Volume 217, number 2, 216–220 FEB 04853 June 1987

Hypothesis

Disulfide bonding as a determinant of the molecular composition of types I, II and III procollagen

Juha Koivu

Collagen Research Unit, University of Oulu, Department of Medical Biochemistry, Kajaanintie 52 A, SF-90220 Oulu, Finland

Received 13 March 1987; revised version received 30 April 1987

Procollagen molecules have amino-terminal and carboxy-terminal propeptides at the respective ends of the collagenous triple helix. The carboxy-terminal propeptides enhance and direct the association of proα-chains into procollagen molecules, but the mechanism of this registration function is still obscure. A hypothesis concerning the function of disulfide bonding in the assembly of types I, II and III procollagen is put forward here.

Procollagen propeptide; Procollagen assembly; Protein disulfide isomerase

1. INTRODUCTION

The precursor molecules of collagens, procollagens, consist of three polypeptide chains, proα-chains, having five domains in their structure: the amino-terminal propeptide is followed by a short telopeptide region, after which the triple helical collagenous domain follows. At the carboxyl end of the triple helix another telopeptide region separates the helical part from the carboxy-terminal propeptide (for a review see [1]). Type I procollagen usually contains proα-chains of two kinds: two proα_1(I)-chains and one proα_2(I)-chain, while type I homotrimers, found in trace amounts in certain tissues, are composed of three proα_1(I)-chains. Types II and III procollagen consist of three identical chains, named proα_1(II)- and proα_1(III)-chains, respectively (see [1]).

The biosynthesis of collagen is characterized by extensive co- and post-translational modification of the newly synthesized polypeptide chains. The intracellular hydroxylating and glycosylating enzymes responsible for these modifications act only on unfolded proα-chains, the triple helix stopping these reactions (see [1]). The degree of intracellular modifications, e.g. proline 3-hydroxylation, lysine hydroxylation and glycosylation of hydroxylysine residues, varies between procollagens, type I heterotrimers and type III trimers being less extensively modified than type II procollagen and type I homotrimers (see [1]). This is generally believed to be mainly due to variation in the time required for the formation of the triple helical molecules, as indicated by in vivo measurements, which give folding times of 10 and 20 min for the procollagens of types I and II, respectively (see [1]).

It has been known for several years that disulfide bond formation between carboxy-

Correspondence address: J. Koivu, Collagen Research Unit, University of Oulu, Department of Medical Biochemistry, Kajaanintie 52 A, SF-90220 Oulu, Finland

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

terminal propeptides is essential for triple helix formation (see [1]). The assembly of type III procollagen starts with the association of folded carboxyl propeptides and is followed by formation of interchain disulfide bonds between the carboxyl propeptides, folding of the triple helix, and formation of disulfide bonds within the amino-terminal propeptides [2]. The triple helix folding seems to proceed from a single nucleus in a zipper-like fashion, in which the *cis/trans* isomerization of peptide bonds adjacent to some proline residues is the rate-limiting step [3].

In a recent study by Koivu and Myllylä [4] it was noticed that the formation of triple helical molecules in vitro took 9.4 min for type I procollagen and 17.1 min for type II procollagen in the presence of protein disulfide isomerase, which is an enzyme capable of catalyzing the formation of native disulfide bonds in various proteins. The formation of interchain disulfide bonds was a more rapid process with half-times of 3.7 and 11.6 min for type I and II procollagen, respectively. The disulfide bonding was followed by a zero-order process having constantly a half-time about 6 min and this was thought to be due to the *cis/trans* isomerization of peptide bonds. Thus, the difference between the times required for the formation of triple helical molecules was due to the different rates of disulfide bond formation within the carboxy-terminal propeptides of these procollagen types.

