Chemical-state-dependent free energy profile from single-molecule trajectories of biomolecular motor: Application to processive chitinase

Kei-ichi Okazaki[1,], Akihiko Nakamura[2,3] and Ryota Iino[2]*

[1]*Department of Theoretical and Computational Molecular Science, Institute for Molecular Science, National Institutes of Natural Sciences, Okazaki, 444-8585, Japan*

[2]*Department of Life and Coordination-Complex Molecular Science, Institute for Molecular Science, National Institutes of Natural Sciences, Okazaki, 444-8787, Japan*

[3]*Department of Applied Life Sciences, Faculty of Agriculture, Shizuoka University, Shizuoka, 422-8529, Japan*

KEYWORDS: single-molecule experiment, biomolecular motor, nonequilibrium trajectory, hidden Markov model, likelihood, Bayesian inference

ABSTRACT

The mechanism of biomolecular motors has been elucidated using single-molecule experiments for visualizing motor motion. However, it remains elusive that how changes in the chemical state during the catalytic cycle of motors lead to unidirectional motions. In this study, we use single-molecule trajectories to estimate an underlying diffusion model with chemical-state-dependent free energy profile. To consider nonequilibrium trajectories driven by the chemical energy consumed by biomolecular motors, we develop a novel framework based on a hidden Markov model, wherein switching among multiple energy profiles occurs reflecting the chemical state changes in motors. The method is tested using simulation trajectories and applied to single-molecule trajectories of processive chitinase, a linear motor that is driven by the hydrolysis energy of a single chitin chain. The chemical-state-dependent free energy profile underlying the burnt-bridge Brownian ratchet mechanism of processive chitinase is determined. The novel framework allows us to connect the chemical state changes to the unidirectional motion of biomolecular motors.

INTRODUCTION

The single-molecule experiment is a powerful tool for studying the mechanism of biomolecular motors via motion visualization. Various single-molecule techniques have been used to elucidate the mechanisms of F_1-ATPase [1–6], V_1-ATPase [7,8], myosin [9–12], kinesin [13–15] and dynein [16–19], the typical motors driven by adenosine triphosphate (ATP). Recently, single-molecule experiments have been used to identify processive cellulase and chitinase as a different type of biomolecular motor, in which the hydrolysis energy of polysaccharides is used for unidirectional motion [20–24]. These experiments have been combined with structural information and molecular dynamics simulations to elucidate how changes in the chemical state during the catalytic cycle are coupled to structural changes and mechanical motions [22,25–30]. However, experiments and simulations are not able to track the structural dynamics of active motors at atomic detail over timescale relevant for unidirectional motions. Thus, simplified mechanisms have been proposed to explain how the motors generate unidirectional motions: the power stroke and the Brownian ratchet mechanisms [31,32]. Although models based on these two mechanisms have been developed [33–41], single-molecule experimental data have not been fully exploited to develop realistic mechanisms.

Attempts have been made to estimate physicochemical models directly from single-molecule trajectories of equilibrium processes such as conformational dynamics. Single-molecule fluorescence (or Förster) resonance energy transfer (FRET) trajectories of DNA, RNA and protein conformational dynamics have been analyzed using the hidden Markov model (HMM) or the maximum likelihood method to estimate the kinetic rates and FRET efficiencies associated with the different conformational states that produce FRET signals [42–47]. FRET trajectories have also been used to estimate Langevin dynamics [48]. Single-particle tracking trajectories of a cell-surface receptor have been analyzed by HMM to estimate multiple diffusion coefficients corresponding to

different conformational states of the receptor [49]. The energy landscape is another key quantity for understanding the motion, kinetics and stability of biomolecules [50–54] that can be estimated from single-molecule trajectories [55,56]. Along the lines of these studies, we aim to develop a method that is applicable to nonequilibrium processes, such as the single-molecule trajectories of biomolecular motors.

Biomolecular motors move unidirectionally along a linear track or rotate relative to a stator by consuming chemical energy. Thus, motor motion can be effectively described by a one-dimensional (1D) model. This simplification has been employed for rotary [57,58] and linear [37] motors. A similar approach has been explored in the field of protein folding. Apparently complex processes can be projected onto a 1D reaction coordinate and modeled as a diffusion process over an energy barrier separating unfolded and folded states [45,59–61]. Methods have been developed for construction of a 1D diffusion model from experimental or simulation trajectories [62,63]. In particular, the method developed by Hummer and coworkers has been well tested with extensive applications to simulation trajectories [59,60,62,64]. However, it is challenging to apply the methods to biomolecular motors, because the unidirectional motor motions generally involve switching among different energy profiles depending on the motor chemical states.

In this study, we develop a method to construct a 1D diffusion model that involves switching among multiple energy profiles from the trajectories of biomolecular motors. In the novel framework, Hummer's method is combined with HMM to estimate the chemical-state-dependent energy profile and the diffusion coefficient. We test the method using trajectories generated from a Brownian dynamics simulation in which two energy profiles switch stochastically. Then, we apply the method to single-molecule trajectories of processive chitinase to obtain the chemical-state-dependent free energy profile triggered by the hydrolysis reaction of a single chitin chain.

METHODS

Likelihood function for diffusion model with multiple energy profiles

Here, we develop a framework for estimation of the diffusion model that involves switching among multiple energy profiles. The underlying concept of the framework is that among different hidden states, each with their own energy profiles and diffusion coefficients, one of these states is adopted at each snapshot of a trajectory. As an example, we consider a trajectory with segments that are either in state 1 or 2 (Fig. 1A, red or blue). In addition to estimating the energy profiles and diffusion coefficients, we must determine the state that is adopted at each snapshot of the trajectory. This type of problem can be solved by HMM [65]. Given a model, the likelihood of a trajectory $X = (x_0, x_1, \cdots, x_N)$ is

$$\mathcal{L}^{mstate} = \sum_{S} \left[p_{s_0}(x_1, t_1, |x_0, t_0) \prod_{i=1}^{N-1} p(s_{i-1} \to s_i) \cdot p_{s_i}(x_{i+1}, t_{i+1} | x_i, t_i) \right] \quad (1),$$

where $p_{s_i}(x_{i+1}, t_{i+1}, |x_i, t_i) \equiv p_{s_i}(bin_{i+1}, t_{i+1}, |bin_i, t_i) = p_{s_i}(bin_{i+1}, \Delta t | bin_i, 0)$ is the positional transition probability in state s_i, discretized into bins with lag time $\Delta t = t_{i+1} - t_i$, and $p(s_{i-1} \to s_i)$ is the switching probability from state s_{i-1} to s_i in time Δt. All possible paths of the hidden states $S = (s_0, s_1, \cdots, s_{N-1})$ are summed over. The positional transition probability $p_{s_i}(bin_{i+1}, \Delta t | bin_i, 0)$ was computed in the same manner as the single energy profile case developed by Hummer and coworkers [60,62,64] (see below). Direct calculation of this likelihood is difficult because the number of the possible hidden-state paths grows exponentially with N (the

number of snapshots in a trajectory). However, the likelihood can be calculated using the HMM forward-backward algorithm at a computational cost ~$O(N)$ (see SI text).

In the HMM context, the observed data are positional transitions between bins, the state transition probability corresponds to the switching probability $p(s_{i-1} \rightarrow s_i)$, and the output probability of the particular transition bin $j \rightarrow$ bin l in state s_i is proportional to $p_{s_i}(l, \Delta t | j, 0)$ (see top panel of Fig. 1B). As the output probabilities must sum to unity, the positional transition probabilities are divided by the number of bins and used as the output probabilities. Given a model (that is, an energy profile and diffusion coefficient for each state as well as switching probabilities), the likelihood $\mathcal{L}^{mstate}$ and the state probability $p_{t_i}(s) \equiv p(s_i = s | X)$ at any point in the trajectory can be computed by the HMM forward-backward algorithm. The state probability $p_t(s)$ indicates which energy profile is likely to be adopted at time t. To deal with multiple trajectories, the log of $\mathcal{L}^{mstate}$ is calculated for each trajectory and summed over to obtain $\log(\mathcal{L}_{tot}^{mstate}) = \sum_{itraj} \log(\mathcal{L}_{itraj}^{mstate})$.

Position-dependent HMM

Thus far, the switching probabilities among different energy profiles have been considered to be independent of the motor position. However, chemical steps, such as catalytic reactions in motors, have been shown to depend on the position of motors [66]. Thus, it is desirable to make the switching probabilities position-dependent [67]. Consequently, we introduce a region called the switching region. Switching can only occur when the motor is in this region (Fig. 1 and Fig. 2A). The switching probability $p(s_{i-1} \rightarrow s_i)$ in Eq. (1) can be replaced with the position-dependent switching probability $p(s_{i-1} \rightarrow s_i | x_i)$ to produce the corresponding likelihood $\mathcal{L}^{mstate-pd}$ ("pd"

stands for position-dependent). If x_i is inside the switching region, it allows the state-switching transition. Otherwise, such transition is not allowed (bottom panel of Fig. 1B).

From the HMM perspective, we effectively consider two different Markov models (or state-transition probabilities) depending on whether the motor is within or outside the switching region (bottom panel of Fig. 1B). This concept is implemented in the HMM forward-backward algorithm (see SI text). In the forward-backward algorithm, the likelihood and the state probability are calculated recurrently with one step in time in either the forward and backward direction. In this step, we introduce the position-dependent switching probabilities, as shown in the bottom panel of Fig. 1B. In this way, we obtain the likelihood $\mathcal{L}^{mstate-pd}$ and the state probability that reflects the position-dependent switching probabilities. This model is termed position-dependent HMM (pd-HMM).

