Supporting Information

Berkovich et al. 10.1073/pnas.1212167109

SI Text

Protein Engineering and Purification. The polyubiquitin Ubi_{12} construct was engineered as previously described (1). The protein was expressed in BLR *Escherichia coli* cells (Novagen). The cells were grown to OD^{600} of 0.6 at 37 °C and the protein expression was induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) for 3 h at 37 °C. The protein was purified using a Cobalt affinity Talon resin (GE Biosciences) and Fast Fluid Liquid Chromatography (FPLC–Superdex 200 HR column). We used standard phosphate buffered saline for both purification and measurements. All the regents used had a molecular biology grading.

Experimental Setup. The measurements were obtained using an atomic force microscope (AFM) customized for high-speed single molecule force spectroscopy. The present AFM configuration distinguishes itself from commonly used AFM spectrometers through the use of smaller cantilevers, faster piezoelectric actuators, a horizontal configuration, and a double-pulse preemptive control system customized for the intended experimental protocol. In order to decouple the force measurements from the main oscillatory mode of the supporting optical table the sample substrate was mounted vertically i.e., the substrate normal and the force measurements were all in a direction parallel to the optical table's surface and therefore perpendicular to the observed dominant perturbations in the vertical axis.

The substrate was attached to a specially designed stage composed of two piezoelectric actuators mounted in series. A slow piezoelectric actuator (PicoCube, Physik Instrumente) with a travel of 6 μm was used for cantilever calibration, to approach the surface to the cantilever and for lateral movement. A second piezoelectric actuator mounted between the first one and the surface (PL055.30 multi-layer piezo stack, Physik Instrumente) with a travel of 2 μm and a nominal resonant frequency higher than 300 kHz was used to obtain the constant force condition with a high response time. In the force clamp mode the force is controlled by a feedback loop. The measured force signal is fed into a well-tuned analog Proportional–Integral–Differential (PID) control system, which drives the fast piezoelectric actuator. Constant force is applied to a polyprotein chain by adjusting the position of the surface through the movement of the piezoelectric actuator. The PID insures a constant tension on the protein by monitoring the deflection of the cantilever.

In order to obtain several recoil traces from each stretched poly-protein we created a multiquench protocol consisting in a fast ramp of force from −100 pN to 180 pN in 20 ms to maximize pick-up rate, followed by 0.2 s of constant force to allow all modules to unfold and finishing with a square wave force pattern lasting up to 1 s. The periodic stretching and quenching protocol is designed to keep the protein in its purely elastic regime, avoiding any enthalpic interactions that occur upon quenching the force to a much lower final value. Contrary to the unfolding events that are stochastic in nature, in this multiquench protocol the stretching and recoil times are known in advance. We take advantage of this knowledge to optimize the control response by driving the actuator closer to the new equilibrium position by a distance equal to $k/\Delta F$ where k is the cantilever spring constant and ΔF is the magnitude of the stepwise change in force. This pre-emptive system leaves the PID control with a smaller task of driving only the change in length of the protein.

The cantilevers were calibrated using the equipartition theorem, and the contact slope was obtained by moving of the piezo-cantilever ensemble. The physical properties of the cantilever used influence the response time and force sensitivity. The spring constant influences the signal to noise ratio of the measurement, while the resonance frequency limits the measurement bandwidth. Furthermore, the surface area of the cantilever perpendicular to the direction of motion induces opposing viscous drag forces. We analyzed three types of cantilevers fitted for the purpose of the experiment: Bruker MLCT (spring constant approximately 20 pN/nm, resonance frequency in water approximately 1 kHz), Olympus Biolever BL-RC150VB (spring constant approximately 40 pN/nm, resonance frequency approximately 9 kHz) and Olympus Biolever BL-AC40TS (spring constant approximately 100 pN/nm, resonance frequency approximately 25 kHz). Fig. S1 shows the drag response of three types of cantilevers as a function of moving velocity close to a surface. In this experiment the force experienced by cantilevers at different approach and redraw velocities was measured. As shown in Fig. S1*A*, the effect of the drag is symmetric around the zero force position (part 2 and 4 vs. part 1 of the trace). In part 3 the cantilever is in contact with the surface. The piezo's position protocol shown in lower box of Fig. S1*A* was applied to measure the deflection force curves (upper box of Fig. S1 *A* and *C*), which were used to evaluate the drag coefficients. Fig. S1*D* shows the way in which the different cantilevers respond to the drag forces with respect to the velocity of the actuator. For our experiments we chose the Olympus Biolever BL-RC150VB, which exhibits good balance between the drag force and the force signal resolution.

