Protein-specific chaperones: **The role of hsp47 begins to gel**

Linda M. Hendershot* and Neil J. Bulleid†

Collagen biosynthesis involves a complex series of post-translational modifications, controlled by a number of general and specific molecular chaperones. A recent study has shed new light on the role played in this process by the procollagen-specific chaperone Hsp47.

Addresses: *Department of Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, Tennessee 38105, USA. †University of Manchester, School of Biological Sciences, 2.205 Stopford Building, Manchester M13 9PT, UK.

Current Biology 2000, 10:R912–R915

0960-9822/00/$ – see front matter

Proteins that are expressed at the surface or secreted from mammalian cells begin their synthesis in the endoplasmic reticulum (ER). The ER of a given cell contains a concentrated mixture of proteins in various stages of folding and assembly. The major functions of this organelle are to provide an environment conducive to protein folding that resembles the extracellular space, to distinguish completely folded and assembled proteins from those still in the process of folding, and to identify incorrectly folded proteins and target them for degradation. In spite of the high concentration of aggregation-prone precursor polypeptides in the ER, more than 95% of newly synthesized proteins attain their proper tertiary and quaternary structure and are transported to their proper destination. This is achieved largely through the interactions of a host of resident ER folding enzymes and molecular chaperones [1,2]. These proteins recognize common features of partially folded proteins, such as hydrophobic patches, free sulfhydryl groups and incompletely processed oligosaccharyl side chains. By virtue of the types of structure they bind, these chaperones are able to interact with a very diverse array of secretory pathway proteins and can be viewed as the primary tier in the quality control machinery.

There are some proteins that have atypical physical structures, or that exist under unusual situations or conditions, which require additional help to achieve their correct mature conformation. Protein-specific chaperones have co-evolved to monitor and aid the proper maturation of these proteins, and these represent a second, more restricted tier in the quality control process (Table 1). The actual functions of many of these protein-specific chaperones is not well understood, however. Collagen is an example of a protein with a particularly complex biosynthesis pathway which requires additional help to mature properly. In addition to interacting with at least six of the general folding enzymes and chaperones [3], procollagen binds to heat shock protein 47 (Hsp47), which appears to be a chaperone specific for procollagen [4,5].

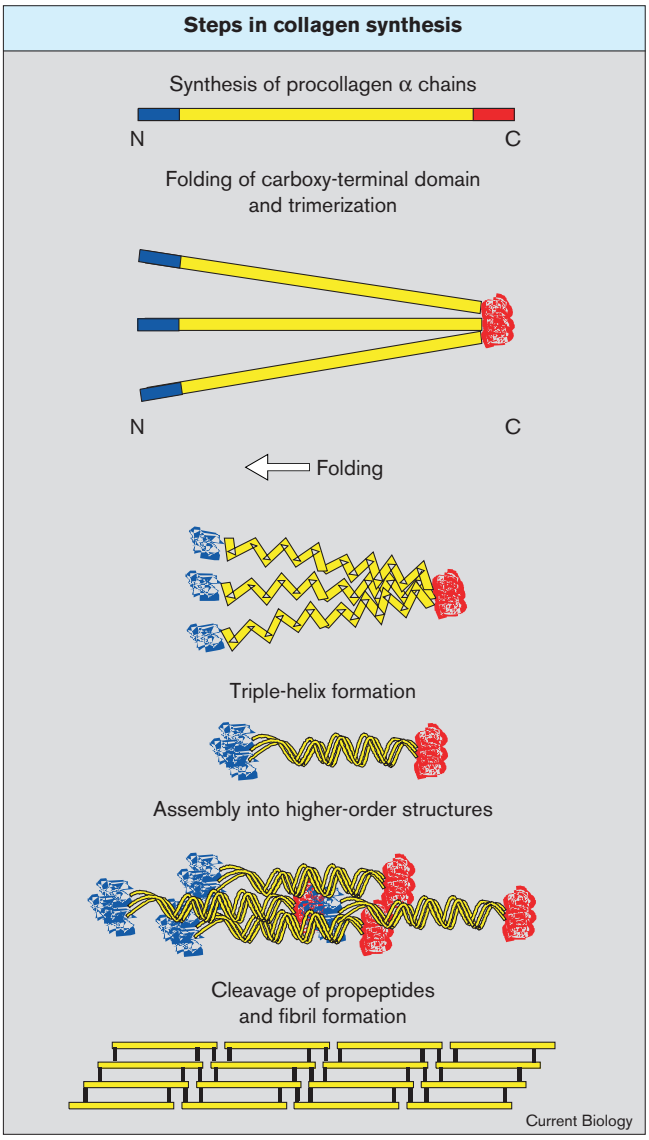
New insight into the role of Hsp47 in collagen biosynthesis has come from the recent work of Nagata *et al.* [6], who disrupted the *hsp47* gene in mice. The null mutant mice were found to be severely deficient in the mature processed form of collagen in mesenchymal tissues and basement membranes. The mice did not survive past day 11.5 postcoitus, and the embryos displayed abnormally orientated epithelial tissues and ruptured blood vesicles. These results demonstrate that Hsp47 is required to facilitate the secretion of stable and correctly processed collagen, and that mice cannot develop without this specialized molecular chaperone.

