Role of Denaturation in Maltose Binding Protein Translocation Dynamics

Marco Bacci,[†] Mauro Chinappi,[‡] Carlo Massimo Casciola,[§] and Fabio Cecconi*,[∥]

[†]Dipartimento di Sistemi e Informatica, Engineering Division, Università degli Studi di Firenze Via di Santa Marta 3, 50139 Firenze, Italy
[‡]Dipartimento di Fisica, Sapienza Università di Roma P.le Aldo Moro 5, 00185 Roma, Italy
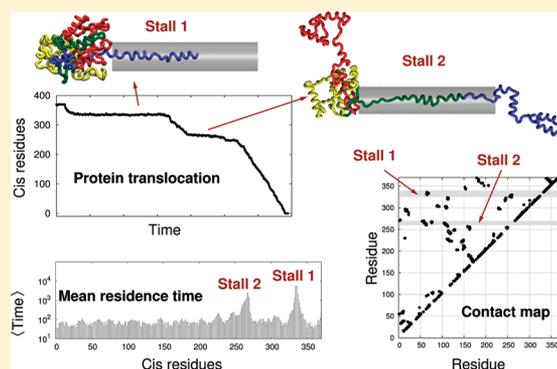[§]Dipartimento di Ingegneria Meccanica e Aerospaziale Sapienza, Università di Roma, Via Eudossiana 18, 00184 Roma, Italy
[∥]Istituto dei Sistemi Complessi (CNR), Via dei Taurini 19, 00185 Roma, Italy

S *Supporting Information*

ABSTRACT: We present a computational study on the driven transport of the Maltose Binding Protein (MBP) across nanochannels in the framework of coarse-grained modeling. The work is motivated by recent experiments on voltage-driven transport of MBP across nanopores exploring the influence of denaturation on translocation pathways. Our simplified approach allows a statistical mechanical interpretation of the process which may be convenient also to the experiments. Specifically, we identify and characterize short and long channel blockades, associated to the translocation of denatured and folded MBP conformations, respectively. We show that long blockades are related to long stall events where MBP undergoes specific and reproducible structural rearrangements. To clarify the origin of the stalls, the *stick-and-slip* translocation is compared to mechanical unfolding pathways obtained via steered molecular dynamics. This comparison clearly shows the translocation pathway to significantly differ from free-space unfolding dynamics and strongly suggests that stalling events are preferentially determined by the MBP regions with higher density of long-range native interactions. This result might constitute a possible criterion to predict a priori some statistical features of protein translocation from the structural analysis.

■ INTRODUCTION

Since the development of the first DNA voltage-driven translocation experiments,[1] the study of driven dynamics of biopolymers across biological and solid-state nanopores has received great interest in view of the possible biomedical applications,[2] thus becoming the subject of an extensive experimental,[3−9] theoretical,[10−15] and computational[16−20] investigation.

At present, nanopore-based technology is believed to be a promising resource offering revolutionary and powerful tools for detection,[21] characterizazion, manipulation, and sequencing of macromolecules.[22] Moreover, a single nanopore provides the possibility to study the folding/unfolding dynamics under confining and controlled conditions, partly simulating a crowding environment. In a classical voltage-driven translocation, a small applied voltage ($V \simeq 100$ mV) induces an ionic current through a membrane-integrated nanopore that can be measured by standard electrophysiological techniques. The mixing of biopolymers to the solution produces a current variation which strongly depends on the chemical and physical properties of the passing biomolecule which temporarily occupies the channel. For this reason, single nanopore systems can work as efficient devices to characterize biological macromolecules.[23]

Recent works[24,25] explored the channel blockade events in voltage-driven translocation of Maltose Binding Protein (MBP)

into α-hemolysin (αHL) and Aerolysin nanopores, respectively, as a function of both denaturing agent concentration and applied voltage. Dynamically, translocation of proteins is interesting as it strongly depends on the denaturation degree and generally couples to an unfolding stage. A folded protein whose gyration radius is larger than the pore narrower section needs a complete or partial unfolding to start the translocation.[4,26] Oukhaled et al.[24] found short and long blockade times of the αHL channel in the presence of MBP. They suggested short blockades due to the passage of fully unfolded MBP; in fact, their frequency increased by adding a protein denaturing agent. Long blockades were associated to partially folded MBP conformations, and their duration grew as the proteins were more compact.

In this paper, we undertake a computational study of MBP translocation across a nanopore to determine which structural properties are mainly responsible for the bottlenecks of transport (long blockades). To this purpose, we carry out simulations at different denaturation states, from fully folded to unfolded conformations.

Structurally, MBP is a monomeric globular protein of 370 residues, resolved by X-ray spectroscopy by Spurlino and co-workers.[27] Mechanical AFM pulling experiments on unbinded MBP structure by Rief et al.[28] identified some structural regions, termed *unfoldons* by the authors, as resistant areas to mechanical denaturation. These domains M1, M2, M3, and M4, shown in Figure 1 in color code on the PDB structure (PDB-ID: 4MBP),
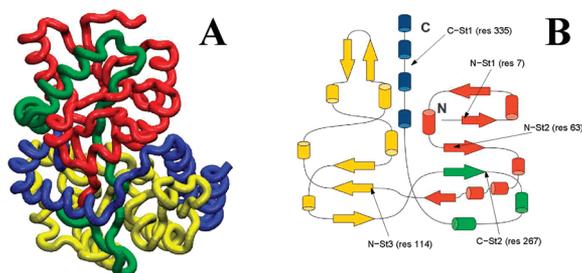

