# Supporting Information

## Aviram et al. 10.1073/pnas.1720448115

### SI Methods

**Protein Expression and Labeling.** A pET15b expression vector containing the *E. coli* AK C77S gene with six-residue histidine tag at the N terminus was used to express the protein. Alanine-to-cysteine and valine-to-cysteine substitutions were introduced at positions 73 and 142 of the protein, respectively, by site-directed mutagenesis. These sites were selected to report on the conformational dynamics of the LID domain of AK. The sequence of AK C77S/V142C/A73C was verified by DNA sequencing. The mutant was overexpressed by transformation into *E. coli* BL21 competent cells. The cells were lysed using high-energy sonication or a French cell press, and the pellet was then separated from the protein-containing supernatant using high-speed centrifugation. The supernatant was separated on a $Ni^{2+}$ Sepharose column (GE Healthcare HisTrap HP). Fractions containing AK were pooled and run on a second gel filtration column (HiLoad 16/60 Superdex 75 prep grade; GE Healthcare) and eluted as a single peak. Labeling reactions were performed by first incubating protein samples with Alexa 594 maleimide (Invitrogen) at a molar ratio of 75%, separating labeled from unlabeled protein on a mono-Q 5/50 GL column (GE Healthcare) and later-on labeling with an excess of Alexa 488 maleimide (Invitrogen).

**Steady-State Enzyme Activity Measurements.** The enzymatic activity of AK variants was measured at 25 °C by monitoring the absorbance of NADH at 340 nm using the coupled reaction assay described by Huss et al. (1):

$$Mg^{2+}ATP + AMP \xrightarrow{AK} Mg^{2+}ADP + ADP$$

$$ADP + phosphoenolpyruvate \xrightarrow{pyruvate\ kinase} ATP + pyruvate$$

$$Pyruvate + NADH \xrightarrow{Lactate\ dehydrogenase} Lactate + NAD^{+}$$

Generally, each reaction was initiated by adding 2 mM $MgCl_2$ to a mixture solution containing 2 nM AK, 1 mM of either ATP or AMP, varying concentrations of the second substrate (i.e., AMP or ATP), 4 mM phosphoenolpyruvic acid, 300 μM NADH, 10 units/mL pyruvate kinase, 15 units/mL lactate dehydrogenase, 0.25 mg/mL BSA, 0.1 M KCl (Merck), all in a 0.1 M Tris buffer at pH 8. All reagents were from Sigma.

The initial velocity of the enzymatic reaction was calculated from the temporal variation of the NADH absorption. Initial velocity values were plotted as a function of substrate concentration to generate Michaelis–Menten curves and to extract the Michaelis constant and the turnover number ($K_m$ and $k_{cat}$, respectively).

**Sample Preparation for Single-Molecule Experiments.** Flow cells were prepared from two coverslip glasses of 24 × 50 mm and 18 × 18 mm (Marienfeld), initially rinsed with a 10% hydrogen fluoride solution (Merc), and then washed with pure water. Dried coverslips were glued together using Teflon strips to form the sample cells. To prevent protein adhesion to the glass, flow cells were coated with a supported lipid bilayer by flowing in a solution containing liposomes prepared in advance from egg phosphatidylcholine (Avanti Lipids) by extrusion through disposable 0.1 μm Anopore syringe filters (Whatman Anotop-10). Flow cells were then filled with a mixture of 15 pM-labeled enzyme, 2 mM $MgCl_2$, 0.01% Tween (Thermo Fisher), and substrates (ATP, ADP, AMP, Ap5A, AMP-PNP; Sigma) at various concentrations. The cells were sealed with silicon grease to prevent water evaporation during measurement and immediately brought to the measurement setup.

For the experiment to be at equilibrium such that the substrate concentrations would not change during the measurement, a fixed concentration of AMP (1 mM) and varied concentrations of ATP were added to the solution. The appropriate ADP concentration to guarantee equilibrium (zero flux) was calculated using the following rate equation (2):

$$v \propto \frac{K_1[M][T]}{K_M K_T} - \frac{K_2[D_1][D_2]}{K_{D1}K_{D2}},$$

where $K_1$, $K_2$ are the equilibrium constants for the reaction, and [M] and [T] are the AMP and ATP concentrations, respectively. $[D_1]$, the concentration of ADP bound to the NMPbind domain can be calculated from the total ADP concentration ([D]) by $[D_1] = [D] - [D_2]$.

