THE AMINO ACID SEQUENCE OF THE CARBOXYTERMINAL NONHELICAL CROSS LINK REGION OF THE α1 CHAIN OF CALF SKIN COLLAGEN

J. RAUTERBERG, P. FIETZEK, F. REXRODT, U. BECKER, M. STARK and K. KÜHN

Max-Planck-Institut für Eiweiss- und Lederforschung, Abteilung Kühn, D-8 München 2, Schillerstrasse 46, W. Germany

Received 6 December 1971

1. Introduction

The collagen molecule terminates at both ends in regions endowed with particular properties. These regions cannot participate in the triple helical conformation since glycine does not occupy every third position as in the central areas of the molecule. The terminal regions contain lysine residues usually converted into α-amino-adipic-semialdehyde by lysine oxidase [1]. Allysine is the essential participant in the crosslinking reaction [2, 3]. Structure and function of the N-terminal region have been determined some time ago [4–9]. In addition, Stark et al. [10] reported the isolation and characterization of the nonhelical peptide of the C-terminal end of the α1-chain of calf skin collagen.

The amino acid sequence of this peptide is reported in the present paper. It was isolated from a chymotryptic digest of the C-terminal cyanogen bromide peptide α1-CB6. The sequence was established applying automated Edman degradation to the whole peptide and manual Edman-dansyl-technique to fragments obtained by proteolytic cleavage.

2. Materials and methods

2.1. *Starting material*

The middle layer of fresh calf skin (900 g) was ground in an electric meat grinder and extracted for 2 hr at room temp (22–24°) with 1500 ml urea (final conc. 8 M). The insoluble material was removed by centrifugation for 30 min at 15,000 *g* (rotor GSA) in an RC-2 Sorvall centrifuge. The supernatant was dialysed against 0.06 M sodium acetate buffer pH 4.8. The α1 chains were isolated on carboxy methyl cellulose as described earlier [11].

2.2. *Preparation of α1-CB6*

The α1-chains were cleaved with CNBr and then α1-CB6 was isolated using the procedure described earlier [5] with the following modification. After digestion, CNBr was removed by desalting on Bio-Gel P-2 (100–200 mesh) in acetic acid. The peptide mixture was lyophilized, then fractionated on Bio-Gel A-1.5 m, 200–400 mesh, equilibrated with 1 M $CaCl_2$ –0.05 M Tris, pH 7.5 (column dimensions 3.6 × 120 cm, elution rate 50 ml/hr). The fractions comprising α1-CB6 were combined, desalted and rechromatographed on Bio-Gel P-150, as described earlier [5].

2.3. *Preparation of α1-CB6-C3*

Digestion of α1-CB6 with chymotrypsin was carried out as described earlier [10]. The chymotryptic peptides were separated on Bio-Gel P-4 (see fig. 1). α1-CB6-$C3^{Lys}$ and $C3^{Ald}$ were rechromatographed on Bio-Gel P-10 (200–400 mesh) in 0.05 M sodium acetate buffer pH 4.5 (column dimensions 2.7 × 110 cm; flow rate 26 ml/hr). Further purification of α1-CB6-$C3^{Ald}$ was achieved by chromatography on phosphocellulose at pH 3.6 in 0.001 M sodium acetate using a linear gradient from 0–0.2 M NaCl over a total vol of 200 ml (column dimension 7 × 1.8 cm, flow rate 80 ml/hr).

