Leading Edge
Commentary
Cell

Molecular Chaperones in Cellular Protein Folding: The Birth of a Field

Arthur L. Horwich[1,*]
[1]Howard Hughes Medical Institute and Department of Genetics, Yale University School of Medicine, Boyer Center, 295 Congress Avenue, New Haven, CT 06510, USA
*Correspondence: arthur.horwich@yale.edu
http://dx.doi.org/10.1016/j.cell.2014.03.029

The early decades of *Cell* witnessed key discoveries that coalesced into the field of chaperones, protein folding, and protein quality control.

In January 1974, at the front of the first issue of *Cell*, Benjamin Lewin provided an opening announcement entitled "A Journal of Exciting Biology," establishing the goal of publishing the elucidation of systems responsible for cellular function and phenotype. For those reading across all or part of the 40 year span currently being celebrated, there can be no question that the goal has been met beyond all expectation. I can think of so many astonishing revelations that were first brought to light in *Cell*. The *Cell* paper from Chow et al. (1977) describing splicing—"An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA"—stands out to me as the most dazzling early paper (coinciding with the equally stunning paper of Berget et al., 1977). In looking back and taking stock of an area close to my own heart, I would say that, as a collective, the papers investigating the molecular machines that govern the folded state of proteins inside of the cell—the chaperones—are equally distinguished in describing biology that was unexpected and exciting. Here, I'll discuss how a number of diverse lines of inquiry, published during the first two decades of *Cell*'s history, coalesced into the field of chaperones, protein folding, and protein quality control as we now know it.

The term "molecular chaperone" hadn't been coined at the time *Cell* was launched. That had to wait until 1978, when Ron Laskey used the term to describe nucleoplasmin, a protein that binds and conveys histones into the nuclear compartment, shielding positive charge of the histones via its own acidic character (Laskey et al., 1978). Obviously the term got repurposed to machines that bind nonnative proteins—I think with Ron's blessing, as indicated by his willingness to attend early meetings of the field, speak about nuclear biology, and strum a few songs in the beer frame. As concerns protein folding at the launch of *Cell*, Christian Anfinsen had recently received the Nobel Prize in Chemistry (1972) for work showing that the primary structure of a protein contains all of the information necessary for folding to the native state, which lies at an energetic minimum (Anfinsen, 1973). Who could have thought at that point that thermodynamics would not be enough to produce the native active form of proteins inside of the cell? Who would have imagined that kinetic assistance by a dedicated group of protein machines, in most cases utilizing ATP, would be essential for the proper folding of a large cohort of proteins?

That realization emerged from two contemporaneous but initially disconnected sets of observations. On one hand, it became clear that many proteins could not spontaneously refold in a test tube in the same way as ribonuclease in Anfinsen's early experiments, lodging instead in insoluble aggregates that could be sedimented to the bottom of the tube. In addition, in the cellular context, as expression of mammalian proteins in *E.coli* was undertaken in the late 1970s and early 1980s, it became clear that many expressed proteins were subject to misfolding, aggregation, and localization into terminal inclusion bodies (Williams et al., 1982; Marston, 1986; Haase-Pettingell and King, 1988; Figure 1). Thus, in these situations, there seemed to be kinetic difficulties during protein folding.

On the other hand, a class of proteins known as heat shock proteins was becoming the subject of considerable scrutiny. In 1962, regions of *Drosophila* salivary gland chromosomes were observed to become "puffed" during heat shock (Figure 2). RNAs induced under these conditions were shown by in situ hybridization to be produced from these regions (Spradling et al., 1975; McKenzie et al., 1975). It became clear with molecular cloning of the abundant heat-induced RNA that one of these regions encoded a 70 kDa heat-shock-induced protein (Schedl et al., 1978). At about the same time, it was observed that a characteristic set of heat-inducible proteins, including a 70 kDa protein, was manifest in both *E.coli* (Lemaux et al., 1978; Yamamori et al., 1978; Bardwell and Craig, 1984) and metazoan fibroblasts (Kelley and Schlesinger, 1978). It seemed likely that these inducible proteins would be protective to the cell under stress. Was there a link between heat-shock-induced proteins and the kinetic challenges of in vivo protein folding?

