Minimum-energy Conformation of Different Tropocollagen Models Calculated for Poly(Gly-Pro-Pro) by Considering Helical Constraints

Kenji Okuyama,* Haruaki Eto, Ikuo Nakaya, and Motowo Takayanagi
Department of Applied Chemistry, Faculty of Engineering, Kyushu University 36, Fukuoka 812
(Received November 27, 1979)

The minimum conformational energies for three helical models of tropocollagen were calculated by using energy functions and values of parameters obtained by Momany *et al.*. Three types of helical models containing 7/1, 10/1, and 12/1 helices were employed. Among the models with the 7/1 helix, only the conformation having the $N(Gly_1)\cdots O(Pro_2)$ type hydrogen bond was found to be energetically stable. Thus all the three helical models had the same type of hydrogen bond. The minimum-energy conformation of the 7/1 helix model was found to be more stable than those of the other two models by about 7 kcal/mol per tripeptide unit. This difference mainly came from the non-bonded energy terms of both intra- and inter-chains. Since the conformation of the 7/1 helix model was widely spread in the conformational space compared with those of the other models, it was concluded that the unit twist of the 7/1 helix model is well suited for peptide chains to conform a triplehelical structure. The above results strongly supported the 7/1 helix model for tropocollagen.

Collagen proteins form the connective tissues, such as bone, skin, tendon, cornea, and basement membrane, and are responsible for the structural integrity of the animal body. The collagen molecule (tropocollagen) is distinctive in containing the glycine residues in every third position, and has been designated as poly(Gly_1–R_2–R_3) in brief. Another feature of collagen is its high content of proline and hydroxyproline residues at the positions of R_2 and R_3. It is well known that the basic structure of tropocollagen has involved three stranded helices,[1] but its definite conformation has not been determined because of the meagre availability of the X-ray fiber patterns obtained from native collagen. So far, many researchers have proposed many structural models for tropocollagen. These models are basically divided into two types. One group contains 30 amino acid residues per turn in each chain of three strands,[2–5] and the other contains 36 residues per turn.[6] Recently two of us (K. O. and M. T.) and others[7] proposed the third type of collagen model having 21 amino acid residues per turn based on the structural analysis of single crystals of $(Pro–Pro–Gly)_{10}$ prepared.

On the other hand, some attempts have been made to determine the most stable conformation of tropocollagen by the method of conformational energy calculation. Tumanyan[8] minimized conformational energies of the triple-stranded helix of $(Gly–Pro–Pro)_n$ from many different combinations of the dihedral angles as starting points, using the Lennard-Jones pairwise potential function and the simple Morse-type potential function of hydrogen bonds. According to his calculations, the reasonable ranges for unit height and unit twist of the repeating unit Gly–Pro–Pro were 8.22—8.76 Å and 25—40°, respectively. Miller and Scheraga[9] calculated the conformational energies of $(Gly–Pro–Pro)_n$ for three different structures, that is, single-chain polymer, triple-stranded parallel-chain complexes and triple-stranded coiled-coils, and found that the last structure was the most stable. The unit heights and the unit twists of 10/1 and 12/1 helix models are within the corresponding ranges proposed by Tumanyan,[8] and the unit twist of the 7/1 helix model is far from this range. Miller and Scheraga,[9] however, showed that their lowest-energy structure was in good agreement with the 7/1 helix structure proposed for $(Pro–Pro–Gly)_{10}$.[10] The above calculations for $(Gly–Pro–Pro)_n$ were made without considering any constraints of helical structure. The present research is aimed to examine what type of helical model for tropocollagen is the most stable in terms of energetics by calculating the conformational energies under conditions with the helical constraints.

Method and Calculation

Three Helical Models for Tropocollagen. The polyproline II[11] type helix, the component of the triple-stranded coiled-coil structure of tropocollagen, has three amino acid residues in one turn (Fig. 1a), and it can be described as a 3/1 helix, since a chemical repeating unit contains one amino acid residue. If the helix axis of 3/1 helix further coils around the central axis, a coiled-coil structure is constructed (Fig. 1b). Here, the 3/1 helix is called the minor helix and the larger one with a longer period is called the major helix. If we consider the one turn of the minor helix (*i.e.*, three amino acid residues) as one repeating unit, the tropocollagen structure can be simply described as composed

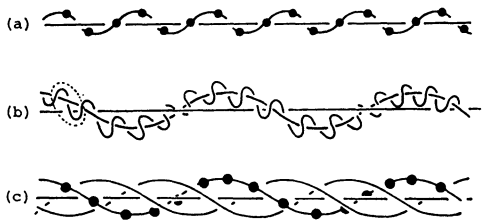

Fig. 1. Schematic representation of a triple-stranded helix of tropocollagen. (a) Simple 3/1 helical structure which contains three repeating units (filled circles) per turn. (b) Coiled-coil structure which is constructed by a further coiling of helix axis of (a). (c) Triple-stranded helix of tropocollagen. The filled circle represents one repeating unit, that is, one turn of the 3/1 helix which is shown by the dotted ellipse in (b).

