ACCOUNTS
of chemical research

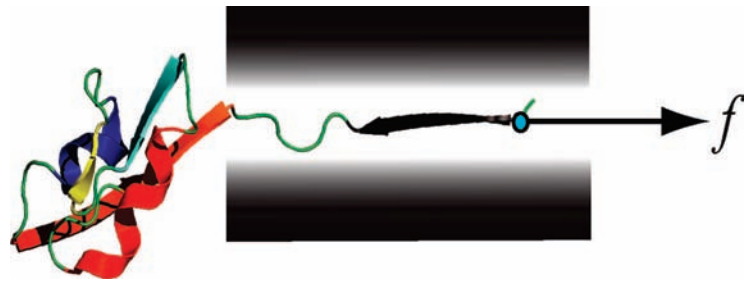
Computer Simulations and Theory of Protein Translocation

DMITRII E. MAKAROV

Department of Chemistry and Biochemistry and Institute for Theoretical Chemistry, the University of Texas at Austin, Austin, Texas, 78712

RECEIVED ON JUNE 3, 2008

CONSPECTUS



The translocation of proteins through pores is central to many biological phenomena, such as mitochondrial protein import, protein degradation, and delivery of protein toxins to their cytosolic targets. Because proteins typically have to pass through constrictions that are too narrow to accommodate folded structures, translocation must be coupled to protein unfolding. The simplest model that accounts for such co-translocational unfolding assumes that both translocation and unfolding are accomplished by pulling on the end of the polypeptide chain mechanically. In this Account, we describe theoretical studies and computer simulations of this model and discuss how the time scales of translocation depend on the pulling force and on the protein structure.

Computationally, this is a difficult problem because biologically or experimentally relevant time scales of translocation are typically orders of magnitude slower than those accessible by fully atomistic simulations. For this reason, we explore one-dimensional free energy landscapes along suitably defined translocation coordinates and discuss various approaches to their computation. We argue that the free energy landscape of translocation is often bumpy because confinement partitions the protein's configuration space into distinct basins of attraction separated by large entropic barriers. Favorable protein−pore interactions and nonnative interactions within the protein further contribute to the complexity.

Computer simulations and simple scaling estimates show that forces of just 2−6 pN are often sufficient to ensure transport of unstructured polypeptides, whereas much higher forces are typically needed to translocate folded protein domains. The unfolding mechanisms found from simulations of translocation are different from those observed in the much better understood case of atomic force microscopy (AFM) pulling studies, in which proteins are unraveled by stretching them between their N- and C-termini. In contrast to AFM experiments, single-molecule experimental studies of protein translocation have just started to emerge. We describe one example of a collaborative study, in which dwell times of β-hairpin-forming peptides inside the α-hemolysin pore were both measured experimentally and estimated using computer simulations. Analysis of the simulated trajectories has explained the experimental finding that more stable hairpins take, on the average, longer to traverse the pore.

Despite the insight we have gained, the general relationship between the structure of proteins and their resistance to mechanically driven co-translocational unfolding remains poorly understood. Future theoretical progress likely will be made in conjunction with single-molecule experiments and will require realistic models to account for specific protein−pore interactions and for solvent effects.

1. Introduction

Translocation of proteins through pores is crucial to many biological phenomena.[1] For example, the mechanism through which protein toxins such as anthrax and botulinum reach the cytosol involves threading of toxin components through pores in intracellular membranes.[2,3] Other examples of protein translocation in the cell include protein

Published on the Web 12/15/2008 www.pubs.acs.org/acr
10.1021/ar800128x CCC: $71.50

degradation by ATP-dependent proteases, mitochondrial protein import, and protein synthesis.[4,5] A common feature of these phenomena is that the pore size is too small to accommodate proteins in their native conformations. For example, the narrowest constriction in the proteasome, the cell machine that degrades proteins, is only about 13 Å in diameter. It is therefore necessary for the proteins to unfold prior to or concurrently with their passage through the pore.

The rate of such co-translocational unfolding is often several orders of magnitude faster than that of chemical or thermal denaturation, suggesting a different unfolding mechanism.[4−6] One appealing view is that this mechanism is mechanical in nature: the unfolding is accomplished simply through pulling at one end of the polypeptide chain.[4−6] This Account provides a theoretical perspective on such coupled translocation and unfolding processes driven by mechanical forces. In the following, we will give theoretical estimates of the forces and free energy barriers involved and discuss how the resistance of proteins to co-translocational unfolding depends on their structure and the applied force.

The organization of the rest of this Account is as follows. Section 2 discusses the computational challenges posed by the protein translocation problem and critically reviews various approaches to attack those. The reader not interested in such technicalities can skip section 2 and proceed directly to section 3, which presents theoretical considerations and describes simulation results. Section 4 discusses open issues and future prospects.

2. Computational Methods

2.1. Brute Force Simulations of Translocation Dynamics. Tian and Andricioaei have recently used atomistic models to study the translocation of barnase through a model pore,[7] as well as the dynamics of a pore itself.[8] Wells et al.[9] have studied translocation of α-helical peptides through the α-hemolysin (αHL) pore using an atomistic model for both the pore and the peptides. Similar all-atom simulations have been performed for nucleic acids in pores, see, for example, refs 10 and 11. Such simulations provide a great deal of information about molecular details of translocation. Their utility is however limited by a several orders of magnitude time-scale gap between simulations and experiments. Typical simulation time scales range from nanoseconds to microseconds, while a typical experimentally measured translocation event, even for a short polypeptide, lasts milliseconds.[12] To make translocation happen on a simulation time scale, one resorts to applying driving forces that are much larger than the biologically or experimentally relevant ones. Typically, simulation forces are on the order of several hundred piconewtons, at least an order of magnitude higher than those in experiments. Use of coarse-grained models can partially alleviate the problem.[13−17] Lattice models have also been used to study protein translocation.[18]

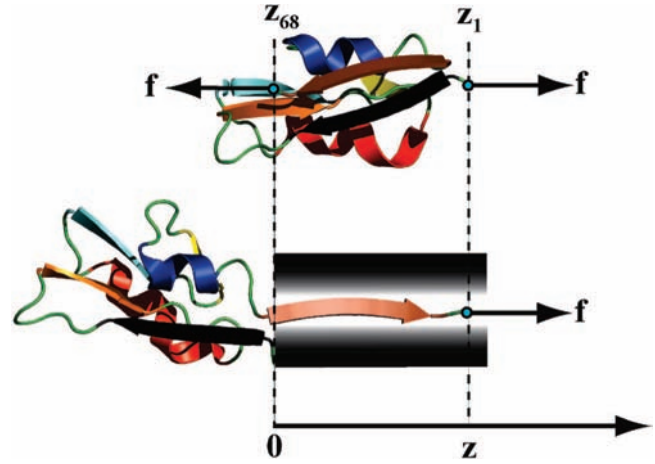


FIGURE 1. Mechanical stretching vs translocation. The reaction coordinate *z* is defined as the end-to-end distance projected onto the stretching force (mechanical stretching) and as the location of the chain end along the pore axis (translocation). This and other protein pictures were generated with the PyMol software.

