analytical chemistry

pubs.acs.org/ac Article

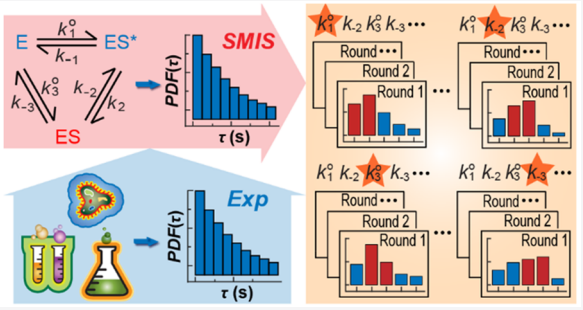
Dissection of Interaction Kinetics through Single-Molecule Interaction Simulation

Manhua Pan, Yuteng Zhang, Guangjie Yan, and Tai-Yen Chen*

Cite This: *Anal. Chem.* 2020, 92, 11582−11589 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The ability to extract kinetic interaction parameters from single-molecule fluorescence resonance energy transfer trajectories without the need for solving complex single-molecule differential equations has the potential to address some of the critical biophysical questions. Here, we provide a noise-free single-molecule interaction simulation (SMIS) tool to give the expected dwell-time distributions and relative populations of each FRET level based on the assigned kinetic model and to dissect kinetic interaction parameters from single-molecule FRET trajectories. The method provides the expected dwell-time distributions, average transition rates, and relative populations of each FRET level based on the assigned kinetic model. By comparing with ground truth data and experimental data, we demonstrated that SMIS is useful to quantify the interaction kinetic rate constants without using the traditional single-molecule analytical solution approach.



Downloaded via UNIV OF HOUSTON MAIN on February 4, 2021 at 23:44:24 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Single-molecule fluorescence resonance energy transfer (or smFRET) is a powerful biophysical technique to dissect stochastic interactions of protein−substrate,[1−3] protein−protein,[3−5] and protein−DNA complexes[6−10] as well as the folding behaviors of DNA,[11,12] RNA,[13] and proteins.[14−17] smFRET typically quantifies the number of FRET levels and the waiting times of each level. These pieces of information provide microscopic insight into the interaction kinetics. Many methodologies have been created to distill out the FRET levels and transitions reliably.[18−22] These approaches enable the experimental determination of the probability density function of waiting times $PDF(\tau)$, which is typically obtained from the histogram of microscopic dwell times. It represents the overall lifetime of species within a FRET level. Potentially, analyzing the PDF(τ) enables the quantification of rate constants,[23] which reveals the mechanism of the interaction process.

Extracting out the interaction rate constants from the PDF(τ) is typically achieved via assigning a kinetic model and fitting the $PDF(\tau)$ by the analytical probability density function $f(\tau)$. For example, Jain's laboratory studied the kinetics of the self-assembled monolayer formation on individual nanoparticles and extracted the formation rate constants of the self-assembled thiol monolayer.[24] Chen's laboratory discovered how transcription factors regulate transcription process *in vitro* and in cells.[25−27] They also provided a detailed single-molecule kinetic theory for heterogeneous catalysis.[28] Scherer's laboratory revealed the kinetics and mechanism of the physical passing of particles in an optical ring trap with an adjustable driving force.[29] Landes' laboratory reported the multistep desorption kinetics of α-lactalbumin from nylon, which provided insight into the mechanisms driving protein−polymer interactions.[30]

However, the determination of the kinetic model and the derivation of analytical solutions for kinetic parameters may not always be straightforward. In general, the process begins with the selection of the kinetic model based on the features of observed FRET levels. This kinetic model generates single-molecule rate equations. With sufficient boundary conditions, one can solve the rate equations to obtain the analytical probability density function $f(\tau)$ for transitions between different FRET levels as well as the analytical solutions of relative populations of each FRET level. Unfortunately, having sufficient boundary conditions could be difficult due to too many species coexisting in the same FRET or the kinetic model containing repeated differential equations. The difficulties getting the $f(\tau)$ and relative populations of each FRET level hinder the mechanistic study of the interaction processes. One possible alternative way to address this issue is to compare the experimental $PDF(\tau)$ with the corresponding simulated results of various kinetic models with wide-range rate constants. This approach does not require any boundary conditions and thus can broadly apply to dissect stochastic interaction kinetics. This idea has been demonstrated at the ensemble level. Software such as Berkeley Madonna has been developed for modeling and visualization of chemical

Received: March 6, 2020
Accepted: August 4, 2020
Published: August 4, 2020

© 2020 American Chemical Society

https://dx.doi.org/10.1021/acs.analchem.0c01014
Anal. Chem. 2020, 92, 11582−11589

