Hydroxylation of Proline in Polytripeptide Models of Collagen: Stereochemistry of Polytripeptide–Prolyl Hydroxylase Interaction†

Rao S. Rapaka, V. Renugopalakrishnan, D. W. Urry, and Rajendra S. Bhatnagar*

ABSTRACT: The unusual imino acid hydroxyproline is synthesized by the enzymatic hydroxylation of proline in the sequence -X-Pro-Gly- in polypeptide precursors and synthetic polytripeptide models of collagen. The degree of hydroxylation is related to the nature of the side-chain residue at X. Polytripeptides with X = Pro were the best substrates, followed, in order of decreasing substrate efficiencies, by X = Ala > Leu > Val > Phe > Gly ≃ Sar. Since substrate–enzyme interaction in this case may result in conformational transitions both in the polytripeptide substrate and the enzyme, and as the conformational states of the various polytripeptide substrates may be influenced by the stereochemical interactions of the side chain X, the conformational states present may determine the extent of hydroxylation. In order to investigate this hypothesis, conformational energy calculations were performed for the X-Pro-Gly tripeptide sequences. Since $(Pro\text{-}Pro\text{-}Gly)_n$ was the most efficient synthetic substrate, its conformation was considered to be the favored conformation for hydroxylation. The calculations bring into perspective three factors that may regulate the relative hydroxylations of the different sequences: (1) the hydroxylatable region of a substrate's ψ_1–ψ_2 plot which relates to the adaptability of the enzyme to a range of conformations of the peptide substrate, (2) the energy difference between the lowest energy conformation of a substrate and its energy in the ψ_1–ψ_2 energy surface, which corresponds to the lowest energy of the preferred substrate, and (3) the multiplicity of low-energy states available to the substrate. Evidence is presented in tripeptides of X-Pro-Gly sequences that where energetically favored conformations predominate with $\psi_1 = 100 \pm 40°$ [$-C_x^{\alpha}-C(=O)-$] and $\psi_2 = 130 \pm 30°$ [$-C_{Pro}^{\alpha}-C(=O)-$] hydroxylation occurs and where energetically preferred conformations are outside this range of ψ_1 and ψ_2 hydroxylation does not occur. It is proposed that for hydroxylation by prolyl hydroxylase to occur the above proposed critical range of conformations is necessary, and the extent of hydroxylation may depend on the relative population of this range of conformations to the sum of allowed conformations.

The structural protein collagen has an unusual repeating primary structure in which glycine appears in every third position. With the exception of short sequences at the N and C terminals, the collagen polypeptides can be represented as polymeric $(X\text{-}Y\text{-}Gly)_n$ (review, Piez, 1976).[1] Approximately one fourth of the residues in collagen are the imino acids proline and hydroxyproline, the latter occurring only at the position Y in the sequence and any occurrence of proline in this position is likely to be due to the reduced hydroxylation of proline, since hydroxyproline is synthesized by the enzymatic hydroxylation of proline after its incorporation in the polypeptide sequence. The hydroxylation of proline in polypeptide precursors or analogues of collagen is catalyzed by the enzyme prolyl hydroxylase in a complex reaction involving the reductive fixation of oxygen and is coupled to the oxidative decarboxylation of α-ketoglutarate (review Prockop et al., 1976). The hydroxylation of proline is a critical step in the synthesis of collagen, since unhydroxylated collagen is not efficiently secreted from cells (Bhatnagar and Prockop, 1966), has a lower denaturation temperature than fully hydroxylated collagen (Rosenbloom et al., 1973; Berg and Prockop, 1973), and is therefore susceptible to nonspecific proteolysis. Hydroxyproline stabilizes the collagen triple helix by participating in hydrogen bonds (Ramachandran et al., 1973, 1975). It is therefore of interest to examine the factors which regulate the formation of this unusual imino acid in collagen.

The prolyl hydroxylase reaction is also a very interesting system for examining enzyme–substrate interactions and for investigating the physicochemical aspects of enzymatic catalysis. The enzyme binds three different substrates, namely, O_2, α-ketoglutarate, and the polypeptide substrate containing susceptible proline residues. An examination of the primary sequence of the α chains of mammalian collagen indicates that in triplets of the type Gly-Pro-X and X-Pro-Gly only the latter are hydroxylated (Piez, 1970). The examination further reveals that hydroxylation occurs when X is Pro, Ala, Leu, Phe, etc. The prolyl substrate specificity of the enzyme is limited, and proline in the -Pro-Gly- sequence is hydroxylated in a large number of polypeptides (Bhatnagar and Rapaka, 1976; Prockop et al., 1976). The conformation of a polypeptide substrate may be expected to play a role in regulating its interaction with the enzyme and in determining the extent of hydroxylation. The available information on the influence of polypeptide conformation on the hydroxylation of proline and on the interaction of the polypeptide with the enzyme is not self-consistent. Thus, the hydroxylation of unhydroxylated collagen and of the collagen analogue $(Pro\text{-}Pro\text{-}Gly)_n$ is less efficient under conditions favoring the development of collagen-like order (Prockop et al., 1976). Strong complex formation with the enzyme without hydroxylation occurs with the synthetic polypeptides $(Gly\text{-}Pro\text{-}Gly)_n$ and $(Sar\text{-}Pro\text{-}Gly)_n$, neither of which exists in ordered conformations under the conditions of hydroxylation (Bhatnagar and Rapaka, 1976). In contrast, $(Ala\text{-}Pro\text{-}Gly)_n$ and $(Leu\text{-}Pro\text{-}Gly)_n$, neither of which forms triple helical structures in aqueous solutions, are hydroxylated (Bhatnagar and Rapaka, 1976; Kivirikko et al.,

† From the Laboratory of Molecular Biophysics, University of Alabama Medical Center, Birmingham, Alabama 35294 (R.S.R., V.R., and D.W.U.), and the Laboratory of Connective Tissue Biochemistry, University of California, San Francisco, California 94143 (R.S.B.). *Received November 2, 1977.* These studies were supported by National Institutes of Health Grants HL-11310 (D.W.U.) and HL-19668 (R.S.B.).

[1] Abbreviations used are: Ala, alanine; Gly, glycine; Leu, leucine; Phe, phenylalanine; Pro, proline; Sar, sarcosine; Val, valine.