Figure 1. (A) Maltose Binding Protein structure (PDB-ID: 4MBP). The color code of the four unfoldons is chosen according to Bertz and Rief:[28] M1, blue (residues: 296–370); M2, green (residues: 244–295); M3, red (residues: 1–113); M4, yellow (residues: 114–243). Panel B reports the positions of the five stall points found in our translocation runs (residues 7, 63, 114, 267, 335) in a 2D-topological view. 3D structures are obtained using the VMD software.[31]

can be expected to be related to long blockade events in MBP translocation. Due to the known capabilities to correctly account for general protein structural properties, we consider a Gō-like computational model of the MBP (already considered in ref 29) as the most natural approach to assess the impact of MBP structural properties on translocation dynamics. Morever, the advantage of a coarse-grained description with respect to atomic scale models of the system (protein, nanopore, solvent) relies on the possibility to explore a large number of denaturation and pulling conditions, thus accumulating robust statistics of translocation events. For recent applications of coarse-grained models to various bio-chemical processes, see the review.[30] As a first step, we show that the model is able to reproduce the general features of MBP chemical denaturation. Then, we study the translocation of MBP in a pore, reproducing the average sizes of the αHL channel. We find that translocation dynamics is strongly affected by the protein denaturation state. In particular, translocation of chemically unfolded MBP conformations requires relatively low forces, and once one protein terminus enters the pore, the transport proceeds uniformly. In contrast, native-like MBP conformations exhibit a much richer translocation phenomenology; larger forces are required to trigger the translocation that, once started, proceeds through bottlenecks and jerky movements due to the rearrangements of the folded part of the protein that has not yet engaged the pore. In this case, the issue is to identify MBP structural motives responsible for the slowing down of the translocation. First we exclude these stalling stages to be related to the unfoldons. Then, by an analysis of static and dynamic native-contact maps, we elucidate that the stall points of the translocation pathway are mainly due to the denser protein regions of long-range native contacts.

METHODS

Coarse-Grained Protein Model. The MBP protein is described by the C_α backbone coarse-grained Gō-like model[32] already applied to protein translocation problems.[29,33] The set of parameter values entering the protein potentials is reported

in those papers, and here we briefly outline the key points of the approach. Each amino acid is conventionally assimilated to a bead whose position $\mathbf{r}_i$ is centered on the corresponding C_α atom. The residue–residue interactions are assigned by the potential

$$
\begin{aligned}
V_{\text{tot}} = & \sum_{i=1}^{N-1} \frac{k_{\text{h}}}{2}(r_{i,i+1} - R_{i,i+1})^2 + \sum_{i=2}^{N-1} \frac{k_\theta}{2}(\theta_i - \Theta_i)^2 \\
& + \sum_{i=3}^{N-2} k_\phi^{(1)}[1 - \cos(\phi_i - \Phi_i)] \\
& + k_\phi^{(3)}[1 - \cos 3(\phi_i - \Phi_i)] + \sum_{i,j>i+3} V_{\text{nb}}(r_{ij})
\end{aligned}
$$

(1)

Nonbonded terms V_{nb} describing the interaction between amino acids i and j, $|j - i| > 3$, in the sequence are defined as

$$
V_{\text{nb}}(r_{ij}) = \varepsilon \begin{cases} \left[5\left(\dfrac{R_{ij}}{r_{ij}}\right)^{12} - 6\left(\dfrac{R_{ij}}{r_{ij}}\right)^{10}\right] & R_{ij} < R_{\text{c}} \\[4mm] \dfrac{10}{3}\left(\dfrac{\sigma}{r_{ij}}\right)^{12} & R_{ij} > R_{\text{c}} \end{cases}
$$

(2)

Residues i and j attract each other when they form a *native contact* in the PDB structure (i.e., their distance is within a cutoff radius R_{c}) and otherwise interact via a soft excluded volume force with core $\sigma = 4.5$ Å. In the above expressions, r_{ij} denotes the distance between residue i and j; θ_i is the bending angle identified by the three consecutive C_α's, $i - 1$, i, $i + 1$; and ϕ_i is the dihedral angle defined by the two planes formed by four consecutive C_α's, $i - 2$, $i - 1$, i, $i + 1$. The capital symbols R_{ij}, Θ_i, and Φ_i are the corresponding native values extracted from the crystallographic structure. In such a Gō-like force field, parameter ε sets the energy scale, and R_{c} determines the number of native contacts (Table S1 in Supporting Information), thus controlling the stability of the native structure (the larger the R_{c} the more stable the structure). Both R_{c} and ε establish the denaturation temperature T_f;[32] in the following, $\varepsilon = 1$, and R_{c} varies in the range (3.0–7.5) Å. A standard leapfrog algorithm is used to integrate the protein dynamics with a Langevin thermostat operating at friction coefficient $\gamma = 0.25$ to control the temperature.

Pore Model. The pore action is reduced to a geometrical confinement enforced by a steplike cylindric potential along the x-axis defined for $x \in [0,L]$ with L being the nominal pore length

$$
V_{\text{pore}}(x, y, z) = V_0 \left(\frac{y^2 + z^2}{R_{\text{P}}^2}\right)^q \Theta[x(L - x)]
$$

(3)

where $\Theta(s) = [1 + \tanh(\alpha s)]/2$ is a smooth steplike function limiting the action of the pore potential in the effective region $[0,L]$. Pore length $L = 100$ Å and radius $R_{\text{p}} = 10$ Å are taken from αHL structural data. A convenient choice of the other parameters is $q = 1$, $\alpha = 3$ Å^{-2}, and $V_0 = 2\varepsilon$ (see ref 29). A homogeneous force F, collinear to the cylinder, acts only in the capture region $[-2,0]$ and inside the pore $[0,L]$ by dragging the foremost protein residue.

4256

dx.doi.org/10.1021/jp300143x | *J. Phys. Chem. B* 2012, 116, 4255–4262

Translocation Simulations. For each cutoff radius R_c, thermalization simulations are performed at $T = 0.75$ to generate initial MBP conformations with one of the two termini constrained near the pore entrance. Protein configurations are sampled at time intervals equal to 10% of the simulation time window, $T_w = 10^5$, to ensure statistical independence of initial states. After thermalization, the protein is driven by applying F into the pore.

Protein translocation is considered accomplished when the last residue leaves the trans-side of the channel. As we will see in the Results section, there are conditions where the proteins escape the capture by diffusion despite the action of the importing force. To save computational time, we stop and discard the simulations when the protein reaches a distance from the pore entrance comparable to the native-state linear size. This criterion is based on a preliminary estimate of a negligible probability for the molecule to reapproach the pore by diffusion and restart the translocation.

To set the nomenclature up, such cases of protein escape will be labeled as *loss* events. Cases where the protein is neither translocated nor *lost* in the time window T_w are indicated as *unsuccess*.

Stretching Simulations. Before each stretching simulation (AFM-like), the protein is equilibrated with no constraints nor additional forces. Stretching is simulated using a constant velocity Steered Molecular Dynamics (SMD) strategy[34] where protein elongation is induced by a spring of elastic constant $k = 0.1\varepsilon$. The protein N-terminus is held fixed, and the C-terminus is attached to the first end of the spring whose second end is dragged at constant velocity V in the direction of the initial end-to-end vector. To test the robustness of the results to simulation protocols, we used different steering velocities and performed AFM-like stretching also from the N-terminus (with the C-terminus blocked). As mentioned in the Introduction, MBP contains four unfoldons (M1, M2, M3, and M4, Figure 1). To monitor the denaturation degree of the k-th unfoldon domain, we consider the number of its internal active native contacts, normalized to the corresponding value in the PDB structure. Two residues originally in contact are considered detached when their distance in the actual molecule conformation exceeds $1.22R_{ij}$.

