Modeling Protein-Based Hydrogels under Force

Kirill Shmilovich and Ionel Popa

Phys. Rev. Lett. **121**, 168101 — Published 19 October 2018

DOI: 10.1103/PhysRevLett.121.168101

Modeling Protein-based Hydrogels under Force

Kirill Shmilovich, Ionel Popa*
Department of Physics, University of Wisconsin-Milwaukee,
3135 North Maryland Ave., Milwaukee, Wisconsin 53211, USA

Hydrogels made from structured polyprotein domains combine the properties of cross-linked polymers with the unfolding phase transition. The use of protein hydrogels as an ensemble approach to study the physics of domain unfolding is limited by the lack of scaling tools and by the complexity of the system. Here we propose a model to describe the biomechanical response of protein hydrogels based on the unfolding and extension of protein domains under force. Our model takes into account the contributions on the network dynamics of the molecules inside the gels, which have random cross-linking points and random topology. This model reproduces reported macroscopic visco-elastic effects and constitutes an important step toward using rheometry on protein hydrogels to scale down to the average mechanical response of protein molecules.

Protein-based hydrogels are a new type of material that retain the main characteristics of polymeric hydrogels, but show a unique visco-elastic response to stress. This response stems from the unfolding and extension of constituent protein domains. The appearance of the unfolding phase transition depends on the experienced force, exposure time, pulling geometry, the nature of the protein used to form the gel. Such a unique response of protein-based hydrogels to external stimuli does not only open new vistas toward designing new biological materials, but also enables a new spectroscopy technique to determine the mechanical response and energy landscape of single proteins from multi-molecule ensemble experiments of protein hydrogels. Rather than gathering single-molecule data through time-consuming atomic force microscopy (AFM) or optical and magnetic tweezers measurements, soft-matter rheometry can probe the force response of a massive number of interconnected proteins [1,2]. Rheometry techniques require the decoupling the force-induced (un)folding of individual proteins from the elastic response related to the cross-linked gel network, an experiment that recently became available through the introduction of force-clamp rheometry [3].

Here we propose a model that describes the macroscopic response of protein-based hydrogels obtained from polyproteins. This model is a critical step toward extracting the average unfolding and extension of single molecules from hydrogel stretching experiments.

Polyproteins have a cylindrical geometry and can be cross-linked into hydrogels using a photo-activated chemical reaction, where exposed tyrosine amino acids produce carbon-carbon bonds between adjacent polyprotein molecules [1]. Once the protein network is formed, its response to force can be analyzed through the dynamics of its cross-linking nodes. An elegant approach was introduced to model the network dynamics of actin filaments under a perturbing force vector [4-6]. While actin domains do not experience any unfolding or refolding transitions, the cylinder like geometry of actin filaments resembles that of polyproteins.

Unique to polyprotein hydrogels is the unfolding transition of constituent domains, which results in a significant increase in the contour length of the molecule. We chose to investigate hydrogels made from polyproteins (repeats of protein L), as this model protein has been extensively studied experimentally by the single molecule force spectroscopy community [2,7]. The domains in polyproteins are arranged as 'beads on a string'. This arrangement is an important characteristic of many proteins that have evolved to operate under force, such as titin in muscles [8] and talin in cellular mechano-transduction [9,10]. Furthermore, an energy landscape model for a polyprotein made of eight repeats of protein L was shown to reproduce the measured unfolding and refolding response of this protein to force, and was adopted herein [11,12]. This model combines the change in the barrier height between the folded and unfolded states due to an applied force with

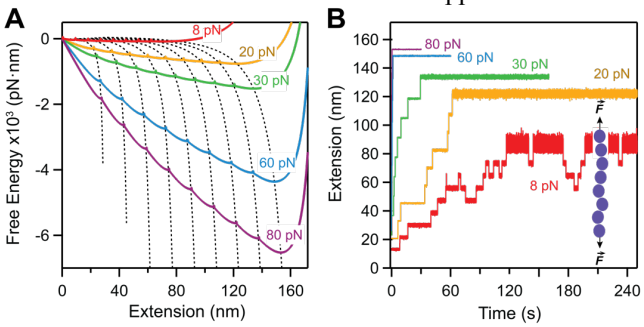

FIG. 1 (color online) **Mechanical response of polyproteins.** (A) Projection of the free energy landscape on the pulling coordinate at various forces for a polyprotein L with 8-repeats, showing an accordion-like shape. Dotted lines follow local energy minima. (B). Single molecule unfolding traces generated with the landscape from (A). Inset: Schematics of an eight-domain polyprotein under force.

