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

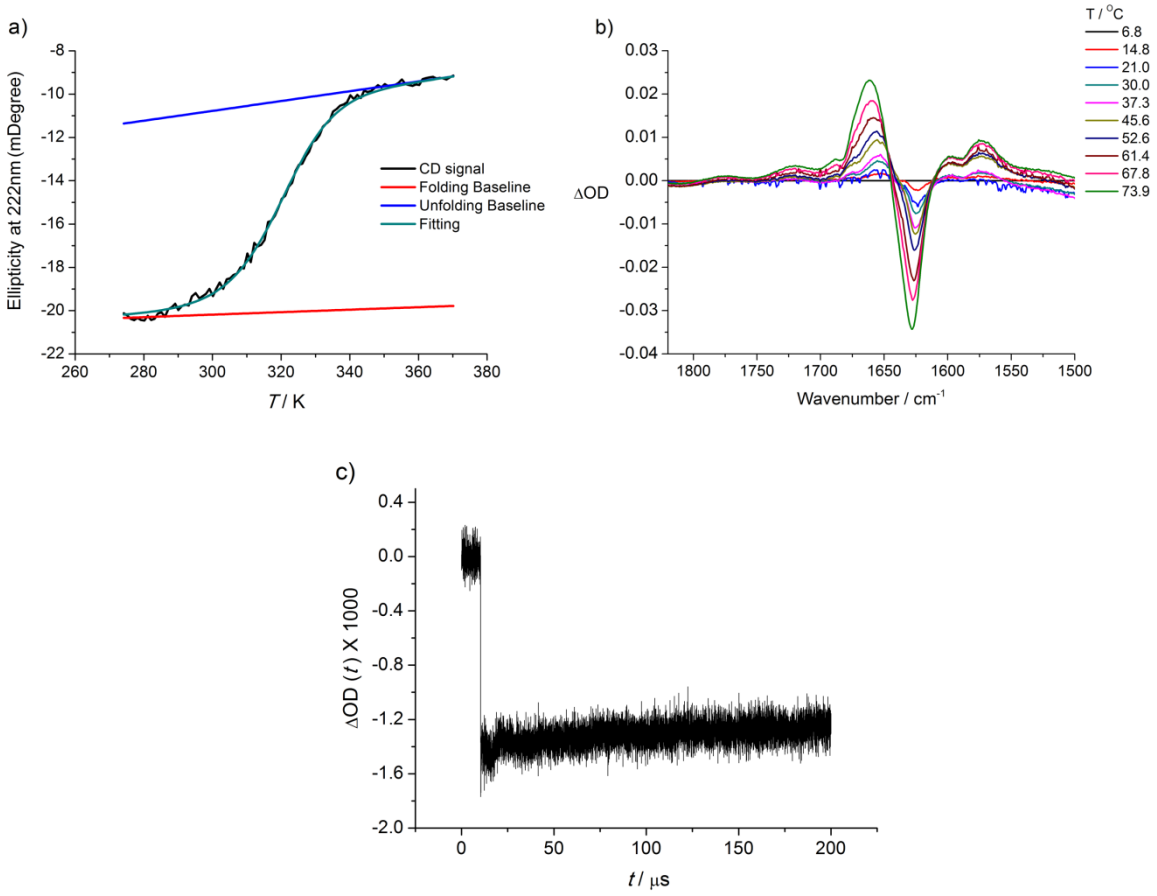


Figure S1. a) Thermal denaturation curve of DS119. Thermal unfolding was measured from 274 K to 370 K in 50 mM PB buffer (pH 2.5). The CD signal at 222 nm was monitored. Fitting to a two-state model (green line) indicated the T_m is 46.2°C at PH 2.5. b) Difference IR spectra of DS119 in 50mM PB buffer (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

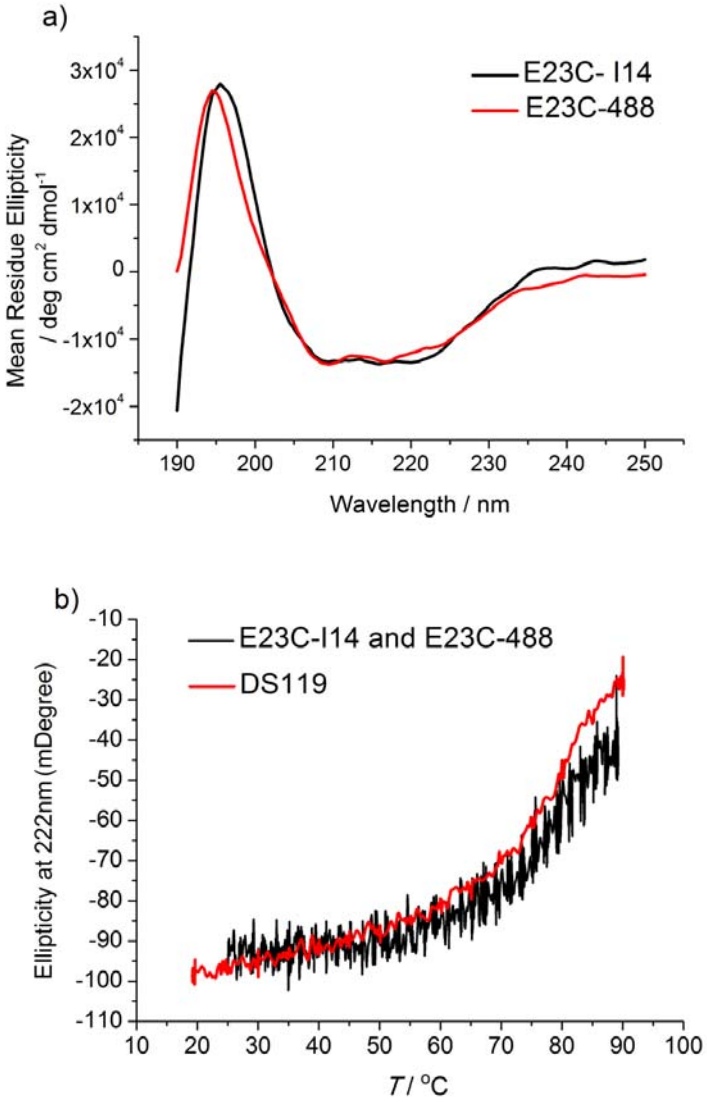


Figure S2. a) Far-UV CD spectra of E23C-I14 and E23C-488 measured in 50 mM PB buffer (pH 7.3). b) Thermal denaturation curves of E23C-I14 and E23C-488. Thermal unfolding was measured from 20°C to 90°C in 50 mM PB buffer (pH 7.3). The CD signal at 222 nm was monitored.

