(13) W. T. Astbury and H. J. Woods, *Philos. Trans. R. Soc. London, Ser. A.*, **232**, 333 (1933).
(14) M. L. Huggins, *J. Org. Chem.*, 1, 407 (1936).
(15) L. Pauling and R. B. Corey, *Proc. Natl. Acad. Sci. U.S.A.*, **37**, 251, 729 (1951); **38**, 86 (1952); **39**, 247 (1953).
(16) R. D. B. Fraser, T. P. MacRae, D. A. D. Parry, and E. Suzuki, *Polymer*, **10**, 810 (1969).
(17) R. D. B. Fraser and T. P. MacRae, "Conformation in Fibrous Proteins and Related Synthetic Polypeptides", Academic Press, New York and London, 1973.
(18) M. L. Huggins, *Proc. Int. Symp. Macromol.*, 289 (1977).
(19) M. L. Huggins, *Colloid Polym. Sci.*, in press.
(20) M. L. Huggins, *Chem.-Ztg.*, **101**, 285 (1977).
(21) R. D. B. Fraser and T. P. MacRae, *J. Mol. Biol.*, **5**, 457 (1962).

Calculation of the Structures of Collagen Models. Role of Interchain Interactions in Determining the Triple-Helical Coiled-Coil Conformation. 2. Poly(glycyl-prolyl-hydroxyprolyl)[1a]

Martha H. Miller, George Némethy, and Harold A. Scheraga*[1b]

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853.
Received December 3, 1979

ABSTRACT: The conformational space of regularly repeating structures of poly(glycyl-prolyl-hydroxyprolyl), $(GPH)_n$, was examined for stable triple-stranded complexes. The three strands were assumed to be equivalent. The structures generated included (a) coiled coils with either screw or rotational symmetry and (b) parallel-chain complexes with either screw or rotational symmetry. The dihedral angles for rotation about the single bonds of the three residues in the GPH unit were treated as the independent variables which repeated in each tripeptide unit. The interchain distance also was an independent variable in structures with rotational symmetry and in parallel-chain structures, with interchain orientation being an additional variable in the latter. Some coiled-coil complexes with screw symmetry were found to have much lower energies than the other structures. Many of the low-energy structures, including the most stable triple-helical coiled-coil complex with screw symmetry, were identical with the analogous structures in $(GPP)_n$ computed in earlier work [M. H. Miller and H. A. Scheraga, *J. Polym. Sci., Polym. Symp.*, **54**, 171 (1976)]. The most stable triple-stranded structure of the two molecules is identical with that which had been proposed for collagen. The hydroxyl group of hydroxyproline does not form any hydrogen bonds in this structure, and its nonbonded and electrostatic interactions with the rest of the molecule are very weak. They do not contribute to the stability of the collagen-like triple helix. Therefore, the observed increased stability of this type of triple helix with the sequence $(GPH)_n$, compared to that with the sequence $(GPP)_n$, must be attributed to interactions with the solvent. There is no difference in the interactions involving the hydroxyl group when hydroxyproline is placed in either position X or Y in the chain. Therefore, the preference of hydroxyproline for position Y in the GXY repeating units in collagen is not due to energetic factors within the triple-stranded structure.

I. Introduction

This paper is part of a series of studies on the conformational properties of collagen-like regular-sequence poly(tripeptide)s.[2-4] In the first paper of the series,[2] dealing with poly(glycyl-prolyl-prolyl), $(GPP)_n$, the complete conformational space was examined for both the single-chain polymer and the triple-stranded complex. Coiled coils with screw symmetry and parallel-chain complexes with either screw symmetry or rotational symmetry were considered. A coiled-coil triple-stranded complex has the lowest potential energy. The dihedral angles of its repeating tripeptide unit are (ϕ_{Gly}, ψ_{Gly}, ω_{Gly}, ϕ_{Pro}, ψ_{Pro}, ω_{Pro}, ϕ_{Pro}, ψ_{Pro}, ω_{Pro}) = (-74°, 170°, 180°, -75°, 168°, 180°, -75°, 153°, 180°). This conformation was used as the reference state in this paper. It is the only minimum-energy structure found which is similar to the models which were proposed for collagen on the basis of X-ray diffraction measurements of fibers.[5-8] The helical parameters (to be defined below) of this structure agree within 22% or better with those of proposed models of collagen,[5-7] with those obtained from X-ray crystallographic measurements of a $(Pro\text{-}Pro\text{-}Gly)_{10}$ crystal,[9] and with those obtained in a recent X-ray diffraction study of stretched tendon collagen.[8]

In the present study, a similar analysis is carried out for poly(glycyl-prolyl-hydroxyprolyl), $(GPH)_n$. The repeating unit of this copolymer is an important structural element of collagen. In the known sequence of the triple-helical part of the α1(I) chain of mammalian skin collagen,[10,11] consisting of 1014 residues (338 tripeptides), the Gly-Pro-Hyp sequence occurs 39 times and is the most frequently occurring Gly-X-Y tripeptide sequence. Considering the two imino acid residues individually, Pro is the residue found most frequently in position X, and Hyp is the most frequent in position Y. Also, the γ-hydroxyprolyl residue is found only in position Y. The enzymatic hydroxylation of prolyl residues in position Y takes place after the synthesis of the polypeptide chain.[12]

Poly(Gly-Pro-Hyp) is among the poly(tripepetide)s which have been studied experimentally as models of collagen.[13-15] Its X-ray fiber diffraction pattern is very similar to that of collagen.[15] Other physical properties, such as optical rotatory parameters, are also very close to those seen for collagen.[16] The triple helix formed by $(GPH)_n$, however, is more stable than that formed by $(GPP)_n$; i.e., the melting transition, corresponding to a triple-helix to random-coil transition, occurs at a higher temperature. The transition temperature is T_m = 297 K for $(Pro\text{-}Pro\text{-}Gly)_{10}$, but T_m = 331 K for $(Pro\text{-}Hyp\text{-}Gly)_{10}$ in aqueous acetic acid.[17-19] An increase of 18 K of T_m is observed in a 1,2-propanediol–acetic acid mixture.[19-21]

A similar effect is seen in natural collagen. The denaturation temperature of chick tendon procollagen before hydroxylation is 297 K. It increases to 311 K upon complete hydroxylation,[22] with no change in the sharpness of the transition curve.[23] Similarly, there is an 8 K difference in T_m of the 36-residue proteolytic fragments α1CB2 obtained from rat skin collagen and rat-tail tendon collagen.[24] The degree of hydroxylation and T_m are higher in the

0024-9297/80/2213-0470$01.00/0

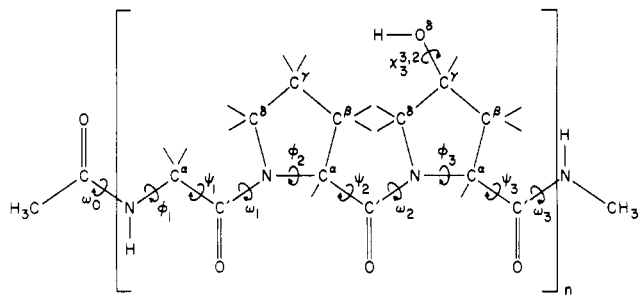

Figure 1. The structure of $(GPH)_n$, indicating the dihedral angles. The end groups used in the computation are shown.

former peptide. The hydrodynamic properties (and, therefore, presumably the entropies) of randomly coiled poly(proline) and poly(hydroxyproline) are very similar,[25] so that differences in T_m must be attributed to properties of the triple helical state. Clearly, the γ-hydroxyl group of Hyp in position Y plays an important role in the stabilization of collagen and of collagen-like poly(tripeptide)s. Therefore, it is of great interest to compare the relative stabilities of various triple-stranded conformations of $(GPH)_n$ with the relative stabilities of corresponding conformations of $(GPP)_n$, in order to see whether hydroxylation in the Y position alters the intrinsic stability of the triple helix.

It has been shown that the hydroxyl group cannot form a direct hydrogen bond to any backbone atom in a collagen-like triple helix which contains only trans peptide bonds.[5,26] It was suggested that a water molecule, hydrogen bonded to the hydroxyl group, could form a bridge to backbone carbonyl groups.[26-28] The difference in thermal stability of the $(GPP)_n$ and $(GPH)_n$ triple helices persists in 1,2-propanediol solutions in the absence of water.[19] If hydrogen bonding involving the solvent is the source of increased thermal stability, then the OH groups of the alcohol presumably contribute hydrogen-bonded bridges in this case.

The calculations in this paper were carried out in the absence of solvent. Therefore, they can be used only to compare the potential energy of $(GPP)_n$ and of $(GPH)_n$ within the triple-stranded complexes. A study of the effects of hydration is to be carried out later, making use of a model for hydration of peptides which was developed recently in our laboratory.[29,30]

In addition to the collagen-like structure, another triple-helical model has been proposed for $(GPH)_n$ with the Gly-Pro peptide bond in the cis conformation.[31,32] It was suggested that such a structure could be stabilized because it contains an OH···O side chain–backbone interchain hydrogen bond.[32] This structure will be compared here with other triple-helical arrangements.