■ RESULTS

Denaturation Characterization. In the Gō-model force field (eq 1), a decrease of the cutoff radius R_c virtually plays the action of a chemical denaturating agent as reducing the number of attractive long-range interactions (eq 2) destabilizes the native-state to thermal fluctuations. Denaturation can also be achieved by either increasing the temperature or equivalently decreasing the overall energy scale ε of the force field (thermal unfolding). Thermal-induced unfolding, however, softens both bonded and nonbonded interactions (eqs 1 and 2), introducing too *global* effects not fully compatible with chemical denaturation experiments.[35] Similarly to a chaotropic agent which destabilizes native states by competing with internal noncovalent protein interactions such as hydrogen bonding and hydrophobic effect, our computer implementation reduces only the number of attractive nonbonded interactions responsible for the collapse onto compact stable native structures.

We thus expect that laboratory denaturation conditions can be effectively taken into account by a suitable choice of the cutoff radius. Experimentally, Ganesh and co-workers[35] analyzed chemical MBP denaturation by guanidine hydrochloride

(GndHCl) at constant temperature, $T = 28\ °C$. They showed that the unfolded protein concentration ρ_u, at different GndHCl concentrations $[D]$, estimated via circular dichroism spectroscopy, could be fitted via the standard two-state model (dashed in Figure 2). The latter compares the denaturation of
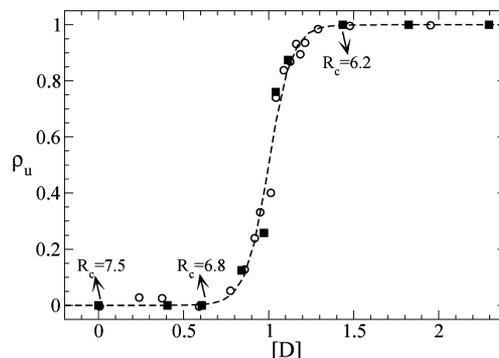

Figure 2. Denaturation plot of MBP obtained by thermal equilibrium simulations at $T = 0.75$. The proper rescaling (eq 4) of rmsd causes simulation data to collapse onto a Ganesh et al. denaturation plot[35] (fraction of unfolded structures by circular dichroism spectroscopy).

the Gō-model MBP as a function of cutoff radius R_c, with denaturation curves of ref 35. As a nativeness indicator, we employ the rmsd (Kabsch distance[36]) with respect to the PDB structure. The mapping between the simulated and experimental denaturation data is established after a baseline subtraction of rmsd values and a normalization to the $[0,1]$ interval. It reads

$$[D] = \frac{a}{R_c - R_c^0} - \frac{a}{R_c^* - R_c^0} \tag{4}$$

where our reference interaction cutoff $R_c^* = 7.5$ Å corresponds to the fully native state $[D] = 0$ in the experiment; $R_c^0 = 4.0$ Å is the cutoff radius for complete denaturation; and a is a tunable parameter adjusted to achieve data collapse.

Translocation Dynamics. We characterize translocation success (top panel of Figure 3) in terms of translocation
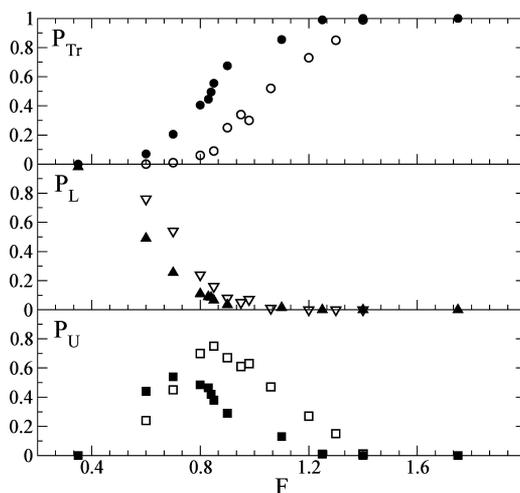

Figure 3. Statistics of translocation as a function of the importing force F at $R_c = 6.8$ Å and $T = 0.75$: translocation probability P_{Tr}, loss event probability P_L, and unsuccessful run probability P_U. Empty symbols, N-pulling; filled symbols, C-pulling.

probability P_{Tr}, the fraction of translocated runs with respect to the total number of attempts, as a function of the importing force F. The figure refers to $T = 0.75$ and $R_c = 6.8$ Å that, as shown in Figure 2, is the lowest R_c value corresponding to native-like conformations for both N-pulling (empty symbols) and C-pulling (filled symbols).

Besides P_{Tr}, Figure 3 provides also the fraction of *loss* events, P_L, and of *unsuccessful* events, $P_U = 1 - P_L - P_{Tr}$, a fraction of proteins that are neither translocated nor *lost* within the time window T_w. The sigmoid shape of P_{Tr} vs F allows a clear-cut definition of the critical force F_c as the value for which $P_{Tr}(F_c) = 1/2$. If $F \ll F_c$, most of the nontranslocated runs are *lost* since the importing force is so weak to be easily overwhelmed by thermal motion. Increasing F, P_L rapidly decays, whereas the number of translocated and unsuccessful runs increases. For $F \simeq F_c$, almost all the untranslocated proteins correspond to unsuccessful runs, and very few are *lost*. The results also indicate an asymmetric translocation process of folded MBP, which depends on the pulling terminus. The difference in F_c and in the number of proteins stuck in the pore between N- and C-pulling suggests that the transport pathways are different in the two cases.

To characterize the role of denaturation in the translocation mechanism, we estimate F_c at different cutoff R_c. The top panel of Figure 4 shows that translocation at low denaturation
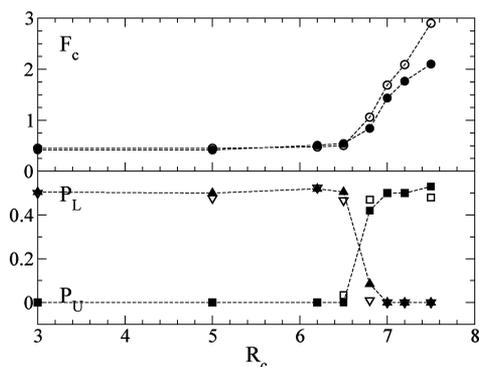

Figure 4. Top Panel: critical force F_c as a function of R_c for $T = 0.75$. Bottom Panel: *loss* and *unsuccess* event probability, P_L and P_U, respectively, at critical force. As in Figure 3, empty and filled symbols denote N-pulling and C-pulling data, respectively.

