Three Weaknesses for Three Perturbations: Comparing Protein Unfolding under Shear, Force and Thermal Stresses

Olivier Languin-Cattoen, Simone Melchionna, Philippe Derreumaux, Guillaume Stirnemann, Fabio Sterpone

HAL Id: hal-01930145

https://hal.sorbonne-universite.fr/hal-01930145v1

Submitted on 21 Nov 2018

Article

Three Weaknesses for Three Perturbations: Comparing Protein Unfolding under Shear, Force and Thermal Stresses

Olivier Languin-Cattoen, Simone Melchionna, Philippe Derreumaux, Guillaume Stirnemann, and Fabio Sterpone

Just Accepted

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Three Weaknesses for Three Perturbations: Comparing Protein Unfolding Under Shear, Force and Thermal Stresses

Olivier Languin-Cattoën,[†] Simone Melchionna,[‡] Philippe Derreumaux,[†] Guillaume Stirnemann,[*,†] and Fabio Sterpone[*,†]

†CNRS Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, Sorbonne Paris Cité, PSL University, 13 rue Pierre et Marie Curie, 75005, Paris, France

‡ISC-CNR, via dei Taurini, 00185, Rome, Italy

E-mail: stirnemann@ibpc.fr; fabio.sterpone@ibpc.fr

Abstract

The perturbation of a protein conformation by a physiological fluid flow is crucial in various biological processes including blood clotting and bacterial adhesion to human tissues. Investigating such mechanisms by computer simulations is thus of great interest but it requires to develop *ad hoc* strategies to mimic the complex hydrodynamic interactions acting on the protein from the surrounding flow. In this study, we apply the Lattice Boltzmann Molecular Dynamics (LBMD) technique built on the implicit solvent coarse-grained model for protein OPEP and a mesoscopic representation of the fluid solvent, to simulate the unfolding of a small globular cold shock protein in shear flow, and to compare it to the unfolding mechanisms caused either by mechanical or thermal perturbations. We show that each perturbation probes a specific weakness of the protein, and causes the disruption of the native fold along different unfolding pathways. Notably, the shear flow and the thermal unfolding exhibit very similar pathways, while because of the directionality of the perturbation, the unfolding under force is quite different. For force and thermal disruption of the native state, the coarse-grained simulations are compared to all-atom simulations in explicit solvent, showing an excellent agreement in the explored unfolding mechanisms. These findings encourages the use of LBMD based on the OPEP model to investigate how a flow can affect the function of larger proteins, e.g. in catch-bond systems.

Introduction

Many biological processes depend on the protein response to the mechanical perturbations arising from a solvent in motion.[1,2] For example, during blood coagulation, a large concatemeric blood protein, the von Willebrand Factor (vWf), senses the modifications in the bloodstream caused by wounds in the vessels.[1] The local increase of shear flow triggers conformational changes that lead to the exposure of anchor domains to platelets and to the subendothelium, thus allowing the formation of a hemostatic plug.[1,3–5] Flow-induced conformational changes also activate the so-called catch-bond proteins like FimH that anchors

bacterial cells to the epithelium in urinary tract infections.[1,6,7] In this case, the shear flow acts as an allosteric effector[8,9] that induces a conformational shift toward protein states of high binding affinity with the membrane-linked substrate.

Beyond these examples of functional activation, it has been debated if and to which extent shear or elongational flows unfold – at least partially – proteins.[10–14] Several experimental studies and numerical estimates concluded that only very high shear rates, $\dot{\gamma} > 10^7\,\mathrm{s}^{-1}$, can unfold mid-size globular proteins ($\sim 10^2$ amino acids, a.a.)[10,15,16] However, a few studies reported conformational perturbations even at low shear rate ($\dot{\gamma} \simeq 10^2\,\mathrm{s}^{-1}$) for proteins such as lyosozyme (~ 130 a.a.) and bovine serum albumin (~ 580 a.a.).[11] Overall, the alteration of a protein native state by external hydrodynamic forces seems to be highly sensitive to its distinct structural features, global shape and internal cohesion.[10,16–18] While physiologically attainable shear rates cannot unfold isolated small globular proteins, for large systems such as the multimeric vWf, the tensile force caused by physiological shear $(10^5\,\mathrm{s}^{-1})$[3] acts down at the nanometer length scale on the small functional domains, e.g. causing the (partial) unfolding of the A2 domain.[19] Moreover, and notably, the fluid perturbation can trigger protein aggregation, including Amyloid-β aggregation in Alzheimer's disease,[20–22] and stirring is used to accelerate and tune the aggregation path possibly by altering the protein's misfolded states.[23] Studying the detailed mechanisms of such processes is thus of theoretical and practical interest.

From the experimental point of view, the study of protein perturbation under shear has been regularly complemented by single-molecule experiments where an external mechanical force is used to perturb the protein fold, using an atomic force microscope (AFM), or optical/magnetic tweezers.[2,24–26] However, the directional pulling force applied to the protein extremities is very distinct from the delocalized and fluctuating tensile forces generated by a shearing fluid.[16,27–29] Interestingly, it was shown experimentally that when subjected to pulling forces along different directions, a model protein (SH3) exhibits two alternative unfolding paths.[30]. It is therefore important to assess how the unfolding mechanism monitored

via mechanical pulling compares to that under shear perturbation for small globular proteins as well as for large multimeric protein chains.

Computational approaches are very helpful to investigate the detailed mechanisms of protein unfolding, and therefore to support the interpretation of single-molecule and ensemble-averaged experiments. For instance, thermal, chemical and force unfolding processes have been compared using all-atom simulations.[31–33] In this respect, all-atom simulations are rarely used to study shear effects,[34,35] mainly because of technical challenges. Therefore, the impact of shear on protein stability has been mostly studied with implicit coarse-grained (CG) models of various complexity coupled to the fluid velocity gradient modelled with several different strategies.[3,16,28,29,36,37]

In this work, following previous investigations on model peptides[16] and proteins,[38,39] we deploy the Lattice Boltzmann Molecular Dynamics technique[40] (LBMD) to investigate the shear unfolding of a cold shock protein, and we compare it to force and thermal unfolding processes. LBMD allows to include in a natural way hydrodynamic interactions (HI) in implicit-solvent, particle-based molecular simulations, and most importantly, it allows to generate a shear flow instantaneously coupled to the protein conformational dynamics. This study is based on the implicit-solvent, CG model for protein OPEP (Optimized Potential for Efficient peptide structure Prediction).[41] For thermal and force unfolding, the results are also compared to previous all-atom simulations in explicit solvent.[33]

Our results show that the unfolding of a model system, the cold shock protein (CspA), occurs on a timescale of several nanoseconds only at high shear rates, $\dot{\gamma} \simeq 10^9\,\mathrm{s}^{-1}$, in agreement with our previous estimations for smaller peptides.[16] Within the framework of our description, we find an equivalence between the unfolding timescales observed at these shear rate values and those occurring under a mechanical force in the range $200-300$ pN. However, shear and force unfolding proceed along very different reaction pathways. In particular, shear unfolding involves protein rotation so that the delocalized action of the perturbing tensile force is more akin to thermal unfolding, since it challenges different weak spots of the protein

scaffold.