The complex collagen biosynthesis pathway

Collagen is the most abundant mammalian protein [7,8]. It is a major component of the extracellular matrix that forms a specialized network around cells and in the interstitial spaces between the cells in a tissue. The extracellular matrix functions to bind cells together, to provide a reservoir for hormones that control growth and differentiation, and to form a lattice through which cells move as they differentiate. Defects in extracellular matrix can lead to developmental abnormalities and to cancer. At least 19 different types of collagen have been identified, which are formed from 33 distinct gene products. Each collagen molecule is composed of three α chains which wind around each other to form the rigid triple-helical structure that is characteristic of collagens [7]. The abundant glycine and proline residues in collagen, which form repeating G-X-Y motifs where X is any amino acid and Y is often proline or hydroxyproline, are responsible for the tight helical structure. Only glycine has a small enough head group to be accommodated in the interior of the collagen helix. Multiple collagen triple helical molecules then assemble with each other through covalent bonds to make the long collagen fibrils that are characteristic of the extracellular matrix and basal lamina.

Collagen's particularly complex biosynthesis pathway begins in the ER, where one or more distinct procollagen α chains are translocated (Figure 1). In addition to the ER targeting signal sequence, the procollagen chain has amino-terminal and carboxy-terminal propeptide sequences that flank the approximately 1050 residue helical region [8]. Most proteins that enter the ER begin folding at their amino termini before protein synthesis and translocation are completed. The procollagen chain, however, begins folding from the carboxy-terminal end of the protein,

Figure 1



Steps in collagen synthesis: a schematic outline of the stages in procollagen folding, assembly and processing. The carboxy-terminal propeptides of three procollagen α chains fold and assemble into a trimer, prior to folding of the helical domain. Folding of the remaining parts of the three molecules continues from the carboxyl terminus to the amino terminus. The folded central domains are wound around each other to form a triple helix. The procollagen molecules may then aggregate to form higher-order structures – procollagen bundles – prior to cleavage of their amino- and carboxy-terminal propeptides. This processing results in a dramatic reduction in solubility, leading to self-assembly into ordered collagen fibrils.

which is the last portion to enter the ER. The Hsp70 class ER chaperone BiP, which recognizes hydrophobic residues on extended polypeptide chains [9], binds to procollagen if the carboxy-terminal region is unable to fold [10]. It is not clear whether BiP also binds transiently to normal procollagen chains that are in the process of folding.

Table 1

Protein-specific quality control factors and their target molecules and functions.

