The primary structure of a triple-helical domain of collagen type VIII from bovine Descemet's membrane

Karlheinz Mann[1], Renate Jander[2], Eberhard Korsching[2], Klaus Kühn[1] and Jürgen Rauterberg[2]

[1]*Max-Planck-Institut für Biochemie, Abt. Bindegewebsforschung, D-8033 Martinsried, FRG* and [2]*Institut für Arteriosklerosеforschung an der Universität Münster, D-4400 Münster, FRG*

Received 4 September 1990

We have isolated and sequenced a fragment of 469 amino acid residues from bovine type VIII collagen. The sequence was composed of a series of Gly-X-Y repeats which was interrupted 8 times by short imperfections. The number and relative location of these interruptions were similar to those of chicken α1(X) and rabbit α1(VIII) chain triple-helical domains. Comparison to published N-terminal sequences to two triple-helical fragments of bovine type VIII collagen and to the cDNA derived sequence of the rabbit α1(VIII) chain showed that this fragment was the triple-helical domain of a second type VIII collagen chain which we designate α2(VIII).

Collagen; Type VIII; Triple-helical domain; Amino acid sequence; Bovine Descemet's membrane

1. INTRODUCTION

Collagen type VIII was first identified as a product of cultured bovine aortic endothelial cells [1] and rabbit corneal endothelial cells [2]. Only proliferating but not quiescent, confluent cells were found to synthesize this collagen [3,4]. The distribution of type VIII collagen in connective tissues is not yet clear [5,6]. A major source for the isolation of collagen VIII in the adult is Descemet's membrane, a specialized basement membrane separating corneal endothelium from corneal stroma [7,8].

The first model of type VIII collagen by Benya [2] and Sage et al. [8,9] proposed a homotrimeric molecule of M_r 560000 in which three short triple-helical domains were connected in tandem by short non-triple-helical domains. This cassette model has been challenged by Benya and Padilla [10], who concluded from their results that type VIII collagen is a short chain collagen consisting of three chains with M_r 61000 each. The latter model was confirmed by cloning and sequencing of a cDNA encoding the rabbit α1(VIII) chain [11]. The derived amino acid sequence contained a triple-helical domain of 454 residues flanked by a non-helical N-terminal domain of 117 residues (NC2) and a non-helical C-terminal domain of 173 residues (NC1). The C-terminal non-triple helical domain was strongly homologous to a similar domain in chicken α1(X), another short chain collagen. Furthermore, both triple-helical domains showed eight short imperfections of the Gly-X-Y triplet structure in the same relative locations within α1(VIII) and α1(X) chains. In the present paper we describe the structure of the triple-helical domain of a second chain, α2(VIII), from bovine Descemet's membrane.

Correspondence address: K. Mann, Max-Planck-Institut für Biochemie, Abt. Bindegewebsforschung, D-8033 Martinsried, FRG

2. MATERIALS AND METHODS

The type VIII collagen triple-helical domain was purified from pepsin treated bovine Descemet's membrane as in [12]. To separate the subunits, the purified preparations were heated to 60°C for 30 min in 6 M guanine, 0.1 M Tris, pH 8.5, then acidified with trifluoroacetic acid to pH 2 and finally subjected to reversed phase HPLC (Vydac C18, 4.6×250 mm column) using 0.1% trifluoroacetic acid (solvent A) and 0.1% trifluoroacetic acid in 70% acetonitrile (solvent B). The gradient was from 0% to 30% B in 10 min and then 30% to 50% B in 80 min, the flow rate was 0.5 ml/min throughout. The separated chains were identified and checked for purity by sequencing the N-termini. The recovery was determined by amino acid analysis.

The α2(VIII) fragment was cleaved with *Clostridium histolyticum* collagenase (Worthington), clostripain (Calbiochem), *Achromobacter lyticus* lysyl endopeptidase (Wako) and *Pseudomonas fragi* endoproteinase Asp-N (Boehringer Mannheim) at an enzyme/substrate ratio of 1:50 (w/w). The reaction buffer was 0.2 M ammonium bicarbonate containing 1 mM dithiothreitol with clostripain. Reaction temperature was 25°C; reaction time was 1 h with collagenase and 15 h with the other enzymes. Proteolysis was stopped by acidification to pH 2 with trifluoroacetic acid and immediate separation of the reaction products.

