SI Methods and Materials

Protein Mutagenesis, Expression, and Purification. To allow for single-molecule FRET experiments, an AK enzyme from *E. coli* was appended with a $(His)_6$ tag for immobilization at the C terminus (1). Using site-directed mutagenesis, cysteine mutations were introduced in the Lid (A127C) and the Core (A194C) domains for site-specific labeling. A further mutation was introduced (C77A) to remove a single native cysteine from the gene. The mutant AK gene was cloned into pET-21a (Novagen) and induced in BL21(DE3)pLysS cells at 37°C for 3 h. Unless otherwise noted, all work with the expressed AK was done in AK reaction buffer: 100 mM Tris HCl (pH 7.5), 100 mM KCl, and 2 mM MgCl2. All chemicals were purchased from Sigma and used as received unless specified. After sonication and centrifugation, the protein was purified over a Ni^{2+} affinity column (HisTrap HP, Amersham Pharmacia) according to the manufacturer's instructions. The protein was further purified over a Q-Sepharose anion exchange column (Amersham Pharmacia) with a linear gradient from 0–500 mM KCl. Finally, aggregates were removed from the sample by running it over a 1-m S-200 gel filtration column (Amersham Pharmacia) (2).

Enzymatic Activity Assay. AK activity was measured following the protocol of Huss *et al.* using the reverse reaction, Mg^{2+}-ADP + ADP→Mg^{2+}-ATP + AMP (3). The coupled assay reaction mixture consisted of 0.1 M Tris·HCl (pH 7.5), 0.1 M KCl, 0.5 mM $NADP^+$, 10 mM glucose, 1 mM TCEP, and 1.2 units hexokinase/0.6 units glucose-6-phosphate dehydrogenase mixture (Sigma). To obtain sufficient data for modeling of kinetic schemes, Mg^{2+} and ADP concentrations were independently varied over a range of 0–3 mM. A total of 48 conditions were measured for each protein sample. Three consecutive measurements were averaged for each condition and error is reported to one standard deviation. Initial-velocity data were fit to a steady-state model assuming a random bi–bi model, where substrate binding is treated with a rapid equilibrium assumption (4). Actual concentrations of ADP and Mg^{2+}-ADP in solution were calculated by using $K_{D,ADP} = [Mg^{2+}][ADP]/[Mg^{2+}ADP] = 250\ \mu M$ (5) for the chemical equilibrium,

Deleted: ???

$Mg^{2+} + ADP \rightleftharpoons Mg^{2+}ADP$. Results from three consecutive experiments were averaged, and error is reported to one SD (SI Table 2 and SI Fig. 7).

Dye-Labeling and Sample Preparation. Thiol-reactive maleimide derivatives of Alexa Fluor 555 and 647 (Molecular Probes/Invitrogen, R_0 = 51 Å) were used as the donor and acceptor respectively for single-molecule FRET experiments. Dyes were dissolved in DMSO and reacted with ≈1 mM AK at a 5- to 10-fold molar excess for 3 h at room temperature. A 5-fold excess of TCEP hydrochloride was added to the reaction to prevent the formation of disulfide bonds. Unreacted dyes were removed from the protein sample with a Superdex 75 gel-filtration column (Amersham Pharmacia). Further enrichment of doubly labeled sample was achieved by reacting the protein with SulfoLink resin (Pierce) for 30 min at room temperature to remove proteins with remaining free thiol groups. After enrichment, dye labeling efficiency was estimated to be 93% by MALDI mass spectrometry. Dye-labeled enzymes were found to retain their activity however we cannot currently quantify the effects dye-labeling on the enzyme's conformational distribution or dynamics on our experimental timescale, if any.

Single Molecule Spectroscopy. Single-molecule FRET experiments with photon-by-photon time stamping were performed on a server-based confocal setup as described previously (6). Individual AK proteins were immobilized following a modified protocol proposed by Pal *et al.* (1). Biotinylated anti-$(His)_6$ antibody was purchased from Rockland and used to anchor the $(His)_6$-AK to a streptavidin functionalized PEG (Nektar Therapeutics) surface on quartz cover slips (7). Experiments were carried out in: 100 mM Tris·HCl, pH 7.5 and 100 mM KCl with 0.5 mM nucleotides as appropriate. Low excitation power (≈1 μW at the sample with an N.A. = 1.4 oil immersion objective, Olympus) was used for locating single-molecules to minimize photobleaching during scanning. To obtain high time-resolution data, high excitation power (≈2.5 μW) was used to acquire single-molecule trajectories. Emissions from both the donor and the acceptor dyes were acquired simultaneously using a pair of single-photon counting APDs where

the chronological arrival time of each detected photon was recorded for subsequent data analysis.

Circular Dichroism Spectroscopy. To investigate if dye labeling caused significant conformational changes in AK, circular dichroism (CD) spectroscopy was performed on the labeled and unlabeled C77A-A127C-A194C mutant proteins and the results were compared with those from untagged wild-type AK (SI Fig. 9). Protein concentration was determined by the method of Edelhoch (8) for unlabeled samples and by Bradford (9) with an unlabeled AK standard for dye-labeled samples. Dye labeled spectra show the same degree of secondary structure as the wild-type protein, which, together with the activity assay, indicates that mutations and dye labeling do not significantly perturb the global fold of the protein.