Methods

Lattice Boltzmann Molecular Dynamics

In the last decades, the Lattice Boltzmann method has become very popular to study fluid simulation problems, since it allows to reproduce the Navier-Stokes dynamics with excellent computational efficiency and scalability.[42,43] The method is based on following a collective group of fluid particles and by characterizing the fluid state by densities of probability f_p to find fluid particles with a given discrete velocity $\vec{c}_p$ at a given place in a 3-dimensional lattice. At each step, the probability densities are propagated and updated according to a collision scheme. The Bhatnagar-Gross-Krook (BGK) scheme[44] is the simplest, yet effective way to recover hydrodynamics behavior and consists of a relaxation to the local equilibrium distribution with a rate τ^{-1} related to the kinematic viscosity of the fluid ν. This scheme can be summarized by the following equation:

$$f_p(\vec{x} + \vec{c}_p \Delta t, t + \Delta t) = f_p(\vec{x}, t) + \tau^{-1}(f_p^{eq} - f_p) + g_p \tag{1}$$

where the first two terms on the right hand side represent the streaming and collision contributions, and g_p is a term incorporating external forces acting on the fluid. In Lattice Boltzmann Molecular Dynamics (LBMD), we couple the molecular motion in continuous space to the fluid evolution in the discrete lattice space.[45,46] This coupling is caused by a drag force

$$\vec{F}_i^D = -\Gamma(\vec{v}_i - \vec{\tilde{u}}(\vec{x}_i)) \tag{2}$$

where Γ is a friction coefficient, $\vec{v}_i$ is the velocity of the i-th particle, and $\vec{\tilde{u}}$ is the fluid velocity extrapolated from the lattice according to the particle position $\vec{x}_i$. The force acts

symmetrically on the fluid, via the term g_p, and on the particles according to the action-reaction principle. For the particle, the drag force adds up with the conservative forces derived from the Hamiltonian of the system $F_i^C = -\nabla_i U(\{r\})$, and a random white noise, F_i^R, that represents thermal fluctuations. A full description of LBMD using an implicit solvent CG model for biomolecules is beyond the scope of this article and the interested reader will find a comprehensive presentation in previous work.[40]

In this study, we used LBMD simulations to study the unfolding of a globular protein under shear flow, and to compare it to force and temperature unfolding. A simulation timestep $\tau_{LBMD} = 1.5\,\text{fs}$ and a friction coefficient $\Gamma = 0.1\,\text{fs}^{-1}$ were used. Particle and fluid dynamics were evolved synchronously. The lattice space supporting the fluid dynamics was $\Delta x = 2.5\,\text{Å}$, and the kinematic viscosity of the fluid was set to the value of pure water at ambient condition $(10^{-6}\,\text{m}^2\,\text{s}^{-1})$. A detailed discussion about the size of the grid spacing used to couple LB to a quasi-atomistic force field for biomolecules can be found elsewhere.[40,47]

The OPEP Force Field

In our simulations, we used the implicit-solvent force field OPEP to model the protein system. OPEP (Optimized Potential for Efficient peptide structure Prediction) was developed to model soluble peptides and proteins with a high structural accuracy while remaining free of any biases constraining the secondary structure.[48–50] It offers the opportunity to efficiently collect large statistics on dynamical and structural features of proteins while keeping a quasi-atomistic resolution for the system of interest. OPEP has been successfully applied to a wide range of problems, from peptide structure prediction to the study of molecular crowding and amyloid aggregation.[41]

System

For the purpose of our study, we selected the major cold shock protein from *E. coli* (CspA, PDB: 1MJC)[51] as a model system. This globular protein consists of 69 residues that fold

into a β-barrel made of five β-strands, hereafter denoted from β_1 to β_5 (Fig. 1). It belongs to a large family of homologous proteins that exhibit multiple examples of adaptation to high-temperatures among thermophilic bacteria. For that reason, mechanical properties of cold shock proteins have been extensively studied both experimentally and in simulations.[33,52–55] Moreover, the unfolding of CspA under force and high temperature was previously investigated by our group using all-atom simulations,[33] providing an interesting reference for the present work.
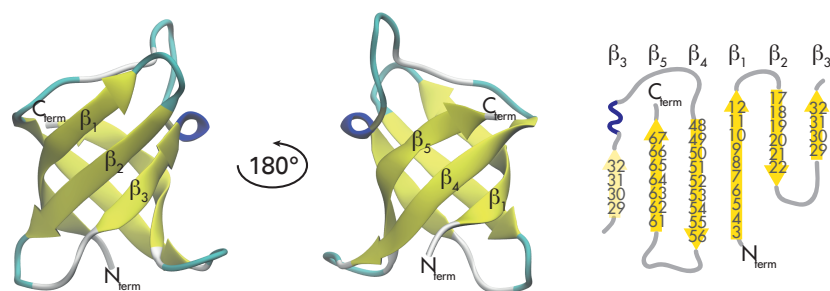
Figure 1: **Cold shock protein A from _E. coli._** In the left part of the figure, the protein is represented by two side-views highlighting the secondary structure motifs. The amino acid sequence along the protein structure is represented on the right.

Results and Discussion

Shear Unfolding

In order to observe and to quantify the effect of a constant shear flow on CspA, we first generated twenty initial configurations representative of the protein native state by performing a MD simulation at ambient temperature (300 K) in the absence of HI and extracting the protein configuration every 300 ps. For each initial configuration, the protein was then placed in a simulation box of size $L_x = 200\,\text{Å}$, $L_y = 100\,\text{Å}$ and $L_z = 150\,\text{Å}$ with a random orientation. Before activating the shear flow, the system was equilibrated for a short time

(5 ns) by including HI.

A shear flow was generated in the simulation box by imposing constant forces in opposite directions to the fluid on the top ($z = L_z$) and bottom ($z = 0$) faces. This simulation setup qualitatively reproduces a Couette cell, where the walls enclosing the fluid move with respect to each other so as to generate a laminar flow with a linear transversal velocity gradient.[16] We have checked that in the absence of the protein, transverse velocities vary linearly along the z-axis, with a slope that directly yields the shear rate $\dot{\gamma}$ (in s^{-1}), and which is proportional to the forces set at the boundaries. In the presence of the protein, we monitored the velocity profile in the molecule's vicinity, which underlines the very natural coupling between the particle and fluid dynamics in LBMD simulations, see Ref. 16. A steady-state profile is observed after a few picoseconds, which is short compared to the protein timescales.

In our simulations, several shear rate values in the range $2 \times 10^9 \, s^{-1}$ to $10^{10} \, s^{-1}$ were employed. It is important to mention that these values correspond to very high rates that are hardly attainable in physiological or experimental conditions. However, such high rates were necessary to explore the unfolding of CspA on the nanosecond timescale (our simulations were extended up to full protein unfolding, which typically occurred in the range 10–100 ns, depending on the shear rate).

In Figure 2, we show a typical unfolding event of the CspA protein under shear. As already described for soft object and polymers in shear flow, the protein experiences two characteristic motions as a result of the decomposition of the velocity field.[16,17,29] The protein is affected by a rotational motion and the fold is altered by elongational forces acting along the shear eigenvector. These features can be understood by the fact that a simple shear flow velocity field can be expressed in terms of rotational and elongational components, which lead to a periodic tumbling movement,[56] where the molecule alternates between compact and extended states as the tensile forces align periodically with its axis of weaker mechanical resistance. During the early stage of shear perturbation, the periodic distortion is somehow reversible, and can be seen as an elastic deformation of the global structure. This charac-
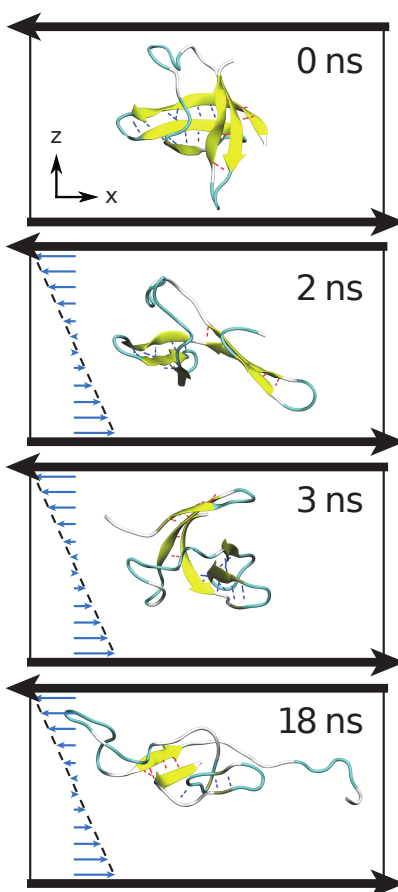
Figure 2: **Protein unfolding under shear stress.** Snapshots showing the protein characteristic tumbling motion, alternating extension (e.g., at 2 and 18 ns) and compaction (3 ns) phases. The progressive loss of secondary structure can be appreciated.

teristic behavior is well represented by the time evolution of gyration radius, R_g, that is marked by regular oscillations with periodicity proportional to the inverse of the shear rate, $\tau_R \propto 1/\dot{\gamma}$ (Fig. 3).[29] The oscillatory behavior can be monitored also in the time evolution of the root mean square displacement (RMSD) (Fig. S1 in the SI).

