Unstable molecules form stable tissues

Anton V. Persikov and Barbara Brodsky*

Department of Biochemistry, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, NJ 08854

Collagens are major structural proteins in the extracellular matrix, making up about one-third of protein mass in higher animals. In addition to their sheer bulk, this protein family is of interest because of their diversity of structural and morphogenetic roles and the attribution of an increasing number of hereditary diseases to mutations in collagens (1–4). All collagens have a distinctive molecular conformation: a triple-helix composed of three supercoiled polyproline II-like helical chains (5–7). This triple-helical conformation places strict constraints on amino acid sequence, requiring Gly as every third residue and a high content of proline and hydroxyproline residues. There are more than 20 distinct genetic types of collagens, and the most abundant are types I, II, and III, found in fibrils with a characteristic 67-nm axial period (1). Type I collagen, a heterotrimer composed of two α1(I) chains and one α2(I) chain, forms the prominent fibrils in tendon, bone, and cornea, whereas type III collagen, a disulfide-linked homotrimer, is found together with type I in fibrils of blood vessels and skin. These fibril-forming collagens are synthesized in a procollagen form, with globular propeptides on each end of a central triple-helix (Fig. 1; ref. 3). Self-association and disulfide cross-linking of three C-propeptides are responsible for the initial events of chain selection and trimer formation, whereas subsequent events include nucleation and zipper-like folding of the triple-helix domain (8). After cleavage of the propeptides, the rod-like triple-helical molecules in the matrix self-associate in a staggered array, forming fibrils and interacting with other matrix molecules to provide the strength, flexibility, or compression required for each tissue. Collagen fibers are inherently stable structures, having lifetimes of at least 6 months, and often much longer. Turnover is accomplished through a specialized family of tightly regulated matrix metalloproteinases, because triple-helices are resistant to digestion by most proteases (9).

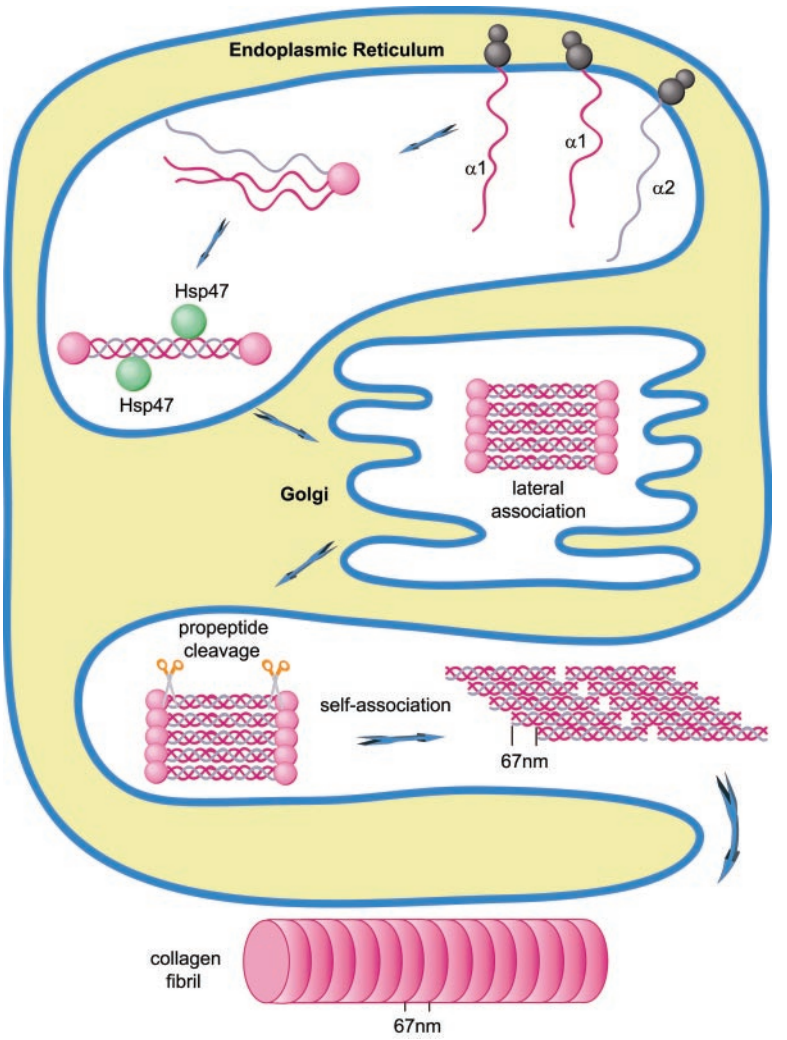

Fig. 1. A schematic illustration of the journey of collagen from its biosynthesis in the endoplasmic reticulum through the Golgi and into the extracellular matrix. The collagen molecule is shown to be complexed with chaperone Hsp47 in the endoplasmic reticulum and self-associated in various aggregates until its final arrival as a stable 67-nm periodic fibril.