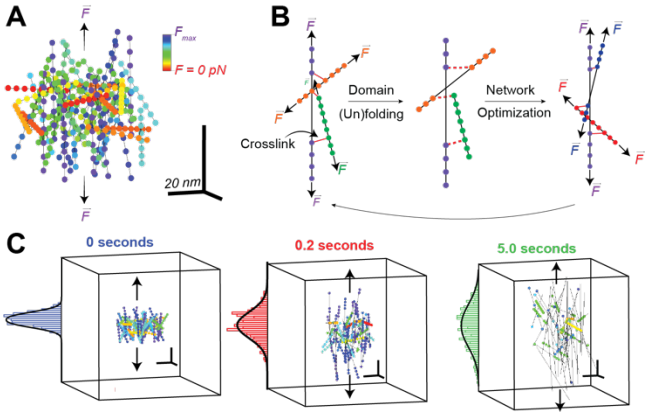

FIG. 2 (color online) **Protein hydrogels under force.** (A) Molecules are depicted as straight lines with 8 spheres along their axis, representing folded domains. A global constant force is applied to the z-axis. (B) Illustration of the unfolding dynamics and orientation change during gel stretching. (C) Snap-shots of the same gel at three different time points. Scale bars are 20x20x20 nm.

standard polymer elasticity models, which account for the entropic extension of the unfolded polypeptide chain (Fig. 1). For a single domain, a step up is seen during its unfolding and extension, while a step down is measured during its refolding and recoil. The energy landscape parameters were chosen to match the unfolding and refolding dynamics of protein L, which were measured experimentally with single molecule magnetic tweezers [13] (Table S1). Therefore, this landscape intrinsically accounts for interactions between the protein molecules and their surrounding solvent molecules.

As hydrogels are over 90% water rheometry[3], it is reasonable to assume that the dynamics of individual molecules inside hydrogels is the same as the dynamics of the polyprotein measured by force spectroscopy in solution. Our model ignores any intermediate states that characterize the folding process, as these are short-lived [14] when compared to the chosen sampling time. Furthermore, as most of the domains composing a molecule will not be part of the cross-link, it is safe to assume that they will experience the force along the polyprotein N-to-C backbone, which is the same pulling coordinate as in single molecule experiments. Those domains that are part of the cross-links will have a different stability [15], but their overall effect on the gel dynamics is limited, as they only partially extend between the cross-link
and either the N or C-terminus (see also below equation (4)).

In our model, each polyprotein L molecule is approximated to a rigid rod, with $m = 8$ domains of radius $r = 2\,nm$ each (PDB code *1hz5*), leading to a total contour length $L = 32\,nm$. This rigid rod structure was described experimentally as a characteristic of folded polyproteins using both electron microscopy [16] and small-angle X-ray scattering [17]. To form the gel network, each polyprotein molecule was assigned a center of mass, following a random distribution inside a square lattice, along with a random orientation. N polyprotein molecules were distributed within a volume $V = N/c$, where c is the molecule number density. The proteins were allowed to diffuse inside a rigid box of a volume $3V$ with a mean square displacement $\langle x^2 \rangle = 2D_t\delta t$ and to rotate with $\langle \varphi^2 \rangle = 2D_r\delta t$, while interacting elastically with the box walls [18,19]. The translation and rotation diffusion coefficients for a single polyprotein molecule are defined as:

$$D_t = \left[\frac{k_B T(\ln(p)+\alpha)}{3\pi\eta L}\right] = 3680\ \mathrm{nm^2 \mu s^{-1}} \qquad (1)$$

