Biophysical Journal, Volume *112*

Supplemental Information

Kinesin Processivity Is Determined by a Kinetic Race from a Vulnerable One-Head-Bound State

Keith J. Mickolajczyk and William O. Hancock

SI Text

Hidden Markov Model for fitting short and long states in high resolution (X,Y,t) data

Nanometric (X,Y) position versus time data returned from point spread function fitting to gold nanoparticle movies were fit with a Hidden Markov Model (HMM) in order to determine points of state transition between the long state on the microtubule binding sites (16.4 nm intervals) and the short state in between the microtubule binding sites. First, the X,Y data were rotated to minimize standard deviation in the X direction, thus aligning Y with the microtubule axis. Next, the experimental noise in Y, σ, was determined using the pair wise differences method described previously (28, 62). Then, the tDetector model-free step-finding algorithm (62) was applied to the Y data, and the Y data as a whole were translated to best align the output staircase with an ideal staircase with 8.2 nm spacing. Next, a transition probability matrix was structured using the state space drawn in **Fig. S1**:

$$A = \begin{bmatrix} 1-g & g/3 & g/3 & g/3 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1-g & 0 & g/2 & g/2 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1-g & g/2 & g/2 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1-g & g & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-g & g/3 & g/3 & g/3 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-g & 0 & g/2 & g/2 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-g & g/2 & g/2 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-g & g & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}$$

Where g is defined as the number of potential 8.2 nm steps in the trace divided by the number of data points in the trace. Row and column indices matched the state numbering in **Fig. S1**. Next, an emission matrix containing parameters for a 2D uncorrelated symmetric Gaussian distribution was structured:

$$B = \begin{bmatrix} 0 & 0 & \sigma \\ 0 & -3\sigma & \sigma \\ 0 & 3\sigma & \sigma \\ 8.2 & 0 & \sigma \\ 16.4 & 0 & \sigma \\ 16.4 & -3\sigma & \sigma \\ 16.4 & 3\sigma & \sigma \\ 24.6 & 0 & \sigma \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Where the first column represents mean Y position values of Gaussian emitters placed according to the diagram in **Fig. S1**, the second column represents mean X position values, and the third column represents standard deviation values. Row indices match the numbering in **Fig. S1**. 3σ was used as the minimum detectable positional change in X. The transition matrix A, emission matrix B, and experimental data were then used as input to the Viterbi algorithm (63), which returned the mostly likely sequence of hidden states. Indices of state transition were then used to call boundaries between successive long and short states. Traces had to be split into ~100-200 nm subtraces to be amenable to the HMM fitting algorithm, as a perfect 8.2 nm register in the experimental data was gradually lost over time due to system drift. The model was implemented in MATLAB software (MathWorks).

Calculation of rate constants for predicting run lengths under first passage model

In order to calculate all the necessary rate constants for modeling run lengths according the model set up in **Fig. 4A**, the following system of equations was solved for each kinesin-1 and kinesin-2 using the cold ATP 1HB duration and GFP run length data:

$$\begin{cases} \boldsymbol{\tau_{1HB}^{14} = \tau_{Hydrolysis} + \tau_{Attach}^{14}} \\ \boldsymbol{\tau_{1HB}^{17} = \tau_{Hydrolysis} + \tau_{Attach}^{17}} \\ \boldsymbol{RL^{14} = (8.2)\,\tau_{Detach}/\tau_{Attach}^{14}} \\ \boldsymbol{RL^{17} = (8.2)\,\tau_{Detach}/\tau_{Attach}^{17}} \\ \boldsymbol{\tau_{Detach} = \tau_{Detach}^{FH} = \tau_{Detach}^{RH}} \end{cases}$$

State times τ are the inverse of the state exit rate k, and 8.2 is the distance per transition in nm. These equations assume that extending the neck linker does not affect $k_{Hydrolysis}$ or k_{Detach}, the latter of which was shown experimentally in **Fig. 3C**. Solutions to these equations are shown in **Table S1**. The following transition probability matrix was then set up:

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ P_{21} & 0 & P_{23} & 0 & 0 & 0 \\ 0 & P_{32} & 0 & P_{34} & P_{35} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ P_{51} & 0 & 0 & 0 & 0 & P_{56} \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Where row and column indices correspond to the state numbers in **Fig. 4A**. Transition probabilities were calculated from the rate constants, for example $P_{34} = k_{Off}^{ADP}/(k_{Detach}^{RH} + k_{Off}^{ADP} + k_{Detach}^{FH})$. The absorbing state (row and column one) was then removed to create a matrix of transient states $\boldsymbol{P^T}$, and the expected number of time periods the motor spends in each transient state per run was calculated as $\boldsymbol{S} = (\boldsymbol{I} - \boldsymbol{P^T})^{-1}$, where $\boldsymbol{I}$ is the identity matrix. S_{15}, the expected number of time periods the motor spent in the ATP waiting state given that it started in the 1HB post-hydrolysis state, was taken as the processivity, and converted to a run length by multiplying by 8.2 nm. Simulations and all data analysis was done in MATLAB (MathWorks).

