

Electronic Supplementary Information

Heterogeneous distribution of kinesin-streptavidin complexes revealed by Mass Photometry

Jing Xu,^{*a} Nathaniel J. S. Brown,^b Yeonee Seol^c and Keir C. Neuman^c

^a Department of Physics, University of California, Merced, CA 95343, USA

^b Department of Quantitative and Systems Biology, University of California, Merced, CA 95343, USA

^c Laboratory of Single Molecule Biophysics, National Heart, Lung and Blood Institute, NIH, Bethesda, MD 20892, USA

*Correspondence: jxu8@ucmerced.edu

Materials and Methods

Proteins

Dimeric kinesin K401-BIO-6xHIS (pWC2, Addgene)¹ carries the first 401 residues of *Drosophila* kinesin heavy chain linked to an 87-residue C-terminus of *E. coli* biotin carboxyl carrier protein followed by a hexahistidine tag. The expression and purification of k401-BIO-H6 were performed as previously described^{1, 2}. In brief, inoculated 2 L of Terrific Broth containing 100 µg/ml Carbenicillin and 25 µg/ml Chloramphenicol was grown overnight at 22°C after induction. Collected cells were lysed by freeze-thaw cycle and high-pressure emulsification (Microfluidics). Following Ni-NTA affinity purification, dimeric and monomeric forms of the protein were separated by size exclusion chromatography using Superdex-200 (Cytiva)^{3, 4}. The purity of kinesin proteins from purification was confirmed with denaturing SDS-PAGE shown in Figs. S2a and S7a. The efficiency of biotinylation of kinesin monomer was determined as ~77% using liquid chromatography-mass spectrometry (Agilent 6224 Accurate-Mass Time-of-Flight LC/MS) (Fig. S7b).

Streptavidin was purchased from Thermo Fisher (catalog # 21122). The purity of streptavidin proteins was confirmed with SDS-PAGE (Figs. S2a and S7a).

Preparation of kinesin-streptavidin complexes

Kinesin-streptavidin complexes were prepared by incubating kinesin dimers and streptavidin tetramers in 1xPEM80 buffer (80 mM PIPES, 2 mM MgCl₂ and 1 mM EGTA, pH 6.9; supplemented with 5 mM DTT) on ice for 30 min². The solution was then flash frozen in liquid nitrogen, stored at -80°C, and thawed immediately before mass photometry measurements.

For data shown in Figs. 2-4, the final concentration of streptavidin tetramer in the mixture was kept constant at 0.6 µM, the final concentration of kinesin dimers was varied between 0.4-2.16 µM.

For data shown in Fig. S9, the molar ratio of kinesin dimers to streptavidin tetramers was kept constant at 2.1:1. The final concentration of streptavidin tetramers in the mixture was varied between 0.3-2.4 µM, and the final concentrations of kinesin dimers was varied between 0.6-4.8 µM.

Concentrations of isolated proteins were estimated via absorption measurements at 280 nm on a NanoDrop 2000C spectrophotometer (Thermo Scientific). The mixing ratio for each preparation of kinesin-streptavidin complexes was determined via quantitative densitometry of proteins stained with Coomassie blue using an Amersham Typhoon imaging system (GE Healthcare Life Sciences) and quantified via ImageJ (<https://imagej.net/ij/>) (Fig. S2). Results of these two independent methods agreed to within ~30%. In our experience, densitometry-based measurements are more reliable and less variable.

Mass photometry experiments

All mass photometry experiments were performed at room temperature on a OneMP instrument (Refeyn)⁵. Sample wells were assembled by adhering a 6-well silicone gasket onto a clean coverglass (Refeyn MP-CON-21004). For each set of measurements, 10 µl of protein-free buffer (1xPEM80, supplemented with 5 mM DTT) was added to a new sample well, and the interface between the buffer and the coverglass surface of the sample well was brought into focus. 10 µL of sample (40 nM total protein

concentration) was then added to the same sample well, resulting in a final 20 nM total protein concentration. Binding of proteins or protein complexes onto the coverglass surface was then imaged within the 3 $\mu\text{m} \times 10 \mu\text{m}$ instrument field of view (Fig. 1b) and recorded for 1 min at a frame rate of 1 kHz. Each sample well was used once. Each preparation of kinesin-streptavidin complexes was measured 12-22 times independently. Reproducibility in mass photometry measurements was verified among independent experiments using the same sample (Fig. S3a) and using different preparations of the same kinesin-streptavidin mixture (Fig. S3b).