The study reported in this paper was carried out in a manner analogous to that reported earlier[2] for $(GPP)_n$. The conformational space was explored for coiled-coil and parallel-chain triple-stranded complexes. Minimum-energy conformations were computed and compared with those obtained for $(GPP)_n$.

II. Computational Procedures

In general, the same assumptions and computational procedures were used as in the first paper of the series,[2] except as noted otherwise. The recommended standard nomenclature was used for peptide conformations.[33] The repeating unit GXY is shown in Figure 1. In this paper, X = Pro and Y = Hyp.

A. Selection of Molecular Geometry. Rigid geometry was assumed, as before.[2] The values of the bond lengths and bond angels correspond to the residue geometries selected by Momany et al.,[34] based on an extensive survey of X-ray and neutron diffraction data.[35-37] Following the first paper,[2] the puckering conformation of the prolyl ring was chosen which is designated as "down",[34,38] with the fixed dihedral angles $\phi = -75.0°$, $\chi^1 = +18.67°$. This conformation was shown to be 0.202 kcal/mol lower in energy[34,38] for *N*-acetyl-*N'*-methylprolinamide than the puckering designated as "up" ($\phi = -67.6°$, $\chi^1 = -6.11°$). In order to test this choice, the potential energy of the lowest-energy triple-stranded conformation of $(GPP)_4$, found in the previous work, was also computed here for the "up" puckering. After minimization, starting from the reference conformation mentioned in section I, the energy remained higher by 1.98 kcal/mol of (GPP) units. Of this energy difference, only 0.40 kcal/mol of (GPP) units is accounted for by the changed interactions within the two prolyl residues. The remaining 1.58 kcal/mol correspond to less favorable intra- and interstrand interactions. About 70% of the increase occurs in the intrastrand energy, and it is due mainly to unfavorable contacts of the $C^\delta H_2$ group of the Pro residues with the H^α of the preceding Gly or Pro residue, respectively. The interstrand energy also increases, primarily because of unfavorable interactions of the glycyl $C^\alpha H_2$ group with carbonyl O atoms of a neighboring strand. This result justified the use of the "down" puckering[39-41] for the most stable conformations of $(GPP)_n$ [and also $(GPH)_n$]. This ring conformation also is denoted[40] as endo.

The only variable dihedral angles during the minimization were ϕ and ψ of glycine, ψ of proline, and ψ and $\chi^{3,2}$ of hydroxyproline, in analogy with earlier work[2] (Figure 1). Both trans and cis conformations for the two peptide bonds preceding Pro or Hyp were considered, but the value of ω was fixed at 180° or 0° for these peptide bonds and at 180° for the peptide bond preceding Gly. As a test of the validity of fixing ω in the computations, the minimization was repeated for the lowest-energy conformation of $(GPP)_n$ with all backbone dihedral angles (including the ω's) treated as variables. Neither the dihedral angles nor the energy changed as a result of this minimization, showing that, at least for the most stable conformation found, ω can be considered as fixed.

It will be shown below that the lowest-energy conformation of $(GPH)_n$ is identical with that of $(GPP)_n$. Thus, the two assumptions, viz., the use of the "down" conformation for the prolyl ring and fixing the values of the ω's, are equally valid for both polypeptides.

B. The Geometry of Triple-Stranded Complexes. Throughout this work, it is assumed that the GXY units along each strand and in the different strands are equivalent. This means that whenever the conformation of a GXY unit is specified, by assigning the values of its dihedral angles, this specification determines the conformation of all three strands.

The parameters characterizing the geometry of coiled-coil structures and those describing the packing arrangements in various triple-stranded complexes are recapitulated here.[2,5,42]

The conformation of a single polypeptide chain is specified by the dihedral angles of the GXY repeat unit. If the residues within one repeat unit do not all have the same backbone dihedral angles, a coiled coil is formed. (If the backbone dihedral angles of each residue are the same, a simple helical structure is formed, with major and minor helical axes coinciding.[42]) The parameters of the minor helix and its handedness depend on the dihedral angles.[42,43] For collagen II and related structures, the minor helix is

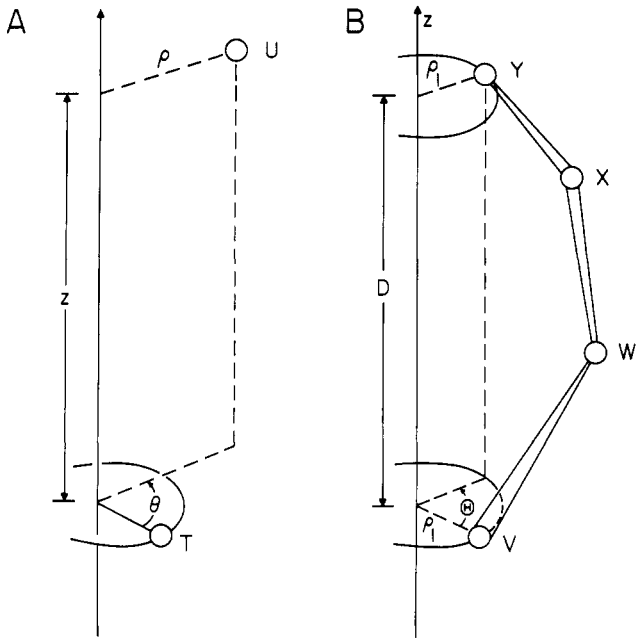

Figure 2. (A) Definition of a cylindrical coordinate system in terms of the variables (ρ, θ, z), where ρ is the distance of any atom U from the z axis, θ is the azimuthal angle of U measured with respect to the reference atom T, and z is the displacement of U along the z axis with respect to the reference atom T. (B) Definition of the helical parameters for a coiled coil in the cylindrical coordinate system shown in Figure 2A. Corresponding atoms (e.g., N atoms) V, W, X, Y of four consecutive residues are shown by circles. V and Y represent equivalent atoms in two successive tripeptide repeating units (e.g., glycyl N atoms). (An analogous diagram, showing all backbone atoms of a GPP tripeptide unit in the same coordinate system, is shown in Figure 1 of ref 2.) D is the translational repeat per tripeptide unit; Θ is the angular repeat per tripeptide unit. ρ_1 has the same value for both atoms V and Y, but ρ of V, W, and X generally are different in coiled coils. z is the axis of the major helix. The axis of the minor helix is defined by atoms V, W, X, Y; the chain makes a complete turn around the minor helical axis in three residues,[42] i.e., from V to Y. In a simple helix, ρ has the same value for all corresponding atoms V, W, X, Y.

left handed.[13] The polypeptide chain and the axis of the minor helix form a helix winding around the major helical axis, denoted as the z axis throughout this series of papers. The major helix in collagen II is right handed.[13] The coordinates of each atom of the polypeptide chain are given in a cylindrical coordinate system, based on the major helix, in terms of (ρ, θ, z), as shown in Figure 2A.

The coiled coil (i.e., the major helix) formed by the polypeptide strand may be described by the helical parameters[2,42,44] D and Θ, in terms of the repeat unit. D is the translational repeat per tripeptide unit (GXY) along the z axis, and Θ is the azimuthal angular repeat per tripeptide unit, measured around the z axis (Figure 2B; see also Figure 1 of ref 2). The handedness of the major axis, i.e., the sense in which the polypeptide chain is wound around the z axis, is given by the sign of the angular repeat: $\Theta > 0$ for a right-handed major helix, and $\Theta < 0$ for a left-handed major helix. $\Theta = 0$ corresponds to no azimuthal displacement from one repeat unit to the next one with respect to the z axis, i.e., to a polypeptide chain running parallel to the major axis, as in polyproline II, for example.[45] The values of D and Θ are functions of the dihedral angles, and they can be obtained from the dihedral angles by means of the procedure of Sugeta and Miyazawa.[44] It should be noted that the helical sense of the minor and of the major helix both are functions of the values of the dihedral angles, but they do not have to be the same, i.e., a right-handed major helix can be formed by both right- and left-handed minor helices and vice versa.

Three polypeptide chains, running in the same direction from the N to the C terminus, can be packed in six geometrically distinct ways to form a triple-stranded structure: (a) triple-stranded coiled coil with screw symmetry, (b) triple-stranded coiled coil with rotational symmetry, (c) three parallel chains with screw symmetry, (d) three parallel chains with rotational symmetry; in addition, two different dispositions of the three stands, viz., counterclockwise and clockwise (to be defined later), are possible for both packings with screw symmetry (Figure 3). Some of these may not exist because of unfavorable interstrand energies. In section IIB we are concerned only with the geometric characterization of various triple-stranded structures and not with their energetic feasibility. Their energies will be compared in section III.