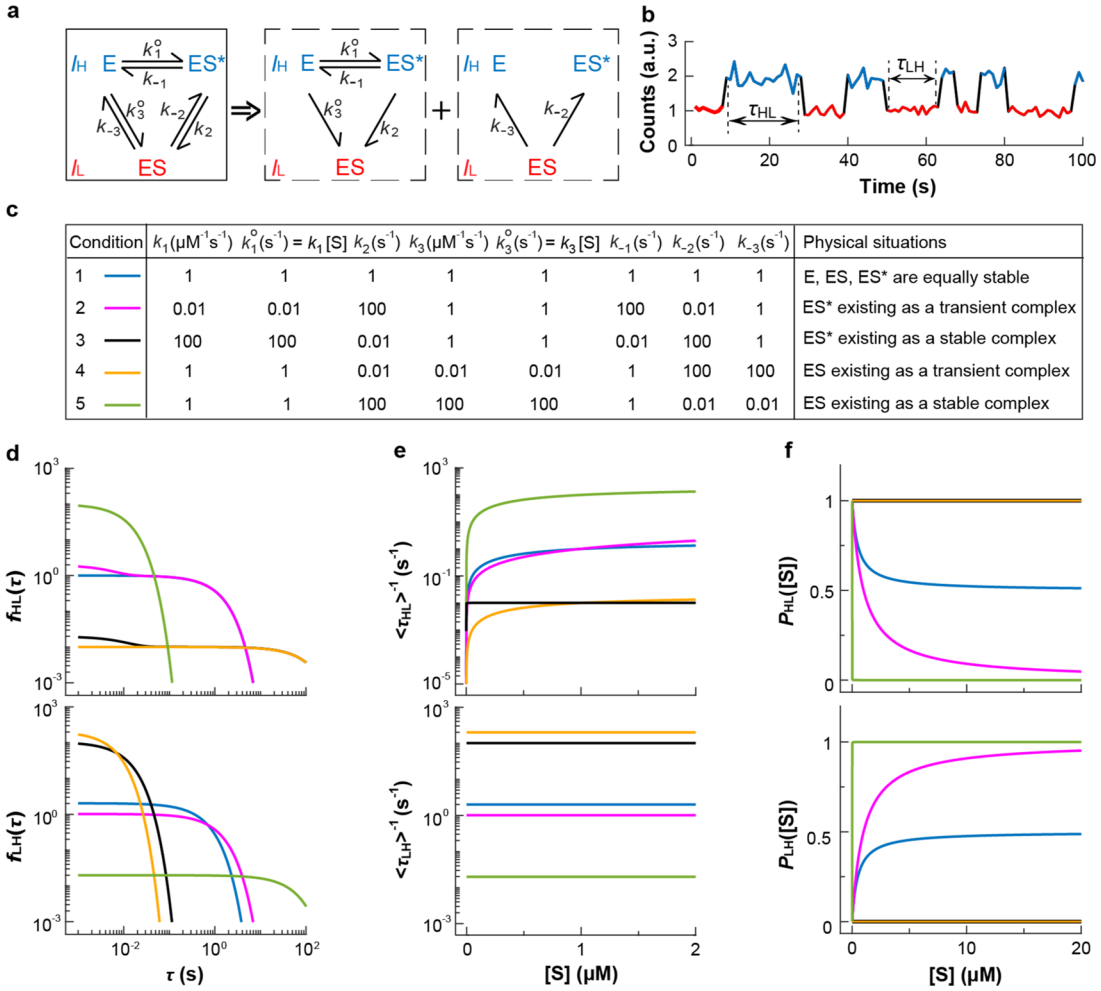


Figure 1. Kinetic scheme, exemplary single-molecule trajectory, and kinetic and thermodynamic results of a two-state FRET model. (a) Kinetic model describes three interacting species (E, ES*, and ES) associated with high (I_H) and low (I_L) states. Note that E and ES* both exist in the I_H state. (b) Single-molecule turnover trajectory shows stochastic transitions between I_H(blue curve) and I_L(red curve) states. (c) Rate constants and physical situations used to obtain $f(\tau)$, $\langle\tau\rangle^{-1}$, and $P([S])$ in parts d–f. (d–f) Analytical solutions of the (d) probability density function $f_{HL}(\tau)$ and $f_{LH}(\tau)$, (e) average transition rates $\langle\tau_{HL}\rangle^{-1}$ and $\langle\tau_{LH}\rangle^{-1}$, and (f) relative populations P_{HL} and P_{LH} under five conditions.

reactions.[31−34] However, a systematic way to simulate these kinetic parameters at the single-molecule level is still lacking.

Here, we provide a single-molecule interaction simulation (SMIS) tool to give the expected kinetic characteristics of each FRET level based on the assigned kinetic model and to dissect interaction kinetics from single-molecule FRET trajectories. In the Experimental Section, using the two-state kinetic model, we derived the analytical probability density function and the relative equilibrated populations of each FRET level. We then introduced SMIS and provided a step-by-step simulation of the single-molecule interaction FRET events, the probability density of dwell times, average transition rates, and the relative population of each FRET level. In the Results and Discussion, the general solutions from the two-level kinetic model were then converted into the analytical solutions of the well-known Michaelis−Menten enzyme kinetic model to validate the SMIS-generated simulations. Accuracy of SMIS-extracted rate constants was examined by using kinetic models with two and three FRET levels. Finally, we demonstrated a successful application of SMIS to simulate and extract kinetic parameters of interactions between a copper transcription regulator, CueR, and DNA.

EXPERIMENTAL SECTION

We chose a system composed of four interacting species: substrate S, enzyme E, loosely bounded intermediate ES*, and final complex ES. The E is labeled with a FRET donor, and S is tagged with the FRET acceptor. In the single-molecule FRET experiment under donor excitation, the highly fluorescent E interacts with S, forms a fluorescent intermediate ES*, and generates a weakly fluorescent product ES. Note that ES* represents a condition that S is loosely bound to E with a distance too far to show distinguishable FRET. This kinetic model assigns three interacting species E, ES*, and ES to two intensity states under dynamic equilibrium conditions (Figure 1a). We demonstrated the derivation of dwell-time distributions, average transition rates, and relative populations of each interacting species through the analytical or the simulation approach. We developed a single-molecule interaction simulation (SIMS) tool to simulate the smFRET trajectories and kinetic parameters. Since many methods have been developed to remove the noise in the smFRET trajectory,[35−37] all simulations were created without noise to reduce the complexity.

Derive Analytical Kinetic Properties of the Kinetic Model. Figure 1a shows the kinetic model describing an enzyme existing as one of the three interacting species (E, ES*, and ES)

https://dx.doi.org/10.1021/acs.analchem.0c01014
Anal. Chem. 2020, 92, 11582−11589

associated with the high (I_H) and low (I_L) fluorescence states with the forward (k_1, k_2, and k_3) and reversed (k_{-1}, k_{-2}, and k_{-3}) rate constants annotated. Figure 1b shows the typical single-molecule trajectory that reflects these processes through the stochastic on−off burst like signals. Each fluorescence intensity increase marks the presence of E or ES*; each decrease marks a formation of ES. The τ_{HL}, the dwell time on the I_H state before transitioning to the I_L state, reports the microscopic dwell times for completing steps that involve k_1° (i.e., $k_1[S]$), k_{-1}, k_2, and k_3° (i.e., $k_3[S]$). τ_{LH}, the dwell time on the I_L state before transitioning to the I_H state, reports the microscopic dwell times for completing steps that involve k_{-2} and k_{-3}. The distributions of these two stochastic observables (τ_{HL} and τ_{LH}) provide valuable kinetic information and molecular insights to the interaction mechanism.

To obtain kinetic rate constants analytically, we derived the probability density function of dwell times $f(\tau)$ by solving the single-molecule rate equations under proper initial conditions. For example, to acquire the probability density function of τ_{HL} $f_{HL}(\tau)$, we wrote out the single-molecule kinetic equations based on the processes (Figure 1a middle) that lead to the transition from the I_H to I_L state. Entering the I_L state can occur either through the E to ES (pathway involves k_3°) or ES* to ES (pathway involves k_2) and relevant to $P_E(t)_1$, $P_{ES^*}(t)_1$, $P_{ES}(t)_1$, $P_E(t)_2$, $P_{ES^*}(t)_2$, and $P_{ES}(t)_2$. The overall probability function of time is a linear combination of both pathways $P_i(t) = C_1P_i(t)_1 + C_2P_i(t)_2$ ($i \in$ [E, ES*, ES]), which C_1 and C_2 are the probability coefficients for two different initial conditions (Supporting Information, section 1.1). We evaluated the probability density function $f_{HL}(\tau)$ from the $P_i(t)$. $f_{HL}(\tau)$ dictates the probability density of finding a dwell time τ_{HL} for the transition from the I_H to the I_L state. The probability for finding a particular dwell time τ_{HL} is $f_{HL}(\tau)\Delta\tau$, which equals the possibility of switching from I_H to I_L between $t = \tau$ and $\tau + \Delta\tau$. Since the transitions from the I_H to the I_L state only occur via E to ES or ES* to ES pathways, $f_{HL}(\tau)\Delta\tau$ can be estimated from the overall $\Delta P_{ES}(\tau)$. In the limit of infinitesimal $\Delta\tau$, the analytical expression of overall probability density function $f_{HL}(\tau)$ is $f_{HL}(\tau) = C_1\frac{dP_{ES}(\tau)_1}{d\tau} + C_2\frac{dP_{ES}(\tau)_2}{d\tau}$ (Supporting Information, section 1.2). Equation 1 shows the final analytical expression of $f_{HL}(\tau)$.