$[D_2]$ is calculated using

$$[D_2] = \frac{\frac{1}{K_{Mg}} + [D] + [Mg] - \sqrt{\left(\frac{1}{K_{Mg}} + [D] + [Mg]\right)^2 - 4 \cdot [D] \cdot [Mg]}}{2},$$

where [Mg] and $K_{Mg}$ are the concentration and dissociation constant for magnesium, respectively.

FRET efficiency histograms from the first and last 15 or 30 min of each measurement were overlapped to validate that indeed the substrate concentration did not change during the measurement (Fig. S1).

**Single-Molecule FRET Data Acquisition.** An inverted confocal microscope (3) was used for data acquisition. A 488-nm laser beam was focused into the sample solution (10 μm deep). Molecules diffusing through the beam created short bursts of photons that were divided into two channels according to their wavelengths, using a dichroic mirror (FF580-FDi01; Semrock) and filtered by band-pass filters (ET-535/70m for the donor channel and ET-630/75m for the acceptor channel, both from Chroma). Arrival times of these photons were registered by two single-photon avalanche photodiodes (SPCM-AQR-15; Perkin-Elmer) coupled to a standalone time-correlated single-photon counting module (PicoHarp 300, PicoQuant). Data were collected continuously at 22 °C for up to 2 h before the sample was replaced. At least three data sets containing a total of ∼15,000 photon bursts were collected for each substrate condition with an average photon flux of ∼450,000 per second.

**Detection of Photon Bursts and FRET Efficiency Histograms.** We detected fluorescent bursts in the single-molecule data using similar methods to those reported previously (4). The task of separating background photons from those related to fluorescent bursts was facilitated by first smoothing the data with a running average of 15 photons. A cutoff time of 5 μs was determined from the histogram of the time lags, and used to effectively find the start and end points of each burst. Only fluorescence bursts with a total of 50 photons or more were selected for further analysis. Data were not corrected for channel cross-talk or background.

FRET efficiency was then calculated for each burst as the number of photons arriving from the acceptor ($I_A$) channel divided

by the total number of photons ($I_A + I_D$). FRET efficiency histograms were constructed from the calculated values.

## Validation Tests

To verify that the $H^2$MM analysis yields parameters that best fit the data, we performed three different validations test, described in detail below.

**Recoloring Method.** We adopted a method first introduced by Gopich and Szabo and termed by them "recoloring" (5). In this method the arrival times of photons in each data set are retained, but their "colors" (i.e., whether they belong to the donor or acceptor) are erased. A stochastic simulation based on the $H^2$MM parameters is used to reassign the photons to the two experimental channels. FRET efficiency values can then be calculated for each fluorescence burst separately, and a FRET efficiency histogram can be plotted and compared with the experimental histogram. Fig. S3 demonstrates recoloring of a series of histograms, showing very good agreement between simulation and experiment.

**Segmentation Analysis.** We used the $H^2$MM parameters and the Viterbi algorithm to assign segments of data in fluorescence bursts that belong to either open or closed state. We then calculated the average FRET efficiency for each segment and plotted histograms of the segments belonging to each state. The results are shown for three ATP concentrations in Fig. S4. In each case, the two histograms now nicely separate from each other (compared with the original histogram), because the averaging effect due to fast exchange between the states has been removed. This is strong evidence for the occurrence of two states in the data. Indeed, in a control calculation we segmented the data randomly (using the $H^2$MM rates but without using the Viterbi algorithm to assign segments), and in this case the two histograms are actually very similar to each other and to the original histogram (Fig. S5).

**Dwell-Time Analysis.** We calculated integrated dwell-time distributions for each of the two states in each data set. For this purpose we used an algorithm developed in-house that takes into account not only the most probable state path for each trajectory, as in the Viterbi algorithm, but all possible state paths (6). The integrated dwell-time distributions (Fig. S6) were fitted to single exponential functions, and the resulting rates are shown in Table S2 to compare very favorably to the rates directly extracted by $H^2$MM.


