Ultrafast folding kinetics and cooperativity of villin headpiece in single molecule autocorrelation force spectroscopy — Supporting Information

Gabriel Žoldák[a,1], Johannes Stigler[a], Benjamin Pelz[a], Hongbin Li[b] and Matthias Rief[a,c,1]

[a] *Physik Department E22, Technische Universität München, James-Franck-Str, 85748 Garching, Germany*

[b] *Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1*

[c] *Munich Center for Integrated Protein Science, 81377 München, Germany*

[1] Corresponding author

Contents

1 Protein sequences

In the following, the cysteines that are used for handle attachment are shown in bold, the sequence corresponding to villin headpiece is shown in italics.

HP35wt, Ubq-HP35wt-Ubq-H_6

MA**C**KMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYN IQKESTLHLVL-
RLRGGELGSSGG*LSDEDFKAVFGMTRSAFANLPLWKQQNLKKEKGLF*GGSSGGT MQIFVKTLTGKTITLEVEPS-
DTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKES TLHLVLRLRGGK**C**LE HHHHHH

HP35stab, Ubq-HP35stab-Ubq-H_6

MA**C**KMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYN IQKESTLHLVL-
RLRGGELGSSGG*LSDEDFKAVFGMTRSAFANLPLWKQQALMKEKGLF*GGSSGGT MQIFVKTLTGKTITLEVEPS-
DTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKES TLHLVLRLRGGK**C**LE HHHHHH

2 SI Methods

2.1 Fluorescence measurements

All fluorescence measurements were performed with a Jasco FP-8500 spectrofluorimeter. The sandwich protein contains a single tryptophan in the villin subdomain and tyrosines in ubiquitin which were used to determine protein domains stability. The fluorescence of W64 was used as a reporter group for HP35 unfolding after selective excitation of tryptophan at 296 nm (3 nm bandwidths) and monitoring an emission monitoring at 355 nm (3 nm bandwidths). Tyrosine fluorescence was used as a specific probe for unfolding of ubiquitin after excitation at 278 nm (3 nm bandwidths) and emission at 302 nm (3 nm). The measurements were performed in 0.1 M M sodium phosphate, 0.15 M NaCl, pH 7.2, 0.75 % glucose, 2 % glycerol at 303 K. The experimental data (HP35wt, HP35stab) were analyzed according to a two-state model by assuming a linear dependence of fluorescence emission on denaturant concentration. A nonlinear least-squares fit of the experimental data was used to obtain the Gibbs free energy of denaturation ΔG_{D} as a function of denaturant [1]. Guanidinium chloride concentrations were determined from the known refractive indices [2].

2.2 Circular dichroism measurements

Circular dichroism measurements were done on a Jasco J-815 (Japan) spectropolarimeter at 20 °C. The peptides were dissolved in 0.1 M sodium phosphate, 0.15 M NaCl to a final concentration of 2 mM. For spectral measurements, the peptides were diluted with 5 mM sodium phosphate to a final concentration of 55 μM. The following parameters were used for measurements in the far UV region (180-250 nm): 1 nm bandwidths, scanning speed of 20 nm/min, 1 nm resolution, 1 s response time, 0.1 cm path length cuvettes and 25 accumulations. Guanidinium chloride titrations were performed in 0.1 M sodium phosphate, 0.15 M NaCl, pH 7.2, 0.65 % glucose, 2 % glycerol at 293 K.

2.3 Experimental determination of distribution moments

Comparison between experimental distributions and theoretical values was based on the distributions' moment expansion. For an experimental trace of bead deflection values $x(t)$ at a sampling rate δ with N data points, the

standard deviation σ_x and skewness γ_x were calculated according to

$$\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}\left[x\left(\frac{i}{\delta}\right) - \langle x\rangle\right]^2} \tag{S1}$$

$$\gamma_x = \frac{1}{N}\sum_{i=0}^{N-1}\left[\frac{x\left(\frac{i}{\delta}\right) - \langle x\rangle}{\sigma_x}\right]^3 \tag{S2}$$

where $\langle x\rangle$ is the time average of the trajectory $x(t)$.

2.4 Elastic linker models

At low forces, when the protein is still folded, the elasticity of the construct is dominated by the behavior of the dsDNA handles. It was modeled using an extensible wormlike chain model [3]. The force as a function of extension is implicitly given by

$$F_{\mathrm{eWLC}}(\xi) = \frac{k_{\mathrm{B}}T}{p_{\mathrm{D}}}\left(\frac{1}{4\left(1-\frac{\xi}{L_{\mathrm{D}}}\right)^2} - \frac{1}{4} + \frac{\xi}{L_{\mathrm{D}}} - \frac{F_{\mathrm{eWLC}}}{K}\right), \tag{S3}$$

where L_{D} is the DNA's contour length, p_{D} its persistence length and K its stretch modulus. In our experiments, typical values were $L_{\mathrm{D}} \approx 360\,\mathrm{nm}$, $p_{\mathrm{D}} \approx 23\,\mathrm{nm}$ and $K \approx 500\,\mathrm{pN}$.

Contributions of the unfolded polypeptide that play a role at high forces were modeled by adding a wormlike chain (WLC) model [4] in series with the eWLC linkers. Here, the force is given by

$$F_{\mathrm{WLC}}(\xi) = \frac{k_{\mathrm{B}}T}{p_{\mathrm{p}}}\left(\frac{1}{4\left(1-\frac{\xi}{L_{\mathrm{p}}}\right)^2} - \frac{1}{4} + \frac{\xi}{L_{\mathrm{p}}}\right). \tag{S4}$$

L_{p} is the contour length of the fully unfolded protein and p_{p} its persistence length. In our experiments, we assumed $p_{\mathrm{p}} = 0.7\,\mathrm{nm}$.

