THE COVALENT STRUCTURE OF COLLAGEN: AMINO ACID SEQUENCE OF THE N-TERMINAL REGION OF α2-CB5 FROM RAT SKIN COLLAGEN

Peter P. FIETZEK and Klaus KÜHN
Max-Planck-Institut für Biochemie,
8033 Martinsried, GFR

Received 25 July 1973
Revised version received 17 August 1973

1. Introduction

The triple helical molecule of skin collagen contains two different peptide chains: two α1(I)- and one α2-chain [cf 1]. However, there are collagen molecules present in other tissues such as cartilage which are composed of three identical chains, designated α1(II)-chains [2]. Collagen from evolutionary primitive species such as sea anemone likewise consists of three identical polypeptide chains [3]. Artificial molecules of composition $[\alpha 1(I)]_3$ which may be obtained upon renaturation of type I α1-chains display properties similar to those of natural skin collagen [4]. Thus, their renaturation temperature is similar to that of $[\alpha 1(I)]_2$ and they may aggregate to form fibrils of the native type. In contrast, much less stability is displayed by molecules consisting entirely of α2-chains [4]. Therefore, the question may be raised, what biological advantage is associated with collagen molecules consisting of two different chains such as the collagens present in skin, bones and tendons. An answer may be expected from comparison of the sequence data of α1- and α2-chains.

The sequence of the α1(I)-chain has been completely elucidated by investigations of calf and rat skin collagen [5–16]. In contrast, investigations of the α2-chain are still in the early stages. So far, only the sequences of the three small CNBr-derived peptides α2-CB1, α2-CB0, and α2-CB2 have been reported. α2-CB1 is the N-terminal peptide of the α2-chain and comprises of 14 amino acid residues. α2-CB0 is the adjacent tripeptide, and α2-CB2 is derived from a central region of the chain and consists of 30 residues [17–19, 22]. The sequence of the 42 N-terminal residues of α2-CB4, the peptide immediately following α2-CB0 has been recently determined [20] by the procedure of automated stepwise degradation according to Edman and Begg [21]. Thus, the sequence of the first 60 residues of the α2-chain is known and may be compared to the homologous segment of the α1-chain.

In the present communication we describe the sequence of the 45 N-terminal residues of the CNBr-peptide α2-CB5 from rat skin collagen. The entire peptide consists of approximately 320 amino acid residues and represents the C-terminal third of the α2-chain. These data now facilitate comparisons between areas of the α1- and the α2-chain in the C-terminal region of the molecule. Our investigations revealed 8 substitutions of amino acid residues between the α1- and the α2-chains.

2. Experimental

The peptide α2-CB5 was prepared from the skins of normal and lathyritic rats by the method previously described [22, 26]. Automated stepwise degradation of the peptide was achieved in a sequencer (Model 890, Beckman Instruments, Palo Alto, Calif., USA) employing a modification of the method of Edman and Begg [21] (Beckman Quadrol double cleavage program No. 111070). Four replicate samples of 0.33–0.38 μmoles were subjected to the procedure; The liberated PTH-amino acid residues were identified by gas–liquid- [23] and thin layer-chromatography [24] as described previously [14, 16]. PTH arginine was identified by thin layer-electrophoresis. Discrimination

between leucine and isoleucine was achieved by hydrolysis of the PTH-derivatives and separation on an amino acid analyser of the free amino acid residues obtained.

3. Results and discussion

The sequence of the 45 N-terminal residues of α2-CB5 as established by automated Edman degradation is depicted in fig. 1. This represents a typical collagen sequence resembling that of the helical areas of the α1-chain. Every third position along the chain is occupied by glycine. Only proline residues in position Y of the tripeptide unit Gly–X–Y are hydroxylated. Phenylalanine (which, in the α1-chain was found exclusively in position X) is present in position X.

The position within the α2-chain of the N-terminal end of α2-CB5 has been fairly accurately determined by electron microscopical investigation of CNBr peptides which had been renatured to yield triple helical fragments of the molecule [25, 26]. The sequence area presently studied may, therefore, be compared to the homologous region of the α1-chain, i.e. position 142–186 of α1-CB7 [16]. This comparison is included in fig. 1. Theoretically, the chains may be shifted against each other by one or two triplets. This would, however, entail considerable reduction in homology. The final decision must be based on investigations of other, overlapping peptides. Of particular interest is the striking similarity of the α1- and α2-chain in this section of the molecule where eight exchanges were found within only 45 residues.

In considering the amino acid exchanges it is important to note the effect of the changes on the stability of the molecule and its interactions in fibril formation. The stability of the triple helix depends to a large extent on the content and distribution of the proline and hydroxyproline residues. The stability increases in the series Gly–X–Pro, Gly–Pro–Y and Gly–Pro–Pro. Since the denaturation temperature of the $[\alpha 2]_3$ molecule is about 10°C lower than that of the $[\alpha 1(I)]_3$ molecule, and since both contain an approximately equal amount of proline or hydroxyproline, one might expect to find a dissimilar distribution of the imino acid residues along the chain [4]. In the section of the molecule investigated here the distribution of imino acid residues is very similar. Only one proline residue of the α1-chain is substituted by an alanine residue in the α2-chain. In contrast, in the N-terminal region of the collagen molecule comprising 48 residues the number of triplets containing proline and/or hydroxyproline was found to be less in the α2-chain compared with the α1-chain [20].