$$f_{HL}(\tau) = D_1\, e^{(B+A)\tau} + D_2\, e^{(B-A)\tau} \tag{1}$$

where

$$A = \frac{\sqrt{(k_{-1} + k_3^\circ + k_2 + k_1^\circ)^2 - 4(k_3^\circ k_{-1} + k_2(k_1^\circ + k_3^\circ))}}{2}$$

$$B = \frac{-(k_3^\circ + k_2 + k_1^\circ + k_{-1})}{2}$$

$$D_1 = \frac{(k_2(A + B + k_1^\circ + k_3^\circ) + k_3^\circ k_{-1})k_{-2} + (k_1^\circ k_2 + k_3^\circ(A + B + k_2 + k_{-1}))k_{-3}}{2A(k_{-2} + k_{-3})}$$

$$D_2 = \frac{(k_2(A + B + k_2 + k_{-1}) - k_3^\circ k_{-1})k_{-2} + (-k_1^\circ k_2 + k_3^\circ(A + B + k_1^\circ + k_3^\circ))k_{-3}}{2A(k_{-2} + k_{-3})}$$

Finally, with the $f_{HL}(\tau)$, we evaluated the average transition time $\langle\tau_{HL}\rangle$ by $\langle\tau_{HL}\rangle = \int_0^\infty \tau f_{HL}(\tau)\, d\tau$, whose reciprocal value reports the average transition rate (eq 2).

$$\langle\tau_{HL}\rangle^{-1} = \frac{(k_{-2} + k_{-3})(k_1k_2 + k_3(k_2 + k_{-1}))[S]}{k_{-3}(k_2 + k_{-1} + k_1[S]) + k_{-2}(k_{-1} + (k_1 + k_3)[S])} \tag{2}$$

Similarly, we evaluated the $P_{LH}(t)$, $f_{LH}(\tau)$, and $\langle\tau_{LH}\rangle^{-1}$ for the transitions from the I_L to I_H state (Supporting Information, section 1.2). Equation 3 and eq 4 summarize the final analytical expression of $f_{LH}(\tau)$ and $\langle\tau_{LH}\rangle^{-1}$, respectively.

$$f_{LH}(\tau) = (k_{-2} + k_{-3})\, e^{-(k_{-2}+k_{-3})\tau} \tag{3}$$

$$\langle\tau_{LH}\rangle^{-1} = \left(\int_0^\infty \tau f_{LH}(\tau)\, d\tau\right)^{-1} = k_{-2} + k_{-3} \tag{4}$$

With these analytical solutions, we can predict the $f_{HL}(\tau)$ and $f_{LH}(\tau)$ and test the effect of substrate concentration, [S], on the $\langle\tau_{HL}\rangle^{-1}$ and $\langle\tau_{LH}\rangle^{-1}$. Figure 1d−f shows these predictions for the below conditions (Figure 1c): (1) all rate constants are similar; (2) ES* exists as a transient complex; (3) ES* existing as a stable complex; (4) ES exists as a transient complex; and (5) ES existing as a stable complex. Since most of the rate constants range from 10^{-2} to 10^2, we used rate constants of 1 s^{-1} to predict the kinetic properties. We then varied the rate constants to simulate ES* and ES existing as the transient or stable complexes. For the stable complex conditions (i.e., formation rate constant is greater than dissociation rate constant), we set the formation rate constants to $\sim 10^2$ s^{-1} and the dissociation rate constants to $\sim 10^{-2}$ s^{-1}. On the other hand, formation rate constants are $\sim 10^{-2}$ s^{-1} and the dissociation rate constants to $\sim 10^2$ s^{-1} for the transient complex condition. The top panels of Figure 1d,e show the $f_{HL}(\tau)$ and $\langle\tau_{HL}\rangle^{-1}$ for these conditions. The $f_{HL}(\tau)$ predicts the multiple exponential binding behaviors, and the $\langle\tau_{HL}\rangle^{-1}$ predicts the hyperbolic dependence of [S] on the ES formation rate. On the other hand, the bottom panels of Figure 1d,e show the $f_{LH}(\tau)$ and $\langle\tau_{LH}\rangle^{-1}$. Since the transitions from the I_L to I_H state represent the dissociation of ES, the log−log plot of $f_{LH}(\tau)$ predicts that the dissociation of ES follows the unimolecular dissociation mechanism.

Thermodynamic parameters such as the relative subpopulation of each species were also derived. $P_E([S])$, $P_{ES}([S])$, and $P_{ES^*}([S])$ are the relative subpopulation of E, ES*, and ES, respectively. Using the two-state kinetic model (Figure 1a), the single-molecule rate equation of each species were drafted based on the generation and consumption of the species (Supporting Information, section 1.3). When reaching equilibrium, the relative population of each species remains constant. This fact

https://dx.doi.org/10.1021/acs.analchem.0c01014
Anal. Chem. 2020, 92, 11582−11589

means that $\frac{dP_i(t)}{dt} = 0 \quad (i \in [E, ES, ES^*])$. Furthermore, the sum of the relative subpopulation of all species is always equal to 1. By setting up experiments on different [S], together with these boundary conditions, we obtained the populations of E, ES*, and ES. One can expect the system reaches equilibrium by considering the time-dependent populations at $t = \infty$. We summarized $P_E([S])$, $P_{ES^*}([S])$, and $P_{ES}([S])$ in eqs S36−S38. Since E and ES* contribute to the I_H while ES to the I_L, we also obtained the $P_{HL}([S])$ and $P_{LH}([S])$, as shown in eq 5 and eq 6, respectively. Figure 1f summarized the [S] dependent $P_{HL}([S])$ and $P_{LH}([S])$.

$$P_{HL}([S]) = \frac{k_2k_{-3} + k_3k_{-2}[S] + (k_{-2} + k_{-3})(k_{-1} + k_1[S])}{k_2(k_{-3} + (k_1 + k_3)[S]) + k_3k_{-2}[S] + k_1(k_{-2} + k_{-3})[S] + k_{-1}(k_{-2} + k_{-3} + k_3[S])} \quad (5)$$

$$P_{LH}([S]) = \frac{(k_1k_2 + k_3(k_2 + k_{-1}))[S]}{k_2(k_{-3} + (k_1 + k_3)[S]) + k_3k_{-2}[S] + k_1(k_{-2} + k_{-3})[S] + k_{-1}(k_{-2} + k_{-3} + k_3[S])} \quad (6)$$

Simulate Kinetic Properties of the Kinetic Model Using Single-Molecule Interaction Simulation (SMIS). We simulated single-molecule interaction trajectories in MATLAB (Supporting Information, section 2) by following the procedures below (Figure 2):

(Step 1) Define the kinetic model and transition probability from the rate constants. We first define the kinetic model by assigning the number of interacting species in each FRET level and assign rate constants for transitions between interacting species. For example, E, ES*, and ES specify the species in the model (Figure 2). The $k_{I,J}$ represents the rate constant for interconversion from the state I to J ($J \neq I$; and $I, J \in$ [E, ES*, ES]). In other words, k_{E,ES^*} is the k_1°, $k_{ES^*,ES}$ is the k_2, $k_{ES,E}$ is the k_{-3}, and vice versa. The transition from species I to J follows the relative probability $k_{I,J}/\sum_I k_{I,J}$.