Since the linker is at equilibrium with the unfolded polypeptide, the forces must be equal. The extension of the combined construct consisting of DNA linker and unfolded polypeptide is hence given by

$$\xi_{\mathrm{construct}}(F) = \xi_{\mathrm{eWLC}}(F) + \xi_{\mathrm{WLC}}(F), \tag{S5}$$

where ξ_{eWLC} and ξ_{WLC} are the inverse of eq. (S3) and eq. (S4), respectively. Consequently, the force-extension relation of the DNA-polypeptide construct is given by the inversion of eq. (S5).

2.5 Calculation of the equilibrium moments for the tethered protein under force

When the kinetics of folding/unfolding is slow, the filter effects that arise due to the damping of the bead motion and/or sampling, are negligible. The expected values for standard deviation and skewness can hence be approximated in equilibrium terms. In the following discussion, F denotes the average force that will be measured experimentally, when the trap centers are a certain distance d apart. It is assumed in the calculatios that the bead size is zero.

We assume that the protein can unfold in N individual substeps and the difference in free energy between the folded

and unfolded states is ΔG_o[1]. The system's Hamiltonian is

$$\mathcal{H}(x_1, x_2, i) = \left(1 - \frac{i}{N}\right) \cdot \Delta G_o + \frac{1}{2} k_1 x_1^2 + \frac{1}{2} k_2 x_2^2 + \int_o^{d - x_1 - x_2} F_{\text{construct}}\left(L_p = \frac{i}{N} \cdot L_{p,\text{tot}}, \xi\right) d\xi. \tag{S6}$$

Here, $F_{\text{construct}}$ is the inverse of eq. (S5).

In experiments, we measured the sum of the absolute displacements of the beads from their center positions $x = x_1 + x_2$. The distribution of x can be calculated based on eq. (S6) by integrating out all other degrees of freedom:

$$p(x_1 + x_2) = \frac{\sum_{i=0}^{N} \exp\left(-\frac{\mathcal{H}(x_1, x_2, i)}{k_B T}\right)}{\int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \sum_{i=0}^{N} \exp\left(-\frac{\mathcal{H}(x_1, x_2, i)}{k_B T}\right)}. \tag{S7}$$

While in principle the entire calculated distribution $p(x) = p(x_1 + x_2)$ can be directly compared to experimental data, we focused on the first three normalized central moments. The n-th moment of the total bead deflection $x = x_1 + x_2$ is given by

$$\mu_n = \langle x^n \rangle = \left\langle (x_1 + x_2)^n \right\rangle = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \, (x_1 + x_2)^n \cdot p(x_1 + x_2). \tag{S8}$$

Note that the μ_n only depend on ΔG_o, d and the mechanical parameters of the linker. A force-distance curve can hence be constructed based on eq. (S8) via

$$F(d) = \langle x \rangle \cdot k_c, \tag{S9}$$

where $k_c = \left(\frac{1}{k_1} + \frac{1}{k_2}\right)^{-1}$ is the combined stiffness of the two traps.

Higher central moments, such as the standard deviation σ and the skewness γ were calculated from eq. (S8):

$$\sigma = \sqrt{\mu_2 - \mu_1^2} \tag{S10}$$

$$\gamma = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{\sigma^3}. \tag{S11}$$

Note that since $\xi_{\text{construct}} = d - x_1 - x_2$, the moments can be equally interpreted as moments of the *extension*. In this case, $\sigma_{\text{extension}} = \sigma$ and $\gamma_{\text{extension}} = -\gamma$.

2.5.1 Force-distance curves under equilibrium conditions

We can use the first moment definition of eq. (S8) to construct fit functions for force-distance curves that were recorded under quasi-equilibrium conditions as shown in Fig. 1 D, F (see eq. (S9)). In the case of a two-state model

[1] Note that $\Delta G_o < 0$ for stably folded proteins.

we arrive at

$$F(d) = \frac{k_c}{\int_{-\infty}^{\infty} dx \left[\exp\left(\frac{\Delta G_o + \frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p=o,\xi\right)d\xi}{k_B T}\right) + \exp\left(\frac{\frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p=L_{p,tot},\xi\right)d\xi}{k_B T}\right)\right]}$$
$$\cdot \int_{-\infty}^{\infty} dx\, x \left[\exp\left(\frac{\Delta G_o + \frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = o,\xi\right)d\xi}{k_B T}\right) + \exp\left(\frac{\frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = L_{p,tot},\xi\right)d\xi}{k_B T}\right)\right]. \tag{S12}$$

In the case of three states, where we assumed that the intermediate state has half the energy of the folded state and half the contour length, we arrive at

$$F(d) = k_c \left\{\int_{-\infty}^{\infty} dx \left[\exp\left(\frac{\Delta G_o + \frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = o,\xi\right)d\xi}{k_B T}\right)\right.\right.$$
$$\left.\left. + \exp\left(\frac{\frac{1}{2}\Delta G_o + \frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = \frac{1}{2}L_{p,tot},\xi\right)d\xi}{k_B T}\right) + \exp\left(\frac{\frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = L_{p,tot},\xi\right)d\xi}{k_B T}\right)\right]\right\}^{-1}$$
$$\cdot \int_{-\infty}^{\infty} dx\, x \left[\exp\left(\frac{\Delta G_o + \frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = o,\xi\right)d\xi}{k_B T}\right)\right.$$
$$\left. + \exp\left(\frac{\frac{1}{2}\Delta G_o + \frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = \frac{1}{2}L_{p,tot},\xi\right)d\xi}{k_B T}\right) + \exp\left(\frac{\frac{1}{2}k_c x^2 + \int_o^{d-x} F_{\text{construct}}\left(L_p = L_{p,tot},\xi\right)d\xi}{k_B T}\right)\right]. \tag{S13}$$

Here, $k_c = \left(\frac{1}{k_1} + \frac{1}{k2}\right)^{-1}$ is the combined trap stiffness of both traps. $F_{\text{construct}}$ is given by the inverse of eq. (S5). The model parameters necessary to construct the fits shown in Fig. 1 D, F are listed in Table S1.