0006-2960/78/0417-2892$01.00/0

1969; Scatturin et al., 1975). Polyproline II, which exhibits a helical form similar to the single chains of collagen, interacts strongly with prolyl hydroxylase but it is not hydroxylated (Prockop and Kivirikko, 1969). $(Gly\text{-}Pro\text{-}Ala)_n$ and $(Gly\text{-}Pro\text{-}Sar)_n$, both of which exhibit considerable order in aqueous solutions, do not interact with prolyl hydroxylase (Bhatnagar and Rapaka, 1976). Several other synthetic polymers with collagen-like sequences have been examined for hydroxylation (Hutton et al., 1968; Okada et al., 1972; Kivirikko et al., 1971), but a comparative evaluation of their ability to serve as substrates for prolyl hydroxylase is difficult because of differences in molecular weights and reaction conditions employed. We compared the hydroxylation of synthetic polypeptides of similar degrees of polymerization and our studies (Bhatnagar and Rapaka, 1974) indicated a relationship between the ability of a polymer $(X\text{-}Pro\text{-}Gly)_n$ to undergo hydroxylation and the nature of the side chain of the residues at X. Polytripeptides in which X = Pro were the most efficient substrates. Increasing complexity of the side chain of other residues at X decreased the efficiency with which hydroxylation occurred.

The side chain of the residue on the N-terminal side of the proline residue can interact with the pyrrolidine ring and thereby restrict the conformational range of the X–Pro peptide bond (Schimmel and Flory, 1968). The residue X preceding proline may influence the conformation, as evidenced by the fact that in X-L-Pro peptides, where X = Gly, L-Ala, L-Leu, L-Phe, D-Ala, D-Leu, and D-Phe, the equilibrium between the cis and trans forms of X–Pro peptide bonds is affected by the side chain of X and the configuration of X (Grathwohl and Wütherich, 1976). Conformational energy calculations by Zimmerman and Scheraga (1977) on Pro-X and X-Pro dipeptides revealed that β turns are not favored on X-Pro dipeptides because the interactions between the X residue and the pyrrolidine ring restrict the conformations of X.

It was shown that interactions between X and Pro in X-Pro peptides play an important role in repeat peptides of elastin (Urry, 1976). In the sequential polypeptide $(Val\text{-}Pro\text{-}Gly\text{-}Gly)_n$ the intramolecular hydrophobic interaction between the valyl and prolyl side chains generated the property of coacervation, whereas in the sequential polypeptide $(Ala\text{-}Pro\text{-}Gly\text{-}Gly)_n$, due to the weakened hydrophobic interaction, the polymer is devoid of coacervation (Rapaka and Urry, 1978; Rapaka et al., 1978b). These observations were further confirmed by nuclear Overhauser studies on these polymers (Urry, Khaled, Rapaka, and Okamoto, in preparation). The present study was undertaken to see the influence of the conformational restrictions imposed by such interactions in regulating the hydroxylation of peptidyl proline as well.

Since enzyme–substrate interactions involve conformational transitions, we considered the possibility that, for the optimal interaction between prolyl hydroxylase and tripeptide substrates, the tripeptide should be able to present a specific recognizable conformation to the enzyme. Hence, conformational energy calculations were performed on X-Pro-Gly tripeptide sequences, where X is Gly, Ala, Leu, Val, Phe, Sar, and Pro. Since $(Pro\text{-}Pro\text{-}Gly)_n$ is the most efficient substrate, its conformation was assumed to be the most favorable for hydroxylation. Our studies have analogy to those of Dafforn and Koshland (1971), where they proposed from transition-state calculations that the large increase in rates in enzymatic reactions is due to the ability of the enzyme to orient the reacting atoms in an optimal manner. This effect was termed "Orbital Steering". In this study, it is proposed that an optimal enzyme–substrate interaction occurs, to the extent that a sufficiently favorable steric conformation is presented by the substrate.

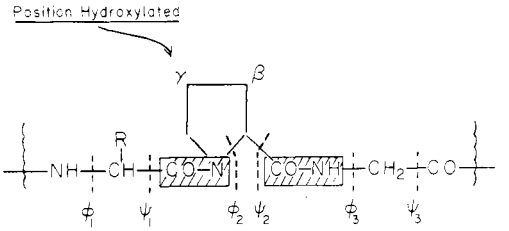

FIGURE 1: A typical -X-Pro-Gly- tripeptide segment as a substrate for hydroxylation is shown. Torsion angles ϕ_1, ψ_1, ϕ_2, ψ_2, ϕ_3, and ψ_3 are illustrated along with the rigid peptide bonds.

Materials and Methods

Synthetic Polypeptides. The polytripeptides $(Ala\text{-}Pro\text{-}Gly)_n$ and $(Pro\text{-}Pro\text{-}Gly)_n$ where purchased from Miles Laboratories, Kankakee, Ill. $(Gly\text{-}Pro\text{-}Gly)_n$ was synthesized by polymerizing the pentachlorophenyl ester of Pro-Gly-Gly. The polytripeptides $(Sar\text{-}Pro\text{-}Gly)_n$, $(Val\text{-}Pro\text{-}Gly)_n$, $(Leu\text{-}Pro\text{-}Gly)_n$, and $(Phe\text{-}Pro\text{-}Gly)_n$ were synthesized by polymerizing the respective active esters as described elsewhere (Rapaka and Bhatnagar, 1975). All polytripeptides were extensively fractionated using gel-filtration procedures, and fractions corresponding to narrow ranges of degree of polymerization were collected for all polymers, to facilitate the comparison of hydroxylation properties of the polypeptides.

Prolyl Hydroxylase Reaction. Chick embryo prolyl hydroxylase was purified using published procedures (Halme et al., 1970). The reaction mixture for hydroxylation consisted of the enzyme, 1.0 mg; α-ketoglutarate, 0.5 mM; ferrous ammonium sulfate, 0.04 mM; ascorbate, 2.0 mM; Tris-HCl, 0.1 mM (pH 7.4); and the polytripeptide, 300 μg; in a total volume of 4.0 mL. Incubation was carried out at 37 °C for 1 h and was terminated by the addition of 4 mL of concentrated HCl. The reaction tubes were sealed under N_2, and the samples were hydrolyzed for 22 h at 110 °C. The hydrolysates were evaporated to dryness in a rotory evaporator under vacuum and the residues were dissolved in appropriate amounts of H_2O. Hydroxyproline content of the hydrolysates was assayed by a microadaptation of Woessner's procedure (Woessner, 1961).

Conformational Energy Calculations. A diagrammatic representation of the X-Pro-Gly segment undergoing hydroxylation is given in Figure 1. The partitioned potential-energy method (Ramachandran and Sasisekharan, 1968; Renugopalakrishnan et al., 1976b) was used for conformational energy calculations on *N*-formyl derivatives of Gly-L-Pro-Gly, L-Ala-L-Pro-Gly, L-Leu-L-Pro-Gly, L-Val-L-Pro-Gly, L-Phe-L-Pro-Gly, and Sar-L-Pro-Gly amides. The fully extended conformations of the tripeptides, except for the sarcosyl tripeptide, were constructed with standard bond lengths and bond angles taken from earlier work (Khaled et al., 1976; Renugopalakrishnan et al., 1976a; Urry et al., 1977) and from Momany et al. (1975). The *N*-formyl end group was constructed with C—N, C=O, and C—H bond lengths of 1.32, 1.24, and 1.09 Å, respectively, and with N—C=O and N—C—H bond angles of 125° and 115°, respectively (Sutton, 1965).