Supplementary Figures

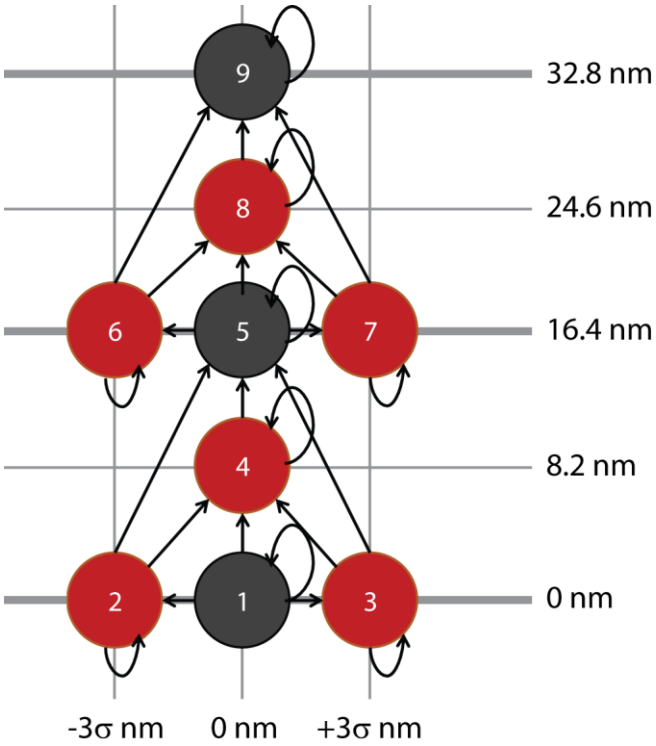

Fig. S1. State space diagram for 2D Hidden Markov Model used to fit long and short states in high resolution tracking data. The nanoparticle-labeled motor domain takes full 16.4 nm steps, with ~8.2 nm substeps that may or may not have a positional component perpendicular to the microtubule. Long states (where the labelled head is bound to the microtubule) are shown as black circles, and short states (labeled head is off the microtubule) comprise the three red circles in between each long state. Arrows denote potential transitions. σ denotes the standard deviation of the particular trace in plateau regions between stepping events; ±3σ is set as the minimum detectable lateral displacement.

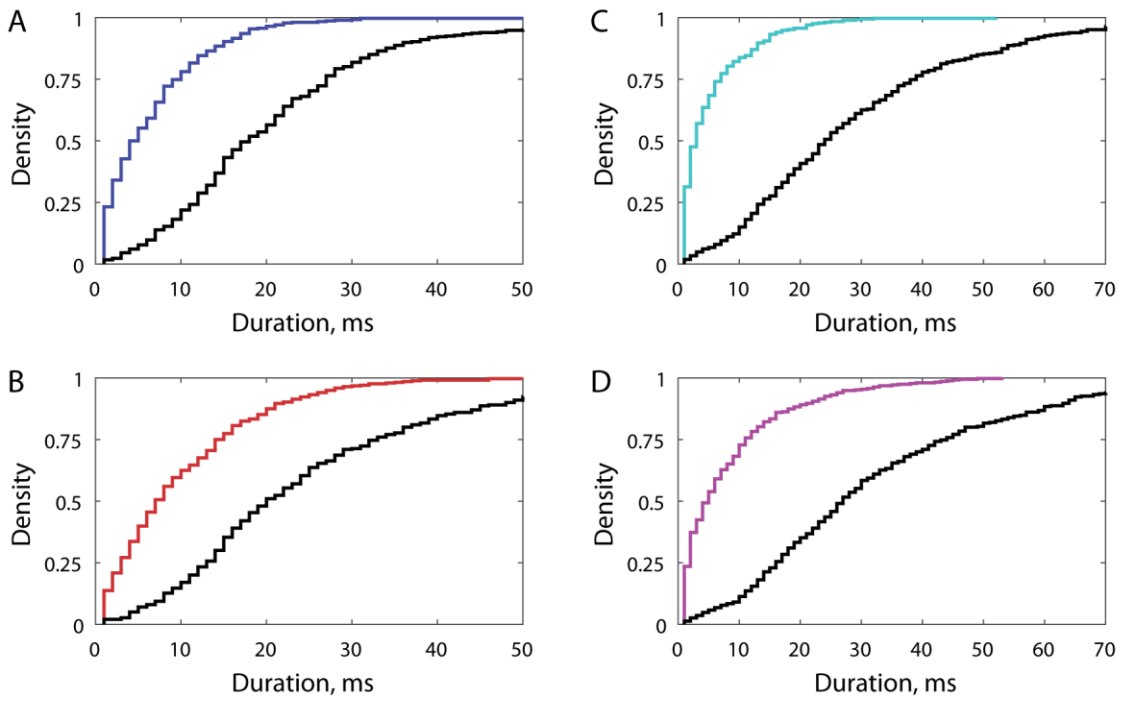

Fig S2. Empirical cumulative density functions of long and short states for all motors. (**A**) Distributions of N=336 long states (black) and N=300 short states (blue) for $K1_{14}$. Data come from 44 molecules. (**B**) Distributions of N=336 long states (black) and N=300 short states (red) for $K1_{17}$. Data come from 37 molecules. (**C**) Distributions of N=370 long states (black) and N=326 short states (cyan) for $K2_{14}$. Data come from 47 molecules. (**D**) Distributions of N=349 long states (black) and 310 short states (magenta) for $K2_{17}$. Data comes from 73 molecules.

