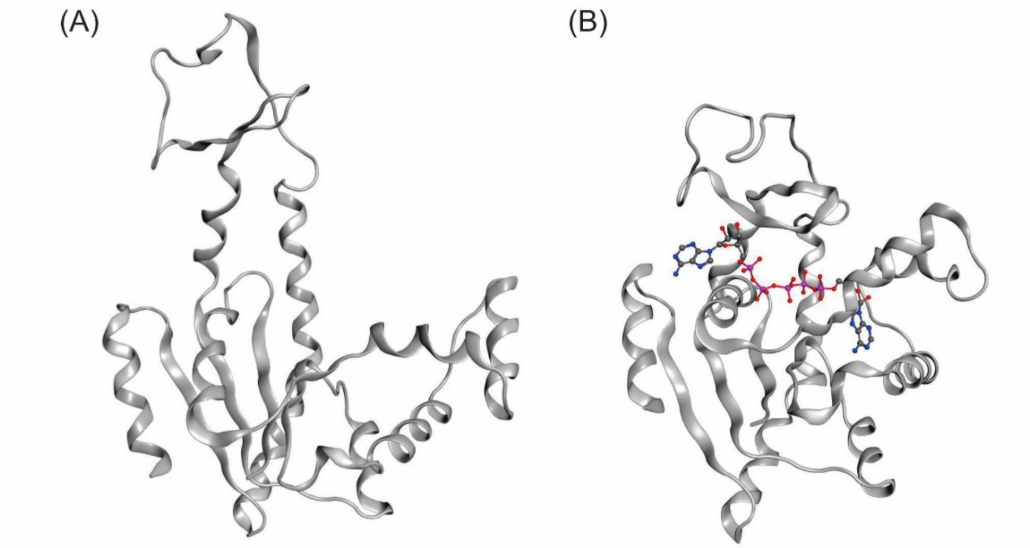


Figure S3. Structures of an open and closed NMP kinase
Crystallographic structures of AK_{eco} are used for illustration of the open and closed states. (**A**) Substrate free and open AK_{eco}[3] (4AKE). (**B**) Ap5A bound and closed AK_{eco}[1] (1AKE).

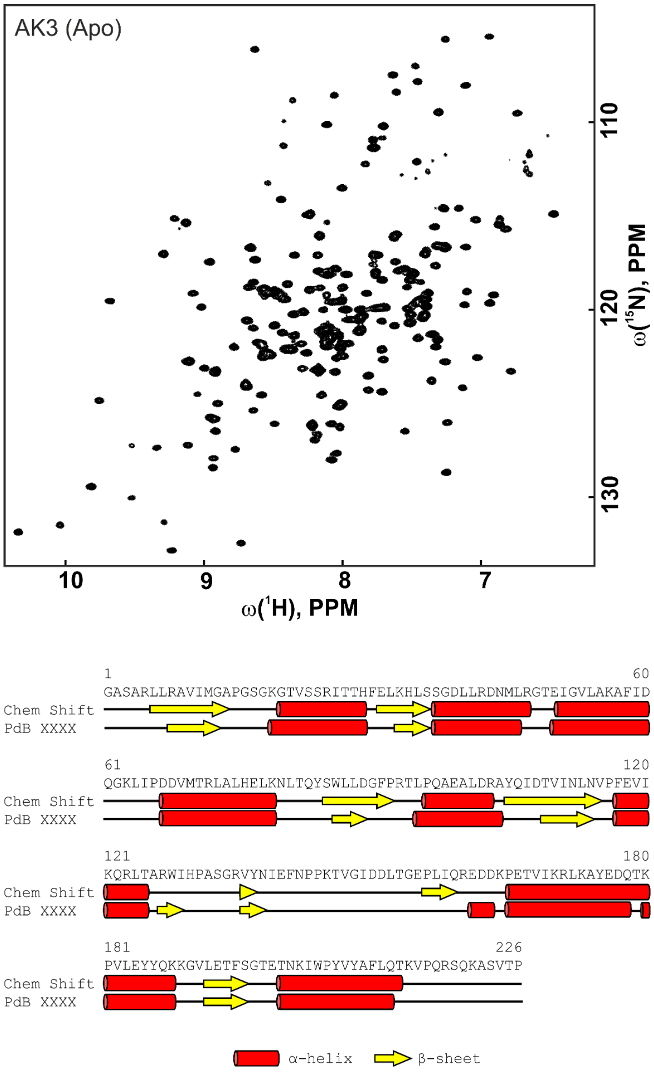


Figure S4. Top: Trosy-HSQC spectra of human AK3 (apo form) and, bottom: a comparison between predicted secondary structure based on the chemical shifts and the secondary structure of the high resolution AK3-Gp5A X-ray structure (6ZJB.pdb).

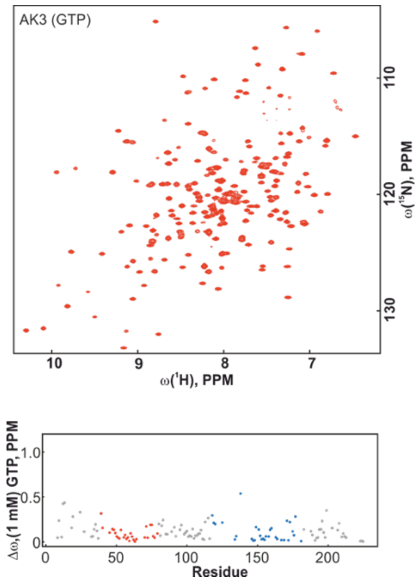


Figure S5 Top; Trosy-HSQC spectra of human AK3 (GTP saturated, 1mM) and bottom; chemical shift differences (|1H|+0.2*|15N|) between apo AK3 and AK3 saturated with 1mM GTP. Residues in the GTPlid are colored red and residues in the AMpdb are colored blue.

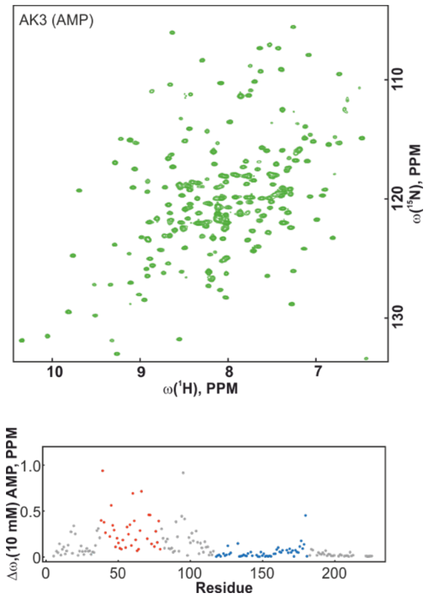


Figure S6 Top: Trosy-HSQC spectra of human AK3 (AMP saturated, 10mM) and bottom; chemical shift differences (|^{1}H|+0.2*|^{15}N|) between apo AK3 and AK3 saturated with 10mM AMP. Residues in the GTPlid are colored red and residues in the AMpdb are colored blue.

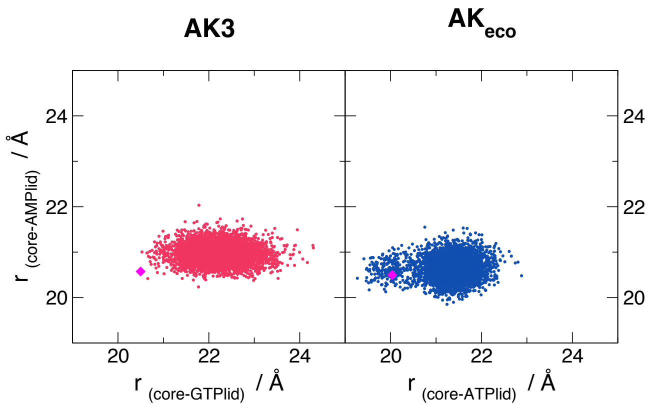


Figure S7. Displacement of the nucleotide binding domains relative to the protein core in AK3 and AK$_{eco}$ during the course of an 500ns MD simulation

The lid-to-core center-of-mass distances were calculated on the atomic coordinates saved every 0.1ns. The elliptical distribution in AK3 (*left panel*) indicates the shifting of the GTPlid by approximately 3Å in the direction towards and away from the protein core while the AMPbd displaces only by about 0.5Å around the 21Å mark. The distribution pattern for AK$_{eco}$ (*right panel*) is more spherical suggesting that both domains exhibit a similar range of motion relative to the core, with the ATPlid domain being at a marginally farther distance from the core. Pink diamonds in both panels mark the values for the respective X-ray structures.

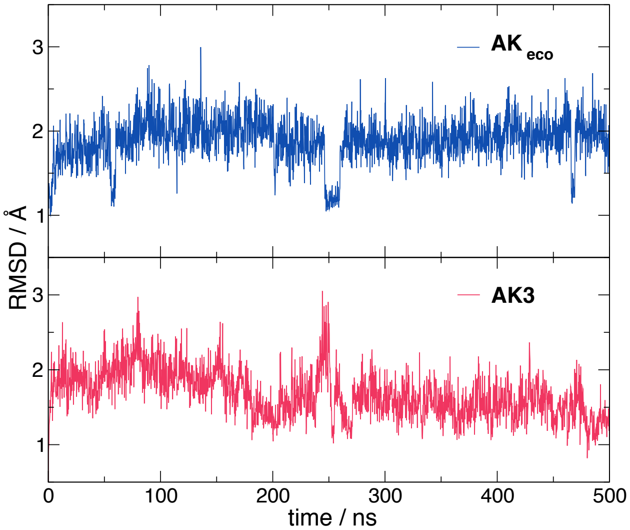
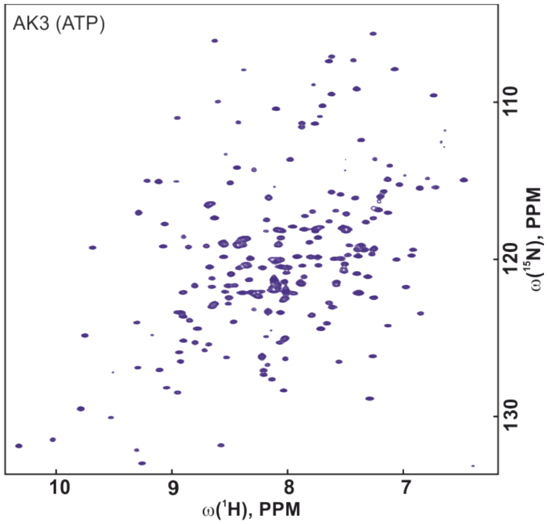


Figure S8. The root-mean-square deviation (RMSD) of the backbone atomic position during the course of a 500ns MD simulation.

The RMSD value quantifies a relative deviation of atomic position per coordinate frame saved every 0.1 ns against the X-ray crystalline reference structure. Both adenylate kinase models demonstrate appropriate structural stability of the equilibrated system with an average displacement of 2Å (AK_{eco}) and 1.5Å (AK3) from the reference conformation.



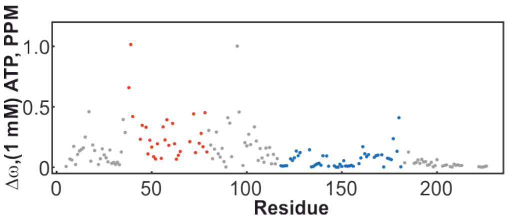


Figure S9 Top; Trosy-HSQC spectra of human AK3 (ATP saturated, 1mM) and bottom; chemical shift differences ($|^{1}H|+0.2*|^{15}N|$) between apo AK3 and AK3 saturated with 1mM ATP. Residues in the GTPlid are colored red and residues in the AMPdb are colored blue.

Appendix I: references

(1) Müller, C. W.; Schulz, G. E., (1992) Structure of the complex between adenylate kinase from Escherichia coli and the inhibitor AP5A refined at 1.9 A resolution - A model for a catalytic transition-State. J. Mol. Biol. 224, 159-177.

(2) Moon, S.; Kim, J.; Bae, E., (2017) Structural analyses of adenylate kinases from Antarctic and tropical fishes for understanding cold adaptation of enzymes. Sci. Rep. 7, 12.

(3) Müller, C. W.; Schlauderer, G.J.; Reinstein, J.; Schulz, G. E., (1996) Adenylate kinase motions during catalysis: An energetic counterweight balancing substrate binding. Structure. 4, 147-156.

Chapter II

Dynamic Connection between Enzymatic Catalysis and Collective Protein Motions

Pedro Ojeda-May[1,2], Ameeq Ul Mushtaq,[1] Per Rogne[1], Apoorva Verma[1], Victor Ovchinnikov[3], Christin Grundström[1], Beata Dulko-Smith[4], Uwe H. Sauer[1], Magnus Wolf-Watz[1], and Kwangho Nam[4]

[1]Department of Chemistry, Umeå University, Umeå SE-90187, Sweden
[2]High Performance Computing Centre North (HPC2N), Umeå University, Umeå SE-90187, Sweden
[3]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA
[4]Department of Chemistry and Biochemistry, University of Texas at Arlington, Arlington, TX 76019, USA

Biochemistry, **2021**, *60*(28), 2246-2258.

Abstract

Enzymes employ a wide range of protein motions to achieve efficient catalysis of chemical reactions. While the role of collective protein motions in substrate binding, product release and regulation of enzymatic activity is generally understood, their roles in catalytic steps per se remain uncertain. Here, molecular dynamics simulations, enzyme kinetics, X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are combined to elucidate the catalytic mechanism of adenylate kinase and to delineate the roles of catalytic residues in catalysis and conformational change of the enzyme. The study reveals that the motions in active site, which occur on a timescale of ps-ns, link the catalytic reaction to the slow conformational dynamics of the enzyme, by modulating the free energy landscapes of subdomain motions. In particular, substantial conformational rearrangement occurs in the active site following the catalytic reaction. This rearrangement not only affects the reaction barrier but also promotes a more open conformation of the enzyme after the reaction, which then results in an accelerated opening of the enzyme as compared to the reactant state. The results illustrate a linkage between enzymatic catalysis and collective protein motions, whereby the disparate time scales between the two processes are bridged by a cascade of intermediate-scale motion of catalytic

residues modulating the free energy landscapes of the catalytic and conformational change processes.

Introduction

Protein motions involved in enzymatic catalysis span a wide range of spatial and temporal scales, from thermal fluctuations that equilibrate rapidly along the catalytic reaction coordinate on the one hand, to the slow collective motions that accompany enzymatic catalysis, ligand binding and allosteric regulation,[4-7] on the other hand. Thermal fluctuations typically take place locally at the active site of an enzyme in response to changes in reacting species and have been studied using various kinetic, spectroscopic and simulation approaches.[9, 10] While the role of collective protein motions in substrate binding, product release and regulation of enzymatic activity is generally understood, their role in catalysis has been challenging to understand due to the difficulty in attributing a direct functional role to such motions or quantifying their effects.[11-14] For this reason, many studies have focused on identifying a phenomenological link between slow protein motions and enzymatic catalysis.[15-18] Furthermore, elucidating the detailed mechanism and sequence of events between slow conformational change, ligand binding/release and catalysis remains a theoretical and experimental challenge. Overcoming this challenge with a combined computational and experimental approach, here, we examine conformational plasticity and dynamics in the active site of *Escherichia coli* (*E. coli*) adenylate kinase (AdK) and explore the mechanisms that link enzymatic catalysis to large-scale collective motion of the enzyme, achieved through the modulation of their respective free energy landscape.

In the cell, AdK regulates cellular adenosine triphosphophate (ATP)/adenosine diphosphate (ADP) homeostasis by catalyzing the reversible conversion of ATP and adenosine monophosphate (AMP) to 2 molecules of ADP.19 X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy have established that during the catalytic cycle, the enzyme undergoes an open-to-close conformational change of two substrate binding subdomains (ATPlid and AMPlid; Figure 1).

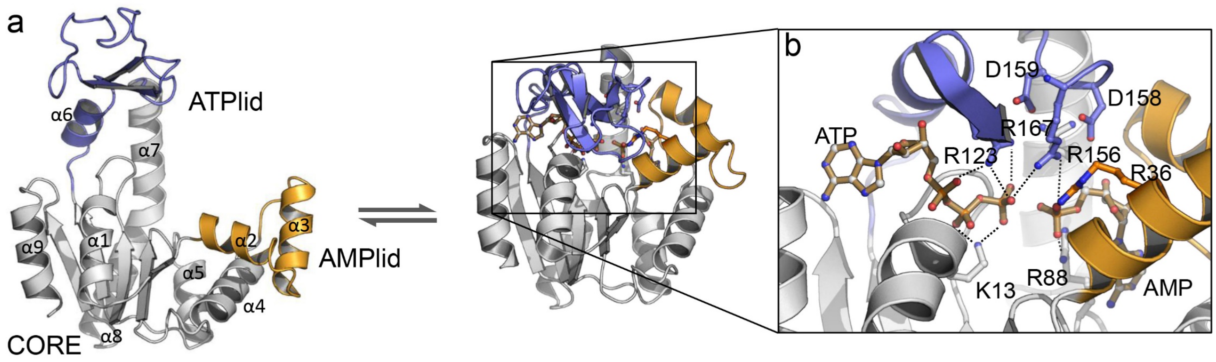


Figure 1. Open-to-close conformational change of *E. coli* adenylate kinase (AdK)
a Ribbon diagram of *E. coli* AdK in the open (left; PDB ID: 4AKE[1]) and closed conformation (right; PDB ID: 7APU). **b** Architecture of the closed-conformation active site with bound ATP and AMP. The five active site arginine residues, one lysine residue and the bound ATP and AMP molecules are shown in a ball and stick model. In the figure, the 2 ADPs in the 7APU structure are modified to ATP and AMP.

A third subdomain, called CORE, is static but crucial for thermal stability of the enzyme.[20] Subdomain opening occurs on the the same time scale as the catalytic turnover,[3] and is thus rate limiting to the overall catalytic turn-over. On this basis, extensive studies, both experimental[8, 21-24] and theoretical,[25-27] were conducted to elucidate the mechanisms of the open-to-close conformational change of the enzyme. However, despite various site-directed mutagenesis studies,[8, 28-31] the catalytic mechanism of the enzyme remain not fully understood at the atomic level, including the roles of active site residues[32] and metal ion in catalysis,[33] and the nature of the reaction transition state.[34] For example, Kerns et al. recently studied the role of the catalytic Mg^{2+} ion as an electrostatic pivot in the phosphoryl transfer, using X-ray crystallography, NMR, enzyme kinetics and molecular dynamics (MD) simulations.[8] Nevertheless, the precise role of Mg^{2+} in the stabilization of the phosphoryl-transfer transition state remains unknown, and its study requires a quantum mechanical description of the catalytic reaction to fully characterize the nature of the transition state (TS). In addition, the coupling between enzymatic catalysis and intrinsic protein motions, which occur at different time scales,[4, 35, 36] remains a fundamental and open question in enzymology. In this study, we present a mechanistic connection between the large-scale opening motion of the enzyme adenylate kinase (i.e., opening of the ATP and AMPlids) and its catalytic phosphoryl transfer reaction, discovered by a combined application of NMR, X-ray crystallography, site-directed mutagenesis and free energy simulations. This connection is achieved through the change of their respective free energy landscapes occurring in response to the change of the other. For example, the free energy barrier of the opening of the enzyme differed between the reactant and product states of the catalytic phosphoryl transfer reaction. Likewise, the free energy of the catalytic reaction was affected by the relative opening of the enzyme. In addition, arginine residues in the active site of the enzyme altered their orientation and interactions along the catalytic reaction. These alterations led to the change in the orientations of the two lids relative to the CORE subdomain, orienting them for faster opening in the product state than in the reactant state. In this way, the catalytic reaction in the active site of the enzyme effects on the slow close-to-open protein motion, without the two different timescale events being directly coupled. Therefore, the coupling mechanism presented in this work is different from the direct coupling mechanism where certain rapid protein motions, such as the rate-promoting vibrations proposed by Schwartz and co-workers,[9, 37] occur concurrently with the crossing of the chemical barrier. In addition, this type of coupling mechanism is not exclusive to this enzyme, but has been proposed in motor proteins, most notably in F1-ATPase, where the catalytic reaction modulates the free energy landscape of the rotation of its central rotor subunit, without the events taking place concurrently.[38] In the following, we presents the materials and methods, followed by the results and discussion.

Materials and Methods

Protein expression and purification

E. coli adenylate kinase (AdK) was over-expressed using a previously published protocol[39] based on a self-inducing plasmid (pEAK91). Plasmids carrying the mutations, R2A, R36A, R36K, R88A, R88K, R123A, R123K, R131A, R156K, R167A and R167K were acquired from Genescript, in which each mutation, except R156K, was designed to examine the impact of side chain removal. The R156K mutant was designed based on Kerns et al.,[8] and in this study it was prepared to provide a consistent set of kinetic parameters. The expressed proteins were purified by affinity chromatography on a Blue sepharose column, followed by a gel filtration step. To obtain uniformly ^{15}N labeled enzyme for the ^{15}N edited NMR experiments, expression cells were grown on minimal media (M9) with $^{15}NH_4Cl$ as the sole nitrogen source.

X-ray crystallography

AdK at a concentration of 18.3 mg/ml (~776 μM) was mixed with 5mM each of AMP and guanosine triphosphophate (GTP) in a 30 mM MOPS buffer at pH 7 containing 50 mM NaCl. Diffraction quality crystals were grown by the hanging drop method in 24 well Linbro plates. Mixed on a cover slip were 2 μl of AdK, preincubated with AMP and GTP, which was a part of screen of different combinations and concentrations of AMP, ADP and ATP with GMP, GDP and GTP, and 2 μl of the buffer containing 0.2 M $NH_4CH_3CO_2$, buffered with 100 mM CH_3CO_2Na adjusted to pH 4.6 and 30 % v/v PEG 4000 as the precipitant. The coverslip with the crystallization drop was inverted and sealed over a well containing 0.5 ml of the precipitant buffer. No Mg^{2+} ions were added at any stage, in order to reduce catalytic reactions in the crystallization set-up. Crystals grew at 291.2 K within 5 days to a size of 0.5 × 0.1 × 0.1 mm^3. Crystals were mounted free floating suspended in a nylon loop and vitrified by chock cooling to 100 K in a nitrogen gas stream (Oxford CryoSystems Ltd, UK). X-ray diffraction data was collected at a wavelength of 0.9750 Å onto a Pilatus 6M detector (Dectris) at beam line ID23-1 of the European Synchrotron Facility (ESRF) in Grenoble, France. The diffraction intensities were indexed, integrated with XDS[40] and prepared for scaling with pointless.[41] Aimless[42] was used to scale and merge the intensities which were converted to structure factors with cTruncate.[43] About 5% of the data was used for R_{free} calculation. Phase determination and structure solution were carried out by molecular replacement (MR) using the *E. coli* AdK (PDB ID: 1AKE[44]) as the search model in Phaser.[45] AdK structure refinement against data extending to 1.36 Å was carried out with phenix.refine[46, 47] (Phenix program package version 1.17.1-3660). Manual model building was carried out with Coot (version 0.8.9.2 EL).[48] Refinement and model building were iterated until the R-free and R-factor converged. The data processing and refinement statistics are listed in Table S1.

$^1H_\varepsilon$-$^{15}N_\varepsilon$ correlation and 15N-relaxation NMR spectroscopy

2D 1Hε-15Nε correlation spectra for side chain arginines and backbone relaxation spectra were recorded on a Bruker Avance III HD spectrometer at 850 MHz using a triple-resonance (TXI 5-mm) cryoprobe equipped with pulsed field gradients along the x, y, and z axes. Side chain arginine spectra were measured in a 30 mM MES, 50 mM NaCl buffer at pH 5.5. Backbone relaxation experiments were measured in 30 mM MOPS, 50 mM NaCl at pH 7.0. D2O [10% (v/v)] was added to all NMR samples for the field-frequency lock. NMR spectra were processed using NMRPipe and NMRDraw softwares49 and analyzed using Topspin 3.6 (Bruker) and SPARKY (v3.113; https://www.cgl.ucsf.edu/home/sparky; UCSF, San Francisco, CA, USA). The side chain correlation spectra were recorded by placing the 15N carrier frequency at 90 ppm and sweep-width was set to 86 ppm for HSQC measurements. The 15N-relaxation experiments, T1, T2, and 15N-{1H}-heteronuclear nuclear Overhauser effects (NOEs), were collected in an interleaved manner. Interleaved 2D NMR experiments based on 1H-15N correlation spectra were performed. Delays of 50, 100(duplicate), 200, 400, 500, 600, 800, 1000(duplicate), 1200 and 1500 ms were used for T1 measurements, and 16.96, 33.92, 50.88, 84.80, 101.76, 118.72, 135.68, 152.64, 169.60 (duplicate), 186.56 and 203.52 ms were used for T2 measurements.50 Relaxation delays of 3 s and 2 s were used for the T1 and T2 measurements, respectively. Steady-state 15N-{1H}-heteronuclear NOE was measured with either 5 s delays between each free-induction decay or 2 s delays, followed by a 3 s series of 120° nonselective 1H pulses.51 T1, T2, and 15N-{1H}-heteronuclear NOE experiments were performed with time-domain sizes of 256 × 2048 complex points and sweep widths of 10204.08 Hz and 2584.62 Hz along the 1H and 15N dimensions, respectively, with 8 scans for T1 and T2 and 32 scans for the 15N-{1H}-heteronuclear NOE experiment. Arginine side chain 15N-{1H}-heteronuclear NOE experiments were recorded as described with 15N carrier frequency and sweep-width set at 86 and 60 ppm, respectively. The details of 15N-relxation data analysis are provided in the Supporting Information. Assignments of the 1Hε-15Nε correlations were accomplished by comparing NMR spectra of apo and P1,P5-di(adenosine-5′) pentaphosphate (Ap5A) bound states with those of several AdK variants. The variants were: R2A, R36A, R26K, R88A, R88K, R119A, R123A, R123K, R131A, R156K and R167K.

System preparation and molecular dynamics (MD) simulations

Three systems were prepared based on the 1AKE wild-type (WT) *E. coli* AdK structure and the 2CDN *Mycobacterium tuberculosis* AdK structure^{33} for ADPs and Mg^{2+}: (1) the reactant state (RS) with ATP and AMP, (2) the transition-state mimic (TSM) and (3) the product state (PS) with

2ADPs. The ATP in RS and one of the two ADPs in PS were protonated. The TSM state was prepared to mimic the geometries of the transition state (TS) of the catalytic reaction (obtained from the QM/MM simulations), and one of the transferred phosphoryl oxygens was protonated. In all systems, we included Mg^{2+} ion, its coordinating waters and all crystal waters of 1AKE. CHARMM22[52] and the CMAP[53] potentials were used to represent the protein and ions, and for ATP, ADP and AMP, CHARMM27[54, 55] force field was used. After solvation with a 72 Å cubic box of TIP3P[56] waters and 150 mM NaCl, each system was minimized and equilibrated for 200 ps at 298 K. Production MD simulations were performed for 300 ns at 298 K and 1 bar with the leap-frog Verlet for 2 fs time integration and SHAKE[57] applied to all bonds involving hydrogen atoms. All MD simulations were performed using the CHARMM[58]/OpenMM[59] program. The particle mesh Ewald (PME) summation method[60] was used and vdW interactions were evaluated up to 12 Å. All MD simulations were triplicated to achieve convergence of simulation results. See the Supporting Information for details of system preparation, force field parameters for TSM and MD simulations.

Hybrid quantum mechanical and molecular mechanical (QM/MM) simulation

The system for the QM/MM simulation was prepared following the procedure described above for the product system (i.e., with 2 ADPs), except that it was solvated with a 69 Å rhombic dodecahedron (RHDO) box of TIP3P waters. After minimization and equilibration, the QM/MM MD simulation was performed at 298 K and 1 bar, in which the two ADP molecules, one Mg^{2+} ion and its coordinated waters were described by the AM1/d-PhoT QM model,[61] which was developed for the catalytic phosphoryl transfer reactions, and the rest by the CHARMM force field. Long-range electrostatic interactions were evaluated with the QM/MM-PME algorithm[62, 63] and vdW interactions with the 10.0 Å cutoff. All QM/MM MD simulations were performed using the CHARMM program (version c43a1),[58, 64] with 0.5 fs time integration and SHAKE[57] applied to non-QM atoms. System temperature and pressure were controlled using the Nose-Hoover thermostat[65] and the extended system pressure algorithm,[66] respectively.

Starting from the path generated by the umbrella sampling (US)[67] simulations, the string method in collective variables (SMCV) simulation[68, 69] was performed using the distances of the forming and cleaving P-O bonds (i.e., P-O_{Nuc} and P-O_{LG}) as CVs. The entire path was described using 48 SM replicas (i.e., MD images). For each string replica, 100 ps QM/MM MD simulation was performed to optimize the path, followed by 400 ps QM/MM MD to calculate the free energy along the path. Thus, a total of 24 ns QM/MM MD simulations were performed for the reaction.

