Topological decoding of biomolecular fold complexity

Scalvini, B.

CHAPTER 4:

TOPOLOGICAL SENSING OF THE CYTOSOLIC MOLECULAR MACHINERY WITH OPTICAL TWEEZERS

Single molecule techniques are uniquely fit to investigate protein folding and chaperone assistance, but current assays provide only a partial view of all the possible ways in which the cellular environment can modulate the folding pathway of a protein. Here, we perform protein and DNA pulling assays in a E. coli cytosolic solution, to test the cumulative topological effect of the cytoplasmic interactome over the folding process. We aim at providing a closer picture of protein folding as it occurs in the cell, and present opportunities and challenges of experiments in quasi-biological environments.

Publications associated to this chapter:
Barbara Scalvini, Laurens W.H.J. Heling, Vahid Sheikhhassani, Vanda Sunderlikova, Sander J. Tans, Alireza Mashaghi, *Cytosolic interactome protects against protein unfolding in a single molecule experiment*, UNDER REVIEW

1. INTRODUCTION

DIn recent years, single molecule techniques have shown their potential in the characterization of folding pathways and molecular interactions [1]–[6], in view of their ability to track transient, heterogeneous phenomena. Protein folding transitions have often been studied for proteins in isolation [7]–[9], although in physiological conditions the protein often interacts with many other molecules present in the crowded cellular environment. In this respect, chaperones play a crucial role in assisting protein folding [10][11], by blocking harmful aggregation [12], rescuing misfolded proteins [13], or even affecting the folding pathway [14][15]. However fundamental, the exact mechanisms of chaperone – protein interactions remain mostly obscure, due in part to the promiscuity and heterogeneity of chaperone functions [16]. While the interaction between specific client and chaperone/co-chaperone systems by optical tweezers assays has been object of growing interest in recent years [12][14][15], we are still lacking the full picture when it comes to protein folding in the cellular environment. This is partly because current single molecule assays are done using buffer solutions including one or a few chaperones, ignoring the complexity of the cytoplasmic content. Besides several major chaperone systems [17], the cytosol contains a wide spectrum of proteins that may exhibit yet unknown transient interactions affecting protein states and folding. Certain single molecule investigations have been designed to probe crowding effect, but these studies lack the specificity of cytoplasmic interactions [18]. Overall, there is a need for assays that probe folding dynamics of proteins when exposed to full complexity of the cellular interactome.

Here, we aim at probing the cumulative structural modifications imposed by the cytosolic molecular machinery on a substrate, in this case, maltose-binding protein (MBP), beyond the passive crowding effect. In this proof-of-concept study, we perform single molecule protein (un)folding experiments in diluted *E. Coli* cell extracts, and observe notable variations in the force profile of intermediate states. We see a stabilization against forced unfolding for partial folds, indicating chaperone action aimed at protecting these structures from aggregation. We also present data on single double stranded DNA (dsDNA) molecules in cell extract, as DNA handles are often integral part of protein folding studies by optical tweezers. To our knowledge, this is the first time the mechanical response to the cytosolic interactome of dsDNA molecules is characterized. We observe a lower elastic modulus and a strong hysteretic response after the 65 pN plateau, where the effect of cytosolic proteins binding stretches of ssDNA is predominant. These variations should be appropriately accounted for when using DNA as handle in this type of assays. In view of these results, we discuss challenges and opportunities of pulling experiments in complex environments.

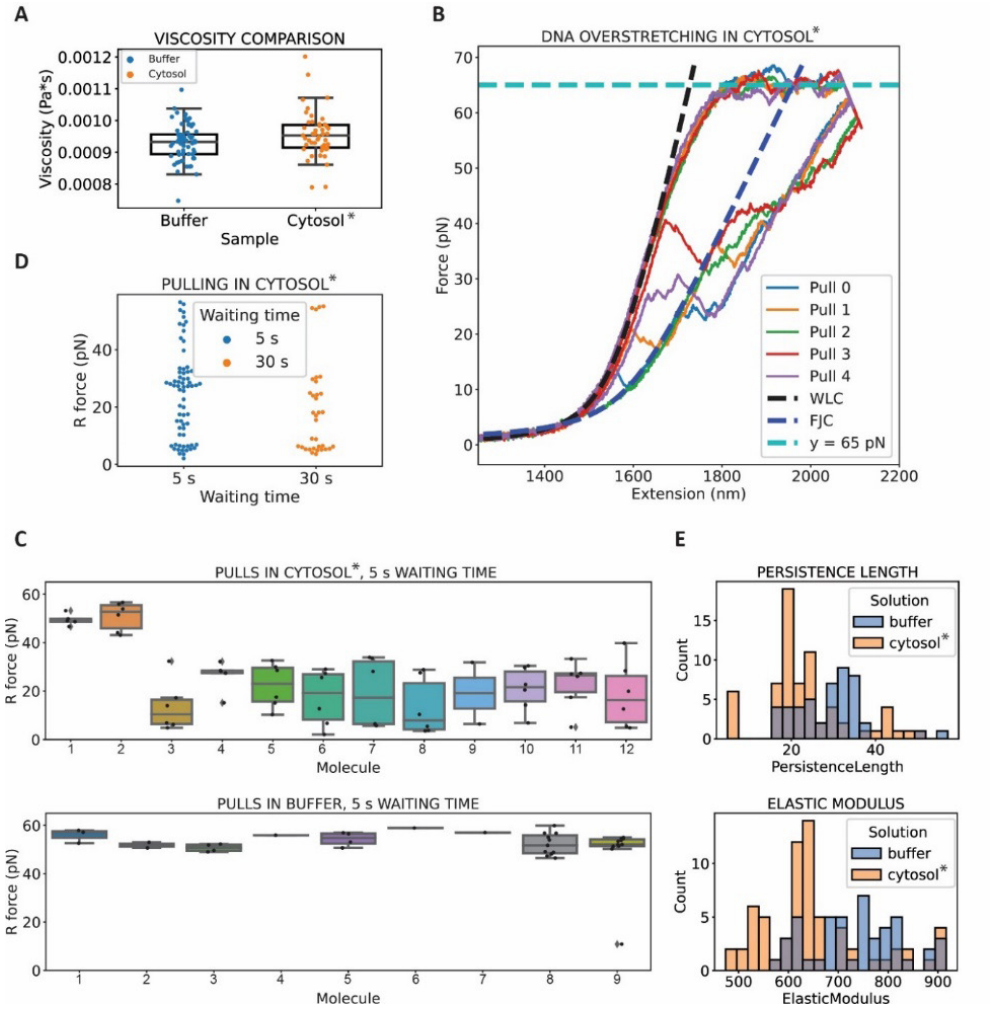

Figure 1. The cytosolic proteome alters the mechanical properties of DNA. ***A*** boxplot of the viscosity of our cytosol solution compared with buffer viscosity, measured by tracking the limit velocity of freely falling beads in solution in the AFS. For buffer, N = 66 beads were measured, and for the cytosol solution, N = 47. ***B*** Example of DNA pulls performed in diluted cytosol solution. The two dashed curves represent the worm-like chain (black) and the freely-joint chain (blue) model. Each pull is represented by two curves, corresponding to the stretching and relaxation of the molecular tether. ***C*** Boxplot of return forces in buffer and in cytosol solution. Whiskers in the boxplot are extended to 1.5 IQR. The data represented was collected from N = 12 DNA molecules (67 pulls) in cytosol solution, and from N = 9 DNA molecules (44 pulls) in buffer. ***D*** Swarm plot of the return force in cytosol, with 5 and 30 s waiting time between pulls. The plot represents data from 67 pulls from N = 12 molecules with a waiting time of 5 s, and data from 43 pulls from N = 14 molecules with a waiting time of 30 s. ***E*** Histogram of the persistence length L_p and elastic modulus K, in buffer and cytosol. The data represented was collected from N = 12 DNA molecules (67 pulls) in cytosol solution, and from N = 9 DNA molecules (44 pulls) in buffer. The label cytosol* in all panels indicates a 4 times dilution of cytosol extract in buffer.

