PRO–GLN: THE PROCOLLAGEN PEPTIDASE CLEAVAGE SITE IN THE α1(I) CHAIN OF DERMATOSPARACTIC CALF SKIN PROCOLLAGEN

D. HÖRLEIN, P. P. FIETZEK+ and K. KÜHN
Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG

Received 28 March 1978

1. Introduction

Collagen is synthesized as a precursor form called procollagen which contains both amino- and carboxy-terminal extensions, the so-called procollagen peptides. During conversion of procollagen to collagen the extensions are removed by procollagen peptidases which are specific either for the N- or C-terminal extensions [1].

Investigation of the structure of the N-terminal extension has been greatly facilitated by the genetic disease, dermatosparaxis, found in cattle [2] and sheep [3], in which the enzyme responsible for removing the N-terminal procollagen peptide is inactive [4,5]. This results in the accumulation of procollagen in the skin which contains pα1 (I) chains.

Here we report the isolation of the N-terminal extension of the pα1(I) chain from a cyanogen bromide digest of dermatosparactic calf skin. The extension was fragmented using bacterial collagenase and peptides separated. The C-terminal collagenase fragment was sequenced and the cleavage site of the procollagen peptidase determined as the peptide bond between proline and glutamine.

2. Materials and methods

Skin from dermatosparactic calves (kindly provided by Professor Charles M. Lapière, University of Liège, Belgium) was minced into 70% formic acid denatured for 2 h at 45°C and cleaved with CNBr (1 g CNBr/100 g skin) for 4 h at 30°C. After centrifugation (1 h, 15 000 × *g*) the supernatant was dialysed exhaustively against 0.1 M acetic acid and lyophilised. The isolation of pα1(I)-CBO, 1 was carried out using DEAE cellulose and agarose A 1.5 chromatography (see fig.1 for experimental details). Serological identification of the peptide was carried out as in [6].

Peptide, 30 mg, was dissolved in 3 ml 0.2 M NH_4HCO_3, pH 7.8 and incubated with bacterial collagenase (Worthington Biochem. Corp., CLSPA-grade, 383 U/mg) for 4 h at 37°C using an enzyme : substrate ratio of 1:50. The digest was directly chromatographed on Sephadex G-50s (see fig.2 for experimental details).

Amino acid analysis was performed as in [7] using a Durrum D 500 analyzer.

The sequence determinations were carried out in a liquid phase Sequencer (Beckman Instr., Palo Alto, CA) using the fast peptide degradation programme no. 111374. For each run about 300 nmol peptide was applied. The identification of the PTH amino acids was carried out using thin-layer and gas–liquid chromatography as in [8]. The yields of the PTH amino acids released in the degradation were determined by quantification of the gas–liquid chromatography data.

3. Results

In order to isolate the N-terminal cyanogen bromide peptide (pα1(I)-CBO,1) of the pα1(I) chain,

+ Present address: Rutgers Medical School, Piscataway, NJ 08854, USA

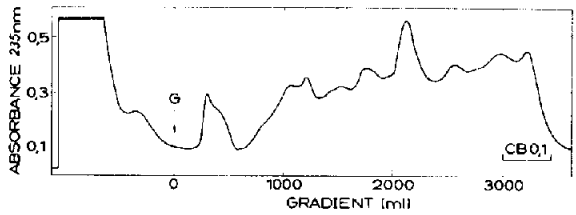

Fig.1. Separation of pα1(I)-CBO,1 from a CNBr digest of dermatosparactic calf skin on DEAE cellulose. The DEAE cellulose (DE-52, Whatman) column (5 × 30 cm) was equilibrated at room temp. with 50 mM Tris–HCl, pH 8.6, containing 2 M urea and eluted with a linear gradient from 0 0.3 M NaCl in total vol. 2 l at flow rate 200 ml/h. The start of the gradient is marked by an arrow. 1.5 g digest as applied.

