Supporting Information

Li *et al.* 10.1073/pnas.0802986106

SI Text

Fitting BBL Folding Dynamics Data to Kinetic Three-State Models. The thermodynamic three-state model (F-I-U) can be constructed with three different kinetic schemes according to the connectivity between the three species: on-pathway, off-pathway and triangular. The first two schemes are linear differing only in whether the intermediate or the unfolded state is the one placed in the middle. In the triangular scheme the three species are interconnected, resulting in 2 additional microscopic constants. The three schemes result in two observable phases and thus, in principle, could be compatible with the BBL data.

On-Pathway Model. The on-pathway model is implemented according to the reaction scheme:

$$U \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} F$$

where U, I, and F are the unfolded, intermediate and folded states, respectively. Ignoring the changes in heat capacity associated to the rates (which are small for a protein of this size, have very little effect on the small perturbations of laser T-jump experiments, and would require 4 additional and ill-defined fitting parameters) each microscopic rate constant (k_x) is determined by the preexponential factor (k_0) and two thermodynamic parameters: the activation enthalpy ($\Delta H^\dagger$) and activation entropy ($\Delta S^\dagger$) following the expression:

$$k_x = k_0 \exp\left(\frac{-\Delta G^\dagger}{RT}\right) = k_0 \exp\left(\frac{-\Delta H^\dagger}{RT}\right) \exp\left(\frac{\Delta S^\dagger}{R}\right)$$

where $\Delta G^\dagger$ is the activation free energy, and R and T are the gas constant and temperature, respectively. The resulting set of microscopic rate constants at each temperature is then used to build the 3×3 rate-matrix. The two nonzero eigenvalues of the rate matrix provide the rates for the two observable phases. From the eigenvectors we obtain the populations in equilibrium. The changes in population upon 10 K temperature jumps for each state in each of the two phases are obtained from the eigenvectors at the final temperature using the equilibrium populations at the initial temperature as initial conditions. The observable FRET and IR amplitudes are obtained as the sum of the signal weighted population changes of the three species. The relative FRET and IR signals for folded and unfolded states are set to 0 and 1, respectively. The relative FRET and IR signals for the intermediate are fitting parameters of the model.

A fully unconstrained fit involves 11 floating parameters (9 physical + 2 for the FRET and IR relative signals of the intermediate), but to avoid convergence problems due to the large degree of correlation between k_0 and $\Delta G^\dagger$ we fixed the preexponential to physically reasonable values. Fig. S2 shows the best fit results to the BBL rate data using a preexponential (k_0) of 10^9 s^{-1}, which results in highly activated kinetics for both phases. The fitting parameters are shown in Table S1. Fig. S2*A* shows the populations for the three states as a function of temperature as predicted by the model, which exhibits a maximum population for the intermediate at ≈305 K. Fig. S2*B* shows a free energy diagram of the three states and two transition states at the two temperatures in which either the FRET (blue) or IR (red) microsecond phase amplitude is maximal. This diagram shows that with this preexponential value both steps cross a barrier, but the U-I barrier is much smaller resulting in rapid equilibration between U and I, and slow equilibration between these two with F. Fig. S2 *C* and *D* show that the model and parameters reproduce *qualitatively* the sign of the signals and the lack of nanosecond phase for the IR probe.

Fig. S2*E* shows the relaxation rates for the nanosecond FRET (green), microsecond FRET (blue) and microsecond IR (red) of BBL as a function of temperature together with the results from the fit shown as lines of the same color. Fig. S2*F* shows the relative amplitudes for the two FRET phases and the IR microsecond phase with the same color coding. This figure clearly shows that to reproduce the rates and qualitatively the FRET and IR decays the three-state model requires that the maximum of the IR microsecond phase amplitude occurs at lower temperature than the maximum of the FRET microsecond phase amplitude (even if minimally after the fit pushes it to the limit). In this fit the statistical weights for the amplitudes were minimized to obtain a reasonable fit to the rates (Fig. S2*E*). If the amplitude weights are increased beyond 3% of the weights for the rates the fit becomes unstable (it does not converge properly) progressively degrading its agreement with the rates without obvious improvement in the agreement with the experimental amplitudes.

We also performed a fit to the same model using a fixed preexponential term (k_0) of 4.31×10^6 s^{-1}, which is similar to the nanosecond FRET relaxation rate. This fit produced similar results. However, in this case the barrier between U and I disappears becoming a downhill process (see Fig. S3). This is analogous to a two-state model in which U changes structurally with temperature with a relaxation rate that sets the preexponential for the slow phase. The fitted parameters are shown in Table S2.