1. Huss RJ, Glaser M (1983) Identification and purification of an adenylate kinase-associated protein that influences the thermolability of adenylate kinase from a temperature-sensitive adk mutant of Escherichia coli. *J Biol Chem* 258:13370–13376.
2. Sheng XR, Li X, Pan XM (1999) An iso-random Bi Bi mechanism for adenylate kinase. *J Biol Chem* 274:22238–22242.
3. Cohen SS, et al. (2015) Probing the molecular origin of native-state flexibility in repeat proteins. *J Am Chem Soc* 137:10367–10373.
4. Sherman E, Haran G (2006) Coil-globule transition in the denatured state of a small protein. *Proc Natl Acad Sci USA* 103:11539–11543.
5. Gopich IV, Szabo A (2009) Decoding the pattern of photon colors in single-molecule FRET. *J Phys Chem B* 113:10965–10973.
6. Pirchi M (2013) Mapping the free-energy landscape of multi-domain proteins by single-molecule FRET spectroscopy. PhD thesis (Weizmann Institute of Science, Rehovot, Israel).

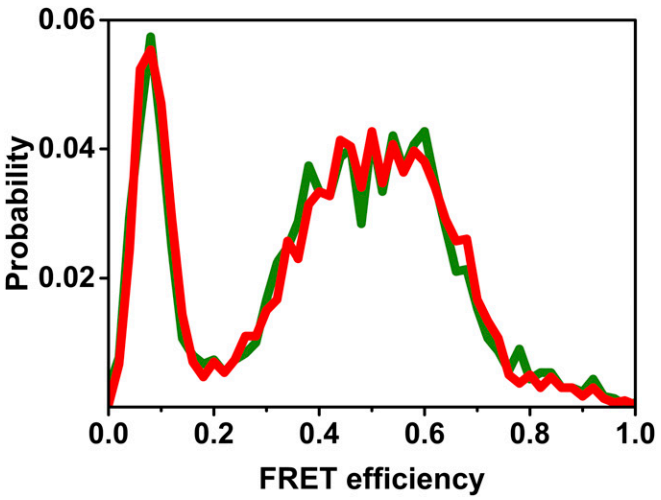




**Fig. S1.** Single-molecule experiments under equilibrium. As discussed in the text, ATP and ADP concentrations were selected to maintain equilibrium in each of the experiments throughout data collection. To validate the equilibrium condition, we calculated FRET efficiency histograms from the first and last part of each measurement and verified that they overlap well. Here we show the result of a 2-h-long measurement, with substrate concentrations selected to guarantee high sensitivity to concentration changes (2.5 μM ATP, 1 mM AMP, and 8.4 μM ADP). Histograms from the first (green) and last (red) half hour of the measurement match each other very well.

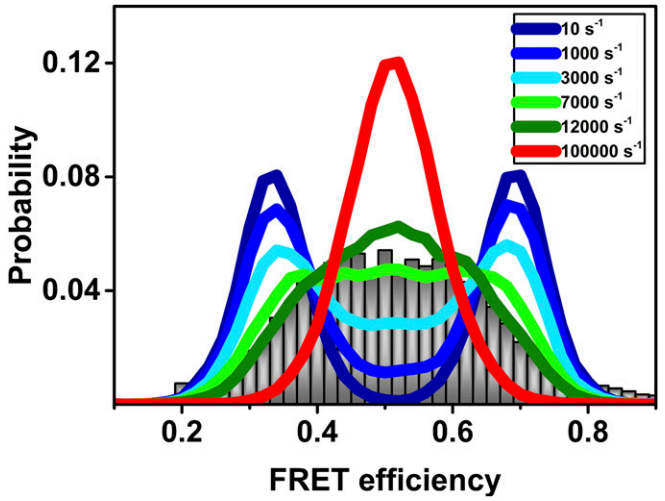



**Fig. S2.** Estimating the rate of domain closure. FRET efficiency histograms were calculated from data generated using a simple Monte Carlo simulation based on a two states model with varying rates. The fluorescent bursts were taken from a smFRET experiment in the presence of 2.5 μM ATP, 1 mM AMP, and 8.4 μM ADP, and photon colors were assigned stochastically according to the model parameters. The closing and opening rates were assumed to be equal for simplification. The generated histograms were then compared with the histogram calculated from the experimental data (gray). A rate of ~10,000 $s^{-1}$ seemed to match the experimental histogram best.

