Edited by Alan Fersht, University of Cambridge, Cambridge, United Kingdom, and approved May 18, 2009 (received for review February 2, 2009)

Transition paths are a uniquely single-molecule property not yet observed for any molecular process in solution. The duration of the rate coefficients and the latter their ratio. The unique transition paths is the thy traction of the time in an equilibrium information in a single-molecule experiment is contained in the single-molecule trajectory when the process actually harmons. Here ungle-molecule trajectory when the process actually happens, nere, we report the determination of an upper bound for the transition path time for protein folding from photon-by-photon trajectories. about folding and unfolding is contained in these so-called transition paths (Fig. 1), which can only be observed for single FRET trajectories were measured on single molecules of the dvelabeled, 56-residue 2-state protein GBI, immobilized on a glass sur face via a biotin-streptavidin-biotin linkage. Characterization of indi vidual emitted photons by their wavelength, polarization, and absolute and relative time of arrival after picosecond excitation allowed the determination of distributions of FRET efficiencies, donor and acceptor lifetimes, steady state polarizations, and waiting times in the folded and unfolded states. Comparison with the results for a malacular charved that immehilization has no do tectable effect on the structure or dynamics of the unfolded protein and only a small affect on the folding/unfolding kinetics. Analysis of and only a small effect on the foliality unfolding kinetics. Analysis of the photon-by-photon trajectories yields a transition path time <200 $\mu$ s, >10,000 times shorter than the mean waiting time in the unfolded state (the inverse of the folding rate coefficient). Szabo's theory for diffusive transition paths shows that this upper bound for the the Kramers preexponential factor for the rate coefficient, and predicts that the transition path time is remarkably insensitive to the tions among the distances (14). In this work we take a major step ing rate, with only a 2-fold difference for rate coefficients that

Alexa 488 | fluorescence | ERET | maximum bladbood function | postein CR1

detailed description and understanding of mechanisms of ehavior (1-7). A 2-state protein has only 2-populations of mole librium and at all times in kinetic experiments-folded nd unfolded. In ensemble folding experiments kinetics are studied by rapidly changing solution conditions, e.g., the temperature or denaturant concentration, and monitoring the relaxation of the 2 populations to their new equilibrium ratio with probes such as fluorescence, circular dichroism or infrared spectroscopy. Single molecule kinetics, however, can be studied at equilibrium. As can be seen from the schematic of a trajectory in Fig. 1, the dynamical ature of equilibrium is dramatically demonstrated when observing forster resonance energy transfer (FRET) in a single-molecule prescence experiment. There are fluctuations due to shot noise nean value in each state, interrupted by what appear to be eous jumps in FRET efficiency signaling folding or unfolding. The residence or waiting times in each state are exponenitally distributed, with the mean time in the unfolded and folded segments of the trajectories corresponding to the inverse of the folding and unfolding rate coefficients, respectively.

Rate coefficients can, albeit with assumptions, be much more easily obtained from a combination of ensemble kinetic and This article contains supporting information online at www.pnas.org/cgi/contain  F. Chen et al., infect. immun. 65, 1626 (1997).
 C. N. Paco, G. R. Grirroley, J. M. Scholtz, J. Mol. Chem. 28-4, 15-85 (AUV). 33. H. Li. & D. Rebertson, T. H. lengen, Proteins 61, 704

Administrator We think A Brown and T Blocky official reading of the manuscript; the staff of beamine 9-2 at the Starford Synchrotron Radiation Lightcource (SSRI) A. M. Kruet, and J. Temblay for excellent nor and R. Uddington and A. Robins for

Infectious Diseases (MAID), NIH, Department of Health and Human Services (HHS), under award number USA 41057159 (C.R.S.). Atomic coordinates and drudum on for the WHH bound M-PTC, M-PTC, and BoAT/M on codes 3VDA, 3VDB, and 3VDC, respectively

complex delikery compositions," that was filed in New mber of 2011. BoAT availability is subject to the restrictions at apply to HHS select agents and NAID Category Supportion Online Material

MIL MILITARE AND THE TOTAL OF T

Single-Molecule Fluorescence **Experiments Determine Protein Folding Transition Path Times** 

Hoi Sung Chung,\* Kevin McHale, John M. Louis, William A. Eaton\*

The transition path is the tiny fraction of an equilibrium molecular trajectory when a transition occurs as the free-energy barrier between two states is crossed. It is a single-molecule property that contains all the mechanistic information on how a process occurs. As a step toward observing transition paths in protein folding, we determined the average transition-path time for a fast- and the FRET efficiency is midway between the a slow-folding protein from a photon-by-photon analysis of fluorescence trajectories in single-molecule Förster resonance energy transfer experiments. Whereas the folding rate coefficients differ by a factor of 10,000, the transition-path times differ by a factor of less than 5, which shows to the average (parsition-path time, \$772) (Fig. 4A). that a fast- and a slow-folding protein take almost the same time to fold when folding actually happens. A very simple model based on energy landscape theory can explain this result.

changes (1). The distribution of those fiding path—
the graph of the statement of those fiding path—
the B1 immunoglobalin-binding domain of protenting from Streptococcup (9). In contrast to a nie lecular dynamics simulations (2-6). Information coefficient, which measures the frequency of a on pathway distributions from experiments must transition, the transition path time is the duption come from measurements on single molecules,
of a successful barrier-crossing event (Fig. 1).
because only average properties are obtained in
The strategy used in this study is to illumibecause only average properties are obtained in tained in what amears to be an instantaneous ump between the two states, called the transition bilized WW domain and protein GB1 molecules v<sub>ac</sub> and v<sub>fin</sub> are vectors that describe the state

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Midney Diseases, National Institutes of

Theory predicts that folding mechanisms are - fast-folding, all-8 protein D9-residue formin-binding heterogeneous, so that an individual un- protein (FBP) WW domain] shown to be two-state Gildred molecule can self-assemble to form in ensemble studies (7, 8), as well as a marked to s biologically active, folded structure by means reduced upper bound compared with our previ of many different sequences of conformational our study for the 56-residue of 8 protein GRU

experiments on the large ensemble of molecules rate dyel-the-field protein molecules at very high or acceptor), and  $\tau_1$  is a time internal between in bulk experiments. A single-molecule, equilibinity intensities to increase the number of detected phorium protein folding-unfolding trajectory is illustons per transition path, to discard the majority S4B (II). The photon color matrix F deper traind in Fig. 1, as monitored by Förster resonance of photons from the less-interesting segments of on the color of a photon as F(acceptor) = E and energy transfer (RRIT) spectoscopy, and its rela-tion to the free-energy barrier as it crosses be-lyze the transition region with a maximum like-with elements that are FRET efficiencies of tween the folked and unfolded states is shown.

