Supplementary Information (SI) to accompany

Observing the mushroom-to-brush transition for kinesin proteins

Emmanuel LP Dumont, Herve Belmas, and Henry Hess*

*Henry Hess

Columbia University

Department of Biomedical Engineering

351L Engineering Terrace

MC 8904

1210 Amsterdam Avenue

New York, NY 10027

Fax: (212) 854-8725

Tel: (212) 854-7749

E-mail: hh2374@columbia.edu

This supplementary information contains three parts:

(1) Detailed experimental protocol

(2) Determination of the width of a microtubule footprint for use in landing rate measurements

(3) Measurement of kinesin grafting density with landing rate measurement

(4) The kinesin grafting density does not saturate up to 4,000 μm^{-2}.

(5) Bayesian algorithm to separate normal distributions

(6) Use of the algorithm to separate populations of microtubules adhering to avidin

(1) Detailed experimental protocol

Microtubules were polymerized by reconstituting a 20 µg aliquot of rhodamine-labeled, lyophilized tubulin (TL331M, Lot 357 from Cytoskeleton Inc, Denver, CO) with 6.25 µL polymerization buffer (BRB80 and 4 mM $MgCl_2$, 1 mM GTP, 5% dimethyl sulfoxide), and incubating it at 37 °C for 30 minutes. The microtubules were then stabilized by diluting them a thousand-fold into BRB80 buffer (80 mM piperazine-N,N′-bis(2-ethanesulfonic acid), 1 mM $MgCl_2$, 1 mM Ethylene Glycol Tetraacetic Acid, pH 6.9 with KOH) with 10 µM paclitaxel (Sigma, St Louis, MO). A kinesin construct consisting of the wild-type, full-length Drosophila melanogaster kinesin heavy chain and a C-terminal His-tag was expressed in Escherichia coli and purified using a Ni-NTA column.[1] Flow cells were constructed using one cover slip, one silicon wafer, and double-sided tape as spacer. Silicon wafers with a 20 nm oxide layer (Siliconsense Inc., 3'' diameter, <100>, SEMI std. flats, one side polished, prime grade, surface roughness < 2 Å, flatness < 9 µm) were used. For microtubules adhering to avidin, a solution of 1 µM avidin (A2667, Life Technologies) in BRB80 buffer was flown into the flow cell. After 5 min, the solution was exchanged with BRB80 to wash out avidin which did not adsorb. After another five minutes, this solution was exchanged with microtubule solution containing an enzymatic antifade system[2] (16 nM tubulin, 10 µM Paclitaxel, 20 mM D-glucose, 20 µg/mL glucose oxidase, 8 µg/mL catalase, 10 mM dithiothreitol, and 1 mM Adenylyl Imidodiphosphate AMP-PNP in BRB80). AMP-PNP (Sigma, St Louis, MO) is an ATP analogue which arrests motor action.[3] After another five minutes, the solution in the flow cell was exchanged with a solution containing the enzymatic antifade system only (10 µM Paclitaxel, 20 mM D-glucose, 20 µg/mL glucose oxidase, 8 µg/mL catalase, 10 mM dithiothreitol in BRB80). For microtubules adhering to kinesins, a solution of 0.5 mg/mL casein in BRB80 buffer was flown into each flow

cell. After 5 min, the solution was exchanged with the kinesin motor solution (kinesin, 0.5 mg/ml casein, 1 mM AMP-PNP in BRB80). After another five minutes, this solution was exchanged with microtubule solution containing an enzymatic antifade system[2] (16 nM tubulin, 0.5 mg/mL casein, 10 µM Paclitaxel, 20 mM D-glucose, 20 µg/mL glucose oxidase, 8 µg/mL catalase, 10 mM dithiothreitol, and 1 mM AMP-PNP in BRB80). After another five minutes, the solution in the flow cell was exchanged with a solution containing the enzymatic antifade system only (0.5 mg/mL casein, 10 µM Paclitaxel, 20 mM D-glucose, 20 µg/mL glucose oxidase, 8 µg/mL catalase, 10 mM dithiothreitol in BRB80). The openings of the flow cells were then sealed with a small amount of vacuum grease to prevent evaporation of the solutions. All experiments were performed at 24 °C.

