Resolution of ångström-scale protein conformational changes by analyzing fluorescence anisotropy

John H. Lewis and Zhe Lu

Conformational changes within typical protein molecules are rapid and small, making their quantitative resolution challenging. These changes generally involve rotational motions and may thus be resolved by determining changes in the orientation of a fluorescent label that assumes a unique orientation in each conformation. Here, by analyzing fluorescence intensities collected using a polarization microscope at a rate of 50 frames per second, we follow the changes of 10°–16° in the orientation of a single bifunctional rhodamine molecule attached to a regulator of conductance to K⁺ (RCK) domain of the MthK channel, and thus, the transitions between its three conformational states, with effective standard deviation (σ) of 2°–5°. Based on available crystal structures, the position of the fluorophore’s center differs by 3.4–8.1 Å among the states. Thus, the present approach allows the resolution of protein conformational changes involving ångström-scale displacements.

Understanding the energetic and kinetic mechanisms of conformational changes of a protein and relating them to known atomic structures of its conformational states require the capability to resolve and monitor the changes of its spatial characteristics, which are unique to individual states, with adequate spatial and temporal resolutions. Conformational changes of protein molecules may occur temporally and spatially on millisecond and ångström scales. On these temporal and spatial scales, resolution of multistate protein conformational changes at a single-protein-molecule level by measuring distance changes remains extremely challenging. However, protein conformational changes generally involve rotational motions of their domains, in which secondary structure elements such as α-helices might adopt unique spatial orientations among individual conformational states. In such cases, resolving the conformational states of a protein could be reduced to resolving the spatial orientation of a single suitable structural element with unique orientation in each state.

In a spherical coordinate system, the orientation of an object, represented here by a vector, can be defined by three inherently nonequivalent parameters, namely, the length (R) and two angles: inclination (θ, from 0° to 180°) and rotation (φ, from 0° to 360°) (Fig. 1a; all notations are defined in Supplementary Note 1). The two angles contain the full information regarding spatial orientation, independent of R. Thus, in order to track orientation changes and relate those tracked conformations to known protein structures, monitoring θ and φ should suffice. The estimated median radius of proteins is approximately 20 Å1–2. A rotation of 5°–10° of a site 20 Å away from the origin would lead to a 1.7°–3.5° Å change in the chord distance. Thus, the capability to resolve angle changes on this scale would in principle allow the study of atomic-scale conformational changes in typical proteins.

One approach to track changes in orientation of a protein region is to monitor the orientation of an attached fluorophore (Fig. 1a) using a polarization microscope. The resolution of the changes in a fluorophore’s emission-polarization (independently of wavelength) is limited in practice by the signal-to-noise ratio (SNR) and cannot exceed what is predicted from the so-called shot noise in photon counting. Fluorescence polarization analysis has long been applied to the study of two-dimensional single-molecule motion in terms of 〈φ²〉. This technique was subsequently extended to tracking three-dimensional motion of the lever arm domain of myosin V labeled with a bifunctional rhodamine in terms of both θ and φ, using a time-multiplexing excitation protocol to show that this 24-nm-long domain undergoes a swinging motion as large as ~80°, which corresponds to its motor domain traversing 36 nm along an actin filament3. In this study and in subsequent studies4, four or eight linearly polarized excitations were used, with each of the resulting emissions split into two polarized emission components.十六 polarization signals allow unambiguous determination of orientation within a full hemisphere, which is required to measure large angle changes but may not be necessary for small angle changes. The σ value (commonly used as a measure of precision) of the distributions of θ and φ were reported as 10° and 14°, respectively, when the measurements comprised a few hundred photons captured from eight polarization channels during a sampling interval on the order of tens of milliseconds5–7. In comparison, the theoretical shot-noise-limited σ values of θ and φ were 6° and 10°, respectively1. In contrast, a σ of 2°–12° has been achieved in imaging analysis studies of individual fluorophore molecules (not attached to protein), which required analyzing several thousand photons under conditions that minimized motions of fluorophores12,13.

Here we aimed to achieve an effective σ of 5° or less to resolve multistate protein conformational changes on ångström and millisecond scales. As a test protein, we chose the cytosolic regulatory module of the prokaryotic Ca²⁺-dependent K⁺ channel MthK14,15. The regulatory module controls the transmembrane channel pore and is structurally four-fold symmetric around the central axis of the pore, comprising two stacked rings, each formed by four RCK domains (Fig. 1b). An RCK domain can bind Ca²⁺ and has a molecular weight of 25 kDa, compared with the average of ~50 kDa for proteins. Crystallographic studies have captured three structural conformations of the RCK domain, dubbed $S_1$, $S_2$, and $S_3$ (Fig. 1c). The $S_2$ conformation was observed in a full MthK channel (PDB 1LNQ)14, whereas $S_0$, either in a pure form (PDB 4RO0) or in a...
mixture with S₁ (PDB 2FY8) 13, was observed with the isolated regulatory module. When the three conformations are aligned for a single RCK domain using the CNS program 14, the locations of individual backbone atoms differ by 0.2–9.2 Å. Thus, a single RCK domain within the regulatory module is an ideal system to develop a strategy for simultaneously achieving the high spatial and temporal resolutions required to examine conformational changes in a typical protein.