2. HYPOTHESIS OF DISULFIDE BONDING IN TYPES I–III PROCOLLAGEN

Newly synthesized proα-chains are in a random coil structure having free thiol groups in their cysteine residues. These are readily oxidized to random disulfides in the cell, resulting in a mixture of disulfide-bonded molecules (see [5]). The random disulfides are then rearranged into native bonds during the folding of the secondary structure of the proα-chains, protein disulfide isomerase catalyzing these rearrangement reactions. The native intrachain disulfide bonds in type I procollagen have been shown to lie between residues 5–8 and 6–7 [6,7] while residues 1 and 2 are linked by interchain bonds (fig.1) [7]. Either intrachain [8] or interchain disulfide bonds [7] have been proposed to be between cysteine residues 3 and 4 (fig.1), but the

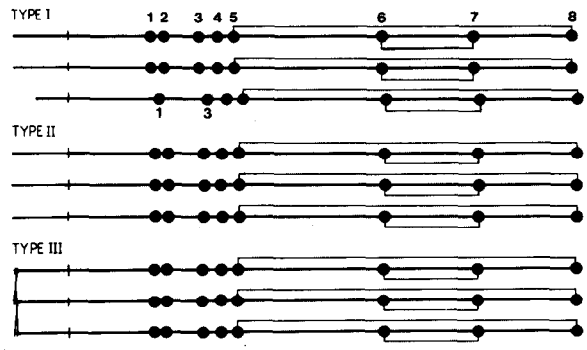

Fig.1. Schematic representation of the carboxy-terminal propeptides of types I, II and III procollagen. The telopeptide region lies to the left of the vertical bar and the carboxy-terminal propeptide to the right. The collagenous domain, not shown here, is to the left of the picture. The solid circles indicate cysteine residues, the numbers of which correspond to their ordinals in the sequence. It should be noted that type I procollagen contains two kinds of chain, the proα_2(I)-chain having 7 cysteine residues as compared with 8 cysteine residues in the proα_1(I)-chains. Type III procollagen has three additional disulfides at the end of the collagenous domain. The intrachain disulfide bonds are shown between residues 5–8 and 6–7, while the interchain disulfides lie between residues 1–2 and 3–4.

nature of these bonds does not affect the hypothesis proposed here. Unfolded proα-chains are probably unable to associate and that is why only intrachain disulfide bonds are formed at this stage (fig.2B), native bonds being formed between residues 5–8 and 6–7 during the folding of the secondary structure of carboxyl propeptides. The second cysteine residue in the proα_2(I)-chain [9] is missing as compared to the proα_1(I)-chain [10] leaving presumably the first residue as a free thiol group (fig.2B). After the association of the folded carboxyl propeptides the negative net charge of this free thiol enhances its attack on the positively charged second cysteine residue in one of the proα_1(I)-chains forming a native interchain disulfide bond (fig.2C) (for the factors affecting disulfide bonding, see [11]). This again leaves the first residue in this chain free, and it attacks the second residue in the other proα_1(I)-chain (fig.2D). In this way the free cysteine residue is able to enhance the formation of native interchain disulfide bonds without any formation of random pairings. The three proα-chains are now correctly

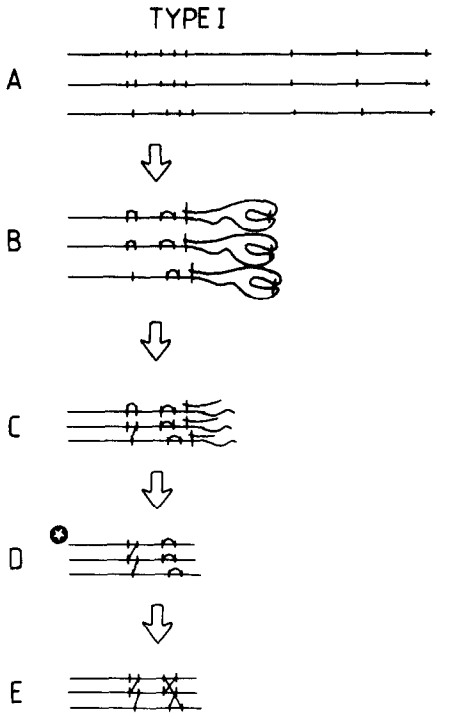

Fig.2. Schematic representation of the association of proα-chains, indicating the mechanism of disulfide bonding in type I procollagen. Only the carboxy-terminal propeptides are shown. (A) Unfolded proα-chains after protein synthesis. (B) Folding of proα-chains to obtain their secondary structure, intrachain disulfides being simultaneously formed, resulting in native disulfides between residues 5–8 and 6–7. Non-native disulfides exist between residues 1–2 and 3–4 at this stage. (C) Association of folded proα-chains. (C–E) Formation of native interchain disulfide bonds. The asterisks indicate the stage at which the proα-chains are in order alignment for triple helix folding. The parts of the propeptides containing intrachain disulfides are omitted at stages C–E for the sake of convenience.

disulfide bonded and the rest of the interchain bonds are formed as the folding of the triple helix begins.

