the matrix of the mitochondrium may keep the translocating polypeptide in an unfolded conformation, interaction with Hsp70 in organelles is also essential for the translocation process (*12–14*). After insertion of the amino-terminal leader sequence across the mitochondrial membranes, the vectorial movement of the bulk of a polypeptide into the matrix requires the binding of mitochondrial Hsp70. A simple model of Hsp70 action in translocation shows that its binding provides a directionality to the translocation process by sterically preventing movement of the polypeptide back toward the cytoplasm. Continued movement into the matrix might be accomplished simply by Brownian motion, with backward movement prevented by the binding of Hsp70 to internal sites on the translocating polypeptide. Such a model would require that mitochondrial Hsp70 have a higher affinity for the translocating polypeptide than that of any cytosolic protein, such as cytosolic Hsp70. However, Hsp70s are thought to bind with particularly high affinity to hydrophobic patches exposed in unfolded polypeptides. Because the polypeptide crosses the membranes in an unfolded state (*15*), such hydrophobic sequences could easily bind to mitochondrial Hsp70.

The binding of Hsp70 to the nascent polypeptide chain can also have important consequences on the earliest events in protein maturation, namely translation (*16*). Yeast strains lacking a class of cytoplasmic Hsp70s that associate with translating ribosomes show defects in protein synthesis, including sensitivity to certain antibiotics that inhibit translation. The growth defect of such strains can be overcome by overproduction of a protein related to the translation elongation factor EF1a, the protein responsible for bringing the aminoacyl tRNA to the ribosome. The binding of Hsp70 to the nascent polypeptide chain may be important for the polypeptide's movement out of the tunnel of the 60S ribosomal subunit, in a way analogous to the binding of mitochondrial Hsp70 in translocation of a polypeptide through the membrane into the matrix. The lack of Hsp70 could result in the nascent chain backing up in the tunnel of the 60S subunit, thus distorting the interactions of peptidyl and aminoacyl tRNAs. The critical importance of chaperones in protein synthesis, translocation across membranes, and folding likely has ensured that the primary sequences of proteins have evolved to contain chaperone binding sites in addition to sequence information for attaining the proper tertiary conformation. Because chaperones can bind a variety of peptide sequences, the need for chaperone binding sites would not have unduly constrained the evolution of protein sequences.

The general outline of the interaction of chaperones with newly synthesized proteins is known. However, many questions remain as to the mechanism of action of these ubiquitous proteins and to the precise nature of their physiological roles in vivo. The answers will be unveiled by a combination of genetic analyses to uncover other components involved in these processes, coupled in vitro translation and folding systems to mimic the in vivo situation, and more sophisticated biochemical analyses of the chaperone-mediated protein folding.

References

1. C. J. Epstein, R. F. Goldberger, C. B. Anfinsen, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 439 (1963).
2. S. Lindquist and E. A. Craig, *Annu. Rev. Genet.* **22**, 316 (1988).
3. R. P. Beckmann, L. A. Mizzen, W. J. Welch, *Science* **248**, 850 (1990).
4. J. Ostermann *et al.*, *FEBS Lett* **277**, 281 (1990).
5. P. Scherer, U. Krieg, S. Hwang, D. Vestweber, G. Schatz, *EMBO J.* **9**, 4315 (1990).
6. U. Manning-Krieg, P. Scherer, G. Schatz, *ibid.* **10**, 3273 (1991).
7. M. Cheng *et al.*, *Nature* **337**, 620 (1989).
8. T. Langer *et al.*, *ibid.* **356**, 683 (1992).
9. E. A. Craig, B. D. Gambill, R. J. Nelson, *Microbiol. Revs.* **57**, 293 (1993).
10. U. Bertsch, J. Soll, R. Seetharam, P. Vitanein, *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 8696 (1992).
11. S. J. Landry, R. Jordan, R. McMacken, L. M. Gierasch, *Nature* **355**, 455 (1992).
12. P.J. Kang *et al.*, *ibid.* **348**, 137 (1990).
13. J. Vogel, L Misra, M. Rose, *J. Cell Biol.* **110**, 1885 (1990).
14. S. Sanders, K. Whitfield, J. Vogel, M. Rose, R. Schekman, *Cell* **69**, 353 (1992).
15. W-J. Chen and M. Douglas, *J. Biol. Chem.* **262**, 15605 (1987).
16. R. J. Nelson *et al.*, *Cell* **71**, 97 (1992).
17. R. J. Nelson and E. A. Craig, *Curr. Biol.* **2**, 487 (1992).