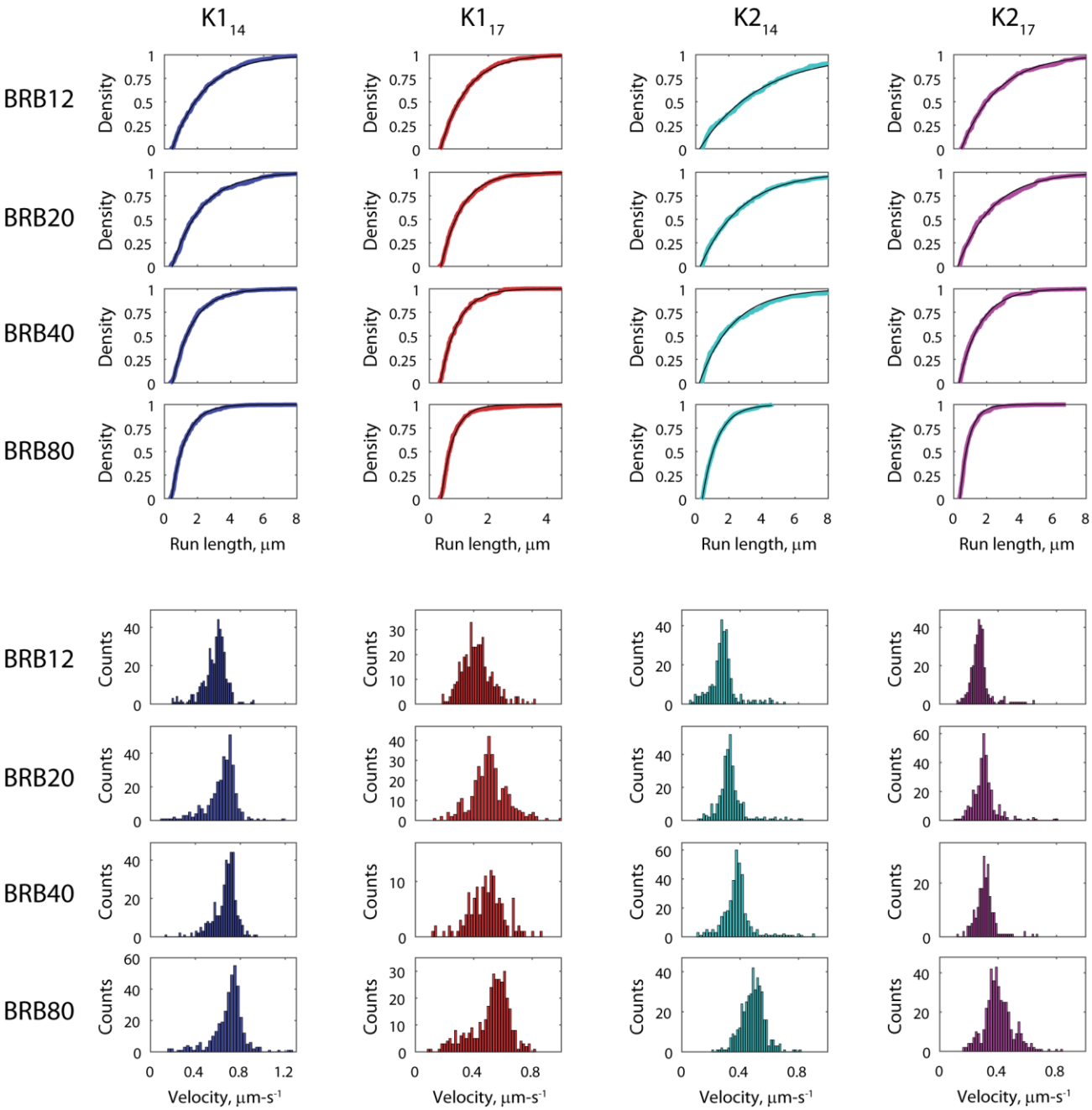

Fig. S3. Velocity and run length distributions for four motor constructs at various ionic strengths. Black lines in run length distributions show exponential plus X-offset fits. **Leftmost column** shows $K1_{14}$, with fit run lengths in μm 2.00±0.21 (N=399) 1.85±0.24 (N=322) 1.21±0.16 (N=350), and 0.86±0.12 (N=365); velocity mean values in μm-s^{-1} 0.57±0.06, 0.63±0.06, 0.67±0.07, and 0.70±0.07 for BRB12, BRB20, BRB40, and BRB80, respectively. **Second column** from left shows $K1_{17}$, with fit run lengths in μm 0.91±0.09 (N=364), 0.72±0.09 (N=350), 0.59±0.12 (N=138), and 0.40±0.04 (N=353); velocity mean values in μm-s^{-1} 0.42±0.04, 0.51±0.05, 0.49±0.05, and 0.52±0.05. **Third column** from left shows $K2_{14}$, with fit run lengths in μm 3.66±0.40 (N=303), 2.56±0.33 (N=308), 2.07±0.30 (N=360), and 0.91±0.11 (N=416); velocity mean values in μm-s^{-1} 0.29±0.03, 0.33±0.03,0.38±0.04, and 0.49±0.05. **Rightmost column** shows $K2_{17}$, with fit run lengths in μm 2.21±0.32 (N=321), 2.10±0.31 (N=334), 1.30±0.23 (N=209), and 0.55±0.07 (N=424); velocity mean values in μm-s^{-1} 0.28±0.03,