Off-Pathway Model. The off-pathway model is implemented according to the reaction scheme:

$$I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} U \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} F$$

where U, I, and F are again the unfolded, intermediate and folded states, respectively. The thermodynamic description for the microscopic rates is the same as the one described above for the on-pathway model, and so are the procedures to calculate populations, amplitudes and rates. The fitting of this model to the BBL data with a fixed preexponential (k_0) of 10^9 s^{-1} produced results identical to those shown in Fig. S2 for the on-pathway model with the parameters shown in Table S3. It is easy to see that the parameters for k_1 and k_{-1} in the table match the reverse on-pathway parameters (k_{-1} and k_1), k_{-2} is almost identical in both models, and k_2 for the off-pathway model is equivalent to $k_1.k_2/k_{-1}$ in the on-pathway model, thus satisfying the thermodynamic relationships. The two models produce identical results because the large difference between the rates for the fast (≈150 ns) and slow (100 to 15 μs) FRET phases imply that the fast phase is the rapid reequilibration between U and I, whereas the slow phase is the equilibration of F with the other two. The slow disappearance of F upon a T-jump is rapidly converted onto the proper U/I ratio regardless of whether the species connected to F is I or U.

Triangular (or Fast-Track) Model. The triangular model is a composite of the off-pathway and on-pathway models so that the

three states are directly connected resulting in 6 microscopic rate constants: k_1 and k_{-1} connecting U and I (the same as in both on- and off-pathway models), k_2 and k_{-2} connecting I to F (equivalent to the on-pathway rates), and k_3 and k_{-3} connecting U with F (equivalent to k_2, k_{-2} in the off-pathway model). The large number of correlated parameters in this model seriously hampers the convergence of the fits, which become undetermined unless specific experimental information about the direct fluxes to U and I is provided. This information cannot be obtained from T-jump experiments.

We thus did not fit the data directly to the triangular model. However, we did perform extensive numerical simulations using the parameters from the on-pathway and off-pathway fits and determined that for the BBL data the predictions of rates and amplitudes for this model are identical to those shown above for on-pathway and off-pathway models. Again the reason is the large separation of time scales between the U-I equilibrium and the equilibration between these two and F that emerges from two experimental rates differing by approximately a factor of 100. The same conclusion is obtained directly using the well-known analytical solution for the three kinetic three-state models (1).

Fitting BBL Folding Dynamics Data to a Kinetic Two-State Model Plus a Nanosecond FRET Process Unrelated to Folding. The alternative to the three-state models that could explain the observation of two phases in the FRET experiments and one by IR is to assume that the nanosecond FRET phase reports end to end fluctuations in the protein that are not directly related to folding. In this case the relaxation time and amplitude of the nanosecond FRET phase are not directly connected to the folding process, and thus can be treated independently. The microsecond phase is the one that monitors true folding, which could be two-state like or downhill. For the two-state case the equilibrium unfolding curve by IR should correspond exactly with the folding-unfolding process (IR does not detect the nonfolding nanosecond phase), and its fit to a two-state model should give the populations for F and U as a function of temperature. The differences between the equilibrium IR and FRET curves are assumed to arise from the nanosecond process that is superimposed to folding in the FRET experiment. This model predicts that the amplitude of the microsecond FRET phase (unlike the equilibrium curve, the kinetic FRET phase is free from signal contamination) should have the same maximum and broadness than the microsecond IR phase. The nanosecond phase will have its amplitude completely determined by the component of the equilibrium FRET curve necessary to correct the true folding curve. Fig. S4 shows the best fit to a model of these characteristics. Fig. S4*B* shows the equilibrium IR curve fitted to a two-state model with free-floating baselines (fitted baselines shown as lines) to obtain the global probability of U and F as a function of temperature. This fit produced T_m of 319.5 K and ΔH_m of 89.5 kJ·mol^{-1}.

Simulation of the kinetic amplitude after T-jumps of 10 K spanning the 280–340 K range is shown in Fig. S4*D* together with the experimentally determined IR amplitude. Fig. S4*D* demonstrates the very good agreement between IR kinetic amplitudes and equilibrium unfolding curve. Once the global probabilities are obtained from the IR equilibrium curve, a fit to the FRET equilibrium curve was carried out using these probabilities and letting the FRET signals for U and F float (Fig. S4*A*). The difference between constrained fit and experimental FRET unfolding curve is then treated as the best estimate of the changes in FRET signal due to the nanosecond nonfolding process (*Inset* to Fig. S4*A*). The model is then fully determined and it can be used to calculate the amplitudes for the microsecond and nanosecond FRET phases by simulating T-jumps of ≈10 K throughout the experimental temperature range (Fig. S4*C*). It is clear from Fig. S4*C* that the experimental amplitudes of the FRET microsecond and nanosecond phases are completely off from the predictions of the model. This discrepancy is not merely quantitative, but qualitative since the trends in the amplitudes of the two phases are reversed. Therefore, this conclusion is independent of the particular values for the U and F FRET signals, which are poorly determined due to the large discrepancies between the experimental FRET curve and the prediction from the two-state model (see blue circles and red curve in Fig. S4*A*). Different FRET values for U and F do change the shape of the predicted amplitude for the nanosecond phase, but the reversed trends for the two phases are maintained.