Results
Labeling, attaching and orienting protein molecules. We chose to track helix αB in the RCK domain because it is the closest helix to the channel gate formed by the second membrane-spanning segment (M2) in the pore module 14,15 (Fig. 1b). From the available crystal structures (PDB 1LNQ and PDB 2FY8), this α-helix is expected to rotate 10–16° among the three conformations; its middle point, ~20 Å from a reference origin, is expected to move 3.4–8.1 Å (Fig. 1c). Residues E146 and L153 in the helix were replaced by cysteine residues to attach a bifunctional rhodamine molecule (Fig. 1a,b and Methods). The two-point attachment ensured alignment of the fluorophore with the helix such that its dipole orientation would report the helix orientation. To minimize the probability of having more than one cysteine-mutation-containing RCK domain in each regulatory module formed by eight RCK domains, we co-transformed bacteria with wild-type and mutant cDNAs in a 15:1 ratio. Additionally, labeling of the purified RCK octamers was performed at a ratio of 1:8.1 bifunctional rhodamine to RCK, to further reduce the chance of having individual octamers labeled with more than one rhodamine molecule. We observed ~10% of fluorescent particles were bleached in more than one step and excluded them from further analysis.

To attach the sample protein to a glass coverslip conjugated with streptavidin, we covalently linked biotin to its N termini. The streptavidin-biotin interaction was exploited to create a stable attachment between the sample protein and the coverslip. An RCK octamer presents two indistinguishable faces normal to the four-fold axis, with four N termini exposed on either face (Fig. 1b). A four-fold attachment between a ring of four RCK domains to the coverslip aligned the regulatory module's central axis with the z axis, allowing us to group θ of each conformational state determined from numerous molecules into a single distribution. From the resulting individual distributions of θ, we calculated the means and used them as a set of benchmarks to relate to the structures of the crystallographic states 14,15 (Fig. 1b,c). The angle φ of the α-helix, relative to the x axis, lies in the plane of the microscope slide (Fig. 1a,b). Because any given regulatory module should be randomly oriented relative to the x axis, its φ value would also vary randomly across different complexes, but the difference in φ between two states should remain the same statistically. Given that the angle change between two conformations in the actual rotation plane (Ω; Fig. 1c) reflects the change in φ and θ, we used Ω values among the three conformations as an additional set of benchmarks.

Specifications and optimization of the microscope. We used a polarization microscope equipped with a single laser, circularly polarized, to excite fluorophores under a total internal reflection condition, and two polarizing beam splitters to separate fluorescence emission 17, which were recorded with an electron-multiplying charge-coupled device (EMCCD) camera (Methods and Supplementary Note 2). This approach was useful in previous studies of proteins labeled with bright fluorescent nanoparticles 16,18, yielding reasonable precision of angle measurement (σ of 12° for φ 18). Here, an effective σ of ~5° or less was required to resolve the predicted 10–16° orientation changes in RCK labeled with rhodamine, which emits far fewer photons per second under the present experimental conditions. Polarizing beam splitters are critical determinants of the microscope’s quality of polarization. Below, we use an idealized system to explain the imperfection of the splitters and necessary corrections for this and other imperfections (details in Supplementary Notes 3 and 4).

A fluorophore can be modeled as a dipole that anisotropically emits photons; the orientation of its net electric field vector E can be defined by θ, φ and |E|, or equivalently by the three Cartesian components Eₓ, Eᵧ, and Ez:  

\[
E_x = |E| \sin \theta \cos \phi \\
E_y = |E| \sin \theta \sin \phi \\
E_z = |E| \cos \theta
\]  

(1)
where \(|\mathbf{E}|\) is the vector magnitude, related to the total emitted fluorescent light by \(I_{\text{tot}} = |\mathbf{E}|^2\). If this light were entirely directed to pass through a 0–90° polarizing beam splitter aligned with the \(x\) and \(y\) axes, the intensities \(I_x\) and \(I_y\), which correspond to the \(E_x\) and \(E_y\) components, would be separated into the 0° and 90° polarization channels, respectively. \(I_z\) would be effectively indistinguishable by these polarization channels and therefore be split evenly between them. The intensities recorded in the 0° \((I_0)\) channel would then be linear combinations of \(I_x\) and \(\frac{1}{2}I_y\), and those of the 90° \((I_90)\) channel would be combinations of \(I_x\) and \(\frac{1}{2}I_y\) (Supplementary Note 3). Thus, additional independent information is required to determine the dipole’s orientation or its three electric-field components. If the intensity emitted from a fluorophore fixed in space were collected by one of \(h\) number of channels with a polarization angle of \(\psi\) (relative to the \(x\) axis), the collected intensity \((I_\psi)\) would be given by:

\[
I_\psi(I_{\text{tot}}, \theta, \phi) = \frac{1}{h} I_{\text{tot}} \left[ \frac{1}{2} \sin^2 \theta \cos 2(\phi - \psi) + 1 \right]
\]

(2)

where according to the dipole model, \(I_{\text{tot}} = 2|\mathbf{E}|^2\) (Supplementary Note 3). The sum of \((I_\psi)\) over all channels would yield \(I_{\text{tot}}\). As we could only split the beam into an even number of components, the three-dimensional information of \(\mathbf{E}\) would be transformed to a minimum of four components, here \(I_4\), \(I_{45}\), \(I_{90}\) and \(I_{135}\), experimentally measurable in a two-dimensional space. In theory, determination of \(\theta\), \(\phi\), and \(I_{\text{tot}}\) requires measuring \(I_\psi\) from a minimum of three polarization channels\(^\text{17}\). However, in practice, the noise and depolarization of the four intensity recordings are expected to be different, and the four possible sets of angle calculations, each using three intensities, would lead to different numerical values. For achieving a unique set of calculated values, we derived, from equation (2), the analytic expression for \(\theta\) or \(\phi\), each of which is a function of all four intensities:

\[
\phi = \tan^{-1} \frac{I_{45} - I_{135}}{I_0 - I_{90}}
\]

(3)

\[
\theta = \sin^{-1} \left( \frac{2}{\sqrt{(I_0 - I_{90})^2 + (I_{45} - I_{135})^2}} \right) \left( \frac{(I_0 + I_{90} + I_{45} + I_{135})}{(I_0 + I_{90} + I_{45} + I_{135})} \right)
\]

(4)

where either angle would be a function of an intensity ratio, and its resolution should thus be limited by the SNR of intensities; \(\theta\) would be defined from 0–90° and \(\phi\) from 0–180°, with degenerate solutions of 180° − \(\theta\) and 180° + \(\phi\), the degeneracy reflecting the underlying symmetry of the fluorophore dipole.

