DIRECT EVIDENCE FOR A CORRELATION BETWEEN AMINO ACID SEQUENCE AND CROSS STRIATION PATTERN OF COLLAGEN

Klaus von der MARK, Peter WENDT, Friedrich REXRODT and Klaus KÜHN
Max-Planck-Institut für Eiweiss- und Lederforschung, Schillerstrasse 46, D-8 München, Germany

Received 2 October 1970

1. Introduction

Upon addition of ATP to an acidic solution of collagen, "Segment long spacing" crystallites called SLS are formed. The SLS crystallites represent an artifical quaternary structure of collagen in which the molecules are aggregated laterally and parallel to one another, with their ends in register and which therefore have the same length as the collagen molecule. After staining with phosphotungstic acid (PTA) or uranyl acetate (UA), the SLS structures show a characteristic cross striation pattern with alternating dark and light bands. Since PTA is bound to the basic amino acids, arginine and lysine [1] and uranyl cations react with the acidic amino acids [2], it has been suggested that the dark bands represent clusters of polar amino acids, whereas the light bands indicate regions rich in neutral amino acids [3].

Recently we have been able to isolate a peptide $\alpha 1$*(780)-CBC from the COOH-terminal region of the $\alpha 1$ chain of calf skin collagen [4] and to determine the sequence of the 112 amino acids present in the peptide [5]. We have observed that the SLS cross-striation pattern of this part of the molecule is especially well resolved. This allowed a direct comparison of the cross-striation pattern with the amino acid sequence of a discrete region in the collagen molecule.

2. Experimental

SLS crystallites were obtained by dialysis of acid soluble calf skin collagen dissolved in 1% acetic acid against a 0.4% ATP solution at pH 2.8. The precipitate which formed during dialysis was applied to a grid for electron microscopic investigations. The preparations were stained on the grid with PTA, and after washing with water they were subsequently stained with UA [6].

The double staining technique provides not only for the binding of PTA and UA to polar amino acids, but in addition, the uranyl anions form insoluble precipitates with ATP and PTA cations present in the polar regions. This intensifies the degree of staining and increases the contrast of polar regions.

The peptide $\alpha 1$*(780)-CBC was isolated from the collagenase-derived peptide $\alpha 1$*(780) after cleavage with CNBr as described earlier [4]. The position of $\alpha 1$*(780)-CBC in the whole collagen molecule is shown in the SLS pattern depicted in fig. 1. The determination of the primary structure of the peptide was initiated by cleavage with trypsin. The tryptic peptides were isolated by ion exchange chromatography on Dowex 1 X 2 and Aminex A6 and degraded stepwise with the combined Edman-Dansyl-procedure [7]. Larger peptides were subjected to enzymatic fragmentation by collagenase and thermolysin and the resulting peptides were sequenced as described above. Digestion of intact $\alpha 1$*(780)-CBC by thermolysin gave overlap peptides, which led to the ordering of the tryptic peptides as shown in fig. 2. Experimental details with regard to amino acid sequencing will be published elsewhere [5].

3. Results and discussion

The amino acid sequence of $\alpha 1$*(780)-CBC is shown in fig. 2. The polar amino acids are restricted to a few regions and are not uniformly distributed

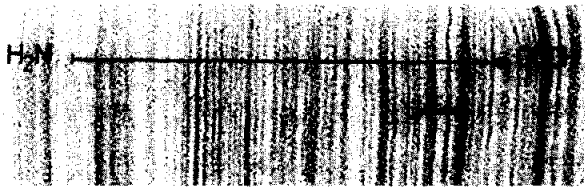

Fig. 1. Segment-long-spacing crystallites from soluble calf skin collagen stained with phosphotungstic acid and uranyl acetate. The position of the peptide α1*(780)-CBC (112 amino acid residues in length) is indicated by the black line in the electron micrograph. Overlapping segments from other SLS crystallites prevent the usual distortion observed in the terminal regions.