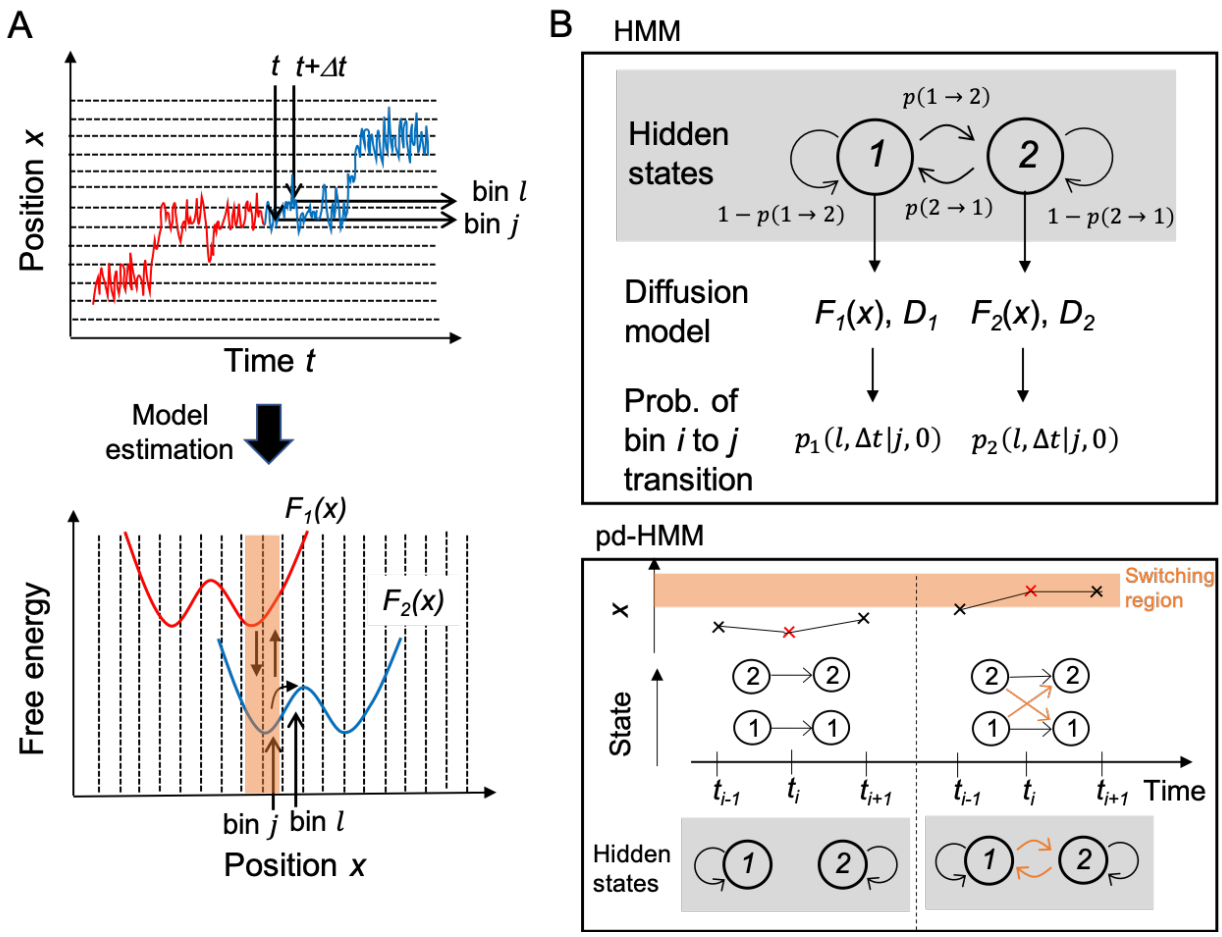


Figure 1. Estimation of chemical-state-dependent free energy profile. (A) Estimation of underlying free energy profiles and diffusion coefficient from one-dimensional trajectories, where red and blue segments of trajectory in top panel correspond to diffusion over red and blue energy profiles in bottom panel, respectively. Orange region represents the switching region. (B) Top panel shows HMM scheme to construct diffusion model with two energy profiles. Bottom panel shows position-dependent HMM scheme that is an extension of traditional HMM.

Positional transition probability for diffusion model with a single energy profile

In this section, we describe the components needed to compute the likelihood function $\mathcal{L}^{mstate}$. We relate the free energy profile and the diffusion coefficient to the positional transition probability, which was originally developed by Hummer and coworkers with extensive applications to simulation trajectories of protein folding [60,62,64]. We use a 1D model to calculate the transition probability between discretized positions (bins) from the underlying energy profile and the diffusion coefficient. The 1D Smoluchowski equation can be spatially discretized to yield the transition rate constant from bin j to bin $j \pm 1$ [68]:

$$k_{j\to j\pm 1} = \frac{D}{\Delta x^2}\exp\left(-\frac{F(j\pm 1)-F(j)}{2k_BT}\right) \qquad (2),$$

where $F(j)$ is the free energy of bin j, D is the diffusion coefficient, and Δx is the bin size. We use this expression to construct a kinetic model in terms of the free energy and the diffusion coefficient. Note that for simplicity we do not consider the diffusion coefficient to be position-dependent in this study. The master equation in matrix form is $\frac{d}{dt}\vec{p} = K\vec{p}$, where K is the rate matrix, and $\vec{p} = (p_1 \quad \cdots \quad p_n)^T$ is the state probability vector for a total number of bins n. The rate matrix K is constructed from the transition rate constants as

$$K_{ij} = \begin{cases} k_{j\to i} & i = j \pm 1 \\ -k_{j\to j+1} - k_{j\to j-1} & i = j \\ 0 & otherwise \end{cases} \qquad (3).$$

Then, the master equation can be formally solved to yield

$$\vec{p}(t) = \exp(Kt)\vec{p}(0) = T(t)\vec{p}(0) \qquad (4).$$

Each element of the transition matrix $T(t) = \exp(Kt)$ represents the positional transition probability, $(T(t))_{ij} = p(i,t|j,0)$, where $p(i,t|j,0)$ is the conditional probability of being in bin

i at time t, given a previous location in bin j at time 0. Note that the subscript specifying the hidden state in F and $p(i,t|j,0)$ is omitted here, because we consider only one state (that is, a single energy profile).

The eigenvalues of the rate matrix K are related to the timescales of the diffusion model [60]. The first non-trivial eigenvalue λ_2 represents the slowest timescale of the model, where $\lambda_1 = 0 > \lambda_2 > \lambda_3 \cdots$. The implied timescale is expressed as $\tau_i = -1/\lambda_i$, where $i \geq 2$.

Translational symmetry and prior probability of model

As biomolecular motors have a catalytic cycle, the motor energy profiles are expected to possess translational symmetry with respect to the distance moved (or step size) l_{step} after one cycle. Thus, the energy profile of a state is copied and shifted by the step size to represent the energy profile of the next state. The diffusion coefficient of the first state is also copied to the next state. With $n_{shift} = \mathrm{int}\left(l_{step}/\Delta x\right)$ and n total number of bins, $F_1(0), \cdots, F_1\left(n - n_{shift} - 1\right)$ are copied to $F_2\left(n_{shift}\right), \cdots, F_2(n-1)$ to ensure translational symmetry. Peripheral regions are defined as the parts of the energy profiles that are outside the lower/upper limit of the coordinate after shifting, that is, $F_1(j)$ with $j = n - n_{shift}, \cdots, n-1$ and $F_2(j)$ with $j = 0, \cdots, n_{shift} - 1$. The peripheral regions, denoted as $j_s \in out$ with the subscript s representing the state, are expected to be rarely visited and have high energies. Thus, we use the prior probability that the energy of the peripheral regions increases at a constant rate:

$$p(\text{model}) \propto \prod_{s=1,2;\ j_s \in out} \exp\left[-\frac{(F_s(j_s) - F_s(j_s^0) - \delta|j_s - j_s^0|)^2}{2\sigma^2}\right] \quad (5),$$

where δ is the increase in the energy per bin, j_s^0 is the last bin before entering the region with $j_1^0 = n - n_{shift} - 1$ and $j_2^0 = n_{shift}$, and σ is standard deviation that determines the strength of the restraint. Then, according to Bayes' theorem, the posterior probability $p(\text{model}|\text{data})$ is

$$p(\text{model}|\text{data}) \propto \mathcal{L}_{tot}^{mstate} \times p(\text{model}) = \mathcal{L}' \qquad (6).$$

The posterior probability $\mathcal{L}'$ is maximized during the optimization process (see below).

Maximizing posterior probability

Our objective is to optimize $F(x)$ and D as well as the switching probabilities to maximize the posterior probability $\mathcal{L}'$ with respect to the input trajectory data. For this objective, Monte Carlo (MC) simulations were performed in the parameter space. Explicitly, the model parameters are $\theta = \left(F_1(0), \cdots, F_1(n-1), F_2(0), \cdots, F_2(n-1), \ln D_1, \ln D_2, k_f, k_b\right)$. The parameters k_f and k_b are related to the switching probabilities as $p(1 \to 2) = 1 - \exp\left(-k_f \Delta t\right)$ and $p(2 \to 1) = 1 - \exp(-k_b \Delta t)$, where Δt is the lag time. In one MC step, each of these parameters is randomly changed as a trial move, $\mathcal{L}'$ of the trial move is computed, and the trial move is accepted or rejected in comparison to the current $\mathcal{L}'$. To ensure translational symmetry as described above, some parameters are copied between states. The widths of the trial moves are adjusted on the fly during the first half of the MC steps to obtain an acceptance ratio of ~0.3, which is then fixed during the second half. To compute the likelihood $\mathcal{L}^{mstate}$, the positional transition probabilities are calculated for each state with given F and D, as described above, and then fed into HMM, as described above and Fig. 1B. The Metropolis criteria is used to accept or reject the trial move, $p_{acc} = min\left(1, e^{\Delta log \mathcal{L}'}\right)$, where $\Delta log L' = log \mathcal{L}'_{trial} - log \mathcal{L}'_{current}$ [49,62]. In this way, the model parameters are guaranteed to be sampled based on the posterior probability $\mathcal{L}'$.

Generation of simulation trajectories

Based on Brownian dynamics, simulation trajectories for test analysis were generated numerically by the forward-Euler integrator,

$$x(t+\delta t) = x(t) - \frac{D}{k_B T}\frac{\partial F_{s_t}}{\partial x}\delta t + g(t) \qquad (7),$$

where F_{s_t} is the potential energy of state s_t, D is the diffusion coefficient, $\delta t = 10^{-6}$ is the integration time step, and $g(t)$ are Gaussian random numbers with mean zero and standard deviation $\sqrt{2D\delta t}$. The model potential is $F_{s_t}(x) = a(x-(s_t-1))^2(x-s_t)^2$, where $s_t = 1$ or 2 and a is a parameter controlling the barrier height. The state s_t, and thus F_{s_t}, is switched by the kinetic MC method with specified rate constants [69] in a position-dependent manner as follows. The forward ($s_t = 1 \rightarrow 2$) and backward ($s_t = 2 \rightarrow 1$) switching events only occur within the range of $0.8 \leq x \leq 1.2$ (Fig. 2A). Switching rates of $k_f = 200$ and $k_b = 1$ were used.