SI Experimental Procedures

Equilibrium Experiments. Naf-BBL and Naf-BBL with a dansyl-lysine at the C terminus were synthesized by solid phase methods as described before (2, 3). Equilibrium FRET experiments were performed at pH 7.0 with 20 mM sodium phosphate and 20 μM protein as described in ref. 2. Equilibrium FTIR experiments were performed on a thermostated FTS-300 IR Spectrometer (BioRad) at 2 cm^{-1} resolution using $CaCl_2$ windows and a 50-μm Teflon spacer. The protein sample was prepared with 2.5 mM protein in 99.9% ^2H_2O, 20 mM sodium phosphate buffer at pH 7.0.

Time-Resolved Experiments. The time-resolved FRET experiments were performed at pH 7.0 with 20 mM sodium phosphate, and 60 μM protein using our custom-built laser-induced temperature jump apparatus following the procedures described in ref. 4. Time-resolved IR experiments were performed on a version of the infrared nanosecond temperature-jump instrument originally developed by Feng Gai and coworkers (5) that has been custom-built in our laboratory. Briefly, the fundamental of a Continuum Surelite I-10 Nd-YAG laser run at 2 Hz is shifted to ≈1.9 μm with a 1 m path length Raman cell (Lightage) filled with a mixture of Ar and H_2 at 1,000 psi to heat deuterated water by vibrational excitation. Heating pulses of ≈20 mJ are used to generate ≈10 K jumps. A CW lead salt diode laser (Laser Components) and a mercury–cadmium–telluride (MCT) detector with 50MHz bandwidth are used to monitor time-dependent changes in IR absorption at 1,631.8 cm^{-1}. ^2H_2O buffer was used for background subtraction and as an internal thermometer to determine the T-jump magnitude. The IR cell and sample preparation were identical to the equilibrium FTIR experiments.

Model and Analysis of Time-Resolved Experiments. Using the property nativeness (n, or the probability of finding a given residue in native conformation) as reaction coordinate, the conformational entropy is defined using the Gibbs entropy formula and defining both the native and nonnative ensembles of a residue microcanonically:

$$\Delta S^{\mathrm{conf}}(n) = -NR[n\ln(n) + (1-n)\ln(1-n)] + N(1-n)\Delta S_{\mathrm{res}}^{n=0}$$

where the term in square brackets is set to 0 for $n = 0$ and $n = 1$, N is the number of residues in the protein, and $\Delta S_{\mathrm{res}}^{n=0}$ is the difference in conformational entropy between a residue being in any nonnative conformation and restricted in native dihedral angles.

The stabilization enthalpy is defined as:

$$\Delta H(n) = N\Delta H_{\mathrm{res}}[1 + (\exp(\kappa_{\Delta H} n) - 1)/(1 - \exp(\kappa_{\Delta H}))]$$

where ΔH_{res} is the difference in enthalpy per residue between the fully nonnative ($n = 0$, all residues in nonnative conformation) and native ($n = 1$) ensembles of the protein at the reference temperature of 385 K. The exponent $k_{\Delta H}$ defines how steeply the stabilization energy decays on nativeness, and thus the fraction of stabilization enthalpy at the intercept with the entropic

contribution to the free energy. The magnitude of the exponent reflects the ratio between nonlocal and local interactions producing from barrierless to two-state-like free energy surfaces as the exponent increases. The model includes a heat capacity term defined as:

$$\Delta C_p(n) = N\Delta C_{p,res}\,[1 + (\exp(\kappa_{\Delta C_p} n) - 1)/(1 - \exp(\kappa_{\Delta C_p}))]$$

where $\Delta C_{\mathrm{p,res}}$ is the difference in heat capacity per residue between $n = 0$ and $n = 1$, and $k_{\Delta C_{\mathrm{p}}}$ is the characteristic exponent as in Eq. **S2**. The solvation free energy is then defined as:

$$\Delta G_{\mathrm{solv}}(T, n) = \Delta C_{\mathrm{p}}(n)[(T - 385) - T\ln(T/385)]$$

The sum of Eqs. **S1** (multiplied by $-T$), **S2**, and **S4** produces the final free energy surface as a function of temperature.

For the analysis of BBL data we fixed the heat capacity parameters to those used previously for other fast-folding proteins: $\Delta C_{p,res} = 58$ J/(mol.K) per residue and $\kappa_{\Delta C_{\mathrm{p}}} = 4.3$ (6). This way the temperature dependent free energy surface is determined with only 3 fitting parameters. T-jump relaxation dynamics were calculated as diffusion on the free energy surface at the final temperature. The relaxation decays were calculated using the matrix method for the diffusion equation (7), the probability distributions calculated from the surfaces at the temperature before and after the T-jump, and $D = D_0 \exp(-E_a\,(1/T - 1/333)/R)$, where D_0 (the diffusion coefficient at 333 K) and E_a (the dynamic activation energy) are the 2 additional fitting parameters. The analysis was done by globally fitting the relaxation times and amplitudes of both FRET and IR data using a least-squares procedure and a simplex minimization algorithm. The 3 fitted parameters that determine the temperature dependent free energy surface were: $\Delta S_{\mathrm{res}} = 61$ J/(mol·K·res), $\kappa_{\Delta \mathrm{H}} = -0.78$, and $\Delta H^{385} = 72$ kJ/(mol·res). ΔH^{385} determines the magnitude of the overall change in enthalpy at 385 K, whereas $\kappa_{\Delta \mathrm{H}}$ determines its curvature as a function of nativeness, and thus the presence or absence of a folding barrier on the surface (6). Since the protein is in the global downhill regime, these parameters do not have any macroscopic meaning, but are just phenomenological parameters that reproduce the required shape for the free energy functional. The fitted kinetic parameters were $D_0 = 680\ n^2 \cdot \mathrm{s}^{-1}$ and $E_a = 60$ kJ/mol.