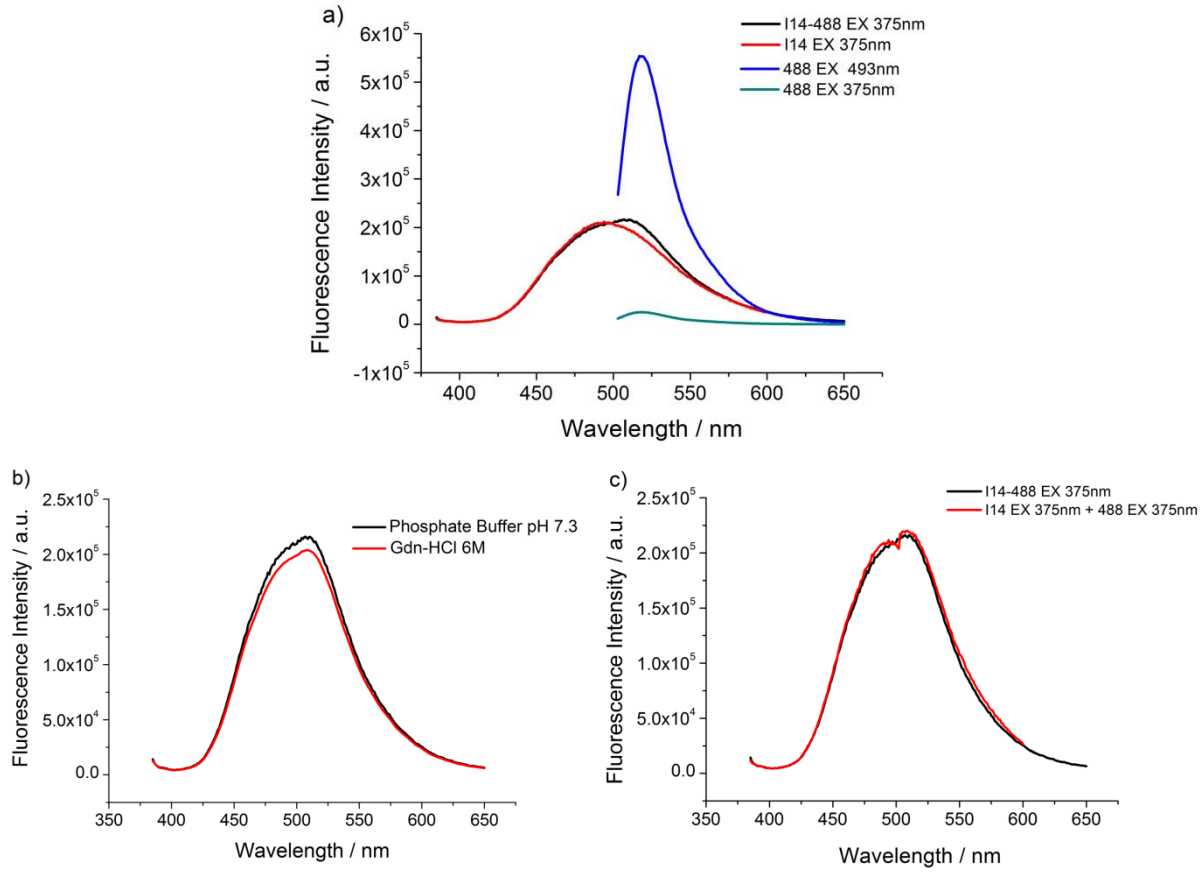


Figure S3. a) Steady-state fluorescence spectra of E23C-I14 and E23C-488. The excitation wavelengths of I14 and Alexa 488 were 375 nm and 493 nm, and their emission peaks were at 492 nm and 518 nm, respectively. Black : the fluorescent spectra of an E23C-I14 (12 μM) and E23C-488 (15 μM) mixture excited 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).

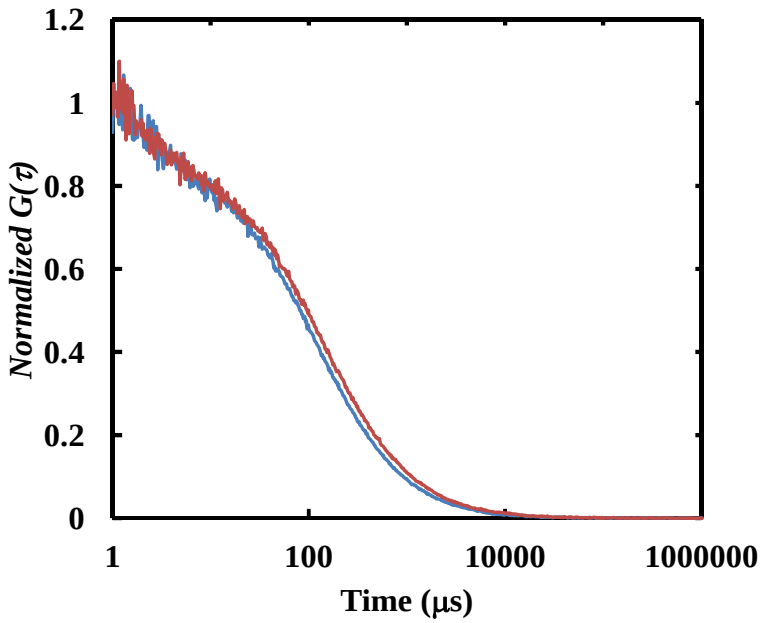


Figure S4. Normalized autocorrelation traces measured with a 16 nM labeled DS119 solution (blue) and a mixture of 8 nM labeled-DS119 and 100 μM DS119 (red). Both samples were prepared in 50 mM phosphate buffer (pH = 7).

