Finding a protein's Achilles heel

Andreas Matouschek and Carlos Bustamante

Recent ensemble and single molecule manipulation studies highlight the mechanical nature of protein unfolding *in vivo* and, thus, the greater need to understand the response of proteins to mechanical force.

Most proteins have to fold into well-defined three-dimensional structures to be active. However, regulated unfolding is also a critical step in several cellular processes. Two clear-cut examples are protein translocation across some membranes and protein degradation by ATP-dependent proteases. In both cases, the requirement for unfolding is due to the dimensions of a proteinaceous channel through which the substrates have to travel to reach their destinations. During protein translocation across membranes, the channel leads to a subcellular compartment, such as the mitochondrial matrix, the lumen of chloroplasts or the endoplasmic reticulum. In the case of ATP-dependent proteases, the final destination is the proteolytic chamber buried deep inside the protease structure. The proteolytic sites in ATP-dependent proteases themselves show very little substrate specificity. Instead selectivity is achieved by controlling access to the active site of degradation. Also in both examples of cellular unfolding, translocation of the polypeptide chain is directly coupled to unfolding[1–3].

Protein folding and unfolding *in vitro* is generally studied by monitoring spectrophotometrically the structural transitions of a protein as it is unfolded by heat or chemical denaturants such as urea. Although these studies have provided much insight on the thermodynamics and kinetics of unfolding, many aspects of the unfolding process *in vivo* required designing specific experiments to follow the structural transitions in substrate proteins as they are unfolded during a particular cellular process. For example, unfolding during degradation by the ATP-dependent protease ClpAP could be measured by following the intrinsic fluorescence of green fluorescent protein[4]. In other more complex systems, such as unfolding during protein import into mitochondria, unfolding was measured indirectly by setting up reaction conditions such that unfolding was the rate-limiting step of the reaction[3,5]. Doing so made it possible to follow the unfolding of precursor proteins at the entrance of the import channel by measuring the accumulation of radioactive protein inside the mitochondria. Both types of studies showed that proteins can unfold during degradation and translocation many hundred times faster than they unfold in *in vitro* studies. Unfolding *in vivo* is catalyzed by specific unfoldases that are able to accelerate the unfolding reaction, apparently, by altering the unfolding pathway of their substrates[1,3]. These earlier studies also found that proteins differ greatly in their susceptibility to unraveling by the various unfoldases, and this susceptibility correlates neither with their thermodynamic nor their kinetic stability as determined by *in vitro* experiments. At the time it was speculated that the susceptibility to unfolding is instead determined by local stability of the substrate proteins immediately adjacent to the sequence that target them to the unfoldases[3].

A recent paper[6] confirms that there is no correlation between protein stability and susceptibility to unfolding by the ATP-dependent protease ClpXP. A systematic study of 13 substrates revealed examples where a protein that is several thousand times more stable than another is nonetheless degraded faster. What determines the susceptibility of proteins to unfolding by ClpXP, if it is not the thermodynamic stability of its substrate, as measured in solvent denaturation experiments? Kenniston *et al.*[6] show that, instead, the unfolding susceptibility of the substrates depends on the local structure and stability of the region immediately adjacent to the tag that targets these proteins to ClpXP. Using an immunoglobulin domain as the test protein, they find that only mutations next to the degradation tag influence unfolding by ClpXP. Mutations elsewhere have no effect. Also, unfolding is not merely induced by the binding of substrate to ClpXP because folded and denatured proteins bind the protease with the same affinity. This finding strongly supports the model in which ClpXP catalyzes unfolding by pulling mechanically on its substrates.

> In mechanical unfolding, local stability and interactions are more important than their global counterparts.

There is an alternative way of measuring protein stability *in vitro*. The last decade has seen enormous progress in the development of single molecule manipulation techniques that make it possible to characterize the physical properties of individual protein molecules one at a time. For example, optical tweezers have been used to measure the forces developed by biological motors such as myosins[7], kinesins[8], or polymerases[9–11]. Optical tweezers[12,13] and atomic force microscopy[14,15] (AFM) have also been used to determine the mechanical properties of proteins and the forces that hold together the molecules themselves. These methods were first applied to mechanically unfold the protein titin, responsible for the generation of passive force in muscle. These and more recent studies[16] clearly suggest that part of the mechanism that confers elasticity to titin fibers is through the reversible unfolding of single domains in titin.

