SI Text

Structural Details of Kinesin. Detailed knowledge of the structure is fundamental to understand the working mechanism of a biological nanomachine. A monomer of the conventional kinesin is structurally categorized into three parts : The head (or motor domain) (residues 2−323), the neck linker (residues 324−338 : β9, β10), and the neck-helix (residue 339− : α7). The head region is composed of eight β-strands flanked with three α-helices [α1, α2, α3 on one side (Fig. 6A), and α4, α5, α6 on the other side (Fig. 6B)] on each side of the β-sheet. One side contains the binding motif (α4, α5, α6, L8, L11, L12) for the microtubule (Fig. 6A) and the other side provides a nucleotide binding pocket (Fig. 6D). The nucleotide binding site in the kinesin head region is structurally homologous across the motor protein and the G-protein superfamilies. The structural motifs around the nucleotide binding site, such as the P-loop (N1) (86−93), switch-1 (N2) (199−204), switch-2 (N3) (232−237), N4 (14-17) are accordingly designated [1, 2]. The crystal structures with the different nucleotide states suggest that the presence or the absence of γ-P_i is sensed by these motifs and that the structural changes of the motifs are related to the allosteric transitions. The ordered state of the neck-linker, which is extended from the N terminus of kinesin head and is composed of two beta strands (β9 and β10), forms contacts with the N-terminal region of the β7-strand. Further extension from the neck-linker leads to the neck-helix (α7-helix), through which two monomers form a dimeric complex.

Computations of Residue Displacement Cross-Correlation Using Elastic Gaussian Network Model and Simulation under SB Potential. For given coordinates of a complex three-dimensional structure, the dynamical property of an object can be extracted at the zeroth order by investigating its topology. The Gaussian network model (GNM) is the simplest possible method to study large biomolecules using the corresponding minimal topology [3–5]. The GNM views the biomolecular construct as a collection of beads connected with harmonic springs with strength γ. The connectivity between the beads is purely determined by the cut-off distance parameter R_C. For a pair of beads i and j, whose distance $R_{ij}(= |\vec{R}_i - \vec{R}_j|)$ satisfies $R_{ij} < R_C$,

the harmonic potential constrains the position of beads via

$$H = \sum_{i<j} \frac{\gamma}{2}(\vec{R}_{ij} - \vec{R}^o_{ij})^2 \Theta(R_c - R_{ij}) = \frac{\gamma}{2}\delta\mathbf{R}^T \cdot \Gamma \cdot \delta\mathbf{R}, \tag{6}$$

where $\Theta(\ldots)$ is the Heaviside function, $\delta\mathbf{R}^T = (\delta\vec{R}_1, \ldots, \delta\vec{R}_N)$ with $\delta\vec{R}_i = \vec{R}_i - \vec{R}^o_i$, and $\Gamma_{ij} = \frac{1}{2}\frac{\partial^2 H}{\partial\delta R_i \partial\delta R_j}$ is the (i,j) element of Kirchhoff matrix Γ. Use of the harmonic potential amounts to the expansion of the potential $H(\{\mathrm{R}\})$ at the potential minimum $\{\mathbf{R}^o\}$ as $H(\{\mathbf{R}\}) = H(\{\mathbf{R^o}\}) + \frac{1}{2}\delta\mathbf{R}^T \frac{\partial^2 H}{\partial\delta\mathbf{R}\partial\delta\mathbf{R}}\delta\mathbf{R} + \cdots$. Since the partition function of GNM is given by $Z_N = \int D[\delta\mathbf{R}] e^{-\beta H} = \left[\det\left(\frac{\gamma\Gamma}{2\pi k_B T}\right)\right]^{-3/2}$, the correlation between the spatial fluctuation of two residues is expressed using the inverse of Kirchhoff matrix,

$$\langle \delta R_i \cdot \delta R_j \rangle = -\frac{2k_B T}{\gamma}\frac{\partial \log Z_N}{\partial \Gamma_{ij}} = \frac{3k_B T}{\gamma}\left(\Gamma^{-1}\right)_{ij}. \tag{7}$$

For $i = j$, the mean square displacement of the i^{th} residue, $\langle \delta R_i^2 \rangle$, corresponds to the B-factor (Debye-Waller temperature factor) as $B_i = \frac{8\pi^2}{3}\langle \delta R_i^2 \rangle$. A comparison between the B-factor and mean square displacement (MSD) from the GNM determines the effective strength of the harmonic potential that stabilizes the structure. Note that the quality of the MSD in GNM is solely controlled by the R_C value, thus we scaled the MSD with $3k_B T/\gamma$ for GNM analysis.