2.2. Extrapolation to Low Driving Forces and Free Energy Calculations. One of the most widespread concepts in chemistry is that of a reaction coordinate, a collective order parameter that quantifies microscopic progress of a chemical reaction from its reactants state to the products state. By chemical reaction, we mean any conformational rearrangement taking the system from one basin of attraction (reactants) to another (products), protein unfolding being one example. For complex biomolecular phenomena, one's choice of the reaction coordinate is often guided by intuition. For mechanically driven processes, it is natural to choose the degree of freedom, *z*, that couples to the driving force.[19,20] This choice is illustrated in Figure 1 for the cases of protein translocation and mechanical stretching. Another convenient choice of the translocation reaction coordinate is the number of monomers that have advanced past the pore entrance. The free energy profile, $G(z)$, along the reaction coordinate *z* is related to its equilibrium probability distribution, $p(z)$

$$G(z) = -k_B T \ln p(z) \quad (1)$$

What does $G(z)$ tell us about translocation dynamics? Consider the simple model potential $G(z)$ shown in Figure 2. It consists of a minimum corresponding to the folded protein and a higher plateau corresponding to the protein unfolded

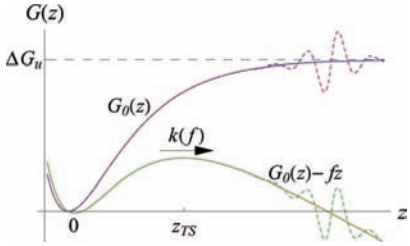


FIGURE 2. A model free energy profile, $G_f(z)$, shown at zero force and at a finite value of the force f.

through either stretching or pore confinement. If a pulling force, f, is applied as shown in Figure 1, the free energy becomes

$$G_f(z) = G(z) - fz \tag{2}$$

which has a barrier separating the folded and the unfolded states.

Variational transition state theory (TST) asserts that the rate constant for crossing the barrier has an upper bound that can be estimated as

$$k_{TST}(f) = \nu_{TST} \exp[-G_f(z_{TS})/(k_B T)] \tag{3}$$

where z_{TS} is the transition-state value of z corresponding to the top of the barrier and ν_{TST} is a prefactor (see ref 21 for further details). If $G(z)$ has more than one minimum, translocation can be viewed as hopping between those minima, the hopping rates being determined by eq 3 for each barrier separately.

Such a one-dimensional view of translocation is commonly invoked to interpret experimental data.[22−24] Computationally, $G(z)$ can be evaluated by using the umbrella sampling method (e.g., see ref 14) from a set of constrained equilibrium trajectories. Fast, nonequilibrium translocation trajectories such as those obtained in brute force simulations (see section) can also be used to estimate $G(z)$. One approach[25−28] is based on the famous Jarzynski identity, which relates nonequilibrium work to equilibrium free energy differences. To compute $G(z)$ using this approach, one basically needs to sample multiple short-time unfolding trajectories produced by fast pulling. In other words, one long-time equilibrium trajectory gets replaced by multiple nonequilibrium ones. Obviously, this does not guarantee any computational savings although this approach has the advantage of being trivially parallel.[27] A recent study[26] shows that practical implementation of the method can be problematic because it presents very stringent requirements for the sampling of the initial state. Specifically, the initial conditions for each nonequilibrium trajectory must be sampled from the equilibrium ensemble, which includes both folded and unfolded conformations of the protein. The

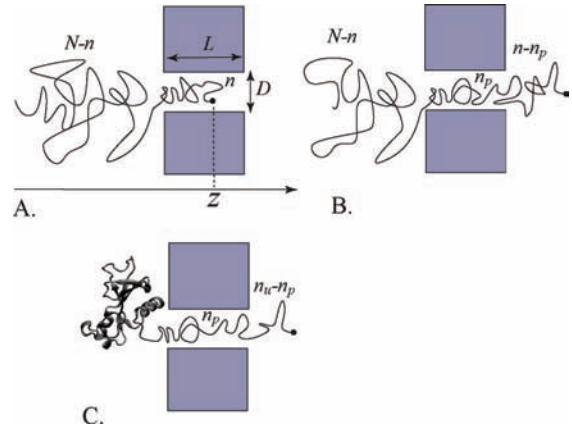


FIGURE 3. (A, B) Two stages of translocation of an unstructured polymer through a pore. During the first stage the chain has not yet emerged on the other side of the pore. (C) Translocation of a folded protein through a pore. An unstructured targeting sequence enters the pore first.

implication is that even though the protein that arrives at the pore entrance (as in Figure 1) is likely to be in the native state, starting each pulling trajectory from the folded initial condition and neglecting the very rare instances where the protein is unfolded prior to its entering the pore may result in large errors in $G(z)$.

Other methods of reconstructing $G(z)$ use the model in which the dynamics along z is described by the Langevin equation

$$\mu\ddot{z} = -\mathrm{d}G_f/\mathrm{d}z - \mu\gamma\dot{z} + R(t) \tag{4}$$

where μ is an effective mass, γ is a friction coefficient, and $R(t)$ is a random force.[21] One approach is based on denoising the trajectory first (i.e., averaging out $R(t)$) and then explicitly subtracting the friction force term $\mu\gamma\dot{z}$ to estimate $\mathrm{d}G_f/\mathrm{d}z$.[29,30] Nummela and Andricioaei[31] have proposed a different method, which reweights Langevin trajectories obtained at high forces to recover correct sampling at a lower force.

