Mechanical unfolding of network nodes drives the stress response of protein-based materials

Joel Nowitzke[1], Sanam Bista[1], Sadia Raman[1], Narayan Dahal[1], Guillaume Stirnemann[2,*], Ionel Popa[1,*]

[1] – Department of Physics, University of Wisconsin-Milwaukee, 3135 N Maryland Ave, Milwaukee, WI 53211, US
[2] – PASTEUR, Département de Chimie, École Normale Supérieure, PSL University, Sorbonne University, CNRS, 75005 Paris, France
* – to whom correspondence may be addressed: guillaume.stirnemann@ens.psl.eu and popa@uwm.edu

Abstract

Protein-based materials synthesized from cross-linked folded proteins have untapped potential for biocompatible, resilient, and responsive implementations, but face challenges due to costly molecular refinement and limited understanding of their mechanical response. Under a stress vector, these materials combine the gel-like response of cross-linked networks with the mechanical unfolding and extension of proteins from well-defined 3D structures to unstructured polypeptides. Yet the nanoscale dynamics governing their viscoelastic response remains poorly understood. This lack of understanding is further exacerbated by the fact that the mechanical stability of protein domains depends not only on their structure, but also on the direction of the force vector. To this end, here we propose a coarse-grained network model based on the physical characteristics of polyproteins and combine it with the mechanical unfolding response of protein domains, obtained from single molecule measurements and steered molecular dynamics simulations, to explain the macroscopic response of protein-based materials to a stress vector. We find that domains are about ten-fold more stable when force is applied along their end-to-end coordinate than along the other tethering geometries that are possible inside the biomaterial. As such, the macroscopic response of protein-based materials is mainly driven by the unfolding of the node-domains and rearrangement of these nodes inside the material. The predictions from our models are then confirmed experimentally using force-clamp rheometry. This model is a critical step toward developing protein-based materials with predictable response and that can enable new applications for shape memory and energy storage and dissipation.

Significance

The mechanical response of biomaterials that have folded proteins as their primary network is dominated by the conformational changes of their nodes in response to local forces. Connecting the nanoscopic unfolding response of protein domains and underlying network dynamics to the macroscopic behavior of protein-based biomaterials is important for mainstreaming these biomaterials, with potential applications for artificial tissue engineering, active delivery systems, biorobots and efficient mechanical batteries. The model proposed here can predict the deterministic macroscopic behavior of protein-based biomaterials from the nanoscopic probabilistic response of individual molecules. The approach shown here for a model alpha-beta polyprotein is universally applicable to any other network building block and represents an important step toward improving the design and mainstreaming discoveries relying on protein-based biomaterials.

Keywords: network science, protein unfolding, soft materials, mechanical response of biomaterials

INTRODUCTION

Protein-based materials synthesized from covalently cross-linked folded proteins are finding transformative bio-inspired applications, as they can harvest from the broad functionality and the unique mechanical response of their structural units. Several implementations of biomaterials using globular folded proteins as their cross-linking unit have been reported. These implementations can replicate the nonlinear elastic response of muscles [1, 2], can respond to chemical and optical stimuli [3, 4] and can provide unique shape-memory mechanisms [5-7]. The main draw-back of these approaches as opposed to polymer-based materials is that molecular refinement through trial-and-error is too costly and more time-consuming. Furthermore, their response to a stress vector is complicated by the fact that the building unit – folded proteins – can unfold under force. Similar to the untying of a knot, unfolding of a single protein domain results in addition of contour length several times its initial diameter. Here we incorporate domain unfolding with network optimization to decipher the intricate response of protein-based materials to stress.

A network produces a simplified representation by prioritizing the fundamental patterns of connection over the intricate details of individual elements. While high dimensionality poses computational limits [8], network physics and dynamics simulations have been successfully employed to develop desirable mechanical properties in colloidal gels [9], polymer networks [10-12], and fibrillar and globular protein cross-links [13-16]. However, to understand the macroscopic response of protein-based materials from the nanoscopic response of its building blocks requires knowledge of the molecular response of proteins to force [2, 17]. Single molecule force spectroscopy measurements have shown that typically globular proteins require forces larger than 10 pN, and in some cases over 100 pN, to unfold under a force vector when pulled along their termini direction [18]. However, the direction of the force vector may play an important role [19-21]. Achieving such elevated forces-per-molecule within materials requires either a low protein density, to reduce the load-bearing molecules per cross-sectional area, or exposure of these materials to relatively high stresses without compromising their structure. For protein-based materials, the minimum concentration at which proteins can be turned into a cross-linked stable network is quite high, typically in the mM range [22]. Furthermore, the maximum strain at which protein-based materials yield is relatively small, below a few tens of kPa, as these materials are relatively soft [23]. And when considering cross-linking density and molecular architecture, the estimated forces-per-molecule achievable before these materials yield is typically below 10 pN [23]. Based on these arguments, one would expect that protein domains will never unfold under stress inside a protein-based material and would act as simple rigid structures.

Several studies using various globular proteins have shown that protein-based materials show a visco-elastic behavior under stress [1, 2, 22]. This visco-elastic response is coming from the hysteresis that appears in stress-strain curves, where the extension and contraction curves do not overlap on top of each other. While naively one would associate such a behavior with network re-alignment and sliding of incompletely cross-linked domains, as is the case for polymeric-based materials, several experiments can be done to assess if protein domains are folded following cross-linking inside the biomaterial and if these domains unfold when stress is applied. For example, X-ray and fluorescence-based methods showed that the protein domains remain folded inside the biomaterials following the cross-linking reaction [22, 24]. But while domains are folded inside the network, do they unfold under stress? To answer this question, scientists took advantage of the fact that protein domains can be chemically unfolded under specific salt conditions, such as concentrated guanidinium hydrochloride [22, 25]. Hence one would expect that if network rearrangement or domain sliding are behind the measured visco-elasticity, that these materials will maintain their hysteresis in stress-strain curves even in the presence of chemical denaturants. Interestingly, addition of chemical denaturants results in a complete disappearance of the viscoelastic behavior, transforming these materials to purely elastic [22, 25]. This change further indicates that predominantly, if not exclusively, the viscoelasticity is a direct consequence

of mechanical unfolding of proteins. Furthermore, this unstructured architecture can be used to develop materials with extreme elastic resilience [16] or high toughness, resembling cartilages [25].

Hence, on one hand, it was proven experimentally that under stress, protein-based materials show a viscoelastic behavior that is directly related to the mechanical unfolding of protein domains; on the other hand, these stresses should not result in molecular forces large enough to unfold and extend proteins. So how can the two conflicting results be reconciled? To answer this question, here we combine a myriad of experimental and computational approaches to study the macroscopic behavior of protein-based materials from the coarse-grained response of its nanoscopic components. We find that the unfolding response of domains inside protein-based materials is mainly driven by the unfolding of node domains, which experience force on different coordinates than along their termini and are mechanically less stable. We then validate this model by comparing our predictions to the creep response of protein-based materials.

RESULTS

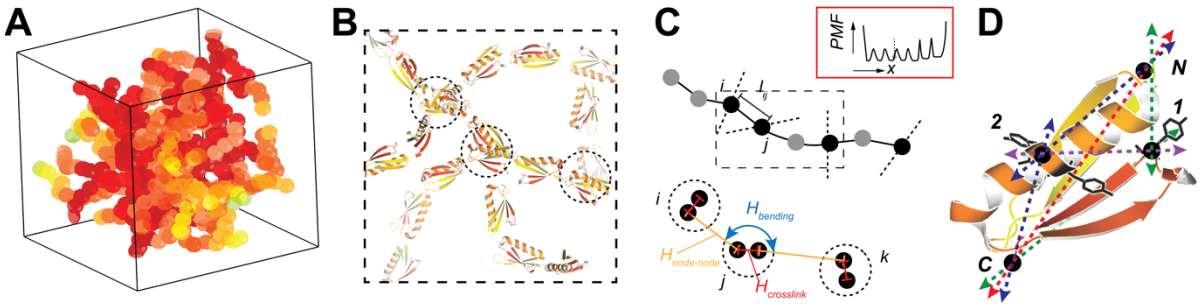

Figure 1. Schematic view of the coarse-graining approach with network optimization to study the mechanical response of protein-based materials made from globular polyproteins. A) Schematics showing randomly crosslinked polyproteins made of eight repeats. B) Detail molecular view showing possible crosslinking conformations between polyproteins. C) Representation of the two-steps approach for finding the response of protein-based materials with force. (top): one step consists in estimating the end-to-end extension of each molecule inside the material based on the number of tethered domains, number of nodes and experienced force. Inset shows the potential-of-mean-force (PMF) for a molecule where six domains are under force, out of which four are nodes (and hence have weak geometries). (bottom): next step is to minimize the network energy by searching for a minimum potential that considers extension between nodes ($H_{node\text{-}node}$), separation between crosslinks forming the nodes ($H_{crosslink}$) and bending energy ($H_{bending}$). D) Ribbon-diagram of a single protein domain showing possible tethering geometries based on the cross-linking reaction (at sites 1 and 2, or along the backbone N/C). The orientation of the force vector is expected to play a key role in domain stability, influencing the overall mechanical response of protein-based materials.

Overview of the approach.