The protein softness and ability to recover quickly its original conformation may strongly depend on the nature of its fold. For example, in the case of a simple β-hairpin peptide, the unfolding is triggered by the tensile force acting on the hydrogen-bond (HB) ramp that stabilizes the parallel β-sheet.[16] The intensity of the fluid drag force decreases periodically the barrier for HB separation, eventually leading to the sequential but fast unzipping of the

all the HBs in the ramp. For the small β-hairpin peptide, no stable intermediate has been observed.[16] On the contrary, in CspA, the β-barrel can be partially opened at the contact between β₃ and β₅, the rest of the protein being intact, which leads to partially-unfolded intermediates.

The sequential separation of secondary structure motifs, and the presence of some intermediate partially unfolded states, has already been observed in the shear unfolding of the A2 domain of the vWf protein.[38] For this protein, the partial unfolding may have functional impact by helping the quick exposure of the cleavage site of the protein under blood shear perturbation.[2] The presence of intermediate states associated to different dissociations of secondary structures is presently a very interesting focus in single molecule experiments that probe unfolding by applying external directional mechanical forces.[55]

In order to quantify the progressive unfolding of a protein under shear, a metric that is mostly insensitive to the periodic tumbling is required. Some frequently used collective variables such as the root-mean-square-displacement (RMSD) or the radius of gyration, albeit suitable for thermal or chemical denaturation, are here dominated by the oscillating extension of the global shape (Fig. 3 and Fig. S1), and do not capture well the irreversible loss of the native structure. Therefore, the unfolding of CspA is better described by considering a native contact analysis based on the average pairwise distances between the backbone C_α extracted from a short reference simulation with no shear. The fraction of native contacts is then computed for each configuration according to

$$Q(t) = \frac{1}{Q^{ref}} \sum_{i}^{N} \sum_{\substack{j \\ |i-j| \geq 4}}^{N} Q_{ij}(t), \tag{3}$$

where the matrix element Q_{ij} for residues i and j is equal to 0 if the reference distance between the two C_αs is $d_{ij}^{ref} > r_c$, $e^{-\frac{1}{2}(d_{ij}(t)-d_{ij}^{ref})^2}$ if $d_{ij}^{ref} < r_c$ and $d_{ij}(t) > d_{ij}^{ref}$, or 1 if $d_{ij}^{ref} < r_c$ and $d_{ij}(t) < d_{ij}^{ref}$, with $r_c = 10$ Å. The normalization factor Q^{ref} is the number of contacts for C_αs within the cut-off r_c. The condition $|i-j| \geq 4$ guarantees that neighboring

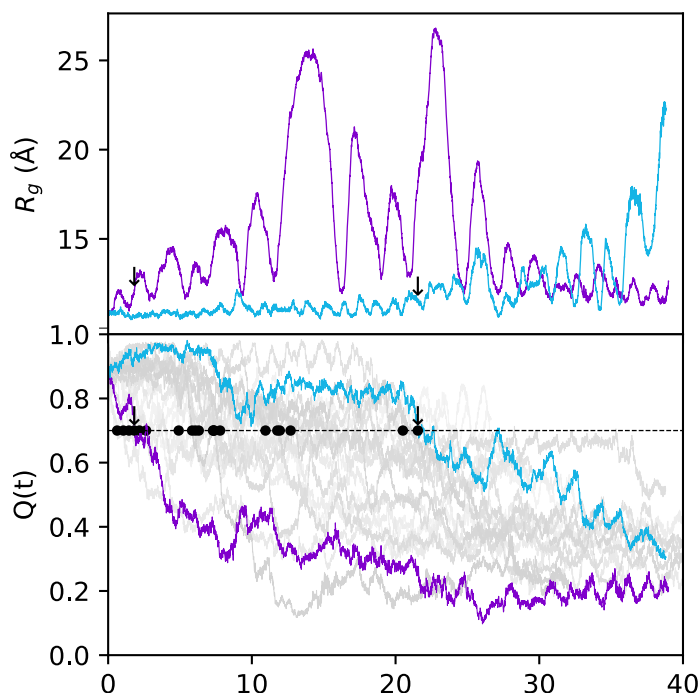amino acids are not taken into account, and the exponential term allows a smoother evolution of $Q(t)$.[57]

Figure 3: **Unfolding under shear.** Top panel. Time evolution of the protein radius of gyration, R_g, as a function of time for two unfolding trajectories, see also curves highlighted in the bottom panel. Bottom panel. Decay of the fraction of native contacts under constant shear. Traces of 20 replicas (light grey) from simulations using a shear rate of $4 \times 10^9 \, \text{s}^{-1}$. The first passage times (FPTs) at the threshold value $Q = 0.7$ are indicated for each replica by black dots. Two trajectories (purple and blue lines) and their respective FPTs (black arrows) were highlighted for clarity purposes.

The decay of the native contact fraction for each individual trajectory reveals highly heterogeneous unfolding paths (Fig. 3). A mean first passage time (FPT) analysis can be used to characterize the unfolding kinetics and the dependence on the shear perturbation. We selected an arbitrary threshold of $Q = 0.7$, which was found to be a good compromise since lower values were attained too slowly at weak shear rates and higher values were too close to the native structure. However, thresholds of $Q = 0.6$ and $Q = 0.8$ yielded qualitatively similar results.

The distributions of inverse FPTs (unfolding rates) and their averages are plotted against

the shear rate in Fig. 4. Overall, the average unfolding rate shows a typical logarithmic dependence with the shear rate $\dot{\gamma}$ that reminds the typical kinetic of unfolding for a protein under a pulling force.[16] Only the highest value $(9.3 \times 10^9 \, \text{s}^{-1})$ departed from the trend. We have already observed for other model peptides/proteins[16] that at high shear rates the kinetics reaches a plateau. An explanation can be proposed for the behavior in this high shear rate regime: the unfolding may no longer be determined by the necessity to overcome energetic barriers but it is limited by the diffusion on the free energy landscape.[16] In fact, a similar behavior is observed when unfolding is monitored by applying an external force and the force value is so high that the limiting factor is diffusion rather than energetics.[27,31,58]

The analogy with the force induced unfolding is strong, but as already noted, the shear drag force acts differently on the protein depending on the tumbling dynamics. As discussed in a previous work,[16] for the simple β-hairpin peptide this rotation exposes the reaction coordinate of unfolding (the axis of the HBs ramp) to the fluid tensile force with a periodicity that is inversely proportional to the shear rate $\dot{\gamma}$. In that case one can describe the mechanical perturbation as a fluctuating external force acting on the reaction coordinate x, e.g. $f_x(x, \dot{\gamma}) = f_0(x, \dot{\gamma}) \cdot cos(\dot{\gamma} t)$. For CspA a clear unfolding reaction coordinate is harder to define, and, as we will show in the following sections, a direct comparison with other unfolding processes, namely under external force or high temperature, is useful.