Analytical gel filtration of DS119

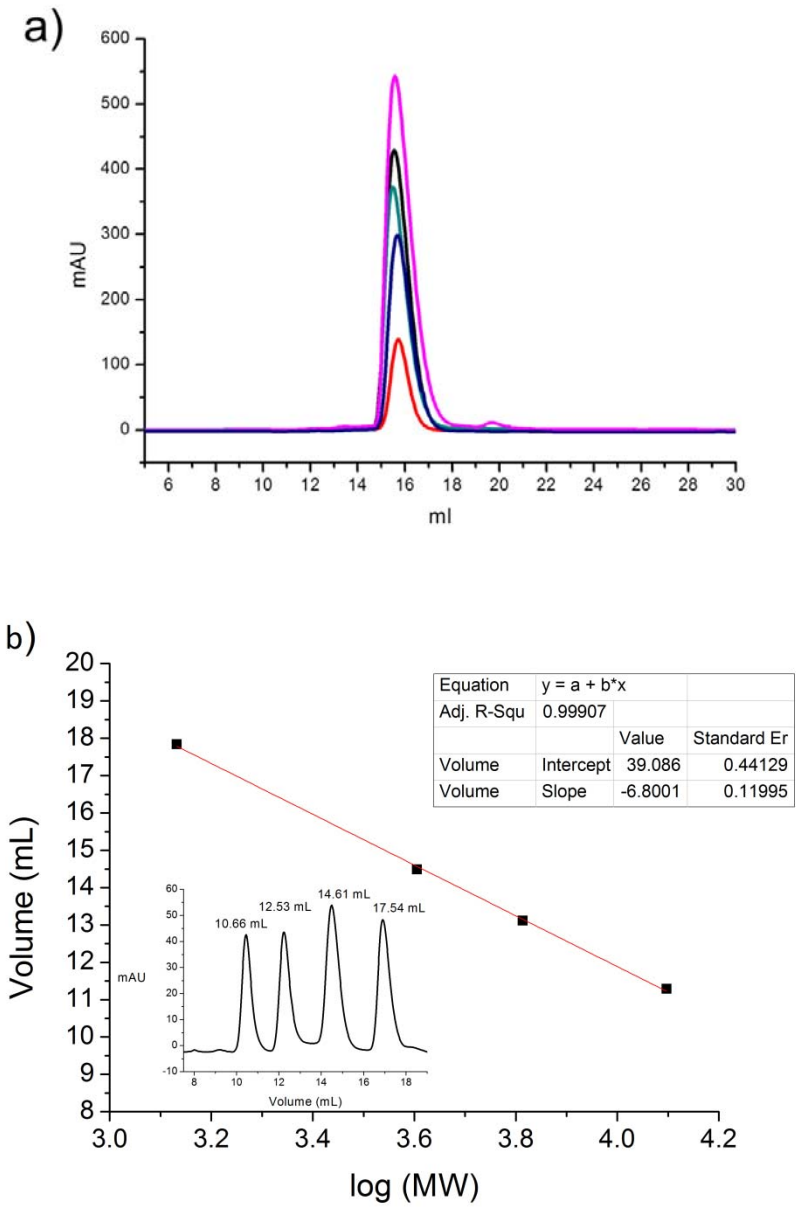


Figure S5. a) Analytical gel filtration of DS119 at different concentrations. Protein samples were dissolved in the buffers containing 50 mM Na_2HPO_4 / NaH_2PO_4 (pH7.3) and 100 mM NaCl, and then loaded on a Superdex^{TM} Peptide 10/300 GL column. The protein concentrations of each curve were (from 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).

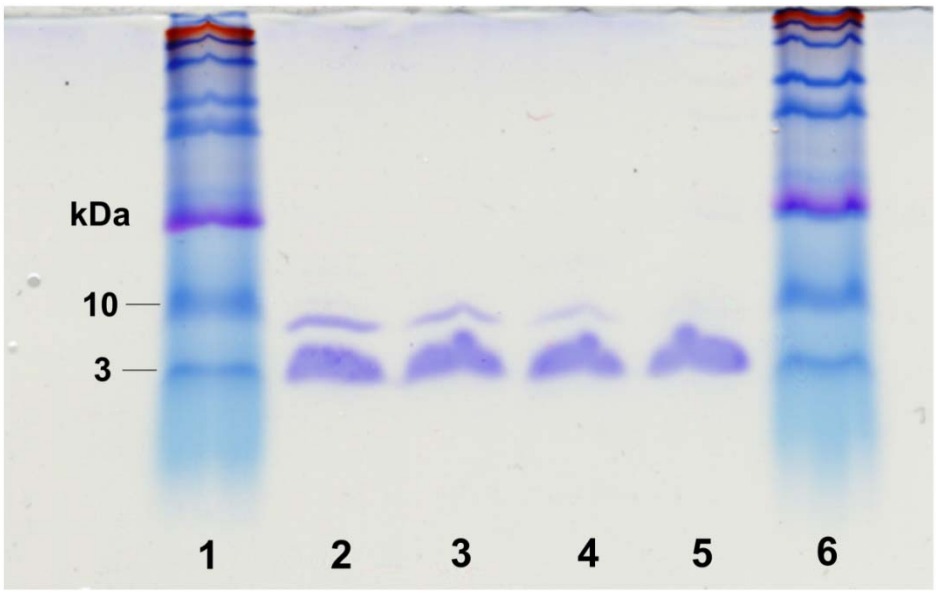


Figure S6. SDS-PAGE gels of E23C submitted to chemical cross-linking in the refolding process (lanes 2- 4) and the steady state (lane 5). In the refolding process cross linking, the denatured E23C was diluted to solutions with crosslinker (thiol-reactive BM-PEG3). In the steady state cross linking, the 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

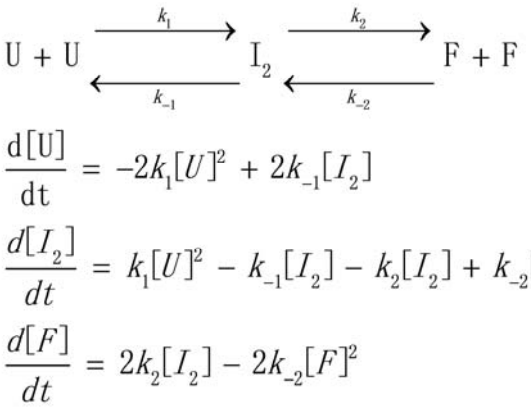


Figure S7. The bimolecular folding model and differential rate equations for the folding process of DS119.

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

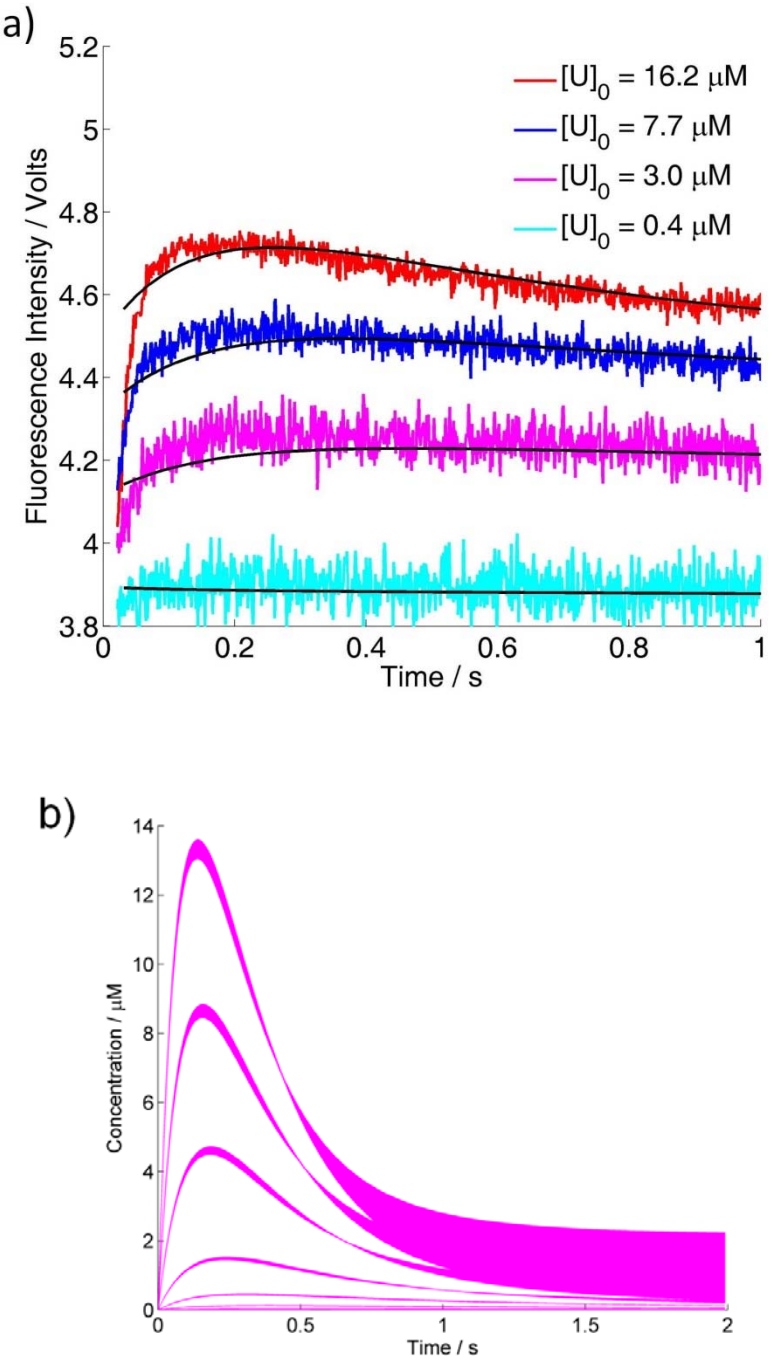


Figure S8. a) The enlarged version of Figure 3b (main text) in the 0-1s range. b)Simulated time course for I_2 at different initial concentrations. (from bottom to top: 6 μM, 12 μM, 25 μM, 50 μM, 75 μM, 100 μM)