Steady-State FRET Spectroscopy. Steady-state FRET was performed on the dye labeled C77A-A127C-A194C mutant AK constructs used in single-molecule experiments to validate that distance measurements from bulk and single-molecule experiments are consistent. Experiments were performed on a fluorometer (Spex Fluorolog) with a 150W Xe Lamp and 2 mm slit widths. Donor dye was excited at 500 nm and emission was scanned from 530 to 750 nm. Acceptor dye was excited directly at 600 nm and emission was scanned from 630 to 750 nm. Energy transfer efficiency was determined using enhancement of acceptor emission presented by Clegg (10). Briefly, energy transfer efficiency is calculated by:

$$E_a = \left(\frac{F^{DA}(\lambda_D^{ex}, \lambda_A^{em}) - F^{DA}\left(\lambda_D^{ex}, \lambda_D^{em}\right)}{F^{DA}(\lambda_A^{ex}, \lambda_A^{em})} - \frac{\varepsilon^A(\lambda_D^{ex})}{\varepsilon^A(\lambda_A^{ex})} \right) \frac{\varepsilon^A(\lambda_A^{ex})}{\varepsilon^D(\lambda_D^{ex})} \frac{1}{f_A^{DA}} \qquad \textbf{[1]}$$

where $F^{DA}(\lambda_D^{ex}, \lambda_A^{em})$ is an emission scan of doubly labeled protein sample excited at λ_D^{ex}, the excitation wavelength for the donor, and scanned over λ_A^{em}, the emission range for the acceptor, $\varepsilon^A(\lambda_D^{ex})$ is the extinction coefficient of the acceptor dye at the emission wavelength used to excite the donor and f_A^{DA} is the fraction of total acceptor dye in

donor/acceptor labeled molecules from the MALDI spectrum and the overall dye concentration assuming random labeling. The extinction coefficients, $\varepsilon^A(\lambda_A^{ex})$ and $\varepsilon^D(\lambda_D^{ex})$ were appropriately scaled to account for differences in lamp intensities at the two excitation wavelengths.

Steady-state FRET experiments were performed with and without 0.5 mM AMP-PNP/AMP in 100 mM Tris (pH 7.4), 100 mM KCl, 2 mM $MgCl_2$ at a protein concentration of ~500 nM. AK was allowed to equilibrate with nucleotides for 30 min at room temperature before measurements were performed. AMP-PNP was purchased from Sigma and further purified by a linear Ammonium Triethylamine gradient on ion-exchange HPLC (Tosoh TSK-gel DEAE-NPR, 4.6 mm i.d., 3.5 cm) before use in these experiments. Control experiments were performed under the same conditions with a dye labeled polyproline (6) to ensure that addition of nucleotides has no effect on measured FRET efficiencies. Results from steady-state FRET experiments compared to mean distances determined from single-molecule distributions are shown in SI Table 5. Mean single-molecule distances were computed from the shot-noise removed distribution. The uncertainties in the single-molecule measurements were estimated from the bootstrapped distributions (outlined below). The relative shift in the mean distance between substrate-free AK and AK with AMP-PNP/AMP are consistent with the mean values determined from single-molecule probability distributions. Assuming an R_0 = 51 Å for the dye pair, the total change in mean distance upon binding substrates is ≈5 Å. This small change in distance is also consistent with steady-state FRET measurements by Bilderback *et al.* (11), who saw a mean ≈3 Å distance change between the Lid and Core of AK upon binding Ap_5A. These authors also reported different distance changes for different labeling positions. These results once again highlight the complexity of a protein, in which a FRET study interrogates the projection of multiple degrees of freedom onto a single coordinate as defined by the labeling positions chosen for FRET dyes on the protein.

Ensenble Steady-State Fluorescence Anisotropy. To verify the dipole randomization assumption in FRET analysis, both ensemble-averaged steady-state anisotropy and

single-molecule polarization experiments (described below) were carried out. Ensemble-averaged steady-state anisotropy experiments were carried out on single- and double-labeled AK mutants using a fluorometer (SPEX Fluorolog) with two rotatable sheet polarizers and slit widths at 2 mm. Donor dye was excited at 500 nm and emission was scanned from 530 to 750 nm. Acceptor dye was excited directly at 600 nm and emission was scanned from 630 to 750 nm. Steady-state anisotropy was collected in the L-format (12) and calculated by:

$$R = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}}, \qquad \textbf{[2]}$$