We applied the GNM analysis with $R_C = 8$ Å on the two-headed kinesin whose both heads fit to the adjacent tubulin binding site, and then computed the cross-correlation matrix as shown in Fig.7A. The cross-correlation value C_{ij}, $\langle \delta R_i \cdot \delta R_j \rangle$ scaled by $3k_B T/\gamma$, shows that except for the neck-helix region the amplitude of correlation in leading head is always larger than that of the trailing head. This is expected since the neck-linker of the leading kinesin is detached from the motor domain. The residues in the network with less coordination number are subject to a larger fluctuation. The relative difference of the cross-correlation between the leading and the trailing kinesin using

$$\delta_{ij} = \frac{C_{ij}(i,j \in L) - C_{ij}(i,j \in T)}{C_{ij}(i,j \in L)} \tag{8}$$

is illustrated in Fig. 7B. Fig. 7C shows the auto-correlations (or mean square displacement), which are the diagonal elements of the C_{ij} matrix.

GNM analysis is useful in analyzing the fluctuation dynamics of the stable structure at the residue level in the basin of attraction where the basin is modeled as a quadratic potential. However, the expansion of the potential minima up to the quadratic term is justified only if the fluctuation δR is small. The amplitude of fluctuation in biological systems at physiological

temperatures ($T \sim 310K$) is most likely to exceed the limit beyond which nonlinear response is no longer negligible. In order to take this effect into account, the Hamiltonian should be expanded beyond the linear response regime. This procedure indeed reverses the simple idea that Tirion [3] and Bahar et al. [4] have proposed in the context of GNM analysis. However, *minimal* inclusion of the nonlinear term can be useful by increasing the susceptibility of the structure. Once the nonlinear term is included, a simple analytical expression such as Eq.7 is not available. Thus, we resort to the simulations.

The analytically obtained quantities, $\langle \delta R_i \cdot \delta R_j \rangle$, δ_{ij} in Fig. 7 can also be calculated over the thermal ensemble of structures obtained from simulations using a nonlinear-Hamiltonian (see Fig. 8). The first conclusion drawn from the simulational analysis is similar to the GNM in that the leading head experiences larger fluctuations. Secondly, the position and relative amplitude of the MSD peaks, reproduced using the simulation results, shows a good agreement with GNM results. However, the direct comparison of C_{ij} (or δ_{ij}) between Fig. 7 and Fig. 8 shows that the simulation results from the nonlinear-Hamiltonian display a more sensitive pattern of cross-correlations. The pronounced amplitude of C_{ij} (or δ_{ij}) suggests a strong spatial correlation between residues i and j.

Alternative Energy Function : SOP Potential. An alternative potential function for the SB potential used in the main text is the self-organized polymer (SOP) potential that was recently adopted for simulations of the mechanical unfolding of large molecules of RNA and proteins [6, 7] as well as the allosteric dynamics of GroEL [8]. The energy Hamiltonian is