Protein-based materials made from globular proteins and cross-linked at specific sites can be visualized through different levels of magnifications (**Figure 1**). At a multi-molecular coarse-grain level, diffusing molecules interact with each other and form covalent connections, limited by steric interactions and the number of cross-linkable sites (**Figure 1A)**. For the specific case considered in this study, polyproteins engineered with eight repeating domains of the B1 domain of protein L (referred from here-on as Protein L_8) can form carbon-carbon bonds at exposed tyrosine amino acids, when cross-linked through a well-established photo-activated reaction [1]. Any protein L domain along the polyprotein molecule can become a network node through cross-linking at one

or more exposed tyrosine amino acid site (**Figure 1B**). Here we will follow a two-step tick-tock simulation approach. In one step, the unfolding of each molecule inside the network is simulated through Brownian dynamics (BD), which takes into account its unique crosslinking pattern and current experienced force. The experienced force results from the molecular orientation in respect to the stress vector (**Figure 1C top**). In the following step, the network nodes are repositioned to minimize the overall energy (**Figure 1C bottom**). The energy minimization algorithm considers the energy of bending between three connected nodes, $H_{bending}$, the elongation energy between nodes, $H_{node\text{-}node}$ and the energy penalty of extending the covalent bonds linking two domains inside a node, $H_{crosslink}$ (see **Methods** section for details) [13, 15]. Importantly and novel, the approach considered here takes into account both the force and tethering geometry of each protein domain that is part of the network. The experienced force of a protein domain results from the relative orientation of a molecule to the stress vector, while its stability depends directly not only on this force, but also on the way the domain is connected. For the case of protein L, which can make two additional connections in the middle of the structure, a total five tethering orientations are possible (**Figure 1D**).

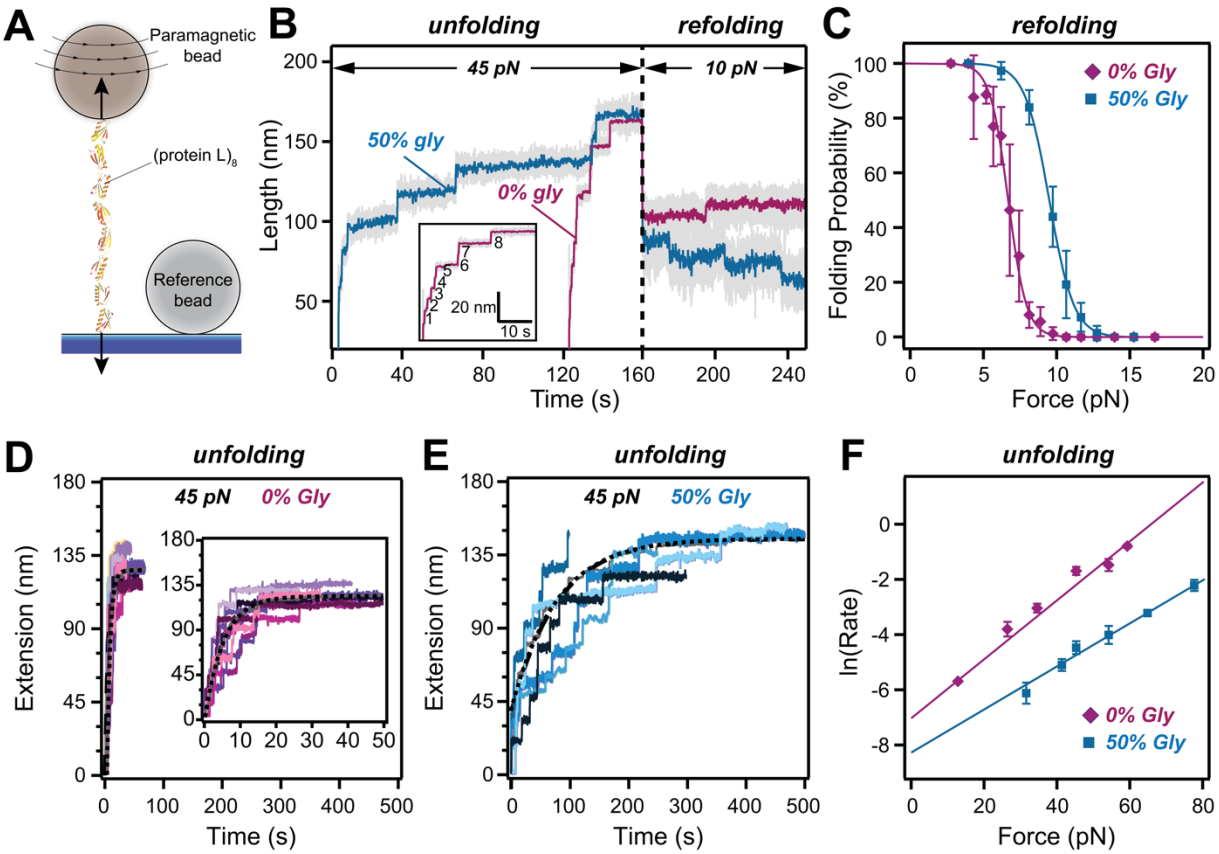

Figure 2. Single molecule unfolding measurement of polyprotein L. A) Schematics of the single molecule magnetic tweezers experiment, showing a octameric polyprotein L tethered between a glass surface and a paramagnetic bead. A non-magnetic reference bead is used to evaluate the location of the end attached to the glass surface and to correct for potential focal drift. B) Trace of the same molecule showing the unfolding and refolding behavior of protein L in regular buffer (magenta) and in glycerol 50% v/v (blue). Glycerol increases the mechanical stability of the protein and shifts its refolding probability. C) Refolding probability of protein L as a function of force. D) Single molecule traces of polyprotein L in regular buffer at a constant force (45 pN) shows the step-like increase in the measured extension, with each step representing the unfolding for a single protein L domain along the N-to-C coordinate. E) Single molecule

traces at the same constant force of 45 pN displaying a slower unfolding kinetics. The dotted lines in panels D and E represent averages, F) Variation in the unfolding rate as a function of force and buffer for protein L.

The mechanical response of protein L along the N-to-C coordinate

To gain insights into the molecular mechanism driving unfolding, we start our journey with characterizing the mechanical unfolding response of single polyproteins made from eight repeats of protein L and exposed to different forces and two solvent conditions: regular buffer and 50% glycerol. In the sampled force-range the proteins used for tethering do not unfold/refold [26] and our polyprotein constructs of eight repeating units not only enable better statistics, but also produce a unique molecular fingerprint. It is well known that osmolytes such as glycerol increase the mechanical stability of proteins [27, 28] and can promote the polymeric collapse of an unfolded polypeptide [29]. This response can be used as a tuning 'knob' to relate nanoscopic unfolding with macroscopic stress response. We use single molecule magnetic tweezers, which tether the protein of interest between a glass surface and a paramagnetic bead (**Figure 2A**). Due to the intrinsic stability of magnetic tweezers, the same molecule can be exposed to repeated force cycles and solvent conditions. **Figure 2B** shows the unfolding and refolding behavior of a polyprotein L molecule in regular buffer and in the presence of 50% glycerol. In the first part of the pulse, a high force is applied, leading to unfolding events, which are measured as step-increments in the molecular extension. In the second part of the pulse, the force is decreased, leading to domain refolding as step-contractions. The folding probability as a function of force can be determined by comparing the number of refolding domains during the low-force pulse with the total number of unfolding domains from the first part of the pulse (**Figure 2C**). For protein L, we found that the half force is 6.7 ± 0.1 pN in regular buffer and increases to 9.5 ± 0.2 pN in buffer with osmolyte. The first part of the pulse was also used to determine the unfolding kinetics of protein L (**Figure 2 D-F**). The unfolding rate was estimated from many single molecule traces at several applied forces, and a linear dependence following Bell's law was seen, as also reported previously [26]. It is obvious when comparing panels **D** and **E** in **Figure 2** that glycerol has a stabilizing effect. The half-folding-force and unfolding kinetics measured from single molecule magnetic tweezers will be used next to determine the potential of mean force for polyprotein L.