1. Ikai A, Tanford C (1973) Kinetics of unfolding and refolding of proteins. 1. Mathematical analysis. *J Mol Biol* 73:145–163.
2. Garcia-Mira MM, Sadqi M, Fischer N, Sanchez-Ruiz JM, Muñoz V (2002) Experimental identification of downhill protein folding. *Science* 298:2191–2195.
3. Naganathan AN, Perez-Jimenez R, Sanchez-Ruiz JM, Muñoz V (2005) Robustness of downhill folding: Guidelines for the analysis of equilibrium folding experiments on small proteins. *Biochemistry* 44:7435–7449.
4. Sadqi M, Lapidus LJ, Muñoz V (2003) How fast is protein hydrophobic collapse? *Proc Natl Acad Sci USA* 100:12117–12122.
5. Huang CY, *et al.* (2002) Helix formation via conformation diffusion search. *Proc Natl Acad Sci USA* 99:2788–2793.
6. Naganathan AN, Doshi U, Muñoz V (2007) Protein folding kinetics: Barrier effects in chemical and thermal denaturation experiments. *J Am Chem Soc* 129:5673–5682.
7. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J (2002) Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: Testing a 1-dimensional diffusion model for peptide dynamics. *J Phys Chem B* 106:11628–11640.

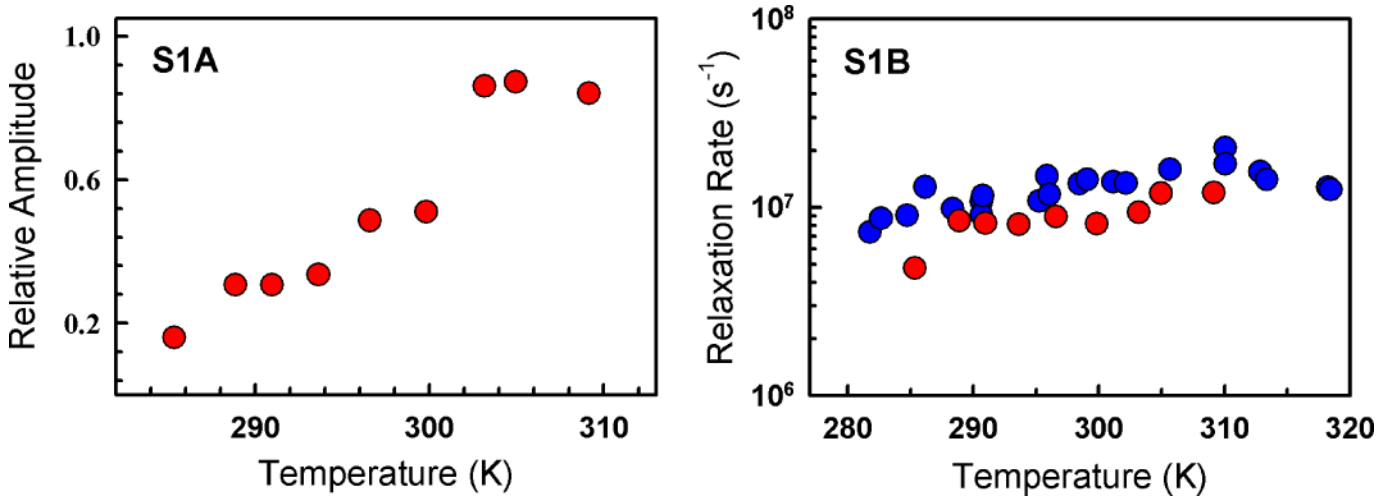

Fig. S1. (*A*) Normalized amplitude of the nanosecond FRET component of BBL at pH 7.0. (*B*) Experimental relaxation rate for the nanosecond FRET component at pH 7.0 (red circles) together with the previously reported single relaxation rate of acid denatured BBL at pH 3 (1) (blue circles).