Considering the heterogeneity of processes and interactions taking place between MBP and the cytosolic environment, we also advance a new way to analyze force spectroscopy data, based on the theoretical framework of circuit topology (CT) [19]–[22]. CT formalizes the arrangement of contacts and loops in a folded chain and thus, disruption and formation of loops as probed in the tweezers assay can be directly modelled by this approach. We can therefore represent such loops and associated contour lengths in the formalism of CT, providing a way to visualize the complete ensemble of unfolding pathways and reconstructing from it the topological fingerprint of the substrate.

2. RESULTS AND DISCUSSION

Pulling experiments in cell extract can present various technical challenges, due to the viscous nature of bacterial cytoplasm [23] and the variety of interactions that can arise between cytosolic biomolecules and the construct under study. In order to limit force contributions arising from crowding effect, we diluted our cell extract 4 times in buffer after extraction (see Methods). In this way, physiological concentration ratios of the cytosolic proteome are conserved, while also allowing us to perform traditional optical tweezers experiments with passive (thermal) calibration [24], since we can expect the viscosity of our solution to be homogeneous. We compared the viscosity of cytosol solution (indicated as cytosol* in the figures) and buffer (Figure 1A) by Acoustic Force Spectroscopy (AFS) [25]. We measured the viscosity by extracting the limit velocity of free-falling 4 μm silica beads in cytosol solution and buffer, from a starting position determined by the acoustic node generated by the AFS. From the boxplot in Figure 1A we can observe the spread of the viscosity distribution measured in the two media: buffer viscosity has mean = 9.3×10^{-4} and standard deviation = 0.6×10^{-4} (Pa x s), while the cytosol solution scores a mean of 9.6×10^{-4} and standard deviation of 0.7×10^{-4} (Pa x s). The two distributions compare with a p value of 0.057, as calculated by Kolmogorov test.

2.1. DNA pulling in cell extract

Many protocols for force spectroscopy assays involve the coupling of the protein(s) of interest to long DNA overhangs [26]–[28]. As such, force-extension assays on DNA represent the baseline for pulling experiments. Therefore, a characterization of the interaction between the DNA molecule and the environment where the experiment takes place is important. Moreover, the mechanical beha-

vior of DNA in a biologically relevant environment is a critical aspect of many cellular processes; its elastic properties play a role in protein-DNA interaction, bending, twisting, but also genome compaction and various other structural transitions [5][29]. This is particularly relevant to bacterial cells, where the DNA is embedded in the cytoplasm.

In order to characterize the environmental interplay between dsDNA molecules and cytosolic environment, we performed DNA pulling experiments in cell extract (Figure 1B). After each pull, the molecule was left for 5 or 30 s in the overstretching region. Subsequently, it was relaxed for other 5/30 s in its relaxed state, before repeating the cycle. After the pull (characterized by the worm-like-chain model [30]) the plateau expected for DNA overstretching (OS) is visible at about 65 pN [31] (Figure 1B). However, substantial changes appear striking in the hysteretic behavior of DNA in the retraction phase, after the OS plateau (see Figure S1 for comparison). DNA overstretching under (close to) physiological conditions has been presented as a highly cooperative transition [31]. Depending on boundary conditions such as salt content and temperature, two competing processes might occur predominantly during DNA overstretching [32][33]: a reversible fast transition to an overstretched form of dsDNA (S DNA) [33], or the force-induced melting of the DNA strands, resulting in the breaking of hydrogen bonds with production of ssDNA regions propagating from either a nick or the ends of the molecule [34]. Considering our experimental setup (25 °C and about 100 mM salt concentration), we expect melting to be visible at overstretching force (65 pN). Therefore, we can assume regions of ssDNA are exposed to the cytosolic environment during overstretching of our DNA molecule. The *E. Coli* cytosol proteome presents a variety of proteins capable of binding to ssDNA [35] [36], and it is reasonable to expect that some of them should interact with the exposed regions of ssDNA during the OS phase. Binding with such proteins might induce the ssDNA regions to delay their transition back to B DNA, which is now energetically unfavorable in the range between 50 and 60 pN, where this return transition generally happens in buffer [33]. We call here *return force* (R_F) the force value for which the retraction curve rejoins the prior pulling curve. We can see from the plots in Figure 1C that, while R_F remains segregated in a narrow interval around 50 pN in buffer, the return force for pulls in cell extract has a wide distribution, populating mostly force ranges below 30 pN. The low force return point becomes more evident when we leave the molecule in the OS phase for longer (Figure 1D). Interestingly, sections of the retraction curves fit a freely-joint chain (FJC) curve, a model generally associated with the force/extension curve ssDNA. What we are observing is probably a ssDNA-dsDNA hybrid complex, somewhat similar to what was obtained by Leila Shokri et al. [37] by stretching dsDNA in presence of glyoxal. This effect is either absent or not relevant enough

to be detected by visual inspection of force-extension curves for dsDNA in buffer (Figure S1).

We could assume that, being the recovery of B DNA structure delayed by ssDNA-protein binding, the retraction curve walks along stretches of curves that are most likely linear combinations between WLC and FJC with weights given by the portion of the chain that is in a ssDNA state in that moment of the transition. We can see a shift towards lower values of the WLC parameters, that is, persistence length LP and elastic modulus K (Figure 1E). This effect might indicate the cytosolic content is interacting with the dsDNA molecule under tension.

These changes in mechanical properties of the dsDNA molecule should be taken into account when measuring in a complex medium such as cell extract, as wider variations in mechanical stiffness of the DNA handle can potentially affect signal to noise ratio of the measurement [38]. It is advisable to perform a separate WLC fit for each pull. Types of OT assays other than force-ramp experiments involving long experimental times might need further adjustments to be able to use DNA effectively as a handle in cytosol.

2.2. MBP pulling in cell extract

Individual MBP molecules were tethered between two polystyrene beads (2 μm diameter) by using a 2500 bp DNA handle on each side of the protein, to reduce risk of protein photodamage and bead-bead interaction. Then, we moved the construct into the microfluidics channel containing diluted cell extract. There, we performed cycles of subsequent stretching and relaxation to low forces for a waiting time of 5 s, to provide the protein the opportunity to refold. A complex picture emerged from the unfolding pathway in cytosol solution (Figure 2), especially when it comes to the distribution of unfolding forces (Figure 2A). Native MBP presents a widely studied and distinctive unfolding curve in buffer, with one minor unfolding event happening at low force (~ 10 pN) and the main unfolding of the MBP core generally measured at about 25 pN [12][14][15][39]–[41]. The upper panel in Figure 2A is compatible with this picture, also including all rare intermediate states, generally unfolding in the range between 10-20 pN. We found unfolding of the MBP core in our dataset to happen at a force of 28,1 ± 0,6 pN, slightly higher than what is generally reported in Optical Tweezers assays[12][14][15][39]–[41], a discrepancy that can be attributed to the pulling speed of the experiments (200 nm/s). However, the histogram of unfolding forces in cytosol solution (Figure 2A, lower panel) presents a more complex process, a wide distribution where it is difficult to identify a preferred pathway. Moreover, the higher force tail of the distribution (>40 pN) appears to be more populated in

