Published in final edited form as:

Cell Mol Life Sci. 2015 November ; 72(22): 4273–4285. doi:10.1007/s00018-015-2002-3.

Fast-Folding Proteins under Stress

Kapil Dave¹ and Martin Gruebele^{1,2}

¹Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

²Departments of Chemistry and Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA

Abstract

Proteins are subject to a variety of stresses in biological organisms, including pressure and temperature, which are the easiest stresses to simulate by molecular dynamics. We discuss the effect of pressure and thermal stress on very fast folding model proteins, whose *in vitro* folding can be fully simulated on computers and compared with experiments. We then discuss experiments that can be used to subject proteins to low and high temperature unfolding, as well as low and high pressure unfolding. Pressure and temperature are prototypical perturbations that illustrate how close many proteins are to instability, a property that cells can exploit to control protein function. We conclude by reviewing some recent in-cell experiments, and progress being made in simulating and measuring protein stability and function inside live cells.

Keywords

Phase diagram; WW domain; Temperature jump; NTL9; Pressure jump; Proton NMR; Fluorescence; Cell cycle; Molecular Dynamics

1. Introduction

Proteins are important biomolecules present in all life forms, where they perform vital cellular functions. Many proteins fold into compact structures that are however filled with voids of various sizes [1–3]. Other proteins have more disordered native states and fold only when they bind [4]. Within cells, membranes or extracellular spaces, proteins are constantly subject to changing solvation environments, external stresses, and crowding by other biomolecules. For example, the pH of the lysosome is about 2 units lower than in the cytosol. Temperature varies, sometimes by small amounts (e.g. in mammalian cells with heat regulation), sometimes by large amounts (e.g. insects cycling from hot to sub-freezing temperatures [5]). Some animal's proteins are subject to large pressure differentials, such as whales diving to pressures of 100s of atmospheres [6]). Some proteins are subject to both compression and tension, such as titin, a gigantic protein responsible for the elasticity of muscles [7]. Organisms have adapted to extreme conditions of temperature or pressure, such

as bacteria in the gabbroic layer of the Earth's crust [8], or fish at the oceanic depth limit of *ca.* –8500 meters, below which osmolytes no longer keep critical proteins stable [9].

When subject to such stresses, proteins display a wide range of stabilities. Some proteins are extremely stable and degrade very slowly, including several proteins in the blood plasma [10]. Others occur in small copy numbers and are easily degraded as part of cellular control [11]. Many enzymes must be quite flexible to adapt to substrates by induced fit or conformational selection, and the price paid off is low stability.

Is it just a price paid, or could cells benefit from marginal stability of some proteins? Cells have a number of ways to regulate function. For example, transcription can be regulated so a protein is not even made. The messenger RNA can be manipulated to avoid its translation on the ribosome. Post-translational modification, such as glycosylation, dramatically alters function in many cases. Subsequent to all of these controls, marginal stability offers the cell another handle on protein regulation. Subjecting a marginally stable protein to stress, either external or due to interactions within the crowded cell, could affect its function. Such 'post-post-translational' regulation could control protein function throughout a protein's life cycle [12]. Much is already known about how specific binding regulates protein function, but little is known about how weaker transient interactions and small external stresses control the cell's proteome. McConkey termed transient but functional association of proteins with other biomolecules 'quinary structure' [13]. Macromolecular crowding, quinary interactions, and applied stresses from the environment are all candidates for more subtle regulation of proteins.

A few examples are already known. The protein lympotactin exists as a GPCR agonist in a pure beta sheet form above 37 °C, and as a GAG binding protein in a beta-alpha structure below 37 °C. A small switch of temperature changes the structure and function of this protein, which metamorphoses between two folds [14]. Another example is RfaH shown in Fig. 1, which switches between a helix bundle and beta sheet domain depending on how it is crowded by a neighboring domain. In one structure, it is a transcription factor, in the other, a translation factor [15]. Many more such multi-tasking proteins undoubtedly exist.

In this review, we discuss in detail how proteins undergo sometimes surprising conformational transitions (folding-unfolding) under external stress. We illustrate the idea with two simple variables, temperature and pressure, and consider four cases: what happens at high/low pressure, and at high/low temperature. We will see that the observed behaviors are universal: Almost every protein will show these based on the underlying physical properties of proteins, although far fewer proteins will show the transitions near physiological conditions, where cells can utilize them to control their proteome. Thus the physical chemistry of protein folding has interesting responses to external stresses 'built in,' and some proteins have evolved to use these responses under physiological conditions. In addition, the response of proteins to stress has been hijacked by researchers who study the folding of small proteins. Fast-folding small proteins have turned out to be very useful for comparing computer simulations directly with experiments [16], and we discuss a few such examples as we look at proteins under stress. The review concludes by returning to the point

we raised at the outset: Is protein folding sufficiently sensitive to the cellular environment and to external stresses for it to make a difference *in vivo*?

2. Proteins: temperature and pressure effects

Protein stability is highly sensitive to temperature and pressure stresses. We all know that proteins unfold at high temperature, and will aggregate when kept at high concentration or at high temperature. The albumin in hard-boiled egg white is a good example. The process is often reversible – indeed even egg whites have been ‘unboiled’ [17]. Fig. 2 shows the phase diagram for protein folding as a function of temperature and pressure. Many experiments over the years have shown that proteins are stable in an elliptical-shaped ‘sweet spot’: not too hot or too cold, not too much pressure or too much tension [18].

Heat unfolding is easily rationalized: higher polypeptide chain disorder is favored at high temperature. Both the protein chain and solvent water molecules can move around more when the polypeptide chain is unfolded at high temperature.

Proteins also unfold at low temperature. This more mysterious process would be somewhat like ice turning back into water when cooled down even more. That of course would violate the second law of thermodynamics. Somehow, the entropy of the protein+solvent system must decrease as energy is removed from the system and the temperature is lowered. So why does the polypeptide chain unfold at low temperature? The widely accepted answer [19] is that water molecules can form cages around exposed amino acid side chains of the unfolded protein, actually increasing the order of the combined protein+solvent system even though the disorder of the protein chain increases. The full answer is of course subtler. There are also enthalpic contributions to the low temperature unfolding of proteins. These contributions happen to be small near around room temperature, but they become more substantial at higher temperature [20].

High pressure unfolding of proteins is equally puzzling at a first glance. We tend to think of proteins as larger once they unfold, but Le Châtelier’s principle states the system should go to a smaller volume when pressure is increased. That is what actually happens [21]: unfolded proteins have a smaller molar volume than folded proteins. The key are the ‘voids’ alluded to in the introduction. Folded proteins are full of various-sized voids because ‘only’ 20 amino acid shapes are available for packing, and because of structural constraints imposed by the polypeptide chain. Even when optimally packed, proteins are actually 2.7 dimensional fractals [3]. When a protein unfolds, its voids become occupied by water molecules. Even though the unfolded chain looks larger, it displaces less solvent volume than the folded protein.