We initially used two Wollaston beam splitters, as described previously\(^\text{11}\) (Fig. 2a). Judging by the polarization ratio (dubbed \(f\)), the light passing through the 0–90° splitter almost completely retained its original polarization \((f_0 = 0.99\) and \(f_90 = 0.99)\), but only ~50% of the light remained polarized after passing through the 45–135° splitter. This depolarization reduced the upper limit for the calculated \(\theta\) to ~45° from the theoretical limit of 90°. To different degrees, the depolarizing effect was shared by all splitters that we tested for the 45–135° configuration. Among them, a wire-grid polarizer yielded \(f_{45} = 0.85\) and \(f_{135} = 0.79\), which had been used previously\(^\text{12}\). With a Wollaston splitter (for 0–90°) and a wire-grid polarizer (for 45–135°), plus numerical correction of the depolarization, we restored \(\theta\) to its full range. Additionally, the beam splitters would separate light into the polarized components according to the actual polarization angle \(\psi\), which may deviate from the expected value. Thus, only the calibrated \(\psi\) value should be used in the calculations (Methods and Supplementary Note 4). To account for the diffusive motion of the probe and incomplete collection of photons by our microscope objective (with a 78.5° half-angle (\(\alpha\)) of the collection cone\(^\text{13}\), as opposed to an \(\alpha\) of 180° for full collection), we implemented additional, well-established corrections (Supplementary Note 3). Equations (53), (61) and (63) are the counterparts of...
We plotted the intensities of the four components of a fluorophore against time (Fig. 3b), where each intensity value was a direct summation of individual pixels (Supplementary Note 2). Intensity $I_{tot}$ defined as the total emitted intensity, was calculated from the four recorded intensities with equation (62) (Supplementary Note 3). Furthermore, from the intensities of the four polarized components, we calculated $\theta$ and $\varphi$ with equations (63) and (61) (Supplementary Note 3) and plotted them in Fig. 3c, along with $\Omega$, calculated using equation (70) (Supplementary Note 6).

Event transition detection, state identification and angle resolution. We detected conformational transitions and identified states in two separate steps. We used the changepoint algorithm$^{21,22}$ to detect event transitions on the basis of concurrent intensity changes, extracting temporal information. This method does not assume any specific kinetic model regarding the protein conformational changes; it simply detects event transitions by a likelihood ratio, that is, testing two hypotheses (with 95% confidence) that, within a given time interval, an intensity change does or does not occur. A Gaussian filter was applied to all four intensity traces to reduce high-frequency noise in a manner that the shortest events, each comprising a single data point, retained more than half of their intensity values. Through calibration, we estimated that on average, 55 photons (out of ~120 emitted photons) were recorded from all four channels during each 20-ms exposure, where the signal-to-background ratio was ~10. Mean $\theta$ and $\varphi$ were 56° and 90°, respectively, whose mean $\sigma_\theta$ and $\sigma_\varphi$ were 8.6° and 9.1°, respectively (the fluorophore’s wobble angle $\delta$ was set at 30°; Supplementary Note 5). To estimate the theoretical limits for $\sigma_\theta$ and $\sigma_\varphi$, we analyzed the variations of $\theta$ and $\varphi$, which were calculated from the intensities simulated for the above conditions, assuming those variations would arise solely from shot noise. The resulting theoretical limits of $\sigma_\theta$ and $\sigma_\varphi$ yielded from this analysis were 8.0° and 6.4°, respectively. Thus, the precision of our angle measurements was close to being shot-noise limited.

Identification of transition points enabled us to average the data points within individual events, effectively increasing precision. Practically, from the original intensity $I_\varphi$ and angles ($\theta$, $\varphi$, $\Omega$) traces (Fig. 3b,c), we generated the corresponding black traces, of which the values of individual points within an event of a given trace were averaged over its lifetime. After averaging (13 data points per event on average), the mean values of $\theta$ and $\varphi$ were reduced to 4.2° and 4.4°, respectively, on the basis of individual events, during which an average estimated 715 photons were collected from four channels. As further illustrated in Fig. 4a, in terms of SNR over a range of 3–13, the values of effective $\sigma_\theta$ and $\sigma_\varphi$ ranged 5–2°. Given that we tried to resolve angle changes as small as ~10°, we set effective $\sigma_\theta$ values of $\leq$5° as a criterion for particle selection.

After intra-event data-point averaging, the distribution of $\theta$ showed three distinct peaks, and the distribution of $\varphi$ showed two peaks, with the broader peak likely reflecting two unresolved peaks (Fig. 3c). These characteristics were expected from the crystal structures, in which the $\theta$ value of the tracked $\alpha$-helix differs between all three conformations by $\geq$10°, and its $\varphi$ values for two conformations are within (unresolvable) 3°, each differing by ~10° from the third. For further illustration, we plotted $\theta$ against $\varphi$ (Fig. 4b), color-coding individual data points by conformational state, identified as described below. As expected from the plot, the distribution of $\Omega$, containing both $\theta$ and $\varphi$ information, showed three distinct peaks (Fig. 3c).