Factor	Target molecule
P24 family	Various secreted proteins, such as invertase and Gas1p in *S. cerevisiae*
ERGIC-53	Glycoproteins, such as cathepsin C and blood clotting factors V and VIII
LST1	Plasma membrane H^+-ATPase Pma1p
Erv14p	Plasma membrane Ax12p
Chs7p	Chs3p catalytic subunit of chitin synthetase III
Vma12p–Vma22p complex	Vacuolar H^+-ATPase
Gsf2p	Hexose transporters Hxt1 and Gal2p
Lag1p and Dgt1p	GPI-anchored proteins Gas1p and Yap3p
Shr3p	Amino-acid permeases, such as Hip1p and Gap1p
ODR-4 and ODR-8	Odorant receptors ODR-10 and STR-2
Prolyl 4-hydroxylase	Procollagen
Hsp47	Procollagen
Bap31	Cellubrevin
RAP	LDL receptor family and other transmembrane receptors
Invariant chain	MHC class II
Tapasin	MHC class I
NinaA	Rh1 and Rh2 rhodopsins
Microsomal triglyceride transfer protein	Apolipoprotein B
Protective protein/cathepsin A	Neuraminidase and β-galactosidase
β-catenin	E-cadherin
Egasyn	β-glucuronidase
Carboxylesterase	C-reactive protein
αA and αB crystallin	Eye lens specific proteins

Adapted from [16].

The nascent procollagen chain is modified by *N*-linked glycans as it enters the ER, which provides the recognition motif for the chaperones calnexin and calreticulin, although their interaction does not appear to be required for the proper maturation of procollagen. Some of collagen's proline and lysine residues are hydroxylated by the resident ER enzyme prolyl 4-hydroxylase, and these residues can be further modified by the addition of galactose and glucose. These hydroxylated residues form hydrogen bonds that stabilize the triple helix [11]. Vitamin C deficiency inhibits hydroxylation of procollagen chains, resulting in the formation of molecules with unstable triple

helices, the degradation of procollagen and the consequently weakened blood vessels, tendons and skin that are characteristic of scurvy [12].

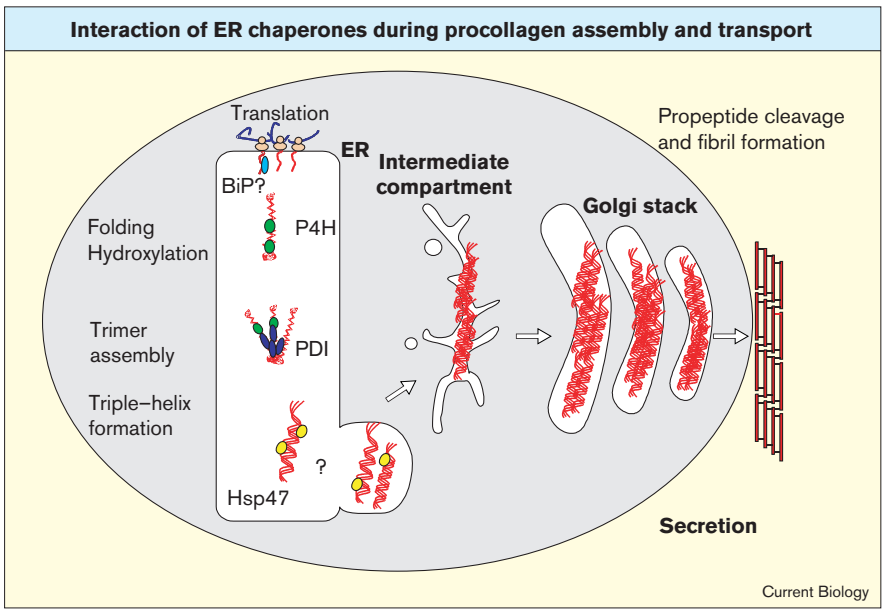
The procollagen α-chains next interact with protein disulfide isomerase (PDI), which promotes folding and catalyzes formation of disulfide bonds between the carboxyl domains to form the procollagen trimers [13]. Each chain continues folding from the carboxyl terminus to the amino terminus. Osteogenesis imperfecta, or brittle bone disease, is caused by single mutations of glycine residues in procollagen α chains; those mutations that occur closer to the carboxyl terminus of the triple-helical region are usually more severe than those near the amino terminus, as they affect the folding of a larger portion of the procollagen chain. The three, folded procollagen chains next zip together to form a triple helix through their helical domains. The amino and carboxyl propeptides are not included in this helical structure, but remain as globular domains at either end. It is thought that these domains prevent the formation of large extended fibrils within the cell. It is the triple-helical form of procollagen that Hsp47 appears to have the highest affinity for [14], although there are reports of it binding to other earlier intermediates [5]. The procollagen molecule is next transported through the intermediate compartment and Golgi and finally to the surface of the cell (Figure 2).

