J|A|C|S ARTICLES

Published on Web 07/22/2010

Navigating the Downhill Protein Folding Regime via Structural Homologues

Athi N. Naganathan,[†,‡] Peng Li,[†,‡] Raúl Perez-Jimenez,[§,||] Jose M. Sanchez-Ruiz,[||] and Victor Muñoz*[,†,‡]

Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas (CSIC), Ramiro de Maeztu 9, Madrid 28040, Spain, Department of Chemistry & Biochemistry, University of Maryland, College Park, Maryland 20742, and Departamento de Química Física, Facultad de Ciencias, Universidad de Granada, 18071-Granada, Spain

Received April 28, 2010; E-mail: vmunoz@cib.csic.es

Abstract: Proteins that fold over free-energy barriers $\leq 3RT$ are classified as downhill folders. This regime is characterized by equilibrium unfolding that is proportionally broader and more complex the lower the folding barrier. Downhill proteins are also expected to fold up in a few microseconds. However, the relationship between rate and equilibrium signatures is affected by other factors such as protein size and folding topology. Here we perform a direct comparison of the kinetics and equilibrium unfolding of two structural homologues: BBL and PDD. BBL folds−unfolds in just $\sim 1\ \mu$s at 335 K and displays the equilibrium signatures expected for a protein at the bottom of the downhill folding regime. PDD, which has the same 3D structure and size, folds−unfolds ~8 times more slowly and, concomitantly, exhibits all the downhill equilibrium signatures to a lesser degree. Our results demonstrate that the equilibrium signatures of downhill folding are proportional to the changes in folding rate once structural and size-scaling effects are factored out. This conclusion has two important implications: (1) it confirms that the quantitative analysis of equilibrium experiments in ultrafast folding proteins does provide direct information about free-energy barriers, a result that is incompatible with the conventional view of protein folding as a highly activated process, and (2) it advocates for equilibrium−kinetic studies of homologous proteins as a powerful tool to navigate the downhill folding regime via comparative analysis. The latter should prove extremely useful for the investigation of sequence, functional, and evolutionary determinants of protein folding barriers.

Introduction

The application of ultrafast protein folding techniques over the last 15 years has led to two general findings. The first one is the discovery of many proteins that fold to completion in a few microseconds,[1] which indicates that ultrafast protein folding is a common phenomenon. The second is the fact that formation of the simplest elements of secondary structure (α-helices and β-hairpins) takes place in similar time scales.[2] This immediately leads to the realization that microsecond folding proteins are close to the folding speed limit and by extension that free-energy barriers to folding (F_b) are in general small.[3] Therefore, some of the fastest proteins may actually fold in a barrierless (downhill) process. The possibility of downhill protein folding is one of the critical predictions of energy landscape theory.[4] It is also of great practical significance because, in contrast to activated two-state-like folding, a downhill folding process can be experimentally resolved with exquisite structural detail in kinetic[5,6] and in equilibrium experiments.[7] Theoretically, the downhill folding regime is defined as the range covered between very small free-energy barriers (maximum F_b of $3RT$) all the way down to one-state downhill ($F_b < 1RT$), in which the free-energy surface approximates a single well.[3,8,9]

The practical question is how to distinguish whether a protein that folds in microseconds does it in a classical barrier crossing or in a downhill fashion and if the latter to what level within the downhill regime. Regarding this point it is important to note that once near or within the downhill folding regime, differences in F_b cannot be accurately determined by simple rate comparison because variability in the effective folding diffusion coefficient between proteins becomes highly significant.[10] Changes in the folding diffusion coefficient have recently become center stage even for very slow folding proteins.[11] In recent years several approaches to experimentally explore the downhill folding

[†] Consejo Superior de Investigaciones Científicas.
[‡] University of Maryland.
[§] Present address: Department of Biological Sciences, Columbia University, New York, NY 10027.
[||] Universidad de Granada.

(1) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76.
(2) Eaton, W. A.; Muñoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R.; Hofrichter, J. *Ann. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327–359.
(3) Muñoz, V. *Ann. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 395.
(4) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167.
(5) Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5897.
(6) Yang, W. Y.; Gruebele, M. *Nature* **2003**, *423*, 193.
(7) Muñoz, V. *Int. J. Quantum Chem.* **2002**, *90*, 1522.
(8) Gruebele, M. In *Protein Folding, Misfolding and Aggregation*; Muñoz, V., Ed.; RSC Publishing: Cambridge, 2008; p 106.
(9) Cho, S. S.; Weinkam, P.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 118.
(10) Muñoz, V.; Sadqi, M.; Naganathan, A. N.; de Sancho, D. *HFSP J.* **2008**, *2*, 342.

regime without prior knowledge of the diffusion coefficient have been developed. Some of these methods have focused on equilibrium manifestations,[7] such as probe-dependent denaturation midpoints,[12,13] characteristically broad unfolding thermograms,[14,15] complex coupling between different denaturing agents,[16] and skewed pre- and post-transition baselines.[17] In terms of kinetic experiments, emphasis has shifted from the observation of nonexponential decays,[18] and the inherent difficulties in their interpretation,[19,20] to identification of quantitative kinetic trademarks for downhill folding. The first of such methods, which was developed for λ-repressor, is based on detecting new kinetic features in rate-increasing mutations on moderately fast folding proteins.[6] A second kinetic method estimates F_b from the response of the folding relaxation rate to temperature[21] or chemical denaturants.[21,22] Another important realization is that the relaxation rate of a downhill folding process should be mostly insensitive to apparent changes in protein stability, whether judged by different probes or caused by experimental conditions.[23] Rate insensitivity is a particularly interesting property because it is fundamentally different from barrier-crossing kinetics in which rate and equilibrium constants go hand in hand.

