Volume 320, number 3, 211–214 FEBS 12292 April 1993

Analysis of the role of the COL1 domain and its adjacent cysteine-containing sequence in the chain assembly of type IX collagen

Lionel Labourdette and Michel van der Rest

Institut de Biologie et Chimie des Protéines, CNRS UPR 412, 7 passage du Vercors, 69367 Lyon Cedex 07, France

Received 19 February 1993

The mechanisms of chain selection and assembly of type IX collagen, a heterotrimer α1(IX)α2(IX)α3(IX), must differ from that of fibrillar collagens since it lacks the characteristic C-propeptide of these latter molecules. We have tested the hypothesis that the information required for this process is contained within the C-terminal triple helical disulfide-bonded region (LMW). The reassociations of the purified LMW fragments of pepsinized bovine type IX collagen were followed by the formation of disulfide-bonded multimers. Our data demonstrate that only three triple helical assemblies form readily, $(\alpha 1)_3$, $(\alpha 2)_3$ and α1α2α3. The information required for chain selection and assembly is thus, at least in part, contained in the studied fragments. Molecular stoichiometries different from the classical heterotrimer may thus also form under certain conditions.

Type IX collagen; Collagen trimerization; Triple helix; Protein folding

1. INTRODUCTION

Collagens are proteins of the extracellular matrix that are characterized by the presence of one or more domains in a specific triple helical conformation with a repetitive -(Gly-Xaa-Yaa)$_n$- primary sequence (for review see [1]). Sixteen different collagen types have been described so far [2]. Collagen molecules can constitute three identical or different polypeptide chains. Folding of the molecule in the proper conformation requires several steps, such as chain selection (in the proper stoichiometry), chain association (in correct registration) and triple helix formation. For fibrillar collagens, the published data suggest that the C-terminal propeptides may play a fundamental role in the initial steps of chain association [3,4] while the triple helix folds by a zipper-like mechanism [5].

This model of assembly, however, cannot be directly applied to the FACITs (fibril associated collagens with interrupted triple helix). For this group of collagen molecules, the size of the non-triple helical C-terminal domain (NC1) is significantly shorter (less than 30 amino acids for type IX collagen, over 260 residues for type I procollagen) and shows no homology with the C-terminal globular domain of the other collagen types. FACIT chains, however, display a remarkable homology in their first triple helical domains, COL1, and in the first five residues of their NC1 domains. The sizes of the COL1 domains are similar (115, 103, 112 and 106 amino acids for collagens IX, XII, XIV and XVI, respectively). Two imperfections in the triple helix of COL1 are similarly located in all chains. At the COL1–NC1 junction, two cysteines, separated by 4 amino-acid residues and responsible for interchain disulfide bond formation, are strictly conserved. All these observations suggest that the COL1 domain may play an essential role in trimeric association of the FACITs and in the triple helical folding of the molecule.

In this study, we have investigated the ability of the chains of the pepsin-resistant low molecular weight (LMW) fragment of bovine type IX collagen to reassemble in vitro into trimers. This fragment is made of the COL1 domain and the cysteine-containing region of NC1.

Correspondence address: M. van der Rest, Institut de Biologie et Chimie des Protéines, 7, Passage du Vercors, 69367 Lyon Cedex 07, France. Fax: (33) 72 72 26 02.

2. EXPERIMENTAL

2.1. *Purification of the LMW fragment of type IX collagen*

The pepsin fragments of type IX collagen were solubilized from fetal calf cartilage as previously described by Ricard-Blum et al. [6]. We have separated the LMW fragment of the other pepsin-resistant fragments using gel permeation chromatography as described by Dayer et al. [7].

2.2. *Separation of individual LMW chains*

The chains of the purified LMW fragments were dissociated after reduction and thermal denaturation (15 min at 60°C in 50 mM Tris-HCl, pH 7.5, 9 M urea, 10 mM dithiothreitol) and separated by HPLC on a C_{18} reverse-phase column (Vydac semi prep) with a linear gradient of acetonitrile (20–50%) in 0.1% trifluoroacetic acid over 60 min. Purified chains were kept at −20°C in 0.1% trifluoroacetic acid and freeze-dried immediately prior to use to avoid possible formation of disulfide bonds during storage.

