Single-molecule force spectroscopy of protein folding

Rafayel Petrosyan, Abhishek Narayan, Michael T. Woodside

Department of Physics, University of Alberta, Edmonton AB T6G 2E1 Canada

The use of force probes to induce unfolding and refolding of single molecules through the application of mechanical tension, known as single-molecule force spectroscopy (SMFS), has proven to be a powerful tool for studying the dynamics of protein folding. Here we provide an overview of what has been learned about protein folding using SMFS, from small, single-domain proteins to large, multi-domain proteins. We highlight the ability of SMFS to measure the energy landscapes underlying folding, to map complex pathways for native and non-native folding, to probe the mechanisms of chaperones that assist with native folding, to elucidate the effects of the ribosome on co-translational folding, and to monitor the folding of membrane proteins.

Introduction

The protein folding problem—understanding how polypeptide chains self-assemble into specific, complex three-dimensional structures—involves at its heart two distinct questions. The first of these is how the structure of a protein is encoded in the amino-acid sequence. Given that the relative ease of sequencing proteins as compared to determining their structure has led to a dramatically widening gap between the number of known protein sequences and solved structures [1], accurate prediction of structure from sequence provides a powerful tool for extending the reach of structural biology. Furthermore, knowledge of the sequence-structure relationship enables proteins with novel structures (and hence novel functions) to be designed, opening new vistas for bionanotechnology. Great strides have been made in recent years in both structure prediction and *de novo* design: by harnessing deep learning methods, AlphaFold2 achieved over 90% global distance test accuracy in the most recent critical assessment of protein structure prediction [2,3], a significant improvement over previous results, and increasing numbers of novel protein folds are being designed, including some with complex functions [4]. These advances reflect the significant progress that has been made in solving the first part of the folding problem, even though challenges remain, such as with the prediction and design of multi-domain and membrane proteins.

The second part of the protein folding problem centers on the question of how the self-assembly works as a dynamic process: what is the mechanism of the folding. A wide range of experimental techniques has been used to probe the dynamics of the polypeptide chain [5], complemented by theoretical studies and computational simulations, leading to the modern picture of folding as a diffusive search for the native structure through multiple microscopic trajectories, all guided by funnel-shaped energy landscapes describing the energy of the protein as a function of its possible conformations [6,7]. Although this framework is conceptually fairly straightforward, the combinatorial complexity of the problem makes predicting the folding dynamics of even a modestly sized protein very challenging. As a result, many interesting questions about folding dynamics and mechanisms remain to be addressed through experiments. Here we review recent progress in understanding folding at the level of single protein molecules, specifically using force spectroscopy.

Single-molecule experiments offer a number of advantages over measurements of the behavior of an ensemble of many molecules, which counteract the additional effort needed to study molecules one-by-one [8]. By monitoring conformational dynamics in individual molecules, rare and/or transient intermediates can be more readily observed [9–11], the succession of intermediates through which the folding progresses can be followed and the available pathways identified [11–13], and distinct sub-populations with different behaviors can be catalogued and studied separately [14–16]. Even the brief trajectories through the high-energy transition states that form the rate-limiting barrier can be observed directly [17–19]. In the case of single-molecule force spectroscopy (SMFS), folding is studied by attaching single protein molecules to a force probe, applying tension to the protein as a mechanical denaturant, and monitoring the change in the extension of the protein as it unfolds or refolds in response (Fig. 1A). The first SMFS studies of protein folding were done using an atomic force microscope (AFM) or optical tweezers as the force probe [20–22], but other force probes such as magnetic tweezers have since also become important tools expanding the technical reach of SMFS. Currently, SMFS measurements can achieve sub-pN force resolution, Å-scale spatial resolution, μs-scale temporal resolution, and measurement durations from seconds to hours or even longer, providing information on the structures of intermediate states (via their extension), the microscopic rates in and out of each state, the free energies of each state, and the pathways between states. Furthermore, SMFS is also excellent way to measure the energy landscapes underlying folding, determining barrier locations and heights, well and barrier curvatures, and landscape roughness [23]. We first discuss SMFS measurements of small globular proteins, which have been the most intensively studied with other techniques and simulations, before considering SMFS studies of multi-domain proteins, chaperone-assisted and co-translational folding, and finally folding of membrane proteins.

SMFS studies of single-domain globular proteins

Single-domain globular proteins have been the focus of much of the work investigating protein folding—regardless of the technique used—because they are small, relatively simple, and often fold reliably without co-factors like chaperones. As a result, they are experimentally and computationally more tractable than multi-domain proteins, even though most proteins have multiple domains. Not surprisingly, a significant amount of the work on protein folding using SMFS has also concentrated on small, single-domain proteins. Early work on tandem repeats of the I27 domain of the muscle protein titin by Rief *et al.* [20] (Fig. 1B) shows some of the characteristic features of SMFS measurements using AFM. Tethering one end of the molecule to the surface and the other end to the AFM tip (Fig. 1B, inset), the tip is moved away to ramp up the force. The force rises nonlinearly with extension until the protein ruptures, generating a "rip" in force-extension curves (FECs) where the extension abruptly increases owing to the unfolding while the force simultaneously drops as the force probe snaps back. Multiple rips are seen owing to unfolding of each of the repeats. These FECs can be fit to models of the polymer elasticity such as a worm-like chain (WLC), to obtain the contour length unfolded during each rip, ΔL_c; in this case, $\Delta L_c \sim 29$ nm, matching the length expected from the number of amino acids involved and the structure in which they are arranged. FECs are the most common type of SMFS measurement for protein folding, but in many cases a complementary modality is used where the force is kept constant using a force clamp. Such an approach was applied for example by Li & Fernandez [24] in studying polyubiquitin (Fig. 1C): here, the force was jumped to a high value and then kept constant as each ubiquitin unit unfolded, generating a staircase-like trajectory, after

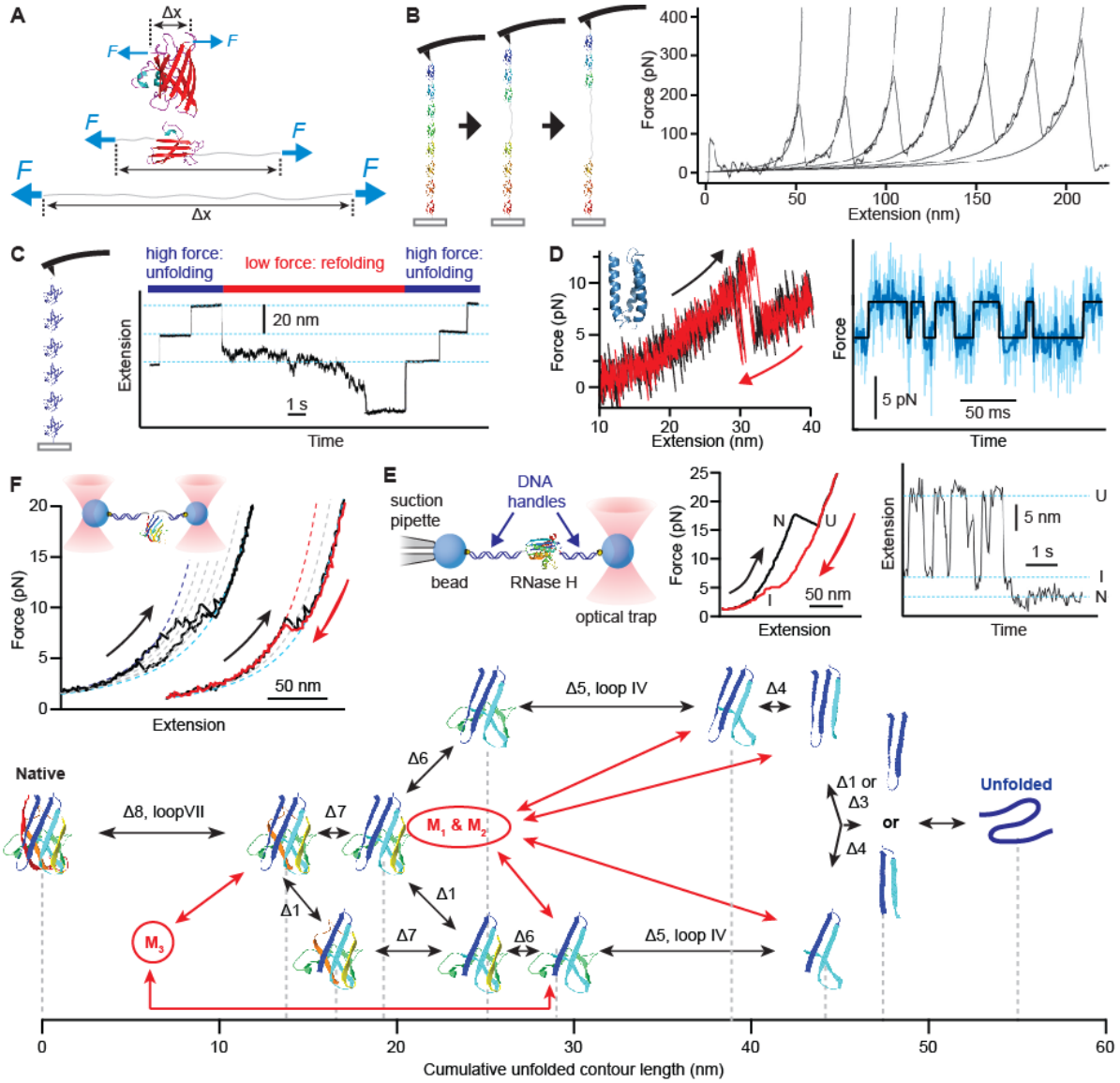

Figure 1: SMFS of folding in small proteins. (A) Schematic of SMFS. Force is applied to two points on a protein molecule, often but not always the termini. When the protein unfolds in response to increasing force, the unfolded part of the protein (grey) is stretched out, changing the extension of the molecule. (B) Under a force ramp applied by AFM, tandem repeats of the titin I27 domain (left) unfold individually in a two-state manner, generating a sequence of discrete rips in the FECs (right). Reprinted from Ref. [20], with permission from AAAS. (C) Under a constant force applied by AFM, individual units in polyubiquitin undergo two-state unfolding at high force (blue), generating a series of same-sized steps, and refold at low force with complex dynamics (red). Reprinted from Ref. [24], with permission from AAAS. (D) The designed protein α_3D hops between folded and unfolded states when the AFM force probe is moved at constant speed to ramp the force (left) or kept at a constant position in equilibrium. Reprinted with permission from Ref. [32]. (E) RNase H pulled with optical tweezers (left) shows two-state unfolding but refolds via an intermediate, seen in both FECs (middle) and constant-force trajectories (right). Reprinted from Ref. [9], with permission from AAAS. (F) FECs of SOD1 monomers show multiple intermediates for native folding and unfolding (top left), reflecting complex pathways (bottom, black arrows), but also several misfolded states (top right) that branch off the native pathways at specific locations (bottom, red arrows). Reprinted from Ref. [14].

which the force was jumped back to a low value to allow refolding to take place, before

repeating the cycle. The advantage of measurements like these is that they allow the lifetimes for unfolding/refolding to be measured directly, giving access to the microscopic rates as a function of force.

The approaches illustrated in these two examples have been applied to study unfolding of a wide range of proteins using AFM in the last two decades, as described in previous reviews [25–27]. Tandem repeats are frequently used in order to generate a characteristic fingerprint for ensuring that only a single tether is being measured at one time; the monomer units used in the tandem repeat typically unfold the same as they do when isolated [28]. In cases where the protein being studied is not naturally a tandem repeat, the protein of interest is usually inserted between repeats of a reference protein, which act effectively as "handles" for applying force. Many of the proteins studied by AFM have been mechanically rigid, unfolding at forces over 50 pN. AFM is particularly well-suited for such proteins because it can typically apply higher forces than optical or magnetic tweezers, owing to the relatively high stiffness of the force probe. Most AFM studies have focused on unfolding rather than refolding, because the lower force sensitivity and stability of AFMs pose a challenge for observing refolding at the low forces where it often occurs. However, recent improvements in stability and force sensitivity [29,30] now allow refolding to be observed by AFM under equilibrium or near-equilibrium conditions [31,32], where the protein hops repeatedly back and forth between folded and unfolded states. Such measurements are illustrated in Fig. 1D for the designed protein α_3D [32].