2.6 Autocorrelation functions

A discrete unnormalized autocorrelation $x^*(\tau)$ at time lag τ of a trace $x(t)$, that was sampled at a sampling rate δ and has N data points, was calculated numerically using

$$x^*\left(\tau = \frac{j}{\delta}\right) = \frac{1}{N}\sum_{i=o}^{N-1}\left(x_i - \mu_x\right)\left(x_{i-j} - \mu_x\right), \tag{S14}$$

where μ_x is the mean of the trace $x\,(t)$. Non-overlapping regions of the expression above were zero-padded. For typical trace lengths in the order of seconds, this zero padding did not affect the typically investigated autocorrelation times.

The autocorrelation functions of the *DNA linkers* could to good approximation be fitted with single exponential functions. We assumed that the kinetics of the fluctuations of the DNA linkers were independent of the fluctuations of the protein tethered to the linkers. To good approximation, the autocorrelation functions of the *tethered protein constructs* could therefore be fitted with double exponential functions of the form

$$f\,(\tau) = A_o \exp\left(-\lambda_o \tau\right) + A_1 \exp\left(-\lambda_1 \tau\right). \tag{S15}$$

When the rate constants λ_o and λ_1 lay close within a factor of $\lessapprox 5$, the unconstrained fitting of autocorrelation functions with eq. (S15) was not robust. We therefore chose to fix the amplitudes A_o and the rates λ_o to values obtained separately in experiments with DNA linkers only. Nevertheless, even when we applied a single exponential model to data of tethered proteins, the obtained rate constants were in good agreement with the data obtained from the more elaborate double exponential model.

Some of our experiments showed a very slow phase in the autocorrelation function. Following a recent suggestion, we attribute the effect to a coupling of the z-(axial-)direction of the instrument into our measurement coordinate [5]. The amplitude of this phase was always less than ≈ 10 % of the full signal and could be accounted for by adding a third exponential decay into the fitting function. Since the time constant of this phase ($\lessapprox 10^4\ \mathrm{s}^{-1}$) was always well separated from the other phases, the identification of the contribution due to z-coupling was unambiguous. This third slow phase was observed in experimental autocorrelation functions of the DNA dimer as well as in HP35stab and HP35wt.

2.7 Model for the kinetics of folding and unfolding under force

The force dependence of the rate constants for *folding* were modeled according to earlier work that accounts for all the energies that are involved during the transition [6, 7]. The rate constants for the corresponding *unfolding* transition were calculated based on the principle of detailed balance as follows:

The folding transition rate constant is given by

$$k_{\mathrm{fold}}(F) = k_{\mathrm{o,fold}} \cdot \exp\left(\frac{\hat{G}_{\mathrm{U}}(F) - \hat{G}_{\mathrm{TS}}\left(F_{\mathrm{U}\to\mathrm{TS}}(F)\right)}{k_{\mathrm{B}}T}\right). \tag{S16}$$

Here, $\hat{G}_{\mathrm{U}}(F)$ is the free energy that is required to stretch all mechanical parts (Hookean displacement of the beads from their positions, stretching of the DNA linker and stretching of the unfolded polypeptide) when the protein is in its unfolded state at force F. Similarly, $\hat{G}_{\mathrm{TS}}$ is the free energy needed to stretch the all mechanical parts when the protein is at the transition state. Generally, when the protein is in a state i that is characterized by the contour length of the unfolded polypeptide L_i, the free energy involved in the stretching mechanics is given by

$$\hat{G}_i(F) = \frac{1}{2}\frac{F^2}{k_c} + \int_{\mathrm{o}}^{\xi_{\mathrm{construct}}(F)} F_{\mathrm{construct}}\left(L_{\mathrm{p}} = L_i, \xi\right) d\xi. \tag{S17}$$

Since the experiments were conducted at constant trap separation rather than constant force conditions, the force changes during the transition. The function $F_{\mathrm{U}\to\mathrm{TS}}(F)$ returns the force that acts when the system is at the transition state, given that is at force F in the unfolded state. It can be readily calculated using the linker models discussed above (eq. (S5)).

The principle of detailed balance requires that

$$\ln\left(\frac{k_{\mathrm{fold}}(F)}{k_{\mathrm{unfold}}\left(F_{\mathrm{U}\to\mathrm{N}}(F)\right)}\right) = -\frac{\hat{G}_{\mathrm{N}}\left(F_{\mathrm{U}\to\mathrm{N}}(F)\right) - \hat{G}_{\mathrm{U}}(F) + \Delta G_{\mathrm{o}}}{k_{\mathrm{B}}T}. \tag{S18}$$

Consequently, the unfolding transition rate constants are given by

$$k_{\text{unfold}}(F) = \frac{k_{\text{o,fold}}}{\exp\left(\frac{-\Delta G_{\text{o}}}{k_{\text{B}}T}\right)} \cdot \exp\left(\frac{\hat{G}_{\text{N}}(F) - \hat{G}_{\text{TS}}\left(F_{\text{N}\to\text{TS}}(F)\right)}{k_{\text{B}}T}\right). \tag{S19}$$

The rate constants calculated by eq. (S16) and eq. (S19) automatically fullfil the principle of detailed balance. The determining parameters for k_{unfold} and k_{fold} are, in addition to the mechanical properties of the DNA linkers and the polypeptide, the zero-force folding rate constant $k_{\text{o,fold}}$, the difference in energy between the folded and unfolded state ΔG_{o} and the contour length of the unfolded polypeptide in the transition state.