The sarcosyl backbone was constructed with the same geometry as a Gly residue. The sarcosyl *N*-methyl group was constructed using an N-C bond length of 1.45 Å and a C^{α}-N-CH_3 bond angle of 117° (Howard et al., 1973). The *N*-methyl group was assumed to be tetrahedral with a C-H bond length of 1.09 Å and an H-C-H bond angle of 109.47°. Peptide groups were assumed to be in a trans planar conformation. L-Proline was assumed to be in C^{γ} endo conformation as found

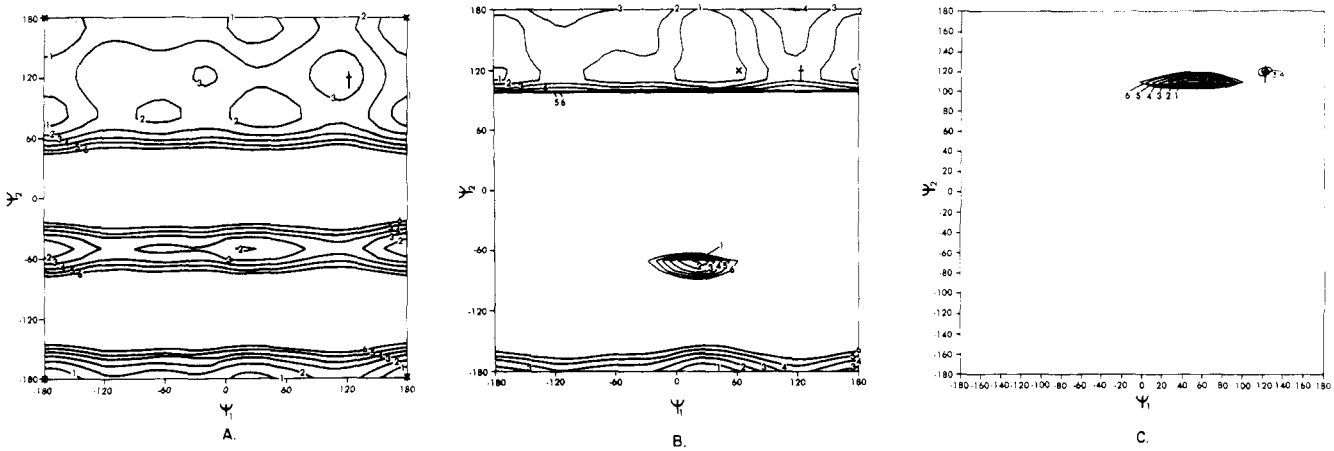

FIGURE 2: ψ_1-ψ_2 energy surfaces in kcal/mol relative to the global minimum marked by the symbol x for: (A) Gly-L-Pro-Gly with $\phi_1 = 80°$, $\phi_2 = -60°$, $\phi_3 = +180°$, and $\psi_3 = +180°$. (B) L-Ala-L-Pro-Gly with $\phi_1 = -100°$, $\phi_2 = -60°$, $\phi_3 = +180°$, and $\psi_3 = \pm 180°$. (C) L-Leu-L-Pro-Gly with $\phi_1 = -100°$, $\phi_2 = -60°$, $\phi_3 = +180°$, and $\psi_3 = +180°$. Symbol † denotes the conformation at 1 = 2 = 120° which corresponds to the ideal conformation of HCO-L-Pro-L-Pro-Gly-NH_2.

in the X-ray study of p-BrCbz-Gly-L-Pro-L-Leu-Gly-OH by Ueki et al. (1969). Side chains were assumed to be in energetically preferred conformations (Ramachandran and Sasisekharan, 1969). Torsion angles ϕ_3 and ψ_3 for Gly_3 were not varied. The torsion angle ϕ_2 for proline was assumed to be $-60°$ (Pullman and Pullman, 1974). Initially, a scan for energetically preferred values of ϕ_1 was carried out by constructing the ϕ_1-ψ_1 energy surface (Renugopalakrishnan and Urry, in preparation).

The total energy was assumed to consist of Van der Waals, electrostatic, and torsional energies. Van der Waals energy was calculated using Buckingham's exponential type potential function with parameters suggested by Ramachandran and Sasisekharan (1968). Electrostatic energy was calculated using ab initio minimal basis set (STO-3G) net charges for glycyl, alanyl, valyl, prolyl, and *N*-formyl moieties (Renugopalakrishnan et al., 1976b; Renugopalakrishnan and Jordan, 1977). CNDO/2 overlap normalized net charges, taken from Momany et al. (1975), were used for leucyl and phenylalanyl moieties. Net charges for atoms in *N*-formylsarcosinamide were calculated using the CNDO/2 molecular orbital method (Renugopalakrishnan and Urry, unpublished). A dielectric constant of unity was used in the calculations of electrostatic energy. Calculations in one case, namely, *N*-formyl-L-Val-L-Pro-Gly-NH_2 using a dielectric constant of 2 showed no significant difference in the ψ_1-ψ_2 energy surfaces. Torsional energies were calculated using a threefold torsional barrier across C^α-N and C^α-C bonds with barrier heights of 0.6 and 0.2 kcal/mol, respectively (Scott and Scheraga, 1966).

Generation of ψ_1-ψ_2 Energy Surfaces. The tripeptides were frozen in a conformation with ϕ_1, at the energetically most preferred conformation, $\phi_2 = -60°$, $\psi_1 = \pm 180°$, and $\psi_2 = \pm 180°$. ψ_1-ψ_2 energy surfaces were then constructed by varying both the torsional angles ψ_1 and ψ_2 from -180 to $+180$ at 10° intervals. ψ_1-ψ_2 energy surfaces corresponding to the preferred ϕ_1 values for the residues X = Gly, Ala, Leu, Val, and Sar are reported in this paper. ψ_1-ψ_2 energy surfaces for X = Gly, Ala, Leu, Val, and Sar are given in Figures 2 and 3. ψ_1-ψ_2 energy surfaces for some of the other allowed values of ψ_1 were also constructed, and it was found that the overall feature of the ψ_1-ψ_2 energy surface is not particularly sensitive to the value of the torsion angle ψ_1 for the residue X. The most preferred values of ψ_1 and ψ_2 so obtained for the residues X = Gly, Ala, Leu, Val, Phe, and Sar are indicated in Figure 4 using the symbols 1, 2, 3, 4, 5, and 6, respectively. Symbol 7 represents polyproline II with $\psi_1 = 85°$ and $\psi_2 = 85°$ (Arnott and Dover, 1968).