There is a high degree of sequence homology between the carboxyl propeptides of types I–III procollagen [9,10,12,13], the sequence homology being best preserved around the location of cysteine residues. Thus, the folding process in type II procollagen is most probably very similar to that in type I procollagen. However, while there is an even number of cysteine residues in the proα_1(II)-chains [12] there may be no free thiols in the proα-chains after the folding of the secondary structure (fig.3B). That is why the formation of interchain disulfide bonds probably proceeds more slowly than in type I procollagen. Moreover, the formation of native interchain disulfides is presumably not so favored as in type I, since the free thiol directing the formation of the native interchain bonds is omitted. As a result a longer time is required for disulfide bonding and triple helix formation than in the case of type I procollagen.

There is an even number of cysteine residues within the carboxyl propeptides of type III procollagen, but there are two additional, adjacent cysteine residues at the end of the triple helical domain (fig.1) [13]. These cannot form an intrachain disulfide bond due to steric constraints and are left free during the folding of the secondary structure of carboxyl propeptides (fig.4B). Association of the proα_1(III)-chains may be facilitated by the hydrophobic telopeptide region, and three disulfide bonds are rapidly formed between these residues (fig.4C). The chains are now aligned for

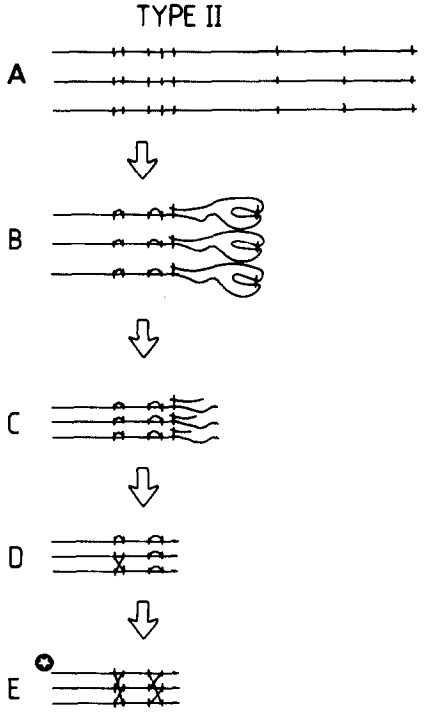

Fig.3. Schematic representation of the association of proα-chains in type II procollagen. See the legend to fig.2 for further details.

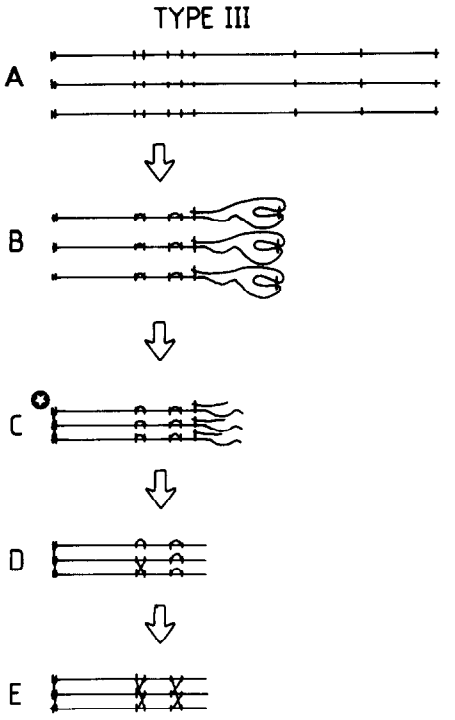

Fig.4. Schematic representation of the association of proα-chains in type III procollagen. The carboxy-terminal propeptide and the telopeptide region are shown. See the legend to fig.2 for further details.

triple helix folding, and the folding begins while the disulfides in the carboxyl propeptides are still in the process of formation.