Table 1. Three helical models for tropocollagen

	Amino acid sequence $(Gly_1–R_2–R_3)_n$	Types of hydrogen bond	Helical parameters h	θ
7/1 helix model				
$(Pro–Pro–Gly)_{10}$[a),7)]	R_2 Pro	$N(Gly_1)\cdots O(R_2)$	8.61 Å	51.4°
	R_3 Pro	or $N(Gly_1)\cdots O(R_3)$		
10/1 helix model				
Collagen I[b),3)]	R_2 any residue	$N(Gly_1)\cdots O(Gly_1)$	8.58 Å	36.0°
	R_3 Gly only			
Collagen II[b),3)]	R_2 any residue	$N(Gly_1)\cdots O(R_2)$	8.58 Å	36.0°
	R_3 any residue			
$(Pro–Gly–Pro)_n$[c),4)]	R_2 Pro	$N(Gly_1)\cdots O(R_2)$	8.61 Å	36.0°
	R_3 Pro			
12/1 helix model				
"standard"	R_2 imino acid impossible	$N(Gly_1)\cdots O(R_3)$	8.73 Å	30.0°
structure[b),6)]	R_3 any residue	$N(R_2)\cdots O(R_2)$		
"one-bonded"	R_2 Pro	$N(Gly_1)\cdots O(R_2)$	8.73 Å	30.0°
structure[b),6)]	R_3 Hypro			

a) The structure is based on three dimensional X-ray diffraction patterns of single crystal of synthetic polypeptide $(Pro–Pro–Gly)_{10}$ whose distribution of molecular weight is monodispersed. b) These structures are based on X-ray fiber diagram of native collagen. c) The structure is based on X-ray fiber diagram of synthetic polypeptide $(Pro–Gly–Pro)_n$ whose molecular weight is about 15000. Throughout this paper 1Å= 0.1 nm, 1 cal=4.184 J.

of three identical helices with a common central axis (Fig. 1c). The helical models proposed so far can be divided into three types according to the number of the repeating units in the major helix (Table 1). The first type includes the structure of Collagen I and II proposed by Rich and Crick[3] and of $(Pro–Gly–Pro)_n$ proposed by Yonath and Traub,[4] which may be represented by a 10/1 helix model since they have 10 repeating units per one turn. The second type of models are Ramachandran's "standard" (or two-bonded) and one-bonded structures,[6] which may be represented by a 12/1 helix model. The third is our model[7] which may be described as a 7/1 helix model. In the case of last model, two possible types are proposed as has been described previously.[10] One type has the $N(Gly_1)\cdots O(Pro_3)$ type hydrogen bond (Model 1) and the other has the $N(Gly_1)\cdots O(Pro_2)$ type hydrogen bond (Model 2). The tropocollagen models are characterized by two helical parameters, h and θ. Here, h means the unit height, a vertical distance of one repeating unit along the helix axis. θ means the unit twist, a rotational angle of the unit around the helix axis. The h and θ values for the present models are listed in Table 1.

In this study, we assumed that the tropocollagen was constructed of three identical helices of poly(Gly–Pro–Pro) for the computational convenience, as had been adopted in energy calculations by Tumanyan[8] and by Miller and Scheraga.[9] Almost all conformations of poly(Gly–Pro–Pro) have been found to be those of a collagen-like triple-helix, since imino acid residues at the positions of R_2 and/or R_3 promote the construction of the triple-stranded helix. In fact, the single crystals of $(Pro–Pro–Gly)_{10}$ molecule were found to exhibit a 7/1 helix.[10] The fiber of $(Pro–Gly–Pro)_n$ gave an X-ray fiber pattern very similar to the native collagen and it was once proposed that the polypeptide chain formed a 10/1 helix.[4] These facts show that the above assumption may be reasonable.

Molecular Geometry. The backbone conformation in the three-residue repeating unit is defined not only by the coordinates of corresponding atoms, but also by bond lengths, bond angles, and dihedral angles of the backbone chain. Since bond lengths and bond angles are not so different among the homologous compounds, we can fix these at the known values of low molecular weight homologues. The latter method is advantageous for constructing conformations of fibrous large molecules. In the case of poly(Gly–Pro–Pro), there are 9 dihedral angles in a repeating unit. These are shown in Figure 2. The dihedral angles ϕ_2 and ϕ_3

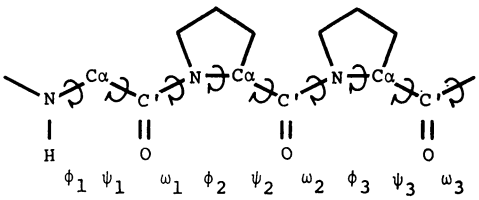

Fig. 2. Structural formula and the notation of dihedral angles in one repeating unit, Gly–Pro–Pro.