As it progresses along this pathway, procollagen begins to form higher-order structures, leading to characteristic distensions of the Golgi apparatus — an observation which supports the idea that the procollagen aggregates are transported via cisternal maturation rather than vesicular transport [15]. At the cell surface, collagen's propeptides are removed by amino- and carboxy-terminal peptidases, leading to a 10,000-fold decrease in solubility of the collagen triple-helical region and consequent self-assembly of collagen α chains into ordered fibrils. Lysyl oxidase catalyzes the formation of highly reactive aldehyde groups on lysine and hydroxlysine groups, which then form covalent aldol cross-links between two triple helical regions to make stable collagen fibrils. So what is Hsp47 doing that is so essential to proper collagen biosynthesis [6]?

Possible roles for Hsp47 in collagen maturation

There are a number of possible roles that Hsp47 might have in collagen maturation. First, it is possible that triple-helical procollagen still requires additional modification or alterations while in the ER, but no longer looks enough like an unfolded protein to be recognized by the general chaperones. Binding of Hsp47 to the triple helix might serve to retain this form in the ER, giving it additional time to achieve the correct structure and protect it from intracellular degradation. In the absence of Hsp47, the procollagen would either be targeted for degradation, explaining the lower steady-state levels observed by Nagata *et al.* [6] in the embryo, or leave the ER prematurely, explaining the presence of unstable and unprocessed extracellular procollagen in the *hsp47*$^{-/-}$ cells [6]. Hsp47 might alternatively act to mask reactive groups that are involved in assembling higher-order complexes, to ensure this does not happen in the ER. The data on the mutant

Figure 2



Interaction of ER chaperones during procollagen assembly and transport. Procollagen interacts with a variety of ER resident proteins during folding and assembly within the ER. The carboxy-terminal propeptide may interact transiently with BiP (light blue) before it folds correctly and assembles covalently with two other pro-α-chains to form a trimer. Protein disulfide isomerase (PDI; dark blue) catalyzes this latter event. Before folding is complete, prolyl 4-hydroxylase (P4H; green) modifies selected proline and lysine residues in the triple helix-forming domains. These hydrolylated residues form hydrogen bonds that stabilize the helix once it forms. This triple helical form of procollagen is the preferred substrate of Hsp47 (yellow). Procollagen chains can then form higher order structures during transport from the ER through the Golgi apparatus. Removal of the carboxy-terminal and amino-terminal propeptides occurs either just before, or immediately after, the chains are secreted into the extracellular matrix, which allows collagen to form ordered fibrils.

mouse [6] are less compatible with this role, as there was no evidence that collagen aggregates accumulate in the *hsp47*$^{-/-}$ cells.

A third possibility is that, by preventing the transport of individual procollagen molecules from the ER, Hsp47 might act to concentrate procollagen in this subcellular compartment. This concentration might be required to initiate an aggregation event, within either the ER or the intermediate compartment, which slows down transport and diverts the procollagen chains into a transport pathway involving cisternal maturation. In the absence of Hsp47, the procollagen chain would be transported normally by vesicular transport. This would explain the lack of propeptide processing in the *hsp47*$^{-/-}$ cells, if the processing event requires formation of higher-order structures of procollagen within the secretory pathway, and is compatible with the finding that some procollagen is secreted from *hsp47*$^{-/-}$ cells. A final possibility is that Hsp47 promotes stability in the triple-helical region of procollagen. The thermal stability of this region is just above body temperature, and it is conceivable that Hsp47 binds to regions of the helix to prevent micro-unfolding, particularly during heat stress. The fact that Hsp47 is inducible by heat shock is compatible with this idea, which also fits with the decreased stability of the procollagen secreted by the *hsp47*$^{-/-}$ cells [6].

