JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Long-Term Observation of Fluorescence of Free Single Molecules To Explore Protein-Folding Energy Landscapes

Kiyoto Kamagata,[†] Toshifumi Kawaguchi,[†] Yoshitomo Iwahashi,[‡] Akinori Baba,[#] Kazuya Fujimoto,[§] Tamiki Komatsuzaki,[¶,∥] Yoshihiro Sambongi,[⊥,∥] Yuji Goto,[§] and Satoshi Takahashi*[,†,∥]

[†]Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan

[‡]Optonica Corporation, 1-26-11, Matsuigaoka, Kyotanabe, Kyoto 610-0353, Japan

[#]RIKEN Center for Developmental Biology, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan

[§]Institute for Protein Research, Osaka University, 3-2, Yamadaoka, Suita, Osaka 565-0871, Japan

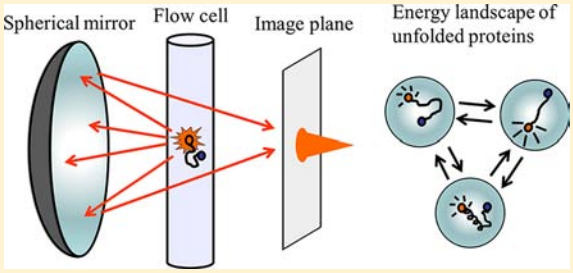
[¶]Molecule & Life Nonlinear Sciences Laboratory, Research Institute for Electronic Science, Hokkaido University, Kita 20 Nishi 10, Kita-ku, Sapporo 001-0020, Japan

[∥]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, 4-1-8, Honmachi, Kawaguchi, Saitama 332-0012, Japan

[⊥]Graduate School of Biosphere Science, Hiroshima University, 1-4-4, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8528, Japan

Supporting Information

ABSTRACT: A method was developed to detect fluorescence intensity signals from single molecules diffusing freely in a capillary cell. A unique optical system based on a spherical mirror was designed to enable quantitative detection of the fluorescence intensity. Furthermore, "flow-and-stop" control of the sample can extend the observation time of single molecules to several seconds, which is more than 1000 times longer than the observation time available using a typical confocal method. We used this method to scrutinize the fluorescence time series of the labeled cytochrome *c* in the unfolded state. Time series analyses of the trajectories based on local equilibrium state analysis revealed dynamically differing substates on a millisecond time scale. This system presents a new avenue for experimental characterization of the protein-folding energy landscape.



INTRODUCTION

Single-molecule fluorescence spectroscopy (SMFS) has evolved as a powerful means to reveal the complex machinery of proteins that fold into unique native conformations[1−8] and to perform various functions in living systems.[9−12] In typical SMFS experiments, the distance between two sites of a molecule labeled by donor and acceptor fluorophores can be deduced by detecting their emissions at the single-molecule level and by analyzing data based on the fluorescence resonance energy transfer (FRET) mechanism.[1−6,8,10,11,13] Structural and dynamical information about individual components of the sample, which can be obtained only minimally using techniques relying on ensemble averaging, can be acquired adequately using SMFS. The conformation and the time constant of the conformational fluctuations of the unfolded state of proteins have been characterized on the basis of SMFS, which provides an advanced understanding of protein-folding phenomena.[1−8]

Despite marked progress in the development of SMFS techniques, the mechanisms of folding dynamics of proteins remain unclear, especially those related to the energy landscape of proteins in the unfolded state. That landscape represents the free energy relationship of various substates in the unfolded ensemble and their kinetic connectivity.[14] Molecular dynamics simulations of small polypeptides and proteins have suggested that the unfolded state comprises a network of several substates, which serves as an important determinant of protein foldability.[14−17] Single-molecule time series should carry crucial information related to the kinetic connectivity among the substates on the energy landscapes.[14,18−22] However, characterizations of the energy landscape of actual proteins in the unfolded state have been hampered by the difficulty in obtaining the time series for long durations using currently available SMFS techniques.

Several SMFS approaches, each presenting distinct advantages and disadvantages, have been proposed.[1,2,7,10,23,33] One approach, confocal microscopy, can detect fluorescence from single molecules in the time range up to milliseconds,[2] but longer time ranges cannot be followed because of diffusion of the molecules out of the focal volume of the microscopic lens.[23] Another approach, total internal reflection (TIRF) microscopy, avoids the diffusion problem by fixing molecules on an optical

Received: March 1, 2012
Published: June 12, 2012

   dx.doi.org/10.1021/ja3020555 | J. Am. Chem. Soc. 2012, 134, 11525−11532

surface. Using this approach, single molecules can be observed for extended periods of time.[10,24] Although several immobilization techniques with small artifacts have been reported,[25,26] fixation sometimes creates artifacts in the properties of the samples.[1] Several attempts[7,27−33] have been made to combine advantages of the two approaches described above. Flow detection of single molecules can monitor single molecules for an extended time without fixation.[7] However, the observation time remains limited to about 100 ms. Recent efforts using the anti-Brownian electrokinetic trap enabled the detection of single molecules for an extended time.[27−29] However, the method cannot trap samples having states with small emissions. Other attempts have trapped molecules inside of a small compartment such as liposomes,[3] water-in-oil emulsions,[30] and nanofabricated devices;[31] convex lens-induced confinement has also been reported.[32] These methods are not applicable to molecules that adsorb on the inner walls of compartments. Therefore, development of a strategy as an alternative to SMFS for the facile observation of single molecules that are diffusing freely for an extended duration is of pivotal importance.

To this end, we developed an experimental protocol to track freely diffusing molecules for an extended duration: the flow-and-stop protocol. This article presents a description of the principles and the features of our new method, followed by investigation of the unfolded state of cytochrome *c*551 (cyt *c*) from *Pseudomonas aeruginosa* labeled with Atto532. Analyses of single-molecule data obtained over an extended duration demonstrate the kinetic connectivity among substates buried in the unfolded state, which is an important first step for experimental characterization of the energy landscapes of proteins.