Strangest perhaps is protein unfolding under tension, as shown at the bottom of the ellipse in Fig. 2. Hydrostatic tension is just negative pressure: force is exerted to expand a liquid (negative pressure), instead of compressing the liquid (positive pressure). How this is done experimentally is explained in section 3. When water is put under tension, it transfers some of that force to any proteins dissolved in it. The adhesion of a protein’s surface to surrounding water molecules literally pulls on the protein in all directions, making it less stable [22]. This is not to be confused with force pulling experiments, where the force is

exerted vectorially instead of hydrostatically, by hooking up the protein to a surface at one end, and an AFM tip at the other end [23].

In section 4, we will look at fast folding and stability of four model proteins subject to those four stresses: heat unfolding, cold unfolding, pressure unfolding and tension unfolding. The phase diagram in Fig. 2 turns out to be very useful for experimental and computational studies: many small, fast-folding proteins have a phase diagram very similar to the one in Fig. 2. It is relatively easy both on the computer and in the test tube to apply high/low temperature/pressure stresses, to see how proteins fold or unfold in response to the stress. Indeed, proteins can now be completely folded into native structures by computer simulations when such stresses are removed [24], and such simulation compare fairly well with experimental data, or even predict how experiments are affected by mutations [25,16].

The sensitivity of proteins to weak stress conditions is a consequence of evolution for function. For substrates to diffuse into enzymes, for signaling proteins too efficiently ‘induce fit’ onto their binding partners, and many other such processes, proteins must be quite flexible. To create flexibility, voids can be a good thing, and easy unfolding upon pressure or temperature stress is the flip side of flexibility and functional dynamics of the protein structure [3]. Pressure, temperature, and many other cellular stresses from low pH to interaction with other bio macromolecules can thus in principle control protein structure and function, as mentioned before in the introduction. The crowding in cells and weak biomolecular interactions can have big effects: for example they can lead to phase separation into a protein-rich and protein-deficient phase. The result is membrane-free ‘micro-organelles’ that behave differently from the surrounding cytoplasm and have been reported to be seen both *in vitro* and *in vivo* studies [26,27]. They are seen in P granules in the germ line cells, where the phase separation is driven by RNA-protein binding [28,29].

3. Methods for protein folding dynamics using pressure and temperature

Experimental techniques used to study fast protein folding or unfolding can be broadly categorized into one of four combinations of: single molecule *vs.* bulk and relaxation *vs.* equilibrium fluctuation measurements. We discuss these methods briefly and non-technically before going through the four case studies in the next section. The technically inclined reader can find methodology details in refs [30–32].

Let us begin with relaxation measurements. In such an experiment, the chemical system is quickly taken out of equilibrium – so fast that the reactants cannot respond. For example, if the pressure or temperature of a protein is jumped in nanoseconds, the polypeptide chain simply cannot fold or unfold that fast. It will relax to the new equilibrium after the jump is complete. This relaxation can be measured by following some spectroscopic variable that is sensitive to folding. Such variables include the infrared absorption of amide linkages in the polypeptide chain (sensitive to secondary structure) [33], or tryptophan fluorescence intensity (sensitive to quenching in the local tertiary structure environment).

Fig. 3 illustrates this approach: A cuvette with a protein solution can be jumped up in temperature by a heating laser (red in Fig. 3B). The jump can be done with nanosecond or even picosecond time resolution [30]. Later, the laser is turned off and the temperature drops

back down; this can be done with millisecond time resolution [34,35]. A probe laser (blue in Fig. 3B) monitors the protein by providing a signal that depends on the amount of folded *vs.* unfolded protein present. The probe signal (Fig. 3A) does not respond to the temperature jumps instantaneously. It responds at the rate at which the polypeptide chain can diffuse through the solvent and cross the activation barrier. The solid curve is for a slow reaction, the dotted curve is for a faster folding reaction. The reaction can also be more complicated and occur in multiple steps. The same effect can be achieved by jumping pressure instead of temperature (Fig. 3C).

Now consider kinetics measurements at equilibrium instead of by inducing relaxation. At a first glance, it seems contradictory that one should be able to measure kinetics while a system is at equilibrium. However, both single molecule and bulk measurement techniques can achieve this feat. In single molecule (or few-molecule) measurements, fluctuations give away the kinetics. Lars Onsager proved in the 1930s that equilibrium fluctuations contains the same information as relaxation [36], as long as the system is not jumped too far from equilibrium in the relaxation experiments.

Fig. 4 illustrates this approach: imagine you have just one protein in equilibrium at a temperature and pressure combination that favors the folded and unfolded states equally. The protein constantly switches back between the folded and unfolded states (Fig. 4 top). With a sensitive detector that rejects background signals, fluorescence from a single protein molecule can be detected [37]. Let us say the folded state emits more red light, and the unfolded state emits more green light. Then the probe signal will switch between more green light if the protein is unfolded, and more red light if it is folded. Fig. 4 (bottom) shows what a time trace might look like. What Onsager proved is: the faster an individual protein molecule switches back and forth in Fig. 4, the faster the decay from many proteins being observed at once as in Fig. 3.

The above examples illustrated relaxation kinetics experiments using an ensemble sample, and equilibrium kinetics experiments using a single molecule sample. However, *vice versa* is also perfectly possible. For example, ensemble protein samples in equilibrium observed by NMR spectroscopy of tyrosine or histidine protons can reveal folding kinetics as well [38] (Fig. 5): There are two peaks in the NMR spectrum, one for the tyrosine in the folded proteins, the other for the equivalent tyrosine in the unfolded proteins. Of course these tyrosine's interconvert all the time as proteins fold and unfold, so any individual protein contributes to both peaks at different times. This leads to a broadening of the peaks: the faster the interconversion rate, the broader the peaks. (Eventually, the peak becomes so broad that the two individual peaks can no longer be distinguished, and finally narrow to a single peak in-between the two original peaks. When this happens, NMR can no longer measure the rate because the rate has gotten too fast.) The width of the peaks can be used to measure protein folding on the microsecond time scale. Finally, consider the other *vice versa*: applying jumps to single-molecule or few-molecule experiments. This has been done, and can be used to observe how a sudden change in environment influences the equilibrium of individual molecules [39].

This brief review of experimental concepts should be sufficient to follow the examples in the next section. The reader interested in technical details should consult the references listed above. Before turning to the case studies, we should also briefly discuss computational methods used to study fast folding.

A general theory for folding based on the concept of the energy landscape exists [40–43]. The theory posits that proteins have evolved to have a sufficiently low enthalpy in the folded state (due to contacts formed), so that the low enthalpy compensates for the unfavorable loss of disorder (decrease of entropy) as the chain folds. In terms of the free energy change $\Delta G = \Delta H - T \Delta S$, this means that the lowering of enthalpy (negative ΔH) compensates the lowering of entropy (positive $-T \Delta S$), so $\Delta G < 0$ and folding is spontaneous. The compensation is not always perfect, so en route to the folded state, a protein can encounter an activation barrier. Sometimes it can be perfect, and then the protein folds downhill without encountering a significant activation barrier (barrier $< 3 kT$) [44–47].

The theory explains folding in general, and experiments provide specific rates and equilibrium constants, but it would be nice to also have detailed structural insight into the process. X-ray crystallography and structural NMR are just not fast enough to follow fast (microsecond) folding reactions. There are also other problems, such as crystals fracturing when the molecule making up the crystal changes its structure too much (e.g. by unfolding). Molecular dynamics (MD) simulations can be very useful in that case.