Lessons from other protein-specific chaperones

Further insights into possible roles for Hsp47 might come from considering the workings of other known protein-specific chaperones (Table 1). The chaperones with known functions work in various ways. Some, such as Vma and triglyceride transfer proteins, promote the assembly of their target protein. Others prevent premature aggregation or ligand binding; this is particularly true of those cases where potential ligands are synthesized in the same compartment as the receptor — examples are the invariant chain, which binds to the peptide binding groove of MHC II molecules, and the lipoprotein receptor associated protein, RAP. Yet other protein-specific chaperones actively incorporate their target proteins into transport vesicles, as in the examples of P24 and ERGIC. All of these functions are compatible with the data on the *hsp47* mutant mouse [6], and can be considered as additional potential roles for Hsp47.

In summary, although the data from the *hsp47* mutant mouse does not tell us exactly what this protein-specific chaperone is doing during the maturation of procollagen, the very dramatic phenotype of this mouse does tell us that Hsp47's function is absolutely critical to the correct maturation of collagen. While interaction of procollagen with other general chaperones may also be required, they are clearly not sufficient to allow animals to produce functionally active collagen. The challenge ahead is to obtain biochemical and cell biological data to elucidate more precisely how Hsp47 acts to facilitate collagen maturation.

References

1. Haas IG: **BiP (GRP78), an essential hsp70 resident protein in the endoplasmic reticulum.** *Experientia* 1994, **50:**1012-1020.
2. Hammond C, Helenius A: **Quality control in the secretory pathway.** *Curr Opin Cell Biol* 1995, **7:**523-529.
3. Lamande SR, Bateman JF: **Procollagen folding and assembly: the role of endoplasmic reticulum enzymes and molecular chaperones.** *Semin Cell Dev Biol* 1999, **10:**455-464.
4. Nakai A, Satoh M, Hirayoshi K, Nagata K: **Involvement of the stress protein HSP47 in procollagen processing in the endoplasmic reticulum.** *J Cell Biol* 1992, **117:**903-914.
5. Nagata K: **Hsp47: a collagen-specific molecular chaperone.** *Trends Biochem Sci* 1996, **21:**22-26.
6. Nagai N, Hosokawa M, Itohara S, Adachi E, Matsushita T, Hosokawa N, Nagata K: **Embryonic lethality of molecular chaperone hsp47 knockout mice is associated with defects in collagen biosynthesis.** *J Cell Biol* 2000, **150:**1499-1506.
7. Kadler KE, Holmes DF, Trotter JA, Chapman JA: **Collagen fibril formation.** *Biochem J* 1996, **316(Pt 1):**1-11.
8. Kadler K: **Extracellular matrix 1: fibril-forming collagens.** *Protein Profile* 1995, **2:**491-619.
9. Blond-Elguindi S, Cwirla SE, Dower WJ, Lipshutz RJ, Sprang SR, Sambrook JF, Gething MJ: **Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP.** *Cell* 1993, **75:**717-728.
10. Chessler SD, Byers PH: **BiP binds type I procollagen pro alpha chains with mutations in the carboxyl-terminal propeptide synthesized by cells from patients with osteogenesis imperfecta.** *J Biol Chem* 1993, **268:**18226-18233.
11. Pihlajaniemi T, Myllyla R, Kivirikko KI: **Prolyl 4-hydroxylase and its role in collagen synthesis.** *J Hepatol* 1991, **13(Suppl 3):**S2-S7.
12. Kivirikko KI: **Collagens and their abnormalities in a wide spectrum of diseases.** *Ann Med* 1993, **25:**113-126.
13. Wilson R, Lees JF, Bulleid NJ: **Protein disulfide isomerase acts as a molecular chaperone during the assembly of procollagen.** *J Biol Chem* 1998, **273:**9637-9643.
14. Tasab M, Batten MR, Bulleid NJ: **Hsp47: a molecular chaperone that interacts with and stabilizes correctly-folded procollagen.** *EMBO J* 2000, **19:**2204-2211.
15. Bonfanti L, Mironov AA, Jr, Martinez-Menarguez JA, Martella O, Fusella A, Baldassarre M, Buccione R, Geuze HJ, Mironov AA, Luini A: **Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation.** *Cell* 1998, **95:**993-1003.
16. Ellgaard L, Molinari M, Helenius A: **Setting the standards: quality control in the secretory pathway.** *Science* 1999, **286:**1882-1888.