Crowding Effects on the Mechanical Stability and Unfolding Pathways of Ubiquitin

David L. Pincus, and D. Thirumalai


Downloaded from http://pubs.acs.org on January 9, 2009

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML
Crowding Effects on the Mechanical Stability and Unfolding Pathways of Ubiquitin

David L. Pincus* and D. Thirumalai

Biophysics Program, Institute for Physical Science and Technology, and Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742

Received: September 1, 2008; Revised Manuscript Received: October 17, 2008

The interiors of cells are crowded, thus making it important to assess the effects of macromolecules on the folding of proteins. Using the self-organized polymer (SOP) model, which is a coarse-grained representation of polypeptide chains, we probe the mechanical stability of ubiquitin (Ub) monomers and trimers ((Ub)₃) in the presence of monodisperse spherical crowding agents. Crowding increases the volume fraction (Φᵥ) dependent average force (<fₘ(Φᵥ)>), relative to the value at Φᵥ = 0, needed to unfold Ub and the polyprotein. For a given Φᵥ, the values of <fₘ(Φᵥ)> increase as the diameter (σᵥ) of the crowding particles decreases. The average unfolding force (<fₘ(Φᵥ)> varies with the radius of gyration of Ub (or (Ub)₃) in the unfolded state. Examination of the unfolding pathways shows that, relative to Φᵥ = 0, crowding promotes reassociation of ruptured secondary structural elements. Both the nature of the unfolding pathways and <fₘ(Φᵥ)> for (Ub)₃ are altered in the presence of crowding particles, with the effect being most dramatic for the subunit that unfolds last. We predict, based on SOP simulations and theoretical arguments, that <fₘ(Φᵥ)> ∼ Φᵥ¹/₃ν, where ν is the Flory exponent that describes the unfolded (random coil) state of the protein.

Introduction

Cells exist in a crowded environment consisting of macromolecules (lipids, mRNA, ribosome, sugars, etc.), making it critical to investigate protein folding in the presence of crowding agents. If the interactions between the crowding agents and the protein of interest are short-ranged and nonspecific (as is often the case), then the volume excluded by the crowding agents prevents the polypeptide from sampling extended conformations. As a consequence, the entropy of the denatured state ensemble (DSE) decreases relative to the case when the crowding agents are absent. These arguments suggest that excluded volume of crowding agents should enhance the stability of the folded state, provided that the crowding-induced changes in the native state are negligible. The entropic stabilization mechanism, described above, has been used in several theoretical models to quantitatively describe the extent of the folded protein as a function of the volume fraction, Φᵥ, of the crowding agents. More recently, a theory whose origins can be traced to the concept of intraprotein attraction due to depletion of crowding agents near the protein predicts that the enhancement in stability is ΔT(Φᵥ) = Tᵢ(Φᵥ) - Tᵢ(Φᵥ = 0) ∼ Φᵥα, where Tᵢ(Φᵥ) is the folding temperature at Φᵥ and α is related to the Flory exponent that characterizes the size of the protein in the DSE. From this prediction, it follows that crowding affects the DSE to a greater extent than that in the folded state. Although the precise theoretical predictions of the power law change in ΔT(Φᵥ) as Φᵥ changes have not been verified, several experiments using a number of proteins have confirmed that indeed Tᵢ(Φᵥ) increases with Φᵥ. It cannot be emphasized enough that the theory described here applies only to cases when the crowding interactions between crowding agents and proteins and between crowding particles themselves are purely repulsive.

While much less is known about the effects of crowding on the folding kinetics, Cheung et al. predicted that the entropic stabilization also suggests that the folding rates should increase at moderate values of Φᵥ. They suggest that crowding can enhance folding rates by a factor of exp(ΔS(Φᵥ)/kT), where ΔS(Φᵥ)(~Φᵥα) is the decrease in the entropy of the DSE relative to its value in the bulk. From the arguments of Cheung et al., it follows that the equilibrium changes in the entropy (ΔS(Φᵥ)) of the DSE, with respect to the bulk, should also determine rate enhancement provided that neither the barriers to folding nor the native state is perturbed significantly by crowding particles.

Single-molecule force spectroscopy, such as atomic force microscopy (AFM) and laser optical tweezers, have been used to monitor the behavior of biopolymers under tension and are ideally suited to probe the enhancement in crowding-induced stability by a direct measure of <fₘ(Φᵥ)> . Indeed, Ping et al. have recently investigated the effect of dextran molecules on the mechanical stability of (Ub)₃. The eight Ub (Figure 1A) modules were N–C linked (i.e., modules i and i + 1 were chemically linked together in a head-to-tail manner). They found that the average force required to unfold a module, <fₘ(Φᵥ)>, increased by 21% as the dextran concentration, ρ, was increased from 0 to 300 g/L at rₙ = 4.2 × 10⁷ pN/s. Similar results have been obtained recently by Yuan et al. at rₙ = 12.5 × 10⁷ pN/s.

Motivated in part by experiments, we used simulations to investigate the effects of crowding agents on the mechanical stability of a protein subject to external tension. We focused on ubiquitin (Ub), a 76 residue protein composed of 5 β strands and 2 α helices (Figure 1A), and confined our investigation to nonequilibrium “force-ramp” experiments. The primary data recorded during such an experiment is a trace of the force exerted on the tip as a function of the extension of the molecule, a force-extension curve (FEC). When the force exceeds some critical value, the FEC displays a sudden increase in length and is often accompanied by a concomitant sharp decrease in force.
Presumably, the sharp change corresponds to the unfolding of the protein. Typical AFM experiments use tandem arrays of proteins which are chemically linked together (often through genetic engineering). We use the term module to denote a protein of the array. The FEC resulting from such an experiment reveals several equally spaced peaks punctuated by sharp increases in the extension of the molecule corresponding to the unfolding of individual modules. The height of these force peaks and their shape depend on the loading rate, \( r_1 = k_s \times \nu \), where \( k_s \) is the cantilever’s spring constant and \( \nu \) is the (constant) speed at which the stage is retracted away from the cantilever.\(^{15}\)

In order to compare to experiments, our simulations are performed using coarse-grained models for which simulations can be done at \( r_1 \) that are comparable to those used in AFM experiments. Our work has led to a number of testable results. (1) At \( \Phi_c = 0.3 \), the average unfolding force for Ub increases by at most only 7% compared to that at \( \Phi_c = 0 \). We find that \( \langle f_d(\Phi_c) \rangle \) in small crowding agents is greater than that in larger particles. (2) In the presence of crowding agents, secondary structural elements re-form multiple times even after initial rupture. (3) Although large crowding particles are predicted to have a smaller effect on \( \langle f_d(\Phi_c) \rangle \) (for a given \( \Phi_c \)), they can profoundly affect the unfolding of polyUb. We predict that \( \langle f_d(\Phi_c) \rangle \) for a given subunit depends on the number of already unfolded portions of the polyprotein. This result is important because many naturally occurring proteins that are subject to tensile stresses exist as tandem arrays of modules. It further suggests that the existence of such redundancy can more properly be understood in the context of a crowded cellular milieu.