The most interesting part of the imjectory is contransition path.

The most interesting part of the imjectory is contransition path. Photon trajectories were measured for immo- that are photon count rates of the three states.

path, which contains all of the information on the with donor and acceptor fluoreobores attached (folded or unfolded) at the beginning and the secharism of folding and unfolding. The first to cysteines incorporated into the proteins (Fig. 2), end of the trajectory. Practically, log-lik sten toward observing transition matis in protein. In these trajectories, two properties of each place, functions were calculated, and the total log-likebiding, which we report here, is the determination ton were recorded—the color, either donor green lihood function of all trajectories was calcu of its memore duration (targetion-each time) for a or acceptor red, and the absolute time of arrival lated by summing the los-likelihood functions the binned fluorescence and photon trajectories,

chundridintidath.cor HSC): nonatheliuth.gov WA.E) of two states. The photon trajectories were ex-

analyzed using the Gorich Szabo maximum like-For a given model, the Gopich-Scabo meth-

can most accumtedy reproduce the photon trajectories (Fig. 3). We adopt a one-step model for the transition path, which may be viewed as the FRET efficiency changes along the path. This nicture can be represented in a kinetic model for a two-state system with a finite transition path

are the same, and therefore, the mass = 1/2 The likelihood function for the 4th photon

 $L_j = \mathbf{v}_{fn}^T \prod_{i=1}^{n} \{ \mathbf{n} \mathbf{F}(c_i) \exp [(\mathbf{K} - \mathbf{n}) \mathbf{t}_i] \} \mathbf{n} \mathbf{F}(c_1) \mathbf{v}_{int}$ 

trajectory, c. is the color of the #h photon (donor

transitions between states are dearly resolved in  $(\ln L = \sum \ln L_j)$  of individual trajectories that and the FRET efficiency distributions (Fig. 3, C contain a transition between folded and unfolded To whom correspondence should be addressed. E-mail: and D) are bimodal, which indicates the presence states. In the likelihood function L, \u03c4s is the only

Chemical Physics 422 (2013) 229-237

21 September 2011: accopted 12 January 2012

tracted from the region near the transitions and

od calculates the parameters of the model that

folded and unfolded states  $[E_N = (E_V + E_V)Y2]$ In this model, the lifetime of S (t<sub>n</sub>) correspond S has the property of a transition state, because the rate mellicients from S to E and S to II (b.)

Here, K is the rate matrix [equation S6 (11)] containing the three rate coefficients &... k...

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 Changesi Physics, National Institute of Disbette and Directive and Kidney Disputes, National Institutes of Health, Rethonds, MD 2010, 0520, USA

ARTICLE INFO ABSTRACT Rolding and unfolding rates for the ultrafact folding villin subdormain were determined from a photon-by-photon analysis of fluorescenic trajectories in single molecule REET experiments. One of the obstacles to measuring fast kinetics in single molecule fluorescenics experiments to blinking of the fluorophores on a Article Matery: Available online 14 August 2012

he power of single molecule fluorescence spectroscopy is the

ability to investigate distributions in molecular behavior for intrin-

sically 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

rently 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 sin-gle-domain protein [2,3], is among the fastest folding proteins

[4,5], and exhibits unusual kinetics such as a denaturant indeper

dent 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

ovide a very demanding test of the accuracy of the mechanism

found in molecular dynamics simulations, but represent a major

round in molecular dynamics simulations, our 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

on this protein is to observe the distribution of transition paths -

Krywords: Single-molecule fluorescence MET Maximum likelihood analysis

lation function and the results of ensemble equilibrium and kinetic experiments using rangescond later

trajectory (8-10) or Markov state modeling of many short

immicale intain to see will apparate or room one process or interest, by incorporating acceptor or intering to we state kinetics model, we show that it is possible to estrate accurate rate coefficients on the micro-emend time scale for folding and unfolding using the maximum likelihood method of Gojich and Sabio, his method yields the most likely parameters of a given model that can reproduce the observed photon rajectories. The extracted parameters agree with both the decay rate of the donor-acceptor cross corre-

Studies of the villin cause 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 reolution in single molecule FRET experiments [13] has been limited by the bin time of the measurement, which is usually 1-10 ms at to avoid photochemical problems such as bleaching and blinking of the dues [14]. For residence times much longer than these hi 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 aver-age 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 can be so easily determined. To extract kinetics from trajectories under these circumstances. Conich and Szaho have developed maximum likelihood methods for analyzing photor

ning [15]. We previously used this method on the ashelical protein or. D and showed that the method produces accurate folding time 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 pro dynamics. Here we show how to extract the kinetics of folding and unfolding from the data that is complicated by blinking, and

E-mill addresses: chunghoidniddi.nih.gov (H.S. Chung). naton@helix.nih.gov (W.A. Eason).

very rapid transitions between the 2 states when the protein is

molecules. The duration of the transition path is the tiny fraction

of the time in a trajectory that it takes for a protein to fold or unfold when it actually happens (8). With the possible exception

of one study of RNA folding (9), transition path times have not

een measured for any molecular process in solution.

A realistic goal for single-molecule FRET experiments is to

neasure transition path times for protein folding and unfolding

fistance versus time trajectories will be totally new kinds

fistance could be measured simultaneously, e.g., by using 3 or mon

dyes (11-13), model-independent information on the width of the microscopic pathway distribution could be derived from correla-

toward these important goals by determining an upper bound for

ne transition path time from single-molecule FRET trajectories of ne 56 residue 2-state protein GB1, immobilized on a glass surface

the difficulty in measuring reliable trajectories is evidenced by the fact that there have been only 3 additional studies since the first measurements on single-immobilized proteins by Hochstrasser and

toworkers almost 10 years ago (15-18). The practical problem has been to immobilize the protein and measure long FRET trajecto-

ies of the protein folding and unfolding, without spurious effect

from dve photophysics or from the immobilization method, until

polarization, and absolute and relative time of arrival after nice

one of the dyes "bleaches," i.e., ceases to emit photons because of

of demanding tests for atomistic molecular dynamics sin

by a biotin-streptswidin-biotin linkage (Fig. 2). by a nonmotrepurvant-rotum manage (Fig. 2).