An exciting recent development from work on the proteins gpW[24] and villin headpiece subdomain,[25] as well as on λ-repressor mutants,[26] is the finding of good agreement between equilibrium and kinetic criteria for estimating F_b. Such agreement is highly encouraging. However, a general scheme to navigate the downhill folding regime concertedly using kinetic and equilibrium experiments is still missing. A pressing issue refers to effects from protein size (number of residues, N) and three-dimensional structure. Folding rates decrease proportional to ($N^{1/2}$),[27,28] and thus, proteins below 50 residues are best candidates for microsecond folding.[28] By the same token, these small proteins are expected to have intrinsically broad equilibrium unfolding since both unfolding entropy and enthalpy increase linearly with size.[29] The debated question is whether broad, apparently complex, unfolding equilibria and microsecond folding rates are indeed interconnected, as expected for the downhill regime,[3] or unrelated manifestations of small protein size. In parallel, folding rates correlate with protein topology,[30] so that α-helical proteins tend to fold faster than β-proteins of similar size. This would suggest that comparison between kinetic and equilibrium estimates of F_b might be topology dependent due, for example, to variability in the pre-exponential term associated to the different intrinsic dynamics of α-helices and β-sheets.[2]

Here we present a strategy to circumvent these issues by comparing the kinetic and equilibrium behaviors of fast folding structural homologues. In particular, we compare the ~40-residue structurally homologous α-helical domains BBL and PDD, which share ~30% sequence identity (Figure 1A and 1B) and perform the same role on two multienzyme complexes involved in glucose metabolism.[31] BBL folds/unfolds in a few microseconds (~1 μs at 335 K[23,32]) and has become a forum for discussing downhill versus barrier-limited folding. On the one hand, BBL has been shown to exhibit the bulk equilibrium[12,13,15,16] and kinetic[23] signatures of one-state downhill folding. These characteristics have been observed in computer simulations as well, both in coarse-grained models[9,33,34] and in atomistic simulations.[35,36] On the other hand, concerns have been expressed about the significance and reliability of the equilibrium analyses[37,38] and the possibility of high sensitivity to small changes in the amino acid sequence and/or experimental conditions.[39] The same discussion has been very recently extended to single-molecule FRET experiments.[40,41] However, an important fact that has gone unnoticed in the midst of discussion is that BBL folds in essentially the same microsecond times regardless of the spectroscopic probe, variant, experimental conditions, and most critically experimental group making the measurements (see Figure S1 in the Supporting Information). Such rate agreement indicates that the controversy is not based on any significant differences between experimental data produced by different laboratories (a point already made for equilibrium experiments[15]) but rather on the interpretation of the data.

In this work BBL serves as a reference for comparison. We use the variant previously termed Naf-BBL[15] (see Figure 1D), which has been thoroughly examined and used to develop and test many of the methods described above. Its counterpart, PDD,[42] appears to fold more slowly according to NMR line-broadening analysis.[43] Further experiments on a mutated form

(11) Wensley, B. G.; Batey, S.; Bone, F. A.; Chan, Z. M.; Tumelty, N. R.; Steward, A.; Kwa, L. G.; Borgia, a.; Clarke, J. *Nature* **2010**, *463*, 685.
(12) Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Muñoz, V. *Science* **2002**, *298*, 2191.
(13) Sadqi, M.; Fushman, D.; Muñoz, V. *Nature* **2006**, *442*, 317.
(14) Muñoz, V.; Sanchez-Ruiz, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17646–17651.
(15) Naganathan, A. N.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Muñoz, V. *Biochemistry* **2005**, *44*, 7435.
(16) Oliva, F. Y.; Muñoz, V. *J. Am. Chem. Soc.* **2004**, *126*, 8596.
(17) (a) Sadqi, M.; Fushman, D.; Muñoz, V. *Nature* **2007**, *445*, E17. (b) Naganathan, A. N.; Muñoz, V. *Biochemistry* **2008**, *47*, 6752.
(18) Sabelko, J.; Ervin, J.; Gruebele, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6031.
(19) Hagen, S. J. *Proteins* **2007**, *68*, 205.
(20) Gruebele, M. *Proteins* **2008**, *70*, 1099.
(21) Naganathan, A. N.; Doshi, U.; Muñoz, V. *J. Am. Chem. Soc.* **2007**, *129*, 5673.
(22) Knott, M.; Chan, H. S. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 373.
(23) Li, P.; Oliva, F. Y.; Naganathan, A. N.; Muñoz, V. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 103.
(24) Fung, A.; Li, P.; Godoy-Ruiz, R.; Sanchez-Ruiz, J. M.; Muñoz, V. *J. Am. Chem. Soc.* **2008**, *130*, 7489.
(25) Godoy-Ruiz, R.; Henry, E. R.; Kubelka, J.; Hofrichter, J.; Muñoz, V.; Sanchez-Ruiz, J. M.; Eaton, W. A. *J. Phys. Chem. B* **2008**, *112*, 5938.
(26) Liu, F.; Gruebele, M. *J. Mol. Biol.* **2007**, *370*, 574.
(27) Thirumalai, D. *J. Phys. I* **1995**, *5*, 1457.
(28) Naganathan, A. N.; Muñoz, V. *J. Am. Chem. Soc.* **2005**, *127*, 480.
(29) Robertson, A. D.; Murphy, K. P. *Chem. Rev.* **1997**, *97*, 1251.
(30) Plaxco, K. W.; Simons, K. T.; Baker, D. *J. Mol. Biol.* **1998**, *277*, 985.
(31) Perham, R. N. *Annu. Rev. Biochem.* **2000**, *69*, 961.
(32) Neuweiler, H.; Johnson, C. M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18569.
(33) Zuo, G.; Wang, J.; Wang, W. *Proteins: Struct., Funct., Genet.* **2006**, *63*, 165.
(34) Badasyan, A.; Liu, Z.; Chan, H. S. *J. Mol. Biol.* **2008**, *384*, 512.
(35) Zhang, J.; Li, W.; Wang, J.; Qin, M.; Wang, W. *Proteins: Struct., Funct., Genet.* **2008**, *72*, 1038.
(36) Pitera, J. W.; Swope, W. C.; Abraham, F. F. *Biophys. J.* **2008**, *94*, 4837.
(37) Ferguson, N.; Sharpe, T. D.; Johnson, C. M.; Schartau, P. J.; Fersht, A. R. *Nature* **2007**, *445*, E14.
(38) Zhou, Z.; Bai, Y. W. *Nature* **2007**, *445*, E16.
(39) Ferguson, N.; Schartau, P. J.; Sharpe, T. D.; Sato, S.; Fersht, A. R. *J. Mol. Biol.* **2004**, *344*, 295.
(40) Huang, F.; Ying, L.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16239.
(41) Campos, L. A.; Liu, J. W.; Muñoz, V. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, E139.
(42) Spector, S.; Kuhlman, B.; Fairman, R.; Wong, E.; Boice, J. A.; Raleigh, D. P. *J. Mol. Biol.* **1998**, *276*, 479.
(43) Spector, S.; Raleigh, D. P. *J. Mol. Biol.* **1999**, *293*, 763.