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}	$1.20\times10^{--2}$

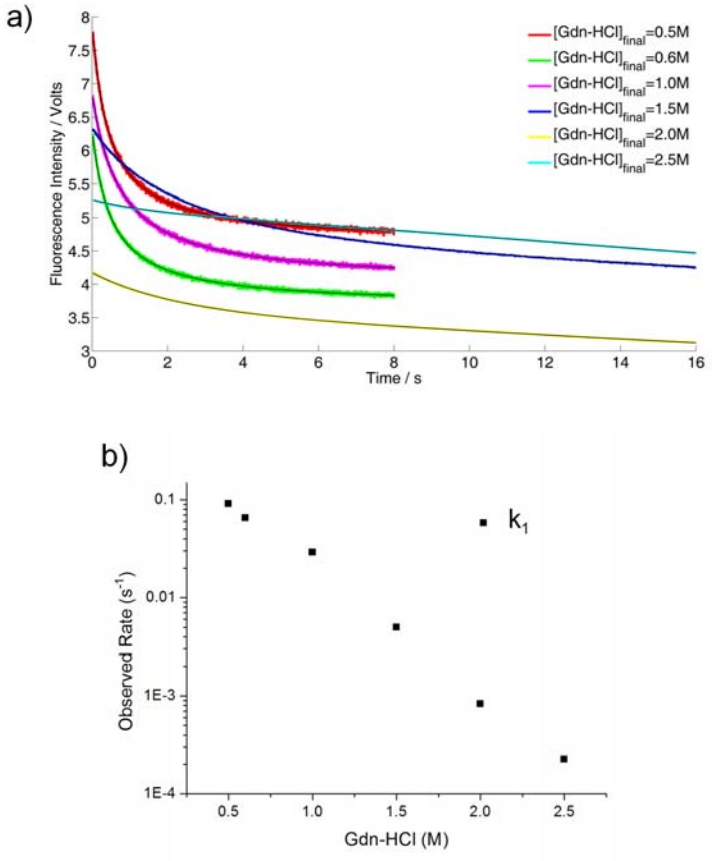


Figure S9. a) Refolding kinetics of DS119 at different denaturant concentrations. The fitting curves are black. b) Semi-logarithmic plot of the simulated rate constant k_1 versus [Gdn-HCl] at pH 7.3, 25°C.

Folding kinetics of DS119 at pH 2.5

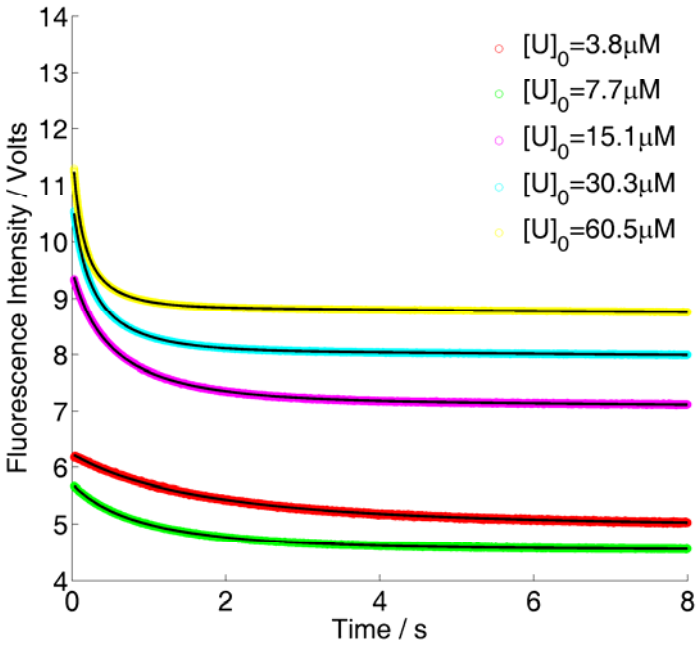
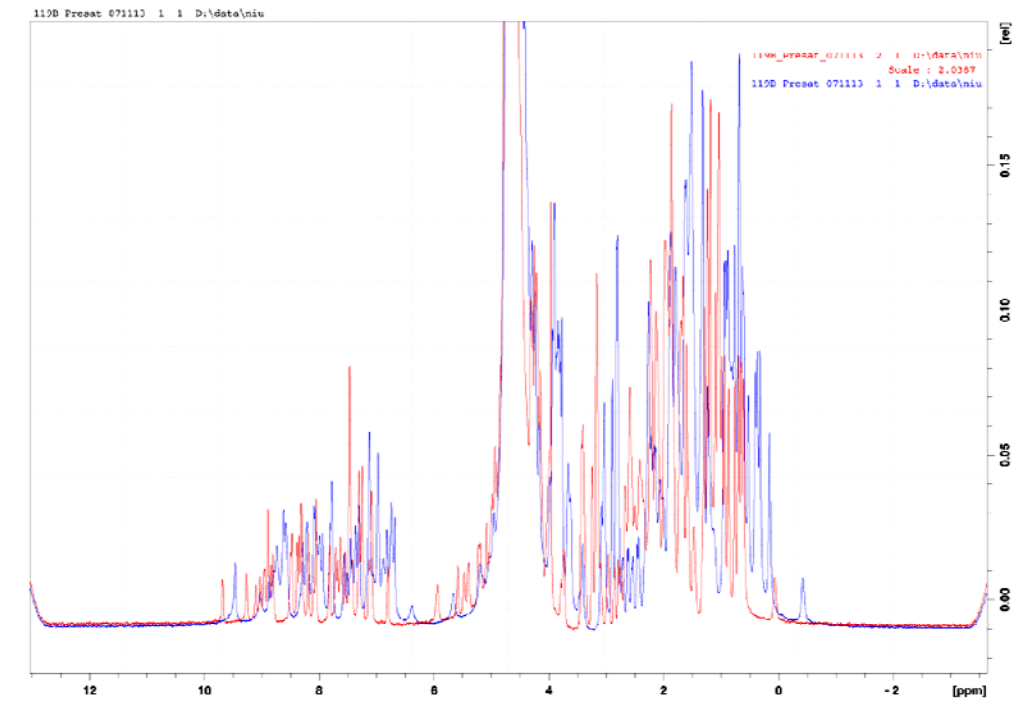