Force Unfolding

In this section, we discuss the unfolding of the CspA protein in an *in silico* simulation setup mimicking typical force spectroscopy experiments. Similarly to the shear unfolding, we generated a set of $\sim$20 initial configurations from a trajectory at ambient condition in the absence of external force. We then fixed the C_α carbon of the N-terminal residue, and we applied a constant pulling force of 400 pN to the C-terminal C_α. Each simulation was extended to observe partial or total unfolding of the initial configuration. Unfolding was followed as in the experiments by monitoring the protein end-to-end distance, see Fig. 5. A
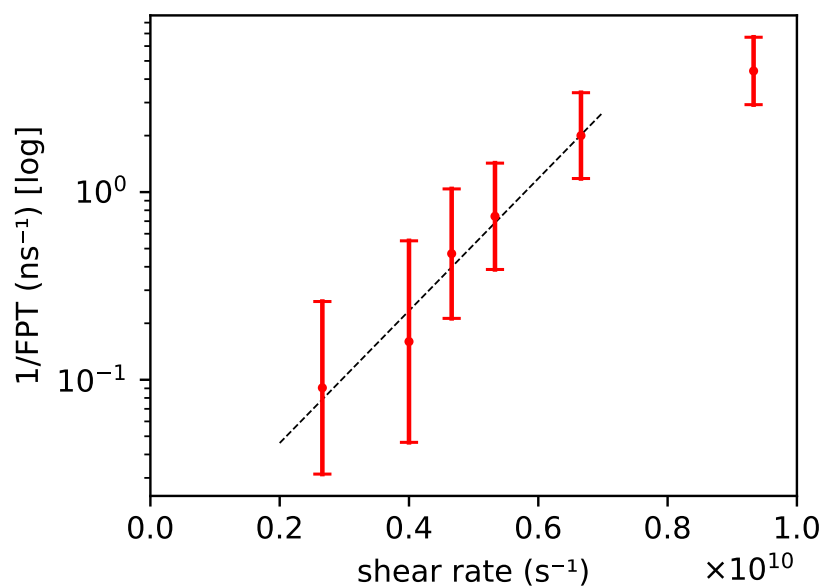
Figure 4: **Kinetics of unfolding as a function of shear rate.** The inverse first passage times for unfolding are extracted from the fraction of native contacts (criterium $Q = 0.7$, see Fig. 3) and plotted against shear rate for sets of replicas. The average measure (red) follows a linear trend at lower rates (dashed line) and seems to reach a plateau above $7 \times 10^9\,\mathrm{s}^{-1}$.

molecular view of a typical unfolding trajectory under force is also shown in Fig. S2. The unfolding proceeds via multiple extensions because of the presence of long-lived intermediates, as illustrated by the cumulative histogram on the right-hand side of Fig. 5. We note that this "staircase" motif, broadly observed among computational studies at constant force,[27,32,55,59] is hardly observed in experimental measurements, where unfolding is generally well-fitted by a two-state model,[60] because of the time resolution of the instrumentation ($\approx$ ms). However, long-lasting intermediate states have recently received experimental support from finely controlled force-clamp and force-ramp experiments, suggesting that unfolding generally occurs along highly heterogeneous paths.[55]

In the specific case of CspA, we had demonstrated in a previous study[33] based on all-atom steered molecular dynamics using the CHARMM36 force field[61] that the presence of a long-lived initial intermediate around $3\,\mathrm{nm}$ was due to a key structural motif, i.e., the loop L_3 connecting the β_3 and β_4 strands. We had shown that when the protein was folded, this

loop hindered the unzipping of the β_4 and β_5 strands. It is only when the L_3 loop moves aside that this obstruction is removed, and that β_4-β_5 unzipping and thus protein extension can occur. The same mechanism was indeed observed in our very different simulation setup based on the implicit solvent CG model OPEP, as shown in Fig. S3, further reinforcing the validity of our modelling.
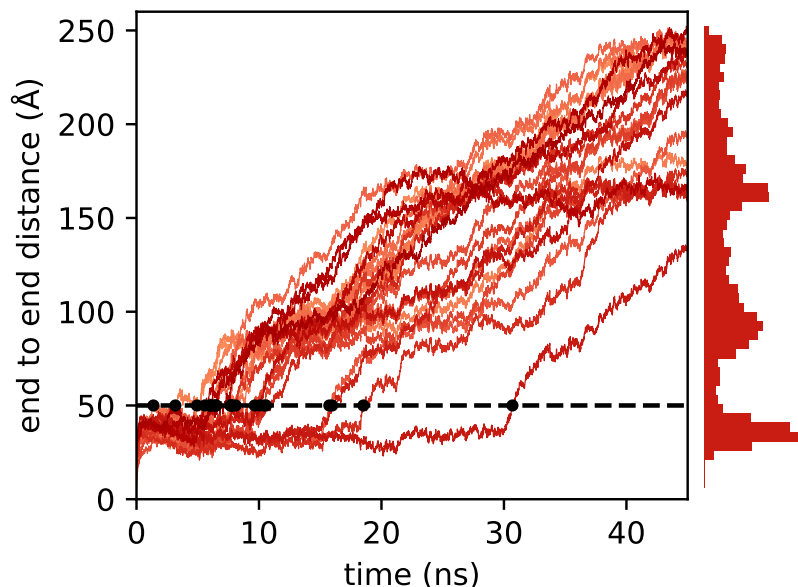

Figure 5: **Multiple-step unfolding under force**. The trace of the CspA unfolding trajectory under a $400\,\mathrm{pN}$ force are plotted, namely we report the time evolution of the end-to-end distance. The threshold $d = 50\,\text{Å}$,is reported as dashed line, and the unfolding time for each trajectory is indicated by a black circle. On the lateral side of the graph, it is reported the accumulated histogram (first 45 ns) of the end-to-end distance during the unfolding event. The presence of multiple intermediated states is indicated by the presence of multiple peaks in the histogram.

To examine the force-dependence of the unfolding kinetics, we repeated the simulations by varying the force in the range 300–$450\,\mathrm{pN}$ (Fig. 6). The first passage time was defined for each trajectory as corresponding to the first crossing of a $50\,\text{Å}$ threshold, which is located beyond an initial, long-lived intermediate that corresponds to a slightly distorted by still folded native state. The average unfolding rate shows an exponential dependence upon force, as expected from the Bell model,[62] $k = k_0 \exp\left(\frac{-\Delta G^{\ddagger} + F\Delta x}{k_B T}\right)$ where Δx is a phenomenological parameter

measuring the distance between the reactant and transition states, assumed not to be affected by the force. For the sake of comparison we have reported in the same plot the results from atomistic simulations[33] performed at 200 and 300 pN. From the model it is possible to extract the phenomenological parameter $\Delta x \simeq 0.3$ Å for both CG and all-atom simulations, which is smaller than the typical value estimated for proteins from experiments, that are usually 1 order of magnitude higher.[53] The prefactor k_0 can be estimated using standard unidimensional transition state theory $(\frac{k_B T}{h})$, and in this case the extrapolated rate constant at zero force leads to a reference free energy barrier $\Delta G^{\ddagger} = 9$ kcal/mol in the absence of force. For all-atom simulations (blue triangles), the value of the free energy barrier is close to 6.9 kcal/mol, in good agreement with the present CG simulations (red circles). An alternative approach is to use Kramers' expression in the high friction limit for the extrapolated rate constant at zero force, in which the reaction rate is expressed after some simplifications (see SI) as $k = \frac{D}{2\pi\sigma^2} \exp\left(\frac{-\Delta G^{\ddagger}}{k_B T}\right)$, where D is the effective diffusion coefficient along the end-to-end coordinate accounting for both the coupling with the solvent and internal contributions, and which can be numerically estimated using harmonic constraints[63] (see SI), while σ represents the standard deviation of the end-to-end distance in the reactant state. The resulting free energy barriers are smaller, $\Delta G^{\ddagger} = 4.4$ kcal/mol and 3.5 kcal/mol, for the CG and all-atom simulations, respectively. It is worth recalling that the energy scale of the CG model OPEP can be modulated so as to tune the propensity to maintain the protein fold stable. The comparison with the all-atom simulations indicates that the actual setup is reasonable for inquiring the mechanical stability of the protein. Experiments on another analogous mesophilic cold-shock protein have reported free-energy barriers of 2.9 kcal/mol for thermal denaturation and 9.1 kcal/mol for the unfolding upon force.[53] Note however that this last value was obtained assuming a given rate constant prefactor, which is largely setup-dependent[64] and that was not specifically determined in these experiments, thereby suggesting that the experimental free-energy barrier might be overestimated.
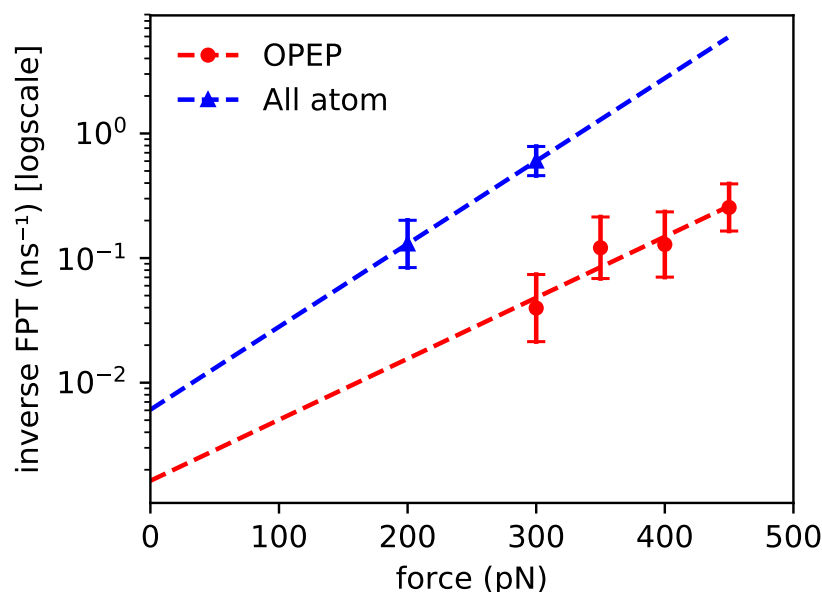
Figure 6: **Kinetics of unfolding as a function of the pulling force.** The inverse first passage times at an end-to-end distance of $50\,\text{Å}$ are plotted against force for CG simulations (red) and atomistic simulation based on CHARMM36 (blue).[33] In both cases the mean FPTs follow a linear trend that matches the Bell model.