We developed a program based on the k-means clustering algorithm$^{23}$ to identify conformational states from $\theta$ and $\varphi$ together, to increase resolvability and confidence. After examination of all possible cases, from one to four states (and more if necessary), the program determined the most probable number of states by comparing all examined cases using an $F$ test$^{34}$. Of 980 particles analyzed, 85% showed three states, whereas the remaining particles showed a

---

**Fig. 3 | Polarized-fluorescence intensities, and calculated $\theta$, $\varphi$ and $\Omega$.**

(a) Consecutive frames of four intensity components ($I_{\psi_0}$, $I_{\psi_1}$, $I_{\psi_2}$ and $I_{\psi_3}$) of a fluorescent particle (boxed in blue, Fig. 2b) captured over the last 0.5 s of an ~4-s recording (Supplementary Video 1). (b) Intensity values of the four components, plotted against time, were integrated from 218 sets of consecutive images, part of which is shown in a, $I_{\psi_0}$ was calculated from $I_{\psi_1}$, $I_{\psi_2}$, and $I_{\psi_3}$ using equation (62) (Supplementary Note 3). (c) Values of $\theta$ and $\varphi$ were calculated from the four intensity components, whereas the value of the angle change $\Omega$, defined relative to conformational state $S_2$, was calculated using equation (70) (Supplementary Note 6). Transition points of the black traces superimposed on the records were determined by the changepoint analysis, whereas the constant values between points were set to the means of the observed data points. Angle histograms generated from the black traces, superimposed with Gaussian fits, are shown on the right.

---

Recorded intensities and calculated angles. Using the EMCCD camera, we simultaneously captured components $I_{\psi_0}$, $I_{\psi_1}$, $I_{\psi_2}$ and $I_{\psi_3}$ of fluorescence emission from numerous spatially separated molecules (Fig. 2b). Figure 3a shows consecutive sets of images of four polarized components from a single attached bifunctional rhodamine (Fig. 2b, boxed spot), recorded at a rate of 50 frames per second. For any given time point, the collected intensities completely defined $\theta$ and $\varphi$. From the simultaneous and relative variations in the four intensities, which reflect angle changes of the tracked helix, the underlying ångström-scale intramolecular motions effectively became visually noticeable in real time (Fig. 3a; Supplementary Video 1).

equations (2–4), with all above corrections (Supplementary Note 3). Using equations (61) and (63), we calculated the angles to avoid any errors that might arise from failure to reach the global minimum during simultaneous fitting of multiple equations to a set of intensities, also markedly shortening the time to analyze the large number of data sets.
SNR, obtained by analyzing a pool of data collected at a range of Ca\(^{2+}\) concentrations\(^{25,26}\). Each point encompasses the data that fall in the SNR range defined by the abscissa values of two consecutive points and is plotted at the position of the smaller value. Mean SNR for each data point was calculated on a per-intensity-channel basis and with the assumption of Gaussian noise, because noise aside from shot noise was likely also present. Number of particles for each SNR value: 3 (n = 259), 4 (n = 217), 5 (n = 207), 6 (n = 128), 7 (n = 76), 8 (n = 39), 9 (n = 32), 10 (n = 14), 11.5 (n = 10) or 13 (n = 4).

Fig. 4 | Angle resolution and distributions obtained through single-particle analysis. a, \(\sigma_\theta\) and \(\sigma_\phi\) (mean ± s.e.m.) plotted against observed SNR, obtained by analyzing a pool of data collected at a range of Ca\(^{2+}\) concentrations\(^{25,26}\). Each point encompasses the data that fall in the SNR range defined by the abscissa values of two consecutive points and is plotted at the position of the smaller value. Mean SNR for each data point was calculated on a per-intensity-channel basis and with the assumption of Gaussian noise, because noise aside from shot noise was likely also present. Number of particles for each SNR value: 3 (n = 259), 4 (n = 217), 5 (n = 207), 6 (n = 128), 7 (n = 76), 8 (n = 39), 9 (n = 32), 10 (n = 14), 11.5 (n = 10) or 13 (n = 4).

b, Distributions of \(\Omega\) (red) and \(\varphi\) (green) of the particle from Fig. 3c, superimposed on the distributions were calculated with the mean and \(\sigma\) values of individual events plotted against their corresponding \(\varphi\) values. Data are color coded by conformational state: S₁ (yellow), S₂ (blue) and S₃ (orange).

Fig. 5 | Ensemble distributions of \(\theta\) or \(\Omega\) angles determined from fluorescence polarization measurements. a, b, Ensemble distributions of three \(\theta\) (a) or \(\Omega\) (b) angles constructed with data obtained from 316 rhodamine-labeled regulatory module particles under nominal Ca\(^{2+}\)-free conditions, each of which was analyzed individually. Gaussian curves superimposed on the distributions were calculated with the mean and \(\sigma\) values of the respective angles. Below each distribution is the comparison of \(\theta\) or \(\Omega\) values (mean ± \(\sigma\), calculated from data used to build the distributions) with those measured from the crystal structures of RCK in three states\(^{43,45}\).

Among the three crystal structural states S₁, S₂ and S₃ of a given RCK domain in the regulatory module, the relative displacements of the \(\alpha\) carbon of residue K150, which is located midway between E146C and L153C used to attach the fluorophore, are between 3.4 Å\(^{43,45}\) and 8.1 Å\(^{43,45}\). These displacements should approximate those of the fluorophore’s center. Thus, by resolving small angle changes, we have successfully detected conformational changes involving displacements on the order of ångströms.

