Assisted Protein Folding*

Raymond W. Ruddon‡§ and Elliott Bedows‡§

From the Eppley Institute for Research in Cancer and Allied Diseases, the Department of Pharmacology, and the Department of Biochemistry and Molecular Biology, University of Nebraska Medical Center, Omaha, Nebraska 68198-6805

Minireview

Historical Perspectives

The pioneering work of the late Christian Anfinsen and his colleagues (1) on the reoxidation of bovine pancreatic ribonuclease (RNase) to a native, biologically active enzyme in vitro after reduction of disulfide bridges and disruption of tertiary structure demonstrated that regeneration of native conformation of a purified protein can occur spontaneously in a test tube without the addition of any other co-factors or helper enzymes. This led to the still valid conclusion that “no special genetic information, beyond that contained in the amino acid sequence, is required for the proper folding of the molecule and for the formation of ‘correct’ disulfide bonds” (2).

Of course, the story of protein folding goes back much further (reviewed in Ref. 3). A number of milestones can be noted. In 1911, Chick and Martin found that proteins could be denatured in vitro, and they distinguished that process from aggregation of the protein. In 1929, Wu postulated that protein denaturation was an unfolding process and that native protein structures involved regular, repeated patterns of folding into a three-dimensional network. Anson and Mirsky in 1931 and Anson (1945) showed that hemoglobin folding is reversible and that hemoglobin could be renatured in vitro to a form that had a native-like absorption spectrum, oxygen binding, and tryptic digestion pattern. Studies in the 1950s by Eisenberg and Schwert and by Schellman demonstrated that denaturation and renaturation are thermodynamic processes, involving a change in free energy and large changes in conformation between the denatured and native states.

Even the early investigators realized that the protein folding processes that occurred in test tubes, although they could reconstruct native structure, were too slow to work inside cells. For example, even under optimized conditions of protein dilution, pH, and temperature, renaturation of RNase takes about 20 min (2), and RNase is a relatively simple monomeric protein. Renaturation of some multidomain proteins may take several hours in vitro, yet it is clear that all possible conformations could not be sampled on the way to native structure. Levinthal (4) summed this up succinctly in the “Levinthal paradox” that can be stated as follows: if a given amino acid can assume approximately 10 different conformations, the total number of possible conformations in a polypeptide chain of 100 residues would be $10^{100}$. The time that this could take would be well beyond the life span of an organism if not of the universe, depending on how many conformations could be sampled before a protein reaches native state. Thus, it was realized early on that cells must have special ways to make the process more efficient.

Experiments to examine the role of the intracellular environment in protein folding involved the renaturation of proteins such as RNase (2), bovine pancreatic trypsin inhibitor (BPTI) (3, 5), or influenza hemagglutinin (6) in isolated microsomal fractions. The results indicated that protein folding can be facilitated by proteins contained in the endoplasmic reticulum (ER) of eukaryotic cells. In the case of disulfide bond-containing proteins such as BPTI (5) or the human chorionic gonadotropin (hCG)-β subunit (7), the key ER protein involved appears to be protein disulfide isomerase (see below).

It was soon realized that many polypeptides can reform native structure easily by themselves in vitro (usually small single domain proteins) while others (more complex, multidomain, or oligomeric proteins) fold and assemble efficiently only in the presence of additional proteins that are not constituents of the final native protein itself. These additional proteins have been called “molecular chaperones.”

The term molecular chaperone was first used by Laskey et al. (8) to describe the role of nucleoplasmin in the assembly of DNA and histones into nucleosomes. The name seemed appropriate because nucleoplasmin promotes histone-histone interactions to form the correct oligomeric form while preventing aggregation. It does so without itself forming part of the nucleosome and without specifying nucleosome structure. Hence nucleoplasmin assumes the role of a chaperone.