Microtubules were imaged using a Nikon TE2000-U Epi-fluorescence microscope (Nikon, Melville, NY) equipped with an X-cite 120 lamp (EXFO, Ontario, Canada), an iXON DU885LC EMCCD camera (Andor, South Windsor, CT) and a 100x oil objective (NA 1.45).

For each flow cell, several fields of view were randomly selected and 150 to 600 measurements of the fluorescence intensity of microtubules were taken using a 100x objective, a cooled CCD camera and an exposure time of 40 s.

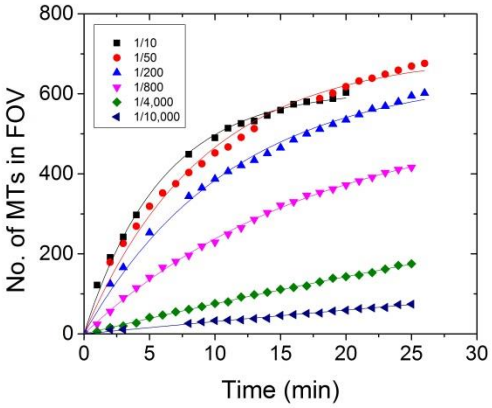
(2) Determination of the width of a microtubule footprint for use in landing rate measurements

In previous landing rate measurements, the interaction area of the microtubule with the kinesin on the surface was assumed to be equal to the "footprint" of the microtubule given by the product of average microtubule length and width (25 nm).[4-5] This was tested by comparing the product of kinesin density and interaction area obtained from landing rate measurements with the kinesin density determined from a measurement of total protein concentration ($C = 0.243$ mg/mL – George Bachand, Center for Integrated Nanotechnology, Sandia National Laboratory, private communication) and a measurement of the relative kinesin content from gel densitometry ($f = 34\%$, gel provided by George Bachand). The implied width is given by:

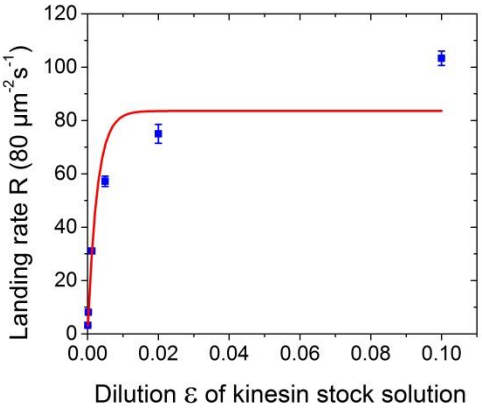
$$w = \frac{2M\rho}{fCHL}$$

where $M = 220$ kDa is the molecular weight of kinesin-1 heavy chains (from the database www.uniprot.org #P17210), $H = 78 \pm 2.4$ μm is the height of a typical flow cell, $L = 2.18 \pm 0.125$ μm is the average length of microtubules used in the landing rate measurements, and $\rho = A\sigma = 375 \pm 68$ is the product of interaction area and kinesin density for the landing rate measurements shown in Supplementary Figures 1 and 2.

Assuming a 5% uncertainty for the concentration C, we obtain $\boldsymbol{w = 20.0 \pm 5.0\ \mathbf{nm}}$, which is the value used for the width of the microtubule "footprint" in this work.



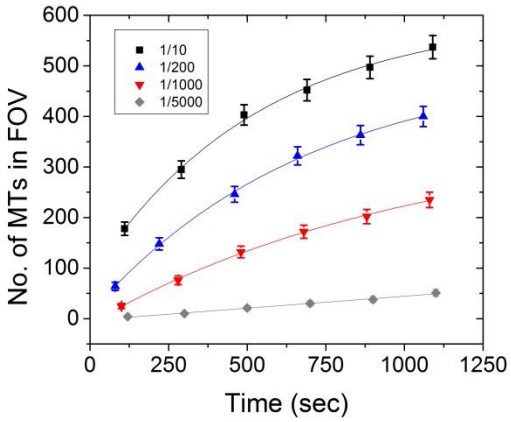
Supplementary Figure 1. The number of microtubules (MTs) attached to the surface as function of time for the casein-coated glass exposed to kinesin solutions diluted from the stock solution. The field of view (FOV) was 80 μm × 80 μm.