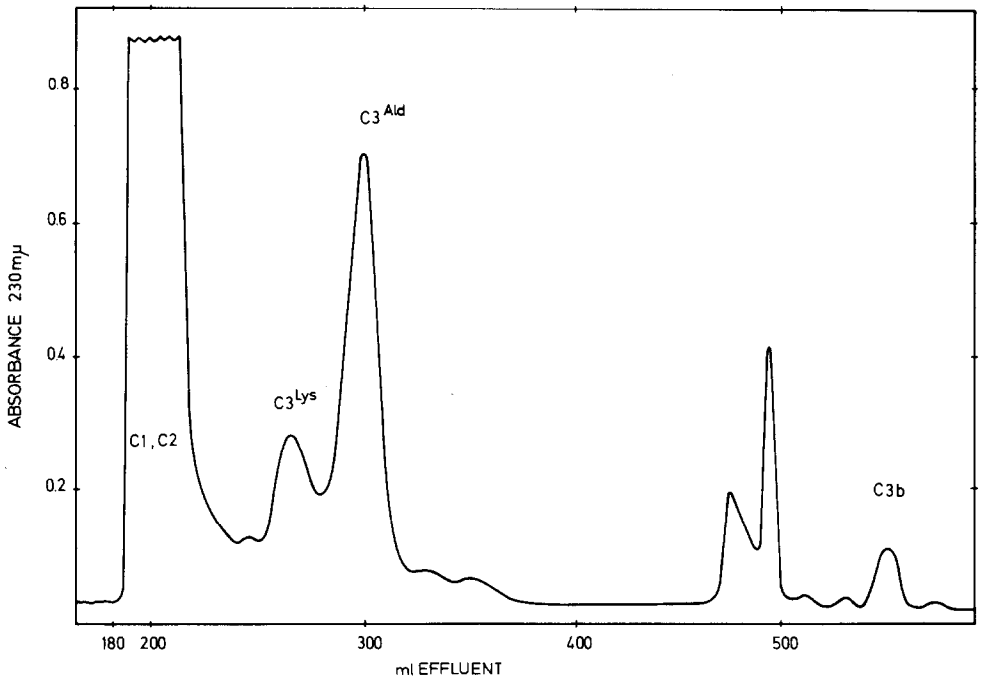

Fig. 1. Molecular sieve chromatography on Bio-Gel P-4 of the chymotryptic digest α1-CB6. 60 mg were applied and eluted with 0.1 M acetic acid at room temp. Column dimensions: 2.5 × 104 cm, elution rate 30 ml/hr.

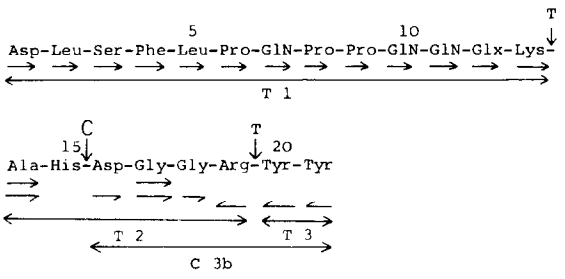

Fig. 2. Amino acid sequence of α1-CB6-C3. T = trypsin cleavage; C = chymotryptic cleavage. The tryptic peptides are indicated. —→ automated Edman degradation, —→ = manual Edman degradation, ←— = carboxypeptidase digestion.

2.4. *Tryptic digestion of α1-CB6-C3*^{Lys}

The digestion with trypsin (2 × cryst., Worthington) was carried out in 0.2 M NH_4HCO_3 containing 10^{-3} M $CaCl_2$ at pH 7.8 for 4 hr, at 37°. The weight ratio of substrate to enzyme was 50:1. The trypsin was pretreated with TPCK ((1-tosylamido-2-phenyl) ethyl chloromethyl ketone). Digestion was terminated by adding acetic acid. The digest was lyophilized and then fractionated on Bio-Gel P-4 (see fig. 2).

2.5. *Treatment of α1-CB6 with carboxypeptidase B*

3 mg α1-CB6 in 1.4 ml of 0.2 M ethylmorpholin acetate, pH 8.5, were incubated at 37° with 0.12 mg carboxypeptidase B (Worthington, COBDFP), dissolved in 0.1 ml 10% LiCl. The digestion was terminated by adding acetic acid. The sample was lyophilized and directly applied to the amino acid analyser (model BC 200 Bio-Cal Instruments, Munich).

2.6. *Sequence determination*

Automated Edman degradation was performed in a Beckman Sequencer, model 890 (Beckman Instruments, Palo Alto, Calif.). Three runs were carried out, using the Quadrol single cleavage or the dimethylallylamine programme supplied by Beckman Instruments. Between 400 and 600 nmoles of peptides were used. In order to prevent cyclization of glutamine, the vacuum and drying steps after adding heptafluorobutyric acid (cleavage step) were shortened. The

Table 1
Amino acid composition of α1-CB6-C3 and its fragments. Results are given as residues per peptide.

	$C3^{Lys}$	$C3^{b}$	T 1	T 2	T 3
Aspartic acid	2 (2.0)	1 (1.0)	1)1.1)	1 (1.1)	
Serine	1 (0.9)		1 (0.6)	–	
Glutamic acid	4 (3.7)		4 (3.5)	–	
Proline	3 (3.2)		3 (3.2)	–	
Glycine	2 (2.3)	2 (2.5)	–	2 (2.5)	
Alanine	1 (1.1)		–	1 (1.0)	
Leucine	2 (2.0)		2 (2.2)	–	
Tyrosine	2 (1.7)	1 (1.5)	–	–	2
Phenylalanine	1 (1.0)		1 (1.1)	–	
Histidine	1 (1.0)		–	1 (0.9)	
Lysine	1 (0.9)		1 (0.9)	–	
Arginine	1	1 (0.9)	–	1 (1.0)	
Total	21	6	13	6	2

manual Dansyl-Edman degradation was performed as described by Gray [12].