are restricted within narrow limits by pyroridine rings and they were fixed at −75° in the calculations. Further, if we confine the conformation to one of the above models, the two dihedral angles (in these calculations, ϕ_1 and ψ_1) will become dependent upon others. Then, the variable parameters were reduced to only five dihedral angles, ω_1, ψ_2, ω_2, ϕ_3, and ω_3. The fixed values of bond lengths and bond angles, and the initial coordinates of pendant atoms (C_β, C_γ, and C_δ atoms in proline residues and all oxygen and hydrogen atoms) were according to Momany *et al.*[12] The definition of dihedral angles followed the recommendation of IUPAC IUB Commission.[13] Five dihedral angles were

changed systematically and the coordinates of all atoms in the unit were represented in the right-handed Cartesian coordinate system.[14] In this system, the origin coincided with a given atom of the back bone chain. One of the neighbouring atoms of the origin lay on the x-axis, and the other axes were chosen arbitrarily. Then the coordinate system was converted into the right-handed Cartesian coordinate system having the z-axis coinciding with the helix axis by the method of Sugeta and Miyazawa.[15] When the coordinates of the jth atom in the nth unit of the standard ($m=0$) chain are expressed in terms of $(x_j^{n,0}, y_j^{n,0}, z_j^{n,0})$, those of the corresponding atom in the next ($(n+1)$th) unit of the same chain $(x_j^{n+1,0}, y_j^{n+1,0}, z_j^{n+1,0})$ may be obtained by the following equations.

$$\boldsymbol{X}_j^{n+i,0} = \boldsymbol{R}_1\boldsymbol{X}_j^{n,0} + i\boldsymbol{T},$$

where

$$\boldsymbol{X}_j^{n,0} = \begin{pmatrix} x_j^{n,0} \\ y_j^{n,0} \\ z_j^{n,0} \end{pmatrix}, \quad R_1 = \begin{pmatrix} \cos(i\theta) & -\sin(i\theta) & 0 \\ \sin(i\theta) & \cos(i\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and

$$\boldsymbol{T} = \begin{pmatrix} 0 \\ 0 \\ h \end{pmatrix}.$$

Similarly, the coordinates of the jth atom in the nth unit of the neighbouring chains $(x_j^{n,m}, y_j^{n,m}, z_j^{n,m})$ are obtained by the following equations.

$$\boldsymbol{X}_j^{n,m} = \boldsymbol{R}_2\boldsymbol{X}_j^{n,0} + \frac{m}{3}\boldsymbol{T},$$

where

$$\boldsymbol{X}_j^{n,m} = \begin{pmatrix} x_j^{n,m} \\ y_j^{n,m} \\ z_j^{n,m} \end{pmatrix}, \quad R_2 = \begin{pmatrix} \cos(m\theta') & \sin(m\theta') & 0 \\ -\sin(m\theta') & \cos(m\theta') & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and $\theta' = 120° - \theta/3$.[10] m is the chain number (0, 1, or 2). When looking down the helical axis from the carbonyl- to the amino-terminus, the value m of the right-side chain is 2 and that of the left-side chain is 1 with respect to the standard chain (0).

Conformational Energies. Electrostatic energy (E_{es}), non-bonded energy (E_{nb}), and torsional energy (E_{ts}) were considered. The total conformational energy for one repeating unit (E_{tot}) was the sum of these energies. The hydrogen bond energy was included in non-bonded energy term. The energy functions and parameter values used were the same as those of Momany *et al.*[12] The torsional energy was calculated only about three peptide bonds in the unit. The same functions and values were used for both intra- and inter-chain energy calculations. The intramolecular energy for a unit was calculated for the region from this unit to the third one of the carbonylterminal side, since the interaction energy between a given unit and any of its three amino-terminal side neighbours was the same as that between the same unit and the corresponding carbonyl-terminal side neighbour. The intermolecular energy was calculated between a central unit of A-chain ($m=0$) and the nearest four unit both up and down along the B-chain ($m=1$) from the central unit of A. Since three chains are identical in each other, the energy between A- and C-chains ($m=2$) equals to the energy between B- and A-chains. In the case of 7/1 helix model, the units which were taken into account of calculations are shown in the radial projection of Fig. 3. The open circle in this figure represents one repeating unit. The solid lines denote the peptide chain direction and a down side of this line points to the amino terminal side. The letters A, B, and C correspond to the peptide chain number in the triple-stranded helix, 0, 1, and 2, respectively. The conformational energies of isolated triple-stranded helix were calculated by summing up the energy in the central unit (filled circle) and the interaction energies between the central unit and all the independent unit (half-filled circles) within about 30 Å from the central unit.

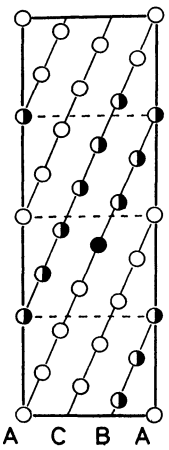

Fig. 3. The range of energy calculations of 7/1 helix model. The open circle denotes one repeating unit and the solid lines A, B, and C denote the peptide chains. The filled circle is the unit which conformational energy is calculated and the half-filled circle is the other unit responsible for the inter-unit conformational energy calculations.

It has been found that several water molecules per tripeptide unit are involved in the triple helices of tropocollagen[16] and model peptides of collagen[17] by way of hydrogen bonding. In addition to the stabilization by triple-helix itself, these hydrogen bonds may also contribute to the stabilization of the collagen-like triple-helix. However, Engel *et al.*[18] have shown that the peptides remain in a triple-helical conformation after dehydration, and concluded that water is not an essential component of forming the triple-helical structure. Therefore, in this study we did not take into account of water in order to simplify the problem, and calculated only the conformational energies of isolated triple-helix in a vaccum.