Although the idea that much could be learned about protein folding mechanisms from such trajectories has been apparent since the very early days of single-molecule spectroscopy, an indication of

The authors declare no conflict of interest

This settlets is a DNAS Direct Submission

See Commentary on page 11829.

and, ultimately, to obtain distance versus time trajectories durin

in the folded and unfolded states. Comparison with the results for tectable effect on the structure or dynamics of the unfolded protein and only a small effect on the folial functioning kinetics. Analysis of the photon-by-photon trajectories yields a transition path time <200  $\mu s$ , >10,000 times shorter than the mean waiting time in the unfolded

trajectory (8-10) or Markov state modeling of many short trajectories [11]. Studies of the villin cause 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

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 aver-age residence times (also called waiting or dwell times). When the average residence time is much shorter than the bin time, howver, transitions occur within the bins, and rate coefficients can be so easily determined. To extract kinetics from trajectories under

likelihood methods for analyzing photon ning [15]. We previously used this method on the ashelical protein or D and showed that the method produces accurate folding times ~1 ms, as judged by the excellent agreement of the sum of the rate coefficients obtained from the maximum likelihood analysis

Contents lists available at SciVerse ScienceDirect

Chemical Physics 422 (2013) 229-237

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\*

Lebensters of Chemical Physics, National Justinute of Diabetes and Diarettee and Kidney Diarase, National Justinute of Health, Rethenda MD 20832-0520, USA

ARTICLE INFO ABSTRACT Rolding and unfolding rates for the ultrafact folding villin subdomain were determined from a photon-by-photon analysis of fluorescence trajectories in single molecule RET operiments. One of the obstacles to measuring fast kinetics in single molecule fluor-oences experiments is bikingle of the fluorophores on a Article Matery: Available online 14 August 2012

annual eithat is not well separated from the process of interest. By incorporating accepter blinking into a brow olatak interior model, we show what it is prossible to extract accurate a rise conflicted into the serior second time scale for kiding and unfolding using the maximum likelinds of section of Coopch and Sasho. This method yields he most likely parameters of a given model that can reproduce the observed printon trajectories. The extracted parameters are given which the decay rate of the donor-acceptor come communities to the contract of parameters are set of the contract of the co

Single molecule fluorescence

MET Maximum likelihood analysis

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 tates [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 sin-gle-domain protein [2,3], is among the fastest folding proteins [4,5], and exhibits unusual kinetics such as a denaturant indepen dent relaxation rate [6] and an apparent increase in the internal friction with temperature in a Kramers description of the barrier

**Experiments Determine Protein** very rapid transitions between the 2 states when the protein is very rapid transitions between the 2 states when the protein is either folding or unfolding. Indeed, all mechanistic information about folding and unfolding is contained in these so-called transition paths (Fig. 1), which can only be observed for single **Folding Transition Path Times** molecules. The duration of the transition path is the tiny fraction Hoi Sung Chung,\* Kevin McHale, John M. Louis, William A. Eaton\* of the time in a trajectory that it takes for a protein to fold or unfold when it actually happens (8). With the possible exception The transition path is the tiny fraction of an equilibrium molecular trajectory when a transition of one study of RNA folding (9), transition path times have not occurs as the free-energy barrier between two states is crossed. It is a single-molecule property that contains all the mechanistic information on how a process occurs. As a step toward observing transition paths in protein folding, we determined the average transition-path time for a fast- and the FRET efficiency is midway between the

A realistic goal for single-molecular process in solution.

A realistic goal for single-molecule FRET experiments is to measure transition path times for protein folding and unfolding and, ultimately, to obtain distance versus time trajectories during the transition paths. The distribution of transition path times and o distance versus time trajectories will be totally new kinds of demanding tests for atomistic molecular dynamics size distance could be measured simultaneously, e.g., by using 3 or mon dyes (11-13), model-independent information on the width of the microscopic pathway distribution could be derived from correla-tions among the distances (14). In this work we take a major step toward these important goals by determining an upper bound for

Experimental determination of upper bound for

transition path times in protein folding from

single-molecule photon-by-photon trajectories

FRET trajectories were measured on single molecules of the dve-

labeled, 56-residue 2-state protein GBI, immobilized on a glass sur face via a biotin-streptavidin-biotin linkage. Characterization of indi

vidual emitted photons by their wavelength, polarization, and

absolute and relative time of arrival after picosecond excitation allowed the determination of distributions of FRET efficiencies, donor

and acceptor lifetimes, steady state polarizations, and waiting times

and only a small effect on the folding/unfolding kinetics. Analysis of

state (the inverse of the folding rate coefficient). Szabo's theory for

diffusive transition paths shows that this upper bound for the

the Kramers preexponential factor for the rate coefficient, and predicts that the transition path time is remarkably insensitive to the

Alexa 488 | fluorescence | ERET | maximum bladbood function | postein CR1

A detailed description and understanding or accommunity of protein folding has been one of the great challenges to bioobysical science. The simplest system to study, and the one that

librium and at all times in kinetic experiment

nd unfolded. In ensemble folding experiments kinetics ar by rapidly changing solution conditions, e.g., the temper denaturant concentration, and monitoring the relaxation

pulations to their new equilibrium ratio with probes orescence, circular dichroism or infrared spectroscopy

be seen from the schematic of a trajectory in Fig. 1, the dy

norrescence experiment. There are fluctuations due to a

nee experiment. There are fluctuations due to se sean value in each state, interrupted by what appo-eous jumps in FRET efficiency signaling folding

nature of equilibrium is dramatically demonstrated when Förster resonance energy transfer (FRET) in a single

folding. The residence or waiting times in each state are

strainly distributed, with the mean time in the unfolded and segments of the trajectories corresponding to the inverse folding and unfolding rate coefficients, respectively.

www.onas.org/cgi/doi/10.1073/onas.0901178106

Rate coefficients can, albeit with assumptions, be much

molecule kinetics, however, can be studied at equilib

detailed description and understanding of mechanisms of

ing molecules showed that immobilization has no de-

ne transition path time from single-molecule FRET trajectories of ne 56 residue 2-state protein GBI, immobilized on a glass surface by a biotin-streptswidin-biotin linkage (Fig. 2). by a nonmatrepurvam-rotom intender (Fig. 2).

Although the idea that much could be learned about protein folding mechanisms from such trajectories has been apparent since the very early days of single-molecule spectroscopy, an indication of he difficulty in measuring reliable trajectories is evidenced by the

come from measurements on single molecules,
of a successful barrier-crossing event (Fig. 1).