Results

In the tripeptide sequence of the type X_1-Pro_2-Gly_3, hydroxylation by prolyl hydroxylase occurs at the C^γ atom of Pro_2. Hence, it is expected that the C^γ atom of Pro_2 be accessible to the enzymatic approach and interaction. A typical X-Pro-Gly segment undergoing hydroxylation is diagrammatically presented in Figure 1. The torsion angles ϕ_1 and ψ_3 are considered to be less important than ψ_1, ϕ_2, and ψ_2, as they are not in close proximity to the reaction site. Their effect would be to indirectly influence ψ_1 and ψ_2. Torsion angle ϕ_2 is an integral part of the proline ring and is nearly frozen. Accordingly, in a tripeptide sequence, ψ_1 and ψ_2 are of critical importance in determining whether a given peptide sequence presents an optimal conformation for recognition by the enzyme. Because of this, the ψ_1-ψ_2 energy surfaces were generated for the peptides under study (Figures 2 and 3). In the present effort, it was assumed that the tripeptide sequence undergoing hydroxylation was in the trans planar conformation. Although conformational energy calculations were performed on tripeptide units, it is assumed that similar conformations are likely to prevail for the polytripeptide. The results of the conformational energy calculations on the individual peptide sequences are discussed below.

Gly-Pro-Gly. Gly-Pro-Gly has several degenerate minima which are equally probable in the ψ_1-ψ_2 energy surface shown in Figure 2A. The global minimum corresponds to $\psi_1 = \psi_2 = \pm 180°$. There are two high-energy ranges, corresponding to $\psi_2 \simeq -80$–$150°$ and $\psi_2 \simeq -30°$–$50°$ for all values of ψ_1, i.e., $-180°$ to $+180°$. The ψ_1-ψ_2 energy surface was generated taking $\phi_1 = 80°$ for Gly_1 and $\phi_2 = -60°$ for Pro_2.

Ala-Pro-Gly. Introduction of a methyl side chain at residue 1 in the tripeptide may be clearly perceived from the ψ_1-ψ_2 energy surface shown in Figure 2B. The global minimum corresponds to $\psi_1 = 60°$ and $\psi_2 = 120°$ and the energy at $\psi_1 = \psi_2 = 120°$ is 2.20 kcal/mol, relative to the global minimum. Except for the regions extending from $\psi_2 \simeq -160°$ to $-180°$ and $\psi_2 \simeq 100°$ to 180° for $\psi_1 = 180°$ to 180° and a small isolated region with $\psi_1 = -30°$ to 60° with $\psi_2 = -70°$ to $-90°$, the rest of the ψ_1-ψ_2 region is completely forbidden. While generating the ψ_1-ψ_2 energy surface, $\phi_1 = -100°$ for Ala_1 and $\phi_2 = -60°$ for Pro_2 residues were assumed. A perspective of a low-energy conformation is given in Figure 5B.

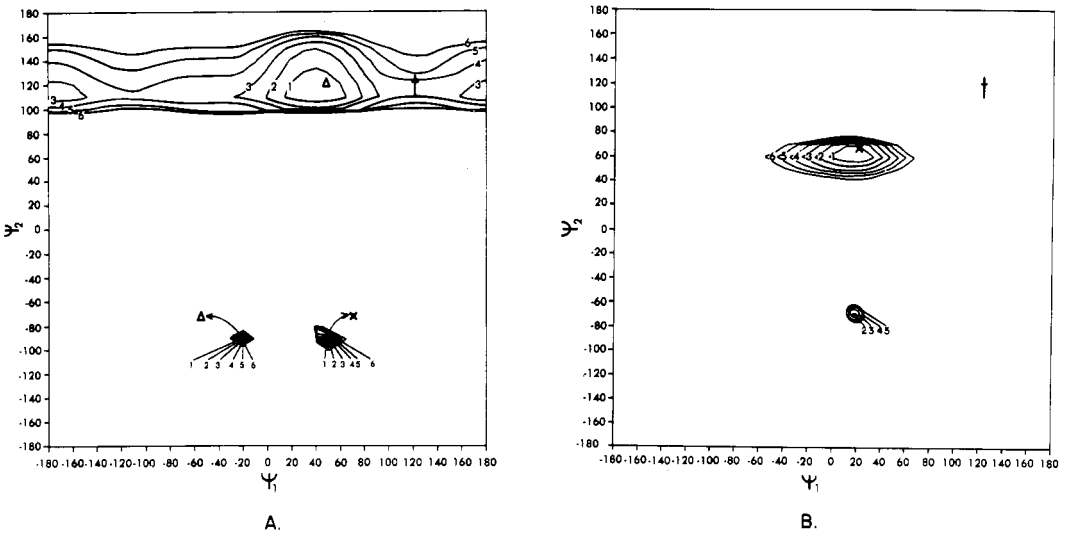

FIGURE 3: ψ_1-ψ_2 energy surfaces in kcal/mol relative to the global minimum marked by the symbol x for: (A) L-Val-L-Pro-Gly with $\phi_1 = -100°$, $\phi_2 = -60°$, $\phi_3 = +180°$, and $\psi_3 = +180°$: (Δ) local minima occurring within the area of hydroxylation. (B) Sar-L-Pro-Gly with $\phi_1 = 100°$, $\phi_2 = -60°$, $\phi_3 = +180°$, and $\psi_3 = +180°$. The symbol † denotes the conformation at $\psi_1 = \psi_2 = 120°$ which corresponds to the ideal conformation of HCO-L-Pro-L-Pro-Gly-NH_2.

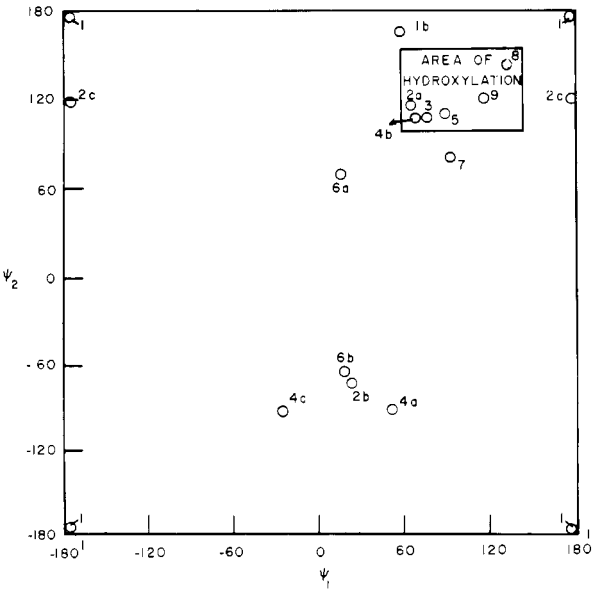

FIGURE 4: Minimum-energy conformations found in the ψ_1-ψ_2 energy surfaces for Gly, Ala, Leu, Val, Phe, and Sar tripeptides are plotted. The tripeptides are designated by the symbols 1-6 for Gly, Ala, Leu, Val, Phe, and Sar, respectively. Notations a, b, and c following the numbers 1-6 denote the global minimum, local minima I and II, respectively. In cases where only global minimum is indicated, notations a, b, c are not given. Symbol 7 represents polyproline II. Peptides that undergo hydroxylation are shown enclosed in a region with ψ_1 ranging from 100 ± 40° and ψ_2 from 130 ± 30°. Symbols 8 and 9 indicate the torsion angles ψ_1 and ψ_2 found by Yonath and Traub in the crystal structure of poly(Pro-Pro-Gly)$_n$ and the calculated values for H-L-Pro-L-Pro-Gly-NH_2.