In the simulation, we protonated one of the γ phosphoryl oxygens of ATP. To determine the site of protonation, the proton transfer from ATP to AMP was simulated using the US technique. The simulation yielded 2.4 kcal/mol as the free energy of reaction (Figure S1), suggesting ATP as the preferred site for protonation for the wild-type enzyme. [This is different in the R167A mutant, in which the protonation of AMP is favored by 3.3 kcal/mol.] In addition, the phosphoryl transfer reaction with fully deprotonated ATP and AMP (using SMCV) was simulated to test the impact of ATP protonation on the catalytic barrier and reaction free energy. The barrier and the reaction free energy determined were 10.8±0.2 kcal/mol and 0.4±0.3 kcal/mol, respectively.

SMCV free energy (FE) simulation of AdK opening

Starting from the close-to-open conformational change path obtained from the MD simulation in PS, the SMCV path optimization was performed in an iterative manner. In each iteration, a 4 ps MD simulation was performed, during which each SM image was harmonically restrained with a force constant of 100.0 kcal/mol-Å^2 to each CV and the average free energy gradients were evaluated for the evolution of path. In the present study, the CORE-ATPlid, CORE-AMPlid and ATPlid-AMPlid distances were used as CVs to describe the path (Figure S2) and the entire path was described with 72 SM images. After the evolution of the path, the CV positions were reparametrized to have an equal distance between neighboring images, completing one path optimization cycle. The path optimization was performed for a total of 2 ns for each image (i.e., 500 iteration). Then, additional 20 ns MD simulations were performed to determine the potential of mean force along the path. Thus, a total of 1.58 µs MD simulations were performed for the conformational change FE profile.

For the opening in the reactant state, 2 ADPs were modified to ATP + AMP for each SM image obtained from the product state SMCV simulations and briefly energy minimized. The rest of the path optimization and the potential mean force simulation followed the same procedure as that used for the product state SMCV simulation. To test the convergence of the results, additional SMCV FE simulations were performed from different initial path with essentially identical results.

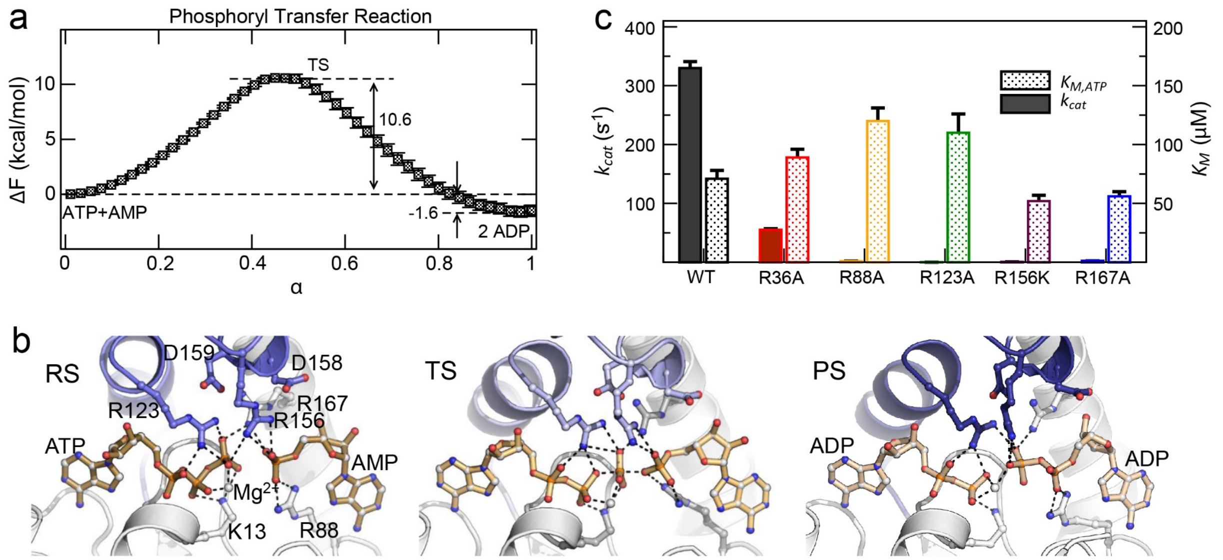


Figure 2. Catalytic mechanism of AdK

a Free energy profile of the catalytic phosphoryl transfer between ATP+AMP and 2 ADPs, where α denotes the normalized reaction coordinate between 0 and 1. **b** Representative structures of the reactant (RS), transition (TS) and product state (PS), obtained from the QM/MM simulations. **c** Kinetic parameters (k_{cat} and $K_{M,ATP}$ for ATP) determined for wild-type AdK and five active site arginine variants

Results

Phosphoryl transfer mechanism and impacts of catalytic residue variants

The catalytic mechanism of AdK was studied by applying the hybrid quantum mechanical and molecular mechanics (QM/MM) method,[64, 70] using the AM1/d-PhoT QM method[61] for the catalytically reactive region and the CHARMM force field[52] for the rest of the enzyme and water solvent. The barrier and free energy of the phosphoryl transfer reaction between ATP + AMP and 2 ADPs were calculated to be 10.6±0.2 kcal/mol and -1.6±0.5 kcal/mol, respectively (Figure 2a). These values are comparable to < 12.4 kcal/mol of catalytic barrier and -1.4 kcal/mol of reaction free energy estimated from kinetic measurements,[8] suggesting the reliability of the employed QM/MM method for the study of the enzyme's catalytic mechanism.

The reaction begins with the nucleophilic attack of an α-phosphoryl oxygen of AMP at the γ-phosphorus of ATP. This is followed by the formation of a transition state and then the reaction product (Figure 2b). Details of the catalytic mechanism and the character of the transition sate are provided in the Appendix II and Figure S3. Importantly, the active site of *E. coli* AdK harbors five conserved arginine residues, i.e., R36, R88, R123, R156, and R167 (Figure 1b), which are well positioned for efficient phosphoryl transfer and, with a bound Mg^{2+} ion, balance the charges of ATP and AMP. In accordance with this, their mutations to alanine, in one case to lysine,[28] substantially reduced the catalytic activity of the enzyme (Figure 2c and Table S2): R88A, R123A, R156K and R167A exhibited less than 1 % of the wild-type activity and R36A 17 % of the wild-type activity. On the other hand, the K_M values were less affected by the mutations and the values determined for ATP ranged between 52 – 120 μM (Table S2). Previously, Wolf-Watz and co-workers established that the K_M value for the wild-type AdK (71 μM for ATP) can serve as a valid proxy for binding affinity, based on its similarity to the 53 μM K_d value for ATP.[71] Assuming the same for the arginine mutants, the small changes in K_M indicate that the thermodynamic contribution to binding exerted by the positive charge of the arginines is small. This was unexpected, because if the effects of the arginine mutation on k_{cat} were primarily electrostatic in nature, equally large impacts on substrate binding would be expected as seen in the mutation of another catalytic residue, lysine 13, to glutamine (K13Q), which affected both k_{cat} and K_M.[28] Then there are contributions of the enzyme, other than electrostatic, to the decrease in catalytic activity in the mutants.

Likewise, significant impacts on catalytic activity have been reported from the mutation of the active site arginines in other adenylate kinases.[8, 29-31] For example, in human adenylate kinase 1 (AK1), the R44A and R97A mutations (equivalent to R36A and R88A of *E. coli* AdK, respectively) retained 21 % and 19 % of the wild-type activity and the R132A, R138A and R149A mutations (i.e., R123A, R156A and R167A of *E. coli* AdK, respectively) < 0.1 % of the wild-type activity, respectively.[29] Similar results were also reported for chicken AK1.[30, 31] However, due to the difference in the ATPlid length, their precise location and interactions are different between

AK1 and *E*. *coli* AdK, and their mutations therefore have different impacts on the catalytic activity. To this end, Table S2 provides a consistent set of kinetics data on mutation for a single and widely investigated enzyme system, i.e. *E. coli* AdK.

Changes in the conformation of active site residues along catalytic reaction

In Figure 2b, we show representative structures of the reactant (RS), transition (TS) and product states (PS) obtained from the QM/MM simulation. These structures illustrate the reorganization of the active site residues along the catalytic reaction trajectory (schematically in Figure S4). On the ATP side, R123, which interacts with ATP in RS, follows the transfer of the γ-phosphoryl group to AMP. It therefore maintains its interaction with both ADP molecules in PS. A conserved lysine (i.e., K13) also interacts with ATP in RS via its β and γ-phosphoryl groups. In PS, its interaction with the γ-phosphoryl group is lost, while the interaction with the β-phosphoryl group is retained. On the AMP side, R156 interacts with ATP and AMP via their γ and α-phosphoryl groups, respectively, in RS. With the transfer of the γ-phosphoryl group, the α-phosphoryl group of AMP turns away from R156, but the residue retains the interaction with the transferred γ-phosphoryl group. R167 undergoes a similar change to R156. Namely, R167 interacts with both ATP and AMP in RS, and with the transfer of the γ-phosphoryl group, its interaction with AMP weakens. In PS, the residue disengages from the α-phosphoryl group of ADP and only interacts with the transferred phosphoryl group in the AMP binding site. In Table S3, we present the hydrogen-bond distances formed between the different arginine residues with ATP and AMP in the RS, TS and PS, respectively, determined from the QM/MM simulations.

In addition, the α6-helix changes its orientation relative to the α7-helix along the catalytic reaction, in which the two helices anchor ATPlid to the reaction substrates (Figure 2b). This change in orientation can be attributed to the change in R123, R156 and R167 interactions during the reaction; note that R123 and R156 are located immediate vicinity of the two helices and R167 in the middle of α7. For example, the α6-helix tilts its orientation toward α7 following the rearrangement of R123 with the transfer of the γ-phosphoryl group. Similarly, the orientation of α7-helix can be affected by R156 and R167. As a result, the interaction between the two helices increases, which then releases the loop connecting the α3 and α4-helices of AMPlid from the α7-helix. Finally, and as discussed in the "*Dual role of the active site arginines*" section, these changes of orientations and interactions ultimately lead to the opening of ATP and AMPlids after the reaction.

Effects of the conformational plasticity of R167 on catalysis

To understand the 137-fold reduction in k_{cat} observed by the R167A variant (Table S2), the QM/MM and (100 ns) MD simulations were carried out for the R167A mutant. Interestingly, the mutation not only raised the reaction barrier by 4.0 kcal/mol (Figure S5a-b), but also enhanced the conformational motion of the two substrate binding subdomains compared to the wild-type enzyme (compare Figure S5c to Figure 3a). In particular, while the wild-type enzyme remained closed during the entire (300 ns) MD simulations, the R167A mutant exhibited an opening, if not entirely, even during the 100 ns MD simulations. For example, in RS, the opening of AMPlid was accompanied by the opening of ATPlid. In PS, AMPlid fluctuated between the closed and (partially) open states, while ATPlid opened slowly. In both states, the α7-helix, harboring R167, displaced slightly away from the loop between the α-3 and α-4 helices of AMPlid. Consistent with the wild-type QM/MM result, the α7 displacement in the R167A mutant can trigger the opening of AMPlid by releasing it from α7. This result suggests that in the wild-type, R167 contributes to the slow-down of the opening of the enzyme through the stabilization of the α7-helix orientation, while it accelerates the catalytic reaction.

Next, to verify the change of the R167 orientation in PS, we analyzed the X-ray crystal structure of *E. coli* AdK with two bound ADPs in the absence of Mg^{2+} determined at a resolution of 1.36 Å (Figure 1a right and Table S1). The structure is essentially identical to other closed AdK structures, for example, with 0.2 Å root-mean-square deviation (RMSD) from the 1AKE structure. In addition, despite the absence of the bound Mg^{2+} ion, the orientation of the active site ADPs and surrounding residues was very close to the structure reported with Mg^{2+} by Kerns et al.,[8] thus providing meaningful details of ADP interaction at both binding sites. In our structure, the AMP binding site is partially occupied by ADP (~90%) and AMP (~10%) based on the occupancy of ADP β-phosphoryl group, while the ATP binding site is fully occupied by ADP. Previously, ATP and AMP were shown to bind independently to AdK, with K_d of 53 μM for ATP and 210 μM for AMP.71 Here, ADP binding to AdK is quantified with isothermal titration calorimetry (Figure S6 and Table S4).

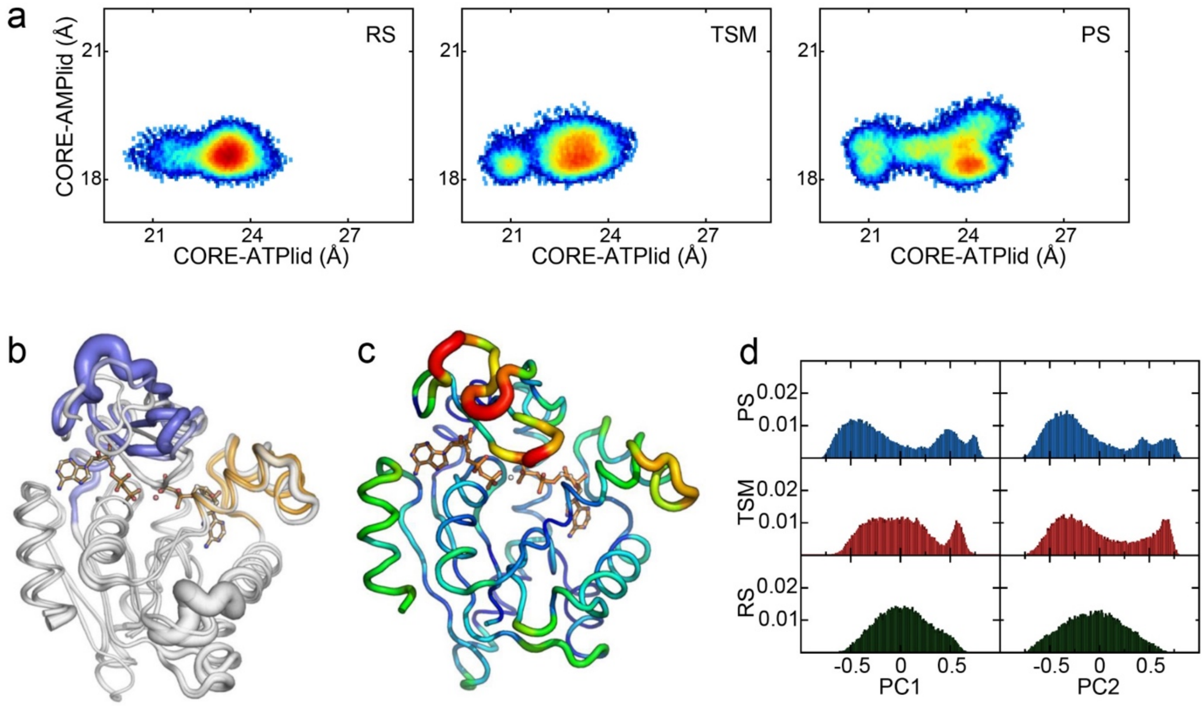


Figure 3. ATPlid dynamics along the catalytic reaction trajectory

a Distribution of CORE-ATPlid and CORE-AMPlid distances between the reactant (RS), transition-state mimic (TSM) and product states (PS). The color varies from blue to red with increasing frequency of the observed distances. **b** Comparison of ATPlid and AMPlid orientations for X-ray structures in different catalytic states. The ATPlid is shown in white for the structure with a bound ATP analog and AMP (1ANK[2]), light blue for the transition state mimic structure (i.e., ADP–AlF_4^-–AMP; 3SR0[8]) and blue for the 2 ADP structure (4JKY[8]), respectively. In all structures, the CORE subdomain is displayed in white and 2ADP+Mg^{2+} are only shown as bound ligands for clarity. The thickness of the tube represents the temperature factor (B-factor) of the protein backbone Cα atoms. **c** The lowest frequency protein motion (PC1) identified from PCA, whose tube thickness and color represent the amplitude of the PC1 motion. **d** Normalized projection of MD simulation trajectories onto the PC1 and PC2 vectors.

The data can be fitted well with a single K_d value of 9 μM, suggesting that the two ADP molecules bind with the same binding affinity for the two binding sites.[72] The 6 and 23-fold changes in the binding affinity for ADP compared to those of ATP and AMP, respectively, suggest some level of cooperativity between the two ADP binding events.

In our ADP bound co-crystal structure, the electron density covering R167 suggests two conformations, designated here respectively by α and β (Figure 4). This heterogeneity is also observed in other X-ray structures of AdK[2, 8, 44, 73] (Figure S7).

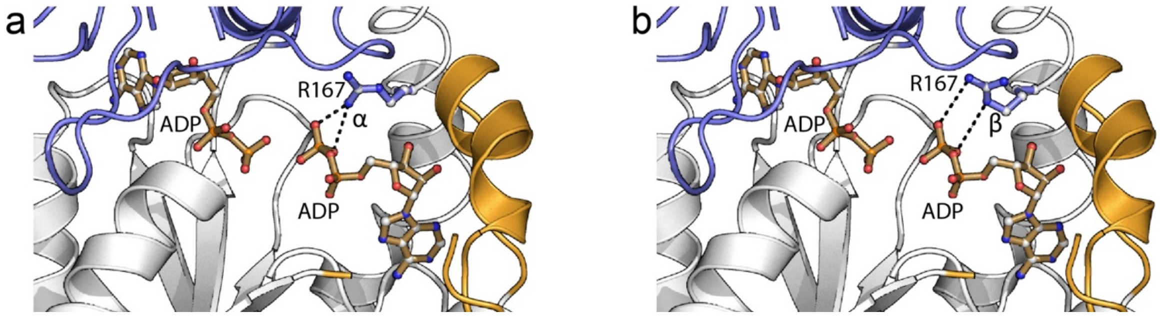


Figure 4. Different orientations of the R167 side chain in the X-ray structure of AdK in complex with 2 ADPs (7APU)
(**a**) the α conformation and (**b**) the β conformation compatible with bound ADP in the AMP binding pocket. The color scheme of the protein cartoon representation is the same as in Figure 1.

To test the effects of the different orientations of R167 on the catalytic reaction, we performed the QM/MM simulations while restraining the C_γ-C_δ-N_ε-C_ζ dihedral angle (i.e., δ angle) of R167 at -66° for the β conformation. The unrestrained QM/MM simulation was close to the α conformation with an average δ angle of -171° at RS. Restraining the R167 side chain orientation raised the catalytic barrier by 4.6 kcal/mol and the reaction free energy by 1.6 kcal/mol, respectively. This suggests the importance of R167 flexibility for the low catalytic reaction barrier, possibly by stabilizing the transition state more effectively than the reactant state. Moreover, the MD simulations, which were performed in the transition state mimic (TSM) and product states (PS) to access the behavior of the active site residues over a longer period of time, showed an enhanced fluctuation of the δ angle towards the β conformation (Figure S8). Likewise, R156 exhibited conformational heterogeneity based on the superposition of multiple X-ray structures.[8] This finding, with the present results, suggests the important of conformational heterogeneity of the two arginine residues (i.e., R156 and R167) in the catalytic turnover of the enzyme.

Role of Mg^{2+} ion in catalysis

The Mg^{2+} ion bound in the active site plays an important role in achieving high catalytic efficiency, as demonstrated by the 4.4×10^{4}-fold reduction of k_{cat} in the absence of Mg^{2+} (from 330 ± 11 s^{-1} to $7.5\times10^{-3} \pm 7.5\times10^{-4}$ s^{-1}) as determined from a real-time ^{31}P NMR assay[74, 75] (Figure S9). Kerns et al. also reported a similar level of reduction in k_{cat} with a removal of Mg^{2+}.[8]

To better understand the reduction of k_{cat}, we carried out the QM/MM and MD simulations in the absence of Mg^{2+}. The QM/MM simulation has yielded a 17.4±1.3 kcal/mol reaction barrier, which is slightly lower than the 18.74 kcal/mol barrier estimated using the forward rate constant of Kerns et al.[8] and 3.1 kcal/mol lower than that of 20.5 kcal/mol estimated on the basis of the k_{cat} value determined in this work. Compared to the reaction in the presence of Mg^{2+}, the barrier is higher by 6.8 kcal/mol. This is understandable because both β and γ-phosphoryl oxygens coordinate with the Mg^{2+} ion in the reactant sate (Figure 2c), and therefore the ion can stabilize the negative charge developed in the leaving group oxygen at TS. In addition, the MD simulations performed in the absence of Mg^{2+} showed more suppressed ATP and AMPlid motions and thus a more closed conformation (Figure S10c). Consistent with this, NMR relaxation dispersion and pre-steady-state kinetics experiments have shown much slower lid opening in the absence of Mg^{2+} (by at least 3,800 fold).[8] Together, the present study and the study by Kerns et al. suggest that Mg^{2+} accelerates both the lid opening and the catalytic reaction rate. This can be compared to the impact of R167, which slow-downs the opening of the enzyme, while it contributes to the acceleration of the catalytic reaction.

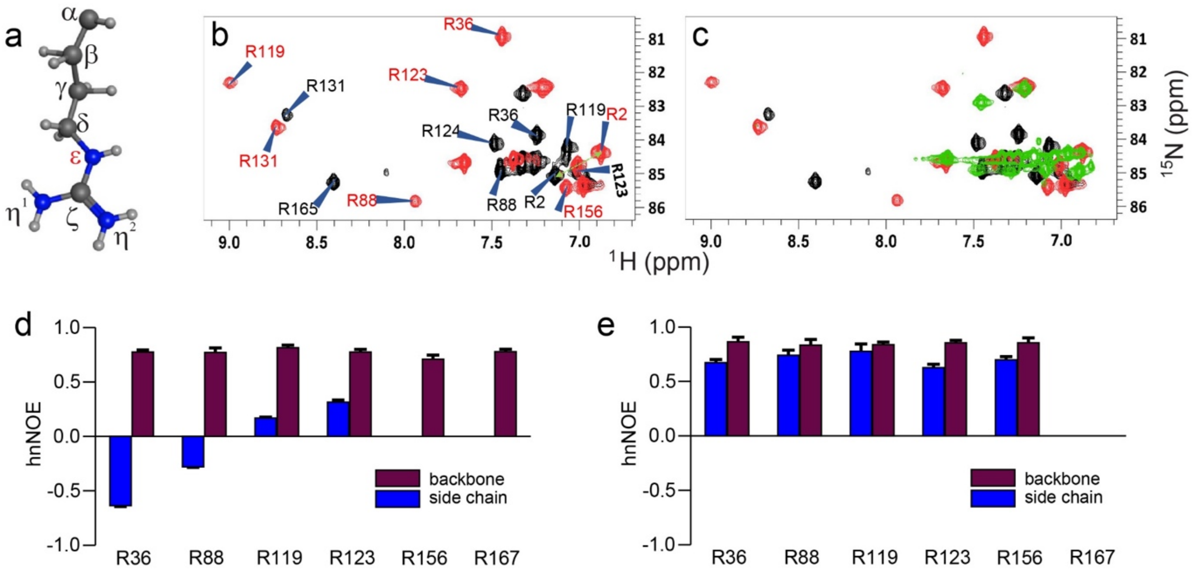


Figure 5. NMR characterization of active site plasticity of *E. coli* AdK

a Arginine side chain and notation of its heavy atoms. **b** Overlay of arginine side chain ^{1}Hε-^{15}Nε HSQC NMR spectra of apo (black) and Ap5A bound (red) states of AdK showing the assignment of the catalytic arginine residues. **c** Overlay of arginine side chain ^{1}Hε-^{15}Nε HSQC NMR spectra of apo (black), Ap5A bound (red) and ADP bound (green) states of AdK. For (**c**), each of the spectra is shown in Figure S13a-c. d-e Plots of the ^{15}N-{^{1}H}-heteronuclear NOEs of catalytic arginine backbone amide (maroon) and arginine side chain ^{15}Nε-^{1}Hε (blue) groups of (**d**) apo and (**e**) Ap5A bound AdK. Heteronuclear NOE (I_{sat}/I_{unsat} ratio of peak intensities) of each catalytic residue are plotted on the y-axis using the same values (-1.0 to 1.0) to compare the apo and Ap5A bound AdK states. Errors associated with each NOE value from measurements in triplicate are placed on top of the bar graphs.

NMR characterization of structural heterogeneity of AdK active site

To probe the structural heterogeneity of the active site arginine residues experimentally, we turned to the ^{1}H-^{15}N correlation NMR experiment. We have previously shown that backbone ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiments are very similar between Ap5A and ADP saturated states, demonstrating that the overall closed conformation is the main conformation in solution for both ligands.[71] In this study, the experiment was applied to the Hε proton at the Nε atom of arginine side chain (Figure 5a). In AdK, five of the 13 arginines are in the active site, and assignment of the catalytic and a few additional arginines was carried out by analyzing the NMR spectra of AdK variants where the arginines were replaced with alanine and/or lysine (Figures S11 and S12). With this approach, we could assign all catalytic arginines with the exception of R156 in the apo state and R167 in both apo and Ap5A bound states.

In the apo state, the catalytic arginines exhibit a relatively narrow chemical shift dispersion (black in Figure 5b), indicating that they are predominantly solvent exposed. Then, in the presence of the inhibitor Ap5A, which is known to stabilize the closed conformation of the enzyme,[44] the chemical shift dispersion of the arginine side chains has increased considerably (Figures 5b and S13a). Analysis of the heteronuclear steady-state NOEs for the side chains shows that molecular motion on the ps-ns time scale[50] is significantly restricted in the Ap5A bound state relative to the apo state (blue bars in Figure 5d-e). In the apo state, the NOE values range from -0.63 for R36 to 0.31 for R123, while the assigned side chains have an average NOE of 0.70 in the Ap5A bound state. This restriction of ps-ns side chain motions is in line with the chemical shift perturbation (CSP) data, and together suggest that the arginine side chains have well defined conformations in the Ap5A complex. This is also consistent with a related analysis of a uridine monophosphate (UMP)/cytidine monophosphate (CMP) kinase.[32]

In order to benchmark the changes in the NOE values, we quantified the corresponding values of the arginine backbone ^{1}H-^{15}N correlations (maroon bars in Figure 5d-e). The average values for the apo and Ap5A bound states were found to be 0.77 and 0.85, respectively. The most dramatic difference is seen in the apo state, where the NOE values indicate that the arginine side chains are dynamically disordered, while their backbones are well ordered. This observation is consistent with the idea that in the apo state, the side chains of all catalytic residues are projected into solvent and that the molecular motion on the ps-ns time scale is, therefore, almost unrestricted. The situation is different in the Ap5A bound state; i.e., the average NOE value for the side chains is well above 0.5 indicative of their order. However, it is still below that of the backbone atoms in the apo state, suggesting that although the side chains are relatively well ordered in the active site in complex with Ap5A, some additional motions in the ps-ns time scale are present relative to the backbone. It is possible that these “additional” motions are an intrinsic property of the enzyme and contribute to its catalytic function as probed by QM/MM simulations in the present manuscript.

In contrast to the well-defined spectrum of AdK in complex with Ap5A, the chemical shift dispersion in the ADP bound state has collapsed relative to the Ap5A state (green in Figures 5c and S13d). However, we were unable to assign the side chain resonances in complex with ADP, and the same difficulty was previously reported for the UMP/CMP kinase.[32] Nevertheless, the narrow chemical shift dispersion is consistent with a significant dynamic disorder of the arginine side chains in the ADP complex. Although the present NMR data in the absence of quantitative analysis cannot provide a molecular-level model for this dynamic disorder, it is still worth correlating this observation with the relatively insignificant changes in the K_M values for the arginine variants in comparison with the wild-type (Figure 2c and Table S2). Taken together, it is possible to envision a scenario in which binding of substrates to AdK is multivalent in nature, allowing the active site arginines to interact with the phosphates (of ATP+AMP versus 2 ADPs) with different binding poses. More specifically, since the catalytic reaction proceeds naturally with ADP during NMR experiments, the observed dynamic disorder may be due to the change in the arginine orientations following the catalytic reaction; that is, they relax to a new orientation on the ps-ns time scale after the phosphoryl transfer reaction. However, more work is required to solidify such a model.

Collective protein motions in catalysis

MD simulations performed at different reaction states (i.e., RS, TSM and PS) showed that while the enzyme stayed in the closed state in all simulations, the product state simulation exhibited a pronounced ATPlid fluctuation towards a relatively more open conformation than the reactant state (Figure 3a). Inspection of X-ray structures of AdK in complex with different ligands revealed a similar variation in ATPlid conformation and dynamics. For example, the structure of AdK with an ATP analog and AMP (PDB ID: 1ANK[2]) is the most closed (white in Figure 3b), followed by the transition-state analog (TSA) structure with ADP, AlF_4^- and AMP (PDB ID: 3SR0;[8] light blue), and the structure with 2 ADP is the most open (PDB ID: 4JKY;[8] blue). In addition, B-factors are largest for amino acids around the ATPlid in the co-crystal structure with 2 ADP, as visualized by the thickness of the tube in Figure 3b, which indicates a large fluctuation of ATPlid residues.