Discussion
We have showcased the application of a light microscope to quantitatively track the orientation changes of a fluorophore attached to a protein, which occur on a millisecond scale, by monitoring the changes in its emission anisotropy, with an effective \(\sigma\) of 2–5°, or
a mean of ~4° (Fig. 4a). This method allows us to detect multistate conformational changes in a single protein molecule and quantify the transitions in real time (50 frames per second) among the conformational states that differ on an ångström scale (Figs. 3 and 5 and Supplementary Video 1).

Because the angle resolution is limited by the SNR of light intensity signals, we experimentally and analytically maximized effective SNR using the following approach. First, we minimized the variance stemming from protein preparation and attachment. Next, we optimized the polarization microscope to achieve near-shot-noise-limited precision, and the expected full range of angle determination by using specific polarizing beam splitters and the implementation of a series of necessary numerical corrections. Then, by detecting the intensity change, we temporally demarcated individual events, extracting kinetic information. Finally, the demarcation of individual events gave us the opportunity to average data points within a given event to mitigate the impact of noise, thereby increasing the resolution of individual events.

We emphasize that the purpose of the present method is to resolve, in real time, individual conformations of a protein, not to determine its atomic structures. This method has two fundamental requirements: first, after successful attachment, the fluorescent label must produce recordable polarized emission with SNR adequate for required resolution; second, the changes in orientations of the label must report those of the individual conformational states.

In two following papers, we demonstrate that this approach enables the investigation of complex mechanisms underlying the protein’s conformational changes and how the resulting dynamic information can be used as a temporal template to link existing structures of the protein.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41594-019-0274-2.

Received: 3 July 2018; Accepted: 3 July 2019; Published online: 5 September 2019

References
1. Broccoli, L. & Karlin, S. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Res.* 33, 3390–3400 (2005).
2. Eriksson, H. P. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. In: Li, S. (Ed.) *Biological Procedures Online* 32–51 (Springer; 2009).
3. Sase, I., Miyata, H., Ishiwata, S. & Kinosita, K. Jr. Axial rotation of sliding actin filaments revealed by single-fluorophore imaging. *Proc. Natl Acad. Sci. USA* 94, 5466–5470 (1997).
4. Warshaw, D. M. et al. Myosin conformational states determined by single fluorophore polarization. *Proc. Natl Acad. Sci. USA* 95, 8034–8039 (1998).
5. Ha, T., Glass, J., Enderle, T., Chemla, D. S. & Weiss, S. Hindered rotational diffusion and rotational jumps of single molecules. *Phys. Rev. Lett.* 80, 2093–2096 (1998).
6. Adachi, K. et al. Stepping rotation of F1-ATPase visualized through angle-resolved single-fluorophore imaging. *Proc. Natl Acad. Sci. USA* 97, 7263–7267 (2000).
7. Sosa, H., Peterman, E. J., Moerner, W. E. & Goldstein, L. S. ADP-induced rocking of the kinesin motor domain revealed by single-molecule fluorescence polarization microscopy. *Nat. Struct. Biol.* 8, 540–544 (2001).
8. Forkey, J. N., Quinlan, M. E., Shaw, M. A., Corrie, J. E. & Goldman, Y. E. Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization. *Nature* 422, 399–404 (2003).
9. Beausang, J. F., Schroeder, H. W. 3rd, Nelson, P. C. & Goldman, Y. E. Twirling of actin by myosins II and V observed via polarized TIRF in a modified gliding assay. *Biophys. J.* 95, 5820–5831 (2008).
10. Rosenberg, S. A., Quinlan, M. E., Forkey, J. N. & Goldman, Y. E. Rotational motions of macro-molecules by single-molecule fluorescence microscopy. *Acc. Chem. Res.* 38, 583–593 (2005).
11. Forkey, J. N., Quinlan, M. E. & Goldman, Y. E. Measurement of single macromolecule orientation by total internal reflection fluorescence polarization microscopy. *Biophys. J.* 89, 1261–1271 (2005).
12. Backlund, M. P. et al. Simultaneous, accurate measurement of the 3D position and orientation of single molecules. *Proc. Natl Acad. Sci. USA* 109, 19087–19092 (2012).
13. Zhang, O., Liu, J., Ding, T. & Lew, M. Imaging the three-dimensional orientation and rotational mobility of fluorescent emitters using the Tri-spot point spread function. *Appl. Phys. Lett.* 113, 031103 (2018).
14. Jiang, Y. et al. Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* 417, 515–522 (2002).
15. Ye, S., Li, Y., Chen, L. & Jiang, Y. Crystal structures of a ligand-free MthK gating ring: insights into the ligand gating mechanism of K+ channels. *Cell* 126, 1161–1173 (2006).
16. Brünger, A. T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* 54, 905–921 (1998).
17. Fourkas, J. T. Rapid determination of the three-dimensional orientation of single molecules. *Opt. Lett.* 26, 211–213 (2001).
18. Ohmachi, M. et al. Fluorescence microscopy for simultaneous observation of 3D orientation and movement and its application to quantum rod-tagged myosin V. *Proc. Natl Acad. Sci. USA* 109, 5294–5298 (2012).
19. Lippert, L. G. et al. Angular measurements of the dynein ring reveal a stepping mechanism dependent on a flexible stalk. *Proc. Natl Acad. Sci. USA* 114, E4564–E4573 (2017).
20. Axelrod, D. Carboxyamine dye orientation in red cell membrane studied by microscopic fluorescence polarization. *Biophys. J.* 26, 557–574 (1979).
21. Chen, J. & Gupta, A. K. On change point detection and estimation. *Commun. Stat. Simul. Comput.* 30, 665–697 (2001).
22. Beausang, J. F., Goldman, Y. E. & Nelson, P. C. Change point analysis for single-molecule polarization total internal reflection fluorescence microscopy experiments. Methods Enzymol. 487, 431–463 (2011).
23. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, 2007).
24. Horn, R. Statistical methods for model discrimination. Applications to gating kinetics and permeation of the acetylcholine receptor channel. *Biophys. J.* 51, 255–263 (1987).
25. Lewis, J. H. & Lu, Z. Energetics of ångström-scale conformational changes in an RCK domain of the MthK K+ channel. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-019-0275-1 (2019).
26. Lewis, J. H. & Lu, Z. Integrating spatiotemporal features of a ligand-regulated, multistate allosteric protein. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-019-0276-0 (2019).