Complementary to SMFS studies based on AFM, studies using optical tweezers have primarily focused on more mechanically labile proteins, making use of the lower stiffness and higher force resolution of the tweezers to study both unfolding and refolding at forces typically below 50 pN. A notable early study of protein folding using optical tweezers showcased RNaseH, whose folding has been investigated extensively by ensemble methods [33]. Holding onto the protein with double-stranded DNA (rather than polyprotein) handles (Fig. 1E, left), Cecconi *et al*. [9] showed that whereas unfolding was two-state (Fig. 1E, middle), refolding occurred hierarchically through a molten-globule-like intermediate state that could not be directly observed in ensemble measurements but was seen in force-clamp measurements, where the molecule hopped back and forth at ~5–6 pN between unfolded and intermediate states before folding into the native state (Fig. 1E, right). This ability to detect transient intermediates that might otherwise escape notice is a key feature of SMFS, allowing complex folding pathways to be mapped. For example, studying calmodulin, Stigler *et al*. identified several intermediates on two distinct native folding pathways as well as an off-pathway non-native state [11], quantifying the populations folding through each pathway. Xi *et al.* showed that a coiled-coil leucine zipper, expected to fold efficiently into the native state, sometimes misfolded with an incorrect, staggered helix registry [34]. Studying superoxide dismutase 1 (SOD1), a β-barrel protein generally considered a two-state folder, Sen Mojumdar *et al*. [14] in fact observed unfolding and refolding through multiple pathways, with intermediates corresponding to formation of individual β-strands in the native structure (Fig 1F). Similar results were found by Schönfelder *et al.* in AFM measurements of the unfolding of cold shock protein B, classically considered a two-state folder, revealing heterogeneous unfolding with multiple intermediates and highlighting the complexity underlying the ensemble behavior [35].

Magnetic tweezers have also been used to study both unfolding and refolding of single protein molecules. A particular advantage of magnetic tweezers over other techniques is that they are more easily able to monitor the folding over extended periods, up to many hours or even

days, owing to very high force stability and low drift. As a result, hopping between states near equilibrium can be measured even for molecules with very slow rates, as seen in work revisiting folding of the titin I27 domain, studied previously by AFM: Chen *et al*. [36] found two-state folding of I27 with rates near equilibrium on the order of 10^{-3} s^{-1}. A similar result was seen in work by Löf *et al*. on the filamin domain ddFLN4 [37], with rates on the order of 10^{-3} s^{-1}, although in this case a short-lived intermediate state was present in the folding trajectories. The latter study also illustrated the ability to measure multiple molecules in parallel, a feature that is much easier to implement with magnetic tweezers than other force probes.

Recent years have seen SMFS studies of small proteins extended to special cases such as proteins that form knots, an unusual but distinctive topological feature in ~1% of solved protein structures [38]. Knots pose interesting conceptual challenges because it is not always obvious how they can form spontaneously, although several cases of spontaneous knotting have been found [39,40]. SMFS studies of the protein UCH-L1 (ubiquitin C-terminal hydrolase isoenzyme L1), which forms a three-twist knot (formally classified as a 5_2 knot), observed many intermediates in the folding and explored how the presence of a knot in the unfolded state affected folding kinetics [40]. By pulling on different parts of the protein to generate unfolded states with varying knot status—a fully-formed 5_2 knot, a simpler trefoil (3_1) knot, or unknotted—Steigler *et al*. showed that the presence of a knot in the unfolded state accelerated the folding substantially, by 3-fold for the trefoil knot and 10-fold for the three-twist knot, confirming that knot formation is a rate-limiting step in folding. A similar result was found by Wang & Li for *E. coli* TrmD protein with a native trefoil knot [41]. The folding of slipknots, which are not formally considered knots because they can be undone by pulling on their ends (like shoelace knots), has similarly been studied by SMFS. Measuring the protein AFV3-109 with optical tweezers, He *et al*. [42] found that unlike the true knots, this slipknotted protein folded rapidly without intermediates, suggesting that the simpler topology of the slipknot poses less of a barrier for folding.

SMFS has also proven useful for studying disease-related misfolding, discerning how non-native structures form at the single-molecule level. For example, the prion protein PrP from hamsters (an early model for mammalian prion disease) was seen by Yu *et al.* to fold natively by a two-state mechanism, while also making frequent excursions from the unfolded state into multiple short-lived misfolded states [10]. Whereas none of these misfolded states were stable in isolated PrP monomers, tandem dimers were found to fold exclusively into a non-native structure that was thermodynamically more stable than two isolated native monomers [43]. Optical tweezers measurements of SOD1 [14] not only observed multiple misfolded states, but moreover showed that misfolding was initiated from partially folded intermediates, and also identified where the misfolding branched off from the native folding pathways (Fig. 1F). SFMS studies of intrinsically disordered proteins (IDPs) related to misfolding diseases such as Aβ peptide and α-synuclein have shown that they can form a wide variety of transient structures with different mechanical stability, even within monomers [44–47]. Such studies have also explored how disordered monomers interact to generate misfolded oligomers [47–49]. The transience and heterogeneity of IDP structures makes them challenging to characterize experimentally, however, motivating the integration of SMFS measurements with computational simulations to extend the reach of both experiment and modeling, as done for Aβ [50] and α-synuclein [51].

SMFS measurements like those described above all share some key characteristics. First, the unfolding forces are determined not so much by the thermodynamic stability of the proteins as

by their mechanical rigidity. Proteins that are rigid and deform only slightly under tension have high unfolding forces and unfold in a brittle manner, often without intermediates. In contrast, proteins that are mechanically compliant (deforming more before unfolding) tend to unfold at lower forces, because the energy barrier for unfolding is reduced more rapidly by the application of force; they also unfold more often through intermediates. These differences are quantified through the properties of the energy landscape, specifically the extension change from the folded state to the energy barrier, $\Delta x^{\ddagger}$; β-rich proteins tend to be more rigid than helical proteins and hence have smaller $\Delta x^{\ddagger}$ [52], because of their denser network of long-range interactions. Conversely, molten globule states, which lack significant tertiary structure, can be distinguished via their larger $\Delta x^{\ddagger}$ even when there is no detectable change in ΔL_c compared to the native state [53]. Second, the unfolded state in SMFS is quite different from that found when using other denaturants, because it is fully extended (low entropy) and has no residual structure, whereas other denaturants lead to random coils that may retain varying levels of residual structure [54]. As a result, the free energy change between folded and unfolded states measured in SMFS assays is generally noticeably larger than that found in other ways; even when corrected for the entropic energy needed to stretch out the unfolded state, it may still remain larger if other methods do not remove all residual structure. Third, because force acts at a local level, rather than globally like other denaturants, the response of the protein depends in part on where the force is applied: pulling from different directions may lead to quite different results, as seen both experimentally [55] and computationally [56]. For example, the unfolding force for GFP varied by ~5–6 fold when pulled from five different directions [57]. Proteins are more rigid under the influence of forces applied in a shearing geometry than an unzipping geometry, as demonstrated in work on the src SH3 domain [58]. As a result, one does not necessarily expect to recapitulate the results seen in studies using other methods to perturb the structure, because the pulling axis may select for different barriers and pathways. This effect was illustrated in elegant experiments by Guinn *et al.* unfolding various mutants of the src SH3 domain with both force and chemical denaturant at the same time [59], which showed that the zero-force pathway and barrier was the same as that for the unzipping geometry but differed from two parallel pathways observed for the shearing geometry.

Measuring energy landscapes and transition paths

SMFS is particularly useful for characterizing the energy landscapes underlying protein folding [23]. The position and height of the barrier energy with respect to the folded and unfolded states can be found by fitting the shape of the unfolding and refolding force distributions, respectively [60,61]; it can also be found from the force-dependence of the unfolding and refolding rates [62,63]. Alternatively, the position of the barrier can be found from the loading-rate-dependence of the most probable force [64]. Illustrating this type of analysis in Fig. 2A, application to the native folding and unfolding of PrP revealed that PrP is very compliant, with a large $\Delta x^{\ddagger}$ from the native state [65]. Independently, the free-energy changes between folded, unfolded, and intermediate states can be found either from FECs [66] or from hopping trajectories [11,67] and they can be combined with the information about the barriers to build an outline of the energy landscape. Such an approach was used, for example, by Gao *et al*. to reconstruct the key features of the energy landscape for SNARE zippering [12], by Choi *et al*. for the membrane protein GlpG [68], and by Yu *et al*. for misfolding of PrP dimers (Fig. 2B) [43].

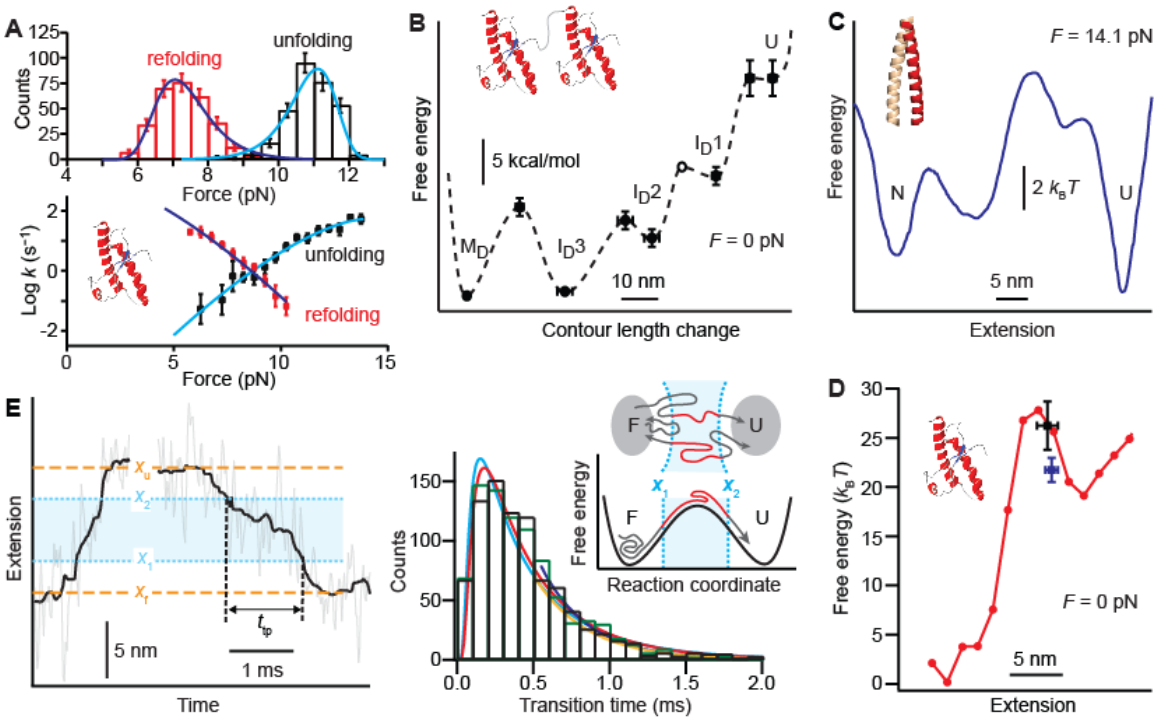

Figure 2: Energy landscapes and transition paths. (A) The distribution of unfolding (top, black) and refolding (top, red) forces can be fit to theories of diffusive barrier crossing to characterize the barrier location and height and the crossing rate at zero force. The force-dependent rates for unfolding (bottom, black) and refolding (bottom, red) can be fit to similar theories to evaluate the same parameters. Illustrated here for native PrP folding, adapted from Ref. [65]. (B) Reconstruction of key landscape features (potential wells and barriers) for misfolding of PrP dimers, which proceeds through multiple intermediate states (I_D1–I_D3) before reaching the misfolded dimer (M_D). Reprinted from Ref. [43]. (C) Full energy profile for folding of GCN4 leucine zipper reconstructed from equilibrium hopping trajectories, showing two intermediate state wells. Reprinted with permission from Ref. [74]. (D) Full energy profile for native folding of PrP, reconstructed from FECs. Reprinted from Ref. [65]. (E) Transition paths involve the part of the folding trajectories that traverse the barrier region (inset, right, shaded blue). Individual transitions for PrP dimer misfolding (left) show durations in the ms range that can be fit by theories of diffusive barrier crossing (right) to determine the diffusion coefficient. Reprinted from Ref. [17].