$$D_r = \left[\frac{3k_B T(ln(p)+\beta)}{\pi\eta L^3}\right] = 20.1\ \mathrm{\mu s^{-1}} \qquad (2)$$

with $p = L/(2r)$ being the shape factor for a rod, and α and β end-effect terms which account for hydrodynamic iterations expressed as second degree polynomials in polynomials in p^{-1} [18,19]. Crosslinking occurs if the center of mass of adjacent protein domains are within a threshold distance $||\vec{\boldsymbol{u}}_i - \vec{\boldsymbol{u}}_j|| \leq 2r$, where $\vec{\boldsymbol{u}}_i$ points to node i of the network.

When cross-linked, clusters of two or more molecules move in tandem with a diffusion coefficient of $D_t^{cluster} = \frac{k_B T}{6\pi\eta R_g}$, where R_g is the radius of gyration [20]. Our approach simulates a popular cross-linking method, which produces C-C bonds at the tyrosine sites [1]. As protein L has three tyrosine amino acids, up to three cross-links per domain were allowed. Following complete cross-linking, the network was allowed to equilibrate using a quasi-Newton algorithm that shifts the position of cross-linking nodes to minimize

$$H_{Crosslink} = K\sum_{<i',j'>}(2r - \|\vec{\boldsymbol{u}}_{i',j'}\|)^2 \qquad (3)$$

The primed indices restrict the summation to all valid inter-protein (i.e. cross-linked) node indices, with $K = 3.72 \cdot 10^5$ pN/nm being the force constant associated with the quadratic approximation of a C_n-C_n bond [21], and $\|\vec{\boldsymbol{u}}_{i,j}\|$ the perturbed bond length, such that $\vec{\boldsymbol{u}}_{i,j} = \vec{\boldsymbol{u}}_j - \vec{\boldsymbol{u}}_i$.

Under a constant force $\vec{\boldsymbol{F}}$ applied to the entire gel, the force experienced by a single molecule $\mathcal{k}$ is calculated based on its orientation of the molecule $|F_{\mathcal{k}}| = \vec{\boldsymbol{F}} \cdot \hat{\boldsymbol{u}}^{\mathcal{k}}$, with $\vec{\boldsymbol{u}}^{\mathcal{k}} = \sum_{i'} \vec{\boldsymbol{u}}_{i',i'+1}$ being determined by the nodes along molecule $\mathcal{k}$ (Fig. 2A). $F_{\mathcal{k}}$ determines the energy landscape on which each protein molecule diffuses at every moment, as the gel stretches (Fig. 1A). Force can lead to unfolding and extension of protein domains along polyprotein molecules (Fig. 1B). An unfolding or refolding event of a domain j on polyprotein molecule $\mathcal{k}$ extends/contracts the total end-to-end length $x_{\mathcal{k}}$ by an amount Δ along $\hat{\boldsymbol{u}}^{\mathcal{k}}$.

$$\vec{\boldsymbol{\ell}}_i^{\mathcal{k}} \xrightarrow{(un)folding} \begin{cases} \vec{\boldsymbol{\ell}}_i^{\mathcal{k}} \pm \hat{\boldsymbol{u}}^{\mathcal{k}}\,\Delta/2\,;\ i > j \\ \vec{\boldsymbol{\ell}}_i^{\mathcal{k}}\,;\ i = j \\ \vec{\boldsymbol{\ell}}_i^{\mathcal{k}} \mp \hat{\boldsymbol{u}}^{\mathcal{k}}\,\Delta/2\,;\ i < j \end{cases} \qquad (4)$$

Where '+' signifies an unfolding event and '–' a refolding transition. To understand how the (un)folding of single