Data analysis

All mass photometry images were processed and analyzed using DiscoverMP (Refeyn). Interference intensities of protein complexes were converted to masses through a calibration with known mass standards⁵⁻⁷ (Fig. S1). The returned molecular masses were presented as mass distribution histograms with 6 kDa bin width. A lower-bound molecular mass of ~50 kDa was used for each mass distribution to account for the detection lower-bound of the OneMP instrument used in this study^{5, 7}. To account for skewness that is common in mass photometry data^{6, 8-11}, mass distributions were fitted to a bi-Gaussian mixture model, $\sum_{i=1}^K H_i \cdot f(x, \mu_i, \sigma_i, \omega_i)$, where K is the number of major species in the fit, μ_i is the molecular mass of the i^{th} mass species, H_i is the peak height of the i^{th} mass species, and $f(x, \mu_i, \sigma_i, \omega_i) = e^{-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2}$ for $x < \mu_i$, and $f(x, \mu_i, \sigma_i, \omega_i) = e^{-\frac{1}{2}\left(\frac{x-\mu_i}{\omega_i}\right)^2}$ for $x > \mu_i$.

The molecular masses of the major species in Fig. 3 were determined via best-fits of mass distributions to the bi-Gaussian mixture model described above. Only mass species with a pronounced peak profile (>40 in peak height, or >400 total counts) were included in the fit (red lines, Fig. 2) to ensure fitting accuracy.

The abundances of the three major complex species (iii-v, Fig. 3) in each mixture were determined via best-fits of mass distributions to the bi-Gaussian mixture model described above. All three major complex species were included in each fit (for example, Fig. S6a). For complex species without a pronounced peak profile, we employed our measurements of molecular masses in Fig. 3 to constrain their peak positions in the fit. The abundance of each major complex species was determined as $\frac{H_i \cdot (\sigma_i + \omega_i)}{2 \cdot \sqrt{2\pi}}$; the associated uncertainty was determined via error propagation of fitting uncertainties.

The abundance of the higher-order complexes was determined as the cumulative counts of masses ≥ 450 kDa (Fig. S4); the associated uncertainty was determined as the counting noise.

Relative abundances of different species of kinesin-streptavidin complexes were determined as the abundance of individual species, normalized by their total abundance. Uncertainties of relative abundances were determined via error propagation. Calculations of relative abundance excluded the isolated proteins. The resulting relative-abundance calculations agreed well between experiments using different preparations of the same kinesin-streptavidin mixture (Fig. S6).

Binding model

Monte Carlo simulations were employed to model the distribution of kinesin-streptavidin complexes. Each streptavidin tetramer was modelled as containing four identical sites. Each kinesin dimer was modelled as having up to two identical biotins. Kinesin dimers without a biotin were assumed to not bind

streptavidin. A kinesin dimer with one biotin was modelled as a single biotin and could bind a single streptavidin site. A kinesin dimer with two biotins was modelled as two identical biotins coupled together: once the first biotin bound the streptavidin, the second biotin would also bind, provided that there was an open site on the streptavidin. A kinesin dimer with two biotins was assumed to be twice as likely to encounter a streptavidin molecule than a single biotin.

Under these assumptions and denoting b as the kinesin-biotinylation efficiency, the fraction of kinesin dimers with two biotins was b^2 , the fraction of kinesin dimers with one biotin was $2b(1 - b)$, the probability that streptavidin encountering a kinesin dimer with two biotins was b , and the probability of streptavidin encountering a kinesin dimer with one biotin was $1 - b$. At each encounter, the probability that the first (or only) biotin on a kinesin dimer bound a streptavidin site was approximated as the molar ratio of biotin to streptavidin monomers present in the mixture, $\frac{1}{2}bRs$, where b is the kinesin-biotinylation efficiency denoted above, R is the molar mixing ratio determined in experiments, and s is an overall scaling factor for the experimentally determined mixing ratios.

Each simulation returned the stoichiometry of a single complex. The simulation was repeated 100000 times to determine the distribution of kinesin-streptavidin complexes for each value of the kinesin-biotinylation efficiency (b), the mixing ratio (R), and the overall mixing-ratio scaling factor (s). Best-fit values of the b and s parameters were determined by globally minimizing the reduced χ^2 between the model and the experimental data over all mixing ratios (Fig. S8).