2.3. *Chain reassociation assays*

The purified chains were dissolved at a concentration of approximately 50 μg/ml in reassociation buffer composed of 50 mM Tris-HCl,

pH 7.5, 100 mM NaCl. The solution was kept at 60°C for 15 min and then quickly cooled to room temperature (23°C). After 15 h (unless otherwise specified) samples were lyophilized and analysed by SDS-PAGE on 12% acrylamide gels. For some experiments the reassociation buffer contained 2 mM of reduced glutathione. After the thermal denaturation, oxidative conditions were re-established by adding an equal volume of 20 mM of oxidized glutathione in the same buffer.

2.4. *Trypsin digestion assays*

The triple helical structure of the reassociation products was tested by digestion with TPCK-treated trypsin under the conditions described by Bruckner et al. [8] Briefly, samples (50 μg/ml) were incubated for 1 h at room temperature with trypsin (10 μg/ml) in 50 mM Tris-HCl, pH 7.6, 100 mM NaCl, immediately dried in a speed-vac and analysed by a SDS-PAGE.

3. RESULTS

Type IX collagen was chosen to study the mechanisms of chain association in the FACITs because it is a heterotrimer. The steps of chain selection can thus be investigated. We have decided to use the bovine molecule because the three chains of the LMW fragment of bovine type IX collagen have very distinct migrations on SDS-PAGE compared, for example, to the chicken fragments [9]. Stoichiometry of the reassociation products can thus be easily determined according to their migration.

When the three purified chains are mixed together and allowed to reassociate, five major bands can be observed by SDS-PAGE (Fig. 1). The upper band migrates at the same distance as the native LMW fragment, and is thus likely to correspond to the α1α2α3 heterotrimer. To better identify the reassociation products, we have tested the ability of the individual chains and of defined mixtures to reassociate. The migrations of the reassociation products could then be established

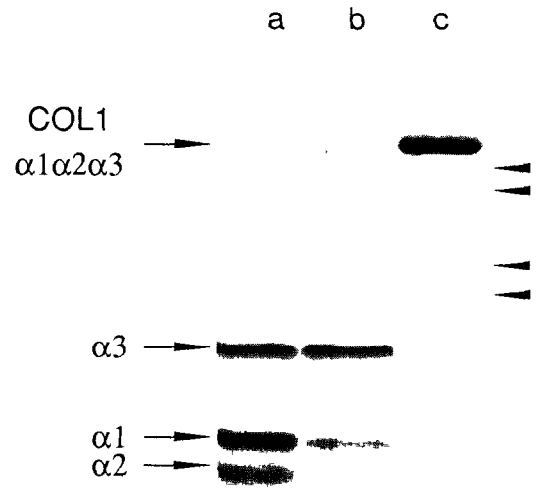

Fig. 1. Reassociation of the mixed LMW polypeptide chains. Lane a, purified native trimeric LMW fragment after reduction with dithiothreitol, lane b, reassociation assay of a mixture of purified individual chains; lane c, unreduced purified native trimeric LMW fragment. The migrations of the heterotrimer and of individual chains are indicated to the left of the figure The identity of the chains has been previously established by N-terminal sequencing [7]. The migrations of the four additional bands observed in the reconstitution experiment are indicated on the right side.