where $G = I_{HV} / I_{HH}$ and I_{VH} is the collected intensity with the analyzer polarizer in a vertical position and the emitter polarizer in a horizontal position. Anisotropy values calculated for donor and acceptor alone while freely diffusing in solution as well as donor and acceptor anisotropies measured during energy transfer in doubly labeled proteins are presented in SI Table 5. The donor exhibits a fluorescence lifetime of ≈0.27 ns (13); therefore, the present case even for dyes free in solution is not in the dynamical averaged regime. Increase in donor anisotropy values of dyes upon attachment to the protein is not unexpected since the protein further restricts the volume in which the dye is able to rotate during its fluorescence lifetime (6). The anisotropy measured during energy transfer (acceptor anisotropy of the doubly labeled sample measured when only the donor is excited) demonstrates that the relative orientation between the dyes is randomized on the timescale of their fluorescent lifetimes. Considerations for dipole orientation in single-molecule experiments are different in that the averaging is not through ensemble, but through time. This aspect has been discussed extensively in refs. 6 and 14. This important difference in fact makes single-molecule FRET more favorable in terms of distance determination. Since the single-molecule measurements were made on a much longer time scale (>1 ms) each distance measurement generally contained >25 photons, a number that is sufficient to approach the rotationally randomized limit required for single-molecule FRET experiments (6).

Deleted: al

Single-Molecule Polarization Experiments. To ensure that there are no transient interactions between the dye and protein on a timescale longer than the experimental time resolution of milliseconds, a polarization modulation technique was applied to monitor the randomization of dyes attached to single molecules. To achieve this, a Pockels cell (Conoptics, 350-50) was used in place of the quarter wave plate in our single-molecule confocal microscope (6) to rotate the polarization of the excitation light by 90° with a period of 1 kHz. Since the absorption of photons by a dye is dependent on the angle between the dye and a polarized excitation source, switching the polarization by 90° will result in a fluctuating intensity observed in the fluorescently emitted photons. If the dye were partially immobilized with a lifetime of ms or greater, this would show up as a spike of positive correlation in the intensity autocorrelation function at the period of the oscillating polarization.

Single-molecule polarization experiments have been performed for AK single cysteine mutants at both labeling positions used in single-molecule fret experiments A127C/C77A and A194C/C77A. The donor dye from FRET experiments (Alexa 555, Molecular Probes) was attached to the protein and proteins were immobilized on biotin-PEG functionalized quartz cover slips as described above. A single channel of fluorescence data were collected from each molecule and an intensity autocorrelation function was applied to the data. As a control an Alexa 555 was spin-coated onto a clean quartz slide and treated in the same method. Intensity autocorrelation functions for two sample single-molecule trajectories are displayed in SI Fig. 10. Intensity autocorrelation functions were calculated for the region of the trajectory before and after dye bleaching. The oscillation frequency used for protein samples was 1 kHz, while a frequency of 20 Hz was used for spin coated dye. To aid visualization of the results, dashed red lines have been plotted at one and two times the Pockels cell oscillation frequency. Error bars were calculated by the method of Zwanzig (15) and are plotted to two standard deviations, centered around zero. For random, uncorrelated emission from a single dye, we expect the intensity autocorrelation function to be centered around zero and fall within the error bars. In the intensity autocorrelation function of the spin coated dye we see a prominent feature with positive correlation at the frequency of the Pockels cell (corresponding to 50 ms) and

decreasing features at integral multiples of the oscillation frequency. Absence of this feature in the intensity autocorrelation function after dye bleaching indicates that this effect arises from the dye only and not from changes in intensity of excitation light leaking through the emission filters. These features are noticeably absent from single-molecule experiments, indicating that the dye is freely rotating on timescales faster than 1 ms. This analysis was performed on more than 500 individual molecule and it was found that more than 95% showed no features in their correlation functions at the oscillation frequency of the polarized light. This also was found to be independent of labeling site and dye. Since the highest time resolution used in single-molecule experiments is 2 ms, randomization of the dyes on timescales longer than 1 ms demonstrates that changes in FRET efficiency or dye emissions which occur on our experimental timescale are not due to transient interactions of the dye with the protein or with the quartz coverslip. The apparent noise in the correlation function also underscores the difficulties in extracting protein dynamics, which may span multiple time scales, from single-molecule time trajectories that are too short to give statistically significance insights. We have thus developed new ways to quantify single-molecule dynamics and to relate them to bulk kinetics.

Uncertainties in single-molecule distance measure due to calibration errors. The maximum-information method extracts FRET distances as a function of time from a single-molecule trajectory (14). The distances were calculated using a maximum-likelihood estimator that included emission filter cross talk and background; both parameters have been independently characterized in separate control experiments. A favorable property for the maximum-likelihood estimator is that it asymptotically converges to the true value with an uncertainty that is normally (Gaussian) distributed. Specifically, for the *j*th distance measurement within the *m*th molecule, the normalized distance is evaluated using,

$$\hat{x}_{m,j} = \frac{R_{m,j}}{R_0} = \left[\frac{I^{\beta}_{a,m} / B_{a,m}}{I^{\beta}_{d,m} / B_{d,m}} \frac{I^{\beta}_{d,m} n_{a,m,j} - I^{\beta}_{a,m} n_{d,m,j} I^{\beta}_{d,m} / B_{d,m}}{I^{\beta}_{a,m} n_{d,m,j} - I^{\beta}_{d,m} n_{a,m,j} I^{\beta}_{a,m} / B_{a,m}} \right]^{1/6}, \qquad \textbf{[3]}$$

