Slow and Bimolecular Folding of a De Novo Designed Monomeric Protein DS119

Cheng Zhu,[†] Ziwei Dai,[‡] Huanhuan Liang,[†¶] Tao Zhang,[†] Feng Gai,[§*] and Luhua Lai[†‡*]
[†]BNLMS, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, Peking-Tsinghua Center for Life Sciences at College of Chemistry and Molecular Engineering, Peking University, Beijing, China; [‡]Center for Quantitative Biology, Peking University, Beijing, China; [§]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania; and [¶]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China

ABSTRACT De novo protein design offers a unique means to test and advance our understanding of how proteins fold. However, most current design methods are native structure eccentric and folding kinetics has rarely been considered in the design process. Here, we show that a de novo designed mini-protein DS119, which folds into a $\beta\alpha\beta$ structure, exhibits unusually slow and concentration-dependent folding kinetics. For example, the folding time for 50 μM of DS119 was estimated to be ~2 s. Stopped-flow fluorescence resonance energy transfer experiments further suggested that its folding was likely facilitated by a transient dimerization process. Taken together, these results highlight the need for consideration of the entire folding energy landscape in de novo protein design and provide evidence suggesting nonnative interactions can play a key role in protein folding.

INTRODUCTION

De novo protein design critically tests our current understanding of protein folding principles (1,2). In particular, it helps delineate the factors that govern the roughness of protein folding energy landscapes (3,4). For example, complex folding kinetics was observed for Top7, a computationally designed α/β protein with a novel $\beta\beta\alpha\beta\alpha\beta\beta$ fold (5). At least three distinct phases were required to describe the folding kinetics of Top7, and this kinetic complexity arose from the topology of Top7, which likely imposes a series of folding barriers (6,7). On the other hand, a de novo designed three-helix bundle protein α_3D was found to fold with an ultrafast rate (8). Molecular dynamic simulations suggested that the transition state ensemble of α_3D was highly heterogeneous and dynamic, allowing fast access to the native state via multiple pathways. In comparison to the folding kinetics of natural proteins, these examples highlight the difficulty in designing specific protein folding energy landscapes or pathways.

To further illustrate this point and to provide new insights into the relationships between protein sequences and folding kinetics, this study focuses on the folding mechanism of DS119, a de novo designed 36-residue protein (9). DS119 folds into a $\beta\alpha\beta$ structure with two parallel β-strands connected by a α-helix (Fig. 1). Because standalone $\beta\alpha\beta$ motifs have not been found in naturally occurring proteins, DS119 was designed rationally with negative design considerations and the introduction of a tryptophan pair. Further NMR studies indicated that DS119 is a well-folded monomeric protein, and its native structural ensemble is in good agreement with the targeted topology (9). In addition, thermal unfolding measurements showed that DS119 has an unusually high thermal melting temperature (T_m >80°C) at neutral pH. A computational study using both all-atom and coarse-grained models has estimated the folding time of DS119 to be 6 to 10 μs (10). A priori estimate of the folding rate based on the Plaxco-Simmons-Baker correlation has led to a similar result (128 μs) (11). Thus, initially we employed an infrared (IR) temperature-jump (T-jump) technique (12) to investigate the folding dynamics of this putative fast folder. In contrast to our expectation and predictions from empirical folding-rate models, there were no detectable population relaxation signals occurring at the 10 ns to 1 ms timescale, suggesting that DS119 is a slow folder compared to natural proteins with similar sizes, and the folding process probably involves either a high free energy barrier or a new mechanism. Based on these initial observations, we present an extensive characterization of the folding mechanisms of DS119 using stopped-flow refolding experiments and mutational studies.

MATERIALS AND METHODS

Cloning and site-directed mutagenesis

The genes encoding DS119 and DS103 were synthesized and cloned into the BamHI and XhoI sites of $pGEX_{4T-1}$ vectors (Invitrogen, Gaithersburg, MD) as reported before (13). Point mutations on DS119 genes (P14A, E23C, W9F, and W9E) were generated with site-directed mutagenesis kits (Saibaisheng, Beijing, China) according to the manufacturer's instructions. All constructs were confirmed by DNA sequencing.

Protein expression and purification

All proteins were expressed and purified as previously described (13). Briefly, they were expressed in *Escherichia coli* as a GST-fusion protein (GST: glutathione S-transferase) and purified on a GST-affinity column

Submitted May 31, 2013, and accepted for publication September 13, 2013.
*Correspondence: lhlai@pku.edu.cn or gai@sas.upenn.edu
Editor: Bertrand Garcia-Moreno.

0006-3495/13/11/2141/8 $2.00
http://dx.doi.org/10.1016/j.bpj.2013.09.014

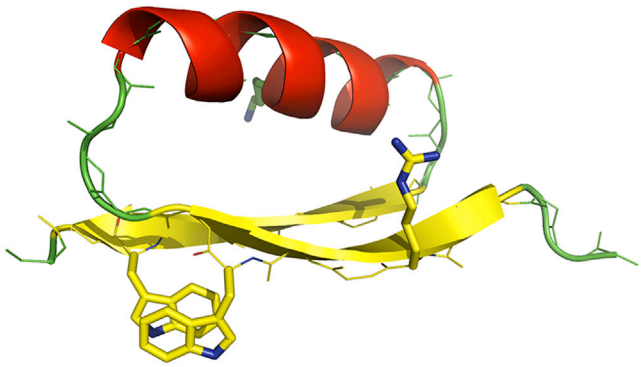


FIGURE 1 Structure of DS119 (Protein Data Bank identification code: 2KI0). Special design features are shown in sticks, including the tryptophan (Trp) pair (W9 and W34) and the negatively designed residues (R6 and K21).

(GE Healthcare, Piscataway, NJ) and then by reversed-phase HPLC. Lyophilized samples were obtained and their molecular weights were confirmed by high-resolution mass spectrometry.

Labeling of fluorescent dyes

E23C was modified with the thiol-reactive fluorophores Alexa Fluor 488 C5-maleimide and 5-((((2-iodoacetyl) amino) ethyl) amino) naphthalene-1-sulfonic acid (1, 5-IAEDANS or I14) (Molecular Probes, Eugene, OR). A fivefold molar excess of fluorophore was used for quantitative modification. Labeling reactions were carried out for 12 h at 4°C in 50 mM Tris-HCl, pH 7.5 with a 10-fold molar excess of Tris (2-carboxyethyl) phosphine (TCEP) (Sigma, St. Louis, MO) to prevent oxidation of cysteine. Labeled proteins were purified by reversed-phase HPLC. All labeled proteins showed the expected increase in mass, as determined by high-resolution mass spectrometry.

Chemical cross-linking

Cross-linking experiments with E23C were conducted with thiol-reactive 1, 11-bis (maleimido) triethylene glycol, or BM-PEG3 (Pierce Biotechnology, Rockford, IL) in two different conditions: steady state and refolding process. In steady state, E23C was diluted to 500 μM BM-PEG3 solutions containing 20 mM PB (Na_2HPO_4/NaH_2PO_4 pH 7.0) and 200 μM TCEP. The final concentration of E23C was 200 μM. In the refolding process, unfolded E23C dissolved in 6 M guanidine hydrochloride (Gdn-HCl) was diluted to 500 μM BM-PEG3 solutions with the same concentrations of PB and TCEP. The final concentration of Gdn-HCl was 0.6 M and different concentrations of E23C were tested in this case from 20 to 200 μM. In both cases the sample was incubated at 25°C for 5 min before SDS-PAGE analysis.

Equilibrium experiments

Circular dichroism (CD) spectra were measured on a MOS 450 AF/CD device (Bio-Logic, Claix, France) at room temperature, using 1-mm quartz cuvettes for the far-ultraviolet (UV) region (190–250 nm) and 10-mm cuvettes for the near-UV region (260–320 nm). Thermal denaturation curves were obtained with Peltier accessory in10-mm quartz cuvettes. Heating was performed at 1°C/min from 1°C to 97°C. In both experiments, the protein concentrations were kept at 0.2 mg/mL in 50 mM PB buffer (Na_2HPO_4/NaH_2PO_4 for pH 7.3 or NaH_2PO4/H_3PO_4 for pH 2.5).

Chemical denaturation experiments were also performed on the MOS 450 AF/CD with a titration accessory in 10-mm quartz cuvettes. The tryptophan fluorescence was monitored by an excitation wavelength of 290 nm and the emission above 320 nm was recorded.

Fluorescence spectra were collected on a FluoroLog 3 spectrofluorometer (HORIBA Jobin Yvon, Longjumeau, France). For fluorescence resonance energy transfer (FRET) measurements, a 1-mm cuvette was used and the buffer was 50 mM PB (pH 7.3). The excitation wavelength was 365 nm, and the spectra between 400 and 650 nm were recorded.

Fourier transform IR spectra were collected on a Magna-IR 860 spectrometer (Nicolet, Madison, WI) with a homemade CaF_2 sample cell. After hydrogen-deuterium exchange, protein samples were dissolved in KD_2PO_4/D_3PO_4 buffer (pH 2.5) to 0.2 mg/mL. Spectra scanning and thermal denaturation were performed as previously described (8).

Fluorescence correlation spectroscopy was conducted with Alexa Fluor 555-labeled E23C at different concentrations (16 nM and 100 μM). The fluorescence correlation spectroscopy (FCS) setup, data collecting, and fitting procedures were described elsewhere (14).

In gel filtration experiments, protein samples were dissolved in 50 mM PB (pH 7.3) buffers with 100 mM NaCl and loaded onto a Superdex Peptide 10/300 GL column (GE Healthcare). The column was eluted using the same buffer with a flow rate of 0.4 mL/min and the UV absorption at 280 nm was monitored.

Proton 1D NMR and 2D NOESY data were collected on a 600 MHz Bruker AVANCE III spectrometer with TXI probe. NOESY spectra were recorded with the mixing time of 200 μs. 1 mM DS119 samples were prepared in 50 mM sodium phosphate buffer (pH 7.3 and pH 2.5) with 10% D_2O. For the samples in pH 7.3 solutions, data were recorded at 20°C and 60°C. For the samples in pH 2.5 solutions, data were recorded at 20°C and 46°C.

Kinetic experiments

The stopped-flow experiments were performed on the MOS 450 AF/CD with a SFM300 module and a 0.8-mm cuvette. The temperature was kept at 25°C by a water bath. Refolding was initiated by an eightfold dilution of protein samples in 6 M Gdn-HCl and 50 mM Na_2HPO_4/NaH_2PO_4 buffer (pH 7.3). The dead time was 0.3 to 0.4 ms and the kinetic traces were recorded 10 ms before the mixing finished. The resulting kinetic curve was an average of 6 to 15 independent measurements. In each measurement the HT value of the detector (photomultiplier tube) was automatically adjusted to obtain the best signal/noise ratio. In stopped-flow fluorescence experiments, the excitation wavelength was 290 nm, and emission above 320 nm was recorded. In stopped-flow FRET experiments, the excitation wavelength was 375 nm, and a filter of 515–555 nm (Chroma, Bellows Falls, VT) was used for detection. For stopped-flow CD, the final concentration of DS119 was 174 μM. For stopped-flow fluorescence, the final concentrations were illustrated (see Fig. 3). Based on the expected amplitude in the equilibrium experiments, the observed signal amplitude in the stopped-flow CD experiments corresponded to ~30% of the total change from unfolded to folded state (−66 deg to −82 deg), and in the stopped-flow fluorescence experiments the observed signal amplitude corresponded to ~90% of the total change (7.5 volts to 4.2 volts) at the highest concentration (99 μM). The details of the T-jump IR setup were described elsewhere (12), except that in the current study a continuous wave quantum cascade laser (Daylight Solutions, San Diego, CA) served as the IR probe.

Data analysis

The kinetic data obtained from stopped-flow experiments were fitted to the bimolecular model described in the main text. The burst phase in each kinetic curve was truncated and only the smooth part was used for the model fitting. Therefore, the start time for the stopped-flow FRET was 22 ms, and for the stopped-flow Trp fluorescence it was 148 ms (see Fig. 3). The changes of protein concentrations with time in the refolding process were described by ordinary differential equations. And the fluorescent emission parameters E (V/μM) were defined for each species. The signal of a specific species, for example U, at a time t was determined by $[U](t) \times QY(U) \times R$,

where [U](t) is the concentration of U at time t, QY is the quantum yield, and R is the response of the instrument. QY (U) × R is the emission parameter E.

The curve fitting toolbox of MATLAB (The MathWorks, Natick, MA) was used to estimate the microscopic rate constants (k_1, k_{-1}, and so on) and emission parameters by modified simulated annealing, an optimization algorithm for nonlinear squared fitting using local Hessian to accelerate convergence. Because the experimental data can be represented by a linear combination of concentrations of species in the ordinary differential equations model, a solver for linear least squared error problems was applied to determine the combinational coefficients. In the cases that different curves share part of the combinational coefficients, the shared and unshared coefficients were iteratively computed by two different linear least squared error solvers.

The asymptotic approximation of 100(1-α) % confidence region was used to evaluate the 95% confidence region:

$$\left\{\boldsymbol{\theta} : \Phi(\boldsymbol{\theta}) < \Phi\left(\widehat{\boldsymbol{\theta}}\right)\left(1 + \frac{p}{n-p}F^{\alpha}_{p,n-p}\right)\right\},$$

where p stands for the number of parameters and n for the number of data points. We used the Metropolis sampling method to search for parameter sets in the confidence region (15).