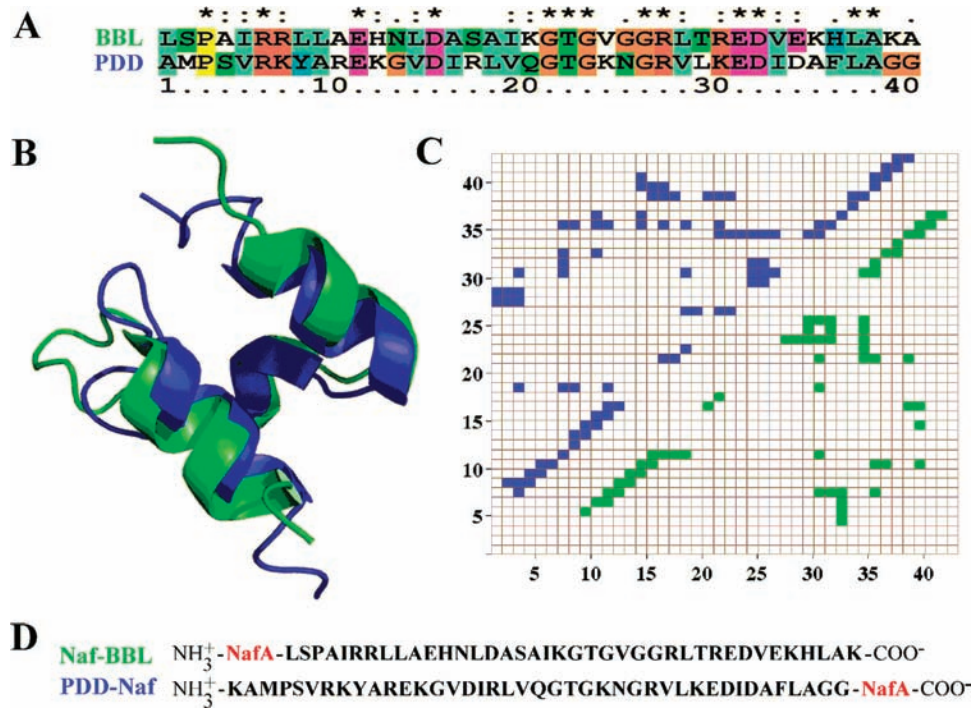

Figure 1. Sequence, structure, and contact-map comparison of BBL (green; 2CYU) and PDD (blue; 2PDD). NafA stands for nafthyl-alanine. The same color code (green/blue for BBL/PDD) is maintained in the following figures.

of PDD also suggest slower folding kinetics[44] and a marginal thermodynamic free-energy barrier.[45] Here we perform an extensive kinetic and equilibrium characterization of PDD folding to compare the various indicators of downhill folding on two microsecond folding proteins of the same size and structure.

Materials and Methods

Protein Synthesis and Buffers. Protein synthesis, purification, and analysis of PDD were carried out as described in Garcia-Mira et al.[12] The Naphthyl-alanine was placed on the C-terminus rather than on the N-terminus used for BBL to facilitate solid-phase synthesis. All experiments, except FTIR, were carried out at pH 7.0 in 20 mM sodium phosphate buffer. Protein concentrations were estimated using the following extinction coefficients (in units of M^{-1} cm^{-1}) at pH 7.0: $\varepsilon_{280} = 5526$ for nafthyl-alanine and $\varepsilon_{280} = 1280$ for tyrosine.

Equilibrium and Kinetic Infrared Experiments. FTIR spectra were recorded in an Excalibur FTS-3000 Spectrometer (BioRad) at a sample concentration of 2.5 mM. Two CaF_2 windows separated with a 50 μm Teflon spacer were used as the sample cell. The sample buffer was prepared in 99.9% ^2H_2O at pH 7.0 20 mM sodium phosphate. The exchangeable protons in the protein sample were substituted with deuterium in two consecutive heating−lyophilization cycles. Samples for IR T-jump kinetic measurements were prepared exactly the same way. The IR T-jump instrument and kinetic experiments are as described before.[23,24]

Far-UV Circular Dichroism. Far-UV CD spectra were collected every 5 K with a 1 mm path length quartz cuvette in a Jasco-815 Spectropolarimeter coupled to a Peltier system. Protein concentrations were ~50 μM. Urea concentrations were measured using a refractometer.

Differential Scanning Calorimetry. DSC experiments were carried out with a VP-DSC calorimeter from MicroCal (Northampton, MA) at a scan rate of 1.5 K/min. Protein solutions were prepared by exhaustive dialysis against the buffer. Calorimetric cells with a volume of ~0.5 mL were kept under an excess pressure of 30 psi to prevent degassing during the scan. The protein concentrations were in the range of 0.2−0.8 mg/mL. Absolute heat capacity values were determined from the concentration dependence of the apparent heat capacities, as previously described.[46,47] All thermograms shown and analyzed in this work are plots of absolute heat capacity versus temperature.

Model Fitting and Parameters. The best fit to the DSC thermogram with the variable-barrier model resulted in the following parameters: $\Sigma\alpha = 139.3$ kJ/mol, $f = 1$, $T_0 = 322.7$ K, β (or F_b) = 1.13 kJ/mol, native baseline slope =0.0307 kJ/(mol·K^2), and native baseline intercept (at 273 K) = 6.65 kJ/(mol·K). For details on the model and significance of the parameters refer to Muñoz and Sanchez-Ruiz.[14] It is important to note that the number of parameters used here is the same as that required for a simple two-state fit (i.e., 6). The relaxation rates as a function of temperature were fit to a phenomenological 1D free-energy surface model to extract estimates of kinetic free-energy barriers based on the curvature of the rate-temperature plot.[21] Following the prescription described before,[24] we fixed the following empirical parameters: entropy cost per residue at 385 K ($\Delta S_{res} = 16.5$ J/(mol·K)), heat capacity change per residue ($\Delta C_p = 8$ J/(mol·K)), and curvature of the heat capacity functional ($\kappa_{\Delta Cp} = 4.3$). The final floating parameters obtained from the fit to the rate data are as follows: the curvature of the enthalpy functional ($\kappa_{\Delta H} = 2.09$), the enthalpy change per residue at 385 K ($\Delta H_{res} = 5.88$ kJ/mol), pre-exponential at 333 K ($k_0 = 10^{3.76}$ s^{-1}), and activation energy per residue for the pre-exponential term ($E_{a,res} = 1.27$ kJ/mol).