Methods Sections

Self-Organized Polymer Model for Ub. We used a coarse-grained model for proteins to investigate crowding effects on the mechanical stability of Ub and (Ub)\(_n\) at loading rates that are comparable to those used in AFM experiments.\(^{12,13}\) We assumed that Ub could be described using the self-organized polymer (SOP) model, a model that has been successfully used to make a number of predictions regarding the unfolding of proteins and RNA,\(^{16,17}\) allosteric transitions in enzymes,\(^{18,19}\) and movement of molecular motors on polar tracks.\(^{20,21}\) Previous studies\(^{22,23}\) have used more standard Go models to probe various aspects of forced unfolding of Ub. The SOP energy function \( (E_p) \) for a protein with \( N \) amino acids, specified in terms of the \( C_a \) coordinates \( r_i \) \((i = 1, 2, \ldots, N)\), is

\[
E_p = E_{FENE} + E_{nb}^{\text{str}} + E_{nb}^{\text{rep}} =
- \sum_{i=1}^{N} \frac{k}{2} \sum_{j=i+1}^{N} R_0^2 \ln \left[ 1 - \frac{(r_{i,j+1} - r_{i,j}^0)^2}{R_0^2} \right] +
\sum_{i=1}^{N-1} \sum_{j=i+3}^{N} \varepsilon_{ij} \left( \frac{r_{ij}^0}{r_{ij}} \right)^{12} - 2 \left( \frac{r_{ij}^0}{r_{ij}} \right)^6 \Delta_{ij} +
\sum_{i=1}^{N} \sum_{j=i+2}^{N} \varepsilon_{ij} \left( \frac{g}{r_{ij}^0} \right)^6 (1 - \Delta_{ij}) \quad (1)
\]

where \( r_{ij} = |r_i - r_j| \), \( r_{ij}^0 = |r_i^0 - r_j^0| \) is the value of \( r_{ij} \) in the native structure, \( k = 2 \times 10^3 \text{ kcal/(mol-nm)}^2 \), \( r_{ij}^0 = 1.4 \text{ kcal/mol} \), \( \varepsilon_{ij} = 1.0 \text{ kcal/mol} \), and \( \sigma = 0.38 \text{ nm} \). Note that \( k_BT \approx 0.6 \text{ kcal/mol} \approx 4.2 \text{ pN-nm} \). In eq 1, \( \Delta_{ij} = 1 \) if \( r_{ij}^0 < 0.8 \text{ nm} \) and \( \Delta_{ij} = 0 \) otherwise. Native coordinates correspond to those of the \( C_a \) atoms of the 1.8 Å resolution Protein Data Bank crystal structure 1UBQ.\(^{24}\) For Ub, \( N = 76 \) and 228 for (Ub)\(_3\). The first term in eq 1 is the FENE potential\(^{25}\) that accounted for chain connectivity. The second (Lennard-Jones) term accounted for the nonbonded interactions that stabilize the native state, and the final (soft-sphere) term accounted for excluded-volume interactions (including those of an angular nature). The SOP model is different from the Go model because there are no angular terms in SOP, and the connectivity is enforced differently as well. The SOP representation of the polypeptide chain is in the same spirit as other coarse-grained models used in polymers.\(^{26}\)
Mechanical Stability and Unfolding Pathways of Ubiquitin

Crowding Particles and Interactions with Ub. We assumed that the crowding particles are spherical with diameter \( \sigma_c = 6.4 \) nm in some simulations, while \( \sigma_c = 1.0 \) nm in others. Crowders interacted among themselves and with the protein, respectively, via the following LJ potentials

\[
E_{cc} = 4\epsilon \left( \left( \frac{\sigma_{cc}}{r} \right)^{12} - \left( \frac{\sigma_{cc}}{r} \right)^{6} + \frac{1}{4} \right) \Theta(r_{\text{min}}^{cc} - r) \tag{2}
\]

\[
E_{cp} = 4\epsilon \left( \left( \frac{\sigma_{cp}}{r} \right)^{12} - \left( \frac{\sigma_{cp}}{r} \right)^{6} + \frac{1}{4} \right) \Theta(r_{\text{min}}^{cp} - r) \tag{3}
\]

where \( \sigma_{cc} = (\sigma_a + \sigma_p)/2 \), \( \sigma_p = \sigma = 0.38 \) nm, \( r_{\text{min}}^{cc} = 2^{1/6} \sigma_c \), and \( r_{\text{min}}^{cp} = 2^{1/6} \sigma_{cp} \). The Heaviside functions truncate the potentials at their minima and thereby ensure that only the repulsive portions of eq 2 and 3 contribute to interactions involving crowding agent.

Mimics of Crowding Using Asakura–Oosawa Theory. Even using a coarse-grained SOP representation of proteins, it is difficult to carry out converged simulations in the presence of crowding agents. The reason is that the number of crowding agents can be large. Moreover, the separation in the spatial and temporal scales of the protein and the crowding particles has to be carefully considered to obtain reliable results. In light of these difficulties, it is of interest to consider the effective attraction between the sites on the protein using the implicit pairwise potential computed by Asakura and Oosawa. The intramolecular attraction arises due to the depletion of crowding particles near the protein. To probe the efficacy of these models, we employed in some simulations the Asakura–Oosawa model of crowding effects. For these simulations, we added the following term to the bare SOP Hamiltonian (eq 1)

\[
E_{AO}(r_{ij}) = -\Phi_c k_B T \sum_{j\neq i+3} \left( \frac{(\sigma + \sigma_c)}{\sigma_c} \right)^3 \times \left( 1 - \frac{3r_{ij}}{2(\sigma + \sigma_c)} + \frac{r_{ij}^2}{2(\sigma + \sigma_c)^3} \right) \sigma < r_{ij} < \sigma + \sigma_c \tag{4}
\]

where \( \sigma = 0.38 \) nm, \( \sigma_c = 6.4 \) nm, \( \Phi_c = 0.3 \), \( k_B T \approx 4.2 \) pN·nm, and \( r_{ij} \) is the distance separating protein beads \( i \) and \( j \).