Going beyond characterizing key points on the landscape, the full profile including the shape of the wells and barriers can be reconstructed from both FECs and hopping trajectories, by applying any of several complementary approaches [67,69–73]. However, it is technically challenging to do so because of the high data quality needed and the confounding effects from the handles used to apply tension. Gebhardt *et al*. [74] reconstructed the full energy landscape for the GCN4 leucine zipper (Fig. 2C), based on the equilibrium occupancy at every point along the reaction coordinate in hopping trajectories. A different approach based on fluctuation theorems was used by Yu *et al*. [65] to reconstruct the native folding landscape for PrP from FECs (Fig. 2D). Edwards *et al.* used multiple methods to reconstruct the full free-energy profile for the protein α_3D, interestingly finding that the barrier was sufficiently small that it was dominated by entropic effects from stretching of the polypeptide chain [32]. Such landscape reconstructions have allowed quantitative experimental tests of the energy landscape theory of protein folding, showing for example that the occupancy at every point along the folding trajectory matches the expectation for diffusive motion over the measured landscape [75] and that folding/unfolding rates can be predicted from measured landscapes using Kramers' theory for diffusive barrier-

crossing kinetics [65]. Landscape reconstructions have also been used to estimate the diffusion coefficient, *D*, which describes the microscopic kinetics of the search over the landscape and reflects the effects of internal friction and landscape roughness. Studies of PrP, for example, found that diffusion was 1000-fold faster for native folding [65] than for misfolding [43], indicating that misfolding involved significantly more internal friction. However, kinetic artifacts from the force probe can lead to unreliable results in some measurement regimes [76–78], which might explain the extremely slow diffusion reported in some studies [79,80].

Fast-folding proteins that have marginal or non-existent barriers at zero force, undergoing so-called downhill folding, present some particular challenges for SMFS measurements. Very high rates may lead to difficulty resolving transitions directly, in which case correlation analyses may be applied to measure rates, as done for the villin headpiece in optical tweezers measurements [81]. The rates observed by SMFS may also differ noticeably from those measured by methods not subject to similar kinetic artifacts, such as temperature jumps. In the case of villin, the rates observed by SMFS were roughly two-fold slower than those reported from temperature jumps [81]; AFM measurements observing two-state hopping of the fast-folding protein gpW found rates that were a few orders of magnitude slower than reported by temperature jump [82]. Another factor that can affect SMFS measurements of these proteins is that the applied tension can induce an entropic barrier from stretching the handles that may be larger than the intrinsic barrier. This effect can obscure the shape of the intrinsic barrier in full landscape reconstructions, as seen in work on α_3D finding that force-dependent rate maps provided a more reliable reflection of the intrinsic barrier position than did full reconstructions [32], and may also contribute to a reduction in the rates observed by SMFS.

Recently, SMFS studies have begun to characterize the transition states within the energy barriers more directly by observing the transition paths—the very brief parts of the folding/unfolding trajectories during which the molecule crosses the barrier (Fig. 2E, inset). Transition paths are inherently a single-molecule property and cannot be observed in ensemble measurements because the folding of different molecules is not synchronized, but they are typically very brief—on the μs scale [83]—making them difficult to observe. An estimate of the average transition-path time for PrP based on landscape reconstructions and measurements of *D* found 2 μs for both folding and unfolding [65], slightly too fast for reliable direct measurements given contemporary time resolution in SMFS measurements. Measurements of transition paths have instead focused on proteins with relatively rough landscapes, which should have longer transition-path times. Indeed, Neupane *et al*. found that transition paths for misfolded PrP (Fig. 2E) had an average duration of 0.5 ms; fitting the distribution of path times returned a value for *D* consistent with the earlier result from the landscape analysis [17]. Examining transition paths in a designed protein with a Rossmann fold, Mehlich *et al*. [84] found an average transition-path time in the ms range. Here the transition paths were slowed down by shallow potential wells in the barrier region that were visible in landscape reconstructions. Transition paths are particularly sensitive to kinetic artifacts from SMFS instrumentation [85], and as a result it remains difficult to make reliable measurements.

Folding of multi-domain proteins

Most proteins (~65–80% in eukaryotes and ~45–65% in prokaryotes [86]) are larger than the proteins discussed so far, containing two or more domains. The folding of multi-domain proteins is more challenging to study than that of smaller proteins, owing to several factors. Many large

proteins have trouble folding spontaneously, needing assistance from chaperones or the ribosome to find their native structure and avoid misfolding and/or aggregation. Furthermore, for those proteins that do fold spontaneously, the complexity of the structures can make it difficult to interpret the observations, when many different assembly pathways are possible. In this regard, the use of single-molecule methods like SMFS, which can resolve the sequence of steps on complex folding pathways with multiple branches, offers an attractive means of characterizing multi-domain folding.

One of the earliest SMFS studies of multi-domain proteins, done with AFM, examined the Lyme disease antigen OspA, which contains a single-layer β-sheet connecting N- and C-terminal domains. Hertadi *et al*. [87] found that the two domains unfolded separately, but ~50% of each domain appeared to unfold at forces below the detection threshold, leading to short ΔL_c. Both domains of OspA were able to refold spontaneously. In contrast, firefly luciferase is generally unable to refold without assistance, a result from ensemble studies that Scholl *et al*. [88] recapitulated using AFM: after unfolding luciferase completely, which occurred via three steps, it could not be refolded. However, when only the C-terminal domain was unfolded, luciferase did refold spontaneously, suggesting that refolding of the full protein was prevented by stable non-native interactions between the domains. Looking at human γD-crystallin, a two-domain protein that aggregates to form cataracts, Garcia-Manyes *et al*. [89] showed that whereas isolated molecules unfolded from the native state, with the two domains unfolding independently similarly to how they unfolded when measured on their own, two adjacent γD-crystallins sometimes formed a domain-swapped misfolded dimer, which might be an early intermediate for aggregation. In contrast to the relatively weak inter-domain interactions in γD-crystallin, Kotamarthi *et al*. found that the two domains of leucine binding protein interacted sufficiently strongly that ~40% of FECs showed two-state unfolding [90]. Although the remainder showed an intermediate, this intermediate had highly variable ΔL_c (from ~20–90 nm) that did not match independent unfolding of each domain, suggesting that the domain interface is stronger than the individual domains, leading to heterogeneous unfolding of variable fractions of the domains. Inter-domain interactions were also explored by Li *et al*., in the two-domain yeast phosphoglycerate kinase [91]. Here the domains unfolded independently, but whereas the C-terminal domain showed two-state unfolding, both in the context of the full protein and on its own, the unfolding of the N-terminal domain was modulated by the interface: an intermediate that occurred in ~4/5 of FECs for the N-terminal domain on its own was reduced to only ~1/8 in the full protein.

Multiple-domain protein folding has also been explored using optical and magnetic tweezers. Using magnetic tweezers, Löf *et al*. examined the dynamics of dimers of von Willebrand factor (VWF), a ~250-kDa protein involved in blood clotting that forms large multimers and is activated by force-induced conformational changes [37]. VWF dimers were found to unzip and rezip at their C-terminal end reversibly at forces of only ~1 pN, passing through intermediates corresponding to (un)zippering of pairs of C-terminal domains. The low force for this unzippering, comparable to physiological forces from shear flow, suggests that it is one of the first conformational changes in the response of VWF to blood flow from wounds. Studying the *E. coli* chaperone DnaK from the Hsp70 family with optical tweezers, Rief and colleagues examined each of its two domains separately. The nucleotide binding domain (NBD) was found to fold through multiple intermediates, one of which acted as an obligate checkpoint ensuring native folding; inserting this folding nucleus into a yeast NBD variant that did not fold without assistance from a chaperone recovered robust spontaneous folding [92]. Binding of the

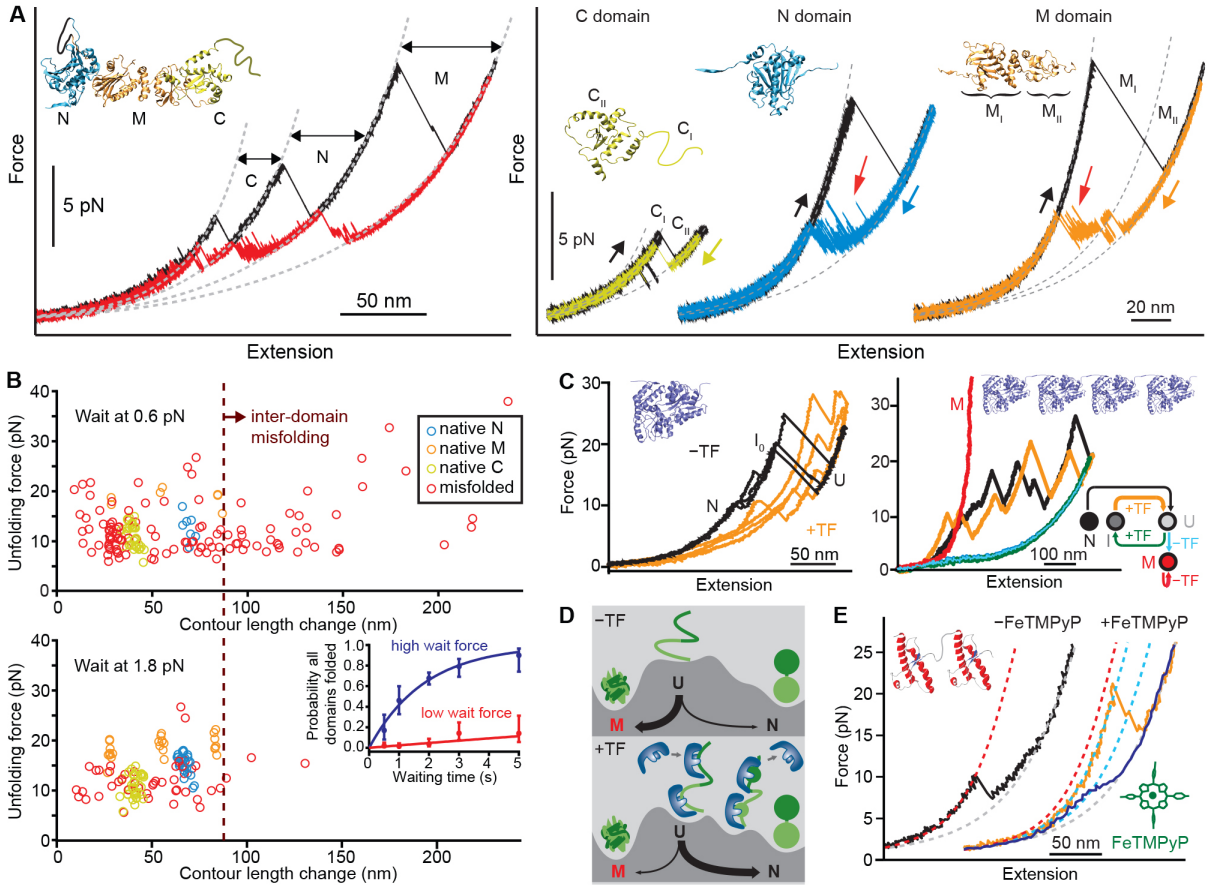

Figure 3: Multi-domain protein folding. (A) Left: Hsp90 shows sequential unfolding (black) and refolding (red) of each of its three domains in FECs. Right: FECs of each domain individually show transient intra-domain misfolding in N and M (red arrows). (B) Full-length Hsp90 shows substantial inter-domain misfolding (large contour lengths) when refolding at very low force (top), but applying a small amount of additional tension reduces the probability of misfolding dramatically (bottom). Panels A and B reprinted with permission from Ref. [94]. (C) Left: FECs of MBP monomer in absence of trigger factor (black) show native unfolding through one intermediate, but addition of trigger factor stabilizes many intermediate states (orange). Right: FECs of MBP tetramer initially show native unfolding (black), but later pulls show stable misfolding in the absence of trigger factor (red), in contrast to stabilization of native-like intermediates (orange) in the presence of trigger factor. (D) Trigger factors reshapes the energy landscape, favoring formation of intermediates on the native pathway. Panels C and D adapted with permission from Ref. [97]. (E) Whereas PrP dimers always misfolds when measured in isolation (left), adding Fe-TMPyP abolishes the thermodynamically stable misfolded state and produces marginally stable misfolding (right, blue) or enables native folding (right, orange). Dashed lines: WLC fits (red: misfolded dimer, cyan: native monomers, grey: unfolded). Adapted from Ref. [104].