Results and Discussion

For this work we chemically synthesized a version of PDD that is equivalent to that studied by Raleigh and co-workers[42] but with the two protein ends free and a nafthyl-alanine attached

(44) Ferguson, N.; Sharpe, T. D.; Johnson, C. M.; Fersht, A. R. *J. Mol. Biol.* **2006**, *356*, 1237.
(45) Naganathan, A. N.; Sanchez-Ruiz, J. M.; Muñoz, V. *J. Am. Chem. Soc.* **2005**, *127*, 17970.
(46) Kholodenko, V.; Freire, E. *Anal. Biochem.* **1999**, *270*, 336.
(47) Guzman-Casado, M.; Parody-Morreale, A.; Robic, S.; Marqusee, S.; Sanchez-Ruiz, J. M. *J. Mol. Biol.* **2003**, *329*, 731.

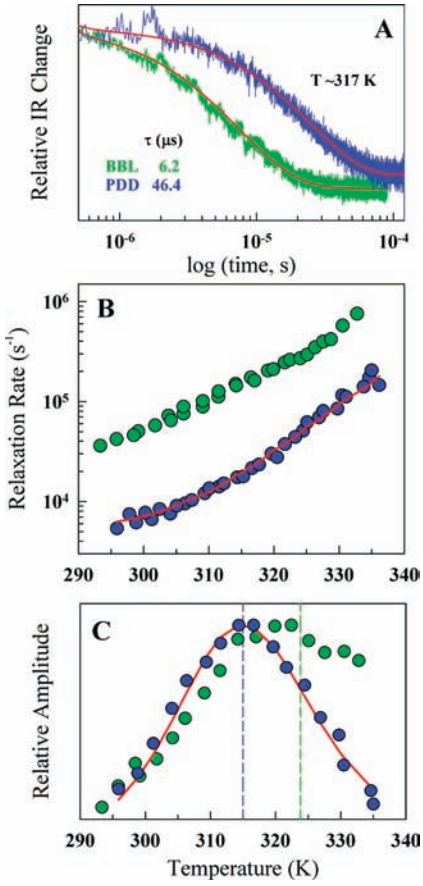

Figure 2. Microsecond folding kinetics of PDD (blue) and BBL (green). (A) Relaxation decays from IR T-jump measurements at a common temperature of 317 K. Single-exponential fits are shown in red. (B) Rate−temperature plots of the two proteins highlighting the differences in relaxation rate as well as in its temperatutre dependence. The red curve is the fit to the 1D free-energy surface model of Naganathan et al.[21] (C) Amplitude from kinetic measurements. The dashed blue and green lines signal the T_m obtained from equilibrium measurements and the red curve the fit from the 1D free-energy surface model.

to the C-terminus (from here onward PDD). The sequence of the exact PDD and BBL variants used here are shown in Figure 1D. Both proteins have free N- and C-termini, but PDD is two residues longer.

As a first step we measured the folding relaxation rate of PDD as a function of temperature using laser-induced temperature-jump experiments with infrared detection (IR) to monitor the peptide bond amide-I band at 1632 cm^{-1}. This technique provides direct information about protein secondary structure.[48] The PDD folding relaxation after T jumps of ~10 K produces single-exponential decays (Figure 2A) with relaxation times ranging from 200 to 5 μs in the explored temperature range of 295−335 K (Figure 2B). This temperature range covers most of the PDD unfolding transition, as indicated by the bell-shaped amplitude of the IR kinetic phase as a function of temperature (Figure 2C). The amplitude maximum at ~315 K provides a kinetic estimate of the midpoint denaturation temperature (T_m) for PDD as monitored by the infrared signal. This kinetic estimate agrees very well with that obtained from the derivative of the Fourier-transformed IR equilibrium unfolding curve (Figure S2, Supporting Information), demonstrating the agreement between equilibrium and kinetic measurements. The data

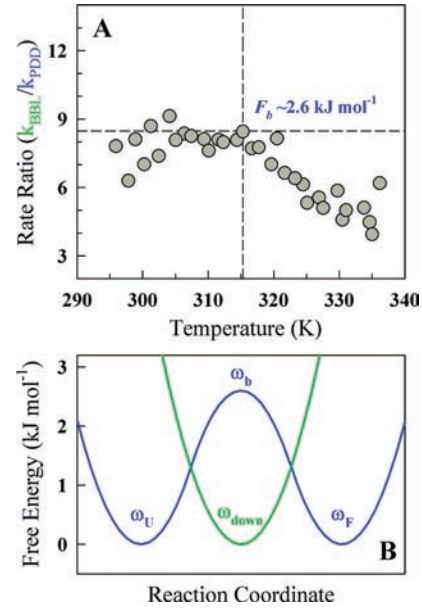

Figure 3. Comparing relaxation rates. (A) Ratio of the relaxation rates of the two proteins as a function of temperature. The dashed lines correspond to a rate ratio of ~8 at the apparent T_m of ~315 K. (B) Idealized folding free-energy surfaces for BBL and PDD representing the same curvature for the single downhill well (ω_{down}) and the curvature of the native well (ω_F), unfolded well (ω_U), and inverted barrier top (ω_b) for marginal barrier-crossing folding.

available on another PDD construct[42,43] allow us to test whether the folding behavior of PDD is affected by specific choices made in variant design (i.e., the fluorescent probe and the presence/absence of charges on the protein ends). In this regard, it is noteworthy that there is excellent agreement between our T-jump relaxation rates and the rates estimated by Raleigh and coworkers in their variant from NMR line-broadening analysis (Figure S1, Supporting Information). The equilibrium thermal unfolding of both variants is also very similar: the unfolding curves at the same experimental conditions have identical apparent cooperativity and pretransition baseline and minimal differences in T_m (~3K) (Figure S3, Supporting Information). These results indicate that the folding properties of PDD, from both equilibrium and kinetic experiments, are essentially unperturbed by the chemical differences between variants.

On the other hand, the comparison between the folding kinetics of PDD and BBL at the same experimental conditions highlights three main differences: (1) PDD folds/unfolds with a rate that is ~8 times slower near the denaturation midpoint (Figure 2A), (2) the logarithm of the folding relaxation rate for BBL increases almost linearly with temperature, whereas for PDD it exhibits slightly more curvature (Figure 2B), and (3) the PDD unfolding curve, as judged by the amplitude of the IR T-jump signal as a function of temperature, is sharper and with a lower apparent T_m for PDD (Figure 2C).