The term molecular chaperone has been applied by Ellis and Hemmingsen (9) to the expanding families of proteins of bacterial and eukaryotic compartments involved in protein folding, assembly, and translocation. The term has stuck, and it is now used to define a wide variety of factors that facilitate generation of native protein and nucleic acid structures.

Protein Folding in Vitro Versus in Vivo

There are some similarities as well as differences between intracellular protein folding and protein folding in test tubes. For instance, for the tailspike protein of Salmonella typhimurium phage P22 (10, 11) and hCG-β subunit (7) intermediates in the folding pathway of the proteins appear to be the same in vitro and in vivo, but the rate and efficiency with which proteins achieve final native state in vivo is higher than that in vitro. It must also be kept in mind that, both in vivo and in vitro, correct folding is in competition with misfolding and aggregation. This depends on the protein concentration used for in vitro folding reactions, and in general, very dilute protein concentrations (0.01–0.02 mg/ml) (2, 12) are needed to prevent aggregation. This has presented a huge problem to the biotechnology industry in attempts to produce useful amounts of recombinant proteins. The efficiency of folding in vitro can frequently be facilitated by appropriate adjustment of the redox potential (13–15) or the addition of factors such as protein disulfide isomerase (PDI) for eukaryotic disulfide-bonded proteins (15–17) or DnaK/DnaJ chaperones for bacterial proteins (reviewed in Refs. 18 and 19). In contrast to what happens in vitro, cells minimize or circumvent the off-pathway events by utilizing molecular chaperones that facilitate the folding process by preventing aggregation and other unfavorable interactions.

There is growing interest in what regulates the folding of mammalian proteins in vivo because of the number of human diseases now known to be related to protein folding defects (reviewed in Refs. 20 and 21). This includes cystic fibrosis, α1-antitrypsin deficiency, Alzheimer’s disease, Creutzfeldt-Jacob disease, neurodegenerative diseases such as Huntington’s chorea, and cancer.

The intracellular folding pathway of only a few proteins has been studied in detail. These include the S. typhimurium phage P22 tailspike protein (11), hCG-β subunit (22–25), luciferase (26), influenza hemagglutinin (27, 28), and the HIV type 1 envelope glycoprotein (29). Where the pathways have been determined both in vitro and in vivo, for example for the phage P22 tailspike protein (10, 11) and the hCG-β subunit (7, 22–25), the in vitro and in vivo
Minireview: Assisted Protein Folding

There is also evidence that disulfide bond rearrangement occurs during protein folding in intact cells. For example, in the intracellular folding pathway of the hCG-β subunit determined by pulse-chase kinetics (35) and by site-directed mutagenesis of cysteines involved in disulfide bonds (36, 37), two of the six disulfide bonds formed during the kinetic folding pathway of hCG-β are different from those seen in the crystal structure of the native protein (38, 39).

Another protein whose in vivo folding has been studied is influenza hemagglutinin (HA) (28). The folding of HA in the ER has also been followed by the formation of intrachain disulfide bonds. Folding of HA starts cotranslationally with some disulfide bonds beginning to form soon after both cysteines that are involved in a disulfide pair enter the ER lumen. However, most disulfide bond formation in HA occurs after nascent polypeptide bond termination. This has also been observed during the in vivo folding of hCG-β (35).

Role of Glycosylation in Protein Folding and Assembly

Since many membrane and secretory proteins are glycoproteins, it is important to consider the role of carbohydrates in protein folding, assembly, and secretion. N-Linked oligosaccharides of the high mannose composition are added cotranslationally to the Asn-X-Ser(Thr) consensus sequence of proteins in the ER. One function of N-linked glycans is to facilitate protein folding and conformational maturation. When N-linked chains are eliminated by site-directed mutagenesis of Asn residues in glycosylation consensus sequences or by treatment of cells with agents that block addition of N-linked glycans or their processing, many ER-synthesized glycoproteins misfold, aggregate, and get degraded within the ER (reviewed in Ref. 40). Some glycoproteins seem to fold and be translocated efficiently without their N-linked glycans. The only rule that seems to emerge here is that larger, more complex glycoproteins have more trouble folding if their N-linked glycans are missing.