Endoproteinase Asp-N digests were separated by reversed phase HPLC using the column and solvents as described above with a linear gradient from 0% to 60% B in 160 min at a flow rate of 0.25 ml/min. Collagenase and clostripain cleavage products were first separated on a Mono-Q HR 5/5 column (Pharmacia-LKB) equilibrated with 0.1% trifluoroacetic acid titrated to pH 2.5 with NaOH. After washing the column for 10 min with the same solvent (solvent A) bound fragments were eluted with 0.1% trifluoroacetic acid containing 2 M NaCl (solvent B) using a linear gradient from 0% to 50% B in 60 min at a flow

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

rate of 0.5 ml/min. Pools of fragments were then further separated by reversed phase HPLC as above. Lysyl endopeptidase digests were separated by size exclusion chromatography on a TSK G3000 SW column (7.5 × 600 mm, Pharmacia-LKB) in a solvent containing 0.2 M ammonium acetate and 0.1% trifluoroacetic acid at a flow rate of 0.15 ml/min. Some of these lysyl endopeptidase derived fragments were further degraded with a second enzyme as described for whole chains. The products were separated by reversed phase HPLC as above.

Amino acid analyses were performed after hydrolysis with 6 M HCl for 24 h at 110°C on a Biotronik LC5000 analyser. For sequence analysis we used gas-liquid phase sequencers (Applied Biosystems, model 470A and 473A) following the manufacturers instructions.

3. RESULTS AND DISCUSSION

The starting material for this investigation was the triple-helical domain previously described as the major product from pepsin cleavage of type VIII collagen [8,12]. After denaturation two chains were separated by HPLC and identified by their N-terminal sequences as the α2(VIII) (50K-A) and α1(VIII) (50K-B) fragments first described by Kapoor et al. [8] (Fig. 1). The ratio of α1(VIII) to α2(VIII) varied between 1.5 and 2 in different preparations. Because the α1(VIII) chain had recently been sequenced at the cDNA level [11] we concentrated our efforts on the α2(VIII) fragment.

The fragment was first cleaved with bacterial collagenase. The complex mixture of degradation products was separated on a Mono-S column (Fig. 2,I) and the pools obtained were rechromatographed by reversed phase HPLC. As examples we show the chromatograms of pools A and I (Fig. 2, II, III, C-peptides in Fig. 3). A second cleavage was performed with clostripain. The degradation products were again separated with the two step procedure (C1 peptides in Fig. 3). In addition it was necessary to cleave the fragment with lysyl endopeptidase. These degradation products were separated according to size with gel permeation HPLC and yielded two large peptides (EL/1 and EL/2), a short peptide (EL/5) and the C-terminal (EL/4) and N-terminal (EL/3) peptides with 44 and 53 amino acid residues, respectively (Fig. 3). The C- and N-terminal peptides were, however, only partially separated and had to be rechromatographed on a Mono-S ion-exchange column. Due to the lack of positively charged residues the C-terminal peptide did not bind to this column under the conditions used. By sequencing the shorter peptides and the N-termini of the larger ones, we could assemble large blocks of continuous sequences. To fill the remaining gaps, two large lysyl endopeptidase derived peptides (EL/1 and EL/2) as well as the entire triple-helical fragment were cleaved with clostripain and endoprotease Asp-N (EL/Cl, EL/EA and EA peptides in Fig. 3). Edman degradation of these peptides, separated by reversed phase HPLC, completed the amino acid sequence of the α2(VIII) fragment.