The differences in relaxation rate are more apparent by plotting them as a ratio (Figure 3A). The BBL/PDD ratio does indeed show that the maximum rate difference occurs near the denaturation midpoint, decreasing at both lower and higher temperatures due to the more curved temperature dependence for PDD. In general, the ratio between rates indicates that PDD crosses a higher folding barrier since it is reasonable to assume that the pre-exponential term is very similar for two homologous proteins with the same size and 3D structure. However, interestingly, the rate factor of ~8 at the denaturation midpoint

(48) Callender, R.; Dyer, R. B. *Curr. Opin. Struct. Biol.* **2002**, *12*, 628.

renders two significantly different scenarios for the interpretation of equilibrium experiments depending on whether BBL folds downhill or over a significant barrier. If BBL crosses a barrier of >3*RT*, the ~2*RT* higher barrier expected for PDD from the rate factor should not affect its equilibrium behavior. On the other hand, the 8-times slower rate of PDD is highly significant in the downhill regime and should translate into a weakening of all its equilibrium manifestations. This argument should be true no matter what level within the downhill regime is BBL. There are, however, additional considerations for the case in which BBL reaches the global downhill (one-state) limit, because there are some fundamental differences between diffusing on concave or convex surfaces. In the simplest terms, a global downhill relaxation can be described with the analytical expression for diffusion on a harmonic well: $k^{down} = \omega^2 D/RT$, where ω represents the well curvature and D the diffusion coefficient.[23] The equivalent Kramers expression for the relaxation rate back and forth over a barrier is $k^{act} = k_U + k_F = 2[\omega_0\omega_b D/2\pi RT]\exp(-F_b/RT)$, where F_b is the free-energy barrier from either of the two ground states (U and F) at isostability conditions (denaturation midpoint), ω_b is the inverted curvature at the barrier top, and ω_0 is the well curvature for both U and F, which for simplicity we assume here to be the same (Figure 3B). Comparing the two expressions shows that even for the idealized situation in which all well curvatures and diffusion coefficients are identical, the pre-exponential term for an activated process is not equal to the downhill relaxation but π times slower. The implication is that in such case the ~8-fold rate ratio corresponds to a PDD barrier of only ~1*RT* (or 2.5 kJ/mol, as illustrated in Figure 3B). Nevertheless, such a smaller barrier for PDD should still produce large differences in equilibrium behavior given that the transition between global downhill and a 1*RT* barrier is in the range of maximal sensitivity.[10,49] Other possible dynamic effects, such as differences in well curvatures and a position-dependent diffusion coefficient, are extremely difficult to take into account but are also likely to partially compensate for each other; the one-state well at the denaturation midpoint may be broader than the barrier top but should be located halfway along the reaction coordinate and thus closer to the unfolded state (potentially faster D).

These arguments are confirmed by analysis of PDD folding kinetics with an empirical 1D free-energy surface model,[21] which has also been used before to interpret the kinetics of BBL[23] and gpW.[24] Fitting the BBL and PDD rate versus temperature data together to this model with a common (constant) diffusion coefficient[21] reproduces the PDD rate data (both magnitude and curvature, see red line in Figure 2B) and the sharper amplitude of PDD (red line in Figure 2C) very well producing a barrier of 1*RT* (red line in Figure 2B). In other words, the slower rate, more curved temperature dependence and sharper amplitude for PDD are all consistent with a 1*RT* barrier.

Building on this idea, we investigated the equilibrium unfolding properties of PDD to compare them with those of BBL on equal footing. Proteins in the downhill regime unfold progressively (more or less depending on how close they are to the global downhill limit), visiting increasingly unstructured conformational ensembles as the denaturational stress rises.[7] This is manifested in a variety of equilibrium unfolding signatures, several of which can be simply detected (if the protein is α-helical) monitoring helix melting by far-UV circular dichroism (CD).[15] The CD unfolding curves of BBL and PDD show clear differences along these lines (Figure 4). The slope of the CD signal at 222 nm in the range of temperatures prior to the main transition (pretransition slope) is an indicator of progressive helix fraying. Figure 4A shows that there is helix fraying on the pretransition of both proteins, but the slope is much steeper for BBL, in compliance with the downhill scenario. The unfolding transition of PDD is also significantly steeper, even though BBL and PDD have similar T_m, size, and 3D structure. The difference in broadness, which is clearly apparent in the first derivative of the non-normalized CD unfolding curve (Figure 4B) and also in the kinetic IR amplitude from T-jump experiments (Figure 2C), indicates that PDD's equilibrium unfolding is slightly more cooperative, as expected for an incipient barrier within the downhill scenario. Another indication that the equilibrium unfolding of PDD is more cooperative is the fact that the residual helical structure at high temperature for PDD is decreased relative to BBL (−3700 deg·cm^2/dmol versus −4600 for BBL, Figure 4A).

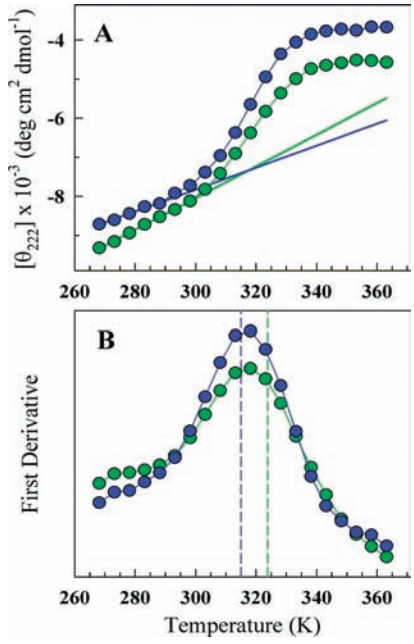

Figure 4. Far-UV CD thermal melts of PDD (blue) and BBL (green). (A) Thermal unfolding curves at 222 nm (circles) exhibiting different pretransition slopes (straight lines obtained from two-state-like fits). (B) Derivatives of the curves in panel A (circles) in which the equilibrium T_m values obtained from FTIR measurements are shown as dashed vertical lines for comparison.