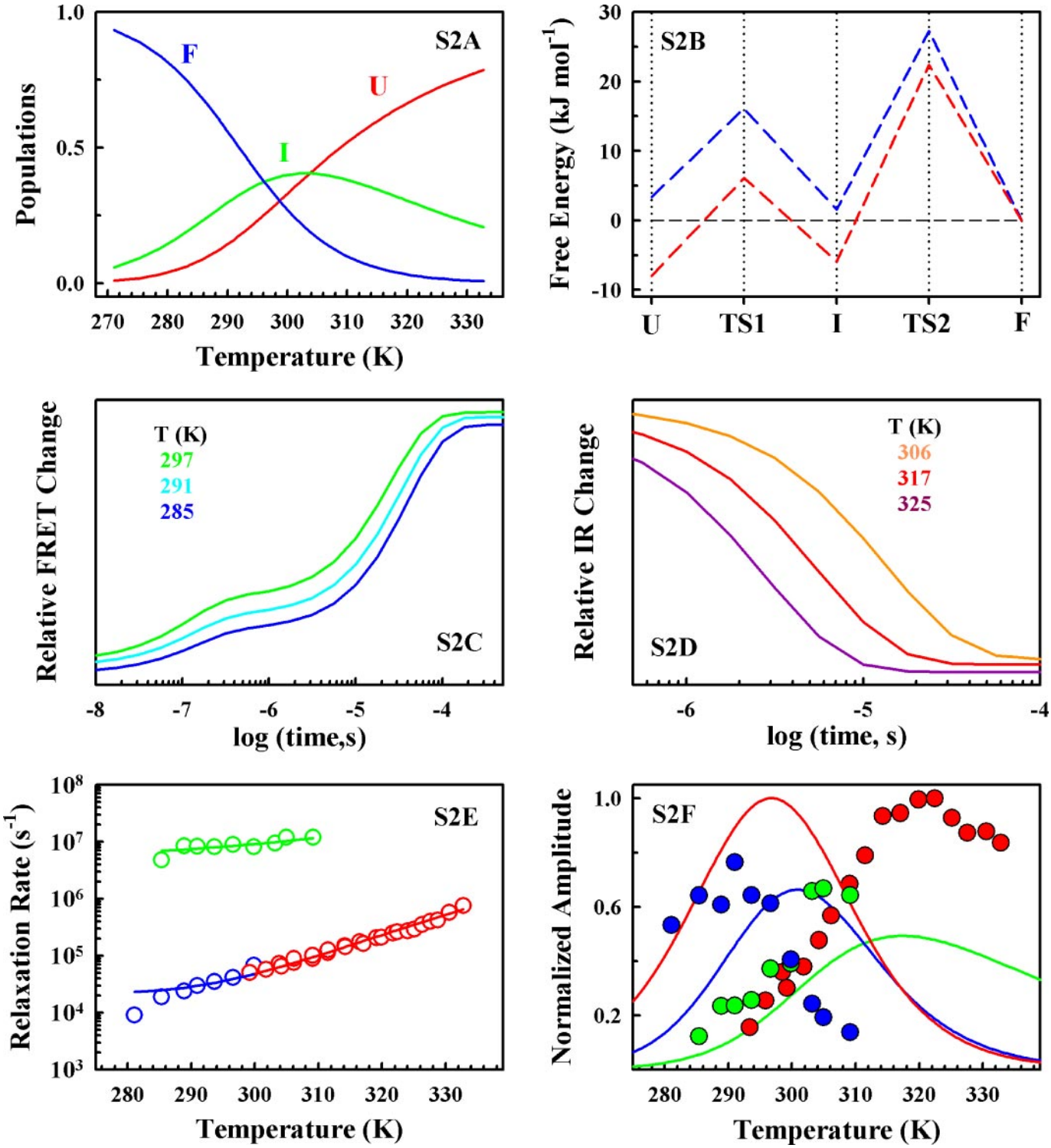

Fig. S2. (*A*) Populations for the three species from the best fit to the BBL rate data. (*B*) Thermodynamic diagram with the 3 species and 2 transition states at 290 K (blue) and 320 K (red). (*C*) FRET decays simulated using the best fit parameters. (*D*) The same as C but for IR decays. (*E*) Experimental rates (circles) and fits (lines) for the FRET nanosecond phase (green), FRET microsecond phase (blue), and IR microsecond phase (red). (*F*) Experimental relative amplitudes (circles) and predictions from the fit (lines) for the nanosecond FRET phase (green), microsecond FRET phase (blue) and microsecond IR phase (red).