In a typical experiment, the cantilever is pushed to the surface with a force of approximately 1 nN for 1 s, it is then retraced to a deflection corresponding to a force of 180 pN. At this force ubiquitin domains unfold and the end-to-end extension increases in steps of 20 nm, corresponding to the elastic extension of the amino acids trapped behind the mechanical transition state corresponding to the unfolding of the protein. Once unfolded, the polypeptide chain is cycled between 250 and 100 pN. After the polyprotein detaches the PID cannot maintain the set-point force anymore and drives the surface away from the cantilever at high velocity.

The final part of the curve, when the protein detaches and the surface is moved away from the cantilever at a PID specific velocity constitutes a second way to measure the drag. Fig. S2 shows such an experiment, where after detaching the cantilever experiences two force regimes. Each regime is characterized by a certain constant force and constant velocity. The drag coefficient is calculated then as the slope of the line fitted to the force and velocity of each stage, plotted one against the other (Fig. S2). The averaged values from all the gathered fits is 2.286± 0.481 pN·ms/nm, which agrees with the drag force measured independently from the cantilever itself.

The use of fast piezoelectric actuators and small cantilevers greatly improves the response time of the feed-back loop. Fig. S3 shows that the time needed to restore the constant force after an unfolding event is typically under 150 μs, while using just the PicoCube actuator and the slower MLCT cantilevers yields a response time of 4 ms.

High Force Approximation. The extension required to stretch an unfolded protein under constant force is typically approximated as resulting mainly from the entropy of the chain. The extension's dependency on the force is described by the worm-like-chain (WLC) model of elastic extension (2–4):

$$F(x) = \frac{k_B T}{p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c}\right], \qquad \textbf{[S1]}$$

where k_B is Boltzmann's constant, T is the absolute temperature, p is the persistence length, L_c is the contour length, and x denotes the end to end extension of the chain, which is also our reaction coordinate. We use a high force approximation to simplify the inversion of Eq. **S1**. According to this approximation, the linear term on the right hand side of Eq. **S1** was taken to be $x/L_c = 0.5$. In this regime, where the force is high, the contribution of the linear term in Eq. **S1** is less than 2%, and the function hardly changes with the extension when looking at the asymptote towards the contour length (Fig. S4).

$$x(F) = L_c\left[1 - \left(\frac{4pF}{k_B T} - 1\right)^{-1/2}\right]. \qquad \textbf{[S2]}$$

From the figure it is apparent that at the intercept between the two curves, at $x/L_c = 0.5$ $F \sim 13$ pN, above which the two curves are very close to each other, until $F \sim 50$ pN, where they coincide, making this approximation reasonable to use as long as we stay above these values.

Force Balance to Determine D_{eff} (Brownian Dynamics). D_{eff} determines the dynamics and time scales measured along the collapse and extension traces. With the application of an external constant force, F_i, the relaxation process is given by the following force balance (under the assumption of high force made above) along the reaction coordinate x, the end-to-end length:

$$\mu\frac{dx}{dt} + F_E(x) - F_i = 0, \qquad \textbf{[S3]}$$

where μ is the damping coefficient, which is related to the diffusion coefficient, D_{eff}, by the Einstein relation $\mu = k_B T/D_{\text{eff}}$. $F_E(x)$ is the elastic force predicted by the WLC polymer elasticity model given by Eq. **S2**. Thus, Eq. **S3** can be fully written:

$$\frac{dx}{dt} = -\frac{D_{\text{eff}}}{4p}\left(1 - \frac{x}{L_c}\right)^{-2} - \frac{D_{\text{eff}}}{4p} + \frac{D_{\text{eff}}}{k_B T}F_i \qquad \textbf{[S4]}$$

which solves to give:

$$\frac{x}{L_C} = -\left|-\frac{x}{L_C} + \left(1 - \frac{1}{q}\right)\right| e^{\kappa} e^{\alpha\left(1 - \frac{x}{L_C}\right)} e^{\lambda(t - t_0)} + \frac{1}{q} + 1, \qquad \textbf{[S5]}$$

where $q = \left(\frac{4pF_i}{k_B T} - 1\right)^{1/2}$, $\kappa = \frac{D_{\text{eff}}}{2pL_c}\left(\frac{4pF_i}{k_B T} - 1\right)^{3/2} C$, $\alpha = 2q$ and $\lambda = \frac{D_{\text{eff}}}{2pL_c}\left(\frac{4pF_i}{k_B T} - 1\right)^{3/2}$ with C being an integration constant. Defining $y = -\frac{x}{L_C} + (1 - \frac{1}{q})$, Eq. **S5** turns into

$$y = |y| e^{(\kappa + \frac{\alpha}{q})} e^{\alpha y} e^{\lambda(t - t_0)} - \frac{2}{q}. \qquad \textbf{[S6]}$$

Taking $y > 0$ returns a false expression: $x < -L_c(1 - 1/q)$ because x, q, and L_c are positive numbers bigger than 1, hence y has to be taken as <0 to give:

$$y e^{\alpha y} = -\frac{2}{q} e^{-(\kappa + 2)} e^{-\lambda(t - t_0)}. \qquad \textbf{[S7]}$$

Interestingly, this expression has both linear and exponential dependencies on x and t, a predicament which is best treated by the *Lambert* W function. This function is defined as the inverse of $f(x) = xe^x$, which basically means that $x = W[f(x)]$ (5). In order to be formalized with the W function, Eq. **S7** has to be rearranged:

$$y = \frac{1}{\alpha}\mathrm{W}\left[-\frac{2}{q} e^{-(\kappa + 2)} e^{-\lambda(t - t_0)}\right] \qquad \textbf{[S8]}$$

and then solved:

$$x = x_0 - A\mathrm{W}[-Be^{-\lambda(t - t_0)}] \qquad \textbf{[S9]}$$

here $x_0 = L_c(1 - 1/q)$ corresponds to the plateau after the relaxation, $A = L_c/\alpha$, $B = (2/q)\exp[-(\kappa + 2)]$ and

$$\lambda \cong \frac{D_{\text{eff}}}{2pL_c}\left(\frac{4pF_i}{k_B T}\right)^{3/2} \qquad \textbf{[S10]}$$

leading to the final expression:

$$x(t) = x_0 - A\mathrm{W}\left\{-B\exp\left[-\frac{D_{\text{eff}}}{2pL_c}\left(\frac{4pF_i}{k_B T}\right)^{3/2}(t - t_0)\right]\right\}. \qquad \textbf{[S11]}$$

This expression was used to fit the traces to get D_{eff}. The simulations were performed in Igor Pro 6 platform (Wavemetrics). The simulated traces were generated by solving Eq. **1** in the main text with a time resolution of $5 \cdot 10^{-8}$ s. Fig. S5 below shows a fitting with Eq. **S11** to a simulated trace (dashed black line) together with a single exponential, $x(t) = x_0 + A\exp(-t/\tau)$ (scattered red line) with their residuals. The single exponential was used to evaluate the relaxations time constants (Fig. 2 in the *main text*). Interestingly, Eq. **S11** fits better the chain's relaxation than the exponential decay.

The fitted averaged D_{eff} value of 1,226 nm^2/s used in the Brownian dynamics simulations remarkably returned a fitted averaged D_{eff} value of $1{,}214 \pm 39.81$ nm^2/s for the collapse and $1{,}223 \pm 29.40$ nm^2/s for the extensions. Interestingly, there is a variation in the spread of the experimentally observed D_{eff}'s fitted values when compared with the simulation's narrow distributions (Fig. 2*C* in the *main text* and Fig. S6). One of the most probable causes for this spread is the persistence length, p, used in the fits, which was assumed to have a constant value of 0.4 nm [a typical value for unfolded proteins (6, 7)]. However, it is evident that p can vary between 0.25 and 2.5 nm (8–11). Unlike the simulated traces, the measured traces shows a length diversity which was demonstrated to be determined to large extent by the strength of the entropic, hydrophobic and electrostatic interactions (8). These interactions result from considerable alterations in the chain's dihedral-space conformation together with side-chain packing of the collapsing structures. All these interactions are giving raise to length deviations from the WLC predictions and resulting with different persistence lengths. Nevertheless, we assumed here a constant value for p to keep the calculation simple and consistent.