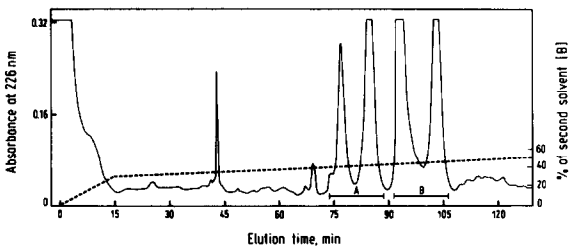

Fig. 1. HPLC separation of α1(VIII) and α2(VIII) chains after pepsin extraction and denaturation. The products were identified by N-terminal sequencing. The two peaks denoted 'A' contained the pure triple-helical region of the α2 chain, the two peaks denoted 'B' contained the pure triple-helical region of the α1 chain. The designation A and B refers to the sequences in [8]. The reason for the splitting into double peaks is not known. Gel electrophoresis did not show any difference in mobility (not shown). Higher column temperatures did not change this peak pattern.

The fragment contains a triple-helical stretch of 460 amino acid residues flanked by short non-triple helical remnants of the N- and C-terminal globular domains which were destroyed during pepsin extraction of

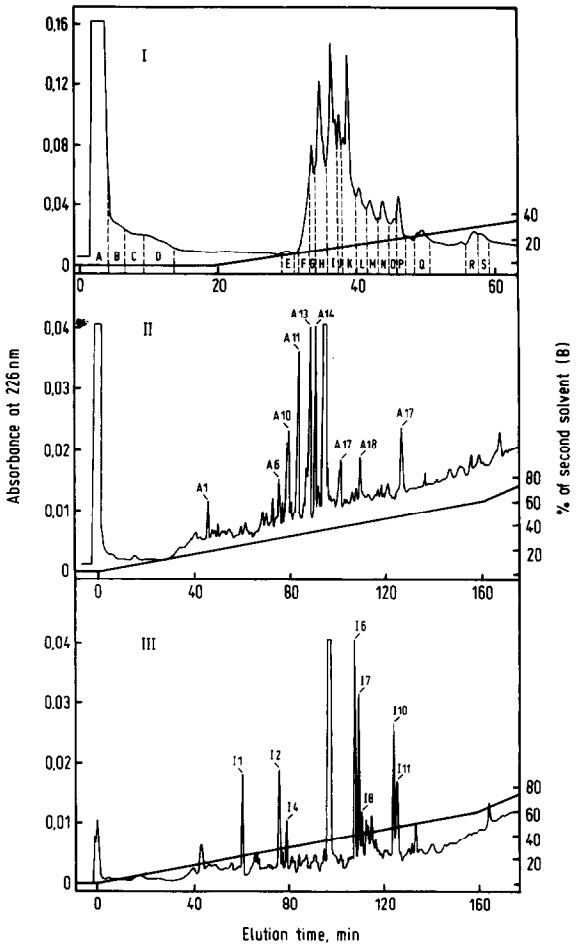

Fig. 2. The separation of the collagenase digest of the α1(VIII) triple-helical fragment on a Mono-S column (I) and further separation of the peptides in pool A (II) and pool I (II) by reversed phase HPLC.