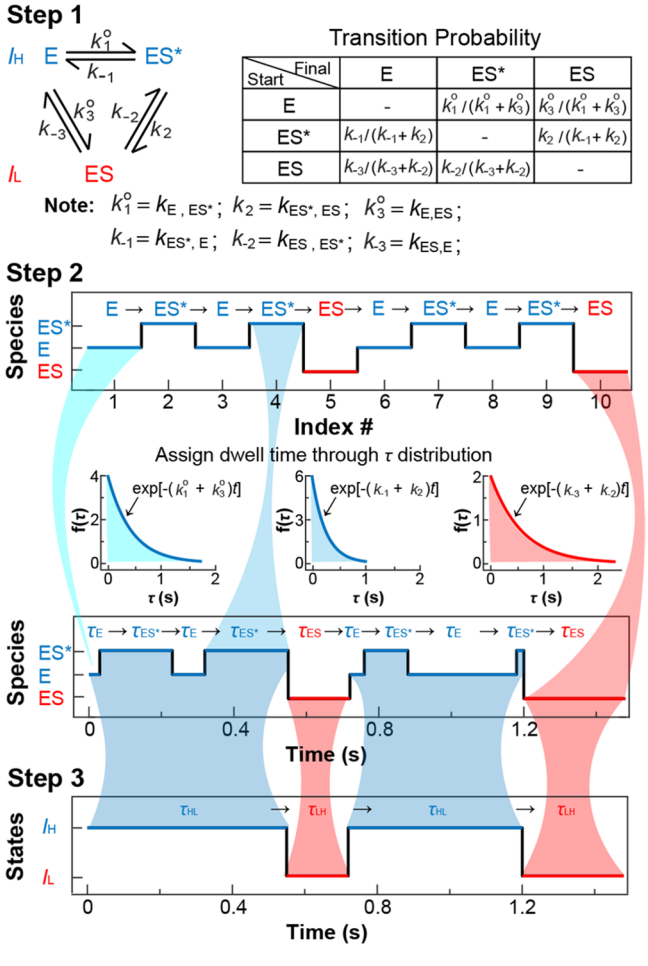


Figure 2. Workflow of SMIS.

(Step 2) Build a sequence of dwell time for each species. With the kinetic model defined, we define the sequence of species (e.g., E → ES* → E → ES* → ES → E → ES* → E → ES* → ES) based on the transition probability. The dwell times of each species follows $\sum_I k_{I,J} \exp(-\sum_I k_{I,J}t)$, where the $\sum_I k_{I,J}$ is a rate-constant sum of all competing pathways leaving from species I to J ($J \neq I$). We randomly sample one dwell time from the dwell-time distribution of $\sum_I k_{I,J} \exp(-\sum_I k_{I,J}t)$ and assign it to each species and generated the transition trajectories (e.g., $\tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_{ES} \rightarrow \tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_{ES}$).

(Step 3) Associate species with FRET levels and generate a single-molecule FRET trajectory. The dwell time for each species in the transition trajectories is associated with FRET levels. For example, by assigning E and ES* in the high, ES in the low FRET level, and combining dwell times belonging the same FRET level, we can convert the transition trajectory ($\tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_{ES} \rightarrow \tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_E \rightarrow \tau_{ES^*} \rightarrow \tau_{ES}$) into single-molecule FRET trajectory of $\tau_{HL} \rightarrow \tau_{LH} \rightarrow \tau_{HL} \rightarrow \tau_{LH}$.

(Step 4) Generate the probability density function of dwell time $PDF(\tau)$, average transition rates, and relative populations. With single-molecule FRET trajectories, we can extract microscopic dwell times (i.e., τ_{HL} and τ_{LH}). Normalizing the histograms of the dwell times by the overall area generates the probability density function of dwell times. The average FRET-state transition time is calculated from $\langle\tau_i\rangle = \frac{\sum \tau_i}{N_i}$ ($i \in$ [HL, LH]; N_i, number of dwell-times), whose reciprocal value reports the average transition rate between FRET levels. Using single-molecule transition trajectories generated from Step 2, we can generate the relative population of each species P_i ($i \in$ [E, ES*, and ES]). P_i is calculated by dividing the sum of microscopic dwell times (i.e., $\sum\tau_E$, $\sum\tau_{ES^*}$, or $\sum\tau_{ES}$) with the length of the trajectory (i.e., $P_i = \frac{\sum \tau_i}{\sum \tau}$). Similarly, we can also generate P_{HL} and P_{LH} (i.e., $P_{HL} = \frac{\sum \tau_{HL}}{\sum \tau}$, $P_{LH} = \frac{\sum \tau_{LH}}{\sum \tau}$).

RESULTS AND DISCUSSION

Validation of SMIS-Generated Simulations Using Michaelis−Menten Enzyme Kinetic Model. We validated the SMIS by comparing our simulation (200 000 dwell times) to

https://dx.doi.org/10.1021/acs.analchem.0c01014
Anal. Chem. 2020, 92, 11582−11589