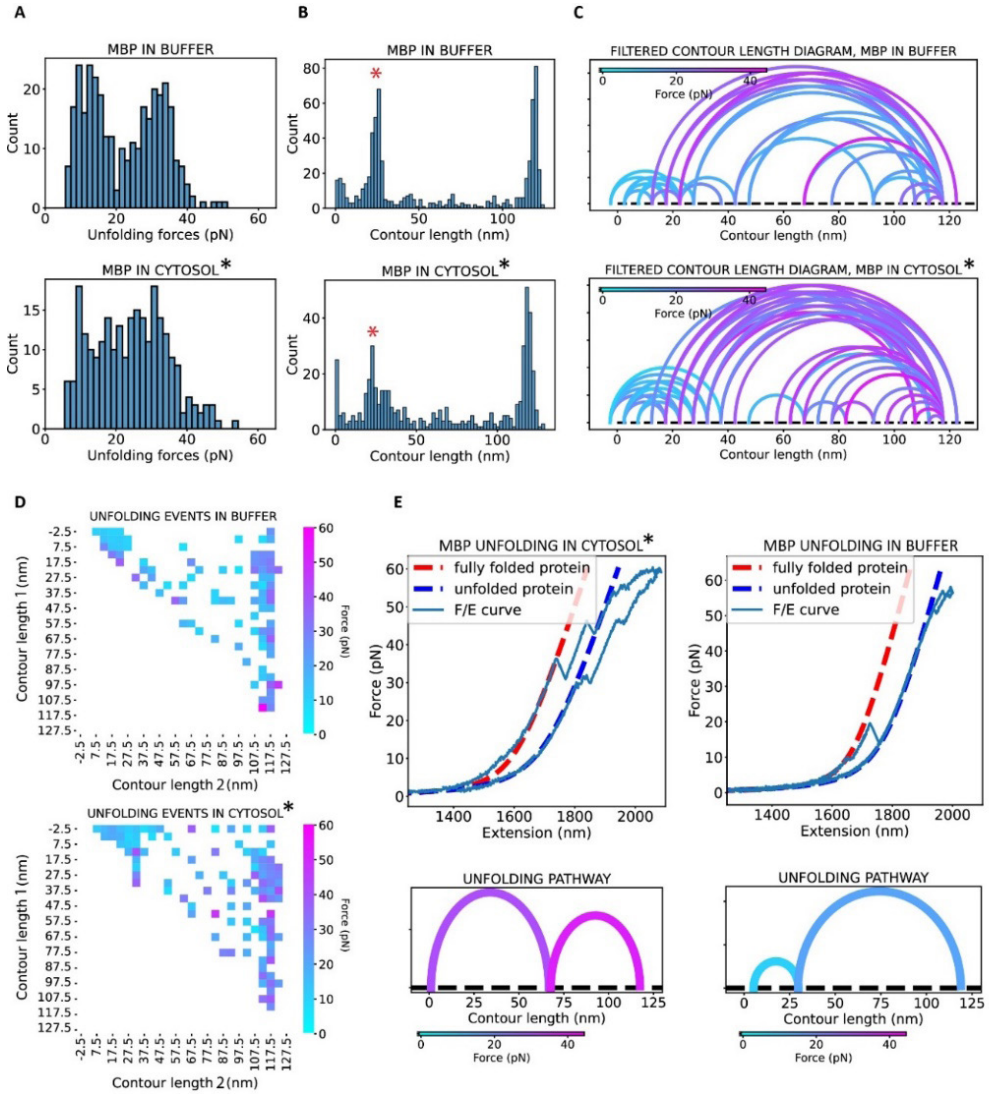

Figure 2. The cytosolic proteome interacts with MBP unfolding intermediates and affects preferred unfolding pathways. Data represented in panel A, B and D of the figure has been collected from N = 42 molecules of MBP (183 pulls) in cytosol, and N = 36 molecules of MBP (221 pulls) in buffer. ***A*** Histogram of MBP unfolding forces in buffer and cytosol. ***B*** Histogram of MBP contour lengths in buffer and cytosol solution. ***C*** Representation of unfolding transitions as circuit diagram. Every arc connects initial and final contour lengths of an unfolding event. Only most frequent unfolding events were kept (last quartile). The color-coding indicates the average force for that specific transition. ***D*** Representation of unfolding events as heatmap. The color-coding indicates the average force for that specific transition. ***E*** Example of force/extension curves for unfolding transitions in cytosol solution and buffer. The label cytosol* in all panels indicates a 4 times dilution of cytosol extract in buffer.

the cytosol solution than in buffer. The histogram of the contour lengths in Figure 2B provides further insight into this complexity. We detect the frequently visited intermediate I state corresponding to the MBP core at 24,5 ± 0,2 nm (see red asterisk in Figure 2B). The same intermediate state in cytosol solution appears to be less prevalent (Figure 2B, lower panel), with a broad peak reaching 40 nm.

In order to make sense of such heterogeneity, we require a smart way to visualize unfolding pathways. We chose to represent transitions on a contour length diagram (Figure 2C). After binning the contour lengths into 5 nm bins, we selected the most frequent unfolding transitions, as defined by initial and final contour lengths. The threshold for frequency is set by the last quartile (0.75) of bin counts, that is, transitions that appear more than 3 times (in buffer) and 2 times (in cytosol solution) are displayed. This filtering allows us to improve the signal to noise ratio of the representation. The color of the arcs represents the average force at which that transition occurs. The same information, without filtering, is shown in Figure 2D in the form of a heatmap. The initial contour length of the transition is displayed along rows, and the final contour length along columns. A few important observations can be drawn from these representations. One main feature is the abundance of high force transitions occurring with initial contour lengths higher than 25 nm (therefore after the I intermediate state), which show to be much more common when the protein is exposed to the cytosolic proteome. An example of these high force transitions is displayed in Figure 2E, alongside the most common pathway for MBP unfolding in buffer. These increased unfolding forces show that cytosolic compounds stabilize intermediate states. This effect is consistent with what previously reported about trigger factor [15], and Dnak and GroEL in the ADP state [14][42].

The circuit diagram also shows a tendency of MBP to present a first unfolding event (in its most common pathway) from its native conformation to contour lengths between 20 and 40 nm in the diluted cytosol (Figure 2C, lower panel). The width of this contour length range might suggest small structural variations affecting the I intermediate corresponding to the MBP core, which has an associated contour length of about 30 nm. A similar picture emerges for rare intermediate states detected at contour lengths between 60 and 80 nm, for which the frequency and unfolding force increase in cytosol solution. In order to provide a guideline to interpret these results, we use the concept of mechanical unfoldons (independently unfolding cores), previously identified in MBP by AFM assays[43]. Four of such structural blocks were identified in MBP unfolding: M1, corresponding to an increase in contour length of ~ 23 nm, attributed to the unfolding of C-terminal alpha helices (residues 296 - 366); M2, corresponding to residues 295 – 244, contributing with a sequential increase of ~ 16 nm, and

	FB (pN)	FC (pN)	state % in buffer	state % in cytosol*
0 nm	11,5 ± 0,5	15,7 ± 1,5	12.5% (N = 41)	13.7% (N = 34)
23 nm	28,1 ± 0,6	27,7 ± 0,9	47.6% (N = 156)	28.2% (N = 70)
39 nm	18,4 ± 3,0	25,2 ± 2,1	3.3% (N = 11)	10.0% (N = 25)
60 nm	12,7 ± 1,4	25,6 ± 3, 0	2.7% (N = 9)	6.04% (N = 15)
81 nm	20,0 ± 2,6	27,2 ± 3, 6	2.1% (N = 7)	3.6% (N = 9)