In MD simulations, the true force field is replaced by an approximate formula, and Newtonian mechanics are used to solve for the motion of the protein. Thus both the force and the motion are treated approximately because a rigorous treatment requires quantum mechanics. MD simulations are not accurate under all conditions, but they can provide a nice picture of how protein structure evolves in time, that can then be checked by experiment. Results from such simulations are shown in Fig. 6. In 6A, the simulated deviation of the protein from the folded structure (RMSD) is shown. When RMSD is small, the protein is folded. When the protein crosses the activation barrier, which happens very rapidly, the trace switches to a large RMSD. This looks very much like the single molecule trace in Fig. 4, and indeed, one can think of the MD simulation of one protein as a ‘single molecule experiment done on the computer.’ Fig. 6B illustrates the free energy ΔG computed by molecular dynamics for an ensemble of proteins. The free energy landscape is plotted as a function of reaction coordinate RMSD. At very high temperature (above the boiling point of water in this example!), ΔG has just a single minimum at $\text{RMSD} \approx 6 \text{ \AA}$, corresponding to an unfolded protein. At low temperature, ΔG has several minima, including the deepest one at the folded state ($\text{RMSD} \approx 0.5 \text{ \AA}$). At even lower temperature (not shown) the folded minimum is much deeper than any of the others, so $\Delta G < 0$. Such ΔG can be measured by bulk thermodynamic or single molecule experiments. Representative computed structures for the small peptide in question (trpzip2) are also shown in Fig. 6B.

4. What can we learn from fast protein folding?

In this section, we present four case studies. Each case involves experiments and simulations of a fast-folding protein under stress, covering one of the four unfolding conditions we

discussed in section 2 and showed in Fig. 2: high temperature unfolding, low temperature unfolding, high pressure unfolding, and tension unfolding.

Section 4a starts off with high temperature unfolding (the right side of the phase diagram in Fig. 2) of a mixed α/β protein NTL9₁₋₃₉. Fast laser temperature jumps and 2-dimensional infrared spectroscopy are combined to yield structural insight. These experiments revealed the presence of multiple pathways of folding even for such a small protein.

Section 4b moves on to low temperature unfolding of apomyoglobin (left side of the phase diagram in Fig. 2). Apomyoglobin refolds in a two-step process. “Functional frustration” is the underlying reason for the existence of a slow and fast step: The heme group that binds oxygen is absent in apomyoglobin, slowing down the folding of the heme-binding part of the protein. The rest of apomyoglobin folds orders of magnitude faster.

Section 4c describes the denaturation of a small β sheet protein (WW Domain) by increasing the pressure (top of the phase diagram in Fig. 2). Refolding is initiated by suddenly dropping the pressure back to 1 atmosphere. Simulations confirm that the protein can choose one of two paths on its way to the folded state: either form one pair of strands first, or the other. One of these possibilities is more likely than the other, and folding reactions often have a dominant pathway even when many paths are taken.

Finally, section 4d explains unfolding of ubiquitin under negative pressure, where tension is applied to the protein (bottom of the phase diagram). It was confirmed using proton NMR spectroscopy that indeed ubiquitin becomes less stable under tension. That experiment was the last one to be carried out chronologically on a protein, completing the elliptical phase diagram shown in Fig. 2.

a. NTL9₁₋₃₉ (as an example of heat high temperature unfolding)

We will start on the right side of the phase diagram in Fig. 2 with high temperature unfolding. But first, let us consider the biology of our example mini-protein NTL9₁₋₃₉.

The ribosome is an important organelle present in all cellular organisms. It is responsible for the synthesis of proteins. This cellular machinery is made up of two subunits, each constituting many RNA and protein molecules. NTL9₁₋₃₉ is a small 39 residue mixed α/β protein derived from the N terminal domain of ribosomal protein L9 situated in the large 50S subunit of eukaryotic cells (Fig. 7). The presence of both alpha/beta secondary structure makes it a good model for formation of both major secondary structure in proteins as compared to the other members of fast folding small proteins which have mostly one type of secondary structure like WW domain (β sheet model), or lambda repressor [49] (α -helical model). The importance of having these miniature proteins is they can still have activity but fold quickly (microseconds) for making comparisons between computational and experimental folding studies. They contain the same motifs commonly found in larger proteins, so we can then use them to understand how larger proteins fold as well.

An interesting aspect of NTL9₁₋₃₉ in the biological context is its gene expression induced by osmotic stress. A study conducted on variants of *Arabidopsis* (a commonly used model

system plant) reported NTL9₁₋₃₉ involvement in leaf senescence by mediating a stress signal. (Leaf senescence is a process involving moving of metabolites from aging leaves to seeds, the last stage of plant development) [50].

NTL9₁₋₃₉ is an independent folding domain. The first thermodynamic analysis of this protein by heat unfolding yielded a surprisingly high melting temperature $T_m = 65.3 \pm 0.8^\circ\text{C}$ [52] for such a small protein. T_m is defined as the temperature at which half of the ensemble is in the folded state and other half is in the unfolded state. The unfolding takes place on the millisecond time scale as shown by fluorescence upon rapid mixing of a bulk sample with denaturant (urea). This rapid mixing ‘jumps’ the ensemble of proteins, which then relaxes to an unfolded state. The study deduced that NTL9₁₋₃₉ folds co-operatively and rapidly in a two state fashion into a well-defined native state characteristic of a normal globular protein.

A more recent study of NTL9₁₋₃₉ combined temperature jump with two-dimensional infrared spectroscopy (2D IR) as the probe [49]. 2D IR spectroscopy is a powerful technique because it combines the ‘fingerprint’ information about chemical groups (e.g. the amides between amino acid residues) possible with IR spectroscopy with structural correlation from one part of the protein to another, like 2D NMR spectroscopy. The results obtained from the combined computational and experimental study showed that NTL9₁₋₃₉ has a compact native state along with a fast-exchanging unfolded state with residual structure. The native state is reached by multiple folding pathways for this protein. The existence of multiple folding pathways (i.e. the flexible order in which subunits can be put together) and residual structure in the unfolded state (which is not just a ‘random coil’) are themes seen in many such experiments. Usually one pathway tends to be more important than the others, though: even a slightly lower free energy ΔG will strongly favor the probability of the lower pathway [34].

What can be learned from the studies conducted on NTL9₁₋₃₉? First, even such small proteins are complicated enough to offer mechanistic discrepancies between simulations and experiments, and are a challenging test of computer simulations. Secondly, current fast folding experiments lack the detailed structural resolution of folding pathway(s), and simulations can step in to provide such structural information. More often than not, the dynamics turn out to be more complex than one would have suspected based on a few averaged experimental variables.

b. Apomyoglobin (as an example of low temperature unfolding)

We now move to the left side of the phase diagram in Fig. 2: low temperature (“cold”) unfolding. We illustrate the idea with apomyoglobin after reviewing a little bit of this protein’s history and biological properties.

Myoglobin is the first protein whose structure was determined by X-ray crystallography by Kendrew and coworkers in 1958 [53]. Myoglobin is an oxygen-binding protein present mostly in the muscle tissue of vertebrates. It is similar to hemoglobin, which is present in red blood cells. Interestingly, when the myoglobin gene is knocked out of mice, they remain viable because hemoglobin acts as a substitute[54].