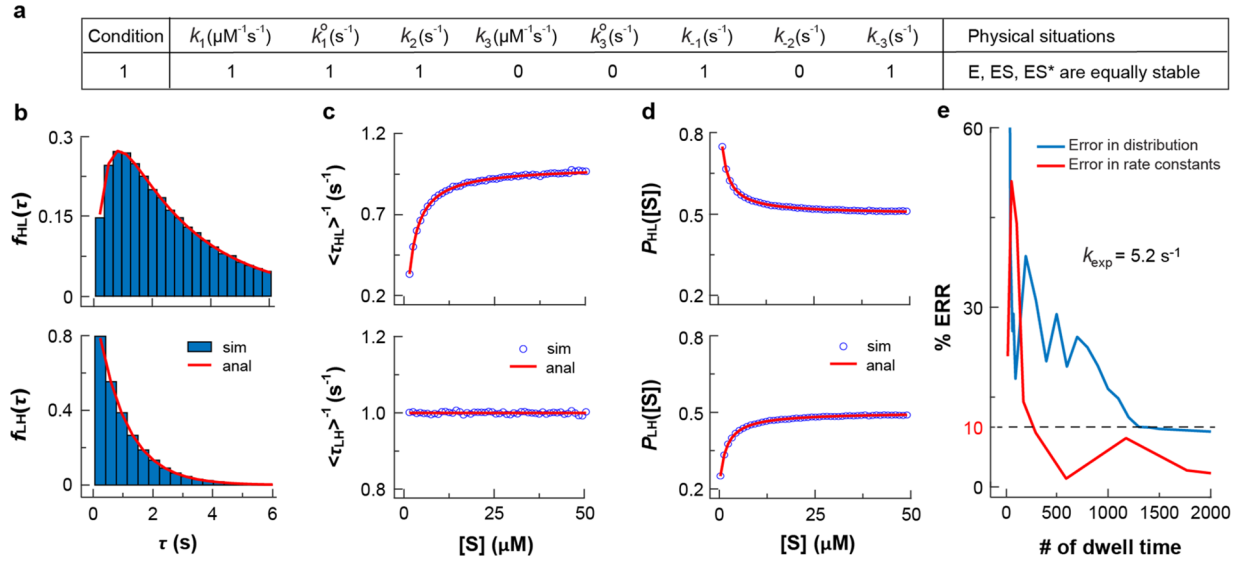


Figure 3. Comparisons between simulated and analytical results for the Michaelis−Menten enzyme kinetic model. (a) Rate constant inputs for the SMIS simulation and analytical solution of the Michaelis−Menten model. (b−d) Comparisons between simulated (blue bar or circle) and analytical (red curve) probability density function $f_{HL}(\tau)$ and $f_{LH}(\tau)$ (b), average transition rate $\langle\tau_{HL}\rangle^{-1}$ and $\langle\tau_{LH}\rangle^{-1}$ (c), and relative populations P_{HL} and P_{LH} (d) under condition 1. (e) Percent error (% ERR) of extracted rate constants (red) and dwell-time distributions (blue) as a function of the number of dwell times.

the well-known Michaelis−Menten enzyme kinetic model with rate constants varying from 0.01 to 100 s^{-1}. The Michaelis−Menten model is a simplified condition of our two-state model where k_{-2} and k_3 are equal to zero. By replacing k_{-2} and k_3 with zeros in eq 1 to eq 6, we obtained the $f_{HL}(\tau)$, $f_{LH}(\tau)$, P_{HL}, P_{LH}, $\langle\tau_{HL}\rangle^{-1}$, and $\langle\tau_{LH}\rangle^{-1}$ of the Michalis−Menten model in eq 7 to eq 12, respectively.

$$f_{HL}(\tau) = \frac{k_1^{\circ}k_2}{2A}(e^{(B+A)\tau} - e^{(B-A)\tau}) \tag{7}$$

where

$$A = \frac{\sqrt{(k_{-1} + k_2 + k_1^{\circ})^2 - 4k_2k_1^{\circ}}}{2}$$

$$B = \frac{-(k_2 + k_1^{\circ} + k_{-1})}{2}$$

$$f_{LH}(\tau) = k_{-3}\, e^{-k_{-3}\tau} \tag{8}$$

$$P_{HL}([S]) = \frac{k_2k_{-3} + k_{-1}k_{-3} + k_1k_{-3}[S]}{k_2(k_{-3} + k_1[S]) + k_1k_{-3}[S] + k_{-1}k_{-3}} \tag{9}$$

$$P_{LH}([S]) = \frac{k_1k_2[S]}{k_2(k_{-3} + k_1[S]) + k_1k_{-3}[S] + k_{-1}k_{-3}} \tag{10}$$

$$\langle\tau_{HL}\rangle^{-1} = \frac{k_1k_2[S]}{k_2 + k_{-1} + k_1[S]} \tag{11}$$

$$\langle\tau_{LH}\rangle^{-1} = \left(\int_0^{\infty} \tau f_{LH}(\tau)\, d\tau\right)^{-1} = k_{-3} \tag{12}$$

Our simulation accurately recovered the characteristics predicted by the Michaelis−Menten equations. We compared the analytical and simulated results using a model where E, ES*, and ES are equally stable (i.e., condition 1 in Figure 1c). The other conditions mimic the ES* and ES existing as transient and stable complexes are summarized in Supporting Information, section 3. Figure 3b−d shows the comparison of dwell-time distributions (Figure 3b), average transition rates (Figure 3c), and relative populations (Figure 3d) of each species. In all conditions, the simulations nicely overlap with the predictions, validating that SMIS successfully generate single-molecule trajectories for the target kinetic model.

To determine how many dwell-times are needed to robustly recovery the rate constants and the dwell-time distributions, we compared how the simulations deviate from the experimentally determined distribution $f_{exp}(\tau)$. The $f(\tau)$ was generated from ∼2 000 dwell times. Fitting the $f_{exp}(\tau)$ gave a decay rate constant k_{exp} of 5.2 s^{-1}. We simulated the dwell-time distributions using k_{exp} with the number of dwell time varies from 30 to 10 000. Figure 3e shows the percent difference of extracted rate constants and dwell-time distributions from $f_{exp}(\tau)$. The percent difference decreases and gets below 10% once the number of dwell time exceeds 150 and 250 for recovering the rate constants and the dwell-time distributions, respectively.

Validation of SMIS-Extracted Rate Constants Using Simulations of Kinetic Models. $f(\tau)$ plays a crucial role in quantifying kinetic rate constants in the typical single-molecule FRET approach. General procedures to derive $f(\tau)$ involve (1) formulating the kinetic model from experimental results, (2) dissecting kinetic model to specific transitions between FRET levels, and (3) solving the differential equations with proper initial conditions (see example in the Supporting Information, section 1). However, solving $f(\tau)$ could be difficult or impossible when having repeated differential equations, which results in a deficit of useful information (i.e., number of useful equations is less than the number of variables). Difficulty can also arise from unclear initial conditions, such as multiple species coexisting in the same FRET level with relative population undefined. Here, we applied SMIS to extract kinetic rate constants. Using a model with analytical solutions, we showed SMIS robustly recaptured the unique solutions as input.

To validate the uniqueness of extracted rate constants, we tested SIMS by using the Michaelis−Menten enzyme kinetic model with rate constants listed in Figure 4a. With the number

https://dx.doi.org/10.1021/acs.analchem.0c01014
Anal. Chem. 2020, 92, 11582−11589

