Journal of Cellular Biochemistry 28:31–37 (1985)
Extracellular Matrix: Structure and Function 183–189

Biosynthesis of Collagen

John H. Fessler, Kurt J. Doege, Keith G. Duncan, and Liselotte I. Fessler

Molecular Biology Institute and Biology Department, University of California, Los Angeles, California 90024

During the biosynthesis and assembly of collagen structures, disulfide links can serve several functions. During biosynthesis they successively stabilize intrapeptide folding and associations of three chains into one molecule. Studies on the refolding and reassociation of reduced and denatured carboxyl propeptides of procollagen I showed that successive interactions of folding and assembly are successively weaker. Disulfide bridges were reestablished within correctly refolded carboxyl propeptides. Rearrangements of disulfide bridges may occur during the processing of type V procollagen molecules as these collagens become incorporated into extracellular matrix. The basement membrane procollagen IV molecules become disulfide linked at each end into networks, and there are indications that further rearrangements of disulfide links may allow additional modulation.

Key words: assembly, biosynthesis, protein folding, disulfide links, collagen, procollagen I, III, IV, V

The biosynthesis of collagen extends from polypeptide synthesis through several intermediate, metastable states to the final connective tissue assembly of many molecules. Although it is convenient to think of discrete steps, such as peptide synthesis, post-ribosomal modification, and several assembly steps, it is unlikely that Nature heeds our conceptual boundaries. In a rapidly growing chick embryo the cells have to function fast and efficiently, and therefore it is unlikely that nascent procollagen chains remain in limbo during the 3–6 min that are required for polypeptide synthesis of various collagens. Yet there is a variety of evidence that the carboxyl propeptides, which will mostly not yet exist for any one nascent proα chain, have a role in chain selection and formation of a collagen helix. We are forced to conclude that the biosynthetic pathway of folding of procollagen chains does not necessarily proceed along the conceptually direct line of flight from random chain to finished molecule. This is concordant with studies of in vitro folding of ribonuclease and bovine trypsin inhibitor that show intermediates that need to be refolded before reaching the final form [1].

Disulfide linkages have been used as naturally occurring indices of folding. If refolding of procollagen chains occurs during their progression from nascent polypep-

K.J. Doege is now at the National Institutes of Dental Research, National Institutes of Health, Bethesda, MD 20205.

Received August 3, 1984; accepted December 1, 1984.

tides to connective tissue matrix, then some disulfide linkages may change correspondingly. Procollagens are structured in domains. Disulfide linkages can only occur in non-collagenous domains or at the boundary of a correctly folded collagen helix. For some collagens disulfide cross-linkages eventually contribute to the stability of the final matrix. We report here in three sections on (A) variations of disulfide linkages in procollagens V, (B) studies on the in vitro refolding of the carboxyl propeptides of procollagen I and their potential relation to the biosynthesis of this material, and lastly, (C) on changes in disulfide bridging of procollagen IV with synthesis and supramolecular association.

Materials and Methods are as given in the original, referenced publications.

DISULFIDE LINKAGE FORMATION DURING PROCOLLAGENS V BIOSYNTHESIS

The principal components of procollagen V are the proα1(V) and proα2(V) chains, which correspond to, respectively, B and A chains of previous nomenclature. A third chain, C, was obtained from placenta and assigned as α3(V) [2]. In chick embryo tendon we found a fourth chain and have traced its processing steps from the proα4(V) form through an intermediate to the final state [3]. This new proα4(V) chain is similar to the proα1(V) chain, but significantly differs from it in its participation in disulfide linkages, as the following will show.