nucleotide switched the relative stability of the two lobes comprising the NBD [93]. Turning to the substrate binding domain (SBD), it unfolded through an intermediate in which the interface between the two subdomains of SBD was broken, similar to the opening of the SBD to accommodate substrate binding that is driven by ATP binding, suggesting that the functional flexibility of DnaK is encoded in its mechanical properties [40]. Looking at Hsp90, a chaperone with three domains (Fig. 3A, insets), Jahn *et al*. [94] showed that the domains unfolded and refolded sequentially and independently, with multiple intermediates for each domain except the two-state C-terminal domain (Fig. 3A). The N-terminal and middle domains both exhibited misfolding when studied on their own, which slowed down their folding, but assembling all three

domains into the full protein introduced additional misfolding between the domains that further reduced the folding rate. Interestingly, applying a small amount of force during refolding to stretch out the unfolded chain led to a significant reduction in inter-domain misfolding by restricting the ensemble of pathways accessible to the protein (Fig. 3B). Such a restriction is conceptually similar to what occurs during co-translational folding on the ribosome owing to vectorial synthesis, suggesting that Hsp90 may avoid misfolding when folding co-translationally. Because of the propensity for multi-domain proteins to misfold, SMFS studies of multi-domain protein folding have also been done with chaperones present or in the context of co-translational folding, as described below.

Probing the effects of chaperones on folding

Although many proteins fold spontaneously into their native structure, others require assistance to find the native state. Larger proteins, especially, often need help to avoid getting stuck in kinetic traps that prevent native folding and may lead to misfolding and aggregation. This assistance is often provided by the class of proteins known as molecular chaperones [95]. By comparing the detailed series of steps followed by individual molecules as they fold in the presence and absence of chaperones, SMFS provides a powerful tool for probing how chaperones act to guide proteins towards the native state. As a result, SMFS is increasingly being used to understand chaperone-mediated folding.

In a series of studies, Tans and colleagues explored how different chaperones interact with maltose binding protein (MBP), which folds spontaneously as a monomer but can form stable aggregates as a tandem tetramer. Looking first at the bacterial chaperone SecB, Bechtluft *et al.* used optical tweezers to show that it bound to MBP only in extended or molten-globule-like states and prevented stable tertiary contacts from forming; notably, it was unable to bind to the rate-limiting intermediate state for MBP folding [96]. In contrast, a bacterial trigger factor chaperone was found to interact with partially folded intermediates of MBP (Fig. 3C), stabilizing them to guide formation of the native state while preventing the formation of long-range contacts between domains that led to stable aggregates (Fig. 3D) [97]. Turning to the small heat shock proteins, classified as holdase chaperones, Ungelenk *et al.* [98] found that Hsp42 also promoted native folding by suppressing stable non-native interactions, but this time binding was primarily to the near-native core structure of MBP. Interestingly, the interactions with Hsp42 decreased the unfolding force of the core structure, suggesting that Hsp42 binding made the protein more compliant. Moayed *et al.* [99] showed that Hsp33 not only stabilized unfolded states as expected for a holdase, inhibiting both native folding as well as aggregation, but it also bound partially folded structures, suggesting that it can act both early and late in the folding cycle. Looking at foldase chaperones, Mashaghi *et al.* [100] showed that the bacterial Hsp70 homolog DnaK binds to states throughout the folding pathway, from near-native and partially folded states to the unfolded state, indicating a broader range of interactions. Using DnaK mutants, the stabilization of folded structures was shown to occur in the ADP-bound state, being driven primarily by the DnaK lid rather than the binding groove. Finally, Avellaneda *et al*. [101] examined the action of the disaggregase chaperone ClpB from the Hsp100 family, using optical tweezers combined with single-molecule fluorescence to show that polypeptide loops from collapsed but unfolded MBP were translocated processively. Translocation occurred on both arms of the loop, but switched to only one arm when impeded by structures that required unfolding, allowing the force applied by ClpB to increase. Notably, the MBP substrate refolded after extrusion from the exit channel, similar to co-translational folding, allowing it to be activated for re-use.

Other work has studied the interactions of some of the same chaperones with different proteins. For example, Haldar *et al.* used magnetic tweezers to probe the effects of trigger factor on the refolding of a tandem octamer of the B1 domain of protein L [102]. The octamer refolded spontaneously without trigger factor at low force, but at intermediate force—where spontaneous refolding was less likely—trigger factor greatly increased the probability of refolding, suggesting that trigger factor acts as a mechanical foldase. The interaction with the refolding protein involved partially folded rather than fully folded states, and weakened at higher forces. Using AFM, Perales-Calvo *et al.* examined the effects of DnaK and its co-chaperones DnaJ and GrpE [103]. DnaJ on its own was actually found to reduce refolding of polyubiquitin and titin I27 domains when the force was jumped to a low value after unfolding the protein at high force, by binding to the unfolded state in the case of ubiquitin and partially folded states in the case of I27. ADP-bound DnaK on its own also reduced refolding, but appeared only to interact with collapsed intermediates, not the unfolded or native states. Reconstituting the full DnaKJE system, however, restored the folding of each protein, even increasing the folding rate. Moreover, DnaKJE induced refolding of tandem-repeat titin Z1 domains, which otherwise were not able to refold.

In addition to the molecular chaperones found in cells, small molecules can also help inhibit misfolding and aggregation as "pharmacological chaperones," and such molecules have therefore been sought as potential therapeutics for protein misfolding diseases. Single-molecule studies have recently begun to decipher the mechanism of action of some of these pharmacological chaperones. Focusing on PrP, whose misfolding causes prion diseases, Gupta *et al*. [104] studied the effects of the iron tetrapyrrole Fe-TMPyP, an anti-prion agent active against multiple prion strains that was proposed to stabilize the native fold. Optical tweezers measurements showed that in addition to stabilizing the native fold thermodynamically, Fe-TMPyP made the protein more rigid. Moreover, it bound to unfolded PrP, such that whereas monomer folding was decreased, the formation of stable misfolded states in dimers was abolished, providing an opportunity for native folding to occur in a process remarkably similar to the mechanism of cellular chaperones (Fig. 3E). Petrosyan *et al.* [16] examined the effects of a different anti-prion agent, pentosan polysulfate (PPS), with very different chemical properties: a negatively charged linear polymer rather than a positively charged planar ring. PPS bound not only to unfolded PrP but also to a wide variety of partially folded states, preventing aggregation, and it significantly rigidified the structures it bound just as for Fe-TMPyP. The similarities in the effects of these very different anti-prion agents suggest that partially or fully unfolded states play a central role in prion propagation and can be targeted to inhibit disease.

We conclude this section with work by Zhang and colleagues using optical tweezers to study the assembly of the SNARE complex that mediates membrane fusion. Composed of three proteins, the quaternary structure of the SNARE complex—a four-helix bundle—is assembled with the help of several accessory proteins. In the absence of accessory proteins, Gao *et al.* [12] used cross-linked proteins to observe folding of the complex through a sequential zippering process, which started with slow formation of the N-terminal domain, passed through a half-zippered intermediate stabilized by moderate force mimicking the tension arising from opposing membranes before fusion, and finished with rapid zippering against load of the remainder of the protein, thereby generating the free energy needed for membrane fusion. Adding in α-SNAP, which enhances SNARE assembly and membrane fusion, Ma *et al*. [105] found that the N-terminal domain was unaffected, but the C-terminal assembly (trans-SNARE complex) was stabilized by α-SNAP binding whereas the linker domain between the two was destabilized.

Munc13-1 and Munc18-1, two other proteins involved in SNARE assembly, were found to chaperone SNARE assembly directly, first generating a weakly bound complex with 2 of the 3 SNARE proteins (syntaxin 1 and VAMP2), followed by recruitment of the third SNARE protein (SNAP-25) via binding to Munc13-1, leading to the ternary SNARE complex [106,107].

SMFS studies of folding on the ribosome

All proteins are synthesized on the ribosome, hence interactions with the ribosome may affect how proteins fold *in vivo*. Indeed, given that protein synthesis is relatively slow (on the order of 0.1 s per residue [108]) compared to typical folding times of μs–s, proteins will generally start to fold even while translation is still under way. The nascent polypeptide chain emerges from the ribosome through a ~80-Å long exit tunnel, whose small diameter severely constrains the protein conformation, allowing only small structures to form [109]. Once enough of the chain exits the tunnel, however, larger structures can begin to form, beginning from the N-terminus and proceeding in tandem with the elongation of the chain towards the C-terminus. The directional synthesis of the polypeptide chain may therefore guide the folding to different pathways than seen when monitoring the folding of the complete chain after synthesis, as done in most folding studies *in vitro*.

In the last decade, folding of proteins as they emerge from the ribosome during translation has begun to be studied directly with SMFS, beginning with the elegant experiment by Kaiser *et al.* [110] in which force was applied to the N-terminal end of a nascent chain exiting a stalled ribosome using optical tweezers, as illustrated in Fig. 4A. Studying the T4 lysozyme and using a 41-residue C-terminal linker that allowed the entire protein to exit the ribosome, unfolding was found to be very similar on and off the ribosome, but refolding was much less frequent on the ribosome and it did not show well-defined transitions in the FECs (Fig. 4B). Refolding was also much slower on the ribosome, modulated by electrostatic interactions with the rRNA. When translation was stalled with only 70–90% of the protein outside the exit tunnel, the protein remained unstructured on the ribosome but formed heterogeneous misfolded structures off the ribosome (Fig. 4C), suggesting the ribosome helps prevent misfolding of the nascent chain. However, studies of the effects of the ribosome on the folding of the src SH3 domain by Guinn et al. [111] using a combination of chemo-mechanical unfolding and mutational phi-value analysis found that the folding mechanism was the same on and off the ribosome, suggesting that the effects of the ribosome are more significant for proteins that have folding intermediates or multiple domains.