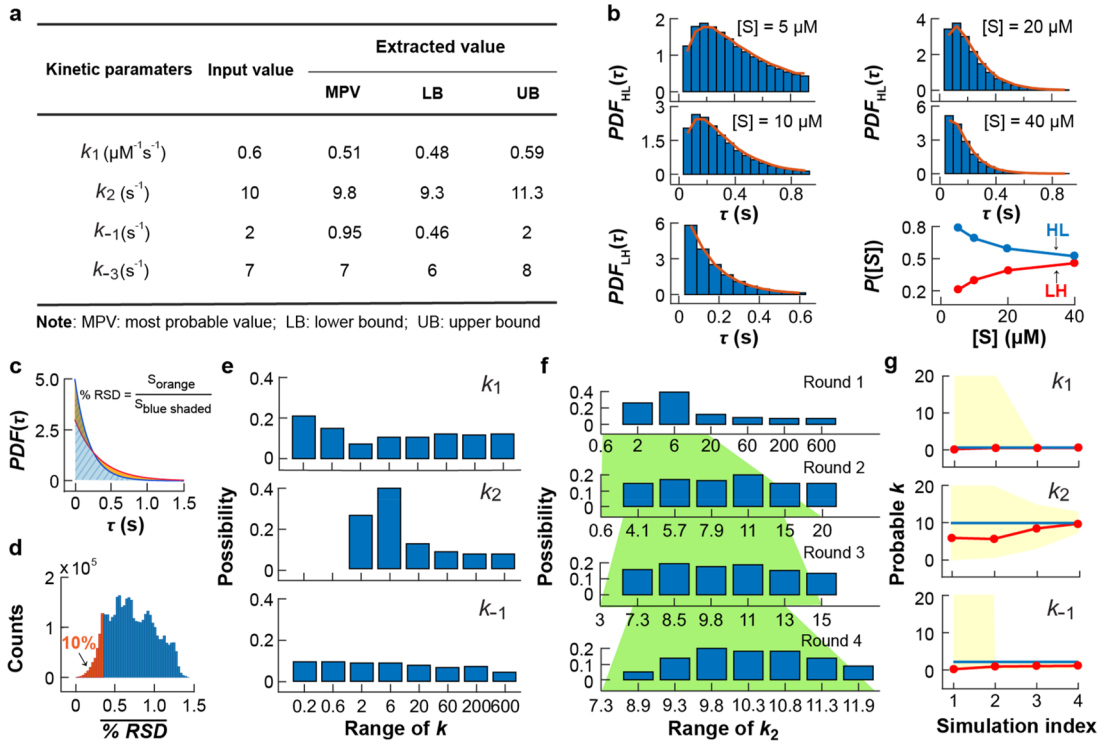


Figure 4. Application of SMIS to kinetic model with analytical solutions. (a) Summary of input and extracted rate constants for SMIS. (b) Experimental data (blue bars) used to compare with the SMIS simulations (red curve). Panels from top to bottom show the $PDF_{HL}(\tau)$, $PDF_{LH}(\tau)$, and [S] dependent population (P_{HL} and P_{LH}). (c) Example of % RSD calculation. The % RSD is estimated by the ratio of the residue area (orange) to the total area under experimental data (blue shaded area). (d) Histogram of simulations based on their $\overline{\%RSD}$ from the first-round screening. The red area highlights the top 10% data with the smallest error. (e) Histogram of rate constants from the top 10% simulations in panel d. (f) Bisection method to extract rate constant through simulations. (g) Progression of identifying correct rate constants. The yellow area highlights the screening range of k in each screening. The blue lines indicate the input rate constants for the experimental data. The red lines show the most probable rate constant extracted from SMIS.

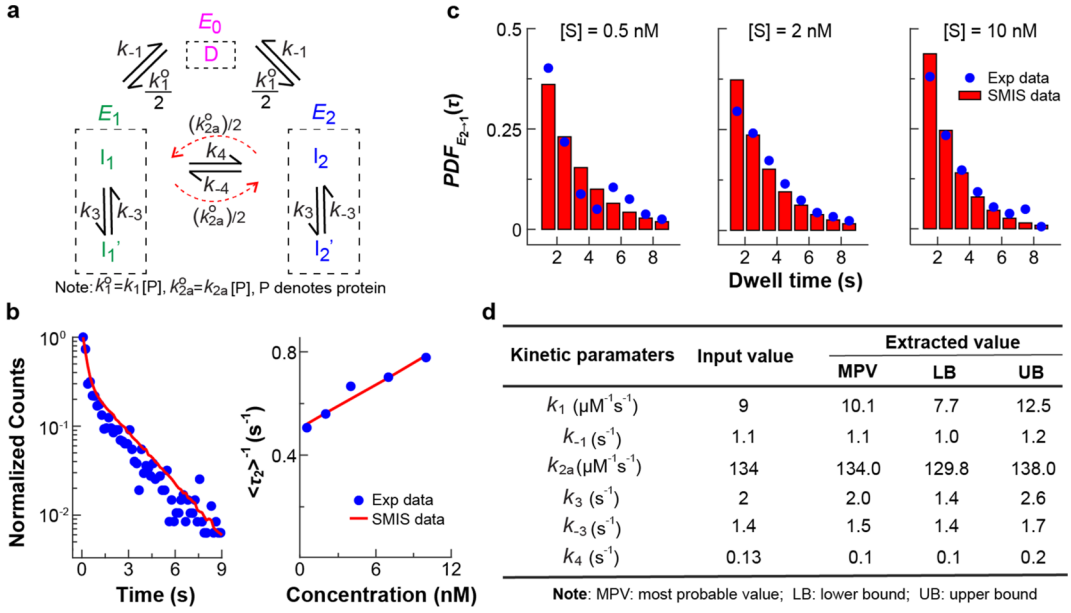
of transition set to 500 000 (i.e., number of dwell time of 217 000), we generated the $PDF_{HL}(\tau)$, $PDF_{LH}(\tau)$, P_{HL}, and P_{LH} under four substrate concentrations ([S] = 5, 10, 20, and 40 μM) as shown in Figure 4b. This set of data serves as the ground truth data, on which we applied SMIS to extract the rate constants.

To extract out the rate constants, we performed an extensive simulation using SMIS and search for most probable rate constants by minimizing the average of percent residue. We adapted SMIS to simulate the $PDF_{HL}(\tau)$, $PDF_{LH}(\tau)$, P_{HL}, and P_{LH} with all four rate constants vary from 0.01 to 1000 s^{-1} in the first search. Deviation of each simulation from the experimental data was individually estimated through the averaged % RSD, $\overline{\%RSD}$. For each simulation, differences of simulated $PDF_{HL}(\tau)$, $PDF_{LH}(\tau)$, P_{HL}, and P_{LH} from the ground truth data were calculated by the ratio of residue to the total area under the curve (Figure 4c). Averaging of all deviations gave the averaged % RSD, $\overline{\%RSD}$, which serves as the goodness of simulation for each input rate constant set. Since the simulation has an ~10% error, we consider the simulations with $\overline{\%RSD}$ < 10% are equally accurate. We sorted and picked the simulations with $\overline{\%RSD}$ < 10% (Figure 4d) and histogrammed each rate constant (Figure 4e). Using these histograms, we bisected the most probable range for each rate constant when the sum of the possibility of simulated values (starting from high to low) is more than 50%. Take k_2 as an example (Figure 4f); we picked the second and the first bin to define a new range for the most probable k_2. This approach provides the boundaries of rate constants for the subsequent screening. By repeating the searching process, we first simulated the PDF_{LH} to extract the k_{-3} and simulated the PDF_{HL} to extract the k_1, k_2, and k_{-1}. Figure 4g shows the searching results for each rate constant after four searches, indicating SMIS recovered the unique solutions of the kinetic model.

The analyses code aiming to solve reactions that involve more than two FRET levels are made available in recent years.[38−40] These multiple-state smFRET analysis codes are based on either hidden Markov modeling (HMM) or change point analysis, which recovers the noise-free single-molecule transition trajectories and reveals the apparent transition rates between FRET levels. However, if more than one species existing in a FRET level, extracting interaction rate constants between them could be challenging. We further performed the same analysis to a kinetic model with three FRET levels and successfully recovered all kinetic rate constants (Supporting Information, section 4).