Acknowledgements
We thank Y. Zhou for technical assistance; Y. Goldman for inspiration and the use of microscopes in his laboratory to test our initial fluorescence-labeled protein samples; Y. Jiang and R. MacKinnon for providing the cDNA encoding MthK; L. Lippert and T. Dadosh for discussions; P. De Weer, T. Hoshi and B. Salzberg for critique of our manuscript at different stages of its development; and E. McCleskey for support. This study was supported by the grant no. GM055560 from the National Institute of General Medical Sciences of the National Institutes of Health to Z.L. During the initial period of the project, Z.L. was supported as an investigator by the Howard Hughes Medical Institute.

Author contributions
J.H.L and Z.L. designed the study; J.H.L assembled the polarization microscope, performed experiments, developed analytical tools, and analyzed the data with input from Z.L.; J.H.L and Z.L. interpreted the results and wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41594-019-0274-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to Z.L.

Peer review information: Ines Chen was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019
Methods

Production of recombinant isolated regulatory module protein of MthK. The cDNA of the RCK of the MthK channel from Methanobacterium thermoautotrophicum” (AA090033E9) was cloned into a pET28c vector containing a kanamycin-resistance gene. A 14-residue-long sequence was added to the N terminus for recognition by bionein ligase, and a 6-His-tag was added to the C terminus for affinity purification, with an intervening thrombin cutting sequence for removing the His-tag. We mutated Glu146 and Leu153 in the surface-exposed helix β6 to cysteines for covalent attachment of bifunctional rhodamine, leaving the three endogenous, buried cysteine residues in place. Mutagenesis was performed using a PCR-based method and verified via DNA sequencing.

Just prior to a recording, 100 mM dithiothreitol (DTT) (Fisher, BP172) was added to the sample coverslip to reduce disulfide bonds. One hundred micromolar biotin ligase (BiRA; Invitrogen B10621) with two iodoacetamide groups to E146C and L153C in the RCK protein was expressed in Escherichia coli (ANDOR iXon Ultra 897), where the four intensities from a given fluorophore appear in the corresponding positions of the four sections (Fig. 2b).

Fluorescence polarization microscopy and intensity recording. We built the fluorescence polarization microscope from a Nikon microscope (model Ti-E). To produce an evanescent field at the surface of the sample coverslip, a linearly polarized laser beam (532 nm) was directed to pass through a 1.4 λ plate, which transformed the linear polarization to circular polarization. After passing through a 100x objective (Nikon Microphot, NA = 1.49), the beam emerged at a 78° angle incident to the coverslip, the so-called critical angle that leads to TIR required for the formation of an evanescent field. The emission of polarized fluorescent light from individual protein molecules was captured by a dichroic mirror (Semrock FF01-593/40-25) that prevents the excitation of out-of-plane fluorescent light. The dichroic mirror was set at 540 nm to block excitation light from entering the detector. The emission of polarized fluorescent light from individual protein molecules was captured by a dichroic mirror (Semrock FF01-593/40-25) that prevents the excitation of out-of-plane fluorescent light. The dichroic mirror was set at 540 nm to block excitation light from entering the detector. The beam was further aimed at a photomultiplier tube (PMT) with an avalanche photodiode (APD) detector. The photon release rate, that is, intensity, changes when one event transitions to another (Supplementary Fig. 1a,b). When a camera is used as a detector, photons emerging from each channel are electronically binned over each frame. A series of consecutive k frames with a constant exposure time (Δt) is acquired as:

$$I_t = \frac{m_t}{\Delta t}$$

where $m_t$ is the number of photons released within frame $i$.

The cumulative distribution, $m_t$, is built by adding the number of photons for the successive time frames:

$$m_t = \sum_{i=0}^{t} n_i$$

The slope (that is, derivative) of the plot of $m_t$ against $t$ reflects the rate of photon release (Supplementary Fig. 1c). In theory, a transition occurs at the frame where the slope significantly changes. Quantitatively, slope changes, which result from relative photon release rates among the four polarization channels, should occur at the same frame for all channels; this redundancy increases the reliability of transition detection (Supplementary Fig. 1d). On the cumulative distribution plot of the logarithmic (log) likelihood for all four intensities together, the transition point occurs at a distinct maximum (Supplementary Fig. 1d), which forms the basis of automated detection.