The strategy used in this study is to illumi-Alexa 594

K<sub>10</sub>C

Alexa 488

F. Chen et al., infect. immun. 65, 1626 (1997).
 C. N. Paco, G. R. Grirroley, J. M. Scholtz, J. Mol. Chem.

284, 13.85 (2009). 33. H. Li, A. D. Robertson, J. H. Jensen, Proteins 65, 704 (2009)

Administrator We think A Brown and T Blocky

Action reading of the manuscript; the staff of beamline 9-2 at the Stanford Synchrotron Radiation Lightnown (SSR) and the NECR staff of the Advanced Photon Source (APS).

L Yao, N. Ricz, A. M. Kruel, and J. Temblay for excellent extended assistance and R. Hattington and A. Rolling for

Single-Molecule Fluorescence

a slow-folding protein from a photon-by-photon analysis of fluorescence trajectories in single-molecule Förster resonance energy transfer experiments. Whereas the folding rate coefficients

that a fast- and a slow-folding protein take almost the same time to fold when folding actually

Theory predicts that folding mechanisms are - fast-folding, all-8 protein D9-residue formin-binding

Colded molecule can self-assemble to form in ensemble studies (7, 8), as well as a marked to

s biologically active, folded structure by means reduced upper bound compared with our previ

lecular dynamics simulations (2-6). Information coefficient, which measures the frequency of a

on pathway distributions from experiments must transition, the transition path time is the duption

of many different sequences of conformational our study for the 56-residue of 8 protein GRU changes (1). The distribution of those fiding path—
the graph of the statement of those fiding path—
the B1 immunoglobalin-binding domain of protenting from Streptococcup (9). In contrast to a nie

happens. A very simple model based on energy landscape theory can explain this result.

centor red, and the absolute tin ons between states are dearly

states. The photon trajectoris

the transition region with a maximum od method by using simple mod oton trajectories were measured WW domain and protein GB1 donor and acceptor fluorophor ines incorporated into the pro se trajectories, two properties

Infectious Diseases (MAID), NIH, Department of Health and Human Services (HHS), under award number

USA 41057159 (C.R.S.). Atomic coordinates and drudum on for the WHH bound M-PTC, M-PTC, and BoAT/M

on codes 3VDA, 3VDB, and 3VDC, respectively

differ by a factor of 10,000, the transition-path times differ by a factor of less than 5, which shows to the average (parsition-path time, \$772) (Fig. 4A).

the FRET efficiency distribution are bimodal, which indicates t

of declaring data and a substantial street of the substantial street

per transition path, to discard the majority

Self (II). The photon color matrix F depends to the color of a photon as F(acceptor) = E and

jectories between transitions, and to area. Eldonories I - E where E is a disconsil matrix

complex delikery compositions," that was filed in New mber of 2011. BoAT availability is subject to the restrictions

MIL MILITARE AND THE TOTAL OF T

at anniv to HHS salet agents and NAID Categor

21 September 2011; accepted 12 January 2012 10 11 Newtonce 12 14 270

tracted from the region near the transitions and

analyzed using the Gorich Szabo maximum likeshood method (10).

For a given model, the Gorich-Szabo meth-

od calculates the parameters of the model that

can most accumtedy reproduce the photon trajectories (Fig. 3). We adopt a one-step model for

the transition path, which may be viewed as the

FRET efficiency changes along the path. This

nicture can be represented in a kinetic model for

a two-state system with a finite transition path

folded and unfolded states  $[E_S = (E_V + E_U)^2]$ . In this model, the lifetime of  $S(t_S)$  corresponds

S has the property of a transition state, because

the rate mellicients from S to E and S to II (b.)

The likelihood function for the 4th photon

 $L_j = \mathbf{v}_{fn}^T \prod_{i=1}^n \{\mathbf{n}\mathbf{F}(c_i) \exp[(\mathbf{K} - \mathbf{n})\tau_i]\} \mathbf{n}\mathbf{F}(c_1)\mathbf{v}_{int}$ 

Here, K is the rate matrix [equation S6 (11)]

containing the three rate coefficients &... k...

and  $k_S$ ), N is the number of photons in the jth trajectory,  $c_j$  is the color of the jth photon (donor

are the same, and therefore, the pass = 1/2.

Supportion Online Material

inned fluorescence and photon to g.org SCIENCE VOL 335

main. blinking occurs on a similar time scale to folding and unfolding, making it problematic to distinguish dye blinking from proteir

dynamics. Here we show how to extract the kinetics of folding and infolding from the data that is complicated by blinking, and

for this two-state system and the decay of the donor-acceptor cross-correlation function [16]. For the faster-folding villin subdo-

the moderate illumination intensities that have been employed to avoid photochemical problems such as bleaching and blinking of the dues [14]. For residence times much longer than these hir

## Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories

Hoi Sung Chung<sup>1</sup>, John M. Louis, and William A. Eaton

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

This Feature Article is part of a series identified by the Editorial Board as reporting findings of exceptional significance.

Edited by Alan Fersht, University of Cambridge, Cambridge, United Kingdom, and approved May 18, 2009 (received for review February 2, 2009)

Transition paths are a uniquely single-molecule property not yet observed for any molecular process in solution. The duration of transition paths is the tiny fraction of the time in an equilibrium single-molecule trajectory when the process actually happens. Here, we report the determination of an upper bound for the transition path time for protein folding from photon-by-photon trajectories. FRET trajectories were measured on single molecules of the dyelabeled, 56-residue 2-state protein GB1, immobilized on a glass surface via a biotin-streptavidin-biotin linkage. Characterization of individual emitted photons by their wavelength, polarization, and absolute and relative time of arrival after picosecond excitation allowed the determination of distributions of FRET efficiencies, donor and acceptor lifetimes, steady state polarizations, and waiting times in the folded and unfolded states Comparison with the results for freely diffusing molecules showed that immobilization has no detectable effect on the structure or dynamics of the unfolded protein and only a small effect on the folding/unfolding kinetics. Analysis of the photon-by-photon trajectories yields a transition path time <200 us. > 10,000 times shorter than the mean waiting time in the unfolded state (the inverse of the folding rate coefficient). Szabo's theory for diffusive transition paths shows that this upper bound for the transition path time is consistent with previous estimates of the Kramers preexponential factor for the rate coefficient, and predicts that the transition path time is remarkably insensitive to the folding rate, with only a 2-fold difference for rate coefficients that differ by 105-fold.