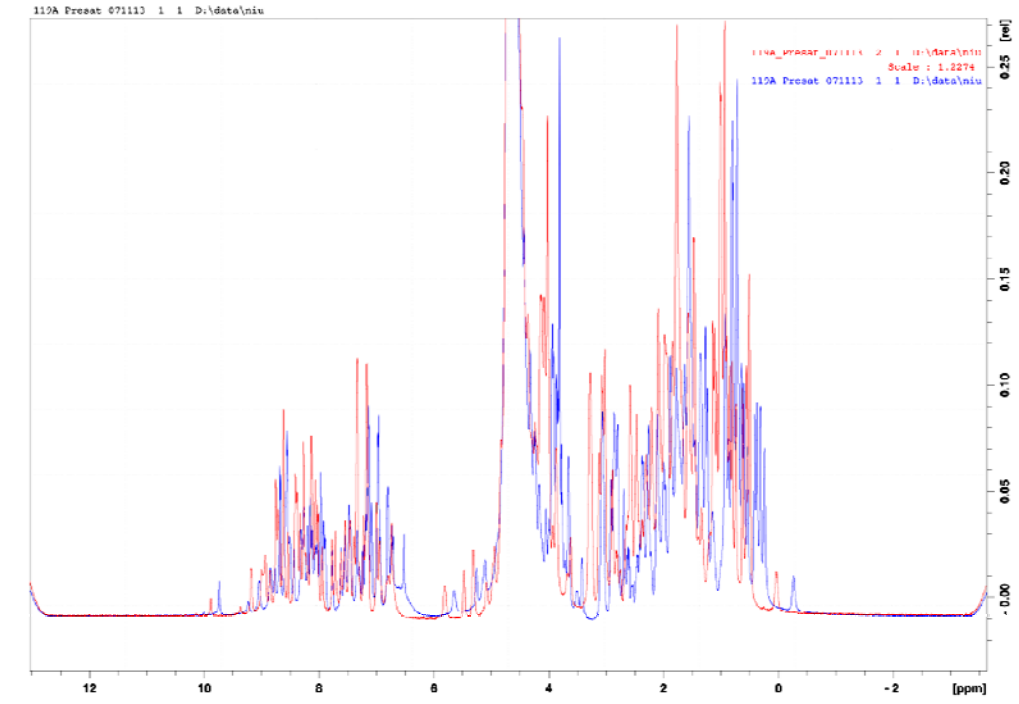
Figure S10. Stopped-flow refolding kinetics of DS103 measured by Trp fluorescence. The fitting curves are shown in black lines and the resulting rate constants are as follows: k_1 = 6.11×10^{-2} $\mu M^{-1}s^{-1}$, k_{-1} = 4.66×10^{-7} s^{-1}, k_2 = 3.72×10^{-2} s^{-1}, k_{-2} = 5.18×10^{-2} $\mu M^{-1}s^{-1}$.

NMR studies of DS119 at pH 7.3 and pH 2.5

a)



b)



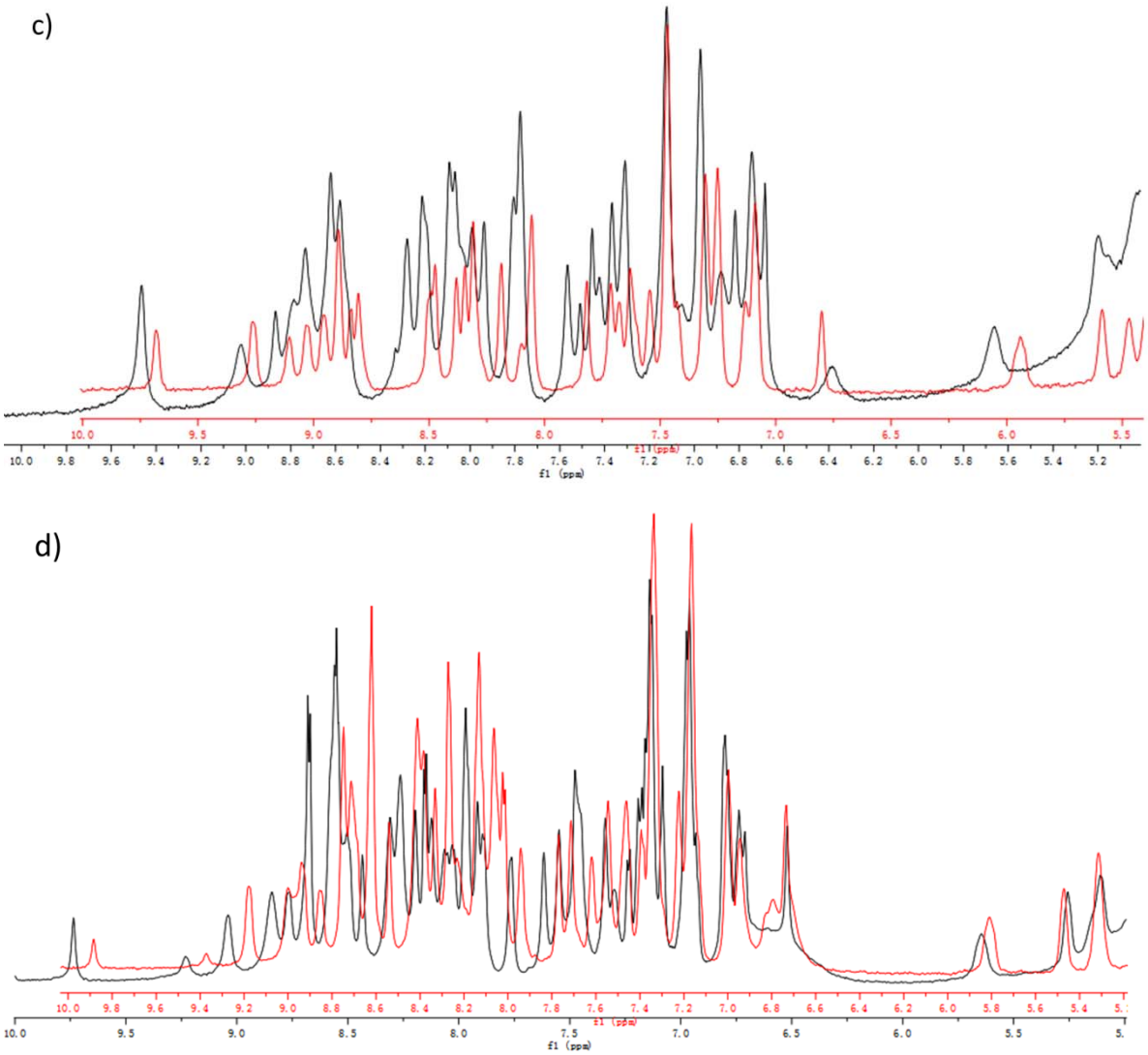


Figure S11. a) Proton 1D-NMR spectra of DS119 at 20°C (blue) and 60°C (red) in the pH 7.3 solution. b) Proton 1D-NMR spectra of DS119 at 20°C (blue) and 46°C (red) in the pH 2.5 solution. c) The enlarged figure of the NH region (5-10 ppm) in the NMR spectra at 20°C (black) and 60°C (red) in the pH 7.3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