In our data analysis we used this model to obtain kinetic and thermodynamic parameters from the force-dependent autocorrelation decay functions (Chevron plot, see Fig. 3A). At a measured average force F the expected measured autocorrelation rate constant from the protein is given by

$$\lambda(F) = k_{\text{fold}}(F) + k_{\text{unfold}}(F), \tag{S20}$$

where $k_{\text{fold}}(F)$ and $k_{\text{unfold}}(F)$ are given by eq. (S16) and eq. (S19), respectively. The parameters determining the mechanical properties of the DNA linker and the polypeptide were determined independently from force-extension traces for each molecule. The symmetry of the Chevron plots suggested a transition state roughly in the middle of the energy landscape. We then adapted the parameters ΔG_{o}, $\Delta L_{\text{U-TS}}$ and $k_{\text{o,fold}}$ to obtain maximal overlap between the model (continuous lines in Fig. 3A) and the measured values (closed circles in Fig. 3A). Results are listed in Table S2.

2.8 Brownian dynamics simulations

The objective of our Brownian dynamics simulations was to estimate the dynamic response in the measured bead deflection signal we are able to pick up when the protein undergoes rapid folding/unfolding events.

We based our simulations on a two (three) state model with the kinetic scheme N $\rightleftharpoons$ U (N $\rightleftharpoons$ I $\rightleftharpoons$ U). The dynamic variables of the model are the current folding state of the protein $i \in \{\text{N}, \text{U}\}$ ($i \in \{\text{N}, \text{I}, \text{U}\}$) and the deflections of the two beads from the trap centers x_1 and x_2.

Each protein state i is defined by its (contour-)length of unfolded polypeptide L_i and its thermodynamic stability $\Delta G_{i\text{-U}}$. Further inputs to the model are the zero-force folding rates $k_{\text{o,fold}} = k_{\text{o,U}\to\text{N}}$ (three state: $k_{\text{o,U}\to\text{I}}$ and $k_{\text{o,I}\to\text{N}}$) and the position of the transition state along the contour length coordinate, yielding the force-dependence of the transition rate constants (i.e. the Chevron plot, also see section 2.7).

For a simulation at trap distance[2] d, we made use of

$$d = \xi_{\text{construct}}\left(F, L_{\text{p}} = L_i\right) + \frac{F}{k_1} + \frac{F}{k_2} \tag{S21}$$

to determine the equilibrium forces acting when the protein is in state i. The linker extension $\xi_{\text{construct}}\left(F, L_{\text{p}}\right)$ is given by eq. (S5). We then used these equilibrium forces to determine transition rate constants between the individual states based on a model fulfilling detailed balance (eq. (S16), eq. (S19)). These fixed transition rate constants were used for a Markov chain generator that determined the folding state i at each time step (see below).

[2] Note that d is defined such that when the bead surfaces touch, $d = 0$.

The dynamics of the system was then simulated by numerically solving the discretized Langevin equations for each of the beads, which were considered the slowest mechanical components in the system. Relaxation times of linkers were considered negligible [8]. Since the bead motion is overdamped, the inertial terms were neglected. At each simulation time step Δt, the one-dimensional positions $x_{1,2}$ of the beads were updated according to

$$x_i\left(t+\Delta t\right) = x_i(t) + \frac{\Delta t}{\gamma_S}\left(-k_i x_i(t) + F_{\text{construct}}(t) + \sqrt{\frac{2k_B T\gamma_S}{\Delta t}}\varepsilon(t)\right) \tag{S22}$$

with the bead's Stokes drag coefficient $\gamma_S = 6\pi\eta R$, the trap spring constant k_i and $\varepsilon(t)$ being a random number from a generator that produces an uncorrelated series with standard deviation $\sigma = 1$. The sign convention is such that high values of x_i mean closer beads. η is the buffer's viscosity and was set to 1 cP = $10^{-9}\,\frac{\text{pN s}}{\text{nm}^2}$. The bead radius R was 500 nm. Additional drag of the solvent that couples to the DNA is small compared to the drag on the beads [9] and was found not to significantly affect the results.

The external force on the beads $F_{\text{construct}}(t)$ is conveyed by the dsDNA linker (depending on the folding state of the protein, possibly in series with an unfolded polypeptide) and was determined at each time step. To this end, we first calculated the current extension of the linker $\xi(t)$ using $d = \xi(t) + x_1(t) + x_2(t)$ and used the inverse of eq. (S5) to calculate the force acting on the linker from ξ. The parameter L_p, which describes the contour length of unfolded polypeptide, was set based on the current folding state i.

The dynamic variable i, which designates the current folding state of the protein, was updated at each time step based on a Markov process with transition rate constants determined as described above. In this model, a folding/unfolding event instantaneously changes the length of the unfolded polypeptide. The simulation result is the response in bead deflection we are able to measure when the protein undergoes rapid folding/unfolding transitions.

Typically, trajectories of the bead motions were simulated at a rate of 9 MHz. Mimicking experimental procedure, the data were then 8-pole Butterworth-filtered at 150 kHz and subsequently subsampled at 300 kHz. The difference signal [10] was then calculated as $x(t) = x_1(t) + x_2(t)$.