(UBj) Intermodule Interactions. For simulations involving (UBj), residues in different modules interacted via

\[
E_{pp} = \epsilon_p \left( \frac{Q}{r} \right)^6 \tag{5}
\]

where \( r \) is the distance separating the two beads. Note that this potential is short-ranged and purely repulsive and that it is the same potential used for non-native intraprotein interactions.17

Simulation Details. \( \Phi_c = 0 \). Hundreds of simulations of 5 \times 10^8 steps (=30 \mu s) at \( T = 300 \) K were used to generate initial structures for use in the pulling simulations. The protein was completely free in solution (i.e., no forces were applied to either terminus), and no crowders were present during the equilibrations. The N-terminus of the protein was subsequently translated to the origin, and the protein was rotated such that its end-to-end vector, \( \mathbf{R} \) (i.e., the vector pointing from the N-terminal bead to the C-terminal bead) coincided with the pulling (+z) direction.

An unfolding trajectory was initiated by selecting a random initial structure from among the set of thermally equilibrated structures and by tethering a harmonic spring to the C-terminal bead. The N-terminal bead was held fixed throughout the simulations. Tension was applied to the protein by displacing the spring along the +z axis and resulted in application of the following force to the C-terminal bead

\[
f_z = -k_c ([z(t) - z(0)] - [z_c(t) - z_c(0)]) \tag{6}
\]

where \( k_c \) is the spring constant, \( z(t) = \mathbf{R}(t)\cdot \hat{z} \), and \( z_c(t) \) corresponds to the displacement of the end of the spring. Note that \( k_c \) was also used to constrain the simulation to the z-axis; \( f_z = -k_c[z(t) - x(0)] \) and \( f_z = -k_c[y(t) - y(0)] \). The displacement of the spring was updated at every time step.

We simulated forced unfolding of monomeric Ub at four different \( r_f \) values (160 \times 10^3, 80 \times 10^3, 20 \times 10^3, and 4 \times 10^3 \) pN/s), while simulations on N–C-linked (Ubj) were performed at \( r_f = 640 \times 10^3 \) pN/s. All overdamped force-ramp simulations were performed at the same speed, \( \nu = 10312 \) nm/s, and spring constants were varied over a range from 0.3879 to 31.032 pN/nm to achieve the aforementioned \( r_f \) (via the relation \( r_f = k_c \nu \)). Our simulations were realistic because they maintained loading rates consistent with experiment and because \( r_f \) is the prime determinant of the unfolding pathway.15

\[ \Phi_c = 0.0 \). Simulations involving explicit crowders were carried out at a fixed volume fraction of \( \Phi_c = 0.3 \) and with a fixed number \( N_c = 100 \) of crowding spheres. \( N_c \) was fixed to render the problem computationally tractable.) By using the relation \( \Phi_c = (N_c \pi \sigma^2 / \alpha \cdot L^3) \), we adjusted the length of a side of the cubic simulation box (\( L \)) to maintain \( \Phi_c = 0.3 \). Thus, \( L = 35.8 \) nm when \( \sigma_c = 6.4 \) nm and \( L = 5.6 \) nm when \( \sigma_c = 1.0 \) nm. Explicit crowders were added to the simulation after loading an equilibrated structure but before the application of tension. Initial crowder positions were chosen randomly and in a serial manner from a uniform distribution. If the distance between an initial crowder position and that of another crowder or protein bead did not exceed the sum of their radii, then the prospective position was rejected and another random position chosen to avoid highly unfavorable steric overlaps.

Periodic boundary conditions (PBC) and the minimum image convention were employed in the simulations. Two sets of coordinates were stored for protein beads at every time step; PBC were applied to one set, and the other was propagated without PBC. Distances between protein beads were calculated from the uncorrected set of coordinates without minimum imaging, while protein–crowder distances were calculated from the PBC coordinates with minimum imaging.

To improve simulation efficiency, a cell list was used to calculate crowder–crowder and protein–crowder interactions. The entire simulation volume was partitioned into 64 subvolumes, and it was only necessary to calculate interactions within a subvolume and between beads of the subvolume and those of 13 of its 26 neighbors. The cell list was updated at every time step to ensure the accuracy of the simulations.

The equations of motion in our overdamped simulations (used in all force-ramp simulations) were integrated with a time step of \( h = 0.01 t_r \) (\( t_r = 2.78 \) ps) using the method of Ermak and McCammon.28 The friction coefficient of the crowders, \( \zeta_c \), was determined via the relation \( \zeta_c = \sigma_i / \alpha_i \), where \( \sigma_i \) is the crowder diameter, \( \sigma = 0.38 \) nm is the diameter of a protein bead, and \( \zeta = 83.3 \times (m / \tau_r) = 9 \times 10^{-9} \) g/s is the friction coefficient associated with a protein bead of mass \( m = 3 \times 10^{-22} \) g. Simulated times were translated into real times using \( \tau_r = (\zeta_c \rho) / (2 \rho) \).
k_{B}T) \times r_{1} \times (r_{1}/m)^{29} \text{At } T = 300 \text{ K, } \tau_{II} = 543.06 \text{ ps, and since } h = 0.01 \times r_{1}, \text{ the real time per step is } 5.4306 \text{ ps.}

**Results and Discussion**

**Monomeric Ub at } \Phi = 0.0. At \Phi = 0.0, \text{ forced unfolding of Ub was simulated at four different } r_{1} \text{ values } (4 \times 10^{3}, 20 \times 10^{3}, 80 \times 10^{3}, 160 \times 10^{3} \text{ pN/s}), \text{ where the average value corresponds approximately to the value used in the pulling experiments of Ping et al.}^{12} \text{ and all } r_{1}'s \text{ are experimentally accessible.}