In triple-stranded coiled coils, the major helical axes (z axes) of the three polypeptide strands coincide, so that all three chains wind around the common axis. In this case, the geometry is fully determined by D and Θ. The interstrand distance is not freely variable. It is fixed by the helical parameters.

In parallel-chain structures, the major helical axes of the three polypeptide strands do not coincide, but they are parallel to each other and to the axis of symmetry. The interchain distance (or equivalently, the distance of each major helical axis from the axis of symmetry) and the orientation of the chains about their major axes, with respect to the axis of symmetry, also are independent variables, in addition to D and Θ.

Another characteristic feature of triple-stranded structures with screw symmetry is the disposition of the three strands with respect to each other, as defined below. It is a function of D and Θ, but it also depends on the type of symmetry. The disposition of the three strands is determined by the screw symmetry operation. The z coordinates of equivalent atoms of chains 2 and 3 are displaced along the z axis by the distances $R_2 = D/3$ and $R_3 = 2D/3$, respectively, relative to the corresponding z coordinate in chain 1. The equivalent positions of the three strands are staggered azimuthally as well. Chains 2 and 3 are displaced with respect to chain 1 by the angles α_2 and $\alpha_3 = 2\alpha_2$, respectively, when viewed along the z axis (see Figure 3 of ref 2).

In coiled coils with screw symmetry, α_2 is a function of Θ. For any given value of Θ, there exist two geometrically possible choices of α, with opposite signs, viz., $\alpha_2 = -(2\pi - \Theta)/3$ and $\alpha_2 = (2\pi + \Theta)/3$. The sign of α_2 defines the relative disposition of the three strands. The case of $\alpha_2 > 0$ is designated in this series of papers as "clockwise disposition" of the three chains.[2] In this case, an imaginary curve laid through structurally equivalent atoms of the three strands describes a right-handed helix around the z axis, i.e., structurally equivalent residues in different chains are related by right-handed screw symmetry (Figure 3B). Conversely, when $\alpha_2 < 0$, the same curve forms a left-handed helix around the z axis. This case is designated here as "counterclockwise disposition" of the three chains (Figure 3A). The arrangement of strands in the collagen II triple-helical models proposed by Rich and Crick[5,6] as well as the models proposed by Ramachandran and Kartha[46] and Yonath and Traub[7] corresponds to the "counterclockwise disposition".[5,13] It must be emphasized that the terms "clockwise" and "counterclockwise disposition", depending on the sign of α_2, *do not* describe the helical sense of either the minor or the major helices. The sense of the major helix is determined by the sign of

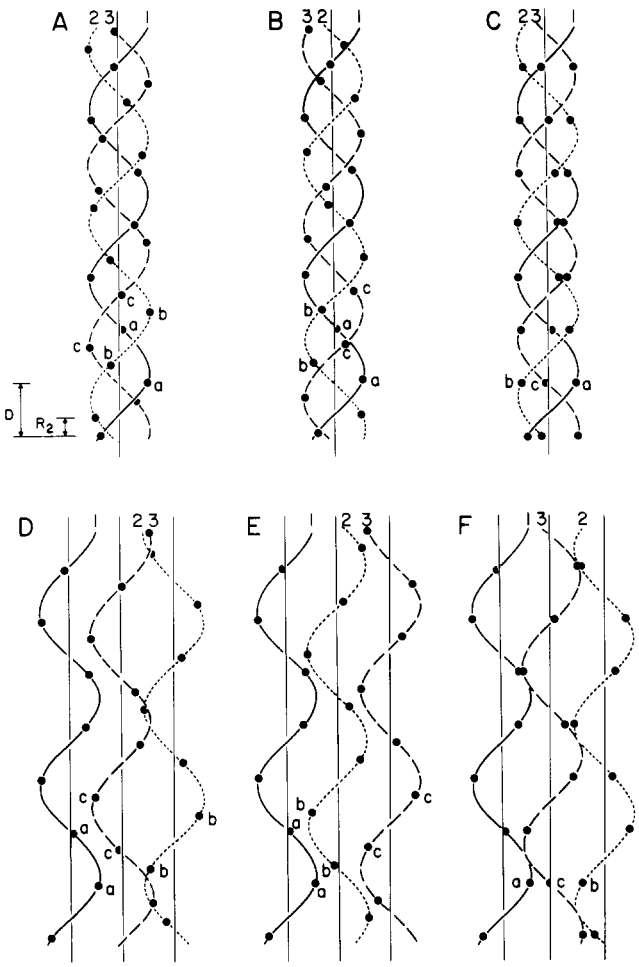

Figure 3. Schematic representation of triple-stranded helical complexes with various kinds of symmetry. Each strand is drawn as a right-handed major helix ($\Theta > 0°$). (A) Coiled coil with three-fold screw symmetry and counterclockwise disposition of the three strands. (B) Coiled coil with threefold screw symmetry and clockwise disposition of the three strands. (C) Coiled coil with threefold rotational symmetry. (D) Parallel-chain complex with threefold screw symmetry and counterclockwise disposition of the three strands. (E) Parallel-chain complex with threefold screw symmetry and clockwise disposition of the three strands. (F) Parallel-chain complex with threefold rotational symmetry. In each diagram, chain 1 is shown with a full line, chain 2 with a dotted line, and chain 3 with a dashed line. The major helical axis of each strand is shown as a thin line. In coiled-coil complexes (parts A–C), the axis of symmetry (z axis) and the major helical axes of all three strands coincide. The common axis of symmetry (z axis) of the parallel-chain complexes (parts D–F) is located directly behind the strand in the middle. Equivalent positions (e.g., points a, b, and c) on the three strands are shown by filled circles. A continuous curve laid through successive equivalent points on chains 1, 2, 3, 1, 2, 3, . . . (a → b → c → a → b → c . . .) describes a left-handed helix around the z axis in complexes with "counterclockwise" disposition of the strands (parts A and D). It describes a right-handed helix in complexes with "clockwise" disposition of the three strands (parts B and E). In complexes with rotational symmetry, equivalent positions on the three strands have the same value of z (parts C and F). The value of D is the same for all six structures shown. The value of R_2, shown in part A, is applicable to parts B, D, and E as well. $R_2 = 0$ in parts C and F.

Θ. In fact, both kinds of dispositions are possible for a given helical sense. The geometrical relations between the backbone dihedral angles of polypeptide chains and the parameters that characterize coiled-coil structures were derived in an earlier paper.[42] It was shown there that severe geometrical restrictions exist for the formation of coiled-coil structures corresponding to a given backbone conformation of the repeat unit.

In three-stranded parallel-chain structures with screw symmetry, $R_2 = D/3$, as described above, but now $\alpha_2 = +120°$ for "clockwise disposition" (Figure 3E) or $\alpha_2 = -120°$ for "counterclockwise disposition" (Figure 3D), and α_2 does not depend on Θ. In this case, the two angular parameters are referred to different axes, viz., Θ to the major axis of each individual polypeptide chain and α_2 to the common axis of symmetry of the triple-stranded complex. (These two axes coincide in coiled coils, as discussed above.) The disposition of the three strands is defined by the sign of α_2, as in coiled coils. It was stated erroneously in the first paper[2] that Θ = 0 for all parallel chains with screw symmetry. Actually, triple-stranded complexes with this symmetry can be generated for any value of Θ.

Both triple-stranded coiled coils and parallel-chain complexes with rotational symmetry are characterized by $R_2 = R_3 = 0$, i.e.,, equivalent atoms are located in a plane which is perpendicular to the z axis (Figure 3C,F). In both cases, $\alpha_2 = \pm 120°$ with respect to the axis of symmetry. The two signs of α_2 correspond to identical structures, and the designation of strands as chain 2 or chain 3 is arbitrary.

In summary, six different packings of three equivalent polypeptide chains can be generated from a given single-stranded coiled-coil conformation (characterized by a given valued of D and Θ). These are either a triple-stranded coiled-coil or a three-stranded parallel-chain structure. Both can occur either with rotational symmetry, or with screw symmetry and counterclockwise disposition of the strands, or with screw symmetry and clockwise disposition of the strands. The geometrical relationships between the parameters of the various packings are summarized in Table I. Some of these packings may not be feasible energetically, because of atomic overlaps.