RESULTS

Equilibrium unfolding measurements of DS119

The conformational stability of DS119 at neutral pH was assessed by guanidine hydrochloride (Gdn-HCl)-induced denaturation, using both CD and tryptophan (Trp) fluorescence as conformational probes. As shown (Fig. 2 *a*), the original Trp fluorescence intensity was low, and it progressively increased along with the Gdn-HCl concentration, indicating that Gdn-HCl-induced protein unfolding alleviated the Trp fluorescence quenching in the native DS119 structure. Furthermore, the CD unfolding curve showed significant deviation from that of Trp fluorescence, suggesting that the secondary structures (i.e., the CD signal) and the tertiary structure (i.e., the fluorescence signal) may not be formed in a cooperative manner, or that the underlying folding/unfolding process involves stable intermediates.

DS119 exhibits slow and concentration-dependent folding kinetics

Because the initial attempt of using time-resolved IR spectroscopy to measure the population redistribution kinetics of DS119 in response to a nanosecond T-jump pulse did not reveal any relaxation events at the submillisecond timescale (Fig. S1 in the Supporting Material), we employed stopped-flow CD and fluorescence to measure the folding kinetics of DS119, via rapid dilution of the protein samples dissolved in a 6 M Gdn-HCl solution (50 mM Na_2HPO_4/NaH_2PO_4 buffer, pH 7.3). Both the stopped-flow CD and fluorescence kinetics indicated that DS119 folded extremely slowly (Fig. 2, *c* and *d*). By fitting the data to a single-exponential function, we roughly estimated the folding time of the secondary and tertiary structures to be 136 ms (CD signal) and 517 ms (fluorescence signal), respectively. More importantly, the folding kinetics of DS119 showed strong concentration dependence in the range of 11 and 99 μM (Fig. 3 *a*). In contrast, the equilibrium unfolding transitions measured at different DS119 concentrations (15–82 μM) overlapped with each other within our experimental uncertainties (Fig. 2 *b*), suggesting that both the folded and unfolded states are monomeric in the concentration range studied.

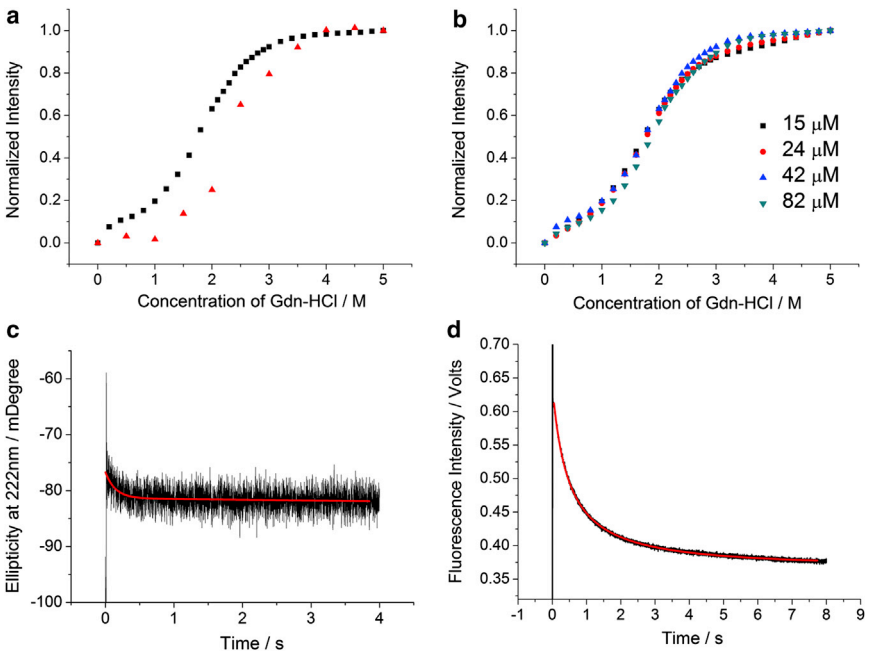


FIGURE 2 Chemical denaturation and refolding kinetics of DS119. (*a*) Chemical denaturation curves measured by CD (*red*) and Trp fluorescence (*black*). (*b*) Chemical denaturation curves of different concentrations of DS119 measured by Trp fluorescence. (*c*) Stopped-flow refolding kinetics measured by CD. (*d*) Stopped-flow refolding kinetics measured by Trp fluorescence. In the kinetic experiments, the final concentration of Gdn-HCl was 0.75 M. Red lines show the fitting to a single-exponential function.

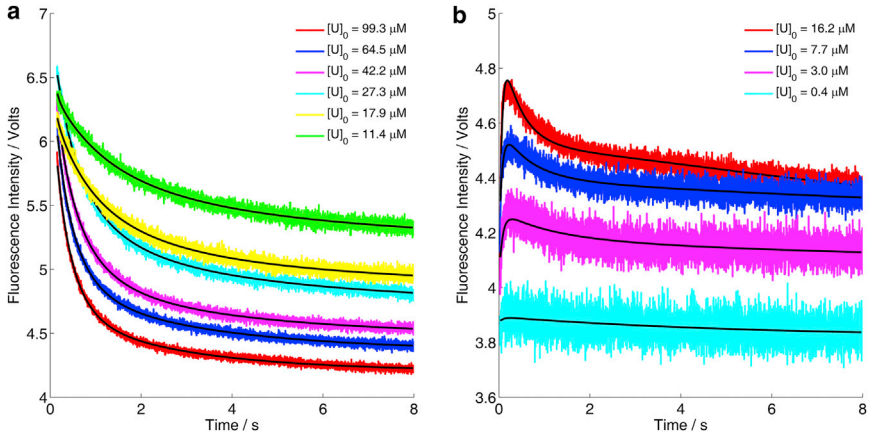


FIGURE 3 (*a*) Refolding kinetics of DS119 measured by stopped-flow fluorescence at different protein concentrations at pH 7.3. (*b*) Stopped-flow refolding kinetics obtained using an E23C-I14 and E23C-488 mixture (1:1). Global fitting to a bimolecular model (Fig. 4 *a*) is shown in black lines.

We further confirmed that DS119 is a monomer in the steady states using FRET, FCS, and analytical gel filtration. We labeled a mutant of DS119 (E23C) with either a fluorescent donor (I14) or an acceptor (Alexa Fluor 488). Far-UV CD and thermal unfolding experiments suggested that the dye labeling did not introduce any significant changes to the structure or thermodynamics properties of the protein (Fig. S2). In equilibrium measurements, no detectable FRET signal was observed for E23C-I14 and E23C-488 mixtures in the unfolded or the folded state (Fig. S3). FCS analysis showed that the dye-labeled DS119 mutant had similar diffusion times at 16 nM and 100 μM (Fig. S4). In the gel-filtration experiments, the elution peaks of DS119 did not show concentration dependence in the range of 5 to 200 μM, and based on the apparent molecular weight calculated from the standard curve, DS119 was determined to be a monomer (Fig. S5). These results are consistent with our previous NMR studies (9).

Next, we employed chemical cross-linking to detect possible oligomeric states of the mutant E23C both in the folded state and in the refolding process (Fig. S6). In the refolding process, a small amount of dimers were trapped for 50 to 200 μM of E23C. For the same concentration of E23C in the steady folded state, however, only the band of monomeric proteins was observed after cross-linking. In the cross-linking experiments, simultaneously linking of three or more proteins is difficult, so we cannot rule out the possibility that other oligomeric forms are involved. Taken together, these results suggest that the folding process of DS119, from a monomeric unfolded state to a monomeric folded state, probably involves a transient dimer formation step.

Stopped-flow FRET kinetics corroborate the transient dimer formation model

To substantiate the previous notion that a dimeric intermediate is populated on the folding pathway of DS119, we carried out stopped-flow FRET experiments with a mixture of E23C-I14 and E23C-488 to detect whether molecular association occurs via detection of intermolecular FRET signals. A transient FRET signal was observed in the kinetic analysis, and the amplitude of this FRET signal showed a strong dependence of the total protein concentration (Fig. 3 *b*). Moreover, within the time window of the stopped-flow experiments, this FRET signal first increased and then decreased. These results provide evidence for the hypothesis that the folding of DS119 involves a partially folded dimeric species, which dissociates to yield the monomeric and folded protein.

Based on the results of the stopped-flow fluorescence and FRET measurements, we proposed a bimolecular folding model to elucidate the refolding mechanism of DS119, which involves a dimeric intermediate species I_2 (Fig. 4 *a*). We used this model to globally fit the stopped-flow kinetic traces by numerically solving the differential rate equations (Fig. S7). The fits were satisfactory and the resulting microscopic rate constants indicate that the rate-limiting step is the formation of the transient dimer (Tables 1, S1 and S2). The time courses of the refolding process were simulated based on the rate constants at different initial concentrations (Fig. 5). At the protein concentration of 50 μM, I_2 reached its maximum concentration at ~150 ms and then dissociated to the folded state F rapidly.

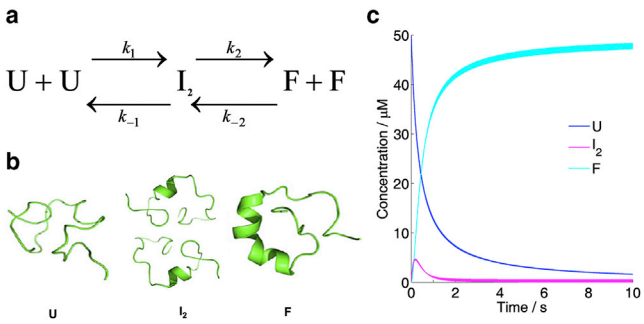


FIGURE 4 (*a*) The bimolecular folding model. (*b*) A cartoon representation of different species in the model. The structural model of I_2 is an illustration of one possible configuration. (*c*) A simulation for the concentration changes of U, I_2, and F in the folding process. The total protein concentration is 50 μM.

TABLE 1 Microscopic rate constants derived from the global fitting of the stopped-flow kinetics data

$k_1/\mu M^{-1}s^{-1}$	2.97×10^{-2}	$[2.89 \times 10^{-2}\ 2.98 \times 10^{-2}]$[a]
k_{-1}/s^{-1}	2.94×10^{-9}	$(0\ 6.36 \times 10^{-2}]$
k_2/s^{-1}	6.57	[6.50 6.91]
$k_{-2}/\mu M^{-1}s^{-1}$	1.01×10^{-4}	$(0\ 1.76 \times 10^{-3}]$

[a]The 95% confidence intervals are shown in brackets.

At ~2 s, >90% of the unfolded protein U changed to F (Fig. 4 *c*). The timescales for the formation of I_2 at different initial concentrations are in the range of 100–200 ms (Fig S8 *b*), which is consistent with the folding time observed in the stopped-flow CD (136 ms), indicating that I_2 probably has a significant amount of secondary structures.

We also calculated the 95% confidence intervals of each microscopic rate constant (Table 1, k_1, k_{-1}, k_2, and k_{-2}). The results revealed that k_1 was more accurately determined than k_{-1}, k_2, and k_{-2}.This is because the strong folding conditions (Gdn-HCl was 0.75 M after mixing) was applied in our experiment and the formation of the transient dimer is the rate-limiting step in the folding kinetics. We further applied different final concentrations of Gdn-HCl in the stopped-flow experiments and we found the resulting rate constant k_1 had a linear relationship with the concentration of Gdn-HCl in the semilogarithmic plot (Fig. S9 and Table S3).

The refolding process of DS119 at pH 2.5

We also studied the refolding process of DS119 in acidic solutions. The T-jump IR experiments were performed at pH 2.5 because the T_m of DS119 (46°C) at this pH was lower than that at neutral pH. The stopped-flow results indicated that at pH 2.5 DS119 also had slow and concentration-dependent folding kinetics (Fig. S10). The kinetic curves were well fitted by our bimolecular folding model and the resulting rate constants indicated that the disaggregation of I_2 to form monomeric F happened much slower in the acidic condition (the underlying rate constant k_2 was ~170 times smaller than that at pH 7.3).

We further applied NMR to probe the structure of DS119 at the temperatures around T_m in both pH 7.3 and pH 2.5 solutions. At pH 7.3 we compared the 1D-NMR and NOSEY spectra of DS119 at 20°C and 60°C, respectively (Fig. S11 and Fig. S12). Multiple peaks were observed at both temperatures, indicating the transition between the folded and unfolded states was a slow process. At pH 2.5 the exact T_m could be reached in our NMR measurement and the results are similar as that in pH 7.3 (Fig. S13). Taken together, these results provided further evidence supporting the proposed slow folding model.

The folding kinetics of the mutants of DS119

We have studied the thermodynamic properties and folding mechanisms of several mutants of DS119, including DS103 (Table 2), W9F, W9E, and P14A. DS103 is an intermediate sequence in the de novo design process of DS119. Compared to DS119, the major differences in DS103 residues are V5E, T13D, and K21L. Gel filtration experiments demonstrated that DS103 folded into a stable dimer (Fig. S14). The CD spectra and thermal denaturation experiments indicated that DS103 has a well-folded α/β structure, and the stopped-flow fluorescence showed that the folding process of DS103 was also slow and concentration-dependent, which was well fitted to a different bimolecular model (Fig. 6 and Fig. S15). For DS103, both the intermediate and native states (I_2 and F_2) are dimers, and the step of forming I_2 was approximately seven times faster than that of DS119 (Table S4), suggesting that the mutated residues accelerate the formation of the transient dimers.