(Fig. 2). All three chains are able to form homotrimeric reassociation products, although to a much lesser extent for α3 compared to the other two chains (lanes h–j of Fig. 2). If the chains are mixed together two by two, homodimeric and homotrimeric reassociation products of the α1 and α2 chains are predominantly observed (Fig. 2, lanes e–g). We conclude that the four unidenti-

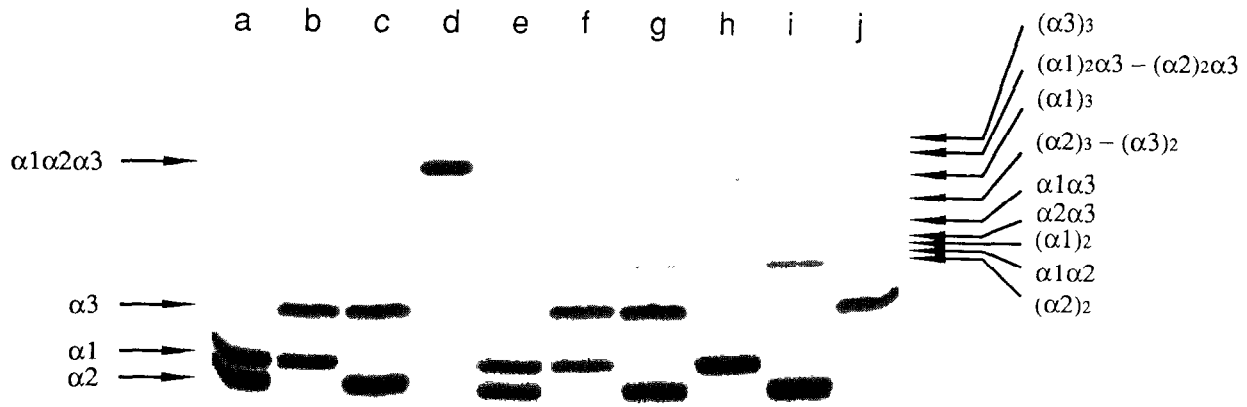

Fig. 2. Reassociation of the purified LMW polypeptide chains. Lanes a–c, mixtures of α1 and α2 fragments, of α1 and α3 fragments, and of α2 and α3 fragments, respectively, reduced with 10 mM dithiothreitol. Lane d, unreduced purified native trimeric LMW fragment. Lanes e–g, reassociation assays of the mixtures of α1 and α2 fragments (e), of α1 and α3 fragments (f), and of α2 and α3 fragments (g). Lanes h–j, homomeric reassociation assays The migrations of the heterotrimer and of individual chains are indicated to the left of the figure The migrations and the deduced stoichiometries of the reassociation products are indicated on the right of the figure

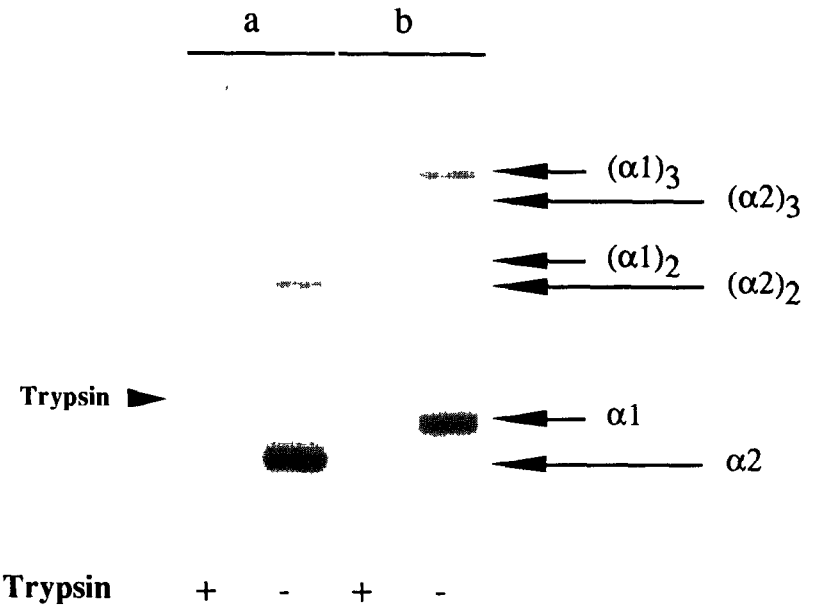

Fig. 3. Resistance of the homomeric α1 and α2 reassociation products to trypsin digestion. The triple helical conformation of reassociation products was evaluated by their resistance to trypsin digestion. Lanes a, homomeric α2 reassociation products with (+) or without (−) trypsin digestion. Lanes b, homomeric α1 reassociation products with (+) or without (−) trypsin digestion. The electrophoretic band corresponding to trypsin is indicated on the left side and the migrations of the individual chains and of their reassociation products are indicated on the right side.