The work of Pelham was particularly telling with respect to heat shock. He observed that *Drosophila* Hsp70 expressed in mouse L cells or monkey COS cells enabled rapid recovery of nucleolar damage following heat shock (Pelham, 1984). He subsequently analyzed release of Hsp70 from the nuclei isolated from heat-shocked cells, observing tight binding of Hsp70 to the nuclei relative to the nonshocked cells and rapid and complete release upon the addition of ATP (Lewis and Pelham, 1985). A model based on these findings was presented in a *Cell* Minireview (Pelham, 1986), proposing a

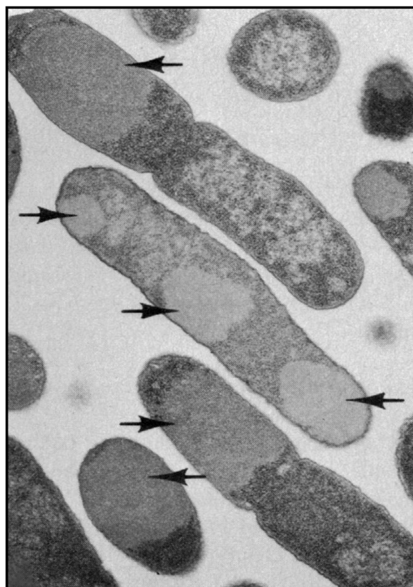

Figure 1. Evidence of Protein Misfolding In Vivo: Formation of Inclusion Bodies
Transmission electron micrograph showing formation of inclusion bodies (arrowed) in *E.coli* expressing a trp-proinsulin fusion protein (from Williams et al., 1982).

cycle of action wherein Hsp70 binds to incipiently aggregating proteins (as produced by heat shock) and pries them apart through recurrent cycles of binding and release associated with ATP binding and hydrolysis. Because Hsp70 was known to strongly bind to hydrophobic column matrices, it was proposed that it recognizes hydrophobic surfaces of the misfolding proteins and prevents them from driving aggregation. It was a prescient model. Indeed, recognition by molecular chaperones generally involves the binding of hydrophobic surfaces specifically exposed in nonnative proteins by hydrophobic surfaces proffered by the chaperones themselves, each chaperone family offering a different geometry of binding surface (Bukau and Horwich, 1998). Subsequent binding of ATP then produces allosterically mediated movement of the binding surface that releases the protein substrate (Zhuravleva et al., 2012; Kityk et al., 2012; Clare et al., 2012).

Pelham's observations concerning the action of Hsp70 fit well with concurrent data that emerged from studies of two other 70 kDa proteins, the immunoglobulin-binding protein (BiP) inside of the endoplasmic reticulum (ER) and the clathrin-uncoating ATPase in the cytosol. In the former case, a 70 kDa protein was found to bind selectively to immunoglobulin heavy chains prior to their association with light chains, indicating once again a protein-protein interaction, here potentially facilitating oligomeric assembly (Haas and Wabl, 1983). In the latter case, studies of Rothman and coworkers (Schlossman et al., 1984; Chappell et al., 1986) and of Ungewickell (1985) indicated that a 70 kDa protein was an ATP-dependent mediator of uncoating clathrin cages from vesicles during endocytosis, releasing clathrin triskelions. This amounted to an action more like that described by Pelham, in which binding of the 70 kDa protein mediates disassembly of a protein complex—in the case of clathrin, an action carried out under normal physiologic conditions (by what we now know to be the constitutively expressed heat shock 70 "cognate" protein, Hsc70 [Xing et al., 2010]).

Shortly thereafter, cytosolic Hsp70 proteins became implicated in transport of protein precursors into ER and mitochondria. The chaperone binds the protein to be transported in the cytosol, apparently preventing its hydrophobic surfaces from producing aggregation and holding it in an unfolded state that could engage with and pass through translocation machinery (Chirico et al., 1988; Deshaies et al., 1988; Eilers and Schatz, 1986). These events were not stress related, indicating a constitutive need for the action of 70 kDa class chaperone proteins.

The apparently disparate worlds of protein folding and molecular chaperones converged in the 1980s, with the characterization of a separate class of heat-inducible ATP-hydrolyzing proteins: oligomeric double-ring protein complexes composed of $\sim$60 kDa subunits, the Hsp60s. These complexes were first implicated as playing a role in oligomeric assembly. Genetic defects in the bacterial *GroE* operon (Georgopoulos et al., 1972; Takano and Kakefuda, 1972) were observed to affect the ability of propagating phages to assemble, but they also affected bacterial cell growth. Likewise, a role in oligomeric assembly was ascribed to the similarly sized Rubisco subunit-binding protein, which could associate with newly translated large subunits of Rubisco inside of the chloroplast stroma, but not with mature Rubisco, formed by assembly of the large subunits with small subunits imported from the cytosol (Barraclough and Ellis, 1980). John Ellis dubbed these ring complexes chaperonins. The homology of GroEL with Rubisco-binding protein was then appreciated upon sequencing of the respective coding regions (Hemmingsen et al., 1988).