Of the three mutations in DS103 (compared with DS119), K21L was the only one representing a change from a charged residue to a highly hydrophobic residue. It is reasonable to assume that Leu-21 in DS103 plays a significant role in stabilizing the dimer both in the folding process and in the folded structure, whereas the Lys-21 in DS119 destabilizes the transient dimer and leads to a monomeric folded structure. We speculate one possible configuration

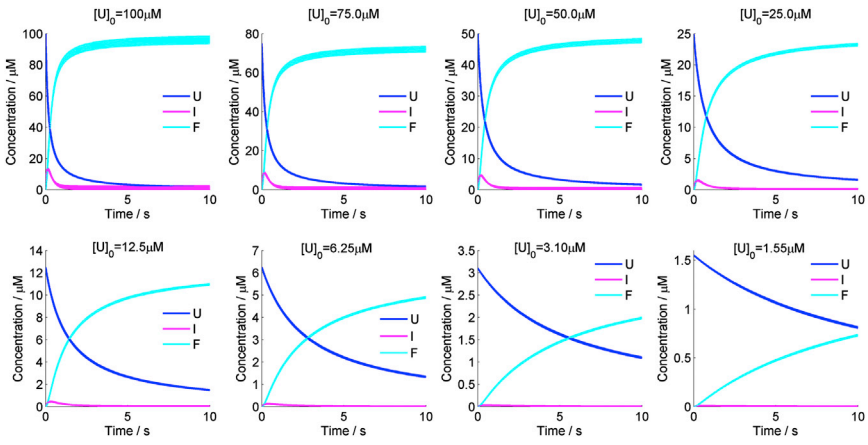


FIGURE 5 Simulated time course of the concentration of each species (U, I_2, and F) in the bimolecular folding model of DS119. The total protein concentrations are labeled as $[U]_0$.

TABLE 2 Sequences of DS119 and DS103

DS119	GSGQV RTIWV GGTPE ELKKL KEEAK KANIR VTFWG D
DS103	GSGQE RTIWV GGDPE ELKKL LEEAK KANIR VTFWG D GG

of the intermediate is that the interfaces of the transient dimers consist of a partially folded hydrophobic core (I29, V31, F33, etc.) and probably the tryptophan residues.

For the other two mutants W9E and W9F, the near-UV CD spectra showed that only W9F preserved a stable tertiary structure (Fig. S16 and Fig. S17). The thermal unfolding process and kinetic refolding results of W9F are similar to those of DS119. These results highlight the importance of π-π stacking of aromatic residues in stabilizing the $\beta\alpha\beta$ structure and in shaping the folding process.

Slow kinetic events observed in the protein folding studies usually arise from proline isomerization, which occurs on a timescale of hundreds of milliseconds (16). To exclude this possibility, we mutated the single proline residue in the N-cap region of DS119 to alanine and measured the folding thermodynamics and kinetics of the resulting mutant P14A. Comparison of the CD spectra of P14A and DS119 suggests that this mutation has not altered the characteristics of the folded structure (Fig. S18). The stopped-flow fluorescence kinetics further showed that just like the native sequence, P14A also folded with an extremely slow rate, indicating that the slow folding behavior of DS119 does not stem from proline isomerization, but rather reflects the uniqueness of the underlying folding energy landscape of DS119.

DISCUSSION

Our results are the first, to our knowledge, to demonstrate the formation of a transient dimer in the folding process of a designed protein. A transient dimer has been observed in the refolding process of cytochrome *c* using small-angle x-ray scattering (17). Furthermore, transient aggregation or partially unfolded oligomers have been reported for maltose-binding protein (18), the SH3 domain of α-spectrin (19), the human spliceosomal protein U1A (20), and most interestingly, a water-soluble protein with an introduced Alzheimer sequence (21). In most studies, the self-associated intermediates were verified by equilibrium experiments, such as gel filtration (17), differential scanning calorimetry (19), or observable precipitates (18). Therefore, they were more stable than the transient dimer we observed and were usually considered to be kinetic traps. In the case of DS119, the lifetime of the transient dimer was much shorter and could not be detected by these techniques. We observed the transient dimer by chemical cross-linking and FRET experiments. However, the structural details of the transient dimer need further investigation.

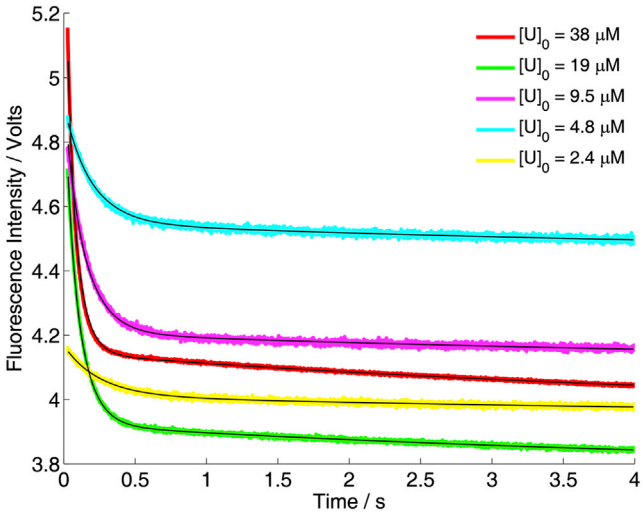


FIGURE 6 Stopped-flow refolding kinetics of DS103 measured by Trp fluorescence at pH 7.3. The fitting curves are shown in black lines.

An interesting question is whether the dimeric intermediate pathway is obligatory and whether the dimeric structure is unique or not. To demonstrate the monomeric folding pathway is kinetically disfavored, we have roughly fit the stopped-flow refolding kinetics to a unimolecular reaction model (Fig. S19 *b*). Although the refolding process cannot be rigorously described by the single-exponential function, this approach nevertheless offers the easiest means to estimate the folding rate of DS119 at nearly zero protein concentration. By extrapolating the rate constants to zero concentration, we estimated that the folding time for monomeric proteins was ~1.8 s (Fig. S20). Comparing this value to the rate constants in Table 1, we think the bimolecular folding pathway is more favored and has a lower free energy barrier. We have also tried to globally fit the kinetic curves in Fig. 3 *a* with a bimolecular reaction model and the fitting was more reasonable (Fig. S19 *a*), indicating the main kinetic process is a bimolecular process.

To probe the possible structure of the intermediate, we have compared the folding mechanisms of DS119 and the mutant DS103. For DS103 (V5E, T13D, and K21L) there are several possible interactions that can accelerate the formation of the transient dimer: 1), nonspecific hydrophobic stacking between K21L and I29, V31, F33; 2), quadrupole-like interactions among 5E, 6R on one monomer, and the same residues on the other monomer, or an intermolecular salt bridge formed between 5E and 6R. The intermolecular β-sheet, which is facilitated by the W-W interactions, is also a possible configuration. Furthermore, it has been reported that nonnative interactions participate in the folding process of several proteins, such as the de novo designed Top7 and the knotted proteins (5,6,22). In our case, the nonnative interactions observed are mainly those at the dimeric interface in the intermediate. The chevron rollover caused by transient nonnative contacts as for Top7 has not been observed in our studies (Fig. S9). Further molecular dynamics simulations and NMR studies are helpful on elucidating the structure of the intermediate.

As a de novo designed protein, DS119 is unique in that it consists of a standalone $\beta\alpha\beta$ motif, in which the two β-strands are parallel to each other. Although the $\beta\alpha\beta$ motif has been predicted to possess high designability (23,24), a tryptophan pair was introduced to stabilize the native fold of DS119. Thus, it is reasonable to assume that the slow folding behavior of DS119 arises from the difficulty of such parallel β-strand arrangement to spontaneously develop within a single protein molecule. In other words, a catalytic event is needed to facilitate the folding of DS119. Our stopped-flow and mutation studies suggest that this catalytic event corresponds to a transient dimerization process. We speculate that the transient dimer of DS119 possesses some secondary structure and the α-helix is probably formed, which means that several charged residues in the helix region are on the protein surface, leading to a relatively rapid disassociation process. At pH 2.5 the repulsive interactions between the charged residues are weakened and consequently the disaggregation is much slower than that at pH 7.3, as suggested by the folding results at pH 2.5. No aggregation was observed for the native state, indicating that the reverse process is rather slow, which is mainly due to the negatively designed lysine and arginine residues.

De novo designed proteins are not subjected to the process of natural selection, which makes most natural proteins fold rapidly and cooperatively. In the living cells, slow-folding kinetics is a drawback because the unfolded region in newly synthesized proteins can easily be digested by various proteases. Our experiments showed that although the thermodynamic properties of designed proteins are optimized to obtain structures with high stability, their kinetic properties are not. This is one possible reason that the designed proteins usually have folding mechanisms distinct from those of natural proteins. De novo designed proteins also provide clues to understand the evolutionary process of protein structures and folding mechanisms. The process of designing DS103 and DS119 can be regarded as an artificial structural evolution. Our results indicate that the differences between DS103 and DS119 originate from the differences in the stabilities of the transient dimers. Thus, in future protein design of novel structures, not only the native states, but also the kinetic properties of proteins need to be considered explicitly.

CONCLUSION

In this study, we used various kinetic techniques including T-jump IR, stopped-flow CD, and stopped-flow fluorescence to probe the folding process of a de novo designed protein DS119. As the first, to our knowledge, designed small protein with a parallel β structure, DS119 provides a model system for understanding the folding mechanisms of the $\beta\alpha\beta$ structures. Our results and analysis showed that the slow folding kinetics of DS119 originates from a transient dimer formation process and intermolecular interactions participate in the folding process of DS119, though the protein is monomeric after being folded. In summary, folding mechanism studies of de novo designed proteins may shed light on the evolutionary process of protein sequence, structure, and folding kinetics.

SUPPORTING MATERIAL

Five tables and twenty-one figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)01034-5.

We thank Professors Zhirong Liu and Zhuqing Zhang for helpful discussions. We acknowledge the help from Dr. Xiaogang Niu and Beijing Nuclear Magnetic Resonance Center on the NMR measurements. We also acknowledge the technical assistance of Arnaldo L. Serrano in the laser T-jump setup and Chun-Wei Lin in the FCS experiments.

This work was supported in part by the Ministry of Science and Technology of China (2009CB918500) and the National Natural Science Foundation of China (11021463, 21173013).

REFERENCES

1. Samish, I., C. M. MacDermaid, …, J. G. Saven. 2011. Theoretical and computational protein design. *Annu. Rev. Phys. Chem.* 62:129–149.
2. Dill, K. A., and J. L. MacCallum. 2012. The protein-folding problem, 50 years on. *Science.* 338:1042–1046.
3. Pitera, J. W., and W. Swope. 2003. Understanding folding and design: replica-exchange simulations of "Trp-cage" miniproteins. *Proc. Natl. Acad. Sci. USA.* 100:7587–7592.
4. Koga, N., R. Tatsumi-Koga, …, D. Baker. 2012. Principles for designing ideal protein structures. *Nature.* 491:222–227.
5. Watters, A. L., P. Deka, …, D. Baker. 2007. The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell.* 128:613–624.
6. Zhang, Z., and H. S. Chan. 2010. Competition between native topology and nonnative interactions in simple and complex folding kinetics of natural and designed proteins. *Proc. Natl. Acad. Sci. USA.* 107:2920–2925.
7. Zhang, Z., and H. S. Chan. 2009. Native topology of the designed protein Top7 is not conducive to cooperative folding. *Biophys. J.* 96: L25–L27.
8. Zhu, Y., D. O. V. Alonso, …, F. Gai. 2003. Ultrafast folding of alpha3D: a de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci. USA.* 100:15486–15491.
9. Liang, H., H. Chen, …, L. Lai. 2009. De novo design of a beta alpha beta motif. *Angew. Chem. Int. Ed. Engl.* 48:3301–3303.
10. Qi, Y. F., Y. Q. Huang, …, L. H. Lai. 2010. Folding simulations of a de novo designed protein with a betaalphabeta fold. *Biophys. J.* 98: 321–329.
11. Capriotti, E., and R. Casadio. 2007. K-Fold: a tool for the prediction of the protein folding kinetic order and rate. *Bioinformatics.* 23:385–386.
12. Huang, C. Y., Z. Getahun, …, F. Gai. 2002. Helix formation via conformation diffusion search. *Proc. Natl. Acad. Sci. USA.* 99:2788–2793.
13. Zhu, C., C. Zhang, …, L. Lai. 2011. Engineering a zinc binding site into the de novo designed protein DS119 with a $\beta\alpha\beta$ structure. *Protein Cell.* 2:1006–1013.
14. Guo, L., P. Chowdhury, …, F. Gai. 2008. Denaturant-induced expansion and compaction of a multi-domain protein: IgG. *J. Mol. Biol.* 384:1029–1036.
15. Srinath, S., and R. Gunawan. 2010. Parameter identifiability of power-law biochemical system models. *J. Biotechnol.* 149:132–140.

16. Jackson, S. E., and A. R. Fersht. 1991. Folding of chymotrypsin inhibitor 2. 2. Influence of proline isomerization on the folding kinetics and thermodynamic characterization of the transition state of folding. *Biochemistry.* 30:10436–10443.

17. Segel, D. J., D. Eliezer, …, S. Doniach. 1999. Transient dimer in the refolding kinetics of cytochrome *c* characterized by small-angle X-ray scattering. *Biochemistry.* 38:15352–15359.