Application of SMIS to Simulate and Extract Kinetic Parameters of a Transcription Regulator, CueR. We applied SMIS to the smFRET data of a Cu^+-responsive MerR-family transcription regulator, CueR, in its holo form (i.e., holo-CueR).[25] Holo-CueR interacts with DNA to activate transcription of downstream Cu-binding proteins. Joshi et al. used smFRET to study the CueR−DNA interactions and quantified the relevant kinetic rate constants. The quantitative analysis revealed interesting CueR concentration-dependent unbinding, which leads to the discovery of transcription deactivation through a multivalency-enabled ternary complexes pathway.[41]

https://dx.doi.org/10.1021/acs.analchem.0c01014
Anal. Chem. 2020, 92, 11582−11589



Kinetic paramaters	Input value	Extracted value		
		MPV	LB	UB
k_1 ($\mu M^{-1}s^{-1}$)	9	10.1	7.7	12.5
k_{-1} (s^{-1})	1.1	1.1	1.0	1.2
k_{2a} ($\mu M^{-1}s^{-1}$)	134	134.0	129.8	138.0
k_3 (s^{-1})	2	2.0	1.4	2.6
k_{-3} (s^{-1})	1.4	1.5	1.4	1.7
k_4 (s^{-1})	0.13	0.1	0.1	0.2

Note: MPV: most probable value; LB: lower bound; UB: upper bound

Figure 5. Application of SMIS to simulate and extract kinetic parameters of CueR−DNA interactions. (a) Kinetic model for five interacting species (D, I_1, I_1', I_2, and I_2') associated with three FRET levels (E_0, E_1, and E_2). (b) Comparison between experimental data with SIMS simulations. (c) Experimental [CueR] dependent dwell-time distributions ($\tau_{2\to1}$) can be well-described by the SMIS results. (d) Summary of rate constants obtained from analytical solution and SMIS approaches.

They proposed a kinetic model associating three FRET levels (E_0, E_1, and E_2) to the unbound and two bound forms of CueR (Figure 5a). The unbinding kinetics of CueR were revealed by the dwell-times of either the E_1 or E_2 state (i.e., τ_1 and τ_2). In the three-level kinetic model, two intermediates (I_1' and I_2') are, respectively, presented in the E_1 and E_2 levels, leading to a challenging derivation of the analytical solutions to extract rate constants (especially k_3 and k_{-3} from the dwell-time distribution of the E_1 or E_2 level). Using the proposed kinetic model, we applied SMIS to simulate the dwell-time distribution of the E_2 level (τ_2) and CueR concentration-dependent $\langle\tau_2\rangle^{-1}$. Figure 5b shows the simulated results agree well with the experimental data. We then globally examined the dwell-time distributions with CueR concentration ([CueR]) varying from 0.5 to 10 nM (Figure 5c). By screening the rate constants from 0.1 to 200 s^{-1}, we extracted the rate constants in agreeing with the reported values (Figure 5d). These results collectively demonstrate that SMIS is useful to quantify the interaction kinetic rate constants even without the analytical solutions.

SMIS approaches the interaction kinetics from an angle that is based more on potential species at each FRET level. This approach allows a direct association between the kinetic model and the targeted biological process as well as the extraction of interaction kinetics between species in the same FRET level. In our opinion, SMIS will be useful toward the systems with moderate complicated kinetic models. We believe SMIS is complementary to the analytical and HMM approach to provide quantitative kinetic information and thus increase the application of smFRET in biological studies.

CONCLUSION

The single-molecule interaction simulation (SMIS) opens new possibilities for objective characterization of interaction kinetics based on the kinetic model of interest, regardless of whether the analytic solutions are available or not. With the two-FRET-level model, we derived, as well as used SMIS to simulate, the probability density function (i.e., $f_{HL}(\tau)$, $f_{LH}(\tau)$), average transition rates ($\langle\tau_{HL}\rangle^{-1}$ and $\langle\tau_{LH}\rangle^{-1}$), and the relative populations of high and low states (i.e., P_{HL} and P_{LH}). These derived analytical solutions justified the feasibility of SMIS to recover important kinetic distributions using the Michaelis−Menten enzyme kinetic model. The same conclusion can be drawn for the three-FRET-level model. To test how effective SMIS can extract rate constants, we applied SMIS to extract rate constants that were responsible for CueR−DNA interactions. These results indicate the SMIS is useful in providing essential characteristics of kinetic parameters for an assigned kinetic model. Comparison between the experimentally determined distributions of kinetic parameters and the simulations crossing a wide range of rate constants can robustly quantify the rate constants. Our findings here contribute to the quantitative analysis of smFRET data, which is an essential step toward understanding biophysical problems using the smFRET approach.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.0c01014.

Derivation of analytical expression of $f(\tau)$, $\langle\tau\rangle^{-1}$ of two FRET levels model: analytical expression of $f_{HL}(\tau)$, $\langle\tau_{HL}\rangle^{-1}$, $P_{ES}(t)$, $P_E(t)$, $P_{ES^*}(t)$ of high (I_H) to low (I_L) model; analytical expression of $f_{LH}(\tau)$, $\langle\tau_{LH}\rangle^{-1}$, $P_{ES}(t)$, $P_E(t)$, $P_{ES^*}(t)$ of low (I_L) to high (I_H) model; analytical expression of $P_{ES}([S])$, $P_E([S])$,$P_{ES^*}([S])$ of two-state model; SMIS validation for Michaelis−Menten model (other 4 conditions): Validation of simulated $f_{HL}(\tau)$ and $f_{LH}(\tau)$ as a function of k; validation of simulated $\langle\tau_{HL}\rangle^{-1}$ and $\langle\tau_{LH}\rangle^{-1}$ as a function of k; validation of simulated $P_E([S])$, $P_{ES^*}([S])$, and $P_{ES}([S])$ as a function of k(PDF)

MatLab scripts of SMIS (ZIP)

https://dx.doi.org/10.1021/acs.analchem.0c01014
Anal. Chem. 2020, 92, 11582−11589

AUTHOR INFORMATION

Corresponding Author

Tai-Yen Chen − *Department of Chemistry, University of Houston, Houston, Texas 77204, United States;* orcid.org/0000-0002-2881-3068; Phone: 1-713-743-9413; Email: tchen37@central.uh.edu

Authors

Manhua Pan − *Department of Chemistry, University of Houston, Houston, Texas 77204, United States*

Yuteng Zhang − *Department of Chemistry, University of Houston, Houston, Texas 77204, United States*

Guangjie Yan − *Department of Chemistry, University of Houston, Houston, Texas 77204, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.0c01014

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from NIH (Grant R35GM133505), the Welch Foundation (Grant E-1947), and the University of Houston.