Table 1. All intermediate states after the I transition happen more frequently and at higher forces in cell extract. Errors on unfolding forces represent standard errors on the mean. The number N of unfolding events for each intermediate state is reported in brackets. The total number of intermediate unfolding events observed for MBP in cytosol solution are N = 152 (from 183 pulls, 42 molecules), while in buffer N = 224 (from 221 pulls, 36 molecules). Percentages of observed states include also the fully unfolded state (120 nm). According to the classification used in this paper [43], the intermediate state at 23 nm corresponds to the release of the first unfoldon (M1). Contour lengths at 39 and 81 nm correspond to unfolding of M2 and M3 respectively.

therefore a cumulative contour length M1 + M2 ~ 39 nm; M3 (N terminal, residues 1 -113), often occurring together with M2 as a single unfolding event, yields M1 + M2 + M3 ~ 81 nm. Finally, M4 unfolds residues 114 to 243, fully unfolding the protein (Figure 3A). These intermediate states (23, 39, 81) can readily be identified in the diagram in Figure 2C; by defining an interval of ± 5 nm around the contour length value, we can calculate the average force at which unfolding occurs in that range. Since an intermediate state around 60 nm seems also to be heavily visited in cytosol solution (and was reported in previous studies of MBP alone and with TF [15]), we shall also consider this contour length in the analysis. The results of this analysis present an interesting trend. The unfolding at 23 nm (*I* intermediate state, MBP core) remains basically unaltered in the cytosol solution, in terms of unfolding forces: $F_{C,23}$ = 27,7 ± 0,9 pN as opposed to an unfolding force in buffer of $F_{B,23}$ = 28,1 ± 0,6 pN. However, as expected, all following intermediate states present a noticeable increase in unfolding force, as reported in Table 1. The observed frequency of such intermediates increases as well roughly 2 to 3 folds in cytosol solution. If we rely on the mapping of unfoldons over MBP sequence (Figure 3A), one might conclude that the biggest variation in stability and frequency is observed from the unfolding at 39 nm onwards, that is, the M3 – M4 independent folding units (1 – 243 residues). Hence, the data indicate that cytosolic components interact mostly with partially folded MBP, specifically with its N-terminal half, increasing its resistance to unfolding. There also seems to be a mild increase in unfolding force for the initial unfolding corresponding to the native conformation of MBP (0 nm, Table 1), although its observed frequency does not increase significantly. These data could suggest interaction between

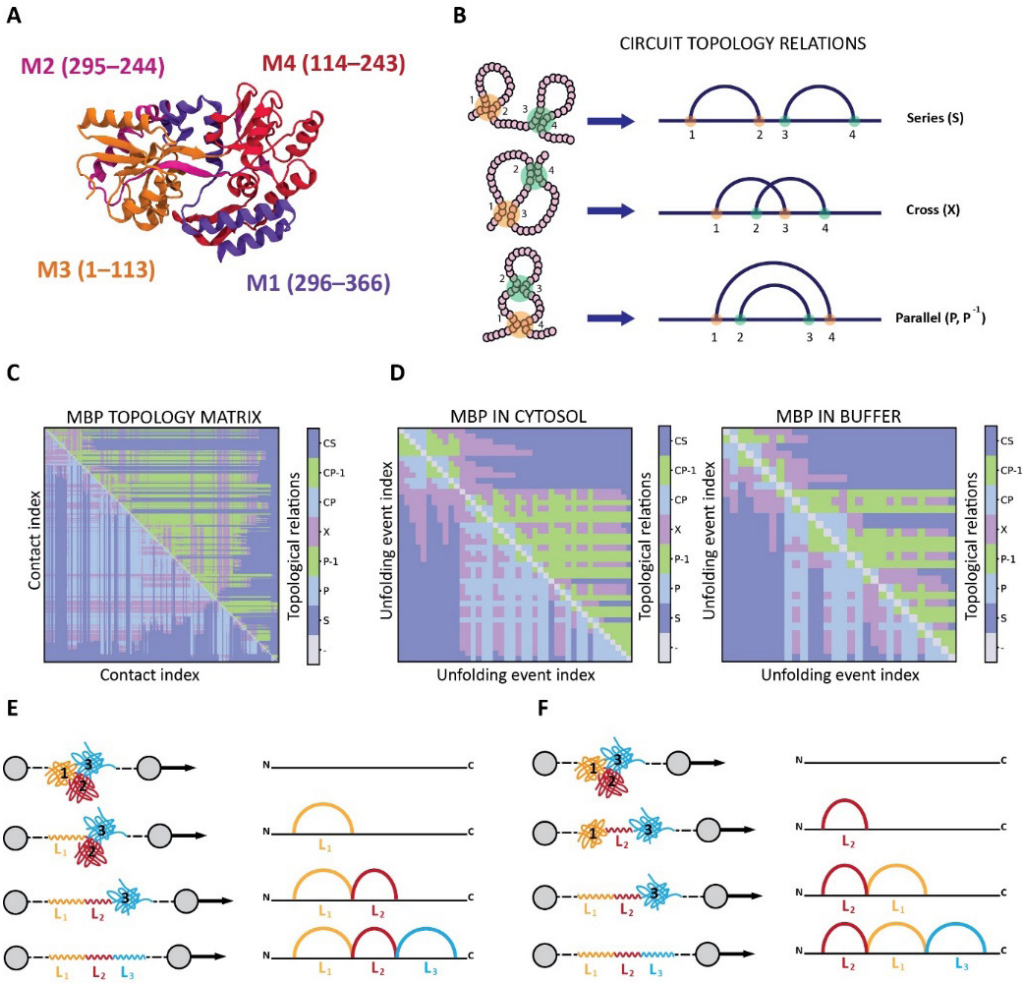

Figure 3. The unfolding pathways can be represented in the topology space. ***A*** Native configuration of MBP. The color coding indicates the subdivision into unfoldons, as identified by Bertz and Rief [43]. ***B*** Summary of the topological relations between loops and their representation as circuit diagram. ***C*** Topology matrix of MBP obtained from its 3D PDB structure. Along rows and columns, specific residue-residue contacts are represented. Each element of the matrix encodes the topological relations between a pair of contacts. In the legend, CP and CS indicate concerted relations: these are a particular class of series and parallel relations, where one contact site is shared among the two contacts. For simplicity, we include these in the count of series and parallel contacts, in the color scheme of the matrix. ***D*** Topology matrix obtained from the unfolding pathway circuit diagram, for transitions in buffer and cytosol solution. The unfolding data used for matrix construction is the same depicted in Figure 2. ***E*** Cartoon of protein pulling where unfolding occurs sequentially, from one end to the other, resulting in a circuit diagram of contour lengths that can be overlapped to protein sequence ($L_1 + L_2 + L_3$). ***F*** Cartoon of protein pulling where unfolding does not occur sequentially, with respective contour length diagram ($L_2 + L_1 + L_3$).

cytosolic chaperones and native MBP.