To characterize the dynamics of the ATPlid along the catalytic reaction coordinate, principal component analysis (PCA)[76] was performed using the coordinates saved in the QM/MM simulations. In agreement with Figure 3b, the two lowest frequency motions identified from the analysis, respectively PC1 and PC2, occurred along the opening direction of the enzyme (Figure 3c and Figure S14). In addition, the projection of the MD simulation trajectories onto the two PCA modes showed a wider conformational space exploration in TSM and PS than in RS (Figure 3d). Interestingly, the projection of PS (and TSM) shows a bimodal distribution, in which the larger peak on the left corresponds to the more open ATPlid and the smaller one to the right to the more closed ATPlid conformation (Figure 3a and d). The QM/MM simulations performed with a restraint on the distance between the ATPlid and CORE subdomains, further showed a pseudo-linear dependence of the reaction free energy on the CORE-ATPlid distance, while the barrier was relatively insensitive (Figure S15). These results suggest that ATPlid dynamics differs between the reactant and product states. This difference in dynamics results in the more open conformation in PS (Figure 3a), as a consequence of the formation of different interactions of the arginine residues in the active site, and at the same time, influences the free energy landscape of the catalytic reaction.

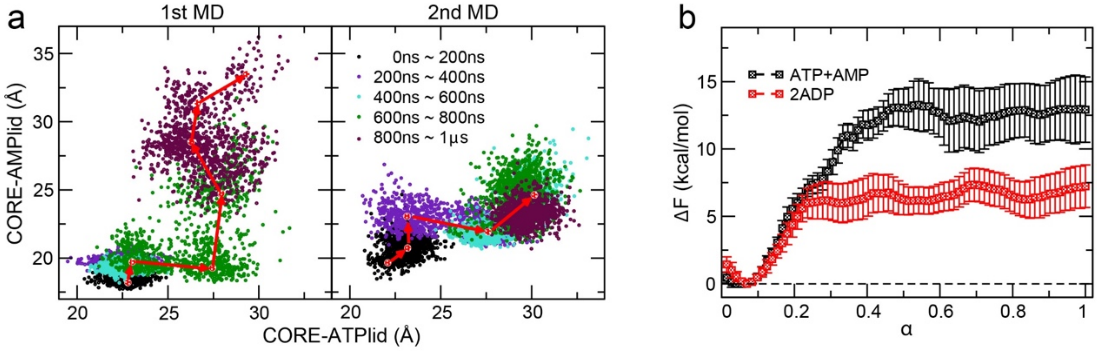


Figure 6. Close-to-open conformational change

a Changes of CORE-ATPlid and CORE-AMPlid distances in the two 1 μs MD simulations performed in the product state. The two distances are shown in different colors with an interval of 200 ns. **b** Free energy (FE) profile of the close-to-open conformational change described by the normalized reaction coordinate α between 0 (closed state) and 1 (open state). The FE profile shown in black is for opening in the reactant state (ATP+AMP) and in red for that in the product state (2 ADP), respectively.

The above results in turn suggest the possibility that the rearrangements of the active site residues and the resultant changes in their interactions trigger the facilitated opening of the enzyme in PS. To test this possibility, the MD simulations performed in RS and PS were extended to 1 μs. In the two MD simulations in PS, a spontaneous opening occurred within 1 μs, while no opening was observed in the RS MD simulations. Figure 6a shows the distributions of the CORE-ATPlid and CORE-AMPlid distances in different colors between the different time segments. In addition, we performed the string method free energy (FE) simulations[69] based on the conformational change path obtained in PS and determined the free energy change along the path optimized. In agreement with the MD simulations, the FE profiles show that the opening of the enzyme has a lower FE barrier in the product state than that in the reactant state (7.4±1.2 kcal/mol in PS vs 13.3±1.9 kcal/mol in RS; Figure 6b). This confirms that the catalytic reaction facilitates the opening of the enzyme in the product state. This is achieved by the rearrangements of the active site residues and the resultant changes in their interactions with the reaction substrates, affecting the free energy landscape of the close-to-open conformational transition. However, we must be careful when comparing the free energy of opening shown in Figure 6b to that determined by single molecule studies.[22, 24] Typically, in single molecule studies, the free energy difference between two conformational states is estimated based on their relative populations, while the FE profile shown in Figure 6b has a contribution of the ligand-enzyme interactions. This difference can be better understood when the free energy value based on the ligand dissociation constant (K_d) is compared with the value based on relative populations. In particular, due to the difference in the definition of the two equilibrium constants, the resulting free energy values differ in the contribution of ligand concentration, and for the same reason, the population-based free energy value depends on the ligand concentration and the ligand binding affinity to the enzyme.

Discussion

Dual role of the active site arginines

This study reveals how conformational heterogeneity around the five arginine residues in the active site of the enzyme adenylate kinase (AdK) facilitates the catalytic reaction. The arginine side chains reorganize rapidly along the catalytic reaction coordinate and, with their ability to form multiple hydrogen bonds, effectively stabilize the transition state of the reaction. For this reason, these residues are highly conserved among NMP kinases, including adenylate kinases[77] and any mutation thereof causes a substantial reduction in catalytic activity (Table S2). Importantly, AdK's R123 occupies a position very similar to that of the so called "arginine finger", R789$^{\mathrm{GAP}}$, of the Ras-RasGAP complex[78] (Figure S16). Based on this finding and also given that AdK inserts its catalytic arginine residues (in particular, R123, R156 and R167) into the active site through conformational change, these active site arginine residues are reminiscent of the arginine finger of RasGAP[78, 79] and other ATP/GTPases.[80-83]

This study also suggests that the active site arginines control the close-to-open conformational dynamics of AdK. As shown in Figure 3a, AdK explores a wider conformational space in the product state (PS) along the opening direction of the enzyme, which leads to a more favorable opening of the enzyme compared to the reactant state (RS). Our analysis of the QM/MM simulations further suggests that the enhanced motion of ATPlid in PS can be caused by the change of hydrogen-bonding interactions of the arginine residues in the active site of the enzyme along the catalytic reaction trajectory (Figures 2b and S4). For example, notable in Figure 2b, which show the rearrangement of the R123, R156 and R167 residues during the catalytic reaction, is the change in the relative orientation of the α6 and α7-helices. Due to the proximity of the three arginine residues to the two α-helices, in the case of R167 in the middle of the α7-helix, their rearrangements during the reaction lead to the change in the orientation of the two helices. In PS, in particular, these changes trigger the release of the loop connecting the α3 and α4-helices of AMPlid relative to the α7-helix. This release then decouples AMPlid from ATPlid, and finally allows the opening of the enzyme. This sequence of events is consistent with that observed from the MD simulations in PS: i.e., a slight opening of AMPlid occurs first, followed by the opening of ATPlid (Figure 6a). If this is the case, the removal of the side chain of these residues, such as, by mutation, would result in an increased motion of ATP and AMPlids and eventual opening of the enzyme. This was indeed observed from the MD simulations of the R167A mutant in both RS and PS (Figure S5c). Combined with the impact on the phosphoryl transfer barrier (Figure S5a-b), this explains the more than 99% reduction in catalytic activity. We expect similar results with the R123A mutant, due to the orientation and interactions formed by the R123 side chain in both RS and PS and the impact of their change on in the orientation of the α6-helix (Figure 2b).

Finally, we note that the change in the time scale of conformational motion in response to the catalytic reaction has been reported in other enzymes, such as dihydrofolate reductase (DHFR)[84] and triosephosphate isomerase (TIM).[85, 86] In DHFR, in addition, a variation in the conformation of active site residues was observed across the catalytic cycle of the enzyme,[87] suggesting the possible generality of the mechanism presented in this study. In these enzymes, however, the region of conformation change is rather localized around the active site of the enzyme compared to that of adenylate kinase.

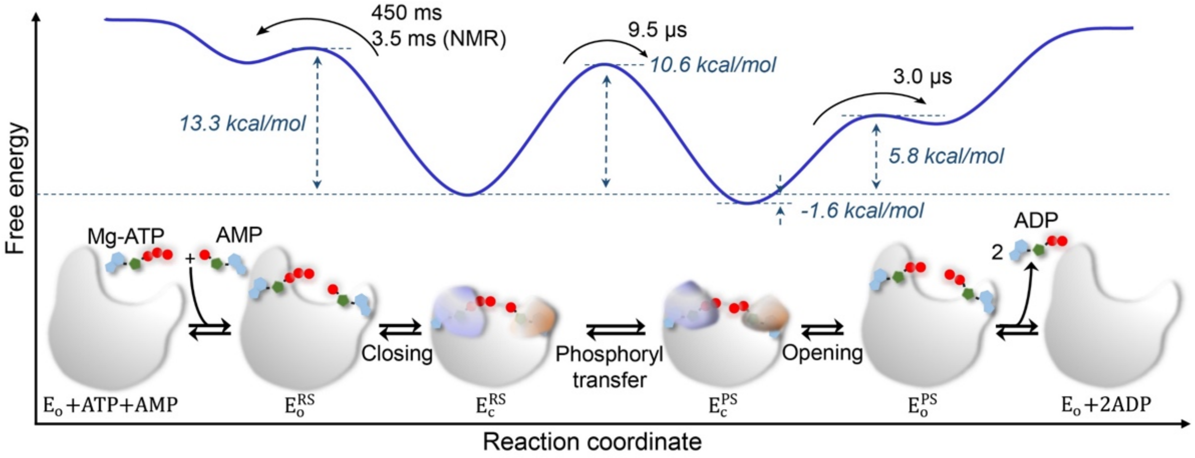


Figure 7. Free energy landscape of AdK along the entire catalytic cycle and schematic representation of corresponding states

Indicated in the figure is the time estimated to cross each barrier: the mean first passage time in the conformational change and the time based on the transition state theory in the chemical reaction step. For the opening in the reactant sate (i.e., $E_c^{RS} \rightarrow E_o^{RS}$), the time determined based on the NMR experiment[3] is also provided.

Implication to the rate-limiting step

The difference in the opening barrier between RS and PS has an important implication in the mechanism of the enzyme. That is, the slowest step of the reaction can be different between the forward and backward reactions (Figure 7). In the forward reaction (i.e., ATP+AMP → 2ADP), the chemical step takes 9.47 µs based on the transition state theory (TST) with $\Delta F^{\ddagger}$ = 10.6 kcal/mol (and < 0.2 ms based on the experimental $\Delta F^{\ddagger}$ of < 12.4 kcal/mol[8]), while the opening step takes 3.0 µs based on the mean first passage time (MFPT) calculation.[88] Therefore, the slowest step in the forward direction is the chemical step. In contrast, it is the opening of the enzyme in the reverse reaction (450 ms MFPT). In this comparison, however, the following is worth noting. First, we used TST to estimate the time of the catalytic reaction (i.e., the rate of the catalytic reaction). In TST, since the reaction is assumed to take place only when the reactants have reached the so-called high energy transition state, the estimated reaction time includes the time that the enzyme spends waiting for the transition state conformation, which is the majority, while the actual reaction occurs on the time scale of the scissile/forming bond vibration at TS. Second, the MFPT rate in RS is 2 orders of magnitude larger than the opening rate measured by NMR (i.e., 3.5 ms).[3] This severe overestimation may be due to errors in the free energy profiles (Figure 6b) and/or limitations of the MFPT method and approximation.[88] For this reason, the precise determination of the rate of slow conformational change has remained a theoretical challenge. Nevertheless, the MFPT-based opening time in PS is comparable to the opening timescale (0.5 ~ 1.0 µs) directly observed from the MD simulations (Figure 6a).

At first glance, the difference in the slowest steps seems incompatible with the rate-limiting conformational change mechanism proposed based on the NMR experiments.[3] The NMR technique applied, however, only probed millisecond protein motions and thus provided no information about protein motions on a faster timescale. Furthermore, recent single-molecule experiment[24] and simulations[89, 90] have shown that the opening can occur on the range of 1.8 µs ~ 77 µs. In contrast, in other single-molecule study,[22] the 6.3 ms opening was observed in the presence of substrate analog (i.e., AMP-PNP and AMP). While this result is consistent with the NMR result, it can be limited by the time resolution of the single-molecule experiment to detect a faster timescale event. Taken together, the two vastly different opening timescales we discovered in this study are consistent with the previously published experimental results and provide a reconciliation between the different sets of experiments.

Conclusion

In the present study, we have applied a combination of multiscale quantum mechanical and classical mechanical simulations, NMR and X-ray crystallography to investigate how the enzyme adenylate kinase achieves a mechanistic coupling between enzymatic catalysis and µs-ms timescale collective protein motions. To achieve the coupling, the enzyme employs the ps-ns side chain motions, which occur locally at the active site, to influence both the catalytic reaction and the slow conformational change of the enzyme. In this way, the two events are connected, thus overcoming the disparity in their time scales. In motor proteins, such as F_1-ATPase, a similar mechanism is adopted to regulate the timing of the catalytic reaction and large scale conformation changes.

References

[1] Müller, C., Schlauderer, G., Reinstein, J., and Schulz, G. (1996) Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding, *Structure 4*, 147-156.

[2] Berry, M. B., Meador, B., Bilderback, T., Liang, P., MGlaser, M., and Phillips, G. N., Jr. (1994) The closed conformation of a highly flexible protein: The structure of E. coli adenylate kinase with bound AMP and AMPPNP, *Proteins 19*, 183-198.

[3] Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., and Kern, D. (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair, *Nat. Struct. Mol. Biol. 11*, 945-949.

[4] Benkovic, S. J., and Hammes-Schiffer, S. (2003) A Perspective on Enzyme Catalysis, *Science 301*, 1196-1202.

[5] Henzler-Wildman, K. A., Lei, M., Thai, V., Kerns, S. J., Karplus, M., and Kern, D. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis, *Nature 450*, 913-916.

[6] Petrović, D., Risso, V. A., Kamerlin, S. C. L., and Sanchez-Ruiz, J. M. (2018) Conformational dynamics and enzyme evolution, *J. R. Soc. Interface 15*, 20180330.

[7] Agarwal, P. K. (2019) A Biophysical Perspective on Enzyme Catalysis, *Biochemistry 58*, 438-449.

[8] Kerns, S. J., Agafonov, R. V., Cho, Y.-J., Pontiggia, F., Otten, R., Pachov, D. V., Kutter, S., Phung, L. A., Murphy, P. N., Thai, V., Alber, T., Hagan, M. F., and Kern, D. (2015) The energy landscape of adenylate kinase during catalysis, *Nat. Struct. Mol. Biol. 22*, 124-131.

[9] Schwartz, S. D. (2013) Protein Dynamics and the Enzymatic Reaction Coordinate, *Top. Curr. Chem. 337*, 189-208.

[10] Hay, S., and Scrutton, N. S. (2012) Good vibrations in enzyme-catalysed reactions, *Nat. Chem. 4*, 161-168.

[11] Hanoian, P., Liu, C. T., Hammes-Schiffer, S., and Benkovic, S. (2015) Perspective on Electrostatics and Conformational Motions in Enzyme Catalysis, *Acc. Chem. Res. 48*, 482-489.

[12] Warshel, A., and Bora, R. P. (2016) Perspective: Defining and quantifying the role of dynamics in enzyme catalysis, *J. Chem. Phys. 144*, 180901.

[13] Cheatum, C. M. (2020) Low-Frequency Protein Motions Coupled to Catalytic Sites, *Annu. Rev. Phys. Chem. 71*, 267-288.

[14] Yang, L.-W., and Bahar, I. (2005) Coupling between Catalytic Site and Collective Dynamics: A Requirement for Mechanochemical Activity of Enzymes, *Structure 13*, 893-904.

[15] Eisenmesser, E. Z., Bosco, D. A., Akke, M., and Kern, D. (2002) Enzyme Dynamics During Catalysis, *Science 295*, 1520-1523.

[16] Boehr, D. D., Dyson, H. J., and Wright, P. E. (2006) An NMR Perspective on Enzyme Dynamics, *Chem. Rev. 106*, 3055-3079.

[17] Masterson, L. R., Shi, L., Metcalfe, E., Gao, J., Taylor, S. S., and Veglia, G. (2011) Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy, *Proc. Natl. Acad. Sci. USA 108*, 6969-6974.

[18] Callender, R., and Dyer, R. B. (2015) The Dynamic Nature of Enzymatic Catalysis, *Acc. Chem. Res. 48*, 407-413.

[19] Tükenmez, H., Magnussen, H. M., Kovermann, M., Byström, A., and Wolf-Watz, M. (2016) Linkage between Fitness of Yeast Cells and Adenylate Kinase Catalysis, *PLOS ONE 11*, e0163115.

[20] Rundqvist, L., Ådén, J., Sparrman, T., Wallgren, M., Olsson, U., and Wolf-Watz, M. (2009) Noncooperative Folding of Subdomains in Adenylate Kinase†, *Biochemistry 48*, 1911-1927.

[21] Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A., Karplus, M., Hübner, C. G., and Kern, D. (2007) Intrinsic motions along an enzymatic reaction trajectory, *Nature 450*, 838-844.

[22] Hanson, J. A., Duderstadt, K., Watkins, L. P., Bhattacharyya, S., Brokaw, J., Chu, J.-W., and Yang, H. (2007) Illuminating the mechanistic roles of enzyme conformational dynamics, *Proc. Natl. Acad. Sci. USA 104*, 18055-18060.

[23] Kovermann, M., Grundström, C., Sauer-Eriksson, A. E., Sauer, U. H., and Wolf-Watz, M. (2017) Structural basis for ligand binding to an enzyme by a conformational selection pathway, *Proc. Natl. Acad. Sci. USA 114*, 6298-6303.

[24] Aviram, H. Y., Pirchi, M., Mazal, H., Barak, Y., Riven, I., and Haran, G. (2018) Direct observation of ultrafast large-scale dynamics of an enzyme under turnover conditions, *Proc. Natl. Acad. Sci. USA 115*, 3243-3448.

[25] Arora, K., and Brooks III, C. L. (2007) Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism, *Proc. Natl. Acad. Sci. USA 104*, 18496-18501.

[26] Lu, Q., and Wang, J. (2008) Single Molecule Conformational Dynamics of Adenylate Kinase: Energy Landscape, Structural Correlations, and Transtion State Ensembles, *J. Am. Chem. Soc. 130*, 4772-4783.

[27] Zeller, F., and Zacharias, M. (2015) Substrate Binding Specifically Modulates Domain Arrangements in Adenylate Kinase, *Biophys. J. 109*, 1978-1985.

[28] Reinstein, J., Schlichting, I., and Wittinghofer, A. (1990) Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of Escherichia coli, *Biochemistry 29*, 7451-7459.

[29] Kim, H. J., Nishikawa, S., Tokutomi, Y., Takenaka, H., Hamada, M., Kuby, S. A., and Uesugi, S. (1990) In vitro mutagenesis studies at the arginine residues of adenylate kinase. A revised binding site for AMP in the x-ray-deduced model, *Biochemistry 29*, 1107-1111.

[30] Yan, H., Dahnke, T., Zhou, B., Nakazawa, A., and Tsai, M. D. (1990) Mechanism of Adenylate Kinase. Critical Evaluation of the X-ray Model and Assignment of the AMP Site, *Biochemistry 29*, 10956-10964.

[31] Dahnke, T., Shi, Z., Yan, H., Jiang, R. T., and Tsai, M. D. (1992) Mechanism of adenylate kinase. Structural and functional roles of the conserved arginine-97 and arginine-132, *Biochemistry 31*, 6318-6328.

[32] Zeymer, C., Werbeck, N. D., Zimmermann, S., Reinstein, J., and Hansen, D. F. (2016) Characterizing Active Site Conformational Heterogeneity along the Trajectory of an Enzymatic Phosphoryl Transfer Reaction, *Angew. Chem. Int. Ed. 55*, 11533-11537.

[33] Bellinzoni, M., Haouz, A., Graña, M., Munier-Lehmann, H., Shepard, W., and Alzari, P. M. (2006) The crystal structure of Mycobacterium tuberculosis adenylate kinase in complex with two molecules of ADP and Mg2+ supports an associative mechanism for phosphoryl transfer, *Protein Sci. 15*, 1489-1493.

[34] Lassila, J. K., Zalatan, J. G., and Herschlag, D. (2011) Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis, *Annu. Rev. Biochem. 80*, 669-702.

[35] Pisliakov, A. V., Cao, J., Kamerlin, S. C. L., and Warshel, A. (2009) Enzyme millisecond conformational dynamics do not catalyze the chemical step, *Proc. Natl. Acad. Sci. USA 106*, 17359-17364.

[36] Kamerlin, S. C. L., and Warshel, A. (2010) At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis?, *Proteins 78*, 1339-1375.

[37] Antoniou, M., and Schwartz, S. D. (2001) Inernal Enzyme Motions as a Source of Catalytic Activity: Rate-Promoting Vibrations and Hydrogen Tunneling, *J. Phys. Chem. B 105*, 5553-5558.

[38] Nam, K., and Karplus, M. (2019) Insights into the origin of the high energy-conversion efficiency of F_1-ATPase, *Proc. Natl. Acad. Sci. USA 116*, 15924-15929.

[39] Reinstein, J., Brune, M., and Wittinghofer, A. (1988) Mutations in the nucleotide binding loop of adenylate kinase of Escherichia coli, *Biochemistry 27*, 4712-4720.

[40] Kabsch, W. (2010) XDS, *Acta Crystallogr. D Biol. Crystallogr. 66*, 125-132.

[41] Evans, P. R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics, *Acta Crystallogr. D Biol. Crystallogr. 67*, 282-292.

[42] Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution?, *Acta Crystallogr. D Biol. Crystallogr. 69*, 1204-1214.

[43] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilsonc, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr. D Biol. Crystallogr. 67*, 235-242.

[44] Müller, C. W., and Schulz, G. E. (1992) Structure of the complex between adenylate kinase from Escherichia coli and the inhibitor Ap5A refined at 1.9 Å resolution: A model for a catalytic transition state, *J. Mol. Biol. 224*, 159-177.

[45] 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. D Biol. Crystallogr. 66*, 213-221.

[46] Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P., and Adams, P. D. (2010) Joint X-ray and neutron refinement with phenix.refine, *Acta Crystallogr. D Biol. Crystallogr. 66*, 1153-1163.

[47] Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine, *Acta Crystallogr. D Biol. Crystallogr. 68*, 352-367.

[48] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr. D Biol. Crystallogr. 66*, 486-501.

[49] Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277-293.

[50] Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology 2 Domain Studied by 15N NMR Relaxation, *Biochemistry 33*, 5984-6003.

[51] Dayie, K. T., and Wagner, G. (1994) Relaxation-Rate Measurements for 15N−1H Groups with Pulsed-Field Gradients and Preservation of Coherence Pathways, *J. Magn. Res., Series A 111*, 121-126.

[52] MacKerell, A. D., Bashford, D., Bellott, Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †, *J. Phys. Chem. B 102*, 3586-3616.

[53] MacKerell, Feig, M., and Brooks, C. L. (2004) Improved Treatment of the Protein Backbone in Empirical Force Fields, *J. Am. Chem. Soc. 126*, 698-699.

[54] MacKerell, A. D., Banavali, N., and Foloppe, N. (2000) Development and current status of the CHARMM force field for nucleic acids, *Biopolymers 56*, 257-265.

[55] Foloppe, N., and MacKerell, J. A. D. (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, *J. Comput. Chem. 21*, 86-104.

[56] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926-935.

[57] Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes, *J. Comput. Phys. 23*, 327-341.

[58] Brooks, B. R., Brooks, C. L., III, MacKerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The Biomolecular Simulation Program, *J. Comput. Chem. 30*, 1545-1614.

[59] Eastman, P., Swails, J., D. Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L.-P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R., and Pande, V. S. (2017) OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics, *PloS Comput. Biol. 13*, e1005659.

[60] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method, *J. Chem. Phys. 103*, 8577-8593.

[61] Nam, K., Cui, Q., Gao, J., and York, D. M. (2007) Specific Reaction Parametrization of the AM1/d Hamiltonian for Phosphoryl Transfer Reactions: H, O, and P Atoms, *J. Chem. Theory Comput. 3*, 486-504.

[62] Nam, K., Gao, J., and York, D. M. (2005) An Efficient Linear-Scaling Ewald Method for Long-Range Electrostatic Interactions in Combined QM/MM Calculations, *J. Chem. Theory Comput. 1*, 2-13.

[63] Nam, K. (2014) Acceleration of ab initio QM/MM calculations under periodic boundary conditions by multiscale and multiple time step approaches, *J. Chem. Theory Comput. 10*, 4175-4183.

[64] Ojeda-May, P., and Nam, K. (2017) Acceleration of Semiempirical QM/MM Methods through Message Passing Interface (MPI), Hybrid MPI/Open Multiprocessing, and Self-Consisten Field Accelerator Implementations, *J. Chem. Theory Comput. 13*, 3325-3536.

[65] Hoover, W. G. (1985) Canonical dynamics: Equilbrium phase-space distributions, *Phys. Rev. A 31*, 1695-1697.

[66] Feller, S. E., Zhang, Y., and Pastor, R. W. (1995) Constant pressure molecular dynamics simulation: The Langevin piston method, *J. Chem. Phys. 103*, 4613-4621.

[67] Torrie, G. M., and Valleau, J. P. (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *J. Comput. Phys. 23*, 187-199.

[68] Vanden-Eijnden, E., and Venturoli, M. (2009) Revisiting the finite temperature string method for the calculation of reaction tubes and free energies, *J. Chem. Phys. 130*, 194103.

[69] Ovchinnikov, V., Karplus, M., and Vanden-Eijnden, E. (2011) Free energy of conformational transition paths in biomolecules: The string method and its application to myosin VI, *J. Chem. Phys. 134*, 085103.

[70] Gao, J. (1996) Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials, In *Rev. Comput. Chem.* (Lipkowitz, K. B., and Boyd, D. B., Eds.), pp 119-185, John Wiley & Sons, Inc.

[71] Åden, J., and Wolf-Watz, M. (2007) NMR Identification of Transient Complexes Critical to Adenylate Kinase Catalysis, *J. Am. Chem. Soc. 129*, 14003-14012.

[72] Brown, A. (2009) Analysis of Cooperativity by Isothermal Titration Calorimetry, *Int. J. Mol. Sci. 10*, 3457-3477.

[73] Berry, M. B., and Phillips, G. N. (1998) Crystal structures of Bacillus stearothermophilus adenylate kinase with bound Ap5A, Mg2+ Ap5A, and Mn2+ Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg2+ and Mn2+, *Proteins 32*, 276-288.

[74] Rogne, P., Sparrman, T., Anugwom, I., Mikkola, J.-P., and Wolf-Watz, M. (2015) Realtime 31P NMR Investigation on the Catalytic Behavior of the Enzyme Adenylate kinase in the Matrix of a Switchable Ionic Liquid, *ChemSusChem 8*, 3764-3768.

[75] Rogne, P., Rosselin, M., Grundström, C., Hedberg, C., Sauer, U. H., and Wolf-Watz, M. (2018) Molecular mechanism of ATP versus GTP selectivity of adenylate kinase, *Proc. Natl. Acad. Sci. USA 115*, 3012-3017.

[76] Pearson, K. (1901) LIII. On lines and planes of closest fit to systems of points in space, *Philos. Mag. Series 6 2*, 559-572.

[77] Szilágyi, A., and Závodszky, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey, *Structure 8*, 493-504.

[78] Scheffzek, K., Ahmadian, M. R., Kabsch, W., Wiesmüller, L., Lautwein, A., Schmitz, F., and Wittinghofer, A. (1997) The Ras-RasGAP Complex: Structural Basis for GTPase Activation and Its Loss in Oncogenic Ras Mutants, *Science 277*, 333-339.

[79] Ahmadian, M. R., Stege, P., Scheffzek, K., and Wittinghofer, A. (1997) Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras, *Nat. Struct. Biol. 4*, 686-689.

[80] Hoff, R. H., Wu, L., Zhou, B., Zhang, Z.-Y., and Hengge, A. C. (1999) Does Positive Charge at the Active Sites of Phosphatases Cause a Change in Mechanism? The Effect of the Conserved Arginine on the Transition State for Phosphoryl Transfer in the Protein-Tyrosine Phosphatase from Yersinia, *J. Am. Chem. Soc. 121*, 9514-9521.

[81] Gorrell, A., Lawrence, S. H., and Ferry, J. G. (2005) Structural and Kinetic Analyses of Arginine Residues in the Active Site of the Acetate Kinase from Methanosarcina thermophila, *J. Biol. Chem. 280*, 10731-10742.

[82] Hanson, P. I., and Whiteheart, S. W. (2005) AAA+ proteins: have engine, will work, *Nat. Rev. Mol. Cell Biol. 6*, 519-529.

[83] Mann, D., Teuber, C., Tennigkeit, S. A., Schröter, G., Gerwert, K., and Kötting, C. (2016) Mechanism of the intrinsic arginine finger in heterotrimeric G proteins, *Proc. Natl. Acad. Sci. USA 113*, E9041-E8050.

[84] Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2010) Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands, *Proc. Natl. Acad. Sci. USA 107*, 1373-1378.

[85] Rozovsky, S., Jogl, G., Tong, L., and McDermott, A. E. (2001) Solution-state NMR Investigations of Triosephosphate Isomerase Active Site Loop Motion: Ligand Release in Relation to Active Site Loop Dynamics, *J. Mol. Biol. 310*, 271-280.

[86] Desamero, R., Rozovsky, S., Zhadin, N., McDermott, A., and Callender, R. (2003) Active Site Loop Motion in Triosephosphate Isomerase: T-Jump Relaxation Spectroscopy of Thermal Activation, *Biochemistry 42*, 2941-2951.

[87] Tuttle, L. M., Dyson, H. J., and Wright, P. E. (2013) Side-Chain Conformational Heterogeneity of Intermediates in the *Escherichia coli* Dihydrofolate Reductase Catalytic Cycle, *Biochemistry 52*, 3464-3477.