I will conclude this section with an example that emphasizes the fundamental difficulty with extrapolating high force data (whether experimental or simulated) to low forces. Consider the solid-line and the dashed-line free energy profiles shown in Figure 2. The only difference between the two is the wiggly part at high values of z. When a sufficiently high force is applied, the tilted potentials $G_f(z) = G(z) - fz$ will have identically the same barriers. The unfolding rate, $k(f)$, becomes insensitive to the features of the potential to the right of the transition state and thus both dashed- and solid-line potentials will have essentially the same unfolding kinetics. The wiggly features of the dashed potential are effectively hidden in a high-force experiment although they matter at lower forces.

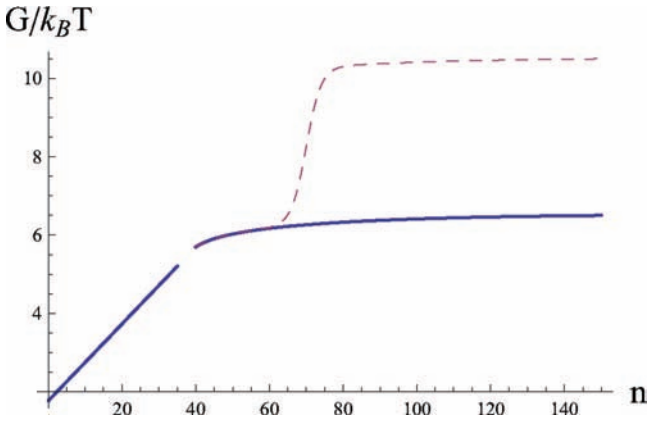


FIGURE 4. The free energy profile $G(n)$ described by eqs 6−8 assuming $N = 300$, $n_p = 35$, $a = 5$ Å, and $D = 20$ Å. The dashed line shows how $G(n)$ is modified when the chain contains a folded domain, as in Figure 3C.

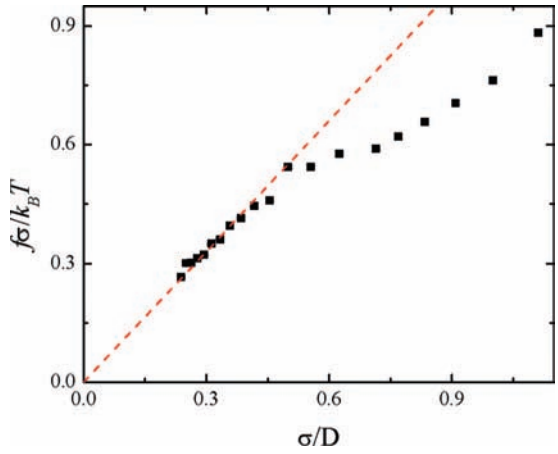


FIGURE 5. The expulsion force acting on an unstructured polypeptide in a pore as a function of the pore diameter (courtesy of Lei Huang). The model of the pore and the peptide is described in ref 16. Here $\sigma = 3.8$ Å is the distance between adjacent α carbons.

2.3. Validity and Limitations of the One-Dimensional Translocation Model. The free energy profile, $G(z)$, is only of value insofar as it provides information about translocation dynamics. Since TST only provides an upper bound for the rate, a poor choice of the reaction coordinate z can render its prediction worthless. While there is no guarantee that a 1D projection onto a single coordinate z would provide a meaningful picture of the dynamics, our recent study of polymer reversal inside a pore[32] provides anecdotal support for 1D models of the type considered here. Consider a polymer chain inside a long narrow pore. There are two equivalent directions, in which the chain can align itself along the pore. Reversing the direction would require the chain to fold onto itself, which clearly involves a large entropic barrier. It is possible to compute the rate of such reversal exactly. On the other hand, one can define a simple reaction coordinate z equal to the relative displacement of the chain ends along the pore and estimate the reversal rate using eq 3. We found that although the numerical value of the TST rate is significantly off, it faithfully reproduces the *exponentially strong* dependence of the rate on the pore diameter and the polymer length, while the transmission factor (i.e., the ratio $\kappa = k/k_{TST}$ of the true rate and the TST rate) shows a much weaker power-law dependence. We have also attempted to improve upon TST and to estimate the transmission factor by assuming that the dynamics along z is governed by eq 4 but found that the scaling properties of κ are inconsistent with the simple Langevin equation.[32]

3. Results

3.1. Forces and Free Energies Involved in Translocation: Simple Theoretical Estimates. We start with the translocation of an unfolded polypeptide as shown in Figure 3. A long chain enters a pore, which we model as a cylinder with a diameter D and length L. Assuming denatured proteins to behave as self-avoiding random walks, their characteristic size in free space is given by the Flory formula,[33]

$$R \approx aN^{\nu} \tag{5}$$

where N is the number of monomers and $\nu \approx 3/5$.

To measure the translocation progress, we count the number of monomers, n, that have advanced past the pore entrance. The free energy of the polymer, as a function of n, can be estimated as

$$G(n)/(k_BT) = \begin{cases} (1-\gamma_1)\ln(N-n) + G_{conf}(n)/(k_BT) & n < n_p \\ (1-\gamma_1)\ln(N-n) + (1-\gamma_1)\ln(n-n_p) + G_{conf}(n_p)/(k_BT) & n \gg n_p \end{cases} \tag{6}$$

where $\gamma_1 \approx 0.69$ and n_p is the number of monomers that can be accommodated inside the pore. The two cases in eq 6 correspond to Figure 3A,B, and the logarithmic terms in each case account for the entropies of the chain segments dangling outside the pore,[34,35] while $G_{conf}(n)$ is the free energy cost of confining n monomers inside the pore. Its estimate can be obtained from de Gennes' blob picture:[33]

$$\Delta G_{conf}(n)/(k_BT) \approx n\left(\frac{a}{D}\right)^{1/\nu} \tag{7}$$

The blob model also provides a scaling estimate for n_p:

$$n_p \approx \frac{L}{D}\left(\frac{D}{a}\right)^{1/\nu} \quad (8)$$

For example, assuming typical values $L = 10$ nm, $D = 2$ nm, and $a = 0.5$ nm, we find $n_p \approx 50$.