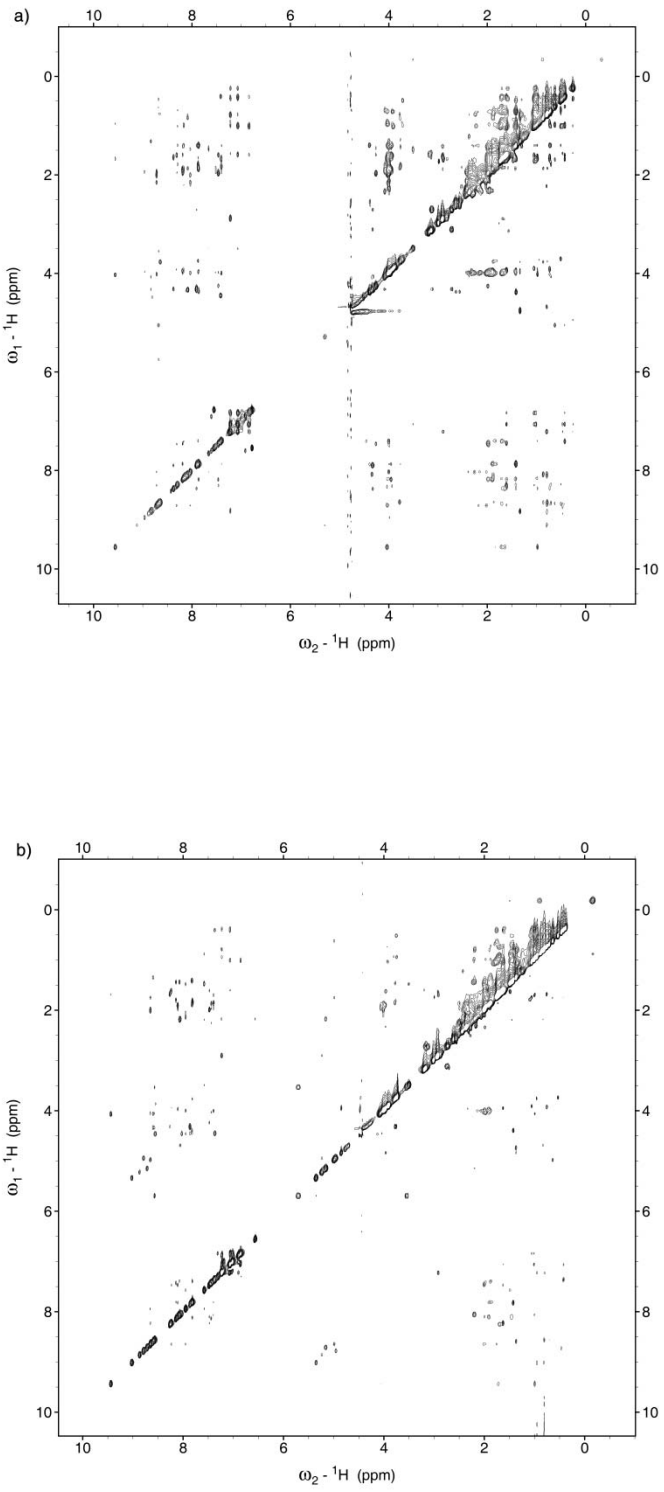


Figure S12. NOSEY spectra of DS119 at 20°C (a) and 60°C (b) in the pH 7.3 solution. The contour levels are set as the same (Positive contours: low 4.00e+06, Negative contours: low -8.00e+07).

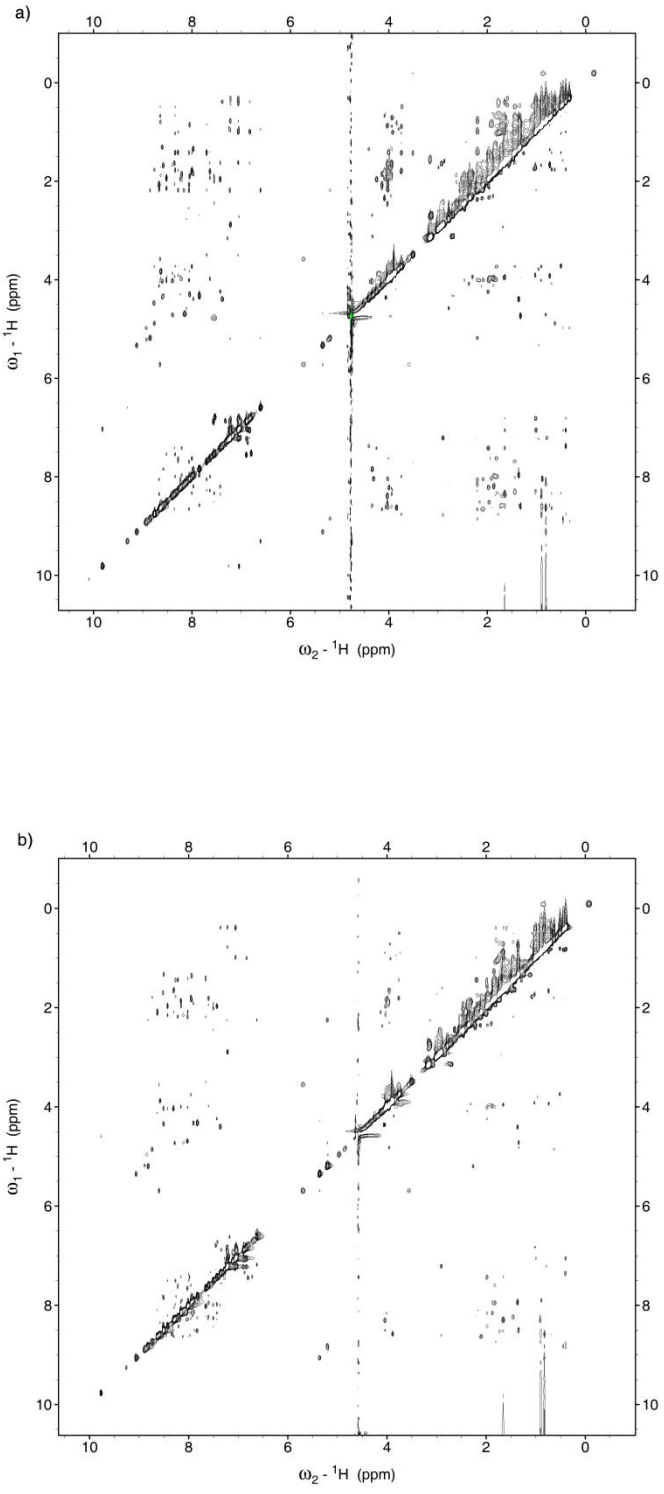


Figure S13. NOSEY spectra of DS119 at 20°C (a) and 46°C (b) in the pH 2.5 solution. The contour levels are set as the same (Positive contours: low 4.00e+06, Negative contours: low -8.00e+07).