Computational Procedure. As described above, the

Table 2. Computational regions of dihedral angles

ω_1	175°—180°——175°
ϕ_2	−75° (fixed)
ψ_2	100°—180°——170°
ω_2	175°—180°——175°
ϕ_3	−75° (fixed)
ψ_3	100°—180°——170°
ω_3	175°—180°——175°

molecular conformation of poly(Gly–Pro–Pro) is defined by five dihedral angles (ω_1, ψ_2, ω_2, ψ_3, and ω_3) and helical parameters (h and θ). So the energy-minimum conformation of each model is included in the five

Table 4. Atomic distances used in the preliminary calculations

Atomic pair	Distance (Å)
C···C	2.50
C···N	2.59
C···O	2.30
N···N	2.67
N···O	2.38
O···O	2.09

Table 3. Number of conformations for which conformational energies are calculated

	Number of grid points[a)]	Number of conformations[b)]	Number of conformations after preliminary calculations
7/1 helix model	9747	4433	34 for Model 1 216 for Model 2
10/1 helix model	9747	4003	66
12/1 helix model	9747	2989	25

a) These were obtained by varing each dihedral angle at 5° intervals within its computational region. b) Number of conformations where the backbone conformation without pendant atoms can be constructed with each helical parameters.

Table 5. Low energy conformations at grid points of five-dimensional conformation space

No.	Dihedral angle/°							H-bond parameter		E_{tot}/[kcal/mol of (Gly–Pro–Pro)]
	ϕ_1[a)]	ψ_1[a)]	ω_1	ψ_2	ω_2	ψ_3	ω_3	r(N···O)/Å	∠HNO/°	
1) 7/1 helix model										
Model 1										
1	−59.9	168.6	180	170	−175	140	−175	4.41	32	−11.53
2	−58.7	165.2	180	175	−175	140	180	4.40	32	−9.99
3	−58.9	168.1	175	180	180	145	175	4.36	26	−3.13
4	−57.8	165.1	−175	175	−175	135	180	4.48	41	−1.51
Model 2										
1	−60.9	174.2	175	165	−175	145	−175	3.10	17	−25.76
2	−61.4	−177.7	180	170	175	140	180	3.45	21	−25.33
3	−62.7	−179.6	175	165	180	145	−175	3.35	22	−25.19
4	−61.9	−176.9	175	170	175	145	180	3.51	22	−25.07
5	−60.3	176.6	180	170	180	140	180	3.34	18	−25.02
6	−62.0	180.0	180	165	180	140	−175	3.22	23	−24.88
7	−60.5	−178.7	−175	170	175	135	180	3.28	22	−24.85
8	−61.1	177.4	175	170	180	145	180	3.45	17	−24.64
9	−64.5	−178.2	175	170	175	145	−175	3.49	23	−24.53
10	−64.1	−173.7	175	165	175	145	−175	3.48	27	−24.49
2) 10/1 helix model										
1	−51.1	162.0	−175	160	−175	125	−175	3.82	9	−17.88
2	−40.6	155.7	−175	165	−175	120	175	3.41	24	−17.61
3	−46.5	167.9	−175	165	175	120	180	4.02	14	−17.29
4	−49.2	157.1	−175	165	−175	125	180	4.11	18	−17.22
5	−51.0	160.6	180	160	−175	130	−175	4.25	29	−17.05
6	−43.2	160.7	180	160	−175	125	180	3.49	16	−16.98
7	−50.3	166.2	−175	155	−175	125	−175	4.14	45	−16.88
8	−42.1	160.7	−175	160	−175	120	180	3.18	35	−16.73
9	−44.5	166.8	−175	160	180	120	180	3.82	40	−16.69
10	−54.0	168.3	−175	160	180	125	−175	4.45	5	−16.63
3) 12/1 helix model										
1	−42.7	158.8	180	160	180	125	175	3.45	34	−18.69
2	−53.1	165.3	−175	155	180	125	−175	3.65	20	−18.57
3	−51.0	160.7	−175	160	180	125	180	3.90	11	−18.09
4	−48.4	158.3	180	155	−175	130	180	3.86	41	−17.78
5	−48.4	155.7	−175	165	180	125	175	4.05	14	−17.72
6	−45.3	163.6	175	155	180	130	180	3.56	32	−17.22
7	−44.5	164.9	180	160	175	125	175	4.16	47	−17.21
8	−51.5	160.2	180	160	180	130	180	4.44	17	−16.73
9	−53.7	164.8	180	155	180	130	−175	4.09	3	−16.64
10	−43.7	159.1	175	160	180	130	175	3.83	13	−16.49

a) ϕ_1 and ψ_1 were calculated numerically from other dihedral angles and helical parameters of the model.

dimensional space of conformation with the range shown in Table 2. All the dihedral angles of the peptide bond, ω_1, ω_2, and ω_3, were assumed to be *trans* conformation. Although it has been known that the peptide bond preceding proline residue is able to have a *cis* conformation (*e.g.*, the structures of poly-L-proline I[19] and of z–Gly–Pro–Leu[20]), these were taken to be *trans* in all the previous models of collagen. Then we considered only *trans* conformation in our calculations. The dihedral angles of C_α–C′ in the proline residue, ϕ_2 and ϕ_3, were assumed to lie in the ranges of 100°—180°——170° based on the structural analyses of oligopeptides[21,22] and the energy calculations of prolyl oligomers.[23]