The heme group that carries the oxygen-binding iron is not covalently bound to the protein. Apomyoglobin is a form of myoglobin without the heme group bound to it. Some of the earliest microsecond temperature jump experiments were conducted on apomyoglobin [55].

Fig. 8 shows what happens to the folded and unfolded protein population as a function of temperature. The protein is most stable around body temperature, but it unfolds below 0 °C and above 60 °C. Thus a temperature jump experiments from -10 to 10 °C (starting in super cooled water) causes refolding from 20% to 85% native protein[56]. The early ($\approx 10 \mu\text{s}$) stage of collapse into a compact state was revealed when tryptophan was used to monitor folding. Apomyoglobin is very sensitive to lower temperatures even in 3 M glycerol solution, which is usually considered a cryo-protectant (Fig. 8).

Apomyoglobin is quite different from NTL9: its folding steps are far from two-state, being separated by orders of magnitude in time (10 μs to 1 s)[57]. The reason is that the heme-binding part of the protein is quite unstable without the heme and folds very slowly, whereas the rest folds very efficiently. A recent study has shown that apomyoglobin can be tweaked to be much more stable and fold much faster overall by filling up its heme pocket with bulky amino acids. This highlights the important point that function of a protein can frustrate its folding [58]. The 'new' apomyoglobin can't bind heme anymore, but it is even more stable than myoglobin and folds faster.

c. WW Domain (a model for fast refolding after pressure unfolding)

WW domains are a family of fast folding protein modules with three anti-parallel beta sheet structure (Fig. 9C, "F" state). The name came along due to the presence of two highly conserved tryptophan amino acids in these small 30–40 residues domains. WW is a binding module involved in apoptosis, among other functions. WW domain's binding to a target protein is mediated by recognition of a proline rich region, which latches onto its loop 1 and hydrophobic pocket to facilitate the binding process. These versatile domains are also involved in transcriptional regulation [59].

Pressure is a thermodynamic variable that can be simulated easily by MD in order to understand protein folding kinetics. One either shrinks the simulation box, or applies a "barostat" on the computer to crank up the pressure. The recent development of a microsecond pressure jump experiment has made it possible to compare experiments with simulations [31,60].

The major difference between the two most common thermodynamics perturbations is that pressure is a mild perturbant as compared to temperature. Pressure unfolded proteins tend to have some residual secondary structure which may not be there in the high temperature counterpart [61,62]. It is worth noting that in Fig. 2, the pressure and temperature-unfolded proteins are all in one zone outside the ellipse of folded proteins. There is evidence for some proteins that the zone outside the ellipse is just a single state [63] with gradually changing structure as temperature and pressure are changed. It is not clear at present whether that is true for most or all proteins. Distinct unfolded states with clear transitions between them may actually exist in the zone outside the ellipse.

Fast downward pressure jump experiments were conducted on Fip35, a WW domain engineered to fold fast and derived from human Pin1 WW domain [64]. The protein refolds in two distinct steps, both of which are on the order of several microseconds. In order to determine the mechanistic details of the process, pressure jump simulations were performed to completely refold the protein following pressure unfolding. The simulations yielded both experimental time scales to near-quantitative accuracy without any adjustments of force field parameters [60]. While this may be unusually lucky, it shows that MD simulations can capture even fine details of the folding mechanism of such model proteins. Analyses based on the coupling of experimental and simulated data reveal the origin of the two different folding steps: The two-step process is due to two intermediates that can be formed in parallel: either beta strands 1+2 form a turn first, beta strands 2+3 do it first (Fig. 9AB). A trap where beta strands 1+3 pair up first and have to un-pair again was also observed. Thus the protein tries out all the possible antiparallel arrangements of the three beta sheets before settling into the native structure, which has the lowest free energy. Although loop 1 is usually the rate limiting steps, the experiment/theory study shows that other combinations are also possible. It is fascinating that how such a simplistic folding model can reveal the subtle details and multiple pathways that the protein takes to reach the native folded form (Fig. 9C), as envisioned by the funneled energy landscape theory [41,65].

The main conclusion from these studies is that pressure as a mild perturbant makes it possible to trace intermediate structures which might provide better insights about the mechanism of folding and on the quality of force fields used in MD simulations. It has by now been shown [66,60] that proteins can be refolded completely after pressure jumps on the computer, joining equilibrium [24,67,68] and temperature jump simulations [69,70].

d. Ubiquitin (a model for proteins under tension)

Finally, we move to the bottom of the phase diagram in Fig. 2, where the pressure is negative, that is, hydrostatic tension is applied to the protein. Strictly speaking, a liquid under tension is in a thermodynamically metastable state, as nature abhors a vacuum. Our example will be ubiquitin.

The aptly named ubiquitin is present throughout the eukaryotic cell and exists as both monomer and polymer, with functions ranging from proteolysis and cell cycle regulation to embryogenesis [71–77]. Depending upon the number of ubiquitin attached to a target protein, its specific fate will be decided (Fig. 10). Mono-ubiquitination is known to have connections to membrane protein endocytosis pathways, and a role in regulation of cytosolic protein localization [72]. On the other hand poly-ubiquitination plays a vital part in degradation of misfolded or misbehaved proteins by the proteasome [78–80]. Poly ubiquitin acts as a tag that signals the proteasome to chop up the protein to which it has been attached[80].

Tension can be attained in bulk aqueous solution by using Berthelot's method [82]. Briefly, the protein solution is put in a sealed tube leaving a small solvent vapor bubble. When the tube is heated, the bubble disappears because the liquid expands faster than the tube. When the tube is cooled, the bubble does not reappear because the liquid sticks to the tube walls. The liquid is now under tension at a density lower than its normal equilibrium state. For

example, water can be stretched to densities well below 1 g/cm^3 , corresponding to a pressure below -1000 atm [83]. Eventually, if the tube is cooled too much, the liquid forms a bubble again and the pressure returns to ambient pressure.

In a study of ubiquitin under tension, proton NMR experiments were conducted at 1 bar and under tension to monitor the thermal stability of ubiquitin. By monitoring the relative concentrations of unfolded and folded ubiquitin by the histidine NMR method illustrated in Fig. 5, ubiquitin was shown to be destabilized under tension as shown by the peak ratios in Figure 10. Panel A shows the protein at ‘physiological’ pressure, where the fraction of folded protein is just slightly larger than 50%. In panel B under a hydrostatic tension of -140 atm , the fraction of folded protein is less than 50% under otherwise identical conditions.

These experimental observations were again supported by MD simulations, which showed that the free energy of the folded protein indeed switches from negative to positive as tension is applied, causing the protein to unfold [18]. Strangely, the outcome of the P-T phase diagram obtained from the MD simulations suggests that at extremely high tension, there exists an ‘island of stability’ where the free energy turns over towards negative values again. Presumably low density cavities in the solvent water under tension favor folded protein inside of them.

In summary, all areas of the phase diagram have been explored in studies of small and fast folding proteins. It is now possible to complement the experiments with computer simulations and make a stronger connection to residual structure in the unfolded state, structure of intermediates that form during folding, and structural fluctuations of the native state that may be functionally relevant.

