Chemical Physics 422 (2013) 229–237

Contents lists available at SciVerse ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

Measuring ultrafast protein folding rates from photon-by-photon analysis of single molecule fluorescence trajectories

Hoi Sung Chung *, Troy Cellmer, John M. Louis, William A. Eaton *

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA

ARTICLE INFO

Article history:
Available online 14 August 2012

Keywords:
Single-molecule fluorescence
FRET
Maximum likelihood analysis
Dye blinking
Ultrafast protein folding rates
Photon trajectories
Villin subdomain

ABSTRACT

Folding and unfolding rates for the ultrafast folding villin subdomain were determined from a photon-by-photon analysis of fluorescence trajectories in single molecule FRET experiments. One of the obstacles to measuring fast kinetics in single molecule fluorescence experiments is blinking of the fluorophores on a timescale that is not well separated from the process of interest. By incorporating acceptor blinking into a two-state kinetics model, we show that it is possible to extract accurate rate coefficients on the microsecond time scale for folding and unfolding using the maximum likelihood method of Gopich and Szabo. This method yields the most likely parameters of a given model that can reproduce the observed photon trajectories. The extracted parameters agree with both the decay rate of the donor–acceptor cross correlation function and the results of ensemble equilibrium and kinetic experiments using nanosecond laser temperature jump.

Published by Elsevier B.V.

1. Introduction

The power of single molecule fluorescence spectroscopy is the ability to investigate distributions in molecular behavior for intrinsically heterogeneous systems. One such system is protein folding, in which theory predicts many different sequences of structural changes in the pathways that connect the folded and unfolded states [1]. The α-helical, 35-residue villin subdomain (Fig. 1) is currently the most extensively studied protein by experiment, theory, and simulations (see bibliography in Supplementary Material). The reasons are that it has equilibrium properties of a much larger single-domain protein [2,3], is among the fastest folding proteins [4,5], and exhibits unusual kinetics such as a denaturant independent relaxation rate [6] and an apparent increase in the internal friction with temperature in a Kramers description of the barrier crossing [7]. Our ultimate goal for single molecule experiments on this protein is to observe the distribution of transition paths – a uniquely single molecule property. Such measurements would provide a very demanding test of the accuracy of the mechanisms found in molecular dynamics simulations, but represent a major challenge since transition paths have not been observed for any molecular system in the condensed phase. New and sensitive tests of simulation are important because, if accurate, everything one would ever want to know about the folding mechanism of a particular protein is contained in a sufficiently long atomistic trajectory [8–10] or Markov state modeling of many short trajectories [11].

Studies of the villin subdomain are particularly challenging because of its very rapid kinetics, with folding times of the wild-type on the order of 5–50 μs [3,6,12]. Until quite recently the time resolution in single molecule FRET experiments [13] has been limited by the bin time of the measurement, which is usually 1–10 ms at the moderate illumination intensities that have been employed to avoid photochemical problems such as bleaching and blinking of the dyes [14]. For residence times much longer than these bin times, two distinct peaks will appear in a histogram of the FRET efficiencies and the rate coefficients can be simply determined from the FRET efficiency trajectories as the reciprocal of the average residence times (also called waiting or dwell times). When the average residence time is much shorter than the bin time, however, transitions occur within the bins, and rate coefficients cannot be so easily determined. To extract kinetics from trajectories under these circumstances, Gopich and Szabo have developed maximum likelihood methods for analyzing photon trajectories without binning [15]. We previously used this method on the α-helical protein α_3D, and showed that the method produces accurate folding times of ~1 ms, as judged by the excellent agreement of the sum of the rate coefficients obtained from the maximum likelihood analysis for this two-state system and the decay of the donor–acceptor cross-correlation function [16]. For the faster-folding villin subdomain, blinking occurs on a similar time scale to folding and unfolding, making it problematic to distinguish dye blinking from protein dynamics. Here we show how to extract the kinetics of folding and unfolding from the data that is complicated by blinking, and

* Corresponding authors.
E-mail addresses: chunghoi@niddk.nih.gov (H.S. Chung), eaton@helix.nih.gov (W.A. Eaton).

0301-0104/$ - see front matter Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.chemphys.2012.08.005

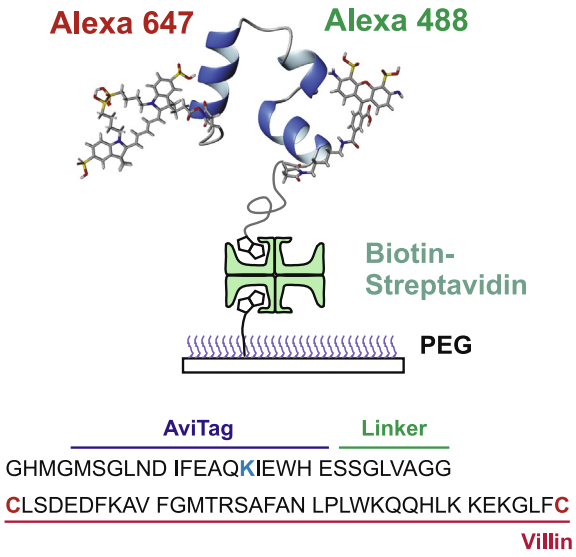

Fig. 1. Schematic of immobilized folded villin subdomain showing donor (green-emitting) and acceptor (red-emitting) fluorophores. Protein molecules were attached to a polyethyleneglycol-coated glass surface via a biotin-strepavidin–biotin linkage. Dyes were attached to cysteine residues (red) at the N- and C-termini of villin and a biotin molecule to the lysine residue (blue) within the AviTag sequence.

confirm the accuracy of the extracted rate coefficients by comparison with both the donor–acceptor cross-correlation functions and the results of ensemble equilibrium and kinetic experiments.