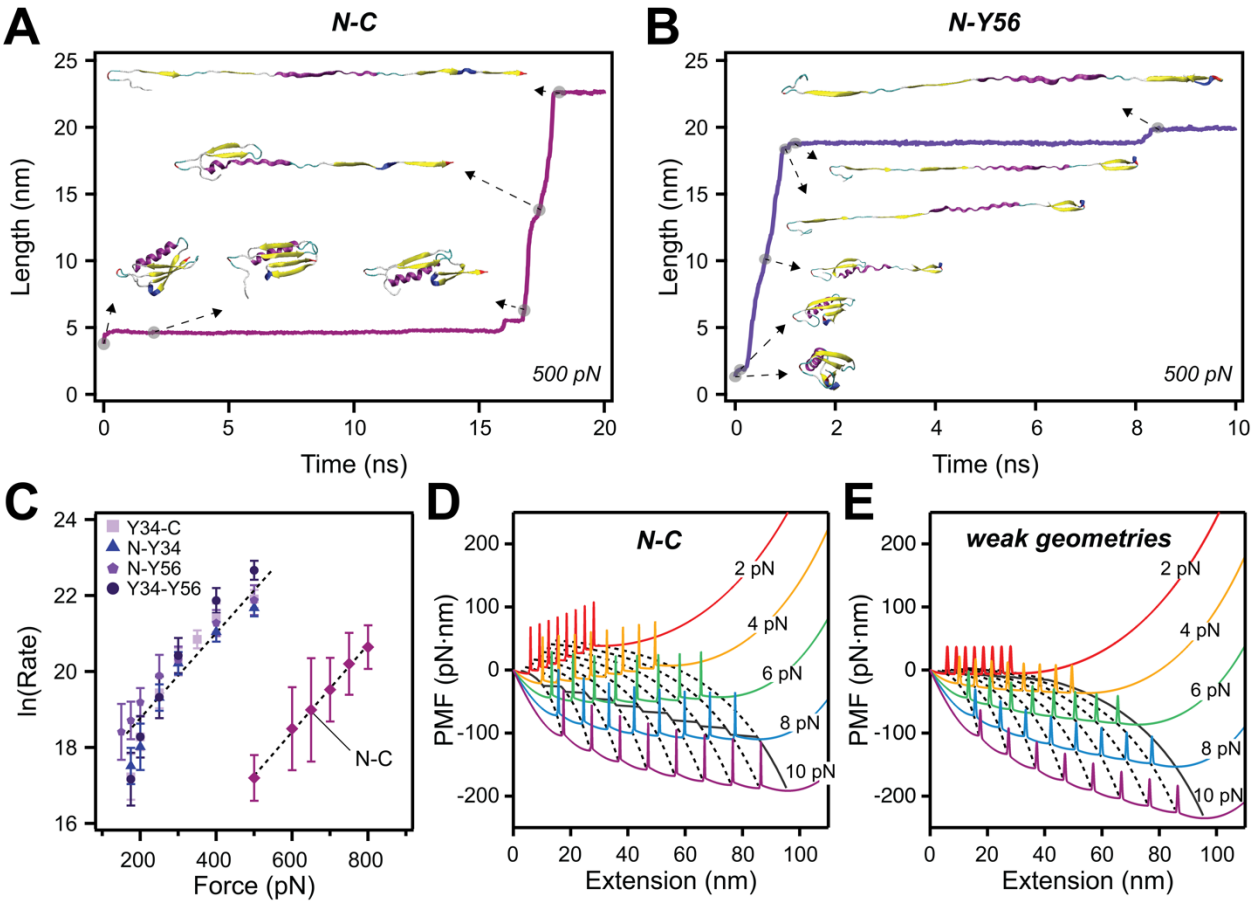

Figure 3. The energy landscape of (protein L)$_8$ as a function of pulling geometry. A) Trace obtained with steered molecular dynamics (SMD) simulations showing the unfolding path of a protein L domain pulled over its N-to-C coordinates at a constant force of 500 pN. B) SMD trace for a protein L domain pulled along the N terminus and tyrosine 56 at a constant force of 500 pN. C) Unfolding rate of protein L as a function of force and pulling geometries. When pulled along its N-to-C coordinate, protein L is significantly more stable. D) Underlying potential of mean force (PMF) reconstructed from measurements and simulation data, for polyprotein made of eight domains and tethered along the N-to-C coordinate, as a function of experienced force. E) Average PMF for the same polyprotein pulled along its weak geometries. The black line in panels D and E shows the location of the global minimum.

The effect of tethering geometry on the mechanical stability of protein L

The direction of the force vector was shown to play as an important role as its magnitude, using enzymatic linking [19] or cysteine mutants [30]. While single molecule approaches can readily assess the mechanical stability of proteins pulled along their N-to-C coordinate, we currently do not have an experimental approach to study the stability along the other possible coordinates when a protein domain is tethered through one or two possible tyrosine-crosslinking sites. To mitigate this limitation, here we employed steered molecular dynamics simulations (SMDS). Single protein L domains were exposed to different constant forces and the extension between the tethering points was recorded as a function of time (**Figure 3A-B** and **Supporting Figure S1, Supporting Movies S1-5**). The unfolding rate was then estimated as a function of pulling geometry (**Figure 3C**). From these simulations, we deduced that protein L is much more stable when force is applied along its N-to-C coordinate than along all the other possible tethering coordinates. Interestingly, the other geometries do not seem to show a large variation among themselves in terms of mechanical

stability. A similar behavior was seen when protein L was unfolded in the presence of glycerol 50% (**Supporting Figure S2 and S3**). As reported previously, molecular dynamics simulations can capture the stabilizing effect of proteins in the presence of osmolytes [29, 31]. In the current case, the unfolding rates too were slower when pulled on the other geometries as opposed to the N-to-C coordinate. Furthermore, the SMDS simulations reproduce the stabilizing effect measured experimentally for protein L in the presence of glycerol, albeit on a different range of forces. Taken together, we can use the SMD simulations to determine the stability ratio between N-to-C and other geometries, while the single molecule experiments using magnetic tweezers produce a direct measurement for the unfolding barriers when force is applied to the protein L termini. From the rates, we find that the barrier of the N-to-C geometry is 1.6 times higher than for the other tethering directions. Using this information, we can produce a potential of mean force (PMF), which represents the energy landscape of the polyprotein L along the pulling coordinate (**Figure 3D** and **E**). As the non N-to-C geometries have similar unfolding rates, we chose to group them together. Due to this similar stability of non-N-to-C geometries, we expect that once part of a domain unfolds, the adjacent part will denature as well. Given that most nodes experience force on both sides, we can assume a constant contour length for these domains. This non-N-to-C energy landscape is constructed as a free energy projection along the pulling coordinate following principles reported earlier [32, 33], where the barriers separating the folded and unfolded states for each domain are spaced apart by the entropic extension of an unfolded chain at a given force. (see **Methods** section for details). Then, Brownian dynamics simulations were used to determine the barriers and diffusion coefficients, that best describe the measured unfolding kinetics (**Supporting Figures S3** and **Tables S1**). While Brownian dynamics simulations do not have the molecular details conferred by SMDS, they allow to easily generate predictable unfolding patterns of many molecules over timescales of seconds, which is desirable given the relatively small force (and slow unfolding rate) experienced inside protein-based materials by individual molecules.

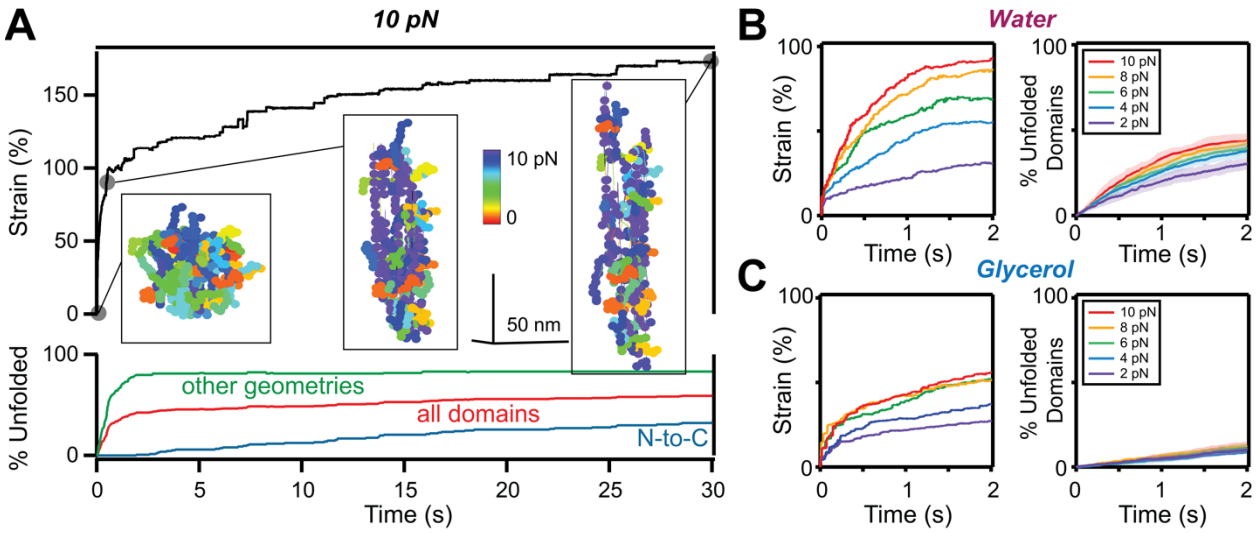

Figure 4. Corse-grained diffusion simulations and experimental comparison of tissue-like materials made from globular polyproteins. A) (top): Change in strain as a function of time under a force of 10 pN per molecule obtained from a gel of 75 molecules. (bottom): Percentage of unfolded domains tethered in N-to-C and in other geometries, and percentage of total domains unfolded. Insets: Snapshots of the gel at three different times. B) Strain vs time (left) and percentage unfolded domains under force (right), of a network made from 75 molecules as a function of force in water. C) Strain vs time (left) and percentage unfolded domains under force (right), of a network made from 75 molecules as a function of force in 50% glycerol.

Predicting mesoscopic response from network dynamics and nanoscopic protein unfolding

Our next task is to use the newly gained characterization of the behavior of polyprotein L at different forces, solvent conditions, and tethering geometry to predict the dynamics of protein-based materials under a stress vector. For this task, we used a coarse-grained diffusive approach to produce networks made from 75 molecules of octameric protein L, where each domain was represented by a hard-sphere [23]. This approach produces networks considering protein concentration, molecular flexibility, and maximum number of cross-links per domain [23, 34]. Once a network was obtained, its dynamics was simulated at a given force, using a two-step procedure. In one step each molecule inside the network was left to diffuse for 100 μs, on a specific PMF which considers the currently experienced force, the total number of tethered domains and the number of network nodes (**Figure 1C Inset**). The new molecular coordinates coming from Brownian diffusion were then used to generate a new frame for the network, which was then followed by a network optimization step. Following the network energy minimization step, the new position of all the nodes was used to reposition all other domains inside the network space and the PMF of each molecule was updated from their new orientation to the force vector, which produces a new local molecular force. These two steps were repeated for a desired number of times to produce a molecular film of the protein network under a force vector. The strain was then estimated for the location of the terminal tethered domains. **Figure 4A** shows a 30 s simulation of a network exposed to a 10 pN force. The strain has three kinetic regimes (**Figure 4A top**): an initial step increase in the first few microseconds, then a rapid network extension in the first few seconds and finally a slower creep behavior thereafter. The initial rise in strain can be attributed mostly to network rearrangement, where the nodes respond to the force vector to minimize the network energy, without interference from unfolding. The last two regimes can be correlated with unfolding of domains that are experiencing force over the weak geometries and over the N-to-C coordinate, respectively (**Figure 4A bottom**). Indeed, the weak domains, which are also part of the network nodes, unfold in an overwhelming percentage at this force in the first few seconds, while the unfolding and extension of the N-to-C domains is much more slow-paced (see also **Supporting Figure S4**). In a domino-like effect, as more and more domains unfold and extend under force, leading to more and more molecules being oriented to the force vector, the network keeps on extending over the course of the simulation.

This approach can then be used to determine the average strain as a function of overall stress, as well as the percentage of domains unfolding inside the network (**Figure 4B** for water and **Figure 4C** for 50% glycerol). The forces used in these simulations were relatively low, between 2 to 10 pN, as such forces were previously estimated for the given concentration needed to produce these protein-based materials [23]. When quantifying the number of unfolding domains based on their geometry, an interesting behavior is observed: most of the domains showing unfolding are tethered on non-N-to-C geometries. These simulations predict that the strain is very sensitive to the applied stress. As glycerol has a strengthening effect on the stability of protein domains, the overall extension of the simulated networks in osmolyte is smaller than in water, while the percentage of unfolded domains is smaller as well (**Figures 4B and 4C**).