18. Ganesh, C., F. N. Zaidi, …, R. Varadarajan. 2001. Reversible formation of on-pathway macroscopic aggregates during the folding of maltose binding protein. *Protein Sci.* 10:1635–1644.

19. Casares, S., M. Sadqi, …, N. A. J. van Nuland. 2004. Detection and characterization of partially unfolded oligomers of the SH3 domain of alpha-spectrin. *Biophys. J.* 86:2403–2413.

20. Silow, M., and M. Oliveberg. 1997. Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl. Acad. Sci. USA.* 94:6084–6086.

21. Otzen, D. E., S. Miron, …, M. Oliveberg. 2004. Transient aggregation and stable dimerization induced by introducing an Alzheimer sequence into a water-soluble protein. *Biochemistry.* 43:12964–12978.

22. Chan, H. S., Z. Zhang, …, Z. Liu. 2011. Cooperativity, local-nonlocal coupling, and nonnative interactions: principles of protein folding from coarse-grained models. *Annu. Rev. Phys. Chem.* 62:301–326.

23. Li, H., R. Helling, …, N. Wingreen. 1996. Emergence of preferred structures in a simple model of protein folding. *Science.* 273:666–669.

24. Miller, J., C. Zeng, …, C. Tang. 2002. Emergence of highly designable protein-backbone conformations in an off-lattice model. *Proteins.* 47:506–512.

Supporting Material

Slow and Bimolecular Folding of a *de Novo* Designed Monomeric Protein DS119

Cheng Zhu,[†] Ziwei Dai,[‡] Huanhuan Liang,[†¶] Tao Zhang,[†] Feng Gai,[§]* and Luhua Lai[†‡]*

[†]BNLMS, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, Peking-Tsinghua Center for Life Sciences at College of Chemistry and Molecular Engineering, Peking University, Beijing, China; [‡]Center for Quantitative Biology, Peking University, Beijing, China; [§]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania; and [¶]National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China

Thermal denaturation, equilibrium and time-resolved IR spectra of DS119 at pH 2.5

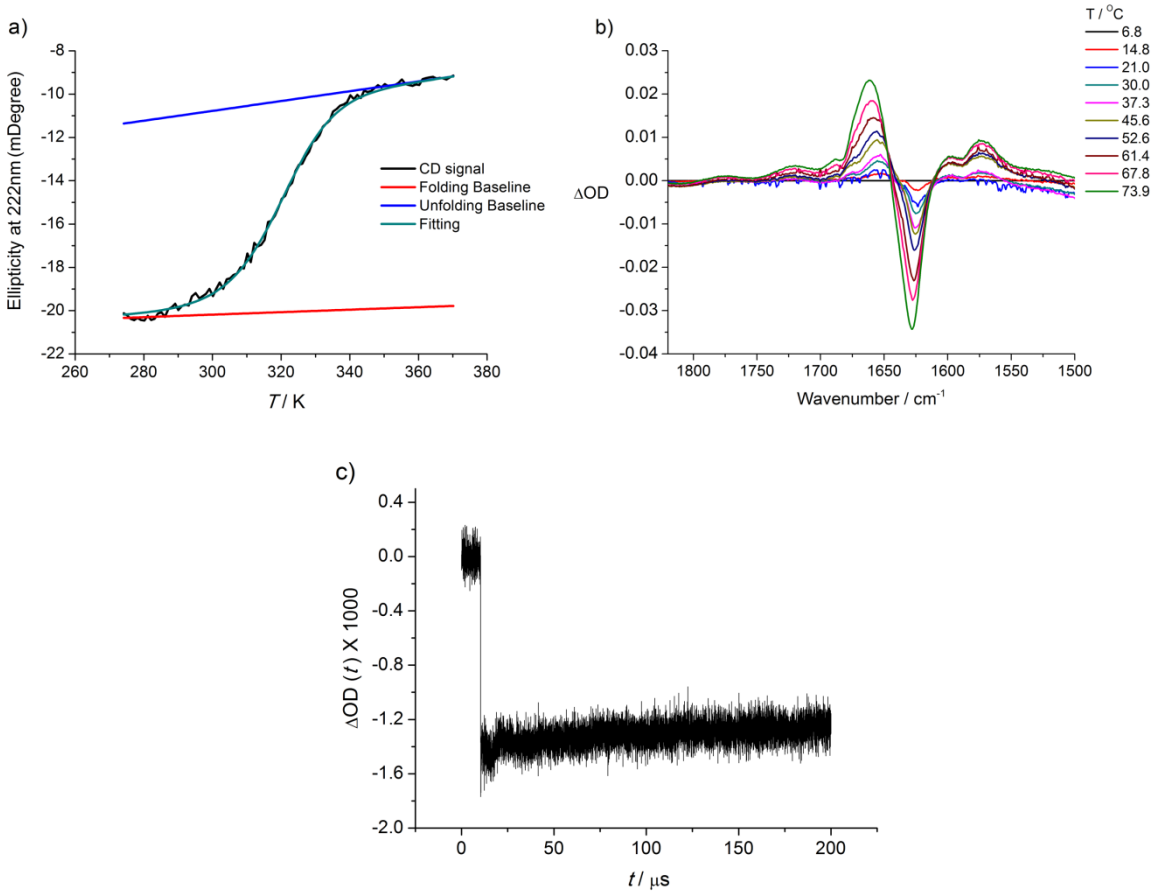


Figure S1. a) Thermal denaturation curve of DS119. Thermal unfolding was measured from 274 K to 370 K in 50 mM PB buffer (pH 2.5). The CD signal at 222 nm was monitored. Fitting to a two-state model (green line) indicated the T_m is 46.2°C at PH 2.5. b) Difference IR spectra of DS119 in 50mM PB buffer (pH2.5) generated by subtracting the spectrum collected at 6.8°C (OD=optical density). c) *T*-jump-induced relaxation kinetics of DS119 measured by time-resolved IR spectroscopy at 1633 cm^{-1} in 50mM PB buffer (pH2.5). The temperature change was from 46°C to 52°C (calibrated by D_2O IR band). The slow phase in the *T*-jump data is the background thermal re-equilibration (not protein folding/unfolding dynamics).

Dye-labeled mutants E23C-I14 and E23C-488

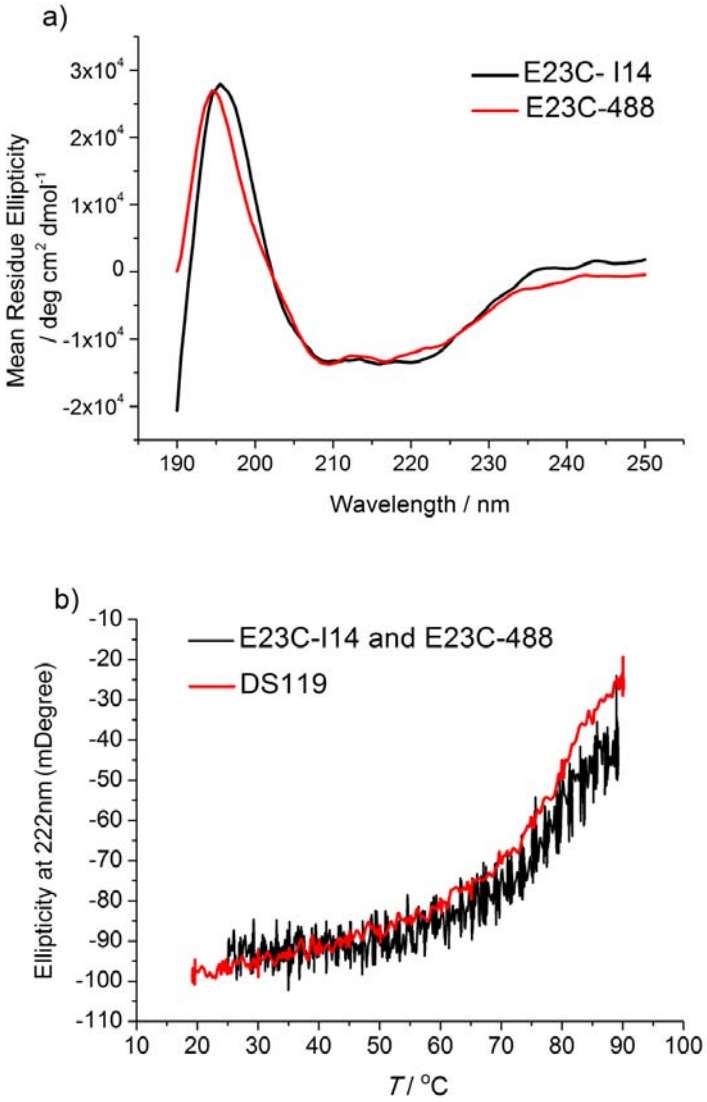


Figure S2. a) Far-UV CD spectra of E23C-I14 and E23C-488 measured in 50 mM PB buffer (pH 7.3). b) Thermal denaturation curves of E23C-I14 and E23C-488. Thermal unfolding was measured from 20°C to 90°C in 50 mM PB buffer (pH 7.3). The CD signal at 222 nm was monitored.

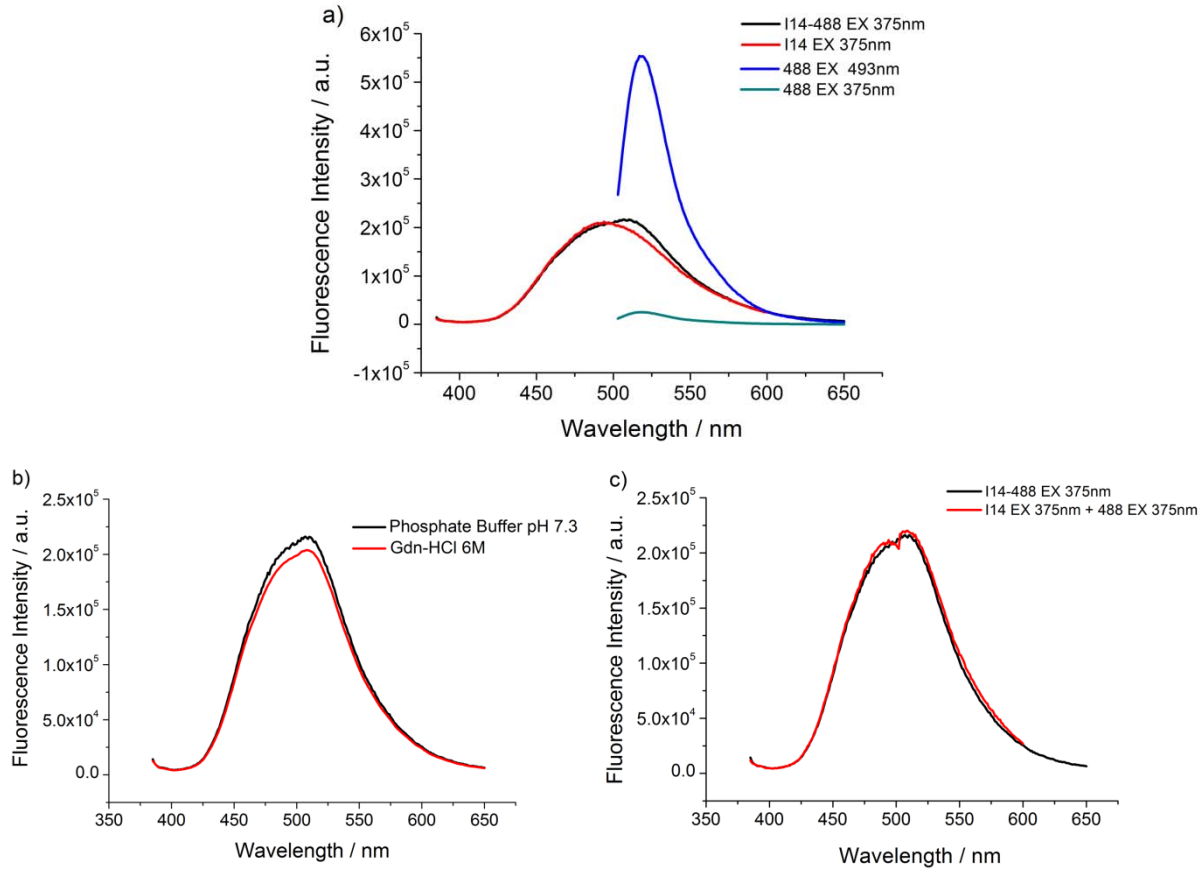


Figure S3. a) Steady-state fluorescence spectra of E23C-I14 and E23C-488. The excitation wavelengths of I14 and Alexa 488 were 375 nm and 493 nm, and their emission peaks were at 492 nm and 518 nm, respectively. Black : the fluorescent spectra of an E23C-I14 (12 μM) and E23C-488 (15 μM) mixture excited at 375nm; Red: E23C-I14 (12 μM) excited at 375nm; Cyan: E23C-488 (15 μM) excited at 375nm; Blue: E23C-488 (15 μM) excited at 493nm. The small peak around 520 nm in the black line was not due to FRET, as shown in Figure S3c.
b) Fluorescence spectra of E23C-I14 and E23C-488 in the native (black, 30 μM proteins and 50 mM phosphate buffer) and denatured (red, 28 μM proteins and 6 M Gdn-HCl) state.
c) Black: the fluorescent spectra of an E23C-I14 (12 μM) and E23C-488 (15 μM) mixture excited at 375nm; Red: Addition of two spectra: E23C- I14 (12 μM) excited at 375nm + E23C-488 (15 μM) excited at 375nm. (EX: excitation wavelength)

FCS Experiments

Two FCS traces were collected, one on a 16 nM dye-labeled (Alexa Fluor 555) DS119 solution (A) and another on a mixture of 8 nM dye-labeled DS119 and 100 μM DS119 (B). Fitting these traces yielded diffusion times (τ_D) of 120 μs and 142 μs, for A and B, respectively. This difference likely arises from the increase in solution viscosity with increasing protein concentration (the τ_D of the free dye in 16 nM and 100 μM DS119 were 47.7 and 51.0 μs).