C. Generation of Chain Conformations. The coordinates of one GXY unit were computed from the fixed residue geometry and from given dihedral angles, using a general procedure for the generation of polypeptide chains.[34] Because of the assumption of equivalence, mentioned at the beginning of section IIB, the procedure of Sugeta and Miyazawa,[44] as modified by McGuire et al.,[47] was then used to generate all three strands in coiled coils, as described in the previous paper.[2] The equations of Sugeta and Miyazawa, however, break down for Θ = 0. In this case, a modification of the equations was used, as derived by McGuire et al.[47] Following their procedure, eq 21 of ref 44 (or eq 13 of ref 47) was replaced by the following equation (eq 17 of ref 47) whenever $|\Theta| < 0.1°$:

$$\mathbf{E}_z = \frac{\mathbf{p}_i^{(i)} + \mathbf{p}_{i+1}^{(i)}}{|\mathbf{p}_i^{(i)} + \mathbf{p}_{i+1}^{(i)}|} \qquad (1)$$

where $\mathbf{E}_z$ is the unit vector in the z direction (see Figure 2), $\mathbf{p}_i^{(i)}$ is the vector from the C^α atom of residue $i-1$ to the C^α of residue i, and $\mathbf{p}_{i+1}^{(i)}$ is the vector from C_i^α to C_{i+1}^α. McGuire et al.[47] have shown that this equation approximates eq 21 of Sugeta and Miyazawa[44] for small values of Θ. They proposed that the use of eq 1 is a good approximation whenever ϕ and ψ are within 0.1° of those values which generate polypeptide chains with Θ = 0.

Parallel-chain complexes were generated by the procedure described earlier,[2] using eq 7, 8, 5a, and 6a of ref 2.

D. Conformational Energies and Minimization Procedure. The intramolecular and intermolecular interaction energies were computed with the energy functions and parameters developed in this laboratory.[34] At large distances, a "united-residue approximation"[48] was used, as in the first paper of this series. It was shown in the first paper[2] that the interactions in the infinite chain

Table I
Geometrical Description of Triple-Stranded Polypeptide Chain Complexes

structure	symmetry	disposition	R_2	α_2, deg	interchain separation and orientation
coiled coil	screw	counterclockwise	$D/3$	<0	fixed by D and Θ
		clockwise	$D/3$	>0	fixed by D and Θ
	rotational		0	±120	fixed by D and Θ
parallel chain	screw	counterclockwise	$D/3$	-120	independent variables
		clockwise	$D/3$	$+120$	independent variables
	rotational		0	±120	independent variables

of $(GPP)_n$ can be represented adequately with $n = 4$. Therefore, the calculations were carried out by using the molecule $CH_3CO(GPH)_4NHCH_3$. Function minimization was performed with the algorithm of Fletcher and Powell,[49] with the same parameters as those used in the first paper.[2]

E. Selection of Starting Conformations and Energy Minimization. Coiled-Coil Structures with Screw Symmetry. In order to compare the minimum-energy conformations of single-stranded $(GPP)_n$ and $(GPH)_n$, potential energy maps as functions of the variable dihedral angles were calculated for the terminally blocked dipeptides *N*-acetyl-*N'*-methyl-L-prolyl-L-hydroxyprolinamide and *N*-acetyl-*N'*-methyl-L-hydroxyprolylglycinamide. The corresponding Gly-Pro dipeptide is common to the two polypeptides, and its low-energy conformations were taken from ref 2. The maps were obtained by computing the energy of each conformation selected on a grid at 10° intervals of all variable backbone dihedral angles, i.e., ψ_1 and ψ_2 for the first dipeptide, and ψ_1, ϕ_2, and ψ_2 for the second dipeptide. ω_0 and ω_1 were kept fixed at either 0 or 180°, and ω_2 and $\chi^{3,2}$ were kept fixed at 180°. The low-energy conformations were identical in all important features with those obtained for the terminally blocked Pro-Pro and Pro-Gly dipeptides.[50] The dihedral angles of all energy minima of the corresponding Pro and Hyp dipeptides are within <1° of each other (see Table II of ref 2). Therefore, it could be assumed that the local energy minima which were found for triple-stranded coiled coil structures of $(GPP)_4$ with screw symmetry in the first paper (Tables V and VI of ref 2) were an adequate sample of starting points for the energy minimization of coiled-coil $(GPH)_4$ as well.

In coiled-coil structures with screw symmetry, the energy was minimized with respect to the five variable dihedral angles ϕ_1, ψ_1, ψ_2, ψ_3, and $\chi_3^{3,2}$. Initially, $\chi_3^{3,2} = 180°$ in all starting conformations. The other two rotamers (g^+ and g^-, with $\chi_3^{3,2} = 60$ and $-60°$, respectively) were not used as starting points. The ω_3 (preceding glycine) was fixed at 180°; ω_1 and ω_2 were fixed at either 180° or 0°.

Parallel-Chain Structures with Screw Symmetry. The only structures with this symmetry that were tested were those having a polypeptide conformation with $\Theta = 0$, both in the case of $(GPP)_n$ in the first paper[2] and here for $(GPH)_n$. (They are a special case of Figure 3D,E.) In the first paper, only one such conformation was generated. All of its dihedral angles were fixed at the values in the regular polyproline II helix,[45] i.e., $(\phi,\psi) = (-75°, 145°)$ was chosen for all residues. This helix was used to generate two three-strand complexes, each with one of the two possible dispositions of the chains. Chain separation and orientation were varied to obtain a minimum-energy structure with these dihedral angles.

A new procedure was introduced in this work, in order to provide a larger sampling of possible single-strand conformations which can be packed as parallel chains with screw symmetry and $\Theta = 0$. Eighty-one chains were generated as variants of the polyproline II helix. The dihedral angles were allowed to vary, in order to generate different chain conformations, subject to the constraint that $\Theta = 0$ for all of them. Each of the four variable backbone dihedral angles (ϕ_{Gly}, ψ_{Gly}, ψ_{Pro}, ψ_{Hyp}) was allowed to take one of three values as starting points: the one cited above for the polyproline II conformation and values differing from it by +10° or −10°. By taking all combinations of these three values of each dihedral angle, 3^4 starting conformations were obtained. Out of these starting conformations, 81 single-chain conformations were generated by varying the dihedral angles, subject to the constraint that $\Theta = 0$. Solutions to the equation $\Theta = 0$ can be obtained numerically by minimizing the quantity Θ^2 with respect to the four variable dihedral angles. All ω's and $\chi_3^{3,2}$ were fixed at 180°. It was necessary to use the geometrical optimization procedure as an initial step, because the systematic gridding procedure carried out for either single or triple chains, based on energy considerations, as described for coiled coils in the first paper,[2] did not yield any structures with $\Theta = 0$. The conformations obtained with this procedure satisfy the geometrical constraint ($\Theta = 0$), but they are not energy minima for the *single strands*. The next step of the procedure, described in the next paragraph, however, yields local energy minima for the packing of *triple-stranded* polyproline-II-like conformations.

The 81 single-stranded conformations obtained by the geometrical optimization were placed into triple strands with counterclockwise disposition of the chains. No packings with clockwise disposition were tested. Minimum energy packings were determined with respect to variation of interchain orientation and distance, with fixed dihedral angles; the latter condition was required to preserve the constraint $\Theta = 0$. No complete energy minimization was carried out with respect to these two variables, because it would have required excessive computer time. Instead, the energy was determined for discrete points on a grid of interchain separation and orientation, using 1.0 Å and 30° intervals, respectively, with 10 Å as the maximal separation. The grid point with lowest energy was selected as the final result for each chain conformation. The potential energy surface is very flat with respect to these two variables, i.e., the energy does not change much between adjacent grid points. Therefore, it is a good approximation to use the lowest-energy grid points instead of the actual minima of the surface.

Coiled-Coil and Parallel-Chain Structures with Rotational Symmetry. The 50 minimum-energy conformations found for $(GPP)_4$ (Table XI of ref 2) were used as starting points. These conformations included both coiled coil and parallel-chain triple-stranded complexes. The former are characterized by small chain separations (near 0.0 Å). Energy minimization was carried out with respect to the backbone dihedral angles (ϕ_1, ψ_1, ψ_2, and ψ_3) and the interchain separation as variables. All ω's were fixed as in Table XI of ref 2, and $\chi_3^{3,2}$ was fixed at 180°.

III. Results

A. Comparison of Triple-Stranded Structures. Energy minimization of the coiled-coil structures with

Table II. Dihedral angles and energies of triple-stranded coiled-coil structures of $(GPH)_4$, with screw symmetry and counterclockwise disposition of the chains

Dihedral angle (degrees)[a]							Energy [kcal/mol of $(GPH)_4$]
ϕ_1	ψ_1	ω_1	ψ_2	ω_2	ψ_3	$\chi_3^{3,2}$	
-79	170	180	167	180	156	-171	-100.0
168	-176	0	158	180	105	-171	-81.8
-159	95	180	168	180	168	-172	-80.8
126	-165	180	164	180	-53	-171	-76.9
-172	-166	180	147	180	170	-44	-76.5
-77	164	180	161	180	77	-171	-70.0
150	-112	0	169	180	79	-175	-66.3
128	169	0	166	180	117	-171	-65.9
-149	-101	180	161	0	175	-172	-61.7
-78	-169	0	166	180	-24	-170	-61.3
102	-174	0	162	180	120	-170	-59.9
96	91	180	168	180	-45	-170	-59.3
65	112	180	155	180	65	-176	-58.5
148	165	180	162	180	-65	-170	-57.0
157	-163	180	118	180	-55	179	-56.6
-79	-172	180	119	180	-178	-165	-56.3
-74	-139	0	169	180	-172	-167	-50.8
82	121	180	167	180	-52	-173	-48.5
-82	-176	180	122	180	88	-171	-48.5
119	-102	0	166	180	78	-169	-48.1
-48	-71	0	172	180	-169	-91	-37.7
-120	-83	0	169	180	-80	-168	-26.6
22	65	180	155	180	58	-169	-17.7

[a] Only ϕ_1, ψ_1, ψ_2, ψ_3 and $\chi_3^{3,2}$ were varied in the energy minimization; ω_1 and ω_2 were kept fixed at the values indicated; $\omega_3 = 180°$ for all conformations.