To Fold or Not to Fold ...

David A. Agard

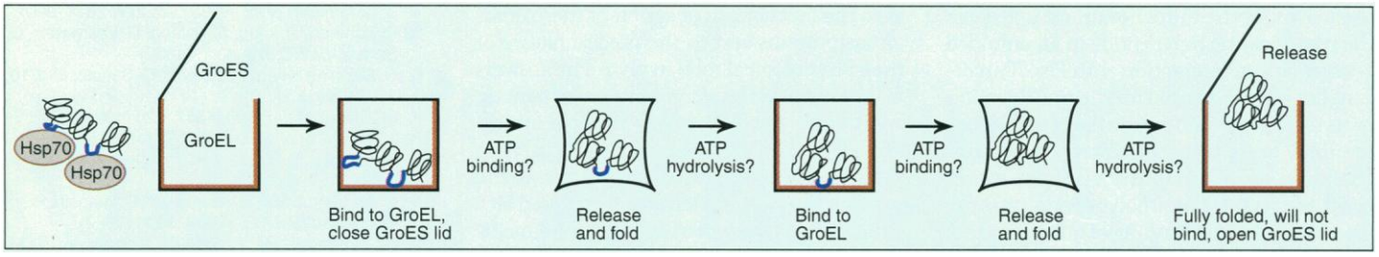
Molecular chaperones are cellular factors that shepherd newly synthesized proteins along the hazardous journey to the folded state. Their discovery has taken the study of protein folding from the arena of the biophysicist into the cell biological limelight. The results of both in vivo and in vitro studies suggest that these molecular chaperones are crucial for the folding and assembly of many multi-domain and multi-subunit proteins. In fact, it is through a combination of biophysical and cell biological approaches that the mechanism of action of the chaperones is now becoming clear.

Protein folding, especially in the extremely concentrated environment of the cell, must be thought of as a kinetic competition between on-pathway reactions leading to the folded state and off-pathway reactions leading to aggregation: $A \rightleftharpoons U \rightleftharpoons F$, where U is the unfolded protein, F is the folded protein, and A represents either reversible or irreversible aggregation. During the course of in vitro refolding reactions, transient states are formed that expose a substantial amount of hydrophobic surface. When such partially folded states are present at moderate concentrations, intermolecular aggregation can occur by way of these hydrophobic patches.

What strategies might the cell use to prevent aggregation? It could (i) block the exposed hydrophobic patches; (ii) create an isolated environment in which each protein molecule can fold independently; or (iii) sufficiently accelerate the on-pathway folding rate to kinetically out-compete aggregation. In fact, nature has exploited each of these approaches. The blockade of exposed hydrophobic patches requires a factor that can reversibly bind to them. The molecular chaperone Hsp70 and its relatives [Hsc70, DnaK, BiP, Kar2p, and Ssa1–4p (*1*)] meet this requirement and are ubiquitous monomeric proteins found in the cytosol, endoplasmic reticulum, and mitochondria. The Hsp70 class of chaperones binds to unfolded and partially folded states of a variety of proteins but shows little interaction with native, folded proteins (*2*). Binding of Hsp70 stimulates an endogenous ATPase, which causes the release of bound protein. How is such preferential binding of unfolded molecules accomplished? The Hsp70s recognize a feature common to the denatured, but not the native, states of proteins: an exposed, extended region of polypeptide rich in hydrophobic residues (*3–5*). This recognition of hydrophobic peptides, which is relatively sequence-independent, could be accomplished by way of a binding site similar to that of the peptide binding domain of the major histocompatibility complex class I molecule (*6*).