The PTH derivatives (from the sequencer) were identified by gas-liquid (GLC) or thin layer chromatography (TLC). GLC was performed in the Beckman GC-45 instrument (Beckman Instruments, Palo Alto, Calif.) using 10% SP 400 as liquid phase. TLC was performed on silica gel plates (DC Fertigplatten F254) from Merck, Darmstadt, Germany, using the H and D system of Edman [13]. The dansyl-amino acids were identified by TLC on polyamide plates F 1700 from Schleicher and Schüll (Dassel, Germany), using the systems described by Wood and Wang [14].

TLC of peptide T3 was performed on silica gel G (Merck, Darmstadt, Germany) using butanol–pyridin–acetic acid–water (60:20:6:24).

The amino acid analyses were performed as described earlier [5]. For detection of lysine derived α-amino-adipic-semialdehyde the samples were treated with performic acid as described by Moore [15]. By this procedure, allysine is oxidised to α-amino adipic acid which can be detected on the amino acid analyser.

3. Results

Fig. 1 shows the separation on Bio-Gel P-4 of the chymotryptic peptides from α1-CB6. The first large peak contains the amino terminal α1-CB6-C1 and the adjacent peptide α1-CB6-C2. The amino acid sequence of the 113 residues long α1-CB6-C1 is known [16], and the sequence of α1-CB6-C2 containing 82 amino acids is under investigation. The carboxyterminal α1-CB6-C3 exists in a lysine and an aldehyde form. In the latter, the lysine residue is converted to the aldehyde allysine. The minor fraction $C3^{b}$ eluting after the salt, contains a hexapeptide. The amino acid composition of C3 and $C3^{b}$ is given in table 1.

The amino acid sequence of α1-CB6-$C3^{Lys}$ is depicted in fig. 2. The residues 1 to 14 and residue 17 were elucidated by automated Edman degradation. However, when α1-CB6-$C3^{Ald}$ was sequenced the PTH derivative of residue 13 could not be detected. The sequence of the remaining section was elucidated after tryptic digestion of α1-CB6-$C3^{Lys}$. Three peptides were obtained by separation on Bio-Gel P-4 (fig. 3), and their amino acid compositions are given in table 1. Peptide T 1 was easily identified as residues 1 to 13 by its amino acid composition. The adjacent hexapeptide T 2 was sequenced manually (fig. 2), and T 3 was identified as tyrosyl-tyrosine as follows. After acid hydrolysis only tyrosine was found in the amino acid analysis. However, when the same peptide was applied to the amino acid analyser without hydrolysis, no amino acid could be detected. By TLC T 3 showed the same R_f value as the synthetic tyrosyl-tyrosine and could clearly be distinguished from free tyrosine. Confirmation of the sequence

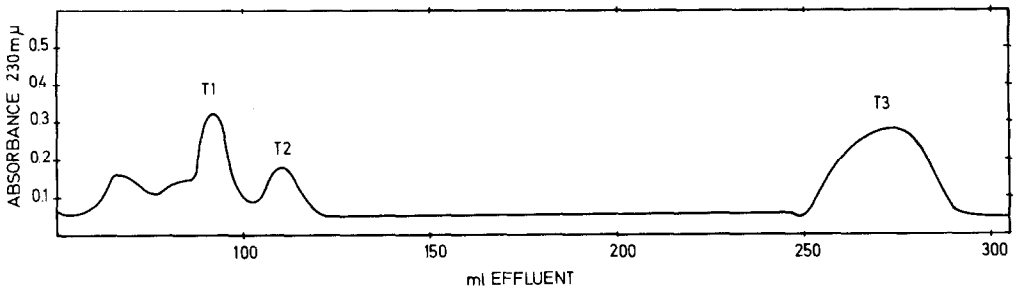

Fig. 3. Molecular sieve chromatography on Bio-Gel P-4 of the tryptic digest of α1-CB6-C3. The column (1.6 × 95 cm) was eluted with 0.1 M acetic acid at room temp. Elution rate was 12 ml/hr.