RESULTS AND DISCUSSION

Flow-and-Stop Protocol with a Spherical Mirror and Pressure-Regulated Pump. The flow-and-stop protocol enables observation of fluorescence from single molecules over an extended period. This developed system comprises a flow system and an optical system. The flow system controls the sample flow in the fused silica capillary cell. A protein of interest labeled with fluorophore(s) is dissolved at a low concentration such that one molecule at a time flows into the cell. The optical system is designed to excite the fluorophores and to collect their emissions in the observation region of the cell (Figure 1a). To prolong the observation time, once the sample reaches the observation region, the flow is stopped and the fluorescence signal is tracked continuously (Figure 1b). The protocol enables us to monitor fluorescence from single molecules until the moment of bleaching (typically several seconds).

The flow-and-stop protocol necessitates imaging of single molecules in a large field of view and requires a large depth of field that covers both the observation region and the cell cross section (Figure 1b). That large field of view is necessary to track the fluorescence from the samples diffusing freely along the cell. The large depth of field enables the usage of standard capillaries with a pore size of about 5 μm. For confocal microscopy based on conventional objectives, the field of view and the depth of field are each less than 1 μm. As we demonstrated previously,[7] a camera lens can enlarge the observation area and the depth of field to several millimeters and about 10 μm, respectively. However, the numerical aperture (NA) of the camera lens is only 0.3, meaning that its efficiency at collecting photons (2.3%) is markedly lower

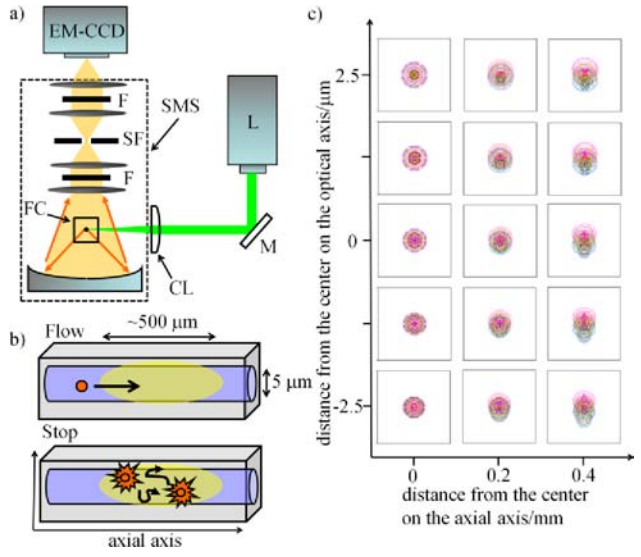


Figure 1. (a) System for the SMFS flow-and-stop protocol. A laser beam (L) is focused along the flow cell (FC) by a cylindrical lens (CL). Protein fluorescence is portrayed on an EM-CCD with a spherical mirror system (SMS) comprising a mirror, a slit spatial filter (SF), lenses, and optical filters (F). (b) Flow-and-stop protocol. Orange circle represents a sample molecule flowing along the cell (upper). It is stopped in the observation area illuminated by laser light (lower). (c) Calculated images of a point light source focused by the spherical mirror system to positions inside the flow cell. The axial axis shows the light source position along the flow cell. The optical axis represents the normal projection from the cell to the detector through the mirror: 0 μm represents the center of the cell; positive values represent positions near the mirror and distant from the lens. The box is 100 μm × 100 μm.

than that of objectives (more than 14.3%) having NA of more than 0.7. Consequently, a new optical system must be used to establish the SMFS flow-and-stop protocol.

To increase the light collection efficiency, we developed a new optics based on a spherical mirror and combination lenses (Figure 1a). It has a high NA value of 0.7, a low magnification of 3.1, and a large working distance of 9.5 mm. Additionally, it provides a field of view of about 2 mm × 2 mm and depth of field of 15 μm, the latter of which exceeds the inner diameter (5 μm) of the cell used. Comparison of our system to a conventional system is described in Supporting Information, Table 1. The spherical mirror enables the collection of photons with high NA while maintaining low magnification. These features are difficult to achieve for refractive objectives. To correct the spatial aberration of the focused image at off-axis points, which is inherent to a spherical mirror, we introduced a set of combination lenses (Supporting Information Figure 1 and supplementary methods). Simulated images from point light sources placed at several off-axis points demonstrate the quality of this developed optics (Figure 1c), which enables viewing of the entire observation region while maintaining high efficiency of light collection, both of which are prerequisite for the SMFS flow-and-stop protocol.

We introduced a pressure-regulated pump, which can maintain a slow passage of molecules on the order of 1 mm/s in the narrow cell with a 5-μm pore. Furthermore, the pump can control the flow speed quickly, with a response time of several seconds. The slow flow speed and the rapid response of the pump can control the sample flow delicately, which is another prerequisite for the flow-and-stop protocol.