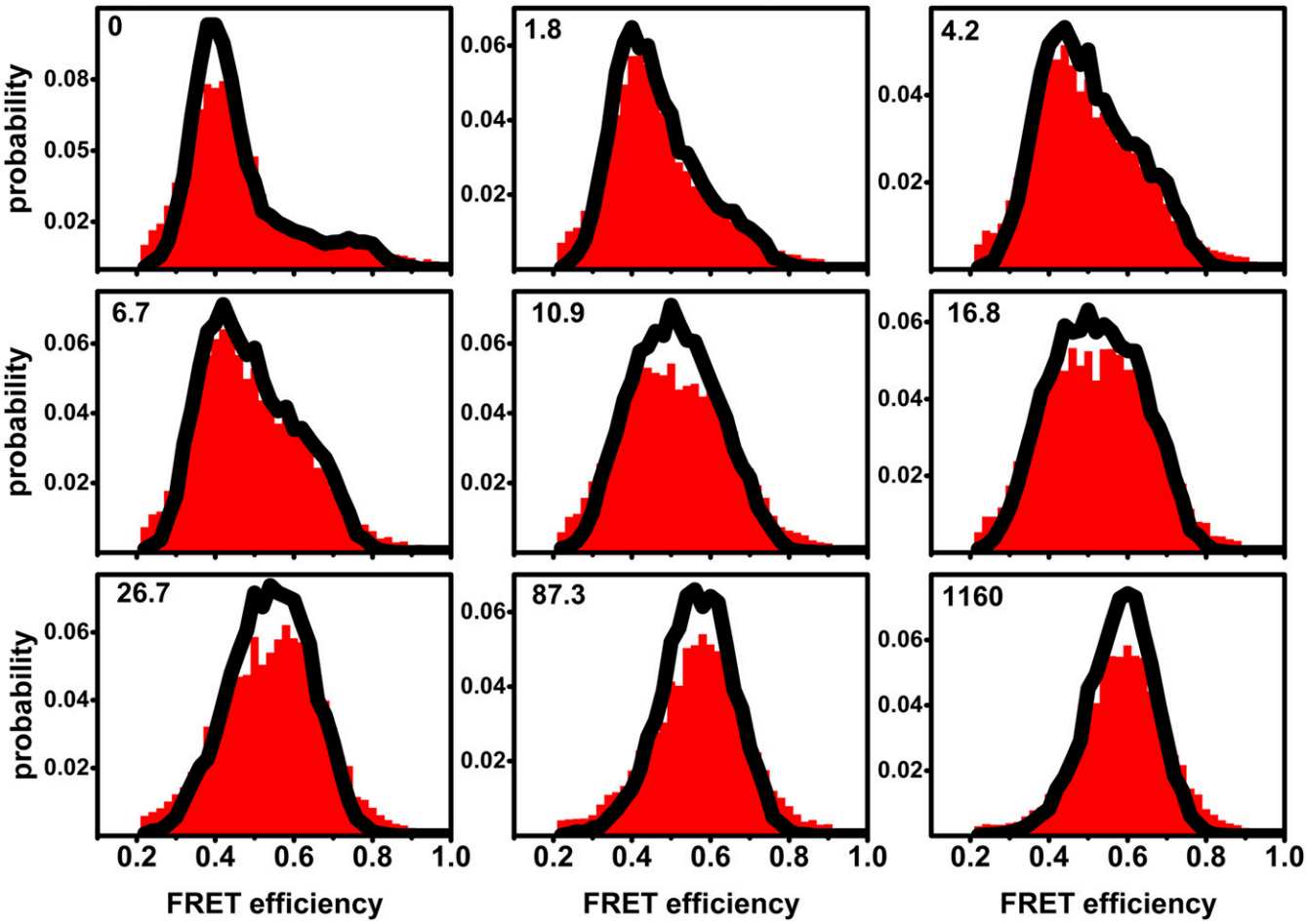



**Fig. S3.** Validating $H^2MM$ models using the recoloring algorithm. Fluorescence bursts with a FRET efficiency larger than 0.2 were used together with the recoloring algorithm described in *Methods* to obtain synthetic data sets from which FRET efficiency histograms were calculated. The figure shows a series of experiments, with the experimental histograms in red and the recolored histograms in black. The substrate concentration (ATP + ADP) in μM is shown in each panel.