A role for chaperonins in polypeptide chain folding, as distinct from oligomeric assembly, soon emerged from studies of a yeast mutant affecting a GroEL homolog in the mitochondrial matrix, mitochondrial Hsp60 (Cheng et al., 1989). In this mutant, proteins entering mitochondria failed to reach native form. Among the first proteins found to be affected in the mutant was a monomeric protein, the Rieske iron-sulfur protein. This suggested that proper polypeptide folding, as opposed to oligomeric protein assembly of already-folded monomers, might be the step facilitated by the chaperonin ring assemblies. This role was further established by the observation that monomeric DHFR imported into mitochondria (by attachment of an N-terminal mitochondrial targeting signal) associated in a nonnative form with the Hsp60 complex and was subsequently released in a native form upon addition of ATP (Ostermann et al., 1989). *Hsp60* proved to be an essential gene in yeast, indicating a requirement for its action under all conditions (Cheng et al., 1989; Reading et al., 1989).

Mechanistic insights were enabled by in vitro reconstitution experiments. The first reconstitution experiment was carried out with the dimeric Rubisco from *R. rubrum*. Denatured subunit diluted from denaturant became bound to GroEL, and subsequent addition of ATP and cochaperonin GroES (a single ring composed of 10 kDa "small" subunits) led to production of native, active Rubisco (Goloubinoff et al., 1989). Further in vitro refolding experiments with monomeric DHFR and rhodanese confirmed that GroEL/GroES could mediate refolding to the native state of these proteins following their dilution from chaotrope, whereas quantitative aggregation occurred upon dilution into buffer solution (Martin et al., 1991). Apparently the chaperonin system

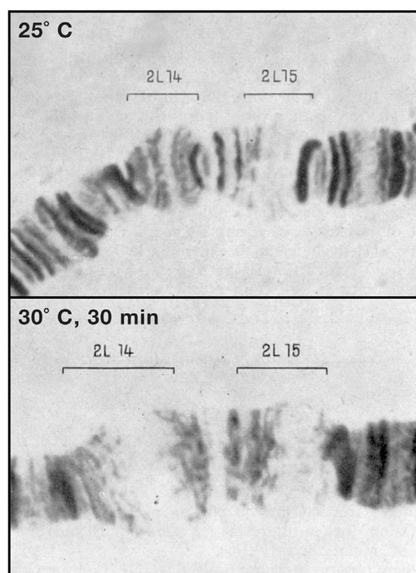

Figure 2. A Transcriptional Response to Heat Stress
Drosophila busckii salivary gland chromosome spreads, showing "puffing" of two regions following temperature shift of larvae from 25°C (top) to 30°C for 30 min (bottom) (from Ritossa, 1962).

could prevent or remove proteins from kinetically trapped states and allow them to reach native form. The chaperonin reaction could be experimentally broken into steps: binding to GroEL prevented wholesale aggregation of the protein substrate, and subsequent addition of GroES/ATP produced the native state over a period of minutes, with GroES appearing to couple the folding reaction to the GroEL ring assembly (Goloubinoff et al., 1989; Martin et al., 1991). We now know that substrate proteins bind to a surrounding hydrophobic surface in the cavity of an open GroEL ring and are, upon ATP/GroES binding, ejected into a now hydrophilic GroES-encapsulated chamber where they proceed to fold in isolation, without the chance of aggregation (Horwich and Saibil, 2011).

Thus, the early experiments summarized here established that there are molecular machines that prevent protein aggregation and use ATP to help adjust the conformation of other proteins. Their abundance increases under stress conditions via transcriptional regulation to provide increased capacity to prevent aggregation. Under nonstress conditions, they provide kinetic assistance to folding, apparently necessary due to ongoing missteps of protein folding even at physiologic temperature in a milieu that has a large concentration of solute. Not only could the chaperones prevent aggregation from occurring, but they could also help proteins to maintain unfolded conformations when necessary, e.g., when emerging from ribosomes or when passage through membranes required an unfolded state. Finally, in the case of the chaperonins, they directly promote the native state of proteins via folding inside of an encapsulated chamber. Thus, the existence of a diverse and dedicated machinery for protein "management" in the cell had been uncovered.

Postscript

A host of other chaperone machines have been identified in a variety of compartments—too many to list here. In addition, the field has expanded to include the characterization of stress-sensing pathways that provide exquisite regulation of the chaperone systems. It is now also apparent that the chaperone systems link to the ubiquitin-proteasome and autophagy systems as part of a global quality control network. Unexpected discoveries are surely still to come.

ACKNOWLEDGMENTS

I'd like to particularly thank Walter Neupert and Ulrich Hartl for a well-timed phone call that triggered a wonderful collaboration to understand the mitochondrial-folding-defective yeast mutant affecting Hsp60. I thank all of the talented people who have worked in my group over the years, as well as my collaborators, with whom I've enjoyed watching our field grow and develop. I thank the NIH and Howard Hughes Medical Institute for generous support of the chaperonin work. And I thank Ben Lewin and his successors for their ongoing interest in our field, which I think has lived up to providing "exciting biology."

REFERENCES

Anfinsen, C.B. (1973). Science *181*, 223–230.