Experimental validation of the proposed model

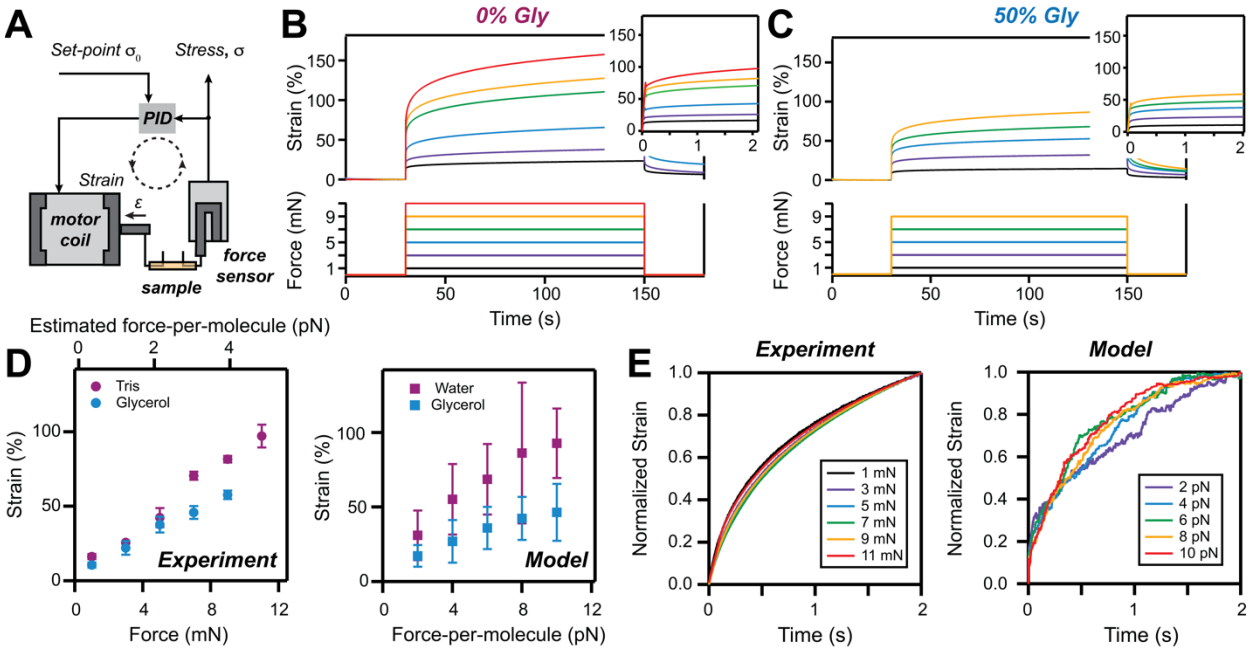

Figure 5. Comparing predicted behavior with experimental results. A) Schematics of the custom-made force-clamp setup developed to study the creep response of protein-based materials. B) and C) Strain measured at constant force (creep behavior) for materials made from 1 mM polyprotein L8 and exposed to different stresses in regular buffer (B) and 50% glycerol (C). Insets show the first two seconds. D) Comparison of average strain after 2 s for the two conditions from experiments (left) and model (right). Bottom-left axes shows the force to which the protein-based materials with 280 μm radii were exposed to, and top-left axes the estimated force-per-molecule. E) Normalized strain at constant stress, showing no significant behavior with force change in both experiments (left) and model (right). Traces were measured/simulated in tris/water. A similar behavior is seen with glycerol conditions (**Supporting Figure S5**).

In order to test the predictions of our model, we used force-clamp rheometry, a technique which was specifically developed to study soft materials such as those investigated here [22, 35]. First, a protein solution was turned into a biomaterial using transparent tubes as a template. The cylindrically shaped protein-material was then extruded from the cast and attached between a motor and a force sensor (**Figure 5A**). An analog proportional-integral-differential (PID) system compared the setpoint force, generated as a step-pulse in time, and adjusted the position of the motor such that the measured force at the opposite end matched this setpoint. The response time of the loop is 30 ms – 100 ms, depending on the applied force and material compliance [22]. Force can then be transformed into engineering stress by dividing by the cross-section area of the cylinder, and in average-force-per-molecule using a simple modeling procedure [23] Following the initial extension due to the change in force (at 30 s) the protein-material extends in a step like fashion (**Figure 5B and 5C)**. As predicted by our model, the initial network rearrangement is much faster than the response time of the PID loop, and hence no kinetic information can be extracted for this part of the curve. Then the strain traces show a two-regime behavior, both for regular buffer and glycerol. When comparing the final strain after a given time, both experiments and simulations show an increase with force, as expected (**Figure 5D**). Furthermore, in glycerol, the strain is consistently lower than in water conditions, and this behavior can be related to the fact that osmolytes increase the mechanical stability of protein domains. An interesting behavior measured for protein L, which we previously also reported for BSA-based protein materials [22], is that under constant force, the creep rate seem time independent when the strain is normalized

(**Figure 5E left)**. This behavior is also reproduced by our model (**Figure 5E right**) and comes from the fact that the kinetic response is dominated by the network rearrangement, while the final strain increases with the applied stress. Hence, due to the continuous orientation of more-and-more molecules to the force vector, the extension kinetics does not change significantly with force.

DISCUSSION

One of the fascinating things when it comes to materials made from folded proteins is how complex their response to force can be [36, 37]. Unlike approaches using man-made polymers or filamentous proteins, these biomaterials have a unique architecture, as their network nodes are represented by folded protein domains. The protein unfolding response leads to release of hidden contour length and is highly dependent on the value and orientation of the experienced force vector. Man-made materials synthetized from folded proteins rely on both enzymatic and radical-driven cross-linking reactions [1, 38]. These biomaterials produce unique molecular architectures where domain unfolding can play a pivotal role, and have been so far used for artificial shape morphing [5, 7], cell and artificial tissue growth [39, 40] and mechanically active [25, 41] and bioactive materials [42, 43]. Similarly, naturally occurring structures have also evolved to take advantage of the unfolding response of polyproteins. At its core, the extracellular matrix (ECM) is formed by randomly cross-linked polyproteins [44]. It is continuously shaped and remodeled by different types of cells and its stiffness and porosity regulates cellular function, including division and apoptosis [45, 46]. One of the many polyproteins forming ECM, fibronectin, is segregated in multi-domains and crosslinked at various sites [47]. An interesting behavior comes from mechanical unfolding of fibronectin by cells, which exposes the cleavage site for a growth factor, to further activate cellular adhesion [48]. Another biological system where cross-linked polyproteins play important roles is represented by muscles. The basic contractile unit of striated muscles, the sarcomere, is formed by the actomyosin assembly in parallel with titin, the largest protein in the human body. It was recently discovered that titin, a polyprotein formed from over 100 domains in series [49], uses the unfolding response to drive muscle contraction and regulate muscle elasticity [50]. Hence, understanding how the nanoscopic unfolding response translates in the macroscopic viscoelasticity is critical toward gaining a mechanistic view on the mechanical response of protein-based networks.

Here we take a multi-scaled approach, which combines the nanoscopic unfolding response of proteins as a function of the value and orientation of a force vector with network dynamics (**Figure 1**). As building blocks for our biomaterials we chose the B1 domains of protein L, a model bacterial protein containing both alpha-helices and beta-strands [51, 52]. Protein L is easy to express and purify in large quantities using E.coli competent cells, and has been previously studied extensively with single molecule approaches, such as atomic force microscopy (AFM) [53-56] and magnetic tweezers [34, 55, 57, 58]. In our protein construct, protein L was repeated eight times to provide a unique unfolding fingerprint in single molecule measurements (eight equidistant unfolding steps) and to allow for complete crosslinking when synthetizing biomaterials, as a single domain would only have two potential binding sites, versus16 potential sites in the polyprotein. Furthermore, due to its antibody binding capabilities [57, 59, 60], biomaterials made from protein L have the potential of finding industrial applications in purifying biological drugs and antibody-bound proteins-of-interest, pending the resolution of a few technical aspects addressing porosity and permeability [61].

Our first step was to measure the unfolding kinetics of single molecules made from eight repeats of protein L. At single molecule level, protein unfolding depends on the applied force and is driven by thermal fluctuations, which makes it a probabilistic process [62]. Furthermore, with the unfolding of every domain the behavior of the remaining folded domains is also influenced by the chain entropy of the denatured part of the protein [63]. While single molecule methods excel at measuring the mechanical unfolding of proteins along the N-to-C direction (**Figure 2**), they lack the flexibility to sample unfolding over other coordinates. To mitigate this inconvenience, we

compared the folding barriers along all possible pulling coordinates via steered molecular dynamics simulations, albeit over a higher force range (**Figure 3A-C**). The single molecule results were then used to estimate energy barriers and produce the potential of mean force (PMF) of the considered polyprotein for all possible tethering configurations, as a function of force (**Figure 3D-F**). Brownian dynamics simulations over these PMFs allowed us to determine the unfolding response of all the tethered protein domains inside the network, on the second time scale. A second important component is network optimization, which considers the changing length between the network nodes as protein domains unfold (**Figure 1C**). By alternating between the mechanical response of protein domains and the network equilibration steps, we can predict the stress response of protein-based materials, both in water and in the presence of osmolytes (**Figure 4**). Our main finding is that the protein domains that are part of the network nodes drive the unfolding response. These predictions agree with the measured response using force-clamp rheometry (**Figure 5**). In these measurements of protein-biomaterials, we see that the measured strain after a given time scales similarly for protein L hydrogels, both in TRIS and 50% glycerol in the low force range (up to 5 mN or ~20 kPa), but then the strain of the materials in TRIS increases more abruptly (**Figure 5A**). This behavior can be conceptualized from the force-per-molecule perspective, which is less than 2 pN at 5 mN. At low forces, below 2 pN, the elasticity is driven by molecular orientation inside the material and some unfolding of network nodes. For the TRIS buffer, above 5 mN we see an accelerated change in strain with stress, as the unfolding of N-to-C domains contribute as well. These domains take longer to unfold in the osmolyte environment, as evidenced also by the single molecule measurements and MD simulations (**Figures 2-4** and **Figure S5**). Another interesting behavior measured for protein L hydrogels is their kinetic response under a constant strain (**Figure 5E**). When single proteins are exposed to increasing force levels, their unfolding kinetics scales accordingly. However, protein materials show similar kinetics, when their strain is normalized, as seen in **Figure 5E**, and also reported for BSA hydrogels [22]. This behavior has been considered indicative of poroelasticity, where the solvent molecules move relative to the polymer network over a long range [64]. However, this behavior can also be explained from the perspective that, as molecules inside the network have different orientations, their combined kinetics does not change significantly with applied stress, which has a more dramatic effect on the final strain (**Figure 5E** model).