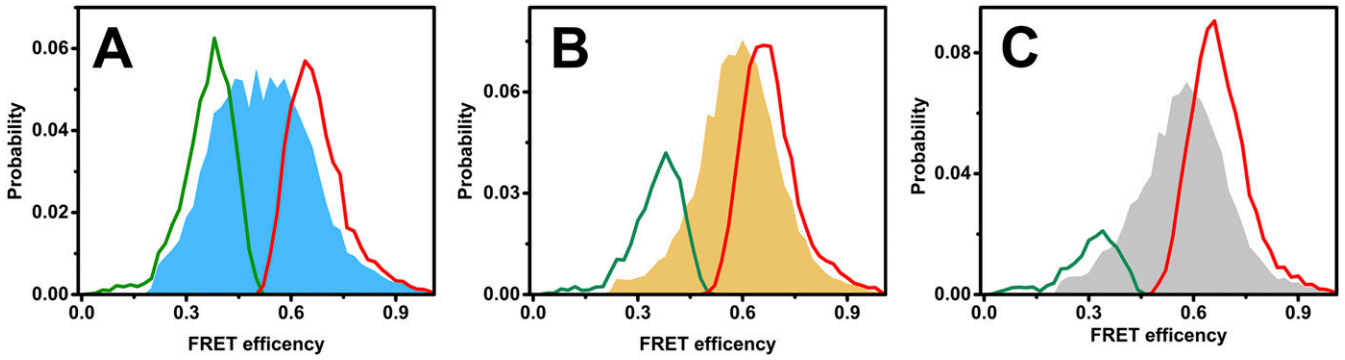



**Fig. S4.** Segmentation analysis of fluorescence bursts. Bursts were segmented as discussed in *SI Text* into segments belonging to either the open or the closed state, and FRET efficiency histograms were plotted together with the original histograms. In each panel the green histogram is for the open state and the red histogram for the closed state; the solid histogram is the original. The two segmented histograms separate well, as opposed to the original, in which fast exchange mixes the states. (*A*) 2.5 μM ATP. (*B*) 10 μM ATP. (*C*) 50 μM ATP.

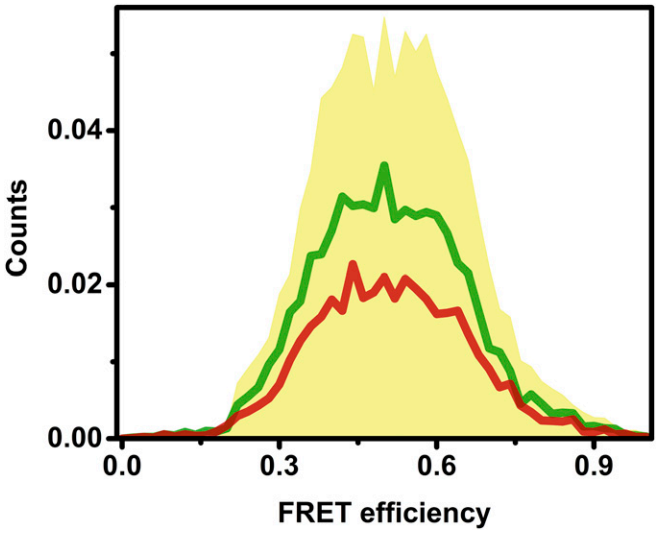



**Fig. S5.** Control segmentation analysis of fluorescence bursts. Bursts were segmented randomly (without using the Viterbi algorithm to assign states) and the FRET efficiency histograms of segments belonging to the two states were plotted. In this case the histograms do not separate and resemble the original histogram.