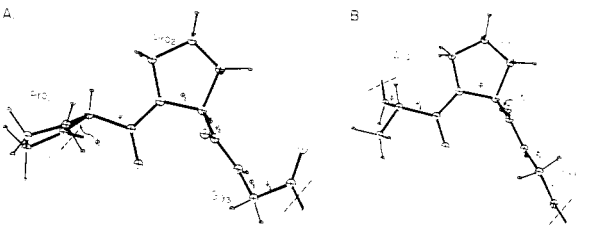

FIGURE 5: (A) A perspective of the minimum-energy conformation of L-Pro-L-Pro-Gly with $\phi_1 = -60°$ and $\psi_1 = 120°$, $\phi_2 = -60°$, $\psi_2 = 120°$, $\phi_3 = -60°$, and $\psi_3 = -170°$. (B) A perspective of a low-energy conformation for -L-Ala-L-Pro-Gly- with $\phi_1 = -100°$, $\psi_1 = 90°$, $\phi_2 = -60°$, $\psi_2 = 120°$, $\phi_3 = +180°$, and $\psi_3 = +180°$.

Leu-Pro-Gly. Leu-Pro-Gly has a pronounced minimum (see Figure 2C) at $\psi_1 = 60°$ to 80° and $\psi_2 = 110°$ to 120°, and the energy at $\psi_1 = \psi_2 = 120°$ is 2.40 kcal/mol relative to the global minimum. Around the global minimum, there are quite a number of local minima within a range of 1 to 2 kcal/mol. In addition, there is a very small isolated region centered around $\psi_1 = 120 \pm 10°$ and $\psi_2 = 120 \pm 10°$ with an energy of approximately 2 to 3 kcal/mol less stable than the global minimum. Torsion angles corresponding to the local minimum are in reasonable agreement with the values of torsion angles ψ_1 and ψ_2 of 153° and 162°, respectively, obtained by Leung and Marsh (1958) for L-Leu-L-Pro-Gly. While generating the ψ_1-ψ_2 energy surface, $\psi_1 = -100°$ for Leu_1 and $\psi_2 = -60°$ for Pro_2 were assumed.

Val-Pro-Gly. The effect of an isopropyl side chain introduces even more drastic changes in the ψ_1-ψ_2 energy surface shown in Figure 3A. The global minimum shifts to the lower right-hand side of the map. The global minimum corresponds to $\psi_1 \simeq 50°$ and $\psi_2 \simeq -90°$, and the energy at $\psi_1 = \psi_2 = 120°$ is 3.60 kcal/mol relative to the global minimum. However, local minima with a range of energies of +1 to +2 kcal/mol, less stable than the global minima, do occur in the region with $\psi_1 = 10$-70°, $\psi_2 = 110$-140°, and at $\psi_1 \simeq 20°$ and $\psi_2 \simeq -98°$. Elsewhere in the ψ_1-ψ_2 energy surface the energies are relatively high and therefore the probability of Val preceding a Pro occurring in these regions is small. While generating the ψ_1-ψ_2 map, a $\psi_1 = -100°$ for Val_1 and $\psi_2 = -60°$ for Pro_2 were assumed.

Phe-Pro-Gly. Phe-Pro-Gly has a broad minimum in the region with $\psi_1 = 60$ to 120° and $\psi_2 = 120$ to 180°. The global minimum occurs at $\psi_1 = 90°$ and $\psi_2 = 110°$, and the energy at $\psi_1 = \psi_2 = 120°$ is 0.60 kcal/mol relative to the global minimum. The region with $\psi_1 = -180$ to +180° and $\psi_2 = -60$ to +160° is completely disallowed.

Sar-Pro-Gly. Sar-Pro-Gly has a pronounced minimum at $\psi_1 = 20°$ and $\psi_2 = 70°$, and to the energy at $\psi_1 = \psi_2 = 120°$ is 14.80 kcal/mol relative to the global minimum (see Figure 3B). Except for a narrow region extending from $\psi_1 = -60°$ to

TABLE I: Hydroxylation of Polytripeptide Models of Collagen by Prolyl Hydroxylase.[a]

polymer	fract of susceptible Pro hydroxylated (%)
$(Pro\text{-}Pro\text{-}Gly)_n$	30.0
$(Gly\text{-}Pro\text{-}Gly)_n$	<2
$(Ala\text{-}Pro\text{-}Gly)_n$	16.0
$(Leu\text{-}Pro\text{-}Gly)_n$	10.2
$(Val\text{-}Pro\text{-}Gly)_n$[b]	5.0
$(Phe\text{-}Pro\text{-}Gly)_n$[c]	<3
$(Sar\text{-}Pro\text{-}Gly)_n$	<2
$(Pro\text{-}Pro\text{-}Pro)_n$	<2

[a] Hydroxylation was carried out as described in the text. All polypeptides had a molecular weight of approximately 4000, on the basis of gel filtration. [b] $(Val\text{-}Pro\text{-}Gly)_n$ was moderately soluble in water. [c] $(Phe\text{-}Pro\text{-}Gly)_n$ was insoluble in water.

+60° and ψ_2 = 40° and 80° and a small region centered around ψ_1 = 20° and ψ_2 = −70°, the rest of the map is disallowed.

Using the above data from the ψ_1–ψ_2 energy surfaces, Figure 4 was constructed, as described under Materials and Methods. For all the peptides that undergo significant hydroxylation, the global minimum occurred within an area designated as the area of hydroxylation. The relative hydroxylation of proline in polytripeptides of comparable molecular weights is given in Table I. Significant amounts of hydroxyproline were synthesized only when $(Ala\text{-}Pro\text{-}Gly)_n$, $(Leu\text{-}Pro\text{-}Gly)_n$, and $(Pro\text{-}Pro\text{-}Gly)_n$ were used as substrates, i.e., points 2a, 3, and 8, respectively (see Figure 4). When $(Gly\text{-}Pro\text{-}Gly)_n$ or $(Sar\text{-}Pro\text{-}Gly)_n$ was used as substrate, only trace amounts of hydroxyproline were synthesized. Low levels of hydroxylation observed with $(Val\text{-}Pro\text{-}Gly)_n$ and much lower levels with $(Phe\text{-}Pro\text{-}Gly)_n$, points 4b and 5 respectively, were attributable to the low solubility of these polymers under the conditions of the reaction and a conformation other than the global minimum. The degree of hydroxylation determined as the ratio of hydroxyproline to total proline in the -Pro-Gly- sequence was greatest for $(Pro\text{-}Pro\text{-}Gly)_n$, followed in order by $(Ala\text{-}Pro\text{-}Gly)_n$ and $(Leu\text{-}Pro\text{-}Gly)_n$. The degree of hydroxylation was consistent with the K_m values for the peptides. When larger molecular weight fractions of $(Leu\text{-}Pro\text{-}Gly)_n$ ranging up to 9200, $(Val\text{-}Pro\text{-}Gly)_n$ up to 25 000, and $(Sar\text{-}Pro\text{-}Gly)_n$ up to 9900 were used, the hydroxylation was better than with lower molecular weight polymers (Rapaka and Bhatnagar, in preparation).