defined as

$$
\begin{aligned}
H(\{\vec{r}_i\}) &= \{H^K_{FENE} + H^K_{nb}\} + H^{K-tub}_{nb} \\
&= -\sum_{i=1}^{N_K-1} \frac{k}{2} R_0^2 \log(1 - \frac{(r_{i,i+1} - r^o_{i,i+1})^2}{R_0^2}) \\
&+ \sum_{i=1}^{N_K-3} \sum_{j=i+3}^{N_K} \epsilon_h \left[\left(\frac{r^o_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r^o_{ij}}{r_{ij}} \right)^{6} \right] \Delta_{ij} \\
&+ \sum_{i=1}^{N_K-2} \epsilon_l \left(\frac{\sigma}{r_{i,i+2}} \right)^6 + \sum_{i=1}^{N_K-3} \sum_{j=i+3}^{N_K} \epsilon_l \left(\frac{\sigma}{r_{ij}} \right)^6 (1 - \Delta_{ij}) \\
&+ \sum_{i=1}^{N_K} \sum_{k=1}^{N_{tub}} \left[\epsilon_h \left(\left(\frac{r^o_{ik}}{r_{ik}} \right)^{12} - 2 \left(\frac{r^o_{ik}}{r_{ik}} \right)^{6} \right) \Delta^*_{ik} + \epsilon_l \left(\frac{\sigma}{r_{ik}} \right)^6 (1 - \Delta^*_{ik}) \right].
\end{aligned} \tag{9}
$$

The first term is for the chain connectivity of the kinesin molecule. The finite extensible nonlinear elastic (FENE) potential [9] is used with $k = 20 kcal/(mol \cdot \text{Å}^2)$, $R_0 = 2$ Å, and $r_{i,i+1}$ is the distance between neighboring interaction centers i and $i+1$. The Lennard-Jones potential interactions stabilize the native topology. A native contact is defined as the pair of interaction centers whose distance is less than $R_C^K = 8$ Å in native state for $|i-j| > 2$. If i and j sites are in contact in the native state, $\Delta_{ij} = 1$, otherwise $\Delta_{ij} = 0$. We used $\epsilon_h = 1.8$ $kcal/mol$ in the native pairs, and $\epsilon_l = 1$ $kcal/mol$ for non-native pairs. To ensure the non-crossing of the chain, we used a 6^{th} power potential in the repulsion terms and set $\sigma = 3.8$Å, which is typical $C_\alpha - C_\alpha$ distance. The parameters determining the native topology, r^o_{ij} and Δ_{ij}, are adopted from the trailing kinesin (X) whose structure is shown in Fig. 2*C*. We transferred the topological information in the trailing head (T) to the leading head (L) by substituting r^o_{ij} and Δ_{ij} from the T to L. Kinesin-tubulin interaction energies are similarly defined as kinesin intramolecular interaction energies with slightly different native contact distances. We set the cut-off distance for the native interactions between the kinesin and the tubulin as $R_C^{K-tub} = 10$ Å. The parameters, r^o_{ik} and Δ^*_{ik}, defining the interface topology between the kinesin head T and the tubulin is transfered to the kinesin head L and the next tubulin binding site. Using the SOP potential, we obtained qualitatively identical results as those obtained from the SB potential. The nucleotide binding pocket of the front head is disrupted in the dimeric kinesin configuration whose both heads are bound to the tubulin binding sites. The figures corresponding to Fig. 3*A* and *C* and Fig. 5 are regenerated using SOP model in Fig. 9.

Master Equations for the Mechanochemical Cycle of Kinesin Described in Fig. 1. In the limit when the dissociation of dimeric kinesin from the microtubule is negligible, the kinetic equation describing the dynamic cycle shown in Fig. 1 is written as

$$\begin{aligned}
\frac{dP_{(i)}}{dt} &= -k_{bi}[ATP]P_{(i)} + k_r P_{(ii)} + k_{dMT}P_{(iv)} + k_a P_{(v)} \\
\frac{dP_{(ii)}}{dt} &= -(k_r + k_D)P_{(ii)} + k_{bi}[ATP]P_{(i)} \\
\frac{dP_{(ii')}}{dt} &= -k_{dADP}P_{(ii')} + k_D P_{(ii)} \\
\frac{dP_{(iii)}}{dt} &= -(k_h + k_{bi}^{(iii)}[ATP])P_{(iii)} + k_{dADP}P_{(ii')} + k_r^{(iii)}P_{(iii')} \\
\frac{dP_{(iv)}}{dt} &= -(k_{dMT} + k_{bi}^{(iv)}[ATP])P_{(iv)} + k_h P_{(iii)} + k_r^{(iv)}P_{(iv')} \\
\frac{dP_{(iii')}}{dt} &= -(k_r^{(iii)} + k_{diss}^{(iii)})P_{(iii')} + k_{bi}^{(iii)}[ATP]P_{(iii)} \\
\frac{dP_{(iv')}}{dt} &= -(k_r^{(iv)} + k_{diss}^{(iv)})P_{(iv')} + k_{bi}^{(iv)}[ATP]P_{(iv)} \\
\frac{dP_{(v)}}{dt} &= -k_a P_{(v)} + k_{diss}^{(iii)}P_{(iii)} + k_{diss}^{(iv)}P_{(iv)},
\end{aligned} \tag{10}$$