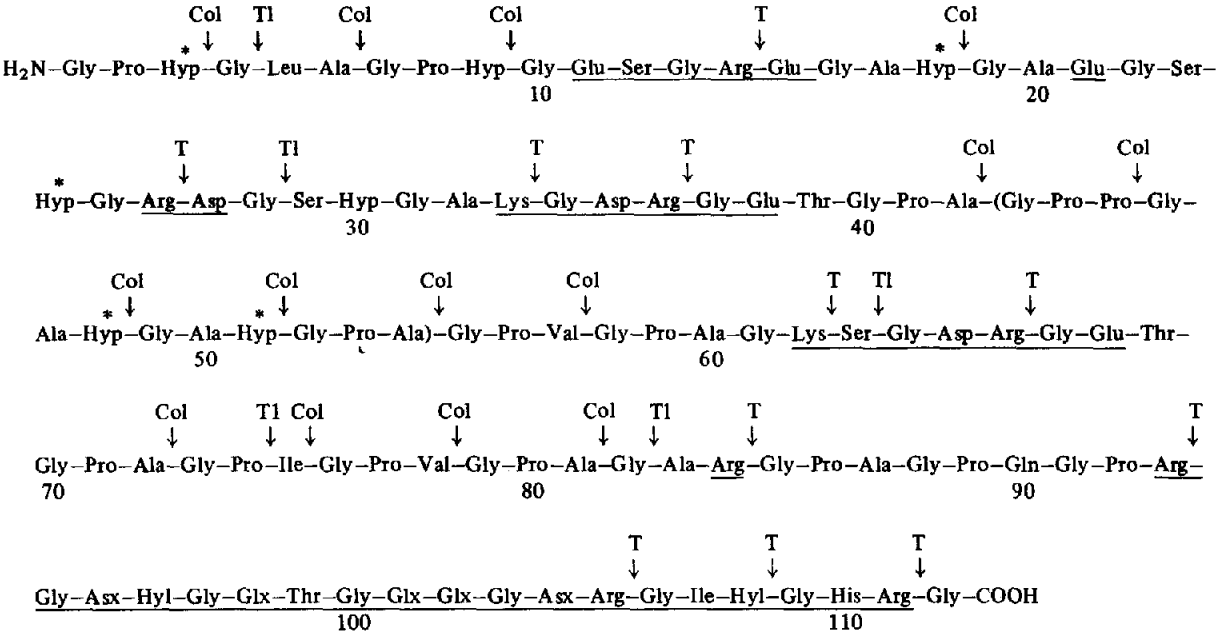

Fig. 2. Amino acid sequence of α1*(780)-CBC. The sequence given in brackets is not as yet completely confirmed. The location of enzyme splits used to establish the sequence is shown by arrows. T = trypsin, T1 = thermolysine, Col = bacterial collgenase. * Hydroxyproline residues only partially hydroxylated. Polar regions are underlined.

along the peptide chain. In only two instances do polar amino acids occur outside of clusters and these are glutamic acid in position 21 and arginine in position 84. For the most part, all polar regions contain acidic amino acids as well as basic amino acids. In this regard, however, the sequence represented by the 19 amino acid residues at the COOH-terminal region is unique. This sequence can be divided into three parts: the first contains an equal amount of acidic and basic amino acids similar to other polar regions appearing on the NH_2-terminal side of residue 92. The second one, however, consists mainly of acidic and the third one exclusively of basic amino acid residues.

The middle part of the entire peptide (position 39–93) is composed chiefly of apolar amino acids. It is divided into two halves by a polar region of seven amino acids (residues 62–68). The apolar region consists mainly of typical collagen sequences, like Gly–Pro–X, Gly–X–Pro and Gly–X–Hyp.