Although these data present a clear trend, our intermediate states are somewhat a simplification of the heterogeneous pathways we can observe in cytosol solution (Figure 2C, 2D). A way to tackle and visualize this complexity could be the formalism of circuit topology (CT, Figure 3). CT has proven to be successful in the characterization of protein folding *in silico* [44]–[46]. The main assumption of CT is that any two pairs of loops created by intra-chain contacts can be in either 1 of 3 topological relations, as exemplified by Figure 3B. An overall view of the topology of a biopolymer is then encoded into a topology matrix, for instance for MBP (Figure 3C), where each element represents the topological relation between a specific pair of contact-induced loops. Whenever a contact is disrupted in a force spectroscopy assay, the associated loop is released, resulting in a jump in the force/extension diagram of the protein. We can build a topology matrix from any circuit diagram (Figure 2C), by assigning to each arch pair in the diagram a topological relation as described in Figure 3B. The topology matrix of the contour length diagram (Figure 3D) shares the general structure with the one extracted from the native structure of MBP (Figure 3C), specifically the presence of roughly 3/4 lobes made of parallel/cross relations elongating from the diagonal. These lobes correspond to compact, entangled structures. However, the two types of topology matrix do not necessarily always look alike. Structural topology matrices are obtained by drawing a diagram of contacts connecting contact sites, that are then encoded in the topology space (according to the CT formalism) by walking along the sequence from one terminus to the other, generally from N to C. If the unfoldons of a protein were to unfold from one terminus to the other, then one could match the observed unfolding contour lengths back to the structure of the protein. In this case, it would be easy to couple the lobes observed in the unfolding-derived topology matrix with the ones observed in the structural one. (Figure 3E, 3F). If the unfoldons do not unfold sequentially, one would most likely observe different patterns in the two types of topology matrix. In this sense, a significant mismatch of the overall structure of the topology matrices could help us understand the order in which unfoldons are released in a force spectroscopy experiment.

In the case of MBP, the unfolding is almost completely sequential, from C terminus to N terminus, although the stretch between 1 to 113 (M3) residues unfolds (when this particular intermediate state is visited) before the 114 -243 stretch (M4). The cross/parallel-rich lobe on the upper left part of the matrix indicates the initial, C-terminal unfolding (between 0 – 40 nm in contour length, approximately). The remaining cross-parallel block underneath the lobe corresponds to the unfolding of the MBP core. Here the buffer matrix presents long series stri-

pes, indicating that intermediate states are disjointed and easy to identify, while in cytosol solution we do not see this separation, indicating that the compounds in solution are compacting the structure of the core.

We show here this type of visualization as a proof-of-concept of the potential of topology as a bridge between force spectroscopy data and structural analysis. Attributing features in force-extension curves to protein structure is one of the biggest challenges in this type of experiments. MBP is a model protein with a relatively simple unfolding pattern. When faced with bigger, multi-state folding proteins, visualization within the topology space might highlight features not readily visible with traditional methods. As such, matching similarities and differences between the topology of the native structure and that of the unfolding pathway might provide clues to understand protein conformations.

3. CONCLUSION

Proteins undergo a variety of conformational changes and interactions in the cellular environment. Current single molecule experiments are performed in buffer, either in isolation [7]–[9] or in presence of specific chaperone systems and ligands [12][14][15][41]. However, we are far from being able to portray protein folding as it happens *in vivo*, in presence of the full cytosolic proteome. Here, we took a first step in this direction, performing DNA and protein pulling in diluted *E. Coli* cell extract. We explored how the cytosolic molecular machinery affects the mechanical properties of DNA, which need to be taken into account when using chimeric molecular constructs. One could speculate that the observed lower values of persistence length L_P and elastic modulus K could be caused by the action of helicases, separating locally the two sdDNA strands, resulting overall in a more stretchable and bendable molecular structure. We performed the first MBP unfolding assay in presence of the full cytosolic proteome, where MBP acts as a substrate to observe the cumulative topological action of the molecular machinery. What we observed is an overall stabilization of the rare folding intermediates situated mostly in the N-terminal half of the protein (1 – 243 residues), indicating most likely the rescuing of partial folds by chaperones to prevent aggregation. Experiments performed in environments aimed at mirroring real biological environments could serve as a validation of protein folding parameters obtained in *in vitro*, while also providing a closer picture to what these processes really look like in the cell. Several next steps can be envisioned moving forward in this research line. Pulling experiments in dense cytoplasm could elucidate the added effect of crowding on protein folding[47]; the use of active calibration [48] would be advisable in this case, to determine the local viscosity perceived by the beads. The CT-based pipeline for visualization and characterization of for-

ce spectroscopy derived folding pathways could prove fundamental for tracking those bacterial processes where chaperone content undergoes critical changes. One example is antibiotic resistance in *E. Coli*, which is directly linked to chaperone function in the cytoplasm [49]–[51]. In this assay, a model protein, MBP, is used as a sensor for the cytosolic interactome, effectively converting topology into a proxy biomarker for chaperone activity. Moreover, performing such experiments with eukaryotic cells and human proteins would increase the complexity of the systems as well as provide crucial information about disease-triggering phenomena such as protein aggregation and misfolding.

4. METHODS

4.1. Protein expression and purification

Protein expression and purification followed the protocol presented in Avellaneda et al. [28]. MBP was modified with cysteine residues using the pET28 vector. Proteins were expressed E. coli BL21(DE3) cells with 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight at 18°C. Cells were cooled, collected by centrifugation at 5000 × g for 20 min. Cell pellets were resuspended in ice-cold buffer A (50 mM potassium phosphate pH 7.5, 0.15 M NaCl, 3 mM chloramphenicol, 50 mM Glu-Arg, 10 mM Complete Protease Inhibitor Ultra from Roche, 10 mM EDTA) and lysed using a French Press homogenizer. The lysate was centrifuged at 50,000 × g for 60 min at 4°C and incubated with Amylose resin (New England Biolabs) in a gravity column for 20 min at 4 °C. The resin was washed with buffer B (50 mM potassium phosphate pH 7.5, 0.2 M NaCl, 1 mM EDTA, 5 mM DTT) three times and bound proteins were eluted in buffer B supplemented with 20 mM maltose. Purified proteins were analysed on an SDS gel, aliquoted, flash-frozen in liquid nitrogen, and stored at −80 °C.

4.2. Protein-anchor coupling protocol

Protein coupling was performed following the procedure presented in Avellaneda et al. [28]. Purified proteins were thawed and passed through a desalting column (PD-10, GE Healthcare) to exchange the buffer to a coupling buffer (Sodium Phosphate 100 mM pH 7.2, NaCl 150 mM, EDTA 10 mM). The proteins were concentrated using 10K Amicon Centrifugal filter at 14000 × g for 10 min. Anchor oligos 5′-modified with maleimide (purchased from biomers.net) were mixed with the protein 4:1 ratio and incubated for 45 min at RT. 5 mM tris(2-car-

boxyethyl) phosphine (TCEP) was added and incubated for a further 30 min to increase the coupling yield. Uncoupled oligos were removed by affinity chromatography using amylose resin (NEB). The protein-oligo complex was diluted and ligated to 2.5 kbp DNA handles (produced by the Tans lab as per Avellaneda et al. [28]) at a 1:1 ratio with T4 ligase for 1-2 hours at room temperature.

4.3. DNA tightrope generation

DNA tethers were produced by the Tans lab; 5kb DNA tightropes were generated by PCR using biotin and digoxigenin functionalized primers from the pOSIP-TT plasmid.

4.4. Cytosol extraction

Our cytosol solution was obtained from 50 ml DH5α cells transformed with pETM14. The cells were centrifuged at 4000 xG for 10-20 minutes at 4°C. The supernatant was discarded and the cell pellet diluted in 2 µl HMK buffer (50 mM HEPES, pH 7.5, 5 mM MgCl2, 100 mM KCl). The bacterial cells were lysed through 5 freeze-thaw cycles: for each cycle, it was flash-frozen in liquid nitrogen and subsequently left to thaw on ice. The solution was subsequently filtered and aliquoted for use in OT experiments.