dx.doi.org/10.1021/ja3020555 | J. Am. Chem. Soc. 2012, 134, 11525−11532

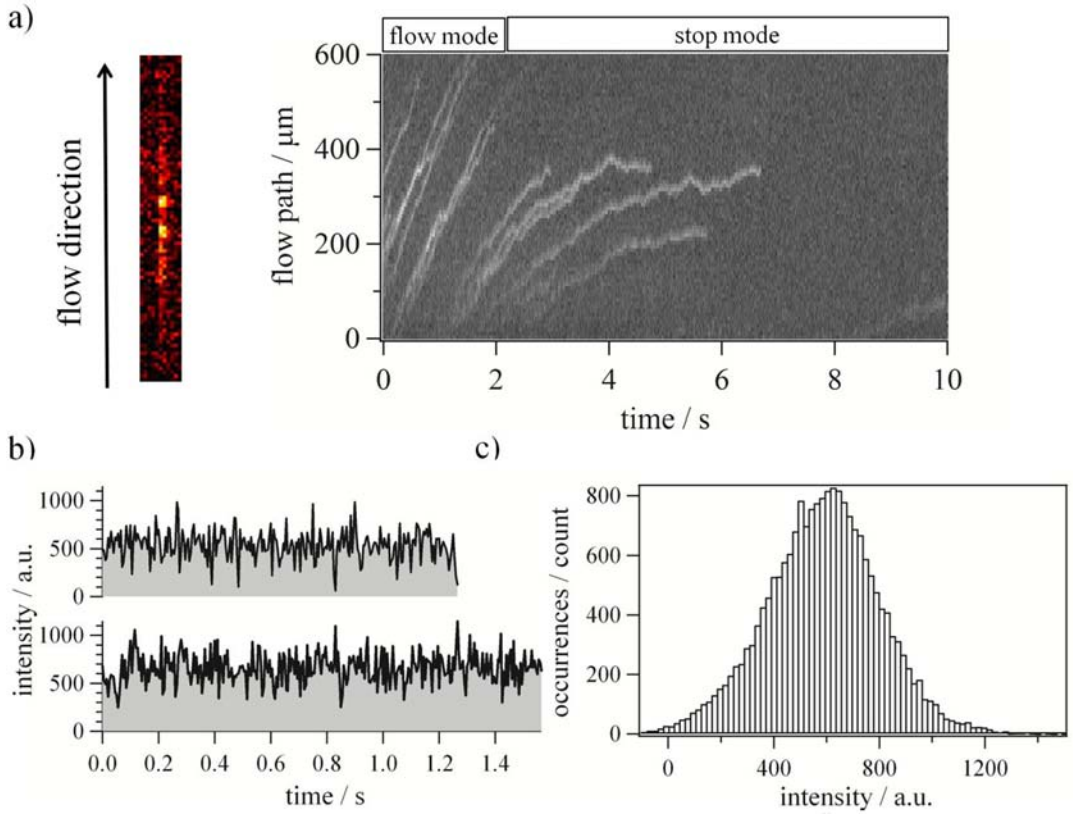


Figure 2. Single-molecule time series for Atto532. (a) Typical fluorescence image of Atto532 in the flow cell (left) and the two-dimensional kymograph of Atto532 during the flow-and-stop protocol (right). (b) Typical single-molecule time series of Atto532. (c) Intensity histogram of traces of Atto532.

Features of Fluorescence Data Detected Using the Flow-and-Stop Protocol. Figure 2 presents fluorescence intensity data obtained using the flow-and-stop protocol for a dye, Atto532. Consecutive images of one measurement are presented in Supporting Information, Video 1. The left of Figure 2a shows a raw image obtained by an electron multiplying charge-coupled device (EM-CCD) with accumulation time of 5 ms, in which the bright spots represent single-molecule fluorescence from Atto532. The kymograph displayed on the right is a two-dimensional representation of the images taken consecutively. The images of the observation region are converted to a set of one-dimensional data by accumulating pixel intensities along the radial axis of the capillary. The one-dimensional data are stacked from left to right, with molecules represented as white lines. In the "flow" mode during 0−2 s, lines represent flowing molecules. In the "stop" mode after 2 s, the flow slows down gradually and the free diffusion of Atto532 becomes apparent. The fluorescence traces were collected from molecules in the stop mode and used for further analysis. In the current condition, there is little adsorption of sample on the inner surface of the cell that scarcely affects the fluorescence intensity, since the histogram of displacement during the time interval of 5 ms calculated from all trajectories of Atto532 in the stop mode shows only a single Gaussian peak, corresponding to molecules diffusing along the capillary in the presence of a small bulk flow (Supporting Information, Figure 2 and supplementary results). The fluorescence traces are detectable until the moment of bleaching, which occurs as a single step, thereby ensuring the observation of a single molecule. Figure 2b presents the traces of Atto532 extracted from the kymograph, demonstrating that the protocol can extend the observation time of free single molecules to several seconds.

Another advantage of our optics is the quantitative acquisition of fluorescence intensity. For confocal microscopy, the illumination intensity and detection efficiency depend substantially on the sample position in the vicinity of the focal spot.[3] In contrast, the current optics uses cylindrical lenses with a long focal length for the homogeneous excitation of the region of about 5 μm × 500 μm (Figure 1b). High power (0.2−0.4 W) of the excitation laser is necessary to illuminate the large volume. The fluorescence intensities are independent of the radial axis of molecules in the capillary cell (Supporting Information, Figure 3b). In contrast, the fluorescence intensities decrease at the periphery of the observation area because the illumination intensity depends on the position of molecules along the axial axis (Supporting Information, Figure 3c). Consequently, the intensities along the axial axis were calibrated using the intensity of the Raman scattering of the solution in the flow path. As presented in Figure 2c, the intensity distribution of fluorescence from Atto532 observed using the current system is nearly Gaussian. The signal-to-noise ratio of the dye at the acquisition time of 5 ms was evaluated as 2.9 (Supporting Information, supplementary results). The standard deviation of the intensity distribution is almost identical to the sum of shot noises originating from signal photons and background photons. Accordingly, the current optics enables us to detect the signal intensity from single molecules quantitatively.

Observation of the Unfolded State Dynamics of a Protein. To explore the energy landscape associated with the