The co-translational folding of a multi-domain protein was explored in a series of studies by Kaiser and colleagues examining elongation factor G (EF-G), a five-domain protein. Liu *et al.* found that although the N-terminal G domain of elongation factor G (EF-G) folds on its own both on and off the ribosome, the ribosome slowed down refolding and destabilized both the native state and misfolded states [112]. Extending the transcript further to include both the G domain and some or all of its neighbour, domain II, the G domain was seen to fold before domain II, but interference between the domains slowed down the folding, presumably by generating off-pathway structures; this effect was even stronger off the ribosome, reflecting alleviation of misfolding owing to the ribosome [113]. A folded G domain was required to fold domain II because of stabilization of the latter by the domain interface, but—surprisingly—an unfolded domain II also destabilized the native G domain (Fig. 4D). This destabilization was observed both on and off the ribosome, suggesting that the chaperoning ability of the ribosome is

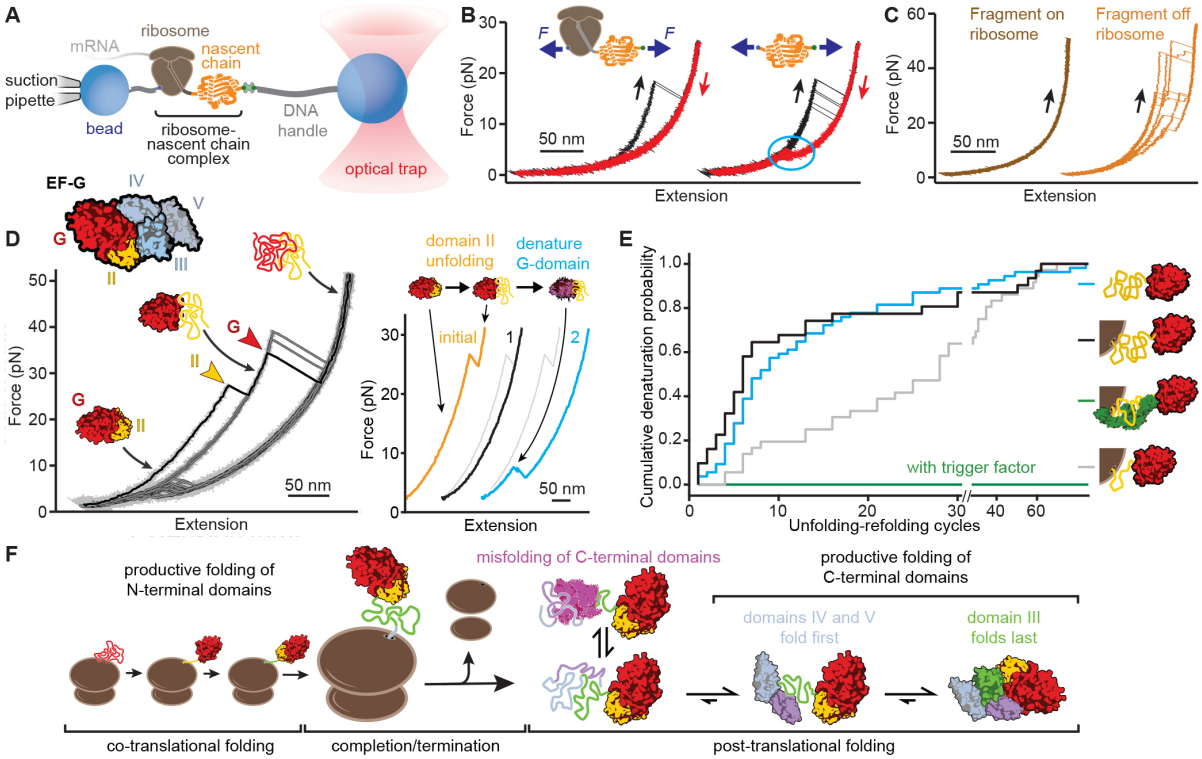

Figure 4: Folding on the ribosome. (A) Force is applied to the nascent polypeptide chain extruded from a stalled ribosome by optical tweezers. (B) FECs of T4 lysozyme on (left) and off (right) the ribosome show similar two-state unfolding (black), but refolding (red) was less frequent and did not show well-defined transitions. (C) Incompletely translated T4 lysozyme remained unstructured on the ribosome (left) but formed a variety of misfolded structure off the ribosome (right). Panels A–C reprinted from Ref. [110], with permission from AAAS. (D) Left: nascent EF-G stalled after translation of domains G and II unfolds domain II (orange arrow) at lower force than domain G (red arrow). Inset: five domains of EF-G. Right: unfolding domain II while keeping domain G folded (orange) led after some period of time with no further change (black) to misfolding of domain G (cyan). Grey curves show expectation if domain II had refolded. (E) The rate of domain-G misfolding induced by unfolded domain II was similar on (black) and off (cyan) the ribosome, slower for partially translated domain II, and misfolding was abolished in the presence of trigger factor chaperone (green). Panels D–E reprinted from Ref. [113], with permission from Elsevier. (F) Domains G and II of EF-G fold co-translationally, whereas domains III–V fold post-translationally, in competition with misfolding of the C-terminal domains. Reprinted with permission from Ref. [114].

limited, but it was prevented entirely by the addition of the trigger factor chaperone (Fig. 4E). Trigger factor also prevented the inter-domain misfolding that frustrated folding of domain G. Extending the transcript to include domains III, IV, and V of EF-G, Liu *et al.* [114] found that domain III did not fold reliably upon its emergence, but instead it required interactions with its natively folded C-terminal neighbors, leading to a post-translational folding mechanism for the remainder of the protein (Fig. 4F). Unfolded domain III interfered with folding of the C-terminal domains through formation of misfolded conformers. Intriguingly, these studies showed that vectorial synthesis does not always lead to co-translational folding in multi-domain proteins, owing to differences in domain stabilities and interactions.

Where most SMFS work on co-translational folding has examined folding of structures outside of the peptide exit channel, some small protein structures can fold within the exit tunnel. Recent work by Wruck *et al*. [115] examined the folding of ADR1a, a 29-residue yeast zinc-finger domain that can fold inside the ribosome. Using combined single-molecule force and

fluorescence spectroscopy, ADR1a folding was found to be faster and more stable inside the ribosome than outside. This effect appeared to arise not from confinement stabilization, but rather from electrostatic interactions with the highly charged exit tunnel that lowered the energy barrier for folding and stabilized the native state.

SMFS of membrane proteins

Membrane proteins are an important class representing over half of all drug targets [116]. However, relatively few membrane protein structures have been solved, and their folding has been much less studied than that of soluble proteins owing to the technical challenge presented by the involvement of a membrane. SMFS provides a convenient way to study membrane protein folding through the use of AFM, which can probe proteins embedded in a membrane deposited on a flat surface like mica. The imaging modality of AFM also allows the location and arrangement of the proteins in the membranes to be determined before and after unfolding. Early work by Oesterhelt *et al*. studying bacteriorhodopsin, an α-helical bundle, imaged protein molecules packed into 2D lattices in supported membranes, then unfolded single molecules while pulling them out of the membrane [117]. Helices were found to unfold most often in pairs, although some unfolded singly, reflecting lower cooperativity than typically seen in globular proteins, a feature that was later confirmed to be characteristic of membrane proteins. Another way in which membrane proteins were found to differ from globular proteins is that pulling on different termini of the protein produced different patterns in the unfolding, as seen for example with bacteriorhodopsin [118] and β_2-adrenergic receptor [119], allowing the free-energy profile to be characterized from each end of the protein. Concerns that the use of supported membranes might introduce artifacts into the AFM results were subsequently laid to rest by measurements of bacteriorhodopsin unfolding using membranes spanning a cavity, showing no significant difference for the protein stability and unfolding pathways compared to measurements on supported membranes [120].

Refolding of membrane proteins has also been studied, by unfolding the protein only partially while leaving the last structural element(s) embedded in the membrane, before relaxing the force to refold the protein, allowing repeated interrogation of the structure. Measurements on bacteriorhodopsin [121] and the Na^+/H^+ antiporter NhaA [122] showed spontaneous refolding of the helical bundles into native-like states, even against tension. Turning to the other major structural class of membrane proteins, β-barrels, measurements of OmpG [123,124], OmpA [125], and FhuA [126] found similar unfolding mechanisms involving sequential unfolding of β-hairpins formed from neighboring antiparallel β-strands. Smaller β-barrels like OmpA and OmpG were able to refold spontaneously [124,125], but larger ones like FhuA required assistance from chaperones for native refolding [126]. Investigating more closely how translocon complexes and insertases help insertion and folding of proteins in membranes, Serdiuk *et al.* [127] found that unfolded LacY, a lactose permease, could insert structural components into a membrane on its own, but this insertion was dominated by misfolding; addition of the YidC insertase, however, stabilized the unfolded state and resolved the misfolding. Surprisingly, the order in which helices were inserted was found to be random, but different inserting orders still led to the native structure with its characteristic signature for unfolding. In contrast, the SecYEG translocon inserted segments in a well-defined order [128].

Recent improvements in AFM cantilevers resulting in greater force stability, lower drift, and higher time resolution [29,30] have yielded new insights from revisiting proteins like

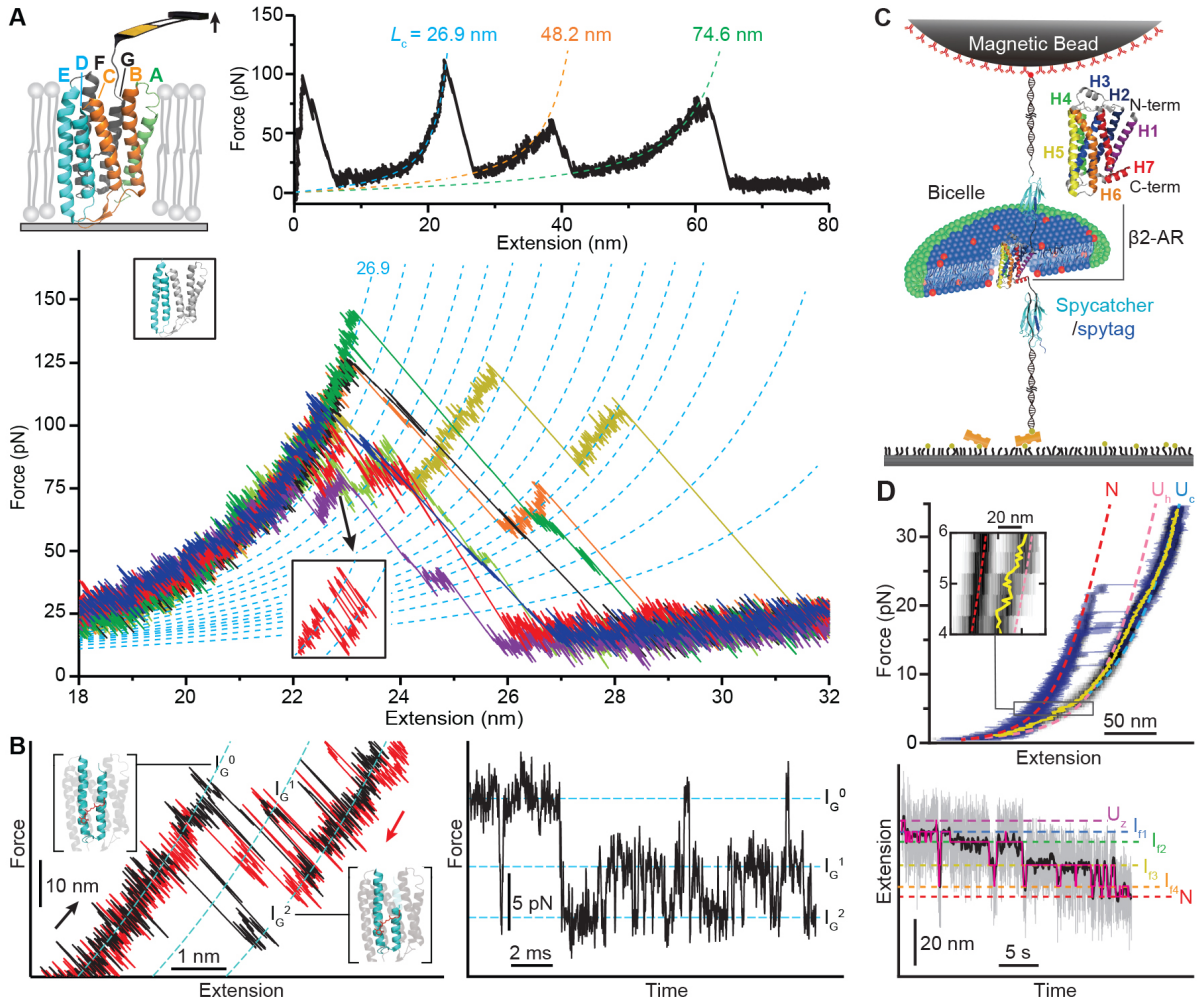

Figure 5: Membrane protein folding. (A) Unfolding of bacteriorhodopsin from a supported membrane by AFM (top left) shows rips in FECs (top right) reflecting the removal of helices in pairs (E and D: blue, B and C: orange, A: green; helices F and G not resolved). Inspecting the unfolding transition for helices E and D at high resolution (bottom) reveals a profusion of intermediate states, some showing quasi-equilibrium hopping in and out of the membrane (lower inset). Reprinted from Ref. [13], with permission from AAAS. (B) Initial unfolding of helix G of bacteriorhodopsin into a retinal-stabilized intermediate I_G^2 is seen in unfolding (black) and refolding (red) FECs (left) and in equilibrium hopping trajectories (right). Reprinted with permission from Ref. [131], copyright (2020) by the American Physical Society. (C) Magnetic tweezers assay for measuring membrane protein unfolding and refolding in a bicelle, illustrated for β_2-adrenergic receptor (β2AR). (D) FECs (top) and constant-force refolding trajectory (bottom) for β2AR. Unfolding FECs (blue) show several intermediates. Refolding trajectories (black) starting from the fully unfolded state (U_c) proceed first through an unfolded state with all helices formed but stretched in a line (U_h) to an intermediate with the helices arranged in a zigzag patter (U_z), followed by multiple intermediates (I_{f1}–I_{f4}) formed by sequential folding of segments from N to C terminus, before reaching the native state. Panels C–D reprinted from Ref. [68], with permission from AAAS.