The comparison between the CD and the FTIR equilibrium unfolding curves provides further evidence of the different equilibrium behavior of these two proteins. For BBL the apparent T_m determined from the maximum of the first derivative of the CD curve differs in ~5 K from that determined by IR experiments, whether from the kinetic amplitude from T-jumps or from equilibrium FTIR experiments (dashed green line in Figure 4B). This probe dependence is another equilibrium signature of downhill folding[12] that reflects that in BBL the native helices melt at apparently lower temperature when monitored by CD (requires helices of at least 3−4 residues[50]) than by IR, which reports the average hydrogen-bonding status of single peptide bonds. However, on PDD the agreement between CD and IR points again to less helix fraying with temperature.

(49) Naganathan, A. N.; Doshi, U.; Fung, A.; Sadqi, M.; Muñoz, V. *Biochemistry* **2006**, *45*, 8466.

(50) Chen, Y.; Yang, J. T.; Chau, K. H. *Biochemistry* **1974**, *13*, 3350.

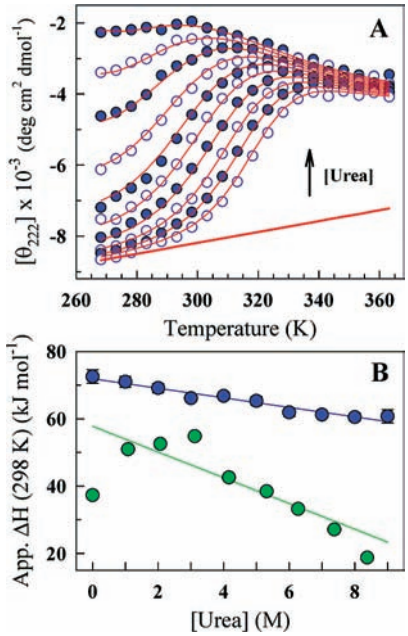

Figure 5. Double-perturbation test for PDD. (A) Thermal unfolding curves at 222 nm and at various urea concentrations (circles). The curved red lines are fits from a global two-state model with a single global native baseline (straight red line). (B) The apparent enthalpy at 298 K (blue circles) together with the prediction from the global two-state fit shown in A (blue straight line). The data for BBL obtained from Naganathan et al.[15] is shown in green for comparison, also including as a straight line the prediction from the global two-state fit to the original BBL data.

Another simple equilibrium test for the downhill scenario is the double-perturbation experiment.[16] Here the idea is to combine two different denaturing agents (e.g., temperature and chemical denaturant) to determine whether their coupling follows the simple Maxwell relationship defined by a two-state model. For a global downhill protein the shifting conformational ensemble induced by one denaturing agent changes the sensitivity to the other agent, and thus, the overall unfolding behavior deviates from the linear dependence between the effects of the two agents (Maxwell relationship) expected for two-state folding.[16] This method was developed using BBL as a model system, but it has also been recently applied to other proteins such as gpW[24] and a WW domain.[51] The temperature−urea equilibrium unfolding of PDD is shown in Figure 5A together with the result from a global two-state fit in which the native baseline is represented by the red line. The global two-state fit for PDD shows some systematic deviations, but these are of much smaller magnitude than for BBL or gpW. Likewise, the apparent unfolding enthalpy at 298 K (obtained from individual two-state fits) as a function of urea concentration is larger than for BBL and even gpW (although this protein is much larger) and closely follows the straight line expected from the simple Maxwell relationship for two-state folding (Figure 5B). In this case, the results of the double perturbation test are strongly modulated by the fact that PDD has a very small change in heat capacity upon unfolding (ΔC_p), as it is common for other thermophilic proteins.[52,53] With an intrinsically small temperature dependence on ΔH it is also more difficult to detect any second-order deviations. Therefore, what we can say at this point is that the double-perturbation test is less sensitive for PDD than for BBL, and within such reduced sensitivity, PDD appears

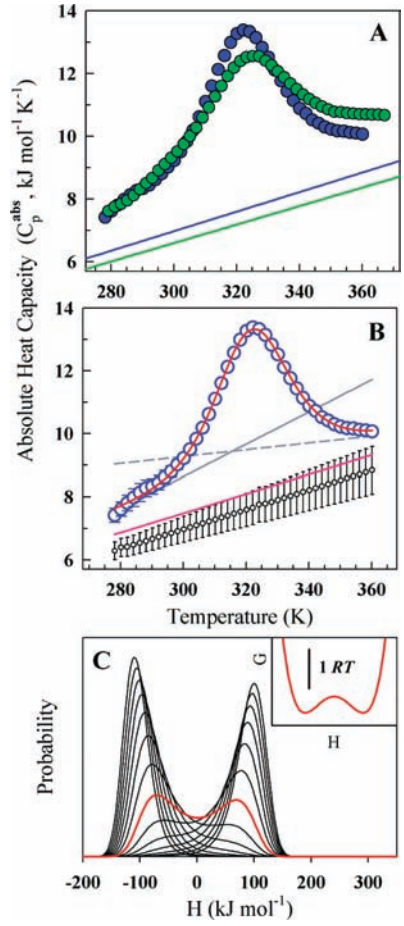

Figure 6. Differential scanning calorimetry of PDD (blue) and BBL (green). (A) Direct comparison of the absolute heat capacity of the two proteins (circles) plotted with the Freire's empirical native baselines (straight lines). (B) The best fit of the PDD data (open circles) to the variable barrier model (red curve) together with the fitted native baseline (magenta). Freire's native baseline with the error estimated from experiments at various protein concentrations is shown for comparison (black symbols). The gray dashed and continuous lines are the unfolded and folded baselines, respectively, obtained from a two-state fit. (C) The probability density from the best fit to the variable barrier model (black curves). The red distribution is the corresponding probability density at $T_0 \approx 323$ K. (Inset) Free-energy surface at T_0 obtained from the best fit.

to be compatible with conventional activated folding according to this test. Since kinetic and other equilibrium criteria point to a marginal barrier for PDD, it is possible that the double-perturbation test requires a significant ΔC_p and/or may be only sensitive to the lowest end of the downhill folding scenario. Along these lines, the linear dependence of the apparent unfolding enthalpy on chemical denaturant concentration measured in the double-perturbation test would not necessarily imply a high folding barrier, but the observation of curvature (e.g., green symbols and line in Figure 5B) would be a strong thermodynamic indicator of downhill folding.