Mechanical denaturation of proteins displays some of the same characteristics observed in protein unraveling by cellular unfoldases. Thus, susceptibility to unfolding by AFM does not appear to correlate with the thermodynamic stability of the protein[17,18]. For example, the protein barnase is

The authors are in the Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208 (A.M.) and the Department of Physics, Department of Molecular and Cell Biology, and Howard Hughes Medical Institute, University of California, Berkeley, California 94720 (C.B.).
e-mail: matouschek@northwestern.edu

very stable against unfolding by solvent or heat but is unfolded by very low pulling forces[19] and shows little resistance against unfolding by some of the cellular machines that have been analyzed[1,3]. As demonstrated in two papers on pages 731 and 738 of this issue of *Nature Structural Biology*, protein resistance to mechanical unfolding strongly depends on the direction of pulling. These two groups were able to control the direction of pulling by selecting specific points of attachment on the proteins using two different biochemical tricks. Fernandez and colleagues[20] used naturally occurring polyubiquitin chains in which ubiquitin molecules are linked through isopeptide bonds between the C-terminal carboxy group of one ubiquitin and the ε-amino group of Lys48 in the other ubiquitin. They compared unfolding of chains linked in this manner to polyubiquitin in which the molecules are simply linked N to C terminus. Radford and co-workers[21] constructed concatemers that consisted of four titin immunoglobulin domain 27 and a domain derived from *E. coli* dihydrolipoyl acetyltransferase. This domain can be covalently modified through the attachment of lipoic acid to the ε-amino group of Lys41. This construct could then be attached to a gold surface in two different geometries, through cysteine residues at the C terminus of the construct or through the dithiolane moiety of the lipoic acid group at Lys41.

Perhaps this dependence on pulling direction is to be expected on purely physical grounds. Proteins are not isotropic materials, and as such their mechanical properties (compressibility and Young's modulus, for example) should depend on the direction and orientation of the forces acting on them. Yet, it is also known that not all residues of proteins contribute equally to the thermodynamic stability of globular proteins and that most of the stability of the protein structure is provided by its hydrophobic core[22,23]. In addition, more often than not, thermal or chemical protein denaturation is a highly cooperative process. Thus, one could have expected that mechanical unfolding of a protein would be relatively insensitive to the pulling direction—mechanical unfolding would ensue only after enough of the nucleus responsible for maintaining the stability of the molecule has been undone by the external force, irrespectively of the pulling direction. This turns out not to be the case. As these papers demonstrate, the direction in which forces are applied to proteins strongly affects their resistance against unfolding. When acted upon by forces, proteins do posses an Achilles' heel. In mechanical unfolding, local stability and interactions are more important than their global counterparts. (A trivial analogy would be a Velcro flap. It opens easily when the two surfaces are peeled apart using forces applied perpendicular to them, but not when pulled along the axis of the strips.)

Despite the similarities between mechanical unfolding *in vitro* and *in vivo*, the processes are not identical. Unfolding in AFM experiments occurs through stretching the substrate protein from both ends of the polypeptide chain, whereas proteases or import machineries would act by pulling their substrate from one end against the entrance to a translocation channel. Although hydrolysis of peptide bonds is thermodynamically favored and can occur rapidly when catalyzed by many simple proteases, efficient degradation by ClpXP requires ATP hydrolysis. ATP powers the unfolding of protein substrates and their translocation from the substrate-binding site at the surface of the protease to the proteolytic site buried deep inside its structure. Depending on the substrate, either one of these two steps can be rate determining. Kenniston *et al.*[6] exploited this fact to determine the ATP-hydrolysis rates during protein unfolding and translocation. Their analysis of ATP hydrolysis during degradation revealed yet another important aspect of the mechanical unfolding by cellular machines. Unlike AFM experiments, where the force acting on the molecules is applied continuously, the ClpXP protease appears to apply mechanical force iteratively to its substrates by tugging at them repeatedly. ClpXP operates at two hydrolysis rates: a fast rate during degradation of unfolded proteins and a slow rate during degradation of native proteins. The ATP hydrolysis rate for the degradation of native proteins did not change with the global stability of the substrate. For hard-to-unfold substrates, the protease would simply engage the protein for longer times at the same slow hydrolysis rate. In this manner, ATP hydrolysis and protein unfolding are directly but loosely coupled and the amount of ATP hydrolyzed per molecule of protein degraded depends on the local stability of the substrate. As described above, this local stability depends on the direction and the location from which the protease pulls on its substrate. In a way, the proteases function like a car engine. The engine can be fully engaged, in which case the motor burns ATP at a high rate and translocates the polypeptide chain at full speed. When the engine runs into an obstacle, a folded domain, the clutch is iteratively applied and released. ATP is then burned at a slower rate while the engine keeps tugging at the polypeptide chain until the substrate unravels. The total amount of ATP consumed in the process is considerable. For unfolded substrates, about one ATP is hydrolyzed per amino acid and for folded proteins it can be an order of magnitude more. The result is that the protease can deal with tightly folded domains simply by tugging at them for longer times.