The charged polar residues and the non-polar residues with large hydrophobic side-chains (such as leucine, isoleucine, valine, phenylalanine and methionine) especially influence the interaction of the collagen molecules [27]. Of the eight exchanges found, only one involved a charged polar amino acid residue. Since glutamic acid is exchanged for aspartic acid the charge pattern is essentially retained. This is in agreement with the very similar cross striation pattern of the long spacing segments obtained from $[\alpha 1(I)]_3$ or $[\alpha 2]_3$ molecules [4].

More exchanges were found in the group of amino acids with large hydrophobic side-chains. In two positions isoleucine in the α2-chain was replaced by an uncharged polar amino acid (asparagine and serine) in the α1-chain. Similarly, one valine residue in the α1-chain was replaced by threonine in the α2-chain. In addition an exchange of alanine (α1-chain) for valine (α2-chain) was found. All those exchanges can be

α1 Thr–Gly–Phe–Hyp–Gly–Ala–Ala–Gly–Arg–[Val]–Gly–Pro–Hyp–Gly–Pro–
α2 Thr–Gly–Phe–Hyp–Gly–Ala–Ala–Gly–Arg–[Thr]–Gly–Pro–Hyp–Gly–Pro–

α1 Ser–Gly–[Asn]–[Ala]–Gly–Pro–Hyp–Gly–Pro–Hyp–Gly–[Pro]–Ala–Gly–Lys–
α2 Ser–Gly–[Ile]–[Thr]–Gly–Pro–Hyp–Gly–Pro–Hyp–Gly–[Ala]–Ala–Gly–Lys–

α1 Glu–Gly–[Ser]–Lys–Gly–Pro–Arg–Gly–[Glu]–[Thr]–Gly–Pro–[Ala]–Gly–Arg–
α2 Glu–Gly–[Ile]–Lys–Gly–Pro–Arg–Gly–[Asp]–[Gln]–Gly–Pro–[Val]–Gly–Arg–

Fig. 1. Comparison of the amino acid sequence of residues 142 to 186 of α1-CB7 from calf skin collagen and residues 1 to 45 of α2-CB5 from rat skin collagen. Non identical residues found in the same position are enclosed in boxes.

considered to be contributing to an alteration of the hydrophobic feature of this region. A more detailed discussion of the effect of the various substitutions on the stability and interactions of the collagen molecule must await further sequence data on the α2-chain.

Acknowledgements

We wish to thank Miss Amira Djonlagic and Miss Ingrid Kell for experimental assistance and the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 51) for financial support.

References

[1] Traub, W. and Piez, K.A. (1971) Advan. Protein Chem. 25, 243–325.
[2] Miller, E.J. and Matukas, V.J. (1969) Proc. Natl. Acad. Sci. U.S. 64, 1264–1268.
[3] Nordwig, A. and Hayduk, U. (1969) J. Mol. Biol. 44, 161–172.
[4] Tkocz, C. and Kühn, K. (1969) European J. Biochem. 7, 454–462.
[5] Bornstein, P. (1967) Biochemistry 6, 3082–3092.
[6] Bornstein, P. (1969) Biochemistry 8, 63–71.
[7] Butler, W.T. (1970) Biochemistry 9, 44–50.
[8] Butler, W.T. and Ponds, S.L. (1971) Biochemistry 10, 2076–2081.
[9] Balian, G., Click, E.M. and Bornstein, P. (1971) Biochemistry 10, 4470–4478.
[10] Rauterberg, J., Timpl, R. and Furthmayer, H. (1972) European J. Biochem. 27, 231–237.
[11] Balian, G., Click, E.M., Hermodson, M.A. and Bornstein, P. (1972) Biochemistry 11, 3798–3806.
[12] Rauterberg, J., Fietzek, P.P., Rexrodt, F.W., Becker, U., Stark, M. and Kuhn, J. (1972) FEBS Letters 21, 75–79.
[13] Wendt, P., Mark, K. v.d., Rexrodt, F.W. and Kuhn, K. (1972) European J. Biochem. 30, 169–183.
[14] Fietzek, P.P., Rexrodt, F.W., Wendt, P., Stark, M. and Kuhn, K. (1972) European J. Biochem. 30, 163–168.
[15] Fietzek, P.P., Wendt, P., Kell, I. and Kühn, K. (1972) FEBS Letters 26, 74–76.
[16] Fietzek, P.P., Rexrodt, F.W., Hopper, K. and Kühn, K. (1973) European J. Biochem., in press.
[17] Kang, A.H., Bornstein, P. and Piez, K.A. (1967) Biochemistry 6, 788–795.
[18] Kang, A.H. and Gross, J. (1970) Biochemistry 9, 796–804.
[19] Highberger, J.H., Kang, A.H. and Gross, J. (1971) Biochemistry 10, 610–616.
[20] Fietzek, P.P., Kell, I. and Kühn, K. (1972) FEBS Letters 26, 66–68.
[21] Edman, P. and Begg, G. (1967) European J. Biochem. 1, 80–91.
[22] Fietzek, P.P. and Piez, K.A. (1969) Biochemistry 8, 2129–2133.
[23] Pisano, J.J. and Bronzert, T.J. (1969) J. Biol. Chem. 244, 5597–5607.
[24] Edman, P. (1970) in: Protein Sequence Determination (Needleman, S.B., ed.), p. 211, Springer Verlag, Berlin, Heidelberg, New York.
[25] Igarashi, S., Kang, A.H. and Gross, J. (1970) Biochem. Biophys. Res. Commun. 38, 697–702.
[26] Fietzek, P.P. and Kühn, K., in preparation.
[27] Hulmes, D.I.S., Miller, A., Parry, D.A.D., Piez, K.A. and Woodhead-Galloway, J. (1973) J. Mol. Biol., in press.