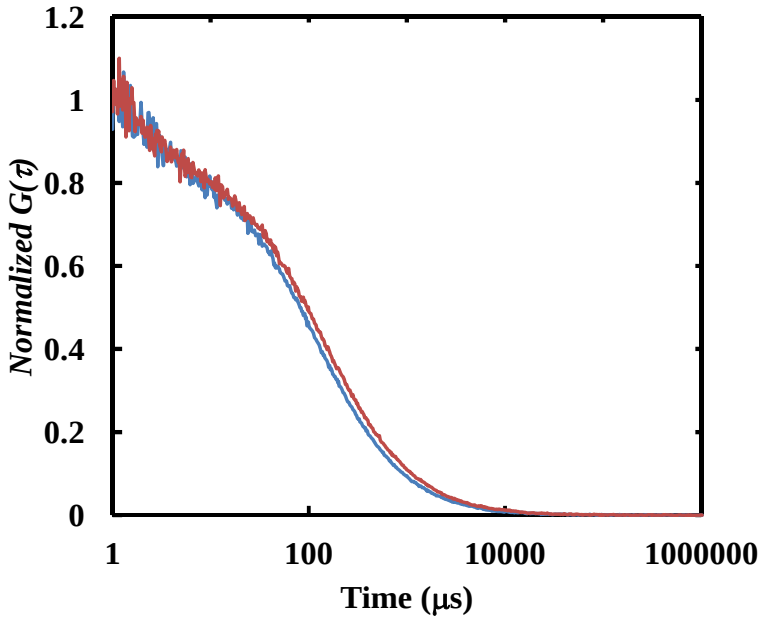


Figure S4. Normalized autocorrelation traces measured with a 16 nM labeled DS119 solution (blue) and a mixture of 8 nM labeled-DS119 and 100 μM DS119 (red). Both samples were prepared in 50 mM phosphate buffer (pH = 7).

Analytical gel filtration of DS119

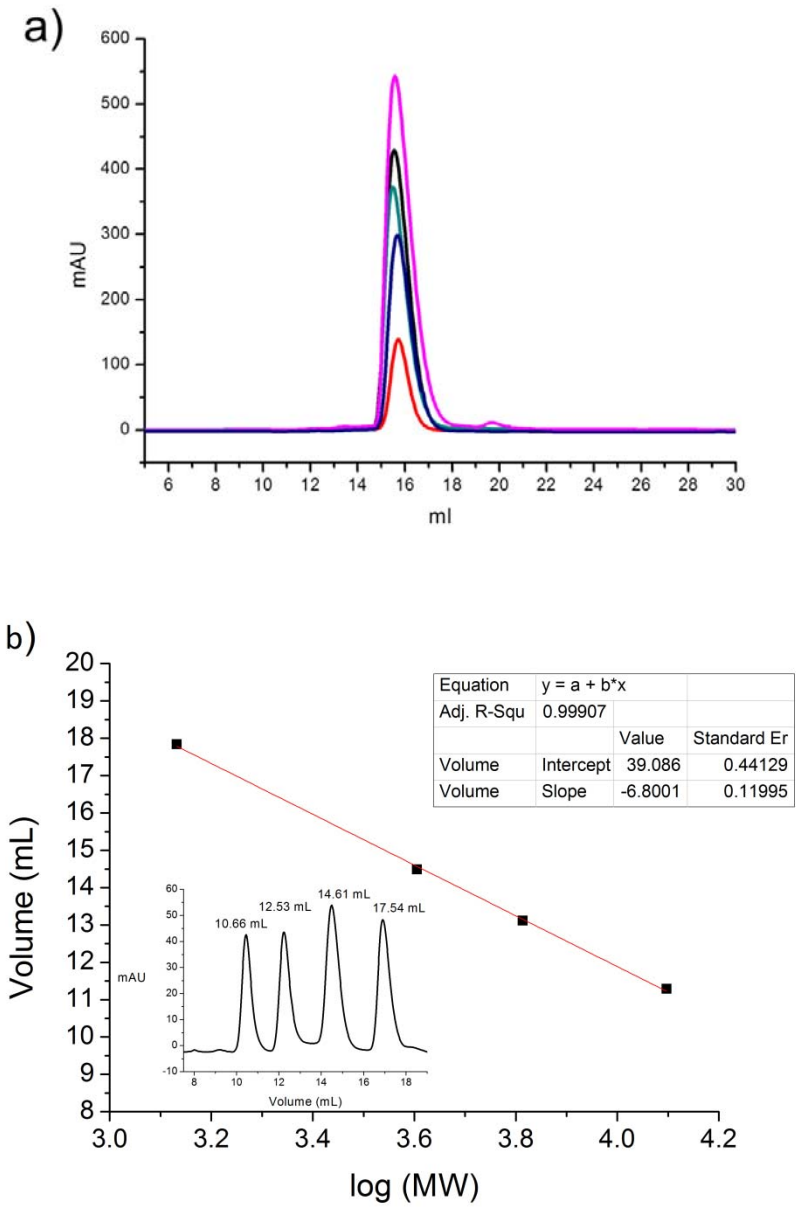


Figure S5. a) Analytical gel filtration of DS119 at different concentrations. Protein samples were dissolved in the buffers containing 50 mM Na_2HPO_4 / NaH_2PO_4 (pH7.3) and 100 mM NaCl, and then loaded on a Superdex^{TM} Peptide 10/300 GL column. The protein concentrations of each curve were (from bottom to top) 5 μM, 20 μM, 50 μM, 100 μM and 200 μM. b) The standard curve of the Superdex^{TM} Peptide 10/300 GL column. The elution peaks (from left to right) are cytochrome C (12.4 kDa), aprotinin (6512 Da), DS119 (4028 Da) and vitamin B12 (1355 Da).

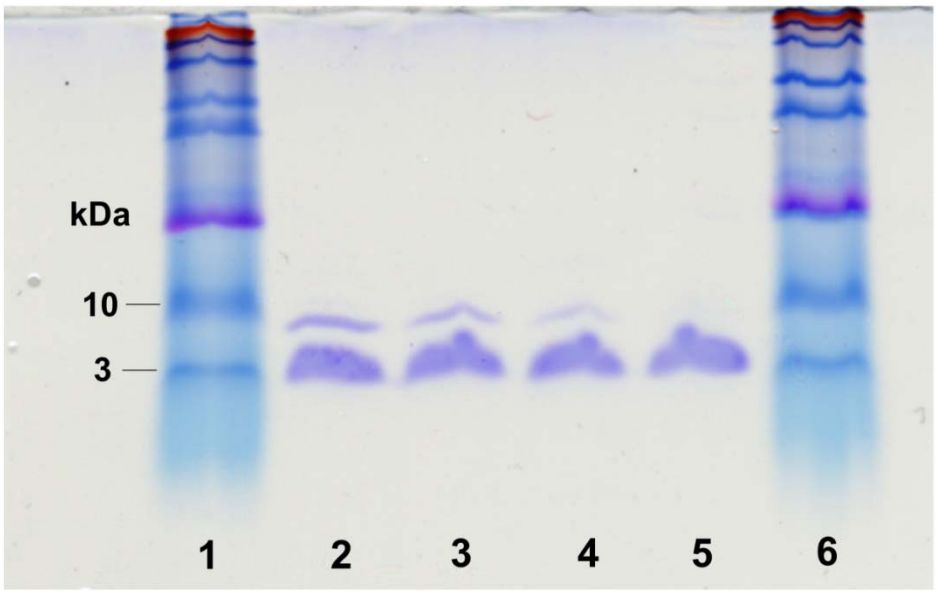
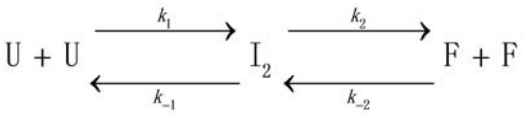


Figure S6. SDS-PAGE gels of E23C submitted to chemical cross-linking in the refolding process (lanes 2- 4) and the steady state (lane 5). In the refolding process cross linking, the denatured E23C was diluted to solutions with crosslinker (thiol-reactive BM-PEG3). In the steady state cross linking, the native E23C was diluted to crosslinker solutions. Lane 1, 6: protein marker ; Lane 2: 200 μM E23C; Lane 3: 100 μM E23C; Lane 4: 50 μM E23C; Lane 5: 200 μM E23C. The cross-linking solutions of lane 2, 3 and 5 were diluted to 50 μM of E23C before loading to the gels.

Global fitting of the stopped-flow Trp fluorescence and FRET data



$$\frac{d[U]}{dt} = -2k_1[U]^2 + 2k_{-1}[I_2]$$

$$\frac{d[I_2]}{dt} = k_1[U]^2 - k_{-1}[I_2] - k_2[I_2] + k_{-2}[F]^2$$

$$\frac{d[F]}{dt} = 2k_2[I_2] - 2k_{-2}[F]^2$$

Figure S7. The bimolecular folding model and differential rate equations for the folding process of DS119.

Table S1. The simulated fluorescent emission parameters for stopped-flow Trp fluorescence of DS119.

E (V/μM) / $[U]_0$ (μM)	E (U)[a]	M1[b]	M2
99.2	0.078	0.95	0.54
64.5	0.109	1.09	0.58
42.2	0.150	1.51	0.64
27.3	0.223	1.80	0.67
17.9	0.311	1.58	0.71
11.4	0.478	1.63	0.75

a Trp fluorescence intesity equals to E(U)×([U]) + E(I_2)×[I_2] + E(F)×[F], E(I_2) = M1 × E(U), E(F) = M2 × E(U). $[U]_0$ is the initial concentration of DS119 after mixing. Since the response of the instrument is not a constant in different experiments (the voltage of the detector was changed), E(U) is a variable in our simulation.

b M is the ratio between E(U) and E(I_2), E(F). Since the response of the instrument is a constant in each experiment, M1 = QY(I_2) / QY (U) and M2 = QY(F) / QY (U) are also constants (QY: quantum yield). In the simulation, we allowed a ±15% variation of M1 and M2.

Table S2. The simulated fluorescent emission parameters for stopped-flow FRET of DS119.

E (V/μM) / $[U]_0$ (μM)	E (U)[a]	N1	N2
16.2	0.28	3.10	0.97
7.7	0.56	3.52	0.99
3.0	1.37	4.34	0.99
0.4	9.73	3.74	0.91

a FRET intesity equals to E(U)×([U]) + E(I_2)×[I_2] + E(F)×[F], E(I_2) = N1 × E(U), E(F) = N2 × E(U). N1 = QY(I_2) / QY (U) and N2 = QY(F) / QY (U) are constants. In the simulation, E(I_2) is a variable and N1, N2 are constants (±15% variation). The reansons are the same as in Table S2.

Simulation of system dynamics

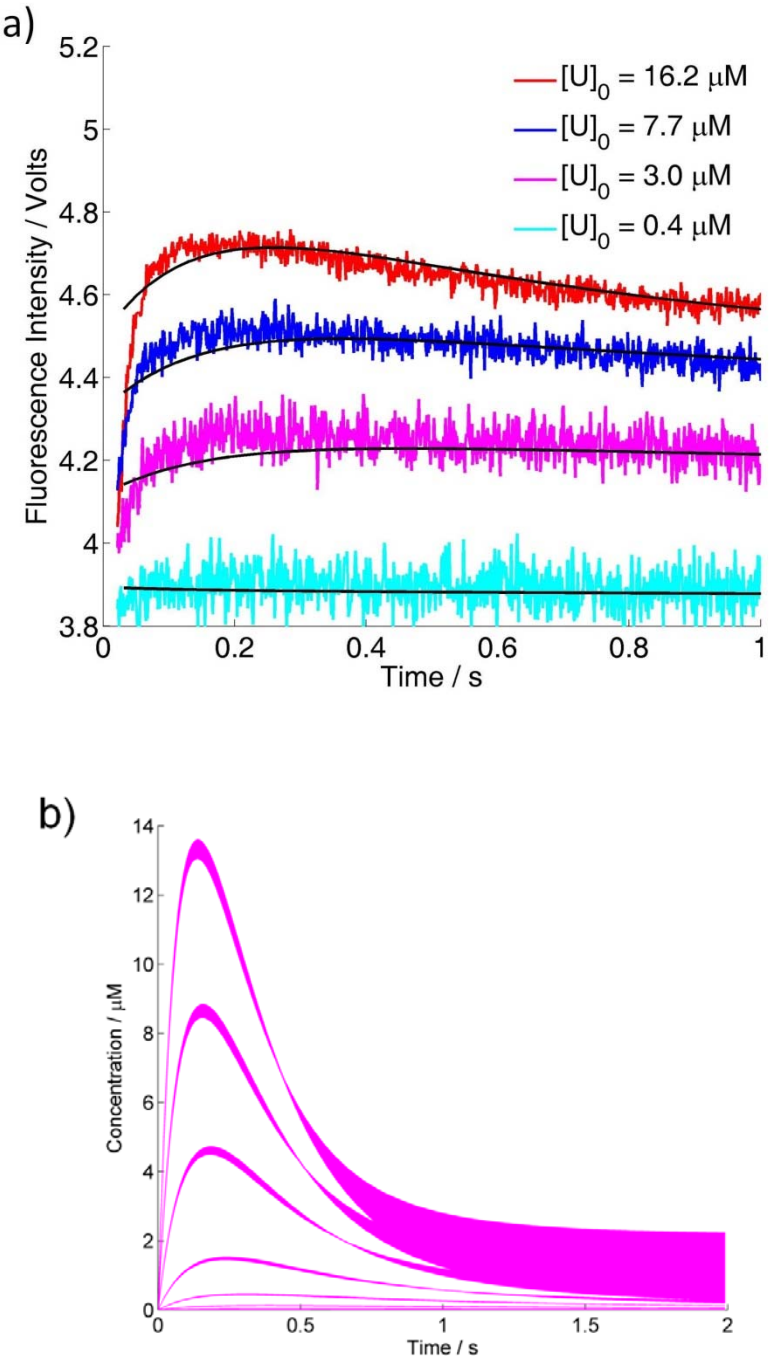


Figure S8. a) The enlarged version of Figure 3b (main text) in the 0-1s range. b)Simulated time course for I_2 at different initial concentrations. (from bottom to top: 6 μM, 12 μM, 25 μM, 50 μM, 75 μM, 100 μM)

Semi-logarithmic plot (Chevron Plot) of the observed rate constant

The refolding processes of DS119 were observed at different denaturant concentrations by the stopped-flow fluorescence. The final Gdn-HCl concentrations are 0.5 M, 0.6 M, 1.0 M, 1.5 M, 2.0 M and 2.5 M. For each denaturant concentration there is one kinetic curve which was fitted to the bimolecular model of DS119 (Figure S7).