Although the binding of an Hsp70 chaperone may keep a partially folded protein from aggregating, the same interactions will also block folding because the hydrophobic regions must be buried in the fully folded molecule. Completion of folding requires the release of the protein from the Hsp70 chaperone, at which point a partially folded protein is again vulnerable to aggregation. Indeed, rather than promote folding to the native

The author is in the Department of Biochemistry and Biophysics and the Howard Hughes Medical Institute Structural Biology Unit, University of California, San Francisco, CA 91413-0448.



Model of protein folding.

state, Hsp70s appear primarily to prevent aggregation or premature folding until the substrate protein can assemble into the appropriate multi-subunit complex (*7*), be translocated across a membrane (*8*), or be passed on to a different chaperone, Hsp60 (*9*).

The completion of folding requires a factor that can sequester partially folded molecules from each other but still allow folding to proceed. This requirement is fulfilled by the Hsp60 class of chaperones, exemplified by the protein GroEL (*10, 11*). The Hsp60s are organized as large assemblies arranged in two rings, each composed of seven 60-kD subunits. Transiently associated with the Hsp60 proteins is the protein GroES, a single ring of seven 10-kD subunits (*12*). In vitro, the Hsp60 protein GroEL forms complexes with unfolded proteins and facilitates their folding in an ATP-dependent fashion. One, or at most two, protein molecule can bind to each GroEL (*13*) in or near the large central cavity of the 14-subunit protein (*14*). As with Hsp70, ATP hydrolysis promotes the release of bound protein. Some proteins, such as rubisco, only require GroEL and ATP hydrolysis to refold (*15*), whereas others such as rhodanese, which is extremely prone to aggregation, require both GroEL and GroES (*16*). The refolding of rhodanese requires the hydrolysis of ~130 molecules of ATP per rhodanese molecule (*16*), indicating that multiple cycles of binding and release occur during folding.

This information can be integrated into a simple functional model (see figure). The central cavity of GroEL could provide a "box" in which an individual polypeptide chain would be free to fold without risk of aggregation. Partially folded proteins, perhaps delivered to GroEL by the Hsp70 chaperones (*9*), would bind to the hydrophobic walls of the box by means of exposed hydrophobic patches. Binding stimulates ATP binding and hydrolysis, causing a conformational change in the box (*17*) and covering up the hydrophobic walls, thereby forcing release of the bound protein. After release, the protein is free to refold, and the GroEL reverts to its normal state of exposed hydrophobic surface. Once released from the walls, partially folded molecules could diffuse out of the box and perhaps aggregate. To prevent this process a "lid," potentially provided by GroES, is required. The binding of protein to the box causes the lid to close, allowing the protein to be refolded through many cycles of binding and release within the protected environment of the box. Once the protein is folded, the box must open to release the protein. A key aspect of this model is that the chaperone does not directly participate in the folding reaction but merely creates an environment of "infinite dilution" in which the protein is free to fold by itself.

A third way to facilitate folding—increasing the on-pathway folding rate—seems to be used by many bacterial and eukaryotic proteases that are synthesized as preproproteins (a hydrophobic signal sequence linked to a proregion linked to a mature protease domain). Although there are now many examples of both amino- and carboxyl-terminal pro regions, the best mechanistic studies have been those of α-lytic protease and subtilisin. These small, bacterial serine proteases have extended amino-terminal pro regions of 166 (*18*) and 77 (*19*) amino acids, respectively, that are not a part of the active proteases. However, these pro regions are required for the proper folding of the mature protease domains (*20–22*). In contrast to the chaperones, they do not suppress aggregation but directly facilitate folding. Indeed, folding intermediates of α-lytic protease (*21*) and subtilisin (*23*) have been isolated under nondenaturing conditions by the omission of the pro region from in vitro refolding reactions. The α-lytic protease intermediate is stable for weeks without aggregating or folding to the native state. The addition of the pro region at any time leads to rapid folding of the intermediate. Thus, the pro region is required for folding without aggregation or other off-pathway reactions. The α-lytic protease pro region directly stabilizes the folding transition state (*21*), thereby accelerating folding by a factor of at least 10^7. In contrast to the chaperones, which use ATP hydrolysis and can interact with a variety of substrates, the α-lytic protease pro region does not require nucleotide triphosphates to promote folding, is highly specific for its substrate, and, most important, directly increases the rate of the forward-folding reaction.