(large R_c) requires large forces. Decreasing R_c, the critical force is reduced, and below $R_c \simeq 6.8$ Å it reaches a plateau. Actually, once denatured ($R_c < 6.8$ Å, see Figure 2), the dynamics of random coil MBP conformations becomes unaffected by a further reduction of R_c. On the contrary, as for $R_c > 6.8$ Å, the MBP structures become more and more compact and stable, and the increase of the critical F_c is naturally expected due to a stronger resistance to unfolding. Therefore, F_c is a quantity able to discriminate folded and unfolded MBP structures, in analogy to the applied voltage in experiments.[4,24,25]

A further difference between denatured and native-like MBP translocation is revealed by the amount of *loss* and *unsuccessful* events at critical force. The bottom Panel of Figure 4 shows a clean transition at $R_c = 6.8$ Å in both P_L and P_U at critical force. At smaller R_c (denatured state), $P_L = 0.5$; hence, none of the unstructured chains gets stuck in the pore. That is compatible with the idea of an importing force competing with thermal fluctuations to insert a residue in the

pore: $F_c \simeq F_{th} = k_B T/d_0 \simeq 0.2$, with $d_0 = 3.8$ Å being the average distance between consecutive residues. Once the first core of residues is imported, denatured structures oppose weak resistance, and their translocation is easily accomplished. On the contrary, at larger R_c, the critical force is significantly greater than F_{th}. Hence, almost all MBP conformations start getting imported, and those unable to finalize the transport within the allotted time end up stuck into the channel. At low denaturation, long pore blockades are entirely due to the structural resistance of folded MBPs to mechanical unfolding.

Residence Time Statistics. Figure 5A reports the average translocation time τ_b as a function of the importing force for
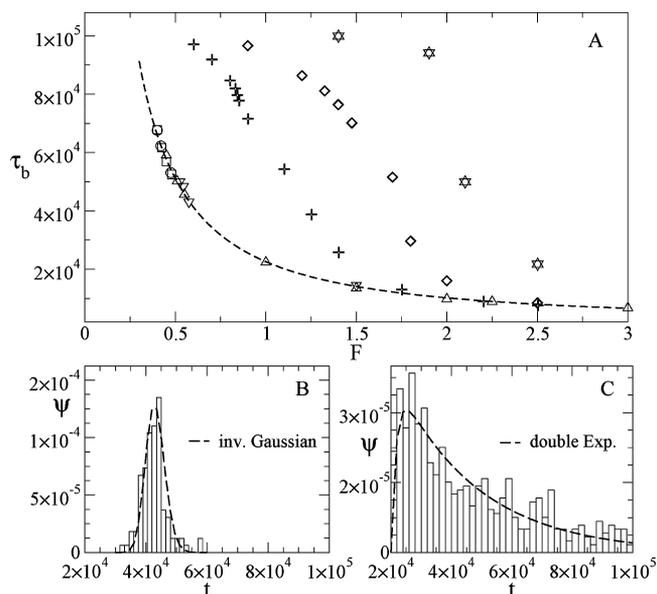

Figure 5. (A) Average translocation time on the subensemble of successful (translocated runs) simulations T_{Tr} vs importing force F. Different symbols represent different cutoff radii, namely, $R_c = \{3.0, 5.0, 6.2, 6.5, 6.8, 7.0\}$ Å and 7.5 Å, squares, circles, downward triangles, upward triangles, cross, diamonds, and stars, respectively. Data for C-pulling simulations. The dashed line represents the fit (eq 5). (B) and (C) Translocation time distribution for $R_c = 6.5$ Å and $F = 0.575$ (denatured) and $R_c = 6.8$ Å and $F = 1.10$ (native), respectively. The dashed lines in Panel B and C are an inverse Gaussian (eq 6) and a double exponential fit (eq 7).

different values of the interaction cutoff R_c. The average is taken over the subensemble of translocated runs. For the sake of clarity, only the values for C-pulling runs are reported, the N-pulling case is shown in the Supporting Information (Figure S1).

For a given cutoff radius, τ_b generally decreases with increasing F as expected. However, interestingly, for $R_c \leq 6.5$ Å, the τ_b-vs-F data collapse onto a single curve (dashed), confirming that MBP structures are completely denatured and react similarly to the importing force. The curve is well fitted by a three-parameter (τ_0, F_0, μ_0) relation