**Force Profiles.** Figure 2A and B provides examples of FECs at the highest and lowest } r_{1}'s. We used a nominal contour length of } (N-1)\sigma = 75 \times 0.38 \text{ nm and unfolding forces, } f_{u}, \text{ to determine contour length increments, } \Delta \zeta' \text{ for each trajectory at } r_{1} = 160 \times 10^{3} \text{ pN/s (Figure 2A). We identified } f_{u} \text{ with the peak of the FEC before the stick--slip transition.}^{30,31} \text{ The average extension } \langle \Delta \zeta' \rangle = 23.991 \pm 0.010 \text{ nm is in excellent agreement with the experimental result of } 24 \pm 5 \text{ nm found by Carrion-Vazquez et al.}^{30} \text{ The projection } (\zeta_{u}) \text{ of the end-to-end vector at } f_{u} \text{ in the } z-\text{direction varied between 4.1 and 4.7 nm, depending on } r_{1}. \text{ Since the native end-to-end distance is } z_{0} = 3.7 \text{ nm, } z_{u} = \zeta_{0} = \Delta \zeta_{u} \text{ ranges from } 0.4 \text{ to } 1.0 \text{ nm. The lower end of this range is slightly larger than the } 0.25 \text{ nm transition-state distance for the mechanical unfolding of the structurally similar titin immunoglobulin domains.}^{32} \text{ Indeed, we expect } \Delta \zeta_{u} > 0.25 \text{ nm because of the nonequilibrium nature of the simulations. Larger } r_{1} \text{ values typically lead to larger } \langle \Delta \zeta_{u} \rangle \text{ (} \sim k_{B}T/\langle f_{u} \rangle \text{ln}(r_{1}) \rangle). \text{ Average unfolding forces, } \langle f_{u}(\Phi_{C}) \rangle \text{, depended approximately logarithmically on } r_{1}^{15} \text{ (Figure 2C), and } \langle f_{u}(r_{1}) = 4 \times 10^{3} \text{ pN/s) } \rangle = 136 \text{ pN is in fair agreement with the experimental value of } 166 \pm 33 \text{ pN observed by Ping et al.}^{12} \text{ at } r_{1} = 4.2 \times 10^{3} \text{ pN/s. The } (\zeta_{u}) \text{ also showed a logarithmic dependence on } r_{1} \text{ but the difference between the value calculated at } r_{1} = 160 \times 10^{3} \text{ pN/s and that calculated at } r_{1} = 4 \times 10^{3} \text{ pN/s is small (} \pm 2 \text{ Å). Although the underlying free-energy landscape is time-depend-} \text{ ent in a nonequilibrium force-} \text{ampull pulling experiment, these results suggest that the distance from the native state to the transition state is small. It is likely that at loading rates that are achieved in laser optical tweezer experiments (} \sim 10 \text{ pN/s), the location to the transition state would increase because of the response of biopolymers to loading rate changes from being plastic (low } r_{1}) \text{ to brittle (high } r_{1}).^{33} \text{ Unfolding Pathways. In the dominant pathway, unfolding proceeded in a fairly Markovian fashion with the primary order of events following the sequence } \beta_{1}/\beta_{5} \rightarrow \beta_{3}/\beta_{5} \rightarrow \beta_{3}/\beta_{4} \rightarrow \beta_{1}/\beta_{2} \text{ (Figures 1B and 3). This is precisely the same sequence as that seen in the simulations by Li et al.}^{23} \text{ Alternative pathways, reminiscent of kinetic partitioning}^{32} \text{ observed in the forced unfolding of GFP}^{35} \text{ and lysozyme,}^{36} \text{ were also infrequently sampled. For example, at } r_{1} = 4 \times 10^{3} \text{ pN/s, } \sim 6\% \text{ of the trajectories unfolded as follows, } \beta_{1}/\beta_{5} \rightarrow \beta_{1}/\beta_{2} \rightarrow \beta_{3}/\beta_{5} \text{ (Figures 1C and 3). Of the 16 trajectories simulated, 4 resulted in larger molecular extensions before significant resistance was encountered. (Yellow arrows correspond to } \Phi \text{ strands and purple cylinders correspond to } \alpha \text{ helices.) (The figures were generated with VMD.}^{25} \text{ (C) } \langle f_{u} \rangle \text{ versus } r_{1} \text{ evaluated at } \Phi_{C} = 0.0. \text{ The red curve corresponds to a linear least-squares fit to the set of basis functions } [1, \text{ln}(r_{1})] \text{ and demonstrates that } \langle f_{u} \rangle \sim \text{ln}(r_{1}). \text{ (Note that the abscissa is a log scale.) Each point is labeled as mean } \pm \text{ standard error. Statistics at } r_{1} = 160 \times 10^{3}, 80 \times 10^{3}, 20 \times 10^{3}, \text{ and } 4 \times 10^{3} \text{ pN/s were calculated from } 50, 49, 50, \text{ and } 16 \text{ trajectories respectively.}

Figure 2. Force-extension curves (FECs) at two different loading rates, (A) } r_{1} = 160 \times 10^{3} \text{ pN/s and (B) } r_{1} = 4 \times 10^{3} \text{ pN/s. Data from the simulation is presented as a red trace. For each trajectory, a black arrow points to the unfolding force, } f_{u}. \text{ The } z_{u} \text{ corresponds to the extension of the molecule along the pulling (i.e., } z)-\text{direction evaluated at } f_{u}. \text{ The } \Delta r_{1} \text{ (dotted blue line) is the contour length increment and is a measure of the amount of chain released in an unfolding event. We measured } \Delta \zeta' \text{ as } \zeta' - z_{u}, \text{ where } \zeta' = (N-1) \times \sigma = 75 \times 0.38 \text{ nm is a nominal contour length of } 28.5 \text{ nm. Stars in each subfigure mark the minimum force observed after an unfolding event, and chain conformations corresponding to the starred points in figures A and B are illustrated at the top of the figure. Unfolding events at smaller } r_{1} \text{ resulted in larger molecular extensions before significant resistance was encountered. (Yellow arrows correspond to } \Phi \text{ strands and purple cylinders correspond to } \alpha \text{ helices.) (The figures were generated with VMD.}^{25} \text{ (C) } \langle f_{u} \rangle \text{ versus } r_{1} \text{ evaluated at } \Phi_{C} = 0.0. \text{ The red curve corresponds to a linear least-squares fit to the set of basis functions } [1, \text{ln}(r_{1})] \text{ and demonstrates that } \langle f_{u} \rangle \sim \text{ln}(r_{1}). \text{ (Note that the abscissa is a log scale.) Each point is labeled as mean } \pm \text{ standard error. Statistics at } r_{1} = 160 \times 10^{3}, 80 \times 10^{3}, 20 \times 10^{3}, \text{ and } 4 \times 10^{3} \text{ pN/s were calculated from } 50, 49, 50, \text{ and } 16 \text{ trajectories respectively.}
Mechanical Stability and Unfolding Pathways of Ubiquitin