5. Folding in the living cell

Cells have many ways in which they can potentially modulate the function of their proteins: through suppression of transcription [84], interference with mRNA [85], different post-translational modifications [86,87], as well as protein transport, storage and degradation [88–92]. Often neglected is the “quinary structure” [13] discussed in the introduction: The cell is a highly crowded environment, and crowders exclude volume but can also interact with proteins (crowders include tRNA, other proteins, osmolytes, complex carbohydrates, microtubules and other large organelles for example). The cell is also a spatially and temporally heterogeneous environment (e.g. different compartments, cell-cycle variations). Finally, cells are subject to variations in temperature, pressure (osmotic or hydrostatic), and solute concentrations. All of these effects together can alter protein stability and function, and could be used by cells to control protein stability and function in more subtle ways than protein synthesis and degradation [93].

An important consequence of the relatively low stability of proteins is that they can unfold and then refold many times inside a cell, not just fold once when they come out of the ribosome [94,95]. A typical protein has a folding equilibrium constant in the range of only 100–10,000, and is therefore unfolded 0.1–1% of the time. A protein thus can spend minutes of a typical weeklong functional lifetime folding and unfolding over and over again inside

the cell. Transient unfolding of cellular proteins is not necessarily a bad thing: it provides the cell an additional regulatory handle for protein degradation.

The inside of the cell is an environment that is highly evolved to facilitate folding, compared to aqueous buffer *in vitro*. Unfolding-refolding experiments inside cells have shown much higher refolding yields than *in vitro* for some proteins [96], and chaperones could play a role in facilitating quick refolding when a protein has become unfolded [97]. Chaperones come at a massive ATP cost, and the interior of the cell may also be conducive to refolding in more passive and energy-efficient ways (e.g. osmolytes such as TMAO), for example.

It is now becoming possible to analyze protein folding inside cells subject to stress, such as osmotic pressure [98]. Single molecule fast protein folding experiments in cells have also become possible thanks to dye labels with high performance in the red or near-infrared, where auto-fluorescence of natural cell components interferes minimally with detection [99]. Protein stability and folding kinetics can even be studied as a function of cell cycle [100].

An example is shown in Fig. 12 [100]. The protein PGK, which has multiple folding steps ranging from 20 microseconds to minutes, is an enzyme that catalyzes ATP formation. When labeled with fluorescent proteins for in-cell detection, a destabilized mutant of PGK shows a clear unfolding signal when the temperature is raised above 40 °C. The signal depends on the cell cycle: when the cell is dividing (red curve in Fig. 12) the cytoplasmic environment somehow changes to make the protein several °C more stable. During interphase (the normal metabolic state), PGK is less stable. At just the right temperature (≈ 40 °C), PGK will unfold and refold as a function of cell cycle even at constant temperature (black arrow). For this enzyme, this is not an issue (the wild type is quite a bit more stable than the mutant), but for signaling or cell control proteins that have partly or fully disordered ‘structure’ [101], such folding/unfolding could lead to timing control over protein function.

In that spirit, we remind the reader of Fig. 1: RfaH C-terminal domain completely changes its secondary structure when crowded by its adjacent domain. And the protein lymphotactin is a GPCR-binding chemokine that rearranges into a glycosaminoglycan binder when the temperature is tuned across 37 °C [102]. These are examples where small perturbations completely reshape a protein’s structure and therefore function. The cell exerts such effects, ranging from the subtle to these two obvious examples, on its whole proteome.

Outlook

The fate of proteins is closely related to our health, and that of all animals and plants. While some malfunctions, such as congenital mutants that turn off or hyper-activate protein function, are very obvious, others may be caused by a much more subtle loss of control of the cell over its proteome. Fast-folding proteins are very useful probes of such subtle malfunction during folding because experiments can be compared directly with all-atom simulations. With single molecule experiments inside cells now pressing into the sub-millisecond range, such experiments can also be extended to the cell. The good news for simulation is that all-atom simulations are not necessarily much more expensive for the cellular milieu than for aqueous solution: representing the other cellular components that influence a protein via “quinary interactions” does not require more atoms than representing

the ubiquitous water molecules, but it results in a much richer and more complex system than a simple aqueous buffer.

Acknowledgments

This work was supported by grants from the NSF (MCB 1413256) and from the NIH (2R01 GM093318). The authors wish to thank Dr. Shahar Sukenik for a critical reading and helpful comments on the manuscript.

References

- Collins MO, Yu L, Campuzano I, Grant SG, Choudhary JS. Phosphoproteomic analysis of the mouse brain cytosol reveals a predominance of protein phosphorylation in regions of intrinsic sequence disorder. *Molecular & Cellular Proteomics*. 2008; 7:1331–1348. [PubMed: 18388127]
- Enright MB, Leitner DM. Mass fractal dimension and the compactness of proteins. *Physical Review E*. 2005; 71:011912.
- Liang J, Dill KA. Are proteins well-packed? *Biophysical Journal*. 2001; 81:751–766. [PubMed: 11463623]
- Iakoucheva LM, Brown CJ, Lawson JD, Obradovi Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. *Journal of molecular biology*. 2002; 323:573–584. [PubMed: 12381310]
- Meister K, Ebbinghaus S, Xu Y, Duman JG, DeVries A, Gruebele M, Leitner DM, Havenith M. Long-range protein–water dynamics in hyperactive insect antifreeze proteins. *Proceedings of the National Academy of Sciences*. 2013; 110:1617–1622.
- Urayama P, Phillips GN Jr, Gruner SM. Probing Substates in Sperm Whale Myoglobin Using High-Pressure Crystallography. *Structure*. 2002; 10:51–60. [PubMed: 11796110]
- Lu H, Schulten K. The Key Event in Force-Induced Unfolding of Titin's Immunoglobulin Domains. *Biophysical Journal*. 2000; 79:51–65. [PubMed: 10866937]
- Mason OU, Nakagawa T, Rosner M, Van Nostrand JD, Zhou J, Maruyama A, Fisk MR, Giovannoni SJ. First Investigation of the Microbiology of the Deepest Layer of Ocean Crust. *PLoS ONE*. 2010; 5:e15399. [PubMed: 21079766]
- Yancey PH, Geringer ME, Drazen JC, Rowden AA, Jamieson A. Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *Proceedings of the National Academy of Sciences*. 2014; 111:4461–4465.
- Lai Z, Colón W, Kelly JW. The acid-mediated denaturation pathway of transthyretin yields a conformational intermediate that can self-assemble into amyloid. *Biochemistry*. 1996; 35:6470–6482. [PubMed: 8639594]
- Mileyko Y, Joh RI, Weitz JS. Small-scale copy number variation and large-scale changes in gene expression. *Proceedings of the National Academy of Sciences*. 2008; 105:16659–16664.
- Gelman H, Platkov M, Gruebele M. Rapid Perturbation of Free-Energy Landscapes: From In Vitro to In Vivo. *Chemistry-a European Journal*. 2012; 18:6420–6427.
- McConkey EH. Molecular evolution, intracellular organization, and the quinary structure of proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 1982; 79:3236–3240. [PubMed: 6954476]
- Tuinstra RL, Peterson FC, Kutlesa S, Elgin ES, Kron MA, Volkman BF. Interconversion between two unrelated protein folds in the lymphotactin native state. *Proceedings of the National Academy of Sciences*. 2008; 105:5057–5062.
- Burmann BM, Knauer SH, Sevostyanova A, Schweimer K, Mooney RA, Landick R, Artsimovitch I, Röscher P. An α helix to β barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell*. 2012; 150:291–303. [PubMed: 22817892]
- Prigozhin MB, Gruebele M. Microsecond folding experiments and simulations: a match is made. *Physical chemistry chemical physics : PCCP*. 2013; 15:3372–3388. [PubMed: 23361200]