Supplementary Figures

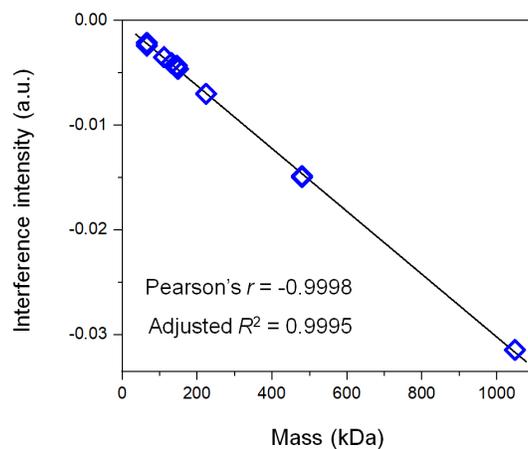


Fig. S1 Calibration of interference intensity versus molecular mass for the OneMP instrument (Refeyn, UK)⁵⁻⁷. Black line, best linear fit. Pearson's $r = -0.9998$ and adjusted $R^2 = 0.9995$.

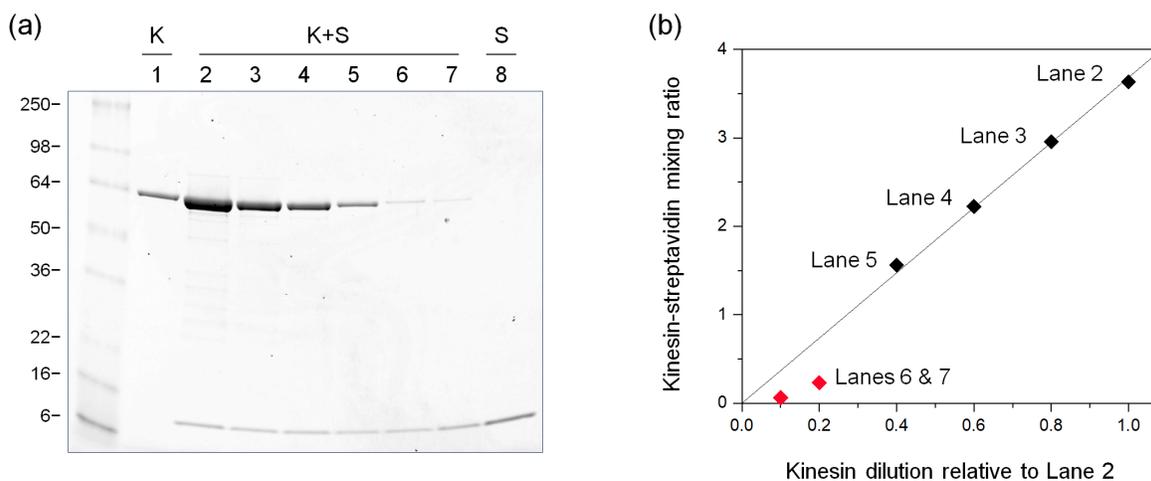


Fig. S2 Kinesin-streptavidin mixing ratio determined via quantitative densitometry of proteins stained with Coomassie blue. (a) Denaturing SDS-PAGE gels of kinesin (K), streptavidin (S), and their mixtures (K+S). Positions of molecular weight standards (in kDa) are indicated. (b) The mixing ratio of kinesin dimer to streptavidin tetramer as a function of kinesin dilution relative to Lane 2. Black line, best linear fit with zero intercept. Pearson's $r = 0.9998$ and adjusted $R^2 = 0.9995$. For complexes in Lanes 6 & 7, the intensities of kinesin bands in the gel (panel a) were below the linear range of densitometry measurements. These densitometry results were therefore excluded from the linear fit (red points, panel b), and the mixing ratios were estimated via the best-fit result (0.8 and 0.4 kinesin dimer per streptavidin tetramer, respectively).