Figure 3. Rupture events at $\Phi_c = 0.0$ (A) and at $\Phi_c = 0.3$ and $\sigma_c = 6.4$ nm (B). The figure illustrates that after the initial rupture event at step 0, subsequent unfolding was affected by the crowding agent. $\beta1/\beta2$ contacts and $\beta3/\beta4$ persisted longer, and there were many more local refolding events in the presence of the crowders (B) than in their absence (A). (C) Snapshots from an unfolding trajectory at $\Phi_c = 0.3$ and $r_i = 4 \times 10^3$ pN/s illustrate the primary difference between unfolding at $\Phi_c = 0.3$ and that at $\Phi_c = 0.0$. The brown dotted arrow shows unfolding without any local refolding events, while the sequence of black arrows illustrates local rupture and refolding events. The hallmark of forced unfolding in a crowded environment is repeated breaking and re-forming of contacts after an initial rupture event. (The figures were generated with VMD.43)

$\rightarrow \beta3/\beta4$ (Figure 1B), while the remaining $\sim 94\%$ followed the dominant pathway. At the highest loading rate, one frequently observed the following sequence of events, $\beta1/\beta5 \rightarrow \beta3/\beta5 \rightarrow \beta1/\beta2 \rightarrow \beta3/\beta4 \rightarrow \beta1/\beta2$ (reform) $\rightarrow \beta3/\beta4 \rightarrow \beta1/\beta2$, where $\beta1/\beta2$ ruptured but then re-formed prior to the rupture of $\beta3/\beta4$. As illustrated in Figure 2A and B, unfolding events at smaller $r_i$ tended to result in larger molecular extensions.

The nonequilibrium character of a pulling experiment decreased with decreasing $r_i$, and smaller $r_i$ resulted in smaller force drops after the unfolding force, $f_u$, is reached. Since the force applied by the spring to the end of the protein did not fall off as sharply at lower $r_i$, more of the protein was extended during an unfolding event. The smaller extensions following unfolding events at higher $r_i$ (relative to those observed at lower $r_i$) were responsible for the $\beta1/\beta2$ unfolding/refolding events mentioned above because the applied tension was very low after the initial rupture event (Figure 2A and B). At lower $r_i$ this situation no longer held because the initial unfolding event resulted in a chain extension that was a significant fraction of the chain’s contour length (Figure 2B).

Crowding Effects on Ub ($\Phi_c = 0.3$). Depletion forces stabilize proteins and shift the folding equilibrium toward more compact states.37 These forces result from an increase in the entropy of the crowding agents that more than compensates for an increase in the free energy of a protein molecule upon compaction. Simulations of forced unfolding of Ub in the presence of explicit crowders of diameters $\sigma_c = 6.4$ and 1.0 nm were used to assess the contribution of the depletion forces to mechanical stability. Sixteen trajectories were collected for each $r_i$ investigated. Three $r_i$’s ($20 \times 10^3$, $80 \times 10^3$, and $160 \times 10^3$ pN/s) were explored for the $\sigma_c = 6.4$ nm sized depletants. Only the two highest $r_i$ values ($80 \times 10^3$ and $160 \times 10^3$ pN/s) were explored for the $\sigma_c = 1.0$ nm sized crowders.

Small Crowding Particles Increase Unfolding Forces. Example FECs collected in the presence of crowders of diameter $\sigma_c = 6.4$ and 1.0 nm are illustrated in Figure 4A and B. The $\sigma_c = 6.4$ nm curves in Figure 4 look qualitatively very similar to those seen at $\Phi_c = 0.0$, while the FECs collected at $\sigma_c = 1.0$ nm look qualitatively different. For example, larger $\langle f_u \rangle$ than those observed at either $\Phi_c = 0.0$ or in the presence of the $\sigma_c = 6.4$ nm crowders are apparent in the FECs collected at $\sigma_c = 1.0$ nm (Figure 4A and B). Indeed, Figure 4C reveals that this observation is quantitatively accurate. Although the $\langle f_u \rangle$ in the presence of the $\sigma_c = 6.4$ nm crowders was statistically indistinguishable from the $\langle f_u \rangle$ at $\Phi_c = 0$ (compare Figure 4C with Figure 2C), the average unfolding forces in the presence of the $\sigma_c = 1.0$ nm crowders were statistically greater than those measured at $\Phi_c = 0.0$. At $r_i = 160 \times 10^3$ pN/s, the $\langle f_u(\Phi_c = 0.3) \rangle$ in the presence of the $\sigma_c = 1.0$ nm crowders exceeded that at $\Phi_c = 0.0$ by 3%, while at $r_i = 80 \times 10^3$ pN/s, the increase was 4% (compare Figure 4C with Figure 2C). In one respect, these results are not surprising; it follows from the AO theory.
and eq 8 that smaller crowders stabilize Ub more than larger ones. On the other hand, the extent of stabilization (as measured by increases in $\langle f_u \rangle$) was small. We should emphasize that although the increase in the unfolding force is small, the stability increase upon crowding is significant. The enhancement in the unfolding force can be accounted for by increases in $\langle f_z \rangle$.