A typical $G(n)$ is shown in Figure 4. In the initial stage of translocation ($n < n_p$), entropic penalty is paid to confine additional monomers inside the pore. The corresponding segment of $G(n)$ is nearly linear and is dominated by eq 7. Once the chain end emerges on the other side of the pore ($n > n_p$), the confinement entropy becomes constant and $G(n)$ has a weaker logarithmic n dependence. The gap between the two segments of $G(n)$ indicates that eq 6 is not applicable when $n \approx n_p$.

Two observations are in order here. First, for typical L, D, and N, the overall free barrier is only a few k_BT, suggesting that denatured proteins can easily traverse biological pores. Second, using eqs 6−8, one finds that for typical biological pores the overall translocation free energy barrier is dominated by the free energy cost of confining the initial n_p monomers inside the pore, except for unrealistically long chains ($N \geq 10^5$).

Expulsion of the chain from the pore can be counterbalanced by applying a pulling force f_e to its end monomer, as in Figure 1. If the position of the monomer along the pore axis is z (see Figure 3A), then this expulsion force is given by

$$f_e = \frac{dG}{dz} = \frac{dG/dn}{dz/dn} \quad (9)$$

The highest expulsion force is achieved when $n < n_p$ and is given by

$$f_e \approx \frac{k_BT}{D} \quad (10)$$

which follows from eqs 7 and 9 and the scaling relationship[33] $n \approx (z/D)(D/a)^{1/\nu}$. For typical pore diameters, $D = 10-30$ Å,[24] this corresponds to $f_e \approx 2-6$ pN. This force is rather small and is readily achieved in single-molecule translocation experiments. Indeed, Oukhaled et al.[36] found maltose binding protein to enter the αHL pore after the molecule was chemically denatured.

The blob estimates used above assume that the pore diameter is much larger than the polymer's persistence length, l_p, a condition that is often violated in narrow biological pores. In the opposite limit of strong confinement ($D \ll l_p$), the expulsion force can be estimated as[37]

$$f_e \approx \frac{k_BT}{\lambda} \quad (11)$$

where $\lambda = l_p^{1/3}D^{2/3}$ is Odijk's "deflection length". If $l_p \simeq D$, both eq 11 and eq 10 predict the same expulsion force. Simulations of a random-coil-like peptide[16] yield the relationship between f_e and D plotted in Figure 5, showing that eq 10 (with a proportionality coefficient close to 1) remains adequate for $D \geq 9$ Å.

Biological translocation of proteins is commonly initiated by unstructured targeting presequences. The above estimates suggest that such sequences should penetrate pores with relative ease, especially in the presence of favorable electrostatic peptide−pore interactions. Indeed, the electrostatic force that acts on a targeting sequence in the case of mitochondrial protein import was estimated to be 2−3 pN per elementary charge,[38] which, depending on the overall charge and the location of the charged groups,[34] may be sufficient to ensure barrierless translocation of the peptides. Recently, such excursions of targeting sequences into protein pores have been demonstrated by single-molecule experiments,[39,40] where unstructured polypeptides fused to protein domains were coaxed into an αHL pore that was modified to contain electrostatic "traps".

We now turn to the situation illustrated in Figure 3C, where an unstructured targeting sequence leads the way to the translocation of an initially folded domain. The domain arrives at the pore at $n = n_u$ and plugs up its entrance. Let us assume, for now, that the domain unfolds in an all-or-none fashion as soon as a small number, Δn_u of its residues become separated from the rest of it by entering the pore. This results in a characteristic free energy "step", ΔG_u, shown as a dashed line in Figure 4. The expulsion force acting on the domain during its unfolding can be estimated as

$$f_e \approx \frac{\Delta G_u}{\Delta z} \approx \frac{\Delta G_u}{\Delta n_u}\frac{\Delta n_u}{\Delta z} \quad (12)$$

If we choose $\Delta z \approx 20$ Å (corresponding to $\Delta n_u \approx 5$ residues) and use $\Delta G_u \approx 10$ kcal/mol comparable with typical free energies of protein unfolding, this gives $f_e \approx 35$ pN. We see that the forces required to accomplish translocation of an initially folded protein are generally much higher than those for denatured proteins and are comparable to those typically encountered in atomic force microscopy (AFM) pulling experiments.[41]

If a weaker pulling force, $f < f_e$, is applied at the chain end, it will lower the step by $f\Delta z$. The TST rate for surmounting this force-modified free energy barrier is roughly estimated as

$$k_{\mathrm{trans}}(f) \approx \nu \exp\left(-\frac{\Delta G_{\mathrm{u}} - f\Delta z}{k_{\mathrm{B}}T}\right) = k_{\mathrm{trans}}(f=0) \exp\left(\frac{f}{f_{\mathrm{c}}}\right) \tag{13}$$

where ν is a prefactor and the characteristic force,

$$f_{\mathrm{c}} = \frac{k_{\mathrm{B}}T}{\Delta z} = f_{\mathrm{e}}\frac{k_{\mathrm{B}}T}{\Delta G_{\mathrm{u}}} \tag{14}$$

reflects how sensitive this rate is to the pulling force.

These considerations show that a more "brittle" protein (small value of the displacement Δz required for unfolding) would produce a higher expulsion force f_{e} but would also be more sensitive to pulling (i.e., a higher value of f_{c}). A less brittle protein domain unraveling continuously over a large displacement, Δz, rather than undergoing an all-or-none denaturation would generally require a lower force f_{e} to induce its barrierless translocation.

3.2. Free Energy Profiles from Simulations: Translocation versus Mechanical Stretching. At the time of writing, force-controlled, mechanically driven co-translocational protein unfolding still remains outside the realm of single-molecule experiments. Consequently, several workers have sought insights from single-molecule AFM pulling studies, where unfolding is induced by stretching proteins between their C- and N-termini.[6,41,42] To compare the two kinds of mechanical unfolding for a model protein,[15] we plot in Figure 6 the free energy profiles, $G_f(z)$, along their respective reaction coordinates z (defined in Figure 1).

In both cases, the zero-force free energy, $G_0(z)$, increases monotonically in a series of steps. Each step corresponds to a structural transition, in which the protein's structure is partially destroyed. For example, the step 1 to 2 in Figure 6B corresponds to the peeling of a β-strand from the rest of the structure. The situation predicted by simulations[15] is therefore more complicated than the single-step scenario discussed in section 3.1.