Comparing Thermal, Force and Shear Unfolding

In this section we compare more extensively the unfolding pathways of CspA under shear, force and thermal perturbations. To this purpose, a set of simulations were performed at $T = 430\,\text{K}$ to observe thermal unfolding on the nanosecond timescale. CspA consists of a β-barrel structure with two groups of interacting β-sheets that are more easily evidenced in Fig. 1, $β_1$-$β_2$-$β_3$ and $β_4$-$β_5$, with some extra HBs between $β_1$ and $β_4$. We therefore decided to project the trajectories of the unfolding events along the two collective coordinates consisting in the number of HBs in the $β_1$, $β_2$, $β_3$, and $β_1$, $β_4$, $β_5$ sheets, denoted $S_β^{123}$ and $S_β^{145}$, respectively:

$$S_β^X(t) = \sum_{i \in X} \sum_{\substack{j \in X \\ |i-j| \geq 4}} S_{ij}(t), \tag{4}$$

where the sum i and j run over the backbone HB donors (the amide N atom) and HB acceptor (the carbonyl oxygen atom) in the β-sheets $X = 123$ or 145. The sums consider

only amino acids separated by four residues, $|i - j| \geq 4$. The term $S_{ij}(t)$ is a smoothed number of HBs, which only depends on the distance between the HB donor and acceptor. It is equal to 1 if this distance is smaller the $d_c = 3\,\text{Å}$, and otherwise is given by $e^{-\frac{1}{2}d_{ij}(t)^2}$.

Given these collective coordinates, we can now compare the three types of perturbation (Fig. 7). In shear and thermal unfolding, the β_1-β_2-β_3 interactions break first, and β_1-β_4-β_5 is more conserved, at least on the explored timescale (50 ns). Under force, the unfolding proceeds along a different path, with the β_1-β_4-β_5 unzipping first, and then a progressive unfolding of β_1-β_2-β_3. Interestingly, because of protein tumbling under shear, the protein experiences a tensile perturbation that does not act along a fixed direction, as in pulling experiments, and the mechanical perturbation is much more distributed on the protein fold. The shear perturbation is therefore more equally distributed onto the protein structure, just as thermal excitation. A simplified scheme of the unfolding paths in the reduced free energy landscape is given in Fig. 7B.

For force and thermal unfolding, we have also reported the data from all-atom simulations.[33] For the force unfolding we have extracted the average unfolding paths considering 25 independent trajectories at $F = 200\,\text{pN}$. For thermal unfolding, we have considered data obtained with replica-exchange simulations performed with the CHARMM36[61] forcefield, each replica evolving for 500 ns at different effective temperatures in the range $[306 - 407]$ K.[33,65] The agreement between the results from the all-atom model and from the OPEP model is very good, suggesting that the CG OPEP model is able to capture the essential aspects of mechanical and thermal response of proteins.

Conclusion

In this work we have compared the unfolding mechanism of a small globular cold shock protein, CspA, under three external perturbations: shear flow, mechanical pulling and thermal excitation. While thermal and force unfolding have been extensively investigated along the
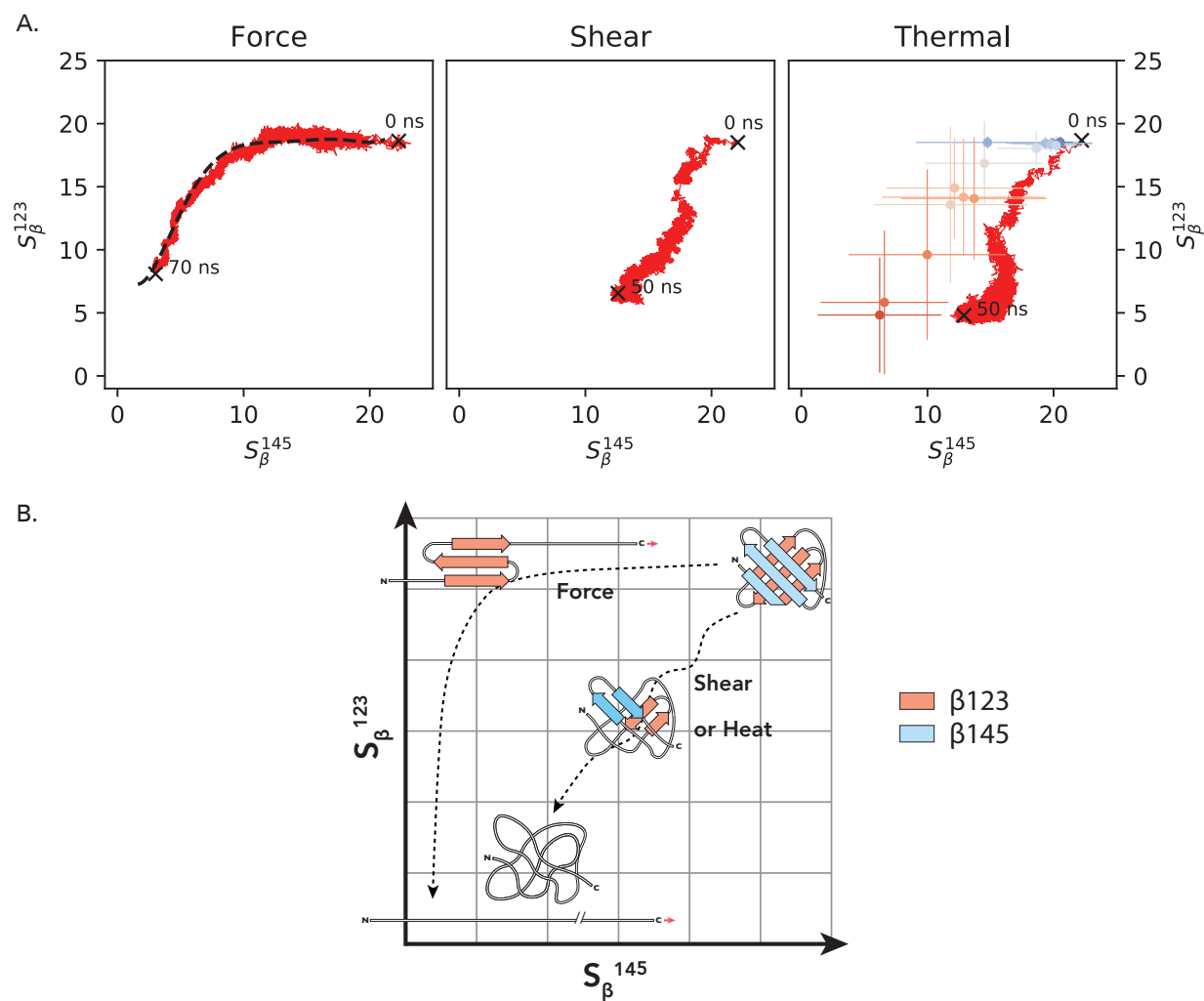
Figure 7: **Comparison of the unfolding paths for three different perturbations.** A. The mean unfolding path trajectories are projected on the two-dimensional space defined by the S_β^{123} and S_β^{145} collective variables. For the thermal unfolding the data extracted from the replica-exchange simulations represent the average value at each associated temperature. B. Schematic representation of the unfolding mechanisms.