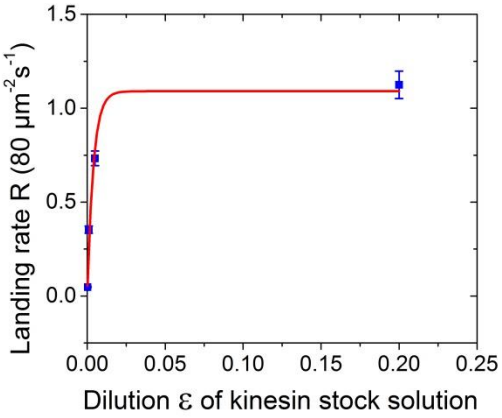
Supplementary Figure 2. Landing rates R computed from the data shown in Supplementary Figure 1 and plotted against the concentration of the kinesin solution for casein-coated glass surfaces. The equation $R = Z(1 - e^{-A\sigma\varepsilon})$ where A is the cross section area of a microtubule and σ is the kinesin grafting density corresponding to undiluted kinesin. We obtained $\rho = A\sigma = 375 \pm 68$.

(3) Measurement of kinesin grafting density with landing rate measurement:

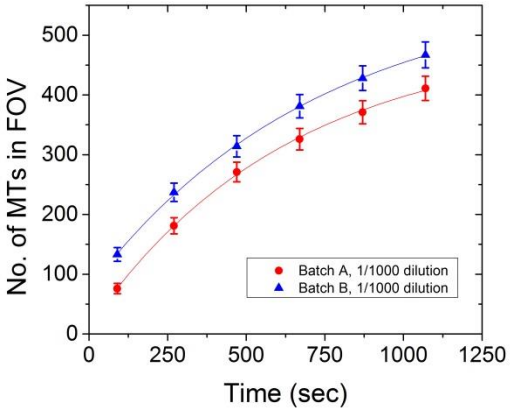
To measure the kinesin grafting densities for batches A and B, we used a third batch C of kinesins to determine the average length of microtubules ($L = 2.7 \pm 0.2$ µm) and the diffusion limited landing rate Z. To do so, we measured the number of microtubules per field of view at several dilutions for the batch C (Figure 3). To each curve, we fit the equation $N = N_0\left(1 - e^{-\frac{R(t-t_{ini})}{N_0}}\right)$ where R, t_{ini} and N_0 are fit parameters. We then plot the landing rate R as a function of the dilution ε as shown in Supplementary Figure 4. To this curve, we fit the equation $R = Z(1 - e^{-A\sigma\varepsilon})$ where $A = wL$ is the cross section area of a microtubule and σ is the kinesin grafting density corresponding to undiluted kinesin. We used $L = 2.7 \pm 0.2$ µm and $w = 20 \pm 5$ nm (determined in section (2) of SI). The fit yields $\rho = A\sigma = 256 \pm 91$ and $Z = 1.09 \pm 0.27$ (80 µm)$^{-2}$s^{-1}. We then ran a landing rate experiment for batch A and B at 1/1000 dilution (Supplementary Figure 5).



Supplementary Figure 3. The number of microtubules (MTs) attached to the surface as function of time for the casein-coated glass exposed to kinesin solutions diluted from the batch C stock solution. The field of view (FOV) was 80 µm × 80 µm.



Supplementary Figure 4. Landing rates *R* computed from the data shown in Supplementary Figure 3 and plotted against the concentration of the kinesin solution for casein-coated glass surfaces.



Supplementary Figure 5. The number of microtubules (MTs) attached to the surface as function of time for the casein-coated glass exposed to kinesin solutions diluted from the batch A and B stock solutions. The field of view (FOV) was 80 μm × 80 μm.