Mechanical Impedance (Kelvin-Voigt Model). Frequency response function in physical systems is conveniently described by the Kelvin-Voigt circuit, which describe a simple mechanical structure consisting of a mass, a spring, and a dashpot. A force, $F(t)$, along one direction is applied and the system responds by changing the end-to-end length displacement, $x(t)$, accordingly. Here, we consider the over-damped scenario by neglecting the inertia term, thus describing the dynamic response of the system (the polypeptide) by its elasticity, k_{pol} and viscous damping coefficient, μ through its equation of motion:

$$F(t) = k_{\text{pol}}x(t) + \mu\dot{x}(t). \qquad \textbf{[S12]}$$

The Kelvin-Voigt model assumes a linear constant parameter system. Therefore its dynamic can be described by a frequency response function, which is the Laplace transform of the system's output: $x(t) => x(\omega)$, where ω is the angular frequency, resulting in:

$$x(\omega) = \frac{F(\omega)k_{\mathrm{pol}}}{k_{\mathrm{pol}}^2 + \mu^2\omega^2} - i\frac{F(\omega)\mu\omega}{k_{\mathrm{pol}}^2 + \mu^2\omega^2} = \mathrm{Re}(\omega) - i\mathrm{Im}(\omega). \quad \textbf{[S13]}$$

Here $\mathrm{Re}(\omega)$ and $\mathrm{Im}(\omega)$ denote the real and imaginary parts of the dynamic response. The output of the system in dB is given by

$$\mathrm{dB} = -10\log\left(\frac{Amp^2(\omega)}{Amp^2(0)}\right). \quad \textbf{[S14]}$$

Where $Amp^2(\omega)$ is the amplitude, given by $Amp^2(\omega) = \mathrm{Re}^2(\omega) + \mathrm{Im}^2(\omega)$ to give

$$\begin{aligned}\mathrm{dB} &= 10\log\left(\frac{1}{1 + (\mu\omega/k_{\mathrm{pol}})^2}\right)\\ &= 10\log\left(\frac{1}{1 + (2\pi f k_B T/k_{\mathrm{pol}}D)^2}\right),\end{aligned} \quad \textbf{[S15]}$$

where $\omega = 2\pi f$, f being the frequency in Hz. The bandwidth frequency, $f_{1/2}$, serves as an important parameter for characterizing the frequency at which the mechanical response drops to half of its amplitude. $f_{1/2}$ is defined at -3 dB by

$$-3 = 10\log\left(\frac{1}{1 + (2\pi f_{1/2}k_{\mathrm{B}}T/k_{\mathrm{pol}}D)^2}\right) \quad \textbf{[S16]}$$

$$f_{1/2} \approx \frac{k_{\mathrm{pol}}D}{2\pi k_{\mathrm{B}}T}. \quad \textbf{[S17]}$$

Using the WLC assumptions with a frictional term, Khatri, et al. introduced an expression for a polypeptide chain elasticity (12):

$$k_{\mathrm{WLC}} = \frac{4}{L_c}\sqrt{\frac{p}{k_{\mathrm{B}}T}}F^{3/2} \quad \textbf{[S18]}$$

when substituted into Eq. **S17**, we get the final expression for the bandwidth frequency dependency on the system's parameters:

$$f_{1/2} = \frac{2D\sqrt{p}}{\pi L_c}\left(\frac{F}{k_{\mathrm{B}}T}\right)^{3/2}. \quad \textbf{[S19]}$$

Molecular Dynamics Simulations. All-atom simulations in explicit solvent were carried out with the software NAMD 2.8, using the CHARMM22 force-field with CMAP corrections for the protein, and the TIP3P water model. We used periodic boundaries conditions and a cutoff of 12 Å for electrostatic and Lennard-Jones interactions. Long-range electrostatic interactions were calculated using the PME method with a grid spacing of 1 Å. All bonds between light and heavy atoms were maintained rigid, while the rest of the protein was flexible. Steered MD simulations of wild-type ubiquitin (PDB ID code 1UBQ) were performed by fixing the Cα of the first residue (MET1) and by applying a constant force on the Cα of the last residue (GLY 76) along the z direction.