dx.doi.org/10.1021/ja3020555 | J. Am. Chem. Soc. 2012, 134, 11525−11532

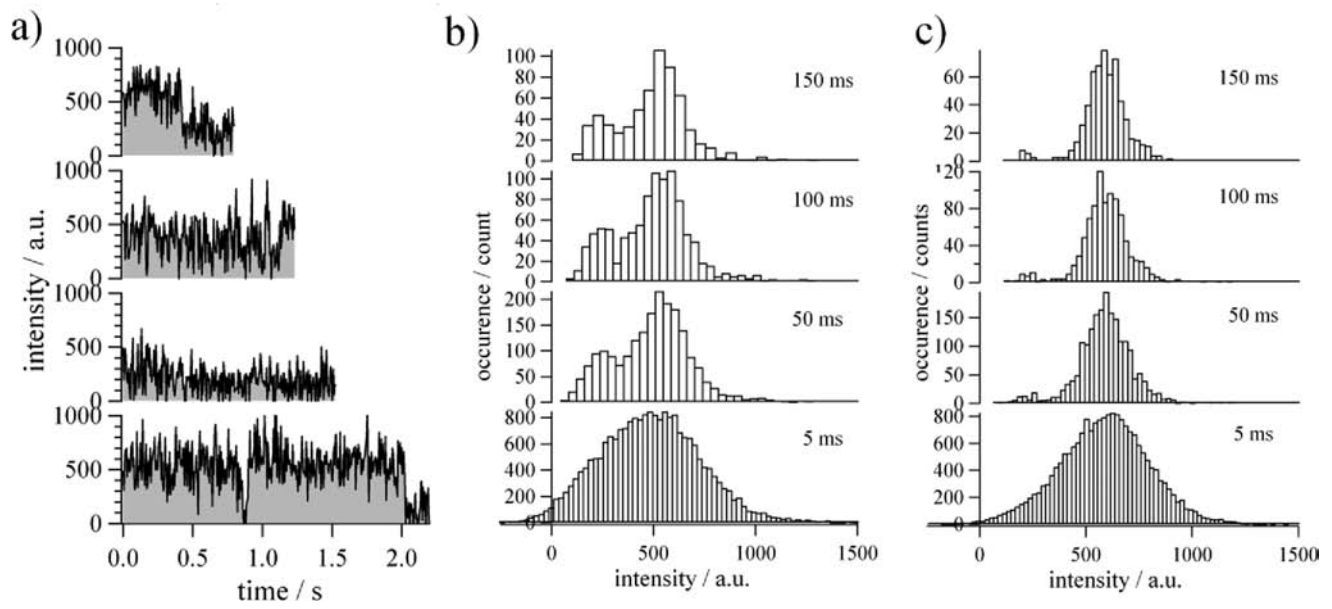


Figure 3. (a) Single-molecule time series for cyt *c*. (b,c) Fluorescence intensity histograms of data averaged for bin sizes of 5, 50, 100, and 150 ms for cyt (b) *c* and (c) Atto532.

unfolded state of a protein, we investigated the unfolded state of cyt *c* from *Pseudomonas aeruginosa* (Figure 3 and Supporting Information, Video 2). Cyt *c* is labeled with Atto532 near the C terminus, with its fluorescence mainly quenched by the heme located near the other terminus. The attachment of Atto532 to cyt *c* does not affect its folding property (Supporting Information, Figures 4 and 5 and supplementary results). In addition to heme, the fluorescence of the labeled iso1-cytochrome *c* was reportedly quenched by amino acid residues such as tryptophan.[34] We confirmed that the heme-cleaved form (apo form) of the labeled cyt *c* from *Pseudomonas aeruginosa* similarly demonstrates the fluorescence quenching at the lower concentration of guanidinium hydrochloride (Supporting Information, Figure 4). Two tryptophan residues, Trp56 and Trp77, are the candidates of the quencher. Accordingly, the fluorescence intensity of the labeled sample is modulated both by heme and aromatic residues, reflecting the global and local conformational changes of cyt *c*.

We examined about 150 traces of cyt *c* with time resolution of 5 ms obtained in a solution containing 2 M guanidinium hydrochloride, which is a midpoint of the equilibrium unfolding transition (Figure 3a and Supporting Information, Figure 4). The intensity histogram of the traces with the time resolution of 5 ms apparently falls into a monomodal distribution (Figure 3b). The standard deviation of the distribution for cyt *c*, 249, is greater than that for Atto532, 214, suggesting the presence of conformational heterogeneity for cyt *c* smeared out by the shot noise (cf. Figure 3b,c). At least two distributions are clearly differentiated in the intensity histograms constructed from the binned traces of cyt *c*, whereas that of Atto532 consists mostly of a monomodal distribution (Figure 3b,c). Furthermore, the averaged autocorrelation function for cyt *c* shows a slower decay than that of Atto532, which indicates the presence of conformational fluctuations. The autocorrelations were fitted with two exponential functions (Supporting Information, Figure 6). We infer that the first phase corresponds to the intensity fluctuation of Atto532, and that the second phase includes the conformational fluctuations in addition to the intensity fluctuation of Atto532 (Supporting Information, Table 2). In addition, the fluorescence lifetime measurements indicate the existence of at least two conformational substates (Supporting Information, Figure 7, Table 3, and supplementary results). Results of these analyses indicate the presence of multiple substates in the unfolded state of cyt *c*.

Local Equilibrium State (LES) Analysis of the Single-Molecule Time Series. To extract the underlying energy landscape of the unfolded state from the single-molecule time series with the limited signal-to-noise ratio, we used LES analysis.[14,22] Conventionally the data analysis has been performed on the whole distribution of the observable (e.g., fluorescence intensity in this study). In contrast, the LES analysis takes into account of the time dependency of the data from the start of the analysis and monitors which state the system visits along the time trace directly (i.e., kinetic connectivity from one state to another).

The key principle of the LES analysis in extracting the states from the time trace is that, when a molecule visits on a local energy basin and is locally equilibrated before going out to another, the distribution of any observable sampled in a certain short time window (τ) should result in a unique distribution inherent to each energy basin, if τ is long enough for the molecule to equilibrate. We define that LESs of the molecule correspond to the local basins on the free energy landscape, in which the molecule achieves the equilibration. Thus, by analyzing the shape of the distribution functions with a proper τ, a set of LESs can be identified from the time series (the collection of all short time distribution functions coincides with the whole distribution). In addition, by counting how often the molecule visits each LES and transitions among different LESs along the time trace, an effective free energy landscape can be constructed.