bacteriorhodopsin. For example, Yu *et al.* [13] observed a plethora of previously undetected intermediates in the unfolding (Fig. 5A), including some involving the unfolding of as few as 3 amino acids (less than one turn of the helix), connected in complex, heterogeneous pathways. Near-equilibrium hopping between states at high force was also seen, allowing reconstruction of the free-energy profile for that portion of the landscape. Heenan *et al.* [129] reconstructed the full free-energy profile from non-equilibrium FECs, revealing variability in the local free-energy

per amino acid in different parts of bacteriorhodopsin. By labeling the protein for specific chemical attachment to the tip, rather than using non-specific tip attachment as in most AFM work, Yu *et al.* [130] characterized the poorly resolved unfolding of the first two helixes (F and G) of bacteriorhodopsin, with the protein starting in a native-like state. These measurements revealed a previously undetected intermediate that was greatly stabilized by retinal binding. Subsequent work explored the energetics of the native state by measuring the free energy change for unfolding and refolding the first eight amino acids in helix G. Multiple analyses were applied to equilibrium hopping measurements and non-equilibrium FECs of the wild-type protein (Fig. 5B) [131] as well as to a series of mutants [132], showing that SMFS provides a more relevant and accurate way to measure free-energy changes in membrane proteins than chemical denaturation, as it fully unfolds the protein and measures changes from a native membrane environment.

SMFS studies of membrane proteins have also been carried out using magnetic tweezers, embedding the proteins in free-floating bicelles or vesicles (Fig. 5C). Choi *et al*. [68] examined the refolding of two evolutionarily distant helical proteins, rhomboid protease GlpG and β_2-adrenergic receptor, after they were unfolded but still associated with the membrane. Folding occurred in the same way for both proteins: first the helices formed but remained unaligned within the membrane, then they aligned into a zigzag pattern, then the helices folded into the native-like conformation sequentially from the N-terminus to the C-terminus (Fig. 5D). Min *et al.* [133] studied a large, two-domain membrane protein, the CIC-ec1 chloride channel, which contains long transmembrane helices as well as shorter helices that only partially penetrate the membrane or are fully embedded. Unfolding of CIC-ec1 was found to proceed first through dissociation of the two domains, before each domain unfolded independently in a two-state manner, suggesting each domain was independently stable. However, refolding attempts were rarely successful, leading mostly to misfolded conformers, suggesting the possible need for assistance from a chaperone.

Lastly, we note that just as the folding of globular proteins designed *de novo* by computational methods has been studied by SMFS, so too has the folding of *de novo*-designed membrane proteins. Lu *et al.* [134] found that the protein scTMHC2, which contains four transmembrane helices, unfolded without intermediates but usually refolded through one intermediate, indicating two dominant energy barriers. Intriguingly, the energetic stability of this designed protein appeared to be higher than the stability measured for natural membrane proteins like bacteriorhodopsin and GlpG.

Outlook

Over the last 25 years, single-molecule force spectroscopy has become an important tool for studying protein folding that can yield unique insights, owing to the high precision it enables for monitoring the folding dynamics of individual protein molecules and the high level of control it offers over the way that the proteins are unfolded and refolded. Even though much has been learned from SMFS studies of protein folding, there are still many exciting opportunities for discoveries, particularly in topics that are well-suited to the capabilities of SMFS. For example, many questions remain to be answered about misfolding, not only in terms of the proteins involved in misfolding diseases, but also more generally in terms of how non-native interactions within and between domains in multi-domain proteins compete with native interactions to influence the folding. SMFS is also well-placed to provide a more physiological understanding

of protein folding by revealing how it is guided during and after translation by factors such as the ribosome and chaperones, which play key roles in the cell. Combining SMFS with live-cell imaging methods can help to probe folding and unfolding *in vivo*, as in recent work linking translocation rates into the cell nucleus measured by confocal imaging to the mechanical stability of the translocating proteins [135]. New frontiers in the microscopic biophysics of folding also remain to explored, too, with studies of transition paths still in their infancy. And of course SMFS is ideally suited to pushing the boundaries of membrane protein folding, with the ability to probe membrane proteins in native-like states under near-physiological conditions.

Looking forward, future SMFS studies will likely take increasing advantage of the opportunity to combine multiple single-molecule, ensemble, and computational methods, as highlighted in some of the work described above, further extending the ability of SMFS to probe folding in detail. Such a trend should be enabled by the increasing availability of multi-mode instrumentation. Continued technical improvements to SMFS methodologies, coupled with deployment of increasingly sophisticated commercial instrumentation, hold great promise for novel insights into protein folding from SMFS in the coming years.

Acknowledgements: This work was supported by Natural Sciences and Engineering Research Council grant number RGPIN-2018-04673. A.N. acknowledges the support of a Banting Postdoctoral Fellowship.

References

[1] T. Schwede, Protein modeling: what happened to the “protein structure gap”?, Structure. 21 (2013) 1531–1540.

[2] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature. (2021). https://doi.org/10.1038/s41586-021-03819-2.

[3] A.N. Lupas, J. Pereira, V. Alva, F. Merino, M. Coles, M.D. Hartmann, The breakthrough in protein structure prediction, Biochem. J. 478 (2021) 1885–1890.

[4] I.V. Korendovych, W.F. DeGrado, *De novo* protein design, a retrospective, Q. Rev. Biophys. 53 (2020) e3.

[5] J. Buchner, T. Kiefhaber, eds., Protein folding handbook, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2005.

[6] P.G. Wolynes, W.A. Eaton, A.R. Fersht, Chemical physics of protein folding, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 17770–17771.

[7] K.A. Dill, H.S. Chan, From Levinthal to pathways to funnels, Nat. Struct. Biol. 4 (1997) 10–19.

[8] W.J. Greenleaf, M.T. Woodside, S.M. Block, High-resolution, single-molecule measurements of biomolecular motion, Annu. Rev. Biophys. Biomol. Struct. 36 (2007) 171–190.

[9] C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Direct observation of the three-state folding of a single protein molecule, Science. 309 (2005) 2057–2060.

[10] H. Yu, X. Liu, K. Neupane, A.N. Gupta, A.M. Brigley, A. Solanki, I. Sosova, M.T. Woodside, Direct observation of multiple misfolding pathways in a single prion protein molecule, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 5283–5288.

[11] J. Stigler, F. Ziegler, A. Gieseke, J.C.M. Gebhardt, M. Rief, The complex folding network of single calmodulin molecules, Science. 334 (2011) 512–516.

[12] Y. Gao, S. Zorman, G. Gundersen, Z. Xi, L. Ma, G. Sirinakis, J.E. Rothman, Y. Zhang, Single reconstituted neuronal SNARE complexes zipper in three distinct stages, Science. 337 (2012) 1340–1343.

[13] H. Yu, M.G.W. Siewny, D.T. Edwards, A.W. Sanders, T.T. Perkins, Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins, Science. 355 (2017) 945–950.

[14] S. Sen Mojumdar, Z. N. Scholl, D.R. Dee, L. Rouleau, U. Anand, C. Garen, M.T. Woodside, Partially native intermediates mediate misfolding of SOD1 in single-molecule folding trajectories, Nat. Commun. 8 (2017) 1881.

[15] Y. Cao, H. Li, Single molecule force spectroscopy reveals a weakly populated microstate of the FnIII domains of tenascin, J. Mol. Biol. 361 (2006) 372–381.

[16] R. Petrosyan, S. Patra, N. Rezajooei, C.R. Garen, M.T. Woodside, Unfolded and intermediate states of PrP play a key role in the mechanism of action of an antiprion chaperone, Proc. Natl. Acad. Sci. U.S.A. 118 (2021) e2010213118.

[17] K. Neupane, D.A.N. Foster, D.R. Dee, H. Yu, F. Wang, M.T. Woodside, Direct observation of transition paths during the folding of proteins and nucleic acids, Science. 352 (2016) 239–242.

[18] K. Neupane, N.Q. Hoffer, M.T. Woodside, Measuring the local velocity along transition paths during the folding of single biological molecules, Phys. Rev. Lett. 121 (2018) 018102.

[19] N.Q. Hoffer, K. Neupane, A.G.T. Pyo, M.T. Woodside, Measuring the average shape of transition paths during the folding of a single biological molecule, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 8125–8130.

[20] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM, Science. 276 (1997) 1109–1112.

[21] M.S. Kellermayer, S.B. Smith, H.L. Granzier, C. Bustamante, Folding-unfolding transitions in single titin molecules characterized with laser tweezers, Science. 276 (1997) 1112–1116.

[22] L. Tskhovrebova, J. Trinick, J.A. Sleep, R.M. Simmons, Elasticity and unfolding of single molecules of the giant muscle protein titin, Nature. 387 (1997) 308–312.

[23] M.T. Woodside, S.M. Block, Reconstructing folding energy landscapes by single-molecule force spectroscopy, Annu. Rev. Biophys. 43 (2014) 19–39.

[24] J.M. Fernandez, H. Li, Force-clamp spectroscopy monitors the folding trajectory of a single protein, Science. 303 (2004) 1674–1678.

[25] A. Borgia, P.M. Williams, J. Clarke, Single-molecule studies of protein folding, Annu. Rev. Biochem. 77 (2008) 101–125.

[26] P.E. Marszalek, Y.F. Dufrêne, Stretching single polysaccharides and proteins using atomic force microscopy, Chem. Soc. Rev. 41 (2012) 3523–3534.

[27] M.L. Hughes, L. Dougan, The physics of pulling polyproteins: a review of single molecule force spectroscopy using the AFM to study protein unfolding, Rep. Prog. Phys. 79 (2016) 076601.

[28] H. Lei, C. He, C. Hu, J. Li, X. Hu, X. Hu, H. Li, Single-molecule force spectroscopy trajectories of a single protein and its polyproteins are equivalent: a direct experimental validation based on a small protein NuG2, Angew. Chem. Int. Ed. 56 (2017) 6117–6121.

[29] D.T. Edwards, J.K. Faulk, M.-A. LeBlanc, T.T. Perkins, Force spectroscopy with 9-μs resolution and sub-pN stability by tailoring AFM cantilever geometry, Biophys. J. 113 (2017) 2595–2600.

[30] D.T. Edwards, T.T. Perkins, Optimizing force spectroscopy by modifying commercial cantilevers: Improved stability, precision, and temporal resolution, J. Struct. Biol. 197 (2017) 13–25.

[31] C. He, C. Hu, X. Hu, X. Hu, A. Xiao, T.T. Perkins, H. Li, Direct observation of the reversible two-state unfolding and refolding of an α/β protein by single-molecule atomic force microscopy, Angew. Chem. Int. Ed. 54 (2015) 9921–9925.

[32] D.T. Edwards, M.-A. LeBlanc, T.T. Perkins, Modulation of a protein-folding landscape revealed by AFM-based force spectroscopy notwithstanding instrumental limitations, Proc. Natl. Acad. Sci. U.S.A. 118 (2021) e2015728118.

[33] T.M. Raschke, J. Kho, S. Marqusee, Confirmation of the hierarchical folding of RNase H: a protein engineering study, Nat. Struct. Biol. 6 (1999) 825–831.