System preparation: To unfold the protein, we first pull on ubiquitin molecule in vacuum at a high force of 800 pN, during 10 ns. A fully extended protein was thus generated, with no remaining secondary structure. It was then solvated using the waterbox module of VMD in a box of 3.5 * 3.5 * 32 nm, comprising 11,499 water molecules and 35,728 atoms total. Energy minimization using the steepest descent method (2,000 steps) was performed before further equilibration, as described below.

Equilibration: The protein was then equilibrated for 6 ns at 250 pN in the isobaric ensemble at 300 K and 1 bar, using a time-step of 2 fs, a Langevin thermostat (damping coefficient of 1 ps^{-1}) for temperature control and the modified NAMD version of the Nose-Hoover barostat with Langevin dynamics (piston period of 0.1 ps and piston decay time of 0.05 ps) for pressure control. This simulation was then propagated for 25 more ns to check that the average end-to-end distance no longer evolved. No dynamical data was extracted from these simulations because of possible bias introduced by the temperature and pressure control.

Collapse simulations: Initial configurations for collapse from 250 pN to 100 pN were chosen along the 25-ns trajectory at 250 pN. The simulations were propagated in the micro-canonical ensemble for 5 ns to avoid spurious effects from pressure and temperature control on the dynamics of collapse. A time-step of 1 fs was used. We performed five such simulations to obtain the average relaxation. Because of the large system size, average temperature and pressure along these trajectories are very close to that targeted during the NPT equilibration.

Extension simulations: The last configuration of one of the collapse trajectory was then propagated for 30 ns using the same setup as for equilibration (NPT ensemble), both to check convergence of the end-to-end distance and to generate starting configurations for extension: finally, five such trajectories were generated by pulling back the protein at 250 pN during 5 ns in the NVE ensemble.

Diffusion coefficient: To estimate the diffusion coefficient along the end-to-end coordinate, we have employed a method described earlier (13, 14). At each given force an additional 3-ns simulation in the micro-canonical ensemble is performed using the collective-variable module of NAMD to add a bias potential on the end-to-end distance. This potential is harmonic and chosen to be much stiffer than the actual PMF on which the protein is moving, so that the resulting PMF is locally harmonic. Under this approximation, it can be shown that the diffusion coefficient can be recovered from the time autocorrelation function of the end-to-end distance L,

$$D = \frac{\langle \delta L^2 \rangle^2}{\int_0^\infty \langle \delta L(t)\delta L(0)\rangle dt}, \quad \textbf{[S20]}$$

where $\delta L = L - \langle L \rangle$ are the fluctuation of L around its average value. We performed simulations at both forces (100 pN and 250 pN), using a force constant of 100 kcal/mol/$\mathrm{Å}^2$ to constraint the system around the average end-to-end distance $\langle L \rangle$ estimated from unperturbed simulations (25.5 nm at 250 pN and 23.7 nm at 100 pN). In both cases, the unperturbed PMF is very smooth in this region and we checked on a smaller, model system (decaalanine in water) using different values for the force constant (20, 50, and 100 kcal/mol/$\mathrm{Å}^2$) that it does not have any significant impact on the obtain value of D.

An example of the average autocorrelation function of L is shown in Fig. S7. Error bars on D are estimated from block averaging and by estimating D independently for each block. A manuscript presenting further details about this approach and other new insights from MD simulations is currently in preparation.