4.5. AFS viscosity measurement

In order to measure the viscosity of our cytosol solution, we added 4.07 µm silica beads (SSD5002, Bangs Laboratories) to cell extract and injected the solution into the AFS chamber. Beads were then subjected to acoustic wave until they detached from the surface (voltage: 10 – 30%). Subsequently, the acoustic wave was turned off, and the beads were let free to settle again on the bottom. The z position of the bead was tracked by observing the diffraction pattern of the bead image and comparing it with a previously calibrated look up table. The terminal velocity V_t of the bead was extracted by fitting the z/time plot in the linear range. The viscosity η′ was then obtained by the following equation:

$$\eta' = \frac{g d^2}{18 v_t}(\rho_S - \rho)$$

Where d and ρ_S are respectively diameter and density of the beads, as reported by the manufacturer, g is the gravitational acceleration and ρ is the density of the solution. The viscosity values were then corrected by a factor accounting for surface proximity [52]:

$$\eta = \eta'\left(1 - \frac{9r}{8z} + \frac{r^3}{2z^3} - \frac{57r^4}{100z^4} + \frac{r^5}{5z^5} + \frac{7r^{11}}{200z^{11}}\right)$$

Where r is the radius of the bead. In order to compare viscosity distributions in buffer and cytosol solution, we used the Kolmogorov-Smirnov test, in view of the non-normality of the cytosolic viscosity distribution, assessed by Shapiro-Wilk test ($p = 0.011$).

4.6. Optical Tweezers assay

MBP-DNA and DNA tethers were incubated for 20 minutes on a rotary mixer (4 °C) with 15 µl HMK buffer (50 mM HEPES, pH 7.5, 5 mM MgCl2, 100 mM KCl) and 2.12 µm Anti-Dig coated Polystyrene beads (DIGP-20-2, Spherotech). The solution was then rediluted in 120 µL HMK buffer. Connection with 2.08 µm Neutravidin coated Polystyrene beads (NVP-20-5, Spherotech), on the other terminus of the construct, was created during the OT assay. Particles were trapped and brought in close proximity for a few seconds, allowing the tether to bind. Pulling was performed at a constant speed of 200 nm/s.

4.7. Data analysis

Tethers were selected in accordance to the following criteria: DNA overstretching in the expected range of ~ 65 pN, and pulling tether size matching the expected construct length. We obtained a total of 44 DNA pulls in buffer, from 9 molecules (5 s waiting time), 67 DNA pulls from 12 molecules in cytosol solution at 5 s waiting time, and 43 pulls from 14 molecules in cytosol solution at 30 s waiting time. For MBP pulling experiments, we measured a total of 221 pulls from 36 molecules in buffer and 183 pulls from 42 molecules in cytosol solution. We performed the WLC fitting with the same procedure presented by Avellaneda et al.[28], using the approximation of an extensible polymer for DNA[53], and the Odijk inextensible approximation[54] to account for protein parameters:

$$x = L'_C\left(\frac{4}{3}\left(1-\frac{1}{\sqrt{\beta'+1}}\right)-\left(\frac{10\exp\left(\sqrt[4]{\frac{900}{\beta'}}\right)}{\sqrt{\beta'}\left(\exp\left(\sqrt[4]{\frac{900}{\beta'}}\right)-1\right)^2}+\frac{\beta'^{1.62}}{3.55+3.8\beta'^{2.2}}\right)+L_c\left(1-\frac{1}{2}\sqrt{\beta}\right)\right)$$

Here, $\beta' = (F L'_p)/(k_B T)$, where T is the temperature and F the force. The constants L'_p, K, and L'_c are persistence length, elastic modulus and contour length of the DNA molecule (or overhangs). We set L'_c to 1700 in both DNA and protein pulling assays, as the two overhangs are 2.5 kbp each, resulting in a total of 5 kbp. The constants L'_p and K were fit separately for each tether. Often these parameters had to be re-fit for each different pull when performing experiments in cytosol solution, to account for the environmental effects on the DNA tether described in this paper. The average and standard deviation of these values in buffer and cytosol solution are reported in Table 2.

For the protein contribution, $\beta = (F L_p)/(k_B T)$, L_p and L_c representing the persistence and the contour length, set respectively to 120 and 0.75 nm. The pulling force was 200 nm/s. For the histograms in Figure 2, all rupture forces happening at all contour lengths were retained, in order to provide a full picture of all differences between unfolding in buffer and in cytosol solution.

4.8. Circuit Topology analysis

The topology matrix was obtained by custom code, published previously by Moes et al.[55]. Cutoffs used for contact identification were: distance cutoff = 4.5 a, number of atom-atom contacts = 5, neighbor exclusion = 4. The contacts were numbered from N to C terminus (residue - residue).

In order to extract the topology matrix from the unfolding transition diagram (Figure 2C), we incorporated into the same loop all transitions happening between the same two 5 nm bins, so that they would account for one contact only. Bins were numbered, and bin indexes were used to assign topological relations between pairs of transitions (defined by initial and final contour length bin) by the following mathematical relations[20]:

$$C_{i,j} S C_{r,s} \Leftrightarrow [i,j] \cap [r,s] = \varnothing$$

$$C_{i,j} P C_{r,s} \Leftrightarrow [i,j] \subset (r,s)$$

$$C_{i,j} X C_{r,s} \Leftrightarrow [i,j] \cap [r,s] \notin \{[i,j],[r,s]\} \cup P(\{i,j,r,s\})$$

	Mean (buffer)	St. dev (buffer)	Mean (cytosol*)	S.dev (cytosol*)
K' (pN)	774	111	577	95
L'p (nm)	25	4	24	7

Table 2. Persistence length and elastic modulus of DNA overhangs from WLC fitting.

$$C_{i,j}CSC_{r,s} \Leftrightarrow (([i,j]\cap[r,s]=\{i\}) \vee ([i,j]\cap[r,s]=\{j\}))$$
$$C_{i,j}CPC_{r,s} \Leftrightarrow (([i,j]\subset[r,s]) \wedge (i=r \vee j=s))$$

where ***P*** denotes the power set, that is, all subsets of a set including the null set (ø). Indexes (i, j, r,s), in this case, correspond to the bin indexes of the transition. The same filtering applied for circuit diagrams (Figure 2C) was retained in the topology matrices in Figure 3.

5. REFERENCES

[1] T. Ha, A. G. Kozlov, and T. M. Lohman, "Single-molecule views of protein movement on single-stranded DNA," Annu Rev Biophys, vol. 41, no. 1, pp. 295–319, 2012, doi: 10.1146/annurev-biophys-042910-155351.

[2] A. A. Rebane, L. Ma, and Y. Zhang, "Structure-Based Derivation of Protein Folding Intermediates and Energies from Optical Tweezers," Biophys J, vol. 110, no. 2, pp. 441–454, 2016, doi: 10.1016/j.bpj.2015.12.003.

[3] C. M. Kaiser, D. H. Goldman, J. D. Chodera, I. Tinoco Jr., and C. Bustamante, "The Ribosome Modulates Nascent Protein Folding," Science, vol. 334, no. December, pp. 1723–1727, 2011.

[4] B. Jagannathan and S. Marqusee, "Protein folding and unfolding under force," Biopolymers, vol. 99, no. 11, pp. 860–869, 2013, doi: 10.1002/bip.22321.

[5] I. Heller, T. P. Hoekstra, G. A. King, E. J. G. Peterman, and G. J. L. Wuite, "Optical tweezers analysis of DNA-protein complexes," Chem Rev, vol. 114, no. 6, pp. 3087–3119, 2014, doi: 10.1021/cr4003006.