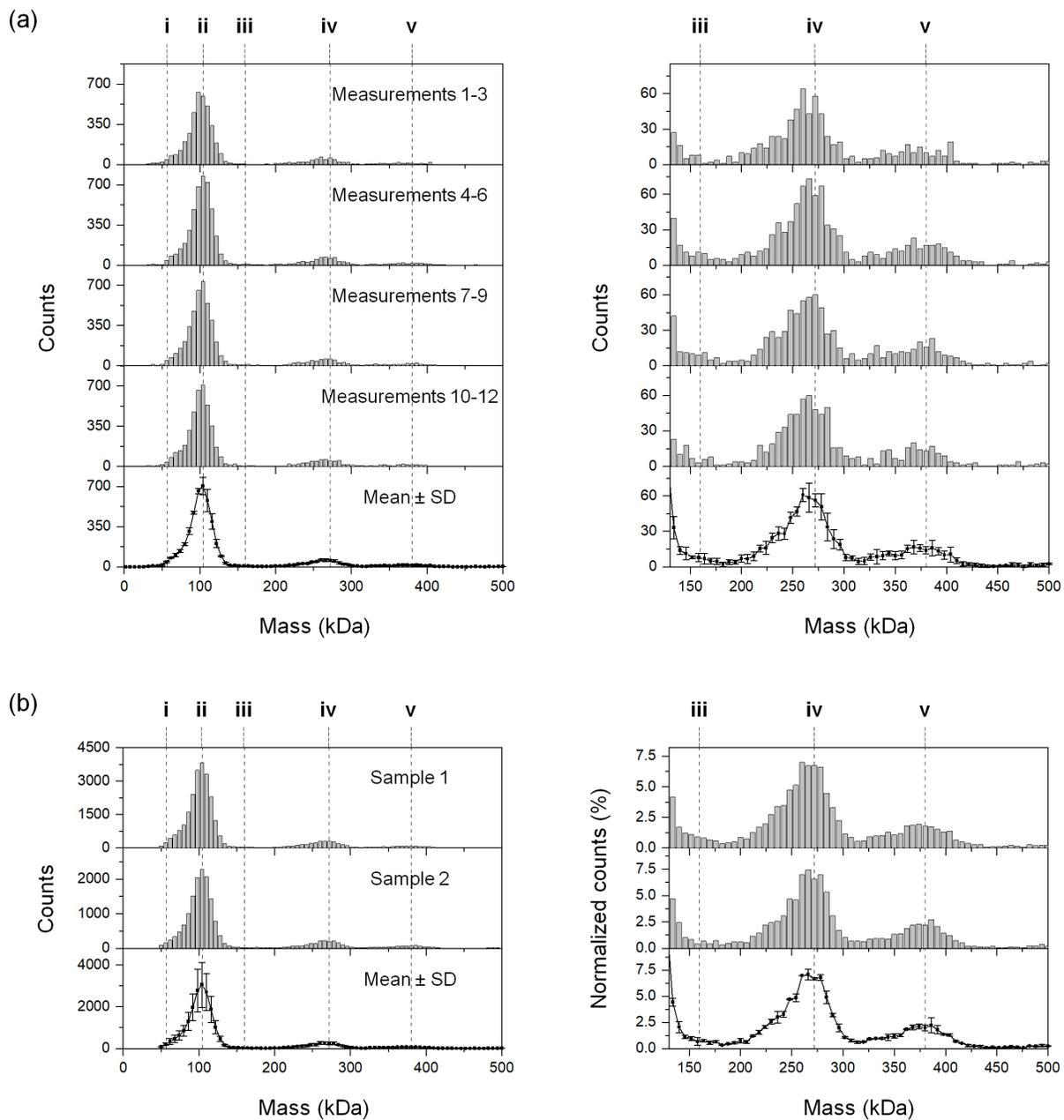


Fig. S3 Reproducibility of mass photometry measurements among independent experiments using the same sample preparation (a) and using different preparations of the same kinesin-streptavidin mixture (b). The concentration of streptavidin was kept constant at $0.6 \mu\text{M}$. The mixing ratio was kept constant at 3.0 kinesin dimers per streptavidin tetramer. Dashed lines indicate mass species identified in the current study. Expanded view of mass species iii-v are shown on the right. (a) Measurements using the same preparation of kinesin-streptavidin mixture (Lane 3, Fig. S2). Top four panels, mass distributions using data pooled from three independent measurements; measurements were pooled in triplets to

increase counting statistics. Bottom panel, mean (\pm standard deviation) of the mass distributions in the top four panels. (b) Measurements using two different sample preparations of the same kinesin-streptavidin mixture. Sample 1 corresponds to data pooled from the twelve independent measurements shown in (a). Sample 2 corresponds to data pooled from fifteen independent measurements of a second preparation using the same mixing ratio. Mass distributions are shown in raw counts (left) and in counts normalized by the total counts of kinesin-streptavidin complexes (right). Error bars indicate standard deviation.