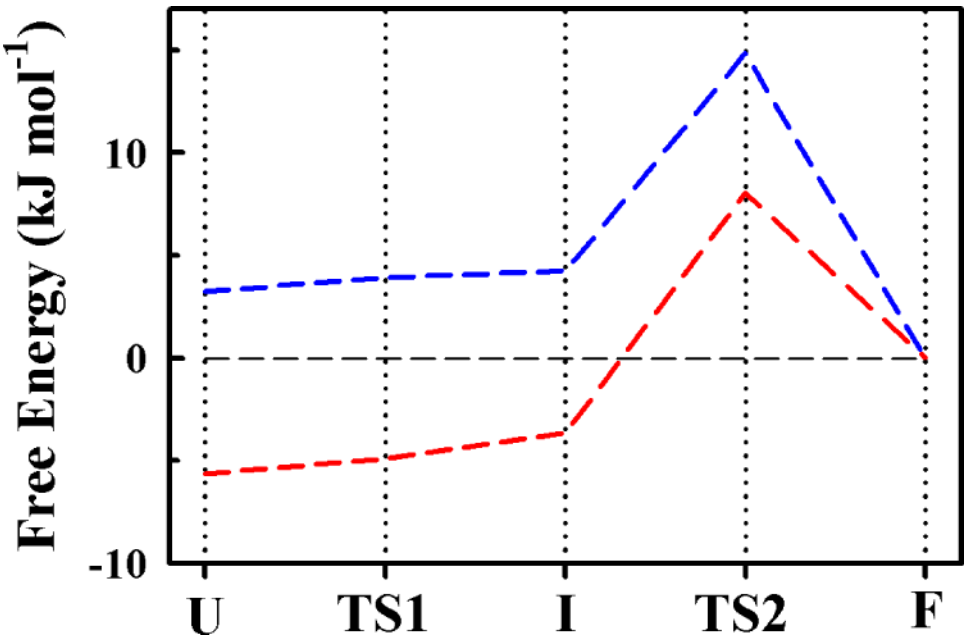

Fig. S3. Thermodynamic diagram for the three states and two transition states resulting from the on-pathway fit to the BBL data using a fixed preexponential 4.31×10^6 s^{-1}. The blue diagram corresponds to ≈290 K and the red diagram to ≈320 K.