In the presence of a pulling force f, $G_f(z)$ exhibits a series of minima corresponding to translocation intermediates. We emphasize that our model[15] assumes purely repulsive protein−pore interactions so the translocation intermediates are created by the pore confinement. Even a protein that displays a perfectly cooperative two-state thermal denaturation can have multiple translocation intermediates. The same conclusion was reached by Tian and Andricioaei,[7] who observed intermediates in the translocation trajectories of barnase. The fact that confinement alone can lead to multiple metastable conformations is not surprising; In fact, in section 2.3, we have discussed a much simpler yet related phenomenon, where the confinement of a polymer in a cylinder partitioned the configuration space into two distinct basins of attraction corresponding to different polymer alignments.

Although the overall free energy cost of squeezing the protein into the pore is the same, the free energy landscape, $G_f(z)$, traversed by a protein that is pulled by a *finite* force f, the translocation intermediates, and thus kinetics depend on which chain end enters the pore first. The authors of ref 43 describe this kind of asymmetry using the example of carrying a Christmas tree through a door, which is obviously easier when the top goes first. Furthermore, Figure 6B,C shows that in each case the domain unfolds from the end by which it is pulled, in agreement with the earlier conclusion reached by the Matouschek group[44,45] in their experimental studies of mitochondrial import and protein degradation.

The intermediate structures and the magnitudes of the barriers observed in the case of translocation are also different from those in the case of mechanical stretching (Figure 6A). Tian and Andricioaei[7] report the same finding for barnase. Translocation generally involves high barriers (e.g., see the barrier 1 to 2 in Figure 6B) that are not present in stretching. This results in the rates of translocation being slower than those for stretching,[15−17] at least for the model systems studied so far. We note that in contrast to translocation of unstructured peptides, the high barrier to translocation of proteins *is not* due to confinement-induced configurational entropy loss. In fact, we found[16] our model protein to *gain* entropy upon its translocation, as the positive entropy of unfolding rather than confinement entropy dominates the overall entropy change.

The difference between the stretching and the translocation pathways is not unexpected. In a mechanical stretching experiment, the force is applied between a pair of residues. Changing these residues may alter the unfolding mechanism and rate.[46−48] In contrast, in the case of translocation, the applied force is distributed over the rim of the pore. To account for this, West et al.[17] proposed a simple albeit ad hoc procedure, where the translocation rate is predicted as a linear combination of the mechanical stretching rates between the residue that is pulled into the pore and each residue situated near the pore rim.

Finally, the force dependence of the translocation rate suggested by Figure 6 is more complex than our eq 13 because the location of the rate-limiting step corresponding to the highest barrier in $G_f(z)$ depends on the magnitude of the force.[15] For example, in Figure 6B, it is the transition 1 to 2 at high forces, while at low forces the complete unraveling of structure 3 is rate-limiting.

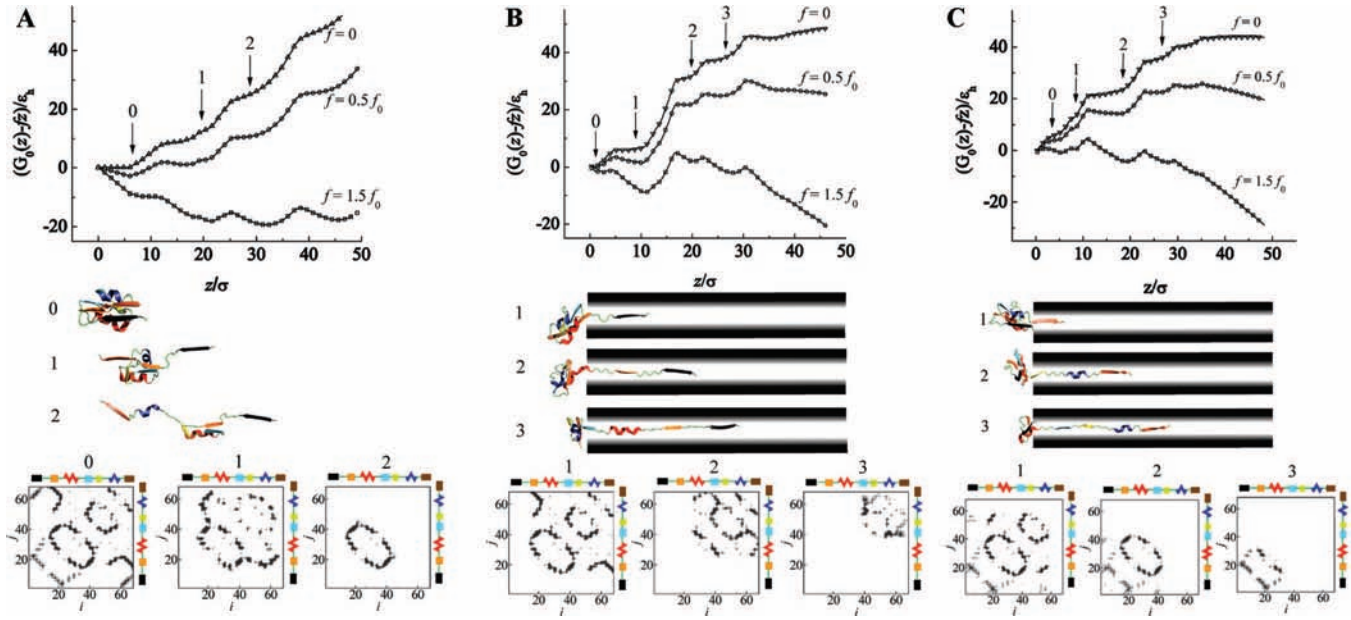


FIGURE 6. Comparison of mechanical stretching (A) with mechanically driven translocation through a narrow pore (B, C) for a ubiquitin-like domain.[15] The force is applied at the N-terminus and the C-terminus of the chain in cases B and C, respectively. The potential of mean force, $G_f(z)$, is plotted as a function of the translocation coordinate defined in Figure 1. The pore radius is $r_{pore} = 3.8$ Å. The free energy is measured in units of ε_h, and the force is measured in units of $f_0 = \varepsilon_h/\sigma$, where ε_h sets the scale of hydrophobic interactions within the protein and σ is the distance between adjacent α-carbons. For $\varepsilon_h = 1$ kcal/mol, $f_0 = 18$ pN. The simulations were performed at $T = 0.26\varepsilon_h/k_B$. The minima of $G_f(z)$ correspond to translocation intermediates, whose structure is shown along with their contact maps. The contact maps were constructed by plotting the contacting pairs of residues $\{i,j\}$ such that $|\mathbf{r}_j - \mathbf{r}_i| < 7.5$ Å and $|i - j| > 3$. The darkness of each point reflects the probability of observing the respective contact in the equilibrium ensemble of conformations for given z, black corresponding to the highest probability.