[34] Z. Xi, Y. Gao, G. Sirinakis, H. Guo, Y. Zhang, Single-molecule observation of helix staggering, sliding, and coiled coil misfolding, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 5711–5716.

[35] J. Schönfelder, R. Perez-Jimenez, V. Muñoz, A simple two-state protein unfolds mechanically via multiple heterogeneous pathways at single-molecule resolution, Nat. Commun. 7 (2016) 11777.

[36] H. Chen, G. Yuan, R.S. Winardhi, M. Yao, I. Popa, J.M. Fernandez, J. Yan, Dynamics of equilibrium folding and unfolding transitions of titin immunoglobulin domain under constant forces, J. Am. Chem. Soc. 137 (2015) 3540–3546.

[37] A. Löf, P.U. Walker, S.M. Sedlak, S. Gruber, T. Obser, M.A. Brehm, M. Benoit, J. Lipfert, Multiplexed protein force spectroscopy reveals equilibrium protein folding dynamics and the low-force response of von Willebrand factor, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 18798–18807.

[38] P. Dabrowski-Tumanski, P. Rubach, D. Goundaroulis, J. Dorier, P. Sułkowski, K.C. Millett, E.J. Rawdon, A. Stasiak, J.I. Sulkowska, KnotProt 2.0: a database of proteins with knots and other entangled structures, Nucleic Acids Res. 47 (2019) D367–D375.

[39] A.L. Mallam, S.E. Jackson, A comparison of the folding of two knotted proteins: YbeA and YibK, J. Mol. Biol. 366 (2007) 650–665.

[40] F. Ziegler, N.C.H. Lim, S.S. Mandal, B. Pelz, W.-P. Ng, M. Schlierf, S.E. Jackson, M. Rief, Knotting and unknotting of a protein in single molecule experiments, Proc Natl Acad Sci USA. 113 (2016) 7533–7538.

[41] H. Wang, H. Li, Mechanically tightening, untying and retying a protein trefoil knot by single-molecule force spectroscopy, Chem. Sci. 11 (2020) 12512–12521.

[42] C. He, S. Li, X. Gao, A. Xiao, C. Hu, X. Hu, X. Hu, H. Li, Direct observation of the fast and robust folding of a slipknotted protein by optical tweezers, Nanoscale. 11 (2019) 3945–3951.

[43] H. Yu, D.R. Dee, X. Liu, A.M. Brigley, I. Sosova, M.T. Woodside, Protein misfolding occurs by slow diffusion across multiple barriers in a rough energy landscape, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 8308–8313.

[44] M. Sandal, F. Valle, I. Tessari, S. Mammi, E. Bergantino, F. Musiani, M. Brucale, L. Bubacco, B. Samorì, Conformational equilibria in monomeric α-synuclein at the single-molecule level, PLoS Biol. 6 (2008) e6.

[45] R. Hervás, J. Oroz, A. Galera-Prat, O. Goñi, A. Valbuena, A.M. Vera, À. Gómez-Sicilia, F. Losada-Urzáiz, V.N. Uversky, M. Menéndez, D.V. Laurents, M. Bruix, M. Carrión-Vázquez, Common features at the start of the neurodegeneration cascade, PLoS Biol. 10 (2012) e1001335.

[46] A. Solanki, K. Neupane, M.T. Woodside, Single-molecule force spectroscopy of rapidly fluctuating, marginally stable structures in the intrinsically disordered protein α-synuclein, Phys. Rev. Lett. 112 (2014) 158103.

[47] K. Neupane, A. Solanki, I. Sosova, M. Belov, M.T. Woodside, Diverse Metastable Structures Formed by Small Oligomers of a-Synuclein Probed by Force Spectroscopy, PLoS ONE. 9 (2014) e86495.

[48] J. Yu, S. Malkova, Y.L. Lyubchenko, α-synuclein misfolding: single molecule AFM force spectroscopy study, J. Mol. Biol. 384 (2008) 992–1001.

[49] Z. Lv, R. Roychaudhuri, M.M. Condron, D.B. Teplow, Y.L. Lyubchenko, Mechanism of amyloid β−protein dimerization determined using single−molecule AFM force spectroscopy, Sci. Rep. 3 (2013) 2880.

[50] Y. Zhang, Y.L. Lyubchenko, The structure of misfolded amyloidogenic dimers: computational analysis of force spectroscopy data, Biophys. J. 107 (2014) 2903–2910.

[51] C.D.M. Churchill, M.A. Healey, J. Preto, J.A. Tuszynski, M.T. Woodside, Probing the basis of α-synuclein aggregation by comparing simulations to single-molecule experiments, Biophys. J. 117 (2019) 1125–1135.

[52] T. Hoffmann, L. Dougan, Single molecule force spectroscopy using polyproteins, Chem. Soc. Rev. 41 (2012) 4781–4796.

[53] P.J. Elms, J.D. Chodera, C. Bustamante, S. Marqusee, The molten globule state is unusually deformable under mechanical force, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 3796–3801.

[54] B.E. Bowler, Residual structure in unfolded proteins, Curr. Opin. Struct. Biol. 22 (2012) 4–13.

[55] D.J. Brockwell, E. Paci, R.C. Zinober, G.S. Beddard, P.D. Olmsted, D.A. Smith, R.N. Perham, S.E. Radford, Pulling geometry defines the mechanical resistance of a β-sheet protein, Nat. Struct. Mol. Biol. 10 (2003) 731–737.

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

[57] H. Dietz, F. Berkemeier, M. Bertz, M. Rief, Anisotropic deformation response of single protein molecules, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 12724–12728.

[58] B. Jagannathan, P.J. Elms, C. Bustamante, S. Marqusee, Direct observation of a force-induced switch in the anisotropic mechanical unfolding pathway of a protein, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 17820–17825.

[59] E.J. Guinn, B. Jagannathan, S. Marqusee, Single-molecule chemo-mechanical unfolding reveals multiple transition state barriers in a small single-domain protein, Nat. Commun. 6 (2015) 6861.

[60] O.K. Dudko, G. Hummer, A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments, Phys. Rev. Lett. 96 (2006) 108101.

[61] C.A. Pierse, O.K. Dudko, Kinetics and energetics of biomolecular folding and binding, Biophys. J. 105 (2013) L19–L22.

[62] O.K. Dudko, G. Hummer, A. Szabo, Theory, analysis, and interpretation of single-molecule force spectroscopy experiments, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 15755–15760.

[63] Y. Zhang, O.K. Dudko, A transformation for the mechanical fingerprints of complex biomolecular interactions, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 16432–16437.

[64] E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds, Biophys. J. 72 (1997) 1541–1555.

[65] H. Yu, A.N. Gupta, X. Liu, K. Neupane, A.M. Brigley, I. Sosova, M.T. Woodside, Energy landscape analysis of native folding of the prion protein yields the diffusion constant, transition path time, and rates, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 14452–14457.

[66] A. Alemany, A. Mossa, I. Junier, F. Ritort, Experimental free-energy measurements of kinetic molecular states using fluctuation theorems, Nat. Phys. 8 (2012) 688–694.

[67] M.T. Woodside, P.C. Anthony, W.M. Behnke-Parks, K. Larizadeh, D. Herschlag, S.M. Block, Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid, Science. 314 (2006) 1001–1004.

[68] H.-K. Choi, D. Min, H. Kang, M.J. Shon, S.-H. Rah, H.C. Kim, H. Jeong, H.-J. Choi, J.U. Bowie, T.-Y. Yoon, Watching helical membrane proteins fold reveals a common N-to-C-terminal folding pathway, Science. 366 (2019) 1150–1156.

[69] G. Hummer, A. Szabo, Free energy reconstruction from nonequilibrium single-molecule pulling experiments, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 3658–3661.

[70] G. Hummer, A. Szabo, Free energy profiles from single-molecule pulling experiments, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 21441–21446.

[71] A.N. Gupta, A. Vincent, K. Neupane, H. Yu, F. Wang, M.T. Woodside, Experimental validation of free-energy-landscape reconstruction from non-equilibrium single-molecule force spectroscopy measurements, Nat. Phys. 7 (2011) 631–634.

[72] M.C. Engel, D.B. Ritchie, D.A.N. Foster, K.S.D. Beach, M.T. Woodside, Reconstructing folding energy landscape profiles from nonequilibrium pulling curves with an inverse weierstrass integral transform, Phys. Rev. Lett. 113 (2014) 238104.

[73] A.P. Manuel, J. Lambert, M.T. Woodside, Reconstructing folding energy landscapes from splitting probability analysis of single-molecule trajectories, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 7183–7188.

[74] J.C.M. Gebhardt, T. Bornschlögl, M. Rief, Full distance-resolved folding energy landscape of one single protein molecule, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 2013–2018.

[75] K. Neupane, A.P. Manuel, M.T. Woodside, Protein folding trajectories can be described quantitatively by one-dimensional diffusion over measured energy landscapes, Nat. Phys. 12 (2016) 700–703.

[76] G.-M. Nam, D.E. Makarov, Extracting intrinsic dynamic parameters of biomolecular folding from single-molecule force spectroscopy experiments, Protein Sci. 25 (2016) 123–134.

[77] P. Cossio, G. Hummer, A. Szabo, On artifacts in single-molecule force spectroscopy, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 14248–14253.

[78] K. Neupane, M.T. Woodside, Quantifying instrumental artifacts in folding kinetics measured by single-molecule force spectroscopy, Biophys. J. 111 (2016) 283–286.

[79] R. Berkovich, R.I. Hermans, I. Popa, G. Stirnemann, S. Garcia-Manyes, B.J. Berne, J.M. Fernandez, Rate limit of protein elastic response is tether dependent, Proc. Natl. Acad. Sci. U.S.A. 109 (2012) 14416–14421.

[80] H. Lannon, J.S. Haghpanah, J.K. Montclare, E. Vanden-Eijnden, J. Brujic, Force-clamp experiments reveal the free-energy profile and diffusion coefficient of the collapse of protein molecules, Phys. Rev. Lett. 110 (2013) 128301.

[81] G. Zoldak, J. Stigler, B. Pelz, H. Li, M. Rief, Ultrafast folding kinetics and cooperativity of villin headpiece in single-molecule force spectroscopy, Proc. Natl. Acad. Sci. U.S.A. 110 (2013) 18156–18161.

[82] J. Schönfelder, D. De Sancho, R. Berkovich, R.B. Best, V. Muñoz, R. Perez-Jimenez, Reversible two-state folding of the ultrafast protein gpW under mechanical force, Commun. Chem. 1 (2018) 59.

[83] H.S. Chung, K. McHale, J.M. Louis, W.A. Eaton, Single-molecule fluorescence experiments determine protein folding transition path times, Science. 335 (2012) 981–984.

[84] A. Mehlich, J. Fang, B. Pelz, H. Li, J. Stigler, Slow transition path times reveal a complex folding barrier in a designed protein, Front. Chem. 8 (2020) 587824.

[85] P. Cossio, G. Hummer, A. Szabo, Transition paths in single-molecule force spectroscopy, J. Chem. Phys. 148 (2018) 123309.

[86] S. Batey, A.A. Nickson, J. Clarke, Studying the folding of multidomain proteins, HFSP J. 2 (2008) 365–377.

[87] R. Hertadi, F. Gruswitz, L. Silver, A. Koide, S. Koide, H. Arakawa, A. Ikai, Unfolding mechanics of multiple OspA substructures investigated with single molecule force spectroscopy, J. Mol. Biol. 333 (2003) 993–1002.

[88] Z.N. Scholl, W. Yang, P.E. Marszalek, Chaperones rescue luciferase folding by separating its domains, J. Biol. Chem. 289 (2014) 28607–28618.

[89] S. Garcia-Manyes, D. Giganti, C.L. Badilla, A. Lezamiz, J. Perales-Calvo, A.E.M. Beedle, J.M. Fernández, Single-molecule force spectroscopy predicts a misfolded, domain-swapped conformation in human γD-crystallin protein, J. Biol. Chem. 291 (2016) 4226–4235.