The ultimate analysis of thermodynamic folding barriers is provided by differential scanning calorimetry (DSC) because the thermogram obtained with this technique is directly connected to the partition function for protein unfolding.[54] Figure 6A shows the DSC thermogram of PDD in absolute heat

(51) Cobos, E. S.; Iglesias-Bexiga, M.; Ruiz-Sanz, J.; Mateo, P. L.; Luque, I.; Martinez, J. C. *Biochemistry* **2009**, *48*, 8712.
(52) Kumar, S.; Nussinov, R. *Biophys. Chem.* **2004**, *111*, 235.
(53) Razvi, A.; Scholtz, J. M. *Protein Sci.* **2006**, *15*, 1569.
(54) Freire, E.; Biltonen, R. L. *Biopolymers* **1978**, *17*, 463.

capacity units compared to that of BBL obtained at the same experimental conditions. The DSC experiments show clearly again that the two proteins have very similar T_m, as indicated by the thermogram maxima. Both thermograms are broad, with heat capacity values at the lowest temperatures that are well above the empirical baselines obtained from Freire's correlation, and also have a much steeper temperature dependence.[55] This indicates that there are large energetic fluctuations even when the two proteins are supposed to be on their native state (below 290 K). This is consistent (although not exclusively) with both BBL and PDD being within the downhill folding scenario.[15] However, the thermogram for PDD is still significantly sharper: the maximum reaches higher heat capacity values, and at low temperatures it is closer to the empirical baseline. This thermogram can be well fitted to a simple two-state model, but the "native" and "unfolded" baselines obtained from the fit cross in the middle of the transition and thus are unphysical (gray lines in Figure 6B). The implication behind the crossing DSC baselines is that the protein unfolding process is too broad to be adequately described as two-state-like unless a significant fraction of the energy fluctuations are trimmed by the fitted baselines. This observation is commonplace for proteins within the downhill regime,[15] as it has been illustrated with theoretical calculations.[49]

However, analysis of the DSC thermogram can be taken to a higher quantitative level and thus obtain an estimate of the free-energy surface for folding and the height of the barrier separating the native and unfolded conformational ensembles.[14] In principle, this information could be directly extracted from the Laplace transform of the thermogram.[56] Practically, this is done by fitting the DSC thermogram to a phenomenological one-dimensional free-energy surface represented by a quartic (Landau) polynomial in which the unfolding enthalpy is the order parameter.[14] This procedure has been shown to produce free-energy surfaces with two conformational ensembles separated by a barrier when applied to analyze DSC thermograms of several proteins that fold slowly and exhibit two-state unfolding thermodynamics.[45] Moreover, the thermodynamic barriers obtained from such DSC analysis correlated well with the experimentally determined folding rates.[45] In parallel, the DSC thermograms of several proteins that fold in few microseconds produce free-energy surfaces that range from globally downhill for BBL[15] to $\sim 2RT$ for the villin headpiece subdomain.[25] For PDD this procedure finds a Landau polynomial that fits the experimental DSC data very closely using a native baseline that is within the experimental error of the one predicted by Freire's empirical correlation (red curve and magenta line in Figure 6B). The resulting free-energy surface has a marginal folding barrier that separates two broad wells representing the unfolded and native ensembles. This incipient barrier reaches a maximal height of just $\sim 0.5RT$ at the denaturation midpoint (inset to Figure 6C). The thermodynamic F_b for PDD is so small that the native and unfolded ensembles move closer together as temperature approaches the denaturation midpoint, effectively merging at the T_m (Figure 6C). The probability density for thermal unfolding shown in Figure 6C shares the general features reported previously for other microsecond folding proteins.

Therefore, the thermodynamic barrier obtained from DSC for PDD is in excellent agreement with the kinetic estimate from the BBL/PDD rate comparison assuming that BBL folds globally downhill. Such quantitative agreement together with the small but consistent differences in equilibrium behavior between the two proteins (Figures 4−6) buttress our understanding of the downhill folding regime. Within this regime the (un)folding process is broad enough to produce characteristic deviations from a conventional equilibrium two-state description that can indeed be used as signatures of downhill folding. Furthermore, the BBL versus PDD comparison demonstrates that for ultrafast folding proteins of the same size and 3D structure the magnitude of such equilibrium deviations are proportional to the differences in rate and thus provide quantitative indicators of F_b. However, our analysis also highlights that once within the downhill regime it becomes difficult to directly correlate folding rates with differences in F_b, especially among structurally unrelated proteins, because ultrafast folding rates can be strongly influenced by dynamic effects on the pre-exponential factor (i.e., diffusion coefficient or shape of the free-energy surface). This issue becomes evident by comparing the four ultrafast folding proteins that have already been analyzed combining DSC and kinetics: BBL, PDD, gpW,[24] and villin.[25] BBL and villin fold in the same time scales ($\tau \approx 4-5\ \mu$s at 320 K) but according to all other tests are on the two extremes of the downhill regime. PDD and gpW fold significantly more slowly (33 and 22 μs at the same temperature). Yet, according to all other criteria these two proteins are in the mid-downhill range, exhibiting very marginal folding barriers. This suggests that such differences in rates are due to protein-specific dynamic effects with villin seemingly having a much faster pre-exponential factor than BBL and PDD. gpW, considering its larger size (62 residues versus 35 for villin and 40−42 for BBL-PDD) and the size-scaling effects expected for the pre-exponential,[21] would be somewhat in between.

Conclusions

We compared the kinetic and equilibrium behaviors of two structural homologues that fold in the few microseconds range: BBL and PDD. BBL folds about ∼8 times faster than its homologue. In principle, this observation could be interpreted with two possible scenarios. In one scenario both proteins fold within the downhill regime, with the faster folding BBL doing it in a globally downhill fashion and the slower PDD folding over an incipient barrier. In this case, both proteins should exhibit the equilibrium signatures of the downhill scenario in degrees proportional to the difference in folding rates. The other scenario considers that both proteins have classical activated folding with a difference in barriers of $\sim 2RT$, in which case they should not exhibit significant differences in equilibrium unfolding since BBL and PDD have essentially the same 3D structure, size, and T_m. From analysis of the equilibrium unfolding we observe the exact pattern predicted by the downhill scenario. PDD displays the equilibrium signatures of downhill folding but to a lesser degree than the faster folding BBL. An advantage of a comparative study between structural homologues is that it naturally eliminates size scaling and structural effects. Another important factor is that direct comparison permits reaching conclusions without performing a detailed quantitative analysis of the experimental data. In this case, the correlation between differences in rate and equilibrium provides very strong evidence that the two proteins fold within the downhill regime with BBL being at the one-state downhill limit. This equilibrium−rate connection is a feature of the downhill regime that is incongruous with interpretations of folding as a highly activated process,[39] and as such it attests the invalidity