1. Carrion-Vazquez M, Marszalek PE, Oberhauser AF, Fernandez JM (1999) Atomic force microscopy captures length phenotypes in single proteins. *Proc Natl Acad Sci USA* 96:11288–11292.
2. Fixman M, Kovac J (1973) Polymer conformaional statistics. 3. Modified Gaussian models of stiff chains. *J Chem Phys* 58:1564–1568.
3. Kovac J, Crabb CC (1982) Modified Gaussian model for rubber elasticity. 2. The wormlike chain. *Macromolecules* 15:537–541.
4. Bustamante C, Marko JF, Siggia ED, Smith S (1994) Entropic elasticity of Lambda-Phage DNA. *Science* 265:1599–1600.
5. Corless RM, Gonnet GH, Hare DEG, Jeffrey DJ, Knuth DE (1996) On the Lambert W function. *Advances in Computational Mathematics* 5:329–359.
6. Carrion-Vazquez M, et al. (1999) Mechanical and chemical unfolding of a single protein: A comparison. *Proc Natl Acad Sci USA* 96:3694–3699.
7. Carrion-Vazquez M, et al. (2003) The mechanical stability of ubiquitin is linkage dependent. *Nature Structural Biology* 10:738–743.
8. Walther KA, et al. (2007) Signatures of hydrophobic collapse in extended proteins captured with force spectroscopy. *Proc Natl Acad Sci USA* 104:7916–7921.
9. Li HB, et al. (2001) Multiple conformations of PEVK proteins detected by single-molecule techniques. *Proc Natl Acad Sci USA* 98:10682–10686.
10. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J (2002) Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: Testing a 1-dimensional diffusion model for peptide dynamics. *J Phys Chem B* 106:11628–11640.
11. Liu RC, Garcia-Manyes S, Sarkar A, Badilla CL, Fernandez JM (2009) Mechanical characterization of protein L in the low-force regime by electromagnetic tweezers/evanescent nanometry. *Biophys J* 96:3810–3821.
12. Khatri BS, et al. (2008) Internal friction of single polypeptide chains at high stretch. *Faraday Discuss* 139:35–51.
13. Hummer G (2005) Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J Phys* 7:516–523.
14. Straub JE, Borkovec M, Berne BJ (1987) Calculation of dynamic friction on intramolecular degrees of freedom. *J Phys Chem* 91:4995–4998.

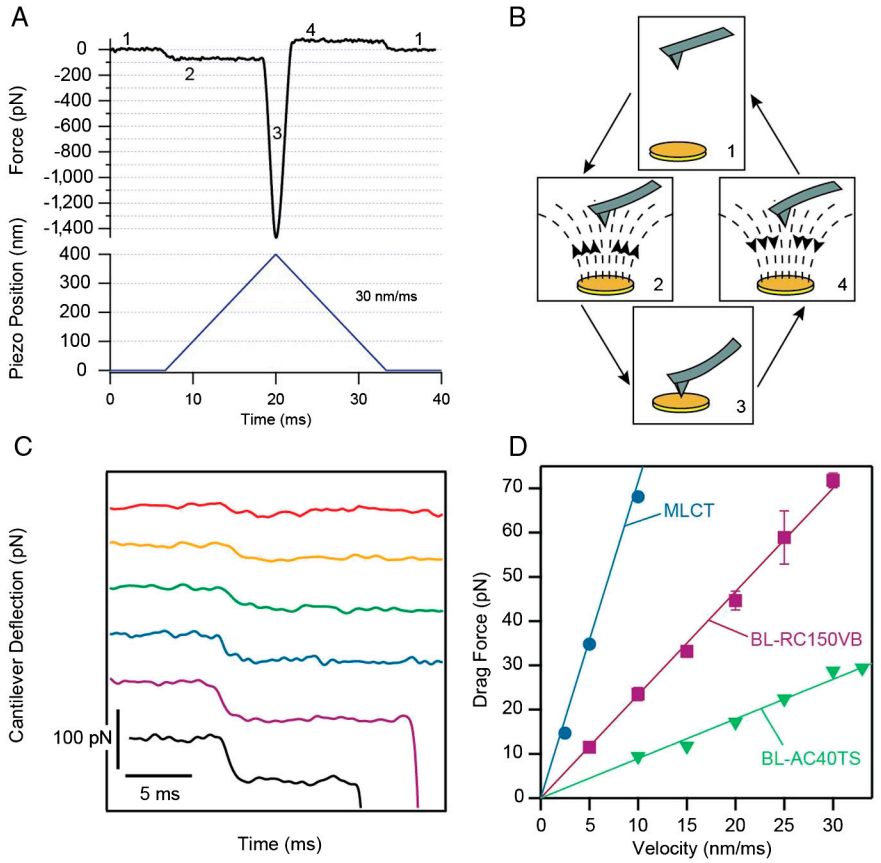

Fig. S1. Cantilever's drag force measurements. (*A*) Traces of the drag force measured on the cantilever according to the position protocol of the Piezo actuator when moved at 30 nm/ms. (*B*) Schematic stages of the position of the piezo and the cantilever numbered according to the different stages shown in a. (*C*) Six force traces taken at velocities varying from 5 to 30 nm/ms. (*D*) Force vs. velocity of each experiment plotted together from three different cantilevers showing a drag coefficient of 5.7 ± 0.1 pN·ms/nm for MLCT, 2.335 ± 0.03 pN·ms/nm for BL-RC150VB and 0.895 ± 0.04 pN·ms/nm for BL-AC40TS.