Table III. Dihedral angles and energies of triple-stranded coiled-coil structures of $(GPH)_4$, with screw symmetry and clockwise disposition of the chains

Dihedral angle (degrees)[a]							Energy [kcal/mol of $(GPH)_4$]
ϕ_1	ψ_1	ω_1	ψ_2	ω_2	ψ_3	$\chi_3^{3,2}$	
-161	175	180	140	180	167	-43	-81.3
158	75	180	158	180	-63	-175	-81.0
164	168	0	167	180	103	-171	-75.8
-81	-174	180	165	180	156	-170	-74.6
167	112	180	167	180	162	-172	-73.2
-85	161	180	161	180	75	-171	-73.1
-131	85	180	168	180	171	-171	-69.6
159	180	0	166	180	75	-170	-67.9
113	167	0	167	180	155	-171	-66.4
-58	140	180	162	180	-178	-175	-65.2
-143	-93	180	162	0	168	-172	-63.1
122	170	180	166	180	-41	-170	-61.1
-92	-90	180	158	0	143	-171	-55.1
129	-103	0	169	180	85	-92	-55.0
166	158	180	161	180	-55	-169	-54.9
-152	-167	0	164	180	-29	-171	-54.3
-64	177	180	121	180	-175	-173	-47.1
135	-113	0	162	180	82	-96	-43.6
64	115	180	151	180	60	-178	-42.7
60	-172	180	118	180	-70	-175	-40.1
-161	-107	0	169	180	-29	-77	-39.6
53	87	180	129	180	61	-165	-22.1

[a] Only ϕ_1, ψ_1, ψ_2, ψ_3 and $\chi_3^{3,2}$ were varied in the energy minimization; ω_1 and ω_2 were kept fixed at the values indicated; $\omega_3 = 180°$ for all conformations.

Table IV. Helical Parameters, Interchain Energy, and Intrachain Energy of Triple-stranded Coiled-coil Structures of $(GPH)_4$, with screw symmetry and counterclockwise disposition of the chains[a]

Helical Parameters			Energy[b] [kcal/mol of $(GPH)_4$]				
D	Θ[c]	ρ_1[d]	Interchain		Intrachain		Total
(Å)	(degrees)	(Å)	E_{ES}	E_{NB}	E_{ES}	E_{NB}	E_{tot}
9.0	43.6	1.6	-6.8	-35.9	-55.0	-4.4	-100.0
7.1	54.9	2.8	-2.5	-26.7	-52.0	-0.7	-81.8
9.0	-71.6	2.8	1.6	-30.6	-55.0	1.1	-80.8
6.5	49.5	3.1	5.5	-31.0	-50.1	-1.4	-76.9
4.0	-66.0	7.6	-2.2	-28.3	-55.0	8.6	-76.5
3.8	-64.0	6.6	-0.7	-14.6	-56.1	1.4	-70.0
4.6	45.3	7.5	-2.0	-28.6	-53.5	17.7	-66.3
4.1	59.3	6.1	-0.3	-18.2	-49.6	2.0	-65.9
6.8	-70.6	3.4	-5.1	-23.4	-47.5	14.1	-61.7
6.7	47.7	5.0	-0.6	-21.4	-40.4	1.0	-61.3
2.8	42.5	9.8	0.2	-10.8	-49.3	-0.3	-59.9
7.5	-63.4	4.4	1.9	-16.2	-46.2	1.1	-59.3
4.1	65.3	6.7	0.7	-30.8	-51.9	23.7	-58.5
8.9	50.7	4.9	0.9	-11.5	-49.5	2.9	-57.0
6.4	64.6	5.4	-1.5	-17.5	-53.0	18.4	-56.6
6.0	67.2	6.1	-1.4	-16.8	-51.5	16.1	-56.3
7.4	-72.6	3.0	-5.4	-35.4	-44.5	32.0	-50.8
8.5	-48.2	4.1	-6.0	-12.9	-44.2	14.5	-48.5
5.1	-60.8	6.5	-9.0	-9.0	-55.7	16.1	-48.5
1.5	-30.1	13.8	-0.2	-2.7	-56.1	10.8	-48.1
6.2	-53.2	5.4	3.0	-15.7	-44.2	17.9	-37.7
5.7	-51.2	6.0	4.1	-6.7	-42.0	17.8	-26.6
4.3	-65.0	6.4	1.6	-16.3	-54.6	51.4	-17.7

[a] The conformations are listed in the same order as in Table II.

[b] The subscript ES refers to the electrostatic energy term, NB refers to the nonbonded energy term. Both of these terms include the contributions to the hydrogen bond energy. In addition to the four energy terms shown here, E_{tot} also contains a torsional energy term E_{TOR}, contributed by the side-chain rotation about the C^{γ}–O^{δ} bond in Hyp, and a prolyl ring energy[34] which depends on whether $\omega = 0°$ or 180° preceding Pro or Hyp.

[c] The sign of Θ determines the handedness of the major helix. See the text.

[d] The distance of the glycyl N atom from the z-axis.

Table V. Helical Parameters, Interchain Energy, and Intrachain Energy of Triple-stranded Coiled-Coil structures of $(GPH)_4$, with Screw Symmetry and Clockwise Disposition of the Chains[a]

Helical Parameters			Energy[b] [kcal/mol of $(GPH)_4$]				
D	Θ[c]	ρ_1[d]	Interchain		Intrachain		Total
(Å)	(degrees)	(Å)	E_{ES}	E_{NB}	E_{ES}	E_{NB}	E_{tot}
5.0	-78.2	6.1	-2.4	-38.0	-55.0	13.7	-81.3
4.9	-73.2	6.4	-0.8	-37.0	-47.5	4.3	-81.0
6.0	66.7	3.6	-1.0	-21.3	-51.8	-1.8	-75.8
8.5	53.7	2.8	-0.8	-14.1	-52.6	-7.2	-74.6
8.2	-101.9	3.1	2.9	-34.2	-52.0	10.4	-73.2
4.6	-72.9	5.7	-1.2	-19.6	-56.6	4.1	-73.1
8.4	-63.3	4.4	0.8	-16.6	-53.4	-0.5	-69.6
3.5	50.1	7.4	-0.2	-11.4	-52.9	-3.5	-67.9
6.7	67.7	3.4	-1.3	-17.1	-48.8	0.7	-66.4
7.6	60.9	6.0	-1.8	-24.3	-52.7	13.6	-65.2
6.7	-64.5	3.9	-3.1	-21.2	-47.6	8.6	-63.1
8.5	33.7	4.6	-2.8	-10.3	-49.4	1.3	-61.1
4.4	-59.4	6.0	-0.4	-14.9	-45.3	5.4	-55.1
2.8	38.5	10.4	0.4	-10.6	-55.0	8.9	-55.0
7.6	52.3	5.7	1.7	-9.5	-51.0	3.7	-54.9
5.3	-53.0	6.5	1.9	-10.4	-48.2	2.1	-54.3
6.2	65.7	6.2	-0.2	-13.7	-51.6	18.3	-47.1
2.9	32.7	12.3	-0.8	-7.1	-54.4	17.2	-43.6
3.7	58.8	7.6	1.1	-26.8	-51.7	34.7	-42.7
5.5	-73.2	5.8	8.9	-24.7	-46.1	21.7	-40.1
6.6	-37.9	5.7	2.2	-12.5	-47.3	17.6	-39.6
4.5	-63.5	6.7	-0.5	-19.8	-53.8	51.6	-22.1

[a] The conformations are listed in the same order as in Table III.

[b] The subscript ES refers to the electrostatic energy term, NB refers to the nonbonded energy term. Both of these terms include the contributions to the hydrogen bond energy. In addition to the four energy terms shown here, E_{tot} also contains a torsional energy term E_{TOR}, contributed by the side-chain rotation about the C^{γ}–O^{δ} bond in Hyp, and a prolyl ring energy[34] which depends on whether $\omega = 0°$ or 180° preceding Pro or Hyp.

[c] The sign of θ determines the handedness of the major helix. See the text.

[d] The distance of the glycyl N atom from the z-axis.