Discussion

Hydroxyproline is synthesized during the posttranslational modification of collagen, and it adds considerable stability to the triple helix. Under the conditions of hydroxylation, (37 °C, pH 7.4–7.8), unhydroxylated collagen exists in random conformations (Rosenbloom et al., 1973; Berg and Prockop, 1973). The triple helical conformation of collagen presents a barrier to the enzymatic reaction (Prockop et al., 1976). The specificity of the enzyme for its polypeptide substrate appears to be largely dependent on the presence of the -Pro-Gly- peptide bond. Our studies indicate that the residue on the N terminal of the proline may regulate the hydroxylation of proline in synthetic polytripeptides. Interactions of the proline ring with the side chains of the residue on its N-terminal side restrict the conformational range of this peptide bond (Schimmel and Flory, 1968; Grathwohl and Wüthrich, 1976; Zimmerman and Scheraga, 1977; Hopfinger and Walton, 1970; Urry, 1976; Rapaka and Urry, 1978; Urry and colleagues, in preparation). It may be postulated that the stereochemical restrictions at the -X-Pro- peptide bond contribute to the regulation of optimal conformational "adjustments", arising as a result of polytripeptide–enzyme interaction. An examination of the conformational features of various tripeptide sequences in relation to their ability to undergo hydroxylation should provide insight into the mechanism of catalysis by prolyl hydroxylase.

When the conformational features of the polytripeptides were compared, it became apparent that in the case of those polypeptides which were hydroxylated the global minima were located within a small area of the ψ_1–ψ_2 plot, defined by ψ_1 = 100 ± 40° and ψ_2 = 130 ± 30°. The polytripeptide $(Pro\text{-}Pro\text{-}Gly)_n$, the best substrate among the polymers examined, has a limited flexibility, and the experimentally determined minimum for the -Pro-Pro-Gly- tripeptide sequence falls within a range of conformations which may be regarded as the area of hydroxylation. This polypeptide appears to be locked in the conformation defined by ψ_1 = 127° and ψ_2 = 148° (Yonath and Traub, 1969). The above values of torsion angles ψ_1 and ψ_2 were obtained by Yonath and Traub from an X-ray study of $(Pro\text{-}Pro\text{-}Gly)_n$. However, our calculations on HCO-L-Pro-L-Pro-Gly-NH_2 predicted a ψ_1 = 120° and ψ_2 = 120°, respectively, for the two residues.

In the case of the tripeptide sequence -Ala-Pro-Gly-, the conformation with the global minimum also occurs within the area of hydroxylation, and this conformation is energetically favored over other conformations. The perspective of a low-energy conformation is given in Figure 5B, and presumably this is the "hydroxylatable" conformation which serves as the "recognition site" for the enzyme and with the above ψ_1 and ψ_2 it bears a striking similarity to the perspective of Pro-Pro-Gly.

The tripeptide sequence Leu-Pro-Gly is a rigid molecule and its conformations with the global minima fall within the designated area of hydroxylation. The solid-state conformation of Leu-Pro-Gly as determined by X-ray crystallography (Leung and Marsh, 1958) is in reasonable agreement with the local minimum calculated.

For the tripeptide sequence Val-Pro-Gly, a low-energy conformation occurs within the area of hydroxylation, with ψ_1 = 60 ± 10° and ψ_2 = 100 ± 10°, although the global minimum of slightly lower energy (<0.2 kcal/mol) occurs outside the area of hydroxylation. The lower levels of hydroxylation in $(Val\text{-}Pro\text{-}Gly)_n$ may be due to the occurrence of two low-energy conformations of approximately equal energy and due to the fact that the global minimum of slightly lower energy conformation is located outside the hydroxylation area. In addition, $(Val\text{-}Pro\text{-}Gly)_n$ has moderate solubility.

The tripeptide sequence Phe-Pro-Gly has a preferred conformation within the area of hydroxylation. However, due to the insolubility of the polymer $(Phe\text{-}Pro\text{-}Gly)_n$, very low levels of hydroxylation were observed. The validity of the concept of permissible conformational range in this case was confirmed by examination of the sequence of the α-1 chain of collagen. In all seven tripeptides of the type Phe-imino-Gly where Phe appeared on the N terminal of the imino residue, the imino residue was invariably Hyp. The high levels of hydroxylation of Phe-Pro-Gly sequences in collagen suggest that if the polytripeptide *were* indeed soluble the tripeptide sequence would present a favorable conformation for interaction with the enzyme.

As evidenced from the contour maps, peptides containing the sequence Gly-Pro-Gly are highly flexible. Due to the inherent flexibility, there are a large number of equally probable low-energy conformations distributed all over the ψ_1–ψ_2 energy surface, both inside and outside the area of hydroxylation, and

the conformations that occur inside the area of hydroxylation are not particularly favored. Hence, the relative number of conformations that occur in the area of hydroxylation are only a small percentage of the total number of low-energy conformations, thus resulting in insignificant levels of hydroxylation.

$(Sar\text{-}Pro\text{-}Gly)_n$ was shown to exist in a non-triple-helical conformation (Ananthanarayanan et al., 1976). The calculated preferred conformations for the tripeptide sequence are outside the area of hydroxylation, and the peptide does not undergo significant hydroxylation. Hence the experimental results are in agreement with those from the predicted conformational-energy calculations. Polyproline II, which may be written as $(Pro\text{-}Pro\text{-}Pro)_n$, has a conformation outside the area of hydroxylation, and it does not get hydroxylated (Prockop and Kivirikko et al., 1969). However, it may be pointed out here that $(Pro\text{-}Pro\text{-}Pro)_n$ is not analogous to an X-Pro-Gly peptide.