3 SI Text

3.1 Protein stability from chemical unfolding

In addition to force spectroscopy measurements, the stability of the constructs HP35wt and HP35stab were also determined by chemical denaturation with GdmCl. The data obtained are summarized in Fig. S1 (A), (B) and Table S2.

3.2 Truncation mutants

We found that the equilibrium unfolding of HP35stab (K70M/N68A HP35) exhibits lower folding cooperativity than HP35wt. While HP35wt shows all-or-none cooperativity, the folding transition of HP35stab is well described by a three-state transition involving a partially folded intermediate. This suggests that the two substitutions in the C terminal helix decrease the folding cooperativity by stabilizing the C terminal substructure. To be able to observe such an intermediate in force spectroscopy experiments, it necessarily must also exhibit measurable stability in solution.

To test this hypothesis, we analysed variants of HP35wt and HP35stab where the N-terminal helix was removed (ΔN-HP35wt – MTRSAFANLPLWKQQNLKKEKGLF), (ΔN-HP35stab – MTRSAFANLPLWKQQALMKEKGLF). For comparison, we also investigated a mutant where the C-terminal helix was removed (ΔC-HP35 – LSDEDFKAVFGMTRSAFANLPLWKQQ), see Fig. S1 (D). The peptides were synthesized and purified by Genscript (>98 % purity).

Far-UV CD spectroscopy confirmed that the mutant ΔN-HP35stab still contains considerable helicity, while the mutant ΔN-HP35wt does not (Fig. S1 (C)). These findings suggest a possible solution to the question why the apparent folding cooperativity of HP35wt is all-or-none, while this model fails for HP35stab.

3.3 Verification of experimental procedures by simulation

For validation of the analysis procedure, we simulated trajectories of a tethered protein with kinetics similar to our experimental results (see SI Methods: Brownian dynamics simulations). We then analyzed these simulated traces using the same methods that we applied to measured traces.

As expected, standard deviation and skewness values of the simulated traces closely resembled those obtained for experimental trajectories (Fig. 2C, D, F, G).

Autocorrelation functions of the simulated trajectories were single-exponential for the DNA-only construct and double-exponential in the case of a tethered protein. We then asked whether we could extract the kinetic information we used as input for the simulations from these autocorrelation functions. By fixing the fast component of the protein double exponential to values we obtained from simulating DNA only, we could robustly obtain information about the protein kinetics from the slow component.

Results of a series of simulations where we used our experimental kinetic results as input parameters (continuous lines) and the obtained kinetic information from autocorrelation analysis (intermediate shaded squares) are shown in Fig. S4. The kinetics obtained from autocorrelation analysis of the simulated traces are in good agreement with their expected values, validating our autocorrelation procedures.

We then asked whether we could distinguish small differences in kinetics using autocorrelation analysis. To this end, we repeated the simulations with faster and slower folding/unfolding rates. Fig. S4 shows the results of simulations with both $10^{0.5} \approx 3.16$ times faster (light colors) and slower (dark colors) kinetics. Again, the results from the simulations agreed well with the expected values. In addition, the kinetic simulation results for both faster and slower kinetics are clearly different from experimental data (panels (A),(C)). We conclude that we can discern differences in folding/unfolding rates of up to a factor of three using autocorrelation analysis.

3.4 Cooperativity of folding/unfolding

The data for HP35stab showed some deviations from a model where the folding/unfolding is fully cooperative. The peak force-dependent distribution standard deviations deviated from what would be theoretically expected from a pure two-state folder under equilibrium conditions.

Experimental data indicated that the stabilized C-terminal part of HP35stab is at least partly structured (Fig. S1 (C)).

To verify the significance of these deviations and to test whether the deviations are possibly due to a populated intermediate, we performed Brownian dynamics simulations of the measurement of HP35stab for the two competing scenarios. In the case of full cooperativity, the kinetic scheme was assumed $N \rightleftharpoons U$. populated intermediate, the

kinetic scheme was N $\rightleftharpoons$ I $\rightleftharpoons$ U. In all cases, each of the transitions N $\rightleftharpoons$ U, N $\rightleftharpoons$ I and I $\rightleftharpoons$ U obeyed detailed balance (eq. (S18)).

3.5 Measurement bandwidth and filtering

The effect of sampling frequency on the obtained distribution moments can be verified using Brownian dynamics simulations and is most pronounced in the standard deviations. Fig. S6 illustrates the importance of high sampling rates when the exact values of the standard deviations are of interest. The exact solution that would be obtained at infinite sampling bandwidth can be calculated using equilibrium theory (eq. (S8)). Nevertheless, even at high sampling frequencies of 9 MHz, deviations from this value are apparent (Fig. S6). Note that sampling at 9 MHz is far above the characteristic frequency of the construct ($f_{3\mathrm{dB}}(17\,\mathrm{pN}) \approx 32\,\mathrm{kHz}$). The effect is even more pronounced at our experimental conditions where the data was filtered at 150 kHz and sampled at 300 kHz. Here, the discrepancy between the standard deviation at infinite bandwidth and at 300 kHz can be as big as ≈ 8 %.

References

[1] M M Santoro and D W Bolen. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, 27(21):8063–8068, October 1988.

[2] C N Pace. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in enzymology*, 131:266–280, 1986.

[3] M D Wang, H Yin, R Landick, J Gelles, and Steven M Block. Stretching DNA with optical tweezers. *Biophysical Journal*, 72(3):1335–1346, March 1997.

[4] Carlos Bustamante, J F Marko, E D Siggia, and Steven B Smith. Entropic elasticity of lambda-phage DNA. *Science*, 265(5178):1599–1600, September 1994.