0.31±0.03, 0.31±0.03, and 0.40±0.04. Run length errors were obtained by bootstrapping (52), and velocity errors we obtained from the standard error of the mean (SEM) plus a 10% added error for 1°C temperature fluctuations.

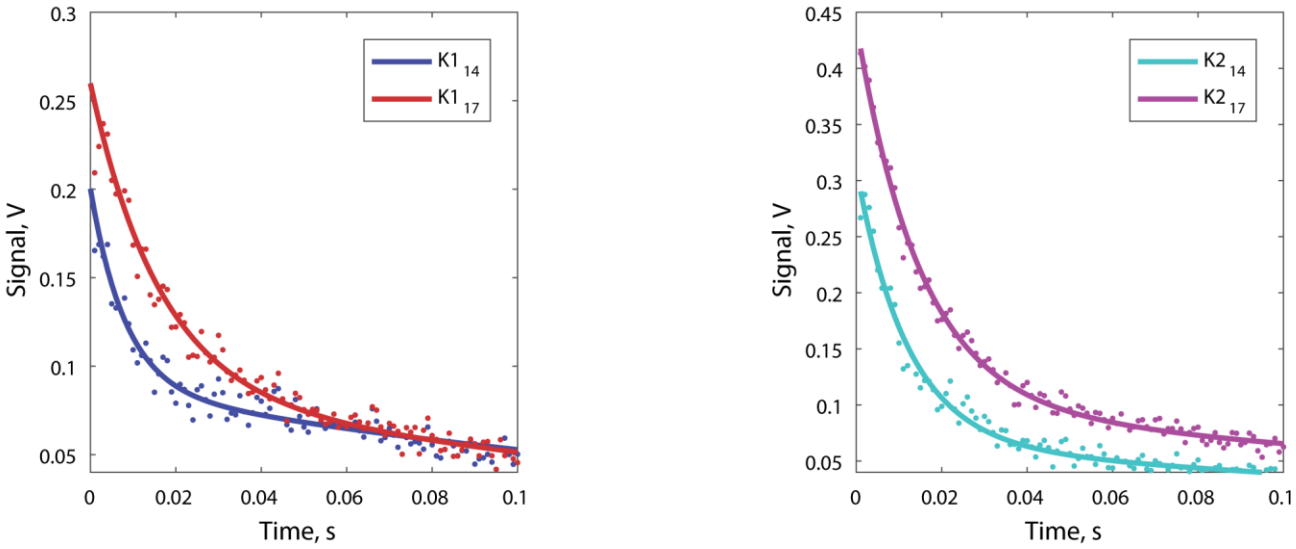

Fig. S4. Example ATP half-site transients from stopped-flow spectrofluorometer. Traces show the drop in fluorescence signal associated with release of mantNDP from the motor domain into solution. Kinesin-2 was observed to have a larger fluorescence enhancement for mant nucleotides than kinesin-1. Kinesin-1 traces were carried out with mantADP, while Kinesin-2 traces were carried with mantGDP due to the motor's high mantADP affinity.