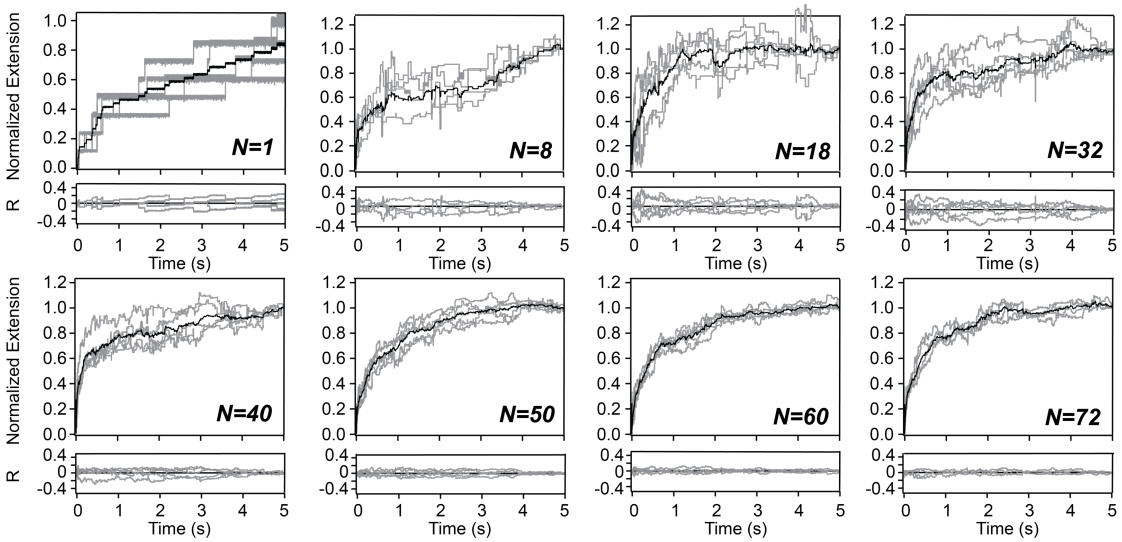

FIG. 3 **Scaling behavior of protein hydrogel extension at constant force.** Hydrogels composed of varying number of proteins (from *N = 1* to *N = 72*) at a concentration of $c = 9.0 \cdot 10^{-4}$ molecules/nm^{3}are subjected to an applied force of 50 pN for 5 seconds. Five normalized traces from separately polymerized hydrogels are simulated for each hydrogel size (grey traces) and averaged (black traces). The residuals, *R*, between the average trace and the individual traces are shown below each graph. *N* = 50 is the minimum number of molecules needed to produce a deterministic behavior.

domains perturb the entire gel network, we adopt a formalism used to describe actin gels [6], which minimizes the stretching and bending terms:

$$H_{Strech} = \frac{\mu}{2}\sum_{<i,j>}\frac{(\vec{u}_{i,j}\cdot\vec{r}_{i,j})^2}{|\vec{r}_{i,j}|^3} \quad (5)$$

$$H_{Bend} = \frac{\kappa}{2}\sum_{<i,j,k>}\frac{[(\vec{u}_{j,k}-\vec{u}_{i,j})\times\vec{r}_{i,j}]^2}{|\vec{r}_{i,j}|^5} \quad (6)$$

where $\vec{u}_i$ and $\vec{r}_i$ point to nodes in the perturbed and stable configurations respectively. The coefficients κ and μ are proportional to the persistence length ℓ_p, and related to the geometry of the network constituents [22]. For a rigid uniform rod-like polyprotein, $\kappa = \ell_p k_B T = 2.4\ \text{pN}\cdot\text{nm}^2$ and $\mu = \frac{4\ell_p k_B T}{\pi r^4} = 0.2$ pN, where $\ell_p = 0.58$ nm for protein L [2]. Each network node is shifted during network optimization to minimize the total energy $H_{Total} = H_{Crosslink} + H_{Bend} + H_{Strech}$ (Fig. 2B). This network optimization step was carried out only following unfolding or refolding events, as only then there is a significant change in the end-to-end length of a molecule inside the gel. Following each optimization step, the force F_k is recalculated to the closest integer, as determined by $\hat{u}^k$, and each molecule is transitioned to the corresponding energy landscape U_k (Fig. 1A). The total gel extension was estimated by projecting all the molecules on the stretching coordinate z, and fitting a higher-order Gaussian function[23]: $F_P(z,t) = A\exp[-(z/\gamma)^{2P}]$, where 2γ represents the gel length (Fig. 2C).

