CHEMISTRY LETTERS, pp. 509-512, 1978. Published by the Chemical Society of Japan

ANALYSIS OF MICROFIBRILLAR STRUCTURE OF COLLAGEN BASED ON α_1(I) CHAIN SEQUENCE AND 7/1 HELIX MODEL OF TROPOCOLLAGEN

Kenji OKUYAMA[1], Nobuhiro GO[2] and Motowo TAKAYANAGI[1]

Department of Applied Chemistry, Faculty of Engineering[1], and Department of Physics, Faculty of Science[2], Kyushu University, Fukuoka 812

Hydrophobic interactions between tropocollagens are analyzed by using the amino acid sequence in α_1(I) chain and the 7/1 helix model of tropocollagen. It is concluded that tropocollagens aggregate to form a five-stranded microfibril. The D/11 periodicity of hydrophobic interactions is understood naturally by a pitch of a tropocollagen.

Collagen type I, the commonest collagen existing in skin, bone and tendon, is unique in having a heterogeneous chain composition, $[\alpha_1(I)]_2\alpha_2$. The sequence of 1052 amino acid residues in α_1(I) chain is known for calf and rat collagen.[1)] After the whole amino acid sequence of α_1(I) chain was determined, many researchers have tried to interpret the collagen structure in terms of its amino acid sequence. Hulmes et al.[1)] studied both interactions between charged residues and between large hydrophobic residues of two α_1(I) chains (the triplet region of 1011 amino acid residues only) as a function of staggered length between them. They found that the interactions were maximal when the chains are staggered by 0D, 1D, 2D, 3D and 4D, where D = 234±1 residues. This provides an explanation of the D ≃ 670 Å period in collagen fibril.

In their calculation, hydrophobic interactions were found to have another clear periodicity of D/11 (∿21 residues). McLachlan[2)] and Hulmes et al.[3)] carried out Fourier analysis of the amino acid sequence in α_1(I) chain, and found independently that the distribution of hydrophobic amino acids shows apparent periodicities of D/5, D/6 and D/11. Significance of these periodicities, if any, should manifest itself in the geometrical structures of the triple helix and in how they assemble into the collagen fibril. However, if we consider the triple helix models of collagen proposed so far, i.e., the 10/1 helix model or Collagen II structure[4)] and the 12/1 helix model or "standard" structure,[5)] amino acid residues of these periodicities do not appear along a definite edge, which runs parallel to the helix axis. Piez and Trus[6)] regarded the periodicity of D/11 as arising from structural constraints imposed by super-coiling of microfibril and deduced the pitch length of a super-coil by using this periodicity.

A new structural model has been proposed recently for collagen by reassigning the layer line indicies to the reflections in the wide angle X-ray diffraction patterns of native collagen.[7)] In this model, tropocollagen is composed of three 7/1 helices with the same helical axis. Each helix has 7 tripeptides (21 residues) in one pitch. It is to be noted that a single crystal of $(Pro\text{-}Pro\text{-}Gly)_{10}$ shows a pseudoperiod of 20.08 Å, which is characteristic of the 7/1 helical structure.[8)] The number of residues in one pitch of this helix is the same as that of the D/11 periodicity of hydrophobic amino

acids in α_1(I) chain. The purpose of this paper is to clarify the structure of microfibril and hydrophobic interactions in it in view of the new structural model of tropocollagen.

Because the whole amino acid sequence of α_2 chain is not known at present, we will study a tropocollagen whose composition is $[\alpha_1(I)]_3$. This tropocollagen does not occur naturally but can be reconstituted from the solution containing only α_1(I) chain. This reconstituted tropocollagen is known to resemble the reconstituted $[\alpha_1(I)]_2\alpha_2$ tropocollagen in its physicochemical properties and electron-microscopic morphology.[9] The 1011 amino acid residues in the triplet region of α_1(I) chain are aligned so as to produce the periodicity of 21 residues. Glycyl residues at every third position of the sequence are located close to the central axis of the tropocollagen. Therefore, 14 residues in one turn of one chain appear on the surface of the tropocollagen. Each of them occupies one of fourteen equally divided azimuthal sections of the surface of tropocollagen. Distribution of residues having large hydrophobic side chains (Met, Val, Leu, Ile and Phe) in one chain (named as chain A, the other two chains being chain B and chain C) in the fourteen azimuthal sections is shown in Fig.1. Uneven distribution of hydrophobic residues is obvious in Fig.1. A number denotes the height of a residue along the helical axis in a residue unit of 2.87 Å. In Fig.1 non-glycyl residues in sequence -Gly-X_i-Y_i-Gly-X_{i+1}-Y_{i+1}- appear in the azimuthal sections in the order of ..., Y_i, X_i, Y_{i+1}, X_{i+1}, ... because of the turn of the minor helix of each chain. As a manifestation of the D/11 periodicity, many hydrophobic amino acid residues appear in section 2 of Fig.1. Chain B is made by rotating chain A about the central axis so as to superimpose section 1 on section 11 of chain A. Chain C is made similarly by superimposing section 1 on section 7. Residue heights in chains B and C are obtained by adding 1 and 2, respectively, to corresponding residue height in chain A. Thereby the surface structure of one triple helix is constructed. Many hydrophobic residues appear in section 2, 8 and 12 in thus constructed surface structure of the triple helix.