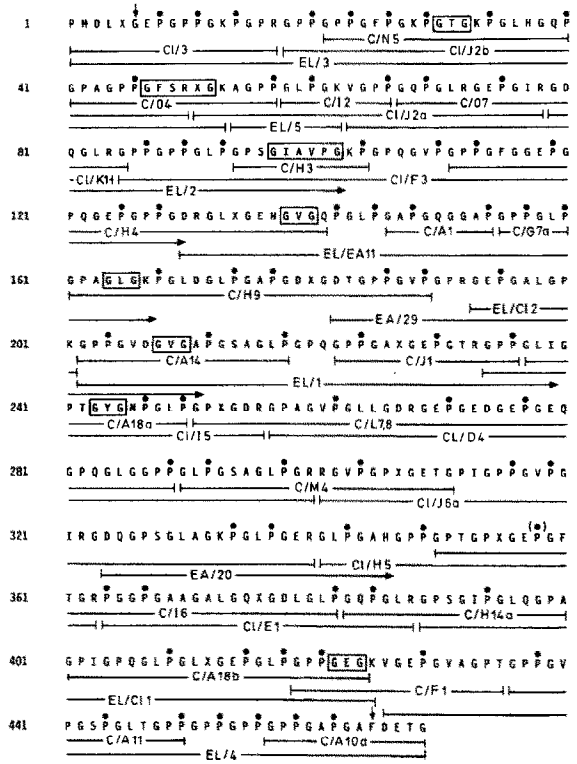

Fig. 3. The sequence of the α2(VIII) chain triple-helical domain. Arrows denote the start and the end of the triple-helical domain. Interruptions of the regular triplet repeat structure are boxed. For nomenclature of peptides see text. Peptide lines ending with an arrow were peptides which were not sequenced to the C-terminus. Dots over P denote hydroxyproline, X denotes unidentified residues. C, Cl, EL and EA denote peptides derived from collagenase, clostripain, lysyl endopeptidase and endoproteinase Asp-N cleavages of the fragment, respectively.

Descemet's membrane. By comparison with the published cDNA derived sequence of the rabbit α1(VIII) chain [11] this represents the entire triple-helical domain of the α2(VIII) chain. The triple-helical sequence is interrupted by 8 short imperfections (Fig. 3). Six of them were of the type Gly-X-Gly and two of the type Gly-X-Y-X-Y-Gly; these imperfections were probably generated by deletion of a residue in X or Y position and by deletion of a glycyl residue, respectively.

Collagens undergo post-translational modifications, such as hydroxylation of propyl and lysyl residues in the Y-position of the Gly-X-Y triplet and glycosylation of the hydroxylysyl residues. All prolines in Y position were found to be completely hydroxylated with the only exception of proline-415, which was hydroxylated to only 50%. The question as to what extent the lysine residues in position Y are hydroxylated and additionally glycosylated cannot be answered directly. Glycosylated hydroxylysines are not detected during Edman degradation and a gap appears instead of the respective phenylthiohydantoin derivative. Ten gaps designated in Fig. 3 as X have been observed during sequence analyses. Since this number correlates well with the number of hydroxylysine residues present in the α2(VIII) fragment (Table I) we assume that all X's are occupied by a glycosylated hydroxylysine. Only one non-modified lysine was found in a Y-position (lysine-201).

Yamaguchi et al. [11] have previously noted the similarity between the rabbit α1(VIII) chain and the chicken α1(X) chain. Comparison of the α2(VIII) sequence to sequences in the MIPSX Database (F. Pfeiffer, Martinsried Institute of Protein Sequences) by the FASTP program [13] revealed an identity of 58–59% with both chicken α1(X) and rabbit α1(VIII). The similarity to other collagens was 45–50%. It is striking that the two mammalian type VIII triple-helical domains were no more similar to each other than each of them was to the chicken type X triple-helical domain.

The fact that the 8 triple-helical imperfections are conserved in all three α-chains shows the importance of this structural feature which introduces flexibility to the rigid triple-helix. As far as the imperfections are concerned, the α2(VIII) chain is more similar to α1(X) than to α1(VIII) since two of the three imperfections of the larger type Gly-X-Y-X-Y-Gly found in type X [14] were matched by similar imperfections in α2(VIII). Only one was substituted by the short type Gly-X-Gly which is the only type found in α1(VIII). Most of the imperfections in all three chains were followed by a lysyl residue. This is similar to type IV collagen, another collagen with frequent interruptions [15].

Table I

The amino acid composition of the α2(VIII) triple-helical domain

	Residues/molecule	
	A	B
Asp/Asn	14	15 ± 3
Thr	10	9 ± 1
Ser	7	8 ± 2
Glu/Gln	33	36 ± 4
Pro	49	51 ± 7
Hyp	76	79 ± 8
Gly	159	142 ± 16
Ala	22	26 ± 3
Val	13	14 ± 2
Met	2	0
Ile	7	8 ± 1
Leu	33	35 ± 4
Tyr	1	0–1
Phe	5	6 ± 1
His	2	4 ± 2
Lys	10	8 ± 2
Hyl	–	9 ± 1
Arg	16	16 ± 2
X	10	–

Comparison of the composition deduced from the sequence (A) to that found with amino acid analyses (B, with standard deviation). X denotes unidentified residues in Fig. 2.