There is a large body of biophysical data on protein refolding in vitro (*24–26*). Although we now appreciate that folding within the cell is vastly more complicated than these simple refolding reactions, the mechanisms derived from in vitro studies are likely to be physiologically valid. Not only are the native states produced in vivo and in vitro identical for almost all proteins, but the importance of the chaperones suggests that the starting points for folding may also be similar. There had been concern that, in vivo, folding could start at the amino terminus and proceed as the molecule is synthesized but that, in vitro, only full-length molecules refolded. However, the binding of Hsp70 to nascent chains (*27*) and its slow release of bound proteins (*4*) indicate that in vivo the temporal order of synthesis is unlikely to correspond to the temporal order of folding.

Despite our limited understanding of protein folding, the exciting synergism that has developed between the cell biologist and the biophysicist will lead to rapid progress.

References

1. M.-J. Gething and J. Sambrook, *Nature* **355**, 33 (1992).
2. M.-J. Gething, K. McCammon, J. Sambrook, *Cell* **46**, 939 (1986).
3. M.-J. Gething, unpublished data.
4. G. C. Flynn, T. G. Chappell, J. E. Rothman, *Science* **245**, 385 (1989).
5. S. J. Landry, R. Jordan, R. McMacken, L. M. Gierasch, *Nature* **355**, 455 (1992).
6. M. F. Flajnik, C. Canel, J. Kramer, M. Kasahara, *Immunogenetics* **33**, 295 (1991).
7. E. Degen and D. B. Williams, *J. Cell Biol.* **112**, 1099 (1991).
8. J. P. Vogel, L. M. Misra, M. D. Rose, *ibid.* **110**, 1885 (1990).
9. T. Langer *et al.*, *Nature* **356**, 683 (1992).
10. S. M. Hemmingsen *et al.*, *ibid.* **333**, 330 (1988).
11. M. B. Yaffe *et al.*, *ibid.* **358**, 245 (1992).
12. C. Georgopoulos and D. Ang, *Semin. Cell Biol.* **1**, 19 (1990).
13. E. S. Bochkareva *et al.*, *J. Biol. Chem.* **267**, 6796 (1992).
14. T. Langer *et al.*, *EMBO J.* **11**, 4757 (1992).
15. P. Goloubinoff *et al.*, *Nature* **342**, 884 (1989).
16. J. Martin *et al.*, *ibid.* **352**, 36 (1991).
17. H. R. Sabil *et al.*, *Curr. Biol.* **3**, 265 (1993).
18. J. L. Silen *et al.*, *Gene* **69**, 237 (1988).
19. J. A. Wells *et al.*, *Nucleic Acids Res.* **11**, 7911 (1983).
20. J. L. Silen and D. A. Agard, *Nature* **341**, 462 (1989).
21. D. Baker, J. L. Sohl, D. A. Agard, *ibid.* **356**, 263 (1992).
22. X. Zhu *et al.*, *ibid.* **339**, 483 (1989).
23. J. Eder, M. Rheinnecker, A. R. Fersht, *Biochemistry* **32**, 18 (1993).
24. C. B. Anfinsen and H. A. Scheraga, *Adv. Protein Chem.* **29**, 205 (1975).
25. K. A. Dill, *Biochemistry* **29**, 7133 (1990).
26. P. S. Kim and R. L. Baldwin, *Annu. Rev. Biochem.* **59**, 631 (1990).
27. R. P. Beckmann, L. A. Mizzen, W. J. Welch, *Science* **248**, 850 (1990).