At first the above conformational space was divided into 9747 grid points which were obtained by varing each dihedral angle at 5° intervals, and remaining dihedral angles (ϕ_1 and ψ_1) at each grid point were calculated numerically by using the helical parameters.[24] The conformations which gave the ϕ_1 and ψ_1 values outside the ranges of −180°——30° and 20°—180°——150°, respectively, were rejected in the following calculations. These criteria for glycyl residue are based on the regions of left-handed polyproline II structure in the energy map of diglycine.[25] In the resulting conformations, the dihedral angle ϕ_1 for three models was in the range of −130°——30°, and the angle ψ_1 for the 7/1, 10/1, and 12/1 helix models was in the range of 145°—180°——165°, 140°—180°——175°, and 140°—175°, respectively. The number of solutions for each model was listed in Table 3. The number of grid points which satisfied the above criteria was assumed to be about one half of solutions, since the most grid points had two solutions.

In order to reduce the time of computation, the number of conformations was further decreased by the preliminary energy calculations and the examinations of inter-chain hydrogen bonds. That is, the conformations were removed, if they exhibit the non-bonded distance between pairs of atoms in different chains which was shorter than the corresponding distance listed in Table 4. The values shown in Table 4 were the distances which gave the non-bonded energies 100 times as large as those at the minimum positions ($<r_g^{kk}>$ or $<r_g^{kl}>$) of Ref. 26, and calculated from the Lennard-Jones 6-12 expression. It has been proposed that in the case of poly(Gly–Pro–Pro), the 10/1 and 12/1 helix models and one of the 7/1 helix models (Model 2) have the $N(Gly_1)\cdots O(Pro_2)$ type hydrogen bond.[3,6,10] Model 1 of the 7/1 helix model has the $N(Gly_1)\cdots O(Pro_3)$.[10] For each of these models the conformations which have the hydrogen bond distance $r(N\cdots O)$ longer than 4.5 Å or the hydrogen bond angle $\angle HNO$ larger than 60° were removed because these conformations would not have hydrogen bonds by only the small change of variable dihedral angles. The number of remaining conformations for each model are given in the last column of Table 3. The conformational energy was calculated for all of these conformations. The dihedral angles, total potential energies, and hydrogen bond parameters $r(N\cdots O)$ and $\angle HNO$, of only 10 low energy conformations of each model are given in Table 5. The conformations of Model 1 of the 7/1 helix model had rather high potential energies than all the other types of models. Of the conformations obtained only four with negative energies are listed. All of the conformations shown in Table 5 which had the $N(Gly_1)\cdots O(Pro_2)$ type hydrogen bond were taken as the starting points for energy minimization procedure according to Fletcher and Powell.[27] The helical parameters for each model were held constant during these procedures. When the energy did not change as much as 0.0001 kcal/mol per tripeptide unit within an iteration of the Fletcher-Powell algorism, or when all of the variable dihedral angles did not change as much as 0.01° within an iteration, it was regarded that the energy-minimum in the conformational space has been reached.

Results and Discussion

Comparison of Models 1 and 2 of the 7/1 Helix Model. According to the type of hydrogen bonding, the 7/1 helix models were divided into Models 1 and 2 (see above). The number of conformations of Model 1 after preliminary calculations was very few as seen in the last column of Table 3, and their total conformational energies were very high compared with those of Model 2 (Table 5). In the four conformations of Model 1 shown in Table 5, the hydrogen bond distances of $N(Gly_1)\cdots O(Pro_3)$ are about 4.4 Å, while those of $N(Gly_1)\cdots O(Pro_2)$ are 2.96, 3.17, 3.61, and 2.86 Å, in an increasing order of the conformational energy. It seems, therefore, that the hydrogen bond of $N(Gly_1)\cdots O(Pro_2)$ is stronger than that of $N(Gly_1)\cdots O(Pro_3)$. Further investigation showed that the 23 conformations out of 34 of Table 3 had the hydrogen bond of $N(Gly_1)\cdots O(Pro_2)$ with distance less than 4.5 Å. Therefore, they were also included in the 216 conformations of Model 2. The resulting conformation after minimization starting from No. 1 of Model 1 (Table 5) was in agreement with that of Model 2 (Table 6) with a deviation of each dihedral angle less than 1.5°. The above findings showed that the global minimum conformation in the conformational space of the 7/1 helix model was included in the region with the hydrogen bond of $N(Gly_1)\cdots O(Pro_2)$, and some conformations within this region had a rather short hydrogen bond of $N(Gly_1)\cdots O(Pro_3)$ and a high conformational energy compared with the global minimum one. Thus it was concluded that the structures of Model 1 were unstable and could not be present. Hereafter, the word "7/1 helix model" may indicate that the conformation has helical parameters of (8.61 Å, 51.4°) and a hydrogen bond of $N(Gly_1)\cdots O(Pro_2)$.