where $P_{(\alpha)}$ is the probability of finding the molecule in a mechanochemical state $\alpha(= i, ii, \cdots v)$ with $\sum_\alpha P_{(\alpha)} = 1$. The steady state solutions by setting $\frac{dP_{(\alpha)}}{dt} = 0$ leads to

$$\begin{aligned}
P_{(i)} &= \frac{1}{k_{bi}[ATP]}\left(1 + \frac{k_r}{k_D}\right)\frac{\mathcal{X}}{\mathcal{Z}}, \quad P_{(ii)} = \frac{1}{k_D}\frac{\mathcal{X}}{\mathcal{Z}}, \quad P_{(ii')} = \frac{1}{k_{dADP}}\frac{\mathcal{X}}{\mathcal{Z}} \\
P_{(iii)} &= \frac{\mathcal{Y}}{\mathcal{Z}}, \qquad P_{(iv)} = \frac{1}{\mathcal{Z}}, \qquad P_{(iii')} = K_m^{(iii)}\frac{\mathcal{Y}}{\mathcal{Z}}, \qquad P_{(iv')} = K_m^{(iv)}\frac{1}{\mathcal{Z}} \\
P_{(v)} &= \frac{1}{k_a}\left(k_{diss}^{(iii)}K_m^{(iii)}\mathcal{Y} + k_{diss}^{(iv)}K_m^{(iv)}\right)\frac{1}{\mathcal{Z}} \\
\mathcal{X} &\equiv k_{dMT}\left(1 + \frac{k_{diss}^{(iii)}}{k_h}K_m^{(iii)}\right)\left(1 + \frac{k_{diss}^{(iv)}}{k_{dMT}}K_m^{(iv)}\right) \\
\mathcal{Y} &\equiv \frac{k_{dMT}}{k_h}\left(1 + \frac{k_{diss}^{(iv)}}{k_{dMT}}K_m^{(iv)}\right) \\
\mathcal{Z} &\equiv \left[1 + \left(1 + \frac{k_{diss}^{(iii)}}{k_a}\right)K_m^{(iii)}\right]\mathcal{Y} + \left[1 + \left(1 + \frac{k_{diss}^{(iv)}}{k_a}\right)K_m^{(iv)}\right] \\
&\quad + \left[\frac{1}{k_{bi}[ATP]}\left(1 + \frac{k_r}{k_D}\right) + k_D^{-1} + k_{dADP}^{-1}\right]\mathcal{X} \\
K_m^{(iii)} &\equiv \frac{k_{bi}^{(iii)}[ATP]}{k_r^{(iii)} + k_{diss}^{(iii)}}, \quad K_m^{(iv)} \equiv \frac{k_{bi}^{(iv)}[ATP]}{k_r^{(iv)} + k_{diss}^{(iv)}}.
\end{aligned} \tag{11}$$

When the average velocity at steady state is computed using $v = d(k_{bi}[ATP]P_{(i)} - k_r P_{(ii)}) =$

$\mathcal{X}/\mathcal{Z}$, one can write the velocity in the form of Michaelis-Menten equation.