Table S3. Microscopic rate constants derived from the global fitting of the stopped - flow kinetics data of DS119 at different final Gdn-HCl concentrations.

[Gdn-HCl] / M	k_1 / $\mu M^{-1}s^{-1}$	k_{-1} / s^{-1}	k_2 / s^{-1}	k_{-2} / $\mu M^{-1}s^{-1}$
0.5	9.10×10^{-2}	2.92×10^{-7}	12.5	1.00
0.6	6.56×10^{-2}	9.68×10^{-5}	37.8	3.04×10^{-1}
1.0	2.90×10^{-2}	2.05×10^{-4}	1.40×10^{-1}	1.53×10^{-4}
1.5	4.98×10^{-3}	5.68×10^{-2}	4.42×10^{-2}	9.33×10^{-6}
2.0	8.28×10^{-4}	2.80×10^{-1}	7.07×10^{-2}	4.05×10^{-9}
2.5	2.26×10^{-4}	2.26×10^{-7}	4.17×10^{-1}	$1.20\times10^{--2}$

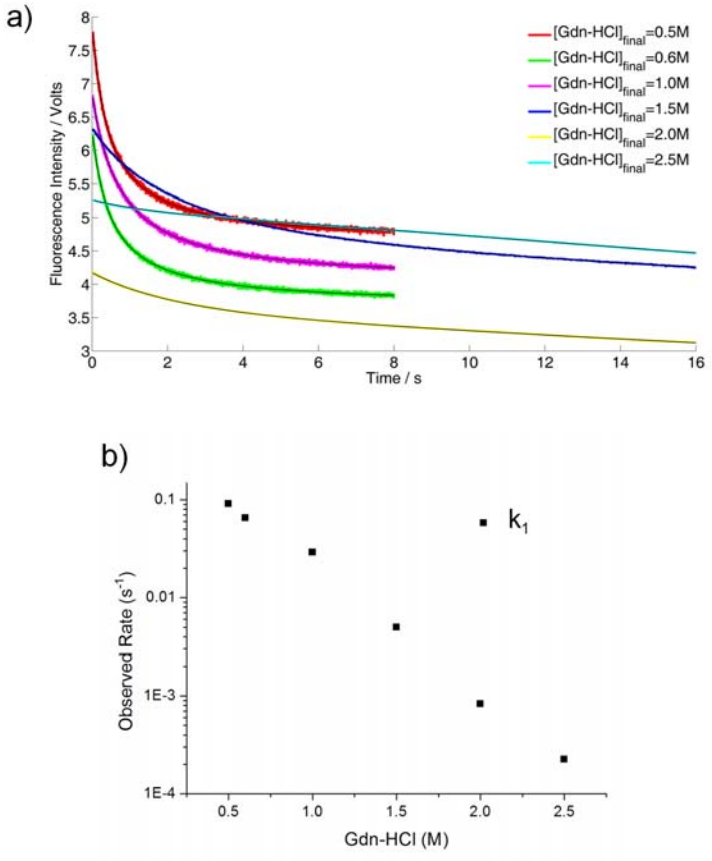


Figure S9. a) Refolding kinetics of DS119 at different denaturant concentrations. The fitting curves are black. b) Semi-logarithmic plot of the simulated rate constant k_1 versus [Gdn-HCl] at pH 7.3, 25°C.

Folding kinetics of DS119 at pH 2.5

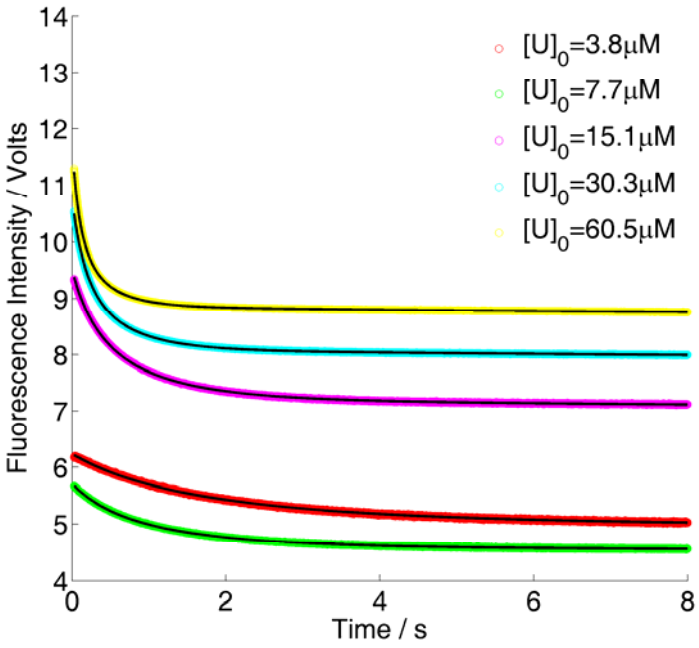
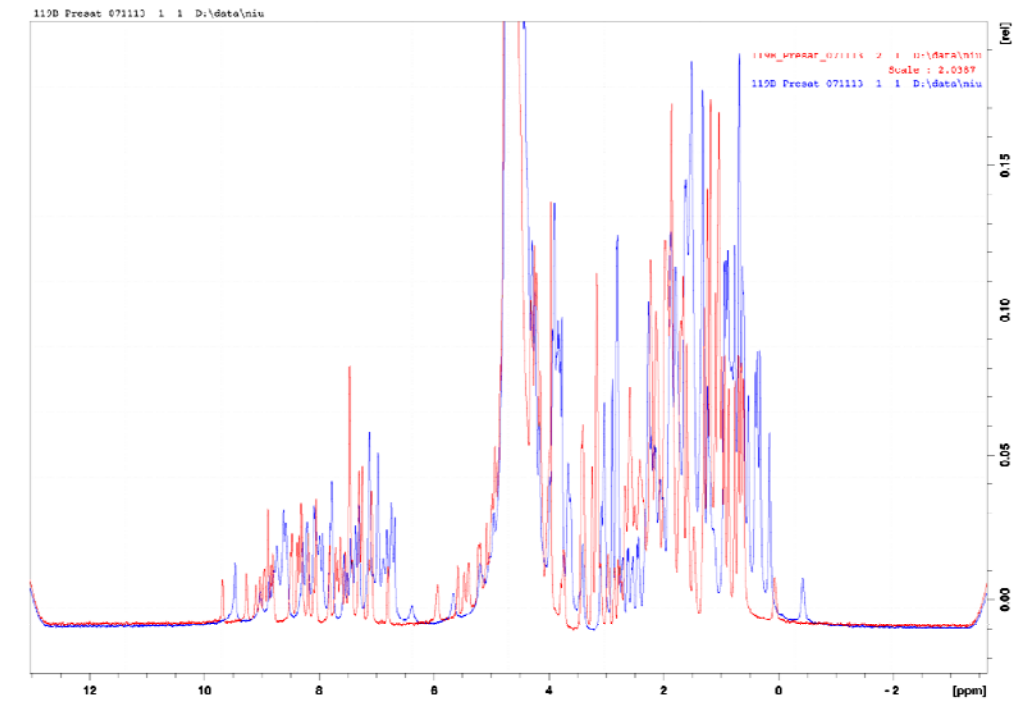


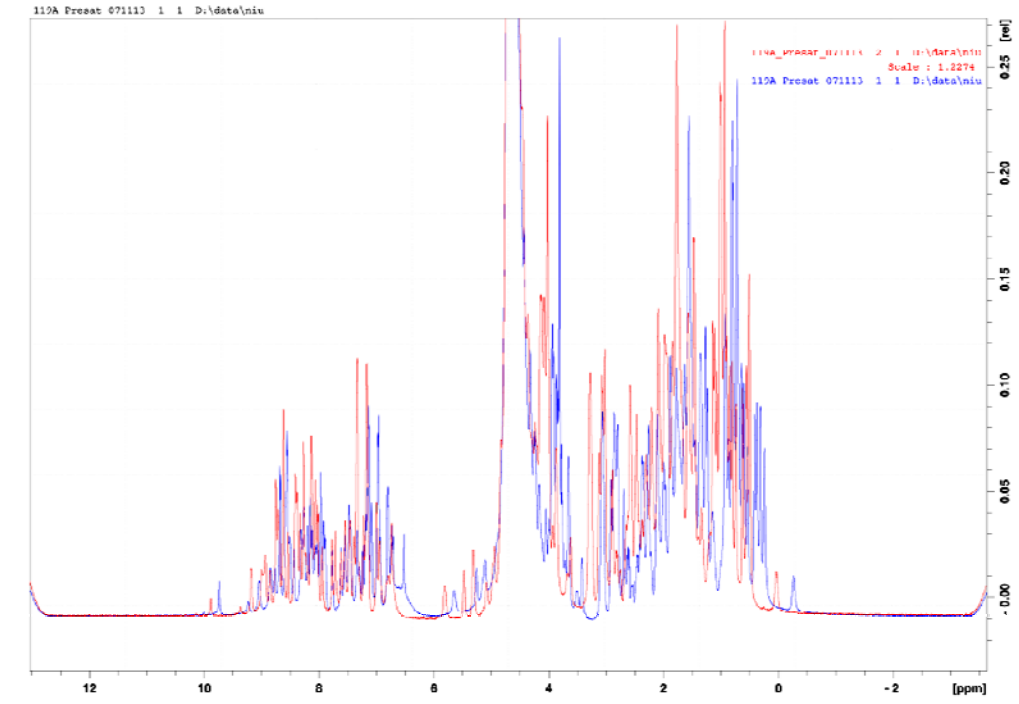
Figure S10. Stopped-flow refolding kinetics of DS103 measured by Trp fluorescence. The fitting curves are shown in black lines and the resulting rate constants are as follows: k_1 = 6.11×10^{-2} $\mu M^{-1}s^{-1}$, k_{-1} = 4.66×10^{-7} s^{-1}, k_2 = 3.72×10^{-2} s^{-1}, k_{-2} = 5.18×10^{-2} $\mu M^{-1}s^{-1}$.