Theoretically, for a given frame $i$, the probability $p_t$ of collecting $n_i$ number of photons, emitted at an average rate $r = n/mΔt$, is given by the Poisson distribution:

$$p_t = \left(\frac{\Delta t}{n}\right)^n e^{-\frac{n}{\Delta t}}$$

The likelihood ($L_t$) of the fluorophore emitting photons at this rate over all frames ($i = 0...T$) is given by:

$$L_t = \prod_{i=0}^{T} p_t = \prod_{i=0}^{T} \left(\frac{n_i}{n} e^{-\frac{n_i}{\Delta t}}\right)$$

The log likelihood ($LL_t$) of photon time is then:

$$LL_t = \ln \left(\prod_{i=0}^{T} p_t \right) = \sum_{i=0}^{T} \ln \left(\frac{n_i}{n} e^{-\frac{n_i}{\Delta t}}\right) = \ln(n_i) + \ln(n) - \ln(n_i) - \sum_{i=0}^{T} \ln(n_i)$$

where $N = \sum_{i=0}^{T} n_i$ is the total number of photons emitted over time $T$. During the interval $T$ the rate changes from $I_i$ to $I_j$ at the time point $t = iΔt$, and the number of emitted photons before this change is $m = \sum_{i=0}^{T} n_i$, and after the change is $N - m$, then the function describing the likelihood of a transition to occur at frame $i$ in the log form is:

$$LL_{i,j} = \ln \left(\prod_{i=0}^{T} p_{I_i} \prod_{j=0}^{T} p_{I_j} \right) = \ln(n_i) + \ln(N - m) - \ln(n_i) - \sum_{i=0}^{T} \ln(n_i)$$

A relation for comparing the two cases is given by the log likelihood ratio, that is $LL_{i,j}/LL_{i,i}$.

$$LL_{i,j} = \ln \left(\prod_{i=0}^{T} p_{I_i} \prod_{j=0}^{T} p_{I_j} \right) = \ln(n_i) + \ln(N - m) - \ln(n_i) - \ln(n_i)$$

This ratio operation eliminates $Δt$, and $LL_{i,j}$ is thus independent of $Δt$. The slope (that is, derivative) of the plot of $m_t$ against $t$ reflects the rate of photon release (Supplementary Fig. 1c). In theory, a transition occurs at the frame where the slope significantly changes. Quantitatively, slope changes, which result from relative photon release rates among the four polarization channels, should occur at the same frame for all channels; this redundancy increases the reliability of transition detection (Supplementary Fig. 1d). On the cumulative distribution plot of the logarithmic (log) likelihood for all four intensities together, the transition point occurs at a distinct maximum (Supplementary Fig. 1d), which forms the basis of automated detection.

Theoretically, for a given frame $i$, the probability $p_t$ of collecting $n_i$ number of photons, emitted at an average rate $r = n/mΔt$, is given by the Poisson distribution:

$$p_t = \left(\frac{\Delta t}{n}\right)^n e^{-\frac{n}{\Delta t}}$$

The likelihood ($L_t$) of the fluorophore emitting photons at this rate over all frames ($i = 0...T$) is given by:

$$L_t = \prod_{i=0}^{T} p_t = \prod_{i=0}^{T} \left(\frac{n_i}{n} e^{-\frac{n_i}{\Delta t}}\right)$$

The log likelihood ($LL_t$) of photon time is then:

$$LL_t = \ln \left(\prod_{i=0}^{T} p_t \right) = \sum_{i=0}^{T} \ln \left(\frac{n_i}{n} e^{-\frac{n_i}{\Delta t}}\right) = \ln(n_i) + \ln(n) - \ln(n_i) - \sum_{i=0}^{T} \ln(n_i)$$

where $N = \sum_{i=0}^{T} n_i$ is the total number of photons emitted over time $T$. During the interval $T$ the rate changes from $I_i$ to $I_j$ at the time point $t = iΔt$, and the number of emitted photons before this change is $m = \sum_{i=0}^{T} n_i$, and after the change is $N - m$, then the function describing the likelihood of a transition to occur at frame $i$ in the log form is:

$$LL_{i,j} = \ln \left(\prod_{i=0}^{T} p_{I_i} \prod_{j=0}^{T} p_{I_j} \right) = \ln(n_i) + \ln(N - m) - \ln(n_i) - \sum_{i=0}^{T} \ln(n_i)$$

A relation for comparing the two cases is given by the log likelihood ratio, that is $LL_{i,j}/LL_{i,i}$.
Further, to increase the reliability of the detection, we combine the likelihood function for each channel ($1 \ldots h$) to give a combined log-likelihood-ratio function:

$$LL_R = \ln \left( \prod_{r=1}^{h} \prod_{i=1}^{m} p(l_{i,r}) \prod_{i=1}^{m} p(l_{i,r}) \right)$$

(13)

Substituting equation (6) into equation (13) yields:

$$LL_R = \sum_{i=1}^{m} \ln \left( \frac{m_i}{\mu_i^2} \right) + (N_i - m_i) \ln \left( \frac{N_i - m_i}{\sigma^2} \right)$$

(14)

We set the threshold of significance for $LL_R$ at the level that limits the false positive events to 5% on the basis of simulation studies, and the resulting false negative events were approximately 1%.

The program started by identifying one transition over the entire trace. If a change point $X$ was identified, it would then search for additional transitions between the start of the trace and point $X$ and between $X$ and the end. This iterative search with successively shortened stretches continued until no more transitions were identified.