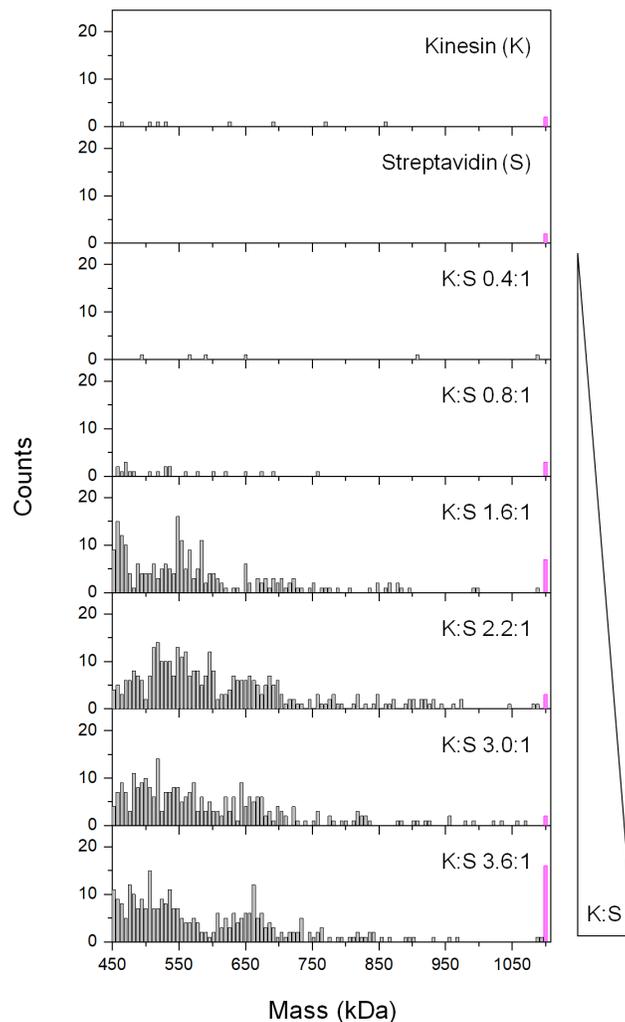


Fig. S4 Expanded view of mass distributions from Fig. 2, showing masses larger than 450 kDa for solutions containing kinesin (K), streptavidin (S), and their mixtures (K:S). K:S indicates the molar mixing ratio of kinesin dimers to streptavidin tetramers in each incubation. The concentration of streptavidin tetramers in each mixture was kept constant at 0.6 μ M, and the concentration of kinesin dimers was varied. Magenta bars indicate cumulative counts of measurements with masses exceeding 1100 kDa.

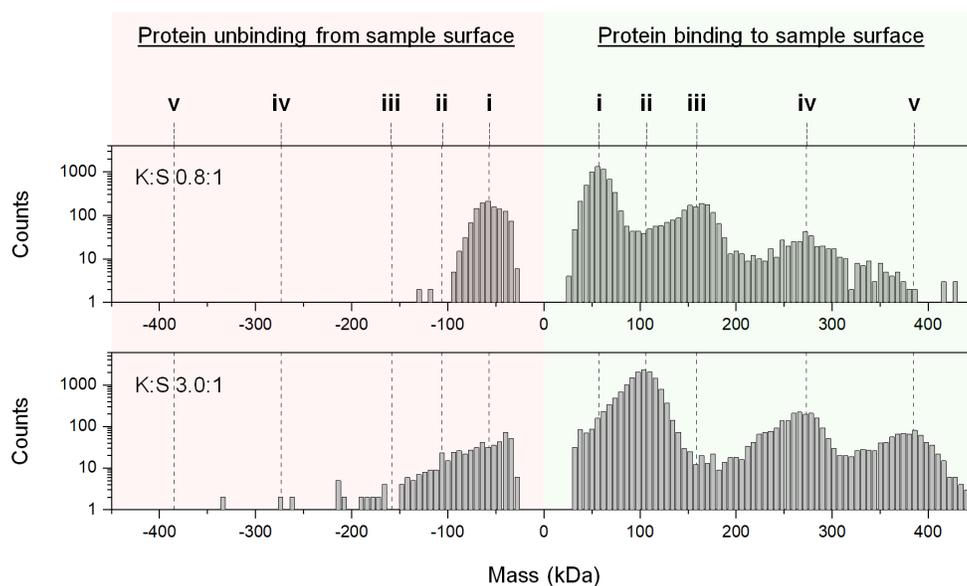


Fig. S5 Example mass photometry measurements indicating substantial unbinding of isolated proteins (i-ii), but not kinesin-streptavidin complexes (iii-v), from the sample surface. Negative mass readings (pink region) indicate unbinding events. Positive mass readings (blue region) indicate binding events. The unbound protein can be counted multiple times, through repeated rebinding and unbinding events, artifactually increasing the molecular counts. We therefore excluded the isolated proteins from calculations of the relative abundance of different complex species.