Even though collagen fibers are long-lived structures, the stability of their constituent collagen molecules is marginal with respect to physiological temperature (10). When heated in physiological buffers, collagen molecules spontaneously self-associate, but if fibril formation is prevented, through use of glycerol or low pH, collagen molecules undergo a thermal transition, from triple-helical trimers to unfolded monomers. For mammals and birds, this denaturation temperature (T_m) appears to be a few degrees higher than body temperature, whereas the T_m correlates with the upper environmental temperature for poikilotherms (10). For example, the T_m of human type I collagen, as determined spectrophotometrically or by proteolytic digestion, is typically cited as approximately 41.5°C (11–13), whereas that from the skin of an ice fish is 6°C (10). This variation in T_m appears to be medi-

See companion article on page 1314.

*To whom reprint requests should be addressed. E-mail: brodsky@umdnj.edu.

ated through the hydroxyproline content, although there is increasing evidence that other amino acids may also play a role (14, 15). The correlation between T_m and body temperature was rationalized as ensuring sufficient stability for collagen molecules before incorporation into fibrils, while allowing sufficient mobility for fiber formation and turnover. In this issue of PNAS, Leikina *et al.* (16) turn the current paradigm on its head and conclude that collagen molecules, which are so important for long-term stability of tissues, are not themselves stable at body temperature. By using extremely slow calorimetry, together with isothermal circular dichroism spectroscopy, the preferential state for type I collagen at body temperature is demonstrated to be random coil rather than a triple helix. After almost half a century of studies, can it be true that collagen molecules are unstable at physiological temperatures, and if so, is it of biological relevance?

Determination of collagen stability is not straightforward. Equilibrium thermodynamics has been used successfully to study denaturation of small globular proteins (17). However, it has proved problematic to apply this theory to the denaturation of collagen, despite its apparent structural simplicity. First, unfolding of collagen molecules does not follow a simple two-state model. The larger value of the calorimetric enthalpy compared with the van't Hoff enthalpy is indicative of the presence of several cooperative units (18). Second, the system is extremely slow to reach equilibrium (19–22). The lack of equilibrium is indicated by hysteresis observed when refolding curves do not retrace unfolding curves (21, 23), the strong dependence of the melting temperature on the heating rate (16), and calorimetric irreversibility (24–26). In 1967, von Hippel noted that "of all macromolecules studied extensively to date, the collagen-gelatin system exhibits by far the slowest rate of helix-coil conversion," a process he described as "agonizingly" slow, on the time scale of hours, days, or even longer (19). Such difficulty in reaching equilibrium is observed for guanidinium hydrochloride-induced, as well as thermal, unfolding and refolding (21, 23). Factors that may contribute to such slow equilibration include association/dissociation for non-cross-linked chains; cis-trans isomerization of the numerous imino acids; misfolding/misalignment of long chains; and the limited conformational mobility at temperatures below T_m (16, 19, 21). Type I collagen, the system studied by Leikina *et al.* (16), is a particularly difficult system for attaining reversibility. After cleavage of its propeptides, type I collagen is not cross-linked. Then, refolding requires chain association to form triple-helices, and its original heterotrimeric nature cannot be recreated, because it was biosynthetically determined by C-propeptides (16, 20). Even type III collagen, which consists of three identical chains disulfide linked at the C terminus of the triple-helix, exhibits hysteresis (21, 23). The difficulty in reaching equilibrium in reasonable time frames leads to an appearance of irreversibility that can be treated by kinetic analysis (27, 28), and Miles and Bailey have applied a kinetic approach to the collagen system (24–26).

Despite the impossibility of observing equilibrium unfolding of type I collagen, the isothermal data presented in Leikina *et al.* clearly show that human collagen is unstable at 36°C (16). In support of this conclusion, equilibrium unfolding curves have been reported recently for type III collagen retaining the N-propeptide, a molecule with three interchain disulfide bonds at both the N and C termini (22). Incubation of this doubly cross-linked collagen for 24 h at each temperature led to an equilibrium T_m value of 35°C, significantly below body temperature.