In fig. 3 the amino acid sequence of α1*(780)-CBC

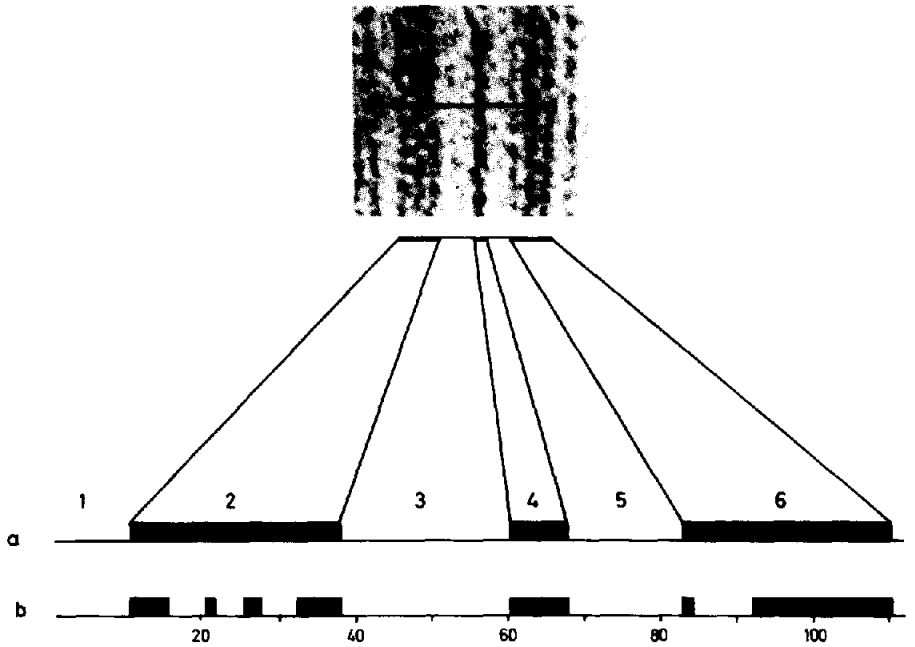

Fig. 3. Correlation of the SLS cross striation pattern with the amino acid sequence of α1*(780)-CBC. (a) Schematic representation of the light and dark bands. The individual bands are indicated with numbers; (b) schematic representation of the apolar and polar regions of the amino acid sequence. Comparison of the length of light and dark bands (in Å) with the corresponding apolar and polar regions of the sequence in numbers of amino acids. See. fig. 1.

Band	Å (SLS-pattern)	Number of amino acids (sequence)
1	30	10
2	80	28
3	65	23
4	22	7
5	43	15
6	80	28

is compared with the corresponding region of the SLS cross striation pattern. Presuming that the length of the entire long spacing segment (fig. 1) is 2.990 Å, which corresponds with 1030 amino acids per α chain, the length of the region comprising α1*(780)-CBC as measured in the electron microscope is approximately 320–330 Å. This value is the expected length for a peptide consisting of 112 amino acids. Electron micrographs of the α1*(780)-CBC region routinely show three dark and three light bands. In fig. 3 the alternating light and dark bands observed in the electron microscope are aligned with corresponding apolar and polar regions as determined from the amino acid sequence of the peptide. In making the alignment, it was presumed that each amino acid residue is 2.9 Å in length and there is an excellent correlation between the lengths of the individual bands and the calculated lengths of apolar and polar regions (see legend of fig. 3).

In high-resolution electron photomicrographs the NH_2-terminal cross striation is divided according to the sequence into three bands. The cross striation at

the –COOH end (no. 6) is also further resolved. In this fashion the arginine residue in position 84 is reflected as a single band. This surprising finding may be due to the accumulation of ATP and PTA around the arginine residue, causing a further precipitation of uranyl cations. In spite of the fact that the polar region (no. 6) comprising positions 93–110 (fig. 2) consists almost exclusively of polar amino acids, two bands are visible electron optically at high resolution. This suggests that the acidic amino acids in positions 101, 102 and 104 are amidated. On the other hand, it is as yet not known to what extent the α2 chain influences the cross striation pattern of SLS, which are built up from $(\alpha 1)_2\alpha 2$ triple helical molecules. Although it has been found that collagen molecules, obtained by renaturation of either α1 or α2 chains, have almost identical cross striation patterns [8], we have not succeeded in obtaining electron micrographs from them with high resolution, comparable to that achieved with the native collagen molecule as utilized in this study.

At the H_2N-terminus of the α1 chain of rat skin collagen, a sequence of 132 amino acid residues is now known (9.10.11.12). A clustering of the polar amino acids within this sequence could not be observed to the same degree as shown here in the case of α1*(780)-CBC. This may be seen in fig. 1 which indicates that the SLS cross striation pattern at the H_2N-terminus does not show a comparably sharp and subdivided cross striation pattern as can be seen at the –COOH terminus of the molecule.

Acknowledgements

We wish to thank Miss Karin Holland and Miss Ingrid Kell for their valuable assistance. This work was supported by the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 51).

References

[1] K.Kühn, W.Grassmann and U.Z.Hofmann, Naturforsch. 13b (1958) 154.
[2] E.Gebhardt and K.Kühn, Z. Anorg. Chem. 320 (1963) 71.
[3] K.Kühn and E.Zimmer, Naturforsch. 16b (1961) 648.
[4] K.von der Mark, F.Rexrodt, P.Wendt and K.Kühn, European J. Biochem. 16 (1970) 143.
[5] K.von der Mark, P.Wendt and K.Kühn, in preparation.
[6] M.Stark and K.Kühn, European J. Biochem. 6 (1968) 648.
[7] R.Gray, Methods Enzymol. 9 (1967) 469.
[8] C.Tkocz and K.Kühn, European J. Biochem. 7 (1969) 454.
[9] H.Kang, P.Bornstein and K.A.Piez, Biochemistry 6 (1967) 788.
[10] P.Bornstein, Biochemistry 6 (1967) 3082.
[11] W.T.Butler, in: Chemistry and Molecular Biology of Intercellular Matrix, Vol. 1, ed. E.Balazs (Academic Press, London, 1970) p. 149.
[12] W.T.Butler, Biochemistry 9 (1970) 44.