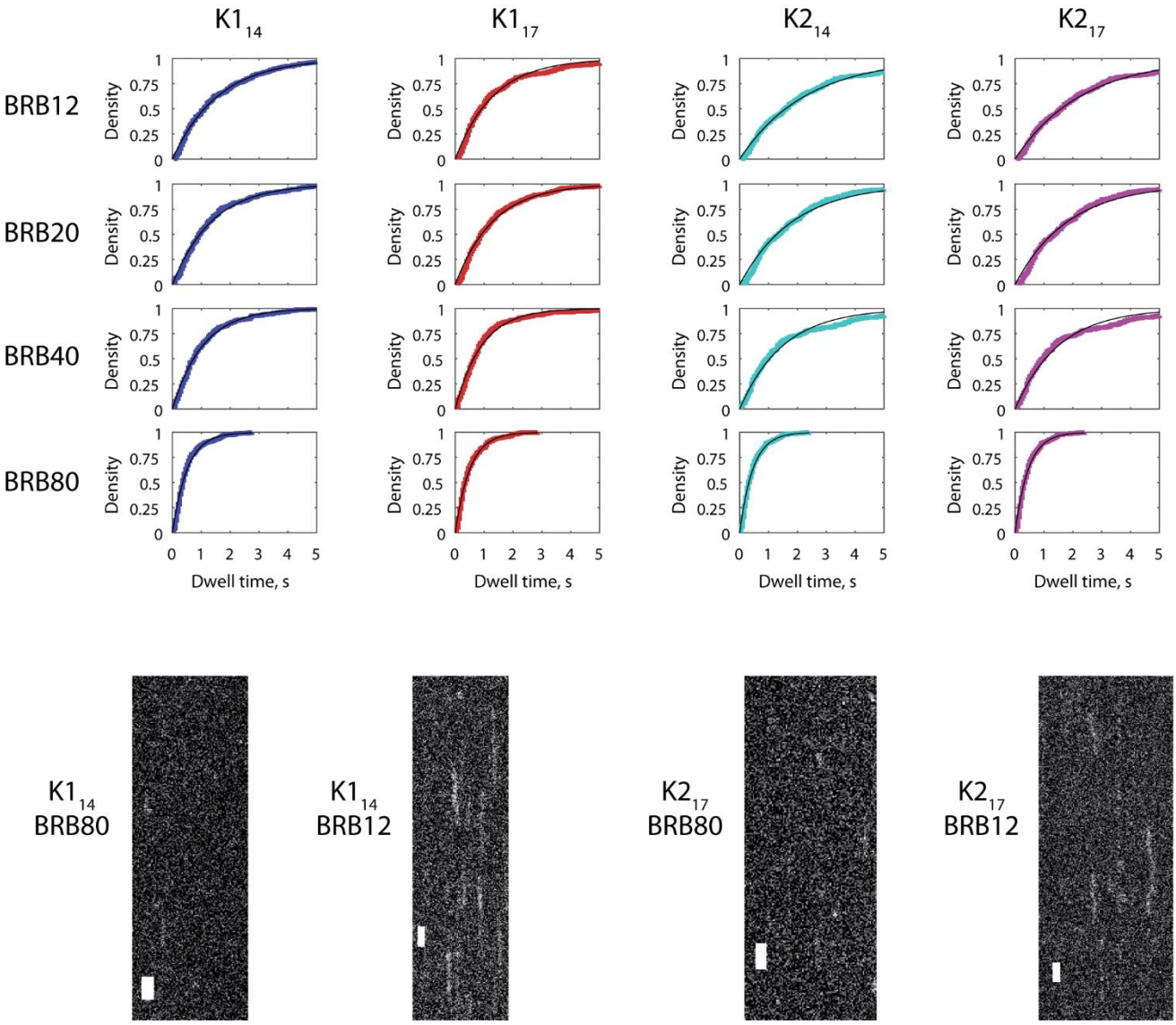

Fig. S5. Raw data for microtubule dwell times for four motors in 2 mM ADP, as a function of ionic strength. **Leftmost column** shows $K1_{14}$ with exponential fits in seconds: 1.59±0.20 (N=311), 1.37±0.18 (N=305), 1.05±0.13 (N=303), and 0.51±0.05 (N=326) for BRB12, BRB20, BRB40, and BRB80, respectively. **Second column** from left shows $K1_{17}$ with exponential fits in seconds: 1.34±0.17 (N=310), 1.32±0.15 (N=309), 0.91±0.11 (N=312), and 0.51±0.06 (N=309). **Third column** from left shows $K2_{14}$ with exponential fits in seconds: 2.30±0.31 (N=389), 1.84±0.20 (N=303), 1.52±0.24 (N=306), and 0.46±0.05 (N=303). **Rightmost column** shows $K2_{17}$ with exponential fits in seconds: 2.52±0.30 (N=314), 2.06±0.26 (N=305), 1.60±0.18 (N=319), and 0.43±0.04 (N=311). All errors were determined by bootstrapping(52). Example kymographs (20 frames per second) shown below. All scale bars are 1 second.