**State identification.** We calculated $\theta$ and $\phi$ with equations (63) and (61) (Supplementary Note 3) and identified multiple states that occurred concurrently in both angle traces, using a $k$-means clustering algorithm, which assumed values in each of $k$ number of states followed a Gaussian distribution$^{11}$. The algorithm operated on a ‘nearest neighbour’ principle, where each data point was assigned to one of $k$ distributions such that the mean value $\mu_k$ of the distribution had the shortest distance $d_{ik}$ to that of the interrogated data point (Supplementary Fig. 2). In terms of $\theta$ and $\phi$ of each event, such a distance was calculated between data point $x_i$ and $\mu_k$ as:

$$d_{ik} = \sqrt{(\mu_k - \theta_i)^2 + (\mu_k - \phi_i)^2}$$

(15)

This process was first performed against initial guess values of distribution means. The distance $d_{ik}$ was then recalculated and reassigned to the compared distributions on the basis of the resulting means. The process was repeated until there were no further reassignments. This algorithm itself offered a simple and adequate means to sort data against two states. However, in the case of multiple states, a $k$-means clustering algorithm was highly sensitive to initial guess values and proved to be extremely challenging, if not practically impossible without being coupled to adequate optimization algorithms. As such, we used a Nelder–Mead downhill simplex routine coupled to a simulated annealing routine$^{10}$ to optimize

the state identification program, which was based on a $k$-means clustering algorithm, in the form of the ratio of log likelihood functions:

$$LL_1 = \sum_i \left[ (\mu_k - \theta_i)^2 + (\mu_k - \phi_i)^2 \right]$$

for the case of one state and:

$$LL_k = \sum_i \left[ (\mu_k - \theta_i)^2 + (\mu_k - \phi_i)^2 \right]$$

(17)

for the case of $k$ number of states. The ratio of these two functions is given by:

$$\frac{LL_R}{LL_1} = \sum_i \left[ (\mu_k - \theta_i)^2 + (\mu_k - \phi_i)^2 \right]$$

(18)

Application of the $k$-means clustering algorithm coupled with the two optimization routines identified the most probable distribution to which each data point belonged, as shown in Supplementary Fig. 2. We have tested this state identification process with series of simulated traces where two states were separated by a varying fold of the standard deviation $\sigma$.

To assess the most probable number of states, we analyzed each set of data, as described above, in the framework of the models assuming a varying number of states, $k = 1, 2 \ldots n$. These models would be nested such that their differences should solely result from a different number of states. We thus judged whether one model of $k$ states was more compatible with the data than another model of $k + 1$ states by performing an $F$ test to determine (with a 95% confidence interval), on the basis of the statistical difference between the respective Chi-squared values ($X^2_k$ to $X^2_{k+1}$) calculated using equation (17). Thus, as a whole, the entire state identification program encompassed a $k$-means clustering algorithm, two optimization routines, and statistical evaluation of $X^2$ functions with an $F$ test$^{24}$.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

Data and materials described here will be made available upon reasonable request.

**Code availability**

Code used in the work will be made available upon reasonable request.

**References**

27. Axelrod, D. Total internal reflection fluorescence. *Annu. Rev. Biophys. Bioeng.* 13, 247–268 (1984).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed                                      |
|-----|------------------------------------------------|
|     | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
|     | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
|     | The statistical test(s) used AND whether they are one- or two-sided |
|     | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
|     | A description of all covariates tested |
|     | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
|     | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
|     | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable. |
|     | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
|     | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
|     | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated |
|     | Clearly defined error bars |
|     | State explicitly what error bars represent (e.g. SD, SE, CI) |

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection

Data collected using an emccd camera in conjunction with NIS elements from Nikon.

Data analysis

For each fluorescent point source, the time-dependent intensities were calculated using a custom ImageJ plug-in written in Java. From these intensities, angles were calculated using a separate program written in Labview. This same program also performs event detection and determines which of three states each event belongs to. Other programs, also written in Labview, were used for viewing the results of this analysis. Algorithms underlying these programs are described in the methods section. The programs themselves can be made available upon request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
Data

Policy information about availability of data.

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw data can be found in Figs. 3 and 5.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences  ☐ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Each distribution described in the paper and shown in figures was formed from events of observed individual MthK regulatory domains under different calcium conditions. Over all seven calcium conditions, ~2300 events were recorded, from which separate distributions were formed for each of the three conformational states. This ensures on average about 750 events per distribution. Sample sizes were not predetermined. However, each normal distribution is symmetric and well defined with standard deviations around 2 degrees, giving standard errors of around 0.3 degrees.

Data exclusions

Data were excluded on the following basis: 1) The model for polarized fluorescence has the expectation that although individual polarized intensities can vary over time, the total combined intensity should not. Therefore, any traces that showed any significant time dependent changes in the total intensity were excluded. 2) For a given recording, at least 10 events are required to obtain a 95% confidence level for state identification, so any trace with less is excluded. 3) For event detection and state identification, a signal to noise ratio greater than 4 is required for a 95% confidence level - traces with less are excluded. 4) States 1, 2 and 3 were identified according to theta - some traces were resolved only in phi and not in theta, and so were excluded from analysis.

Replication

These experiments were performed over 3 separate days. Each day’s experiments had comparable results.

Randomization

Randomization as is described here is not relevant. Because randomization is inherent to the process being observed (i.e. stochastic movement of a single protein) it is not under our control. This is a physical study of the response of a single type of protein (MthK) to a single type of ligand (calcium) of varying concentrations.

Blinding

Blinding is not relevant for the same reason that randomization is not relevant.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a

☑ Involved in the study

☐ Unique biological materials

☐ Antibodies

☐ Eukaryotic cell lines

☐ Palaeontology

☐ Animals and other organisms

☐ Human research participants

Methods

n/a

Involved in the study

☐ ChiP-seq

☐ Flow cytometry

☐ MRI based neuroimaging