fied bands observed in lane c of Fig. 1 correspond to homodimers and homotrimers of the α1 and α2 chains ($(\alpha1)_2$, $(\alpha2)_2$, $(\alpha1)_3$, $(\alpha2)_3$).

In order to study their conformation, the reassociation products have been digested with trypsin under conditions where denatured type II collagen chains, used as control, are completely degraded (not shown). The bands corresponding to the dimers and the trimers appear much more resistant to this treatment than monomers (Fig. 3), indicating that they are in a helical conformation. Whether the resistance of dimers is due to their association with non-disulfide-bonded chains or whether the dimers are by themselves resistant to this type of proteolysis has not been established. Only very little $(\alpha3)_3$ homotrimer could be obtained. Resistance of this trimer to trypsin digestion was also noted, although the weakness of the staining has made reproduction of this gel worthless.

4. DISCUSSION

The data presented in this communication demonstrate that disulfide-bonded trimeric molecules can spontaneously form when the purified component chains of the LMW fragment of type IX collagen are mixed and allowed to refold. The published sequences of the COL1 and NC1 domains of FACITs show that they contain only 2 cysteines located at the COL1–NC1 junction and separated by 4 amino acid residues [10–14]. It is known that cysteinyl residues can interact and form disulfide bonds only if they are in close enough proximity and if the biochemical environment, in particular the pH, is correct [15]. The reducibility of the observed reassociation products demonstrates a close interaction between these cysteines and, consequently, a correct folding of the COL1–NC1 region.

In the in vitro system used, we have never obtained more than ~25% disulfide-bonded trimer assembly (see Fig. 1). Comparable yields have been reported by others in similar reconstitution studies [3]. This value is obtained after less than 4 h incubation and is not increased even after 48 h (data not shown). We speculate that intrachain disulfide bond formation may limit availability of free SH groups for interchain crosslinking.

These results suggest that the information needed for a right registration of the chains and for their triple helical folding is contained in the primary sequence. Previous work on collagen-like synthetic peptides (such as $-(GPP)_n-$) demonstrated their ability to associate in triple helical structures with biophysical properties identical to those of native collagen triple helix (circular dichroism spectra, X-ray diffraction data, thermal stability...) [16–18]. It is therefore not surprising that chain reassociation occurred in our system. Three factors may have contributed to the proper registration of the chains demonstrated by the formation of the right interchain disulfide bonds: (i) peptides with all Gly-Xaa-Yaa in a registered triple helix are energetically favored over peptides with overhangs at their extremities [19]; (ii) the presence of imperfections in the repetitive sequences may favor an assembly where these are at the same position in the triple helix formed; (iii) the conformation

of the cysteine-containing region at the COL1–NC1 junction may serve as a nucleation site for helix formation.

In our in vitro assay, a chain selection process clearly exists but stoichiometries that have not been described so far in vivo are formed in addition to the $\alpha1\alpha2\alpha3$ heterotrimer. If the chains were to associate randomly, 16 species of polymers would be obtained (6 dimers and 10 trimers). However, the predominant assemblies that have been obtained in the various mixtures that were tested are homodimers and homotrimers of the $\alpha1$ and $\alpha2$ chains and the heterotrimer $\alpha1\alpha2\alpha3$. This indicates that the simultaneous presence of the three chains facilitates heteromeric combinations.