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

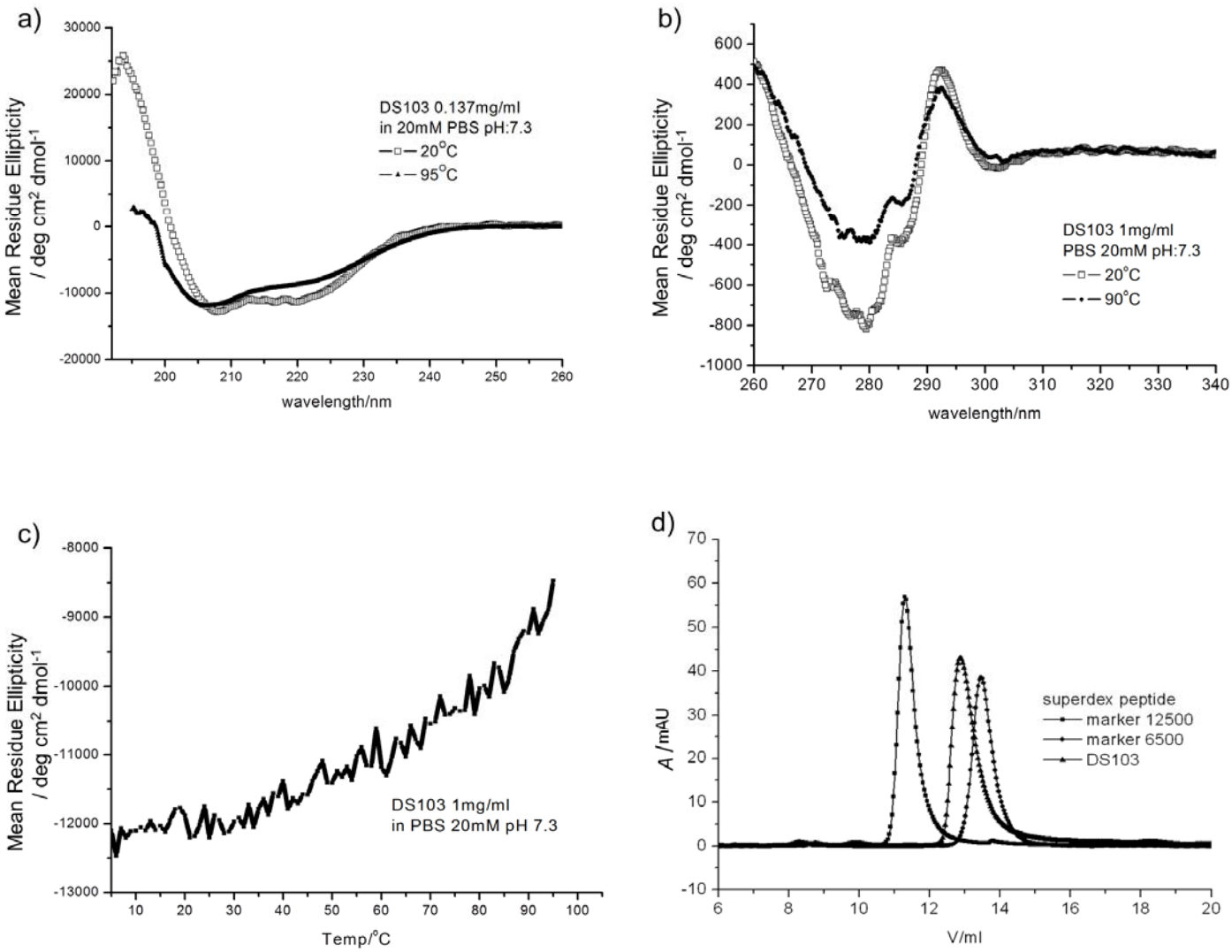
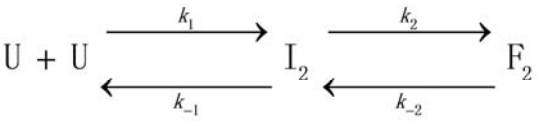


Figure S14. a) Far-UV CD spectra of DS103. b) Near-UV CD spectra. c) Thermal denaturation curves. The CD signal at 222 nm was monitored. d) Gel filtration curves. The molecular weight of a monomer is 4172.6. The apparent molecular weight of DS103 is about 8200 on the Superdex^{TM} Peptide 10/300 GL column.



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

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

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

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

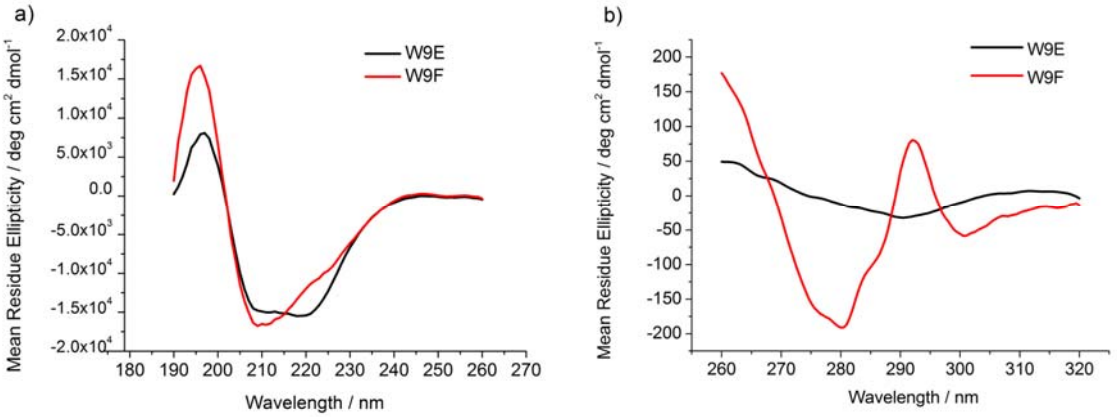


Figure S16. a) Far-UV CD spectra of W9E and W9F. b) Near-UV CD spectra. Both experiments were measured in 50 mM PB buffer (pH 7.3) at 25°C. The protein concentrations are 50μM.

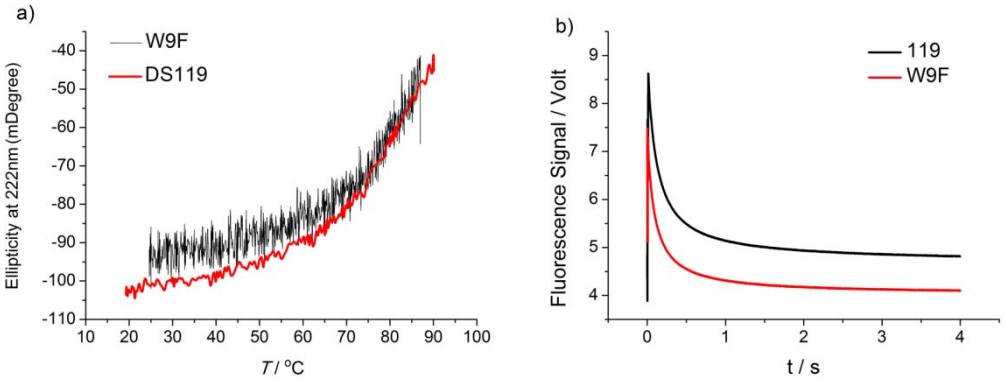


Figure S17. a) Thermal denaturation curves of DS119 and W9F. b)Stopped-flow refolding kinetics of W9F and DS119 measured by fluorescence. The protein concentrations after mixing are 35μM.