17. Yuan TZ, Ormonde CFG, Kudlacek ST, Kunche S, Smith JN, Brown WA, Pugliese KM, Olsen TJ, et al. Shear-Stress-Mediated Refolding of Proteins from Aggregates and Inclusion Bodies. *Chembiochem*. 2015; 16:393–396. [PubMed: 25620679]
18. Larios E, Gruebele M. Protein stability at negative pressure. *Methods*. 2010; 52:51–56. [PubMed: 20433926]
19. Haymet ADJ, Silverstein KAT, Dill KA. Hydrophobicity Reinterpreted As Minimisation Of the Entropy Penalty Of Solvation. *Farad Disc*. 1996; 103:117–124.
20. Southall NT, Dill KA, Haymet A. A view of the hydrophobic effect. *The Journal of Physical Chemistry B*. 2002; 106:521–533.
21. Hawley SA. Reversible pressure-temperature unfolding of chymotrypsinogen. *Biochemistry*. 1971; 10:2436–2442. [PubMed: 5557794]
22. Dadarlat VM, Post CB. Adhesive-cohesive model for protein compressibility: An alternative perspective on stability. *Proceedings of the National Academy of Sciences of the United States of America*. 2003; 100:14778–14783. [PubMed: 14638940]
23. Rief M, Gautel M, Oesterhelt F, Fernandez JM, Gaub HE. Reversible unfolding of individual titin immunoglobulin domains by AFM. *science*. 1997; 276:1109–1112. [PubMed: 9148804]
24. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How Fast-Folding Proteins Fold. *Science*. 2011; 334:517–520. [PubMed: 22034434]
25. Piana S, Sarkar K, Lindorff-Larsen K, Guo MH, Gruebele M, Shaw DE. Computational Design and Experimental Testing of the Fastest-Folding beta-Sheet Protein. *Journal of Molecular Biology*. 2011; 405:43–48. [PubMed: 20974152]
26. Muschol M, Rosenberger F. Liquid-liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *The Journal of Chemical Physics*. 1997; 107:1953–1962.
27. Li P, Banjade S, Cheng H-C, Kim S, Chen B, Guo L, Llaguno M, Hollingsworth JV, et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature*. 2012; 483:336–340. [PubMed: 22398450]
28. Weber Stephanie C, Brangwynne Clifford P Getting RNA and Protein in Phase. *Cell*. 149:1188–1191. [PubMed: 22682242]
29. Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, Gharakhani J, Jülicher F, Hyman AA. Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. *Science*. 2009; 324:1729–1732. [PubMed: 19460965]
30. Ballew RM, Sabelko J, Reiner C, Gruebele M. A single-sweep, nanosecond time resolution laser temperature-jump apparatus. *Review of Scientific Instruments*. 1996; 67:3694–3699.
31. Dumont C, Emilsson T, Gruebele M. Reaching the protein folding speed limit with large, sub-microsecond pressure jumps. *Nat Meth*. 2009; 6:515–519.
32. Schuler B, Lipman EA, Eaton WA. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*. 2002; 419:743–747. [PubMed: 12384704]
33. Dyer RB, Gai F, Woodruff WH, Gilmanshin R, Callender RH. Infrared Studies of Fast Events in Protein Folding. *Acc Chem Res*. 1998; 31:709–716.
34. Prigozhin MB, Gruebele M. The Fast and the Slow: Folding and Trapping of λ (6–85). *Journal of the American Chemical Society*. 2011; 133:19338–19341. [PubMed: 22066714]
35. Abe H, Go N. Noninteracting Local-structure Model Folding and Unfolding Transition in Globular Proteins. II. Application to Two-dimensional Lattice Proteins. *Biopolymers*. 1981; 20:1013–1031. [PubMed: 7225529]
36. Onsager L. Reciprocal Relations in Irreversible Processes. I. *Physical Review*. 1931; 37:405–426.
37. Deniz AA, Laurence TA, Belligere GS, Dahan M, Martin AB, Chemla DS, Dawson PE, Schultz PG, et al. Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc Natl Acad Sci USA*. 2000; 97:5179–5184. [PubMed: 10792044]
38. Huang GS, Oas TG. Submillisecond folding of monomeric λ repressor. *Proceedings of the National Academy of Sciences of the United States of America*. 1995; 92:6878–6882. [PubMed: 7624336]

39. Benjamin Schuler EAL, Willam A, Eaton. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature*. 2002; 419:4.
40. Frauenfelder H, Sligar SG, Wolynes PG. The Energy Landscapes and Motions of Proteins. *Science*. 1991; 254:1598–1603. [PubMed: 1749933]
41. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. *Proteins: Struct, Funct, Genet*. 1995; 21:167–195. [PubMed: 7784423]
42. Onuchic JN, Socci ND, Luthey-Schulten Z, Wolynes PG. Protein Folding Funnels: the Nature of the Transition State Ensemble. *Folding Des*. 1996; 1:441–450.
43. Bohr HG, Wolynes PG. Initial Events of Protein Folding From an Information-processing Viewpoint. *Phys Rev A*. 1992; 46:5242–5248. [PubMed: 9908746]
44. Naganathan AN, Perez-Jimenez R, Sanchez-Ruiz JM, Munoz V. Robustness of downhill folding: Guidelines for the analysis of equilibrium folding experiments on small proteins. *Biochemistry*. 2005; 44:7435–7449. [PubMed: 15895987]
45. Yang WY, Gruebele M. Folding lambda-repressor at its speed limit. *Biophysical Journal*. 2004; 87:596–608. [PubMed: 15240492]
46. Garcia-Mira MM, Sadqi M, Fischer N, Sanchez-Ruiz JM, Munoz V. Experimental identification of downhill protein folding. *Science*. 2002; 298:2191–2195. [PubMed: 12481137]
47. Cho SS, Weinkam P, Wolynes PG. Origin of barriers and barrierless folding in BBL. *Proc Nat Acad Sci USA*. 2008; 105:118–123. [PubMed: 18172203]
48. Yang WY, Pitera J, Swope W, Gruebele M. Heterogeneous folding of the trpzip hairpin: full atom simulation and experiment. *J Mol Biol*. 2004; 336:241–251. [PubMed: 14741219]
49. Baiz CR, Lin YS, Peng CS, Beauchamp KA, Voelz VA, Pande VS, Tokmakoff A. A Molecular Interpretation of 2D IR Protein Folding Experiments with Markov State Models. *Biophysical Journal*. 2014; 106:1359–1370. [PubMed: 24655511]
50. Hye-Kyung Y, Sang-Gyu K, Sun-Young K, Chung-Mo P. Regulation of Leaf Senescence by NTL9-mediated Osmotic Stress Signaling in Arabidopsis. *Mol Cells*. 2008; 25:438–445. [PubMed: 18443413]
51. Cho JH, Meng W, Sato S, Kim EY, Schindelin H, Raleigh DP. Energetically significant networks of coupled interactions within an unfolded protein. *Proc Natl Acad Sci U S A*. 2014; 111:12079–12084. [PubMed: 25099351]
52. Horng JC, Moroz V, Raleigh DP. Rapid cooperative two-state folding of a miniature alpha-beta protein and design of a thermostable variant. *Journal of Molecular Biology*. 2003; 326:1261–1270. [PubMed: 12589767]
53. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-ray Analysis. *Nature*. 1958; 181:662–666. [PubMed: 13517261]
54. Garry DJ, Ordway GA, Lorenz JN, Radford NB, Chin ER, Grange RW, Bassel-Duby R, Williams RS. Mice without myoglobin. *Nature*. 1998; 395:905–908. [PubMed: 9804424]
55. Ballew RM, Sabelko J, Gruebele M. Observation of distinct nanosecond and microsecond protein folding events. *Nat Struct Biol*. 1996; 3:923–926. [PubMed: 8901868]
56. Ballew RM, Sabelko J, Gruebele M. Direct observation of fast protein folding: The initial collapse of apomyoglobin. *Proc Natl Acad Sci USA*. 1996; 93:5759–5764. [PubMed: 8650166]
57. Jennings P, Wright P. Formation of a Molten Globule Intermediate Early in the Kinetic Folding Pathway of Apomyoglobin. *Science*. 1993; 262:892–895. [PubMed: 8235610]
58. Goodman JS, Chao SH, Pogorelov TV, Gruebele M. Filling Up the Heme Pocket Stabilizes Apomyoglobin and Speeds Up Its Folding. *The Journal of Physical Chemistry B*. 2014; 118:6511–6518. [PubMed: 24456280]
59. Yagi R, Yagi L-F, Chen K, Shigesada Y, Murakami Y, Ito A. A WW domain-containing Yes-associated protein (YAP) is a novel transcriptional co-activator. *EMBO journal*. 1999; 18:2551–2562. [PubMed: 10228168]
60. Wirth AJ, Liu Y, Prigozhin MB, Schulten K, Gruebele M. Comparing Fast Pressure Jump and Temperature Jump Protein Folding Experiments and Simulations. *Journal of the American Chemical Society*. 2015