Table VI. Cartesian Coordinates[a] of One Tripeptide Unit[b] of the Minimum-Energy Structure[c] of $(GPH)_n$

Residue	Atom	X	Y (Angstrom units)	Z
GLYCINE	N	1.60	0.0	0.0
	HN	0.78	-0.57	0.04
	CA	1.56	1.20	0.82
	HA[d]	1.87	1.98	0.29
	HA[d]	0.62	1.38	1.11
	C'	2.45	1.04	2.06
	O	2.93	-0.05	2.35
PROLINE	N	2.65	2.18	2.77
	CA	3.47	2.18	3.97
	HA[d]	4.29	1.63	3.83
	CB	3.80	3.64	4.22
	HB	3.75	3.88	5.28
	HB	4.81	3.87	3.89
	CG	2.78	4.44	3.43
	HG	2.05	4.90	4.10
	HG	3.26	5.25	2.89
	CD	2.09	3.49	2.46
	HD	1.01	3.51	2.60
	HD	2.28	3.77	1.43
	C'	2.73	1.53	5.14
	O	1.52	1.30	5.07
HYDROXYPROLINE	N	3.51	1.24	6.22
	CA	2.94	0.62	7.41
	HA[d]	2.30	-0.10	7.15
	CB	4.13	0.07	8.17
	HB	4.02	0.23	9.24
	HB	4.23	-1.01	8.02
	CG	5.35	0.80	7.63
	HG	5.74	1.43	8.30
	OD1	6.32	-0.26	7.23
	HD1	7.20	0.15	7.00
	CD2	4.94	1.50	6.34
	HD2	5.15	2.57	6.39
	HD2	5.48	1.10	5.48
	C'	2.13	1.63	8.22
	O	2.39	2.83	8.16

[a] Expressed in a right-handed coordinate system whose z-axis coincides with the major helical axis (the horizontal axis in Fig. 5 of ref. 2).

[b] The coordinates of the other tripeptide units in the same and in neighboring strands can be obtained by the use of symmetry operations as described in ref. 2.

[c] This is the first conformation listed in Tables II and IV.

[d] See footnote 36.

screw symmetry resulted in 23 minimum-energy structures with counterclockwise disposition of the chains and 22 structures with clockwise disposition of the chains. The energies ranged from −100.0 to −17.7 kcal/mol. This energy is expressed per mole of $(GPH)_4$, i.e., in terms of *one* polypeptide chain, as in earlier work.[2] The dihedral angles for the minima with the two dispositions are given in Tables II and III, respectively, listed in order of increasing energy. Tables IV and V, respectively, list the helical parameters and the inter- and intramolecular energy components for these conformations. The handedness of the triple helices of either disposition of the chains is indicated by the sign of Θ, as discussed above. In Tables IV and V, only the relative values of *intrachain* energies between conformations listed in the two tables are of physical significance, and not the magnitudes of these energies, because of differences in the conformationally independent intraresidue interactions when various sequences are compared.[34] For example, the large numerical values listed here and in ref 2 for the *intrachain* electrostatic energies are artifacts. They are a result of the omission of the invariant energy terms between atoms which are separated by one or two covalent bonds. The inclusion of these terms would merely change all energies in Tables IV and V by a constant amount. For the same reasons, the relative energies of various conformations of $(GPP)_4$ and $(GPH)_4$, with respect to a reference conformation, can be compared, but not the total energies (cf. Tables IX and X of ref 2 and Tables IV and V of this paper). On the other hand, the values of the *interchain* energies can be compared directly for various structures and for different sequences, because the zero of the interchain energy scale corresponds to well-defined physical conditions, viz., to noninteracting chains at infinite separation.

The total *intrachain* nonbonded energy is relatively small in magnitude in the low-energy structures, because there are few close approaches in the nearly stretched single-strand conformations. By contrast, the *interchain* nonbonded interaction energy dominates in most cases, because of the large number of pairs of atoms found at short distances in the intertwined triple-stranded helices. Interchain energies influence not only the packing of the chains but also the conformation which they adopt after energy minimization in the triple-stranded complex.

All of the 81 parallel-chain triple helices with screw symmetry were much higher in energy than the coiled-coil triple helices. The energies ranged from −40.0 to 0.0 kcal/mol of $(GPH)_4$. Twenty-eight minimum-energy conformations were found for triple-stranded coiled-coil and parallel-chain structures with rotational symmetry, with energies ranging from −72.3 to −35.3 kcal/mol of $(GPH)_4$. Detailed data are not tabulated here for either of these two kinds of structures because all conformations are of very high energy.

All results are closely similar to those obtained for $(GPP)_4$. The lowest-energy conformation computed for both molecules is a coiled-coil structure with screw symmetry and a counterclockwise disposition of the strands, which is very close to proposed structures[5,8,13] for collagen. The energy difference between this conformation and the next-lowest energy conformation (line 2 of Table II) is 4.6 kcal/mol of GXY repeat units. Thus, the lowest-energy conformation is the only significant triple-stranded complex in both cases, in agreement with experiment.

Table VII
Dihedral Angle $\chi^{3,2}$ for Rotation around the C^{γ}-O^{δ} Bond and Relative Energies of the Three Staggered Minimum-Energy Side-Chain Conformations[a] of the Hydroxyprolyl Residue

	staggered minima[b]					
	gauche$^+$		trans		gauche$^-$	
structure	χ	ΔE	χ	ΔE	χ	ΔE
N-acetyl-*N'*-methylhydroxyprolinamide	59.0	0.47	−171.1	0.02	−70.2	0.00
single-stranded $(GPH)_4$	58.9	0.45	−171.0	0.09	−69.6	0.00
single-stranded $(Gly\text{-}Ala\text{-}Hyp)_4$	58.9	0.59	−171.0	0.11	−69.7	0.00
triple-stranded $(GPH)_4$	58.6	0.45	−170.9	0.13	−69.1	0.00
single-stranded $(GHP)_4$	58.7	0.45	−170.7	0.15	−69.3	0.00
single-stranded $(Gly\text{-}Hyp\text{-}Ala)_4$	58.7	0.43	−170.7	0.14	−69.4	0.00
triple-stranded $(GHP)_4$	59.1	0.48	−172.1	−0.08	−69.9	0.00

[a] The dihedral angles in the backbone were kept fixed at the values listed in the first line of Table II. Only $\chi^{3,2}$ was treated as a variable during this minimization. The ring was placed in the "down" puckering geometry. [b] Dihedral angles in degrees; energies, relative to the energy in the gauche$^-$ conformation, in kcal/mol of (GXY) repeat units.

Five of the six conformations of $(GPH)_4$ with counterclockwise disposition and energies ≤ -70.0 kcal/mol of $(GPH)_4$, i.e., with $\Delta E \leq 30.0$ kcal relative to the lowest energy conformation (Table II), are identical (to within 2° in terms of dihedral angles and 0.9 kcal/mol in terms of ΔE) with the corresponding five lowest-energy conformations of $(GPP)_4$ (Table V of ref 2). In addition, most of the higher-energy conformations in Table II correspond to one of the higher-energy conformations of $(GPP)_4$, although the ordering in terms of energies is different and there are changes of a few degrees in dihedral angles. The same observation can be made in Table III. There are only two, relatively high energy, conformations of $(GPH)_4$ (with energies −81.3 and −73.2 kcal/mol) which are not close to corresponding conformations of $(GPP)_4$ (Table VI of ref 2). All parallel-chain triple-stranded conformations, as well as coiled coils with rotational symmetry, have very high energies in both polypeptides.

The lowest-energy conformation (first line of Tables II and IV) has an energy −100.0 kcal/mol of $(GPH)_4$. The dihedral angles of its repeating unit are (ϕ_1, ψ_1, ω_1, ψ_2, ω_2, ψ_3, $\chi_3^{3,2}$) = (−74°, 170°, 180°, 167°, 180°, 153°, −171°); its helical parameters are D = 9.0 Å and Θ = 43.6°. This conformation is practically identical with that found for $(GPP)_4$ in the previous paper (Figure 5 of ref 2). It consists of right-handed major helices in a counterclockwise disposition. The same $N_1H_1\cdots O_2C_2$ interchain backbone hydrogen bond occurs in both molecules. Based on the coordinates of all backbone atoms (including the C^{β}), the root-mean-square deviation between the two structures is 0.02 Å.

B. Interactions of the Hydroxyl Group with the Backbone. As discussed in the first paper, the lowest-energy triple-stranded backbone conformation found there for $(GPP)_4$ and here for $(GPH)_4$ is the one closest to the proposed conformation[5-8,13] of the collagen triple helix. In view of the reported stabilization of the triple helix when Hyp occupies position Y, as discussed in section I, it is important to see whether intra- or interchain interactions of the hydroxyl group of Hyp with other groups in this triple-stranded structure make a stabilizing contribution. In the lowest-energy conformation, in which all peptide groups are *trans*, the OH groups are on the outside of the triple helix and therefore they cannot form a hydrogen bond with backbone carbonyl groups.[28] The interchain energy contributions, shown in the first line of Table IV for $(GPH)_4$ and in the first line of Table IX of ref 2 for $(GPP)_4$, are nearly equal. This indicates that the OH group does not make a stabilizing contribution in this conformation when it is placed in the minimum-energy position for the *t* rotamer, corresponding to $\chi^{3,2} = -171°$.