Biomaterials that harvest the protein unfolding response have the potential to produce transformative solutions, while also being used as tools to answer physics questions. The model developed here holds the potential of mainstreaming the development of protein-based materials, as it enables their development using rational design. By combining coarse-grained modeling of the protein network with Brownian dynamics simulations of the unfolding kinetics, our model will enable the study of scaling principles and will improve our understanding on how to correlate the unfolding response of proteins as molecular components with macroscopic responses seen at biomaterial level. We foresee that the model proposed here could be further used to develop biomaterials that take advantage of proteins known to catalyze bioreactions, with activity control for binding-release as a function of force, biomaterials optimized for easier fluid exchange, as well as biomaterials with applications as responsive cell-growth matrix, better resembling the ECM environment. Moving forward, we envision that the current model will be applied to other protein systems. We also see a potential improvement in simulation performance and number of molecules considered through machine learning approaches [65, 66]. From this perspective, the current model would be the starting point for training, by generating molecular movement inside the network and unfolding probabilities of domains that are tethered along their N-to-C coordinate, or part of a node. Once trained, a machine learning algorithm would then be able to quickly scale to many thousands of molecules and enable coarse-grained studies at mesoscale.

MATERIALS AND METHODS

Protein expression and purification

The main protein under investigation in this study is the B1 domain of protein L, sourced from the *Finegoldia magna* bacterium, henceforth referred to simply as protein L [34]. The protein L domain was repeated eight times by generating new restriction sites through compatible cohesive ends [67]. Distinct constructs of protein L are employed depending on whether they are utilized for single-molecule experiments or protein hydrogel investigations. For single-molecule experiments, the L_8 polyprotein was ligated into a modified HaloTag pFN18a vector (Promega), which had incorporated a SpyTag at the C-terminus [57]. These molecular tags enable attachment to the fluid chamber and paramagnetic beads, respectively. Also, a HaloTag-SpyCatcher protein was purified and used for attachment. Conversely, for protein hydrogel experiments, the protein L_8 fragment was ligated into a pQ80E vector, which also had a Histidine tag at the N-terminus. The vectors were transformed into BLR DE3 competent cells, which were grown at 37°C in Luria Broth (LB) in the presence of antibiotics (50 µg/mL carbenicillin) until an optical density of 0.6 at 600 nm (OD_{600}) was reached. Protein overexpression was induced by adding 1 mM Isopropyl ß-D-1-thiogalactopyranoside (IPTG), followed by overnight incubation at room temperature. Subsequently, cells were pelleted then resuspended in elution/wash (E/W) buffer (300 mM NaCl, 50 mM NaH_2PO_4, pH 7.0). Cell lysis was performed by incubating the resuspended slurry with lysozyme, deoxyribonuclease (DNase), ribonuclease (RNase), and protease inhibitors for 40 minutes. The resulting lysate was subjected to sonication, utilizing 7-10 cycles of 10 seconds on at 50% power, followed by 20 seconds off for 10 pulses with 3-minute intervals to mitigate heat-induced damage and protein aggregation. Subsequently, the cell slurry is centrifuged at 20,000 RPM at 4°C for 40 minutes to segregate soluble protein from insoluble cellular debris. Purification of the protein solution was achieved through a two-step process. Initially, chemical affinity purification is conducted by passing the solution through a Nickel Nitriloacetic acid (NiNTA) column, where the HisTag binds to the nickel, separating it from the solute. The column is subsequently washed with approximately 50 mL of E/W buffer to eliminate residual solute. Elution of the protein from the column is achieved with E/W buffer containing 250 mM imidazole, and the eluate is fractionated into 200 µL aliquots, with the six aliquots of highest protein concentration retained. Finally, size exclusion chromatography is employed, utilizing an Akta GE system with elution in HEPES buffer (50 mM, NaCl 150 mM, pH 7.2) with and without 5% (v/v) glycerol for single-molecule and hydrogel experiments respectively. Fractions corresponding to the protein's size are collected and concentrated if being used for protein hydrogels. All the buffers used for HaloTag proteins also contained 5% v/v glycerol.

Single molecule magnetic tweezers measurements

Single molecules of polyprotein L_8 molecules were tethered between a functionalized glass coverslip and a superparamagnetic bead (M-270 Dynabeads) using the HaloTag and SpyTag/AviTag ends [57]. Glass coverslips (Ted Pella) were first cleaned with 1% Helmanex III solution (Sigma-Aldrich) for 20 min using a sonication bath warmed at 60°C, followed by extensive rinsing with double-distilled (DD) water. The slides were then sonicated 20 min at a time in acetone, followed by methanol and dried in an oven. The bottom slides were activated for 20 min using air-plasma (Harrick Plasma) and silanized with 3-aminopropyl trimethoxysilane 0.1% v/v in methanol, for another 20 min. After rinsing with methanol, the slides were left in an oven at 110 °C for at least 1 hour. Fluid chambers were assembled using parafilm strips as spacers, which were melted to fuse the chambers over a hot plate. A ring of silicone (Sylgard 184 Dow Corning) was applied around each edge of the coverslip, creating fluid basins. Prior to use, chambers were functionalized to facilitate attachment of single molecules to their surfaces. Each chamber was

filled with a 1% v/v glutaraldehyde and 0.025% w/v reference beads mixture in PBS buffer at pH 7.2 for 1 hour (3 μm Amino/Streptavidin Polystyrene Particles, Spherotech). Following washing with PBS buffer, a solution of 10 μg/mL amine-terminated chloroalkane ligand (HaloTag O4 Ligand, Promega) in PBS buffer (pH 7.2) was added and allowed to react for 4 hours, facilitated by the glutaraldehyde bridge molecules. Following incubation, each chamber is washed with PBS and a TRIS blocking buffer (20 mM Tris, 150 mM KCl, pH 7.2). Before an experiment, the chambers were incubated overnight with 1% w/v sulfhydryl blocked-BSA (Lee Biosolutions) to prevent non-specific attachment of paramagnetic beads. Amine-terminated Superparamagnetic M-270 Dynabeads were functionalized in a similar way with the chloroalkane ligand and later reacted with the HaloTag-SpyCatcher protein. The fluid chambers were incubated for 10 min with the HaloTag-L8-SpyTag protein, washed with PBS buffer, and the functionalized paramagnetic beads were added inside the fluid chamber for 10-30 min. The chamber was placed on the magnetic tweezers' microscope and a low 4 pN force was applied to remove the non-specifically sedimented paramagnetic beads from the field of view.

Our magnetic tweezers is built ontop of an Olympus IX-81 microscope, as detailed elsewhere [57]. For each experiment, a pair of beads (one magnetic and one non-magnetic) was selected by having 128x128 pixels region-of-interest (ROIs) drawn around them. high-speed camera capturing was initiated (operating at 1 kHz). Traces showing the unfolding fingerprint of protein L, as a staircase-like length increase with eight steps, were then saved and used to measure the unfolding rates as a function of force. Both regular PBS and glycerol 50% v/v in PBS were used as buffers. The measured protein L unfolding traces were cut and averaged to determine the rates, while the error was estimated using a bootstrapping algorithm.

Protein hydrogel synthesis and Force-clamp rheometry measurements

Protein L_8 hydrogels were synthesized employing a photoactivated chemical crosslinking procedure. This approach involves concentrating the purified protein to 1 mM using Amicon Ultra 0.5 mL Centrifugal Filters (EMD Millipore). Typically 15 μL of concentrated protein L was mixed with 1 M ammonium persulfate (APS) and 6.67 mM tris(bipyridine) ruthenium(II) chloride $[Ru(bpy)_3]^{2+}$ in a volumetric ratio of 15:1:1. The resulting mixture was loaded into polytetrafluoroethylene (PTFE) tubes (Cole-Parmer ID 0.022") that was previously passivated with Sigmacote for 5 minutes, to prevent the adhesion of the protein hydrogel to the tube walls. Subsequently, the loaded tubes were exposed to a 100W mercury lamp with a 400 nm long-pass filter for 30 minutes. Following the light exposure, the protein hydrogels were extracted from the PTFE tube. The extracted hydrogel was then sectioned into 7 mm segments and affixed to the measuring instrument.