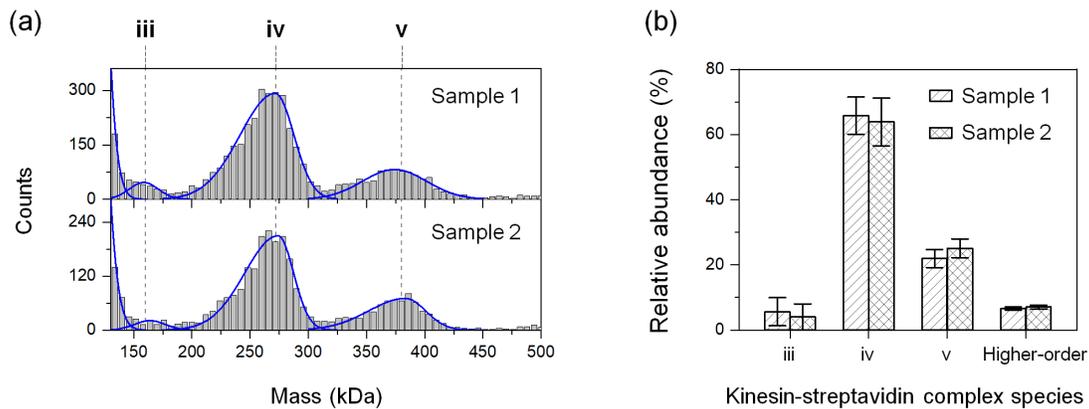


Fig. S6 Reproducibility in relative-abundance calculations between experiments using two independent preparations of the same kinesin-streptavidin mixture. The concentration of streptavidin was kept constant at $0.6 \mu\text{M}$. The mixing ratio was kept constant at 3.0 kinesin dimers per streptavidin tetramer. (a) Distributions of molecular masses. Blue lines indicate the best-fitted bi-Gaussian peaks for individual mass species. Mass distribution of sample 1 is shown in Fig. 2 (K:S 3.0:1). (b) Relative abundance of complex species determined for the two samples. Calculations of relative abundance excluded the isolated proteins. For complex species iii-v, relative abundances were calculated based on best-fitting results in panel a (blue lines); error bars indicate the associated fitting uncertainties. For higher-order complexes, relative abundances were estimated as the cumulative counts of masses ≥ 450 kDa; error bars indicate the associated counting noise.

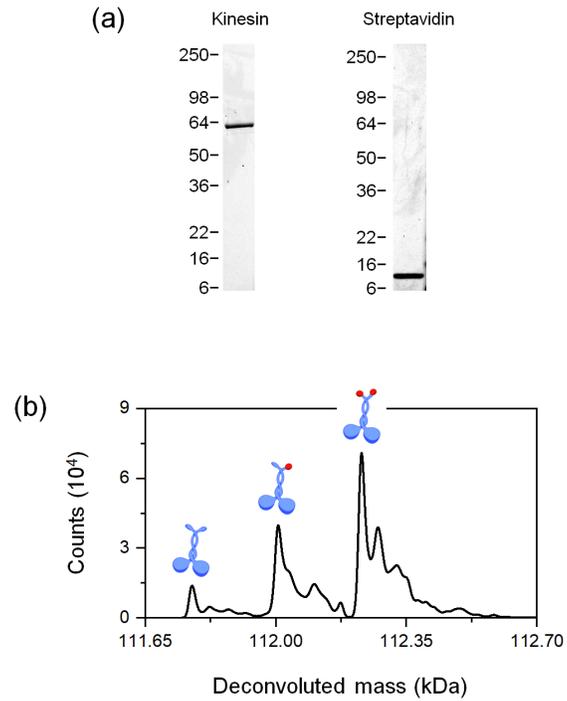


Fig. S7 (a) Representative denaturing SDS-PAGE of kinesin and streptavidin proteins used in this study. (b) Liquid chromatography-mass spectrometry (LC-MS) spectra of kinesin. Cartoons illustrate dimeric kinesins (blue) containing 0, 1, and 2 biotins (red). Relative abundances of kinesins with 0 biotin (8%), 1 biotin (30%), and 2 biotins (62%) correspond to ~77% efficiency of biotinylation of the kinesin monomer.