[88] Ovchinnikov, V., Nam, K., and Karplus, M. (2016) A Simple and Accurate Method to Calculate Free Energy Profiles and Reaction Rates from Restrained Molecular Simulatoins of Diffusive Processes, *J. Phys. Chem. B 120*, 8457-8472.

[89] Wang, Y., Gan, L., Wang, E., and Wang, J. (2013) Exploring the Dynamic Functional Landscape of Adenylate Kinase Modulated by Substrates, *J. Chem. Theory Comput. 9*, 84-95.

[90] Zheng, Y., and Cui, Q. (2018) Multiple Pathways and Time Scales for Conformational Transitions in apo-Adenylate Kinase, *J. Chem. Theory Comput. 14*, 1716-1726.

Appendix II

Table S1. Crystallographic data collection and refinement statistics.

	AdK in complex with 2 ADP
Data collection	Beam line ID23-1 ESRF, France
Wavelength [Å]	0.97498
Space group (Nr.)	$P2_1\,2_1\,2$ (18)
Unit cell dimensions	
a, b, c [Å]	72.75, 82.23, 78.78
α, β, γ [°]	90.0, 90.0, 90.0
Molecules per asymmetric unit	2
Resolution range (highest shell) [Å]	44.81 – 1.36 (1.41 – 1.36)
Accepted reflections	730201 (72337)
Unique reflections	99726 (9898)
Completeness [%]	97.8 (95.8)
Multiplicity	7.3 (7.3)
Mean (I/σI)	16.0 (1.90)
R_{merge} [%]	6.2 (99.7)
R_{meas} [%]	6.6 (107.3)
R_{pim} [%]	2.5 (39.7)
$CC_{1/2}$ [%]	99.8 (69.2)
Wilson B-factor	16.6
Refinement	
Resolution range (highest shell) [Å]	39.39 - 1.36 (1.38 – 1.36)
Reflections used for refinement	99720 (2995)
Reflections used for R_{free}	5028 (169)
R_{work} [%]	17.7 (30.1)
R_{free} [%]	19.4 (34.3)
Overall map CC (Fc, 2mFo-DFc) [%]	89.4
R. m. s. deviations	
Bond length [Å]	0.010
Bond angles (°)	1.30
Ramanchandran plot statistics	
Residues in favored region [%]	99.76
Residues in allowed region [%]	0.24
Residues in outlier region [%]	0.0
All-atom clash score	2.1
No. of non-hydrogen atoms	3506
Protein (non-hydrogen)	3398
2 ADP	108
Solvent (H_2O) (non-hydrogen)	594
Average B-factors ($Å^2$):	21.6
Protein chains	20.6
Water molecules	28.8

Table S2. Michaelis-Menten parameters of the wild-type AdK and its variants measured from the present study.

AdK variant	metal	k_{cat} (s^{-1})	$K_{M,ATP}$ (μM)[a]
Wild-type	Mg^{2+}	330 ± 11	71 ± 7
	None	$7.5\times10^{-3} \pm 7.5\times10^{-4}$ [b]	
R36A	Mg^{2+}	55 ± 2	89 ± 7
R88A	Mg^{2+}	1.8 ± 0.04	120 ± 11
R123A	Mg^{2+}	0.28 ± 0.01	110 ± 16
R156K[c]	Mg^{2+}	0.74 ± 0.01	52 ± 5
R167A	Mg^{2+}	2.4 ± 0.05	56 ± 4

[a] All reported K_M values are for the binding of ATP in the presence of saturating AMP concentration.

[b] The k_{cat} estimated from NMR by Kerns at al. was < 1 s^{-1}, while the opening rate of lids was estimated to 0.05±0.01 s^{-1}.[18]

[c] Kerns et al. reported the rate of the catalytic reaction: The forward chemical reaction rate was 0.7± 0.2 s^{-1} from the quenched-flow experiment and the rate of lid opening was 300±75 s^{-1} based on the NMR relaxation data, respectively.[18]

Table S3. Hydrogen-bond distances formed between the ATP and AMP phosphoryl groups and the catalytic arginine residues, in the reactant (RS), transition (TS) and product (PS) states. All values are obtained from the QM/MM simulation of the wild-type AdK. All distances are in the unit of Å.

	RS	TS	PS
ATP $O_{2\alpha}$ ---- R123 N_{H1}	2.752 ± 0.092	2.779 ± 0.102	2.767 ± 0.099
ATP $O_{3\beta}$ ---- R123 N_{H1}	2.938 ± 0.129	2.892 ± 0.122	2.771 ± 0.100
ATP $O_{1\gamma}$ ---- R123 N_{H2}	2.843 ± 0.121	2.787 ± 0.097	2.801 ± 0.108
ATP $O_{1\gamma}$ ---- R156 N_{H2}	3.196 ± 0.220	3.797 ± 0.257	3.252 ± 0.336
AMP $O_{1\alpha}$ ---- R156 N_{H1}	2.868 ± 0.167	3.099 ± 0.292	3.389 ± 0.429
AMP $O_{1\alpha}$ ---- R156 N_{H2}	2.822 ± 0.114	2.775 ± 0.106	4.302 ± 0.446
ATP $O_{1\gamma}$ ---- R167 N_{H2}	2.977 ± 0.198	2.815 ± 0.102	2.770 ± 0.099
AMP $O_{3\alpha}$ ---- R167 N_{H2}	2.776 ± 0.088	3.286 ± 0.197	4.321 ± 0.259

Table S4.Thermodynamic parameters for ADP binding to *E. coli* AdK determined by ITC at 298.15 K.

n	K_d (μM)	ΔH^0 (kcal/mol)	$-T\Delta S^0$ (kcal/mol)	ΔG^0 (kcal/mol)
0.9 ± 0.08	9.0 ± 0.7	-6.76 ± 0.9	-0.7 ± 0.5	-7.4 ± 1.1

Table S5. Diffusion tensors derived from the ^{15}N relaxation data in different AdK states.

AdK state	τ_c (ns)	α (deg)	β (deg)	$D_x = D_y$ (s^{-1})	D_z (s^{-1})
Apo	13.77(±0.14)	102(±32)	122(±16)	$1.10(\pm0.03)\times10^7$	$1.43(\pm0.04)\times10^7$
ADP bound	11.36(±0.06)	114(±70)	52(±24)	$1.43(\pm0.03)\times10^7$	$1.55(\pm0.03)\times10^7$
Ap5A inhibited	11.38(±0.05)	13(±27)	24(±9)	$1.52(\pm0.03)\times10^7$	$1.35(\pm0.05)\times10^7$

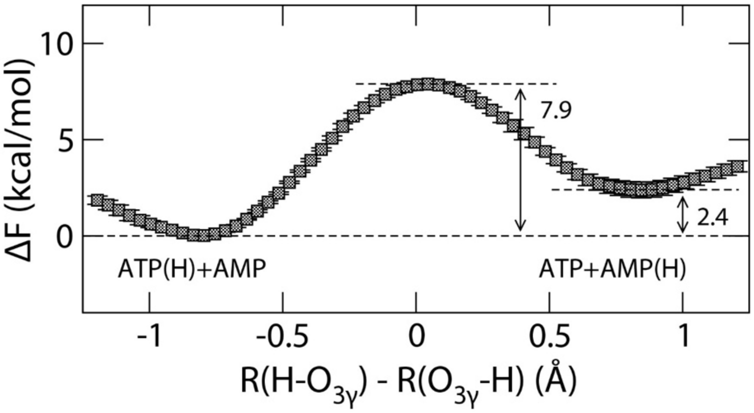


Figure S1. Umbrella sampling (US) free energy profile of the proton transfer of the wild-type AdK between ATP and protonated AMP to yield protontated ATP and AMP, which was the reactant state of the phosphoryl transfer reaction studied in the Main text. In each free energy value, the error is also indicated.

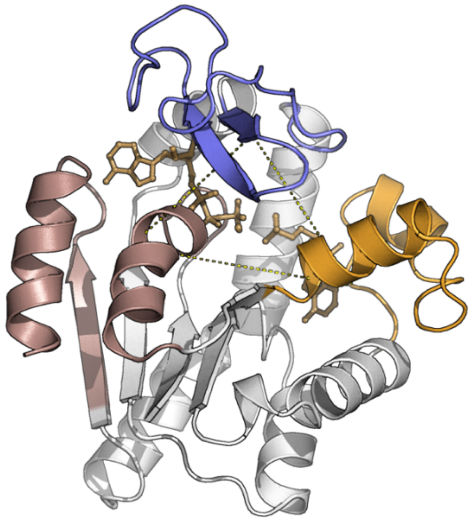


Figure S2. Definition of collective variables (CVs) for the close-to-open conformational change. The residues used to define the center of mass of the ATPlid, AMPlid and CORE subdomains (dark salmon) are shown in blue, orange and dark salmon, respectively, and their distances (i.e., CVs) are indicated with dashed yellow lines.

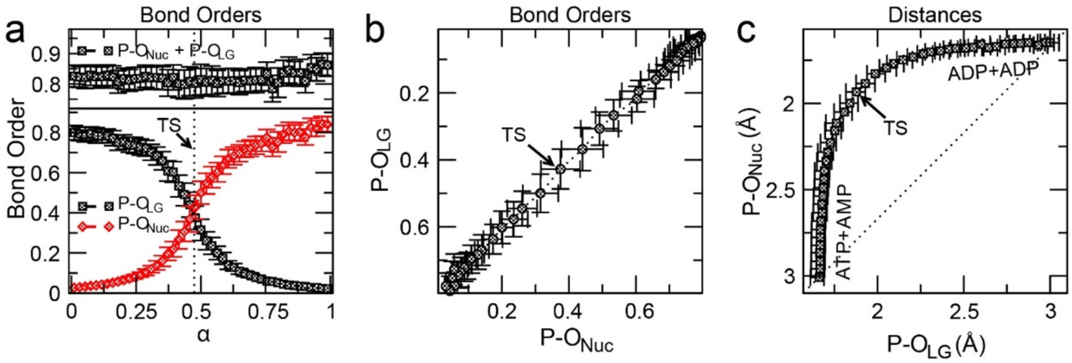


Figure S3. a Changes in P-O bond orders along the catalytic reaction. In the lower panel, the black and red lines represent the bond order of the cleaving and forming P-O bonds, denoted as P-O_{LG} and P-O_{Nuc}, respectively. The upper panel shows the sum of the two bond orders along the reaction. **b-c** Two-dimensional (2D) bond order and distance diagrams of the forming and cleaving P-O bonds. In (**c**), the change of the two bond orders along the diagonal line indicates the concerted reaction between the P-O bond cleavage and formation.

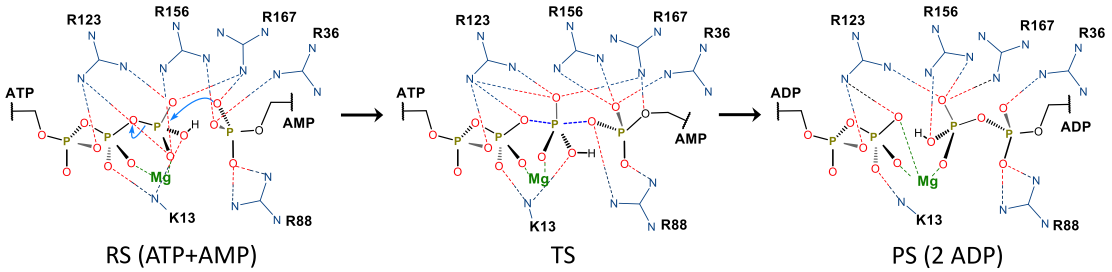


Figure S4. Reaction scheme of *E. coli* AdK derived from QM/MM simulation. Changes of the active site arginine and lysine residues are shown along with their respective hydrogen-bonding interactions with the reaction substrates. For clarity, only heavy atoms are shown, except for the proton of the γ-phosphoryl group.

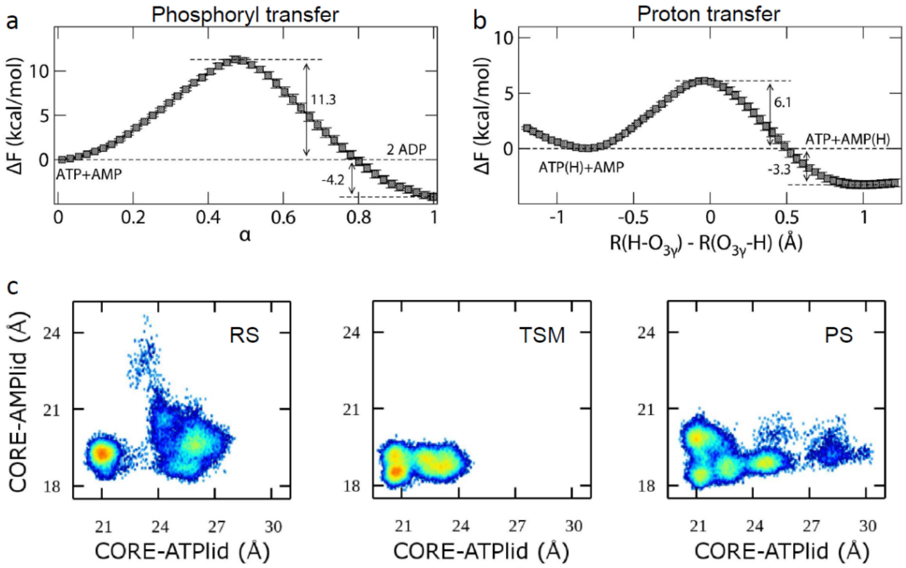


Figure S5. The free energy profiles of the R167A mutant determined from the QM/MM simulations: (**a**) the phosphoryl transfer between (protonated) ATP and AMP and (**b**) the proton transfer between ATP and protonated AMP (denoted as AMP(H)) to yield protontated ATP (denoted as ATP(H)) and AMP. In the proton transfer step, ATP and protonated AMP are favored by 3.3 kcal/mol compared to protonated ATP and AMP, which is the reactant state of the phospohryl transfer in (**a**). Thus, the barrier of the catalytic reaction is estimated as the sum of the phosphoryl transfer barrier (11.3 kcal/mol) and the free energy of proton transfer to ATP(H) and AMP (3.3 kcal/mol) to take into account the energy penalty to protonate ATP. (**c**) Distribution of CORE-ATPlid and CORE-AMPlid distances for the reactant (RS), transition-state mimic (TSM) and product state (PS) from the (100 ns) R167A mutant MD simulations. The color varies from blue to red with increasing frequency of the observed distances.

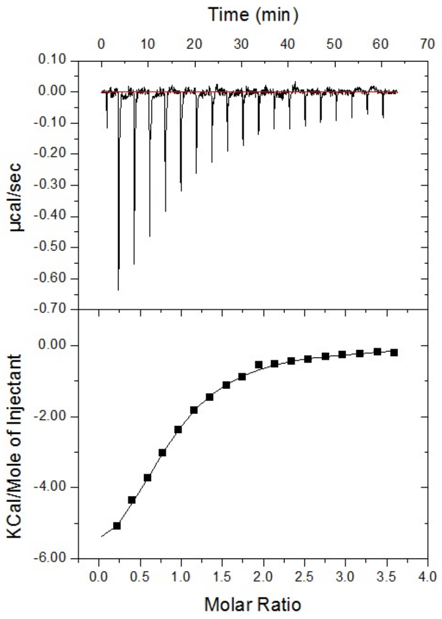


Figure S6. ITC titration isotherm of ADP binding to *E. coli* AdK.

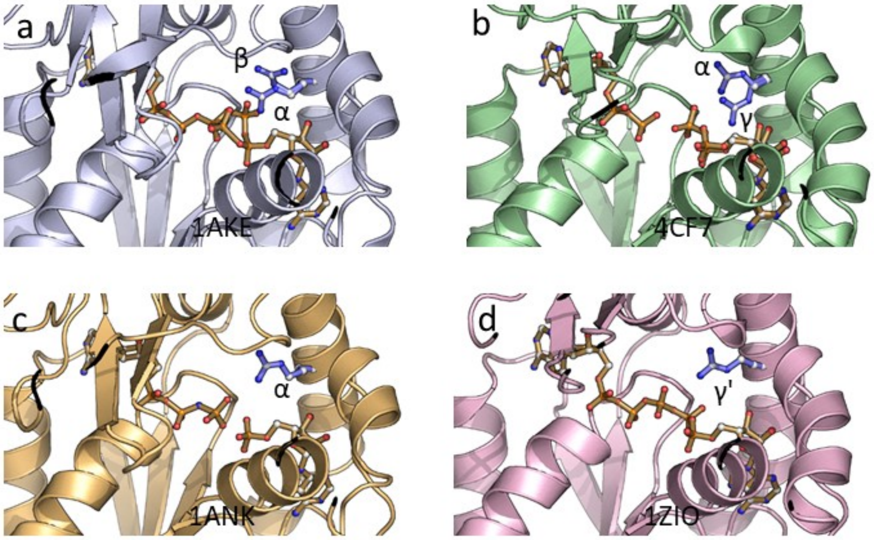


Figure S7. R167 side chain orientations in different X-ray structures of AdK: (**a**) 1AKE,[9] (**b**) 4CF7,[18] (**c**) 1ANK[31] and (**d**) 1ZIO.[12] The α and β orientations indicated are defined in **Figure 4**, and the γ and γ' orientations are shown in (**b**) and (**d**). The different Arg side chain orientations are defined based on their respective δ angles: The α and β orientations in (**a**) have 138.7° and -87.7° δ angles; the α and γ orientations in (**b**) have 138.7° and -87.7° δ angles; the α orientation in (**c**) shows 165.4°; and the γ' orientation in (**d**) has -161.7° δ angle, respectively. In (**a**) and (**b**), the different orientations are related with the different orientations of bound ligands, e.g., Ap5A in (**a**), and ADP and AMP in (**b**).

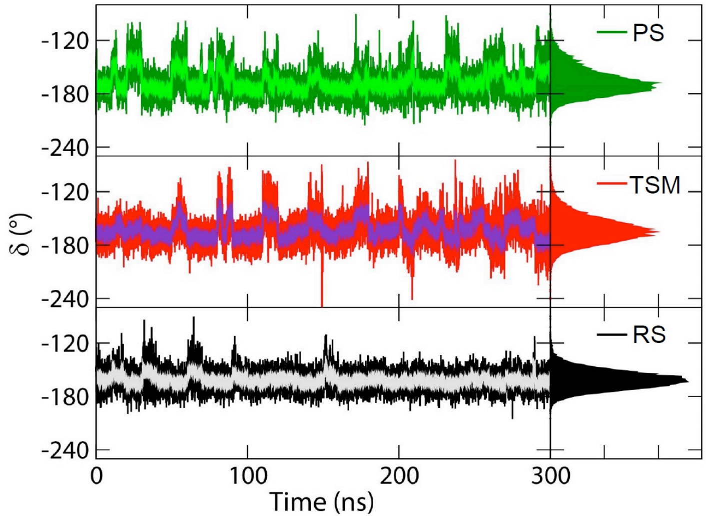


Figure S8. Time course of the R167 δ angle in the reactant (RS), transition-state mimic (TSM) and product states (PS), determined from the 300 ns MD simulations. In RS, the δ angle remains close to -163°, while in TSM and PS, the δ angle fluctuates more frequently toward smaller angles, thus exhibiting a broader angle distribution.

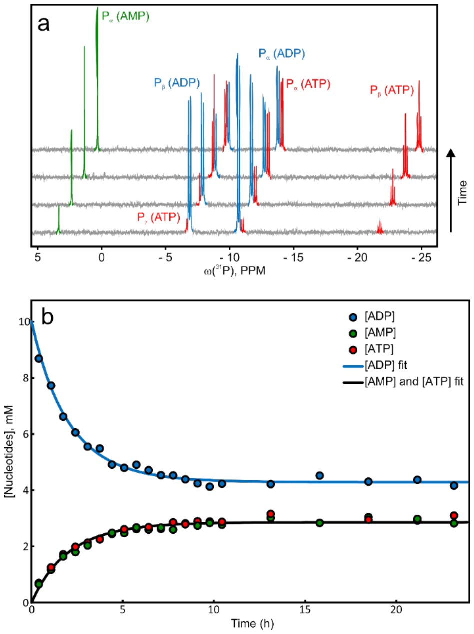


Figure S9. Real time ^{31}P NMR activity assay for quantification of residual activity of AdK in the absence of Mg^{2+}. **a** 1D ^{31}P NMR experiments were run at equal intervals in the absence of Mg^{2+} and with [ADP] = 10 mM. The area under each peak is directly proportional to the concentration of each phosphorous species: AMP shown in green: ADP in blue; and ATP in red. **b** Integrated data for each species and their fit using an analytical function[1] to yield the initial reaction rate (k_{cat}) and the equilibrium constant (K).

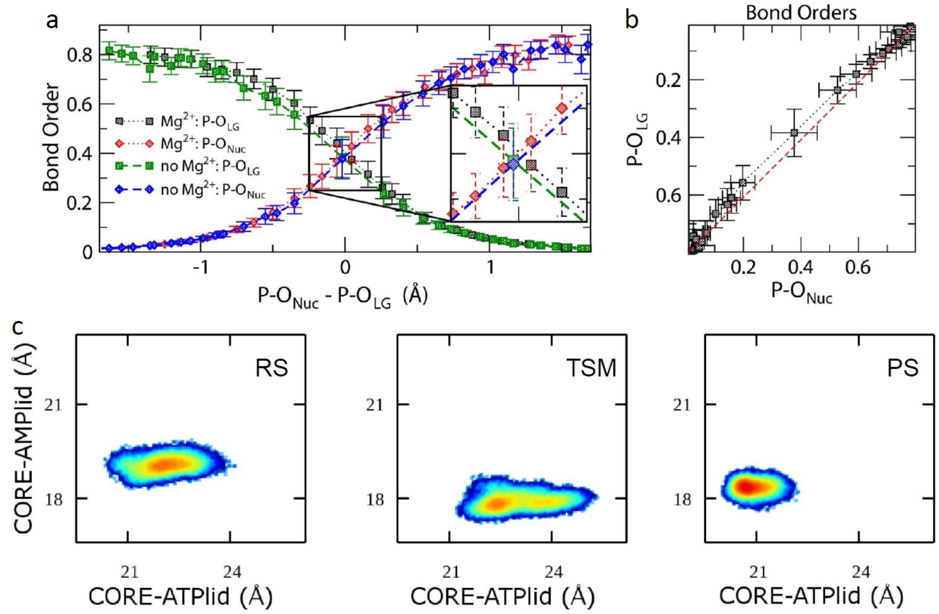


Figure S10. **a** Progressin of the cleaving and leaving P-O bond orders along the phosphoryl transfer reaction coordinate, compared between the QM/MM simulations in the presence and absence of Mg^{2+} ion. **b** 2D diagram of the forming and cleaving P-O bond orders in the reaction without Mg^{2+}. At the transition state, the two P-O bond orders are slightly smaller than those of the reaction in the presence of Mg^{2+} (**Figure S3b**). **c** Distribution of the CORE-ATPlid and CORE-AMPlid distances between the reactant (RS), transition-state mimic (TSM) and product states (PS) obtained from the (300 ns) MD simulations perfromed in the absence of the active site Mg^{2+} ion. The color varies from blue to red with increasing frequency.

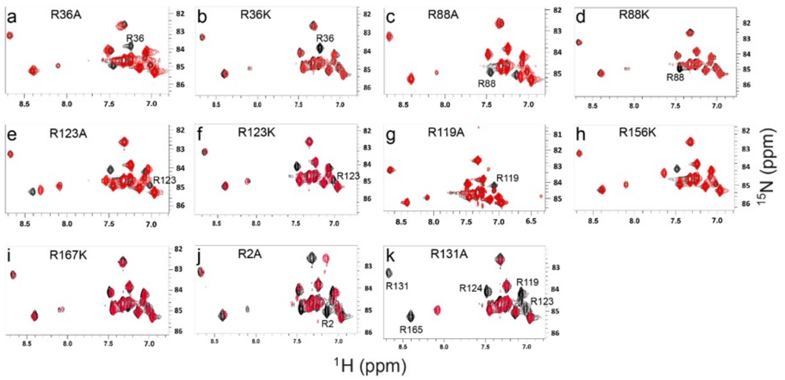


Figure S11. Overlay of arginine side chain $^{1}H\varepsilon$-$^{15}N\varepsilon$ HSQC NMR spectra of apo WT AdK (black) and AdK Arg mutants (red): (**a**) R36A, (**b**) R36K, (**c**) R88A, (**d**) R88K, (**e**) R123A, (**f**) R123K, (**g**) R119A, (**h**) R156K, (**i**) R167K, (**j**) R2A and (**j**) R131A. All spectra were acquired at 298K at 850MHz.

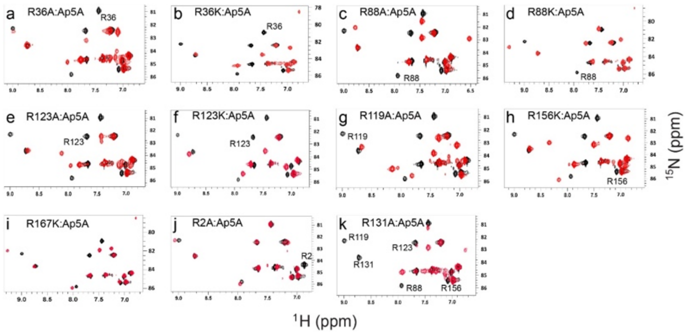


Figure S12. Overlay of arginine side chain ^{1}Hε-^{15}Nε HSQC NMR spectra of Ap5A bound WT AdK (black) and AdK Arg mutants (red): (**a**) R36A, (**b**) R36K, (**c**) R88A, (**d**) R88K, (**e**) R123A, (**f**) R123K, (**g**) R119A, (**h**) R156K, (**i**) R167K, (**j**) R2A and (**j**) R131A. All spectra were acquired at 298K at 850MHz.

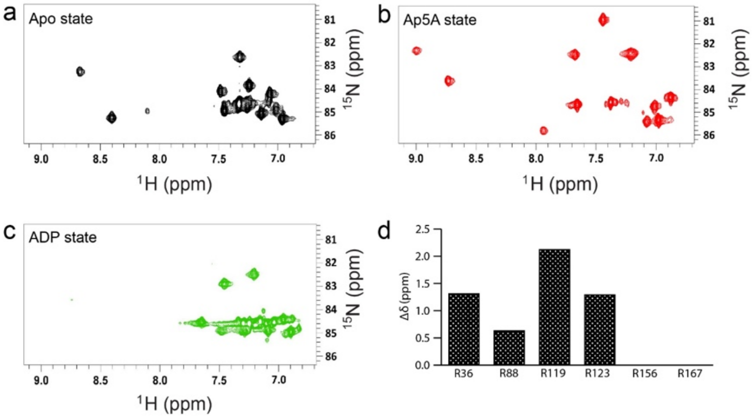


Figure S13. Arginine side chain ^{1}Hε-^{15}Nε HSQC NMR spectra of (**a**) apo (black), (**b**) Ap5A bound (red) and (**c**) ADP bound (green) states of AdK. All spectra were acquired at 298K in 850MHz field. **d** Weighted ^{1}H-^{15}N chemical shift differences ($\Delta\delta$ (ppm) = $[0.5 \times ((\delta H)^2 + 0.2 \times (\delta N)^2)]^{1/2}$) between the Ap5A bound and apo AdK catalytic arginine side chain ^{1}Hε-^{15}Nε groups plotted along with the residue number.

Figure S14. The second lowest PCA mode (PC2) determined from the wild-type QM/MM simulation trajectories. The color and cross-section of the tube represent the amplitude of the PC2 motion. The color code follws the rainbow colors from blue (low amplitude) to red (large amplitude).

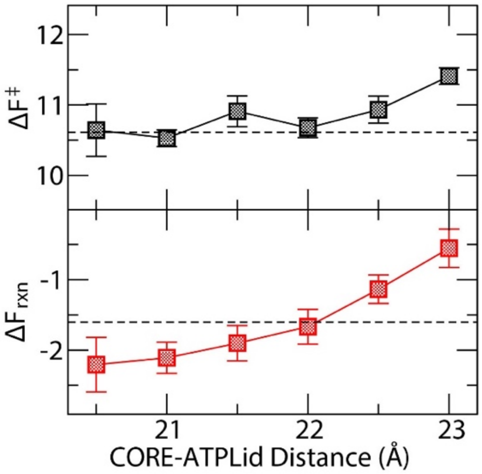


Figure S15. Changes in the phosphoryl transfer barrier and reaction free energy with increasing CORE-ATPlid distance determined from the QM/MM SMCV simulations with a restraint applied to the CORE-ATPlid distance. In each figure, the dash line represents the free energy value of the unrestrained QM/MM simulations. Free energies are in kcal/mol.