To measure the viscoelastic response of protein L hydrogels, we used a custom-built protein hydrogel force-clamp (FC) rheometer, as described elsewhere in detail [35]. Briefly, the hydrogel biomaterial was tethered between the hooks linked to a voice coil and a force sensor using double knotted sutures. Our instrument uses an analog proportional- integral- differential (PID) feed-back loop to match the forces applied to the sample and reported by the force sensor with a setpoint provided by the computer through a National Instruments (NI) data acquisition card (DAC) by continuously adjusting the other end of the sample with the help of a voice coil (Equipment Solutions). Before each trace, a slack curve was performed, where the tension was quickly increased and decreased to determine the point where the tethered gel starts experiencing force. This slack curve was then used during the analysis part to determine exactly the sample length. Following the slack curve, a force-ramp trace was performed where the stress was increased and decreased linearly with time (0.04 kPa/s) to a low value (4 kPa) and the Young's modulus of the protein was measured. This Young's modulus was later used to assess if any defects were present, as these defects would result in a lower stiffness. Then a constant force protocol was

applied which consisted in having the sample for 30 s at a low 0.1 mN/0.4 kPa force, followed by exposure for 120 s at a given force, followed by a relaxation back to 0.4 kPa for 300 s. After this pulse, the sample was discarded. Force clamp experiments are performed for various applied loads up to 11 mN, with a minimum of three hydrogels tested per force. The PID controller governing the voice coil was optimized for each force, ensuring that the transition from no load to a stable full load occurs in less than 100 ms. During analysis, the samples that showed more than 10% difference from average in the measured Young's modulus during the force-ramp measurements, or final strain during the constant force measurements, were discarded. Defects can be seen if air bubbles formed inside the cast, or if the gel was damaged during the attachment to the hooks. Also some samples, especially those measured at high forces, can show slippage from hooks, which prevents us from determining the strain from the measured length.

Steered Molecular Dynamics Simulations

All-atom simulations were carried out in explicit solvent utilizing the GROMACS version 2019.2 molecular dynamics engine [68] along with PLUMED version 2.5.2 [69].The force field used was AMBER99SB-ILDN [70] with the TIP3P water model. Periodic boundary conditions were utilized as well as electrostatic and Lennard Jones interactions being cutoff over 12 Å. Long-range electrostatic interactions were calculated using the particle mesh Ewald method with a grid spacing of 1.6 Å. Steered molecular dynamics simulations of Protein L (PDB 1HZ5) were carried out by applying a constant force between the C_α's of the residues used along the z-direction. Five pairs of residues were investigated. The residue pairs are 1MET-64GLY (N-C), 1MET-34TYR (N-Y34), 1MET-56TYR (N-Y56), 34TYR-56TYR (Y34-Y56), and 34TYR-64GLY (Y34-C). *System preparation:* A single protein molecule was placed in a 5 X 5 X 30 nm box and subsequently solvated with water. This comprised of 27,823 water molecules and 84730 atoms total. Energy minimization was then performed using the steepest descent method (2,000 steps) before equilibration. A second system was prepared for simulations in a glycerol environment. Glycerol molecules were placed in the simulation volume such that 50% of the solution is glycerol (1985 glycerol and 20157 water molecules). Energy minimization was again performed using the steepest descent method (2,000 steps) before equilibration. *Equilibration:* In both systems, the protein was then equilibrated for 1 ns in the isothermal-isobaric ensemble at 300K and 1 bar with a time step of 2 fs. During the equilibration time, a moving harmonic potential with a force constant was applied between the pair of residues that we will be pulling on during the constant force simulations. This moving potential gradually rotates the protein domain to align the line between residues to the long axis of the simulation box. This equilibration was performed for each of the five pulling trajectories. *Unfolding simulations:* Constant force simulations were performed by applying a linear potential (i.e., a constant force) between the two residues used for pulling. Depending on force, each simulation was run for either 20 ns or 40 ns for higher and lower forces respectively. The distance between the z-coordinates of the pulling residues was recorded for plotting the end-to-end distance as a function of time. Each force for each system and pulling geometry was performed 10 times. The forces ranged from 200 pN to 800 pN. *Analysis:* The time at which the end-to-end distances is halfway between its folded and unfolded state was recorded for each trace at the same force for each geometry and system. Taking the average of these times gives the average rupture time which, when inverted, gives the unfolding rate of that configuration. We take the natural log of the rate and plot it verses force.

Brownian Dynamics Simulations

Coarse-grained network assembly: Simulated protein networks were created in Igor Pro (Wavemetrics) using the physical dimensions of polyprotein L, as previously described [23]. Briefly, starting with 75 molecules placed randomly inside a box to produce an equivalent concentration of 1 mM. Molecules were then randomly placed inside the box such that no molecules were interacting with each other. After placement, molecules were left to diffuse and whenever two domains came within two radii and did not have more than three connections, they were considered as being crosslinked. The diffusion coefficients for each cluster were recalculated and the movement adjusted accordingly. Once all molecules were connected into a network, the simulation was stopped. Several 75-molecules networks of fully crosslinked protein L were generated and subsequently used for the network dynamics simulations.

Network Dynamics Simulations under a Force Vector. The network dynamics simulations were run in Igor Pro, with calls for network energy minimization to a Matlab script. To speed up the simulation, the network optimization was only called when a domain inside the network changed its folded state. Also a library of PMF profiles was generated in increments of 0.1 pN for molecules ranging from having all eight domains as N-to-C to having all eight domains as weak (nodes). The PMFs were constructed as previously reported [32, 33]. Briefly, two components were incorporated to generate each PMF profile for a given force: the first one describing the free energy barrier separating the folded and unfolded states (**Equation 1**), which was constructed onto of the force-dependent entropic energy, calculated based on the worm-like-chain model for polymer elasticity (**Equation 2**):

$$U_{barrier} = M_0\left[\left(1 - e^{-2\frac{b}{r}(x-r)}\right)^2 - 1\right] + G_0\left[e^{-\frac{(x-x_b)^2}{2\sigma^2}}\right] \quad (1)$$

$$U_{entropy} = \frac{k_B T}{\ell_p}\left[\frac{\Delta L_c}{4}\left(\left(1 - \frac{x}{L_c}\right)^{-1} - 1\right) - \frac{x}{4} + \frac{x^2}{2L_c}\right] - Fx \quad (2)$$

where M_0 is the Morse potential depth, b is the Morse barrier width, r is the domain radius, G_0 and σ the gaussian height and width respectively, x_b equivalent distance to transition state, x extension, $L_c = 18.6\ nm$ is the contour length for a protein L domain[34], and F the current force.

The single molecule traces of polyproteins under a given force were then computed by assuming an overdamped Langevin diffusion process [71]:

$$\frac{dx}{dt} = \frac{D}{k_B T}\left(\Gamma(t) - \frac{\partial U}{\partial x}\right) \quad (3)$$

where D is the diffusion coefficient over the energy landscape (PMF) U, and $\Gamma(t)$ is a random force representing the thermal fluctuations. The sum for the Morse and Gaussian potentials, together with the diffusion coefficient D, were estimated to match the unfolding rates, by minimizing the difference between simulated and measured unfolding rates. Then, the Morse potential alone was calculated from the half folding force. The values used to construct the free energy landscapes for both strong and weak domains, in water and glycerol, are given in **Supplementary Table 1**, and were estimated from measurements and SMDS, as explained in the main text. For each force a total of nine PMFs were generated for molecules having all domains experiencing force over the N-to-C coordinates (denoted as 'strong') to having all domains as nodes (denoted as 'weak'). At the start of each simulation, a connection matrix was generated for each molecule inside the gel. Throughout the simulation, each molecule received a PMF based on its orientation, number of tethered domains between the first and last node and number of domains part of a node. The force was assigned as the product between the overall applied force and cosine of the angle to the z-axis, which was considered the force axis (and

rounded to the nearest 0.1 pN). The force of each molecule was recalculated for each simulation step.

Each simulation step is composed of two sub steps. First, each molecule was diffused on the free energy landscape for 100 μs after which the position of each domain in the molecule is altered such that the distance between the terminal domains matches its location on the energy landscape. Since the global force is applied in the *z*-direction, any extension associated with traveling along the energy landscape is applied to the *z*-coordinates resulting in a rotation to align to that force. These minor changes in length are typically less than a nanometer. The larger changes in extension occurred when an unfolding event happens. If an unfolding event occurred, then the locations of the nodes in the network were saved and are passed to Matlab for optimization. Optimization took the location of the nodes of the network, the domains that each node was connected to, and which nodes were part of a single molecule. Stretching and bending energies, $H_{node-node}$ and $H_{bending}$ were calculated for intermolecular crosslinks and crosslink energies $H_{crosslink}$ were calculated for intramolecular node locations (see also **Figure 1C**). A weak potential kept the nodes in their initial location (**Equation 4**), with the stretching energy is given by [13]:

$$H_{node-node} = \frac{\mu}{2}\sum_{<i,j>}\frac{(\vec{\boldsymbol{u}}_{\boldsymbol{i,j}}\cdot\vec{\boldsymbol{r}}_{\boldsymbol{i,j}})^{\boldsymbol{2}}}{|\vec{\boldsymbol{r}}_{\boldsymbol{i,j}}|^3} \quad (4)$$

where $\mu = \frac{4\ell_p k_B T}{\pi r^4} = 0.2$ pN for a persistence length $\ell_p = 0.58$ nm for protein L[34], with $k_B T$ being the thermal energy, and the vectors $\vec{\boldsymbol{u}}_{\boldsymbol{i}}$ and $\vec{\boldsymbol{r}}_{\boldsymbol{i}}$ point to nodes in the perturbed and stable configurations respectively.

The intermolecular bending energies were calculated for any node that was connected to two others. The vectors from the middle node to the other two were calculated and a cross product is taken. The cross product returns the angle between these two vectors and the bending energy is proportional to the square of this value (**Equation 5**).

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

with the bending constant for a rigid rod as $\kappa = \ell_p k_B T = 2.4$ pN · nm^2 [15].

Crosslinking energies followed a simple spring like potential where any perturbation from the rest length, the energy will drive it back towards this minimum (**Equation 6**):

$$H_{crosslink} = K\sum_{<i',j'>}(2r - \|\vec{\boldsymbol{u}}_{i',j'}\|)^2 \quad (6)$$

where $K = 3.72 \cdot 10^5$ pN/nm is the force constant associated with the quadratic approximation of a C_n-C_n bond [72] and $\|\vec{\boldsymbol{u}}_{i,j}\|$ is the perturbed bond length between two adjacent bonds i and j, such that $\vec{\boldsymbol{u}}_{\boldsymbol{i,j}} = \vec{\boldsymbol{u}}_j - \vec{\boldsymbol{u}}_i$. For intermolecular crosslinks, this rest length is equal to two radii *r* (4 nm), which was determined for a protein L domain from its crystal structure. The values for each potential are also given in **Supporting Table S2**.