where R is the donor-acceptor distance, R_0 is the Förster radius. Omitting the m and j indices, I_d^β (I_a^β) is the donor (acceptor) intensity including background contribution for that particular molecule, B_d (B_a) is the donor (acceptor) background level, and n_d (n_a) is the number of donor (acceptor) photons in a time interval T that gives a distance estimate with a predetermined uncertainty δx. In the present application, δx is set to be 0.1, or 10% uncertainty in measuring the normalized distance x. The parameters required to estimate the distance, $I_{d,m}^\beta$, $I_{a,m}^\beta$, $B_{d,m}$, and $B_{a,m}$, are evaluated for each single-molecule trajectory (6). Since these parameters are also experimental quantities, they are subjected to measurement uncertainties. To the first-order approximation, the uncertainties in these parameters result in shifts in the measured absolute distances. For a given molecule m, the shift is constant although out all j distance measurements within the same single-molecule time trace, but the shift varies randomly from molecule to molecule. Therefore, the mean value of a molecular quantity such as intra-molecular distances converges to the true value if it is evaluated from a collection of many single-molecule measurements. The remainder of this section outlines the derivation of this result.

The primary uncertainties in determining $I_{d,m}^\beta$, $I_{a,m}^\beta$, $B_{d,m}$, and $B_{a,m}$ comes from photon statistics. The relative error is inversely proportion to the square root of the number of photons used in the evaluation, $\delta \propto \left(\sqrt{n}\right)^{-1}$. Since both $B_{d,m}$ and $B_{a,m}$ are determined from the time trace after both donor and acceptor photobleach, they can be evaluated relative accurately by continue to take single-molecule trajectory after they bleach. Therefore, the main source of error comes from $I_{d,m}^\beta$ and $I_{a,m}^\beta$, where errors in the latter dominate under most experimental conditions in which the acceptor tends to shorter lived compared to the donor dye. Focusing on the uncertainties due to $I_{d,m}^\beta$ and $I_{a,m}^\beta$, one starts by expanding $\hat{x}_{m,j}$ in terms of the deviations in these intensities, $\delta I_{d,m}^\beta$ and $\delta I_{a,m}^\beta$,

$$x_{m,j} = x_{m,j}^{(0)} + \frac{\partial x_{m,j}}{\partial I_{d,m}^{\beta}}\left(\delta I_{d,m}^{\beta}\right)\left(R_{d,m}\right) + \frac{\partial x_{m,j}}{\partial I_{a,m}^{\beta}}\left(\delta I_{a,m}^{\beta}\right)\left(R_{a,m}\right), \quad \textbf{[4]}$$

where $R_{d,m}$ and $R_{a,m}$ are Normal random variables with zero mean and unit variance, denoting the molecule-to-molecule uncertainties. The partial derivatives can be easily computed to give

$$\frac{1}{x_{m,j}}\frac{\partial x_{m,j}}{\partial I_{d,m}^{\beta}} = -\frac{n_{d,m,j}}{6B_{d,m}n_{a,m,j} - 6I_{a,m}^{\beta}n_{d,m,j}}, \quad \textbf{[5]}$$

and

$$\frac{1}{x_{m,j}}\frac{\partial x_{m,j}}{\partial I_{a,m}^{\beta}} = \frac{n_{a,m,j}}{6B_{a,m}n_{d,m,j} - 6I_{d,m}^{\beta}n_{a,m,j}}. \quad \textbf{[6]}$$

Thus, to the first order, the normalized distance becomes

$$x_{m,j} = x_{m,j}^{(0)}\left[1 - \frac{n_{d,m,j}}{6B_{d,m}n_{a,m,j} - 6I_{a,m}^{\beta}n_{d,m,j}}\left(\delta I_{a,m}^{\beta}\right)\left(R_{a,m}\right) + \frac{n_{a,m,j}}{6B_{a,m}n_{d,m,j} - 6I_{d,m}^{\beta}n_{a,m,j}}\left(\delta I_{d,m}^{\beta}\right)\left(R_{d,m}\right)\right].$$

Since $\left(\delta I_{a,m}^{\beta}\right)\left(R_{a,m}\right)$ and $\left(\delta I_{d,m}^{\beta}\right)\left(R_{d,m}\right)$ randomly vary from molecule to molecule, they eventually cancel out when x is evaluated over a collection of molecules. That is,

$$\left\langle x\right\rangle_{m,j} \sim \sum_{m}\sum_{j} x_{mj} \to x^{(0)}. \quad \textbf{[7]}$$

To gain further insights about the effects of calibration errors, the above expression can be simplified with reasonable approximations. As discussed previously, one generally expects that $\delta I_{a,m}^{\beta} >> \delta I_{d,m}^{\beta}$, therefore, the error is dominated by the first term.