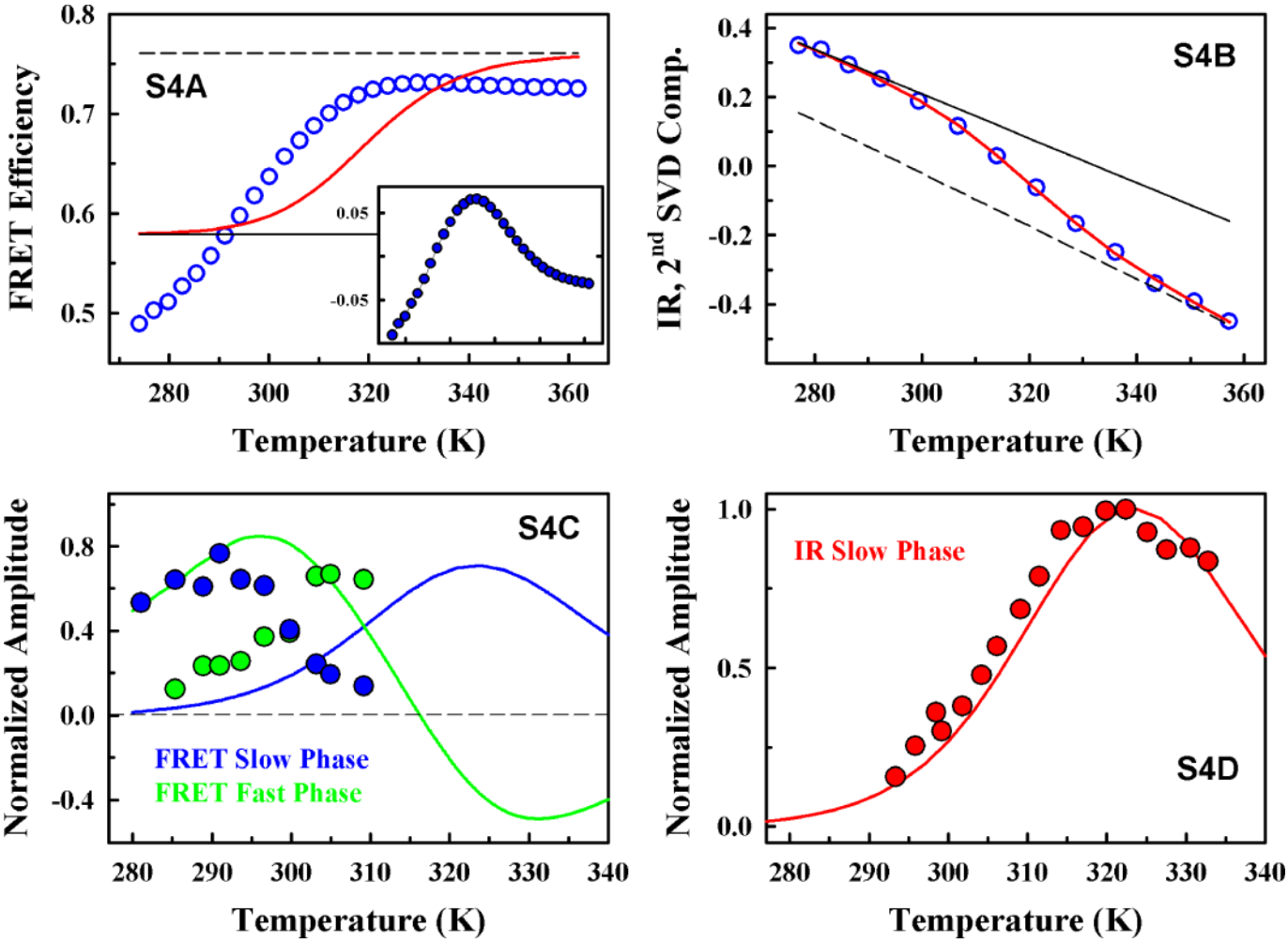

Fig. S4. Fit to a two-state model plus a nonfolding nanosecond FRET phase. (*A*) FRET equilibrium unfolding curve with best fit assuming $T_m = 319.5$ K and $\Delta H_m = 89.5$ kJ·mol^{-1}. (*Inset*) difference between the experimental equilibrium FRET curve and the prediction from the best fit using the parameters from the fit to the IR curve. This curve is taken as the best estimate of the changes in FRET signal due to the nanosecond nonfolding process. (*B*) IR equilibrium unfolding curve plus best fit to a two-state model, which rendered $T_m = 319.5$ K and a $\Delta H_m = 89.5$ kJ·mol^{-1}. The fitted baselines are shown as straight lines in the figure, with the continuous one for F and the dashed one for U. (*C*) Experimental FRET amplitudes (circles) for nanosecond (green) and microsecond (blue) phases together with the prediction from the model (lines) in the same color coding. (*D*) As in *C* but for the microsecond IR phase.

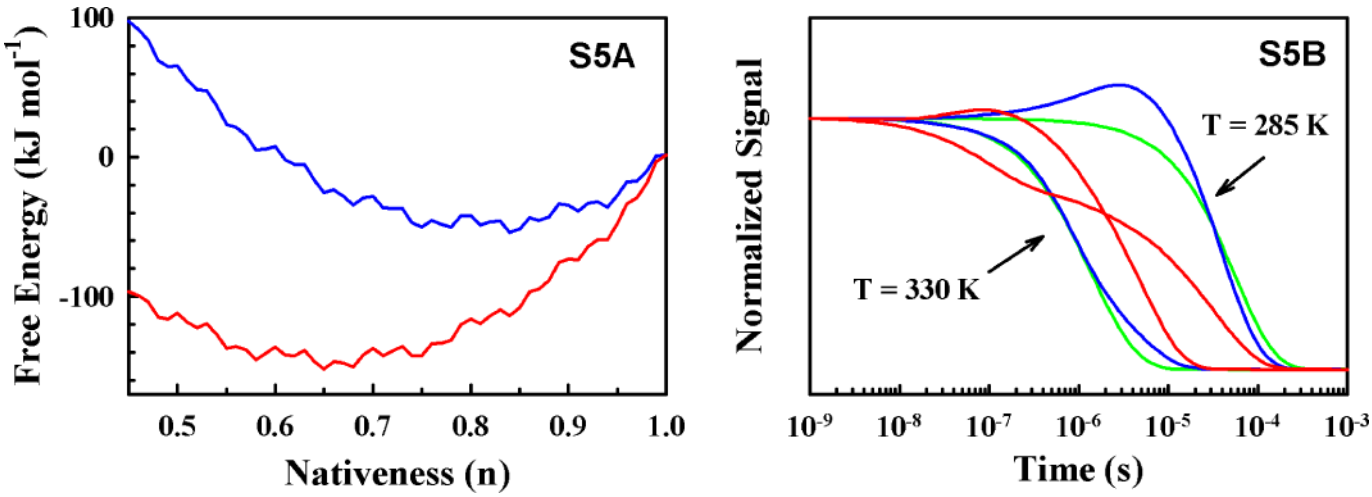

Fig. S5. (*A*) Globally downhill free energy surfaces for BBL at 285 K (blue) and 330 K (red) calculated with the fitted model (see below) and with 4 kJ/mol of added roughness. A blow out of the surface from $n = 0.45$ to $n = 1$ is shown to facilitate visualization of the degree of roughness used in the calculations. Roughness was introduced as a combination of dephased cosine and sine functions following the Zwanzig procedure [Diffusion in a rough potential. *Proc Natl Acad Sci USA* 85:2029–2030]. Particularly, we used the formula:

$$\Delta E^2(n) = \delta E(\cos(96x) + \sin(34x))$$

in which x bound within $[-1,1]$ is defined as $x \equiv 2n - 1$, and δE is 5 kJ/mol. For simplicity we assumed that the position of peaks and troughs in the landscape is temperature independent. The degree of roughness was chosen so that the temperature dependence of the FRET relaxation would match the experimental results assuming that the diffusion coefficient (D) only depends on solvent viscosity (i.e., setting the activation energy for D to 16 kJ/mol rather than the 60 kJ/mol required in fitting the smooth surface). (*B*) T-jump relaxations at 285 and 330 K obtained for the FRET probe (blue) and IR probe (red) on the rough surfaces depicted in panel S5A. The exponential relaxation decays obtained on the smooth surfaces are shown in green for comparison. The IR relaxation on the rough surface at 285 K is approximately biexponential with a fast phase of ≈35 ns that would be close to the detection limit of our IR T-jump instrument. The FRET relaxation on the rough surface at 285 K is ''squeezed'' (i.e., $\beta \approx 1.2$). At this temperature the relaxation times are already different for the two probes. At 330 K the relaxation times for the two probes become significantly more divergent. Both FRET and IR decays have multiple components being clearly nonexponential (compare with the green almost exponential decays). Fitting of the FRET and IR decays at 330 K to a stretched exponential function resulted in reasonable, but no perfect, phenomenological fits with $\beta \approx 0.75$ for both probes.