12/1 Helix Model. The 12/1 helix model, which was proposed by Ramachandran[6] as a modification of 10/1 helix model, is somewhat different from the 7/1 and 10/1 helix models. If the amino acid sequence of tropocollagen is represented by poly(Gly_1–R_2–R_3), the true repeating period in the 7/1 and 10/1 helix models is one-third the pitch length of the peptide chain, as we see in the radial projections (a) and (b) of Fig. 4. This period is 20.1 Å for the 7/1 helix model and 28.6 Å for the 10/1 helix model. In this figure the open circle denotes a tripeptide unit, Gly_1–R_2–R_3, and the solid

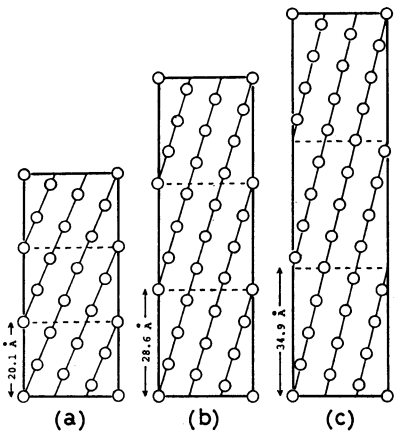

Fig. 4. Radial projections for three different kinds of tropocollagen models. The open circle denotes one repeating unit, Gly_1–R_2–R_3, and the solid line denotes the peptide chain direction. (a) 7/1 helix model, (b) 10/1 helix model, and (c) 12/1 helix model.

line denotes a direction of the peptide chain. Because of the finite molecular length of tropocollagen (about 3000 Å) and irregularities of occurrence of amino acids at the positions of R_2 and R_3 in the real sequence, these periods were usually called "pseudo-period."

On the other hand, the period obtained from X-ray fiber patterns of native collagen was taken to be 20.1 Å in the case of the 7/1 helix model or 28.6 Å in the case of 10/1 helix model. The above difference of the periods occurs from the ambiguity of indexing the layer lines which appear in the fiber patterns of native collagen. That is, if we take the meridional reflection of the spacing 2.86 Å as the 7th layer, the fiber period will be 20.1 (2.86×7) Å. And if we take that reflection as the 10th layer, the fiber period will be 28.6 (2.86×10) Å. In both 7/1 and 10/1 helix models, each pseudo-period of the models is naturally the same as that obtained from X-ray fiber patterns, because each model is constructed by using the periods obtained from the fiber pattern itself. In the case of 12/1 helix model, however, the one-third the pitch length is not a perfect repeating period as we see in Fig. 4(c), even though the amino acid sequence is represented by poly(Gly_1-R_2-R_3). Further, this value, 34.9 Å (2.91×12), does not agree with the period 29.1 Å obtained by Ramachandran.[6] Despite the apparent discrepancy, the conformational energy of the 12/1 helix model could be calculated in a similar manner as other models, since this did not affect the results obtained in the present study.

Comparison of Minimum-energy Conformations of the Three Models. The conformations with the minimum-energy for the three models are listed in Table 6 together with their energy terms and hydrogen bonding parameters. Tumanyan[8] carried out the minimization of conformational energies from many quadraplets of the dihedral angles (ϕ_1, ψ_1, ϕ_2, and ψ_3) as starting points without constraints of helical parameters, and concluded that the minimum-energy conformation was obtained with helical parameters within the ranges of $8.22\text{ Å}<h<8.76\text{ Å}$ and $25°<\theta<40°$. It was also suggested by

Table 6. Minimum-energy conformations of three helical models

	7/1 helix	10/1 helix	12/1 helix
Dihedral angle/°[a]			
ϕ_1[b]	−62.3	−48.3	−53.0
ψ_1[b]	179.4	158.2	164.3
ω_1	−179.6	−172.2	−174.6
ψ_2	167.0	164.1	155.0
ω_2	180.0	−175.3	−179.2
ψ_3	139.6	121.6	125.0
ω_3	−175.0	−178.4	−175.0
H-bond parameter[c]			
r(N···O)/Å	3.17	3.54	3.50
∠HNO/°	16	18	20
Energy/[kcal/mol of (Gly–Pro–Pro)][d]			
E_{nb}(intra)	−0.30	3.59	3.53
E_{nb}(inter)	−9.18	−6.63	−6.62
E_{es}(intra)	−15.06	−15.28	−15.26
E_{es}(inter)	−1.70	−1.03	−1.11
E_{ts}	0.15	0.52	0.33
E_{tot}	−26.09	−18.83	−19.12

a) $\phi_2=\phi_3=-75.0°$. b) ϕ_1 and ψ_1 were calculated numerically from other dihedral angles and helical parameters of each model. c) These parameters are for N(Gly_1)···O(Pro_2) type of hydrogen bond. d) E_{nb} refers to the non-bonded energy term, E_{es} to the electrostatic energy term, E_{ts} to the torsional energy term: intra in the parenthese to intrachain, inter to interchain. $E_{tot}=E_{nb}(\text{intra})+E_{nb}(\text{inter})+E_{es}(\text{intra})+E_{es}(\text{inter})+E_{ts}$

