# EVIDENCE FOR A NON-HELICAL REGION AT THE CARBOXYL TERMINUS OF THE COLLAGEN MOLECULE

M. STARK, J. RAUTERBERG and K. KÜHN
*Max-Planck-Institut für Eisweiss- und Lederforschung, Abteilung Kühn, D-8 München 15, Schillerstrasse 46, Germany*

Received 28 December 1970

## 1. Introduction

The triple helical structure of collagen requires that glycine occupies every third position in the amino acid sequence of its constituent polypeptide chain. No deviation from this pattern has been detected so far in those regions of the chains which are known to be part of the triple helical structure. The *N*-terminal sequence of approximately 15 amino acids of the α1- and α2-chains of several vertebrate collagens, however, does not contain glycine in every third position [1–5]. Consequently, this region cannot be part of the helical structure and was, therefore, designated 'non-helical'. This conclusion was confirmed by the observation that this region is susceptible to the attack of several proteases under conditions in which the helical structure of the molecule is not disturbed.

First evidence for an additional non-helical region at the *C*-terminus originated from immunologic studies of calf skin collagen. Pontz et al. [6] found the predominant antigenic site to lodge within the *C*-terminal cyanogen bromide peptide α1-CB6. Antigenic activity was abolished by proteolytic treatment under conditions which left the helical structure of the molecule unaffected. Thus, the antigenic determinant is likely to belong to a non-helical region.

Cyanogen bromide cleavage of the α1-chain of calf skin collagen yielded two species of α1-CB6, $CB6^a$ and $CB6^b$ [5], both showing serological activity. After chymotryptic pretreatment of the native collagen, only one kind of CB6 could be isolated, designated $\alpha1\text{-}CB6^{Chy}$, which very closely resembles $\alpha1\text{-}CB6^b$ [7]. However, no antigenic determinant could be detected on this peptide [8].

This paper describes the isolation and characterization of a chymotryptic peptide of $\alpha1\text{-}CB6^a$ containing the *C*-terminal non-helical region of the α1-chain.

## 2. Methods

Cyanogen bromide peptide $\alpha1\text{-}CB6^a$ was isolated from a cyanogen bromide digest of acid soluble calf skin collagen as described earlier [5]. The peptide $\alpha1\text{-}CB6^{Chy}$ was isolated in the same manner from acid soluble collagen, pretreated with chymotrypsin as follows: 200 mg of acid soluble calf skin collagen were dissolved in 100 ml 0.1 percent acetic acid and dialysed against 0.05 M tris buffer, pH 7.5, containing 0.1 M $CaCl_2$. After addition of 20 mg α-chymotrypsin (48 U/mg, Worthington) the mixture was incubated for 24 hr at 20°. The solution was acidified to pH 3.5 by addition of acetic acid and then dialysed against 0.1 percent acetic acid. The collagen was precipitated by addition of a concentrated NaCl solution to a final NaCl concentration of 6 percent. The precipitate was collected by centrifugation and dissolved in 0.1 percent acetic acid. Following an additional precipitation it was redissolved, dialysed exhaustively against 0.1 percent acetic acid, and lyophilized.

Chymotryptic cleavage of $\alpha1\text{-}CB6^a$ and $\alpha1\text{-}CB6^{Chy}$ was carried out on 50 mg of peptide dissolved in 10 ml 0.2 M $NH_4HCO_3$ solution by addition of 1 mg

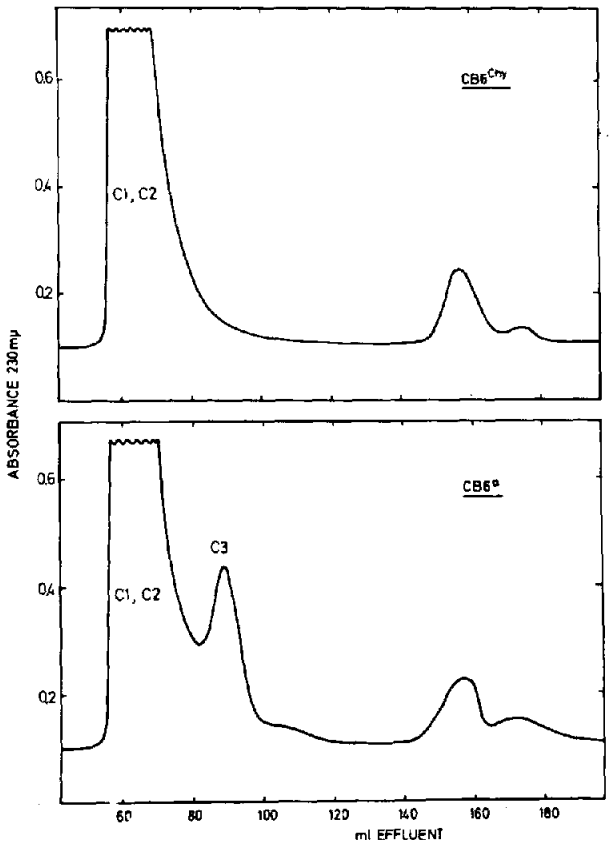


Fig. 1. Molecular sieve chromatography on Bio-Gel P-4 of the chymotryptic digests of $CB6^{Chy}$ and $CB6^{a}$. The column was eluted with 0.1 M acetic acid.

chymotrypsin dissolved in 1 ml 0.02 percent solution of $CaCl_2$ and incubation for 40 min at 37°. After cooling and addition of a few drops of glacial acetic acid the digestion products were lyophilized.