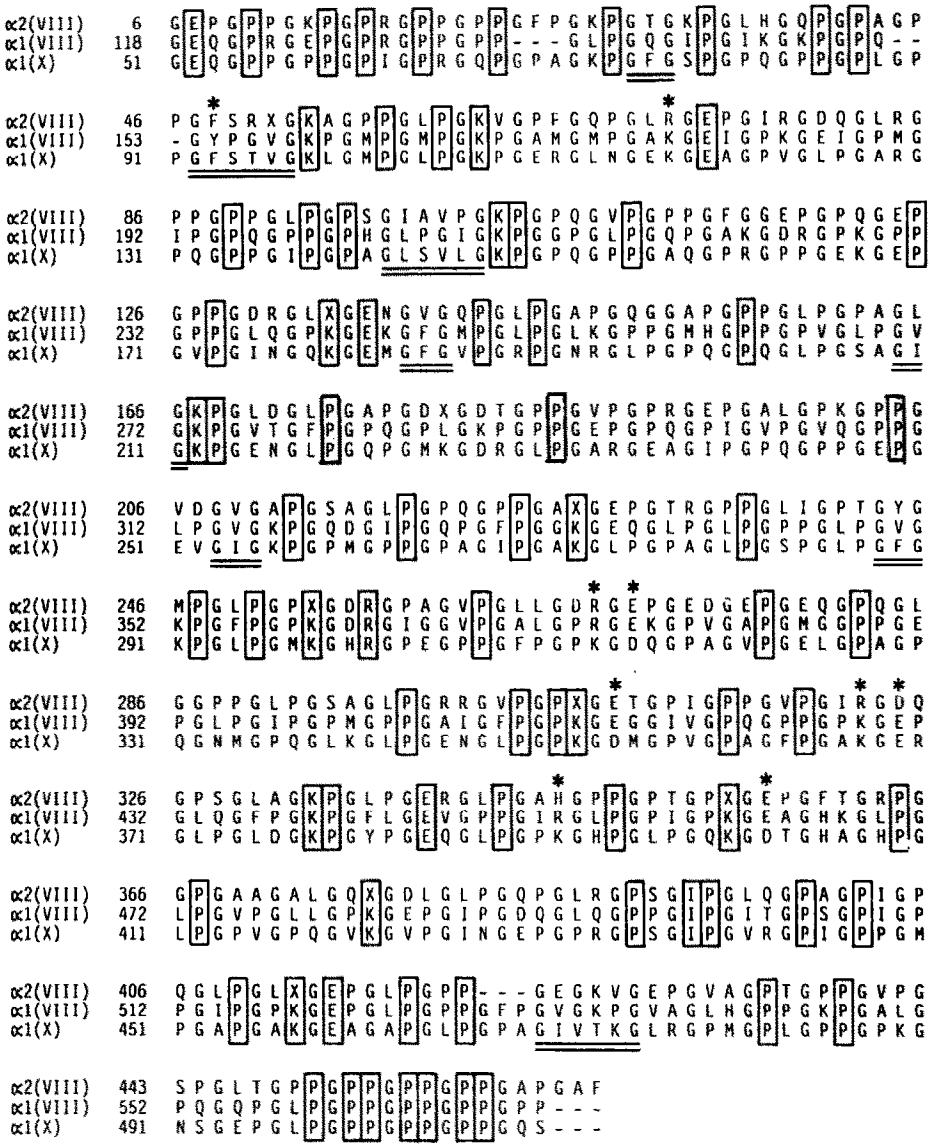

Fig. 4. Alignment of the triple-helical domains of α2(VIII), α1(VIII) and α1(X). Double bars denote imperfections of the Gly-X-Y repeat collagen structure. Residues identical in all three chains are boxed (with the exception of glycines). *denotes non-identical but similar residues (charged aa, aromatic aa, etc.) in all three sequences. Modified residues are not indicated because the comparison is to cDNA-derived sequences. The numbering of α1(VIII) and α1(X) residues is according to [11] and [14], respectively.