NMR studies of DS119 at pH 7.3 and pH 2.5

a)



b)



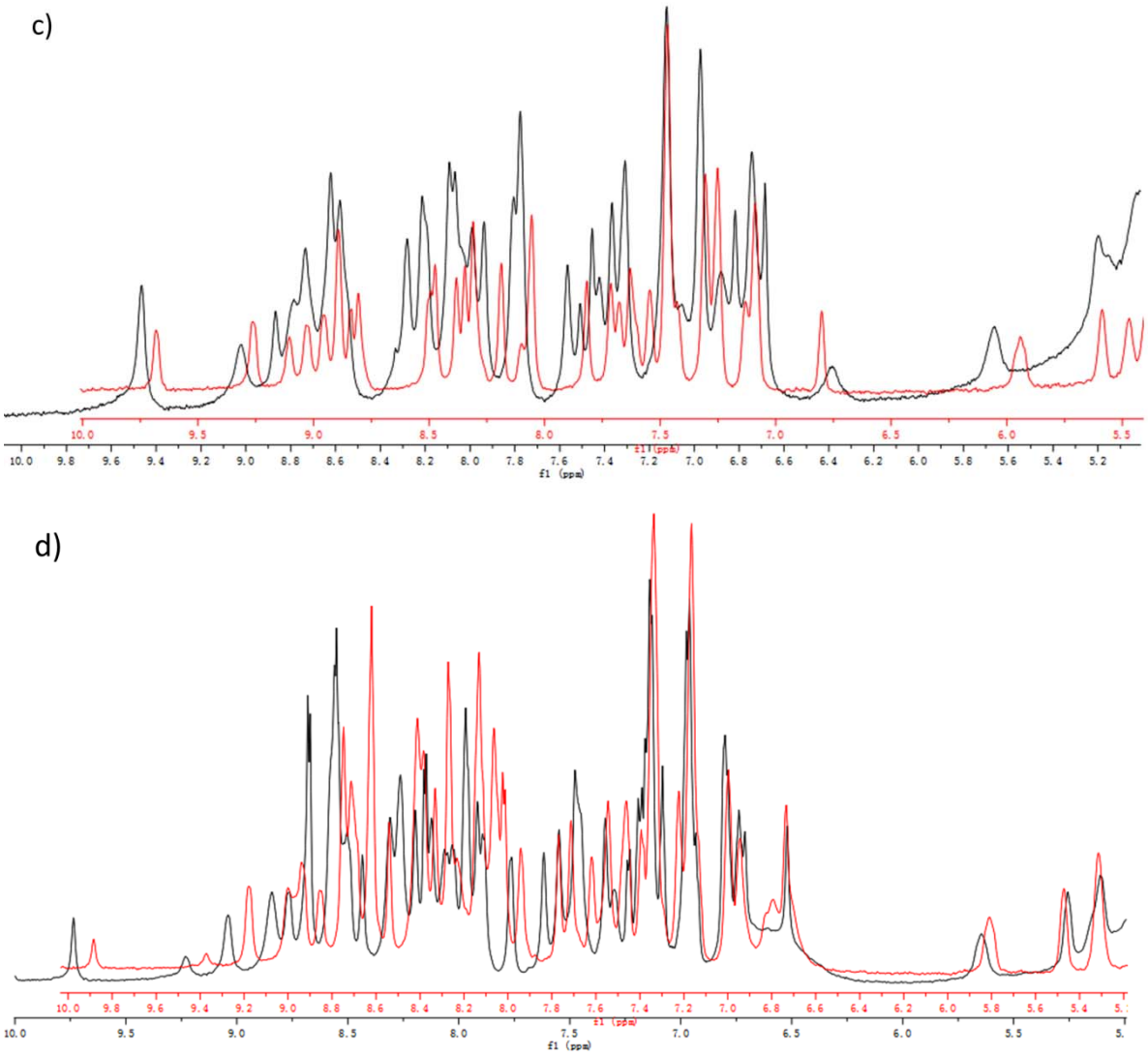


Figure S11. a) Proton 1D-NMR spectra of DS119 at 20°C (blue) and 60°C (red) in the pH 7.3 solution. b) Proton 1D-NMR spectra of DS119 at 20°C (blue) and 46°C (red) in the pH 2.5 solution. c) The enlarged figure of the NH region (5-10 ppm) in the NMR spectra at 20°C (black) and 60°C (red) in the pH 7.3 solution. d) The enlarged figure of the NH region (5-10 ppm) in the NMR spectra at 20°C (black) and 46°C (red) in the pH 2.5 solution.

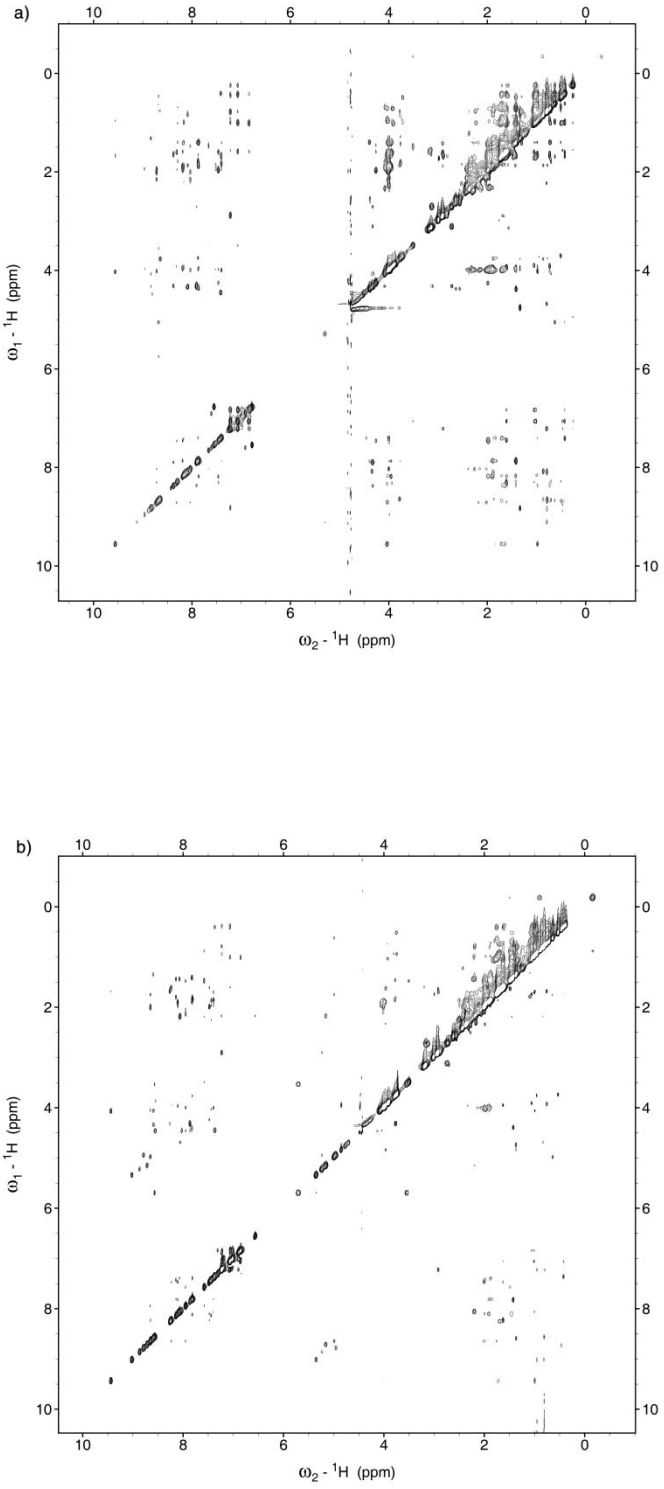


Figure S12. NOSEY spectra of DS119 at 20°C (a) and 60°C (b) in the pH 7.3 solution. The contour levels are set as the same (Positive contours: low 4.00e+06, Negative contours: low -8.00e+07).

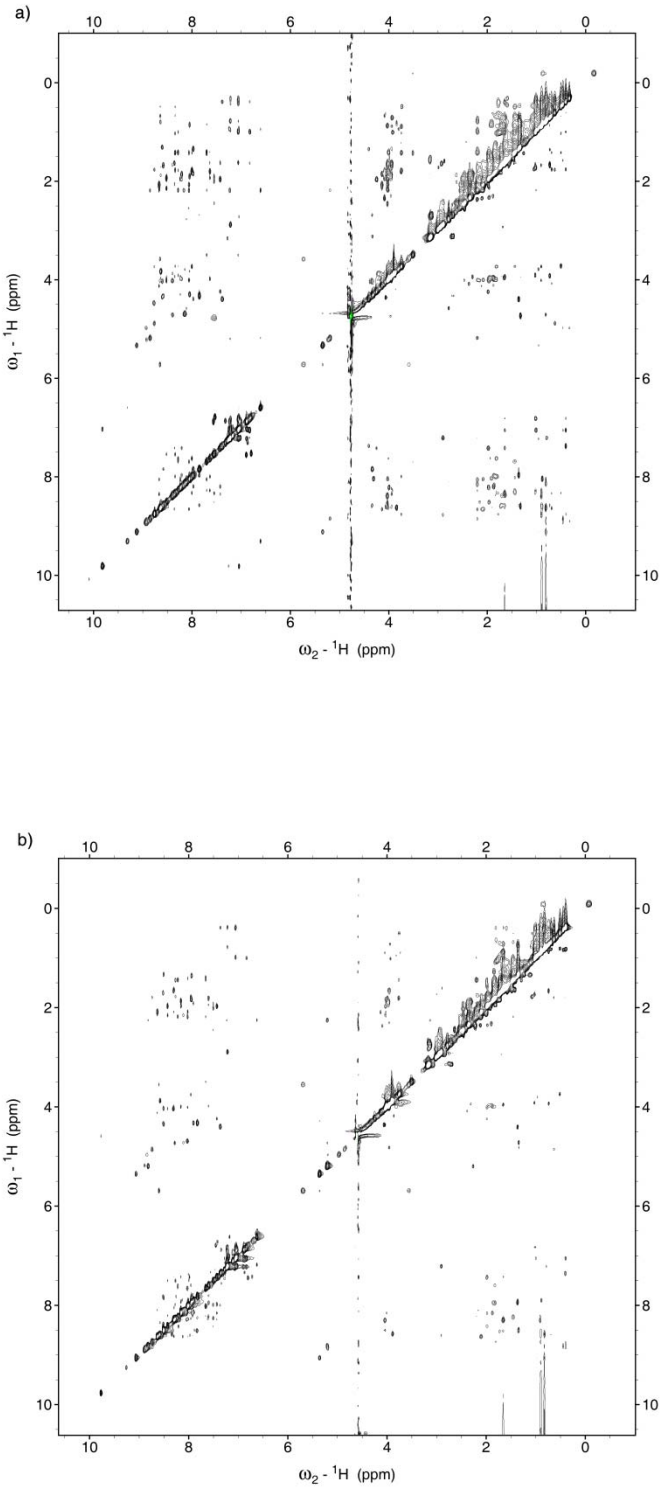


Figure S13. NOSEY spectra of DS119 at 20°C (a) and 46°C (b) in the pH 2.5 solution. The contour levels are set as the same (Positive contours: low 4.00e+06, Negative contours: low -8.00e+07).

Thermodynamic properties and folding mechanisms of DS119 mutants (DS103, W9F, W9E and P14A)

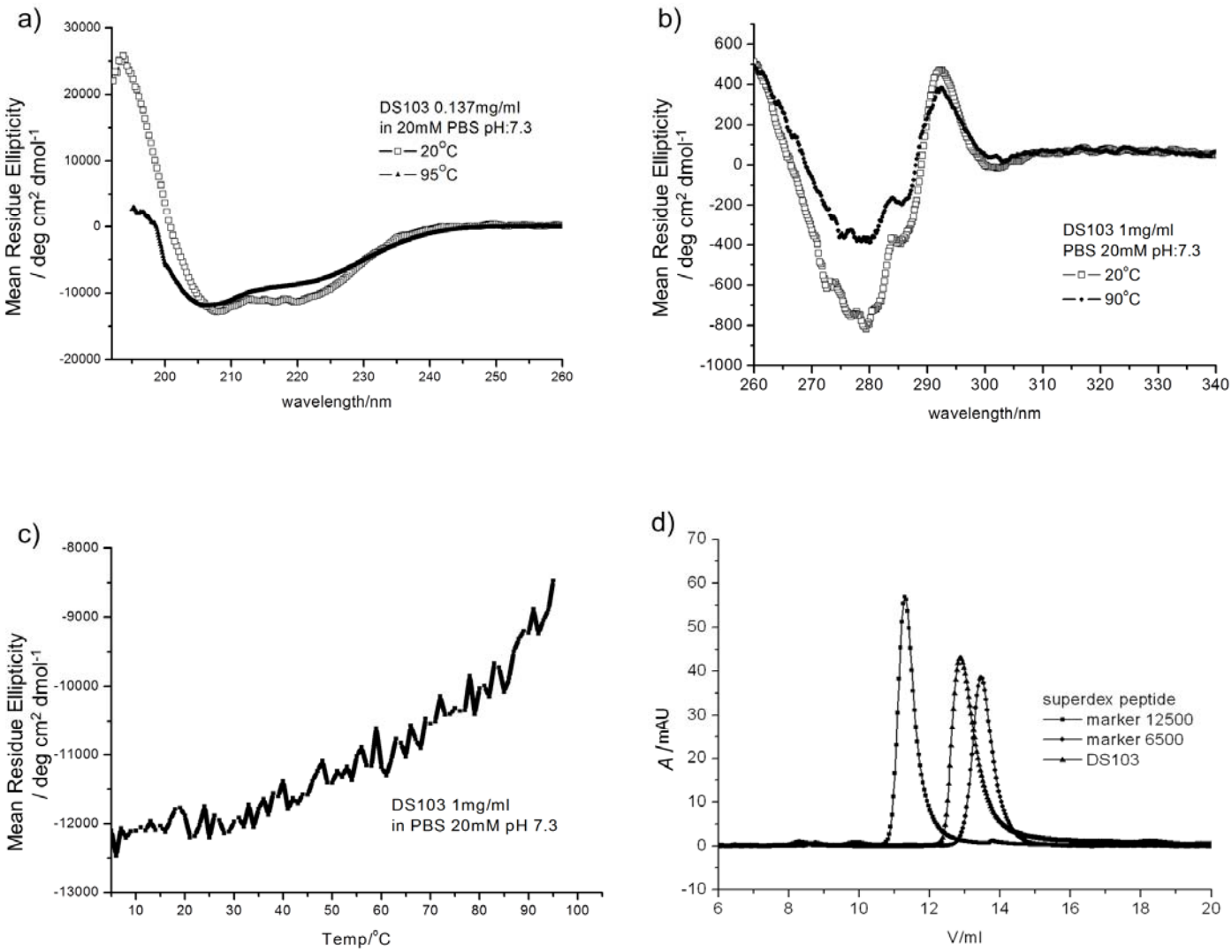
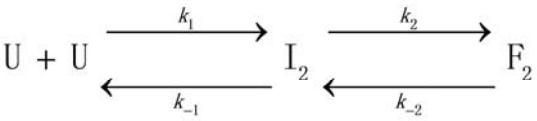


Figure S14. a) Far-UV CD spectra of DS103. b) Near-UV CD spectra. c) Thermal denaturation curves. The CD signal at 222 nm was monitored. d) Gel filtration curves. The molecular weight of a monomer is 4172.6. The apparent molecular weight of DS103 is about 8200 on the Superdex^{TM} Peptide 10/300 GL column.



$$\frac{d[U]}{dt} = -2k_1[U]^2 + 2k_{-1}[I_2]$$

$$\frac{d[I_2]}{dt} = k_1[U]^2 - k_{-1}[I_2] - k_2[I_2] + k_{-2}[F_2]$$

$$\frac{d[F_2]}{dt} = k_2[I_2] - k_{-2}[F_2]$$

Figure S15. The bimolecular folding model and differential rate equations for the folding process of DS103.