Since the same microtubules were used for Batch A, B, and C and since the cross section area of a microtubule *A* and the diffusion limited landing rate *Z* were previously determined with Batch C, the kinesin surface densities which would be obtained from undiluted stocks for Batch

A and B can be determined by fitting $N = N_0\left(1 - e^{-\frac{R(t-t_{ini})}{N_0}}\right)$ to their landing rate measurements (Supplementary Figure 5), finding the fit value for *R*, and using the following equation:

$$\sigma = -\frac{1}{A\varepsilon}\ln\left(1 - \frac{R}{Z}\right)$$

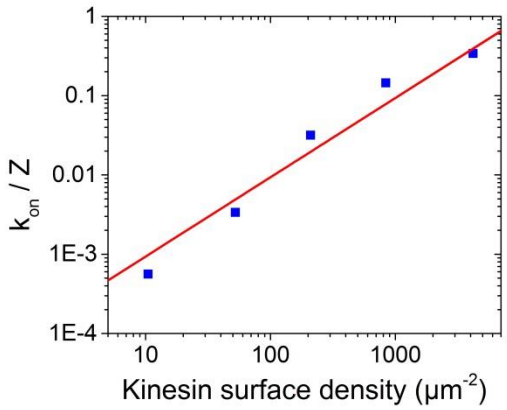
The error on the kinesin densities is given by:

$$\left(\frac{\delta\sigma}{\sigma}\right)^2 = \left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta\varepsilon}{\varepsilon}\right)^2 + \frac{(\delta Z/Z)^2 + (\delta R/R)^2}{[\ln(1 - R/Z)]^2}$$

We found the kinesin density to be 23,761 ± 7,841 μm^{-2} for batch A and 24,651 ± 8,022 μm^{-2} for batch B as the kinesin densities which would be obtained from undiluted stock solution.

(4) The kinesin grafting density does not saturate up to 4,000 μm^{-2}:

Agarwal et al. showed that the kinesin grafting density is proportional to the kinesin concentration in solution if the adsorption time is 5 min (Supplementary Figure 6).[6]



Supplementary Figure 6. Relative attachment rate constants computed by dividing attachment rate constants k_{on} with the diffusion limited maximal landing rate Z determined on a bare glass surface. The error bars are roughly the size of the data points and represent the standard error. Adapted from Agarwal et al.[6]

(5) Bayesian algorithm to separate normal distributions:

The fluorescence measurements of microtubules adhering to avidin are represented by N observations $y = (y_1, \dots, y_n)$. We assume that microtubules are either singles or aggregate to each other in doublets, triplets, etc. We designed a Bayesian clustering algorithm to differentiate these different populations. We assume that they are K populations of microtubules and that the fluorescence intensity of each population "k" follows a normal law of mean μ_k and standard deviation σ_k. We assume that a fluorescence measurement has the probability η_k to belong to the population "k". The probability density function f for the N observations y_n is given by:

$$\forall\, n \in [\![1, N]\!], \qquad f(y_n|\nu) = \sum_{k=1}^{K} \eta_k f_{\mathcal{N}_k}(y_n|\, \mu_k, \sigma_k^2)$$

Where $f_{\mathcal{N}_k}$ is the probability density function of a normal distribution of mean μ_k and standard deviation σ_k and $\nu = (\, \mu_1,\, \sigma_1^2,\, \eta_1, \dots,\, \mu_K,\, \sigma_K^2,\, \eta_K)$. We also have:

$$\forall\, k \in [\![1, K]\!],\, \eta_k \geq 0 \qquad \text{and} \qquad \sum_{k=1}^{K} \eta_k = 1$$

A common and convenient formulation introduces a latent unobserved random variable $S = (S_1, \dots, S_K)$ such as:

$$\forall\, k \in [\![1, K]\!], \qquad P(S_t = k) = \eta_k$$

The random variable S plays the role of an "indicator variable" and we can re-write the random variable Y that generates the observed data:

$$Y \sim \sum_{k=1}^{K} \mathbb{1}_{S_t=k} \, \mathcal{N}(\mu_k, \sigma_k^2)$$

Using Bayes' law, we obtain:

$$f(\nu, S \mid y) = f(y \mid \nu, S) \times f(S \mid \nu) \times f(\nu)$$

Where $f(\nu, S \mid y)$ is the distribution of posteriors, $f(y \mid \nu, S)$ is the likelihood function, $f(S \mid \nu)$ is the distribution of allocations conditional on knowing ν, and $f(\nu)$ is the distribution of priors. We can re-write the distribution of posteriors as:

$$f(\nu, S \mid y) \propto \left(\prod_{k=1}^{K} \prod_{n:S_n = k} \left(\frac{1}{\sigma_k^2} \right)^{\#_k/2} \exp\left\{ -\frac{1}{2} \sum_{n:S_n = k} \frac{(y_n - \mu_k)^2}{2\sigma_k^2} \right\} \right) \times \left(\prod_{k=1}^{K} \eta_k^{\#_k} \right) \times f(\nu)$$