Different hypotheses can be proposed to explain the formation of chain stoichiometries, in our in vitro system, that are different from the common physiological one. Firstly, we have used pepsin fragments, and further information for chain selection may be contained in the missing parts of the chains, particularly in the NC1 domain since triple helix formation proceeds from the C-terminus to the N-terminus. Except for the cysteine-containing region, which is of course conserved in the pepsin fragments, the remaining part of this domain shows no homology between the three chains [10–12] and it is thus less likely that it has a function in chain assembly. Secondly, protein chaperones may be involved in the intracellular chain selection. Evidence for a direct role of the triple helical COL1 region on chain assembly in FACIT molecules has been recently obtained in our laboratory in a study of the expression of a minigene encoding the COL1 and NC1 domains of collagen XII [20]. The data presented here further support the involvement of the COL1 domain in the assembly of this class of collagen molecules.

Acknowledgements The authors express their gratitude to their colleagues, Dr. M. Mazzorana, Pr. Peter Bruckner and J.-P. Kleman, for their help and their interest in this work and for critical reading of this manuscript. They thank A. Bosch for his help in the preparation of the figures.

REFERENCES

[1] van der Rest, M. and Garrone, R. (1990) Biochimie 72, 473–484.
[2] Pan, T.-C., Zhang, R.-Z., Mattei, M.-G., Timpl, R. and Chu, M.-L. (1992) Proc. Natl. Acad. Sci. USA 89, 6565–6569.
[3] Doege, K.J. and Fessler, J.H. (1986) J. Biol. Chem. 261, 8924–8935
[4] Brass, A., Kadler, K.E., Thomas, J.T., Grant, M.E. and Boot-Handford, R.P. (1992) FEBS Lett. 303, 126–128
[5] Engel, J. and Prockop, D.J. (1991) Annu. Rev. Biophys. Biophys. Chem. 20, 137–152.
[6] Ricard-Blum, S., Hartmann, D.J., Herbage, D., Payen-Meyran, C. and Ville, G. (1982) FEBS Lett. 146, 343–347
[7] Dayer, J.M., Ricard-Blum, S., Kaufmann, M.T. and Herbage, D. (1986) FEBS Lett. 198, 208–212
[8] Bruckner, P. and Prockop, D.J. (1981) Anal. Biochem. 110, 360–368.
[9] Mayne, R., van der Rest, M., Bruckner, P. and Schmid, T.M. (1993) Extracellular Matrix Molecules, chapter 2, IRL Press, Oxford (in press).
[10] Ninomiya, Y. and Olsen, B.R. (1984) Proc. Natl. Acad. Sci. USA 81, 3014–3018.
[11] Ninomiya, Y., van der Rest, M., Mayne, R., Lozano, G. and Olsen, B.R. (1985) Biochemistry 24, 4223–4229.
[12] Brewton, R.G., Ouspenskaia, M.V., van der Rest, M. and Mayne, R. (1992) Eur. J. Biochem. 205, 443–449
[13] Gordon, M.K., Gerecke, D.R., Dublet, B., van der Rest, M. and Olsen, B.R. (1989) J. Biol. Chem. 264, 19772–19778
[14] Dublet, B. and van der Rest, M. (1991) J. Biol. Chem. 266, 6853–6858.
[15] Creighton, T.E. (1984) Methods Enzymol. 107, 305–329.
[16] Lazarev, Y.A., Lobachov, V.M., Grishkovski, B.A., Shibney, V.A., Grechishko, V.S., Finogenova, M.P., Esipova, N.G. and Rogulenkova, V.N. (1978) Biopolymers 17, 1215–1234.
[17] Brown, F.R., Di Corato, A., Lorenzi, G.P. and Blout, E.R. (1972) J. Mol. Biol. 63, 58–99.
[18] Germann, H.P. and Heidmann, E. (1988) Biopolymers 27, 157–163.
[19] Piez, R.A. and Sherman, M.R. (1970) Biochemistry 9, 4134–4140.
[20] Mazzorana, M., Gruffat, H., Sergeant, A. and van der Rest, M. (1993) J. Biol. Chem. (in press).