Furthermore, if the background is negligible, $B_{d,m}n_{a,m,j} << I^{\beta}_{a,m}n_{d,m,j}$, the distance becomes

$$x_{m,j} \approx x^{(0)}_{m,j}\left[1+\frac{1}{6}\frac{\delta I^{\beta}_{a,m}}{I^{\beta}_{a,m}}\left(R_{a,m}\right)\right] \approx x^{(0)}_{m,j}\left[1-\frac{1}{\sqrt{n}}\left(R_{a,m}\right)\right], \qquad \textbf{[8]}$$

where n can be understood as (approximately) the number of photons used to evaluate the acceptor intensity. While it is evident that the more photons there are to evaluate the acceptor intensity, the more precise $I^{\beta}_{a,m}$ can be determined; it is nevertheless instructive to estimate the number of photons needed to reduce calibration error to a certain extent, say $0.01x$, 10% of the preset uncertainties ($0.1x$) in photon-counting error. Using the approximation in Eq. **8**, one estimates that at least 277 photons are needed to reduce this calibration error to $0.01x$.

The foregoing discussion also points to experimental strategies to improve accuracy and precision in measurements. For example, the rejection of single-molecule trajectories that are too short is justified in view of the above discussion. To achieve the best signal-to-background ratio, the data acquisition was carried out at the maximum of the Gaussian centroid on a confocal single-molecule image. This has been achieved using a newly installed, automated data acquisition scheme where the location of each molecule is accurately determined using computer-aided image analysis. Finally, examination of the linear relationship between the donor and the acceptor intensities further helps to reject potential complications from overlapping molecules (not perceptible from the confocal image), or from impurities on the coverslip. These were essential steps to ensure that parameter calibration does not contribute significantly to the measured distance distribution.

To ensure that the harvested trajectories obey Poisson statistics, a Cox-Oakes test (16) was applied to the portion of donor trajectory after the acceptor photobleached. Passing the test meant the data can be treated as Poisson with 95% confidence. Details of this

statistical test will be discussed elsewhere. This step was necessary so that the advanced statistical analyses used in the study could be applied. All 763 single-molecule trajectories included in this report passed this test.

Two-State Motional Narrowing Model. Here we present the details of the motional narrowing model used to fit single molecule data in the main text. The probability density for a two-state model under this model is:

$$p(\xi) = p_0\delta(\xi) + p_1\delta(1-\xi) + p_2(\xi), \quad \textbf{[9]}$$

where ξ is the state indicator with 0 being the closed or nearly closed conformation and 1 being the open conformation. $\delta(\xi)$ is the Dirac delta function and is equal to 1 only when $\xi = 0$. Thus, p_0 is the probability density of finding the molecule at closed conformation and p_1 at the open conformation at a given time resolution, τ. Denoting $k = k_{\text{open}} + k_{\text{close}}$ and $p = k_{\text{open}} / k$, the right-hand-side terms in Eq. **9** are expressed as

$$p_0 d\xi = (1-p)e^{-pk\tau}, \quad \textbf{[10]}$$

$$p_1 d\xi = pe^{-(1-p)k\tau}, \quad \textbf{[11]}$$

$$p_2(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dy e^{i\xi y}\gamma_2(y), \quad \textbf{[12]}$$

where $\gamma_2(y) = \gamma(y) - p_0 - p_1 e^{-iy}$ and $\gamma(y) = e^{-(k\tau+iy)/2}\left[\cosh\phi + (\beta/\phi)\sinh\phi\right]$ with $\phi = \sqrt{(k\tau/2)^2 - i(p-1/2)k\tau y - (y/2)^2}$, $\beta = k\tau/2 - i(p-1/2)y$, and $i \equiv \sqrt{-1}$. Note that the expressions for p_0 and p_1 in Eqs. **10** and **11**, respectively, have the correct unit for probability density whereas in the original Geva-Skinner paper they were expressed in terms of probability.

To fit the model to single-molecule data, τ in this model is taken as the average time resolution, $T_{\text{avg}} = \sum_j T_j / n$, at a fixed relative uncertainty $\alpha \equiv \left(\delta x / x\right)$ from the maximum-information analysis of n maximum-likelihood estimated $x \equiv R / R_0$. The time resolution (τ) and the measurement uncertainty (α) are related by an expression based on the information theory (14). In this analysis, each *j*-th maximum-likelihood estimate of x, denoted $\hat{x}_j$, is carried out at constant Fisher information, $J(\hat{x}_j)$ (or relative uncertainties):

$$J(\hat{x}_j) = \frac{36\hat{x}_j^{10}}{\left(1+\hat{x}_j^6\right)^3}\left[I_d^{\beta} T_j \frac{\left(1-\beta_d^{-1}\right)^2}{\left(\hat{x}_j^6+\beta_d^{-1}\right)} + I_a^{\beta} T_j \frac{\left(1-\beta_a^{-1}\right)^2}{\left(\hat{x}_j^6+\beta_a^{-1}\right)} \right] = \frac{1}{\text{var}(\hat{x}_j)} = \frac{1}{\alpha^2}, \qquad \textbf{[13]}$$

where I_d^{β} (I_a^{β}) is the detected emission intensity of the donor (acceptor) in the absence of the acceptor (donor), β_d (β_a) is the signal-to-background ratio of the donor (acceptor). T_j can thus be understood as the time resolution associated with the *j*-th measurement with a relative uncertainty α. For this particular experiment, the relative FRET distance uncertainty α and the corresponding average time resolution are tabulated in SI Table 6. Eq. **13** gives a quantitative assessment of the extent to which one can make a statement about a FRET measurement give the number of photons included in the analysis. On one hand, it cautions one from over interpret the experimental results. On the other hand, it allows one to extract the FRET distance with minimal number of photons. In the present case, one needs about ≈20 photons to extract a FRET distance with an uncertainty of ≈5 Å.

