

Distinguishing between cooperative and unimodal downhill protein folding

Fang Huang*, Satoshi Sato*, Timothy D. Sharpe*, Liming Ying^{†‡}, and Alan R. Fersht*^{†§}

*Medical Research Council Centre for Protein Engineering, Hills Road, Cambridge CB2 2QH, United Kingdom; [†]Cambridge University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, United Kingdom; and [‡]Biological Nanoscience Section, National Heart and Lung Institute, Imperial College London, London SW7 2AZ, United Kingdom

Contributed by Alan R. Fersht, November 1, 2006 (sent for review October 10, 2006)

Conventional cooperative protein folding invokes discrete ensembles of native and denatured state structures in separate free-energy wells. Unimodal noncooperative (“downhill”) folding, however, proposes an ensemble of states occupying a single free-energy well for proteins folding at $\geq 4 \times 10^4 \text{ s}^{-1}$ at 298 K. It is difficult to falsify unimodal mechanisms for such fast folding proteins by standard equilibrium experiments because both cooperative and unimodal mechanisms can present the same time-averaged structural, spectroscopic, and thermodynamic properties when the time scale used for observation is longer than for equilibration. However, kinetics can provide the necessary evidence. Chevron plots with strongly sloping linear refolding arms are very difficult to explain by downhill folding and are a signature for cooperative folding via a transition state ensemble. The folding kinetics of the peripheral subunit binding domain POB and its mutants fit to strongly sloping chevrons at observed rate constants of $> 6 \times 10^4 \text{ s}^{-1}$ in denaturant solution, extrapolating to $2 \times 10^5 \text{ s}^{-1}$ in water. Protein A, which folds at 10^5 s^{-1} at 298 K, also has a well-defined chevron. Single-molecule fluorescence energy transfer experiments on labeled Protein A in the presence of denaturant demonstrated directly bimodal distributions of native and denatured states.

BBL | denaturation | kinetics | T jump

A currently controversial subject in protein folding is unimodal “downhill” versus classical cooperative folding. It is generally accepted that proteins fold on a free-energy landscape in which there are ensembles of states separated by free-energy barriers (1). Accordingly, there are cooperative transitions between those ensembles of states, as, for example, the native N and denatured D ensembles, which have a bimodal distribution of properties. However, when there is an extreme energetic bias toward the native state, the protein may fold “downhill,” without an energy barrier (1). This conventional (“chemical”) view of folding has been challenged by Muñoz and colleagues (2, 3), who claim that for a protein NapBBL, a truncated and naphthylalanine-labeled derivative of the BBL peripheral subunit binding domain (PSBD) from *Escherichia coli*, in particular, and for all proteins that fold faster than $40,000 \text{ s}^{-1}$ at 298 K (4), the D and N states are not separated by an energy barrier, but slowly merge into each other with changing conditions; Muñoz and colleagues (2) call this mechanism “downhill” folding (Fig. 1). We use the term noncooperative or unimodal (5) for this downhill folding. Various criteria derived from equilibrium experiments have been proposed to be signatures of unimodal folding (2, 6). Downhill folding is proposed to be important, in particular as a means for the PSBDs to adjust their sizes as “molecular rheostats,” and in general because it is suggested (2, 4, 7) that it opens up the exciting prospect of examining the whole pathway of folding from equilibrium spectroscopic observations.

Until recently, the proteins that were studied folded slowly compared with the time scales of observation by spectroscopy, and it was easy to show cooperative folding transitions because separate D and N states could be directly observed. NMR studies

on the denaturation of such slow folding proteins clearly show two sets of spectra, corresponding to the D and N states that are in slow exchange (8). For example, NMR-monitored kinetics of refolding of the two proteins barnase and chymotrypsin inhibitor 2 (CI2), show the signals of the denatured state disappearing concomitantly with the cooperative formation of the native state (9). A single-molecule FRET experiment demonstrates the distinct ensembles of denatured and native states for CI2 (10). The two-state folding of cold shock protein B, a relatively fast folder (11), is seen by single-molecule FRET studies to have a bimodal distribution (12). The problem of mechanistic distinction resides with the new generation of fast folding proteins that equilibrate on a rapid time scale relative to NMR spectroscopy or single-molecule studies (13–19, ¶).