The actual procedure of the LES analysis is as follows. First, the time series of the observable is transformed into a series of short time distribution functions with a window τ at every instant along the time series. Second, the distribution functions are classified into a set of clusters, in each of which the shape of the distribution functions is "identical" within a statistical bound (the statistical bound is defined to take into account the

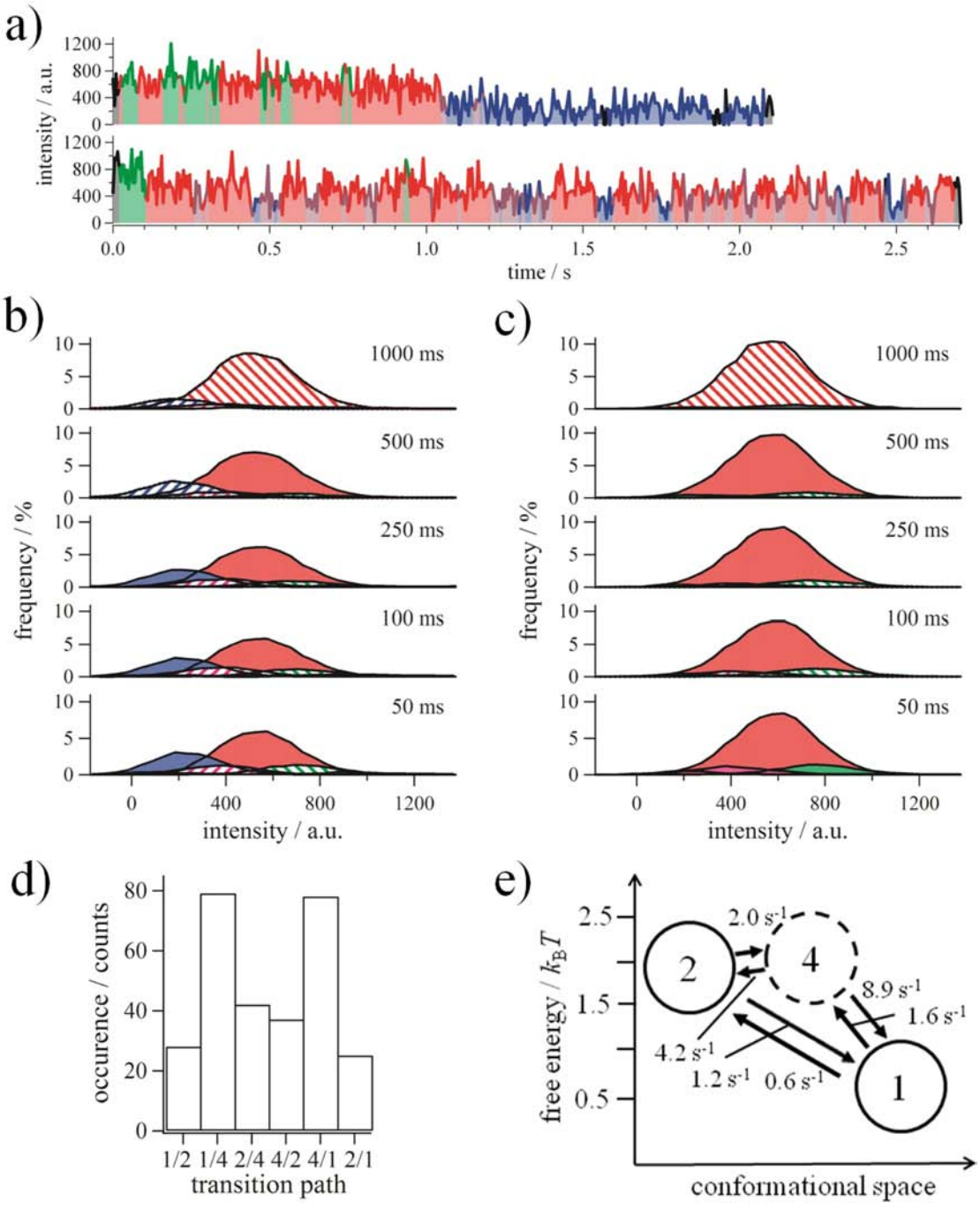


Figure 4. (a) Single-molecule time series obtained for cyt *c*. Red, blue, green, pink, and black data respectively represent states 1, 2, 3, and 4, and UA obtained from LES analysis with τ = 50 ms. (b,c) Intensity histograms of individual states of (b) cyt *c* and (c) Atto532 obtained from LES analysis with various τ. Filled and shaded distributions respectively represent LES and non-LES. (d) The number of transitions between states. The transition path i/j represents the transition from state *i* to state *j*. (e) Free energy landscape for the unfolded state of cyt *c*. Values represent the kinetic rate constants obtained by the analysis of the number of transitions.

finiteness of the number of the sampled data points in τ and the time correlation if it exists within τ[22]). These clusters provide the candidates of LES. Third, the classification of each cluster is performed so that a cluster from which the molecule escapes with an escape time (τ_{esc}) longer than τ is classified as an LES, whereas a cluster whose τ_{esc} is shorter than τ as a non-LES. Some distribution functions might not possess the same shape as any of the clusters representing LES or non-LES, and are classified as "unassigned" (UA). The non-LES and UA represent either the short-lived states with a lifetime shorter than τ, or transient phenomena related to rare events. Fourth, the relative free energy of the states, and the free energy barrier linking two states are evaluated, respectively, by counting the residential probability dwelling in the individual state, and the transition probabilities between the two states.