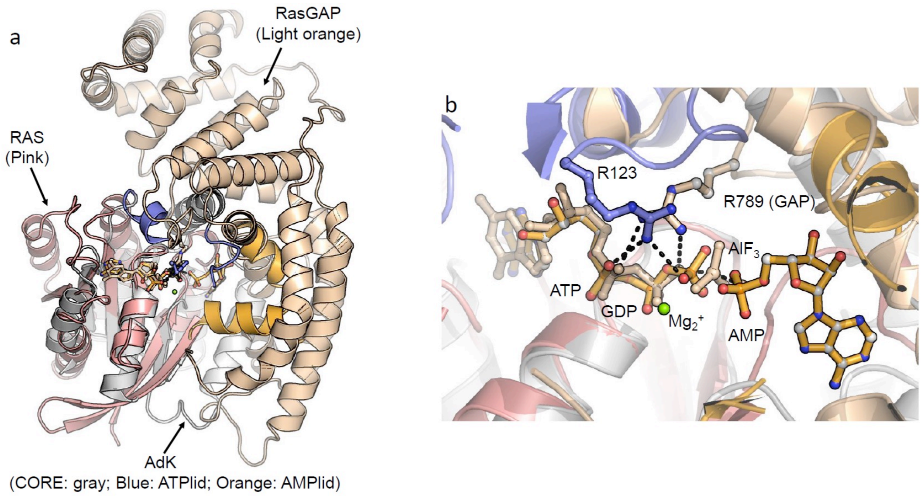


Figure S16. **a** Superposition of Ras-RasGAP complex (PDB: 1WQ1[32]) to AdK (PDB: 1AKE) using the residues around the P-loop. AdK is shown in gray for CORE subdomain, blue for ATPlid and orange for AMPlid. For ATPlid, we only show a part of the subdomain for clarity. Ras and RasGAP are shown in pink and light orange, respectively. b Blow-up view of the active site of AdK and Ras-RasGAP complex. For AdK, we remove the middle phosphoryl group of Ap5A inhibitor, leaving ATP and AMP groups shown in orange. For Ras-RasGAP complex, GDP and AlF_3 are shown in light gray and Mg^{2+} ion in green, respectively. Interestingly, the location of the R123 headgroup occupies the similar position of the "arginine finger" R789 of RasGAP.

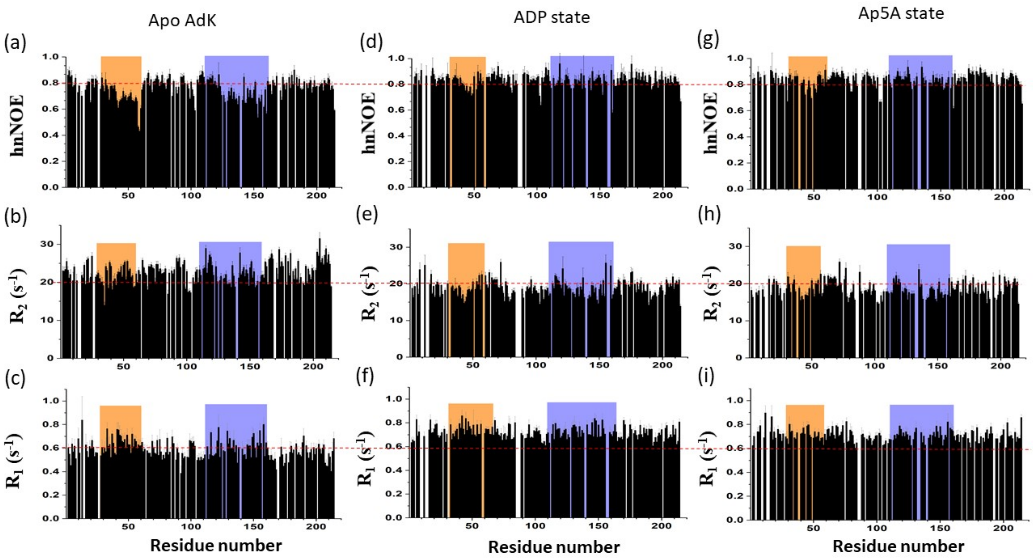


Figure S17. Plot of backbone amide ^{15}N-^{1}H heteronuclear steady-state nuclear Overhauser effects (NOE), R_2 and R_1 relaxation parameters. Shown are the NOE, R_2 and R_1 values as a function of residue number for (**a-c**) the apo state of AdK, (**d-f**) the ADP bound state of AdK and (**g-h**) the Ap5A inhibited state of AdK, respectively. Error bars are indicated on top of the bar graphs. AMP and ATPlid subdomains are highlighted in orange and light blue colors, respectively.

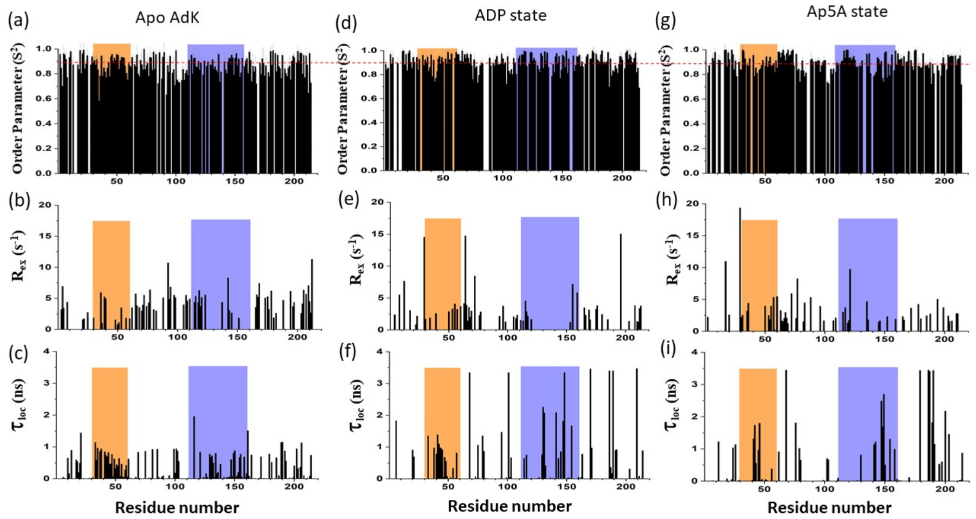


Figure S18. Plot of backbone amide group order parameter (S^2), conformational exchange rate R_{ex} and internal correlation time τ_{loc} versus AdK residue numbers. Shown are the S^2, R_{ex} and τ_{loc} values versus the residue number for (**a-c**) the apo AdK, (**d-f**) the ADP bound state of AdK and (**g-i**) the Ap5A bound state of AdK, respectively. The root-mean-square error associated with each value of S^2 is indicated on each bar graph. AMP and ATPlids subdomains are highlighted in orange and light blue colors, respectively.

Appendix II: references

[1] Rogne, P., Sparrman, T., Anugwom, I., Mikkola, J.-P., and Wolf-Watz, M. (2015) Realtime 31P NMR Investigation on the Catalytic Behavior of the Enzyme Adenylate kinase in the Matrix of a Switchable Ionic Liquid, *ChemSusChem 8*, 3764-3768.

[2] Rhoads, D. G., and Lowenstein, J. M. (1968) Initial Velocity and Equilibrium Kinetics of Myokinase, *J. Biol. Chem. 243*, 3963-3972.

[3] Dayie, K. T., and Wagner, G. (1994) Relaxation-Rate Measurements for 15N−1H Groups with Pulsed-Field Gradients and Preservation of Coherence Pathways, *J. Magn. Res., Series A 111*, 121-126.

[4] Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, *J. Am. Chem. Soc. 104*, 4546-4559.

[5] Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, *J. Am. Chem. Soc. 104*, 4559-4570.

[6] Fushman, D. (2012) Determining Protein Dynamics from 15N Relaxation Data by Using DYNAMICS, In *Protein NMR Techniques* (Shekhtman, A., and Burz, D. S., Eds.), pp 485-511, Humana Press, Totowa, NJ.

[7] Berlin, K., Longhini, A., Dayie, T. K., and Fushman, D. (2013) Deriving quantitative dynamics information for proteins and RNAs using ROTDIF with a graphical user interface, *J. Biomol. NMR 57*, 333-352.

[8] Müller, C., Schlauderer, G., Reinstein, J., and Schulz, G. (1996) Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding, *Structure 4*, 147-156.

[9] Müller, C. W., and Schulz, G. E. (1992) Structure of the complex between adenylate kinase from Escherichia coli and the inhibitor Ap5A refined at 1.9 Å resolution: A model for a catalytic transition state, *J. Mol. Biol. 224*, 159-177.

[10] Diehl, C., Engström, O., Delaine, T., Håkansson, M., Genheden, S., Modig, K., Leffler, H., Ryde, U., Nilsson, U. J., and Akke, M. (2010) Protein Flexibility and Conformational Entropy in Ligand Design Targeting the Carbohydrate Recognition Domain of Galectin-3, *J. Am. Chem. Soc. 132*, 14577-14589.

[11] Bellinzoni, M., Haouz, A., Graña, M., Munier-Lehmann, H., Shepard, W., and Alzari, P. M. (2006) The crystal structure of Mycobacterium tuberculosis adenylate kinase in complex with two molecules of ADP and Mg2+ supports an associative mechanism for phosphoryl transfer, *Protein Sci. 15*, 1489-1493.

[12] Berry, M. B., and Phillips, G. N. (1998) Crystal structures of Bacillus stearothermophilus adenylate kinase with bound Ap5A, Mg2+ Ap5A, and Mn2+ Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg2+ and Mn2+, *Proteins 32*, 276-288.

[13] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926-935.

[14] MacKerell, A. D., Bashford, D., Bellott, Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins †, *J. Phys. Chem. B 102*, 3586-3616.

[15] MacKerell, Feig, M., and Brooks, C. L. (2004) Improved Treatment of the Protein Backbone in Empirical Force Fields, *J. Am. Chem. Soc. 126*, 698-699.

[16] MacKerell, A. D., Banavali, N., and Foloppe, N. (2000) Development and current status of the CHARMM force field for nucleic acids, *Biopolymers 56*, 257-265.

[17] Foloppe, N., and MacKerell, J. A. D. (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, *J. Comput. Chem. 21*, 86-104.

[18] Kerns, S. J., Agafonov, R. V., Cho, Y.-J., Pontiggia, F., Otten, R., Pachov, D. V., Kutter, S., Phung, L. A., Murphy, P. N., Thai, V., Alber, T., Hagan, M. F., and Kern, D. (2015) The energy landscape of adenylate kinase during catalysis, *Nat. Struct. Mol. Biol. 22*, 124-131.

[19] Brooks, B. R., Brooks, C. L., III, MacKerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The Biomolecular Simulation Program, *J. Comput. Chem. 30*, 1545-1614.

[20] Eastman, P., Swails, J., D. Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L.-P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R., and Pande, V. S. (2017) OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics, *PloS Comput. Biol. 13*, e1005659.

[21] Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes, *J. Comput. Phys. 23*, 327-341.

[22] Brünger, A., Brooks Iii, C. L., and Karplus, M. (1984) Stochastic boundary conditions for molecular dynamics simulations of ST2 water, *Chem. Phys. Lett. 105*, 495-500.

[23] Chow, K.-H., and Ferguson, D. M. (1995) Isothermal-isobaric molecular dynamics simulations with Monte Carlo volume sampling, *Comput. Phys. Commun. 91*, 283-289.

[24] Åqvist, J., Wennerström, P., Nervall, M., Bjelic, S., and Brandsdal, B. O. (2004) Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm, *Chem. Phys. Lett. 384*, 288-294.

[25] Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, *J. Chem. Phys. 98*, 10089-10092.

[26] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method, *J. Chem. Phys. 103*, 8577-8593.

[27] Yao, J., Wang, X., Luo, H., and Gu, P. (2017) Understanding the Catalytic Mechanism and the Nature of the Transition State of an Attractive Drug-Target Enzyme (Shikimate Kinase) by Quantum Mechanical/Molecular Mechanical (QM/MM) Studies, *Chem. Eur. J. 23*, 16380-16387.

[28] Recabarren, R., Osorio, E. H., Caballero, J., Tuñón, I., and Alzate-Morales, J. H. (2019) Mechanistic insights into the phosphoryl transfer reaction in cyclin-dependent kinase 2: A QM/MM study, *PLOS ONE 14*, e0215793.

[29] Lassila, J. K., Zalatan, J. G., and Herschlag, D. (2011) Biological Phosphoryl-Transfer Reactions: Understanding Mechanism and Catalysis, *Annu. Rev. Biochem. 80*, 669-702.

[30] Guthrie, R. D., and Jencks, W. P. (1989) IUPAC recommendations for the representation of reaction mechanisms, *Acc. Chem. Res. 22*, 343-349.

[31] Berry, M. B., Meador, B., Bilderback, T., Liang, P., MGlaser, M., and Phillips, G. N., Jr. (1994) The closed conformation of a highly flexible protein: The structure of E. coli adenylate kinase with bound AMP and AMPPNP, *Proteins 19*, 183-198.

[32] Ahmadian, M. R., Stege, P., Scheffzek, K., and Wittinghofer, A. (1997) Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras, *Nat. Struct. Biol. 4*, 686-689.

Chapter III

Mechanistic basis for a connection between the catalytic step and slow opening dynamics of adenylate kinase.

Beata Dulko-Smith[1], Pedro Ojeda-May[2], Jörgen Åden[3], Magnus Wolf-Watz[3] and Kwangho Nam[1]

[1] Department of Chemistry and Biochemistry, University of Texas at Arlington, Arlington, Texas 76019, United States.
[2] High Performance Computing Centre North (HPC2N), Umeå University, Umeå SE-90187, Sweden
[3] Department of Chemistry, Umeå University, Umeå SE-90187, Sweden

(manuscript)

Abstract

E. coli adenylate kinase (AdK) is a small, monomeric enzyme which synchronizes its catalytic reaction with conformational dynamics to optimize a phosphoryl transfer reaction and the subsequent release of the product. Guided by experimental measurements of low catalytic activity in seven single-point mutation AdK variants (K13Q, R36A, R88A, R123A, R156K, R167A, and D158A), we utilized classical mechanical simulations to probe mutant dynamics linked to product release, and quantum mechanical and molecular mechanical calculations to compute a free energy barrier for the catalytic event. The goal was to establish a mechanistic connection between the two activities. Our calculations of the free energy barriers in AdK variants were in line with those from experiments and conformational dynamics consistently demonstrated an enhanced tendency towards enzyme opening. This indicates that the catalytic residues in the wild-type AdK serve a dual role in this enzyme's function: one to lower the energy barrier for the phosphoryl transfer reaction and another to delay enzyme opening and maintain it in a catalytically active closed conformation long enough to enable the subsequent chemical step. Our results also indicate that while each catalytic residue individually contributes to facilitate the catalysis, R36, R123, R156, R167, and D158 are organized in a tightly coordinated interaction network and collectively modulate AdK's conformational transitions. Unlike the existing notion of

the product release being rate-limiting, our findings suggest a mechanistic linkage between the chemical step and the conformational dynamics of the enzyme collectively acting as the bottleneck of the catalytic process.

Introduction

Adenylate kinase (AdK) is an essential[1] small enzyme ubiquitously found in organisms across all domains of life[2]. Its function is to catalyze interconversion of adenine nucleosides through a magnesium-dependent reversible transfer of the γ-phosphoryl group from the donor adenosine triphosphate (ATP) to an acceptor adenosine monophosphate (AMP) yielding two molecules of adenosine diphosphate (ADP). It therefore plays a fundamental role in regulating adenylate turnover to maintain adenosine phosphate concentrations in equilibrium[3,4].

From X-ray crystallographic studies, the tertiary structure of AdK comprises of three subdomains: the ATP-binding lid, the AMP-binding lid, and the CORE subdomain (Figure 1A and B). Crystal structures of unliganded AdK or in complex with native substrates, analogues, or inhibitors[5-8] have established that the enzyme exhibits a domain reorganization upon substrate binding, manifested by the closure of both lids. This conformational rearrangement positions substrates favorably for efficiency of the chemical step, which has been a subject of numerous kinetic and single-point mutation studies[9-15]. Nuclear magnetic resonance (NMR)[16-18] and small-molecule Förster resonance energy transfer (FRET) experiments, performed on AdK from mesophilic and thermophilic prokaryotes found a correlation in the rates of the catalytic turnover (k_{cat} = 263 ± 30 s^{-1}) and the ATP lid opening (k_{open}= 286 ± 85 s^{-1} under the turnover condition)[17]. This, comparing to a much faster lid-closing rate of k_{close} = 1,374 ± 110 s^{-1} (17) led to a conclusion that re-opening of the ATP lid after the catalytic step is rate-limiting. The chemical reaction is assumed to take place multiple times in both directions before enzyme re-opening[15]. On the other hand, thermostability studies conducted on the prokaryotic and eukaryotic forms of AdK noted that catalytic abilities of eukaryotic AdK carrying an abbreviated ATP lid domain depended on the stability of the core region, which is unexpected based on the rate-limiting re-opening, a correlation unobserved in the prokaryotic form[5]. Then Mehaffey *et al*. 2018[21] followed conformational changes of the eukaryotic AdK throughout the catalytic cycle and at various ligation states using ultraviolet-photodissociation mass spectrometry and recorded the largest conformational fluctuations within the region of the AMP lid suggesting its prominent role during the catalysis in the eukaryotic form. Furthermore, our recent arginine sidechain ^{1}H-^{15}N HSQC NMR study has found the structural heterogeneity of the *E*. *coli*'s active site residues between the apo-state enzyme, a tight inhibitor bound enzyme and natural substrate bound enzyme[15], suggesting that the enzyme exhibits more complex ATP lid and AMP lid dynamics during catalysis and ligand binding/release than the simple two-state conformational transition model. In contrast, in earlier NMR studies, the two-state conformational transition model was assumed to fit experimental data and determine the relative population and residence time of the open versus closed states of the enzyme in earlier NMR studies[17,20].

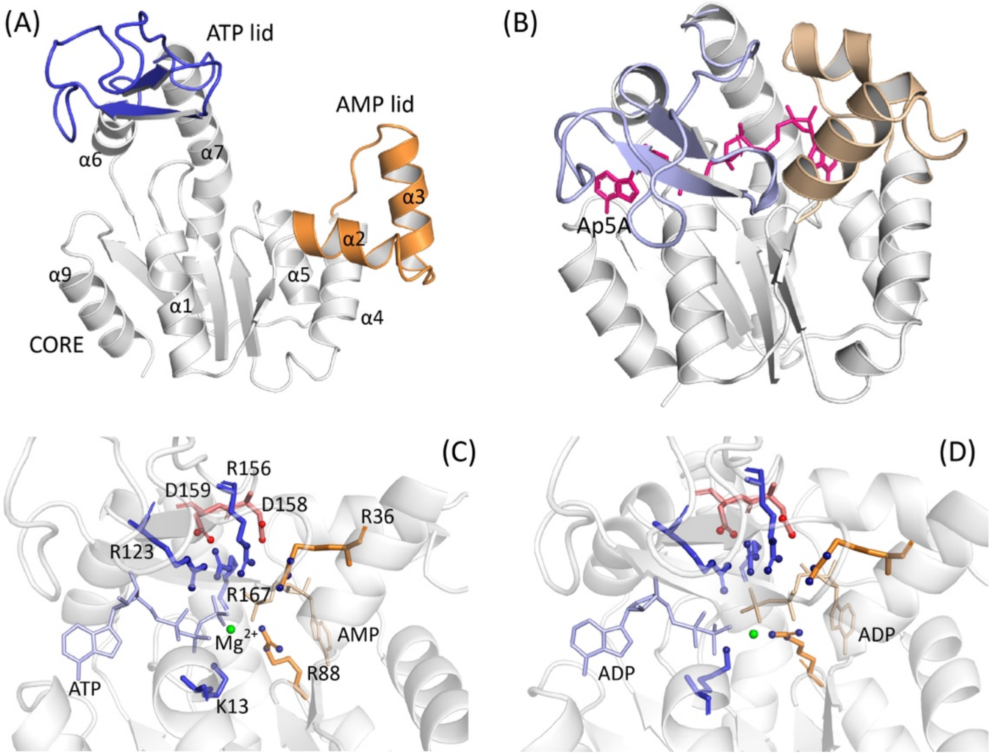


Figure 1. X-ray structures of (**A**) open (PDB ID: 4AKE[8]) and **(B)** closed conformations (PDB ID: 1AKE[7]) of AdK. ATP lid (*blue)* comprises residues 120 to 160, AMP lid *(orange)* residues 31 to 62, and the reminder constitutes the CORE domain *(grey)*, respectively. In (**B**), the Ap5A (purple) inhibitor bound in the structure is shown in purple. **(C)** A close-up of AdK's active site in the reactant state, bound with ATP and AMP and **(D)** in the product state bound with two ADP molecules. Key catalytic residues are shown in the ball-and-stick model: licorice for the residues on the ATP binding and CORE domains and orange on the AMP binding domain, respectively. The two aspartates that coordinate R123, R156 and R167 are shown in pink. Key α helices are indicated in (**A**).

Considering that unravelling mechanistic details of the catalytic step through experiment remains just as elusive as proving a potential link between the enzyme's catalytic rate and dynamics of a particular subdomain, computational investigation has shown its power to model enzymatic behavior on an atomic scale to demystify its inner workings. Coarse-grained simulations on AdK by Cui and colleagues[22,23] as well as Zeller and Zacharias[24] probed global dynamics of the enzyme in relation to substrate binding, focusing particularly on motions of the lids. Additionally, in a separate study[25], all-atom simulations of open-to-close transitions of unliganded AdK were modeled. It was found that the apo AdK can form a catalytically competent conformation in solution with the ATP and AMP lids displaying dynamically asymmetric behavior, *i.e.* both lids shift between open and closed conformations independently. Open-to-closed transitions are also important for the substrate binding step. In particular, it has very recently been proposed that repeated cycles of conformational shifts are necessary for the substrates to adjust themselves into the correct orientation within the binding pocket[26]. This happens when the enzyme is partially open and allows enough mobility for the substrate to relax into its native binding pose. Additionally, Henzel-Wildman *et al.*[27] who focused on the mechanics behind AdK motional fluctuations, identified eight hinges within the structure of AdK that control the direction and amplitude of functionally significant large-scale segmental shifts. However, assuming the accepted interpretation of kinetic data whereby a causal relation is implied between the frequency of ATP lid opening and the catalytic turnover of AdK, the mechanism of such a coupling remains to be understood.

Given the small size of AdK and the location of the catalytic residues on the ATP and AMP lids, it is conceivable that the domain motion and catalytic activity are somehow linked, if not coupled, and mediated by the same active site residues (Figure 1C and D), which while facilitating the chemical step may also modulate dynamics of the lids. Under this hypothesis, we envisioned that mutation of these residues might affect the dynamics of the two lids as well as the chemical step. We proceeded to test this hypothesis using molecular dynamics (MD) and hybrid quantum mechanical/molecular mechanical (QM/MM) free energy simulations. In such simulation study, our goal was to observe mechanistic details of interactions between catalytic residues as compared to the wild type and ascertain the impact that each of the catalytic residue may have on the free energy barrier for the chemical step and enzyme dynamics.

On the basis of the kinetic data from single-point mutation studies of the six catalytic residues – five arginines and one lysine – and one coordinating aspartate that in each case showed a varying degree of diminished enzyme activity (Table 1)[15], we conducted MD simulations and calculated the chemical barrier of each variant state. These seven cases, K13Q, R36A, R88A, R123A, R156K, D158A, and R167A, were investigated at four orders of complexity. First, the overall dynamics of the two lid domains are compared to identify global and large-scale impact of the mutations. Secondly, correlated motions between distant side chains are examined on a residue level to find more nuanced abnormalities induced by mutation. From that we proceed to a detailed, mechanistic analysis of interaction networks within the active site, whose disruption might be responsible for extinguishing catalytic activity in mutant AdK. Finally, QM/MM simulations compare the changes of interactions during catalysis and the chemical barrier between the wild-type and mutant enzymes against experimental kinetic parameters.

Table 1. Experimental kinetic parameters (k_{cat} and K_M) for the AdK wild-type and its mutant variants

Variant	k_{cat} (s^{-1})	$K_{M,ATP}$ (μM)	$k_{cat}/K_{M,ATP}$ (s^{-1}/μM)	Fold change[d]	
				k_{cat}	$k_{cat}/K_{M,ATP}$
Wild-type[a]	330 ± 11	71 ± 7	4.6	-	-
K13Q[b]	0.016	1400	1.1×10^{-5}	4.8×10^{-5}	2.5×10^{-6}
R36A[a]	55 ± 2	89 ± 7	6.2×10^{-1}	1.7×10^{-1}	1.3×10^{-1}
R88A[a]	1.8 ± 0.04	120 ± 11	1.5×10^{-2}	5.5×10^{-3}	3.2×10^{-3}
R123A[a]	0.28 ± 0.01	110 ± 16	2.5×10^{-3}	8.5×10^{-4}	5.5×10^{-4}
R156K[a]	0.74 ± 0.01	52 ± 5	1.4×10^{-2}	2.2×10^{-3}	3.1×10^{-3}
D158A[c]	5.7 ± 0.06	59 ± 3	9.6×10^{-2}	1.7×10^{-2}	2.1×10^{-2}
R167A[a]	2.4 ± 0.05	56 ± 4	4.3×10^{-2}	7.3×10^{-3}	9.2×10^{-3}

[a] Reported by Ojeda-May *et al.*[15]

[b] Reported by Reinstein *et al.*[28] without a value for the margin of error.

[c] This work.

[d] Fold changes of k_{cat} and $k_{cat}/K_{M,ATP}$ relative to their corresponding wild-type values.

Methods

MD simulations.

The wild type and all mutant enzyme systems were build based on atomic coordinates from the X-ray AdK structure (PDB code: 1AKE)[7] co-crystallized with an Ap5A inhibitor in the closed conformation. Reconstruction of the ligands and Mg^{2+} followed the procedure as described in Ojeda-May *et al.*[15]. Protonation states of titratable amino acid residues were assigned based on their pK_a values in water; the histidine imidazole ring was protonated to maintain the optimal hydrogen bond network within its neighborhood. The systems were modelled in 73 Å cubic boxes under periodic boundary conditions with an explicit water model, TIP3P[29] and 0.15 M NaCl buffer to neutralize the proteins' charge; all ions were distributed randomly throughout the solvent phase. The systems, sized at 36 200 atoms (wild type), were described with the all-atom CHARMM27[30,31] force field. The CMAP[32,33] backbone dihedral correction was used to ensure protein's structural stability.

Initial equilibration was carried out in CHARMM version c42a2[34] and the production MD was handled by OpenMM[35] with GPU acceleration. Simulations were run in the isothermal-isobaric ensemble with a constant target temperature of 300 K and pressure of 1 bar maintained with the Langevin thermostat and piston methods[36]. The leapfrog-Verlet scheme was used to propagate the systems with a 2-fs integration step, justified when constraining all bonds to hydrogen with SHAKE[37]. Electrostatic interactions were computed with particle mesh Ewald summation (PME) scheme[38] where real-space interactions were computed up to 12 Å. The van der Waals interactions were smoothed out with a switching function within the interval of 10 to 12 Å. All mutant state trajectories were computed in duplicate or triplicate (WT) for at least 500 ns each and up to 1 µs. The mutant systems, number of replications, and length of each simulation are summarized in Table S1. In Appendix III, we also describe the details of analysis performed in this study.

Kinetic analysis of the D158A variant.

Expression and purification of the D158A protein was performed as previously described for wild-type Adk from *E. coli*[39] based on the self-inducing plasmid (pEAK91). The D158A substitution was created using site-directed mutagenesis by applying the QuikChange approach (Stratagene) with primers purchased from GenScript® (Leiden, the Netherlands). Verification of the plasmid DNA sequence (Eurofins Genomics, Germany) confirmed the presence of the D158A mutation. The kinetic analysis was accomplished with a coupled ATPase assay as described generally in (Rhoads and Lowenstein, 1968)[40] and specifically for the Adk application[41]. The kinetic data was fitted using the Michaelis-Menten equation (equation 1):

$$V = \frac{k_{cat} * [S]}{K_M + [S]} \qquad \text{(eq. 1)}$$

Where V is the reaction velocity, [S] is the substrate concentration, k_{cat} is the catalytic turn-over number and K_{M} is the Michaelis constant.

Results

Center of mass distances

Center of mass (COM) distance analysis provides a global view of large-scale, close-to-open dynamics of AdK's two ligand-binding domains, the ATP and AMP lids, used to assess the impact of a single-point mutation on the active site (Figures 2A and B). The COM distances are defined in Figure S1. Additionally, since the pattern of motion in the ligand-binding domains is relatively complex as observed in this study as well as recently by Zheng and Cui[25], to discern the details of its distortion, we converted the COM coordinates into the polar coordinate system and thus visualized them in terms of angular displacements (θ and φ angles; Figure S1). For example, we observe that while the ATP lid straightens upwards and the AMP lid bends downwards, both domains also extend sidewise away from each other during the opening event. Therefore, radial distances on their own do not provide a complete description of lid dynamics as the enzyme opens. The picture is completed by adding contributions from the lateral motion as captured by the azimuthal angle, θ, and the simultaneous vertical displacement described by the polar angle,

φ. The results are presented in Figure S2 for the reactant state (RS) and Figure S3 for the product state (PS), respectively, and discussed below.