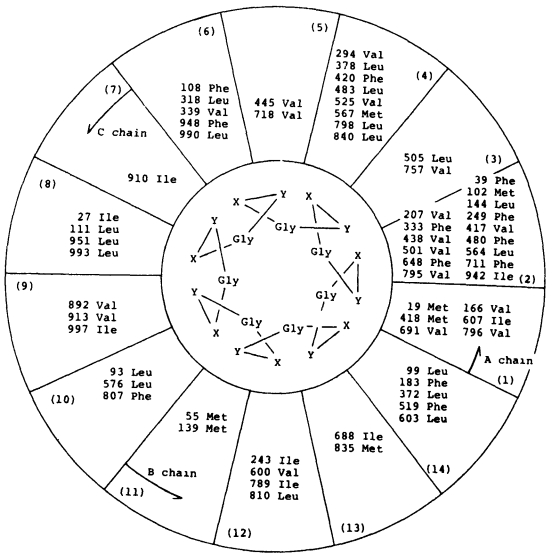

Fig.1 Distribution of hydrophobic residues in one α_1(I) chain on the surface of tropocollagen.

We now inspect the hydrophobic interactions between triple helices in a microfibril. For this purpose we consider three triple helices, TA, TB and TC. TA and TB are arranged in parallel and with staggering of 1D. TC is related to TB by a translation of -5D along the helical axis (Fig.2). Magnitude of the hydrophobic interactions is calculated by a similar scoring system as that of Hulmes et al.[1] They did not consider the triple-helical structure of tropocollagen, but studied interactions between hypothetical linear tropocollagens. Because we are considering the triple-helical structure in this paper, we apply the same scoring system as theirs only to those residues appearing along two edges which are facing each other. That is, when edges corresponding to j-th section of TA and i-th section of TB and TC are facing

each other, hydrophobic amino acid pairs from these sections with difference in height less than 3 residue units are given a score of one. All other pairs are scored zero. Scoring is carried out for all j-i pairs. Heights of residues which appear in the same azimuthal section of one tropocollagen differ in an integral multiple of the pitch length. Therefore, length of staggering D between adjacent tropocollagens must be very close to an integral multiple of the pitch length. By taking into account of the experimental values of 230∿235, we use D=231 (21x11) residue units.

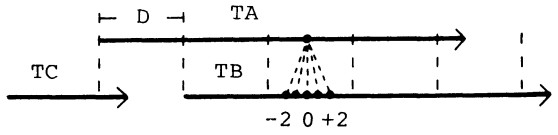

Fig.2 Relative position of tropocollagens TA, TB and TC.

Table 1 Scores of hydrophobic interactions between tropocollagens.

Section number j of tropocollagen TA	Section number i of tropocollagens TB and TC													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0	2	0	2	0	2	0	2	0	0	0	2	0	0
2	3	[10]	0	0	1	4	0	[9]	1	3	3	[9]	1	0
3	0	0	0	1	0	0	0	0	0	1	0	0	0	1
4	2	0	0	3	0	0	1	5	0	3	2	5	0	4
5	1	0	0	0	0	1	0	0	0	2	1	0	0	0
6	1	3	2	0	1	3	0	0	1	2	0	2	1	1
7	0	0	0	2	0	0	0	2	0	2	0	2	0	2
8	3	[10]	0	4	1	0	2	[9]	0	0	3	[9]	0	5
9	1	0	0	0	0	1	0	0	0	2	0	0	0	2
10	0	3	1	2	0	3	2	0	0	4	0	0	0	3
11	0	2	0	2	0	0	0	2	0	0	0	2	0	2
12	3	[10]	0	4	1	4	2	[9]	1	0	3	[9]	0	0
13	0	0	0	2	0	1	0	0	0	2	0	0	0	2
14	0	0	1	3	0	2	2	5	0	3	3	0	0	4

The result is shown in Table 1. It is clearly seen that 9 pairs formed by combinations of 2, 8 and 12 sections have significantly large scores than others. This is because these three sections have many hydrophobic residues. Let us study structures of microfibril that can be constructed from a repetitive use of inter-tropocollagen interaction of the same j-i pair. In three cases of 2-2, 8-8 and 12-12 pairs, two-stranded structure would result. In four cases of 2-12, 12-2, 8-12 and 12-8 pairs, five-stranded structure is most plausible one, because the local optimal angle 102.9° of rotation of adjacent tropocollagens for hydrophobic interactions is close to an internal angle of pentagon 108°. For a similar reason, fourteen-stranded structure is obtained in two cases of 2-8 and 8-2 pairs (Fig.3). However, this fourteen-stranded structure is not expected to be stable because of the large central vacancy. The five-stranded structure has been already proposed by Smith[10] as a microfibril structure and supported by an X-ray diffraction study.[11] Two microfibril structures constructed from j-i and i-j pair, respectively, are different in that a relative stagger between adjacent tropocollagens is D or -D. Therefore, helical arrangements of gap regions of the two microfibril structures are enantiomorphic to each other. It is not possible to conclude from the analysis in this paper which one of the four five-stranded structures is the most stable one.