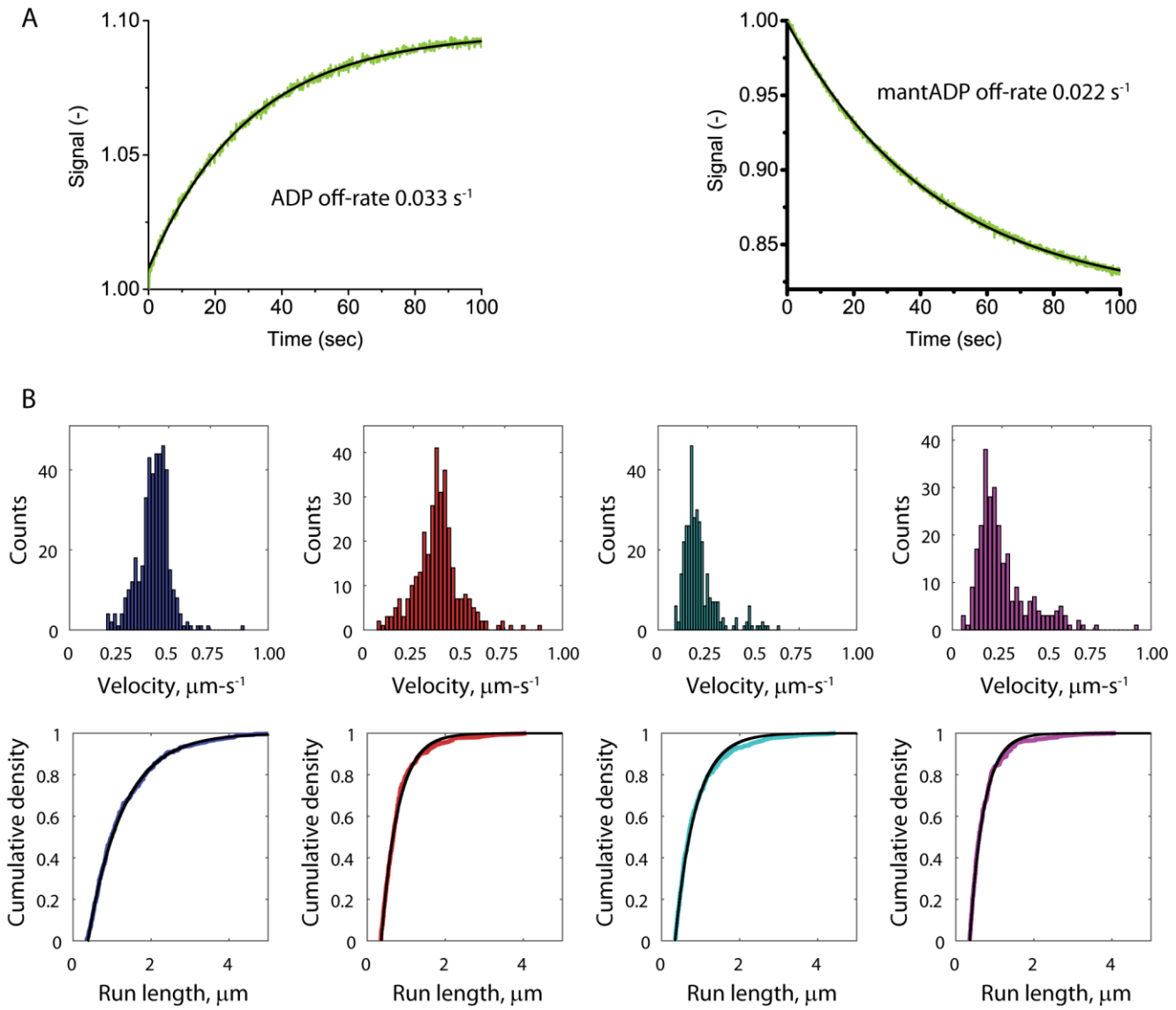

Fig. S6. Kinesin-1 solution ADP/mantADP exchange and Kinesin-1 and -2 motor performance in mantATP. (**A**) At left, kinesin-1 motors incubated in cold ADP were flushed against excess mantADP. Data were fit to a rising exponential with rate constant of 0.033 s^{-1}, corresponding to the (rate-limiting) off-rate for cold ADP. At right, motors incubated in mantADP were flushed against excess cold ADP. Data were fit by a falling exponential with a rate constant of 0.022 s^{-1}, corresponding to the off-rate for mantADP. The similarity in these off-rates demonstrates that kinesin-1 has a similar affinity for mantADP and cold ADP. (**B**) GFP motor run lengths and velocity distributions in mantATP. **Leftmost column** (blue) shows $K1_{14}$ velocity 0.42±0.04 μm-s^{-1} and run length 0.90±0.11 μm (N=425). **Second column** (red) shows $K1_{17}$ velocity 0.37±0.04 μm-s^{-1} and run length 0.52±0.05 μm (N=326). **Third column** (cyan) shows $K2_{14}$ velocity 0.21±0.02 μm-s^{-1} and run length 0.52±0.07 μm (N=325). **Rightmost column** (magenta) shows $K2_{17}$ velocity 0.25±0.03 μm-s^{-1} and run length 0.36±0.06 μm (N=268). Run length fits (black lines) were to an exponential distribution with X-offset, and errors were obtained by bootstrapping (52). Velocity errors were obtained from the SEM plus a 10% added error due to 1°C temperature fluctuations.