RMSE of energy profile

The root-mean-squared error (RMSE) of an energy profile $Y(x)$ relative to the true profile $Y_0(x)$ is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_i (Y(x_i) - Y_0(x_i))^2} \qquad (8),$$

where n is the number of bins, excluding the peripheral regions. We shifted the entire profile $Y(x_i)$ by a constant C by the least squares method before calculating the RMSE, $C = -\frac{1}{n}\sum_i (Y(x_i) - Y_0(x_i))$.

Alignment of chitinase trajectory segments

Each single-molecule trajectory of chitinase was divided into trajectory segments of two forward steps. The positional histogram of the segments was used to identify a high-frequency position for aligning the trajectory segments. A bin size of 0.15 nm was used to calculate the positional histogram for each trajectory segment. The high-frequency position was defined as a bin that was more frequent than the average frequency and the neighboring bins. The origin of the trajectory segments was set by subtracting the minimum of the identified high-frequency positions from the entire trajectory segment.

Diffusion coefficient of gold nanoparticles

The diffusion coefficient of a 40-nm gold nanoparticle was estimated using the Stokes-Einstein relation,

$$D_{AuNP} = \frac{k_B T}{6\pi\eta a} \sim 1.2 \times 10^7\ nm^2\ s^{-1} \qquad (9),$$

where $\eta = 8.9 \times 10^{-4} kg\ m^{-1}\ s^{-1}$ is the viscosity of water at 298 K, and $a = 20\ nm$ is the probe radius.

RESULTS

Reconstruction of diffusion model from simulation trajectories

We tested the developed method using simulation trajectories for which the true model parameters are known. Fig. 2A shows the two energy profiles used in the Brownian dynamics simulation with stochastic switching between these two profiles. The switching is position-dependent and only

occurs in the shaded region ($0.8 \leq x \leq 1.2$). Two out of thirty independent trajectories generated by the simulation are also shown. The double-well potential is intended to mimic a scenario wherein chitinase moves back and forth before a catalytic reaction occurs [22]. Considering the size and time constant of chitinase stepping motions, the units of length and time in Fig. 2 can be interpreted as nanometers and seconds, respectively. The trajectories were generated by Brownian dynamics simulation for diffusive motions, and kinetic MC was used to switch between the two potentials with specified rate constants (see Methods). The generated trajectories were used to reconstruct the underlying energy profiles ($F(x)$) and the diffusion coefficient (D). In the estimation, the trajectories were discretized into bins with a bin size $\Delta x = 0.1$ to produce 27 bins over the range $-0.35 \leq x \leq 2.35$, and the transitions among bins after a lag step of 400 were used as the observed data. In the Bayesian inference framework, the posterior probability $\mathcal{L}'$ was maximized in terms of $F(x)$ and D as well as the switching rates (see Methods). Note that the relative height of the two energy profiles (denoted as $\Delta\mu$ in Fig. 2A) is arbitrarily shifted in the plots. Although the relative height $\Delta\mu$ might be determined from the switching rates based on the detailed balance condition [70,71], it is controversial that if nonequilibrium dynamics of biomolecular motors should satisfy the condition [32].

The estimated energy profiles and diffusion coefficient are shown in Figs. 2B and 2C, respectively. In Fig. 2B, the estimated energy profiles averaged over the last 20% of the MC optimization steps are shown as red and blue lines, and the true profile is shown by a black line. The estimated diffusion coefficient along the MC optimization steps is shown in Fig. 2C, where the true value is shown as a broken line. Both the energy profiles and the diffusion coefficient agree quite well with the true profile (value). Furthermore, the state probabilities estimated using pd-HMM correctly reproduce the timing of the potential switching among the energy profiles (Fig.

2D). Considering that chemical steps, such as the hydrolysis reaction of chitin in chitinase, depend on the motor position, we look into position-dependency of the potential switching. The switching position was defined as the position at which the state probability of state 2 becomes larger than that of state 1, that is, $p_t(s = 2) > p_t(s = 1)$, and its histogram was plotted in Fig. 2E. All the switching positions predicted by pd-HMM were inside the switching region ($0.8 \leq x \leq 1.2$), as one would expect (Fig. 2E, top). By contrast, the switching positions predicted by HMM without the position-dependent switching resulted in a few cases wherein the predicted switching position lay outside the switching region (Fig. 2E, bottom). This problem can be attributed to the relatively large lag step of 400, because the problem was almost fixed when a shorter lag step of 20 was used (Fig. S1). However, a short lag step resulted in an overestimation of the diffusion coefficient (as we discuss below). Moreover, the switching rates were reproduced by pd-HMM but underestimated by HMM (Fig. S2).

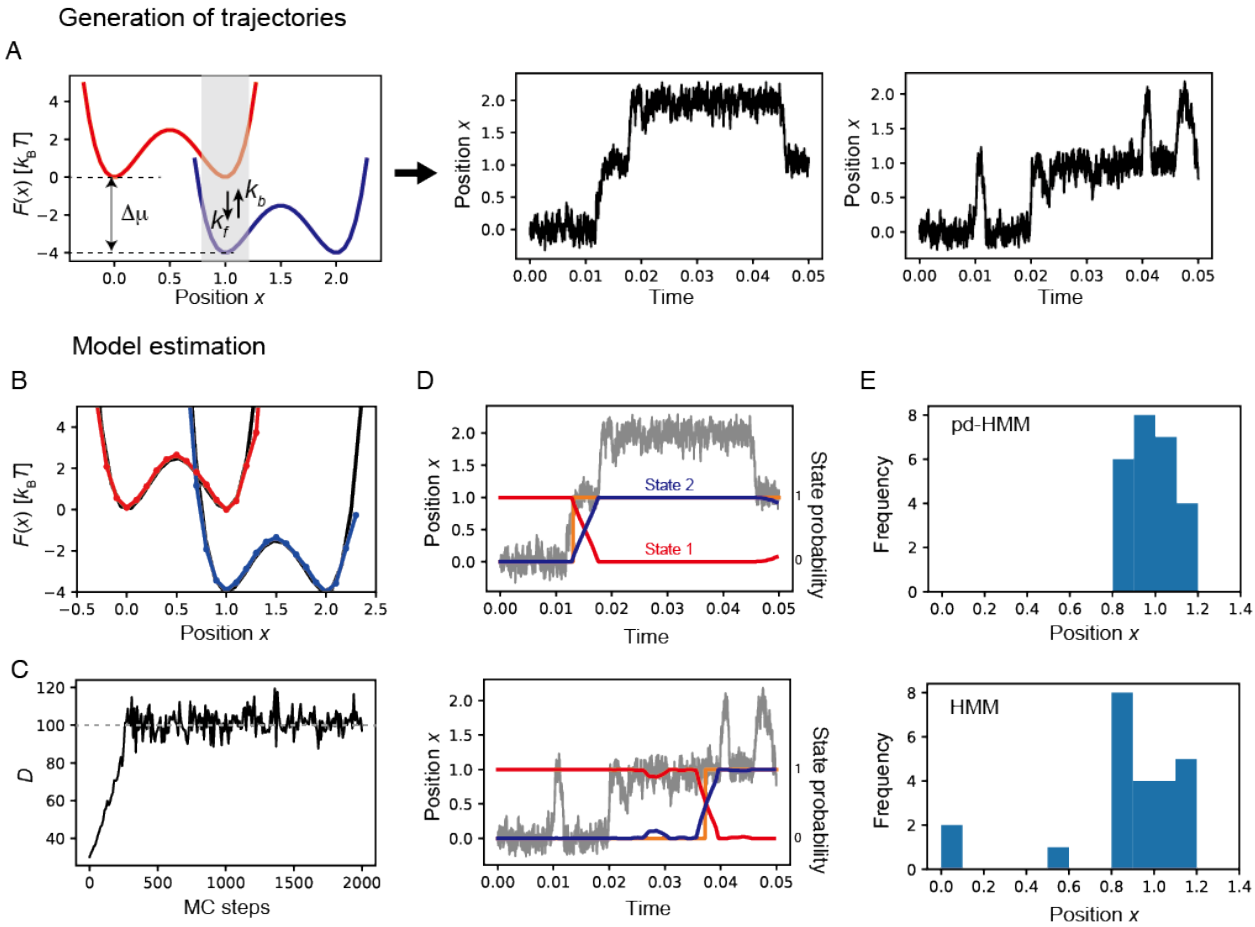


Figure 2. Generation of simulation trajectories and model estimation using the trajectories. (A) Two energy profiles used in Brownian dynamics simulation with stochastic switching (left panel) and two trajectories generated by the simulation (center and right panels). Switching only occurs in the shaded region of the left panel. (B) Estimated energy profiles shown as red and blue lines and true energy profile shown as black line. (C) Estimated diffusion coefficient during MC optimization steps (black line) and true diffusion coefficient (gray broken line). (D) pd-HMM predictions of state probabilities shown as red and blue lines and trajectories shown as gray lines. Orange lines represent actual timing of potential switching, where value 0 (1) corresponds to red (blue) potential. (E) Histograms of switching position obtained using pd-HMM (top panel) and

HMM (bottom panel). Length and time units are arbitrary and can be interpreted as nanometers and seconds, respectively, from comparison with experiments.

Then, we investigated how the accuracy of the estimated energy profiles and the diffusion coefficient depends on the lag time Δt (or lag step) used to collect the positional transition data among bins. In Fig. 3A, the RMSE of the estimated profile (value) relative to the true profile (value) are plotted for different lag steps. We found that RMSE of the energy profile takes low values until lag step of 500 and then increases, whereas the RMSE of the diffusion coefficient converges at a lag step of approximately 400. Note that the diffusion coefficient was overestimated for short lag steps. This result suggests that it is necessary to use an appropriately long lag step, although an excessive length results in insufficient transition data. To determine the optimal lag time without knowledge of the true profiles, we performed a lag-time test to see how the implied timescale converges with the lag time [60,72]. The first non-trivial eigenvalue λ_2 of the rate matrix estimated from the transition data with a prescribed lag time was converted into the so-called implied timescale using $\tau_2 = -1/\lambda_2$ (see Methods). The implied timescale along the lag step converged at a lag step of approximately 500 (Fig. 3B). We only focus on the first non-trivial mode, because there is a clear timescale gap after this mode. Low RMSEs were obtained near the convergence point of the implied timescale (Fig. 3A and B). Thus, in the absence of knowledge of the true profiles, high accuracy can be achieved by using the lag time near the convergence point of the implied timescale.