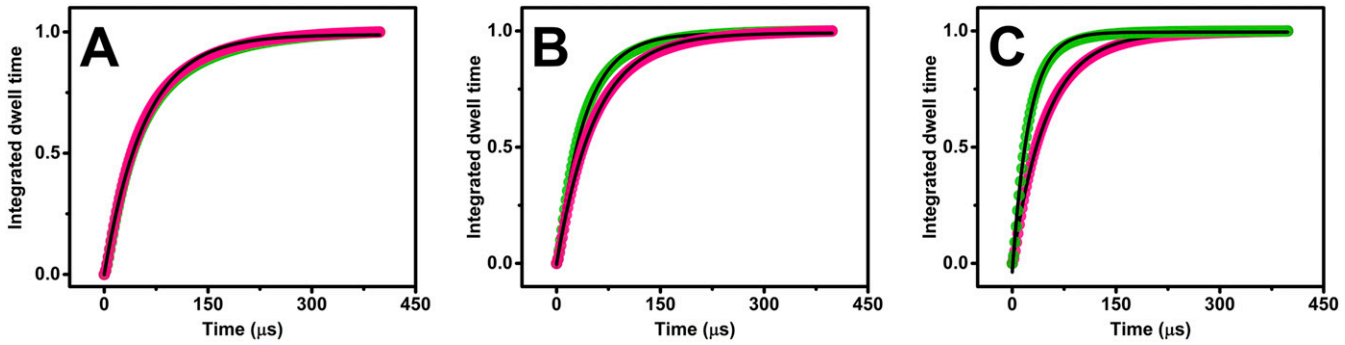



**Fig. S6.** Integrated dwell-time distributions. Using an algorithm that takes into account all possible state sequences, integrated dwell-time distributions were calculated for the open state (green) and for the closed state (red). Black lines are fits to single-exponential functions, and extracted rates are reported in Table S2. (*A*) 2.5 μM ATP (curves for open and closed states overlap). (*B*) 10 μM ATP. (*C*) 50 μM ATP.

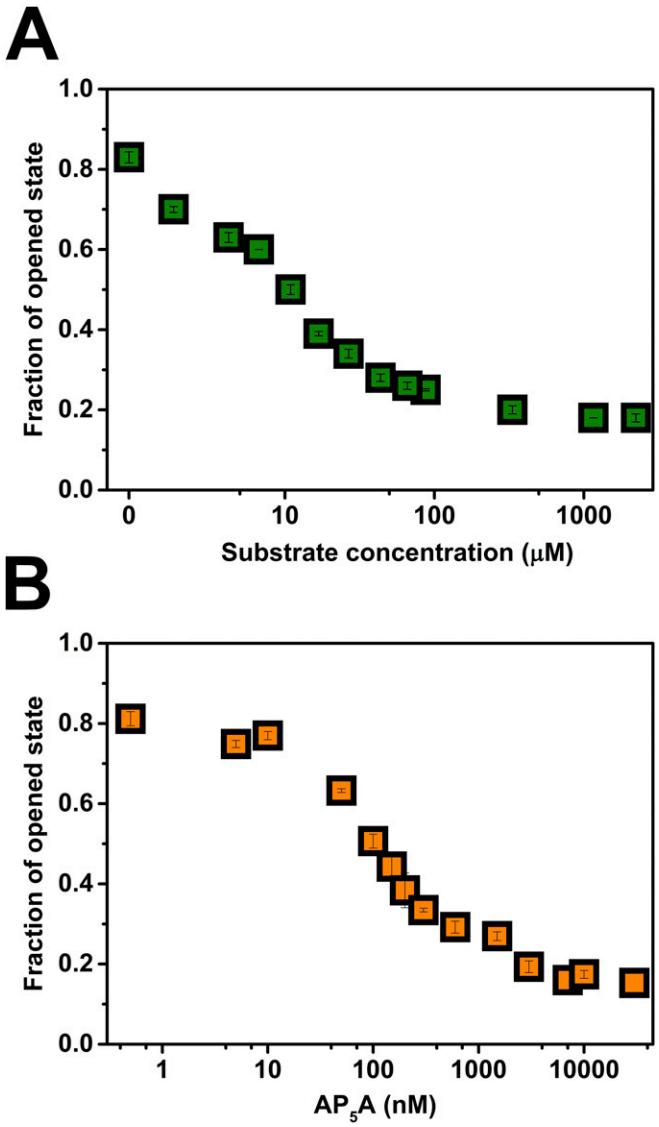



**Fig. S7.** Gradual shift from open to closed conformation. (*A*) The population of the open state of AK in the presence of increasing substrate concentrations (ATP + ADP), as obtained from the $H^2$MM analysis, are shown in green squares. (*B*) The population of the open state of AK in the presence of increasing concentrations of $AP_5A$ were obtained from the $H^2$MM analysis, and are shown in orange squares. Standard errors of the mean, based on two (*B*) and three (*A*) repeats of the experiments, are shown.