The minimization was run by summing the three energy components using the “fminunc” MATLAB feature and ran until it reaches a global minimum or 1000 were performed. After minimization, the optimized network was passed back to IgorPro (Wavemetrics) to continue the simulation. The distance between the intramolecular nodes was recalculated and compared to the distance that that molecule is at on the energy landscape. If there was a difference between the two distances, the location of the nodes was adjusted to match the separation on the PMF. After the distances were matched, we returned to the diffusion on the energy landscape and repeated these two steps until the simulation time was reached.

Five simulated gels were prepared, and each were simulated under the global applied forces of 2, 4, 6, 8, and 10 pN using energy landscapes for water and glycerol. The five simulations for each force were taken and the strains and number of unfolded domains were averaged. Analysis was done through measuring the length of the simulated gel. The length was determined by taking

the average of the 10 furthest node domains on each end and taking the difference. This calculation was done for each frame of the simulation. The normalized strain of the gel was calculated by dividing the length at each frame by the initial length at frame three (when the force was applied). From the strain we plotted the final strain versus the applied global force. We also recorded how many domains unfolded through the simulation separated by whether they were nodes or not.

ACKNOWLEDGEMENTS
This work was supported through a National Science Foundation (NSF) award to IP (award number MCB-1846143 and DMR-2423251) and by Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States (to JN).

AUTHOR CONTRIBUTIONS
JN, SR and ND acquired and analyzed the single molecule data, JN and SB acquired and analyzed the rheometry data, JN did the SMDS, IP and JN developed the network dynamics model, GS and IP conceived and supervised the research project, IP wrote the manuscript with input from all authors.

REFERENCES

1. Lv, S.; Dudek, D. M.; Cao, Y.; Balamurali, M. M.; Gosline, J.; Li, H., Designed biomaterials to mimic the mechanical properties of muscles. *Nature* **2010,** *465* (7294), 69-73.
2. Perez-Benito, A.; Huerta-Lopez, C.; Alegre-Cebollada, J.; Garcia-Aznar, J. M.; Hervas-Raluy, S., Computational modelling of the mechanical behaviour of protein-based hydrogels. *J Mech Behav Biomed Mater* **2023,** *138*, 105661.
3. Khoury, L. R.; Slawinski, M.; Collison, D. R.; Popa, I., Cation-induced shape programming and morphing in protein-based hydrogels. *Sci Adv* **2020,** *6* (18), eaba6112.
4. Wu, X.; Huang, W. M.; Wu, W. H.; Xue, B.; Xiang, D. F.; Li, Y.; Qin, M.; Sun, F.; Wang, W.; Zhang, W. B.; Cao, Y., Reversible hydrogels with tunable mechanical properties for optically controlling cell migration. *Nano Res* **2018,** *11* (10), 5556-5565.
5. Khoury, L. R.; Popa, I., Chemical unfolding of protein domains induces shape change in programmed protein hydrogels. *Nat Commun* **2019,** *10* (1), 5439.
6. Dong, Y.; Ramey-Ward, A. N.; Salaita, K., Programmable Mechanically Active Hydrogel-Based Materials. *Adv Mater* **2021,** *33* (46), e2006600.
7. Bian, Q.; Fu, L.; Li, H., Engineering shape memory and morphing protein hydrogels based on protein unfolding and folding. *Nat Commun* **2022,** *13* (1), 137.
8. Montanari, A. N.; Duan, C.; Aguirre, L. A.; Motter, A. E., Functional observability and target state estimation in large-scale networks. *Proc Natl Acad Sci U S A* **2022,** *119* (1).
9. Nabizadeh, M.; Nasirian, F.; Li, X.; Saraswat, Y.; Waheibi, R.; Hsiao, L. C.; Bi, D.; Ravandi, B.; Jamali, S., Network physics of attractive colloidal gels: Resilience, rigidity, and phase diagram. *Proc Natl Acad Sci U S A* **2024,** *121* (3), e2316394121.
10. Parada, G. A.; Zhao, X., Ideal reversible polymer networks. *Soft Matter* **2018,** *14* (25), 5186-5196.
11. Park, J.; Kim, W. J.; Kim, Y.; Lee, E. K.; Kim, H., Threading Subunits for Polymers to Predict the Equilibrium Ensemble of Solid Polymer Electrolytes. *J Phys Chem Lett* **2024,** *15* (5), 1227-1233.
12. Khare, E.; A, C. S. d. A.; Lee, N.; Skaf, M. S.; Buehler, M. J., Crosslinker energy landscape effects on dynamic mechanical properties of ideal polymer hydrogels. *Mater Adv* **2024,** *5* (5), 1991-1997.
13. Broedersz, C. P.; Mao, X.; Lubensky, T. C.; MacKintosh, F. C., Criticality and isostaticity in fibre networks. *Nature Physics* **2011,** *7* (12), 983-988.
14. Wu, J.; Li, P.; Dong, C.; Jiang, H.; Bin, X.; Gao, X.; Qin, M.; Wang, W.; Bin, C.; Cao, Y., Rationally designed synthetic protein hydrogels with predictable mechanical properties. *Nat Commun* **2018,** *9* (1), 620.
15. Shmilovich, K.; Popa, I., Modeling Protein-Based Hydrogels under Force. *Phys Rev Lett* **2018,** *121* (16), 168101.
16. Mu, X.; Yuan, J. S. K.; Choi, J.; Zhang, Y. X.; Cebe, P.; Jiang, X. C.; Zhang, Y. S.; Kaplan, D. L., Conformation-driven strategy for resilient and functional protein materials. *Proceedings of the National Academy of Sciences of the United States of America* **2022,** *119* (4).
17. Popa, I., Modeling and Simulations of Multicomponent Hydrogels for Biomedical Applications. In *Multicomponent Hydrogels: Smart Materials for Biomedical Applications*, Dodda, J. M.; Deshmukh, K.; Bezuidenhout, D., Eds. The Royal Society of Chemistry: 2023; p 0.
18. Bujalowski, P. J.; Oberhauser, A. F., Tracking unfolding and refolding reactions of single proteins using atomic force microscopy methods. *Methods* **2013,** *60* (2), 151-60.
19. Carrion-Vazquez, M.; Li, H.; Lu, H.; Marszalek, P. E.; Oberhauser, A. F.; Fernandez, J. M., The mechanical stability of ubiquitin is linkage dependent. *Nat Struct Biol* **2003,** *10* (9), 738-43.

20. Stirnemann, G.; Kang, S. G.; Zhou, R.; Berne, B. J., How force unfolding differs from chemical denaturation. *Proc Natl Acad Sci U S A* **2014,** *111* (9), 3413-8.
21. Bernardi, R. C.; Durner, E.; Schoeler, C.; Malinowska, K. H.; Carvalho, B. G.; Bayer, E. A.; Luthey-Schulten, Z.; Gaub, H. E.; Nash, M. A., Mechanisms of Nanonewton Mechanostability in a Protein Complex Revealed by Molecular Dynamics Simulations and Single-Molecule Force Spectroscopy. *J Am Chem Soc* **2019,** *141* (37), 14752-14763.
22. Khoury, L. R.; Nowitzke, J.; Shmilovich, K.; Popa, I., Study of Biomechanical Properties of Protein-Based Hydrogels Using Force-Clamp Rheometry. *Macromolecules* **2018,** *51* (4), 1441-1452.
23. Nowitzke, J.; Popa, I., What Is the Force-per-Molecule Inside a Biomaterial Having Randomly Oriented Units? *Journal of Physical Chemistry Letters* **2022,** *13* (31), 7139-7146.
24. Hughes, M. D. G.; Cussons, S.; Hanson, B. S.; Cook, K. R.; Feller, T.; Mahmoudi, N.; Baker, D. L.; Ariens, R.; Head, D. A.; Brockwell, D. J.; Dougan, L., Building block aspect ratio controls assembly, architecture, and mechanics of synthetic and natural protein networks. *Nat Commun* **2023,** *14* (1), 5593.
25. Fu, L.; Li, L.; Bian, Q.; Xue, B.; Jin, J.; Li, J.; Cao, Y.; Jiang, Q.; Li, H., Cartilage-like protein hydrogels engineered via entanglement. *Nature* **2023,** *618* (7966), 740-747.
26. Popa, I.; Berkovich, R.; Alegre-Cebollada, J.; Badilla, C. L.; Rivas-Pardo, J. A.; Taniguchi, Y.; Kawakami, M.; Fernandez, J. M., Nanomechanics of HaloTag tethers. *J Am Chem Soc* **2013,** *135* (34), 12762-71.
27. Garcia-Manyes, S.; Dougan, L.; Fernandez, J. M., Osmolyte-induced separation of the mechanical folding phases of ubiquitin. *Proc Natl Acad Sci U S A* **2009,** *106* (26), 10540-5.
28. Popa, I.; Kosuri, P.; Alegre-Cebollada, J.; Garcia-Manyes, S.; Fernandez, J. M., Force dependency of biochemical reactions measured by single-molecule force-clamp spectroscopy. *Nat Protoc* **2013,** *8* (7), 1261-76.
29. Mondal, J.; Halverson, D.; Li, I. T.; Stirnemann, G.; Walker, G. C.; Berne, B. J., How osmolytes influence hydrophobic polymer conformations: A unified view from experiment and theory. *Proc Natl Acad Sci U S A* **2015,** *112* (30), 9270-5.
30. Dietz, H.; Berkemeier, F.; Bertz, M.; Rief, M., Anisotropic deformation response of single protein molecules. *Proc Natl Acad Sci U S A* **2006,** *103* (34), 12724-8.
31. O'Brien, E. P.; Ziv, G.; Haran, G.; Brooks, B. R.; Thirumalai, D., Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. *Proc Natl Acad Sci U S A* **2008,** *105* (36), 13403-8.
32. Berkovich, R.; Hermans, R. I.; Popa, I.; Stirnemann, G.; Garcia-Manyes, S.; Berne, B. J.; Fernandez, J. M., Rate limit of protein elastic response is tether dependent. *Proceedings of the National Academy of Sciences of the United States of America* **2012,** *109* (36), 14416-14421.
33. Valle-Orero, J.; Eckels, E. C.; Stirnemann, G.; Popa, I.; Berkovich, R.; Fernandez, J. M., The elastic free energy of a tandem modular protein under force. *Biochem Biophys Res Commun* **2015,** *460* (2), 434-8.
34. Popa, I.; Rivas-Pardo, J. A.; Eckels, E. C.; Echelman, D. J.; Badilla, C. L.; Valle-Orero, J.; Fernandez, J. M., A HaloTag Anchored Ruler for Week-Long Studies of Protein Dynamics. *J Am Chem Soc* **2016,** *138* (33), 10546-53.
35. Khoury, L. R.; Nowitzke, J.; Dahal, N.; Shmilovich, K.; Eis, A.; Popa, I., Force-Clamp Rheometry for Characterizing Protein-based Hydrogels. *J Vis Exp* **2018,** (138).
36. Huerta-Lopez, C.; Alegre-Cebollada, J., Protein Hydrogels: The Swiss Army Knife for Enhanced Mechanical and Bioactive Properties of Biomaterials. *Nanomaterials (Basel)* **2021,** *11* (7).
37. Kim, M.; Tang, S.; Olsen, B. D., Physics of engineered protein hydrogels. *Journal of Polymer Science Part B: Polymer Physics* **2013,** *51* (7), 587-601.
38. Gao, X. Y.; Fang, J.; Xue, B.; Fu, L. L.; Li, H. B., Engineering Protein Hydrogels Using SpyCatcher-SpyTag Chemistry. *Biomacromolecules* **2016,** *17* (9), 2812-2819.