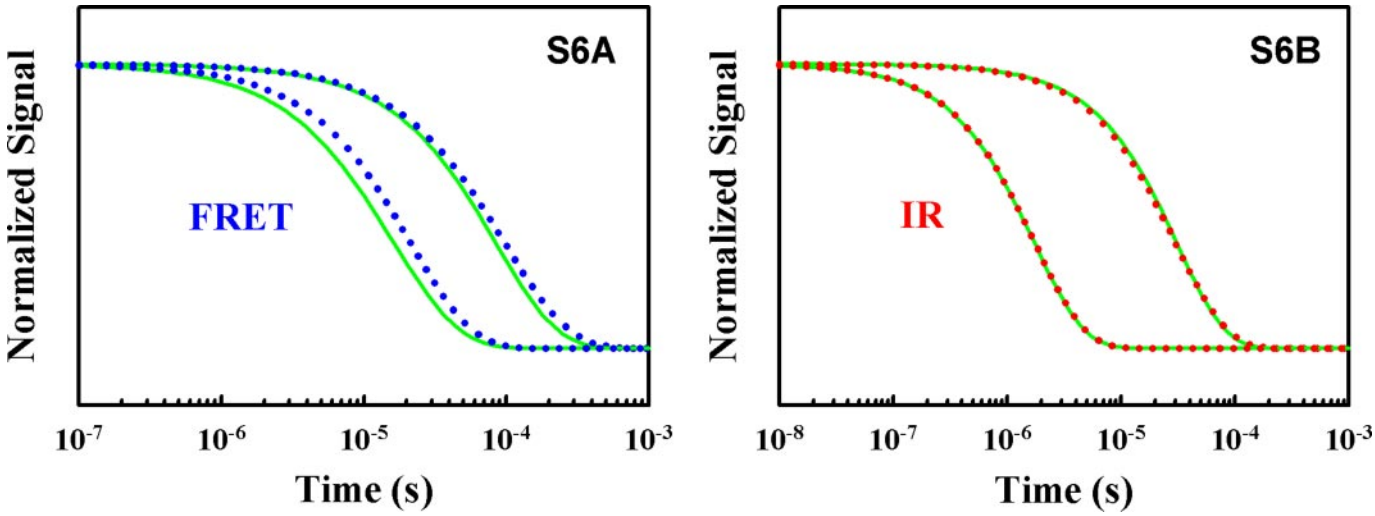

Fig. S6. Dynamic calculations on the downhill free energy surfaces of BBL produced by the fitted model (Fig. 3*C*) and adding a position dependent diffusion coefficient (*D*). *D* was assumed to decrease linearly by a factor of 2.5 in the interval between $n = 0.64$ and $n = 0.85$. Therefore, in these calculations all values of *n* between 0 and 0.64 have the original fitted values (see below) and all values above $n = 0.85$ have the same *D*. The activation energy for *D* was decreased from 60 to 48 kJ/mol to reproduce the temperature dependence of the IR experiments. (*A*) FRET relaxations at the lowest and highest experimentally accessible temperatures for this probe (281 K and 300 K). (Green lines) original calculations with position independent *D*; (blue dotted lines) calculations with position dependent *D*. The position dependent *D* produces a slightly stretched exponential decay at the lowest temperature (curve to the right) and a ''squeezed'' exponential at the highest accessible temperature (curve to the left). (*B*) IR relaxations at lowest and highest temperatures accessible to this probe (293 and 333 K). (Green lines) original calculations with position independent *D*; (red dotted lines) calculations with position dependent *D*. The position dependent *D* reproduces the observed temperature dependence of the IR relaxation with a 48 kJ/mol activation energy, and results in still almost perfectly exponential decays (small deviations from the green curves).

Table S1. Fitting parameters for on-pathway model

Microscopic rate constants	$\Delta H^{\dagger}$ (kJ mol^{-1})	$\Delta S^{\dagger}$ (J mol^{-1} K^{-1})
k_1	0.0	−44.1
k_{-1}	38.7	83.3
k_2	0.0	−88.6
k_{-2}	74.0	161.3

†, transition state.

Table S2. Fitting parameters for on-pathway model with slow preexponential

Microscopic rate constants	$\Delta H^{\dagger}$ (kJ mol^{-1})	$\Delta S^{\dagger}$ (J mol^{-1} K^{-1})
k_1	0.0	−2.3
k_{-1}	8.4	30.2
k_2	0.0	−36.6
k_{-2}	80.3	226.0

†, transition state.

Table S3. Fitting parameters for off-pathway model

Microscopic rate constants	$\Delta H^{\dagger}$ (kJ mol^{-1})	$\Delta S^{\dagger}$ (J mol^{-1} K^{-1})
k_1	39.1	85.2
k_{-1}	−0.1	−44.8
k_2	−37.2	−212.0
k_{-2}	72.9	157.8

†, transition state.