When fitting the experimental single-molecule data, the integration in Eq. **12** was carried out using vectorized Gaussian quadrature implemented in Matlab 7 (Mathworks, Natick, MA). The computed distribution is then rescaled using $x = a\xi + b$ to match with the experimentally measured x, where b can be understood as the x-position at the closed conformation and $(a+b)$ as the x-position at open conformation.

Deconvolution of Single-Molecule Probability Density Functions. To gain information about the number of conformational states accessible to AK, a probability density function (PDF) was created from the distances measured in single-molecule FRET experiments. As previously discussed, the Maximum Information Method for analyzing single-molecule FRET data yields a maximum likelihood estimator $\hat{x}_{m,j}$ for the *j*th distance measured on the *m*th molecule (Eq. **3**) such that each distance has a constant variance α^2 (Eq. **13**) (14). Each measured distance, $\hat{x}_{m,j}$, is normally distributed around it's true value, which allows the use of a Gaussian kernel estimator for the construction of a time-weighted, raw PDF, $\hat{r}(x)$,

$$\hat{r}(x) = \frac{1}{T}\sum_m \sum_j \Delta t_{m,j} k(\hat{x}_{m,j}; \alpha^2) \quad \textbf{[14]}$$

Deleted: (
Formatted: Font: Bold
Deleted:)

where $T = \sum T_m$ is the total lenth of all trajectories used, Δt_j is the time duration of the *j*th measurement and $k(\hat{x}_j; \alpha^2)$ is a Gaussian kernel centered at $\hat{x}_j$ with a variance α^2 (6).

As illustrated by the dashed line in Fig. 2*c*, the raw PDF is significantly broadened and nearly featureless due to the presence of noise. A significant contribution to the broadening of $\hat{r}(x)$ is the uncertainty in the distance measurements, $\hat{x}_{m,j}$, which are normally distributed around their true values with a variance α^2 due to the use of the maximum likelihood estimator. The Gaussian kernel used to create the raw PDF also contributes to broadening, again with a variance of α^2. Mathematically, the true PDF, $h(x)$, can be related to the raw PDF by convolution with two sources of Gaussian noise

$$\hat{r}(x) = \left[h(x) \otimes k(x; \alpha^2) \right] \otimes k(x; \alpha^2) = h(x) \otimes k(x; 2\alpha^2) \quad \textbf{[15]}$$

Deleted: (
Formatted: Font: Bold
Deleted:)

where $\otimes$ denotes the convolution operation (6).

To deconvolve the true PDF, we have previously reported the use of a method based on the maximum-entropy method which minimizes the test metric Q

$$Q(h(x),\lambda) = \chi^2 + \lambda H \quad \textbf{[16]}$$

where H is the negative entropy of the proposed true PDF

$$H = \int_{-\infty}^{\infty} h(x) \ln h(x) dx \quad \textbf{[17]}$$

λ is a Lagrange multiplier and χ^2 is a measure of the goodness of fit between the raw PDF and the proposed true PDF convoluted with Gaussian error.

$$\chi^2 = \int_{-\infty}^{\infty} \frac{\left[\hat{r}(x) - h(x) \otimes k(x; 2\alpha^2)\right]^2}{\sigma_r^2(x)} dx \quad \textbf{[18]}$$

The variance in the raw PDF, $\sigma_r^2(x)$, is calculated with the method of Efron's bootstrap (17). In this case, the entire collection of maximum likelihood distance estimators, $\hat{x}_{m,j}$, used to create the raw PDF was re-sampled with replacement to make a bootstrapped set of distance measurements which is the same size as the original dataset and can be used to calculate a corresponding bootstrapped $\hat{r}(x)$. This procedure was repeated 25 times to obtain a bootstrapped set of $\hat{r}(x)$ from which the x-dependent variance in the raw PDF was calculated.

An algorithm which minimizes the metric Q in Eq. **16** was implemented following the methods of Skilling and Bryan (18). Briefly, this method iteratively converges toward a minimum for both χ^2 and H in a three dimensional subspace defined by the vectors

$$e_1 = h(x)\left(\nabla H\right)$$
$$e_2 = h(x)(\nabla\chi^2) \qquad \textbf{[19]}$$
$$e_3 = \frac{h(x)\left(\nabla\nabla\chi^2\right)h(x)\left(\nabla H\right)}{\left|\nabla H\right|} - \frac{h(x)\left(\nabla\nabla\chi^2\right)h(x)\left(\nabla\chi^2\right)}{\left|\nabla\chi^2\right|}$$