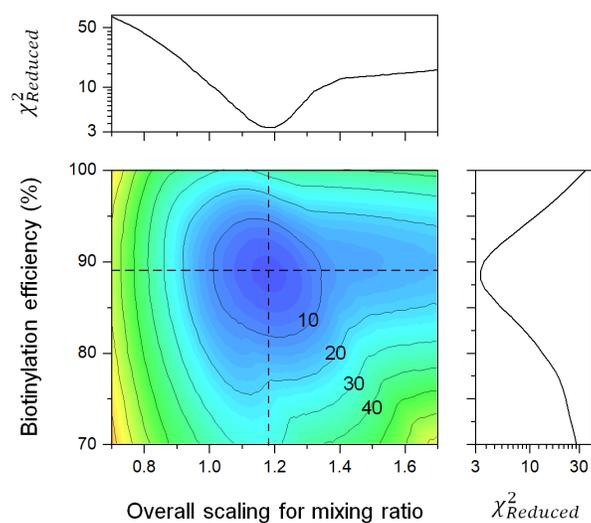


Fig. S8 Contour plot of reduced χ^2 as a function of the kinesin-biotinylation efficiency and the overall scaling for the kinesin-streptavidin mixing ratio in the binding model. Reduced χ^2 contours of 10, 20, 30, 40 are as indicated. Dashed lines, parameter values that minimize the reduced χ^2 between the model and the experiments, corresponding to an 88% efficiency in kinesin biotinylation and an overall scaling factor of 1.19 for the kinesin-streptavidin mixing ratio.

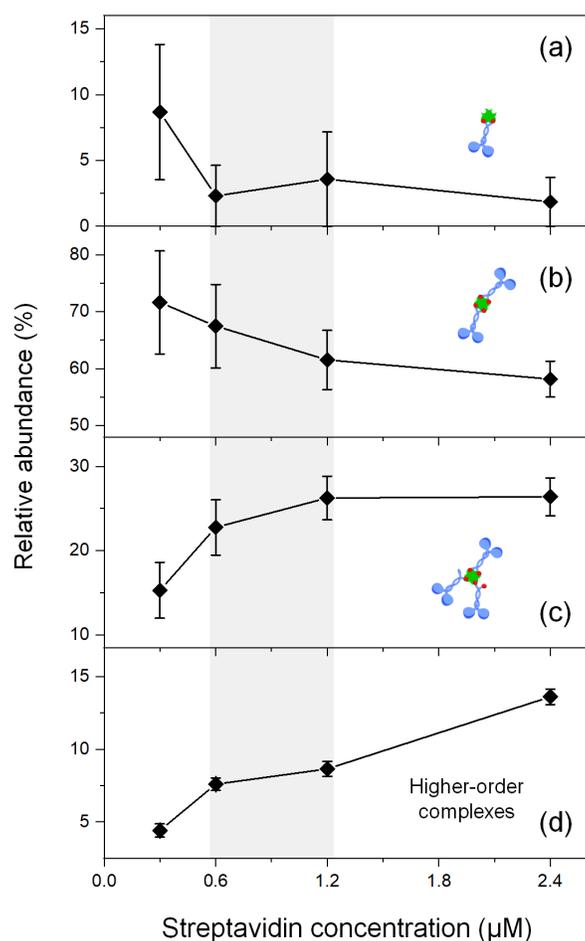


Fig. S9 Relative abundance of kinesin-streptavidin complex species as function of the concentration of streptavidin tetramers in the mixture. The mixing ratio was kept constant at 2.1 kinesin dimers per streptavidin tetramer. Calculations of relative abundance excluded the isolated proteins. (a-c) Complexes with 1:1, 2:1, and 3:1 stoichiometry. Relative abundances were determined via best-fits of mass distributions to a bi-Gaussian mixture model; error bars indicate the associated fitting uncertainties. (d) Higher-order complexes. Relative abundances were estimated as the cumulative counts of masses ≥ 450 kDa in the mass distribution; error bars indicate the associated counting noise. Grey region, the relative abundance of complex species remained largely unchanged for streptavidin concentrations ranging between ~ 0.6 - 1.2 μM .