An important feature of the LES analysis is its sensitivity in differentiating multiple states from a time series of scalar variables. In general, it is difficult to assign fluorescence intensity at each instant to either state, because the intensity for different states with ambiguity caused by noise frequently

dx.doi.org/10.1021/ja3020555 | J. Am. Chem. Soc. 2012, 134, 11525−11532

degenerates. Since the LES analysis assigns states in terms of the shape of the distribution functions, the degeneracy among the states with different shapes can be lifted. Consequently, the LES analysis can characterize the free energy landscape of the unfolded proteins taking full advantage of the time series analysis.[14,22]

LES Analysis of the Single-Molecule Time Series for the Unfolded State of Cyt *c*. We conducted LES analysis for the single-molecule time series of cyt *c* and Atto532 with various τ. Figure 4a shows typical time series of cyt *c*, with colors representing four states obtained through analysis at τ = 50 ms. The distribution at the bottom of Figure 4b presents an intensity histogram of the four states. Two LESs have higher and lower intensities, assigned respectively as state 1 (red, 52 ± 4%) and state 2 (blue, 24 ± 3%). Two non-LESs are assigned as state 3 (green, 11 ± 2%) and state 4 (pink, 7 ± 4%) (Figure 4b, bottom). The number of the state index is presented in descending order of population of the respective states. The additional minor components (black) are regarded as rare states and UA (not visible in Figure 4b, but cf. Figure 4a). In contrast, the distributions of Atto532 at τ = 50 ms possess a major LES assigned as state 1 (red, 75 ± 4%) and two minor LES assigned as states 3 (green, 11 ± 2%) and 4 (pink, 10 ± 2%) (Figure 4c, bottom). The LES analysis for an artificial time series, created by shuffling the data while preserving the total intensity distribution, indicates that the assignment of LES and non-LESs does not arise from experimental noise or some artifacts arising from the finiteness of the time series (Supporting Information, Figure 8 and supplementary results). The results of the LES analysis manifestly demonstrate its ability to uncover states that are sometimes difficult to recognize in the intensity histogram (Figure 3b,c).

One beneficial feature of LES analysis is its capability of obtaining dynamic information about the individual states. In Figure 4b,c and Supporting Information, Figures 9−12, the dynamic information is apparent in terms of different time scales τ. For Atto532, states 3 and 4 merge to the single dominant state, state 1, as τ increases, and the two major states in cyt *c* persist as different states. We therefore interpret that states 3 and 4 for Atto532 are mostly attributable to fluorescence intensity fluctuations, which are expected to result from photochemical processes and/or from diffusion of molecules under a slightly non-uniform illumination (Supporting Information, Figures 13 and 17 and supplementary results). Similarly, state 3 for cyt *c* is likely caused by the intensity fluctuations of the dye, rather than the conformational change of cyt *c*, because the population, the distribution, and τ_{esc} for state 3 of cyt *c* are similar to those in Atto532. In contrast, the presence of state 2 in cyt *c* is distinct from Atto532, suggesting that the state reflects the conformational substate in the unfolded state of cyt *c*. As τ increases, state 2 converts from LES to non-LES for cyt *c* with its decrease in population. State 4 of cyt *c* might represent the conformational substate of the unfolded protein and the photodynamical processes of the dye, since the population of state 4 of cyt *c* does not diminish even at a longer τ such that state 4 for Atto532 does (Supporting Information, Figure 9). Consequently, based on the LES analysis, we obtained detailed dynamical information related to the unfolded state of cyt *c*.

To obtain insight into the property of the unfolded states, we compared the result of cyt *c* to that of Atto532. State 1 is interpreted as an extended unfolded state prevented from quenching by quenchers inside the protein. This is because the distribution of state 1 resembles that of state 1 in Atto532, indicating the absence of collisions between Atto532 and heme located respectively near the C- and N-termini. Furthermore, we investigated the unfolded state of cyt *c* in the apo form labeled with Atto532 by the same procedure, and identified multiple states whose intensity distributions are similar to those identified in the holo form (Supporting Information, Figures 14 and 15 and supplementary results). The result suggests that states 1 and 2 of cyt *c* are different in the conformations around heme and amino acids such as tryptophan. We suggest that state 2 is an extended unfolded state with the N-terminus distant from the C-terminus and with a structured region near Atto532 whose fluorescence is quenched by tryptophan. It is noted that we cannot detect single-molecule fluorescence from the native state of cyt *c*, because of the quenching of fluorescence intensity by heme.

Free Energy Landscape of the Unfolded State of Cyt *c*. For further investigation of the stability and the kinetic connectivity between the states, the free energy landscape of unfolded cyt *c* is constructed from the single-molecule traces. The (relative) free energy of state *i*, G_i, is calculated using $G_i = -k_B T \log(N_i/N)$, where k_B, T, N_i, and N respectively represent Boltzmann's constant, absolute temperature, the number frequency to visit state *i* ($N_i = \sum_j N_{i\to j}$, where $N_{i\to j}$ is the number of transitions from state *i* to state *j*), and the total number of transitions ($N = \sum_{i,j} N_{i\to j}$). Based on Kramers theory, the free energy of the barrier linking state *i* and state *j*, G_{ij}, is calculated as

$$G_{ij} = -k_B T \ln\left(\frac{k_{i\to j}}{A}\right) = -k_B T \ln\left(\frac{N_{i\to j}}{AN}\frac{1}{\tau}\right)$$

where $k_{i\to j}$, A, and τ respectively denote the rate constant from state *i* to state *j*, the pre-exponential factor, and the time window of the LES analysis, which corresponds to the time scale to count the transitions (Supporting Information, supplementary methods). We use 1 μs^{-1} as A^2. $N_{i\to j}$ is calculated using a sequence of states assigned by the LES analysis with τ = 50 ms (Figure 4d). Results of binomial tests showed that the number of transitions from one state to another for any pair of the states is not significantly different from that for the inverse reaction, which validates the detailed balance at the single-molecule level with the chosen time scale (Supporting Information, supplementary methods). The estimated G_i and G_{ij} are listed in Supporting Information, Table 4. In Figure 4e, the free energy landscape of the unfolded cyt *c* is presented, demonstrating the existence of a parallel pathway between state 1 and state 2.

Using conventional ensemble techniques such as equilibrium and kinetic re/unfolding experiments, it has been very difficult to determine the connectivity between states for systems including more than two states. This is because the assignment of the transitions between the states is in most cases a laborious process, causing difficulty in determining the kinetic models. In contrast, our method based on the analysis of single-molecule time series can directly reveal the energy landscape of the system with the connectivity between the states. We emphasize that the current result is one of the first demonstrations of the complexity for the energy landscape of proteins in the unfolded state supported by the experimental evidence.