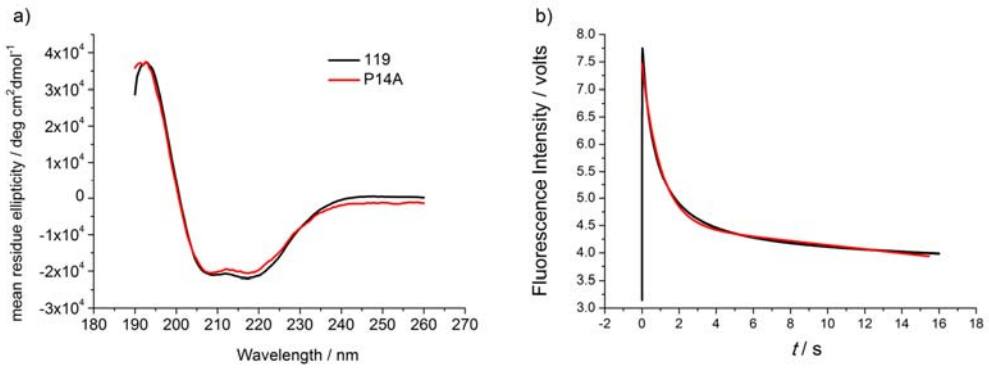


Figure S18. a) Far-UV CD spectra of DS119 and P14A measured in 50 mM PB buffer (pH 7.3) at 25 °C. b) Refolding kinetics of P14A measured by stopped-flow fluorescence. Fitting these data to a single-exponential function (red line) yielded a time constant of 426 ms. The final concentration of P14A was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

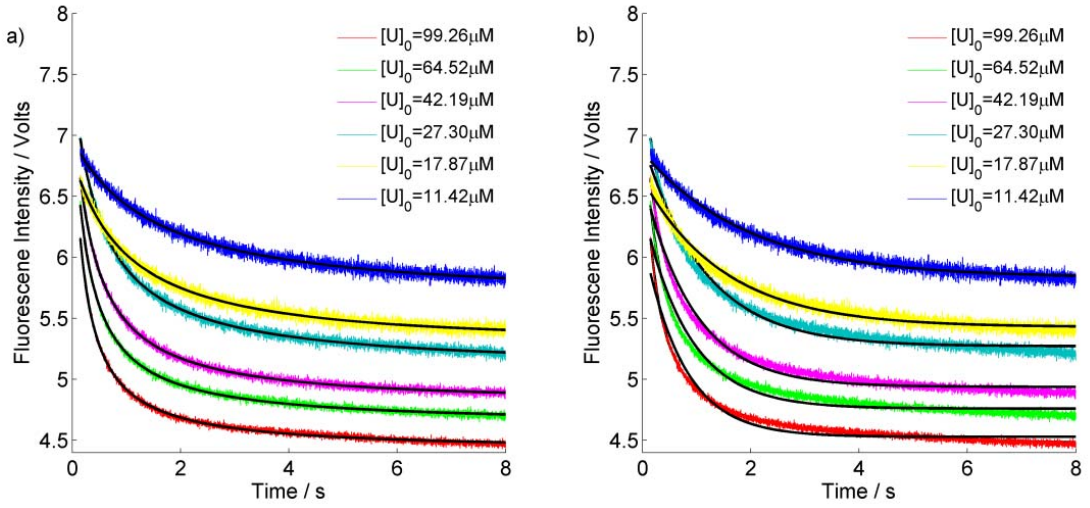


Figure S19. Global fitting of the stopped-flow Trp fluorescence data of DS119 with a a) 2nd reaction model ($y = At + B + \frac{C}{1+Dt}$); b) 1st reaction model ($y = At + B + Ce^{-kt}$).

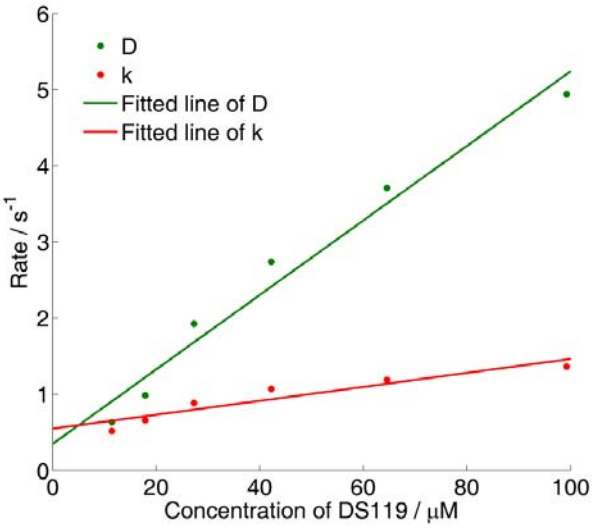


Figure S20. The relationship between the folding rate constants (s^{-1}) and the protein concentration (D: 2nd reaction rate constants; k: 1st reaction rate constants). The extrapolation to zero concentrations led to D (0 μM) = 0.35 s^{-1} and k (0 μM) = 0.55 s^{-1}.

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.

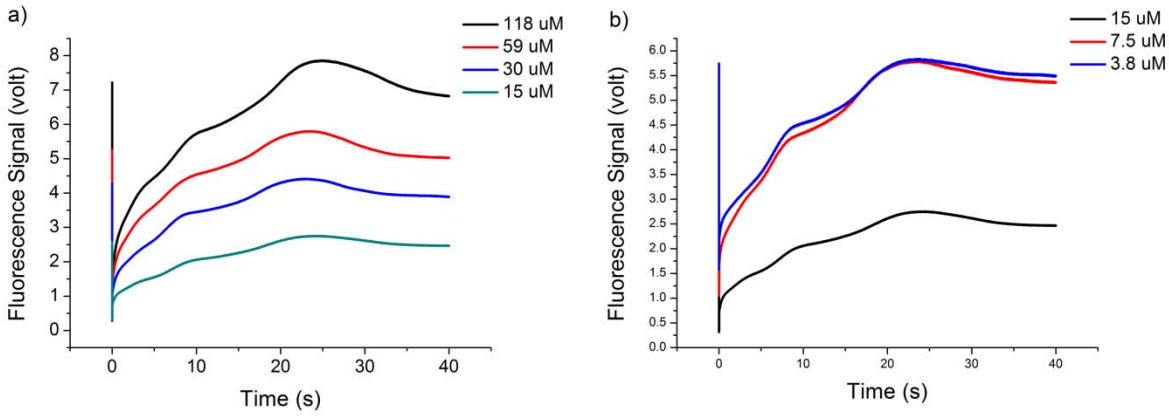


Figure S21. Concentration dependence of the unfolding process of DS119. The final concentrations of DS119 were a) 118 – 15 μM; b) 15 – 3.8 μM. The final concentration of Gdn-HCl was 5.1M.