years and provide a conceptual reference, lesser is known about the behavior of proteins under shear.[10–14] However this aspect is key in order to understand the functionality of special types of proteins that work under physiological fluid flow.[1,2]

In this study we have used the Lattice Boltzmann Molecular Dynamics technique, that allows to self-consistently generate the fluid flow perturbing the protein dynamics,[16,40] combined with the implicit-solvent CG OPEP model for the protein.[41] The validity of our results for the force and thermal unfolding was assessed by comparing the current approach with data obtained from more accurate, but computationally more expensive, all-atom simulations.[33]

As already observed for other systems, the unfolding of an isolated small globular protein can occur under Couette flow only at high shear rates, $\dot{\gamma} \sim 10^9\,\mathrm{s}^{-1}$, much higher that those in physiological conditions.[10,16] At these high shear rates, CspA unfolds on the nanosecond timescale, and the unfolding rate is proportional to the shear rate $\dot{\gamma}$. Moreover, we observed that unfolding under a directional mechanical force of $\sim 300\,\mathrm{pN}$ occurred on similar timescales. Extrapolation of the force-extension kinetics data at zero force allowed to extract the free-energy barrier for unfolding, which was found to be in very good agreement with that obtained from atomistic simulations.

It is worth noting, however, that for very long multimeric proteins such as vWf, the extension of the chain under shear occurs at physiological shear rates up to $\dot{\gamma} \sim 10^5\,\mathrm{s}^{-1}$. The resulting tensile force acts at the nanometer scale on the individual domains and is expected to cause partial unfolding of the weaker ones – in the case of vWf, the functional A2 domain. It was shown[19] that using a simple mechanical model and assuming an unfolding force of $\sim 10\,\mathrm{pN}$, the shear rate required to cause the chain extension and domain unfolding was estimated to be $10^5 - 10^6\,\mathrm{s}^{-1}$, close to the physiological conditions. Experimentally the force that would cause the partial unfolding of the A2 domain is estimated to be $20\,\mathrm{pN}$. Therefore the magnitude of the unfolding force for A2 compares to the low force regime in single molecule experiments where the characteristic unfolding time is generally within

seconds.

We have finally individuated two collective variables that allow to compare the unfolding mechanism under the three perturbations. We have clearly shown that the unfolding pathways under shear have strong similarities with that observed during thermal unfolding. This is to contrast with the mechanical pulling perturbation where the external force acts anisotropically and oriented on the end-to-end distance of the protein. *Vice versa*, the tensile force generated by the shear acts isotropically on average because of the protein tumbling. Therefore, force and shear perturbations probe very different weaknesses of the protein fold, and lead to distinct unfolding paths. In the future, we will assess how, when moving to a longer multimeric protein like vWf, the shearing acts down at the level of the individual domains, and how the elongation of the protein chain affects the isotropic action of the tensile force.

In conclusion we have demonstrated that the combination of a very effective CG model with a realistic treatment of hydrodynamic interactions via Lattice Boltzmann allows to successfully explore the response of a protein under perturbing fluid conditions, and to compare its behaviour under thermal and mechanical stresses. This methodology is ready to be deployed to extensively study at the molecular level some peculiar proteins such as catch-bonds or vWf, to characterize how the shear flow imposes local reorganization of the functional domains, and try to extrapolate this behavior at low-shear physiological conditions.

Supporting Information Available

The Supporting Information is available free of charge on the ACS Publications website. An additional plot of the time evolution of the RMSD under shear; snapshots of force unfolding trajectories illustrating the unfolding mechanism and the key role of the L_3 loop; a detailed description of the method employed to estimate the free-energy barriers along with related graphs and data.

Acknowledgement

References

(1) Thomas, W. E.; Vogel, V.; Sokurenko, E. Biophysics of Catch Bonds. *Ann. Rev. Biophys.* **2008**, *37*, 399–416.

(2) Springer, T. A. von Willebrand Factor, Jedi Knight of the Bloodstream. *Blood* **2014**, *124*, 1412–1425.

(3) Schneider, S. W.; Nuschele, S.; Wixforth, A.; Gorzelanny, C.; Alexander-Katz, A.; Netz, R. R.; Schneider, M. F. Shear-Induced Unfolding Triggers Adhesion of von Willebrand Factor Fibers. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7899–7903.

(4) Zhang, Q.; Zhou, Y.-F.; Zhang, C.-Z.; Zhang, X.; Lu, C.; Springer, T. A. Structural Specializations of A2, a Force-Sensing Domain in the Ultralarge Vascular Protein von Willebrand Factor. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 9226–9231.

(5) Fu, H.; Jiang, Y.; Yang, D.; Scheiflinger, F.; Wong, W. P.; Springer, T. A. Flow-Induced Elongation of von Willebrand Factor Precedes Tension-Dependent Activation. *Nat. Commun.* **2017**, *8*, 324.

(6) Cusumano, C. K.; Pinkner, J. S.; Han, Z.; Greene, S. E.; Ford, B. A.; Crowley, J. R.; Henderson, J. P.; Janetka, J. W.; Hultgren, S. J. Treatment and Prevention of Uri-

nary Tract Infection with Orally Active FimH Inhibitors. *Sci. Transl. Med.* **2011**, *3*, 109ra115.

(7) Hartmann, M.; Lindhorst, T. K. The Bacterial Lectin FimH, a Target for Drug Discovery Carbohydrate Inhibitors of Type 1 Fimbriae-Mediated Bacterial Adhesion. *Eur. J. Org. Chem.* **2011**, *2011*, 3583–3609.

(8) Sauer, M. M.; Jakob, R. P.; Eras, J.; Baday, S.; Eris, D.; Navarra, G.; Berneche, S.; Ernst, B.; Maier, T.; Glockshuber, R. Catch-bond Mechanism of the Bacterial Adhesin FimH. *Nat. Commun.* **2015**, *7*, 10738.

(9) Kalas, V.; Pinkner, J. S.; Hannan, T. J.; Hibbing, M. E.; Dodson, K. W.; Holehouse, A. S.; Zhang, H.; Tolia, N. H.; Gross, M. L.; Pappu, R. V. et al. Evolutionary Fine-Tuning of Conformational Ensembles in FimH During Host-Pathogen Interactions. *Sci. Adv.* **2017**, *3*, e1601944.

(10) Jaspe, J.; Hagen, S. J. Do Protein Molecules Unfold in a Simple Shear Flow? *Biophys. J.* **2006**, *91*, 3415–3424.

(11) Bekard, I. B.; Asimakis, P.; Teoh, C. L.; Ryan, T.; Howlett, G. J.; Bertolini, J.; Dunstan, D. E. Bovine Serum Albumin Unfolds in Couette Flow. *Soft Matter* **2012**, *8*, 385–389.

(12) Bekard, I. B.; Dunstan, D. E. Shear-Induced Deformation of Bovine Insulin in Couette Flow. *J. Phys. Chem. B* **2009**, *113*, 8453–8457.