where $\nabla H \equiv \frac{\delta H\left(h(x)\right)}{\delta h(x)}$, gradient of the negative entropy with respect to the probability density function. After orthonorrmalizing this basis set, $\widetilde{\chi}^2$ and $\widetilde{H}$ can be defined in the new subspace by a Taylor expansion and simplified to the form

$$\widetilde{\chi}^2 = \chi^2 + \chi^2_\mu x_\mu - \tfrac{1}{2} x_\mu x_\mu, \quad \widetilde{H} = H + H_\mu x_\mu - \tfrac{1}{2} M x_\mu x_\mu \qquad \textbf{[20]}$$

where $\chi^2_\mu = e^T_\mu \nabla\chi^2$, $H_\mu = e^T_\mu \nabla H$ and $M = e^T_\mu \nabla\nabla\chi^2 e_\mu$. The metric $\widetilde{Q}$ in the new subspace is then

$$\widetilde{Q} = \widetilde{\chi}^2 + \lambda\widetilde{H} \qquad \textbf{[21]}$$

which by minimizing $\widetilde{Q}$ gives

$$x_\mu = \frac{\lambda H_\mu - \chi^2_\mu}{M + \lambda} \qquad \textbf{[22]}$$

where x_μ are the coefficients for the search direction that are used to define the new trial true PDF, $h_{new}(x)$, to be used in the subsequent iteration

$$h_{new}(x) = h(x) + x_\mu e_\mu \qquad \textbf{[23]}$$

Deleted: χ^2 and H

Deleted: $\chi^2 = \chi^2 + \chi^2_\mu\ x_\mu - (1/2)x_\mu x_\mu$, $H = H + H_\mu x_\mu - (1/2)M x_\mu x_\mu$

Deleted: Q

Deleted: $Q = \chi^2 + \lambda H$

Deleted: Q

The initial guess used for $h(x)$ in the algorithm is $h(x) = \hat{r}(x)$. The solution is considered to have converged either when $\chi^2 < 10^{-5}$ or $\Delta\chi^2 / \chi^2 < 10^{-4}$ where $\Delta\chi^2$ is the change in χ^2 between iterations. The algorithm was implemented in a C program. In this application of the Skilling and Bryan algorithm, the condition $\lambda = 0$ was found be sufficient for convergence.

Bootstrap Estimation of Uncertainties in the Distribution and Fitting Parameters. Efron's bootstrap method (17) was used to estimate the confidence intervals of the deconvolved conformational distribution and the fitting parameters (19). The collection of 300+ single-molecule trajectories was re-sampled with replacement to make up a bootstrapped collection of trajectories that was of the same size as the original set (20), to form an instance of bootstrapped experiment. Trajectories in a bootstrapped experiment were subjected to exactly the same data reduction procedures including maximum-information analysis and entropy-regularized deconvolution to compute a bootstrapped conformational distribution. The above procedure was repeated 300 times to obtain a set of bootstrapped conformation distribution, Ω. The *x*-dependent confidence intervals were computed from the collection of distributions in Ω. For example, the ~2-ms distributions shown in Fig. 2f in the main text have the error distribution as shown below (SI Fig. 11). The dashed lines are one-standard deviation error bounds for these distributions, indicating that the number of single-molecule trajectories used for building these distributions was satisfactory, and that the observed two-state distribution was statistically significant.

Similarly, to estimate the uncertainties in the parameters for the two-state model outlined in the previous section, each set of trajectory in the Ω was re-analyzed systematically at different α values and fitted to the motional narrowing model. This generates a set of bootstrapped parameters, from which the standard deviations were computed and reported as the uncertainties to one standard deviation in the main text and in SI Table 1).

Simulation Methods. To develop a structural model of AK with the attachment of dye molecules, short molecular dynamics simulations of both the open and closed forms of AK were performed with the presence of dye molecules. Following the FRET experiments, mutations A127C, A194C, and C77A were made; the Alexa Fluor dyes were then attached to the mutated C127 and C197 residues. The structure for the acceptor dye, AlexaFluor 647, was obtained from the US Patent database (21). Since the chemical structure of the donor dye is not available, a structure similar to that of the acceptor dye was assumed with two fewer -CH- groups in the alkene chain, drawing analogies from Cy3-Cy5 dyes. The guessed Alexa Fluor 555 dye was attached via the succinimide group to C194 and the Alexa Fluor 647 dye was attached to C127. Force fields of the dye molecules were constructed using standard CHARMM (1, 2) geometries and point charges of the constituent residues. Ligand molecules were not included in either simulation.

The structure of Muller *et al.* (22) (PDB ID code 4AKE) was used as the initial configuration for the simulation starting from the open state, and the structure of Muller and Schulz (2) (PDB ID code 1AKE) was used as the initial configuration for the simulation starting from the closed state. Coordinates that are missed in the x-ray structures were generated using the internal coordinate facilities of charmm (23). The initial positions of the dye molecules were adjusted to prevent overlapping with AK, and then 800 adaptive-basis Newton-Raphson (ABNR) minimization steps were performed with the protein fixed. Potassium chloride was used to neutralize the system (28 K^+ and 18 Cl^-), and the positions of ions were obtained using the Solvate 1.0 program (24). The proteins and ions were then placed in a preequilibrated water box with the TIP3P explicit water model (25); water molecules within 2.4 Å of protein atoms or ions were removed. The final systems for both the open and closed forms contain a total of 32,144 atoms, with 3544 atoms for the modified AK, 46 ions, and 28,554 water atoms.