Where $\#_k = \mathrm{Card}(\{n: S_n = k\})$. We then introduce the "prior" distributions of $f(\nu)$, that is μ_k, σ_k, and η:

$$\begin{cases} \forall\, k \in [\![1, K]\!] \;\; f(\mu_k | \sigma_k^2) \sim \mathcal{N}(a_0(k), A_0(k)\sigma_k^2) \\ \forall\, k \in [\![1, K]\!] \;\; f(\sigma_k^2) \sim \mathcal{G}^{-1}\big(b_0(k), B_0(k)\big) \\ f(\eta) \sim \mathcal{D}\big(c_0(1), \dots, c_0(K)\big) \end{cases}$$

Where $\mathcal{N}$ is a normal distribution, $\mathcal{G}^{-1}$ is an inverse-gamma distribution, and $\mathcal{D}$ is a Dirichlet distribution and where $a_0(k), A_0(k), b_0(k), B_0(k), \; c_0(k)$ are the prior parameters for $k \in [\![1, K]\!]$. The priors μ_k and σ_k are chosen to be dependent, a commonly used approach in Bayesian statistics because the "posterior" distributions $f(\mu_k, \sigma_k^2 \mid y)$ turn out to be in closed forms:

$$\begin{cases} \forall\, k \in [\![1,K]\!]\ \ f(\mu_k \mid \sigma_k^2, S, y) \sim \mathcal{N}(a_N(k), A_N(k)\sigma_k^2) \\ \forall\, k \in [\![1,K]\!]\ \ f(\,\sigma_k^2 \mid S, y) \sim \mathcal{G}^{-1}\big(b_N(k), B_N(k)\big) \\ f(\eta \mid S, y) \sim \mathcal{D}\big(c_N(1), \dots, c_N(K)\big) \\ \forall\, k \in [\![1,K]\!]\ \ \mathbb{P}(S_n = k \mid \nu, \mathrm{S}) = \mathbb{P}(S_n = k \mid \eta, \mathrm{S}) \propto \dfrac{1}{\sigma_k^2} \exp\left\{\dfrac{(y_n - \mu_k)^2}{2\sigma_k^2}\right\} \eta_k \end{cases}$$

And where the parameters $a_N(k)$, $A_N(k)$, $b_N(k)$, $B_N(k)$, $c_N(k)$ are the following:

$$\begin{cases} \forall\, k \in [\![1,K]\!]\ \ a_N(k) = \dfrac{A_0(k)}{\#_k A_0(k) + 1}\left(\dfrac{a_0(k)}{A_0(k)} + \#_k \bar{y}(k)\right) \\ \forall\, k \in [\![1,K]\!]\ \ A_N(k) = \dfrac{A_0(k)}{\#_k A_0(k) + 1} \\ \forall\, k \in [\![1,K]\!]\ \ b_N(k) = b_0(k) + \dfrac{\#_k}{2} \\ \forall\, k \in [\![1,K]\!]\ \ B_N(k) = B_0(k) + \dfrac{1}{2}\left(\#_k s_k^2(k) + \dfrac{\#_k / A_0(k)}{\#_k + A_0(k)}\big(\bar{y}(k) - b_0(k)\big)^2\right) \\ \forall\, k \in [\![1,K]\!]\ \ \bar{y}(k) = \dfrac{1}{\#_k} \sum_{n:S_n = k} y_n \\ \forall\, k \in [\![1,K]\!]\ \ s_k^2(k) = \dfrac{1}{\#_k} \sum_{n:S_n = k} \big(y_n - \bar{y}(k)\big)^2 \end{cases}$$