Under constant force, polyproteins show probabilistic unfolding events in single molecule experiments, resulting in a stair-case like extension, rather than one large step at a well-defined time [2]. As more proteins participate in the overall mechanical response, this probabilistic response is expected to be smeared out. Indeed, when increasing the number of molecules that are used to form the hydrogel network, we observe a decrease in the variance between individual extension traces and the average for a given force protocol (Fig. 3). In this case, five separately polymerized hydrogels composed of *N* = 1, 8, 18, 32, 40, 50, 60, and 72 proteins at a concentration $c = 9.0 \cdot 10^{-4}$ molecules/nm^3 (~1.5 mM) were exposed to a constant applied force of *F* = 50 pN for 5 seconds. The residual of the extension between the average and individual traces decreases with an increasing number of molecules, and stabilizes to ~4% for *N* ≥ 50. These results agree with the correspondence principle and tend toward the strictly deterministic behavior observed in tissues and biomaterials [24]. Indeed, rheometry measurements of protein hydrogels show very little change between measured elastic responses under identical conditions [3].

To investigate the extension of hydrogels to mechanical forces using our model, we simulated networks made of *N* = 50 molecules at a concentration $c = 9.0 \cdot 10^{-4}$ molecules/nm^3, which we find optimal in terms of probabilistic hydrogel response and computation time. Our simulations reproduce the measured behavior of protein hydrogels at constant force, which showed an initial elastic response, followed by a slower visco-elastic regime (Fig. 4A) [3]. As previously reported [25], a single exponential law describes poorly the visco-elastic regime. The macroscopic parameters of complex visco-elastic materials with multiple mechanisms underlying their force response can be obtained using the Maxwell–Wiechert model [26]: a parallel assemblage of separately parameterized springs and dashpot

Maxwell elements $x(t) = a_0 + \sum_i a_i e^{-\tau_i t}$, where τ_i is the corresponding rate.

Here, we also find that a two-term exponential model ($i = 2$) is sufficient to capture the hydrogel response (Fig. S4 and Table S2). Now we also have a clearer picture of what each exponential represents. The fast rate constant is dominated by the initial alignment of molecules to the applied force and the unfolding events taking place in molecules already aligned to the direction of the force vector (triangles in Fig. 4B). The slow rate constant, on the other hand, are dominated by individual protein unfolding events as network rearrangements take place less often (squares in Fig. 4B).

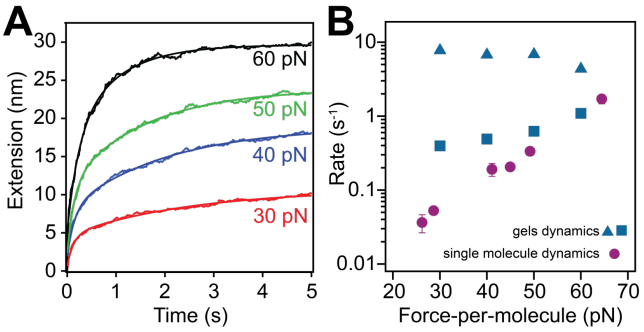

FIG. 4. (color online) **Force dependent behavior of protein-based hydrogels.** (A) Average of 15 different hydrogels traces of $N = 50$ molecules at a concentration of $c = 9.0 \cdot 10^{-4}$ molecules/nm^3, subjected to a constant force of 30, 40, 50, and 60 pN. (B) Comparison between rate constants of single molecule unfolding (circles) and rate constants from gels, determined by two-term exponential fits from (A), plotted as a function of the applied force. By correlating simulated and measured rates of protein hydrogels, one can determine the underlying average single molecule kinetics.