There are two fundamental problems in falsifying the uni- and multimodal mechanisms for fast folding proteins; one inherent in the very nature of proteins, the other a problem of mechanistic equivalence. First, the quantitative biophysical analysis of protein structure and energetics is particularly difficult because all of the states involved are dynamic ensembles of structures that can change according to conditions (Fig. 2). The denatured state is a very loose ensemble of structures, the compactness of which varies considerably from protein to protein under conditions that favor folding. Folding intermediates have more restricted ensembles. However, even native structures have a rich variety of dynamic processes that can lead to small or larger changes in structure with changing conditions (20–22). Accordingly, all states are ensembles, occupying a range of conformations and free-energy levels within their wells. Thus, there is the basic problem of distinguishing one set of conformations occupying a very broad free-energy well from two sets of conformations in two adjacent free-energy wells, one of which is broad and the other narrower (Fig. 2). This situation is further complicated when there are folding intermediates. Second, rapid equilibration between states leads to a classical mechanistically equivalent situation, which often occurs in spectroscopy: if the time scale of observation is much longer than that of the interconversion, then it is very difficult, and usually impossible, to determine whether there are separate states or a just single one that has the weighted average properties of the individual states. Thus, it is inherently problematic to prove for rapidly folding and unfolding proteins whether there are separate N and D states that are in rapid equilibrium or a single unimodal species that has the weighted mean properties of N and D. For example, if the changes in

Author contributions: F.H., S.S., L.Y., and A.R.F. designed research; F.H., L.Y., S.S., and T.D.S. performed research; A.R.F. contributed new reagents/analytic tools; F.H., L.Y., T.D.S., and A.R.F. analyzed data; and A.R.F. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: PSBD, peripheral subunit binding domain; SASA, solvent accessible surface area; BDPA, B domain of Protein A.

[§]To whom correspondence should be addressed. E-mail: arf25@cam.ac.uk.

[¶]Brewer, S. H., Vu, D. M., Tang, Y. F., Raleigh, D. P., Dyer, R. B., Franzen, S. (2005) *Biophys J* 88:561A (abstr.).

© 2006 by The National Academy of Sciences of the USA

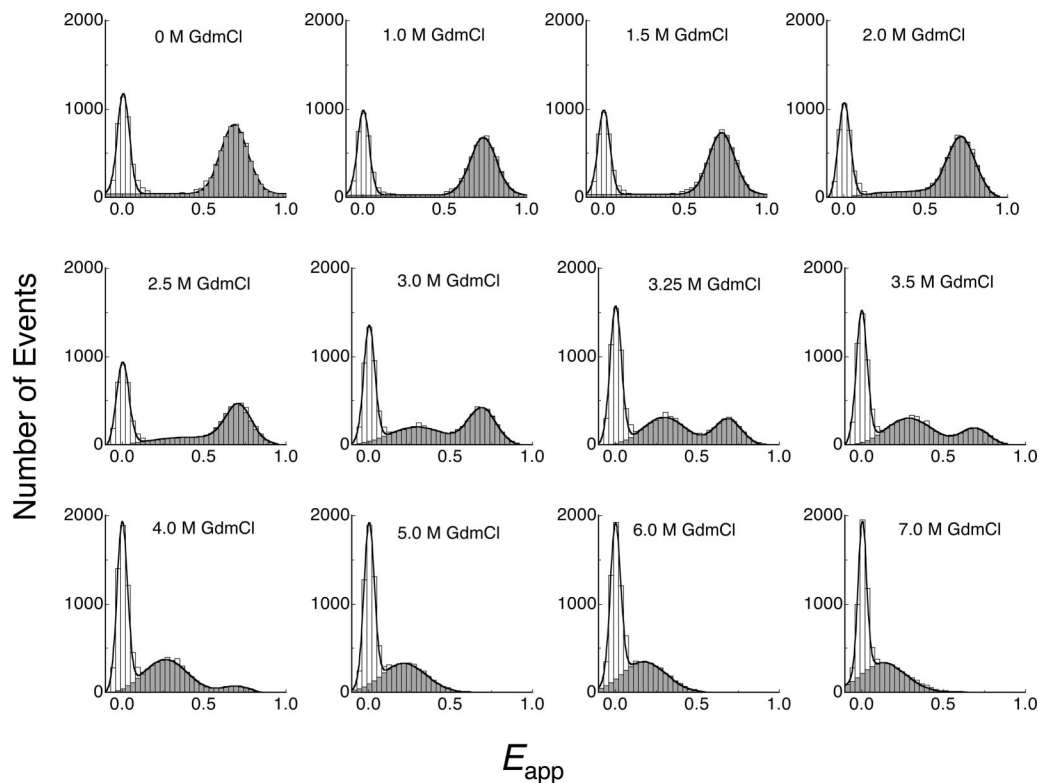


Fig. 6. Single molecule histogram of apparent FRET efficiency (E_{app}) for the double-labeled BDPA. The solid line is the Gaussian fit of the histogram, with the “zero” peak included in feint.

erative transition at rate constants well in excess of the predicted transition to unimodal downhill folding (4).