Alexa 488 | fluorescence | FRET | maximum likelihood function | protein GB1

A detailed description and understanding of mechanisms of biophysical science. The simplest system to study, and the one that has produced the most insights, is a protein exhibiting 2-state behavior (1-7). A 2-state protein has only 2-populations of molecules in equilibrium and at all times in kinetic experiments-folded and unfolded. In ensemble folding experiments kinetics are studied by rapidly changing solution conditions, e.g., the temperature or denaturant concentration, and monitoring the relaxation of the 2 populations to their new equilibrium ratio with probes such as fluorescence, circular dichroism or infrared spectroscopy. Single molecule kinetics, however, can be studied at equilibrium. As can be seen from the schematic of a trajectory in Fig. 1, the dynamical nature of equilibrium is dramatically demonstrated when observing Förster resonance energy transfer (FRET) in a single-molecule fluorescence experiment. There are fluctuations due to shot noise about a mean value in each state, interrupted by what appear to be instantaneous jumps in FRET efficiency signaling folding or unfolding. The residence or waiting times in each state are exponentially distributed, with the mean time in the unfolded and folded segments of the trajectories corresponding to the inverse of the folding and unfolding rate coefficients, respectively.

Rate coefficients can, albeit with assumptions, be much more easily obtained from a combination of ensemble kinetic and

equilibrium experiments, where the former measure the sum of the rate coefficients and the latter their ratio. The unique information in a single-molecule experiment is contained in the very rapid transitions between the 2 states when the protein is either folding or unfolding. Indeed, all mechanistic information about folding and unfolding is contained in these so-called transition paths (Fig. 1), which can only be observed for single molecules. The duration of the transition path is the tiny fraction of the time in a trajectory that it takes for a protein to fold or unfold when it actually happens (8). With the possible exception of one study of RNA folding (9), transition path times have not been measured for any molecular process in solution.

A realistic goal for single-molecule FRET experiments is to measure transition path times for protein folding and unfolding and, ultimately, to obtain distance versus time trajectories during the transition paths. The distribution of transition path times and of distance versus time trajectories will be totally new kinds of demanding tests for atomistic molecular dynamics simulations of folding (10), which, if accurate, contain everything one would ever want to know about a protein folding mechanism. If more than one distance could be measured simultaneously, e.g., by using 3 or more dyes (11-13), model-independent information on the width of the microscopic pathway distribution could be derived from correlations among the distances (14). In this work we take a major step toward these important goals by determining an upper bound for the transition path time from single-molecule FRET trajectories of the 56 residue 2-state protein GB1, immobilized on a glass surface by a biotin-streptavidin-biotin linkage (Fig. 2).

Although the idea that much could be learned about protein folding mechanisms from such trajectories has been apparent since the very early days of single-molecule spectroscopy, an indication of the difficulty in measuring reliable trajectories is evidenced by the fact that there have been only 3 additional studies since the first measurements on single-immobilized proteins by Hochstrasser and coworkers almost 10 years ago (15-18). The practical problem has been to immobilize the protein and measure long FRET trajectories of the protein folding and unfolding, without spurious effects from dye photophysics or from the immobilization method, until one of the dyes "bleaches," i.e., ceases to emit photons because of an irreversible photochemical change. To overcome this hurdle we have characterized individual emitted photons by their wavelength, polarization, and absolute and relative time of arrival after pico-

Author contributions: H.S.C. and W.A.E. designed research; H.S.C. performed research; H.S.C. and J.M.L. contributed new reagents/analytic tools; H.S.C. analyzed data; and H.S.C. and W.A.E. wrote the paper.

The authors declare no conflict of interest.

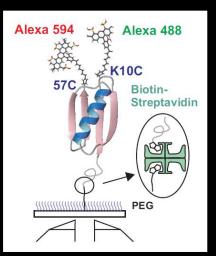
This article is a PNAS Direct Submission See Commentary on page 11823.

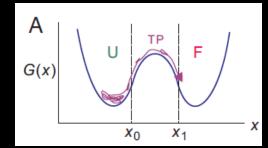
<sup>1</sup>To whom correspondence may be addressed. E-mail: chunghoi@niddk.nih.gov or

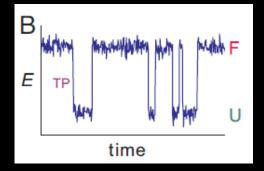
This article contains supporting information online at www.pnas.org/cgi/content/full/

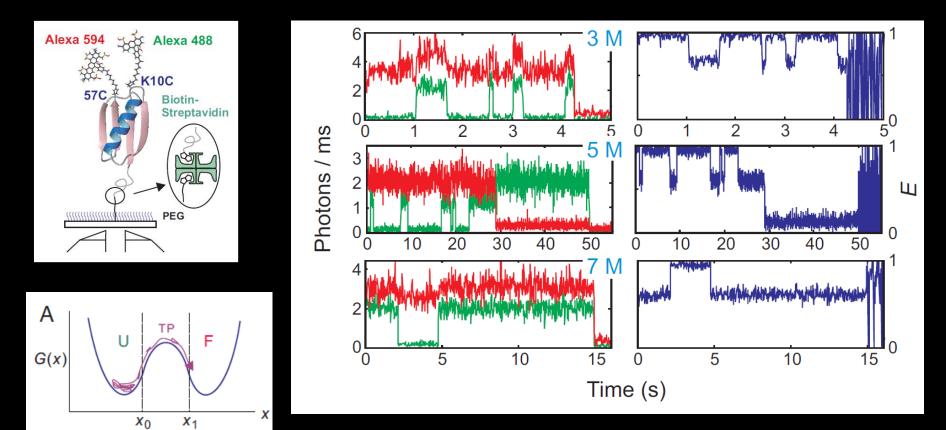
# Experimental determination of upper bound for transition path times in protein folding from single-molecule photon-by-photon trajectories

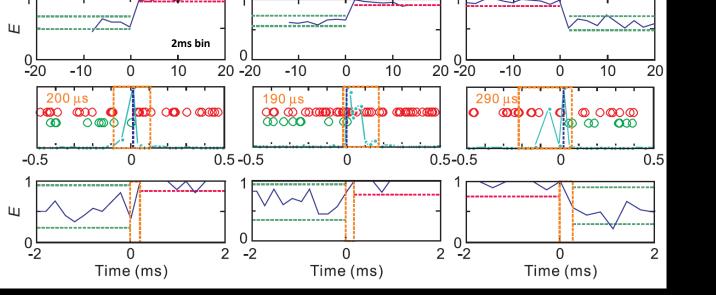
Hoi Sung Chung<sup>1</sup>, John M. Louis, and William A. Eaton<sup>1</sup>











$$f_{F(U)}(n) = E_{F(U)}^{n_A} \left(1 - E_{F(U)}\right)^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