2. Methods

2.1. Materials

The structure of the 35-residue long villin headpiece subdomain (PDB 1YRF), together with the amino acid sequence of the construct used in the single molecule experiments, is shown in Fig. 1 [17]. For the laser temperature jump measurements, the chemically-synthesized 35-residue sequence at a purity >98% was purchased from California Peptide (Napa, CA). For single molecule studies, the complete sequence in Fig. 1 (avitag + linker + cysteine + villin subdomain sequence + cysteine) was expressed in *Escherichia coli*. Details of the expression, purification and dye labeling are given in the Supplementary Material.

2.2. Nanosecond laser temperature-jump measurements

Ensemble kinetic measurements were carried out on solutions containing 300 μM of the 35-residue sequence (Fig. 1) or N-acetyl tryptophanamide (NATA) using a nanosecond-laser-temperature-jump instrument previously described [5]. All solutions were buffered to pH = 4.9 with 20 mM sodium acetate and flowed through the illuminated region to eliminate effects of tryptophan photodamage. Temperature jumps of ~7–10 °C to a final temperature of 20 °C were generated by Raman shifting pulses of a Nd:YAG fundamental at 1064 nm to 1560 nm using D_2 gas. To ensure a consistent temperature jump in the presence of changing solvent conditions, the temperature jump was calibrated using NATA. A frequency-doubled Kr laser with an output at 284 nm was used to excite Trp fluorescence. In each experiment, four to eight traces of 512 laser shots were collected. Rate constants and amplitudes were calculated by a least-squares exponential fit of the data at times >3 μs and baseline from a NATA trace.

2.3. Single molecule spectroscopy

Single molecule FRET experiments were performed using a confocal microscope system (MicroTime200, Picoquant). The CW mode of a dual mode (CW/pulsed) 485 nm diode laser (LDH-D-C-485, PicoQuant) was used to excite donor dyes (Alexa Fluor 488) through an oil-immersion objective (PlanApo, NA 1.4, ×100, Olympus). Donor and acceptor (Alexa Fluor 647) fluorescence was collected by the same objective, split into two channels, and focused through optical filters (ET525/50m for the donor and E600LP for the acceptor, Chroma Technology) onto photon-counting avalanche photodiodes (SPCM-AQR-15, PerkinElmer Optoelectronics). Additional details for the optical setup and single molecule experiments can be found elsewhere [18,19].

In the free diffusion experiment, where the fluorescence bursts of single freely diffusing molecules entering into the focal volume were measured, the protein concentration was 40 pM in 50 mM HEPES buffer (pH 7.6). A surfactant, 0.01% Tween-20 (Thermo Scientific) was used to prevent sticking of proteins on a glass surface. The illumination power of the laser was 29 μW.

In the immobilization experiment, protein molecules were immobilized on a biotin-embedded, polyethyleneglycol (PEG)-coated glass coverslip (Bio_01, Microsurfaces Inc.) via a biotin (surface)-streptavidin–biotin (protein) linkage. To reduce dye bleaching and blinking, 2 mM L-ascorbic acid (A92902, Sigma) and 2 mM methyl viologen (856117, Sigma) were added [20] to the 50 mM HEPES buffer (pH 7.6) solution.

Fluorescence trajectories were collected using an automated data collection scheme. An area of 10 × 10 μm^2 was raster scanned at low intensity and the location of molecules was determined. To ensure the single molecule detection, molecules immobilized too close to one another were identified by an image larger than a threshold size and were excluded. The piezo-controlled stage was then moved to locate each molecule and the trajectory was collected at high illumination intensity. The laser was turned off during movement of the stage to prevent photobleaching. After the collection of trajectories for all identified molecules was completed, the procedure was repeated for the next 10 × 10 μm^2 area. Before raster scanning each area, the focus along z-axis (perpendicular to the surface) was set at the position with minimum variance of the reflected image from the surface recorded by a CCD camera. The illumination intensity for the raster scan (0.4 kW/cm^2, the average power entering the microscope was 1.3 μW) and for the trajectory collection (4 kW/cm^2) were adjusted by inserting a neutral density filter (OD = 1) in the laser path using a home-built mechanical shutter.

3. Results

Fig. 2 shows the ensemble equilibrium and kinetic data at 20 °C on the 35-residue sequence that was not labeled with the donor and acceptor fluorophores. The population versus GdmCl concentration in Fig 2(b) was obtained from a two-state fit to the circular dichroism and fluorescence unfolding curves. The relaxation rate, measured from the time course of the tryptophan quantum yield in nanosecond laser temperature jump experiments, is nearly independent of GdmCl concentration at ~50 ms^{-1} in the range 2–4 M GdmCl (Fig. 2(c)). Since we did not anticipate a major effect from either the attachment of the fluorophores or the additional residues of the complete construct in Fig. 1, the residence times in single molecule experiments were expected to be sub-100 μs.

To measure single molecule equilibrium folding/unfolding trajectories using FRET spectroscopy, the recombinant villin construct was labeled with two fluorescent dyes, Alexa Fluor 488 as donor and Alexa Fluor 647 as acceptor, and immobilized on a PEG-coated