Single-Molecule Experiments. Single-molecule experiments can give conclusive answers. The B domain of Protein A (BDPA) is an ultrafast folding protein (17, 36, 37). Its refolding rate constant at 298 K in water is 10^5 s^{-1} (24). A Φ -analysis using multiple optical probes shows that it folds cooperatively (38). BDPA is an excellent model to distinguish between unimodal folding and conventional bimodal folding because it folds ultrarapidly in water and so falls into predicted unimodal downhill folding time range ($>40,000 \text{ s}^{-1}$; ref. 4). The FRET-labeled derivative prepared here folds at $40,000 \text{ s}^{-1}$ (Fig. 5). However, its D–N relaxation rate in GdmCl slows down to $\approx 200 \text{ s}^{-1}$ between 2 and 4 M GdmCl, which is in the single-molecule detection time range of ms. This feature allows the direct observation of the conformational distribution at single molecule level.

The single-molecule FRET studies showed directly that there is a bimodal distribution of native and denatured states for a suitable double-labeled sample of BDPA that folds at a similar rate to wild type (Fig. 6). There was only one peak with high FRET efficiency (centered at $E_{app} \approx 0.7$) in the histogram at 0 M GdmCl, corresponding to the native state of BDPA. With increasing of GdmCl concentration, another peak corresponding to the denatured state centered at $E_{app} \approx 0.3$ appeared. This peak had a much lower FRET efficiency and suggests a very expanded ensemble. Although the relative population of the native ensemble decreased with GdmCl concentration, the native peak does not shift obviously across the whole range of GdmCl concentration. A very slight shift was observed, which is due to the change of fluorescence quantum yield and refractive index. The invariant peak position of the native peak implies a constant native state, not changing with environment. In contrast, the

denatured peak shifted from $E_{app} \approx 0.3$ to <0.2 , indicating a more expanded denatured ensemble at higher concentration of GdmCl, which is consistent with previous reports (39). The single molecule FRET experiments show the existence of two distinct ensembles that are separated by a considerable free-energy barrier, which is inconsistent with unimodal folding.

For proteins folding faster than milliseconds under denaturing conditions, for example BBL, better time resolution will be required to distinguish between bimodal and unimodal folding mechanisms. Although it is extremely difficult to falsify downhill folding by equilibrium measurements, single-molecule studies and kinetics strongly imply that a genuine example of unimodal or downhill folding has yet to be found.

In conclusion, the experimental data reported so far for the destabilized, chemically labeled truncated PSBD, NapBBL, are ambiguous and do not adequately distinguish between cooperative and unimodal folding. The full-length PSBDs E3BD, BBL, and POB fold cooperatively with barrier-limited kinetics faster than that predicted for the onset of downhill folding. The sets of proteins that we have studied with rate constants of up to 10^5 s^{-1} , as exemplified by the PSBDs and BBPA, may be adequately analyzed by conventional kinetics and are suitable for Φ -analysis. A 35-residue subdomain of the chicken villin headpiece folds by barrier-limited kinetics at $\approx 10^6 \text{ s}^{-1}$ (40), suggesting that the onset of downhill folding may be far higher still than that predicted.

Experimental Methods

Protein Labeling. Two cysteine residues were introduced into Protein A at position 10 and 59, respectively, and purified as described (38). The labeling reaction was carried out in Tris buffer (pH 7.4, 50 mM Tris and 100 mM NaCl) with protein concentration of 100 μM , 4-fold excess of dye, and 2-fold excess of Tris(2-carboxyethyl)phosphine hydrochloride (TCEP).

Cys-59 was labeled with Alexa Fluor 488. To achieve site-specific labeling, 2-fold excess of IgG was applied to protect Cys-10, which is at the binding site of Protein A-IgG complex (41). Single-labeled protein was purified on HPLC and followed by labeling of Cys-10 with Alexa Fluor 647.