3. DISCUSSION

The above-mentioned hypothesis is able to explain some questions that have puzzled scientists in collagen biosynthesis. The variable degree of certain post-translational modification in different types of procollagen may be explained, at least partially, by variable amounts of enzymes catalyzing these modifications in cells synthesizing different collagen types [14]. However, this explanation alone is not sufficient since while type I heterotrimers and homotrimers are synthesized by the same cells, there are significant differences in the degree of post-translational modification between these molecules [15]. The proα_2(I)-chain is missing in type I homotrimers and there are no free thiol groups at the moment of chain association (fig.3B,C). Thus interchain disulfide bonding is probably delayed as in type II procollagen and the enzymes responsible for the co- and post-translational modification have more time to operate resulting in a higher degree of these modifications as has been found by Uitto [15]. These findings clearly indicate the significance of the unpaired cysteine residue in the proα_2(I)-chain during the assembly of type I heterotrimers.

Another unresolved problem in collagen biosynthesis is that heterotrimers composed of proα-chains of both type I and type III procollagen have never been found, although several cells are able to synthesize simultaneously these collagen types (see [1]). The difference in the association and interchain disulfide bonding between procollagen types I and III however may explain this finding, offering a rather simple mechanism for the determination of proα-chains that are to be assembled with each other.

For a more definitive review of the structure and function of the carboxy-terminal propeptides of the major procollagens see [16].

ACKNOWLEDGEMENTS

The author wishes to thank Dr Raili Myllylä and Dr Deborah Kaska for critical reading of the manuscript.

REFERENCES

[1] Kivirikko, K.I. and Myllylä, R. (1984) in: Extracellular Matrix Biochemistry (Piez, K.A. and Reddi, A.H. eds) pp.83–118, Elsevier, Amsterdam, New York.
[2] Bächinger, H.P., Fessler, L.I., Timpl, R. and Fessler, J.H. (1981) J. Biol. Chem. 256, 13193–13199.
[3] Bächinger, H.P., Bruckner, P., Timpl, R., Prockop, D.J. and Engel, J. (1980) Eur. J. Biochem. 106, 619–632.
[4] Koivu, J. and Myllylä, R. (1987) J. Biol. Chem., in press.
[5] Anfinsen, C.B. (1973) Science 181, 223–230.
[6] Olsen, B.R. (1982) in: Trends in Basement Membrane Research (Kühn, K. et al. eds) pp.225–236, Raven, New York.
[7] Koivu, J. (1987) FEBS Lett. 212, 229–232.
[8] Kühn, K. (1986) in: Connective Tissue: Biological and Clinical Aspects (Kühn, K. and Krieg, T. eds) Rheumatology, An Annual Review, vol.10, pp.29–69, Karger, Basel.

[9] Bernard, M.P., Myers, J.C., Chu, M.-L., Ramirez, F., Eikenberry, E.F. and Prockop, D.J. (1983) Biochemistry 22, 1139–1145.
[10] Bernard, M.P., Chu, M.-L., Myers, J.C., Ramirez, F., Eikenberry, E.F. and Prockop, D.J. (1983) Biochemistry 22, 5213–5223.
[11] Snyder, G.N., Cennerrazzo, M.J., Karalis, A.J. and Field, D. (1981) Biochemistry 20, 6509–6519.
[12] Sangiorgi, F.O., Benson-Chanda, V., De Wet, W.J., Sobel, M.E., Tsipouras, P. and Ramirez, F. (1985) Nucleic Acids Res. 13, 2207–2225.
[13] Loidl, H.R., Brinker, J.M., May, M., Pihlajaniemi, T., Morrow, S., Rosenbloom, J. and Myers, J.C. (1984) Nucleic Acids Res. 12, 9383–9394.
[14] Keller, H., Eikenberry, E.F., Winterhalter, K.H. and Bruckner, P. (1985) Collagen Rel. Res. 5, 245–251.
[15] Uitto, J. (1979) Arch. Biochem. Biophys. 192, 371–379.
[16] Dion, A.S. and Myers, J.C. (1987) J. Mol. Biol. 193, 127–143.