(13) Bekard, I. B.; Barnham, K. J.; White, L. R.; Dunstan, D. E. a-Helix Unfolding in Simple Shear Flow. *Soft Matter* **2011**, *7*, 203–210.

(14) Ashton, L.; Dusting, J.; Imomoh, E.; Balabani, S.; Blanch, E. W. Shear-Induced Unfolding of Lysozyme Monitored In Situ. *Biophys. J.* **2009**, *96*, 4231–4236.

(15) Brückl, L.; Schröder, T.; Scheler, S.; Hahn, R.; Sonderegger, C. The Effect of Shear on the Structural Conformation of rhGH and IgG1 in Free Solution. *J. Pharm. Sci.* **2016**, *105*, 1810–1818.

(16) Sterpone, F.; Derreumaux, P.; Melchionna, S. Molecular Mechanism of Protein Unfolding under Shear: A Lattice Boltzmann Molecular Dynamics Study. *J. Phys. Chem. B* **2017**, *122*, 1573–1579.

(17) Alexander-Katz, A.; Schneider, M. F.; Schneider, S. W.; Wixforth, A.; Netz, Shear-Flow-Induced Unfolding of Polymeric Globules. *Phys. Rev. Lett.* **2006**, *97*, 138101.

(18) Alexander-Katz, A.; Netz, R. R. Dynamics and Instabilities of Collapsed Polymers in Shear Flow. *Macromolecules* **2008**, *41*, 3363–3374.

(19) Morabito, M.; Dong, C.; Wei, W.; Cheng, X.; Zhang, X.; Oztekin, A.; 3rd, E. W. Internal Tensile Force and A2 Domain Unfolding of von Willebrand Factor Multimers in Shear Flow. *Biophys. J.* **2018**, *https://doi.org/10.1016/j.bpj.2018.09.001* .

(20) Hamilton-Brown, P.; Bekard, I.; Ducker, W. A.; Dunstan, D. E. How Does Shear Affect Ab Fibrillogenesis? *J. Phys. Chem. B* **2008**, *112*, 16249–16252.

(21) Trumbore, C. N. Shear-Induced Amyloid Formation in the Brain: I. Potential Vascular and Parenchymal Processes Shear-Induced Amyloid Formation in the Brain: I. Potential Vascular and Parenchymal Processes Shear-Induced Amyloid Formation in the Brain: I. Potential Vascular and Parenchymal Processes. *J. Alzheimers Dis.* **2016**, *54*, 457–470.

(22) Byington, M. C.; Safari, M. S.; Conrad, J. C.; Vekilov, P. G. Protein Conformational Flexibility Enables the Formation of Dense Liquid Clusters: Tests Using Solution Shear. *J. Phys. Chem. Lett.* **2016**, *7*, 2339–2345.

(23) Kinoshita, M.; Lin, Y.; Dai, I.; Okumura, M.; Markova, N.; Ladbury, J. E.; Sterpone, F.; Lee, Y.-H. Energy landscape of polymorphic amyloid generation of b2-microglobulin revealed by calorimetry. *Chem. Commun.* **2018**, *54*, 7995–7998.

(24) Neuman, K. C.; Nagy, A. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Meth.* **2008**, *5*, 491–505.

(25) Jagannathan, B.; Marqusee, S. Protein folding and unfolding under force. *Biopolymers* **2013**, *99*, 860–869.

(26) Javadi, Y.; Fernandez, J. M.; Perez-Jimenez, R. Protein Folding Under Mechanical Forces: A Physiological View. *Physiology* **2013**, *28*, 9–17.

(27) Szymczak, P.; Cieplak, M. Stretching of Proteins in a Force-Clamp. *J. Phys. Condens. Matter* **2006**, *18*, L21.

(28) Szymczak, P.; Cieplak, M. Stretching of Proteins in a Uniform Flow. *J. Chem. Phys.* **2006**, *125*, 164903.

(29) Szymczak, P.; Cieplak, M. Protein in Shear Flow. *J. Chem. Phys.* **2007**, *127*, 155106.

(30) Jagannathan, B.; Elms, P.; Bustamante, C.; Marqusee, S. Direct Observation of a Force-Induced Switch in the Anisotropic Mechanical Unfolding Pathway of a Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17820–17825.

(31) Sun, L.; Noel, J. K.; Sulkowska, J. I.; Levine, H.; Onuchic, J. N. Connecting Thermal and Mechanical Protein (Un)folding Landscapes. *Biophys. J.* **2014**, *107*, 2950–2961.

(32) 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*, 3413–3418.

(33) Stirnemann, G.; Sterpone, F. Mechanics of Protein Adaptation to High Temperatures. *J. Phys. Chem. Lett.* **2017**, *8*, 5884–5890.

(34) Chen, Z.; Lou, J.; Zhu, C.; Schulten, K. Flow-Induced Structural Transition in the b-Switch Region of Glycoprotein Ib. *Biophys. J.* **2008**, *95*, 1303–1313.

(35) Lou, J.; Zhu, C. Flow Induces Loop-to-b-Hairpin Transition on the b-Switch of Platelet Glycoprotein Iba. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 13847–13852.

(36) Sing, C. E.; Alexander-Katz, A. Elongational Flow Induces the Unfolding of von Willebrand Factor at Physiological Flow Rates. *Biophys. J.* **2010**, *98*, 35–37.

(37) Alexander S. Lemak, J. R. L.; Chen, J. Z. Y. Unfolding Proteins in an External Field: Can We Always Observe the Intermediate States? *Phys. Rev. E* **2003**, *67*, 031910.

(38) Chiricotto, M.; Sterpone, F.; Derreumaux, P.; Melchionna, S. Multiscale Simulation of Molecular Processes in Cellular Environments. *Philos. Trans. Royal Soc. A* **2016**, *374*, 20160225.

(39) Sterpone, F.; Doutreligne, S.; Tran, T. T.; Melchionna, S.; Baaden, M.; Nguyen, P.; Derreumaux, P. Multi-Scale Simulations of Biological Systems Using the OPEP Coarse-Grained Model. *Biochem. Biophys. Res. Commun.* **2018**, *498*, 296–304.

(40) Sterpone, F.; Derreumaux, P.; Melchionna, S. Protein Simulations in Fluids: Coupling the OPEP Coarse-Grained Force Field with Hydrodynamics. *J. Chem. Theory Comput.* **2015**, *11*, 1843–1853.

(41) Sterpone, F.; Melchionna, S.; Tuffery, P.; Pasquali, S.; Mousseau, N.; Cragnolini, T.; Chebaro, Y.; St-Pierre, J.-F.; Kalimeri, M.; Barducci, A. et al. The OPEP Protein Model: from Single Molecules, Amyloid Formation, Crowding and Hydrodynamics to DNA/RNA Systems. *Chem. Soc. Rev.* **2014**, *43*, 4871–4893.

(42) Succi, S. *The Lattice Boltmzann Equation for Fluid Dynamics and Beyond*; Clarendon Press, 2001.

(43) Benzi, R.; Succi, S.; Vergassola, M. The Lattice Boltzmann Equation: Theory and Applications. *Phys. Rep.* **1992**, *222*, 145–197.

(44) Bhatnagar, P.; Gross, E.; Krook, M. A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems. *Phys. Rev.* **1954**, *94*, 511–525.

(45) Ahlrichs, P.; Dünweg, B. Simulation of a Single Polymer Chain in Solution by Combining Lattice Boltzmann and Molecular Dynamics. *J. Chem. Phys.* **1999**, *111*, 8225–8239.

(46) Ahlrichs, P.; Dünweg, B. Lattice-Boltzmann Simulation of Polymer-Solvent Systems. *Int. J. Mod. Phys. C* **1998**, *9*, 1429–1438.

(47) Chiricotto, M.; Melchionna, S.; Derreumaux, P.; Sterpone, F. Hydrodynamic Effects on b-Amyloid (16-22) Peptide Aggregation. *J. Chem. Phys.* **2016**, *145*, 035102.

(48) Derreumaux, P.; Mousseau, N. Coarse-grained Protein Molecular Dynamics Simulations. *J. Chem. Phys.* **2007**, *126*, 025101.