No additional stabilizing contributions arise in this triple-stranded backbone conformation for the other stable conformations of the OH group either. The energy is not changed significantly when the OH group is placed into the g^+ or g^- conformations, as shown in the fourth row of Table VII. Energy minimization was carried out for several sequences and structures as a function of $\chi^{3,2}$, with the backbone dihedral angles kept fixed at the values given[2] for the lowest-energy $(GPP)_4$ conformation, and starting from all three staggered OH conformations. The results are shown in Table VII. The final dihedral angles and relative energies are identical for *N*-acetyl-*N'*-methylhydroxyprolinamide and for single- and triple-stranded $(GPH)_n$. The same results were obtained when Hyp was placed in position X, i.e., for single- and triple-stranded $(GHP)_4$, and even when alanine was substituted for the neighboring proline, i.e., in single-stranded $(Gly\text{-}Ala\text{-}Hyp)_4$ and $(Gly\text{-}Hyp\text{-}Ala)_4$, with the same fixed backbone conformation. The change in relative energies is very small even in the case of $(GHP)_4$. Table VII shows that the energy difference between the three staggered conformations is practically independent of the polypeptide sequence and of the presence and absence of the other two strands. Consequently, this energy difference must be due entirely to local intrastrand interactions of the hydroxyl side chain with the hydroxyprolyl ring and the two peptide groups flanking it. The effect of other amino acid substitutions in positions X and Y upon intra- and interchain interaction energies will be presented in a more detailed survey of side-chain interactions in collagen.[51]

These results demonstrate (a) that the collagen-like triple-helical conformation of $(GPH)_n$ is not stabilized by interactions involving only the hydroxyl group of Hyp and the rest of the triple helix (cf. ref 19), and (b) that the presence of Hyp in position Y rather than in position X does not result in increased intrinsic stability of the collagen triple helix. Therefore, these results support the suggestions[26-28,52] that the increased thermal stability of the $(GPH)_n$ triple helix, as compared with the $(GPP)_n$ triple helix, is due to interactions with the solvent. A study of the effect of hydration on the energetics of these triple helices and a comparison of the energies of hydrated Hyp in positions X and Y is in progress.

The preceding discussion applies only to the lowest-energy conformation. It was demonstrated[31,32] that a triple helix containing cis peptide units is possible geometrically in which the hydroxyl group forms a hydrogen bond with a carbonyl oxygen of a neighboring strand. Such a hydrogen bond occurs in some of the conformations found in this study as well. These conformations, however, are

much higher in energy than the triple helix of lowest energy.

It is possible that OH···backbone hydrogen bonds could be formed in some of the other high-energy conformations listed in Tables II and III, if the OH group were to be rotated into other staggered conformations instead of that with $\chi^{3,2}$ near 180°. Such a hydrogen bond, however, would contribute about −5 kcal/mol of GPH units at best.[34] This is not sufficient to lower the energy of the higher-energy conformations below that of the lowest-energy conformation found here.

C. Comparison with Other Studies. X-ray diffraction studies of $(GPH)_n$ fibers[15,53] gave a collagen-like diffraction pattern, with a reflection corresponding to the helical parameter D = 8.46 Å. No value of θ was reported. We computed D = 9.0 Å and θ = 43.6° for the lowest-energy triple-stranded coiled-coil structure. The extent of agreement is similar to that obtained for $(GPP)_n$ earlier,[2] as discussed in section I.

Based on their X-ray crystallographic study[9] of $(Pro\text{-}Pro\text{-}Gly)_{10}$, Okuyama et al.[54] proposed a model for collagen in the form of a 7/1 helix. Their proposed structure, with D = 8.6 Å and θ = 51.4°, is close to the one we computed[2] for $(GPP)_n$. They reported[54] the same pseudoperiod, i.e., the same D for $(Pro\text{-}Hyp\text{-}Gly)_9$ as for $(Pro\text{-}Pro\text{-}Gly)_{10}$.

In a recent X-ray diffraction study of stretched kangaroo-tail tendon, Fraser et al.[8] obtained the helical parameters $\alpha = -107.1 \pm 0.6°$ and $h = D/3 = 2.983 \pm 0.003$ Å. In terms of the parameters used in this paper, this corresponds to $D = 8.949 \pm 0.01$ Å and $\theta = 38.7 \pm 1.8°$. Fraser et al.[8] state that this value of α (or θ) is inconsistent with the model we proposed in the first paper,[2] for which D = 8.95 Å and θ = 44°. The differences in D and θ are similar in magnitude, however, to differences between various models[5-9] based on diffraction studies. Such differences may be due to differences in physical conditions between various studies, as discussed in the next paragraph. Fraser et al.[8] carried out a least-squares refinement of atomic coordinates, using the sequence $(Gly\text{-}Pro\text{-}Hyp)_n$ and the molecular geometry defined by Momany et al.[34] They obtained the best fit with ring puckering "down" for Pro and "up" for Hyp. However, the fit that they obtained is not very different with "down" puckering for both residues, i.e., with the conformation that we found to be of lowest energy.

The values of both D and θ which we compute for $(GPP)_n$ and $(GPH)_n$ are higher than those derived for the collagen models which were based on X-ray studies of fibers[5-8,13] but not far from those reported by Fraser et al.[8] There might be several reasons for the differences. First, the assumption of rigid prolyl ring geometry, made in the present work, may influence the computed helical parameters. Second, packing in fibers may cause small changes in the geometry of the triple helix. This possibility is suggested by the even larger differences between the observed helical parameters of the models based on fibers[5-8,13] and the parameters derived from single crystals[9,54] of $(PPG)_{10}$ and $(PHG)_9$. Third, the substitution of various amino acids in place of proline in collagen may alter slightly the helical parameters from those computed[2] and observed[9] for poly(Gly-Pro-Pro). This possibility is suggested by recent potential energy calculations, to be reported elsewhere,[51] on a polypeptide with a sequence corresponding to a fragment of the α1(I) chain of collagen.

Bansal et al.[32] proposed a low-energy structure for the $(GPH)_n$ triple helix, with "up" puckering of the prolyl rings, in which the peptide bond preceding the prolyl residue is in the cis conformation and a hydrogen bond is formed between the hydroxyl group and a glycyl carbonyl oxygen of a neighboring strand. We have minimized the energy of the triple helix, starting from their dihedral angles.[55,56] We found that the energies of the minimized structures are −63.6 kcal/mol of $(GPH)_4$ for "down" puckering and −57.6 kcal/mol of $(GPH)_4$ for "up" puckering, as compared with −100.0 kcal/mol of $(GPH)_4$ for the collagen-like structure (Table II). Expressed in terms of the energy per (GPH) repeat unit, the two minima we computed are 9.1 and 10.6 kcal/mol higher in energy, respectively, than the collagen-like structure, in contrast to the findings of Bansal et al. An exact comparison is not possible because of differences of the residue geometries used in the two studies. The helical parameters that we compute for both minimized structures, however, differ considerably from those cited by Bansal et al.,[32] indicating that energy minimization leads to different structures.

IV. Conclusions

In the use of energy computations to determine stable structures of globular proteins, the most difficult problem arises from the existence of many minima in multi-dimensional conformational space.[57] Even for small molecules, various strategies must be resorted to in order to circumvent this problem.[57,58] In the case of collagen-like repeating poly(tripeptide)s discussed here and in the other papers of this series, the problem could be overcome because the structure is regular. Each strand as well as the neighboring strands can be generated by symmetry operations from the conformation of the GXY repeating unit. This unit is sufficiently small so that the available conformational space can be explored adequately. The choice of many starting conformations and their testing in triple-helical structures of various symmetries ensures that the global minimum in the conformational energy is not missed.

The energies of the computed lowest-energy triple-stranded structures with different kinds of symmetry were as follows: −100.0, −40.0, and −72.3 kcal/mol of $(GPH)_4$ for coiled coil with screw symmetry, for parallel chains with screw symmetry, and for structures with rotational symmetry, respectively.

The lowest-energy conformation of triple-helical $(GPH)_n$ was found to be identical with that of $(GPP)_n$. It is a coiled-coil structure with screw symmetry. It is similar to a deformed polyproline II conformation and close to that proposed[5-8,13] for the structure of collagen. It is stabilized by interchain interactions, just as the analogous structure[2] of $(GPP)_n$. The results indicate that this conformation can accommodate hydroxyproline and proline with equal ease. The energy of all other computed triple-stranded complexes is much higher. Thus, the collagen-like conformation is shown by the computations to be the most stable one.