Ramachandran[6] that the conformations whose unit twist is 50° or more cannot be stable because of short contacts between atoms in adjacent chains and worse situations about hydrogen bonds. According to our calculations as shown in Table 6, however, the conformations of the 7/1 helix model, whose unit twist is 51.4°, can exist stably and further, the minimum conformational energy of this model was the lowest among the three helical models. The energy difference of 7 kcal/mol per tripeptide unit between the 7/1 and the other two helix models is attributable to the non-bonded energy terms. With respect to the hydrogen bond, the distance of N(Gly_1)···O(Pro_2) of the 7/1 helix model is 3.17 Å and those of 10/1 and 12/1 helix models are somewhat longer (3.54 and 3.50 Å, respectively). Further, as seen in Table 5, the conformations in considerably wide regions of conformational space of the 7/1 helix model tend to have lower energies and more stable hydrogen bonds in comparison with those of the 10/1 and 12/1 helix models. It seems that these preponderance of the 7/1 helix model results from its large unit twist compared with those of the 10/1 and 12/1 helix models. (Unit heights are almost the same among these models (Table 1).) That is, in the case of 7/1 helix model, the number of conformations (4433), where the backbone conformation with no pendant atoms can be constructed, is much larger than those of the other models (4003 for 10/1 and 2989 for 12/1 helix model). And the last column shows the resulting numbers of conformation after preliminary calculations,

in which the conformations were removed if they have very short contacts between atoms or a hydrogen bonding distance longer than 4.5 Å. The differences between the numbers of conformation for the three models clearly showed that the conformational space of the 7/1 helix model at this stage is increasingly wider than those of the others, although the absolute space is narrower than that of the previous stage.

The difference of non-bonded energies between the 7/1 and the other two (10/1 and 12/1) helix models was the sum of the differences of inter- and intra-chain non-bonded energies, 2.5 and 4.0 kcal/mol per tripeptide unit, respectively (Table 6). As there were no short contact between atoms of different chains in each model structure, the stabilization of interchain non-bonded energy of the 7/1 helix model showed that the chain molecule with this conformation was suited to be packed into a collagen-like triple-stranded structure. On the other hand, the difference between intrachain non-bonded energies resulted from short contacts between atoms in the same chain. For example; for the 7/1 helix model, $H_\alpha(Pro_2)\cdots H_\delta(Pro_3)$, 2.09 Å; for the 10/1 helix model, $H_\alpha(Pro_2)\cdots H_\delta(Pro_3)$, 2.06; $O(Pro_3)\cdots C'(Gly_1)$, 2.74; $H_\alpha(Gly_1)\cdots C_\delta(Pro_2)$, 2.56; $H_\alpha(Gly_1)\cdots H_\delta(Pro_2)$, 2.02 Å; and for the 12/1 helix model, $H_\alpha(Pro_2)\cdots C_\delta(Pro_3)$, 2.45; $H_\alpha(Pro_2)\cdots H_\delta(Pro_3)$, 1.98; $H_\alpha(Gly_1)\cdots H_\delta(Pro_2)$, 2.08 Å. The atomic coordinates of the minimum-energy conformation of the 7/1 helix model are shown in Table 7.

The molecular structures with the minimum-energy conformations for the three models are illustrated in Fig. 5, in which the chains are projected along the direction perpendicular to the helix axis. As already mentioned, the periods along the helix axis for the 7/1 and 10/1 helix models are 20.1 and 28.6 Å, respectively. The period of the 12/1 helix model, however, is neither 29.1 nor 34.9 Å. In this case the true period is the pitch length itself, 2.91×36 Å.

Table 7. Atomic coordinates of the minimum-energy conformation with 7/1 helix

Residue	Atom	x/Å	y/Å	z/Å
Gly	N	1.61	0.00	0.00
	C_α	1.60	1.28	0.69
	C′	2.48	1.24	1.94
	O	3.09	0.22	2.24
Pro	N	2.58	2.36	2.65
	C_α	3.39	2.46	3.85
	C_β	3.58	3.95	4.07
	C_γ	2.49	4.64	3.27
	C_δ	1.90	3.62	2.31
	C′	2.71	1.75	5.03
	O	1.54	1.39	4.95
Pro	N	3.50	1.55	6.11
	C_α	2.99	0.89	7.30
	C_β	4.23	0.45	8.07
	C_γ	5.38	1.29	7.52
	C_δ	4.92	1.93	6.23
	C′	2.09	1.82	8.10
	O	2.37	3.01	8.22

The results obtained from above calculations showed that the minimum-energy of the 7/1 helix model is less than those of the other two models with 10/1 and 12/1 helices by about 7 kcal/mol per tripeptide unit. It is unknown that the present 7/1 helix model may represent the true conformation of the native tropocollagen.