This strategy makes perfect physical sense. When a force is applied to a chemical bond, its effect is to lower the activation energy associated with the kinetic barrier to bond breaking and to make the broken bond the thermodynamically preferred state. But the crossing of the barrier and the breaking of a bond is itself always a thermally activated process. If the force applied to the bond is weak compared to the resistance of the bond, one may have to wait for a long time before thermal fluctuations carry the system over the still-high barrier. Alternatively, one can apply this comparatively weak force iteratively and repeatedly, hoping for the fluctuation event that will allow the system to cross the barrier at that low force. This seems to be the strategy utilized by the proteases.

Why doesn't the protease keep pulling at the substrate with a constant force instead of applying multiple tugs? We do not know the answer. Tugging at the substrate repeatedly allows the protease to reset itself at the end of each cycle and substrates that are too stable to be unfolded to be released. This way, the protease does not clog but remains available for other substrates[7].

Unfolding can be a committing step in two important biological processes, protein translocation across some membranes and protein degradation by ATP-dependent proteases. It appears that cells utilize differences in the susceptibility of folded domains to unfolding to control the specificity of these processes[24]. The papers discussed here bring us closer to understanding the many ways in which proteins respond to unfolding forces.

1. Huang, S., Ratliff, K.S., Schwartz, M.P., Spenner, J.M. & Matouschek, A. *Nat. Struct. Biol.* **6**, 1132–1138 (1999).
2. Reid, B.G., Fenton, W.A., Horwich, A.L. & Weber-Ban, E.U. *Proc. Natl. Acad. Sci. USA* **98**, 3768–3772 (2001).
3. Lee, C., Schwartz, M.P., Prakash, S., Iwakura, M. & Matouschek, A. *Mol. Cell* **7**, 627–637 (2001).
4. Weber-Ban, E.U., Reid, B.G., Miranker, A.D. & Horwich, A.L. *Nature* **401**, 90–93 (1999).
5. Matouschek, A. *et al. EMBO J.* **16**, 6727–6736 (1997).
6. Kenniston, J.A., Baker, T.A., Fernandez, J.M. & Sauer, R.T. *Cell* in the press (2003).
7. Finer, J.T., Simmons, R.M. & Spudich, J.A. *Nature* **368**, 113–119 (1994).

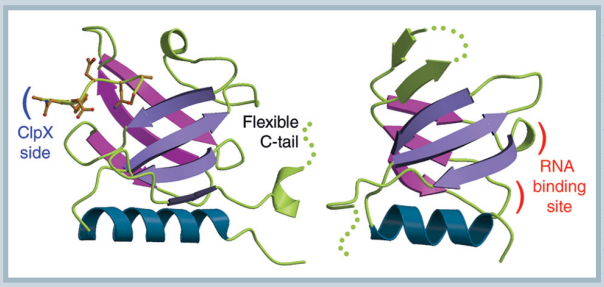
8. Block, S.M., Goldstein, L.S. & Schnapp, B.J. *Nature* **348**, 348–352 (1990).
9. Yin, H. *et al. Science* **270**, 1653–1657 (1995).
10. Wuite, G.J., Smith, S.B., Young, M., Keller, D. & Bustamante, C. *Nature* **404**, 103–106 (2000).
11. Davenport, R.J., Wuite, G.J., Landick, R. & Bustamante, C. *Science* **287**, 2497–2500 (2000).
12. Kellermayer, M.S.Z., Smith, S.B., Granzier, H.L. & Bustamante, C. *Science* **276**, 1112–1116 (1997).
13. Tskhovrebova, L., Trinick, J., Sleep, J.A. & Simmons, R.M. *Nature* **387**, 308–312 (1997).
14. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M. & Gaub, H.E. *Science* **276**, 1109–1112 (1997).
15. Carrion-Vazquez, M. *et al. Proc. Natl. Acad. Sci. USA* **96**, 3694–3699 (1999).
16. Kellermayer, M.S., Bustamante, C. & Granzier, H.L. *Biochim. Biophys. Acta* **1604**, 105–114 (2003).
17. Rief, M., Pascual, J., Saraste, M. & Gaub, H.E. *J. Mol. Biol.* **286**, 553–561 (1999).
18. Li, H., Carrion-Vazquez, M., Oberhauser, A.F., Marszalek, P.E. & Fernandez, J.M. *Nat. Struct. Biol.* **7**, 1117–1120 (2000).
19. Best, R.B., Li, B., Steward, A., Daggett, V. & Clarke, J. *Biophys. J.* **81**, 2344–2356 (2001).
20. Carrion-Vazquez, M. *et al. Nat. Struct. Biol.* **10**, 738–743 (2003).
21. Brockwell, D.J. *et al. Nat. Struct Biol.* **10**, 731–737 (2003).
22. Serrano, L., Kellis Jr., J.T., Cann, P., Matouschek, A. & Fersht, A.R. *J. Mol. Biol.* **224**, 783–804 (1992).
23. Eriksson, A.E. *et al. Science* **255**, 178–183 (1992).
24. Matouschek, A. *Curr. Opin. Struct. Biol.* **13**, 98–109 (2003).