$$f_{F(U)}(n) = \frac{1}{2} \sum_{j=1}^{n_A} (1 - E_{F(U)})^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

$$f_{F(U)}(n) = \frac{1}{2} \sum_{j=1}^{n_A} (1 - E_{F(U)})^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

$$f_{F(U)}(n) = \frac{1}{2} \sum_{j=1}^{n_A} (1 - E_{F(U)})^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

$$f_{F(U)}(n) = \frac{1}{2} \sum_{j=1}^{n_A} (1 - E_{F(U)})^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

$$f_{F(U)}(n) = \frac{1}{2} \sum_{j=1}^{n_A} (1 - E_{F(U)})^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

$$f_{F(U)}(n) = \frac{1}{2} \sum_{j=1}^{n_A} (1 - E_{F(U)})^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

$$f_{F(U)}(n) = \frac{1}{2} \sum_{j=1}^{n_A} (1 - E_{F(U)})^{n-n_A} \prod_{j=1}^{n-1} P_{F(U)}(d_j)$$

: photon counting rates

d: the time interval

- 31 F Chen et al. Indext Immun 65, 1626 (1997). 32. C. N. Pace, G. R. Grimsley, J. M. Scholtz, J. Biol. Chem.
- 33. H. U. A. D. Robertson, 1. H. Jensen, Proteins 61, 704

Acknowledgments: We thank A Brunger and T. Birg for critical reading of the manuscript: the staff of hearning 9-2 at the Stanford Synchrotron Radiation Lightsource (SSR) and the NE-CAT staff of the Advanced Photon Source (APS). particularly K. Rajashankar, for a stistance in data collection; G. Yao, N. Krez, A. M. Kruel, and J. Tremblay for excellent technical assistance; and R. Uddington and A. Bobkov for

assistance with ITC and ALIC This work was partly supported by an Alfred P. Sloan Research Fellowship (R.J.), by the Douts the East chun accompline to the CDEC Expellent in Italian GSC 108 to S.R.), by the Robert-Koch-Institut (1362/I-979 to A.R.). and by grants from the National Institute of Alleroy and Infectious Diseases (NIAD), NIH, Department of Health and Human Services (HHS), under award number USA AIOS7159 (C.R.S.). Atomic coordinates and drudure fadors for the WHH-bound M-PTC, M-PTC and BoNT/Al have been deposited with the Protein Data Bank under accession codes 3VOA, 3VOB, and 3VOC, respectively. Sanford-Burnham Medical Research Inditate has a pending patent application, titled "Botalinum neurotoxin protective

complex delivery imminestions," that was filled in November of 2011. BoNT availability is subject to the restrictions that apply to HHS select agents and NIAD Category

Supporting Online Material www.sciencemag.org/cdi/mrtentfulV335.6071.977/DCL Materials and Methods Flor: \$1 to \$13

Tables S1 to S4 References (34-51)

21 September 2011; accepted 12 January 2012 10.1126/science.1214270

## Single-Molecule Fluorescence **Experiments Determine Protein Folding Transition Path Times**

Hoi Sung Chung,\* Kevin McHale, John M. Louis, William A. Eaton\*

The transition path is the tiny fraction of an equilibrium molecular trajectory when a transition occurs as the free-energy barrier between two states is crossed. It is a single-molecule property that contains all the mechanistic information on how a process occurs. As a step toward observing transition paths in protein folding, we determined the average transition-path time for a fast- and a slow-folding protein from a photon-by-photon analysis of fluorescence trajectories in singlemolecule Förster resonance energy transfer experiments. Whereas the folding rate coefficients differ by a factor of 10,000, the transition-path times differ by a factor of less than 5, which shows that a fast- and a slow-folding protein take almost the same time to fold when folding actually happens. A very simple model based on energy landscape theory can explain this result.

Theory predicts that folding mechanisms are fast-folding, all-8 protein 139-residue formin-binding heterogeneous, so that an individual un- protein (FBP) WW domain) shown to be two-state folded molecule can self-assemble to form in ensemble studies (7, δ), as well as a markedly its biologically active, folded structure by means reduced upper bound compared with our previof many different sequences of conformational changes (1). The distribution of these folding nathways can now be calculated from atomistic molecular dynamics simulations (2-6). Information on efficient, which measures the frequency of a on nathway distributions from experiments must come from measurements on single molecules, because only average properties are obtained in experiments on the large ensemble of molecules in bulk experiments. A single-molecule, equilibrium protein folding-unfolding trajectory is illustrated in Fig. 1, as monitored by Förster resonance energy transfer (FRET) spectroscopy, and its relation to the free-energy barrier as it crosses between the folded and unfolded states is shown. The most interesting part of the trajectory is contained in what appears to be an instantaneous iumn between the two states called the transition path, which contains all of the information on the mechanism of folding and unfolding. The first step toward observing transition paths in protein folding, which we report here, is the determination of its average duration (transition-math time) for a propertor red, and the absolute time of arrival

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Midney Diseases, National Institutes of Health (NH), Retherch, MD, 20892-0520, USA

To whom correspondence should be addressed. E-mailchunghoiganiddunh.gov (RSC); exton@helxxnh.gov (WA.E) of two states. The photon trajectories were ex- variable parameter (II).

tracted from the region near the transitions and analyzed using the Gorich-Szabo maximum likelihood method (10) For a given model, the Gonich-Szaho meth

od calculates the parameters of the model that can most accurately reproduce the photon trajectories (Fig. 3). We adopt a one-step model for the transition path, which may be viewed as the simplest discrete representation of how the FRET efficiency changes along the path. This nicture can be represented in a kinetic model for a two-state system with a finite transition rath by introducing a third virtual state, S, for which the FRET efficiency is midway between the folded and unfolded states  $[E_S - (E_F + E_U)^2]$ . In this model, the lifetime of S (ts) corresponds to the average transition-path time, (fp) (Fig. 4A). S has the property of a transition state, because the rate coefficients from S to F and S to U (kg) are the same, and therefore, the  $p_{total} = \frac{1}{2}$ .