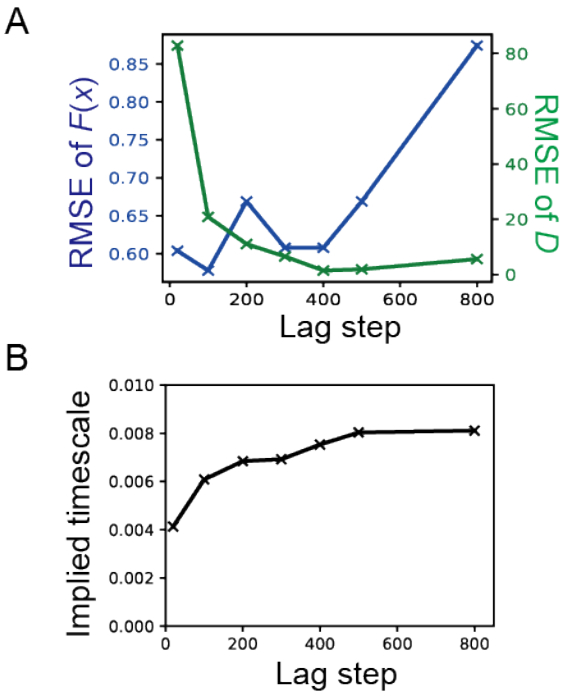


Figure 3. Effect of lag time on estimation accuracy. (A) Dependence of RMSEs of energy profile ($F(x)$) and diffusion coefficient (D) on lag step and (B) dependence of slowest implied timescale on lag step.

Effect of background noise

In the previous section, we showed that our method accurately estimates the underlying energy profiles and diffusion coefficient from simulation trajectories. However, experimental single-molecule trajectories always contain some background noise. We labeled chitinase with a 40-nm gold nanoparticle (AuNP) and observed 1-nm stepping motions using total internal reflection dark-field microscopy [22,73]: a localization precision of ~0.3 nm was obtained at a 0.333 - 0.5-ms time resolution. The localization precision is restricted by the detector saturation and can be improved [73]. However, the limited localization precision acts as background noise and should be addressed.

We added background noise to the simulation trajectories to investigate its effect on our predictions. Gaussian noise with zero mean and standard deviation of 0.3 nm was added to the x position saved every 500 steps ($\Delta t = 0.5$ ms) to simulate the experimental background noise. A median filter with various window sizes was applied to reduce the noise level (Fig. 4A). The trajectories were then analyzed by our method. To address the noise, we used a larger bin size $\Delta x = 0.2$ to produce a total of 13 bins over the range $-0.3 \leq x \leq 2.3$. Note that using this bin size in the analysis without noise accurately reproduces the energy profiles and diffusion coefficient (Fig. S3). It was prohibitively difficult to fully recover the true energy profiles and diffusion coefficient from the trajectories containing noise; however, partial recovery was achieved by adjusting the window size of the median filter. In the estimated energy profiles, the overall double-well shape with a lower energy barrier was recovered (Fig. 4B). Using trajectories median-filtered with a larger window size resulted in more accurate energy profiles. By contrast, increasing the window size produced an underestimated diffusion coefficient (Fig. 4C). Without applying the median filter, the diffusion coefficient was overestimated. Thus, this type of analysis can be used to determine lower and upper bounds on the diffusion coefficient. The timing of the potential switching was recovered reasonably well (Fig. 4D). Note that the standard deviation of 0.3 nm of the background noise is close to the step size of 1 nm, which makes an accurate estimation difficult. Irrespective of the limited estimation accuracy in the presence of background noise, we obtained the essential features of the energy profiles and a rough estimate of the diffusion coefficient.

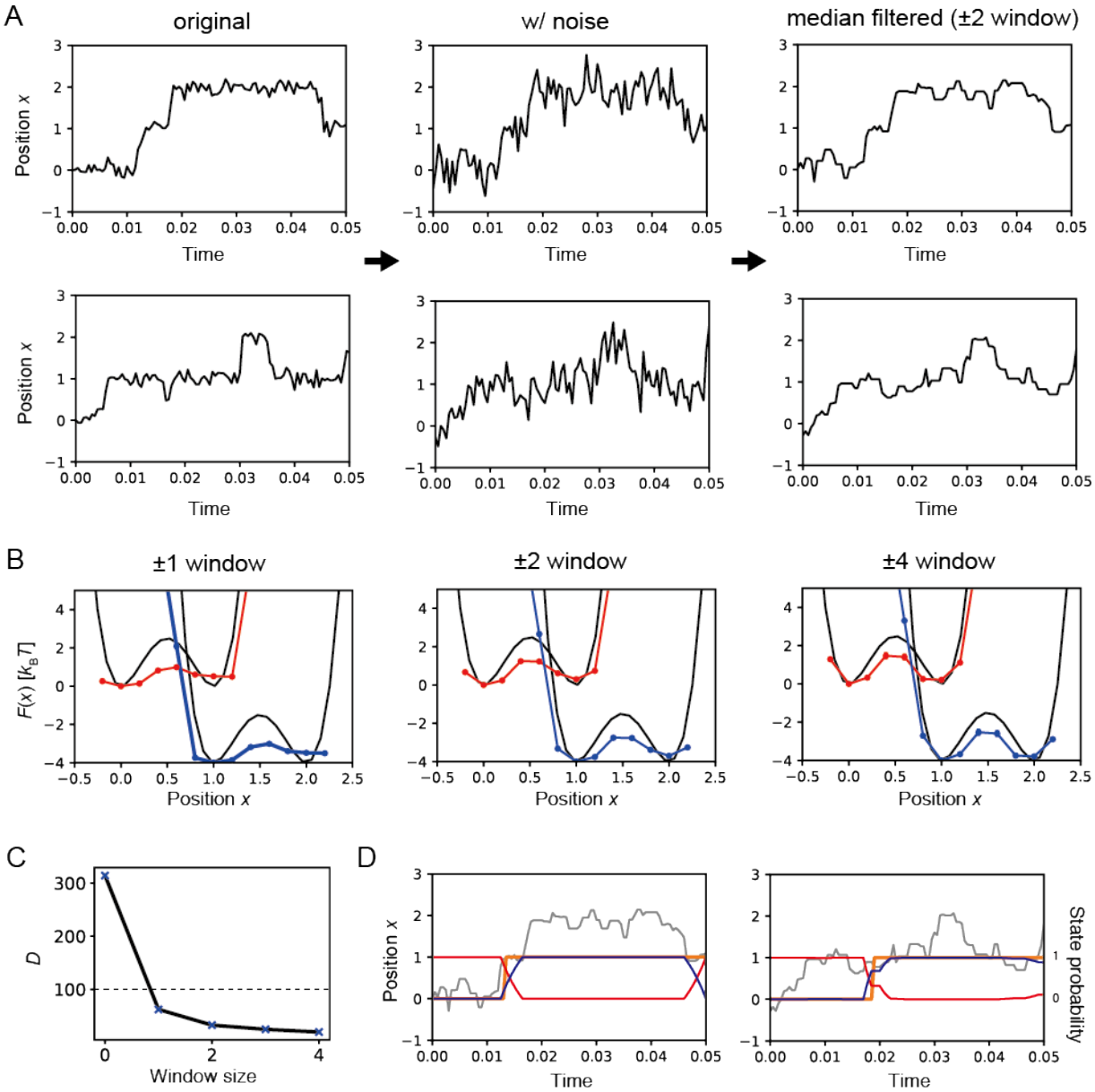


Figure 4. Effect of background noise. (A) Background noise was added to the original simulation trajectory and then reduced by a median filter. (B) Estimated energy profiles from trajectories median-filtered using window sizes of 1, 2 and 4, shown in left, center and right panels, respectively. (C) Estimated diffusion coefficients using trajectories median-filtered with different window sizes, where zero window size represents unfiltered trajectories, and dotted line represents the true value. (D) Two trajectories median-filtered with window size 2 are shown in gray lines.

Orange lines represent potential switching, where value 0 (1) corresponds to red (blue) potential. Estimated state probabilities of state 1 (red energy profile) and 2 (blue energy profile) shown by red and blue lines, respectively.

The prior probability appears to be an important factor in producing reliable estimates by the model from trajectories containing background noise. Using a weak prior probability ($\sigma = 1\ k_B T$, $\delta = 1\ k_B T$, see Methods) for trajectories containing background noise resulted in mixing of the two states, as evidenced by the state probabilities along the trajectories (Fig. S4). Consequently, the timing of switching was obscured for some trajectories. To rectify this situation, we used a stronger prior probability to analyze trajectories containing background noise ($\sigma = 0.1\ k_B T$, $\delta = 6\ k_B T$ for Fig. 4).

Application to single-molecule trajectories of processive chitinase

We applied the developed method to single-molecule trajectories of processive chitinase. Processive chitinase is a unique biomolecular motor that moves unidirectionally on a crystalline chitin surface using the hydrolysis energy of a single chitin chain [74]. We previously used single-molecule experiments to clarify the mechanochemical coupling scheme of processive chitinase (Fig. 5A) and proposed that it operates under a burnt-bridge Brownian ratchet mechanism, which is driven by a fast catalytic reaction and biased Brownian motions [22]. The burnt-bridge Brownian ratchet, has also been proposed for other motors, such as collagenase, that move processively on a collagen fibril with proteolysis energy [75,76]. Here, we aim to estimate the chemical-state-dependent free energy profile underlying the burnt-bridge Brownian ratchet mechanism of processive chitinase, directly from single-molecule trajectories.

In the single-molecule experiments, chitinase A from bacteria *Serratia marcescens* was labeled with a 40-nm gold nanoparticle (Fig. 5B), and total-internal reflection dark-field microscopy was used to observe the unidirectional motion of the particle on the chitin surface with a 0.3-nm localization precision and a 0.5-ms temporal resolution [22]. We used trajectories that were median-filtered to reduce positional noise. As the experimental trajectories were few and rather long in duration, the original trajectories were divided into segments to increase the sample size. In Fig. 5C, we show how the trajectory segments were generated. Each trajectory segment has two forward steps that were identified by the step-finding algorithm [22,77]. The segments were then aligned by matching the zero-points of the segments (see Methods).