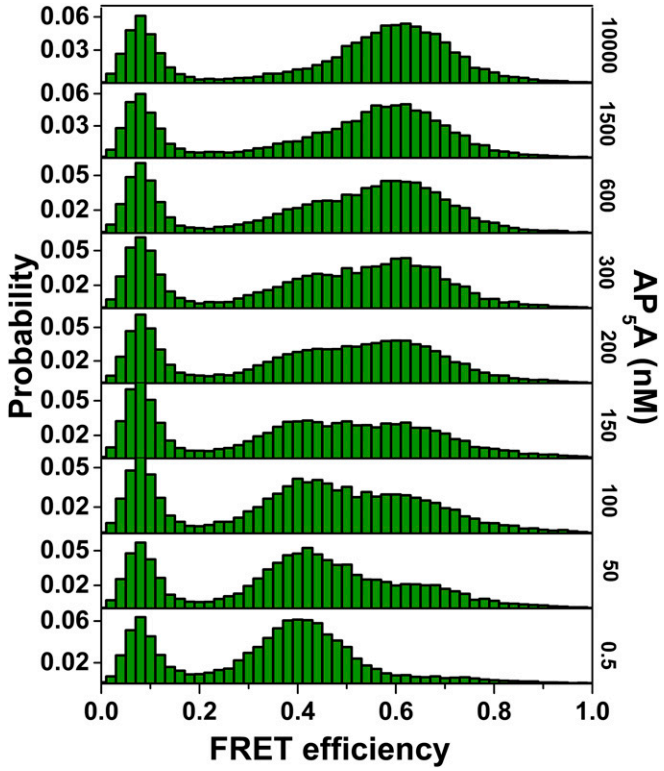



**Fig. S8.** FRET histograms of AK with $Ap_5A$ as substrate analog. Histograms were generated from experiments measured at increasing $AP_5A$ concentrations, indicated on the right-hand side. The gradual increase of the population of the closed state with increasing substrate concentrations shows a similar trend to that of the native substrates, although it occurs in a different range of concentrations.

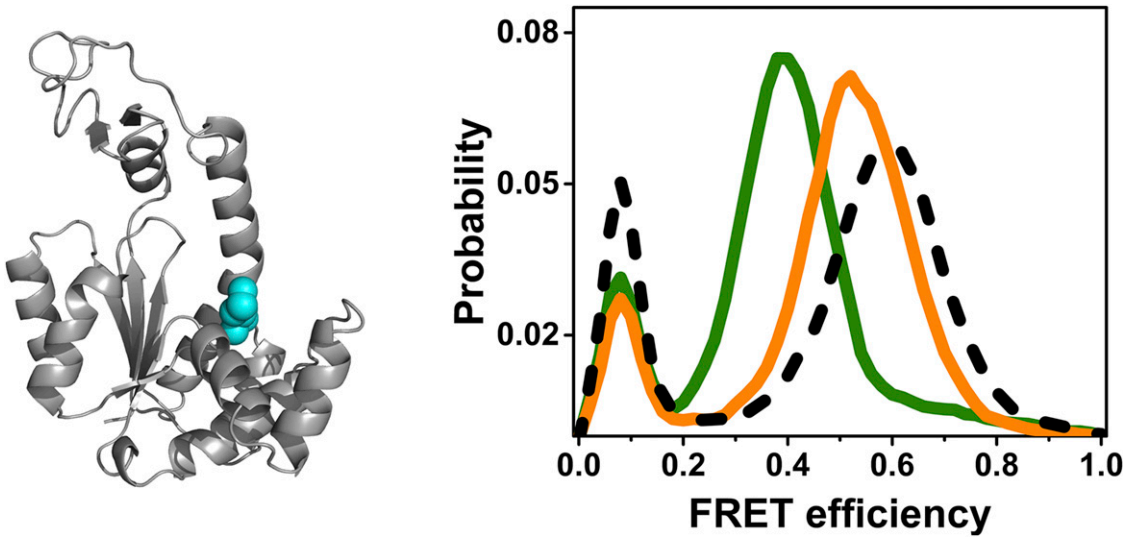
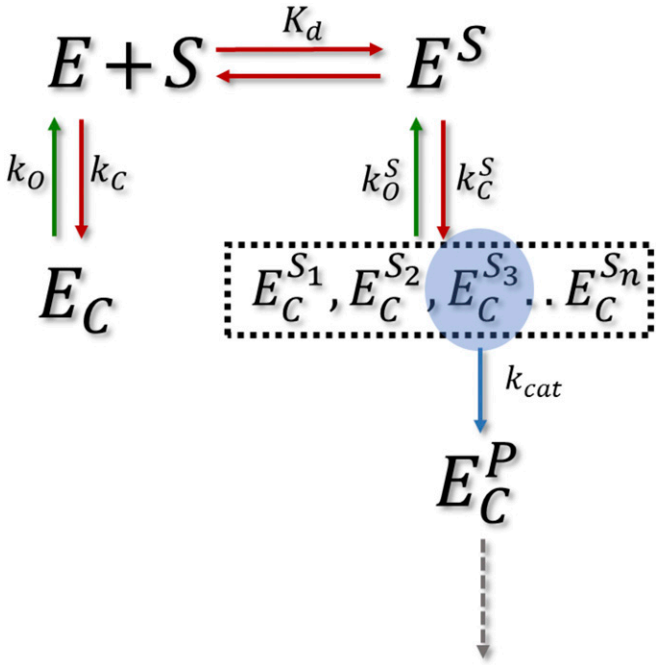