The existence of an unstable structural molecule that forms some of the most stable tissue structures in the body seems like a paradox. However, collagen molecules are not in equilibrium *in vivo*, and "unstable" native collagen molecules do successfully navigate the biosynthetic process, getting secreted and incorporated into durable fibrils. There is evidence that collagen in the cell is more stable than expected from *in vitro* studies (29), and this increased stability is likely to come from binding energy. Such binding includes the interaction of chaperone Hsp47 with triple-helical procollagen in the endoplasmic reticulum (Fig. 1), although its contribution to stability is controversial (30, 31). In the Golgi, lateral aggregation of procollagen molecules is observed (32–34), and longer aggregates are seen in secretory vesicles (35). Early forms of fibrils with an axial 67-nm periodicity are found in cell invaginations, followed by growth to mature fibrils (36). Mammalian collagen fibrils have a thermal stability near 60°C, much higher than body temperature, and stability is no longer an issue. During this entire complex process, there is little evidence of soluble procollagen or collagen molecules being present in an unbound form for any significant time in the intracellular or extracellular space. Analogous to the situation in fibrillar collagens, the presence of supramolecular structures in non-fibrillar collagens (1) and the collagenous domains of host-defense proteins (37) may promote stabilization. For example, the 78-residue repeating Gly-X-Y domain of C1q has a $T_m = 46$°C, well above body temperature, which is likely to come from the hexamer association of triple-helical molecules (38).

Type I collagen is a particularly difficult system for attaining reversibility.

The presence of collagen in complexes and aggregates eliminates the biological need to have individual molecules stable for long times at body temperature, yet the observation of collagen T_m values below body temperature may have physiological relevance. Having a global stability just below body temperature makes it more likely that regions of lower stability undergo "microunfolding," defined as reversible local structural perturbation (18, 39, 40). Such locally mobile regions would favor self-association and interactions in collagen molecules and fibrils. Local flexibility of the triple-helix has been implicated in recognition of the matrix metalloproteinase cleavage site in collagen (41), binding of a monoclonal antibody to type III collagen (42), heparin binding to the collagen tail of acetylcholinesterase (43), and the temperature-dependent ligand binding by the triple-helix domain of the macrophage scavenger receptor (44). Less stable sites could also play a role in degradation, either of newly synthesized collagen in proteosomes or of collagen fibrils by matrix metalloproteinases.

In addition to a physiological role of unstable local regions in binding and breakdown, inherent instability could lead to pathological consequences when there are mutations in the collagen triple-helix (16). Mutations in collagens have been implicated in a range of hereditary connective tissue diseases, including osteogenesis imperfecta (type I collagen), various chondrodysplasias (type II collagen), Ehlers-Danlos syndrome (type III collagen), dystrophic form of epidermolysis bullosa (type VII collagen), and Alport syndrome (type IV collagen; ref. 2–4). Missense mutations leading to a Gly substitution that breaks the $(Gly\text{-}X\text{-}Y)_n$ repeating pattern in the collagen triple-helix are common molecular defects, and such mutant collagens appear to have defective folding and a small decrease in thermal stability (8, 13, 45). Leikin points out that even a small decrease in stability could have a dramatic effect on the collagen unfolding rate at body temperature (16). Detailed investigations of the rate-dependent folding and unfolding of mutant collagens will clarify whether molecular instability is an important consideration in understanding the etiology of connective tissue diseases.

We thank Drs. Robert Trelstad and Peter Kahn for helpful comments.