REFERENCES

(1) Li, I. T.; Pham, E.; Truong, K. *Biotechnol. Lett.* **2006**, *28* (24), 1971−1982.
(2) Sharma, S.; Chakraborty, K.; Muller, B. K.; Astola, N.; Tang, Y. C.; Lamb, D. C.; Hayer-Hartl, M.; Hartl, F. U. *Cell* **2008**, *133* (1), 142−153.
(3) Lamichhane, R.; Solem, A.; Black, W.; Rueda, D. *Methods* **2010**, *52* (2), 192−200.
(4) Liu, R.; Hu, D.; Tan, X.; Lu, H. P. *J. Am. Chem. Soc.* **2006**, *128* (31), 10034−10042.
(5) Brunger, A. T.; Strop, P.; Vrljic, M.; Chu, S.; Weninger, K. R. *J. Struct. Biol.* **2011**, *173* (3), 497−505.
(6) Ha, T. *Curr. Opin. Struct. Biol.* **2001**, *11* (3), 287−292.
(7) Heilemann, M.; Hwang, L. C.; Lymperopoulos, K.; Kapanidis, A. N. *Methods Mol. Biol.* **2009**, *543*, 503−521.
(8) Feng, Q.; Liang, S.; Jia, H.; Stadlmayr, A.; Tang, L.; Lan, Z.; Zhang, D.; Xia, H.; Xu, X.; Jie, Z.; Su, L.; Li, X.; Li, X.; Li, J.; Xiao, L.; Huber-Schonauer, U.; Niederseer, D.; Xu, X.; Al-Aama, J. Y.; Yang, H.; Wang, J.; Kristiansen, K.; Arumugam, M.; Tilg, H.; Datz, C.; Wang, J. *Nat. Commun.* **2015**, *6*, 6528.
(9) Hwang, W. L.; Deindl, S.; Harada, B. T.; Zhuang, X. *Nature* **2014**, *512* (7513), 213−217.
(10) Sarkar, S. K.; Andoy, N. M.; Benitez, J. J.; Chen, P. R.; Kong, J. S.; He, C.; Chen, P. *J. Am. Chem. Soc.* **2007**, *129* (41), 12461−12467.
(11) McKinney, S. A.; Declais, A. C.; Lilley, D. M.; Ha, T. *Nat. Struct. Biol.* **2003**, *10* (2), 93−97.
(12) Chen, J.; Poddar, N. K.; Tauzin, L. J.; Cooper, D.; Kolomeisky, A. B.; Landes, C. F. *J. Phys. Chem. B* **2014**, *118* (42), 12130−12139.
(13) Keller, B. G.; Kobitski, A.; Jaschke, A.; Nienhaus, G. U.; Noe, F. *J. Am. Chem. Soc.* **2014**, *136* (12), 4534−4543.
(14) Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Haustein, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (26), 15516−15521.
(15) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. A. *Science* **2012**, *335* (6071), 981−984.
(16) McLoughlin, S. Y.; Kastantin, M.; Schwartz, D. K.; Kaar, J. L. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (48), 19396−19401.
(17) Uphoff, S.; Holden, S. J.; Le Reste, L.; Periz, J.; van de Linde, S.; Heilemann, M.; Kapanidis, A. N. *Nat. Methods* **2010**, *7* (10), 831−836.
(18) Chen, J.; Pyle, J. R.; Sy Piecco, K. W.; Kolomeisky, A. B.; Landes, C. F. *J. Phys. Chem. B* **2016**, *120* (29), 7128−7132.
(19) Hadzic, M.; Borner, R.; Konig, S. L. B.; Kowerko, D.; Sigel, R. K. O. *J. Phys. Chem. B* **2018**, *122* (23), 6134−6147.
(20) Bauer, M.; Li, C.; Mullen, K.; Basche, T.; Hinze, G. *J. Chem. Phys.* **2018**, *149* (16), 164104.
(21) Borner, R.; Kowerko, D.; Hadzic, M.; Konig, S. L. B.; Ritter, M.; Sigel, R. K. O. *PLoS One* **2018**, *13* (4), e0195277.
(22) Danial, J. S.; Garcia-Saez, A. J *Curr. Opin. Struct. Biol.* **2017**, *46*, 24−30.
(23) Bokinsky, G.; Zhuang, X. *Acc. Chem. Res.* **2005**, *38* (7), 566−573.
(24) Smith, J. G.; Jain, P. K. *Phys. Chem. Chem. Phys.* **2016**, *18* (34), 23990−23997.
(25) Joshi, C. P.; Panda, D.; Martell, D. J.; Andoy, N. M.; Chen, T. Y.; Gaballa, A.; Helmann, J. D.; Chen, P. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (38), 15121−15126.
(26) Chen, T. Y.; Santiago, A. G.; Jung, W.; Krzeminski, L.; Yang, F.; Martell, D. J.; Helmann, J. D.; Chen, P. *Nat. Commun.* **2015**, *6*, 7445.
(27) Martell, D. J.; Joshi, C. P.; Gaballa, A.; Santiago, A. G.; Chen, T. Y.; Jung, W.; Helmann, J. D.; Chen, P. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (44), 13467−13472.
(28) Shen, H.; Zhou, X.; Zou, N.; Chen, P. *J. Phys. Chem. C* **2014**, *118* (46), 26902−26911.
(29) Figliozzi, P.; Peterson, C. W.; Rice, S. A.; Scherer, N. F. *ACS Nano* **2018**, *12* (6), 5168−5175.
(30) Wang, W.; Shen, H.; Moringo, N. A.; Carrejo, N. C.; Ye, F.; Robinson, J. T.; Landes, C. F. *Langmuir* **2018**, *34* (23), 6697−6702.
(31) Dunn, I. J.; Heinzle, E.; Ingham, J.; Prenosil, J. E. *Biological Reaction Engineering*; John Wiley & Sons, 2003.
(32) Robeva, R.; Kirkwood, J. R. *Laboratory Manual of Biomathematics*; Academic Press, 2007.
(33) Snape, J. B.; Dunn, I. J.; Ingham, J.; Prenosil, J. E. *Dynamics of Environmental Bioprocesses: Modeling and Simulation*; John Wiley & Sons, 2008.
(34) Padovani, D.; Labunska, T.; Palfey, B. A.; Ballou, D. P.; Banerjee, R. *Nat. Chem. Biol.* **2008**, *4* (3), 194−196.
(35) Haran, G. *Chem. Phys.* **2004**, *307* (2−3), 137−145.
(36) Taylor, J. N.; Landes, C. F. *J. Phys. Chem. B* **2011**, *115* (5), 1105−1114.
(37) Lee, H. C.; Tu, I. P.; Lin, B. L.; Chang, W. H. *J. Biomed. Opt.* **2012**, *17* (1), 011007.
(38) McKinney, S. A.; Joo, C.; Ha, T. *Biophys. J.* **2006**, *91* (5), 1941−1951.
(39) Bronson, J. E.; Fei, J.; Hofman, J. M.; Gonzalez, R. L., Jr; Wiggins, C. H. *Biophys. J.* **2009**, *97* (12), 3196−3205.
(40) Watkins, L. P.; Yang, H. *J. Phys. Chem. B* **2005**, *109* (1), 617−628.
(41) Chen, T. Y.; Cheng, Y. S.; Huang, P. S.; Chen, P. *Acc. Chem. Res.* **2018**, *51* (4), 860−868.