39. Caliari, S. R.; Burdick, J. A., A practical guide to hydrogels for cell culture. *Nat Methods* **2016,** *13* (5), 405-14.
40. Davari, N.; Bakhtiary, N.; Khajehmohammadi, M.; Sarkari, S.; Tolabi, H.; Ghorbani, F.; Ghalandari, B., Protein-Based Hydrogels: Promising Materials for Tissue Engineering. *Polymers (Basel)* **2022,** *14* (5).
41. Doolan, J. A.; Alesbrook, L. S.; Baker, K.; Brown, I. R.; Williams, G. T.; Hilton, K. L. F.; Tabata, M.; Wozniakiewicz, P. J.; Hiscock, J. R.; Goult, B. T., Next-generation protein-based materials capture and preserve projectiles from supersonic impacts. *Nat Nanotechnol* **2023,** *18* (9), 1060-1066.
42. Hudalla, G. A.; Sun, T.; Gasiorowski, J. Z.; Han, H.; Tian, Y. F.; Chong, A. S.; Collier, J. H., Gradated assembly of multiple proteins into supramolecular nanomaterials. *Nat Mater* **2014,** *13* (8), 829-36.
43. 'Wang, S. T.; Minevich, B.; Liu, J.; Zhang, H.; Nykypanchuk, D.; Byrnes, J.; Liu, W.; Bershadsky, L.; Liu, Q.; Wang, T.; Ren, G.; Gang, O., Designed and biologically active protein lattices. *Nat Commun* **2021,** *12* (1), 3702.
44. Goult, B. T.; von Essen, M.; Hytonen, V. P., The mechanical cell - the role of force dependencies in synchronising protein interaction networks. *J Cell Sci* **2022,** *135* (22).
45. Hytonen, V. P.; Wehrle-Haller, B., Mechanosensing in cell-matrix adhesions - Converting tension into chemical signals. *Exp Cell Res* **2016,** *343* (1), 35-41.
46. Popa, I.; Gutzman, J. H., The extracellular matrix–myosin pathway in mechanotransduction: from molecule to tissue. *Emerging Topics in Life Sciences* **2018,** *2* (5), 727-737.
47. Patten, J.; Wang, K., Fibronectin in development and wound healing. *Adv Drug Deliv Rev* **2021,** *170*, 353-368.
48. Dong, X.; Zhao, B.; Iacob, R. E.; Zhu, J.; Koksal, A. C.; Lu, C.; Engen, J. R.; Springer, T. A., Force interacts with macromolecular structure in activation of TGF-beta. *Nature* **2017,** *542* (7639), 55-59.
49. Eckels, E. C.; Tapia-Rojo, R.; Rivas-Pardo, J. A.; Fernandez, J. M., The Work of Titin Protein Folding as a Major Driver in Muscle Contraction. *Annu Rev Physiol* **2018,** *80*, 327-351.
50. Rivas-Pardo, J. A.; Eckels, E. C.; Popa, I.; Kosuri, P.; Linke, W. A.; Fernandez, J. M., Work Done by Titin Protein Folding Assists Muscle Contraction. *Cell Rep* **2016,** *14* (6), 1339-47.
51. Kim, D. E.; Fisher, C.; Baker, D., A breakdown of symmetry in the folding transition state of protein L. *J Mol Biol* **2000,** *298* (5), 971-84.
52. O'Neill, J. W.; Kim, D. E.; Baker, D.; Zhang, K. Y., Structures of the B1 domain of protein L from Peptostreptococcus magnus with a tyrosine to tryptophan substitution. *Acta Crystallogr D Biol Crystallogr* **2001,** *57* (Pt 4), 480-7.
53. Brockwell, D. J.; Beddard, G. S.; Paci, E.; West, D. K.; Olmsted, P. D.; Smith, D. A.; Radford, S. E., Mechanically unfolding the small, topologically simple protein L. *Biophys J* **2005,** *89* (1), 506-19.
54. Sadler, D. P.; Petrik, E.; Taniguchi, Y.; Pullen, J. R.; Kawakami, M.; Radford, S. E.; Brockwell, D. J., Identification of a mechanical rheostat in the hydrophobic core of protein L. *J Mol Biol* **2009,** *393* (1), 237-48.
55. Liu, R.; Garcia-Manyes, S.; Sarkar, A.; Badilla, C. L.; Fernandez, J. M., Mechanical characterization of protein L in the low-force regime by electromagnetic tweezers/evanescent nanometry. *Biophys J* **2009,** *96* (9), 3810-21.
56. Valle-Orero, J.; Rivas-Pardo, J. A.; Popa, I., Multidomain proteins under force. *Nanotechnology* **2017,** *28* (17), 174003.
57. Dahal, N.; Nowitzke, J.; Eis, A.; Popa, I., Binding Induced Stabilization Measured on the Same Molecular Protein Substrate using Single Molecule Magnetic Tweezers and Hetero-Covalent Attachments. *The Journal of Physical Chemistry B* **2020**.

58. Stannard, A.; Mora, M.; Beedle, A. E. M.; Castro-Lopez, M.; Board, S.; Garcia-Manyes, S., Molecular Fluctuations as a Ruler of Force-Induced Protein Conformations. *Nano Lett* **2021,** *21* (7), 2953-2961.
59. Graille, M.; Stura, E. A.; Housden, N. G.; Beckingham, J. A.; Bottomley, S. P.; Beale, D.; Taussig, M. J.; Sutton, B. J.; Gore, M. G.; Charbonnier, J. B., Complex between Peptostreptococcus magnus protein L and a human antibody reveals structural convergence in the interaction modes of Fab binding proteins. *Structure* **2001,** *9* (8), 679-87.
60. Beckingham, J. A.; Housden, N. G.; Muir, N. M.; Bottomley, S. P.; Gore, M. G., Studies on a single immunoglobulin-binding domain of protein L from Peptostreptococcus magnus: the role of tyrosine-53 in the reaction with human IgG. *Biochem J* **2001,** *353* (Pt 2), 395-401.
61. Slawinski, M.; Khoury, L. R.; Sharma, S.; Nowitzke, J.; Gutzman, J. H.; Popa, I., Kinetic Method of Producing Pores Inside Protein-Based Biomaterials without Compromising Their Structural Integrity. *Acs Biomater Sci Eng* **2022,** *8* (3), 1132-1142.
62. Berkovich, R.; Garcia-Manyes, S.; Klafter, J.; Urbakh, M.; Fernandez, J. M., Hopping around an entropic barrier created by force. *Biochem Biophys Res Commun* **2010,** *403* (1), 133-7.
63. Chetrit, E.; Sharma, S.; Maayan, U.; Pelah, M. G.; Klausner, Z.; Popa, I.; Berkovich, R., Nonexponential kinetics captured in sequential unfolding of polyproteins over a range of loads. *Curr Res Struct Biol* **2022,** *4*, 106-117.
64. Biot, M. A., General Theory of Three-Dimensional Consolidation. *Journal of Applied Physics* **1941,** *12* (2), 155-164.
65. Doffini, V.; Liu, H.; Liu, Z.; Nash, M. A., Iterative Machine Learning for Classification and Discovery of Single-Molecule Unfolding Trajectories from Force Spectroscopy Data. *Nano Letters* **2023,** *23* (22), 10406-10413.
66. Shmilovich, K.; Mansbach, R. A.; Sidky, H.; Dunne, O. E.; Panda, S. S.; Tovar, J. D.; Ferguson, A. L., Discovery of Self-Assembling π-Conjugated Peptides by Active Learning-Directed Coarse-Grained Molecular Simulation. *The Journal of Physical Chemistry B* **2020,** *124* (19), 3873-3891.
67. Popa, I.; Berkovich, R., *Mechanical Unfolding Response of Proteins*. 2023.
68. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1-2*, 19-25.
69. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **2014,** *185* (2), 604-613.
70. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E., Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics* **2010,** *78* (8), 1950-1958.
71. Berkovich, R.; Garcia-Manyes, S.; Urbakh, M.; Klafter, J.; Fernandez, J. M., Collapse dynamics of single proteins extended by force. *Biophys J* **2010,** *98* (11), 2692-701.
72. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society* **1996,** *118* (45), 11225-11236.