19–21 was obtained by digestion of α1-CB6 with carboxypeptidase B. After 6 hr incubation, 1.6 moles of tyrosine and 0.8 moles of arginine were obtained per mole of peptide. The residue in position 15 could not be identified either by the automated or manual Edman degradation. On the other hand, both α1-CB6-C3 and T 2 contained one histidine residue which could not be placed in the sequence. Therefore, histidine could be located in position 15. This conclusion was supported by the existence of $C3^b$ (table 1), which covers residues 16–21 and contained no histidine. This peptide is probably the result of chymotryptic cleavage of the peptide bond His–Asp.

4. Discussion

Due to lability of certain bonds in the region of the α1 chain under investigation, special experimental attention was required. For instance, the nonhelical peptide C3 is partially lost during normal extraction with neutral salts or citrate buffer, possibly due to the action of proteolytic enzymes present in connective tissue. However, when skin was extracted with the strongly denaturing agent 8 M urea, breakage of C3 could be prevented. Furthermore, the peptide α1-CB6 from salt- or acid-extracted collagen always appears in several forms, predominantly as $\alpha1\text{-CB6}^a$ and $\alpha1\text{-CB6}^b$. The former contains 1 tyrosine less, and the latter contains 16 amino acids less than α1-CB6 obtained from urea extracted collagen [5, 17]. In contrast, the N-terminal region is obtained unimpaired.

Residue 12 is shown as Glx, because in this particular case we could not distinguish between glutamine or glutamic acid for the following reasons. When a glutamine-containing peptide is subjected to Edman degradation, some of the glutamine is always deaminated during the conversion. In addition, there is some overlap from one degradation step to the next.

Sequence studies on α1-CB6-C2 now in progress have demonstrated glycine to occupy every third position. This suggests that α1-CB6-C3 contains the whole nonhelical C-terminal region [18].

Comparative studies on the intermolecular cross links of collagens from different species suggest that other collagens contain similar nonhelical regions at their C-terminal ends. Further studies now in progress are necessary to confirm this.

Parallel to these studies the sequence of the N-terminal cross link region of the calf α1 chain has also been elucidated [19]. This is the first instance in which the structure of both the N-terminal as well as the C-terminal region of an α1-chain has been completely determined.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft, SFB 51, and by the Arbeitsgemeinschaft Industrieller Forschungsvereinigungen. We wish to thank Dr. E. Wünsch for supplying the synthetic tyrosyl-tyrosine, Miss I. Langer, W. Wohland and I. Kell for their excellent technical assistance.

References

[1] S.R. Pinnell and G.R. Martin, Proc. Natl. Acad. Sci. U.S. 61 (1968) 708.
[2] P. Bornstein, A.H. Kang and K.A. Piez, Proc. Natl. Acad. Sci. U.S. 55 (1966) 417.
[3] P. Bornstein and K.A. Piez, Biochemistry 5 (1966) 3450.
[4] A.H. Kang, P. Bornstein and K.A. Piez, Biochemistry 6 (1967) 788.
[5] A.H. Kang and J. Gross, Biochemistry 9 (1970) 796.
[6] E.M. Click and P. Bornstein, Biochem. J. 9 (1970) 4699.
[7] P. Bornstein and R. Nesse, Arch. Biochem. Biophys. 138 (1970) 443.
[8] J. Rauterberg and K. Kühn, European J. Biochem. 19 (1971) 398.
[9] F.H. Epstein, R.D. Scott, E.J. Miller and K.A. Piez, J. Biol. Chem. 246 (1971) 1718.
[10] M. Stark, J. Rauterberg and K. Kühn, FEBS Letters 13 (1971) 101.
[11] K.A. Piez, E.A. Eigner and M.S. Lewis, Biochemistry 2 (1963) 58.
[12] W.R. Gray and J.F. Smith, Analyt. Biochem. 33 (1970) 36.
[13] P. Edman, in: Protein Sequence Determination, ed. Saul B. Needleman (Springer Verlag, Berlin, Heidelberg, New York, 1970).
[14] K.R. Wood and K. Wang, Biochim. Biophys. Acta 133 (1967) 369.
[15] S. Moore, J. Biol. Chem. 238 (1963) 235.
[16] K. von der Mark, P. Wendt, F. Rexrodt and K. Kühn, FEBS Letters 11 (1970) 105.
[17] U. Becker, R. Timpl and K. Kühn, in preparation.
[18] P. Fietzek, F. Rexrodt, P. Wendt, M. Stark and K. Kühn, unpublished results.
[19] J. Rauterberg, R. Timpl, H. Furthmayr and K. Kühn, in preparation.