The role of N-linked oligosaccharide chains in intracellular folding of the hCG-β subunit has been determined by examining the kinetics of folding in Chinese hamster ovary cells transfected with wild-type or mutant hCG-β genes lacking one or both of the asparagine glycosylation sites (41). Folding of hCG-β lacking both N-linked glycans was inefficient and correlated with the slow formation of the last three disulfide bonds (i.e. disulfides 23–72, 93–100, and 26–110) to form in the hCG-β folding pathway. Unglycosylated hCG-β was slowly secreted from Chinese hamster ovary cells, and β subunit folding intermediates retained in cells for more than 5 h were degraded into a smaller hCG-β fragment. However, coexpression of hCG-α, which is required for formation of the biologically active αβ heterodimer, enhanced folding and formation of disulfide bonds 23–72, 93–100, and 26–110 of hCG-β lacking N-linked glycans, suggesting that the presence of its heterodimeric companion subunit fosters β subunit folding and assembly, perhaps because the α subunit can act like a chaperone for β subunit folding. In addition, the molecular chaperones BiP, ERP72, and ERP94, were found in a stable complex with unglycosylated, unfolded hCG-β and may be involved in the folding of this β form (41). These data indicate that N-linked oligosaccharides assist hCG-β subunit folding by facilitating disulfide bond formation, perhaps by increasing the stability and solubility of the native structure that fosters disulfide formation.

Molecular Chaperones

The role of molecular chaperones in protein folding, assembly, and intracellular translocation has been the subject of a number of recent reviews (19, 42–45). ER chaperones play a key role in protein folding and quality control. Cytosolic chaperones play a key role in folding, transport, and biological activity of a number of proteins targeted for transport to specific organelles such as the nucleus and mitochondria. Examples of some eukaryotic molecular chaperones are shown in Table I.

Important members of the ER family of chaperones include BiP, originally characterized as an immunoglobulin binding protein (hence the name (46), GRP (or ERP) 72 (47), GRP (or ERP) 94 (48).
The observation that chaperones are needed to assist protein folding in living cells does not negate the findings of Anfinsen and others that proteins can fold spontaneously in solution based only on information contained in their primary amino acid sequence. Indeed, the data comparing the in vitro versus in vivo folding pathways for proteins that have been studied in this regard, for example the S. typhimurium phase P22 tailspike protein (10, 11) and the hCG-β subunit (7), indicate that proteins go through the same folding steps in vitro and intracellularly. What then do chaperones do and why do cells need them?

**Mechanisms of Chaperone Action**

The best evidence for the mechanisms by which chaperones assist protein folding comes from Escherichia coli in which the chaperones DnaJ/DnaK and the chaperoina GroEL/GroES form in concert to facilitate folding of proteins. Recent biochemical evidence (reviewed in Refs. 19 and 45) and crystallographic data (58, 59) provide a fairly clear and fascinating story on this subject, although there is still some controversy on some key points. GroEL is made up of 14 identical 60-kDa subunits that form two heptameric ring structures with a pocket in the middle that can accommodate proteins up to about 60 kDa. GroES is a single heptameric subunit. Some mutant forms of the hCG-β that do not fold properly are degraded intracellularly while other mutant forms that remain incompletely folded are secreted (37). Nevertheless, there are a number of examples where protein misfolding leads to protein accumulation in the ER and degradation. Some molecular chaperones appear to be involved in targeting irreversibly misfolded proteins for degradation in the ER. Bip is one of these. For example, immunoglobulin light chains that are slowly folding and retained in the ER of cultured mouse cells are quantitatively bound to Bip as partially disulfide-bonded forms and then degraded, whereas light chains that are more rapidly folded and secreted only transiently interact with Bip (57).