First we summarize earlier conclusions on proα1(V) and proα2(V) participation in triple helical type V molecules. The homotrimer (proα1[V])$_3$ is made by hamster lung cells [4,5] but is not interchain disulfide-linked. Probably this trimer is also made in chick tissue. The heterotrimer [(proα1[V])$_2$(proα2[V])] also exists without interchain disulfide linkage, and in addition there are forms in which one or both proα1(V) chains are disulfide-linked to the proα2(V) chain. Various attempts to increase the relative amount of trimerically disulfide-linked molecules failed, suggesting that partial interchain disulfide linkage was a characteristic of procollagens V. All interchain disulfide links between proα1(V) and proα2(V) are lost in the first proteolytic processing step, which removes an approximately 30,000–35,000-dalton peptide from proα1(V). This peptide resembles established carboxyl propeptides of other procollagens, contains complex carbohydrate, and is internally disulfide-linked.

Results

In contrast to this relative paucity of interchain disulfide links of proα1(V), the new proα4(V) chain mostly occurs disulfide-linked either to itself, in the homotrimer (proα4[V])$_3$, or linked to a proα2(V) chain in the heterotrimer [(proα4[V])-S-S-(proα2[V]),proα1(V)]. In the latter heterotrimer the third component is a proα1(V) chain, which is not attached by disulfide bridges. Again, all interchain disulfide bridges are removed in the first proteolytic processing step, and this cleaves a 30,000–35,000-dalton peptide from proα4(V).

The new proα4(V) chain undergoes the same sequence of proteolytic cleavage as the proα1(V) chain. However, throughout this sequence the proα4(V) chain and its derivates behave as smaller molecules than the proα1(V) chain and derivative counterparts. The differences are shown in SDS-acrylamide gel electrophoresis and in velocity sedimentation. Only part of the greater mass of the proα1(V) chain and derivates is due to some complex carbohydrate that is attached to the non-collagenous

region that is retained after the last processing step, and this is absent, throughout, from the proα4(V) chain and derivates.

Thus the proα1(V) and proα4(V) chains differ at least in two non-collagenous domains at opposite ends of the collagen helix. Peptide maps show both similarities and differences, and further characterization remains to be done. The maps are unlike those expected for putative chick α3(V) chains. The products of pepsin digestion of tissue containing both proα1(V) and proα4(V) in native form are indistinguishable by SDS-acrylamide gel electrophoresis. These materials are either products of two different genes, or differently spliced transcripts of a single gene. We have not found proα4(V) in chick crop or blood vessels, and it may be a specialization for whatever role collagen V has in the ordered fibrillar matrix of tendon.

Discussion

The improved ability to form interchain disulfide bridges might be related to an intracellular event, as proteolytic cleavage removes these linkages relatively early after synthesis. However, something more complex, which we do not understand, happens at this first processing step. Independently of whether the proα2(V) chains were matched in triplets to proα1(V) or proα4(V) chains in procollagen V, and independently of the state of interchain disulfide-linkage in the procollagen V, all the pα2(V) chains of the immediately following p-collagen V have become disulfide-linked to two or more so-called P peptides of approximately the same size as the peptides that have been cleaved off the proα1(V) and proα4(V) chains. As the ion-exchange chromatographic properties of procollagens V and p-collagens V are substantially different, there must have been a greater change than just splitting of some peptide linkages without separation of the cleavage products. We know neither the origin nor the identity of these P peptides. It seems that a rearrangement of disulfide linkages accompanies the processing of proα(V) chains to p-collagen V.

DISULFIDE LINKAGE FORMATION DURING BIOSYNTHESIS AND REFOLDING OF PROCOLLAGENS I AND III

Appreciable evidence suggests that interchain disulfide linkage between carboxyl propeptides contributes to efficient chain selection and molecular assembly of procollagens I and III. When helix formation of procollagen I is blocked by inhibiting proline hydroxylation, the carboxyl propeptides of procollagen I still become disulfide-linked to each other as heterotrimers of correct composition [6]. The carboxyl propeptides of procollagen III also form correct, disulfide-linked homotrimers under these conditions in tissue (and probably cells) that make both procollagen I and procollagen III [7]. In contrast, disulfide linkages between the amino propeptides of the three proαIII chains cannot form under these conditions, though they are made after hydroxylation and helix formation are resumed in previously synthesized molecules. This suggests that while disulfide links between carboxyl propeptides could help to initiate correct collagen helix formation, disulfide links between amino propeptides III could only stabilize a previously formed helix. In support of this view, it was shown that in normal, uninhibited tissue the disulfide links between the carboxyl propeptides III are made before those between the amino propeptides III [8]. The concept that correct association at the carboxyl end is vital to correct helix formation is consistent with in vitro refolding studies on collagen III, pN collagen III,