**Fig. S9.** A point mutation at arginine 88 barely affects the conformational dynamics of AK. (*Left*) Arginine 88 was replaced by methionine to generate mutant R88M (marked in cyan spheres). Unlike other mutations (Fig. 4), this mutation reduced the turnover rate (Table S2) without preventing switching from the unbound FEP to the bound FEP (*Right*). At saturating concentrations of the substrates (10 mM of ATP, 15 mM of AMP together with 0.6 mM of ADP), the population of the closed conformation in the FRET efficiency histogram (orange) was almost the same as that of the wild-type enzyme (dashed black line). Plotted in green is the FRET efficiency histogram of the apo mutant enzyme.




**Scheme S1.** Expanded kinetic model for domain closure of bound and unbound enzyme molecules. In this scheme, $E$ and $E_C$ are the unbound enzyme in its open and closed conformations, $E^S$ is the bound enzyme in its open conformation, and the $E_C^{S_i}$ terms refer to the bound enzyme in its closed conformation with different potential conformations of the substrates. $E_C^P$ is the enzyme in its closed state with bound products. $K_d$ is the substrate dissociation constant, $k_C$ and $k_O$ are the domain closing and opening rates of the unbound enzyme, and $k_C^S$ and $k_O^S$ are the domain closing and opening rates for the bound enzyme. Finally, $k_{cat}$ is the rate of the chemical step. The reaction can only occur from one substrate conformation (or perhaps a few conformations) out of the many possible ones in the closed state (colored blue). Achieving the right conformation requires multiple domain opening and closing cycles. The dashed arrow leads to all of the states that follow product formation, which are not explicitly depicted.

**Table S1. Substrate concentrations used in the experiments**

| Experiment | ATP, μM | ADP, μM | Total substrate, μM |
|---|---|---|---|
| 1 | 0.1 | 1.7 | 1.8 |
| 2 | 0.5 | 3.7 | 4.2 |
| 3 | 1 | 5.7 | 6.7 |
| 4 | 2.5 | 8.4 | 10.9 |
| 5 | 5 | 11.8 | 16.8 |
| 6 | 10 | 16.7 | 26.7 |
| 7 | 20 | 23.6 | 43.6 |
| 8 | 35 | 31 | 66 |
| 9 | 50 | 37.3 | 87.3 |
| 10 | 250 | 83 | 333 |
| 11 | 1,000 | 160 | 1,160 |
| 12 | 2,000 | 230 | 2,230 |

AMP concentration was fixed at 1 mM. For each ATP concentration, the concentration of ADP was calculated so as to maintain the system at equilibrium (the calculation is described in *Methods*).

**Table S2. Comparison of rates from H$^2$MM analysis with rates from integrated dwell-time distributions**

| ATP concentration, μM | Rate O → C H$^2$MM, Hz | Rate O → C dwell time, Hz | Rate C → O H$^2$MM, Hz | Rate C → O dwell time, Hz |
|---|---|---|---|---|
| 2.5 | 16,130 ± 1,200 | 16,500 ± 100 | 16,830 ± 100 | 17,200 ± 600 |
| 10 | 25,390 ± 300 | 24,800 ± 200 | 18,400 ± 800 | 17,800 ± 100 |
| 50 | 42,970 ± 2,000 | 40,350 ± 200 | 22,370 ± 1,000 | 20,200 ± 100 |

O, open state; C, closed state.