3.3. Electrophoretically Driven Translocation of Short Peptides. Translocation of short β-hairpin-forming peptides across the αHL pore provides an example of a system well-characterized both by single-molecule experiments and by simulations.[12] It was experimentally found that less stable hairpins (i.e., ones that are more likely to be unfolded) tend, on average, to traverse the pore faster.[12] This finding was somewhat counterintuitive because no significant entropic or enthalpic barrier was expected to impede translocation of folded hairpins, whose dimensions are comparable with the pore size (see Figure 7B). A plausible explanation was provided by coarse-grained Langevin dynamics simulations, which used a C_α-only representation for both the pore and the peptides. We have designed sequences of four β-hairpin peptides, which differed in their thermodynamic stability.[12] The fraction of folded molecules in the equilibrium ensemble was 82%, 70%, 54%, and 32% for peptides 1−4, respectively. Residues in each of the peptides were assigned charges corresponding to those in one of the peptides studied experimentally.[12] The simulation results are summarized in Figure 7A, which shows that in the low force regime the average translocation time decreases with the decreasing fraction of folded peptides.

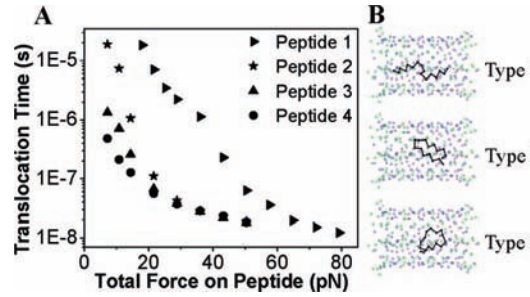


FIGURE 7. (A) The average translocation time as a function of the total force acting on β-hairpin-forming peptides.[12] The peptide stability decreases from peptide 1 to peptide 4. (B) Typical conformations of the peptide inside the β-barrel part of the αHL pore corresponding to three types of observed translocation trajectories.

This dependence can be understood if one considers the typical translocation trajectories illustrated in Figure 7B. In the trajectories of type 1, the peptide enters the pore in an extended conformation and its residues are threaded through the pore in a single file. This type of trajectory corresponds to the fastest observed translocation times. In type 2 trajectories, corresponding to longer translocation times, the peptide remains folded as it goes through the pore. Type 3 trajectories, which were only observed for one of the peptides, correspond to the slowest translocation events, in which a

peptide−pore complex is formed and the peptide remains trapped inside the pore in a misfolded conformation until it either refolds or attains an extended conformation to exit the pore. Less-stable hairpins are more likely to enter the pore in an open conformation, resulting in a fast type 1 translocation event. In contrast, translocation of more stable peptides is more likely to occur via the type 2 or 3 trajectories.

The statistics of the hairpin translocation times is analyzed in ref 16. At low forces, their distribution is close to exponential, and the mean translocation time exhibits a nearly exponential force dependence (see eq 13) indicating that translocation involves barrier crossing. This agrees with the experimental voltage dependence.[12] At high forces, the translocation process eventually becomes downhill in free energy (see Figures 2 and 6B,C), leading to a much weaker force dependence. In this limit, the mean translocation time is inversely proportional to the force $t_{trans} \propto Af^{-1}$, where the constant A is determined by the hairpin's mobility. The transition from barrier crossing to a downhill regime has been observed in another simulation study.[49]

4. Outlook

Despite the studies reviewed here, the general relationship between the resistance of proteins to mechanically driven cotranslocational unfolding and their structure remains an open issue. In the context of mechanical unfolding through stretching, this question has been largely resolved over the past several years and the key structural motifs maximizing the mechanical resistance have been identified both theoretically and experimentally.[50,51] The translocation case is however more difficult because of the distributed character of the force exerted by the pore and the complexity of the unfolding mechanism.

Such complexity manifests itself in translocation intermediates[7,15,16] created through several mechanisms: Confinement alone can partition the configuration space into distinct basins of attraction or induce nonnative interactions.[7,12] In addition, attractive protein−pore interactions can trap proteins.[12,40] This means that a protein's translocation free energy landscape may be bumpy even if its conventional "folding" landscape is smooth.

In the case of AFM stretching, all-atom free energy calculations benefit considerably from the fact that mechanical unfolding transition states are often very native-like.[30] In contrast, translocation transition states may be very different from native conformations,[15] thus presenting the computational problem of accurate sampling of partially unfolded states. Furthermore, details of protein−pore interactions and solvent effects[52] are important thus limiting the utility of simple models. The inverse problem of extracting the underlying free energy profiles from experimental translocation data[23] is also more challenging than in the AFM stretching case because single-well model potentials are often inapplicable.

Initial theoretical studies[7,14,15] were motivated by the lack of experimental information about mechanistic details of protein translocation. Experiments are however quickly catching up, and it is likely that much of the further progress in this area will be made in conjunction with single-molecule experiments.[2,3,12,24,39,40,53,54] Recently, optical tweezers have been used to pull on DNA molecules in pores.[55,56] Although similar experiments have not yet been done for peptides, this new approach holds great promise because it provides a wider range of tunable forces and the ability to spatially locate peptides.

Another largely open issue is concerned with the physical mechanisms through which pulling forces are actually generated *in vivo*. Several models describing such pulling have recently been proposed.[38,57,58] Given the complexity of the problem, developing a coherent picture of biological translocation will likely require a cooperative effort between experiment and theory.

I am grateful to former members of my group, Lei Huang and Serdal Kirmizialtin, who have carried out much of the work described here. I am also thankful to our collaborators, Venkat Ganesan and Liviu Movileanu, for their contributions and numerous discussions. This work was supported in part by the Robert A. Welch Foundation (Grant F-1514) and by the National Science Foundation (Grant CHE 0347862).