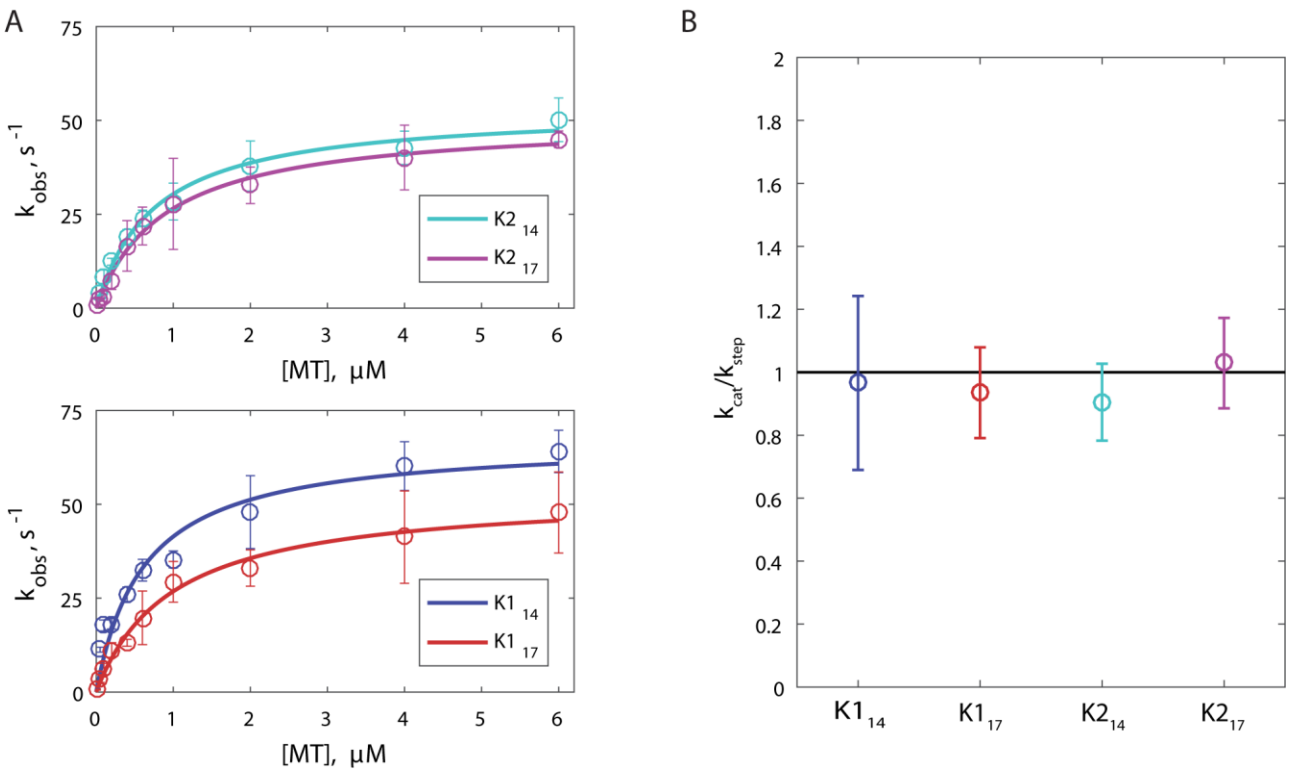

Fig. S7. Elongating the neck linker does not disrupt chemomechanical coupling. (A) Solution ATPase rates normalized to the active kinesin concentration as determined by mantADP exchange. All data points show mean plus or minus SEM for N=5 measurements. The k_{cat}/K_M values for $K1_{14}$, $K1_{17}$, $K2_{14}$, and $K2_{17}$ in $s^{-1}$and µM are 67.1±10.6/0.62±0.32, 53.1±5.8/0.98±0.32, 53.2±4.3/0.75±0.19, and 50.2±4.4/0.89±0.24, respectively. (**B**) Fit k_{cat} values from solution ATPase normalized to motor stepping rates. Stepping rates were determined by the inverse of the sum 1HB and 2HB durations reported in **Fig. 1E**. A value of 1.0 indicates that one ATP turnover event occurs per step. Thus, no motors showed evidence of futile hydrolysis cycles.