**Minireview: Assisted Protein Folding**

| Organelle                | Target polyepitope         | Chaperone          | Role                                      |
|-------------------------|----------------------------|--------------------|-------------------------------------------|
| Mitochondria            | Mitochondrial precursors   | Hsp70              | Completion of translocation               |
|                         | Precursors in matrix       | Hsp60              | Stabilization of prefolded structures     |
| Endoplasmic reticulum   | Nascent secretory proteins | BiP (GRP78)        | Re-export of precursors to intermembrane space |
|                         | Unfolded proteins          | GRP72              | Completion of folding                     |
|                         | Unfolded proteins          | GRP94              | Completion of folding                     |
|                         | Unfolded proteins          | Calreticulin       | Completion of folding (ER membrane bound) |
|                         | Mutant or foreign proteins | Bip                | Stabilization of unfolded structures; target for degradation |
| Subunits of T cell receptor | TRAP or p28             |                    | Receptor assembly                         |
| Cytosol                 | Nascent polypeptides       | Hsc70              | Stabilization of newly synthesized heavy chains |
|                         | Mitochondrial and          | Hsc70              | Catalysis of slow protein folding reactions |
|                         | secretary precursors       |                    |                                           |
|                         | Unfolded proteins          | Hsp40              | Functions in cooperation with Hsc70 to stabilize non-native protein conformations |
| Clathrin-coated vesicles |                            |                    | Binds exposed loop of clathrin light chain to promote uncoating |
| Aged proteins           | Hsc70 (Pep73)             |                    | Targeting to lysosomes for degradation    |
| Steroid receptors       | Hsp90                     |                    | Stabilizes inactive form of receptor      |
| Retroviral transforming proteins | Hsp70                  |                    | Stabilizes inactive receptor              |
| Nucleus                 | Preribosomes               | Hsp70/Hsc70        | Stabilization of prefolded structures     |
|                         | Histones                   | Nucleosomes         | Protection of heat denatured proteins     |

**TABLE I**

Intracellular roles of eukaryotic protein chaperones (modified from Gething and Sambrook (30))

---

2 The GroEL/GroES/Trc family of folding factors are called chaperonins. Functionally, they are members of the larger family of chaperones. The chaperonins have in common the ability to promote protein folding in their central cavity.
otide exchange factor GrpE, complete folding to a native state of some proteins. Other proteins require the additional actions of the GroEL/GroES system to complete folding (Fig. 2) (18).

Analogous systems exist in eukaryotic cells and most likely act through mechanisms similar to the E. coli folding systems. The cytosolic Hsc70 family and the Hsp70 ER analogue BiP appear to act like the DnaK/DnaJ chaperones in E. coli. The TCP-1 family of eukaryotic cytosolic chaperonins, of which the chaperonin TRiC, appears to function like the GroEL/ES system to complete folding (Fig. 2) (18).

One key question that remains is whether proteins fold to native structure while they are bound to chaperones or whether they have to be released into solution to complete folding. Hartl and his colleagues (53) support the former hypothesis. They have data to show that when proteins are synthesized by a more in vitro-like translational system (reticulocyte lysate), sequential binding of folding intermediates by Hsc70/Hsp40 and then by TRiC occurs and that native state is achieved while a protein substrate remains bound to TRiC. Lorimer and his colleagues (60), on the other hand, believe that a polypeptide dissociates completely from GroEL during the folding process and may rebind several times before it is folded correctly to native state (so called "iterative annealing") but that final folding events occur in solution. Whichever model turns out to be correct, it may be substrate-dependent, it is clear that proteins folding at the high concentration and in the highly compact state of the intracellular environment require chaperones in order to assist their folding and prevent their spontaneous aggregation.

REFERENCES
1. Anfinsen, C. B., and Haber, E. (1961) J. Biol. Chem. 236, 1961–1963
2. Goldberg, R. E., Epstein, C. J., and Anfinsen, C. B. (1963) J. Biol. Chem. 238,