[6] C. J. Bustamante, Y. R. Chemla, S. Liu, and M. D. Wang, "Optical tweezers in single-molecule biophysics," Nature Reviews Methods Primers, vol. 1, no. 1, 2021, doi: 10.1038/s43586-021-00021-6.

[7] C. Cecconi, E. A. Shank, F. W. Dahlquist, S. Marqusee, and C. Bustamante, "Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers," European Biophysics Journal, vol. 37, no. 6, pp. 729–738, 2008, doi: 10.1007/s00249-007-0247-y.

[8] G. Cecconi, E. A. Shank, C. Bustamante, and S. Marqusee, "Direct observation of the thre-

e-state folding of a single protein molecule," Science (1979), vol. 309, no. 5743, pp. 2057–2060, 2005, doi: 10.1126/science.1116702.

[9] A. F. Oberhauser and M. Carrión-Vázquez, "Mechanical biochemistry of proteins one molecule at a time," Journal of Biological Chemistry, vol. 283, no. 11, pp. 6617–6621, 2008, doi: 10.1074/jbc.R700050200.

[10] A. Mashaghi, G. Kramer, D. C. Lamb, M. P. Mayer, and S. J. Tans, "Chaperone action at the single-molecule level," Chemical Reviews. 2014. doi: 10.1021/cr400326k.

[11] H. Saibil, "Chaperone machines for protein folding, unfolding and disaggregation," Nat Rev Mol Cell Biol, vol. 14, no. 10, pp. 630–642, 2013, doi: 10.1038/nrm3658.

[12] P. Bechtluft et al., "Direct observation of chaperone-induced changes in a protein folding pathway," Science (1979), vol. 318, no. 5855, pp. 1458–1461, 2007, doi: 10.1126/science.1144972.

[13] R. Imamoglu, D. Balchin, M. Hayer-Hartl, and F. U. Hartl, "Bacterial Hsp70 resolves misfolded states and accelerates productive folding of a multi-domain protein," Nat Commun, vol. 11, no. 1, 2020, doi: 10.1038/s41467-019-14245-4.

[14] A. Mashaghi et al., "Alternative modes of client binding enable functional plasticity of Hsp70," Nature, vol. 539, no. 7629, pp. 448–451, Nov. 2016, doi: 10.1038/nature20137.

[15] A. Mashaghi et al., "Reshaping of the conformational search of a protein by the chaperone trigger factor," Nature, vol. 500, no. 7460, pp. 98–101, Aug. 2013, doi: 10.1038/nature12293.

[16] M. J. Avellaneda, E. J. Koers, M. M. Naqvi, and S. J. Tans, "The chaperone toolbox at the single-molecule level: From clamping to confining," Protein Science, vol. 26, no. 7, pp. 1291–1302, 2017, doi: 10.1002/pro.3161.

[17] J. C. Young, V. R. Agashe, K. Siegers, and F. U. Hartl, "Pathways of chaperone-mediated protein folding in the cytosol," Nature Reviews Molecular Cell Biology, vol. 5, no. 10. pp. 781–791, Oct. 2004. doi: 10.1038/nrm1492.

[18] J. Yuan et al., "The effects of macromolecular crowding on the mechanical stability of protein molecules," Protein Science, vol. 17, no. 12, pp. 2156–2166, Dec. 2008, doi: 10.1110/ps.037325.108.

[19] A. Mashaghi, "Circuit Topology of Folded Chains," Not. Am. Math. Soc., vol. 68, pp. 420–423, 2021, doi: 10.1090/noti2241.

[20] A. Mashaghi, R. J. van Wijk, and S. J. Tans, "Circuit topology of proteins and nucleic acids," Structure, vol. 22, no. 9, pp. 1227–1237, 2014, doi: 10.1016/j.str.2014.06.015.

[21] D. Moes, E. Banijamali, V. Sheikhhassani, B. Scalvini, J. Woodard, and A. Mashaghi, "ProteinCT: An implementation of the protein Circuit Topology framework," MethodsX, p. 101861, Sep. 2022, doi: 10.1016/j.mex.2022.101861.

[22] B. Scalvini et al., "Topology of Folded Molecular Chains: From Single Biomolecules to Engineered Origami," Trends Chem, vol. 2, no. 7, pp. 609–622, 2020, doi: 10.1016/j.trechm.2020.04.009.

[23] I. Golding and E. C. Cox, "Physical Nature of Bacterial Cytoplasm," vol. 098102, no. March, pp. 14–17, 2006, doi: 10.1103/PhysRevLett.96.098102.

[24] Y. Jun, S. K. Tripathy, B. R. J. Narayanareddy, M. K. Mattson-Hoss, and S. P. Gross, "Cali-

bration of optical tweezers for in vivo force measurements: How do different approaches compare?," Biophys J, vol. 107, no. 6, pp. 1474–1484, 2014, doi: 10.1016/j.bpj.2014.07.033.

[25] T. M. J. Evers, V. Sheikhhassani, M. C. Haks, C. Storm, T. H. M. Ottenhoff, and A. Mashaghi, "Single-cell analysis reveals chemokine-mediated differential regulation of monocyte mechanics," iScience, vol. 25, no. 1, 2022, doi: 10.1016/j.isci.2021.103555.

[26] E. Pfitzner et al., "Rigid DNA Beams for High-Resolution Single-Molecule Mechanics," Angewandte Chemie International Edition, vol. 52, no. 30, pp. 7766–7771, Jul. 2013, doi: 10.1002/anie.201302727.

[27] Y. Hao, C. Canavan, S. S. Taylor, and R. A. Maillard, "Integrated Method to Attach DNA Handles and Functionally Select Proteins to Study Folding and Protein-Ligand Interactions with Optical Tweezers," Sci Rep, vol. 7, no. 1, pp. 1–8, 2017, doi: 10.1038/s41598-017-11214-z.

[28] M. J. Avellaneda, E. J. Koers, D. P. Minde, V. Sunderlikova, and S. J. Tans, "Simultaneous sensing and imaging of individual biomolecular complexes enabled by modular DNA–protein coupling," Commun Chem, vol. 3, no. 1, pp. 1–7, 2020, doi: 10.1038/s42004-020-0267-4.

[29] A. Aggarwal, S. Naskar, A. K. Sahoo, S. Mogurampelly, A. Garai, and P. K. Maiti, "What do we know about DNA mechanics so far?," Curr Opin Struct Biol, vol. 64, pp. 42–50, 2020, doi: 10.1016/j.sbi.2020.05.010.

[30] C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block, and V. Croquette, "Estimating the persistence length of a worm-like chain molecule from force-extension measurements," Biophys J, 1999, doi: 10.1016/S0006-3495(99)77207-3.

[31] S. B. Smith, Y. Cui, and C. Bustamante, "Overstretching B-DNA: The Elastic Response of Individual Double-Stranded and Single-Stranded DNA Molecules," Science (1979), vol. 271, no. February, pp. 795–800, 1996.

[32] G. A. King, P. Gross, U. Bockelmann, M. Modesti, G. J. L. Wuite, and E. J. G. Peterman, "Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy," Proc Natl Acad Sci U S A, vol. 110, no. 10, pp. 3859–3864, 2013, doi: 10.1073/pnas.1213676110.

[33] H. Fu, H. Chen, J. F. Marko, and J. Yan, "Two distinct overstretched DNA states," Nucleic Acids Res, vol. 38, no. 16, pp. 5594–5600, 2010, doi: 10.1093/nar/gkq309.