Tag, you're degraded

Protein degradation in the cytosol is a regulated process. The destruction of unwanted polypeptides, whether folded or unfolded, is mostly performed by large proteolytic machinery, but how does it know which proteins must be annihilated? It identifies these substrates by the tags they carry.

Eukaryotes most often attach a 76-residue ubiquitin to the free amino groups on a protein, a modification that potentially marks the protein for destruction by the proteosome. In contrast, eubacteria add an 11-residue peptide, the SsrA tag, to the C terminus of the protein that directs the protein to specific multisubunit protease complexes, ClpXP or ClpAP. The tag, which is encoded by an RNA molecule with properties of both transfer and messenger RNA, is coupled to the protein when the ribosome stalls. Protein factors modulate the recognition of the SsrA tag by the appropriate protease complex. For example, in *Escherichia coli*, stringent starvation protein B (SspB) binds to the SsrA tag and enhances its recognition by ClpXP. In an effort to understand how the SspB binds the SsrA, Hyun Kyu Song and Michael Eck (*Mol. Cell* **12**, 75–86; 2003) determined the crystal structures of the protein alone and in complex with the 11-amino acid tag.

The most surprising aspect of the SspB structure (left ribbon diagram) is the topology. Despite the lack of sequence similarity to RNA binding proteins, the overall fold resembles that of a small nuclear ribonucleoprotein, Sm D2, and more distantly to a portion of the ribosome-associated protein L1. The structurally related regions are visually apparent when comparing the purple and pink strands and the blue helix of SspB (left) and Sm D2 (right). The structural similarity to RNA-binding proteins and the observation that SspB copurifies with ribosomes, suggests the possibility that SspB might recognize an RNA component of the ribosome. Comparison of SspB and Sm D2 structures suggests that recognition of the ribosomal RNA might occur through the β-turns located on the opposite side of the structure from the ClpX recognition face. In effect, SspB could link the protein synthesis and degradation machinery, and in so doing, this protein would promote the binding of ClpX to SsrA-tagged polypeptides before they are released into the cytoplasm to cause damage. Alternatively, if the SsrA and ribosomal RNA binding sites overlap, the interactions with the tagged polypeptide could release SspB from the ribosome. Identification of the exact site of RNA binding in SspB will be essential to determining if and how this protein binds to the ribosomal RNA.



Modified with permission from *Molecular Cell*.

The structure of the complex (left) of SspB with the SsrA tag (ball and stick) explains why SspB is crucial for the specific recognition of SsrA-marked proteins by the protease complex. ClpX recognizes the C-terminal three residues in the SsrA tag, which does not provide sufficient specificity for or result in efficient degradation of the tagged protein. In contrast, SspB binds the eight N-terminal amino acids of the tag, six of which make specific interactions with the protein. Hence, it is SspB that makes the necessary interactions with the SsrA-marked proteins that results in their degradation.

These structural results open the way to additional studies of how SspB associates with ClpX, as well as of the mechanisms by which SspB feeds the SsrA tag into the protease and if SspB, through a possible association with the ribosome, maintains the balance between protein degradation and synthesis.

Evelyn Jabri