Table S4. Microscopic rate constants derived from the global fitting of the stopped - flow kinetics data 0f DS103.

k_1 / $\mu M^{-1}s^{-1}$	2.10×10^{-1}	[0.205 0.216]
k_{-1} / s^{-1}	2.30	[2.23 2.37]
k_2 / s^{-1}	1.35×10^{-2}	[0.0132 0.0225]
k_{-2} / s^{-1}	1.78×10^{-1}	[0.164 0.191]

a The 95% confidence intervals are shown in the brackets.

Table S5. The simulated fluorescent emission parameters for stopped-flow Trp fluorescence of DS103.

E (V/μM) / $[U]_0$ (μM)	E (U)[a]	M1[b]	M2
38	0.153	1.16	0.0027
19	0.266	1.24	0.046
9.5	0.522	1.36	0.49
4.8	1.02	1.58	0.29
2.4	1.74	1.69	0.19

a Trp fluorescence intesity equals to $E(U)\times([U]) + E(I_2)\times[I_2] + E(F_2)\times[F_2]$, $E(I_2) = M1 \times E(U)$, $E(F_2) = M2 \times E(U)$. $[U]_0$ is the initial concentration of DS103 after mixing. Since the response of the instrument is not a constant in different experiments (the voltage of the detector was changed), E(U) is a variable in our simulation.

b M is the ratio between E(U) and $E(I_2)$, $E(F_2)$. Since the response of the instrument is a constant in each experiment, $M1 = QY(I_2) / QY(U)$ and $M2 = QY(F_2) / QY(U)$ are also constants. In the simulation, we allowed a ±15% variation of M1 and M2

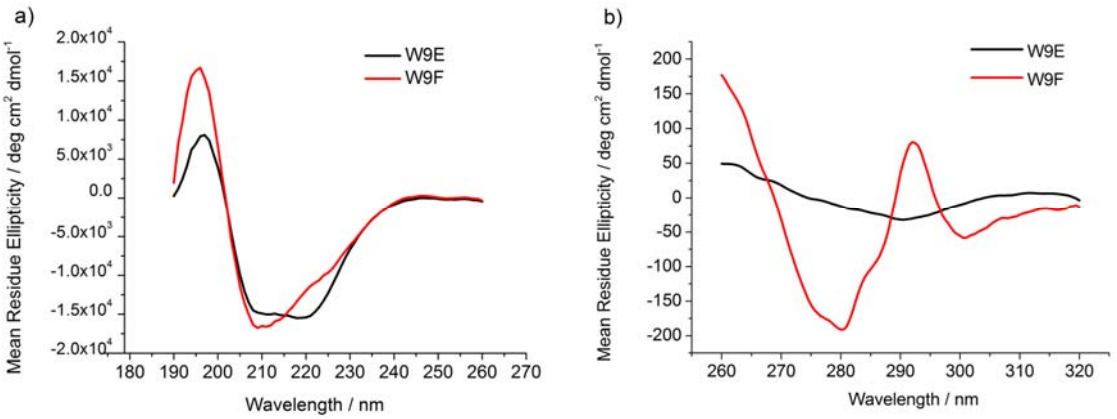


Figure S16. a) Far-UV CD spectra of W9E and W9F. b) Near-UV CD spectra. Both experiments were measured in 50 mM PB buffer (pH 7.3) at 25°C. The protein concentrations are 50μM.

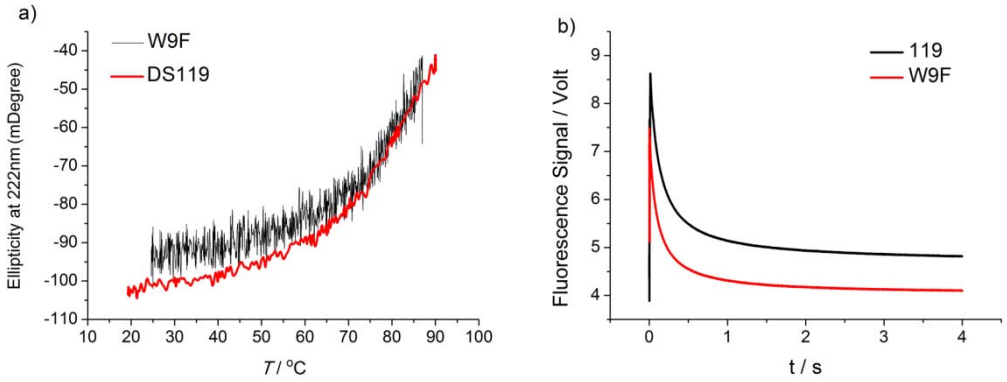


Figure S17. a) Thermal denaturation curves of DS119 and W9F. b)Stopped-flow refolding kinetics of W9F and DS119 measured by fluorescence. The protein concentrations after mixing are 35μM.

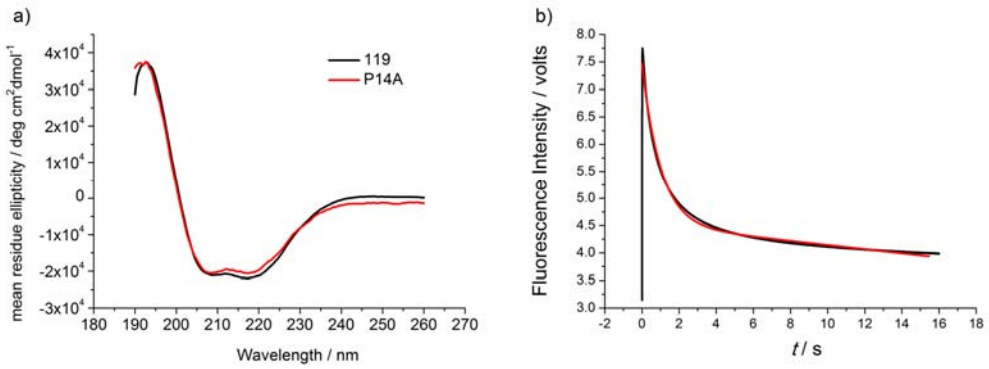


Figure S18. a) Far-UV CD spectra of DS119 and P14A measured in 50 mM PB buffer (pH 7.3) at 25 °C. b) Refolding kinetics of P14A measured by stopped-flow fluorescence. Fitting these data to a single-exponential function (red line) yielded a time constant of 426 ms. The final concentration of P14A was 31 μM. Other conditions were the same as used for DS119.

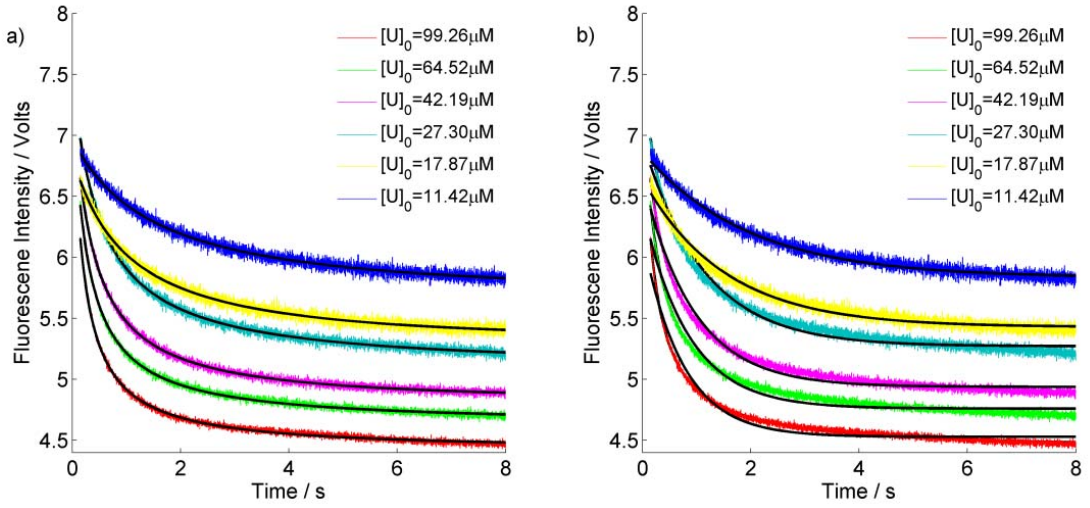


Figure S19. Global fitting of the stopped-flow Trp fluorescence data of DS119 with a a) 2nd reaction model ($y = At + B + \frac{C}{1+Dt}$); b) 1st reaction model ($y = At + B + Ce^{-kt}$).

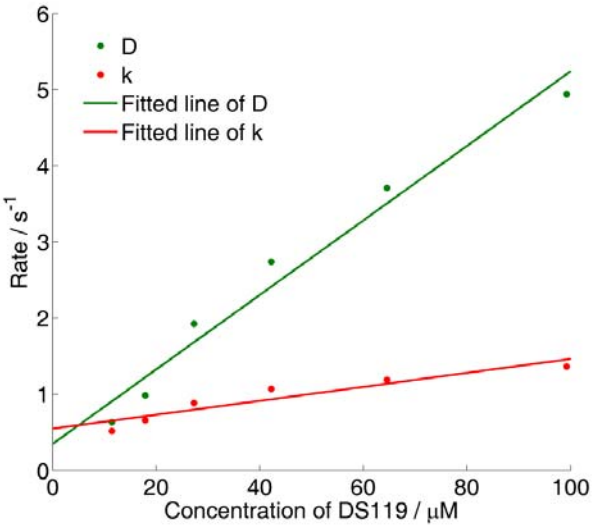


Figure S20. The relationship between the folding rate constants (s^{-1}) and the protein concentration (D: 2nd reaction rate constants; k: 1st reaction rate constants). The extrapolation to zero concentrations led to D (0 μM) = 0.35 s^{-1} and k (0 μM) = 0.55 s^{-1}.

Unfolding kinetics of DS119

The unfolding kinetics of DS119 was measured by mixing DS119 in the native state (Na_2HPO_4 / NaH_2PO_4 pH 7.3 solutions) and high concentration of Gdn-HCl (6M). The Trp fluorescence was recorded. We found that in the higher concentration range (118 – 15 μM) the unfolding rate is strongly concentration-dependent (Fig. 21a), while in the lower range (7.5 – 3.8 μM) the concentration-dependence became less obvious (Fig. 21b), indicating the unfolding process probably have a more complex mechanism and at low concentrations a monomeric unfolding pathway became more dominant. In addition, the relative poor quality of the data likely results from a combination of incomplete mixing (at shorter time, due to high viscosity of the solution) and instability of the instrumentation (at longer time). We think the detailed unfolding mechanism needs further investigation.

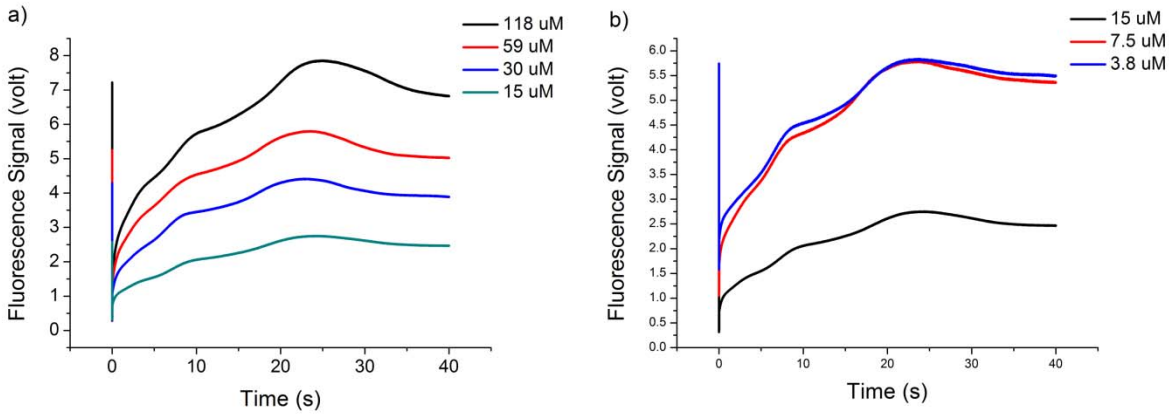


Figure S21. Concentration dependence of the unfolding process of DS119. The final concentrations of DS119 were a) 118 – 15 μM; b) 15 – 3.8 μM. The final concentration of Gdn-HCl was 5.1M.