(49) Chebaro, Y.; Pasquali, S.; Derreumaux, P. The Coarse-Grained OPEP Force Field for Non-Amyloid and Amyloid Proteins. *J. Phys. Chem. B* **2012**, *116*, 8741–8752.

(50) Sterpone, F.; Nguyen, P.; Kalimeri, M.; Derreumaux, P. Importance of the Ion-Pair Interactions in the OPEP Coarse-Grained Force Field: Parametrization and Validation. *J. Chem. Theory Comput.* **2013**, *9*, 4574–4584.

(51) Schindelin, H.; Jiang, W.; Inouye, M.; Heinemann, U. Crystal Structure of CspA, the Major Cold Shock Protein of Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 5119–5123.

(52) Tych, K. M.; Hoffmann, T.; Brockwell, D. J.; Dougan, L. Single Molecule Force Spectroscopy Reveals the Temperature-Dependent Robustness and Malleability of a Hyperthermophilic Protein. *Soft Matter* **2013**, *9*, 9016.

(53) Tych, K. M.; Batchelor, M.; Hoffmann, T.; Wilson, M. C.; Paci, E.; Brockwell, D. J.; Dougan, L. Tuning Protein Mechanics through an Ionic Cluster Graft from an Extremophilic Protein. *Soft Matter* **2016**, *12*, 2688–2699.

(54) Tych, K. M.; Batchelor, M.; Hoffmann, T.; Wilson, M. C.; Hughes, M. L.; Paci, E.; Brockwell, D. J.; Dougan, L. Differential Effects of Hydrophobic Core Packing Residues for Thermodynamic and Mechanical Stability of a Hyperthermophilic Protein. *Langmuir* **2016**, *32*, 7392–7402.

(55) Schönfelder, J.; Perez-Jimenez, R.; Muñoz, V. A Simple Two-State Protein Unfolds Mechanically via Multiple Heterogeneous Pathways at Single-Molecule Resolution. *Nat. Commun.* **2016**, *7*, 11777.

(56) Guyon, E.; Hulin, J.-P.; Petit, L.; Mitescu, C. D. *Physical Hydrodynamics*; Oxford University Press, 2001.

(57) Best, R. B.; Hummer, G.; Eaton, W. A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 17874–17879.

(58) West, D. K.; Brockwell, D. J.; Olmsted, P. D.; Radford, S. E.; Paci, E. Mechanical Resistance of Proteins Explained Using Simple Molecular Models. *Biophys. J.* **2006**, *90*, 287–297.

(59) Beedle, A. E. M.; Lezamiz, A.; Stirnemann, G.; Garcia-Manyes, S. The Mechanochemistry of Copper Reports on the Directionality of Unfolding in Model Cupredoxin Proteins. *Nat. Commun.* **2015**, *6*, 7894.

(60) Garcia-Manyes, S.; Brujić, J.; Badilla, C. L.; Fernández, J. M. Force-Clamp Spectroscopy of Single-Protein Monomers Reveals the Individual Unfolding and Folding Pathways of I27 and Ubiquitin. *Biophys. J.* **2007**, *93*, 2436–46.

(61) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone phi, psi and Sidechain chi1 and chi2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(62) Evans, E.; Ritchie, K. Dynamic Strength of Molecular Adhesion Bonds. *Biophys. J.* **1997**, *72*, 1541–1555.

(63) Stirnemann, G.; Giganti, D.; Fernandez, J. M.; Berne, B. J. Elasticity, structure, and relaxation of extended proteins under force. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 3847–3852.

(64) 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. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14416–14421.

(65) Stirnemann, G.; Sterpone, F. Recovering Protein Thermal Stability Using All-Atom Hamiltonian Replica-Exchange Simulations in Explicit Solvent. *J. Chem. Theory Comput.* **2015**, *11*, 5573–5577.
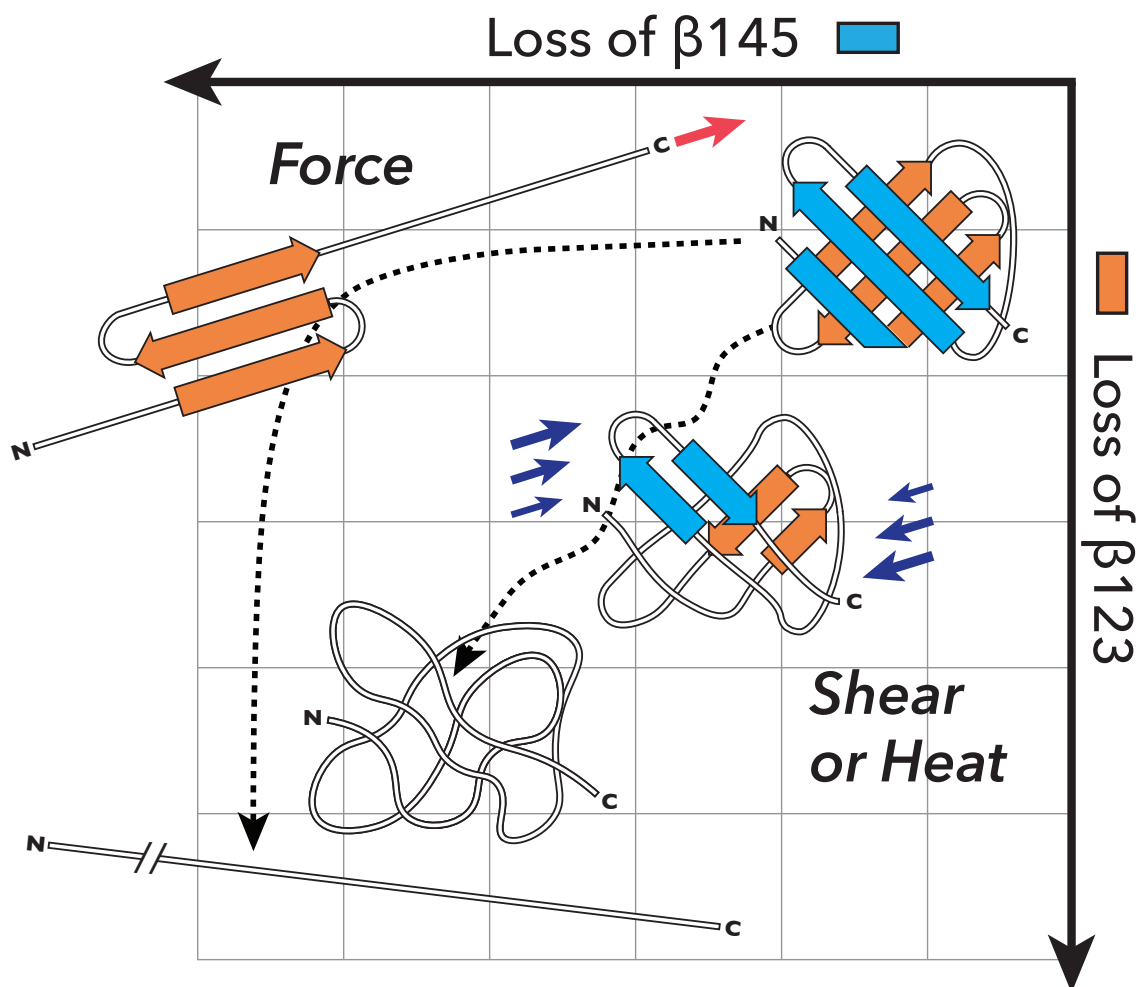
Graphical TOC Entry

Some journals require a graphical entry for the Table of Contents. This should be laid out "print ready" so that the sizing of the text is correct. Inside the `tocentry` environment, the font used is Helvetica 8 pt, as required by *Journal of the American Chemical Society*.

The surrounding frame is 9 cm by 3.5 cm, which is the maximum permitted for *Journal of the American Chemical Society* graphical table of content entries. The box will not resize if the content is too big: instead it will overflow the edge of the box.

This box and the associated title will always be printed on a separate page at the end of the document.