Protein hydrogels allow to sample several billion molecules in one pull, providing a simple experimental approach to measure the mechanical response of proteins. We anticipate that protein hydrogel experiments will enable the screening for bioactive compounds by studying the same biomolecules in different solution conditions, with tremendous speed and accuracy. While macroscopic engineering models developed to study the elongation of materials, such as steel, are currently being applied for hydrogels [27-29], proteins do not behave as simple Hookean springs [11]. Other fractal-based scaling approaches have been used, but they also do not account for the randomness of the cross-linking and of the molecular orientation, which are specific for the gelation process [30,31]. Our model provides the first theoretical framework to obtain the corresponding average mechanical (un)folding behavior of a single polyprotein, that would otherwise be determined from much more tedious single molecule experiments. To predict the mechanical response of single molecules, one would need to use this model to build a library of traces obtained with different unfolding and refolding barriers and correlate the experimentally measured gel dynamics with the rates obtained from simulations (Fig. 4 B).

Here we assume that the volume of the hydrogel is the same as the volume of the protein solution, before cross-linking, as it was recently reported experimentally for protein L in this concentration range [32]. But solvation forces can play a significant role, especially when moving a gel from a poor to a good solvent [33]. For some protein hydrogels, a change in gel volume was measured following the cross-linking reaction [34,35], which was estimated to a few pN per molecule [31]. Such changes in volume can be readily incorporated in this model, as they represent a simplified case, where the forces experienced by the proteins do not have a predefined directionality ($|F_{\not k}|$ = constant, Figure S5). However, this approach cannot currently account for aggregation effects, which can be avoided experimentally by working with sufficiently diluted protein solutions [3].

In summary, we propose a model to describe the mechanical properties of protein-based hydrogels. This model builds on an established approach which describes the unfolding response of polyproteins to a force along their end-to-end coordinate [11,12]. Our model assumes no breaking of covalent bonds and utilizes the constraint imposed by domain cross-linking to optimize the network dynamics. The network equilibrates at the cross-linking points, following unfolding/refolding events. As the number of molecules forming a gel is increased, this model successfully recovers the probabilistic to deterministic scaling behavior expected in aggregating the stochastic process of polyprotein (un)folding. Using the minimum number of molecules where the gel shows deterministic behavior ($N = 50$), we have investigated the force-dependency protein-based hydrogels to better understand the contributions from individual protein (un)folding events and the deformation mechanics of the crosslinked hydrogel network. Extension traces fit the multi-term exponential behavior commonly attributed to viscoelastic materials [26]. Our model can now explain from a molecular perspective the mechanical response of protein hydrogels from the unfolding and extension of constituent protein domains. The relative ease of applying the presented formalism to hydrogels of generic protein composition furnishes an exciting new approach to probe the nanoscale behavior of proteins and offers a new way to extract unfolding and refolding dynamics from hydrogel rheometry measurements.

This work was funded by Research Growth Initiative (Award No. 101X340), National Science Foundation, Major Research Instrumentation Program (Grant No. PHY-1626450), Greater Milwaukee Foundation (Shaw Award) and University of Wisconsin System (Applied Research Grant). K.S. also acknowledges funding from National Science Foundation (Grant No. UBM-1129056).