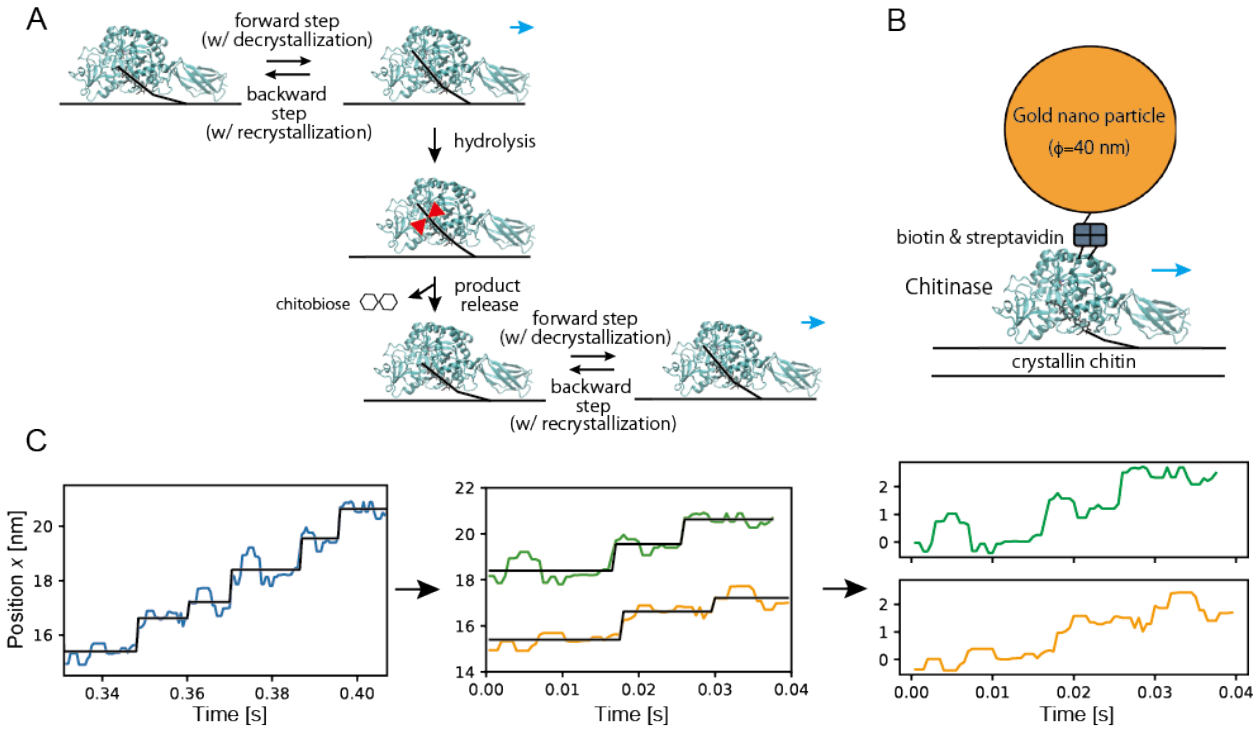


Figure 5. Single-molecule trajectories of processive chitinase. (A) Mechanochemical coupling scheme of chitinase A from bacteria *Serratia marcescens*, where cyan arrows represent direction of chitinase motion, and red saw teeth represent hydrolysis of a single chitin chain, shown as a

black line. (B) Experimental setup of chitinase labeled with gold nanoparticle on crystalline chitin. (C) Single-molecule trajectory of chitinase at 2,000 frames per second (0.5-ms temporal resolution, left panel) was divided into two trajectory segments with two forward steps (center panel). Black lines represent steps identified by step-finding algorithm [77]. Trajectory segments were shifted to start from the origin (right panels)

The trajectory segments were analyzed in the same way as the test cases above. Since the trajectory segments have two forward steps, we assume two states (free energy profiles) that correspond to before and after the chemical state change (the hydrolysis reaction and the subsequent product release). This assumption is supported by molecular insights that each forward step corresponds to the sliding of the chitin chain through the chitinase cleft, after which the catalytic reaction and the subsequent product release reset the chitin-chitinase interaction (Fig. 5A). In the calculation, a bin size $\Delta x = 0.2$ nm was used to produce 15 bins over the range of $-0.5 \leq x \leq 2.5$. To ensure translational symmetry, a shifting length of 1 nm was used, based on the chitinase step size (see Methods). The switching region $0.6 \leq x \leq 1.4$ and the forward and backward switching rates of 100 s^{-1} and 0 s^{-1}, respectively, were used. The switching region was based on the molecular insight that the hydrolysis reaction occurs at approximately $x = 1$. The forward switching rate was determined using the experimental time constants for the catalytic reaction (2.9 ms) and the subsequent product release (6.8 ms) [22]. The backward switching which is induced by the chitobiose binding and glycosidic bond formation at the catalytic site is unlikely to happen, because chitobiose concentration in solution is very low in our experimental condition. Therefore, the backward switching rate was set to zero.

The estimated diffusion model is shown in Fig. 6. The estimated free energy profiles capture the essential features of chitinase motions (Fig. 6A). Minima are located at approximately 0 and 0.6-1.0 nm for the pre-switching free energy profile and at approximately 1 and 1.6-2.0 nm for the post-switching free energy profile. These minima have similar free energies separated by a low barrier and are reflected in the chitinase stepping motions. After the chemical state change, the high-energy region $x > 1.5$ nm of the pre-switching profile decreases, resulting in a new minimum around the region of the post-switching free energy profile. The broad minimum after the barrier may explain variations in the step size. The estimated diffusion coefficient depends on the window size of the median filter (Fig. 6B). The larger the window size is, the smaller the estimated diffusion coefficient is. This result suggests a lower bound of ~200 nm^2 s^{-1} and an upper bound of ~1000 nm^2 s^{-1} on the diffusion coefficient. Performing the same analysis on trajectories at a 0.333-ms temporal resolution increases the lower bound to ~500 nm^2 s^{-1} (Fig. S5). These values are four orders of magnitude smaller than the calculated diffusion coefficient of the 40-nm gold probe D_{AuNP} (see Methods), and explain the experimental observation that the chitinase velocity is not affected by the probe size [22]. The estimated time series of the state probabilities provide further information on the predictive ability of the model for switching between the two free energy profiles (Fig. 6C). These results suggest that chitinase moves forward via diffusive motion over a barrier, which facilitates the catalytic reaction for the forward switching of the free energy profile. The free energy profiles without a large driving gradient provide further evidence of the burnt-bridge Brownian ratchet mechanism of chitinase.

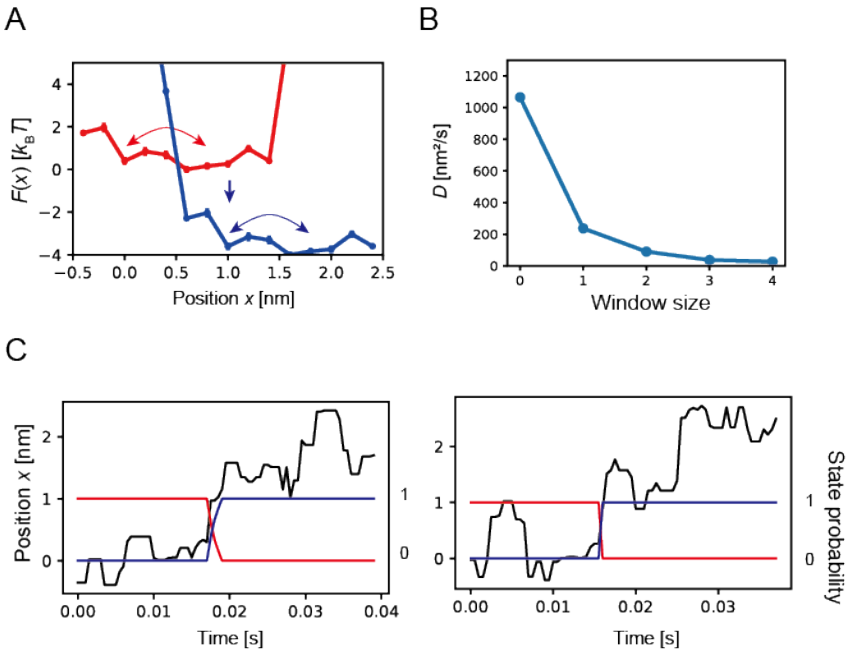


Figure 6. Estimated diffusion model of processive chitinase from trajectories at 0.5-ms temporal resolution. (A) Estimated free energy profiles before and after chemical state change are shown in red and blue, respectively, where arrows represent diffusive motion over a barrier and switching of energy profiles. (B) Estimated diffusion coefficients for different window sizes of median filter. (C) Examples of trajectory segments (black) and estimated state probability before (red) and after (blue) chemical state change are shown.

DISCUSSION

The accuracy of the estimated chemical-state-dependent free energy profile and diffusion coefficient depends on the localization precision and temporal resolution of the single-molecule trajectory data. We used single-molecule trajectories of processive chitinase that were observed with a 0.3-nm localization precision and a 0.333 – 0.5-ms temporal resolution [22]. We selected trajectory segments showing clear stepping motions for the estimation. The clear stepping motions were used to align the trajectory segments with two forward steps used in the estimation (Fig. 5C).

We smoothed background noise by using a median filter with various window sizes. The test analysis of simulation trajectories containing background noise shows that our method would underestimate the barrier height and diffusion coefficient (Fig. 4B and C). Indeed, the estimated free energy profiles have a low barrier (Fig. 6A, Fig. S5A), which is likely caused by the background noise. Measurements with a higher localization precision and temporal resolution could be used in the future to improve the quality of the predictions [73].