Molecular sieve chromatography was carried out on a 1.6 × 100 cm column of Bio-Gel P-4, which was eluted at room temperature with 0.1 M acetic acid, and on a 2.5 × 100 cm jacketed column of Bio-Gel P-150 which was eluted at 30° with 0.05 M sodium acetate buffer, pH 4.5.

Amino acid analysis was performed as described [5]. For detection of lysine derived α-aminoadipic-δ-semialdehyde (allysine) the samples were oxidized with performic acid as described by Moore [9].

For electronmicroscopy α1-$CB6^{Chy}$ and the chymotryptic peptides α1-CB6-C1 and α1-CB6-C2 were renatured and precipitated to segment long spacing fragments as described earlier [10].

## 3. Results

### 3.1. *Chymotrypic cleavage of α1-$CB6^{a}$ and α1-$CB6^{Chy}$*

Alpha 1-$CB6^{a}$ and α1-$CB6^{Chy}$ were cleaved by chymotrypsin, and the digests were chromatographed on Bio-Gel P-4. Both chromatograms are depicted in fig. 1. The first peak in both chromatograms contained two peptides with 115 and 82 amino acid residues respectively, which could be separated from each other on Bio-Gel P-150. They were designated α1-CB6-C1 and α1-CB6-C2. With respect to their chromatographic behaviour and amino acid composition, C1 as well as C2 from $CB6^{a}$ were identical to their respective analogs from α1-$CB6^{Chy}$. The second peak was present only in chymotryptic digests of $CB6^{a}$. It contained a single peptide of 20 amino acid residues which was termed C3.

The two minor peaks emerging with the total volume contained in both digests only small quantities of a tripeptide. The sequence of this tripeptide was determined by the Edman dansylation procedure and found to be Arg–Gly–Phe. This peptide could also be obtained by chymotryptic treatment of the isolated C1. Therefore, this peptide appears to be the result of an additional cleavage with C1.

### 3.2. *Amino acid composition of the chymotryptic peptides*

The amino acid composition of the three chymotryptic peptides of α1-$CB6^{a}$ is recorded in table 1. C1 was identical with α1*(78)-CBC, except for one additional phenylalanine in C1. Alpha 1*(78)-CBC had been isolated from the cyanogen bromide digest of a collagenase derived peptide by v.d. Mark et al., who had also determined its amino acid sequence [11, 12].

C3 was composed of 20 amino acid residues and contained only three glycines. This amino acid composition precludes participation of this peptide in the triple helical structure. After oxidation with performic acid and acid hydrolysis a new amino acid appeared which, from its chromatographic position, was identified as α-aminoadipic acid. This demonstrated the presence of α-aminoadipic-δ-semialdehyde

Table 1
Amino acid composition of the chymotryptic peptides of α1-CB6[a].

| | C1 | C2 | C3 | Total | α1-CB6[a]** |
|---|---|---|---|---|---|
| 3-Hydroxyproline | – | 1 (0,9) | – | 1 | 1 |
| 4-Hydroxyproline | 6 (6,4) | 9 (9,0) | – | 15 | 14 |
| Aspartic Acid | 5 (4,7) | 3 (2,6) | 2 (1,8) | 10 | 10 |
| Threonine | 3 (2,8) | 1 (1,1) | – | 4 | 4 |
| Serine | 4 (3,7) | 5 (5,1) | 1 (1,0) | 10 | 9 |
| Glutamic Acid | 9 (9,4) | 3 (3,4) | 3 (3,2) | 15 | 15 |
| Proline | 17 | 17 | 3 (3,4) | 37 | 37 |
| Glycine | 38 | 26 | 3 (2,7) | 67 | 68 |
| Alanine | 14 | 6 (6,3) | 1 (1,1) | 21 | 21 |
| Valine | 2 (2,1) | 1 (0,8) | | 3 | 3 |
| Isoleucine | 2 (2,0) | 1 (0,8) | | 3 | 3 |
| Leucine | 1 (1,1) | 3 (2,9) | 2 (2,0) | 6 | 6 |
| Tyrosine | – | 1 (0,7) | 1 (1,4) | 2 | 2 |
| Phenylalanine | 1 (0,6) | – | 1 (1,0) | 2 | 2 |
| Hydroxylysine | 2 (2,2) | – | – | 2 | 2 |
| Histidine | 1 (0,9) | – | 1 (1,0) | 2 | 2 |
| Lysine | 2 (2,2) | 2 (1,6) | (1)* | 4 + (1)* | 4 |
| Arginine | 8 (7,6) | 3 (3,2) | 1 (1,0) | 12 | 12 |
| Total | 115 | 82 | 20 | 217 | 215 |