Single-Molecule FRET Measurements. A home-built dual-channel confocal fluorescence microscope was used to detect freely diffusing single molecules (42). FRET pair (Alexa Fluor 488/Alexa Fluor 647) labeled Protein A sample was excited by an Argon ion laser (Model 35LAP321–230; Melles Griot, Didam, The Netherlands) with 50 μ W at 488 nm. The donor and acceptor fluorescence were collected simultaneously through an oil-immersion objective (Apochromat \times 60, numerical aperture 1.45; Nikon, Surrey, U.K.) as the protein molecules diffuse through the laser focus, separated by a dichroic mirror (585DRLP; Omega Optical, Brattleboro, VT), filtered by long-pass and bandpass filters and detected separately by two photon-counting modules (SPCM-AQR14; PerkinElmer, Fremont, CA). The output of the two detectors was recorded by two computer-implemented multichannel scalar cards (MCS-PCI; EG&G, Quebec, QC, Canada). Sample solutions of 50–100 pM diluted in \approx 1 μ M unlabelled Protein A in PBS buffer (pH 7.4, 10 mM phosphate and 100 mM NaCl) were used to achieve single-molecule detection while reducing surface adsorption. All experiments were carried out at 20°C. A threshold of 30 counts per ms bin for the sum of the donor and acceptor fluorescence signals was used to differentiate single molecule bursts from the background. Because we use Alexa-647 as the acceptor, direct excitation of the acceptor is negligible. Apparent FRET efficiencies, E_{app} , of each burst were calculated according to $E_{app} = n_A/(n_A + n_D)$, where n_A and n_D are the background corrected acceptor and donor counts, respectively. Measurements were repeated at different denaturant concentrations and single molecule FRET histograms were built accordingly.

Single-molecule experiments at different excitation laser power proved that the “zero” peak in the apparent FRET histogram is predominantly due to the photobleaching (photoisomerization) of the FRET acceptor Alexa Fluor 647. Donor-only labeled protein shows that the “zero” peak disappears at $E_{app} > 0.2$.

Kinetics. Temperature jump studies were performed at 298 K as described for protein A (38, 43). Chevron data for POB Y166W (pseudo wild-type POB) and POB Y166W A131G (POB A131G) were acquired by using a modified Hi-Tech PTJ-64 temperature-jump apparatus with a 3 mm by 3 mm or 5 mm by 5 mm cell. Solutions of 300–600 μ M protein in 50 mM sodium acetate (pH 5.7, ionic strength adjusted to 150 mM, with appropriate concentrations of denaturant), were degassed with stirring for \approx 40 min before the experiment. Temperature jumps of 1–3.5 K to a final temperature of 298 K were used, and folding was monitored by fluorescence emission at >335 nm. Twenty to forty traces were acquired and averaged for each measurement, data from within the heating time of the instrument was discarded, and the resulting transient was fitted to a single exponential function. The dependence of rate constant on concentration of denaturant was fitted to a modified two-state chevron equation (23) with the denaturation midpoint constrained to the value from equilibrium chemical denaturation (for POB wild-type) or the value calculated from equilibrium thermal stability and the average equilibrium m value for 20 mutants (for POB A131G, where the equilibrium chemical denaturation had a truncated native baseline). The fits for protein A had no constraints. In all cases, there was excellent agreement between the kinetic and equilibrium m values, as is required for two-state kinetics (44).

L.Y. is a recipient of a Biotechnology and Biological Sciences Research Council David Phillips Research Fellowship.

1. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) *Proteins* 21:167–195.
2. Garcia-Mira MM, Sadqi M, Fischer N, Sanchez-Ruiz JM, Muñoz V (2002) *Science* 298:2191–2195.
3. Sadqi M, Fushman D, Muñoz V (2006) *Nature* 442:317–321.
4. Naganathan AN, Sanchez-Ruiz JM, Muñoz V (2005) *J Am Chem Soc* 127:17970–17971.
5. Knott M, Chan HS (2006) *Proteins* 65:373–391.
6. Naganathan AN, Perez-Jimenez R, Sanchez-Ruiz JM, Muñoz V (2005) *Biochemistry* 44:7435–7449.
7. Naganathan AN, Doshi U, Fung A, Sadqi M, Muñoz V (2006) *Biochemistry* 45:8466–8475.
8. Zeeb M, Balbach J (2004) *Methods* 34:65–74.
9. Killick TR, Freund SMV, Fersht AR (1998) *FEBS Lett* 423:110–112.
10. Deniz AA, Laurence TA, Belligere GS, Daham M, Martin AB, Chema DS, Dawson PE, Schultz PG, Weiss S (2000) *Proc Natl Acad Sci USA* 97:5179–5184.
11. Jacob M, Holtermann G, Perl D, Reinstein J, Schindler T, Geeves MA, Schmid FX (1999) *Biochemistry* 38:2882–2891.
12. Schuler B, Lipman EA, Eaton WA (2002) *Nature* 419:743–747.
13. Huang GS, Oas TG (1995) *Proc Natl Acad Sci USA* 92:6878–6882.
14. Spector S, Raleigh DP (1999) *J Mol Biol* 293:763–768.
15. Mayor U, Johnson CM, Daggett V, Fersht AR (2000) *Proc Natl Acad Sci USA* 97:13518–13522.
16. Myers JK, Oas TG (2002) *Annu Rev Biochem* 71:783–815.
17. Dimitriadis G, Drysdale A, Myers JK, Arora P, Radford SE, Oas TG, Smith DA (2004) *Proc Natl Acad Sci USA* 101:3809–3814.
18. Kubelka J, Hofrichter J, Eaton WA (2004) *Curr Opin Struct Biol* 14:76–88.
19. Zhu Y, Alonso DOV, Maki K, Huang CY, Lahr SJ, Daggett V, Roder H, DeGrado WF, Gai F (2003) *Proc Natl Acad Sci USA* 100:15486–15491.
20. Fersht A (1998) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York).
21. Korzhnev DM, Salvatella X, Vendruscolo M, Di Nardo AA, Davidson AR, Dobson CM, Kay LE (2004) *Nature* 430:586–590.
22. Karplus M, McCammon JA (2002) *Nat Struct Biol* 9:646–652.
23. Ferguson N, Sharpe TD, Schartau PJ, Sato S, Allen MD, Johnson CM, Rutherford TJ, Fersht AR (2005) *J Mol Biol* 353:427–446.
24. Ferguson N, Schartau PJ, Sharpe TD, Sato S, Fersht AR (2004) *J Mol Biol* 344:295–301.
25. Religa TL, Markson JS, Mayor U, Freund SMV, Fersht AR (2005) *Nature* 437:1053–1056.
26. Ferguson N, Sharpe TDS, Johnson CM, Schartau PJ, Fersht AR (2006) *Nature*, in press.
27. Tanford C (1970) *Adv Prot Chem* 24:1–95.
28. Matouschek A, Kellis JT, Serrano L, Bycroft M, Fersht AR (1990) *Nature* 346:440–445.
29. Matouschek A, Fersht AR (1993) *Proc Natl Acad Sci USA* 90:7814–7818.
30. Shen TY, Hofmann CP, Oliveberg M, Wolynes PG (2005) *Biochemistry* 44:6433–6439.
31. Schatzle M, Kiefhaber T (2006) *J Mol Biol* 357:655–664.
32. Kaya H, Liu ZR, Chan HS (2005) *Biophys J* 89:520–535.
33. Ferguson N, Sharpe TD, Johnson CM, Fersht AR (2006) *J Mol Biol* 356:1237–1247.
34. Spector S, Young P, Raleigh DP (1999) *Biochemistry* 38:4128–4136.
35. Allen MD, Broadhurst RW, Solomon RG, Perham RN (2005) *FEBS J* 272:259–268.
36. Arora P, Oas TG, Myers JK (2004) *Protein Sci* 13:847–853.
37. Vu DM, Myers JK, Oas TG, Dyer RB (2004) *Biochemistry* 43:3582–3589.
38. Sato S, Religa TL, Fersht AR (2006) *J Mol Biol* 360:850–864.
39. Kuzmenkina EV, Heyes CD, Nienhaus GU (2005) *Proc Natl Acad Sci USA* 102:15471–15476.
40. Kubelka J, Chiu TK, Davies DR, Eaton WA, Hofrichter J (2006) *J Mol Biol* 359:546–553.
41. Nakanishi T, Miyazawa M, Sakakura M, Terasawa H, Takahashi H, Shimada I (2002) *J Mol Biol* 318:245–249.
42. Ying L, Green JJ, Li H, Klenerman D, Balasubramanian S (2003) *Proc Natl Acad Sci USA* 100:14629–14634.
43. Sato S, Religa TL, Daggett V, Fersht AR (2004) *Proc Natl Acad Sci USA* 101:6952–6956.
44. Jackson SE, Fersht AR (1991) *Biochemistry* 30:10428–10435.