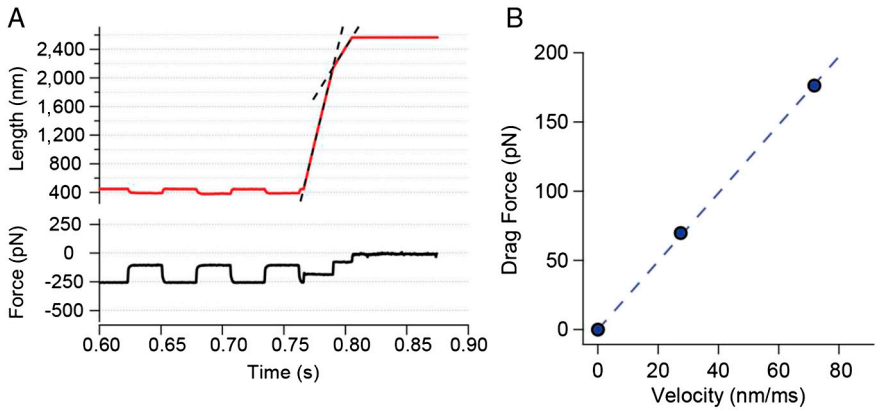

Fig. S2. Drag force measurements from constant force AFM experiments. (*A*) An experiment example, in which an unfolded poly-ubiquitin chain is stretched and relaxed between 250 and 100 pN. The piezo actuator is rapidly moved away from the cantilever as the chain ruptures at approximately 0.77 s. (*B*) Force vs. velocity of the two stages detected after the detachment, out of which the drag coefficient is estimated per molecule from the slope of the linear fit between the points.

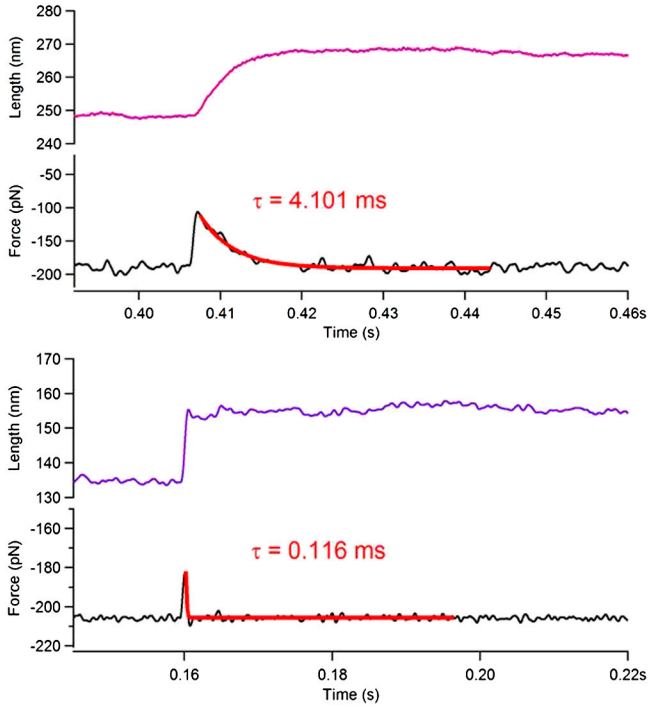

Fig. S3. Two force traces of a single ubiquitin unfolding. The control system responds to the unfolding event driving the actuator withdraw the protein and recover the condition of constant force. Force trace obtained by a standard AFM that shows decay with a time resolution of 4 ms (upper box). The new fast-AFM used in this work shows an almost instantaneous force recovery with a time constant smaller than 150 μs under the same conditions (low box).