To draw the "posterior" distribution $f(\nu, S \mid y)$, we use Gibbs sampling, a Markov Chain Monte Carlo algorithm widely used in Bayesian inference.[7-9] The only requirement of Gibbs sampling is to be able to draw from the conditional distributions $f(\mu_k, \sigma_k^2 \mid S, y)$, $f(\eta \mid S, y)$, and $f(S \mid \eta, y)$ for $k \in [\![1,K]\!]$ which we know. The Gibbs sampler we use is the following:

Initialize: Start with an initial classification $S^{(0)}$:
for $d = 1, \dots, D_0, \dots, D$
  Update the parameters $\big(a_N(k)\big)_d, \big(A_N(k)\big)_d, \big(b_N(k)\big)_d, \big(B_N(k)\big)_d, \big(c_N(k)\big)_d$ using $S^{(d-1)}$
  Draw $\eta^{(d)}$ from $f\big(\eta^{(d)} \mid S^{(d-1)}, y\big) \sim \mathcal{D}\big(c_0(1), \dots, c_0(K)\big)$
  for k = 1,…, K
    Draw $\sigma_k^{2\,(d)}$ from $f\left(\sigma_k^{2\,(d)} \,\middle|\, S^{(d-1)}, y\right) \sim \mathcal{G}^{-1}\big(b_N(k), B_N(k)\big)$

Draw $\mu_k{}^{(d)}$ from $f\left(\mu_k{}^{(d)} \,\middle|\, \sigma_k^{2\,(d)},\ S^{(d-1)},\ y\right) \sim \mathcal{N}\left(a_N(k), A_N(k)\,\sigma_k^{2\,(d)}\right)$

End

Update the probabilities $\mathbb{P}\left(S^{(d)} = k \,\middle|\, \eta^{(d)}, \mathrm{y}\right)$

Draw a new classification $S^{(d)}$ using $\mathbb{P}\left(S^{(d)} = k \,\middle|\, \eta^{(d)}, \mathrm{y}\right)$

End

Drop the D_0 draws

Result: D draws $(\nu, S)^{(D_0+1)}, \ldots, (\nu, S)^{(D)}$

We chose to iterate the algorithm five times to obtain stable results. In the first iteration, we use prior parameters drawn from data points:

$$\begin{cases} \forall\, k \in [\![1, K]\!],\ a_0(k) = \bar{y} = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} y_n \\ \forall\, k \in [\![1, K]\!],\ A_0(k) = 1 \\ \forall\, k \in [\![1, K]\!], b_0(k) = 2 \\ \forall\, k \in [\![1, K]\!], B_0(k) = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} (y_n - \bar{y})^2 \\ \forall\, n \in [\![1, K]\!], c_0(k) = 1 \end{cases}$$

Each iteration leads to a set of posterior parameters and *D* drawings from the posterior distributions of $\hat{\mu}_k$, $\hat{\sigma}_k^2$, $\hat{S}$, and $\hat{\eta}_k$. We use these drawings to update the priors of the next iteration such as:

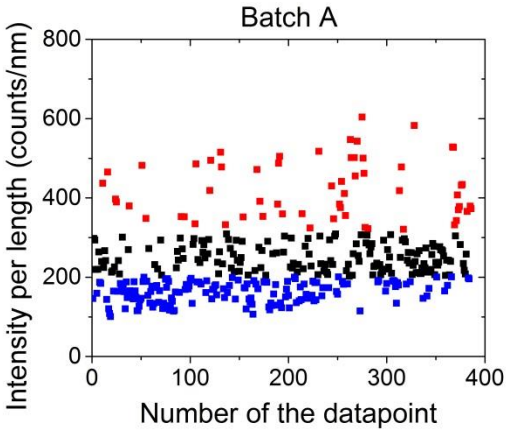
$$\begin{cases} \forall\, k \in [\![1, K]\!],\ a_0(k) = \dfrac{1}{D}\displaystyle\sum_{d=1}^{D} \hat{\mu}_k^{(d)} \\ \forall\, k \in [\![1, K]\!],\ A_0(k) = 1 \\ \forall\, k \in [\![1, K]\!], b_0(k) = 2 \\ \forall\, k \in [\![1, K]\!], B_0(k) = \dfrac{1}{D}\displaystyle\sum_{d=1}^{D} \hat{\sigma}_k^{(d)} \\ \forall\, k \in [\![1, K]\!],\ c_0(n) = \dfrac{1}{D}\displaystyle\sum_{d=1}^{D} \hat{\eta}_k^{(d)} \end{cases}$$

This algorithm leads to the results shown in the next section of supporting information.