$$v = d\frac{\frac{k^*}{(1+\mathcal{Q}([ATP]))}[ATP]}{\frac{k^*}{1+\mathcal{Q}([ATP])}(1+\frac{k_r}{k_D})}{k_{bi}} + [ATP]} = d\frac{u_1^o[ATP]}{\frac{u_1^0+w_1^0}{k_0^0} + [ATP]} = \frac{V_{max}[ATP]}{K_M + [ATP]} \tag{12}$$

where $(k^*)^{-1} = k_D^{-1} + k_{dADP}^{-1} + k_{dMT}^{-1} + k_h^{-1}$,

$$\begin{aligned}\mathcal{Q}([ATP]) \equiv \frac{k^*}{k_h}&\left[\frac{1+\left(1+\frac{k_{diss}^{(iii)}}{k_a}\right)K_m^{(iii)}}{1+\frac{k_{diss}^{(iii)}}{k_h}K_m^{(iii)}} - 1\right] \\ &+ \frac{k^*}{k_{dMT}}\left[\frac{1+\left(1+\frac{k_{diss}^{(iv)}}{k_a}\right)K_m^{(iv)}}{\left(1+\frac{k_{diss}^{(iii)}}{k_h}K_m^{(iii)}\right)\left(1+\frac{k_{diss}^{(iv)}}{k_h}K_m^{(iv)}\right)} - 1\right],\end{aligned} \tag{13}$$

and $d = 8.2\ nm$ (the gap between neighboring tubulin binding sites). If the dissociation from the microtubule is suppressed by small $K_m^{(iii)}$ and $K_m^{(iv)}$ then $\mathcal{Q} \to 0$. Depending on the rate constant, $\mathcal{Q}$ can be either positive or negative. Although the large dissociation constant reduces the processivity, the presence of dissociation can increase the effective velocity of kinesin if $\mathcal{Q} < 0$. The second and the third expressions following the equality sign are given to compare our result with the (N=2)-model of Fisher et. al. [10] and Michaelis-Menten kinetics, respectively. If $\mathcal{Q} = 0$ the (N=2)-model analysis on the experimental data by Block and coworkers [11] predicts $k^* = u_1^o = 108s^{-1}$, $k_{bi} = k_0^0 = 1.80\mu M^{-1}s^{-1}$, and $k^*k_r/k_D = w_1^0 = 6.0s^{-1}$ ($k_D = 18 \times k_r$). This sets the lower bound for the parameters as k_D, k_{dADP}, k_{dMT}, $k_h > 108\ s^{-1}$.

The average run length of kinesin, L, is calculated using

$$L = d \times \langle l \rangle = d \times \sum_{l=1}^{\infty} l(1-P_{(v)})^l P_{(v)} = d \times \frac{1-P_{(v)}}{P_{(v)}}, \tag{14}$$

where l is the number of mechanical steps of kinesin. If the probability of the dissociated kinesin is small ($P_{(v)} \approx 0$), (i.e., if $[(ADP)_Y - (ADP)_X]_{(v)}$ is negligible in Fig. 1), then $L \approx d/P_{(v)} = v/k_{diss}$ where k_{diss} is the dissociation rate.

[1] Sack, S., Muller, J., Alexander, M., Manfred, T., Mandelkow, E. M, Brady, S. T., & Mandelkow, E. (1997) *Biochemistry* **36**, 16155 –16165.

[2] Sack, S., Kull, F. J., & Mandelkow, E. (1999) *Eur. J. Biochem.* **262**, 1–11.

[3] Tirion, M. M. (1996) *Phys. Rev. Lett.* **77**(9), 1905–1908.

[4] Haliloglu, T., Bahar, I., & Erman, B. (1997) *Phys. Rev. Lett.* **79**(16), 3090–3093.

[5] Tama, F., Valle, M., Frank, J., & Brooks III, C. L. (2003) *Proc. Natl. Acad. Sci.* **100**, 9319–9323.

[6] Hyeon, C., Dima, R. I., & Thirumalai, D. (2006) *Structure* **14**, 1633–1645.

[7] Hyeon, C. & Thirumalai, D. (2007) *Biophys. J.* **92**, 731-743.

[8] Hyeon, C., Lorimer, G. H., & Thirumalai, D. (2006) *Proc. Natl. Acad. Sci.* **103**, 18939–18944.

[9] Kremer, K. & Grest, G. S. (1990) *J. Chem. Phys.* **92**, 5057–5086.

[10] Fisher, M. E. & Kolomeisky, A. B. (2001) *Proc. Natl. Acad. Sci.* **98**, 7748–7753.

[11] Visscher, K., Schnitzer, M. J., & Block, S. M. (1999) *Science* **400**, 184–187.