It should be noted that in the five-stranded structures constructed from 2-12 or 12-2 pairs, 8 points out of 9 and 10 points, respectively, are scored when chain A appears in section 2 and chain B appears in section 12. Similarly, in the five-stranded structures constructed from 8-12 or 12-8 pairs, again 8 points out of 9 points are scored when chain C appears in section 8 and chain B appears in section 12. Therefore, the heights of 4 peaks in Table 1 will remain distinguished when α_1(I) chain is replaced by α_2 chain in the former cases at the position of chain C, and in the latter cases at the position of chain A. The sequence of α_2 chain is known about

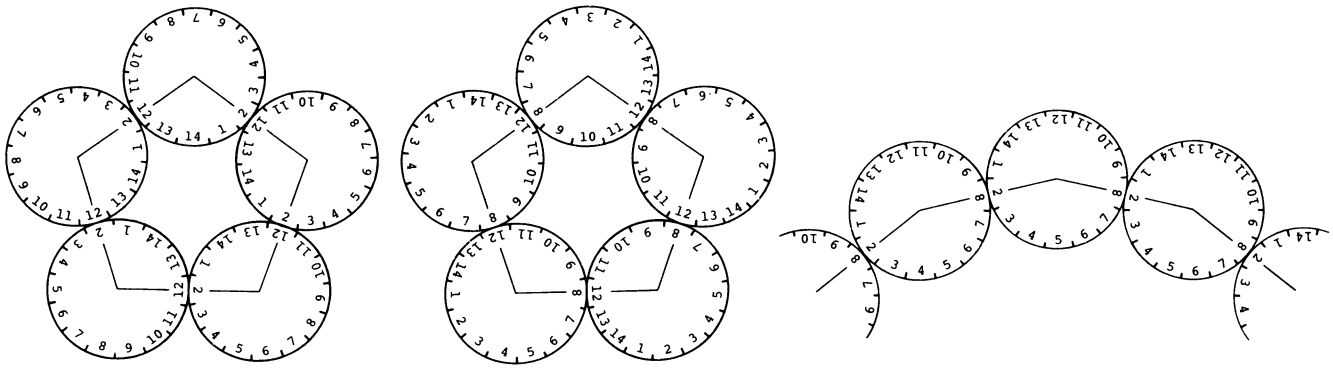

Fig.3 Cross sections of possible microfibril structures. (a) Five-stranded structure constructed from inter-tropocollagen interactions of 2-12 or 12-2 pairs. (b) Five-stranded structure constructed from 8-12 or 12-8 pairs. (c) Fourteen-stranded structure constructed from 2-8 or 8-2 pairs.

80 % at present.[12] Analyses of explicitly including this sequence of α_2 chain were also carried out. However, no results qualitatively different from those described above were obtained. The question of why there are two kinds of α-chains in collagen type I still remain open.

It is concluded that the five-stranded structure of microfibril, proposed previously by Smith,[10] is supported by an analysis based on the amino acid sequence of α_1(I) chain and the recently proposed 7/1 helix model of tropocollagen. If we assume other models of tropocollagen, 10/1 helix model or 12/1 helix model, no specific structure of microfibril can be deduced from similar analysis. Therefore, the present analysis can also be regarded as supporting the 7/1 helix model of tropocollagen, when the five-stranded structure of microfibril is more firmly established.

References

1) D. J. S. Hulmes, A. Miller, D. A. D. Parry, K. A. Piez and J. Woodhead-Galloway, J. Mol. Biol., 79, 137 (1973)
2) A. D. McLachlan, Biopolymers, 16, 1271 (1977)
3) D. J. S. Hulmes, A. Miller, D. A. D. Parry and J. Woodhead-Galloway, Biochem. Biophys. Res. Commun., 77, 574 (1977)
4) A. Rich and F. H. C. Crick, J. Mol. Biol., 3, 483 (1961)
5) G. N. Ramachandran, "Treatise on Collagen", Vol.1, G. N. Ramachandran Ed., Academic Press, New York, (1967) p.103
6) K. A. Piez and B. L. Trus, J. Mol. Biol., 110, 701 (1977)
7) K. Okuyama, M. Takayanagi, T. Ashida and M. Kakudo, Polymer J., 9, 341 (1977)
8) K. Okuyama, N. Tanaka, T. Ashida and M. Kakudo, Bull. Chem. Soc. Jpn., 49, 1805 (1976)
9) C. Tkocz and K. Kuhn, Eur, J, Biochem., 7, 454 (1969)
10) J. W. Smith, Nature, 219, 157 (1968)
11) A. Miller and J. S. Wray, Nature, 230, 437 (1971)
12) P. P. Fieztek and K. Kuhn, Int. Rev. Connect. Tissue Res., 7, 1 (1976)

(Received January 20, 1978)