The nature of the 10 unidentified residues in Fig. 3 was previously discussed. Seven of them match with lysines in the Y-position in the cDNA derived sequences from both other chains, supporting our identification as glycosylated hydroxylysines (Fig. 4). Most of these modified lysines in all 3 chains are followed by a negatively charged amino acid in position X of the next triplet. Such clusters were also found in other collagens (see type VI collagen as an example [16]). Another well-preserved feature of all 3 chains is the block of Gly-Pro-Hyp triplets at the C-terminal end. Similar blocks have also been found in several other collagens and are supposed to serve as nucleation centres in helix formation and to stabilize the end region of the triple helix [17–19]. All 3 chains contain more hydroxyproline than proline. This is similar to type IV collagen and different from the fiber-forming collagens [15]. The hydroxyl group of 4-hydroxyproline in position Y stabilizes the triple-helical structure [20] and the high content of this residue may thus compensate for the presence of the 8 non-triple-helical interruptions.

Comparison of our results with those of Yamaguchi et al. [11] shows that we have sequenced the triple-helical domain of a second chain of type VIII collagen, α2(VIII). The similarity of the triple-helical areas of α1(VIII) and α2(VIII) in length, location and number of imperfections suggests that both chains together form the 120–140 nm long triple helical domain [12] of

the molecule. The fact that we have not found any evidence for another chain as well as our observation that the ratio of α1(VIII):α2(VIII) is between 1.5 and 2, speaks for a chain composition of [α1(VIII)]$_2$[α2(VIII)] for the type VIII collagen molecule.

REFERENCES

[1] Sage, H., Pritzl, P. and Bornstein, P. (1980) Biochemistry 19, 5747–5755.
[2] Benya, P.D. (1980) Renal Physiol. 3, 30–35.
[3] Alitalo, K., Bornstein, P., Vaheri, A. and Sage, H. (1983) J. Biol. Chem. 258, 2653–2661.
[4] Sage, H., Tupper, Y. and Bramsen, R. (1986) J. Cell Physiol. 127, 373–387.
[5] Kapoor, R., Sakai, L.Y., Funk, S., Roux, E., Bornstein, P. and Sage, H. (1988) J. Cell Biol. 107, 721–730.
[6] Kittelberger, R., Davis, P.F., Flynn, D.W. and Greenhill, N.S. (1990) Connect. Tiss. Res. 24, 303–318.
[7] Labermeier, U. and Kenney, M.C. (1983) Biochem. Biophys. Res. Commun. 116, 619–625.
[8] Kapoor, R., Bornstein, P. and Sage, H. (1986). Biochemistry 25, 3930–3937.
[9] Sage, H., Trüeb, B. and Bornstein, P. (1983) J. Biol. Chem. 258, 13391–13401.
[10] Benya, P.D. and Padilla, S.R. (1986) J. Biol. Chem. 261, 4160–4169.
[11] Yamaguchi, N., Benya, P.D., van der Rest, M. and Ninamya, Y. (1989) J. Biol. Chem. 264, 16022–16029.
[12] Jander, R., Korsching, E. and Rauterberg, J. (1990) Eur. J. Biochem. 189, 601–607.
[13] Lipman, D.J. and Peerson, W.R. (1985) Science 227, 1435–1441.
[14] Ninomiya, Y., Gordon, M., van der Rest, M., Schmid, T., Linsenmayer, T. and Olsen, B.R. (1986) J. Biol. Chem. 261, 5041–5050.
[15] Brazel, D., Pollner, R., Oberbäumer, I. and Kühn, K. (1988) Eur. J. Biochem. 172, 35–42.
[16] Chu, M.-L., Conway, D., Pan, T., Baldwin, C., Mann, K., Deutzmann, R. and Timpl, R. (1988) J. Biol. Chem. 263, 18601–18606.
[17] Sakakibara, S., Inouye, K., Skudo, K., Kishida, Y., Kobayashi, Y. and Prockop, D.J. (1973) Biochem. Biophys. Acta 303, 198–202.
[18] Engel, J., Chen, H.-T. and Prockop, D.J. (1977) Biopolymers 16, 601–622.
[19] Germann, H.-P. and Heidemann, E. (1988) Biopolymers 27, 157–163.
[20] Berg, R.A. and Prockop, D.J. (1973) Biochem. Biophys. Res. Commun. 52, 115–120.