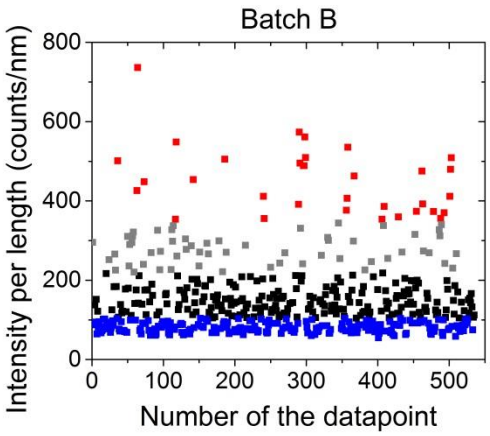
(6) Use of the algorithm to separate populations of microtubules adhering to avidin:

Taking the algorithm described in section (4) of the supporting information with $K = 5$, we ran the algorithm five times. If the algorithm only finds 4 Gaussians, we would re-run five times the algorithm with $K = 4$, etc.

For batch A, the algorithm found 3 normal distributions (Supplementary Figure 7). For the fluorescence of single microtubules adhering to avidin, the average of population 1 (blue squares) is 162.6 ± 4.0 counts/nm. For batch B, the algorithm found 4 normal distributions as seen below (Supplementary Figure 8). For the fluorescence of single microtubules adhering to avidin, the average of population 1 (blue squares) is 80.8 ± 1.5 counts/nm.



Supplementary Figure 7. Results of the Bayesian algorithm for microtubules adhering to avidin from Batch A. Blue squares are population 1, black squares are population 2, red squares are population 3.



Supplementary Figure 8. Results of the Bayesian algorithm for microtubules adhering to avidin from Batch B. Blue squares are population 1, black squares are population 2, grey squares are population 3, red squares are population 4.

References

1. Coy, D. L.; Hancock, W. O.; Wagenbach, M.; Howard, J., Kinesin's tail domain is an inhibitory regulator of the motor domain. *Nat Cell Biol* **1999,** *1* (5), 288-292.

2. Wetterma.G; Borglund, E.; Brolin, S. E., A Regenerating System for Studies of Phosphoryl Transfer from ATP. *Anal Biochem* **1968,** *22* (2), 211-218.

3. Vale, R. D.; Schnapp, B. J.; Reese, T. S.; Sheetz, M. P., Organelle, Bead, and Microtubule Translocations Promoted by Soluble Factors from the Squid Giant-Axon. *Cell* **1985,** *40* (3), 559-569.

4. Agarwal, A.; Luria, E.; Deng, x.; Lahann, J.; Hess, H., Landing Rate Measurements to Detect Fibrinogen Adsorption to Non-fouling Surfaces. *Cellular and Molecular Bioengineering* **2012,** *5* (3), 320-326.

5. Katira, P.; Agarwal, A.; Fischer, T.; Chen, H. Y.; Jiang, X.; Lahann, J.; Hess, H., Quantifying the performance of protein-resisting surfaces at ultra-low protein coverages using kinesin motor proteins as probes. *Adv Mater* **2007,** *19* (20), 3171-3176.

6. Agarwal, A.; Luria, E.; Deng, X. P.; Lahann, J.; Hess, H., Landing Rate Measurements to Detect Fibrinogen Adsorption to Non-fouling Surfaces. *Cellular and Molecular Bioengineering* **2012,** *5* (3), 320-326.

7. Diebolt, J.; Robert, C. P., Estimation of Finite Mixture Distributions through Bayesian Sampling. *J Roy Stat Soc B Met* **1994,** *56* (2), 363-375.

8. Raftery, A. E., Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika* **1996,** *83* (2), 251-266.

9. Bensmail, H.; Celeux, G.; Raftery, A. E.; Robert, C. P., Inference in model-based cluster analysis. *Stat Comput* **1997,** *7* (1), 1-10.