The intra- and interstrand energy of this triple-stranded complex is not changed by the substitution of hydroxyproline for proline in the Y position. Presumably, the entropies of $(GPP)_n$ and $(GPH)_n$ in the random coil state are very similar, as discussed in section I. Therefore, the observed stabilization of the triple-stranded structure in collagen and collagen-like polypeptides, with increasing hydroxyproline content, is likely to be due to interactions with the solvent.

Acknowledgment. We thank Dr. S. S. Zimmerman for helpful discussions and M. S. Pottle and S. Rumsey for extensive help with the computations. This work was supported by research grants from the National Science

Foundation (PCM75-08691), the National Institute of General Medical Sciences (GM-14312), and the National Institute on Aging (AG-00322) of the National Institutes of Health.

Miniprint Material Available: Full-sized photocopies of Tables II–VI (5 pages). Ordering information is given on any current masthead page.

References and Notes

(1) (a) Presented before the Division of Cellulose, Paper, and Textile at the 178th National Meeting of the American Chemical Society, Washington, D.C., September 1979, Abstract No. CELL 13. (b) To whom requests for reprints should be addressed.
(2) Miller, M. H.; Scheraga, H. A. *J. Polym. Sci., Polym. Symp.* **1976**, *54*, 171.
(3) Paper 3: Miller, M. H.; Némethy, G.; Scheraga, H. A. *Macromolecules*, in press.
(4) Paper 4: Némethy, G.; Miller, M. H.; Scheraga, H. A. *Macromolecules*, in press.
(5) Rich, A.; Crick, F. H. C. *Nature (London)* **1955**, *176*, 915.
(6) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483.
(7) Yonath, A.; Traub, W. *J. Mol. Biol.* **1969**, *43*, 461.
(8) Fraser, R. D. B.; MacRae, T. P.; Suzuki, E. *J. Mol. Biol.* **1979**, *129*, 463.
(9) Okuyama, K.; Tanaka, N.; Ashida, T.; Kakudo, M. *Bull. Chem. Soc. Jpn.* **1976**, *49*, 1805.
(10) Hulmes, D. J. S.; Miller, A.; Parry, D. A. D.; Piez, K. A.; Woodhead-Galloway, J. *J. Mol. Biol.* **1973**, *79*, 137.
(11) Fietzek, P. P.; Glanville, R. W., personal communication.
(12) Berg, R. A.; Kishida, Y.; Sakakibara, S.; Prockop, D. J. *Biochemistry* **1977**, *16*, 1615.
(13) Traub, W.; Piez, K. A. *Adv. Protein Chem.* **1971**, *25*, 243.
(14) Bhatnagar, R. S.; Rapaka, R. S. In "Biochemistry of Collagen"; Ramachandran, G. N., Reddi, A. K., Eds.; Plenum Press: New York and London, 1976; p 479.
(15) Millionova, M. I.; Andreyeva, N. S.; Lebedev, L. A. *Biophysics (Russ.)* **1963**, *8*, 478.
(16) Millionova, M. I.; *Biophysics (Russ.)* **1964**, *9*, 149.
(17) Kobayashi, Y.; Sakai, R.; Kakiuchi, K.; Isemura, T. *Biopolymers* **1970**, *9*, 415.
(18) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, H.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta* **1973**, *303*, 198.
(19) Engel, J.; Chen, H.-T.; Prockop, D. J.; Klump, H. *Biopolymers* **1977**, *16*, 601.
(20) It has also been suggested[21] that poly(Ala-Hyp-Gly) may form a triple helix in water which is more stable than triple-helical poly(Ala-Pro-Gly).
(21) Rao, N. V.; Adams, E. *Biochem. Biophys. Res. Commun.* **1979**, *86*, 654.
(22) Rosenbloom, J.; Harsch, M.; Jimenez, S. *Arch. Biochem. Biophys.* **1973**, *158*, 478.
(23) Berg, R. A.; Prockop, D. J. *Biochem. Biophys. Res. Commun.* **1973**, *52*, 115.
(24) Ward, A. R.; Mason, P. *J. Mol. Biol.* **1973**, *79*, 431.
(25) Clark, D. S.; Mattice, W. L. *Macromolecules* **1977**, *10*, 369.
(26) Ramachandran, G. N.; Bansal, M.; Bhatnagar, R. S. *Biochim. Biophys. Acta* **1973**, *322*, 166.
(27) Traub, W. *Isr. J. Chem.* **1974**, *12*, 435.
(28) Bansal, M.; Ramakrishnan, C.; Ramachandran, G. N. *Proc. Indian Acad. Sci., Sect. A* **1975**, *82*, 152.
(29) Némethy, G.; Hodes, Z. I.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5760.
(30) Hodes, Z. I.; Némethy, G.; Scheraga, H. A. *Biopolymers* **1979**, *18*, 1565.
(31) Berg, R. A.; Kishida, Y.; Kobayashi, Y.; Inouye, K.; Tonelli, A. E.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1973**, *328*, 553.
(32) Bansal, M.; Ramakrishnan, C.; Ramachandran, G. N. *Biopolymers* **1975**, *14*, 2457.
(33) IUPAC-IUB Commission on Biochemical Nomenclature: *Biochemistry* **1970**, *9*, 3471.
(34) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.
(35) Momany, F. A.; Carruthers, L. M.; McGuire, R. F.; Scheraga, H. A. *J. Phys. Chem.* **1974**, *78*, 1595.
(36) The residue geometry in our computer program[34] defines the C^{α}–H bond length as 1.00 Å. All other C–H bond lengths are fixed at 1.09 Å. Neutron diffraction data[37] published since the time of completion of the study used to establish the residue geometry[34] indicate that the C^{α}–H distance also is 1.09 Å. This was found out after the completion of the computations reported here. Repetition of the computations for several of the lowest-energy structures, using the new bond length, indicated that the dihedral angles and the energies did not change noticeably. Therefore, the conclusions in the text are unaffected by the correction of the C^{α}–H bond length. The tables in this paper were computed with the old value of 1.00 Å.
(37) (a) Al-Karaghouli, A. R.; Koetzle, T. F. *Acta Crystallogr., Sect. B* **1975**, *31*, 2461. (b) Power, L. F.; Turner, K. E.; Moore, F. H. *Ibid.* **1976**, *32*, 11. (c) Kvick, Å.; Al-Karaghouli, A. R.; Koetzle, T. F. *Ibid.* **1977**, *B33*, 3796.
(38) Tanaka, *S.*; Scheraga, H. A. *Macromolecules* **1974**, *7*, 698.
(39) Ramachandran, G. N.; Lakshminarayanan, A. V.; Balasubramanian, R.; Tegoni, G. *Biochim. Biophys. Acta* **1970**, *221*, 165.
(40) Ramachandran, G. N.; Bansal, M. *Curr. Sci.* **1976**, *45*, 647.
(41) Madison, V. *Biopolymers* **1977**, *16*, 2671.
(42) Nishikawa, K.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 395.
(43) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95.
(44) Sugeta, H.; Miyazawa, T. *Biopolymers* **1967**, *5*, 673.
(45) Cowan, P. M.; McGavin, S.; North, A. C. T. *Nature (London)* **1955**, *176*, 1062.
(46) Ramachandran, G. N.; Kartha, G. *Nature (London)* **1954**, *174*, 269.
(47) McGuire, R. F.; Vanderkooi, G.; Momany, F. A.; Ingwall, R. T.; Crippen, G. M.; Lotan, N.; Tuttle, R. W.; Kashuba, K. L.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 112.
(48) Pincus, M. R.; Scheraga, H. A. *J. Phys. Chem.* **1977**, *81*, 1579.
(49) Fletcher, R.; Powell, M. J. D. *Comput. J.* **1963**, *6*, 163.
(50) Zimmerman, S. S.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 811.
(51) Némethy, G.; Scheraga, H. A.; in preparation.
(52) Privalov, P. L.; Tiktopulo, E. I.; Tischenko, V. M. *J. Mol. Biol.* **1979**, *127*, 203.
(53) Rogulenkova, V. N.; Millionova, M. I.; Andreeva, N. S. *J. Mol. Biol.* **1964**, *9*, 253.
(54) Okuyama, K.; Takayanagi, M.; Ashida, T.; Kakudo, M. *Polym. J.* **1977**, *9*, 341.
(55) We have learned from Dr. Bansal[56a] that the list of dihedral angles in ref 32 contains a typographical error. We have used the corrected value of $\psi_2 = 162°$ in our calculation, as supplied by Dr. Bansal.[56a,b]
(56) (a) Bansal, M.; personal communication. (b) Bansal, M.; Ph.D. Dissertation, Indian Institute of Science, Bangalore, 1976.
(57) Némethy, G.; Scheraga, H. A. *Q. Rev. Biophys* **1977**, *10*, 239.
(58) Scheraga, H. A. *Pept., Proc. Am. Pept. Symp., 5th* **1977**, 246.