CONCLUSIONS

We developed the flow-and-stop protocol of SMFS using new spherical optics and a pressure-regulating pump. The protocol enables observation of single molecules for more than several seconds without tethering the sample to an optical surface. Furthermore, the fluorescence intensity from single molecules can be obtained quantitatively using this simple protocol. Numerous single-molecule traces are easily gathered. The developed protocol is readily applicable to investigations of numerous processes, including protein folding. As observed in the single-molecule investigation of the NADH:flavin oxidoreductase complex,[19] the protein might feel a rugged energy landscape transiently, whose morphological feature is dependent on the time scale. Application of LES analysis to the time series demonstrated its feasibility in investigating how a biological macromolecule experiences the energy landscape at the single-molecule level with different time scales. With these features first achieved in the current experimental protocol, long-term observation of free single proteins and LES analysis are expected to provide a powerful means to explore the dynamic mechanisms of the molecular machinery associated with protein folding and functions.

EXPERIMENTAL SECTION

Materials. The S80C mutant of cytochrome *c*551 from *Pseudomonas aeruginosa* was prepared using a previously described method.[36] The S80C mutant was used to label with Atto532. Guanidinium hydrochloride was purchased from Nacalai Tesque Inc. (Kyoto, Japan), and Atto532 maleimide from ATTO-TEC GmbH (Siegen, Germany). All chemicals were of the highest grade available, and were used without further purification. A spherical mirror system was designed and constructed (Optonica, Kyoto, Japan). Ion-exchange chromatography was used for the preparation of cyt *c* labeled with Atto532.

Single-Molecule Measurements. Buffer containing 50 mM citric acid−NaOH, 2 M guanidinium hydrochloride, and 2 mM Trolox was adjusted to pH 4.7. Buffer solutions were passed through a filter with 0.22-μm pores and were used without degassing. To prevent sample adsorption to the capillary cell, 0.001% polyoxyethylene sorbitan monolaurate (Surfact-Amp20, Pierce Chemical Co., Rockford, IL) was added after filtering.

The capillary cell and the cell holder design were adopted from an earlier study[7] with some modification, enabling the quick introduction of fresh solutions and minimization of the nonspecific absorption of samples inside the flow path. We used a fused-silica capillary with inner diameter of 5 μm (TSP005375, Polymicro Technologies LLC, Phoenix, AZ). Although the system performs best with rectangular capillaries, as presented in Figure 1b, a spherical capillary was used to obtain data with optical attachments designed to correct image deformation. Fittings, tubes, and ferrules were purchased from Upchurch Scientific Inc. (Harbor, WA) and GL Science Inc. (Tokyo, Japan).

The excitation laser (Juno5000, Showa Optronics Co., Ltd., Tokyo, Japan) illuminated the observation area (5 μm × 500 μm) of the capillary from the side using a cylindrical lens (f = 40 mm) (Figure 1a,b). The excitation power was 440 mW. The spherical mirror system was used to collect the fluorescence photons on EM-CCD (amplification ×1380, iXon; Andor Technology, Belfast, Northern Ireland). Two notch filters (Notch-plus, Kaiser Optical Systems, Inc., Ann Arbor, MI, and NF01-532U-25, Semrock Inc., Rochester, NY), a red dichroic filter (Edmund Optics Japan, Tokyo, Japan), and a long-pass filter (XF593, Asahi-bunko, Tokyo, Japan) were used.

For single-molecule measurements, we used 1−10 pM solutions of the labeled cyt *c* or ca. 100 fM solution of Atto532. The sample flow was maintained at ca. 0.3 nL/min by controlling the inlet and outlet pressures using a N_2-gas regulated pump (MFCS-4C, Fluigent, Paris, France). To stop the flow in the capillary cell, we reduced the inlet pressure suddenly. Repeating the flow-and-stop protocols roughly one hundred times produced sufficient quantities of single-molecule traces. Finally, we extracted the time series of the fluorescence intensity of single molecules from the consecutive images detected by EM-CCD. Details of data extraction are described in Supporting Information, methods.

LES Analysis. Detailed procedures for the LES analysis were presented in earlier reports.[14,22] The minimum value of the time window τ of the analysis, 50 ms, was selected to have the lowest minimal number (=10) of data points to define the distribution function. If we use data points less than 10, the differentiation of the four states for cyt *c* becomes difficult. We used the Kantorovich metric as the measure of similarity of the distribution functions. It is calculated by integrating the differences between the cumulative distribution functions. To identify a cluster of the distribution functions having the same shape within a statistical bound, we first defined an appropriate radius of the maximally populated cluster in the space of Kantorovich metric in consideration of the finite number of data points of the cluster and the autocorrelation time scale in the cluster. The same procedure was repeated cluster-by-cluster until the data points of a cluster became fewer than 100. Thus, we identified clusters with more than 100 data points and assigned them as states. This number corresponds to ca. 0.5% of all data sets. The minimum percentage of the states observed in this study was 2%. All residual distribution functions that cannot be categorized as clusters or which are scattered through the Kantorovich metric space were classified as UA. For clusters with more than 100 data points, we estimated the escape time (τ_{esc}) for single molecules using a single exponential fitting of the survival functions for $t \geq \tau$. For τ_{esc} longer than τ, we assigned the cluster as an LES. We confirmed that the time scale of the autocorrelation in a cluster assigned as an LES was always much shorter than τ. In contrast, for τ_{esc} shorter than τ, we assigned such a cluster as a non-LES. To confirm the validity of the present assignment of the states with the time window of 50 ms, we scrutinized the same analysis for surrogate data with randomly shuffled order (Supporting Information, supplementary results). To estimate the standard errors of parameters of the LES analysis, the time series data obtained in one condition were split into three data sets, and each of which was analyzed by the same LES method. The standard errors were obtained using the set of the parameters calculated.