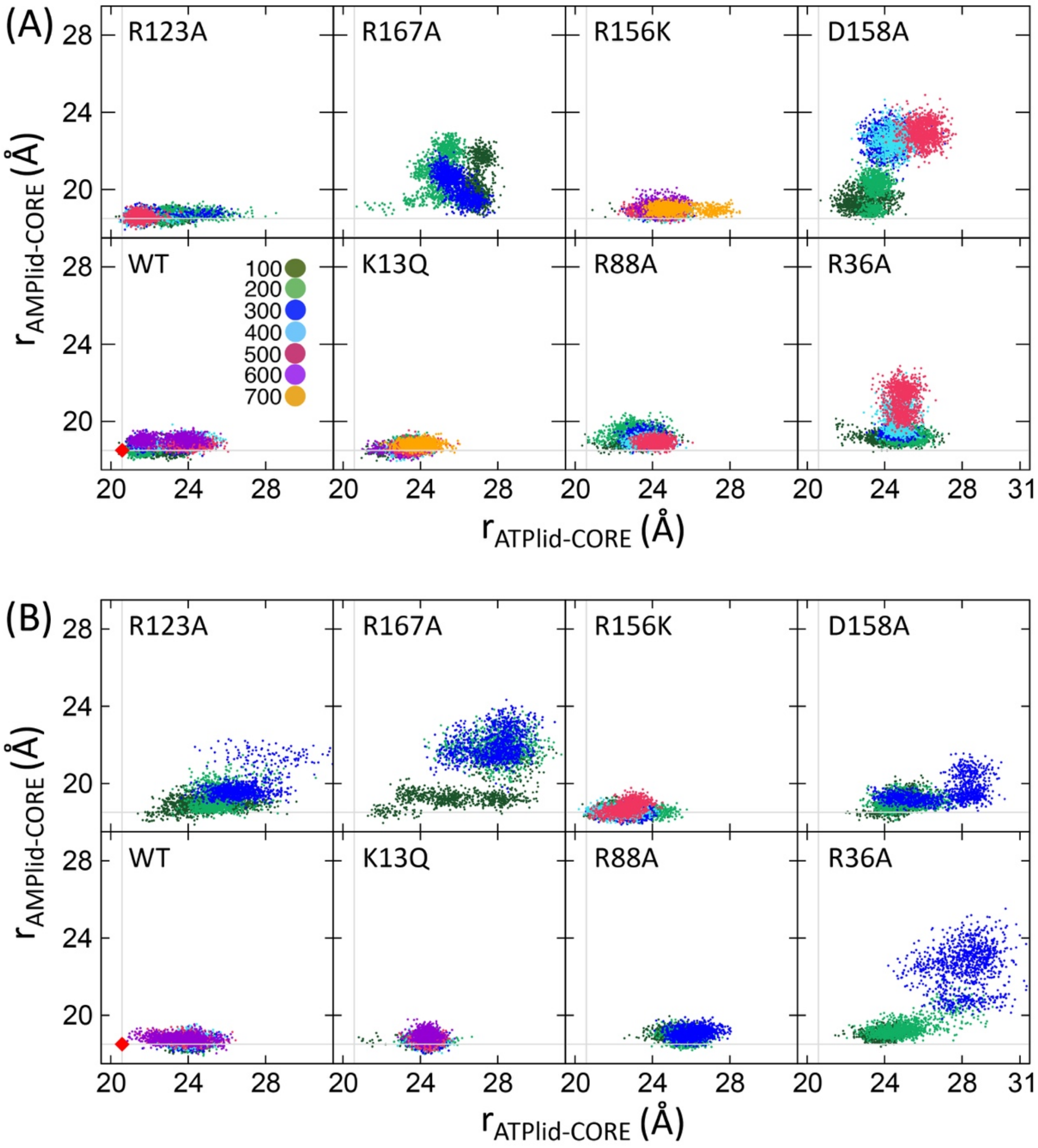


Figure 2. Distribution of the ATP and AMP lid fluctuations, shown with their center of mass (COM) domain displacement with respect to the CORE domain (Figure S1): (**A**) for the reactant state (RS) and (**B**) for the product state (PS). Data points are collected into 100 ns color-coded segments to show their time evolution during MD simulations. Each system was followed up until the opening event or up to 700 ns if no earlier opening was observed. The length of the different MD simulations is summarized in Table S1.

First, we established enzyme's benchmark behavior using triplicated wild-type (WT) MD simulations in both the reactant (*i.e.* ATP+AMP) and product states (*i.e.* 2ADP) (Figure S4). The analysis shows that, consistently with previously reported observations[15], the product state (PS) is more labile overall than the reactant state (RS). Of the three replicates, only one can be considered fully closed in the PS but two in the RS (Figures S4A-B); however, only one PS replicate fully opens while none does so in the RS (Figures S4C-D). Any deviations from the closed conformation in the PS, which signify a potential progression towards an opening state, appear to be initiated by the ATP lid fluctuating sidewise (along the θ angle), while the AMP lid tends to tighten upwards (smaller φ angle) against the ATP lid. In the RS, we observe a reversed trend where φ angle for both lids is larger than that seen in the PS. The larger φ angle means tighter closing for the ATP lid but evolution towards the opening for the AMP lid. Based on these observations, we postulate that the PS exhibits an overall greater tendency towards enzyme opening than the RS, likely initiated by ATP lid fluctuations.

For the mutant states, whose trajectories have been computed for at least 500 ns in duplicate, we consider a mutant state to show a tendency towards rapid opening in the PS if the event is observed in both replicates, and in the RS if the event is observed in at least one replicate. Following from this we find that four out of seven mutations, *i.e.* R36A, R123A, R167A, and D158A, result in an increased frequency of opening in the PS (Figure 2B), which occur within the first 300 ns of the simulation. Even those systems that remain closed, *i.e.* R156K and K13Q, open up when mutated to alanine, which is in line with the observed trend (Figure S5). In the RS (Figure 2A), fewer systems open which are the D158A, R36A and R167A mutations. Interestingly, the R167A system, which opens within 100 ns both in PS and RS, reveals incontinent enzyme dynamics perpetrated on a global scale and exhibits not an orderly transition towards an open conformation.

Enzyme opening is invariably marked by a downward bend of the AMP lid (larger φ_M angle, see Figures S2 and S3), which in the PS is preceded by increased fluctuations of the ATP lid. In the RS, those fluctuations are substantially suppressed, and enzyme opening is then driven by deflection of the AMP lid. Some insight as to why the discrepancy happens can be offered by the plots in Figures S2 and S3. In most of the mutant PS cases, when they open, both lids move away from each other in a concerted fashion. In the RS, we observe that the opening on the AMP lid is initiated by the ATP lid swinging towards it on a collision course (larger θ_T angle) but then rebounding (e.g. Figure S2, R167A and D158A). These sidewise fluctuations of ATP lid do not impact AMP lid dynamics in the PS because, as stated above, the latter tends to tighten upwards and incurvate and thus avoids the approaching ATP lid. In the RS, on the other hand, because the AMP lid is relatively stationary and closed, atypical behavior of the ATP lid results in uncooperative lid dynamics. We note that these details in the opening in the PS versus RS are distinctive from the opening observed in the apo enzyme. In particular, unlike the apo state where both lids move relatively independently from each other[25], the dynamics of the two lids are partly connected in the holo states (*i.e.* RS and PS), through bound ligands, which result in the observed behaviors.

Covariance analysis

Following from the large domain motion quantified by COM distance plots, we set out to obtain a residue-level view of the enzyme dynamics between the product and reactant states. To this end, we performed a covariance analysis of correlated motions between distant residues using Pearson cross-correlation function[42], which measures the strength of correlation in directional motion between two residues. In order to compare the behavior of side chain motions in the systems that remained closed despite the mutation to those that opened rapidly, we selected the early regions of the MD trajectory, where no conformational transition was yet observable, *i.e.* coordinate frames up to 200 ns (or 30 ns in the case of the R167A mutant).

We first benchmarked our observations by performing covariance analysis on the wild-type reactant and product states within the closed conformation, taking one each from the cases where the enzyme remained in the closed conformation throughout the 600 ns MD simulation (e.g., PS replicate 1) and where it eventually opened (PS replicate 2 in Figure S4). For the wild-type RS, since no (full) opening was observed, the most dynamics system, *i.e.* replicate 3, was used as a proxy for an open case (Figure 3). Note that all covariance analyses were performed within the closed conformation (*i.e.* 100 ns MD simulations for WT). Through the comparison of the trajectories that remained closed versus those that eventually opened at a later time, we sought to visualize the difference in the correlated/dynamic behaviors of the enzyme between the two cases, through which we aimed to provide the range of correlations expected from the mutant systems.

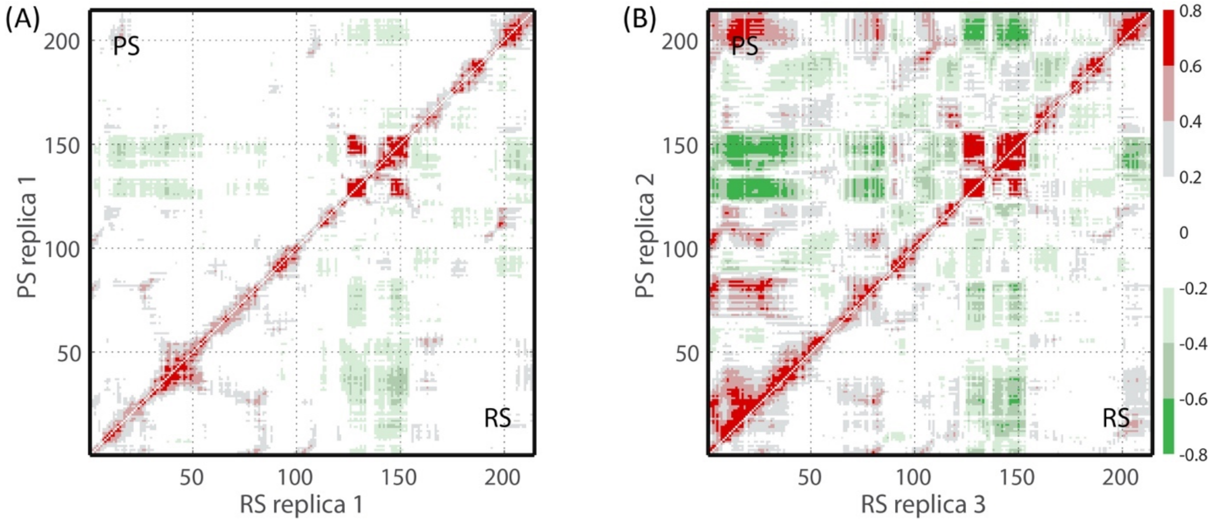


Figure 3. Covariance and pattern of side chain fluctuations in wild-type AdK, which differ between MD replicas (**A**) that remain closed and (**B**) that undergo a conformational transition to the open state. The x and y axes refer to the residue ID of the enzyme, the PS plot is displayed in the upper left corner and the RS in the lower right corner, as indicated in each figure. When the enzyme is not destined to open as in (**A**), side chain motion is either random (white) or regionally ordered to move in an anti-correlated (green) or correlated (red) fashion. When the opening event is imminent as in (**B**), large areas and residues distant from the active site begin to move in a concerted fashion. For the WT RS, replicate 3 was used in the analysis even though the simulation did not undergo an opening event, but it is the most dynamic system within the WT RS replicate MDs. Note that all covariance analyses were performed only for MD segments where the enzyme remained in the closed conformation, e.g., 100 ns in all WT systems, regardless of their ultimate open vs closed conformational states.

In the WT cases that remained closed throughout the simulation, we observe that vast regions of the protein show uncorrelated motion with only the ATP lid region moving in a correlated fashion (Figure 3A, see the region of residue 120 ~ 160 for ATP lid and residue 8 ~ 26 for a part of CORE and 30 ~ 60 for AMP lid). On the other hand, those cases that were destined to open show side chain motion as almost universally ordered, and mostly anticorrelated (negative correlation coefficient) (Figure 3B). This suggests that prior to the opening, the randomized fluttering of the enzyme’s side chains is transformed into a concerted movement. To provide insights into the difference in covariance pattern between the two cases, we compared the average structures of the 100 ns MD trajectories in Figure S6. Interestingly, the average structure of PS replicate 2 adopts a conformation whose ATP and AMP lids are slightly more open compared to the PS replicate 1 and all RS replicates. This conformational difference is consistent with our recent publication, where PS populated two different ATP lid orientations: one tightly closed and the second one slightly open, whereas the closed conformation was only observed from the WT RS MD simulations. Moreover, this difference in the ATP lid orientation influenced the free energy of the phosphoryl transfer reaction, *i.e.* the free energy increased with the increase in the distance between the ATP lid and the CORE domain. Based on these findings and the present results, we propose that the open ATP lid conformation is en route to opening, where the difference in covariance is perhaps the signature as well as a point at which the protein breaks from moving as a whole unit and compartmentalizes into smaller modules, a step needed to shift the large lid subdomains into an open conformation.

Between RS and PS, the PS is more dynamically active in terms of opening motion in the ATP and AMP lids than the RS, for both mutant and wild-type cases. Consistent with this, we observed enhanced anticorrelation of the ATP lid against the remainder of the (WT) enzyme in PS for the trajectory that eventually opened (Figure 3B); the RS system with enhanced COM motions also showed similarly enhanced correlation. On the other hand, in those cases where the enzyme remained closed, no significant difference is observed between RS and PS. Similar trend is also observed for the mutant systems. Based on this, we propose that the intensified correlation can be used as a predictor of imminent opening. For example, in R36A and R123A, there is a significant enhancement of the positive and negative correlations in the product state (Figure S7), which is consistent with the rapid opening of the enzyme (Figure 2). Similarly, for D158A and R167A, despite a weaker correlation than in R36A and R123A cases, the pattern has intensified in comparison with other mutants that remained closed such as K13Q and R156K. In the K13Q and R156K cases, the concerted movement is diminished, if not lost, between the ATP lid and the farther regions of the AMP lid (residues 122-160 and 55-75, respectively). Only the correlation of ATP lid to CORE (residues 122-160 and 10-55, respectively) is retained. This suggests that the ATP lid may not be in contact with the body of the enzyme sufficiently to fully occult the active site. The only exception to this trend is the product state of R88A, which is expected to open rapidly based on the correlation pattern but has remained closed during the 500 ns MD simulations. Together, the correlation analysis suggests that each of the catalytic residues affects the behavior of distant regions of the enzyme and through them influences enzyme's global dynamics.

Behavior of AdK variants

Single point mutation experiments have demonstrated that the removal of an individual active site residue has a detrimental impact on the catalytic activity of the AdK enzyme (Table 1). Here, we repeat the experiment *in silico* to develop a mechanistic explanation behind the loss of activity. In particular, MD simulations are performed to investigate the change in the close-to-open protein motions and their time scales, and QM/MM simulations to investigate the variations in the barrier for the phosphoryl transfer chemical step. We present the results of the QM/MM simulations in the "*QM/MM simulations and barriers*" section.

Initially, our hypothesis was that upon mutation of one residue, the remaining residues re-arrange to compensate for that loss thus revealing what role the mutated-out residue might play in the wild-type AdK. Indeed, some level of compensation by the unmutated active site residues was observed (see below for the description of each mutant). However, the analysis revealed a different picture regarding the organization of the catalytic residues and their coordinated function, for example, as Figures 1C and D illustrate, arginine residues 123, 156, and 167, which cluster above the substrates/products, can be considered main players in the catalytic step as they all coordinate, throughout the reaction, to the γ-phosphoryl group of ATP that is being transferred to AMP. They also form an interaction network with each other, mediated by aspartate residues at positions 158 and 159 (D158 and D159; Figure 1C and D). Specifically, R123 and R167 are both coordinated to D159 whereas R156 only to D158. Based on this organization, we propose that these five residues (R123, R156, R167, D158 and 159) form a catalytic core of AdK, where the two aspartate residues provide an interaction platform to coordinate the orientation of the three arginine residues. The significance of these residues becomes clear from their impacts on the activity of the enzyme (Table 1). Consistent with this, their *in silico* mutations are disruptive to the network (Figure 4), to the point that the D158A mutation has even prompted disengagement of the arginine residues from the ligand, a possible contributing factor behind catalytic impairment of AdK mutants.

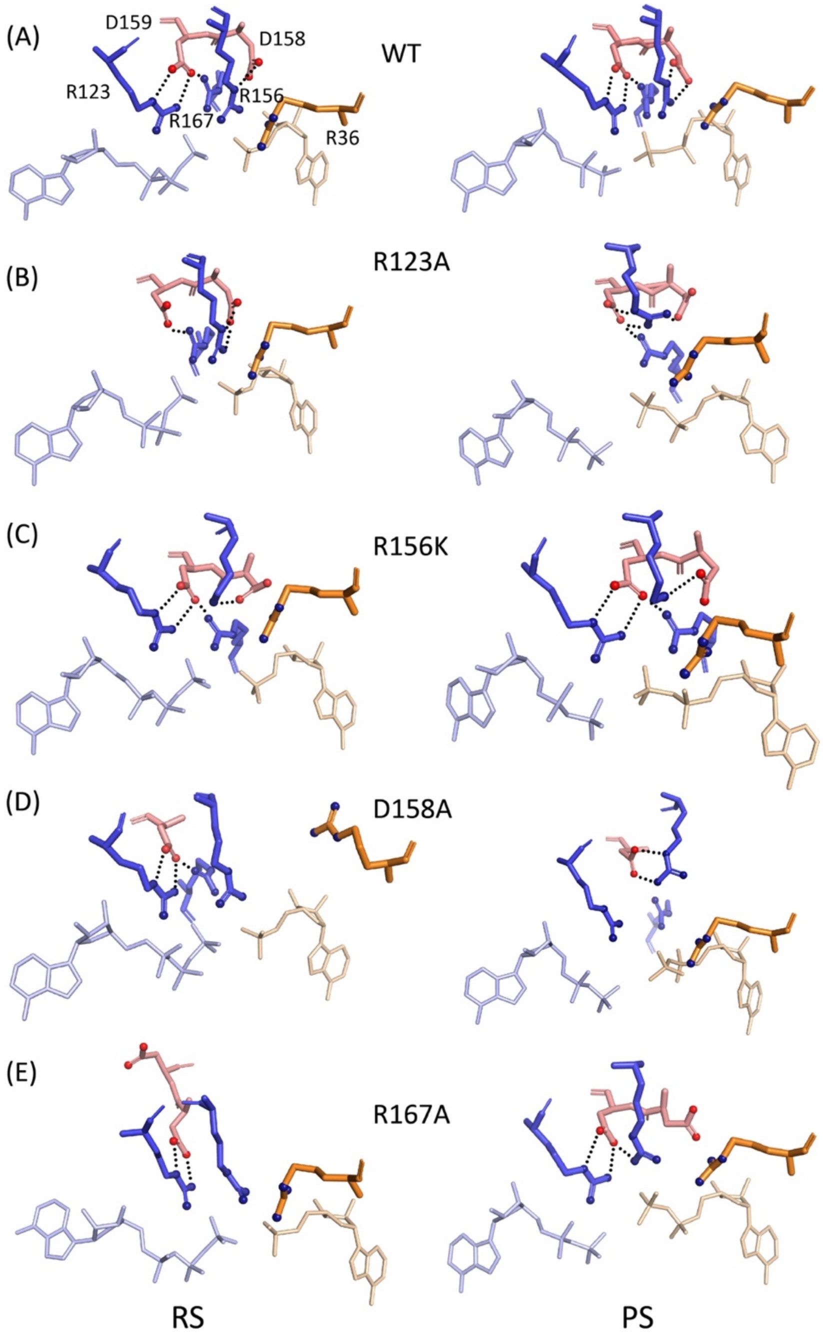


Figure 4. Interaction network of residues within the lid domain sequence and their coordination pattern: (**A**) WT (**B**) R123A, (**C**) R156K, (**D**) D158A, and (**E**) R167A. The left panel shows the interactions in the reactant state and the right panel in the product state, respectively. (**B-E**) The mutant systems also show disruptions in the interaction between catalytic residues preceding enzyme opening. Residues and ligands within the ATP lid are colored blue (ATP, ADP, arginines) and salmon (aspartates); those within the AMP lid are shown in orange (R36) and gold (AMP, ADP). For key arginine residues, interactions with the two aspartate residues are also shown in dashed line.

R123A variant

Among the different mutant systems, R123A is unique in that its dynamic behavior differs between product and reactant states (Figures 2 and S3). In PS, R123A opens rapidly, whereas in the RS, the enzyme remains closed during 500 ns (duplicated) MD simulations except for the typical fluctuations of the ATP lid. Mechanistic reasons for this difference stem from the shift in orientation of R156 and R167 relative to the ligand between the reaction states. In WT RS, R167 coordinates solely to ATP's γ-phosphate and is therefore skewed towards the ATP lid, optimally to maintain its coordination to D159. In WT PS, R167 having followed the transferred phosphate to the product ADP, is now turned towards the AMP lid, its interaction with D159 now being more precarious and looser, but stable when D159 is held steady through a double coordination to R123 (Figure 4A); We have also discussed in our recent publication the change of R167 orientation with the reaction[15]. Therefore, when R123 is replaced with alanine, while D159 can maintain its coordination to R167 in RS, it is quickly broken in the PS leaving the aspartate without a partner and competing with the neighboring D158 for the coordination to R156 (Figure 4B). In our simulation, we observed a brief "tug-of-war" event, lasting *circa* 5 ns, during which both aspartates (158 and 159) became coordinated to R156 (Figure 4B) causing the latter to detach from the β-phosphate of the product ADP. While these changes may have an impact on the closed-to-open protein dynamics especially in PS, they do not seem to be a direct cause behind an impaired catalytic activity of the R123A mutant. During the phosphoryl transfer reaction, R123 is the only residue that directly interacts with the oxygen leaving group of ATP (*i.e.* β-phosphoryl oxygen of ATP), providing stabilization of the negative charge developed on it. Therefore, its removal through mutation results in a significant increase in the phosphoryl transfer reaction barrier as well as the reaction free energy (Table 2).

R156K and D158A variants

The aspartate residues 158 and 159 ensure that the catalytic arginine residues 123, 156 and 167 are positioned above the correct phosphate. Nominally, R156 resides within the sequence of the ATP lid, but topologically its side chain is positioned above the crevice between the ATP and AMP lids, with the guanidinium moiety bridging the two substrates by coordinating to their terminal phosphates (*i.e.* γ- and α-phosphate of ATP and AMP, respectively; Figure 1C).

Adding to it that its interacting partners are D158 (ATP lid) and R36 (AMP lid) through hydrogen bond and π stacking, respectively, this residue, as well as being involved in the catalysis, also bridges the two subdomains. Thus, a loss of the R156—R36 contact is a harbinger of the enzyme's imminent shift to an open conformation.

Given that experimental studies looked at arginine-to-lysine mutation, which still retained the positive charge of the side chain, its impact on protein dynamics was obscured and we saw no difference from the behavior of the WT, that is the R156K mutant remained closed throughout the 700 ns simulation (Figures 2, S2 and S3). Nevertheless, the mutation reveals that the only nitrogen in the R156K side chain is invested in the interaction to D158 rather than participating in ligand coordination (Figure 4C) and therefore in the catalytic step (Table 2). On the other hand, when the positive charge is removed entirely via a to-alanine mutation, MD simulations observed an enhanced closed-to-open dynamics of the enzyme like the other arginine-to-alanine mutations (Figure S5B). These results suggest that the interaction within the group of residues clustered above the ligands (*i.e.* R123, R156, R167, D158 and D159) must be undisturbed for the enzyme to remain in a closed conformation.

One important finding of the present study is the central role of D158 in coordinating one of the key, catalytic arginine residue in the active site (*i.e.* R156) during the catalytic phosphoryl transfer and opening/closing of the ATP lid. To test this idea, we performed a single-point D158 to alanine mutation experiment and measured its kinetic parameters (*i.e.* k_{cat} and K_M; Table 1). As expected, the mutation substantially reduced the catalytic activity of the enzyme (58-fold in k_{cat}), while the change in K_M is relatively small. This result suggests its involvement only in the catalytic reaction and thus corroborates our prediction from MD simulations. Importantly, since D158 does not directly interact with the reaction substrates and products, thus being a second-shell interactions residue, a significant impact on enzymatic activity is not anticipated. To understand at the atomic level how the mutation affects the conformational dynamics of the ATP and AMP lids and locally at the active site, we performed the MD simulations on the D158A mutant (Figures 2 and S3). As described above, in the R123A variant, D158 was too wobbly to stabilize R156. When it is absent altogether, the same scenario plays out: we observe abstraction of R156 by D159, its departure from the ligand and a gradual lifting of the ATP lid until the interdomain contacts break and the enzyme opens. In particular, the pull of D159 on R156 skews it towards the ATP lid (Figure 4D). This releases the π-stacking R36 residue and drags away the adjacent K157 from its D54 contact on the AMP lid (Figure S8A) and the enzyme then cracks open immediately. This event is clearly seen in the 100-200 ns segment on the D158A COM plot (Figure 2A) as a discontinuity in the COM distance distribution in the RS. Another key interaction identified by our MD simulations is the K57-E170 coordination on the interface between the ATP and AMP lid (Figure S8A); when it is broken, the enzyme is destined for a transition towards an open conformation. Consistent with this, when the E170 residue was mutated to alanine, the conformational equilibrium was shifted towards the open state[43]. In the X-ray structure (PDB ID: 1ADK [7]), however, E170 coordinates instead to the nitrogen in L58, which is not sustained in the MD.

R167A variant

MD simulations of the R167A mutant provided an overall dramatic picture of a disorderly dynamics of the entire enzyme showing signs of impending loss of structural integrity of the two lids. This was particularly pronounced in the reactant state. R167 resides on the α7 helix, the largest helical structure in AdK, with a hinge half-way through, which affords flexibility to the N-terminal part of the segment where it connects to the ATP lid (Figure 1A and B). R167 is held in position via a single hydrogen bond to D159 (Figure 1C). Interestingly, as in the R123A case, the R167A mutation also interferes with the coordination of R156 to the ligand (Figure 4E),), by destabilizing its hydrogen bond to the coordinating partner, D158. In the WT, the α7 helix is rather flexible, which is partly counteracted by R167 anchoring it to the ligands. We speculate that without this contact in the mutant system, an appreciable degree of α7 flexibility is imparted onto the C-terminal part of the ATP lid, which hosts residues R156, K157, D158, and D159 arranged into a loop, which we denote as the catalytic loop in this paper (Figure 5A). Neither PS nor RS of the R167A system maintains that loop formation (Figure 5B) or the R156-to-D158 coordination. Nevertheless, the PS manages to linger in the closed state for *circa* 100 ns while the RS opens immediately (Figure S4).

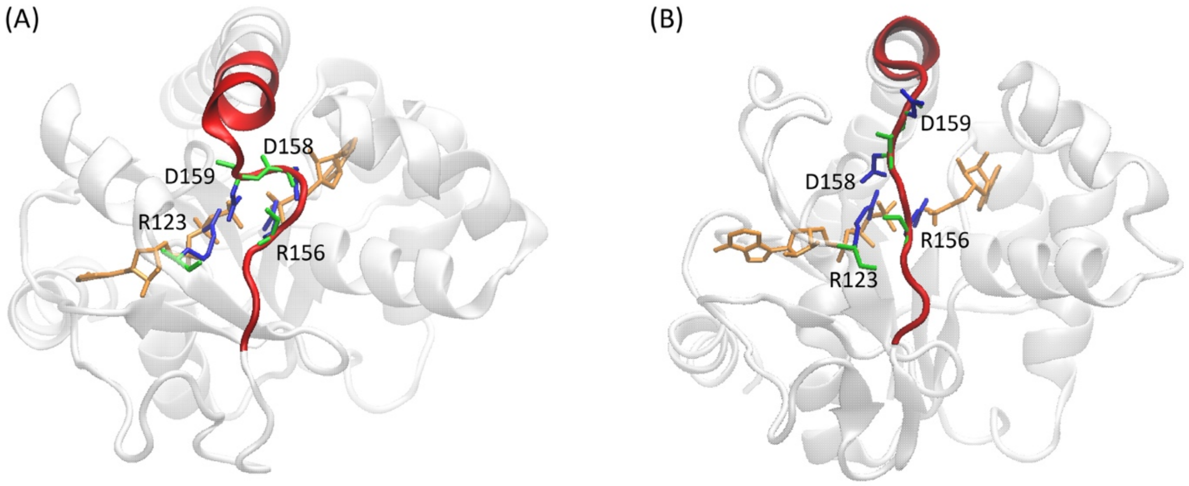


Figure 5. Catalytic loop structure of the ATP lid, which comprises of the residues between 156 and 159 and is positioned precisely above the site of the phosphate transfer: (A) wild-type and (B) R167A mutant both in the RS. The loop and a part of the α7 helix are shown in red and the residues involved in shown in green and blue cartoon. Structural distortion of R167A is substantial and visible in the lost of the secondary structure in helix α6 as well as in the straightening of the loop feature. The loss of the curvature disables interactions between R123-D159 and R156-D158 and breaks an interdomain contact between K157 (ATP lid) and D54 (AMP lid).