[34] J. Van Mameren et al., "Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging," Proc Natl Acad Sci U S A, vol. 106, no. 43, pp. 18231–18236, 2009, doi: 10.1073/pnas.0904322106.

[35] M. J. McCauley, L. Shokri, J. Sefcikova, Č. Venclovas, P. J. Beuning, and M. C. Williams, "Distinct Double- and Single-Stranded DNA Binding of E. coli Replicative DNA Polymerase III α Subunit," ACS Chem Biol, vol. 3, no. 9, pp. 577–587, Sep. 2008, doi: 10.1021/cb8001107.

[36] R. R. Meyer and P. S. Laine, "The Single-Stranded DNA-Binding Protein of Escherichia coli," vol. 54, no. 4, pp. 342–380, 1990.

[37] L. Shokri, M. J. McCauley, I. Rouzina, and M. C. Williams, "DNA overstretching in the presence of glyoxal: Structural evidence of force-induced DNA melting," Biophys J, vol. 95,

no. 3, pp. 1248–1255, 2008, doi: 10.1529/biophysj.108.132688.

[38] J. R. Moffitt, Y. R. Chemla, D. Izhaky, and C. Bustamante, "Differential detection of dual traps improves the spatial resolution of optical tweezers," Proceedings of the National Academy of Sciences, vol. 103, no. 24, pp. 9006–9011, Jun. 2006, doi: 10.1073/pnas.0603342103.

[39] F. Moayed, A. Mashaghi, and S. J. Tans, "A Polypeptide-DNA Hybrid with Selective Linking Capability Applied to Single Molecule Nano-Mechanical Measurements Using Optical Tweezers," PLoS One, vol. 8, no. 1, 2013, doi: 10.1371/journal.pone.0054440.

[40] D. B. Ritchie and M. T. Woodside, "Probing the structural dynamics of proteins and nucleic acids with optical tweezers," Curr Opin Struct Biol, vol. 34, pp. 43–51, 2015, doi: 10.1016/j.sbi.2015.06.006.

[41] V. Aggarwal, S. R. Kulothungan, M. M. Balamurali, S. R. Saranya, R. Varadarajan, and S. R. K. Ainavarapu, "Ligand-modulated parallel mechanical unfolding pathways of maltose-binding proteins," Journal of Biological Chemistry, vol. 286, no. 32, pp. 28056–28065, 2011, doi: 10.1074/jbc.M111.249045.

[42] M. M. Naqvi et al., "Protein chain collapse modulation and folding stimulation by GroEL-ES," Sci Adv, vol. 8, no. 9, pp. 1–10, 2022, doi: 10.1126/sciadv.abl6293.

[43] M. Bertz and M. Rief, "Mechanical Unfoldons as Building Blocks of Maltose-binding Protein," J Mol Biol, vol. 378, no. 2, pp. 447–458, 2008, doi: 10.1016/j.jmb.2008.02.025.

[44] M. Heidari, H. Schiessel, and A. Mashaghi, "Circuit Topology Analysis of Polymer Folding Reactions," ACS Cent Sci, vol. 6, p. 839–847, May 2020, doi: 10.1021/acscentsci.0c00308.

[45] B. Scalvini, V. Sheikhhassani, and A. Mashaghi, "Topological principles of protein folding," Physical Chemistry Chemical Physics, vol. 23, no. 37, pp. 21316–21328, 2021, doi: 10.1039/d1cp03390e.

[46] A. Mugler, S. J. Tans, and A. Mashaghi, "Circuit topology of self-interacting chains: Implications for folding and unfolding dynamics," Physical Chemistry Chemical Physics, vol. 16, no. 41, pp. 22537–22544, 2014, doi: 10.1039/c4cp03402c.

[47] N. Tokuriki et al., "Protein folding by the effects of macromolecular crowding," Protein Science, vol. 13, no. 1, pp. 125–133, 2004, doi: 10.1110/ps.03288104.

[48] M. Fischer, A. C. Richardson, S. N. S. Reihani, L. B. Oddershede, and K. Berg-Sørensen, "Active-passive calibration of optical tweezers in viscoelastic media," Review of Scientific Instruments, vol. 81, no. 1, 2010, doi: 10.1063/1.3280222.

[49] K. I. Wolska, E. Bugajska, D. Jurkiewicz, M. Kuć, and A. Jóźwik, "Antibiotic Susceptibility of Escherichia coli dnaK and dnaJ Mutants," Microbial Drug Resistance, vol. 6, no. 2, pp. 119–126, Jun. 2000, doi: 10.1089/107662900419429.

[50] L. Goltermann, M. v. Sarusie, and T. Bentin, "Chaperonin GroEL/GroES Over-Expression Promotes Aminoglycoside Resistance and Reduces Drug Susceptibilities in Escherichia coli Following Exposure to Sublethal Aminoglycoside Doses," Front Microbiol, vol. 6, Jan. 2016, doi: 10.3389/fmicb.2015.01572.

[51] F. Chiappori, M. Fumian, L. Milanesi, and I. Merelli, "DnaK as Antibiotic Target: Hot Spot Residues Analysis for Differential Inhibition of the Bacterial Protein in Comparison with the Human HSP70," PLoS One, vol. 10, no. 4, p. e0124563, Apr. 2015, doi: 10.1371/journal.pone.0124563.

[52] V. Romanov, G. Silvani, H. Zhu, C. D. Cox, and B. Martinac, "An Acoustic Platform for Single-Cell, High-Throughput Measurements of the Viscoelastic Properties of Cells," Small, vol. 17, no. 3, p. 2005759, Jan. 2021, doi: 10.1002/smll.202005759.

[53] R. Petrosyan, "Improved approximations for some polymer extension models," Rheol Acta, vol. 56, no. 1, pp. 21–26, 2017, doi: 10.1007/s00397-016-0977-9.

[54] T. Odijk, "Stiff Chains and Filaments under Tension," Macromolecules, vol. 28, pp. 7016–7018, 1995.

[55] D. Moes, E. Banijamali, V. Sheikhhassani, B. Scalvini, J. Woodard, and A. Mashaghi, "ProteinCT: An implementation of the protein circuit topology framework," MethodsX, vol. 9, p. 101861, 2022, doi: 10.1016/j.mex.2022.101861.

6. SUPPLEMENTARY

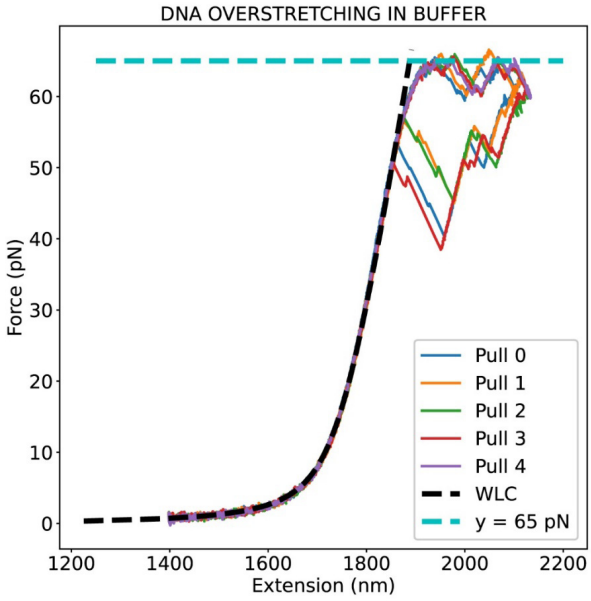

Figure S1. Example of dsDNA pulls performed in buffer. The dashed black line represents the worm-like-chain (WLC) fit, with parameters: persistence length L_P = 30 nm, elastic modulus K = 800 pN.