Bardwell, J.C.A., and Craig, E.A. (1984). Proc. Natl. Acad. Sci. USA *81*, 848–852.

Barraclough, R., and Ellis, R.J. (1980). Biochim. Biophys. Acta *608*, 19–31.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Proc. Natl. Acad. Sci. USA *74*, 3171–3175.

Bukau, B., and Horwich, A.L. (1998). Cell *92*, 351–366.

Chappell, T.G., Welch, W.J., Schlossman, D.M., Palter, K.B., Schlesinger, M.J., and Rothman, J.E. (1986). Cell *45*, 3–13.

Cheng, M.-Y., Hartl, F.-U., Martin, J., Pollock, R.A., Kalousek, F., Neupert, W., Hallberg, E.M., Hallberg, R.L., and Horwich, A.L. (1989). Nature *337*, 620–625.

Chirico, W.J., Waters, M.G., and Blobel, G. (1988). Nature *332*, 805–810.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977). Cell *12*, 1–8.

Clare, D.K., Vasishtan, D., Stagg, S., Quispe, J., Farr, G.W., Topf, M., Horwich, A.L., and Saibil, H.R. (2012). Cell *149*, 113–123.

Deshaies, R.J., Koch, B.D., Werner-Washburne, M., Craig, E.A., and Schekman, R. (1988). Nature *332*, 800–805.

Eilers, M., and Schatz, G. (1986). Nature *322*, 228–232.

Georgopoulos, C.P., Hendrix, R.W., Kaiser, A.D., and Wood, W.B. (1972). Nat. New Biol. *239*, 38–41.

Goloubinoff, P., Christeller, J.T., Gatenby, A.A., and Lorimer, G.H. (1989). Nature *342*, 884–889.

Haas, I.G., and Wabl, M. (1983). Nature *306*, 387–389.

Haase-Pettingell, C.A., and King, J. (1988). J. Biol. Chem. *263*, 4977–4983.

Hemmingsen, S.M., Woolford, C., van der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W., and Ellis, R.J. (1988). Nature *333*, 330–334.

Horwich, A.L., and Saibil, H.R. (2011). The GroEL/GroES chaperonin machine. In Molecular Machines in Biology, Joachim Frank, ed. (New York: Cambridge University Press), pp. 191–207.

Kelley, P.M., and Schlesinger, M.J. (1978). Cell *15*, 1277–1286.

Kityk, R., Kopp, J., Sinning, I., and Mayer, M.P. (2012). Mol. Cell *48*, 863–874.

Laskey, R.A., Honda, B.M., Mills, A.D., and Finch, J.T. (1978). Nature *275*, 416–420.

Lemaux, P.G., Herendeen, S.L., Bloch, P.L., and Neidhardt, F.C. (1978). Cell *13*, 427–434.

Lewis, M.J., and Pelham, H.R.B. (1985). EMBO J. *4*, 3137–3143.

Marston, F.A.O. (1986). Biochem. J. *240*, 1–12.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., and Hartl, F.-U. (1991). Nature *352*, 36–42.

McKenzie, S.L., Henikoff, S., and Meselson, M. (1975). Proc. Natl. Acad. Sci. USA *72*, 1117–1121.

Ostermann, J., Horwich, A.L., Neupert, W., and Hartl, F.-U. (1989). Nature *341*, 125–130.

Pelham, H.R.B. (1984). EMBO J. *3*, 3095–3100.

Pelham, H.R.B. (1986). Cell *46*, 959–961.

Reading, D.S., Hallberg, R.L., and Myers, A.M. (1989). Nature *337*, 655–659.

Ritossa, F.M. (1962). Experientia *18*, 571–573.

Schedl, P., Artavanis-Tsakonas, S., Steward, R., Gehring, W.J., Mirault, M.-E., Goldschmidt-Clermont,

Cell

M., Moran, L., and Tissières, A. (1978). Cell *14*, 921–929.

Schlossman, D.M., Schmid, S.L., Braell, W.A., and Rothman, J.E. (1984). J. Cell Biol. *99*, 723–733.

Spradling, A., Penman, S., and Pardue, M.L. (1975). Cell *4*, 395–404.

Takano, T., and Kakefuda, T. (1972). Nat. New Biol. *239*, 34–37.

Ungewickell, E. (1985). EMBO J. *4* (13A), 3385–3391.

Williams, D.C., Van Frank, R.M., Muth, W.L., and Burnett, J.P. (1982). Science *215*, 687–689.

Xing, Y., Böcking, T., Wolf, M., Grigorieff, N., Kirchhausen, T., and Harrison, S.C. (2010). EMBO J. *29*, 655–665.

Yamamori, T., Ito, K., Nakamura, Y., and Yura, T. (1978). J. Bacteriol. *134*, 1133–1140.

Zhuravleva, A., Clerico, E.M., and Gierasch, L.M. (2012). Cell *151*, 1296–1307.