[5] M Ribezzi-Crivellari and Felix Ritort. Force Spectroscopy with Dual-Trap Optical Tweezers: Molecular Stiffness Measurements and Coupled Fluctuations Analysis. *Biophysical Journal*, 103(9):1919–1928, November 2012.

[6] J Christof M Gebhardt, Thomas Bornschlögl, and Matthias Rief. Full distance-resolved folding energy landscape of one single protein molecule. *Proceedings of the National Academy of Sciences of the United States of America*, 107(5):2013–2018, February 2010.

[7] Michael Schlierf, Felix Berkemeier, and Matthias Rief. Direct observation of active protein folding using lock-in force spectroscopy. *Biophysical Journal*, 93(11):3989–3998, 2007.

[8] M Manosas, Jin-Der Wen, Pan T X Li, Steven B Smith, Carlos Bustamante, Ignacio Tinoco, and Felix Ritort. Force unfolding kinetics of RNA using optical tweezers. II. Modeling experiments. *Biophysical Journal*, 92(9):3010–3021, May 2007.

[9] J Howard. *Mechanics of Motor Proteins and the Cytoskeleton*. Sinauer Associates, Publishers, 2001.

[10] Jeffrey R Moffitt, Yann R Chemla, David Izhaky, and Carlos Bustamante. Differential detection of dual traps improves the spatial resolution of optical tweezers. *Proceedings of the National Academy of Sciences of the United States of America*, 103(24):9006–9011, June 2006.

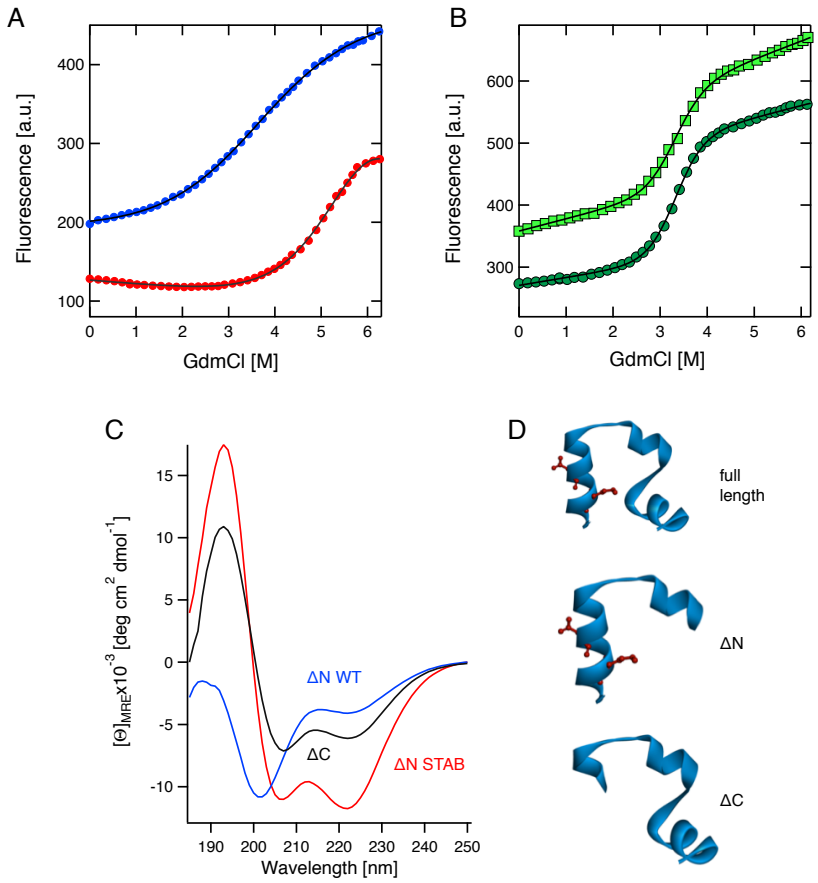

Figure S1: (A), (B) Protein stability from GdmCl-induced unfolding.
(A) GdmCl-induced unfolding transition of 2 μM Ubq-HP35wt-Ubq (blue circles) and Ubq-HP35stab-Ubq (red circles) at 30 °C measured by protein fluorescence at 355 nm after excitation at 295 nm in 0.1 M Na-phosphate, 0.65 % glucose, 0.15 M NaCl, 2 % glycerol at pH 7.2. The two-state analysis [1] (continuous lines) gave values for HP35wt: $\Delta G_D(30\,^\circ\mathrm{C}) = 10.1\pm0.2\,\mathrm{kJ\,mol^{-1}}$ ($\approx -4.1\,k_BT$), $m = 3.1\pm0.1\,\mathrm{kJ\,mol^{-1}\,M^{-1}}$ and $[\mathrm{GdmCl}]_M = 3.26\,\mathrm{M}$ and for HP35stab: $\Delta G_D(30\,^\circ\mathrm{C}) = 22.0 \pm 2.5\,\mathrm{kJ\,mol^{-1}}$ ($\approx -8.9\,k_BT$), $m = 3.8 \pm 0.1\,\mathrm{kJ\,mol^{-1}\,M^{-1}}$ and $[\mathrm{GdmCl}]_M = 5.8\,\mathrm{M}$. Data are summarized in table S2.
(B) GdmCl-induced unfolding transition of 2 μM Ubq-HP35wt-Ubq (light green squares) and Ubq-HP35stab-Ubq (dark green circles) at 30 °C measured by protein fluorescence at 302 nm after excitation at 278 nm in 0.1 M Na-phosphate, 0.65 % glucose, 0.15 M NaCl, 2 % glycerol at pH 7.2. The two-state analysis (continuous lines) gave values for Ubq-HP35wt-Ubq: $\Delta G_D(30\,^\circ\mathrm{C}) = 26 \pm 1\,\mathrm{kJ\,mol^{-1}}$, $m = 7.8 \pm 0.3\,\mathrm{kJ\,mol^{-1}\,M^{-1}}$ and $[\mathrm{GdmCl}]_M = 3.33\,\mathrm{M}$ and for Ubq-HP35stab-Ubq: $\Delta G_D(30\,^\circ\mathrm{C}) = 27 \pm 1\,\mathrm{kJ\,mol^{-1}}$, $m = 8.2 \pm 0.2\,\mathrm{kJ\,mol^{-1}\,M^{-1}}$ and $[\mathrm{GdmCl}]_M = 3.29\,\mathrm{M}$.
(C) Far-UV region CD spectra of 55 μM ΔN-HP35wt (blue), ΔN-HP35stab (red), ΔC-HP35wt (black) in 2 mM sodium phosphate (pH 7.5) at 20 °C.
(D) Representation of the mutants in (C), based on the crystal structure 1YU5.