ASSOCIATED CONTENT

Supporting Information

Detailed description of the spherical mirror system; preparation of the apo form of cyt *c* labeled with Atto532; details of data analysis; experimental information for diffusions of single molecules in the capillary cell, estimation of noise in single-molecule traces, effects of the labeling of Atto532 to cyt *c*, single-molecule experiment for the apo form of cyt *c*, importance of sequences of single-molecule traces in the LES analysis, origin of minor states observed in single-molecule traces of the dye, LES analysis of cyt *c* from different species, dynamics within individual states, and fluorescence lifetime measurements; videos showing long-time observation of Atto532 and cyt *c*, obtained using the flow-and-stop protocol. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

st@tagen.tohoku.ac.jp

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

K. Kamagata, T. Komatsuzaki, and S. Takahashi thank Prof. Chun Biu Li, Dr. Irina V. Gopich, Prof. Akihiko Ishijima, and Dr. Hiroyuki Oikawa for helpful discussions, Mr. Shion Ando for bead experiments, Dr. Tetsuichi Wazawa and Prof. Makoto Suzuki for fluorescence lifetime measurements, and Dr. Kazuhiro Iwashita for mass spectroscopy. This work was partially supported by JST/CREST and Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to K. Kamagata, T. Komatsuzaki, Y. Sambongi, and S. Takahashi).

■ REFERENCES

(1) Talaga, D. S.; Lau, W. L.; Roder, H.; Tang, J.; Jia, Y.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13021−13026.
(2) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743−747.
(3) Rhoades, E.; Gussakovsky, E.; Haran, G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3197−3202.
(4) Kuzmenkina, E. V.; Heyes, C. D.; Nienhaus, G. U. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15471−15476.
(5) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528−1533.
(6) Nettels, D.; Gopich, I. V.; Hoffmann, A.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2655−2660.
(7) Kinoshita, M.; Kamagata, K.; Maeda, A.; Goto, Y.; Komatsuzaki, T.; Takahashi, S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10453−10458.
(8) Nettels, D.; Hoffmann, A.; Schuler, B. *J. Phys. Chem. B* **2008**, *112*, 6137−6146.
(9) Noji, H.; Yasuda, R.; Yoshida, M.; Kinosita, K., Jr. *Nature* **1997**, *386*, 299−302.
(10) Arai, Y.; Iwane, A. H.; Wazawa, T.; Yokota, H.; Ishii, Y.; Kataoka, T.; Yanagida, T. *Biochem. Biophys. Res. Commun.* **2006**, *343*, 809−8015.
(11) Zhao, Y.; Terry, D.; Shi, L.; Weinstein, H.; Blanchard, S. C.; Javitch, J. A. *Nature* **2010**, *465*, 188−193.
(12) Uemura, S.; Aitken, C. E.; Korlach, J.; Flusberg, B. A.; Turner, S. W.; Puglisi, J. D. *Nature* **2010**, *464*, 1012−1017.
(13) *Single Molecule Biophysics: Experiments and Theories*; Advances in Chemical Physics *146*; Komatsuzaki, T., Takahashi, S., Kawakami, M., Yang, H., Silbey, R., Eds.; John Wiley and Sons: New York, 2011.
(14) Baba, A.; Komatsuzaki, T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19297−19302.
(15) Rao, F.; Caflisch, A. *J. Mol. Biol.* **2004**, *342*, 299−306.
(16) Hori, N.; Chikenji, G.; Berry, R. S.; Takada, S. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 73−78.
(17) Gfeller, D; De Los Rios, P; Caflisch, A.; Rao, F. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1817−1822.
(18) Li, C. B.; Yang, H.; Komatsuzaki, T. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 536−541.
(19) Yang, H.; Luo, G.; Karnchanaphanurach, P.; Louie, T. M.; Rech, I.; Cova, S.; Xun, L.; Xie, X. S. *Science* **2003**, *302*, 262−266.
(20) Schuetz, P.; Wuttke, R.; Schuler, B.; Caflisch, A. *J. Phys. Chem. B* **2010**, *114*, 15227−15235.
(21) Miyazaki, M.; Harada, T. *J. Chem. Phys.* **2011**, *134*, 085108.
(22) Baba, A.; Komatsuzaki, T. *Phys. Chem. Chem. Phys.* **2011**, *13*, 1395−1406.
(23) Deniz, A. A.; Laurence, T. A.; Dahan, M.; Chemla, D. S.; Schultz, P. G.; Weiss, S. *Annu. Rev. Phys. Chem.* **2001**, *52*, 233−253.
(24) Funatsu, T.; Harada, Y.; Tokunaga, M.; Saito, K.; Yanagida, T. *Nature* **1995**, *374*, 555−559.
(25) Chung, H. S.; Louis, J. M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 11837−11844.
(26) Kuzmenkina, E. V.; Heyes, C. D.; Nienhaus, G. U. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15471−15476.
(27) Cohen, A. E.; Moerner, W. E. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4362−4365.
(28) Cohen, A. E.; Moerner, W. E. *Opt. Exp.* **2008**, *16*, 6941−6956.
(29) Goldsmith, R. H.; Moerner, W. E. *Nature Chem.* **2010**, *2*, 179−186.
(30) Chiu, D. T.; Lorenz, R. M.; Jeffries, G. D. M. *Anal. Chem.* **2009**, *81*, 5111−5118.
(31) Levy, S. L.; Craighead, H. G. *Chem. Soc. Rev.* **2010**, *39*, 1133−1152.
(32) Leslie, S. R.; Fields, A. P.; Cohen, A. E. *Anal. Chem.* **2010**, *82*, 6224−6229.
(33) Nie, S.; Chiu, D. T.; Zare, R. N. *Science* **1994**, *266*, 1018−1021.
(34) Choi, J.; Kim, S.; Tachikawa, T.; Fujitsuka, M.; Majima, T. *Phys. Chem. Chem. Phys.* **2011**, *13*, 5651−5658.
(35) Gianni, S.; Travaglini-Allocatelli, C.; Cutruzzolà, F.; Brunori, M.; Shastry, M. C.; Roder, H. *J. Mol. Biol.* **2003**, *330*, 1145−1152.
(36) Kobayashi, Y.; Sonoyama, T.; Takeda, T.; Sambongi, Y. *Biosci. Biotechnol. Biochem.* **2009**, *73*, 1227−1229.