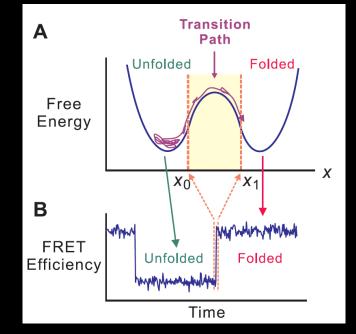
The likelihood function for the ith photon trajectory is (10):

$$L_j = \mathbf{v}_{f_n}^T \prod_{t=2}^N \{ \mathbf{n} \mathbf{F}(c_t) \exp \left[ (\mathbf{K} - \mathbf{n}) \mathbf{r}_t \right] \} \mathbf{n} \mathbf{F}(c_1) \mathbf{v}_{int}$$
(1)

Here, K is the rate matrix [equation S6 (11)] containing the three rate coefficients  $(k_{tr}, k_{tr})$ and  $k_S$ ), N is the number of photons in the jth trajectory, c, is the color of the ith photon (donor or acceptor), and  $\tau_c$  is a time interval between the ith and (i-1)th photons as shown in fig. S4B (11) The photon color matrix E depends on the color of a photon as F(acceptor) = E and F(donor) = I - E, where E is a diagonal matrix lyze the transition region with a maximum like- with elements that are FRET efficiencies of the three states (F, S, and U), and I is the unit matrix, n is a diagonal matrix with elements that are photon count rates of the three states. hilized WW.domain and protein GR1 molecules  $v_{tot}$  and  $v_{tht}$  are vectors that describe the state (folded or unfolded) at the beginning and the end of the trajectory. Practically, log-likelihood functions were calculated, and the total log like-Bhood function of all trajectories was calculated by summing the log-likelihood functions

to within 
$$-0.5$$
 ns. As shown in Fig. 3, A and B,  
transitions between states are dentry resolved in (ln  $L = \sum_{j} \ln L_{jj}$ ) of individual trajectories that  
the binned fluorescence and photon trajectories,

and the FRET efficiency distributions (Fig. 3, C contain a transition between folded and unfolded and D) are bimodal, which indicates the presence states. In the likelihood function L, to is the only



# **Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times**

Hoi Sung Chung,\* Kevin McHale, John M. Louis, William A. Eaton\*

ous study for the 56-residue, a/B protein GB1

(the B1 immunoglobulin-binding domain of pro-

transition, the transition noth time is the duration

The strategy used in this study is to illumi-

nate dve-labeled protein molecules at very high

intensities to increase the number of detected pho-

tons ner transition nath, to discard the majority

of photons from the less-interesting segments of

the trajectories between transitions, and to ana-

lihood method by using simple models for the

Photon trajectories were measured for immo-

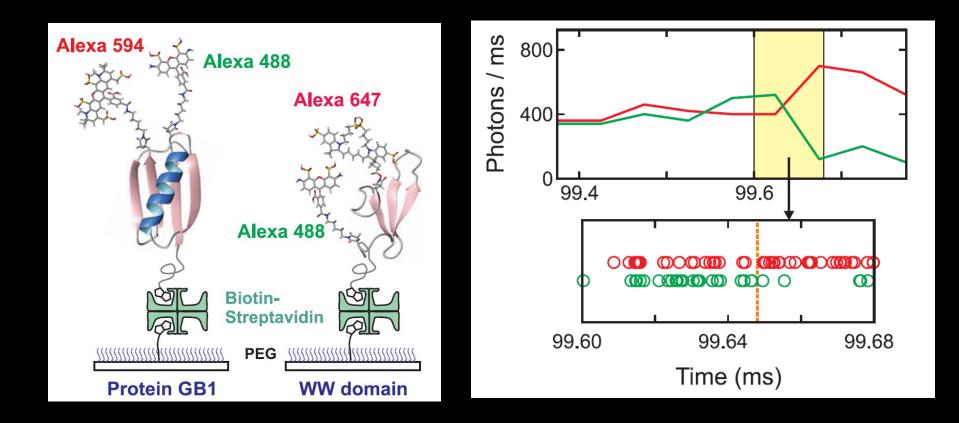
with donor and acceptor fluorophores attached

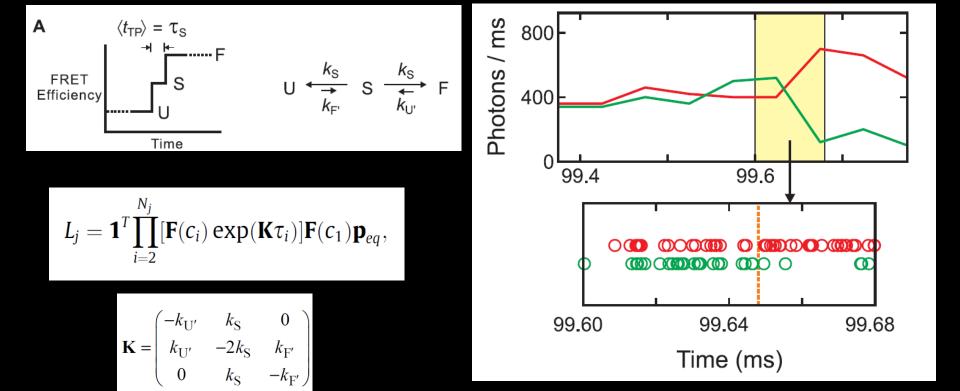
to cysteines incorporated into the proteins (Fig. 2).

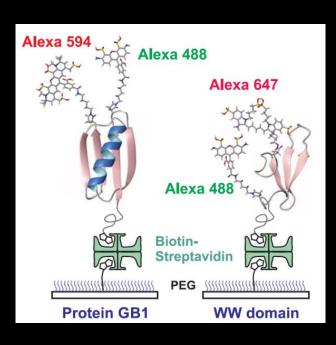
In these trajectories, two properties of each pho-

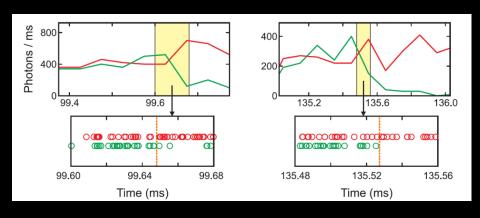
ton were recorded—the color, either donor green

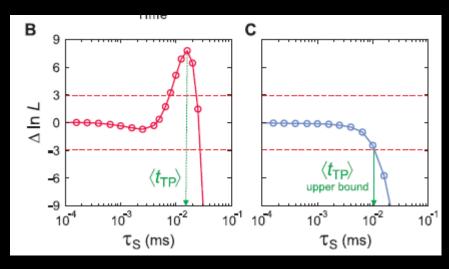
of a successful barrier-crossing event (Fig. 1).