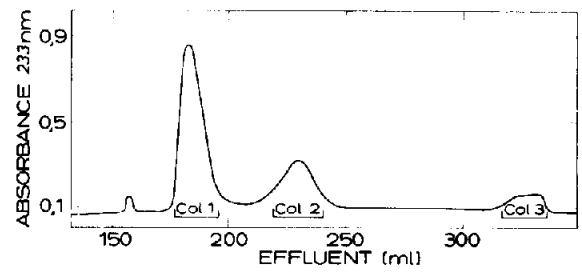

Fig.2. Chromatography of the collagenase-derived peptide mixture of pα1(I)-CBO,1 on Sephadex G-50s. The three peptide fractions are marked. The Sephadex G-50s (Pharmacia) column (3.5 × 150 cm) was equilibrated with 0.2 M NH_4HCO_3, pH 8.5, and eluted at room temp. at flow rate 50 ml/h. 30 mg peptide mixture was applied.

Table 1
Amino acid composition of the collagenase-derived fragments of pα1(I)-CBO,1

	Col 1	Col 2	Col 3[b]	Col 1 + Col 2 + Col 3	CBO,1
Hydroxyproline	–	–	6.8	6.8	7.0
Aspartic acid	15.3	2.1 (2)[a]	1.1	18.5	18.4
Threonine	5.8	1.1 (1)	–	6.9	6.6
Serine	0.9	2.9 (3)	–	3.8	3.9
Homoserine	–	0.8 (1)	–	0.8	0.9
Glutamic acid	18.7	2.0 (2)	0.9	21.8	22.3
Proline	10.1	3.2 (3)	8.2	21.5	21.8
Glycine	9.4	6.3 (6)	11.0	26.7	27.2
Alanine	0.8	1.3 (1)	1.2	3.3	3.4
Valine	10.2	0.9 (1)	–	11.1	10.5
Isoleucine	2.6	1.0 (1)	0.9	4.5	4.5
Leucine	3.0	2.1 (2)	0.9	6.0	6.1
Tyrosine	0.8	2.0 (2)	–	2.8	2.5
Phenylalanine	–	1.2 (1)	–	1.2	1.2
Histidine	1.1	–	–	1.1	1.1
Lysine	4.2[c]	0.4 (1)	–	4.6	4.3
Arginine	3.0	–	2.0	5.0	5.2
Aminoethylcysteine	9.8	–	–	9.8	9.5
Total	95.7	27.3 (27)	33.0	155.0	156.4

[a] Values in brackets are those found by sequence analysis
[b] Col 3 was calculated on the basis of the molecular weight difference between CBO,1 and Col 1 plus Col 2
[c] Lysine is in part converted to the lysine-derived aldehyde which cannot be seen in the amino acid analysis – see also [9]

Values are given as residues/peptide. A dash indicates less than 0.2 residue

the peptide mixture was chromatographed on a DEAE cellulose column (fig.1). The fractions obtained were assayed with antisera specific for the N-terminal of pα1(I) from dermatosparactic calf. The serological activity was found in the last peak which was then rechromatographed on a Bio-Gel A 1.5 column (not shown).

The peptide pα1(I)-CBO,1 was cleaved with collagenase and subsequently chromatographed on a Sephadex G-50s column. The elution pattern (fig.2) consisted of 3 peaks containing Col 1, Col 2 and Col 3, whose amino acid compositions are given in table 1. Col 1 represents the globular cysteine-containing portion of pα1(I)-CBO,1. The homoserine peptide Col 2, composed of the C-terminal region of the procollagen peptide and the N-terminal CNBr peptide of the α1(I) chain, contains the cleavage site of the procollagen peptidase. The Col 3 fraction consists of a mixture of small peptides which represent the triple helical portion of the extension.

The sequence elucidated by the automated Edman degradation procedure is depicted in fig.3. The yield of the PTH amino acids obtained are given in table 2. Presumably the distinct yield of PTH alanine in step 8 is due to an overlap from step 7. A dramatic reduction of the yield of PTH proline has been obtained in step 8, a fact which we have observed several times if the following position is occupied by glutamine. Considerable differences in the yields of PTH proline in the sequence have been described as depending on the nature of the amino acid following proline in the sequence [9]. 15 Col 2 residues were determined. The first 8 residues belong to the C-terminal part of the procollagen peptide. The following sequence is identical with the beginning of α1(I)-CBO,1 [10]. From these results one can deduce that the procollagen peptidase splits the peptide bond between proline in position 8 and glutamine in position 9.

```
                5                   10
Gly-Leu-Gly-Gly-Asn-Phe-Ala-Pro-Gln-Leu-Ser-Tyr-Gly-Tyr-

 15                  20                  25
Asp-Glu-Lys-Ser-Thr-Gly-Ile-Ser-Val-Pro-Gly-Pro-Hse
```

Fig.3. Amino acid sequence of pα1(I)-CBO,1-Col 2 as determined by automated Edman degradation. Positions 9–27 are identical with α1(I)-CBO,1 [9].