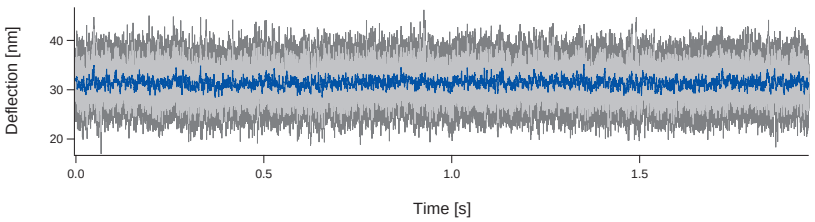

Figure S2: Shown is a time trace of deflection at 8 pN for HP35wt at 300 kHz (dark grey), 20 kHz (light grey) and at 1 kHz (blue).

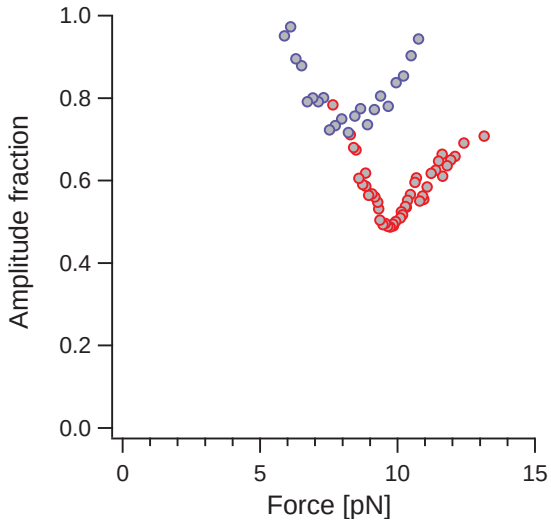

Figure S3: Relative amplitudes of the *fast* component in the autocorrelation signal for HP35wt (blue circles) and HP35stab (red circles).

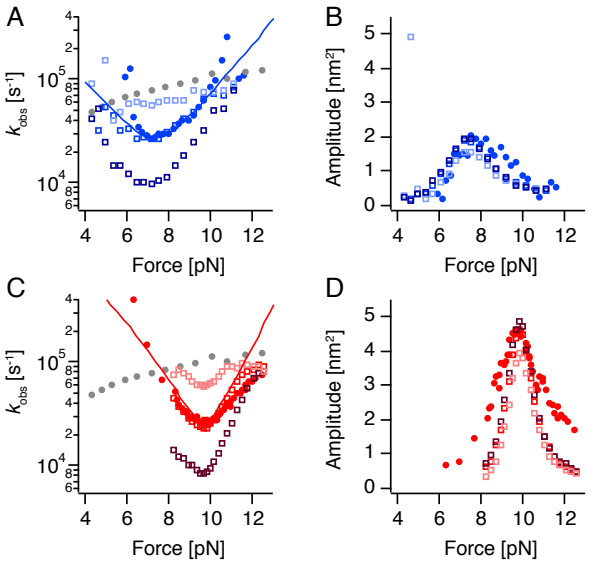

Figure S4: Simulations for HP35wt and HP35stab at faster and slower kinetics. (A), (B) data for HP35wt, (C), (D) data for HP35stab. Data for the DNA only construct are shown in grey. Continuous lines are the best fits to the experimental data (filled circles). Empty squares are results from brownian dynamics simulations. Plain red and blue correspond to simulations of the best fit model. Light colors are results from a simulation with a folding rate that is $10^{0.5} \approx 3.16$ times faster than the best fit. Dark colored squares are results from a simulation with a folding rate that is $10^{0.5} \approx 3.16$ times slower.

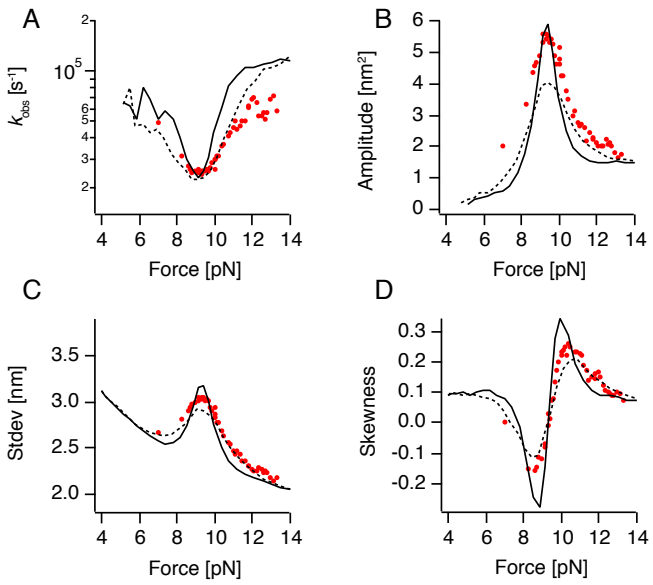

Figure S5: Brownian dynamics simulations for HP35stab with full cooperativity of folding (two-state model, continuous lines) and an assumed intermediate (three-state model, dashed lines). Red dots are experimental values. The intermediate was assumed to have a free energy of $-\Delta G_o/2$ and a length of $L/2$, where $-\Delta G_o$ is the free energy of the native state and L the contour length difference between native and unfolded (see Table S2). In the case of full cooperativity, the zero-force folding rate k_o was $10^{6.0}$ s^{-1} for this molecule. In the case of a populated intermediate, we assumed $k_{o,U\to I} = k_{o,I\to N} = 10^{5.2}$ s^{-1}.