(55) Gomez, J.; Hilser, V. J.; Xie, D.; Freire, E. *Proteins: Struct., Funct., Genet.* **1995**, *22*, 404.
(56) Kaya, H.; Chan, H. S. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 637.

of this conventional view to describe ultrafast protein folding. Strictly speaking, this conclusion applies to the BBL and PDD versions that we use for this study (Naf-BBL and PDD-Naf, see Figure 1). However, given that the relaxation rates we report for BBL and PDD are essentially the same as those obtained by other authors with various experimental methods and in variants with slightly different chemistry (see Figure S1, Supporting Information), we can invoke a simple rate criterion to make this conclusion general and independent of the protein variant and of the experimental conditions tested so far.

In addition, the combination of kinetic and equilibrium analyses of ultrafast folding structural homologues provides an empirical test for the validity and sensitivity of the various equilibrium signatures that have been proposed for the downhill folding regime. The procedure is conceptually identical to the analysis of rate-accelerating mutations in fast folding proteins.[57,58] The use of homologues offers the advantage of comparing natural proteins with large differences in sequence due to divergent evolution, whereas mutations provide the opportunity to further engineer folding barriers. These two approaches thus complement each other. From the existing data on several ultrafast folding proteins we can conclude that the thermodynamic free-energy surfaces obtained from DSC experiments are in reasonable agreement with the marginal barrier heights estimated using a variety of kinetic arguments that rely on rate comparisons. Moreover, we show here that the DSC analysis successfully detects differences in equilibrium unfolding of proteins that fold with slightly different rates but share size, 3D structure, and over 30% of their amino acid sequence. There is, however, an emerging trend in which the estimates of thermodynamic barriers seem to be slightly, but consistently, lower than the kinetic estimates. This could be error in either analysis or indicative of constitutive differences due to dynamic effects and/or surface roughness.[10] Another interesting result is that simple equilibrium methods, such as standard CD thermal unfolding curves, provide sensitive indicators of barrier heights within the downhill folding regime to the extent that they could be used to predict changes in ultrafolding rates resulting from mutation[26] or comparisons between structural homologues. We see, for example, that the degree of helix fraying of the "native" baseline, the broadness of the unfolding transition, and the degree of residual structure in the high-temperature unfolded ensemble correspond with the differences in relaxation rate between structural homologues. The double-perturbation analysis could turn out to be less sensitive, requiring that the protein undergoes a significant change in heat capacity upon folding or perhaps detecting only cases in which the thermodynamic barrier is nearly zero. In such case the observation of curvature in the correlation between the apparent equilibrium unfolding enthalpy and the concentration of chemical denaturant could be considered a strong indicator of global downhill folding, whereas the lack of curvature would still be consistent with the higher end of the downhill regime.

Finally, the comparison between kinetic and equilibrium experiments on ultrafast folders opens the opportunity to experimentally dissect the structural, sequence, and evolutionary determinants of folding free-energy barriers and cooperativity. One critical question is what makes PDD fold more slowly and cooperatively than its structural homologue BBL? These proteins have the same folding topology (Figure 1B), but their 3D structures have some differences in packing. The available structure of PDD is more compact and thus has more long-range contacts (see Figure 1C). In molecular simulations using Go models these differences in packing do indeed produce lower cooperativity for BBL relative to PDD.[33,34] There is similar packing variability between 3D structures of BBL variants, which also produce different free-energy barrier and cooperativity in simulations with Go models.[9] However, the critical difference is that the latter is not reflected on equivalent changes in relaxation rate between BBL variants (Figure S1, Supporting Information). It is thus unclear whether such differences in structure are indeed connected to different heights of the folding barrier or rather reflect alternative NMR structure calculation protocols.[10] At the level of the amino acid sequence there is a related pattern that stands out. The two helical regions of BBL have a high intrinsic propensity to form α-helix structure, especially the first one (25% according to the algorithm AGADIR[59]). The same regions in PDD have only marginal helical propensity (∼2%). The lower helical propensity of PDD implies a larger relative stabilization from nonlocal interactions, which provides a simple explanation for its increased folding cooperativity and incipient free-energy barrier. Another related pattern is the fact that BBL does not have any aromatic residues, whereas PDD has a tyrosine and a phenylalanine involved in many hydrophobic contacts.

Navigating the downhill folding regime via structural homologues may also shed light into the biological implications of any observed differences. Different degrees of downhill folding could be related to adaptation to different environments or may have important functional implications, especially if the broad conformational ensembles characteristic of downhill folding are exploited to perform roles as molecular rheostats.[49] The BBL versus PDD story may have components of both since these domains perform similar roles in two different multienzymatic complexes and come from microorganisms that live at quite different temperatures (310 K for *E. coli* and 333 K for *B. stearothermophilus*, respectively). Along these lines it is interesting to note that the folding relaxation rates of BBL and PDD happen to be nearly identical at their respective physiological temperatures (∼100 000 s^{-1}, see Figure 2B). However, this coincidence could be just an accident. Further work combining equilibrium, DSC, and kinetic measurements on equivalent domains from the two multienzymatic complexes of other organisms will be critical to further explore these interesting issues.

Acknowledgment. This work was supported by the Marie Curie Excellence Grant MEXT-CT-2006-042334 to V.M. and grants BFU2008-03237 (V.M.), BIO2009-09562 (J.M.S.R), and CONSOLIDER CSD2009-00088 (V.M. and J.M.S.R.) from the Spanish Ministry of Science and Innovation.

Supporting Information Available: Figures with the comparison between BBL and PDD rates measured on different variants by different authors with different methods, comparison between equilibrium unfolding curves for PDD variants, and FTIR experiments for BBL and PDD. This material is available free of charge via the Internet at http://pubs.acs.org.

JA103612Q

(57) Yang, W. Y.; Gruebele, M. *Biophys. J.* **2004**, *87*, 596.
(58) Liu, F.; Du, D.; Fuller, A. A.; Davoren, J. E.; Wipf, P.; Kelly, J. W.; Gruebele, M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 2369.
(59) Muñoz, V.; Serrano, L. *Nat. Struct. Biol.* **1994**, *1*, 399.