The friction coefficient estimated using the Einstein-Smoluchowski relation $\xi = k_B T/D$ provides an insight into the energy conversion mechanism of processive chitinase. A stepping speed $\overline{v}$ of processive chitinase can be estimated ~1000 nm s^{-1}, considering that it takes ~1 ms (two frames with 0.5-ms temporal resolution) for the 1-nm stepping transitions [22]. Note that this is a rough estimate and for an accurate estimate one would need an improved localization precision and temporal resolution. Then, the viscous drag force $\xi\overline{v}$ over a 1-nm stepping distance results in ~2 $k_B T$ of work. This work, compared to the hydrolysis energy of chitin (see below), leads to a low Stokes efficiency [78], which is clearly different from the motion of the rotary motor F_1-ATPase that shows high Stokes efficiency [79–82]. This result suggests that a considerable amount of the input hydrolysis energy of chitin is used for activities other than the mechanical movement of the motor. The important difference between the chitinase motor and typical ATPase motors is that the chitin rail shortens by one chitobiose unit during every catalytic cycle. Thus, it is essential to take the rail into consideration. Hypothetically considering the chitin rail alone, two reactions occur for every cycle: decrystallization of the chitobiose unit and the hydrolysis reaction (Fig. 5A). As decrystallization is an energy-consuming process, the net energy balance is maintained by using the input hydrolysis energy for the decrystallization of one chitobiose unit from crystalline β-chitin (the type of chitin used in the experiments). Although the free energies of hydrolysis of the β-1,4-

glycosidic bond in chitin and decrystallization per chitobiose for β-chitin are not available to our best knowledge, the approximate free energies for similar reactions are available: the free energy of hydrolysis of the β-1,4-glycosidic bond in cellulose is −16 kJ mol^{-1} ≈ $-6.5\ k_BT$ [83], and the free energy of decrystallization per chitobiose for α-chitin is 5 kcal mol^{-1}≈ $8\ k_BT$ [84]. The net energy balance of the entire chitinase-chitin system can be understood in a stepwise manner. In the catalytic cycle, starting from the state before decrystallization, processive chitinase decrystallizes the chitin chain and makes a forward step, where the chitin chain slides and attaches to the product binding site of chitinase. This binding free energy compensates the decrystallization energy, resulting in an equally stable state after the forward step [22]. Then, the hydrolysis reaction of the glycosidic bond and chitobiose release reset the system to the initial state, wherein the hydrolysis energy is liberated and used to expel the bound product. In summary, processive chitinase develops a stabilization mechanism for the high-energy intermediate formed in decrystallization, which is subsequently reset by the hydrolysis reaction.

Lastly, we mention previous methods used to estimate free energy profile(s) from nonequilibrium trajectories of the motors. Methods have been developed to estimate an effective single energy profile [31,85]. Although the dynamics generated by an effective single energy profile have been shown to reproduce some quantities, such as the mean number of steps per unit time [85], these methods have limited applicability because chemical state changes that are central to the mechanochemical coupling of motors are not considered [3,86]. Toyabe and colleagues developed a pioneering method to estimate chemical-state-dependent free energy profile [87]. Although the method developed by Toyabe *et al*. is conceptually similar to our method, the likelihood function, the positional transition probability and the optimization process differed from those used in our method. In particular, the likelihood function (the path probability) used by Toyabe *et al*. only

considers the best hidden-state path (as obtained by the Viterbi algorithm). Considering a computational cost of $\sim O(N)$ in both HMM and the Viterbi algorithm, the likelihood considering all the hidden-state paths used in our method should be more reliable. Another advantage offered by our method is that the free energy profiles and diffusion coefficient are calculated simultaneously, whereas the method of Toyabe *et al*. only optimizes free energy profiles with a fixed diffusion coefficient, which is estimated separately. In addition, our method can explicitly consider position-dependent switching rates.

CONCLUSION

A chemical-state-dependent free energy profile can explain how biomolecular motors generate unidirectional motions. In this study, we introduced a framework to estimate the chemical-state-dependent free energy profile and the diffusion coefficient from single-molecule trajectories of biomolecular motors. The method was tested using simulated trajectories and successfully reproduced the energy profiles and the diffusion coefficient. The method was then applied to single-molecule trajectories of processive chitinase. The estimated chemical-state-dependent free energy profile and diffusion coefficient provide a physical basis for the previously proposed burnt-bridge Brownian ratchet mechanism. The double-well free energy profile is shifted forward by the chemical state change to drive the unidirectional motion of chitinase (Fig. 6A). This mechanism is clearly different from the power-stroke mechanism, where a large driving gradient in the free energy profile is assumed. The estimated diffusion coefficient provides an insight into the energy conversion mechanism of processive chitinase.

ASSOCIATED CONTENT

Supporting Information

The following files are available free of charge. The SI text describes the HMM algorithm. Figures S1-S5 show the model predictions.

AUTHOR INFORMATION

Corresponding Author

*(K.O.) E-mail: keokazaki@ims.ac.jp, phone: +81-564-55-7303

Author Contributions

K.O., A.N. and R.I. conceived the study. K.O. designed the algorithm and performed the research. K.O., A.N. and R.I. discussed the results. K.O. wrote the manuscript with contributions from A.N. and R.I.

ACKNOWLEDGMENT

K.O. would like to thank Dr. Yohei Nakayama for discussions and Prof. Shoji Takada for comments on the manuscript. This study was supported by the Interdisciplinary Research Promotion Project (J281002 to K.O.) of the National Institutes of Natural Sciences (NINS), JSPS KAKENHI (JP18H02418 to R.I. and JP18H02415 to K.O.), and by Grant-in-Aid for Scientific Research on Innovative Areas "Molecular Engine" (grant number JP18H05424 to R.I.). K.O. was supported by the Building of Consortia for the Development of Human Resources in Science and Technology, MEXT, Japan.

ABBREVIATIONS

HMM, hidden Markov model; pd-HMM, position-dependent hidden Markov model; 1D model, one-dimensional model; MC, Monte Carlo; RMSE, root-mean-squared error

REFERENCES

(1) Noji, H.; Yasuda, R.; Yoshida, M.; Kinosita, K. Direct Observation of the Rotation of F1-ATPase. *Nature* **1997**, *386* (6622), 299–302. https://doi.org/10.1038/386299a0.

(2) Nishizaka, T.; Oiwa, K.; Noji, H.; Kimura, S.; Muneyuki, E.; Yoshida, M.; Kinosita, K. Chemomechanical Coupling in F1-ATPase Revealed by Simultaneous Observation of Nucleotide Kinetics and Rotation. *Nat. Struct. Mol. Biol.* **2004**, *11* (2), 142–148. https://doi.org/10.1038/nsmb721.

(3) Adachi, K.; Oiwa, K.; Nishizaka, T.; Furuike, S.; Noji, H.; Itoh, H.; Yoshida, M.; Kinosita, K. Coupling of Rotation and Catalysis in F1-ATPase Revealed by Single-Molecule Imaging and Manipulation. *Cell* **2007**, *130* (2), 309–321. https://doi.org/10.1016/j.cell.2007.05.020.

(4) Watanabe, R.; Iino, R.; Noji, H. Phosphate Release in F1-ATPase Catalytic Cycle Follows ADP Release. *Nat. Chem. Biol.* **2010**, *6* (NOVEMBER), 814–820. https://doi.org/10.1038/nchembio.443.

(5) Toyabe, S.; Watanabe-Nakayama, T.; Okamoto, T.; Kudo, S.; Muneyuki, E. Thermodynamic Efficiency and Mechanochemical Coupling of F 1-ATPase. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (44), 17951–17956. https://doi.org/10.1073/pnas.1106787108.

(6) Sugawa, M.; Okazaki, K.; Kobayashi, M.; Matsui, T.; Hummer, G.; Masaike, T.; Nishizaka, T. F 1 -ATPase Conformational Cycle from Simultaneous Single-Molecule FRET and Rotation Measurements. *Proc. Natl. Acad. Sci.* **2016**, *113* (21), E2916–E2924. https://doi.org/10.1073/pnas.1524720113.

(7) Imamura, H.; Takeda, M.; Funamoto, S.; Shimabukuro, K.; Yoshida, M.; Yokoyama, K. Rotation Scheme of V1-Motor Is Different from That of F1-Motor. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (50), 17929–17933. https://doi.org/10.1073/pnas.0507764102.

(8) Iida, T.; Minagawa, Y.; Ueno, H.; Kawai, F.; Murata, T.; Iino, R. Single-Molecule Analysis Reveals Rotational Substeps and Chemo-Mechanical Coupling Scheme of Enterococcus Hirae V1-ATPase. *J. Biol. Chem.* **2019**, *294* (45), 17017–17030. https://doi.org/10.1074/jbc.RA119.008947.

(9) Funatsu, T.; Harada, Y.; Tokunaga, M.; Saito, K.; Yanagida, T. Imaging of Single Fluorescent Molecules and Individual ATP Turnovers by Single Myosin Molecules in Aqueous Solution. *Nature* **1995**, *374* (6522), 555–559. https://doi.org/10.1038/374555a0.

(10) Ishijima, A.; Kojima, H.; Funatsu, T.; Tokunaga, M.; Higuchi, H.; Tanaka, H.; Yanagida, T. Simultaneous Observation of Individual ATPase and Mechanical Events by a Single Myosin Molecule during Interaction with Actin. *Cell* **1998**, *92* (2), 161–171. https://doi.org/10.1016/S0092-8674(00)80911-3.

(11) Yildiz, A.; Forkey, J. N.; McKinney, S. A.; Ha, T.; Goldman, Y. E.; Selvin, P. R. Myosin V Walks Hand-over-Hand: Single Fluorophore Imaging with 1.5-Nm Localization. *Science (80-.).* **2003**, *300* (5628), 2061–2065. https://doi.org/10.1126/science.1084398.

(12) Kodera, N.; Yamamoto, D.; Ishikawa, R.; Ando, T. Video Imaging of Walking Myosin v by High-Speed Atomic Force Microscopy. *Nature* **2010**, *468* (7320), 72–76. https://doi.org/10.1038/nature09450.

(13) Svoboda, K.; Schmidt, C. F.; Schnapp, B. J.; Block, S. M. Direct Observation of Kinesin Stepping by Optical Trapping Interferometry. *Nature* **1993**, *365* (6448), 721–727. https://doi.org/10.1038/365721a0.

(14) Isojima, H.; Iino, R.; Niitani, Y.; Noji, H.; Tomishige, M. Direct Observation of Intermediate States during the Stepping Motion of Kinesin-1. *Nat. Chem. Biol.* **2016**, *12* (4), 290–297. https://doi.org/10.1038/nchembio.2028.

(15) Ariga, T.; Tomishige, M.; Mizuno, D. Nonequilibrium Energetics of Molecular Motor Kinesin. *Phys. Rev. Lett.* **2018**, *121* (21), 218101. https://doi.org/10.1103/PhysRevLett.121.218101.

(16) Reck-Peterson, S. L.; Yildiz, A.; Carter, A. P.; Gennerich, A.; Zhang, N.; Vale, R. D. Single-Molecule Analysis of Dynein Processivity and Stepping Behavior. *Cell* **2006**, *126* (2), 335–348. https://doi.org/10.1016/j.cell.2006.05.046.

(17) DeWitt, M. A.; Chang, A. Y.; Combs, P. A.; Yildiz, A. Cytoplasmic Dynein Moves through Uncoordinated Stepping of the AAA+ Ring Domains. *Science (80-.).* **2012**, *335* (6065), 221–225. https://doi.org/10.1126/science.1215804.