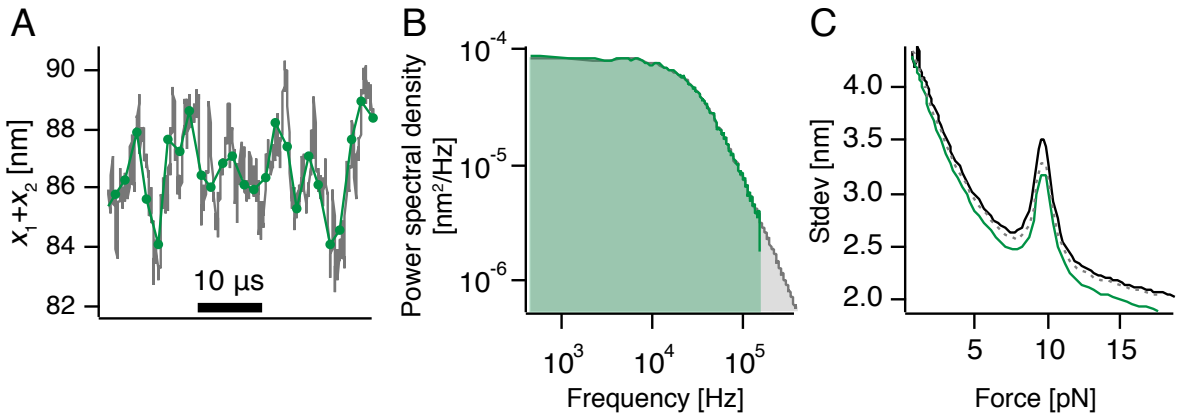

Figure S6: The effect of filtering on the standard deviation of the data. (A) Simulated trace of a HP35stab tethered protein at 17 pN. Mimicking experimental procedures, a simulated trace at a sampling rate of 9 MHz (grey) was filtered at 150 kHz and subsampled at 300 kHz (green). The corresponding power spectra are shown in (B). (C) Corresponding standard deviations at various forces. The black line is the calculated standard deviation (eq.(S8)) for HP35stab from equilibrium theory, i.e. at infinite bandwidth. A kinetic simulation (see (A)), that was sampled at the experimental bandwidth of 300 kHz (green continuous line) shows a significant offset. Kinetic simulations at higher bandwidths (represented by the grey dotted line, 9 MHz) approach the equilibrium limit.

	HP35wt	**HP35stab**
p_D (nm)	24 ± 4	
T (K)	303.15	
L_D (nm)	360 ± 10	
K (pN)	450	
p_p (nm)	0.7	
$L_{p,tot}$ (nm)	11 ± 1	
$-\Delta G_o$ ($k_B T$)	4.9 ± 1.0	9.8 ± 1.0
$-\Delta L_{N\text{-}I}$ (nm)	–	5.5 ± 1.0
$-\Delta L_{I\text{-}U}$ (nm)	–	5.5 ± 1.0
$-\Delta G_{N\text{-}I}$ ($k_B T$)	–	4.9 ± 1.0
$-\Delta G_{I\text{-}U}$ ($k_B T$)	–	4.9 ± 1.0

Table S1: Parameters of the equilibrium folding/unfolding model (Figure 1 D, F) for a two-state model (HP35wt) and for a three-state model (HP35stab).

	HP35wt	HP35stab
Chemical unfolding		
ΔG_D at 30 ° C (kJ mol^{-1})	−10.1 ± 1.1	−22.0 ± 2.5
ΔG_D at 30 ° C ($k_B T$)	−4.1 ± 0.5	−8.9 ± 1.0
m (kJ mol^{-1} M^{-1})	3.1 ± 0.1	3.8 ± 0.1
$[\mathrm{GdmCl}]_M$ (M)	3.26 ± 3.1	5.8 ± 0.2
Force unfolding		
ΔG_o at 30 ° C (kJ mol^{-1})	−12.1 ± 2.5	−24.3 ± 2.5
ΔG_o at 30 ° C ($k_B T$)	−4.9 ± 1.0	−9.8 ± 1.0
L_{tot} (nm)	10.6 ± 0.8	10.4 ± 1.0
$F_{1/2}$ (pN)	8.3 ± 0.6	10.3 ± 0.7
$\Delta L_{U\text{-}TS}$ (nm)	5.8 ± 1.0	6.3 ± 1.0
$\Delta L_{N\text{-}TS}$ (nm)	4.8 ± 1.0	4.1 ± 1.0
$\log_{10} k_{o,fold}$	5.4 ± 0.1	6.1 ± 0.1
$\log_{10} k_{o,unfold}$	3.3 ± 0.5	1.8 ± 0.5

Table S2: Stabilities and kinetic model parameters for the HP35stab and HP35wt variants. The kinetic values for HP35stab assume a two-state mechanism (eq. (S20)).