$$\tau_b(F) = \tau_0 e^{-F/F_0} + \frac{L}{\mu_0 F} \tag{5}$$

which is a linear combination of an activation term $\exp(-F/F_0)$ and a constant velocity drift, $1/F$. The origin of the activation term can be explained as follows: an irreversible translocation can occur only after a stable core of residues is established inside the channel. This sort of capture process requires the overcoming of
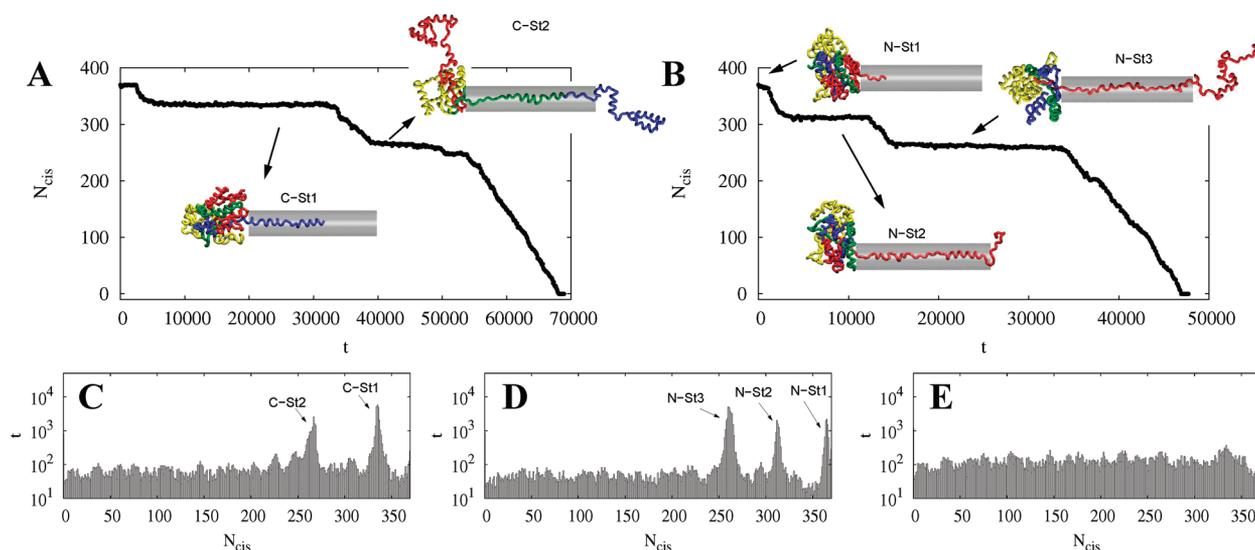
Figure 6. Time evolution of residue number on the cis side of the pore N_{cis} for a pulling run at R_c = 6.8 Å and critical force (Panel A: C-terminus pulling, Panel B: N-terminus pulling). Snapshots of representative configurations (VMD software[31]) are reported for the plateau corresponding to the protein blockades. The color code of the four unfoldons is chosen according to ref 28: M1, blue; M2, green; M3, red; M4, yellow. The average time spent by the protein in different N_{cis}, at critical force, is reported in Panels C, D, and E for R_c = 6.8 Å C-pulling, R_c = 6.8 Å N-pulling, and R_c = 6.5 Å C-pulling, respectively.

the initial entropic barrier, due to the strong confinement of the chain induced by channel, that even unstructured polypeptides experience.[29]

For $R_c \geq 6.8$ Å, average translocation time grows with decreasing F (crosses, diamonds, stars in Figure 5A), indicating that translocation is drastically slowed down. At large enough F, overwhelming the stability of the native-like MBP, τ_b tends to collapse on curve (eq 5), regardless of the value of the cutoff radius. This curve thus constitutes a sort of baseline for all translocation times. In low force regimes, the bending of data toward the horizontal line $\tau = T_w$ is an effect due to the time window finiteness. In Figure 5A, that effect has been partially corrected by using the formula

$$\tau_b = \frac{1}{P_{Tr} + P_U}(P_{Tr}\langle t \rangle + P_U T_w)$$

(whose derivation is sketched in Section S1 of Supporting Information), where $\langle t \rangle$ is the average time from the runs and P_{Tr} is the corresponding probability (Figure 3).

Panel B shows a typical translocation time distribution $\psi(t)$ of a denatured MBP, which looks well localized around its average τ_b, and it is properly fitted by an inverse Gaussian, with parameters D_0, μ_0

$$\psi(t) = \frac{L}{\sqrt{4\pi D_0 t^3}} \exp\left\{-\frac{(L - \mu_0 F t)^2}{4D_0 t}\right\} \qquad (6)$$

that is the first-passage time distribution of biased random walkers on the interval $[-\infty, L]$ emitted in 0 and absorbed at $-\infty$ and L.[37] An instance of the time distribution $\psi(t)$ of native-like MBP translocation is plotted in Panel C. In this case, $\psi(t)$ has a fatter large-time tail, and it is well fitted by a double exponential

$$\psi(t) = \frac{k_1 k_2}{k_2 - k_1}[e^{-k_1(t-t_0)} - e^{-k_2(t-t_0)}] \qquad (7)$$

with rates k_1, k_2, and t_0 being an offset time interpreted as the time taken by a denatured MBP to cross the pore. As we will see below, the double exponential is justified by the presence of two successive stall points, corresponding to two energy barriers whose overcoming can be seen as activated processes.

Translocation phenomenology is better characterized by addressing the time evolution of the number of residues N_{cis} on the cis-side of the pore. Panel A of Figure 6 shows $N_{cis}(t)$ for one successful run at critical force for R_c = 6.8 Å, pulled from the C-terminus. Clearly, most of the time is spent by the MBP in two particular stalling states, $N_{cis} \simeq 335$ (C-St1) and $N_{cis} \simeq 267$ (C-St2), corresponding to configurations where either residue 335 or residue 267 is, respectively, located at the pore entrance. Representative snapshots of the two states are included in the figure for illustration. To gain more quantitative information, we compute the time $T_r(N_{cis})$ spent by the r-th run in a given N_{cis} state during translocation. Time $T(N_{cis})$ defined as the average of $T_r(N_{cis})$ over the ensemble of translocated runs is plotted in Panel C of Figure 6, for C-pulling translocations at critical force and R_c = 6.8 Å. The two peaks in the histogram correspond to the two blockade events shown in Panel A. Clearly translocation is far from being uniform and looks more like a steplike (or stick and slip) process, where the protein has to overcome successive free energy barriers associated to specific structural rearrangements. The analysis of N-pulling (same cutoff radius and corresponding critical force), Panels B and D in Figure 6, confirms the overall picture: also in this case the protein spends most of the time in specific conformations. The stalling events are three and take place at different positions, namely, N_{cis} = 363 (N-St1), corresponding to residue 7 at the pore mouth, N_{cis} = 307 (N-St2, residue 63 at the mouth), and N_{cis} = 256 (residue 114, N-St3). The scenario both for N- and C-pulling is robust under changes of both R_c (provided $R_c \geq 6.8$ Å, i.e., native-like MBP) and pulling force range F, with only slight differences in peak intensity (Figure S2 in the Supporting Information). Concerning the *unsuccessful* translocations, we find that the already mentioned stall points still occur (Figure S2 and S3 in Supporting Information).

To summarize, Panel B of Figure 1 sketches in a 2D-topological view the positions of the five blockage points detected for structured protein translocation runs (residues 7, 63, 114, 267, 335).

To complete the phenomenology, we note that the transport of denatured proteins ($R_c < 6.8$ Å) looks much more uniform (peakless) (Panel E of Figure 6).