[90] H.C. Kotamarthi, R. Sharma, S. Narayan, S. Ray, S.R.K. Ainavarapu, Multiple unfolding pathways of leucine binding protein (LBP) probed by single-molecule force spectroscopy (SMFS), J. Am. Chem. Soc. 135 (2013) 14768–14774.

[91] Q. Li, Z.N. Scholl, P.E. Marszalek, Unraveling the mechanical unfolding pathways of a multidomain protein: phosphoglycerate kinase, Biophys. J. 115 (2018) 46–58.

[92] D. Bauer, S. Meinhold, R.P. Jakob, J. Stigler, U. Merkel, T. Maier, M. Rief, G. Žoldák, A folding nucleus and minimal ATP binding domain of Hsp70 identified by single-molecule force spectroscopy, Proc. Natl. Acad. Sci. U.S.A. 115 (2018) 4666–4671.

[93] D. Bauer, D.R. Merz, B. Pelz, K.E. Theisen, G. Yacyshyn, D. Mokranjac, R.I. Dima, M. Rief, G. Žoldák, Nucleotides regulate the mechanical hierarchy between subdomains of the nucleotide binding domain of the Hsp70 chaperone DnaK, Proc. Natl. Acad. Sci. U.S.A. 112 (2015) 10389–10394.

[94] M. Jahn, J. Buchner, T. Hugel, M. Rief, Folding and assembly of the large molecular machine Hsp90 studied in single-molecule experiments, Proc. Natl. Acad. Sci. U.S.A. 113 (2016) 1232–1237.

[95] F.U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis, Nature. 475 (2011) 324–332.

[96] P. Bechtluft, R.G.H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H.L. Tepper, A.J.M. Driessen, S.J. Tans, Direct observation of chaperone-Induced changes in a protein folding pathway, Science. 318 (2007) 1458–1461.

[97] A. Mashaghi, G. Kramer, P. Bechtluft, B. Zachmann-Brand, A.J.M. Driessen, B. Bukau, S.J. Tans, Reshaping of the conformational search of a protein by the chaperone trigger factor, Nature. 500 (2013) 98–101.

[98] S. Ungelenk, F. Moayed, C.-T. Ho, T. Grousl, A. Scharf, A. Mashaghi, S. Tans, M.P. Mayer, A. Mogk, B. Bukau, Small heat shock proteins sequester misfolding proteins in near-native conformation for cellular protection and efficient refolding, Nat. Commun. 7 (2016) 13673.

[99] F. Moayed, S. Bezrukavnikov, M.M. Naqvi, B. Groitl, C.M. Cremers, G. Kramer, K. Ghosh, U. Jakob, S.J. Tans, The anti-aggregation holdase Hsp33 promotes the formation of folded protein structures, Biophys. J. 118 (2020) 85–95.

[100] A. Mashaghi, S. Bezrukavnikov, D.P. Minde, A.S. Wentink, R. Kityk, B. Zachmann-Brand, M.P. Mayer, G. Kramer, B. Bukau, S.J. Tans, Alternative modes of client binding enable functional plasticity of Hsp70, Nature. 539 (2016) 448–451.

[101] M.J. Avellaneda, K.B. Franke, V. Sunderlikova, B. Bukau, A. Mogk, S.J. Tans, Processive extrusion of polypeptide loops by a Hsp100 disaggregase, Nature. 578 (2020) 317–320.

[102] S. Haldar, R. Tapia-Rojo, E.C. Eckels, J. Valle-Orero, J.M. Fernandez, Trigger factor chaperone acts as a mechanical foldase, Nat. Commun. 8 (2017) 668.

[103] J. Perales-Calvo, D. Giganti, G. Stirnemann, S. Garcia-Manyes, The force-dependent mechanism of DnaK-mediated mechanical folding, Sci. Adv. 4 (2018) eaaq0243.

[104] A.N. Gupta, K. Neupane, N. Rezajooei, L.M. Cortez, V.L. Sim, M.T. Woodside, Pharmacological chaperone reshapes the energy landscape for folding and aggregation of the prion protein, Nat. Commun. 7 (2016) 12058.

[105] L. Ma, Y. Kang, J. Jiao, A.A. Rebane, H.K. Cha, Z. Xi, H. Qu, Y. Zhang, α-SNAP enhances SNARE zippering by stabilizing the SNARE four-helix bundle, Cell Rep. 15 (2016) 531–539.

[106] T. Shu, H. Jin, J.E. Rothman, Y. Zhang, Munc13-1 MUN domain and Munc18-1 cooperatively chaperone SNARE assembly through a tetrameric complex, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 1036–1041.

[107] R.V. Kalyana Sundaram, H. Jin, F. Li, T. Shu, J. Coleman, J. Yang, F. Pincet, Y. Zhang, J.E. Rothman, S.S. Krishnakumar, Munc13 binds and recruits SNAP25 to chaperone SNARE complex assembly, FEBS Lett. 595 (2021) 297–309.

[108] K. Boström, M. Wettesten, J. Borén, G. Bondjers, O. Wiklund, S.-O. Olofsson, Pulse-chase studies of the synthesis and intracellular transport of apolipoprotein B-100 in Hep G2 cells., 261 (1986) 13800–13806.

[109] J.A. Farías-Rico, F. Ruud Selin, I. Myronidi, M. Frühauf, G. von Heijne, Effects of protein size, thermodynamic stability, and net charge on cotranslational folding on the ribosome, Proc. Natl. Acad. Sci. U.S.A. 115 (2018) E9280–E9287.

[110] C.M. Kaiser, D.H. Goldman, J.D. Chodera, I. Tinoco, C. Bustamante, The ribosome modulates nascent protein folding, Science. 334 (2011) 1723–1727.

[111] E.J. Guinn, P. Tian, M. Shin, R.B. Best, S. Marqusee, A small single-domain protein folds through the same pathway on and off the ribosome, Proc. Natl. Acad. Sci. U.S.A. 115 (2018) 12206–12211.

[112] K. Liu, J.E. Rehfus, E. Mattson, C.M. Kaiser, The ribosome destabilizes native and non-native structures in a nascent multidomain protein, Protein Sci. 26 (2017) 1439–1451.

[113] K. Liu, K. Maciuba, C.M. Kaiser, The ribosome cooperates with a chaperone to guide multi-domain protein folding, Mol. Cell. 74 (2019) 310–319.

[114] K. Liu, X. Chen, C.M. Kaiser, Energetic dependencies dictate folding mechanism in a complex protein, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 25641–25648.

[115] F. Wruck, P. Tian, R. Kudva, R.B. Best, G. von Heijne, S.J. Tans, A. Katranidis, The ribosome modulates folding inside the ribosomal exit tunnel, Commun. Biol. 4 (2021) 523.

[116] R. Santos, O. Ursu, A. Gaulton, A.P. Bento, R.S. Donadi, C.G. Bologa, A. Karlsson, B. Al-Lazikani, A. Hersey, T.I. Oprea, J.P. Overington, A comprehensive map of molecular drug targets, Nat. Rev. Drug Discov. 16 (2017) 19–34.

[117] F. Oesterhelt, D. Oesterhelt, M. Pfeiffer, A. Engel, H.E. Gaub, D.J. Müller, Unfolding pathways of individual bacteriorhodopsins, Science. 288 (2000) 143–146.

[118] M. Kessler, H.E. Gaub, Unfolding barriers in bacteriorhodopsin probed from the cytoplasmic and the extracellular side by AFM, Structure. 14 (2006) 521–527.

[119] M. Zocher, J.J. Fung, B.K. Kobilka, D.J. Müller, Ligand-specific interactions modulate kinetic, energetic, and mechanical properties of the human β_2 adrenergic receptor, Structure. 20 (2012) 1391–1402.

[120] R. Petrosyan, C.A. Bippes, S. Walheim, D. Harder, D. Fotiadis, T. Schimmel, D. Alsteens, D.J. Müller, Single-molecule force spectroscopy of membrane proteins from membranes freely spanning across nanoscopic pores, Nano Lett. 15 (2015) 3624–3633.

[121] M. Kessler, K.E. Gottschalk, H. Janovjak, D.J. Muller, H.E. Gaub, Bacteriorhodopsin folds into the membrane against an external force, J. Mol. Biol. 357 (2006) 644–654.

[122] A. Kedrov, H. Janovjak, C. Ziegler, W. Kuhlbrandt, D.J. Muller, Observing folding pathways and kinetics of a single sodium-proton antiporter from *Escherichia coli*, J. Mol. Biol. 355 (2006) 2–8.

[123] K.T. Sapra, M. Damaghi, S. Köster, Ö. Yildiz, W. Kühlbrandt, D.J. Muller, One β hairpin after the other: exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG, Angew. Chem. Int. Ed. 48 (2009) 8306–8308.

[124] M. Damaghi, S. Köster, C.A. Bippes, Ö. Yildiz, D.J. Müller, One β hairpin follows the other: exploring refolding pathways and kinetics of the transmembrane β-barrel protein OmpG, Angew. Chem. 123 (2011) 7560–7562.

[125] P.D. Bosshart, I. Iordanov, C. Garzon-Coral, P. Demange, A. Engel, A. Milon, D.J. Müller, The transmembrane protein KpOmpA anchoring the outer membrane of *Klebsiella pneumoniae* unfolds and refolds in response to tensile load, Structure. 20 (2012) 121–127.

[126] J. Thoma, B.M. Burmann, S. Hiller, D.J. Müller, Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins, Nat. Struct. Mol. Biol. 22 (2015) 795–802.

[127] T. Serdiuk, D. Balasubramaniam, J. Sugihara, S.A. Mari, H.R. Kaback, D.J. Müller, YidC assists the stepwise and stochastic folding of membrane proteins, Nat. Chem. Biol. 12 (2016) 911–917.

[128] T. Serdiuk, A. Steudle, S.A. Mari, S. Manioglu, H.R. Kaback, A. Kuhn, D.J. Müller, Insertion and folding pathways of single membrane proteins guided by translocases and insertases, Sci. Adv. 5 (2019) eaau6824.

[129] P.R. Heenan, H. Yu, M.G.W. Siewny, T.T. Perkins, Improved free-energy landscape reconstruction of bacteriorhodopsin highlights local variations in unfolding energy, J. Chem. Phys. 148 (2018) 123313.

[130] H. Yu, P.R. Heenan, D.T. Edwards, L. Uyetake, T.T. Perkins, Quantifying the initial unfolding of bacteriorhodopsin reveals retinal stabilization, Angew. Chem. Int. Ed. 58 (2019) 1710–1713.

[131] H. Yu, D.R. Jacobson, H. Luo, T.T. Perkins, Quantifying the native energetics stabilizing bacteriorhodopsin by single-molecule force spectroscopy, Phys. Rev. Lett. 125 (2020) 068102.

[132] D.R. Jacobson, T.T. Perkins, Free-energy changes of bacteriorhodopsin point mutants measured by single-molecule force spectroscopy, Proc. Natl. Acad. Sci. U.S.A. 118 (2021) e2020083118.

[133] D. Min, R.E. Jefferson, Y. Qi, J.Y. Wang, M.A. Arbing, W. Im, J.U. Bowie, Unfolding of a ClC chloride transporter retains memory of its evolutionary history, Nat. Chem. Biol. 14 (2018) 489–496.

[134] P. Lu, D. Min, F. DiMaio, K.Y. Wei, M.D. Vahey, S.E. Boyken, Z. Chen, J.A. Fallas, G. Ueda, W. Sheffler, V.K. Mulligan, W. Xu, J.U. Bowie, D. Baker, Accurate computational design of multipass transmembrane proteins, Science. 359 (2018) 1042–1046.

[135] E. Infante, A. Stannard, S.J. Board, P. Rico-Lastres, E. Rostkova, A.E.M. Beedle, A. Lezamiz, Y.J. Wang, S. Gulaidi Breen, F. Panagaki, V. Sundar Rajan, C. Shanahan, P. Roca-Cusachs, S. Garcia-Manyes, The mechanical stability of proteins regulates their translocation rate into the cell nucleus, Nat. Phys. 15 (2019) 973–981.