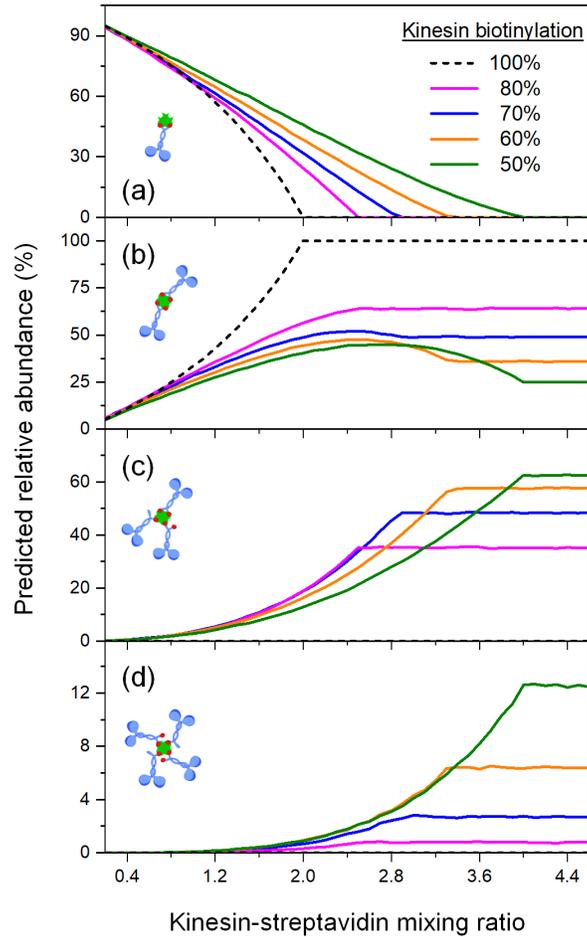


Fig. S10 Predicted relative abundance of kinesin-streptavidin complex species as function of the kinesin-streptavidin mixing ratio for five different kinesin-biotinylation efficiencies. Note that our simple binding model makes explicit the relative abundances of complexes with well-defined stoichiometries (a-d), but does not consider higher-order complexes. Calculations of relative abundance excluded the isolated proteins.

Supplementary References

- 1 R. Subramanian and J. Gelles, *J. Gen. Physiol.*, 2007, **130**, 445-455.
- 2 A. M. Tayar, L. M. Lemma and Z. Dogic, *Methods Mol. Biol.*, 2022, **2430**, 151-183.
- 3 L. C. Kapitein, B. H. Kwok, J. S. Weinger, C. F. Schmidt, T. M. Kapoor and E. J. Peterman, *J. Cell Biol.*, 2008, **182**, 421-428.
- 4 J. Xu, B. J. Reddy, P. Anand, Z. Shu, S. Cermelli, M. K. Mattson, S. K. Tripathy, M. T. Hoss, N. S. James, S. J. King, L. Huang, L. Bardwell and S. P. Gross, *Nat. Commun.*, 2012, **3**, 754.
- 5 Refeyn Ltd, *OneMP (version 1.3) User Manual*.
- 6 G. Young, N. Hundt, D. Cole, A. Fineberg, J. Andrecka, A. Tyler, A. Olerinyova, A. Ansari, E. G. Marklund, M. P. Collier, S. A. Chandler, O. Tkachenko, J. Allen, M. Crispin, N. Billington, Y. Takagi, J. R. Sellers, C. Eichmann, P. Selenko, L. Frey, R. Riek, M. R. Galpin, W. B. Struwe, J. L. P. Benesch and P. Kukura, *Science*, 2018, **360**, 423-427.
- 7 D. Wu and G. Piszczek, *Anal. Biochem.*, 2020, **592**, 113575.
- 8 M. Liebel, J. T. Hugall and N. F. van Hulst, *Nano Lett.*, 2017, **17**, 1277-1281.
- 9 A. Fineberg, T. Surrey and P. Kukura, *J. Mol. Biol.*, 2020, **432**, 6168-6172.
- 10 Y. Li, W. B. Struwe and P. Kukura, *Nucleic Acids Res.*, 2020, **48**, e97.
- 11 E. Bertosin, P. Stommer, E. Feigl, M. Wenig, M. N. Honemann and H. Dietz, *ACS Nano*, 2021, **15**, 9391-9403.