Stretching vs Translocation. The above picture of MBP translocation raises the natural question as to why the transport of initially compact structures becomes temporarily stalled at well-defined stages. In this section, we relate these "rate-limiting steps" of translocation to MBP structural properties. As mentioned in the Introduction, targeted AFM experiments[28] have shown that MBP mechanical stretching occurs via a sequence of events corresponding to the successive breakdown of specific domains, termed unfoldons. The analogy of pulled translocation with AFM mechanical stretching suggests that unfoldons might be involved also in MBP translocation bottlenecks. A similar comparison between translocation and stretching processes has been carried out by Huang et al.[38] for ubiquitin. Preliminarily, we checked, via standard Steered Molecular Dynamics stretching protocol (see Methods), that the Gō-model of MBP is able to reproduce the "unfoldon picture" observed in the experiments.[28] We plot in Figure 7a
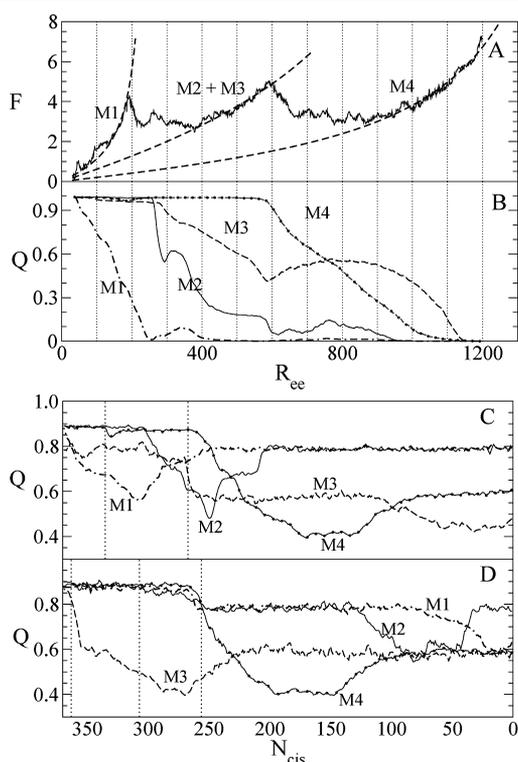

Figure 7. Upper: force (A) and average value of fraction of native contacts $Q(Mk)$, $k = 1,4$, (B) vs end-to-end distance R_{ee} in stretching simulation at pulling velocity $V = 0.05$, temperature $T = 0.5$, and cutoff radius $R_c = 7.5$ Å. Lower: average value of $Q(Mk)$ as a function of N_{cis} for $R_c = 6.8$ and $T = 0.75$ for N-pulling (C) and C-pulling (D) translocated runs at critical force.

the average over 50 runs of the force–extension (end-to-end distance R_{ee}) curve of the MBP in our numerical stretching experiment. The three branches, denoted by M1, M2 + M3, and M4 according to the nomenclature of ref 28, separated by the three worm-like-chain curves[39] identify the same unfoldon

opening sequence described in ref 28, including the coupled unfolding of M2 and M3. The picture becomes clear when plotting the average fraction of native contacts $Q(Mk)$, $k = 1, ..., 4$, in each unfoldon as a function of the end-to-end distance R_{ee} of the MBP. For instance, in the initial stage of the elongation, corresponding to the curve branch labeled as M1, the structural parameter $Q(M1)$ (dash-dotted line) decreases signaling the main involvement of M1 in this stage. The simultaneous opening of M2 and M3, solid and dashed line, respectively, is also evident in the second branch. Finally, M4 opens. Snapshots of the detachment sequence are shown in Figure S4 in the Supporting Information.

To understand the role of unfoldons in the MBP translocation process, first we locate the five blockage points with respect to the unfoldons in the protein structure (Figure 1). Then, again we plot unfoldon structural parameters $Q(Mk)$ as a function of N_{cis} for C-pulled and N-pulled translocations, respectively (Figure 7c and d). In the C-pulling case, the sequence of unfoldon opening is essentially the same as in the mechanical stretching. First M1 breaks down, followed by M2 and M3, whose dynamics is again correlated, and finally the M4 opening concludes the process. The vertical dotted lines in Figure 7c and d highlight the two stall points C-St1 and C-St2 in the C-pulling and the three other ones N-St1, N-St2, and N-St3 in the N-pulling (see discussion of Figure 6). A very weak correlation emerges between unfoldon dynamics and blockades. Specifically, the two configurations reported in Panel A of Figure 6 show that the first blockage point for C-pulling (residue 267) takes places amid unfoldon M1 (blue in the figure), while the second one (residue 335) lies in the initial part of M2 (green). In the pulling from the N-terminus, we observe that M3 is the first to be broken followed by M4, then M2, and finally M1. Again, the blockade points do not appear directly related to unfoldon boundaries (Panel B of Figure 6), with the only exception of the stall N-St3 located at the boundary between M3 and M4.

■ DISCUSSION

A crucial result of our analysis is a sharp change in the translocation dynamics upon varying the cutoff radius from $R_c = 6.5$ Å to $R_c = 6.8$ Å. We recall that according to eq 4 we can establish a mapping between R_c and MBP denaturation degree (Figure 2).

It is convenient to split the discussion of the results into two parts: translocation of denatured MBP ($R_c \leq 6.5$ Å) and translocation of native-like MBP ($R_c \geq 6.8$ Å).

Denatured-MBP Translocation ($R_c \leq 6.5$ Å). Below $R_c = 6.5$ Å, the MBP behaves as a random coil, and its translocation dynamics is basically independent of R_c as revealed by the critical force behavior (Figure 4) that remains almost constant for $R_c \leq 6.5$ Å and by the blockage time data (Figure 5) that collapse onto the same curve for every R_c, thus exhibiting a common behavior (eq 5). In this R_c range, not only the critical force F_c but also the probability of *lost* proteins at F_c exhibit a plateau (Figure 4). Moreover, the pulling terminus (C or N) does not influence significantly any of these quantities. The only barrier in the translocation dynamics is due to the pore entrance; once the protein is captured, the average time spent at different stages of the process is similar (Panel E Figure 6). Thus, random-coil MBP translocation can be interpreted as a capture (activated) stage followed by an elementary first passage process (FPP), where random walkers under a constant bias are injected from the cis side of the channel and absorbed to the trans side.[29]