61. Royer CA. The nature of the transition state ensemble and the mechanisms of protein folding: A review. *Archives of Biochemistry and Biophysics*. 2008; 469:34–45. [PubMed: 17923105]
62. Vidugiris GJA, Markley JL, Royer CA. Evidence for a Molten Globule-like Transition State in Protein Folding from Determination of Activation Volumes. *Biochemistry*. 1995; 34:4909–4912. [PubMed: 7711012]
63. Yang WY, Gruebele M. Kinetic equivalence of the heat and cold structural transitions of lambda6–85. *Phil Trans Royal Soc London B*. 2005; 43:13018–13025.
64. Jäger M, Zhang Y, Bieschke J, Nguyen H, Dendle M, Bowman ME, Noel JP, Gruebele M, et al. Structure–function–folding relationship in a WW domain. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103:10648–10653. [PubMed: 16807295]
65. Brooks CLI, Gruebele M, Onuchic JN, Wolynes PG. *Frontiers in Science 1997 - Chemical Physics of Protein Folding*. *Proc Natl Acad Sci USA*. 1998; 95:11037–11038. [PubMed: 9736683]
66. Prigozhin MB, Liu Y, Wirth AJ, Kapoor S, Winter R, Schulten K, Gruebele M. Misplaced helix slows down ultrafast pressure-jump protein folding. *Proceedings of the National Academy of Sciences*. 2013; 110:8087–8092.
67. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science*. 2010; 330:341–346. [PubMed: 20947758]
68. Piana S, Lindorff-Larsen K, Shaw DE. Protein folding kinetics and thermodynamics from atomistic simulation. *Proc Natl Acad Sci U S A*. 2012
69. Prigozhin MB, Gruebele M. Microsecond folding experiments and simulations: a match is made. *Physical Chemistry Chemical Physics*. 2013; 15:3372–3388. [PubMed: 23361200]
70. Piana S, Sarkar K, Lindorff-Larsen K, Guo M, Gruebele M, Shaw DE. Computational design and experimental testing of the fastest-folding beta-sheet protein. *J Mol Biol*. 2011; 405:43–48. [PubMed: 20974152]
71. Hochstrasser M. Origin and function of ubiquitin-like proteins. *Nature*. 2009; 458:422–429. [PubMed: 19325621]
72. Galan JM, Haguenuer-Tsapis R. Ubiquitin lys63 is involved in ubiquitination of a yeast plasma membrane protein. *Embo j*. 1997; 16:5847–5854. [PubMed: 9312043]
73. Terrell J, Shih S, Dunn R, Hicke L. A function for monoubiquitination in the internalization of a G protein-coupled receptor. *Molecular cell*. 1998; 1:193–202. [PubMed: 9659916]
74. Hofmann RM, Pickart CM. Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell*. 1999; 96:645–653. [PubMed: 10089880]
75. Deng L, Wang C, Spencer E, Yang L, Braun A, You J, Slaughter C, Pickart C, et al. Activation of the I κ B kinase complex by TRAF6 requires a dimeric ubiquitin-conjugating enzyme complex and a unique polyubiquitin chain. *Cell*. 2000; 103:351–361. [PubMed: 11057907]
76. Spence J, Gali RR, Dittmar G, Sherman F, Karin M, Finley D. Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell*. 2000; 102:67–76. [PubMed: 10929714]
77. Hicke L. Protein regulation by monoubiquitin. *Nat Rev Mol Cell Biol*. 2001; 2:195–201. [PubMed: 11265249]
78. Clague MJ, Urbé S. Ubiquitin: Same Molecule, Different Degradation Pathways. *Cell*. 2010; 143:682–685. [PubMed: 21111229]
79. Ye Y, Rape M. Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol*. 2009; 10:755–764. [PubMed: 19851334]
80. Thrower JS, Hoffman L, Rechsteiner M, Pickart CM. Recognition of the polyubiquitin proteolytic signal. *Embo j*. 2000; 19:94–102. [PubMed: 10619848]
81. Vijay-Kumar S, Bugg CE, Cook WJ. Structure of ubiquitin refined at 1.8 Å resolution. *J Mol Biol*. 1987; 194:531–544. [PubMed: 3041007]
82. Berthelot M. Sur quelques phénomènes de dilatation forcée des liquides. *Ann Chim Phys*. 1850; 30:232–237.
83. Netz PA, Starr FW, Stanley HE, Barbosa MC. Static and dynamic properties of stretched water. *The Journal of Chemical Physics*. 2001; 115:344–348.