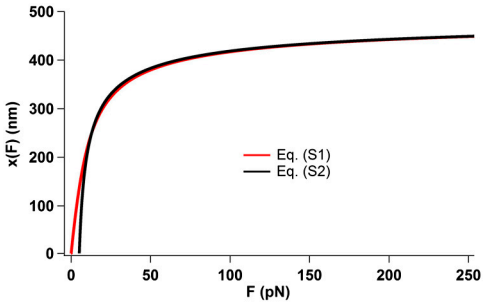

Fig. S4. Comparison between the phenomenological WLC model, given by Eq. **S1** (red) and its high force modification, Eq. **S2** (black). From forces above 50 pN the two curves coincide.

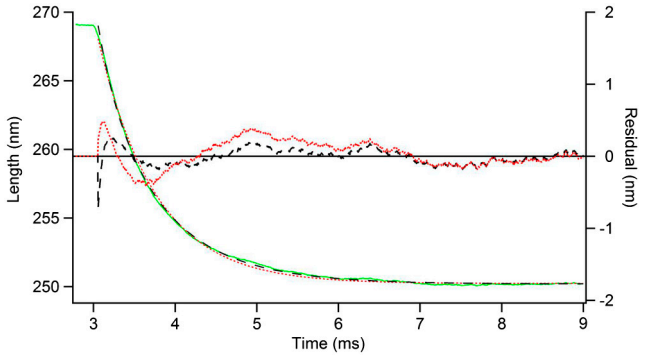

Fig. S5. Fitting of the model to a simulated trace. A simulated trace (green line; $L_c = 300$ nm, $p = 0.4$ nm and $D_{eff} = 1{,}226$ nm^2/s) was fitted with the model presented in this study (Eq. **S11**—dashed black line) and with a single exponential (scattered red line). The right axis shows the residuals of the fits.

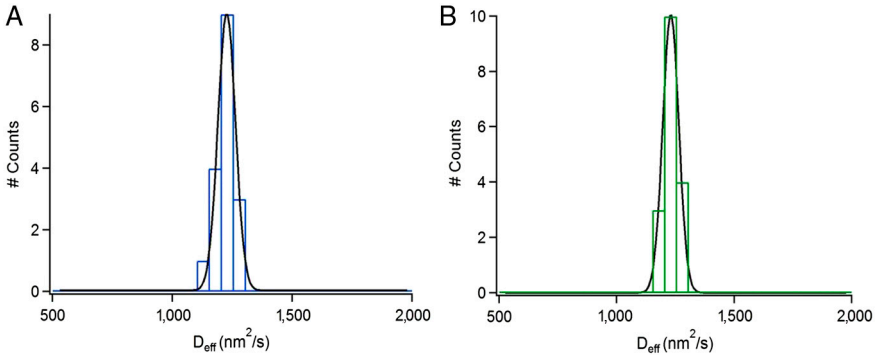

Fig. S6. Histograms of values for D_{eff} measured from Brownian Dynamic simulated traces for (*A*) collapse and (*B*) extension.

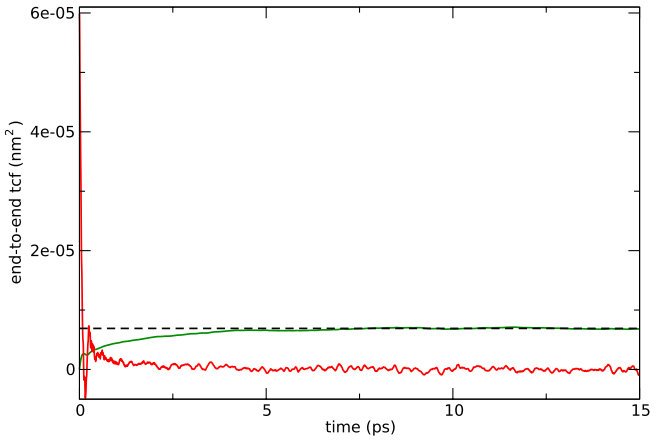

Fig. S7. Time correlation function of the normalized end-to-end length δL as a function of time (red curve) for ubiquitin at 100 pN. Its integral is shown in green, together with the extrapolated value used in Eq. **S20** (black dashes). The corresponding value of D is $(5.3 \pm 1.4)\,10^8$ nm^2/s.