Native-Like MBP Translocation ($R_c \geq 6.8$ Å). Folded structures exhibit a much more complex dynamics characterized by a transport getting stalled in the pore due to specific MBP conformations. Translocation times are greater than the unfolded case. Their average increases with R_c (Figure 5a), and their distributions show fat tails (Figure 5c). These findings agree with experiments by Oukhaled et al.[24] indicating that, in the explored forcing regime, only denatured proteins can translocate, whereas very long blockades pertain to partially folded ones. The capability of numerical techniques to explore the dynamical details evidences the presence of stall points, C-St1, C-St2, N-St1, N-St2, and N-St3, in native-like MBP transport (Figures 1 and 6). Natural candidates for such stall points are the boundaries between MBP unfoldons; however, our results clearly indicate that the AFM-unfolding dynamics is not correlated to the stalls. Rather, the analysis of MBP native-contact maps allows us to shed light on the actual mechanism underlying the stalls. Indeed, the blocking C-St1 and C-St2 can be interpreted in terms of the specific sequence of native contact breaking that a MBP subdomain undertakes to engage the pore in an almost linear conformation. In fact, the *stalling domains* of the protein are those with the greater amount of *external native contacts*, where external means "excluding all the inner contacts of the domain". In formulas, if B is the domain, then its contact density reads

$$\rho(B) = \frac{1}{N(B)} \sum_{i \in B} \sum_{j \notin B} \Delta_{ij} \tag{8}$$

where $N(B)$ denotes the number of residues in B and $\Delta_{ij} = 1$ if i and j form a native contact and $\Delta_{ij} = 0$ otherwise. In Figure 8,

we consider the contact map of native MBP obtained for $R_c = 6.8$ Å. The lower panel shows the average of external contacts formed by ten-residue-long consecutive regions of the MBP, and the peaks identify protein regions with the greater number of contacts that are thus the most probable candidates to cause stall points once engaging the pore entrance. However, considering dynamics data (average blocking times in Figure 6C), we can be more precise by selecting the two critical regions around the peaks C-St1 and C-St2, C−B1 = [328,338] and C−B2 = [260,270], respectively, and the complementary regions X1 = [339,370], X2 = [271,327], and X3 = [1,259], to patch the whole molecule structure. Table 1 summarizes the

Table 1. Contact Density of MBP Subdomains Described in the Text[a]

	C-pulling			N-pulling	
name	region	ρ	name	region	ρ
X1	[339,370]	0.75	N−B1	[1,11]	3.64
C−B1	[328,338]	2.00	Y1	[12,54]	0.47
X2	[271,327]	1.07	N−B2	[55,63]	2.33
C−B2	[260,270]	5.09	Y2	[64,104]	0.93
X3	[1,259]	0.00	N−B3	[105,115]	5.45
—	—	—	Y3	[116,370]	0.00

[a]Col.1: domain nomenclature, C−B1, C−B2 and N−B1, N−B2, N−B3 stand for critical blocks in C- and N-pulling, respectively. X1, X2, X3 and Y1, Y2, and Y3 are the complementary regions. Col.2: residues involved in each sub-domain. Col.3: density of external native contacts as defined in eq 8. Critical blocks have a contact density larger than complementary regions.

density of external contacts (eq 8), and these regions form with that part of the molecule that runs from the boundary of the issued region to the free terminus. For instance, external contacts pertaining to C−B1 bond this domain with residues from 327 to 1 (N-Terminus), as here C-terminus pulling blockades are considered. This approach allows us to track only effective external contacts that come into play when a region is facing the pore, namely, the contacts formed with that part of the molecule which has not yet engaged the pore. The higher values pertain to C−B1 and C−B2 subdomains, confirming the picture emerging from Figure 8. These areas, being dense of contacts, oppose the maximal resistance to unfolding once in front of the pore. In other words, translocation bottlenecks are determined by those subdomains that, still partially folded when approaching the pore cis side, carry with themselves other distal regions of the molecule tightly bonded to by native interactions. The contact analysis leads to the same conclusion also for N-terminus pulling, where the critical regions are N−B1 = [1,11], N−B2 = [55,63], and N−B3 = [105,115].

CONCLUSIONS

We have simulated Maltose Binding Protein translocation across the αHL nanopore via a coarse-grained computational model for both the MBP and the pore. As the channel is narrow, translocation properties strongly depend on the denaturation state of the MBP. In our Gō-model of the MBP, molecule denaturation is controlled by the parameter R_c determining the number of native attractive interactions. In the region $(6.5 < R_c < 6.8)$ Å, a transition is observed from random-coil MBP (denatured) to native-like structures. The transition emerges from both equilibrium (Figure 2) and transport simulations (Figure 3 and Figure 4). In particular, translocation
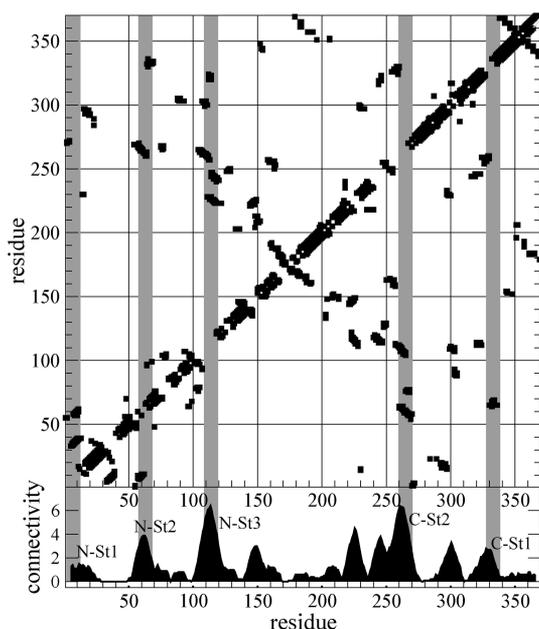

Figure 8. Contact map of native MBP obtained for $R_c = 6.8$ Å. Bottom panel depicts the average number of long-range contacts each residue establishes. The peaks identify the regions of the chain with the largest density of external contacts which are responsible for putative stall points. Vertical gray bands highlight the stalling regions as obtained by the analysis of Figure 6. These regions form a significant number of contacts with distal parts of the molecule, and only when one of these contact clusters is broken translocation can proceed until to the next bottleneck (next cluster).

of denatured MBP is almost uniform and consists of a capture stage followed by a simple driven diffusion process. The passage in the channel of folded MBP is more critical and interesting, and it looks like a "stick–slip" dynamics characterized by constant-velocity transport broken by stalling events in the channel (Figure 6). For instance, the C-terminus translocation occurs via two long stall events resulting in a double-exponential tail behavior of the translocation time distribution (Figure 5C). This is presumably associated to the presence of two successive free-energy barriers the MBP has to overcome to complete the passage. Our analysis moreover shows that stall events are related to those MBP regions with a larger density of native external contacts. Thus, long blockade events and stall points can be predicted by looking directly at MBP PDB-structure. On the contrary, a weak correlation is found between stall points and unfoldons, the structural blocks through which the MBP reacts to mechanical stretching.[28] The latter result is a strong indication that despite the analogy between pulled translocation and mechanical unfolding the pathways gathered from mechanical pulling are not sufficient to make inference on translocation mechanisms. The action of the pore, indeed, drastically modifies the unfolding pathway during translocation with respect to a free pulling process, in agreement with the result by Huang et al.[38]

ASSOCIATED CONTENT

Ⓢ Supporting Information

Supporting Table S1 includes the number of native contacts as a function of the cutoff radius R_c. Supporting Section S1 illustrates the derivation of the correction for the average translocation time used to set data in Figure 5. Figures S1–S4 include additional numerical results to confirm and enrich the main results of our work. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: fabio.cecconi@roma1.infn.it.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13770–13773.