Contents lists available at SciVerse ScienceDirect

### Chemical Physics





## 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 20852-0520, USA

## ARTICLE INFO

Article history: Available online 14 August 2012

\_\_\_\_

Single-molecule fluorescence FRET Maximum likelihood analysis Dye blinking Ultrafast protein feiding rates Photon trajectories

#### ABSTRACT

iolding and unfolding rates for the ultrafast folding will in subdamain were determined from a photon-byphoton analysis of interessent trajectories in single molecule Ritt? experiments. (see of the obstacles is measuring fast kinetics in single molecule fluore occurs experiments is blinking of the fluorophores on a simuscular that is now the lequarated from the process of interest, blinkings of the fluorophores on a two state hieries model, we show that it is possible to extract accruate rate coefficients on the intertor of the process of the process of the process of interest, blinking of the interest of the process of the interest that is nettically delice from the 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 convetation function and the results of ensemble equilibrium and storiest experiments using nanoescond laters.

bublished by Elsevier R.V

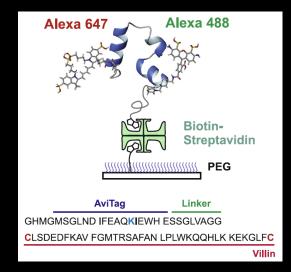
#### 1. Introduction

Willin subdomain

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  $\alpha$ -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 paid laterics, with folding times of the wist-type on the order of 5-50 µs [3,612]. Until quite recently the time resbution in single molecule REIT experiments [3] has been insitted by the bin time of the measurement, which is usually 1-10 ms at the moderate infundation intensities that have been employed of the dyes [14]. For residence times much longer than these bin times, two distinct peaks will appear in a histogram of the REIT



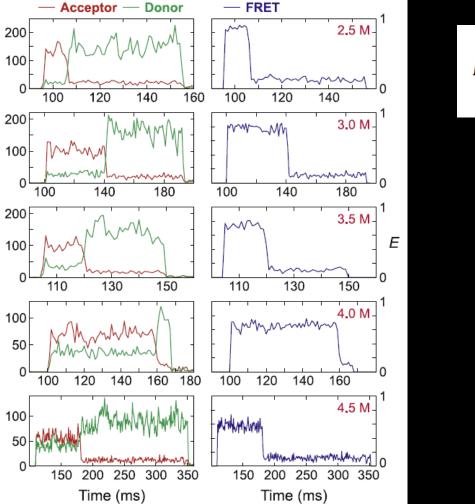
Measuring ultrafast protein folding rates from photon-by-photon analysis of single molecule fluorescence trajectories

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

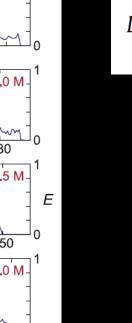
Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA

<sup>\*</sup> Corresponding authors

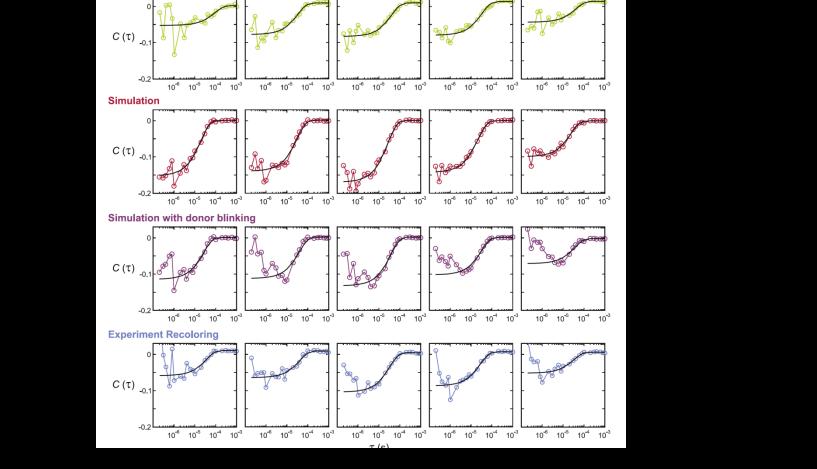
E-mail addresses: chunghoi@niddk.nih.gov (H.S. Chung), eaton@helix.nih.gov (W.A. Eston).



Photons / ms



 $\mathbf{T}[\mathbf{F}(c_i)\exp(\mathbf{K}\tau_i)]\mathbf{F}(c_1)\mathbf{p}_{eq},$ 



3.5 M

4.0 M

4.5 M

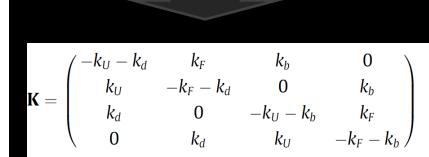
2.5 M

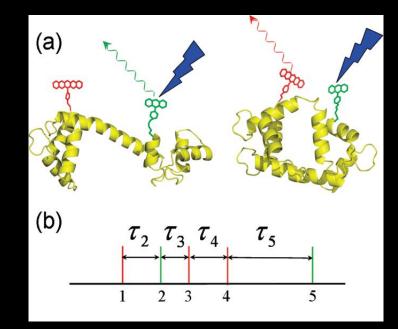
Experiment

3.0 M

 $k_U$ 

 $\overline{k_U}$ 





$$L = \mathbf{1}^{\mathrm{T}} (\boldsymbol{I} - \boldsymbol{E}) e^{\boldsymbol{K}\tau_{5}} \boldsymbol{E} e^{\boldsymbol{K}\tau_{4}} \boldsymbol{E} e^{\boldsymbol{K}\tau_{3}} (\boldsymbol{I} - \boldsymbol{E}) e^{\boldsymbol{K}\tau_{2}} \boldsymbol{E} \boldsymbol{p}_{\mathrm{eq}}$$

For a two-state case & two detected photon,

$$L = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 - E_1 & 0 \\ 0 & 1 - E_2 \end{bmatrix} \begin{bmatrix} e^{-k_{21} \cdot \tau_2} & e^{k_{12} \cdot \tau_2} \\ e^{k_{21} \cdot \tau_2} & e^{-k_{12} \cdot \tau_2} \end{bmatrix} \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}$$

$$L = (1 - E_1) e^{-k_{21} \cdot \tau_2} E_1 P_1 + (1 - E_1) e^{k_{12} \cdot \tau_2} E_2 P_2$$

$$+(1-E_2)e^{k_{21}\cdot\tau_2}E_1P_1+(1-E_2)e^{-k_{12}\cdot\tau_2}E_2P_2$$

$$L = \mathbf{1}^{\mathrm{T}} \prod_{k=2}^{N_{\mathrm{ph}}} (\mathbf{F}(c_k) \mathbf{e}^{\mathbf{K}\tau_k}) \mathbf{F}(c_1) \mathbf{p}_{\mathrm{eq}}$$