BIOGRAPHICAL INFORMATION

Dmitrii E. Makarov received his Ph.D. in 1992 at the Institute of Chemical Physics in Moscow. After postdoctoral research at the University of Illinois at Urbana−Champaign and the University of California, Santa Barbara, he joined the faculty at the University of Texas, where he is currently an Associate Professor of Chemistry. His current research interests are mostly in the area of theory and computer simulations of the dynamics of biomolecules.

REFERENCES

1 Wickner, W.; Schekman, R. Protein translocation across biological membranes. *Science* **2005**, *310*, 1452–1456.

2 Krantz, B. A.; Melnyk, R. A.; Zhang, S.; Juris, S. J.; Lacy, D. B.; Wu, Z.; Finkelstein, A.; Collier, R. J. A phenylalanine clamp catalyzes protein translocation through the anthrax toxin pore. *Science* **2005**, *309*, 777–781.

3 Fischer, A.; Montal, M. Single molecule detection of intermediates during botulinum neurotoxin translocation across membranes. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10447–10452.

4 Matouschek, A. Protein unfolding - an important process in vivo. *Curr. Opin. Struct. Biol.* **2003**, *13*, 98–109.

5 Prakash, S.; Matouschek, A. Protein unfolding in the cell. *Trends Biochem. Sci.* **2004**, *29*, 593–600.

6 Matouschek, A.; Bustamante, C. Finding a protein's Achilles heel. *Nat. Struct. Biol.* **2003**, *10*, 674–676.

7 Tian, P.; Andricioaei, I. Repetitive pulling catalyzes co-translocational unfolding of barnase during import through a mitochondrial pore. *J. Mol. Biol.* **2005**, *350*, 1017–1034.

8 Tian, P.; Andricioaei, I. Size, motion, and function of the SecY translocon revealed by molecular dynamics simulations with virtual probes. *Biophys. J.* **2006**, *90*, 2718–2730.

9 Wells, D. B.; Abramkina, V.; Aksimentiev, A. Exploring transmembrane transport through alpha-hemolysin with grid-steered molecular dynamics. *J. Chem. Phys.* **2007**, *127*, 125101.

10 Yeh, I. C.; Hummer, G. Nucleic acid transport through carbon nanotube membranes. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12177–12182.

11 Heng, J. B.; Aksimentiev, A.; Ho, C.; Marks, P.; Grinkova, Y. V.; Sligar, S.; Schulten, K.; Timp, G. Stretching DNA using the electric field in a synthetic nanopore. *Nano Lett.* **2005**, *5*, 1883–1888.

12 Goodrich, C. P.; Kirmizialtin, S.; Huyghues-Despointes, B. M.; Zhu, A.; Scholtz, J. M.; Makarov, D. E.; Movileanu, L. Single-molecule electrophoresis of beta-hairpin peptides by electrical recordings and Langevin dynamics simulations. *J. Phys. Chem. B* **2007**, *111*, 3332–3335.

13 Friedel, M.; Sheeler, D. J.; Shea, J. E. Effects of confinement on the thermodynamics and kinetics of folding of a minimalistic $\mathbf{B}$-barrel protein. *J. Chem. Phys.* **2003**, *118*, 8106–8113.

14 Kirmizialtin, S.; Ganesan, V.; Makarov, D. E. Translocation of a beta-hairpin-forming peptide through a cylindrical tunnel. *J. Chem. Phys.* **2004**, *121*, 10268–10277.

15 Huang, L.; Kirmizialtin, S.; Makarov, D. E. Computer simulations of the translocation and unfolding of a protein pulled mechanically through a pore. *J. Chem. Phys.* **2005**, *123*, 124903.

16 Kirmizialtin, S.; Huang, L.; Makarov, D. E. Computer simulations of protein translocation. *Phys. Status Solidi b* **2006**, *243*, 2038–2047.

17 West, D. K.; Brockwell, D. J.; Paci, E. Prediction of the translocation kinetics of a protein from its mechanical properties. *Biophys. J.* **2006**, *91*, L51−L53.

18 Contreras Martinez, L. M.; Martinez-Veracoechea, F. J.; Pohkarel, P.; Stroock, A. D.; Escobedo, F. A.; DeLisa, M. P. Protein translocation through a tunnel induces changes in folding kinetics: A lattice model study. *Biotechnol. Bioeng.* **2006**, *94*, 105–17.

19 Best, R. B.; Paci, E.; Hummer, G.; Dudko, O. K. Pulling direction as a reaction coordinate for the mechanical unfolding of single molecules. *J. Phys. Chem. B* **2008**, *112*, 5968–5976.

20 Kirmizialtin, S.; Huang, L.; Makarov, D. E. Topography of the free energy landscape probed via mechanical unfolding of proteins. *J. Chem. Phys.* **2005**, *122*, 234915.

21 Hanggi, P.; Talkner, P.; Borkovec, M. 50 years fter Kramers. *Rev. Mod. Phys.* **1990**, *62*, 251–341.

22 Movileanu, L.; Schmittschmitt, J. P.; Scholtz, J. M.; Bayley, H. Interaction of peptides with a protein pore. *Biophys. J.* **2005**, *89*, 1030–1045.

23 Dudko, O. K.; Mathe, J.; Szabo, A.; Meller, A.; Hummer, G. Extracting kinetics from single-molecule force spectroscopy: nanopore unzipping of DNA hairpins. *Biophys. J.* **2007**, *92*, 4188–4195.

24 Movileanu, L. Squeezing a single polypeptide through a nanopore. *Soft Matter* **2008**, *4*, 925–931.

25 Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* **2003**, *119*, 3559–3566.

26 West, D. K.; Olmsted, P. D.; Paci, E. Free energy for protein folding from nonequilibrium simulations using the Jarzynski equality. *J. Chem. Phys.* **2006**, *125*, 204910.

27 Xiong, H.; Crespo, A.; Marti, M.; Estrin, D.; Roitberg, A. E. Free energy calculations with non-equilibrium methods: Applications of the Jarzynski relationship. *Theor. Chem. Acc.* **2006**, *116*, 338–346.

28 Hummer, G.; Szabo, A. Free energy surfaces from single-molecule force spectroscopy. *Acc. Chem. Res.* **2005**, *38*, 504–513.