R36A and R88A variants

Both R36 and R88 reside within the AMP lid and coordinate to the α-phosphate of the AMP substrate (Figure 1C). Additionally, R36 interacts with R156 via a π stacking (Figure S8B) to keep the enzyme closed along the ATP–AMP lid interface. Its absence forces enzyme opening along the AMP lid, with a rapid (within 100 ns) and substantial bent of 6Å (PS) and 4Å (RS) while the fluctuation of the ATP lid remains steady (RS) (Figures S2 and S3). On the other hand, the R88A mutant remains in the closed conformation throughout the 500 ns (duplicated) MD simulations in both RS and PS (Figure 2). Unlike this, residue-level analysis of the correlated motion reveals that side-chain fluctuation (Figure S7) follows a pattern in the PS that signals domain opening, suggesting that the lack of R88 prevents the expected transition to the open state. In addition, despite the relatively weaker interaction with the reaction substrates than other arginine residues in the active site, the impact of mutation on the catalytic activity is substantial (310-fold reduction relative to the wild-type; Table 1). Compared to this, the R38A mutation has the smallest impact on the catalytic activity (8-fold reduction) among the tested arginine-to-alanine mutations. This difference between the two mutants suggests that the impact on opening dynamics does not always have to be related to the change in catalytic activity and that other factors, such as the phosphoryl transfer reaction step and overall conformation, may play a more important role in the catalytic activity loss.

K13Q variant

The nitrogen of the K13 side chain coordinates two oxygen atoms of ATP, one on a β-phosphate and the other on the γ-phosphate (Figure 1C). In addition, it interacts with the carbonyl oxygen of alanine 8 within the P-loop. The mutation to glutamine (*i.e.* K13Q) shortens the side chain by one methylene group and removes the positive charge of the residue. Then, the mutated side chain may not be long enough to reach the coordinating oxygens of ATP and the removed positive charge is unable to stabilize negatively charged phosphates of the nucleotides. Despite this, our simulations show that the K13Q mutant is exceptionally rigid, *i.e.* overall dynamics of the protein is suppressed compared to the WT on a domain-motion and residue-correlation levels in both PS and RS (Figures 2, S3 and S4).

To see what might happen if the entire side chain of residue 13 was removed, we constructed a K13A mutation and simulated it for 300 ns in the PS. We observed that this mutant opened within the first 100 ns (Figure S5A) in line with the behavior as seen in the other mutant cases which also opened on timescales shorter than that of the WT. Notably, both ADPs in K13A were shifted upwards from their reference positions in the WT and no longer nestled in the P-loop (Figure S9), indicating a more persistent interaction with the guanidinium groups of R123, R156, and R167. This led us to propose that the role of K13 in AdK dynamics may be not so much a facilitation of enzyme opening as providing an electrostatic counterbalance to weaken the coordination from catalytic arginines, especially those located above the bound ligands, and enable nucleotide release. There is yet another species, symmetrically across the active site, which appears to be burdened with a similar function – the Mg^{2+} cofactor. Kerns and coworkers[20] demonstrated that while a divalent cation cofactor must be uniquely a magnesium for the chemical step to proceed at an optimal rate, it can be any other divalent cation, e.g. Ca^{2+}, to preserve a timely lid opening. Likewise, Ojeda-may *et al.* observed in the simulations that in the absence of the Mg^{2+} ion, the WT assumes a more compact, closed conformation[15] like what we observed in the K13Q simulations.

Together the present results paint a picture of the active site as an electrostatic sandwich: the negative phosphates of the nucleotides in the middle, flanked by the positive charge from the catalytic arginines on one side (residues above the ligand in Figures 1C and D) and K13, Mg^{2+} and R88 on the other (residues below the ligand). Therefore, a removal of any of these residues via mutation results in a loss of the balanced interactions between them and leads to a facilitated opening of the enzyme. For example, when any of the arginines is mutated away, that anchoring interaction to the nucleotides is lost and the enzyme opens; that is, the ATP lid opens first after being detached from the ligands as in R123A, D158A and R167A mutant cases. In contrast, when the lysine or Mg^{2+} is removed, the nucleotides remain electrostatically held by arginines above (*i.e.*123, 156, and 167) and the ATP lid ends up lifting with the ligands in tow. Therefore, maintaining a fine balance (of positive charges) between the two sides of the enzyme is critical for a slow opening of the two lids and to allow a sufficient time for reaction.

QM/MM simulations and barrier

Next, the QM/MM simulations were performed for each mutant to determine their catalytic mechanism and free energy barrier. The results are summarized in Table 2 and in Figure 6, where we compare the effective barriers calculated from the QM/MM simulations to the barriers estimated based on the k_{cat} values (Table 1). For each system, we determined both the barrier of the phosphoryl transfer step and the free energy of proton transfer between ATP and AMP, where the latter determines the lowest free energy value in the reactant state. Briefly, the proton transfer in the reactant state depended on the system. For example, while the WT enzyme favored the proton on ATP, four mutants (*i.e.* R88A, R156K, D158A and R167A) favored the proton on AMP. In these mutants, the calculated barriers presented in Figure 6 included the contribution of the free energy of proton transfer (from AMP to ATP), while in all other cases they were the free energy barrier of the phosphoryl transfer step. In addition, for the WT enzyme, the QM/MM simulations were performed at the density functional theory (DFT) level[44] to correct the phosphoryl-transfer barrier determined at the semi-empirical AM1/d-PhoT QM/MM level. The determined correction factor of the barrier (*i.e.* 3.8 kcal/mol) was applied to the phosphoryl transfer barriers of all mutant systems.

Table 2. Free energies of proton and phosphoryl transfers of the wild-type AdK and mutant variants. All free energy values are in kcal/mol.

	WT[a]		K13Q[b]	R36A[b]	R88A[b]	R123A[b]	R156K[b]	D158A[b]	R167A[a,b]
	Mg^{2+}	no-Mg^{2+}							
$\Delta F^{\ddagger}_{p-trs}$ [c]	10.6	17.4	13.4	11.0	6.4	16.3	12.4	12.1	11.3
ΔF_{p-trs} [c]	-1.6	1.3	-1.3	-1.7	-5.4	0.3	-0.9	-1.3	-4.2
$\Delta F_{H^{+}-trs}$ [d]	-2.4	-	-4.9	-0.3	2.5	-5.0	1.0	0.3	3.3
$\Delta F^{\ddagger}_{eff}$ [e]	10.6 (14.4)	17.4 (21.2)	13.4 (17.1)	11.0 (14.8)	8.9 (12.6)	16.3 (20.0)	13.4 (17.2)	12.3 (16.1)	14.6 (18.3)
$\Delta F^{\ddagger}_{expt}$ [f]	14.1	20.4	20.0	15.2	17.2	18.3	17.7	16.5	17.0

[a] The barrier and reaction free energy are determined by Ojeda-May *et al.*[15]

[b] All mutant values are in the presence of the Mg^{2+} ion.

[c] The free energy of reaction and barrier for the phosphoryl transfer from ATP to AMP.

[d] The free energy of proton transfer from AMP to ATP, thus $ATP + AMP(H) \rightarrow ATP(H) + AMP$, where (H) indicates the molecule with a proton on its terminal phosphoryl group.

[e] Effective free energy barrier. For $\Delta F_{H^{+}-trs} < 0$, *i.e.* when the proton on ATP is favored, $\Delta F^{\ddagger}_{eff} = \Delta F^{\ddagger}_{p-trs}$, and for $\Delta F_{H^{+}-trs} > 0$, *i.e.* when the proton on AMP is favored, $\Delta F^{\ddagger}_{eff} = \Delta F^{\ddagger}_{p-trs} + \Delta F_{H^{+}-trs}$. The values in parenthesis are after the DFT-level correction (3.8 kcal/mol) to $\Delta F^{\ddagger}_{p-trs}$.

[f] Free energy barrier estimated from the experimental k_{cat} (Table 1) and the transition state theory.

Overall, the calculated barriers are well correlated with the experimental estimation of the barrier heights (based on the k_{cat} values in Table 1). A few outliers include K13Q and R88A mutants; in particular, our QM/MM and MD simulations have failed to explain the diminished catalytic activity of the R88A mutant. Except these two mutants, the linear correlation observed in Figure 6 shows that the active site mutations have significant impacts on the phosphoryl transfer reaction step as well as the closed-to-open protein dynamics observed from MD simulations. Importantly, our results not only capture the relative change of the catalytic barrier but also well reproduce the absolute catalytic barrier. This implies that the rate-limiting step changes from the re-opening step to the chemical step for mutants. Alternatively, it is the chemical step in all systems including WT. However, in our recent publication[15] we challenge the rate-determining step model centered on protein dynamics. Consistent with this, our results show that at least for the mutants that show enhanced re-opening motion, the rate-determining step is the chemical phosphoryl transfer. In addition, the present results can be used to deduce the role of these residues in the wild-type enzyme: they facilitate on the one hand the phosphoryl transfer reaction and on the other hand contribute to the retention of the closed conformation, during which the reaction ensues. Therefore, their mutations lead to slowed phosphoryl transfer and facilitated re-opening.

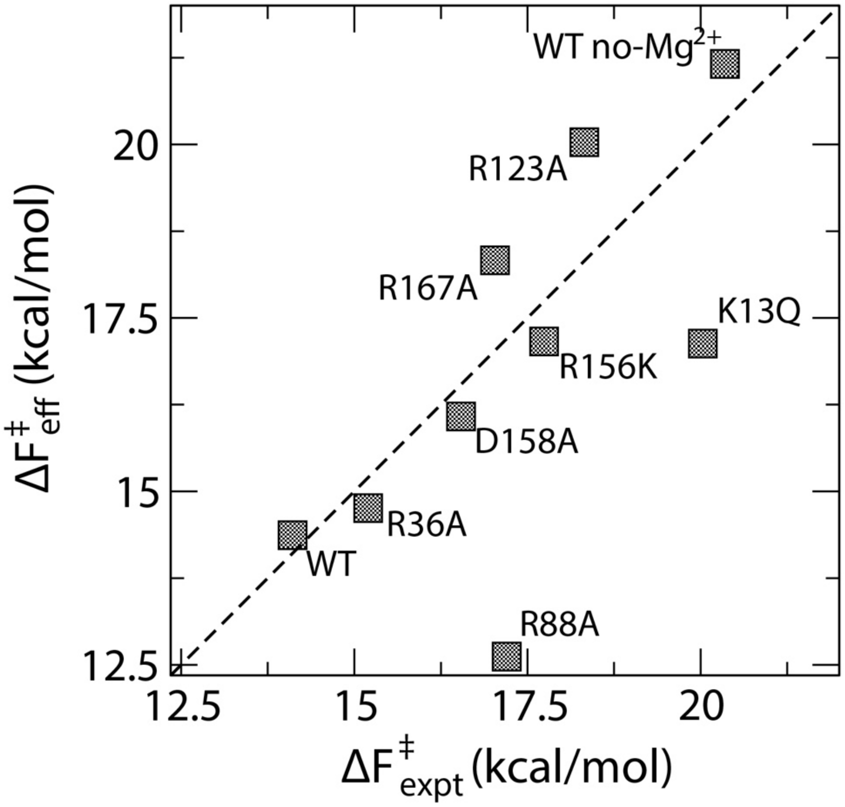


Figure 6. Comparison of free energy barriers of the catalytic phosphoryl transfer between the ATP and AMP: the y-axis represents the effective barriers computed and the x-axis the experimental estimation of them based on k_{cat} (Table 1). Each free energy value is provided in Table 2 and the dashed line refers to a theoretical line of a slope 1.

Second, the mutation affected the phosphoryl transfer barrier as large as 5.7 kcal/mol (for R123A), and interestingly, the same level of impact was also observed on the proton transfer (e.g., 5.7 kcal/mol increase for R167A). In particular, for the R88A, R156K, R158A and R167A mutants, the proton on AMP is favored to the proton on ATP. In these mutants, the effective barriers ($\Delta F_{eff}^{\ddagger}$) reported in Table 2 are after adding the proton transfer free energy to the phosphoryl transfer barrier. For example, for R167A, the phosphoryl transfer barrier ($\Delta F_{p-trs}^{\ddagger}$) is 11.3 kcal/mol and the proton transfer free energy to ATP ($F_{H^{+}-trs}$) is 3.3 kcal/mol, totally 14.6

kcal/mol as the effective barrier (Table 2). We can attribute the missing interaction of R167 with AMP to the shift in proton transfer free energy from -2.4 kcal/mol of WT to 3.3 kcal/mol for R167A; thus, the negative charge of the (fully deprotonated) AMP not being fully stabilized in RS. This is consistent with the rapid opening of the AMP lid in the absence of the arginine. Thus, the role of WT R167 is to stabilize the reactant state (*i.e.* ATP + protonated AMP) relative to the product state (ADP + protonated ADP). Similarly, the roles of other catalytic residues, including Mg^{2+} ion, in the phosphoryl transfer reaction can be deduced from the present QM/MM results. For example, no-Mg^{2+} ion and R123A systems have significantly increased phosphoryl transfer barrier, indicating their role in the stabilization of the reaction transition state. On the other hand, other mutants, such as K13Q, R36A, R156K, D158A, have moderate impacts on the phosphoryl transfer barrier.

Discussion

By introducing *in silico* single point mutations of each catalytic residue, this study probed the role of the catalytic residues in the wild-type adenylate kinase (AdK) enzyme. In all mutants, the QM/MM calculations showed that the energy barrier for the phosphoryl transfer increased. At the same time, most of the mutant enzymes underwent a conformational transition to the open state sooner than the WT. It follows that catalytic residues in the WT delay opening of the enzyme for a sufficient time to firstly enable the catalytic phosphoryl transfer reaction and secondly, facilitate the catalytic reaction through stabilization of the transition state. They are also involved in the relative stability of the product versus reactant states as well as the protonation state in the reactant state.

Mechanistically, a higher free energy barrier in mutants could be attributed to residue's weaker coordination to the substrate phosphate groups. As shown in Figures 1C and D, each catalytic residue interacts precisely with a specific phosphoryl oxygen. While residues K13, R36, and R88 coordinate to oxygen in fixed phosphate groups in both the product and reactant states, R123, R156, and R167 escort the γ-phosphate as it is passed from ATP to AMP. Nam and coworkers[15] noted that the mutation at position 167 (*i.e.* R167A) resulted in disturbed orientation of helices α6 and α7 as well as loss of native coordination to the substrate by the neighboring catalytic residues. This study adds to the previous findings with insight regarding the remaining residues lining the active site.

The pattern of catalytic residues' side chain motion highlights an interplay of two factors: the necessity of maintaining the enzyme in catalytically active state prior to the chemical step and coordinating the opening event for product release afterwards, both apparently characterized by some order-disorder dynamics in the internal behavior spanning the whole enzyme as evidenced by the covariance analysis. It is unlikely that a single residue could have such a considerable control over the behavior of the entire protein. Alternatively, we are looking at an organization of the catalytic residues into a higher-level structure acting collectively, as a network, to govern the enzyme dynamics and the catalytic reaction. If any of those residues is missing, it interrupts the phosphoryl transfer step and disturbs the interaction network between the remaining catalytic residues, simultaneously impairing the chemistry and dynamics of the enzyme. As we demonstrate in this work, one such network is formed by R36, R123, R156, R167 and the coordinating aspartates 158 and 159 (Figure 1C). Among them, the side chains of arginines 123, 156, and 167 follow the γ-phosphate during the transfer reaction as well as they are unique to connect the AMP lid to AMP lid through the interaction with ligands. Therefore, through their movement, the two lids communicate between the reactant and product states. Our MD simulations show that any mutation (to alanine) within the group of residues R36, R123, R156, R167, and D158 affects them all as a collective – their electrostatic interconnectivity becomes unbalanced, invariably culminating in R156 and R167 breaking off from the ligand, soon after which time the enzyme transitions into the open conformation. The fact that the lysine in the R156K mutant (Figure 4C) uses its ammonium nitrogen preferentially to maintain the inter-residual connectivity suggests that this interaction is sufficiently robust that it is favored over the coordination to the phosphoryl oxygen. It therefore follows that those residues act to ensure that the enzyme remains in the closed conformation. The loss in the catalytic activity of the R156K mutant can then be contributed to the weaker interaction with the reaction transition state. The role of D158 transpires as vital to the functioning of this catalytic collective. When missing, it so drastically destabilizes one of the most important players that is R156, which in turn is promptly snatched by D159 as a coordinating partner and drawn away from the γ-phosphoryl group that it arrests the catalysis. Another important interaction among these residues is the π-stacking interaction formed by R156 and R36, which move as a collective following the transfer of the γ-phosphoryl group.

The second group of residues include K13, R88 and Mg^{2+} ion, which lie below the ligands in Figure 1C and interact with them from the core domain direction. In our MD simulations, their mutation (*i.e.* K13Q, R88A and no-Mg^{2+} ion systems) remains in the closed conformation. Based on this, we propose that their role in the wild-type enzyme is to facilitate re-opening of the enzyme, a role opposite to that of the first cohort of residue groups above the ligand. Thus, they exert a fine balance of interactions between the two residue groups: one, provided by the ATP lid interacts with the bound ligands, allowing the ATP lid to remain closed long enough during the reaction, while the second group interacts with the ligands from the core domain side and by this interaction, it weakens the interaction of the ligands with the first residue group. Therefore, this second group not only helps to maintain the enzyme in the closed conformation through the three-way, thus balanced, interactions between this group-ligands-first group, but also enables a robust re-opening of the enzyme by allowing the first group of residues to easily detach from the ligands. For this reason, any mutation of the first group leads to an accelerated re-opening and the mutation of the second group has opposite effects on the re-opening of the enzyme. Nevertheless, while the two groups may play a different role in the closed-to-open enzyme dynamics, all of these residues contribute to the catalytic phosphoryl transfer reaction likely via electrostatic interactions between reactant, transition and product states, which is apparent from measured kinetic parameters and computed phosphoryl transfer barriers.

Conclusion

Our kinetic experiments demonstrated that the removal of a single catalytic residue from the active site of adenylate kinase incapacitates its catalytic activity. Based on that, we introduced the same mutations *in silico* to deduce the role of the targeted residues (K13, R36, R88, R123, R156, R167, and second-coordination shell's D158) in the catalytic reaction and re-opening dynamics of the enzyme, and to investigate a link connecting the two events. From the quantum mechanical calculations, we found that mutations increase the reaction barrier for the phosphoryl transfer. At the same time, all but three mutant cases (R88, K13Q and R156K) show an enhanced transition to the open conformation. This suggest that the catalytic residues lower the free energy barrier for the reaction and delay the opening event (thus product release) showing that at least in these mutants, the chemical step must be rate-limiting. To achieve this, the residues R36, R123, R156, R167, and D158 are organized into a tight-knit network, which, as a collective, acts to maintain the enzyme in the closed, catalytically competent conformation for the completion of the chemical step. In contrast, K13 and Mg^{2+} ion likely counterbalance the action of the arginines by contributing to enzyme opening through electrostatics and thus ultimately promoting product release.

References

(1) Tükenmez, H., Magnussen, H. M., Kovermann, M., Byström, A., & Wolf-Watz, M. (2016). Linkage between fitness of yeast cells and adenylate kinase catalysis. *PloS One*, *11*(9), e0163115.

(2) Fukami-Kobayashi, K., Nosaka, M., Nakazawa, A., & Go, M. (1996). Ancient divergence of long and short isoforms of adenylate kinase: molecular evolution of the nucleoside monophosphate kinase family. *FEBS Letters*, *385*(3), 214–220.

(3) Patel, A., Malinovska, L., Saha, S., Wang, J., Alberti, S., Krishnan, Y., & Hyman, A. A. (2017). ATP as a biological hydrotrope. *Science (New York, N.Y.)*, *356*(6339), 753–756.

(4) De la Fuente, I. M., Cortés, J. M., Valero, E., Desroches, M., Rodrigues, S., Malaina, I., & Martínez, L. (2014). On the dynamics of the adenylate energy system: homeorhesis vs homeostasis. *PloS One*, *9*(10), e108676.

(5) Moon, S., Kim, J., & Bae, E. (2017). Structural analyses of adenylate kinases from Antarctic and tropical fishes for understanding cold adaptation of enzymes. *Scientific Reports*, *7*(1), 1–12

(6) Kovermann, M., Grundström, C., Sauer-Eriksson, A. E., Sauer, U. H., & Wolf-Watz, M. (2017). Structural basis for ligand binding to an enzyme by a conformational selection pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(24), 6298–6303.

(7) Müller, C. W., & Schulz, G. E. (1992). Structure of the complex between adenylate kinase from Escherichia coli and the inhibitor Ap5A refined at 1.9 Å resolution: A model for a catalytic transition state. *Journal of Molecular Biology*, *224*(1), 159–177.

(8) Müller, C. W., Schlauderer, G. J., Reinstein, J., & Schulz, G. E. (1996). Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure (London, England: 1993)*, *4*(2), 147–156.

(9) Byeon, I.-J. L., Shi, Z., & Tsai, M.-D. (1995). Mechanism of adenylate kinase. The essential lysine helps to orient the phosphates and the active site residues to proper conformations. *Biochemistry*, *34*(10), 3172–3182.

(10) Dahnke, T., & Tsai, M. D. (1994). Mechanism of adenylate kinase. The conserved aspartates 140 and 141 are important for transition state stabilization instead of substrate-induced conformational changes. *The Journal of Biological Chemistry*, *269*(11), 8075–8081.

(11) Shi, Z., Byeon, I. J., Jiang, R. T., & Tsai, M. D. (1993). Mechanism of adenylate kinase. What can be learned from a mutant enzyme with minor perturbation in kinetic parameters? *Biochemistry*, *32*(25), 6450–6458.

(12) Yan, H., & Tsai, M. D. (1991). Mechanism of adenylate kinase. Demonstration of a functional relationship between aspartate 93 and magnesium by site-directed mutagenesis and proton, phosphorus-31, and magnesium-25 NMR. *Biochemistry*, *30*(22), 5539–5546.

(13) Yan, H. G., Shi, Z. T., & Tsai, M. D. (1990). Mechanism of adenylate kinase. Structural and functional demonstration of arginine-138 as a key catalytic residue that cannot be replaced by lysine. *Biochemistry*, *29*(27), 6385–6392.

(14) Tsai, M. D., & Yan, H. G. (1991). Mechanism of adenylate kinase: site-directed mutagenesis versus X-ray and NMR. *Biochemistry*, *30*(28), 6806–6818.

(15) Ojeda-May, P., Mushtaq, A. U., Rogne, P., Verma, A., Ovchinnikov, V., Grundström, C., Dulko-Smith, B., Sauer, U. H., Wolf-Watz, M., & Nam, K. (2021). Dynamic connection between enzymatic catalysis and collective protein motions. *Biochemistry*, *60*(28), 2246–2258

(16) Olsson, U., & Wolf-Watz, M. (2010). Overlap between folding and functional energy landscapes for adenylate kinase conformational change. *Nature Communications*, *1*(1), 111.

(17) Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., & Kern, D. (2004). Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nature Structural & Molecular Biology*, *11*(10), 945–949.

(18) Ådén, J., Weise, C. F., Brännström, K., Olofsson, A., & Wolf-Watz, M. (2013). Structural topology and activation of an initial adenylate kinase-substrate complex. *Biochemistry*, *52*(6), 1055–1061.

(19) Hanson, J. A., Duderstadt, K., Watkins, L. P., Bhattacharyya, S., Brokaw, J., Chu, J.-W., & Yang, H. (2007). Illuminating the mechanistic roles of enzyme conformational dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(46), 18055–18060.

(20) Kerns, S. J., Agafonov, R. V., Cho, Y.-J., Pontiggia, F., Otten, R., Pachov, D. V., Kutter, S., Phung, L. A., Murphy, P. N., Thai, V., Alber, T., Hagan, M. F., & Kern, D. (2015). The energy landscape of adenylate kinase during catalysis. *Nature Structural & Molecular Biology*, *22*(2), 124–131.

(21) Mehaffey, M. R., Cammarata, M. B., & Brodbelt, J. S. (2018). Tracking the catalytic cycle of adenylate kinase by ultraviolet photodissociation mass spectrometry. *Analytical Chemistry*, *90*(1), 839–846.

(22) Daily, M. D., Phillips, G. N., Jr, & Cui, Q. (2010). Many local motions cooperate to produce the adenylate kinase conformational transition. *Journal of Molecular Biology*, *400*(3), 618–631.

(23) Daily, M. D., Makowski, L., Phillips, G. N., Jr, & Cui, Q. (2012). Large-scale motions in the adenylate kinase solution ensemble: coarse-grained simulations and comparison with solution X-ray scattering. *Chemical Physics*, *396*, 84–91.

(24) Zeller, F., & Zacharias, M. (2015). Substrate binding specifically modulates domain arrangements in adenylate kinase. *Biophysical Journal*, *109*(9), 1978–1985.

(25) Zheng, Y., & Cui, Q. (2018). Multiple pathways and time scales for conformational transitions in apo-adenylate kinase. *Journal of Chemical Theory and Computation*, *14*(3), 1716–1726.

(26) Lu, J., Scheerer, D., Haran, G., Li, W., & Wang, W. (2022). Role of repeated conformational transitions in substrate binding of adenylate kinase. *The Journal of Physical Chemistry. B*, *126*(41), 8188–8201.

(27) Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A., Karplus, M., Hübner, C. G., & Kern, D. (2007). Intrinsic motions along an enzymatic reaction trajectory. *Nature*, *450*(7171), 838–844.

(28) Reinstein, J., Schlichting, I., & Wittinghofer, A. (1990). Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of Escherichia coli. *Biochemistry*, *29*(32), 7451–7459.

(29) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *79*(2), 926–935.

(30) Foloppe, N., & MacKerell, A. D., Jr. (2000). All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry*, *21*(2), 86–104.

(31) MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., … Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry. B*, *102*(18), 3586–3616.

(32) Mackerell, A. D., Jr, Feig, M., & Brooks, C. L., 3rd. (2004). Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, *25*(11), 1400–1415.

(33) MacKerell, Alexander D., Jr, Feig, M., & Brooks, C. L., 3rd. (2004). Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society*, *126*(3), 698–699.

(34) Brooks, B. R., Brooks, C. L., III, Mackerell, A. D., Jr, Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., … Karplus, M. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, *30*(10), 1545–1614.

(35) Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L.-P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R., & Pande, V. S. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Computational Biology*, *13*(7), e1005659.

(36) Feller, S. E., Zhang, Y., Pastor, R. W., & Brooks, B. R. (1995). Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics*, *103*(11), 4613–4621.

(37) Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, *23*(3), 327–341.

(38) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, *103*(19), 8577–8593.

(39) Reinstein, J., Brune, M., & Wittinghofer, A. (1988). Mutations in the nucleotide binding loop of adenylate kinase of Escherichia coli. *Biochemistry*, *27*(13), 4712–4720.

(40) Rhoads, D. G., & Lowenstein, J. M. (1968). Initial velocity and equilibrium kinetics of myokinase. *The Journal of Biological Chemistry*, *243*(14), 3963–3972.

(41) Rogne, P., Andersson, D., Grundström, C., Sauer-Eriksson, E., Linusson, A., & Wolf-Watz, M. (2019). Nucleation of an activating conformational change by a cation-π interaction. *Biochemistry*, *58*(32), 3408–3412.

(42) Pearson, K. (1895). VII. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, *58*(347–352), 240–242.

(43) Ådén, J., Verma, A., Schug, A., & Wolf-Watz, M. (2012). Modulation of a pre-existing conformational equilibrium tunes adenylate kinase activity. *Journal of the American Chemical Society*, *134*(40), 16562–16570.

(44) Nam, K. (2014). Acceleration of ab initio QM/MM calculations under periodic boundary conditions by multiscale and multiple time step approaches. *Journal of Chemical Theory and Computation*, *10*(10), 4175–4183.

Appendix III

Center of mass distance and angle definitions

The AdK protein was divided into three regions: the ATP lid (residues 120 to 157), the AMP lid (residues 39 to 69), and the CORE containing those residues, which were not assigned to either lid. Coordinates of the regions' mass-weighted centers were computed every 100 picoseconds (symbolically depicted in Figure S1 as points T (ATP lid), M (AMP lid) and the origin, for the CORE).

The angles defined in Figure S1 follow the convention used in spherical coordinate system, where the azimuthal angle, θ, measures the lids' sidewise motion from the reference x-axis to the orthogonal projection of the lid-core vector and runs from 0° to 360° counterclockwise. The polar angle, φ, representing the lids' vertical motion, measures the angle between the zenith direction (z-axis) and the lid-core vector running from 0° to 180° away from the zenith. In this representation, because of the arbitrariness of the origin and direction of each axis and to compare the pattern of the two domain motions between the wild-type and each active site mutant, we first superimposed each coordinate set saved during the molecular dynamics (MD) simulations to a "common" reference frame and determined the two angles. The results are presented in Figure S2 for the reactant state (RS) and Figure S3 for the product state (PS), respectively.

Covariance analysis

A degree to which motions of an individual residue is correlated with each of the other residues was quantified using CHARMM's long tail correlation (cross-correlation) based on the Pearson correlation coefficient calculations (Equation 1).

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

(Equation 1), where *x* and *y* are two sets of variables representing residue-averaged positions of the side chains.

A reference set of coordinates of an average structure was computed first, against which all coordinate frames used in the analysis were aligned before covariance calculations. Covariance calculations were performed on the first 200 ns trajectory segments (or 30 ns for R167A), containing 2000 frames.