It is suggested above that the ideal substrate is $(Pro\text{-}Pro\text{-}Gly)_n$ and hence it is reasonable that it presents an ideal conformation initially to be recognized by the enzyme and following that conformational changes may be induced in the peptide prior to hydroxylation. We also investigated the possibility that after a peptide presents a low energy conformation to be recognized by the enzyme in the designated ψ_1–ψ_2 area, conformational changes may be induced in the peptide to mimic a conformation corresponding to that of Pro-Pro-Gly. To that end, ΔE, i.e., the difference in conformational energy of the tripeptide at the global minimum, arbitrarily set equal to 0 kcal/mol (E_1) and the energy of the same tripeptide at the preferred conformation of Pro-Pro-Gly (E_2) with $\psi_1 = 120°$ and $\psi_2 = 120°$ was calculated (see Table II). The above values of ψ_1 and ψ_2 were taken from our calculations. In other words, this energy difference would be a measure of the conformational energy required by any given tripeptide in order to attain or mimic the ideal conformation assumed by Pro-Pro-Gly. The results are shown in Table II. However, these data are to be interpreted cautiously. As seen in Table II, Sar-Pro-Gly and Pro-Pro-Pro, where ΔE is very high, do not get hydroxylated. For Gly-Pro-Gly, which has several low-energy conformations, and hence ΔE cannot be calculated, also does not undergo hydroxylation. As expected, ΔE is small in all the other peptides that undergo hydroxylation. The order of the relative energies is Phe-Pro-Gly < Ala-Pro-Gly < Leu-Pro-Gly < Val-Pro-Gly. It is in agreement with the experimental data, as the relative degrees of hydroxylation of the polytripeptides follow the order Ala-Pro-Gly > Leu-Pro-Gly > Val-Pro-Gly. From the above data, it is also expected that the sequence Phe-Pro-Gly should be a very good substrate for hydroxylation; however, this fact could not be verified experimentally due to the insolubility of $(Phe\text{-}Pro\text{-}Gly)_n$. The high level of hydroxylation observed for the Pro residue in the Phe-Pro-Gly sequences in collagenous chains is consistent with the results predicted from the conformational-energy calculations.

The involvement of conformational changes in the mechanism of enzyme-catalyzed reactions is well recognized (Citri, 1973). According to the induced-fit model, binding of the substrate to the enzyme active site results in conformational changes which align the reactive and catalytic groups, thus facilitating the reaction (Citri, 1973). In the case of the prolyl hydroxylase reaction, the substrate polypeptide may also undergo conformational changes. This is suggested by the requirement that the substrate be in a specific conformation for interaction to occur between the polypeptide and the enzyme. A conformational change in the polypeptide substrate may help in aligning the prolyl residue to be hydroxylated, in apposition

TABLE II: Conformational Energy Differences, ΔE, in kcal/mol for the Tripeptide Sequences.

tripeptide	E_1 energy	E_2 energy at ideal conformation[a]
Pro-Pro-Gly	0	0
Ala-Pro-Gly	0	2.20
Leu-Pro-Gly	0	2.40
Val-Pro-Gly	0	3.60
Phe-Pro-Gly	0	0.60
Gly-Pro-Gly	0	*c*
Sar-Pro-Gly[b]	0	14.80
Pro-Pro-Pro	0	*d*

[a] Ideal energy conformation was assumed to be that of Pro-Pro-Gly with $\psi_1 = 120°$ and $\psi_2 = 120°$. [b] A trans conformation was assumed. [c] Cannot be estimated due to multiple conformational energy minima. [d] Rigid conformation high energy difference.

to the other two reactants, O_2 and α-ketoglutarate, both of which are presumably bound to appropriate groups on the active site. Alterations in the conformation of the substrate have been considered in explaining the mechanism of lysozyme and other enzymes (Citri, 1973; Koshland and Neet, 1967; Blake et al., 1967; Levitt, 1974; Eyring et al., 1954; Johnson et al., 1974; Jencks, 1976; Dafforn and Koshland, 1971). Labilization of a susceptible group by polarization of an atom pair, leading to a charge separation, is a unique aspect of catalysis. The selective penetration of the substrate into the catalytic cavity followed by the polarization of the susceptible atom pair by a highly localized electric field within the cavity may be driven by forces including conformational energy which may be components of the overall potential energy of the enzyme as proposed by Green (1974). Quantum mechanical calculations suggest that an electrophilic substitution mechanism may be involved in the hydroxylation of proline (Zaharadnick et al., 1971). Polarization at the γ carbon of the susceptible proline residue by mechanisms discussed above would facilitate hydroxylation by electrophilic substitution. Another aspect of induced conformational change in the substrate may relate to the entropic effects which make major contributions to the acceleration of reaction rates by enzymes, utilizing substrate binding forces to act as "entropy traps" (Jencks, 1976; Green, 1974; Lumry, 1974). Such a mechanism has been used to explain the hydrolysis of the peptide bond by chymotrypsin (Lumry, 1974; Lumry and Rajender, 1971; Martinek et al., 1972). The formation of the activated enzyme–substrate complex results in a large loss of entropy. Increasing α-carbon side-chain size increases the negative entropy and decreases the enthalpy of activation and thereby increases the catalytic rate. The differences in the hydroxylatability of different tripeptide sequences may well be related to conformational features in the same way as the conformational changes in peptides relate to their hydrolysis by chymotrypsin. In case of the substrates of chymotrypsin, the α-carbon side chains play a role both in binding and in bond rearrangement steps. It has been proposed (Lumry, 1974; Lumry and Rajender, 1971; Martinek et al., 1972) that these steps involve a binding mechanism and possibly include the "charge-relay" system (Lumry, 1974). The electronic properties of the charge-relay system are conformationally controlled (Llinas and Klein, 1975).

It is interesting to speculate that the hydroxylation of proline in synthetic polytripeptide analogues of collagen may proceed by a mechanism in which conformational changes induced in the substrate contribute to catalysis by placing the three reacting molecular species, namely, the γ carbon of the peptidyl

proline to be hydroxylated, O_2 and α-ketoglutarate, in juxtaposition, and "activating or destabilizing" the substrate γ carbon by polarization of certain bonds, resulting in the increased susceptibility to electrophilic substitution.

References

Ananthanarayanan, V. S., Brahmachari, S. K., Rapaka, R. S., and Bhatnagar, R. S. (1976), *Biopolymers 15,* 707.

Arnott, S., and Dover, S. D., (1968), *Acta Crystallogr. Sect. B 24,* 599.

Berg, R. A., and Prockop, D. J. (1973), *Biochem. Biophys. Res. Commun. 52,* 115.

Bhatnagar, R. S., and Prockop, D. J. (1966), *Biochim. Biophys. Acta 130,* 383.

Bhatnagar, R. S., and Rapaka, R. S. (1974), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 33,* 1596.

Bhatnagar, R. S., and Rapaka, R. S. (1976), in Biochemistry of Collagen, Ramachandran, G.N., and Reddi, A. H., Ed., Plenum Press, New York, N.Y., p 479.

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., and Sarma, V. S. R. (1967), *Proc. R. Soc. (London) Ser. B 167,* 378.