(2) Desai, T. A.; Hansford, D. J.; Kulinsky, L.; Nashat, A. H.; Rasi, G.; Tu, J.; Wang, Y.; Zhang, M.; Ferrari, M. *Biomed. Microdevices* **1999**, *2*, 11–40.

(3) Meller, A.; Nivon, L.; Branton, D. *Phys. Rev. Lett.* **2001**, *86*, 3435–3438.

(4) Talaga, D.; Li, J. *J. Am. Chem. Soc.* **2009**, *131*, 9287–9297.

(5) Movileanu, L.; Schmittschmitt, J. P.; Scholtz, J. M.; Bayley, H. *Biophys. J.* **2005**, *89*, 1030–1045.

(6) Pastoriza-Gallego, M.; Gibrat, G.; Thiebot, B.; Betton, J. M.; Pelta, J. *Biochim. Biophys. Acta, Biomembr.* **2009**, *1788*, 1377–1386.

(7) Li, J.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. A. *Nat. Mater.* **2003**, *2*, 611–615.

(8) Storm, A. J.; Chen, J. H.; Ling, X. S.; H. W. Zandbergen, C. D. *Nat. Mater.* **2003**, *2*, 537–540.

(9) Bates, M.; Burns, M.; Meller, A. *Biophys. J.* **2003**, *84*, 2366–2372.

(10) Sung, W.; Park, P. J. *Phys. Rev. Lett.* **1996**, *77*, 783–786.

(11) Lubensky, D. K.; Nelson, D. R. *Biophys. J.* **1999**, *77*, 1824–1838.

(12) Berezhkovskii, A. M.; Pustovoit, M. A.; Bezrukov, S. M. *J. Chem. Phys.* **2002**, *116*, 9952–9956.

(13) Berezhkovskii, A. M.; Gopich, I. V. *Biophys. J.* **2003**, *84*, 787–793.

(14) Metzler, R.; Klafter, J. *Biophys. J.* **2003**, *85*, 27762779.

(15) Slonkina, E.; Kolomeisky, A. B. *J. Chem. Phys.* **2003**, *118*, 7112–7118.

(16) Aksimentiev, A.; Heng, J. B.; Timp, G.; Schulten, K. *Biophys. J.* **2004**, *87*, 2086–2097.

(17) Makarov, D. E. *Acc. Chem. Res.* **2008**, *42*, 281–289.

(18) Muthukumar, M.; Kong, C. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 5273–5278.

(19) Mathé, J.; Aksimentiev, A.; Nelson, D. R.; Schulten, K.; Meller, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12377–12382.

(20) Matysiak, S.; Montesi, A.; Pasquali, M.; Kolomeisky, A. B.; Clementi, C. *Phys. Rev. Lett.* **2006**, *96*, 118103(1–4).

(21) Braha, O.; Walker, B.; Cheley, S.; Kasianowicz, J.; Song, L.; Gouaux, J.; Bayley, H. *Chem. Biol.* **1997**, *4*, 497–505.

(22) Robertson., J. W. F.; Rodrigues, C. G.; Stanford, V. M.; Rubinson, K. A.; Krasilnikov, O. V.; Kasianowicz, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8207.

(23) Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1079–1084.

(24) Oukhaled, G.; Mathe, J.; Biance, A. L.; Bacri, L.; Betton, J.; Lairez, D.; Pelta, J.; Auvray, L. *Phys. Rev. Lett.* **2007**, *98*, 15810(1–4).

(25) Pastoriza-Gallego, M.; Rabah, L.; Gibrat, G.; Thiebot, B.; van der Goot, F. G.; Auvray, L.; Betton, J. M.; Pelta, J. *J. Am. Chem. Soc.* **2011**, *133*, 2923–2931.

(26) Oukhaled, A.; Cressiot, B.; Bacri, L.; Pastoriza-Gallego, M.; Betton, J. M.; Bourhis, E.; Jede, R.; Gierak, J.; Auvray, L.; Pelta, J. *ACS Nano* **2011**, *5*, 3628–3638.

(27) Quiocho, F. A.; Spurlino, J. C.; Rodseth, L. E. *Structure* **1997**, *5*, 997–1015.

(28) Bertz, M.; Rief, M. *J. Mol. Biol.* **2008**, *378*, 447–458.

(29) Chinappi, M.; Cecconi, F.; Casciola, C. M. *Philos. Mag.* **2011**, *91*, 2034–2048.

(30) Hyeon, C.; Thirumalai, D. *Nat. Commun.* **2011**, *2*, 487.

(31) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics Modell.* **1996**, *14*, 33–38.

(32) Clementi, C.; Nymeyer, H.; Onuchic, J. N. *J. Mol. Biol.* **2000**, *298*, 937–953.

(33) Ammenti, A.; Cecconi; Marini-Bettolo-Marconi, U.; Vulpiani, A. *J. Phys. Chem. B* **2009**, *113*, 10348–10356.

(34) Lu, H.; Schulten, K. *Chem. Phys.* **1999**, *247*, 141–153.

(35) Ganesh, C.; Shah, A. N.; Swaminathan, C. P.; Surolia, A.; Varadarajan, R. *Biochemistry* **1997**, *36*, 5020–5028.

(36) Kabsch, W. *Acta Crystallogr., Sect. A* **1976**, *32*, 922–923.

(37) Redner, S. *A guide to first-passage processes*, 1st ed.; Cambridge University Press: Cambridge, 2001.

(38) Huang, L.; Kirmizialtin, S.; Makarov, D. *J. Chem. Phys.* **2005**, *123*, 124903.

(39) Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759–8770.