(18) Qiu, W.; Derr, N. D.; Goodman, B. S.; Villa, E.; Wu, D.; Shih, W.; Reck-Peterson, S. L. Dynein Achieves Processive Motion Using Both Stochastic and Coordinated Stepping. *Nat. Struct. Mol. Biol.* **2012**, *19* (2), 193–201. https://doi.org/10.1038/nsmb.2205.

(19) Ando, J.; Shima, T.; Kanazawa, R.; Shimo-Kon, R.; Nakamura, A.; Yamamoto, M.; Kon, T.; Iino, R. Small Stepping Motion of Processive Dynein Revealed by Load-Free High-Speed Single-Particle Tracking. *Sci. Rep.* **2020**, *10* (1), 1–11. https://doi.org/10.1038/s41598-020-58070-y.

(20) Igarashi, K.; Uchihashi, T.; Koivula, A.; Wada, M.; Kimura, S.; Okamoto, T.; Penttilä, M.; Ando, T.; Samejima, M. Traffic Jams Reduce Hydrolytic Efficiency of Cellulase on Cellulose Surface. *Science (80-.).* **2011**, *333* (6047), 1279–1282. https://doi.org/10.1126/science.1208386.

(21) Igarashi, K.; Uchihashi, T.; Uchiyama, T.; Sugimoto, H.; Wada, M.; Suzuki, K.; Sakuda, S.; Ando, T.; Watanabe, T.; Samejima, M. Two-Way Traffic of Glycoside Hydrolase Family 18 Processive Chitinases on Crystalline Chitin. *Nat. Commun.* **2014**, *5*, 3975. https://doi.org/10.1038/ncomms4975.

(22) Nakamura, A.; Okazaki, K. ichi; Furuta, T.; Sakurai, M.; Iino, R. Processive Chitinase Is Brownian Monorail Operated by Fast Catalysis after Peeling Rail from Crystalline Chitin. *Nat. Commun.* **2018**, *9* (1), 1–12. https://doi.org/10.1038/s41467-018-06362-3.

(23) Nakamura, A.; Tasaki, T.; Okuni, Y.; Song, C.; Murata, K.; Kozai, T.; Hara, M.; Sugimoto, H.; Suzuki, K.; Watanabe, T.; et al. Rate Constants, Processivity, and Productive Binding Ratio of Chitinase A Revealed by Single-Molecule Analysis. *Phys. Chem. Chem. Phys.* **2018**, *20* (5), 3010–3018. https://doi.org/10.1039/c7cp04606e.

(24) Visootsat, A.; Nakamura, A.; Vignon, P.; Watanabe, H.; Uchihashi, T.; Iino, R. Single-Molecule Imaging Analysis Reveals the Mechanism of a High-Catalytic-Activity Mutant

of Chitinase A from Serratia Marcescens. *J. Biol. Chem.* **2020**, *295* (7), 1915–1925. https://doi.org/10.1074/jbc.RA119.012078.

(25) Ma, J.; Flynn, T. C.; Cui, Q.; Leslie, A. G. W.; Walker, J. E.; Karplus, M. A Dynamic Analysis of the Rotation Mechanism for Conformational Change in F1-ATPase. *Structure* **2002**, *10* (7), 921–931. https://doi.org/10.1016/S0969-2126(02)00789-X.

(26) Koga, N.; Takada, S. Folding-Based Molecular Simulations Reveal Mechanisms of the Rotary Motor F1-ATPase. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (14), 5367–5372. https://doi.org/10.1073/pnas.0509642103.

(27) Okazaki, K.; Hummer, G. Phosphate Release Coupled to Rotary Motion of F1-ATPase. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (41), 16468–16473. https://doi.org/10.1073/pnas.1305497110.

(28) Mukherjee, S.; Warshel, A. Electrostatic Origin of the Unidirectionality of Walking Myosin v Motors. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (43), 17326–17331. https://doi.org/10.1073/pnas.1317641110.

(29) Cecchini, M.; Alexeev, Y.; Karplus, M. Pi Release from Myosin: A Simulation Analysis of Possible Pathways. *Structure* **2010**, *18* (4), 458–470. https://doi.org/10.1016/j.str.2010.01.014.

(30) Hwang, W.; Lang, M. J.; Karplus, M. Force Generation in Kinesin Hinges on Cover-Neck Bundle Formation. *Structure* **2008**, *16* (1), 62–71. https://doi.org/10.1016/j.str.2007.11.008.

(31) Wang, H.; Oster, G. Ratchets, Power Strokes, and Molecular Motors. *Appl. Phys. A Mater. Sci. Process.* **2002**, *75* (2), 315–323. https://doi.org/10.1007/s003390201340.

(32) Hwang, W.; Karplus, M. Structural Basis for Power Stroke vs. Brownian Ratchet Mechanisms of Motor Proteins. *Proc. Natl. Acad. Sci.* **2019**, *116* (40), 19777–19785. https://doi.org/10.1073/pnas.1818589116.

(33) HUXLEY, A. F. Muscle Structure and Theories of Contraction. *Prog. Biophys. Biophys. Chem.* **1957**, *7*, 255–318.

(34) HUXLEY, A. F.; SIMMONS, R. M. Proposed Mechanism of Force Generation in Striated Muscle. *Nature* **1971**, *233* (5321), 533–538. https://doi.org/10.1038/233533a0.

(35) Vale, R. D.; Oosawa, F. Protein Motors and Maxwell's Demons: Does Mechanochemical Transduction Involve a Thermal Ratchet? *Adv. Biophys.* **1990**, *26* (C), 97–134. https://doi.org/10.1016/0065-227X(90)90009-I.

(36) Astumian, R. D.; Bier, M. Fluctuation Driven Ratchets: Molecular Motors. *Phys. Rev. Lett.* **1994**, *72* (11), 1766–1769. https://doi.org/10.1103/PhysRevLett.72.1766.

(37) Jülicher, F.; Ajdari, A.; Prost, J.; Julicher, F. Modeling Molecular Motors. *Rev. Mod. Phys.* **1997**, *69* (4), 1269–1282. https://doi.org/10.1103/RevModPhys.69.1269.

(38) Howard, J. Motor Proteins as Nanomachines: The Roles of Thermal Fluctuations in Generating Force and Motion. *Prog. Math. Phys.* **2010**, *60*, 47–59. https://doi.org/10.1007/978-3-0346-0428-4_3.

(39) Wagoner, J. A.; Dill, K. A. Molecular Motors: Power Strokes Outperform Brownian Ratchets. *J. Phys. Chem. B* **2016**, *120* (26), 6327–6336. https://doi.org/10.1021/acs.jpcb.6b02776.

(40) Takano, M.; Terada, T. P.; Sasai, M. Unidirectional Brownian Motion Observed in an in Silico Single Molecule Experiment of an Actomyosin Motor. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (17), 7769–7774. https://doi.org/10.1073/pnas.0911830107.

(41) Kanada, R.; Shinagawa, R.; Sasaki, K. Diffusion Enhancement of Brownian Motors Revealed by a Solvable Model. *Phys. Rev. E* **2018**, *98* (6), 1–11. https://doi.org/10.1103/PhysRevE.98.062110.

(42) McKinney, S. A.; Joo, C.; Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **2006**, *91* (5), 1941–1951. https://doi.org/10.1529/biophysj.106.082487.

(43) Gopich, I. V.; Szabo, A. Decoding the Pattern of Photon Colors in Single-Molecule FRET. *J. Phys. Chem. B* **2009**, *113* (31), 10965–10973. https://doi.org/10.1021/jp903671p.

(44) Keller, B. G.; Kobitski, A.; Jäschke, A.; Nienhaus, G. U.; Noé, F. Complex RNA Folding Kinetics Revealed by Single-Molecule FRET and Hidden Markov Models. *J. Am. Chem. Soc.* **2014**, *136* (12), 4534–4543. https://doi.org/10.1021/ja4098719.

(45) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. A. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science (80-.).* **2012**, *335* (6071), 981–984. https://doi.org/10.1126/science.1215768.

(46) Chung, H. S.; Eaton, W. A. Protein Folding Transition Path Times from Single Molecule FRET. *Curr. Opin. Struct. Biol.* **2018**, *48*, 30–39. https://doi.org/10.1016/j.sbi.2017.10.007.

(47) Matsunaga, Y.; Sugita, Y. Linking Time-Series of Single-Molecule Experiments with Molecular Dynamics Simulations by Machine Learning. *Elife* **2018**, 1–19. https://doi.org/10.7554/eLife.32668.

(48) Haas, K. R.; Yang, H.; Chu, J. W. Expectation-Maximization of the Potential of Mean Force and Diffusion Coefficient in Langevin Dynamics from Single Molecule FRET Data Photon by Photon. *J. Phys. Chem. B* **2013**, *117* (49), 15591–15605. https://doi.org/10.1021/jp405983d.

(49) Das, R.; Cairo, C. W.; Coombs, D. A Hidden Markov Model for Single Particle Tracks Quantifies Dynamic Interactions between LFA-1 and the Actin Cytoskeleton. *PLoS Comput. Biol.* **2009**, *5* (11). https://doi.org/10.1371/journal.pcbi.1000556.

(50) Onuchic, J. N.; Wolynes, P. G. Theory of Protein Folding. *Curr. Opin. Struct. Biol.* **2004**, *14* (1), 70–75. https://doi.org/10.1016/j.sbi.2004.01.009.

(51) Okazaki, K.; Koga, N.; Takada, S.; Onuchic, J. N.; Wolynes, P. G. Multiple-Basin Energy Landscapes for Large-Amplitude Conformational Motions of Proteins: Structure-Based Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (32), 11844–11849. https://doi.org/10.1073/pnas.0604375103.

(52) Okazaki, K.-I.; Takada, S. Dynamic Energy Landscape View of Coupled Binding and Protein Conformational Change: Induced-Fit versus Population-Shift Mechanisms. *Proc.*

Natl. Acad. Sci. U. S. A. **2008**, *105* (32), 11182–11187. https://doi.org/10.1073/pnas.0802524105.

(53) Whitford, P. C.; Onuchic, J. N.; Sanbonmatsu, K. Y. Connecting Energy Landscapes with Experimental Rates for Aminoacyl-TRNA Accommodation in the Ribosome. *J. Am. Chem. Soc.* **2010**, *132* (38), 13170–13171. https://doi.org/10.1021/ja1061399.