Citri, N. (1973), *Adv. Enzymol. 37,* 397.

Dafforn, A., and Koshland, D. E. Jr. (1971), *Proc. Natl. Acad. Sci. 68,* 2463.

Eyring, H., Lumry, R., and Spikes, J. (1954) in Mechanisms of Enzyme Action, Glass, B., and McElroy, W., Ed., The Johns Hopkins Press, Baltimore, Md.

Grathwohl, C., and Wüthrich, K. (1976), *Biopolymers 15,* 2025.

Green, D. E. (1974), *Ann. N.Y. Acad. Sci. 227,* 6.

Halme, J., Kivirikko, K. I., and Simons, K. (1970), *Biochim. Biophys. Acta 198,* 460.

Hopfinger, A. J., and Walton, A. G. (1970), *Biopolymers 9,* 29.

Howard, J. C., Momany, F. A., Andreatta, F. H., and Scheraga, H. A. (1973), *Macromolecules 6,* 535.

Hutton, J. J., Jr., Margolin, A., Witkop, B., Kurtz, J., Berger, A., and Udenfriend, S. (1968), *Arch. Biochem. Biophys. 125,* 779.

Jencks, W. P. (1976), *Adv. Enzymol. 43,* 219.

Johnson, F. H., Eyring, H., and Stover, B. J. (1974), in The Theory of Rate Processes in Biology and Medicine, Wiley, New York, N.Y., p 1.

Kendrew, J. C., Kline, W., Lisson, S., Miyazawa, T., Nemethy, G., Philips, D. C., Ramachandran, G. N., and Scheraga, H. A. (1970), *Biochemistry 9,* 3671.

Khaled, A. M., Renugopalakrishnan, V., and Urry, D. W. (1976), *J. Am. Chem. Soc. 98,* 7547.

Kivirikko, K. I., Prockop, D. J., Lorenzi, G. P., and Blout, E. R. (1969), *J. Biol. Chem. 244,* 2755.

Kivirikko, K. I., Suga, K., Kishida, Y., Sakakibara, S., and Prockop, D. J. (1971), *Biochem. Biophys. Res. Commun. 45,* 1591.

Koshland, D. E., Jr., and Neet, K. E. (1967), *Annu. Rev. Biochem. 37,* 359.

Leung, Y. C., and Marsh, R. I. (1958), *Acta Crystallogr. 11,* 17.

Levitt, M. (1974), in Peptides, Polypeptides and Proteins, Blout, E. R., Bovey, F. A., Goodman, M., and Lotan, N., Ed., Wiley, New York, N.Y., p 99.

Llinas, M., and Klein, M. P. (1975), *J. Am. Chem. Soc., 97,* 4371.

Lumry, R. (1974), *Ann. N.Y. Acad. Sci. 227,* 46.

Lumry, R., and Rajender, S. (1971), *J. Phys. Chem. 75,* 1387.

Martinek, K., Dorovska, V., Varfolomeyer, S., and Berezin, I. (1972), *Biochim. Biophys. Acta 271,* 80.

Momany, F. A., McGuire, R. F., Burgess, A. W., and Scheraga, N. A. (1975), *J. Phys. Chem. 79,* 2361.

Okada, K., Kikuchi, Y., Kawashiri, Y., and Hiramoto, M. (1972), *FEBS. Lett. 28,* 226.

Piez, K. A. (1976), in Biochemistry of Collagen, Ramachandran, G. N., and Reddi, A. H., Ed., Plenum Press, New York, N.Y., p 1.

Prockop, D. J., Berg, R. A., Kivirikko, K. I., and Uitto, J. (1976), in Biochemistry of Collagen, Ramachandran, G. N., and Reddi, A. H., Ed., Plenum Press, New York, N.Y., p 163.

Prockop, D. J., and Kivirikko, K. I. (1969), *J. Biol. Chem., 244,* 4838.

Pullman, B., and Pullman, A. (1974), *Adv. Protein Chem. 28,* 248.

Ramachandran, G. N., Bansal, M., and Bhatnagar, R. S. (1973), *Biochim. Biophys. Acta 322,* 166.

Ramachandran, G. N., Bansal, M., and Ramakrishnan, C. (1975), *Curr. Sci. 44,* 1.

Ramachandran, G. N., and Sasisekharan, V. (1968), *Adv. Protein Chem. 23,* 283.

Rapaka, R. S., and Bhatnagar, R. S. (1975), *Int. J. Pept. Protein Res. 7,* 119.

Rapaka, R. S., Long, M. M., Urry, D. W., Ryu, R. K. N., and Bhatnagar, R. S. (1978a), submitted for publication.

Rapaka, R. S., Okamoto, K., and Urry, D. W. (1978b), *Int. J. Pept. Protein Res.* (in press).

Rapaka, R. S., and Urry, D. W. (1978), *Int. J. Pept. Protein Res. 11,* 97.

Renugopalakrishnan, V., and Jordan, F. (1977), *Am. Pept. Symp. 5,* 590.

Renugopalakrishnan, V., Khaled, M. A., and Urry, D. W. (1976a), *J. Chem. Soc., Perkin Trans. 2* (in press).

Renugopalakrishnan, V., Nir, S., and Rein, R. (1976b), in Environmental Effects on Molecular Structure and Properties, Pullman, B., Ed., D. Reidel Publishing Co., Dordrecht, The Netherlands, p 109.

Rosenbloom, J., Harsch, M., and Jiminez, J. A. (1973), *Arch. Biochem. Biophys. 158,* 478.

Scatturin, A., Tamburro, A. M., Del Pra, A., and Brodignon, E. (1975), *Int. J. Pept. Protein Res. 7,* 424.

Schimmel, P. R., and Flory, P. R. (1968), *J. Mol. Biol. 34,* 105.

Scott, R. A., and Scheraga, H. A. (1966), *J. Chem. Phys. 45,* 2101.

Sutton, D. W. (1965), in Table of Interatomic Distances and Configurations in Molecules and Ions, Special Publication No. 18 Chemical Society, London, U.K.

Ueki, T., Ashida, T., Kakudo, M., Sasda, U., and Katsuba, Y. (1969), *Acta Crystallogr., Sect. B 25,* 1840.

Urry, D. W. (1976), *Faraday Discuss. Chem. Soc. 61,* 205.

Urry, D. W., Khaled, M. A., Renugopalakrishnan, V., and Rapaka, R. S. (1978), *J. Am. Chem. Soc. 100,* 697.

Woessner, J. F., Jr. (1961), *Arch. Biochem. Biophys. 93,* 440.

Yonath, A., and Traub, W. (1969). *J. Mol. Biol. 43,* 479.

Zahradnick, R., Hobza, P., and Hurych, J. (1971), *Biochim. Biophys. Acta 251,* 314.

Zimmerman, S. S., and Scheraga, H. (1977), *Biopolymers 16,* 811.