84. Bird AP, Wolffe AP. Methylation-induced repression—belts, braces, and chromatin. *Cell*. 1999; 99:451–454. [PubMed: 10589672]
85. Brummelkamp TR, Bernards R, Agami R. A system for stable expression of short interfering RNAs in mammalian cells. *science*. 2002; 296:550–553. [PubMed: 11910072]
86. Karve TM, Cheema AK. Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *Journal of Amino Acids* 2011. 2011
87. Deribe YL, Pawson T, Dikic I. Post-translational modifications in signal integration. *Nat Struct Mol Biol*. 2010; 17:666–672. [PubMed: 20495563]
88. Rapoport TA. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature*. 2007; 450:663–669. [PubMed: 18046402]
89. Rothman JE, Wieland FT. Protein sorting by transport vesicles. *Science*. 1996; 272:227–234. [PubMed: 8602507]
90. Rothman JE. Mechanisms of intracellular protein transport. *Nature*. 1994; 372:55–63. [PubMed: 7969419]
91. Jiang L, Phillips TE, Hamm CA, Drozdowicz YM, Rea PA, Maeshima M, Rogers SW, Rogers JC. The protein storage vacuole: a unique compound organelle. *The Journal of Cell Biology*. 2001; 155:991–1002. [PubMed: 11739409]
92. Varshavsky A. Discovery of cellular regulation by protein degradation. *Journal of Biological Chemistry*. 2008; 283:34469–34489. [PubMed: 18708349]
93. Wirth AJ, Gruebele M. Leaving the test-tube behind: Quinary protein structure and the consequences of crowding in living cells. *BioEssays*. 2013; 35:984–993. [PubMed: 23943406]
94. Bai Y, Englander JJ, Mayne L, Milne JS, Englander SW. Thermodynamic parameters from hydrogen exchange measurements. *Methods in enzymology*. 1995; 259:344–356. [PubMed: 8538461]
95. Bai Y, Englander SW. Future directions in folding: the multi-state nature of protein structure. *Proteins*. 1996; 24:145–151. [PubMed: 8820481]
96. Ebbinghaus S, Dhar A, McDonald JD, Gruebele M. Protein folding stability and dynamics imaged in a living cell. *Nat Meth*. 2010; 7:319–323.
97. Guo M, Gelman H, Gruebele M. Coupled Protein Diffusion and Folding in the Cell. *PLoS ONE*. 2014; 9:e113040. [PubMed: 25436502]
98. Dhar A, Girdhar K, Singh D, Gelman H, Ebbinghaus S, Gruebele M. Protein Stability and Folding Kinetics in the Nucleus and Endoplasmic Reticulum of Eucaryotic Cells. *Biophysical Journal*. 2011; 101:421–430. [PubMed: 21767495]
99. Borgia MB, Borgia A, Best RB, Steward A, Nettels D, Wunderlich B, Schuler B, Clarke J. Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins. *Nature*. 2011; 7353:662–665. [PubMed: 21623368]
100. Wirth AJ, Platkov M, Gruebele M. Temporal Variation of a Protein Folding Energy Landscape in the Cell. *Journal of the American Chemical Society*. 2013; 135:19215–19221. [PubMed: 24304346]
101. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: Introducing the D(2) concept. *Annual Review of Biophysics*, vil 37. *Annual Review of Biophysics*. Annual Reviews, Palo Alto. 2008; 37:215–246.
102. Volkman, BF.; Liu, TY.; Peterson, FC. Chapter 3 Lymphotactin Structural Dynamics. In: Tracy, MH.; Damon, JH., editors. *Methods in Enzymology*. Vol. 461. Academic Press; 2009. p. 51-70.doi:[http://dx.doi.org/10.1016/S0076-6879\(09\)05403-2](http://dx.doi.org/10.1016/S0076-6879(09)05403-2)

Fig. 1. Segment of the protein RfaH C-terminal domain. The segment completely changes secondary structure depending on whether it is separated from the adjacent domain, or crowded by it. Its function switches from transcription to translation factor.

Fig. 2.

The phase diagram of proteins as a function of temperature and pressure. Proteins are folded within the ellipse, where the four different proteins discussed in section 4 are shown in their folded state. Proteins become unfolded outside of the ellipse. At the boundary of the ellipse, the free energy of the folded state is the same as of the unfolded state ($\Delta G=0$). Four stresses are shown: Applying high pressure, applying tension (negative pressure), raising the temperature, and lowering the temperature. All four of these stresses cause proteins to unfold.

Fig. 3. Relaxation experiments. (A) Signal collected from protein responds to a jump up and down in temperature slowly (solid) or rapidly (dashed) as the protein first unfolds and then refolds (representative structures are shown). (B): Temperature jump in three steps: (1) heating laser 'jumps' sample; (2) probe laser excites protein; (3) fluorescence emission, infrared absorption or another signal is collected. (C) Simple instrumental setup for a pressure downward jump to refold a protein. In practice, mechanical or piezoelectric valves, or burst membranes can be used.

Fig 4.

Top: successive frames in time of a single proteins randomly switching between the unfolded and folded states. The protein is shown labeled with two fluorophores, one emitting green light, and the other red light. The green fluorophore is excited by a laser. When the protein is unfolded, mostly the green fluorophore emits green light. When the protein is folded, the green fluorophore can donate its energy to the red fluorophore, which then emits red light. The race at the bottom shows how the green and red fluorescence are anticorrelated as a result.

Fig. 5. Schematic of experimental data from Oas and coworkers [38], showing how folding kinetics can be measured at equilibrium by NMR spectroscopy. The faster interconversion between folded and unfolded protein shown at the top produces 'fat' peaks in the NMR spectrum, the slower interconversion shown at the bottom produces narrower peaks in the NMR spectrum.

Fig. 6.

(A) Long molecular dynamics trajectory of WW domain fast folding protein [24]. “RMSD” is the average root mean square deviation of alpha carbons in the protein from a folded reference structure. The protein switches between unfolded (large RMSD) and folded (small RMSD) states in a way similar to the single molecule experiment in Fig. 4. (B) The free energy landscape of the peptide trpzip2 at two temperatures. The free energy is plotted as a function of RMSD, which is close to 0 for the folded state, and larger for the unfolded state(s). At a high temperature the peptide is unfolded, and at a low temperature the folded peptide (all the way to the left) becomes stable. [48].

Fig. 7. Schematic showing the large 60S subunit of ribosome and L9 protein. The blow-up shows the truncated N terminal of the ribosomal protein L9 known as NTL9 (pdb id 2hbb) [51].

Fig. 8.
The thermal unfolding curve of apomyoglobin with (blue curve) and without (black curve) glycerol. Below 0 °C the protein unfolds, as well as above 60 °C [56].

Fig. 9.

Panel A and B are the MD simulation trajectories showing the formation of loop 1 first, or of loop 2 first. Panel C depicts the simplest possible mechanism of folding for Fip35 WW domain consistent with the simulation results. A trap with a loop between strands 1 and 3 can also form and is not shown [58].

Fig. 10.

Panels A and B shows the fate of the target protein if it is singly labeled or multiply labeled by ubiquitin. (PDB id = 1ubq) [81].

Fig. 10.
NMR Signal showing change in histidine proton peaks of unfolded (U) to folded (F) protein at 1 atm and under a tension of -140 atm.

Fig 12.

Thermal stability of the mutated version (Y122W W308F W333F) of ATP-synthesizing enzyme PGK from yeast. Melting curves are shown during two different parts of the cell cycle. The protein unfolds at several °C higher temperature in cells that divide. The black arrow shows that at just the right temperature, the cell cycle can thus cause a large switch of unfolded vs. folded protein concentration inside the cell.