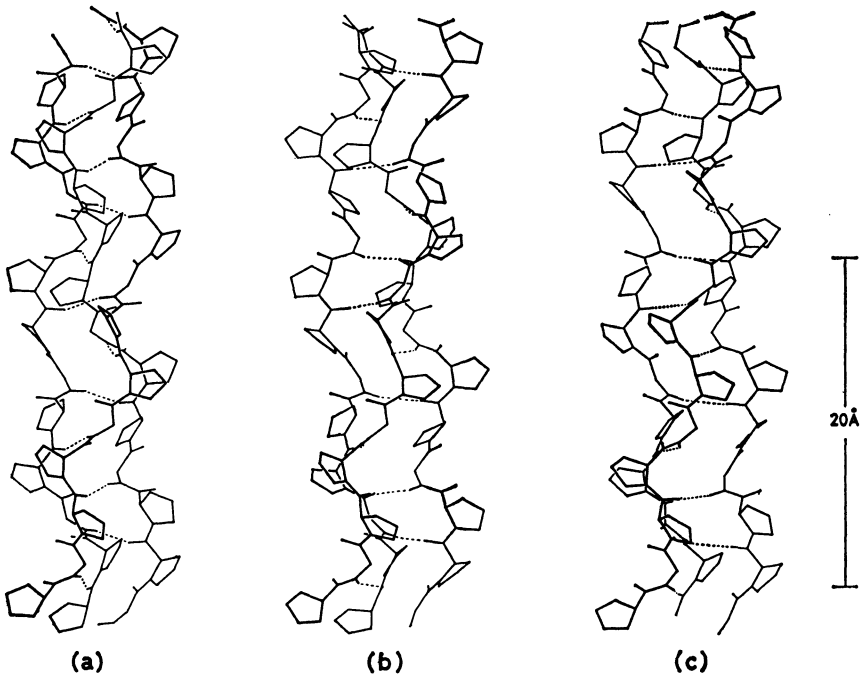

Fig. 5. Conformations with minimum-energy as projected along the direction perpendicular to helix axis. The dashed line denotes a hydrogen bond of $N(Gly_1)\cdots O(Pro_2)$. (a) 7/1 helix model, (b) 10/1 helix model, and (c) 12/1 helix model.

However, several lines of evidence suggest that this is the case. First, the synthesized molecule of (Pro–Pro–Gly)$_{10}$ as a collagen model can form a 7/1 helix in its single crystals.[10] Second, the hydrophobic interactions between tropocollagen molecules can form the stable microfibril structure only when the tropocollagen has a 7/1 helix model.[28] Third, the X-ray fiber diagram obtained from native collagen can be interpreted in terms of not only a 10/1 helix model but also a 7/1 helix model.[7] Taken these facts together with the present results, we may conclude that the 7/1 helix model is the best of the models so far presented for tropocollagen.

References

1) R. D. B. Fraser and T. P. MacRae, "Conformation in Fibrous Proteins," Academic, New York and London (1973), pp. 344—402.
2) G. N. Ramachandran and G. Kartha, *Nature*, **176**, 593 (1955).
3) A. Rich and F. H. C. Crick, *J. Mol. Biol.*, **3**, 483 (1961).
4) A. Yonath and W. Traub, *J. Mol. Biol.*, **43**, 461 (1969).
5) R. D. B. Fraser, T. P. MacRae, and E. Suzuki, *J. Mol. Biol.*, **129**, 463 (1979).
6) "Treatise on Collagen," ed by G. N. Ramachadran, Academic, New York (1967), Vol. 1, pp. 103—183.
7) K. Okuyama, M. Takayanagi, T. Ashida, and M. Kakudo, *Polymer J.*, **9**, 341 (1977).
8) V. G. Tumanyan, *Biopolymers*, **9**, 955 (1970).
9) M. H. Miller and H. A. Scheraga, *J. Polym. Sci., Symposium*, **54**, 171 (1976).
10) K. Okuyama, N. Tanaka, T. Ashida, and M. Kakudo, *Bull. Chem. Soc. Jpn.*, **49**, 1805 (1976).
11) V. Sasisekharan, *Acta. Crystallogr.*, **12**, 897 (1959.)
12) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
13) IUPAC IUB Commision on Biochemical Nomenclature, *J. Mol. Biol.*, **52**, 1 (1970).
14) S. Arnott and A. J. Wonacott, *Polymer*, **7**, 157 (1966).
15) H. Sugeta and T. Miyazawa, *Biopolymers*, **5**, 673 (1967).
16) R. S. Bear, *Adv. Protein Chem.*, **7**, 69 (1952).
17) S. Sakakibara, K. Inouye, K. Shudo, Y. Kishida, Y. Kobayashi, and D. J. Prockop, *Biochem. Biophys. Acta.* **303**, 198 (1973).
18) J. Engel, H. Chen, D. J. Prockop, and H. Klump, *Biopolymers*, **16**, 601 (1977).
19) V. Sasisekharan, *J. Polym. Sci.*, **47**, 373 (1960).
20) T. Yamane, T. Ashida, K. Schimonishi, M. Kakudo, and Y. Sasada, *Acta Crystallogr., Sect. B*, **32**, 2071 (1976).
21) T. Ashida and M. Kakudo, *Bull. Chem. Soc. Jpn.*, **47**, 1129 (1974).
22) R. Balasubramanian, A. V. Lakshminarayanan, M. N. Sabesan, G. Tegoni, K. Venkatatesan, and G. N. Ramachandran, *Int. J. Protein Res.*, **3**, 25 (1971).
23) A. J. Hopfinger and A. G. Walton, *J. Macromol. Sci. Phys.*, **B3**, 171 (1969).
24) N. Go and K. Okuyama, *Macromolecules*, **9**, 867 (1976).
25) A. J. Hopfinger, Ph. D. Thesis, Case Western Reserve University, Cleveland, Ohio, 1970.
26) F. A. Momany, L. M. Carruthers, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, **78**, 1595 (1974).
27) R. Fletcher and M. J. D. Powell, *Computer J.*, **6**, 163 (1963).
28) K. Okuyama, N. Go, and M. Takayanagi, *Chem. Lett.*, **1978**, 509.