and procollagen I monitored by optical and protease-resistance methods [9–11]. A statistical analysis of proteolysis of newly synthesized collagens was consistent with the proposal that the same direction of folding occurs in cells [12], and chains that lack carboxyl ends fail to be exported from cells as helical molecules [13].

Results

The carboxyl propeptides must have the necessary ability for mutual selection, and only after this has occurred can the arrangement be stabilized by disulfide links between the propeptides. The folding of each separate carboxyl propeptide is also stabilized by other, intra-peptide disulfide links.

To investigate this further, carboxyl propeptides were fully reduced and denatured in 6 M urea and then allowed to refold at successively decreasing urea concentrations. The state of refolding was monitored by ascertaining which disulfide bridges could be reformed. Tryptic peptide mapping of optimally refolded carboxyl propeptides indicated restoration of the original, biosynthetically formed intra-peptide disulfide bridges. The investigation showed that stronger interactions drive the internal folding of each carboxyl propeptide than the selective association between them. While these disulfide linkages within individual peptides could be restored in 2 M urea buffer, as judged by electrophoretic mobility, no disulfide-linked pairs or trimers of carboxyl propeptides could be formed until the concentration of urea was decreased to 0.5 M or less. This is consistent with a scheme for assembly of triple-chain molecules in which each newly made carboxyl propeptide folds itself first, and thereby creates the requisite mutual recognition sites, and only subsequently interacts with two other carboxyl propeptides. However, the biosynthetic process will differ from the in vitro study in relative timing. It is well known that the efficiency of reforming oligomeric proteins from fully denatured protomers is lowered by non-specific aggregation into insoluble polymers as the concentration of denaturant is decreased. In contrast, as the newly made proα chain is extruded into the lumen of the rough endoplasmic reticulum, successive portions of each carboxyl propeptide will have a preferential opportunity to interact with preceding segments of the same polypeptide, and this will minimize non-specific aggregation at this stage. By the time of completion of peptide synthesis, the folding of the recognition sites for two other carboxyl propeptides may be ready.

Discussion

These results suggest that in sequence of folding and assembly steps, the later interactions will be weaker. In other words, after the right interactions between folded individual carboxyl propeptides have occurred, subsequently established disulfide links may be necessary to maintain the association. This concept was supported by the results of mild reduction of the trimeric carboxyl propeptide [$(C1)_2C2(I)$], which first cleaves only interchain disulfide links [14–16]: this readily causes dissociation under conditions of ultracentrifugal analysis [15], and if the monomeric carboxyl propeptides are carboxymethylated, they do not reassociate [14]. If the propeptides are not carboxymethylated and are retained in each other's vicinity by attached, partly denatured collagen chains, then the correct disulfide-linked state can be readily restored.

We do not know the conditions of the newly formed proα chains within the lumen of the rough endoplasmic reticulum, but several factors could assist the

selective interactions of the carboxyl propeptides. They are likely to be at a high concentration, might be mutually aligned by the inner surface of the endoplasmic reticulum, and could become loosely associated by partial interactions of the nascent collagenous sequences. These have been reported to fold helically even while still associated with polyribosomes [17,18], and to form pepsin-resistant structures during cell-free translation [19].