Table S1. The mutant and wild-type systems computed for this study. The combined length of calculated MD totals 20.71 µs.

Mutant system	Reaction state	Replicate	Length / ns
K13Q	PS	1	620
		2	700
	RS	1	700
		2	500
K13A	PS	1	300
R36A	PS	1	300
		2	300
	RS	1	500
		2	500
R88A	PS	1	700
		2	700
	RS	1	700
		2	700
R123A	PS	1	300
		2	300
	RS	1	500
		2	500
R156K	PS	1	1000
		2	300
	RS	1	1200
		2	700
R156A	PS	1	300
	RS	1	300
D158A	PS	1	300
		2	300
	RS	1	500
		2	500
R167A	PS	1	300
		2	300
		3	300
	RS	1	300
		2	300
		3	300
Wild-type	PS	1	1000
		2	600
		3	600
	RS	1	1290
		2	600
		3	600

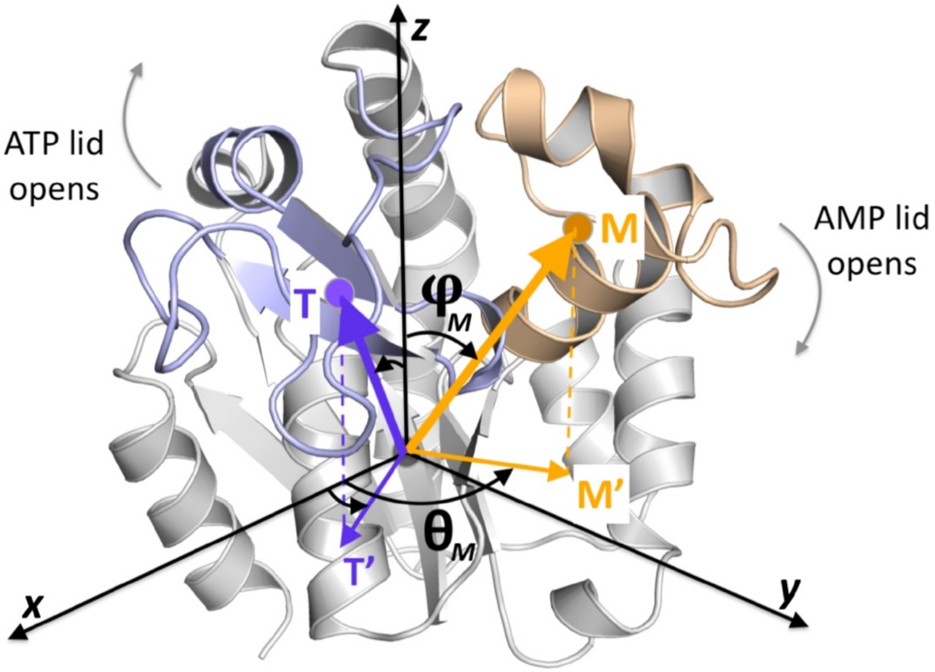


Figure S1. An illustration of the reference points used as domain centroids for the center-of-mass (COM) distance plots, where point *T* is the ATPlid COM, point *M* is the AMPlid COM, and the core's COM is at the origin (*Org*). When the COM distances are converted into polar coordinates, the lids' horizontal motion within the xy-plane is described by angle θ, measured as the angle between the x-axis and the projection of the *Org-T/M* vector onto the xy-plane (*Org-T'/M')* and angle ϕ represents the vertical displacement perpendicular to the xy-plane, measured as the angle between the *Org-T* vector and the z-axis. Subscripts $_M$ and $_T$ stand for polar angles associated with the ATP and AMP lid, respectively.

The grey arrows indicate a typical pattern of lid motion upon enzyme opening, which can be helpful to note when viewing the COM plots in polar coordinates as it will become apparent that those can be interpreted intuitively.

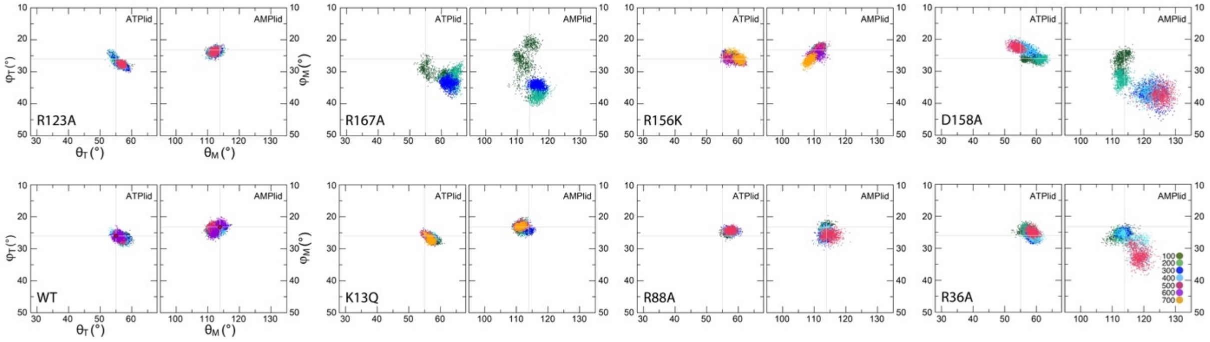


Figure S2. Centre-of-mass positions converted into the polar coordinate system to capture a pattern of lid domain motions in the reactant state, RS. Data points are organized by color into 100 ns segments, totaling up to 700 ns of the MD. The cross hairs indicate the reference values from 1AKE X-ray structure. In mutant cases which open, a variety of modes of lid motion can be observed. For instance, in R167A (RS), within ~100 ns the ATP lid collides against the AMP lid (θ_T increases by ~30 °). Rigidity of the K13Q mutant is interpreted as a smaller amplitude of fluctuations, comparing to the WT, and tightening of both lids against each other with respect to the reference X-ray position.

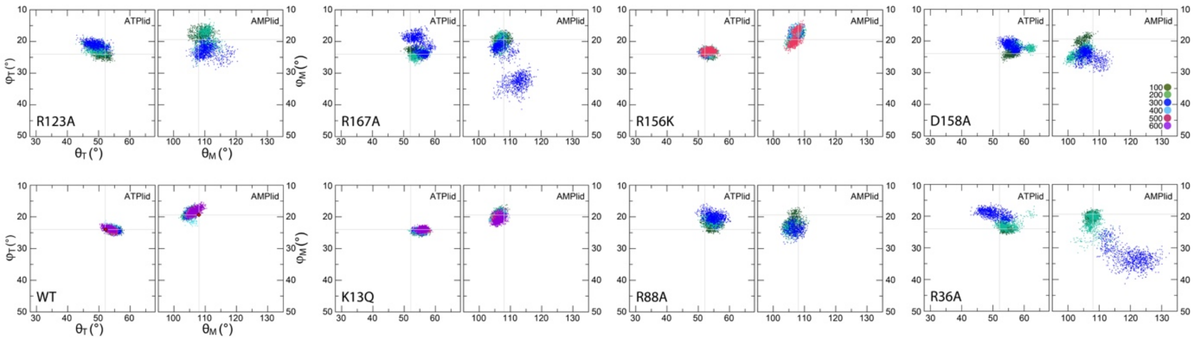


Figure S3. Centre-of-mass positions converted into the polar coordinate system to capture a pattern of lid domain motions in the product state, PS. Data points are organized by color into 100 ns segments, totaling up to 700 ns of the MD. The cross hairs indicate the reference values from 1AKE X-ray structure. In mutant cases which open, a variety of modes of lid motion can be observed. For instance, in R36A (PS) and D158A (PS), between ~100-200 ns the ATP lid collides against the AMP lid, while R167A shows an erratic opening behavior.

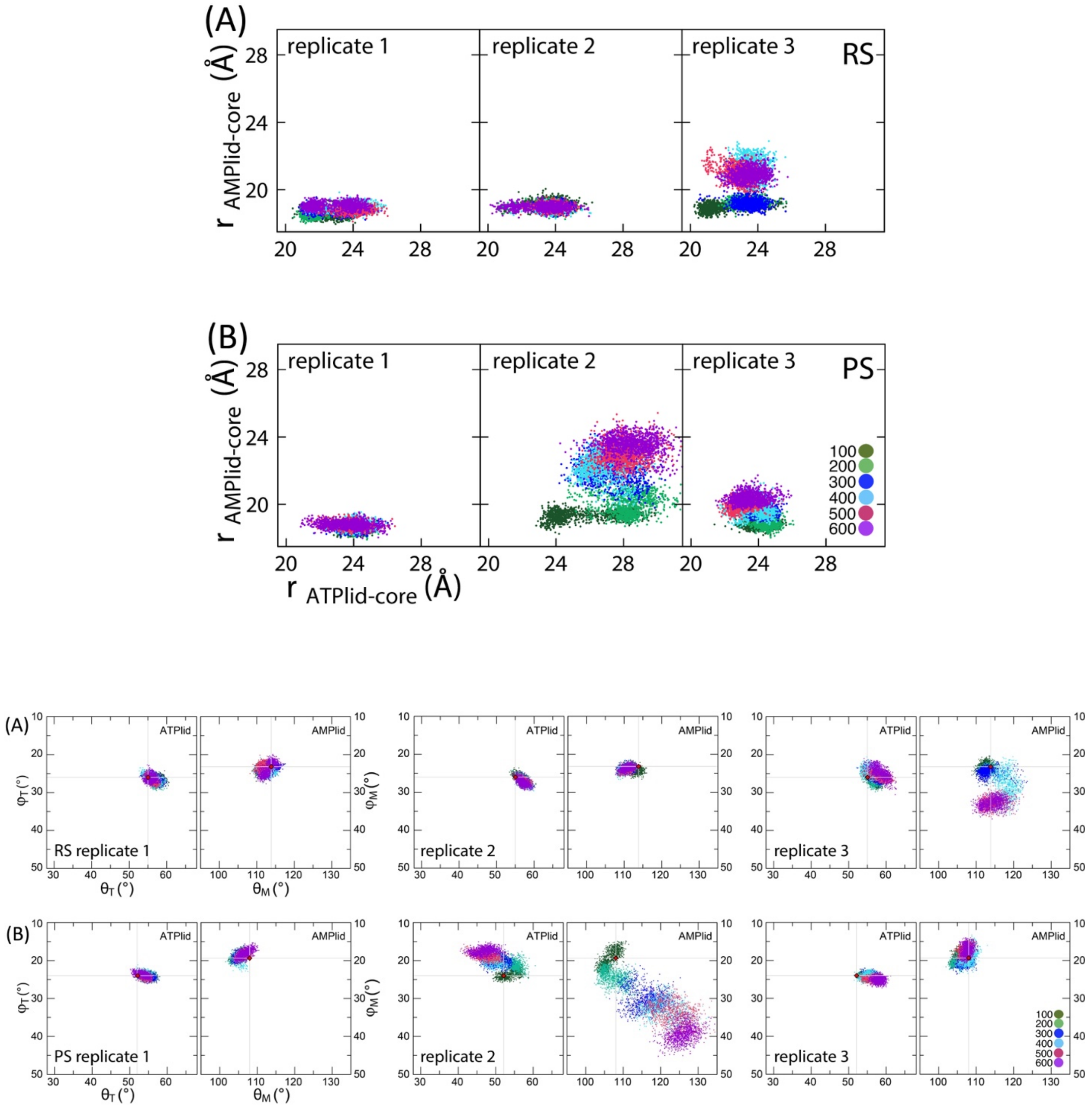


Figure S4. Center-of-mass graphs for the triplicated molecular dynamics (MD) simulations of the wild-type AdK in reactant **(A)** and product **(B)** states plotted as lid domain to core distances *(top two rows)* to show the magnitude of lid opening, and angular displacement in spherical coordinates *(bottom two rows)* to show its pattern. Trajectories are color-coded into 100 ns segments, totaling 600 ns of MD for each system. The cross hairs indicate the reference values from 1AKE X-ray structure. Only replicate 1 in PS fully opens, where the ATP lid lifts by ~6Å, and the AMP lid deflects by ~5Å. Displacement of the AMP lid with respect to the core by ~3-4Å in replicate 2, of the PS and RS, does not develop into an opening event, reflecting only the AMP lid fluctuations.

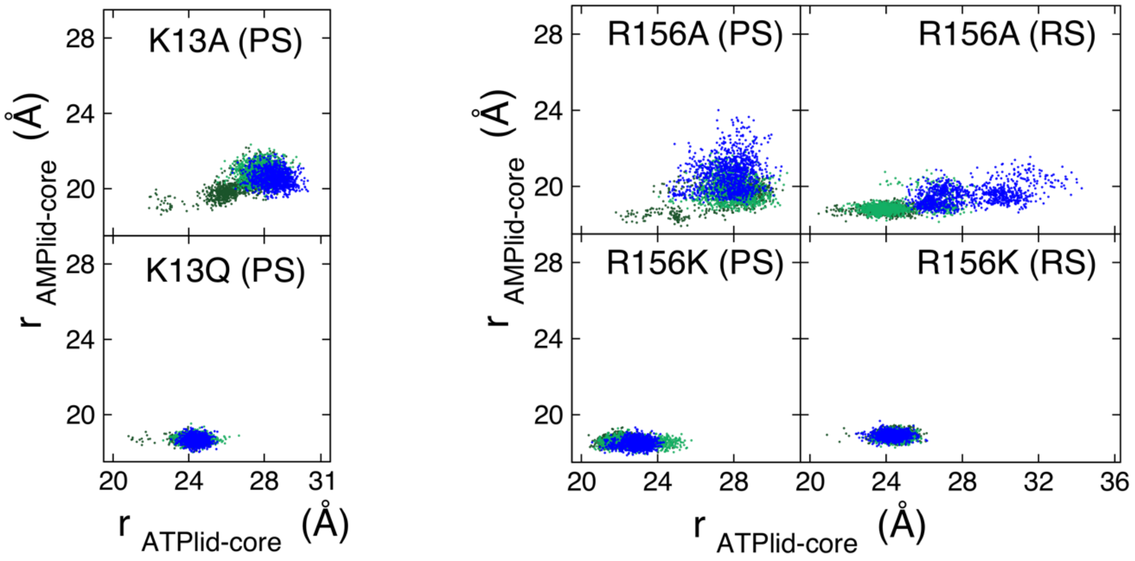


Figure S5. A proof-of-concept comparison of the dynamic behavior between K13Q and R156K *bottom panels)* with their to-alanine mutations *(top panels*).Removing any interactions from the amide Q13 and ammonium K156 groups, the new mutants open within the time frame of >300 ns, which is similar to what we observe in other to-alanine mutant cases. (RS: reactant state, PS: product state)

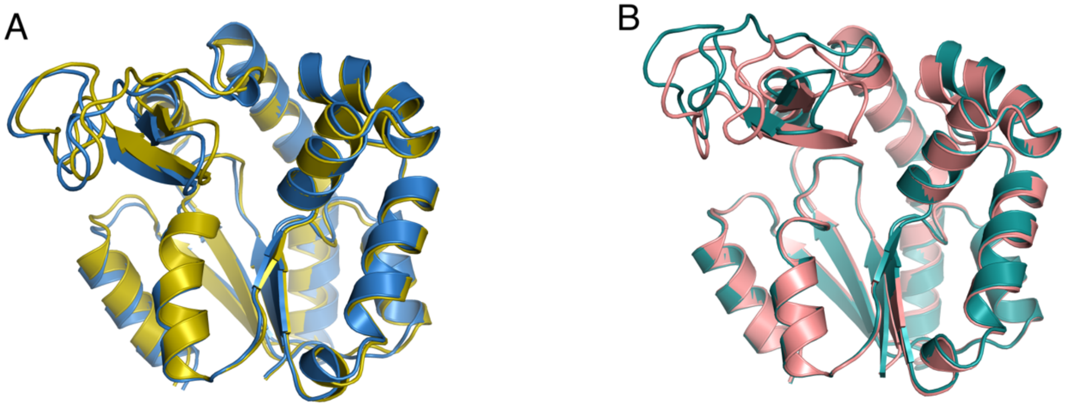
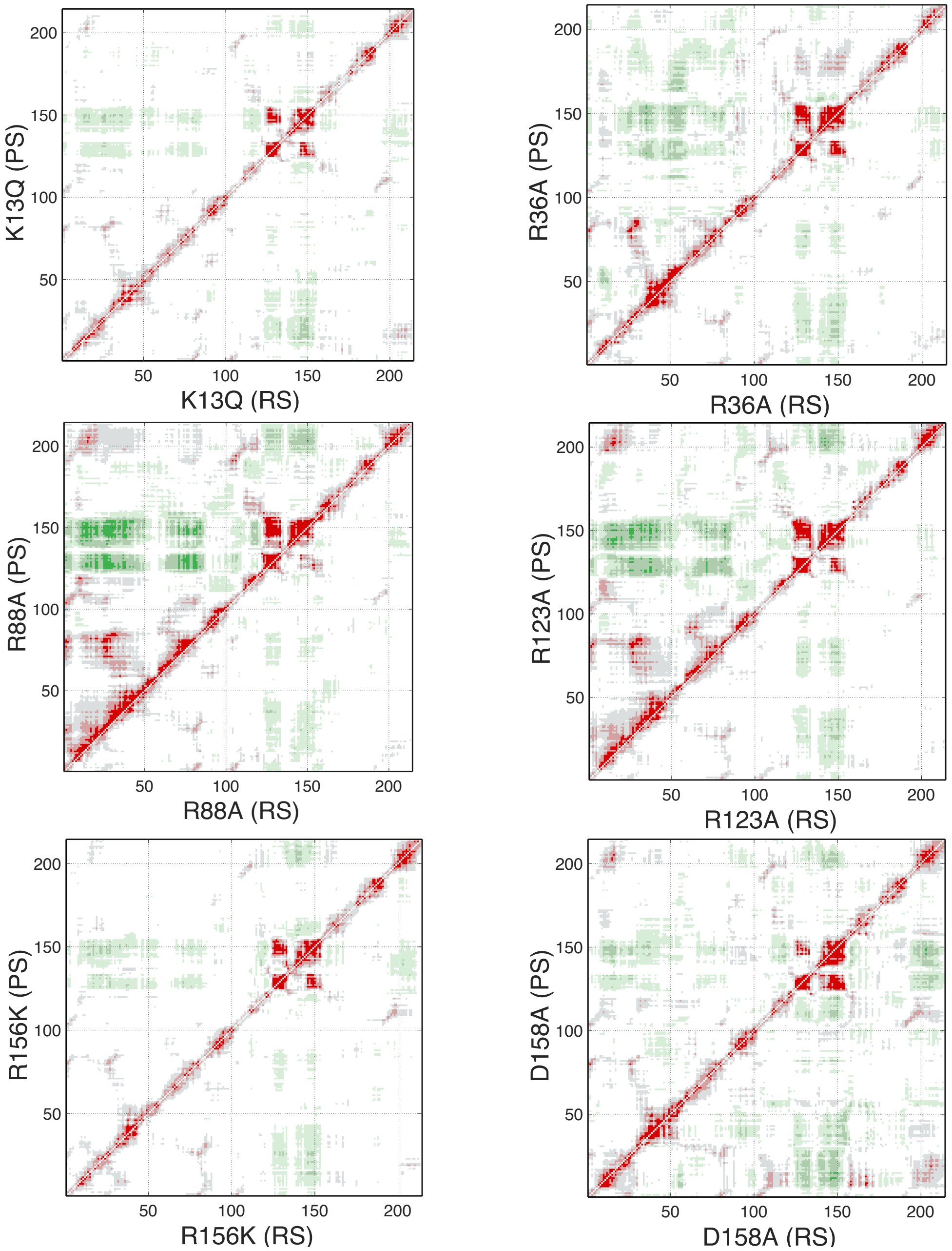


Figure S6. Comparison of average structures from WT MD simulations: (**A**) RS between replicate 1 (olive) and replicate 3 (blue) and (**B**) PS between replicate 1 (pink) and replicate 2 (teal). The PS replicate 2, which eventually opens during MD simulation (Figure S4), the ATP and AMP lids are slightly in an open conformation compared to the closed-conformation replica 1 as well as all RS replicates.



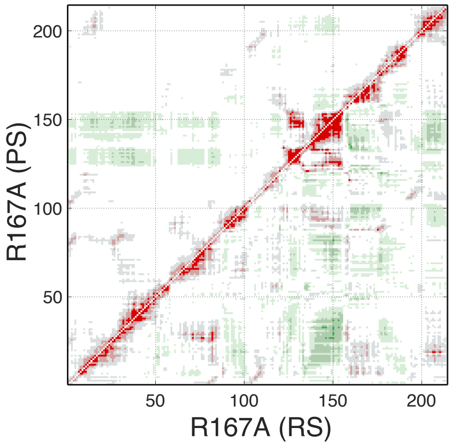


Figure S7. Covariance analysis of the side-chain fluctuations in AdK mutant variants in product (PS) and reactant (RS) states. Of importance is the intensity of color indicating a greater statistical significance of the correlation (red)/anticorrelation (green) in residue fluctuations as well as the extent of the protein affected by the ordered motion. Notably, in the more dynamic PS the residue motion tends to be more ordered than in the RS.

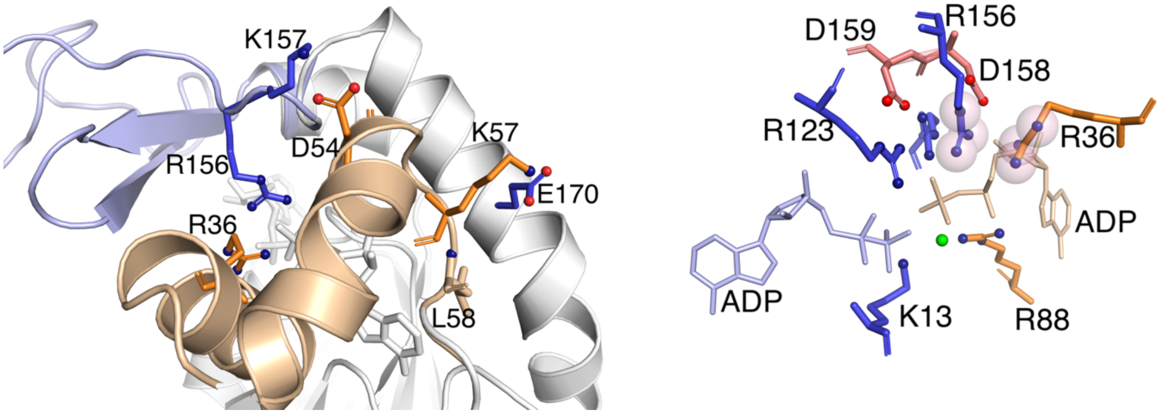


Figure S8. ***(left)*** The three most significant electrostatic interactions on the interface between the ATP and AMP lid domains important to the maintenance of the AdK's closed conformation. They are present in all mutant and wild-type simulations to a varying degree of robustness. As the enzyme transitions into an open conformation, the R156-R36 π-stacking *(shown on the* ***right****)* is lost first, followed by the disengagement of the K157-D54 pair. The remaing pair, E170-K57, is the last one to break, after which point the enzyme has moved past a point of no return and will infallibly open. L58 is included to acknowledge the fact that in the X-ray structure E170 coordinates to its backbone nitrogen atom instead. Under MD conditions, this interaction is not sustained and preferably shifts to K57 side chain.

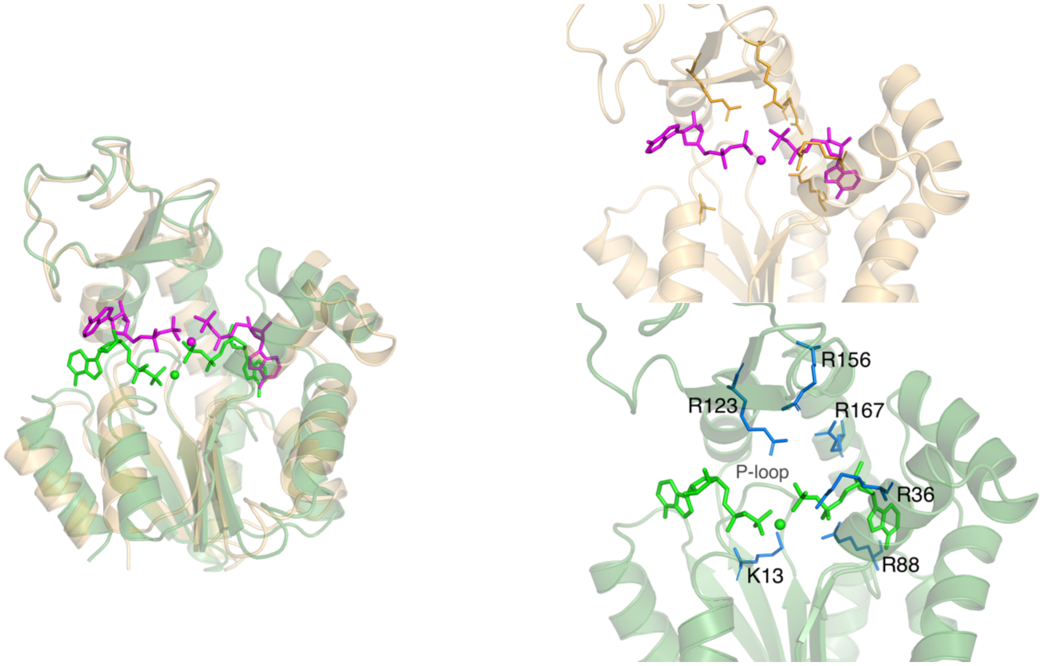


Figure S9. (***left***) Snapshots from an MD trajectory showing superimposed AdK wild-type *(protein: green; ligands/Mg^{2+}: light green)* and the K13A system *(protein: gold; ligands/Mg^{2+}: purple)* whose ATP lids are starting to lift, shortly before enzyme opening . Given the same arrangement of the ATP lid in both, it can be seen that the loss of coordination from K13 allows the ADP (in the ATP-binding domain) to drift away from the P-loop and be lifted together with the ATP lid. (***right***) The orientations of the catalytic residues in the K13A mutant *(top)* and the wild-type AdK *(bottom)*. R123, R156, and R167 hold on to the ligands in K13A while the ATP lid goes up. R88 is under visible strain to maintain its coordination. In contrast, the wild-type ATP lid severs all interactions to the ligands, which are held in place by K13, R36 and R88.

Conclusion

Computer-based research complements the experiment by providing insight at levels often not accessible to direct investigation. In the case of my work that would mean attempting to provide a mechanistic explanation of the inner workings of enzyme's active site residues down to the atomic level: Which one interacts with which other? What is the role of this interaction? How does this affect observable qualities recorded and quantified by the experiment? All that happening on the periphery of the real conundrum, the epistemological elephant in every scientist's room: how do I know if what I know is true? My short experience as a doctoral student made obvious to me that it is the questions that are the true Holy Grail of academic research, not the answers. Answers are but ephemeral entities, useful only in as much as they can provide a decent springboard for a new question, with subsequent answers to which solidly corroborating any previous findings. And such proved to be the nature of this research into the coupling between the catalytic step of phosphoryl transfer from ATP to AMP and the mechanism of enzyme opening in the adenylate kinase (AK) protein: jumping from one question to another, getting bruised by contortionist methods of data analysis.

With Chapter 1 I was dipping my toe into the field by performing molecular dynamics simulations on a mitochondrial, AK3 isoform of the adenylate kinase. The question was to identify a molecular reason behind the ability of AK3 to recognize its substrate GTP, and to preferentially select for it over ATP. My contribution was to demonstrate that GTP was able to sandwich itself between two residues in the catalytic lid domain, Arg124 and Phe142, which held the aromatic system of the guanosine moiety in place. This stacking interaction was possible due to the formation of a hydrogen bond between the GTP's carbonyl oxygen and the amide proton on the Thr201.

My contribution to research presented in Chapter 2 is foundational and involved qualitative data analysis on the simulations of the wild-type AK and its Arg167Ala variant. This was to report a survey of interactions within the enzyme's active site, upon which observations was built the subsequent project in Chapter 3. Chapter 3 carries my contribution to the development of ideas and explanation of the roles that the catalytic residues play in the wild-type AK enzyme. Propped by generous input from Drs Pedro Ojeda-May and Kwangho Nam who calculated the values for energy barriers in the catalytic step using the QM/MM approach, I used MD to discern the functions of each catalytic residue and elucidate how they trigger the enzyme to open. I proposed, contrary to our initial hypothesis that each residue might have a clearly defined task, that they in fact grouped together and the opening was an event that stemmed from their emergent action as a collective. The group in question comprises of arginines 123, 156, 167, and 36 together with their coordinating partners that are aspartates 158 and 159. The biggest unknow at the time of writing is the role of arginine 88. The data are inconclusive, and the topic needs further investigation. Two possible avenues to explore would be (i) how/if the absence of Arg88 may strengthen interdomain interactions between the Lys157—Asp54 and Lys57—Glu170 pairs (Figure S8, Appendix III) and keep the enzyme in the closed position and (ii) how/if its absence may cause the hosting α-helix 5 (Figure 1A, Chapter 3) to lift and push the AMP-binding domain and α-helix 7 to apart to precipitate enzyme opening.