(54) Whitford, P. C.; Onuchic, J. N. What Protein Folding Teaches Us about Biological Function and Molecular Machines. *Curr. Opin. Struct. Biol.* **2015**, *30*, 57–62. https://doi.org/10.1016/j.sbi.2014.12.003.

(55) Baba, A.; Komatsuzaki, T. Construction of Effective Free Energy Landscape from Single-Molecule Time Series. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (49), 19297–19302. https://doi.org/10.1073/pnas.0704167104.

(56) Okuno, D.; Iino, R.; Noji, H. Stiffness of γ Subunit of F1-ATPase. *Eur. Biophys. J.* **2010**, *39* (12), 1589–1596. https://doi.org/10.1007/s00249-010-0616-9.

(57) Elston, T.; Wang, H.; Oster, G. Energy Transduction in ATP Synthase. *Nature* **1998**, *391* (6666), 510–513. https://doi.org/10.1038/35185.

(58) Wang, H.; Oster, G. Energy Transduction in the F1 Motor of ATP Synthase. *Nature* **1998**, *396* (6708), 279–282. https://doi.org/10.1038/24409.

(59) Best, R. B.; Hummer, G. Coordinate-Dependent Diffusion in Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (3), 1088–1093. https://doi.org/10.1073/pnas.0910390107.

(60) Best, R. B.; Hummer, G. Diffusion Models of Protein Folding. *Phys. Chem. Chem. Phys.* **2011**, *13* (38), 16902. https://doi.org/10.1039/c1cp21541h.

(61) Neupane, K.; Manuel, A. P.; Woodside, M. T. Protein Folding Trajectories Can Be Described Quantitatively by One-Dimensional Diffusion over Measured Energy Landscapes. *Nat. Phys.* **2016**, No. March, 10–14. https://doi.org/10.1038/nphys3677.

(62) Hummer, G. Position-Dependent Diffusion Coefficients and Free Energies from Bayesian Analysis of Equilibrium and Replica Molecular Dynamics Simulations. *New J. Phys.* **2005**, *7*, 0–14. https://doi.org/10.1088/1367-2630/7/1/034.

(63) Manuel, A. P.; Lambert, J.; Woodside, M. T. Reconstructing Folding Energy Landscapes from Splitting Probability Analysis of Single-Molecule Trajectories. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (23), 7183–7188. https://doi.org/10.1073/pnas.1419490112.

(64) Buchete, N. V.; Hummer, G. Coarse Master Equations for Peptide Folding Dynamics. *J. Phys. Chem. B* **2008**, *112* (19), 6057–6069. https://doi.org/10.1021/jp0761665.

(65) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*; Cambridge University Press, 2007.

(66) Watanabe, R.; Okuno, D.; Sakakihara, S.; Shimabukuro, K.; Iino, R.; Yoshida, M.; Noji, H. Mechanical Modulation of Catalytic Power on F 1-ATPase. *Nat. Chem. Biol.* **2012**, *8* (1), 86–92. https://doi.org/10.1038/nchembio.715.

(67) Cossio, P.; Hummer, G.; Szabo, A. Transition Paths in Single-Molecule Force Spectroscopy. *J. Chem. Phys.* **2018**, *148* (12). https://doi.org/10.1063/1.5004767.

(68) Bicout, D. J.; Szabo, A. Electron Transfer Reaction Dynamics in Non-Debye Solvents. *J Chem Phys* **1998**, *109* (6), 2325–2338. https://doi.org/10.1063/1.476800.

(69) Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **1977**, *81* (25), 2340–2361. https://doi.org/10.1021/j100540a008.

(70) Astumian, R. D. Microscopic Reversibility as the Organizing Principle of Molecular Machines. *Nat. Nanotechnol.* **2012**, *7* (11), 684–688. https://doi.org/10.1038/nnano.2012.188.

(71) Kawaguchi, K.; Sasa, S.; Sagawa, T. Nonequilibrium Dissipation-Free Transport in F1-ATPase and the Thermodynamic Role of Asymmetric Allosterism. *Biophys. J.* **2014**, *106* (11), 2450–2457. https://doi.org/10.1016/j.bpj.2014.04.034.

(72) Chodera, J. D.; Noé, F. Markov State Models of Biomolecular Conformational Dynamics. *Curr. Opin. Struct. Biol.* **2014**, *25*, 135–144. https://doi.org/10.1016/j.sbi.2014.04.002.

(73) Ando, J.; Nakamura, A.; Visootsat, A.; Yamamoto, M.; Song, C.; Murata, K.; Iino, R. Single-Nanoparticle Tracking with Angstrom Localization Precision and Microsecond Time Resolution. *Biophys. J.* **2018**, *115* (12), 2413–2427. https://doi.org/10.1016/j.bpj.2018.11.016.

(74) Beckham, G. T.; Ståhlberg, J.; Knott, B. C.; Himmel, M. E.; Crowley, M. F.; Sandgren, M.; Sørlie, M.; Payne, C. M. Towards a Molecular-Level Theory of Carbohydrate

Processivity in Glycoside Hydrolases. *Curr. Opin. Biotechnol.* **2014**, *27*, 96–106. https://doi.org/10.1016/j.copbio.2013.12.002.

(75) Saffarian, S.; Collier, I. E.; Marmer, B. L.; Elson, E. L.; Goldberg, G. Interstitial Collagenase Is a Brownian Ratchet Driven by Proteolysis of Collagen. *Science (80-.).* **2004**, *306* (5693), 108–111. https://doi.org/10.1126/science.1099179.

(76) Saffarian, S.; Qian, H.; Collier, I.; Elson, E.; Goldberg, G. Powering a Burnt Bridges Brownian Ratchet: A Model for an Extracellular Motor Driven by Proteolysis of Collagen. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **2006**, *73* (4), 1–5. https://doi.org/10.1103/PhysRevE.73.041909.

(77) Kerssemakers, J. W. J.; Laura Munteanu, E.; Laan, L.; Noetzel, T. L.; Janson, M. E.; Dogterom, M. Assembly Dynamics of Microtubules at Molecular Resolution. *Nature* **2006**, *442* (7103), 709–712. https://doi.org/10.1038/nature04928.

(78) Wang, H.; Oster, G. The Stokes Efficiency for Molecular Motors and Its Applications. *Europhys. Lett.* **2002**, *57* (1), 134–140. https://doi.org/10.1209/epl/i2002-00385-6.

(79) Yasuda, R.; Noji, H.; Kinosita, K.; Yoshida, M. F1-ATPase Is a Highly Efficient Molecular Motor That Rotates with Discrete 120°steps. *Cell* **1998**, *93* (7), 1117–1124. https://doi.org/10.1016/S0092-8674(00)81456-7.

(80) Kinosita, K.; Adachi, K.; Itoh, H. Rotation of F1-ATPase: How an ATP-Driven Molecular Machine May Work. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 245–268. https://doi.org/10.1146/annurev.biophys.33.110502.132716.

(81) Toyabe, S.; Okamoto, T.; Watanabe-Nakayama, T.; Taketani, H.; Kudo, S.; Muneyuki, E. Nonequilibrium Energetics of a Single F1-ATPase Molecule. *Phys. Rev. Lett.* **2010**, *104* (19), 1–4. https://doi.org/10.1103/PhysRevLett.104.198103.

(82) Okazaki, K.; Hummer, G. Elasticity, Friction, and Pathway of γ-Subunit Rotation in F_oF_1-ATP Synthase. *Proc. Natl. Acad. Sci.* **2015**, *2015*, 201500691. https://doi.org/10.1073/pnas.1500691112.

(83) Sørensen, T. H.; Cruys-Bagger, N.; Borch, K.; Westh, P. Free Energy Diagram for the Heterogeneous Enzymatic Hydrolysis of Glycosidic Bonds in Cellulose. *J. Biol. Chem.* **2015**, *290* (36), 22203–22211. https://doi.org/10.1074/jbc.M115.659656.

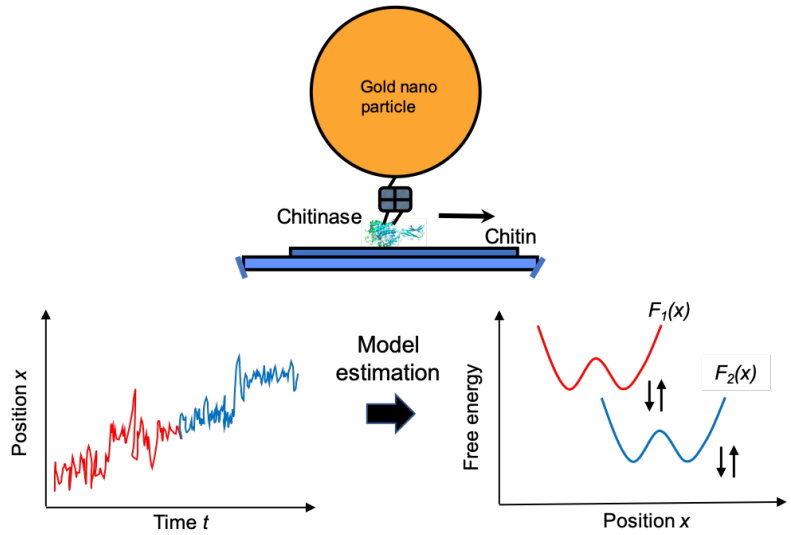
(84) Beckham, G. T.; Crowley, M. F. Examination of the ??-Chitin Structure and Decrystallization Thermodynamics at the Nanoscale. *J. Phys. Chem. B* **2011**, *115* (15), 4516–4522. https://doi.org/10.1021/jp200912q.

(85) Berezhkovskii, A. M.; Makarov, D. E. From Nonequilibrium Single-Molecules Trajectories to Underlying Dynamics. *J. Phys. Chem. Lett.* **2020**. https://doi.org/10.1021/acs.jpclett.9b03705.

(86) Astumian, R. D.; Bier, M. Mechanochemical Coupling of the Motion of Molecular Motors to ATP Hydrolysis. *Biophys. J.* **1996**, *70* (2 I), 637–653. https://doi.org/10.1016/S0006-3495(96)79605-4.

(87) Toyabe, S.; Ueno, H.; Muneyuki, E. Recovery of State-Specific Potential of Molecular Motor from Single-Molecule Trajectory. *EPL (Europhysics Lett.* **2012**, *97* (4), 40004. https://doi.org/10.1209/0295-5075/97/40004.

TOC IMAGE


Gold nano particle
Chitinase
Chitin
Position x
Time t
Model estimation
Free energy
F1(x)
F2(x)
Position x