Short pulse-labeled calvaria contain associations of proα chains that are not disulfide-linked and have properties expected for incomplete associations of three chains. Their behavior clearly differs from quenched, denatured procollagen chains and they sediment as an entity slightly faster than native procollagen molecules [15]. They are probably imperfectly fitted, trimeric procollagen molecules, and could represent forms that just precede the correctly aligned and disulfide-fixed arrangement. Although they decrease during the chase phase of an experiment, they could also represent aberrant forms that are progressing more slowly than the majority of the molecules toward the final form. With the discovery of partial deletion mutations in human procollagens by Prockop and associates, the question of how cells handle aberrant molecular arrangements of procollagen chains has become acute. The extensive studies on degradation of newly made chains to dialyzable, hydroxyproline-containing peptides are beyond the scope of this paper.

DISULFIDE LINKAGE FORMATION DURING BIOSYNTHESIS AND SUPRAMOLECULAR ASSEMBLY OF PROCOLLAGEN IV

Procollagen IV is exported even when helix formation is deranged, as it is by very high synthetic rates, inadequate ascorbic acid supplementation, or inhibition of hydroxylation. While newly made procollagen IV molecules are normally interchain disulfide-linked at more than one location [20], these linkages are missing in such deranged molecules, and subsequent, supramolecular association of four molecules fails to occur [21]. Thus there appears to be a reciprocal relationship between correct folding and interchain disulfide links of procollagen IV, and the establishment of both is much slower than in procollagen I or III. Although the carboxyl propeptides of procollagen IV contain intra-peptide disulfide links, they are not disulfide-linked to each other [22]. Furthermore, carboxyl propeptides IV isolated by us showed relatively weak mutual affinity. We suspect that chain selection may also be less stringent than for procollagen I, and this may account for some of the conflicting evidence of hetero and homotrimers of proα1(IV) and proα2(IV). In other words, the intracellular assembly mechanism for procollagen IV may be less refined, and slower than for procollagen I. Additional difficulties for helix folding are likely to arise from the interruptions of the Gly-X-Y sequence monotony in collagen IV.

The formation of disulfide links by procollagen IV chains seems to be an amusing and complicated game of musical chairs that proceeds beyond intracellular synthesis to supramolecular association, as far as we have obtained experimental glimpses of it. That disulfide links between procollagen I chains can be formed intracellularly has been shown in various pulse-chase studies [15,23–25], and we presume that the intra- and inter-peptide disulfide links of procollagen IV are also initially made within cells. However, either some free SH groups remain and/or the linkage system can be altered by disulfide interchange.

Results

We find that the small amounts of cysteine-cystine mixture present in standard Dulbecco's modified Eagle's medium influence the state of disulfide linkage of procollagen IV secreted by cultured cells [21]. Only after disulfide exchange, catalyzed by a mixture of reduced and oxidized glutathione, can the molecules proceed to the next stage of disulfide linkage, namely, four molecules with overlapping and disulfide-linked amino ends, ie, to form the "7S collagen" junctional complex. If cystine is omitted from the medium, the covalently linked tetrameric complex forms spontaneously, especially on warming [21]. However, the tetramers formed under the influence of glutathione sediment 8% faster than those formed in cystine-free medium. The electron microscopic length of the "7S collagen" junctional complex is almost the same in the two tetramers, so that differences in sedimentation velocity are not due to differences in the disulfide-linked, overlapping portion of the molecules. The electron microscopic contour lengths of the four projecting "arms" or molecules are also not significantly different. However, we find that the relationship between contour length and the direct end-to-end distance between the carboxyl end of an arm and its point of exit from the 7S junctional complex is changed. This is equivalent to increasing the flexibility of the thread-like projecting molecules or arms and accounts well for the decreased hydrodynamic resistance of the tetramer formed under the action of glutathione.