* Lysine derived aldehyde (allysine)
** from [5]

(Allysin) in C3. This lysine derived aldehyde is already known to be present in α1-CB1 of several collagens. No C3 with unaltered lysine could be detected so far. The *N*-terminal cyanogen bromide peptide α1-CB1, on the other hand, occurred in two forms, containing either lysine or lysine derived aldehyde.

### 3.3. *The order of the chymotryptic peptides*

The position of C1 could be inferred from its amino acid composition which is almost identical with that of α1*(78) CBC, a peptide known to be the *N*-terminal part of α1-CB6. C3 could be obtained from the native molecule by chymotryptic digestion, whereas C1 and C2 were not affected by such treatment. Therefore, the position of C3 would be *C*-terminal. The order of the chymotryptic peptides, therefore, is 1-2-3.

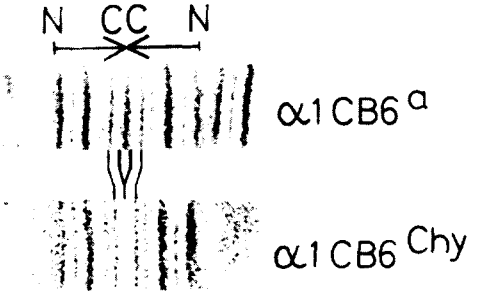


Fig. 2. Electron micrographs of segment long spacing fragments of CB6[a] and CB6[Chy], stained with phosphotungstic acid and uranyl acetate. The segment fragments are aggregated at their *C*-ends.

This sequence was confirmed by electron microscopy. C1 and C2 were renatured and precipitated with ATP to segment long spacing fragments. From its cross striation pattern C2 was identified as the *C*-terminal region of $\alpha$1-CB6$^{Chy}$. C1 exhibited antiparallel aggregation within the segment fragments and, therefore, its position could not be determined on the basis of its cross striation pattern. The *C*-terminal position of C3 in $\alpha$1-CB6$^{a}$ was confirmed by comparison of segment long spacing elements of $\alpha$1-CB6$^{a}$ and $\alpha$1-CB6$^{Chy}$. The segment fragments shown in fig. 2 are aggregated at their *C*-terminals. The broad, dark band in the *C*-terminal aggregation zone of CB6$^{a}$ is present as only a very small band in the aggregated fragments of CB6$^{Chy}$.

## 4. Conclusions

The present study reveals a striking similarity between the *N*-terminal and the *C*-terminal regions of the $\alpha$1 chains of calf skin collagen. They both contain a non-helical region of approximately 15 amino acids, which is susceptible to proteolytic attack under conditions which leave the helical part of the molecule unaffected. Both regions have a relatively high content of aromatic amino acids. In addition, each of them contains a lysine derived aldehyde.

The presence of the aldehyde suggests that the *C*-terminal region may be equally important for the formation of cross links as the non-helical region at the *N*-terminus of the collagen molecule.

## Acknowledgements

This work was supported by the Deutsches Forschungsgemeinschaft, SFB51. We wish to thank Miss Ingeborg Langer and Miss Waltraud Wohland for their expert technical assistance.

## References

[1] P. Bornstein and K.A. Piez, Biochemistry 5 (1966) 3460.
[2] P. Bornstein, A.H. Kang and K.A. Piez, Biochemistry 5 (1966) 3803.
[3] E.J. Miller, J.M. Lane and K.A. Piez, Biochemistry 8 (1969) 70.
[4] P. Bornstein, Science 161 (1968) 592.
[5] J. Rauterberg and K. Kühn, European J. Biochem., submitted for publication.
[6] B. Pontz, W. Meigel, J. Rauterberg and K. Kühn, European J. Biochem. 16 (1970) 50.
[7] S. Moore, J. Biol. Chem. 238 (1963) 235.
[8] J. Rauterberg and K. Kühn, FEBS Letters 1 (1968) 230.
[9] R. Timpl, U. Becker, H. Furthmayr and K. Kühn, Immunochemistry 7 (1970) 876.
[10] K.v.d. Mark, F. Rexrodt, P. Wendt and K. Kühn, European J. Biochem. 16 (1970) 143.
[11] K.v.d. Mark, P. Wendt, F. Rexrodt and K. Kühn, FEBS Letters 11 (1970) 105.