*popa@uwm.edu

[1] S. Lv, D. M. Dudek, Y. Cao, M. M. Balamurali, J. Gosline, and H. Li, Nature **465**, 69 (2010).
[2] I. Popa, J. A. Rivas-Pardo, E. C. Eckels, D. J. Echelman, C. L. Badilla, J. Valle-Orero, and J. M. Fernandez, J Am Chem Soc **138**, 10546 (2016).
[3] L. R. Khoury, J. Nowitzke, K. Shmilovich, and I. Popa, Macromolecules (2018).
[4] C. Heussinger and E. Frey, Phys Rev Lett **97**, 105501 (2006).
[5] C. P. Broedersz, M. Depken, N. Y. Yao, M. R. Pollak, D. A. Weitz, and F. C. MacKintosh, Phys Rev Lett **105**, 238101 (2010).
[6] C. P. Broedersz, X. Mao, T. C. Lubensky, and F. C. MacKintosh, Nature Physics **7**, 983 (2011).
[7] R. Liu, S. Garcia-Manyes, A. Sarkar, C. L. Badilla, and J. M. Fernandez, Biophys J **96**, 3810 (2009).
[8] H. Li, W. A. Linke, A. F. Oberhauser, M. Carrion-Vazquez, J. G. Kerkvliet, H. Lu, P. E. Marszalek, and J. M. Fernandez, Nature **418**, 998 (2002).
[9] A. W. M. Haining, M. von Essen, S. J. Attwood, V. P. Hytonen, and A. D. Hernandez, Acs Nano **10**, 6648 (2016).
[10] M. Yao, B. T. Goult, B. Klapholz, X. Hu, C. P. Toseland, Y. Guo, P. Cong, M. P. Sheetz, and J. Yan, Nat Commun **7**, 11966 (2016).
[11] J. Valle-Orero, E. C. Eckels, G. Stirnemann, I. Popa, R. Berkovich, and J. M. Fernandez, Biochemical and Biophysical Research Communications **460**, 434 (2015).
[12] J. Valle-Orero, J. A. Rivas-Pardo, and I. Popa, Nanotechnology **28**, 174003 (2017).
[13] J. Valle-Orero, R. Tapia-Rojo, E. C. Eckels, J. A. Rivas-Pardo, I. Popa, and J. M. Fernandez, J Phys Chem Lett **8**, 3642 (2017).
[14] R. Berkovich, J. Mondal, I. Paster, and B. J. Berne, The Journal of Physical Chemistry B **121**, 5162 (2017).
[15] H. Dietz and M. Rief, Phys Rev Lett **100**, 098101 (2008).
[16] T. E. Fisher, A. F. Oberhauser, M. CarrionVazquez, P. E. Marszalek, and J. M. Fernandez, Trends Biochem Sci **25**, 6 (2000).
[17] M. A. da Silva, S. Lenton, M. Hughes, D. J. Brockwell, and L. Dougan, Biomacromolecules **18**, 636 (2017).
[18] A. Ortega and J. García de la Torre, The Journal of Chemical Physics **119**, 9914 (2003).
[19] A. Einstein and R. Fürth, *Investigations on the theory of Brownian movement* (Dover Publications, New York, N.Y., 1956).
[20] I. Teraoka, in *Polymer Solutions* (John Wiley & Sons, Inc., 2002), pp. 1.
[21] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, Journal of the American Chemical Society **118**, 11225 (1996).
[22] L. D. Landau and E. M. Lifshitz, in *Statistical Physics (Third Edition, Revised and Enlarged)* (Butterworth-Heinemann, Oxford, 1980), pp. 333.
[23] A. Parent, M. Morin, and P. Lavigne, Optical and Quantum Electronics **24**, S1071 (1992).
[24] R. Day and V. Daggett, Proc Natl Acad Sci U S A **102**, 13445 (2005).
[25] O. Chaudhuri *et al.*, Nature Materials **15**, 326 (2015).
[26] W. Kaiser, ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik **58**, 524 (1978).
[27] M. A. Meyers and K. K. Chawla, *Mechanical Behavior of Materials* (Cambridge University Press, 2009).
[28] R. Kocen, M. Gasik, A. Gantar, and S. Novak, Biomed Mater **12**, 025004 (2017).
[29] B. Jozwiak, M. Orczykowska, and M. Dziubinski, Plos One **10** (2015).
[30] M. Pouzot, T. Nicolai, D. Durand, and L. Benyahia, Macromolecules **37**, 614 (2004).
[31] J. Wu *et al.*, Nat Commun **9**, 620 (2018).
[32] L. R. Khoury, J. Nowitzke, N. Dahal, K. Shmilovich, A. Eis, and I. Popa, J. Vis. Exp., e58280 (2018).
[33] F. Brochard and P. G. de Gennes, Macromolecules **10**, 1157 (1977).
[34] J. Fang *et al.*, Nat Commun **4**, 2974 (2013).
[35] N. Kong and H. B. Li, Biophysical Journal **110**, 40a (2016).