Table 2
Yields of PTH amino acids released during automated degradation of pα1(I)-CBO,1-Col 2 as determined by gas–liquid chromatography

Cycle no.	Amino acid released[a]	PTH-derivative (nmol)
0		300
1	Gly	213
2	Leu	278
3	Gly	213
4	Gly	199
5	Asn	n.d.
6	Phe	128
7	Ala	152
8	Pro/Ala	48/43
9	Gln/Pro	n.d./11
10	Leu	38
11	Ser	n.d.
12	Tyr	n.d.
13	Gly	28
14	Tyr	n.d.
15	Asp	n.d.

[a] Qualitative identification was carried out using thin-layer chromatography

4. Discussion

The amino acid composition of the 3 collagenase-derived peptides led to the conclusion that the procollagen peptide is built of 2 structurally different domains an N-terminal cysteine-containing globular region (Col 1) and a collagen-like portion with a high proline and hydroxyproline content (Col 3) which is connected with the α1(I) chain by a short non-collagen sequence. Similar results have been obtained for the N-terminal procollagen peptide of dermatosparactic sheep type I procollagen [6]. The sequence of Col 1 and Col 3 will be published (D.H. and P.P.F., in preparation).

According to the sequence of Col 2 presented here, the procollagen peptidase is endowed with a rather unique specificity. It acts as a proline-specific endopeptidase. The enzyme has an approx. 1000-fold higher affinity to native procollagen than to the denatured chains but is also capable of removing the extension from denatured chains [5] suggesting that the three dimensional structure of the cleavage site is important. However, the α1(I) chain also contains

several Pro-Gln bonds that are resistant to cleavage by the enzyme indicating that the linear amino acid sequence must play a significant role.

There are two differences between the cleavage site and the stable Pro–Gln bonds in the α1(I) chain. First, it is not surrounded by glycine residues as is the case for the Pro–Gln bonds in the triple helical region. Second, in the region of the cleavage site there are several amino acids with large hydrophobic side chains, which are not observed in the neighbourhood of the stable Pro–Gln bonds.

Little can be said about the three dimensional structure of the enzyme binding site on native type I procollagen as the cleavage site sequence of the pα2 chain is not yet known. It also remains to be clarified whether the conversion of the N-terminal glutamine to the pyrroglutamic acid which is found as N-terminus of the collagens so far described is carried out by the enzyme or by another enzyme or perhaps occurs spontaneously.

Acknowledgements

We wish to acknowledge the expert technical assistance of Mrs Veronique Heinemann and Miss Gudrun Landrath. The study was supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 51, project B/12.

References

[1] Prockop, D. J., Berg, R. A., Kivirikko, K. J. and Uitto, I. (1976) in: Biochemistry of Collagen (Ramachandran, G. and Reddi, A. H. eds) Plenum, p. 163, New York.
[2] Lenaers, A., Ansay, M., Nusgens, B. V. and Lapière, C. M. (1971) Eur. J. Biochem. 23, 533–543.
[3] Fjolstad, M. and Helle, O. (1974) J. Path. 112, 183–188.
[4] Lapière, C. M., Lenaers, A. and Kohn, L. D. (1971) Proc. Natl. Acad. Sci. USA 68, 3054–3058.
[5] Kohn, L. D., Isersky, C., Zupnik, J., Lenaers, A., Lee, G. and Lapière, C. M. (1974) Proc. Natl. Acad. Sci. USA 71, 40–44.
[6] Becker, U., Timpl, R., Helle, O. and Prockop, D. J. (1976) Biochemistry 15, 2853–2862.
[7] Rauterberg, J. and Kühn, K. (1971) Eur. J. Biochem. 19, 398–407.
[8] Fietzek, P. P. and Rexrodt, F. W. (1975) Eur. J. Biochem. 59, 113–118.
[9] Brandt, W. F., Edman, P., Henschen, A. and Von Holt, C. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 1505–1508.
[10] Rauterberg, J., Timpl, R. and Furthmayr, H. (1972) Eur. J. Biochem. 27, 231–237.