Crowding Leads to Transient Local Refolding. The unfolding pathways in the presence of crowding agents of both sizes were very similar to those seen at $\Phi_c = 0.0$. Nevertheless, at $\Phi_c = 0.3$, there tended to be more unfolding/refolding events as the molecule extended past $z_u$ (Figure 3) than that at $\Phi_c = 0.0$. As illustrated, strand pairing between $\beta_1$ and $\beta_2$ and between $\beta_3$ and $\beta_4$ persisted to a greater extent after the initial unfolding event at $\Phi_c = 0.3$ than that at $\Phi_c = 0.0$. As Ub passed through the point $(z_u, f_u)$, its termini were extended to distances greater than the length of even the larger crowders. Depletion forces resulting from the presence of 6.4 nm crowders act on these larger length scales. Assuming that the tension applied to the C-terminus is small enough (e.g., after rupture events at higher loading rates), then such depletion forces can promote re-formation of contacts between secondary structural elements several times during the course of a trajectory. If this is indeed the case, then it suggests that unfolding polyUb may be different from the unfolding of monomeric Ub because depletion effects should increase with the number of modules in the tandem (see below).

**AO Model For Forced Unfolding in the Presence of Crowders.** We used the Asakura–Oosawa AO model\(^5\) (eq 4) of the depletion interaction to model the effects of a crowded environment on Ub. The AO theory has been successfully used to model the effects of a crowded environment on polymers.
and colloids, and we found that it gives qualitatively accurate results for Ub. We used eq 4 to approximate the effective interaction between spherical protein beads immersed in a crowded solution of volume fraction $\Phi_c = 0.3$. From the form of the AO potential between protein beads, it follows that the range of the potential is proportional to $\sigma_c$, but the strength is greater for smaller crowders. Indeed, as revealed in the previous section, simulations in the presence of explicit crowders showed that 1.0 nm crowders resulted in larger average unfolding forces than 6.0 nm crowders (Figure 4C).

Although the AO potential yields qualitatively accurate results, use of the AO potential of eq 4 to implicitly model nonbonded interactions in Ub did not yield quantitatively accurate results. At $r_f = 160 \times 10^3$ pN/s, $\langle f_u(\Phi_c = 0.0) \rangle = 175.78 \pm 1.52$ pN, while $\langle f_u(\Phi_c = 0.3) \rangle = 266.26 \pm 2.70$ pN. Thus, simulations with the AO potential led to a mean unfolding force that is roughly 50% greater than that in its absence. This disagrees sharply with the $\langle f_u(\Phi_c = 0.3) \rangle = 173.64 \pm 3.49$ pN resulting from our own simulations in the presence of an explicit crowding agent at $\Phi_c = 0.3$ and $r_f = 160 \times 10^3$ pN/s (Figure 4C). Indeed, the result also stands in marked contrast to the experimental results of Ping et al. on octameric Ub, which saw a maximum increase in $\langle f_u \rangle$ of 21% (at $\Phi_c > 0.3$, $r_f = 4200$ pN/s, and with $\sigma_c = 7.0$ nm) over the $\langle f_u \rangle = 166$ pN at $\Phi_c = 0.12$.

There are a couple of origins to the discrepancy. First, the AO potential was derived to understand the equilibrium of colloidal spheres and plates in the presence of smaller-sized spherical crowding agents. Our experiments were of a nonequilibrium nature; therefore, it is somewhat unreasonable to expect such simulations to yield quantitatively accurate unfolding forces or dynamics. Second, as pointed out by Shaw and Thirumalai, three-body terms are required to properly model depletion effects even in good solvents let alone in the poor solvent conditions of our simulations. To elaborate, let us consider the volume excluded to crowders by a Ub molecule, $V_{ex}(Ub)$, to be the volume enclosed by a union of spheres of radii $S_i = (\sigma + \sigma_c)/2$. With an AO potential, the volume excluded to the spherical crowding agents is

$$V_{ex}(Ub) = \sum_{i=1}^{N} V(S_i) - \sum_{j<i} V(S_i \cap S_j) \quad (7)$$

### Table 1: Module Unfolding Order Frequencies at $\Phi_c = 0$ and 0.3

<table>
<thead>
<tr>
<th>unfolding order</th>
<th>frequency observed at $\Phi_c = 0$</th>
<th>frequency observed at $\Phi_c = 0.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C $\rightarrow$ B $\rightarrow$ A</td>
<td>0.44</td>
<td>0.56</td>
</tr>
<tr>
<td>C $\rightarrow$ A $\rightarrow$ B</td>
<td>0.38</td>
<td>0.13</td>
</tr>
<tr>
<td>B $\rightarrow$ A $\rightarrow$ C</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>B $\rightarrow$ C $\rightarrow$ A</td>
<td>0.13</td>
<td>0.25</td>
</tr>
<tr>
<td>A $\rightarrow$ B $\rightarrow$ C</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>A $\rightarrow$ C $\rightarrow$ B</td>
<td>0.06</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Figure 7. Illustration of the stochastic nature of module unfolding. In (A), module C (the most proximal to the applied force) unfolds first, followed by module B, and finally by rupture of module A. In (B), the order of module unfolding events is A $\rightarrow$ C $\rightarrow$ B. The two trajectories illustrated here were both collected at $\Phi_c = 0$ and $r_f = 640 \times 10^4$ pN/s. Table 1 provides the frequencies at which the different possible orders were observed at $\Phi_c = 0$ and at 0.3. The diameter of the crowders is $\sigma_c = 6.4$ nm.
Despite these limitations, the AO model, which is simple, can be used to provide qualitative predictions. Finally, we note that experiments typically use polyproteins to study force-induced unfolding (e.g., Ping et al.12 used (Ub)₈ in their experiments). For polyproteins, the quantitative accuracy of the AO theory for unfolding in the presence of crowders of diameter $\sigma_c = 6.4$ nm is likely to increase because the volume excluded to the crowders will change significantly with changes in the configuration of the polyprotein.

$(\text{Ub})₃$ at $\Phi_c = 0.0$ and $0.3$. From arguments based on volume exclusion that lead to crowding-induced entropic stabilization of the folded structures, it follows that crowding effects should be more dramatic on polyUb than that on the monomer. In order to illustrate the effect of crowding on the stretching of $(\text{Ub})₃$, we chose $\sigma_c = 6.4$ nm, which had negligible effect on $\langle f(t) \rangle$ for the monomer. However, we found significant influence of the large crowding particles when $(\text{Ub})₃$ was forced to unfold in their presence. The FEC (Figure 5) shows three peaks that corresponded to unfolding of the three domains when simulations were performed at crowder volume fractions of $\Phi_c = 0.0$ and $0.3$ at $r_l = 640 \times 10^4$ pN/s. Figure 6 presents average unfolding forces as a function of the unfolding event number. Although the first two unfolding events were statistically indistinguishable at $\Phi_c = 0.0$ and $0.3$, the final event occurred at much larger $\langle f(t) \rangle$ in the presence of crowders than in their absence.