1. Kielty, C. M., Hopkinson, I. & Grant, M. E. (1993) in *Connective Tissue and Its Hereditable Disorders*, eds. Royce, P. M. & Steinmann, B. (Wiley–Liss, New York), pp. 103–147.
2. Byers, P. H. (1993) in *Connective Tissue and Its Hereditable Disorders*, eds. Royce, P. M. & Steinmann, B. (Wiley–Liss, New York), pp. 351–407.
3. Prockop, D. J. & Kivirikko, K. I. (1995) *Annu. Rev. Biochem.* **64,** 403–434.
4. Myllyharju, J. & Kivirikko, K. I. (2001) *Ann. Med.* **33,** 7–21.
5. Rich, A. & Crick, F. H. C. (1961) *J. Mol. Biol.* **3,** 483–506.
6. Ramachandran, G. N. (1964) in *Treatise on Collagen,* ed. Ramachandran, G. N. (Academic, New York), Vol. 1, pp. 103–183.
7. Bella, J., Eaton, M., Brodsky, B. & Berman, H. M. (1994) *Science* **266,** 75–81.
8. Engel, J. & Prockop, D. J. (1991) *Annu. Rev. Biophys. Biophys. Chem.* **20,** 137–152.
9. Nagase, H. & Woessner, J. F., Jr. (1999) *J. Biol. Chem.* **274,** 21491–21494.
10. Mathews, M. B. (1975) *Connective Tissue: Macromolecular Structure and Evolution* (Springer, New York).
11. Baker, A. T., Ramshaw, J. A. M., Chan, D., Cole, W. G. & Bateman, J. F. (1989) *Biochem. J.* **261,** 253–257.
12. Bachinger, H. P., Morris, N. P. & Davis, J. M. (1993) *Am. J. Med. Gen.* **45,** 152–162.
13. Raghunath, M., Bruckner, P. & Steinmann, B. (1994) *J. Mol. Biol.* **236,** 940–949.
14. Burjanadze, T. V. (2000) *Biopolymers* **53,** 523–528.
15. Persikov, A. V., Ramshaw, J. A., Kirkpatrick, A. & Brodsky, B. (2000) *Biochemistry* **39,** 14960–14967.
16. Leikina, E., Mertts, M. V., Kuznetsova, N. & Leikin, S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 1314–1318.
17. Privalov, P. L. (1979) *Adv. Protein Chem.* **33,** 167–241.
18. Privalov, P. L. (1982) *Adv. Prot. Chem.* **35,** 1–104.
19. von Hippel, P. H. (1967) in *Treatise on Collagen*, ed. Ramachandran, G. N. (Academic, London), Vol. 1, pp. 253–338.
20. Traub, W. & Piez, K. A. (1971) *Adv. Prot. Chem.* **25,** 243–352.
21. Engel, J. & Bachinger, H. P. (2000) *Matrix Biol.* **19,** 235–244.
22. Bachinger, H. P. & Engel, J. (2001) *Matrix Biol.* **20,** 267–269.
23. Davis, J. M. & Bachinger, H. P. (1993) *J. Biol. Chem.* **268,** 25965–25972.
24. Miles, C. A. (1993) *Int. J. Biol. Macromol.* **15,** 265–271.
25. Miles, C. A., Burjanadze, T. V. & Bailey, A. J. (1995) *J. Mol. Biol.* **245,** 437–446.
26. Miles, C. A. & Bailey, A. J. (2001) *Matrix Biol.* **20,** 263–265.
27. Potekhin, S. A. & Kovrigin, E. L. (1998) *Biophys. Chem.* **73,** 241–248.
28. Sanchez-Ruiz, J. M. (1992) *Biophys. J.* **61,** 921–935.
29. Bruckner, P. & Eikenberry, E. F. (1984) *Eur. J. Biochem.* **140,** 397–399.
30. Tasab, M., Batten, M. R. & Bulleid, N. J. (2000) *EMBO J.* **19,** 2204–2211.
31. Macdonald, J. R. & Bachinger, H. P. (2001) *J. Biol. Chem.* **276,** 25399–25403.
32. Trelstad, R. L. & Hayashi, K. (1979) *Dev. Biol.* **71,** 228–242.
33. Hulmes, D. J., Bruns, R. R. & Gross, J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 388–392.
34. Bonfanti, L., Mironov, A., Jr., Martinez-Menarguiez, J. A., Martella, O., Fusella, A., Baldassarre, M., Buccione, R., Geuze, H. J., Mironov, A. A. & Luini, A. (1998) *Cell* **95,** 993–1003.
35. Trelstad, R. L. (1971) *J. Cell Biol.* **48,** 689–694.
36. Birk, D. E. & Trelstad, R. L. (1984) *J. Cell Biol.* **99,** 2024–2033.
37. Hoppe, H. J. & Reid, K. B. (1994) *Protein Sci.* **3,** 1143–1158.
38. Brodsky-Doyle, B., Leonard, K. R. & Reid, K. B. (1976) *Biochem. J.* **159,** 279–286.
39. Ryhanen, L., Zaragoza, E. J. & Uitto, J. (1983) *Arch. Biochem. Biophys.* **223,** 562–571.
40. Kadler, K. E., Hojima, Y. & Prockop, D. J. (1988) *J. Biol. Chem.* **263,** 10517–10523.
41. Fields, G. B. (1991) *J. Theor. Biol.* **153,** 585–602.
42. Shah, N. K., Sharma, M., Kirkpatrick, A., Ramshaw, J. A. & Brodsky, B. (1997) *Biochemistry* **36,** 5878–5883.
43. Deprez, P., Doss-Pepe, E., Brodsky, B. & Inestrosa, N. C. (2000) *Biochem. J.* **350,** 283–290.
44. Doi, T., Higashino, K., Kurihara, Y., Wada, Y., Miyazaki, T., Nakamura, H., Uesugi, S., Imanishi, T., Kawabe, Y., Itakura, H., *et al.* (1993) *J. Biol. Chem.* **268,** 2126–2133.
45. Kuivaniemi, H., Tromp, G. & Prockop, D. J. (1991) *FASEB J.* **5,** 2052–2060.