29 Balsera, M.; Stepaniants, S.; Izrailev, S.; Oono, Y.; Schulten, K. Reconstructing potential energy functions from simulated force-induced unbinding processes. *Biophys. J.* **1997**, *73*, 1281–1287.

30 Li, P.-C.; Makarov, D. E. Theoretical studies of the mechanical unfolding of the muscle protein titin: Bridging the time-scale gap between simulation and experiment. *J. Chem. Phys.* **2003**, *119*, 9260–9268.

31 Nummela, J.; Andricioaei, I. Exact low-force kinetics from high-force single-molecule unfolding events. *Biophys. J.* **2007**, *93*, 3373–3381.

32 Huang, L.; Makarov, D. E. The rate constant of polymer reversal inside a pore. *J. Chem. Phys.* **2008**, *128*, 114903.

33 De Gennes, P. G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1979.

34 Mohan, A.; Kolomeisky, A. B.; Pasquali, M. Effect of charge distribution on the translocation of an inhomogeneously charged polymer through a nanopore. *J. Chem. Phys.* **2008**, *128*, 125104.

35 Muthukumar, M. Polymer translocation through a hole. *J. Chem. Phys.* **1999**, *111*, 10371–10374.

36 Oukhaled, G.; Mathe, J.; Biance, A. L.; Bacri, L.; Betton, J. M.; Lairez, D.; Pelta, J.; Auvray, L. Unfolding of proteins and long transient conformations detected by single nanopore recording. *Phys. Rev. Lett.* **2007**, *98*, 158101.

37 Bicout, D. J.; Burkhardt, T. W. Simulation of a semiflexible polymer in a narrow cylindrical pore. *J. Phys. A.: Math. Gen.* **2001**, *34*, 5745–5750.

38 Shariff, K.; Ghosal, S.; Matouschek, A. The force exerted by the membrane potential during protein import into the mitochondrial matrix. *Biophys. J.* **2004**, *86*, 3647–3652.

39 Mohammad, M. M.; Movileanu, L. Excursion of a single polypeptide into a protein pore: simple physics, but complicated biology. *Eur Biophys J* **2008**, *37*, 913–925.

40 Mohammad, M. M.; Prakash, S.; Matouschek, A.; Movileanu, L. Controlling a single protein in a nanopore through electrostatic traps. *J. Am. Chem. Soc.* **2008**, *130*, 4081–4088.

41 Sato, T.; Esaki, M.; Fernandez, J. M.; Endo, T. Comparison of the protein-unfolding pathways between mitochondrial protein import and atomic-force microscopy measurements. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17999–18004.

42 Wilcox, A. J.; Choy, J.; Bustamante, C.; Matouschek, A. Effect of protein structure on mitochondrial import. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15435–15440.

43 Lua, R. C.; Grosberg, A. Y. First passage times and asymmetry of DNA translocation. *Phys. Rev. E* **2005**, *72*, 061918.

44 Huang, S.; Ratliff, K. S.; Schwartz, M. P.; Spenner, J. M.; Matouschek, A. Mitochondria unfold precursor proteins by unraveling them from their N-termini. *Nat. Struct. Biol.* **1999**, *6*, 1132–1138.

45 Lee, C.; Schwartz, M. P.; Prakash, S.; Iwakura, M.; Matouschek, A. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol. Cell* **2001**, *7*, 627–637.

46 Brockwell, D. J.; Paci, E.; Zinober, R. C.; Beddard, G. S.; Olmsted, P. D.; Smith, D. A.; Perham, R. N.; Radford, S. E. Pulling geometry defines the mechanical resistance of a beta-sheet protein. *Nat. Struct. Biol.* **2003**, *10*, 731–737.

47 Carrion-Vazquez, M.; Li, H.; Lu, H.; Marszalek, P. E.; Oberhauser, A. F.; Fernandez., J. M. The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Biol.* **2003**, *10*, 738–743.

48 Li, P.-C.; Makarov, D. E. Simulation of the mechanical unfolding of ubiquitin: Probing different unfolding reaction coordinates by changing the pulling geometry. *J. Chem. Phys.* **2004**, *121*, 4826–4832.

49 Matysiak, S.; Montesi, A.; Pasquali, M.; Kolomeisky, A. B.; Clementi, C. Dynamics of polymer translocation through nanopores: Theory meets experiment. *Phys. Rev. Lett.* **2006**, *96*, 118103.

50 Eom, K.; Li, P.-C.; Makarov, D. E.; Rodin, G. J. Relationship between the mechanical properties and topology of cross-linked polymer molecules: Parallel strands maximize the strength of model polymers and protein domains. *J. Phys. Chem. B* **2003**, *107*, 8730–8733.

51 Sharma, D.; Perisic, O.; Peng, Q.; Cao, Y.; Lam, C.; Lu, H.; Li, H. Single-molecule force spectroscopy reveals a mechanically stable protein fold and the rational tuning of its mechanical stability. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9278–9283.

52 Sorin, E. J.; Pande, V. S. Nanotube confinement denatures protein helices. *J. Am. Chem. Soc.* **2006**, *128*, 6316–6317.

53 Jung, Y.; Bayley, H.; Movileanu, L. Temperature-responsive protein pores. *J. Am. Chem. Soc.* **2006**, *128*, 15332–15340.

54 Wolfe, A. J.; Mohammad, M. M.; Cheley, S.; Bayley, H.; Movileanu, L. Catalyzing the translocation of polypeptides through attractive interactions. *J. Am. Chem. Soc.* **2007**, *129*, 14034–14041.

55 Keyser, U. F.; Koeleman, B. N.; van Dorp, S.; Krapf, D.; Smeets, R. M.; Lemay, S. G.; Dekker, N. H.; Dekker, C. Direct force measurements on DNA in a solid-state nanopore. *Nat. Phys.* **2006**, *2*, 473–477.

56 Trepagnier, E. H.; Radenovic, A.; Sivak, D.; Geissler, P.; Liphardt, J. Controlling DNA capture and propagation through artificial nanopores. *Nano Lett.* **2007**, *7*, 2824–2830.

57 Goloubinoff, P.; De Los Rios, P. The mechanism of Hsp70 chaperones: (Entropic) pulling the models together. *Trends Biochem. Sci.* **2007**, *32*, 372–80.

58 Simon, S. M.; Peskin, C. S.; Oster, G. F. What drives the translocation of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3770–3774.