**Order of Unfolding was Stochastic.** $(\text{Ub})₃$ has three chemically identical modules. In the pulling simulations, the N-terminus of module A was held fixed, while force was applied to the C-terminus of module C. Figure 7 illustrates the time dependence of contacts between secondary structure elements of modules A, B, and C. It is clear from Figure 7A that module C unfolded first, followed by module B, and finally by module A. On the other hand, the order of events in Figure 7B was $A \rightarrow C \rightarrow B$. The frequency with which the $3! = 6$ possible permutations of these orders at $\Phi_c = 0$ and $\Phi_c = 0.3$ were observed is presented in Table 1. It is clear that at both $\Phi_c = 0$ and $0.3$, the most probable order of events was $C \rightarrow B \rightarrow A$. At $\Phi_c = 0$, this order was only marginally more probable than the order $C \rightarrow A \rightarrow B$, while at $\Phi_c = 0.3$, $C \rightarrow B \rightarrow A$ became overwhelmingly more probable than any other unfolding order. In none of the simulations at $\Phi_c = 0$ or $0.3$ did the rupture of module C occur as the final event.

**Unfolding Within a Module Depended on the Proximity to the Point of Force Application.** At $r_l = 640 \times 10^4$ pN/s, $(\text{Ub})₃$ is fairly brittle, and the rupture of contacts within a module occurs nearly simultaneously. Nevertheless, by carefully examining time-dependent contact maps such as those illustrated in Figure 7, we were able to determine that (1) at both $\Phi_c = 0$ and $0.3$ $\beta1/\beta5$ contacts were the first to rupture and (2) only for module C was this rupture event invariably followed by the loss of $\beta3/\beta5$ contacts. When other modules ruptured, loss of $\beta1/\beta5$ contacts was occasionally followed by loss of the $\beta1/\beta2$ strand pair contacts.

$(\text{Ub})₈$ Must Achieve a Larger $R_z$ to Rupture at $\Phi_c = 0.3$. Figure 8A illustrates the time dependence of $\langle R_g \rangle$ at $\Phi_c = 0$ and $0.3$. Interestingly, the plot reveals that after the second
rupture event, the \( \langle R_g(\Phi_c, t) \rangle \) increased more rapidly in the presence of crowders than in their absence. This is likely a reflection of the fact that at \( \Phi_c = 0 \) modules A and C were the first two modules to unfold in 44% of the trajectories, while at \( \Phi_c = 0.3 \), these two modules were the first to unfold in only 19% of trajectories. Thus, \( \langle R_g(\Phi_c, t) \rangle \) increased more rapidly in the presence of crowders because the \( R_g \) of (Ub)₃ with two adjacent modules unfolded is larger than that with two unfolded but nonadjacent modules. Interestingly, these differences are masked in the time-dependent increase of the end-to-end distance (Figure 8B). Figure 8A also reveals that the horizontal inflection points marking the third unfolding event occur at different times at \( \Phi_c = 0.3 \) and in the absence of crowders. The difference between these two times was responsible for the difference in average unfolding forces of \( \approx 14 \) pN illustrated in Figure 6. Thus, it is clear that despite the highly nonequilibrium nature of this pulling experiment, depletion effects were substantial, and it required a much greater force to reach \( f_u \) in the presence of these crowders than in their absence.

We predict that systematic experiments will reveal that polyUb molecules composed of larger numbers of modules will show greater increases in \( \langle f_u(\Phi_c) \rangle \) relative to \( \langle f_u(\Phi_c = 0) \rangle \) than polyUb molecules composed of fewer repeats. The size of these differences should increase with decreasing loading rate. Finally, it may even be possible to observe differences in the \( \langle f_u \rangle \) as a function of the unfolding event number (as in Figure 6). The increase in \( \langle f_u(\Phi_c) \rangle \), at a fixed \( \Phi_c \), for polyUb is likely to be even more significant for small crowding agents as shown in eq 8 (see below for further discussion).

Conclusions

General theory, based on the concept of depletion effects (see eqs 4 and 8), shows that crowding should enhance the stability of proteins and hence should result in higher forces to unfold proteins. However, predicting the precise values of \( \langle f_u(\Phi_c) \rangle \) is difficult because of the interplay of a number of factors such as the size of the crowding agents and the number of amino acid residues in the protein. Despite the complexity, a few qualitative conclusions can be obtained based on the observation that when only excluded volume interactions are relevant then the protein or polypeptide would prefer to be localized in a region devoid of crowding particles.⁴⁰ The size of such a region is \( D \approx \sigma_c(\pi/6\Phi_c)^{1/3} \). If \( D \gg R_g \), then the crowding would have negligible effect on the unfolding forces. The condition \( D \gg R_g \) can be realized by using large crowding particles at a fixed \( \Phi_c \). In the unfolded state, \( R_g \approx 0.2N^{1/3} \) Å,¹¹ which for Ub leads to \( R_g \approx 2.7 \) nm. Thus, \( D/R_g \approx 40\sigma_c \). These considerations suggest that the crowder with \( \sigma_c \approx 6.4 \) nm would have negligible effect on the unfolding force, which is in accord with the simulations. On the other hand, \( D/R_g \approx 0.4 \) when \( \sigma_c = 1 \) nm, and hence, we expect that the smaller crowders would have measurable effect on the unfolding forces. Our simulations are in harmony with this prediction. We expect that for the smaller crowding agent, \( \langle f_u(\Phi_c) \rangle \) would scale with \( \Phi_c \) in a manner given by eq 8. In general, the appreciable effect of crowding on the unfolding forces can be observed only for large proteins or for polypeptides using relatively small crowding agents. Although we have only carried out simulations for Ub and (Ub)₃ at one nonzero \( \Phi_c \), theoretical arguments can be used to predict the changes in \( \langle f_u(\Phi_c) \rangle \) as \( \Phi_c \) increases. The expected changes in the force required to unfold a protein can be obtained by using a generalization of the arguments of Cheung et al.¹³ In the presence of crowding agents, the protein is localized in a region that is largely devoid of the crowding particles.⁴⁰ The most probable size of the region is \( D \approx \sigma_c \Phi_c^{-1/3} \), where \( \sigma_c \) is the size of the crowding agent. If the structures in the DSE are treated as a polymer with no residual structure, then the increase in entropy of the DSE upon confinement is \( \Delta S/k_B \approx (R/cD)^{10/3} \), where \( R_c \) is the dimension of the unfolded state of the protein. If native-state stabilization is solely due to the entropic stabilization mechanism, we expect

\[
\langle f_u(\Phi_c) \rangle \sim T\Delta S/L_c = \left( \frac{R_c^{10/3}}{\sigma_c^{1/3}} \Phi_c^{1/3} \right) \left( \frac{k_BT}{L_c} \right)\]  