Discussion

We can interpret this apparent change in flexibility in a more specific way. As reported elsewhere [20], a type IV specific collagenase cuts each molecule at a point abount 90 nm from the point of exit from the junctional complex. The electron microscopic relationship of end-to-end distance and contour length of the major, cut-off portion of the thread-like molecule that retains the carboxyl propeptides attached at the "peripheral" end indicates the same flexibility as that measured for collagen I helices [26], and this peripheral portion of the molecule is "stiffer" than the uncut arms of the original tetramers. Therefore the residual, 90-nm central portion must contain a considerably more flexible region. This is also consistent with the results of Hofman et al [26], who used a different approach. We infer that the disulfide interchanges brought about by reaction with glutathione caused changes in this more central portion, which also contains interruptions of the collagen helix. The usefulness of a "hinge region" in facilitating formation of a branched, network aggregate is known from the structure and function of immunoglobulins, and may readily have an equivalent role in basement membrane collagen networks. Our results suggest that disulfide interchange may further modulate this hinge region.

ACKNOWLEDGMENTS

We thank the U.S. Public Health Service for support through grants AM-13748 and AG-02128. K.J.D. and K.G.D. were supported by USPHS Cellular and Molecular Biology Training Grant GM-07185.

REFERENCES

1. Kim PS, Baldwin RL: Ann Rev Biochem 51:459–489, 1982.

2. Bornstein P, Sage H: Ann Rev Biochem 49:957–1003, 1980.
3. Fessler LI, Shigaki N, Fessler JH: Fed Proc 42 (abs. 753):1886, 1983.
4. Fessler LI, Robinson WJ, Fessler JH: J Biol Chem 256:9646–9651, 1981.
5. Haralson MA, Mitchell WM, Rhodes RK, Kresina TF, Gay R, Miller EJ: Proc Natl Acad Sci USA 77:5206–5210, 1980.
6. Fessler LI, Fessler JH: J Biol Chem 240:7637–7646, 1974.
7. Fessler LI, Fessler JH: J Biol Chem 254:233–239, 1979.
8. Bächinger HP, Fessler LI, Timpl R, Fessler JH: J Biol Chem 256:13195–13199, 1981.
9. Bächinger HP, Bruckner P, Timpl R, Prockop DJ, Engel J: Eur J Biochem 106:619–632, 1980.
10. Engel J, Bächinger HP, Bruckner P, Timpl R: In Jaenicke R (ed): "Protein Folding." New York: Elsevier-North Holland Biomedical Press, 1980, pp 345–368.
11. Gerard S, Puett D, Mitchell WM: Biochemistry 20:1857–1865, 1981.
12. Bruckner P, Eikenberry EF, Prockop DJ: Eur J Biochem 118:607–613, 1981.
13. Rosenbloom J, Endo R, Harsch M: J Biol Chem 251:2070–2076, 1976.
14. Olsen BR, Guzman NA, Engel J, Condit C, Aase S: Biochemistry 16:3030–3036, 1977.
15. Doege KJ: Ph.D. Thesis, UCLA, 1983.
16. Kosen PA, Franzblau C: Biochemistry 21:4278, 1982.
17. Veis A, Brownell AG: Proc Natl Acad Sci USA 74:902–905, 1977.
18. Brownell AG, Veis A: J Biol Chem 251:7137–7143, 1976.
19. Monson JM: Collagen Rel Res 3:1–12, 1983.
20. Fessler LI, Salo T, Fessler JH, Tryggvasson K: J Biol Chem 259:9783–9789,1984.
21. Duncan KG, Fessler LI, Bächinger HP, Fessler JH: J Biol Chem 258:5869–5877, 1983.
22. Fessler LI, Fessler JH: J Biol Chem 257:9804–9810, 1982.
23. Schofield JD, Uitto J, Prockop DJ: Biochemistry 13:1801–1806, 1974.
24. Morris NP, Fessler LI, Weinstock A, Fessler JH: J Biol Chem 250:5719–5726, 1975.
25. Kao WW-Y, Mai SH, Chou KL, Ebert J: J Biol Chem 258:7779–7787, 1983.
26. Hofmann H, Voss T, Kuhn K, Engel J: J Mol Biol 172:325–343, 1984.