The NAMD package (26) was used to perform the molecular dynamics simulations of both systems with periodic boundary conditions. The charmm22 (27) all-atom force field with the CMAP cross-terms (1, 9) was used to compute the potential energy and forces of

the system. Long-range electrostatic interactions were computed using the particle mesh Ewald algorithm (28). Covalent bonds in the water molecules were constrained to their equilibrium values by the SHAKE algorithm (29), thus allowing an integration time step of 2 fs. The structures of both systems were minimized with 60,000 conjugate gradient steps, with protein backbone atoms restrained to their original positions using harmonic potentials with a force constant of 1.0 kcal/mol/Å. The system was then heated to 300 K by reassigning velocities every 0.2 ps at a rate of 30 K/ps. After reaching 300 K, the systems were allowed to equilibrate at 300 K and 1 atm for 1 ns in the presence of restraint potentials. The temperature was controlled with a Langevin thermostat with a 0.5 ps^{-1} damping coefficient, and the pressure was controlled by a Langevin-piston barostat (30) with a piston period of 2 ps and a damping time constant of 2 ps. The restraint potentials were then removed for both systems to further equilibrate. In all of the aforementioned steps, no restraint potential was applied to the dye molecules to assume any favorable orientation with the protein molecule. A total of 25 ns of trajectory was collected for the simulation starting from the open state and 5 ns for the simulation starting from the closed state. Both trajectories indicate that AK is flexible in an aqueous environment. However, the structures sampled in both simulations remain close to their initial structures; neither open-to-close nor close-to-open transition has been observed. More sampling is required to better compare the results of MD simulations with FRET measurements; these results and more detailed analysis will be reported later.

1. Pal P, Lesoine JF, Lieb MA, Novotny L, Knauf PA (2005) *Biophys J* 89:L11–L13.

2. Berry MB, Bae EY, Bilderback TR, Glaser M, Phillips GN (2006) *Proteins* 62:555-556.

3. Huss RJ, Glaser M (1983) *J Biol Chem* 258:13370-13376.

4. Segel IH (1993) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems* (Wiley, New York), pp 274-276.

5. Sheng XR, Li X, Pan XM (1999) *J Biol Chem* 274:22238-22242.

6. Watkins LP, Chang H, Yang H (2006) *J Phys Chem A* 110:5191-5203.

7. van Oijen AM, Blainey PC, Crampton DJ, Richardson CC, Ellenberger T, Xie XS (2003) *Science* 301:1235-1238.

8. Edelhoch H (1967) *Biochemistry* 6:1948-1954.

9. Bradford MM (1976) *Ann Biochem* 72:248-254.

10. Clegg RM (1992) *Methods Enzymol* 211:353-388.

11. Bilderback T, Fulmer T, Mantulin WW, Glaser M (1996) *Biochemistry* 35:6100-6106.

12. Lakowicz JR (1999) *Principles of Fluorescence Spectroscopy* (Springer, New York), pp 298-301.

13. Luong AK, Gradinaru CC, Chandler DW, Hayden CC (2005) *J Phys Chem B* 109:15691-15698.

14. Watkins LP, Yang H (2004) *Biophys J* 86:4015-4029.

15. Zwanzig R, Ailawadi NK (1969) *Phys Rev* 182:280-283.

16. Henze N, Meintanis SG (2005) *Metrika* 61:29-45.

17. Efron B (1979) *Ann Stat* 7:1-26.

18. Skilling J, Bryan RK (1984) *Monthly Notice R Astronom Soc* 211:111-124.

Deleted: ???

19. DiCiccio TJ, Efron B (1996) *Stat Sci* 11:189-212.

20. Scott DW (1992) *Multivariate Density Estimation* (Wiley, New York), pp 257-260.

21. Terpetschnig EA, Patsenker LD, Tatarets A (2004) US Patent 20040166515.

22. Muller CW, Schlauderer GJ, Reinstein J, Schulz GE (1996) *Structure (London)* 4:147-156.

23. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) *J Comput Chem* 4:187-217.

24. Grubmüller H (1996) (Max Planck Institute for Biophysical Chemistry, Göttingen).

25. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) *J Chem Phys* 79:926-935.

26. Nelson MT, Humphrey W, Gursoy A, Dalke A, Kale LV, Skeel RD, Schulten K (1996) *Int J Supercomput Ap* 10:251-268.

27. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, *et al.* (1998) *J Phys Chem B* 102:3586-3616.

28. Darden T, York D, Pedersen L (1993) *J Chem Phys* 98:10089-10092.

29. Allen MP, Tildesley DJ (1987) *Computer Simulations of Liquids* (Oxford Univ. Press, New York)

30. Feller SE, Zhang YH, Pastor RW, Brooks BR (1995) *J Chem Phys* 103:4613-4621.