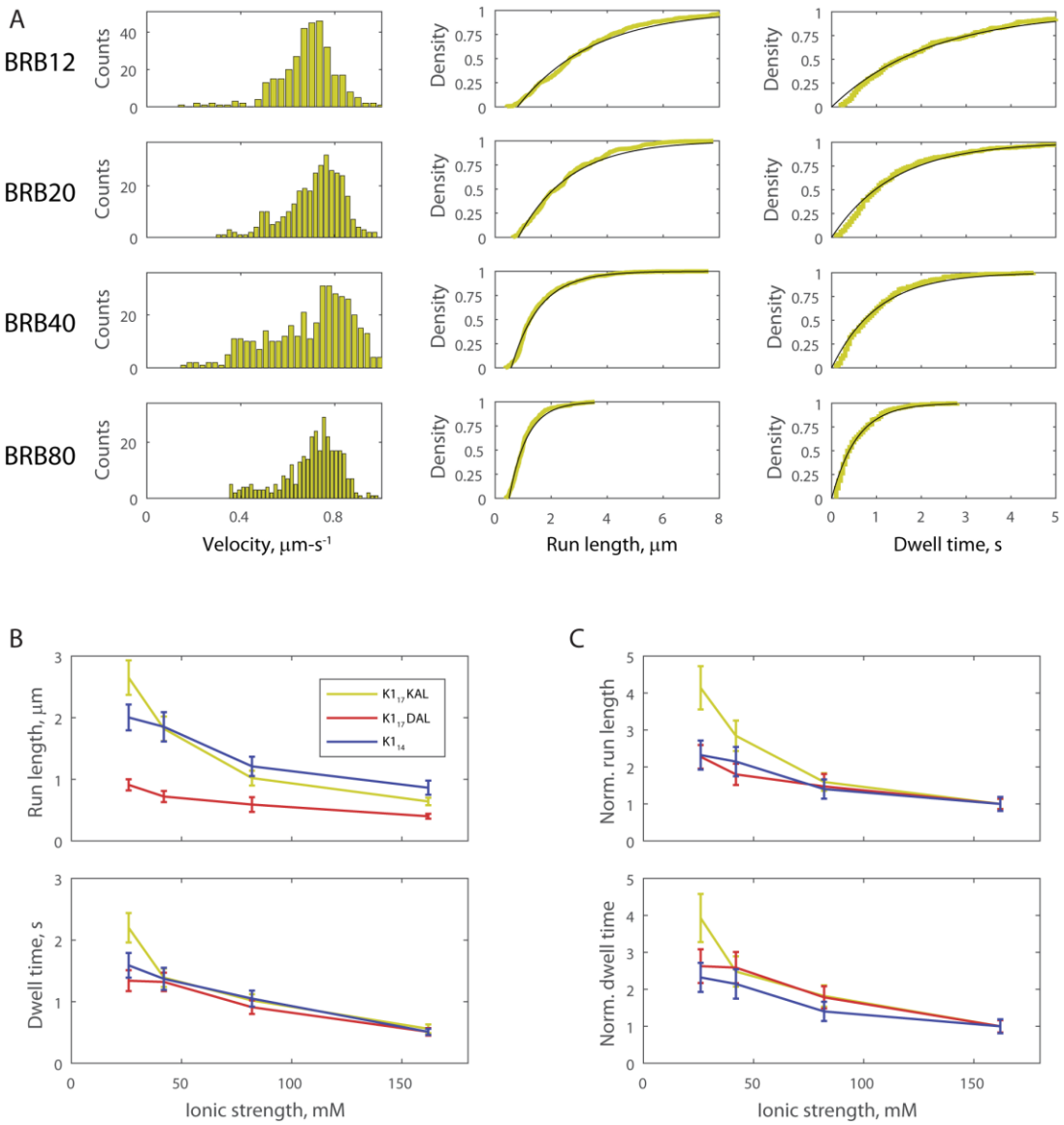

Fig. S8. Adding lysines to the neck linker changes the ionic strength dependence of run length through changes to k_{Detach}. (**A**) Raw data for kinesin-1 with an extended NL including a lysine (K_{17}KAL) at various ionic strengths. Leftmost column shows velocity distributions in 2 mM ATP, with mean values in μm-s^{-1} 0.69±0.07 (N=328), 0.72±0.07 (N=330), 0.70±0.07 (N=392), and 0.70±0.07 for BRB12, BRB20, BRB40, and BRB80, respectively. Center column shows run length distributions in 2 mM ATP, with fit values in μm 2.65±0.28, 1.82±0.20, 1.02±0.12, and 0.64±0.06. Rightmost column shows dwell time distributions in 2 mM ADP, with fit values in s 2.20±0.24, 1.39±0.15, 1.02±0.10, and 0.56±0.07. (**B**) Compiled raw data for wildtype kinesin ($K1_{14}$) versus kinesin with DAL inserted into the NL ($K1_{17}$DAL) and kinesin with KAL inserted into the NL ($K1_{17}$KAL). At high ionic strength, all three motors have an equal ADP dwell time and thus an equal k_{Detach}, but different run length values and thus different k_{Attach} values. At low ionic strength, K_{17}KAL had a much larger increase in run length and ADP dwell time than K_{17}DAL and K_{14}. This is emphasized in the normalized data (**C**): K_{17}KAL had a 4-fold increase in both values, whereas K_{17}DAL and K14 had only a 2.5-fold increase. We propose that this difference in scaling is due to the positive charge on the lysine residue. This difference in scaling means that at low ionic strength, K_{17}KAL has a similar run length to K_{14}—its k_{Attach} value is still lower, but its strongly depressed k_{Detach} value compensates.

	Kinesin-1	**Kinesin-2**
k_{Attach}^{14}, s^{-1}	329.2±291.1	218.4±176.4
k_{Attach}^{17}, s^{-1}	153.0±135.6	131.9±106.5
k_{Detach}, s^{-1}	3.13±2.75	1.96±1.57
$k_{Hydrolysis}$, s^{-1}	280.7±214.5	1,218±5,505

Table S1. Calculated rate constants used for ADP/mADP processivity model in **Fig 4A**.

SUPPORTING REFERENCES

28. Mickolajczyk, K. J., Deffenbaugh, N. C., Ortega Arroyo, J., Andrecka, J., Kukura, P., and Hancock, W.O. 2015. Kinetics of nucleotide-dependent structural transitions in the kinesin-1 hydrolysis cycle. Proc. Natl. Acad. Sci. U. S. A. 112: E7186–E7193.

52. Thorn, K.S., J.A. Ubersax, and R.D. Vale. 2000. Engineering the processive run length of the kinesin motor. J. Cell Biol. 151: 1093–1100.

62. Chen, Y., N.C. Deffenbaugh, C.T. Anderson, and W.O. Hancock. 2014. Molecular counting by photobleaching in protein complexes with many subunits: best practices and application to the cellulose synthesis complex. Mol. Biol. Cell. 25: 3630–42.

63. Viterbi, A. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Trans. Inf. Theory. 13: 260–269.