(8)

where \( f_u(\Phi_c) \) is the critical force for unfolding the protein, \( L_c \) is the gain in contour length at the unfolding transition, and \( v \) (\( \approx 0.588 \)) relates \( R_c \) to the number of amino acids through the relation \( R_c \approx 0.5N^2 \) Å (\( v \) varies between 2 and 4 Å). A few comments regarding eq 8 are in order. (1) The \( \Phi_c \) dependence in eq 8 does not depend on the nature of the most probable region that is free of crowding particles. As long as the confining region, which approximately mimics the excluded volume effects of the macromolecule, is characterized by a single length \( D \), we expect eq 8 to be valid. (2) The additional assumption used in eq 8 is that \( N \gg 1 \), and hence, there may be deviations due to finite size effects. (3) The equivalence between crowding and confinement breaks down at large \( \Phi_c \) values. Consequently, we do not expect eq 8 to fit the experimental data at all values of \( \Phi_c \). (4) It follows from eq 8 that, for a given \( R_g \), small crowding agents are more effective in stabilizing proteins than large ones. Thus, the prediction based on eq 8 is supported by our simulations. (5) From the variation of \( \langle f_u(\Phi_c) \rangle \) with \( \Phi_c \), Ping et al.¹² suggest that \( \langle f_u \rangle \sim \Phi_c \). However, the large errors in the measurements cannot rule out the theoretical prediction in eq 8. We have successfully fit their experimental results using eq 8 (Figure 8C). Additional quantitative experiments are required to validate the theoretical prediction.

It is difficult to map the concentrations in g/L used in the study of Ping et al.¹² to an effective volume fraction because of uncertainties in the molecular weight of dextran used in the study. Hence, a quantitative comparison between theory and experiments is challenging. A naive estimate may be obtained by using the values reported by Weiss et al.⁴² Given that the dextran used in the study is thought to have an average molecular weight of 40 kDa and an estimated average hydrodynamic radius of 3.5 nm,¹² we find that \( \rho = 300 \) g/L corresponds to a volume fraction of \( \Phi_c = 0.8 \), which is very large. Nevertheless, \( \Phi_c \) must be large when \( \rho = 300 \) g/L. Alternatively, we estimated \( \sigma_c/2 \) for dextran using \( \sigma_c/2 \approx a_0N^{1/3} \), where \( N \) is the number of monomers in a 40 kDa dextran, which is 0/2/0.162 \( \approx 147 \). If the monomer size is \( a_0 = 0.4 \)–0.45 nm, then we find \( \Phi_c \approx 0.3–0.4 \). If we assume that \( \rho = 300 \) g/L corresponds to \( \Phi_c = 0.4 \) and that \( \langle f_u(\Phi_c) \rangle \sim \Phi_c^{-2} \) (eq 8), then at \( \Phi_c = 0.3 \), we would expect nearly 18% increase in \( \langle f_u \rangle \). Similarly, if \( \rho = 300 \) g/L corresponds to \( \Phi_c = 0.3 \), then we would expect an increase in \( \langle f_u \rangle \) of approximately 21%. In any case, we can say that at physiologically relevant volume fractions (\( \Phi_c \approx 0.1,0.3 \)), the percent increase in \( \langle f_u \rangle \) is likely to be \( \approx 20\% \). Our simulations for \( \sigma_c = 1 \) nm predict an increase of \( 3–4\% \), which shows that a more detailed analysis is required to obtain an accurate value of \( \sigma_c \) for dextran before a quantitative comparison with experiments can be made. The larger increase seen in experiments may also be a reflection of the use of (Ub)₃ rather than a monomer.

Regardless of the crowder size, we find that the unfolding pathways are altered in the presence of crowding agents. It is
normally assumed that the rupture of secondary structure elements is irreversible if the applied force exceeds a threshold value. However, when unfolding experiments are carried out in the presence of crowding particles that effectively localize the protein in a smaller region than when \( \Phi_c = 0 \), reassociation between already ruptured secondary structures is facilitated as shown here. Thus, forced unfolding cannot be described using one-dimensional free-energy profiles with \( z_0 \) as the reaction coordinate.\(^3\,3\)

We find that the average unfolding force for the final rupture event of the unfolding of (Ub)\(_1\) occurred at much larger values in the presence of crowders than in their absence. With \( \alpha_t = 6.4 \) nm, which has practically no effect on the unfolding force of the monomer, and \( \Phi_c = 0.3 \), even with unfolding of two modules, the interactions between the stretched modules and protein are small (\( D \approx 1.2\alpha_t \)). Only upon unfolding of the third Ub do crowding effects become relevant, which leads to an increase in \( \langle f_d(\Phi_c) \rangle \). Our results suggest that \( \langle f_d(\Phi_c) \rangle \langle f_d(0) \rangle \) should increase with the number of modules in the array and that it may be possible to detect differences in \( \langle f_d \rangle \) which are conditional on the unfolding event number. We speculate that naturally occurring polyproteins that are subject to mechanical stress have evolved to take advantage of precisely such enhanced depletion effects.

**Acknowledgment.** This work was supported by a grant from the National Science Foundation (CHE 05-14056). D.L.P. is grateful for a Ruth L. Kirschstein Postdoctoral Fellowship from the National Institute of General Medical Sciences (F32GM077940). Computational time and resources for this work were kindly provided by the National Energy Research Scientific Computing (NERSC) Center.

**References and Notes**


JP807755B