THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 50, Issue of December 12, pp. 31441–31446, 1997

Printed in U.S.A.

Positional Preferences of Ionizable Residues in Gly-*X*-*Y* Triplets of the Collagen Triple-helix*

(Received for publication, August 8, 1997, and in revised form, October 3, 1997)

Virginia C. Chan‡, John A. M. Ramshaw§, Alan Kirkpatrick§, Konrad Beck‡, and Barbara Brodsky‡¶

From the ‡Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854-5636 and the §Commonwealth Scientific and Industrial Research Organization, Division of Molecular Science, 343 Royal Parade, Parkville, 3052, Australia

Collagens contain a high amount of charged residues involved in triple-helix stability, fibril formation, and ligand binding. The contribution of charged residues to stability was analyzed utilizing a host-guest peptide system with a single Gly-*X*-*Y* triplet embedded within Ac(Gly-Pro-Hyp)$_3$-Gly-*X*-*Y*-(Gly-Pro-Hyp)$_4$-Gly-Gly-NH$_2$. The ionizable residues Arg, Lys, Glu, and Asp were incorporated into the *X* position of Gly-X-Hyp; in the *Y* position of Gly-Pro-*Y*; or as pairs of oppositely charged residues occupying *X* and *Y* positions. The Gly-*X*-Hyp peptides had similar thermal stabilities, only marginally less stable than Gly-Pro-Hyp, whereas Gly-Pro-*Y* peptides showed a wide thermal stability range (T_m = 30–45 °C). The stability of peptides with oppositely charged residues in the *X* and *Y* positions appears to reflect simple additivity of the individual residues, except when *X* is occupied by a basic residue and *Y* = Asp. The side chains of Glu, Lys, and Arg have the potential to form hydrogen bonds with available peptide backbone carbonyl groups within the triple-helix, whereas the shorter Asp side chain does not. This may relate to the unique involvement of Asp residues in energetically favorable ion pair formation. These studies clarify the dependence of triple-helix stability on the identity, position, and ionization state of charged residues.

The triple-helix is the characteristic domain of all collagens and has also been found as a motif in a variety of host defense proteins such as the complement factor C1q, mannose-binding protein, and the macrophage scavenger receptor (1–4). The conformation of the triple-helix consists of three extended polyproline II-like polypeptide chains that are supercoiled around a common axis (5–8). To accommodate close packing of the three staggered chains, every third residue must be glycine. A high content of imino acids, Pro and Hyp,[1] is also required to stabilize the conformation. These sequence constraints give rise to a (Gly-*X*-*Y*)$_n$ repeating sequence with Gly-Pro-Hyp as the most frequent tripeptide unit (triplet). In this repeating structure, Gly residues are buried and solvent inaccessible, whereas residues in the *X* position are highly exposed to solvent (8, 9). Residues in the *Y* position are somewhat less accessible to solvent than those in the *X* position because of their proximity to the neighboring chain (8–11). As in all proteins, the backbone carbonyl groups as well as the amide groups of residues other than proline and hydroxyproline are available for hydrogen bonding. One direct interchain N–H··O=C hydrogen bond per Gly-*X*-*Y* triplet is formed in the triple-helix, leaving two carbonyl groups of each triplet available for hydrogen bonding. These unsatisfied carbonyl groups are seen to be involved in a highly ordered hydration network in the crystal structure of a peptide with Gly-Pro-Hyp units (8, 12).

All collagen triple-helices have a high content of ionizable residues that are well conserved among different types of collagen (13). Intra- and interchain ion pairs have been suggested to contribute to molecular stability (14–18). Both experimental evidence and theoretical approaches support the involvement of electrostatic interactions in the association of triple-helices to form fibrils and in the binding of various molecules to collagen (19–22). Ionizable residues are asymmetrically distributed with respect to the *X* and *Y* positions of Gly-*X*-*Y* triplets (13). For example, Glu is preferentially found in the *X* position, whereas Arg and Lys are preferentially found in the *Y* position. It has been suggested that these positional preferences may be related to ion pair formation (14, 15).

In this report, a host-guest peptide set is used to assess the contributions of single charged residues and ion-pairs to triple-helix stability. Peptides with glycine as every third residue and a high content of imino acids have been shown to form stable triple-helices (2, 23–25). Recently it has been shown that peptides with one guest Gly-*X*-*Y* triplet introduced into a host structure of Gly-Pro-Hyp units form a stable triple-helix in which the stability varies with the guest triplet (11). The four ionizable residues commonly found in collagens, Glu, Asp, Lys, and Arg, were introduced into the *X* and *Y* positions of a guest triplet individually or in oppositely charged pairs. The thermal stability of the triple-helix was found to depend strongly on the identity of the residue in the *Y* position but shows only a weak dependence on the specific residue in the *X* position. Ion-pairs do not appear to contribute a stabilizing effect in general but appear to be important when there is a basic residue in the *X* position and an Asp residue in the *Y* position.

MATERIALS AND METHODS

Peptide Synthesis—Peptides were purchased from SynPep Corp. (Dublin, CA) or were synthesized in house on an Applied Biosystems 430A peptide synthesizer using the standard FastMoc method on *N*-(9-fluorenyl)methoxycarbonyl-RINK resin. Side chain protection was *tert*-butyl for Hyp, *tert*-butyl ester for Asp and Glu, benzyloxycarbonyl for Lys, and pentamethylchroman-sulfonyl for Arg. Acetylation was performed with acetic anhydride and triethylamine in dimethylformide.

* This work was supported by National Institutes of Health Grant AR 19626 (to B. B.), a National Science Foundation U. S.-Australia International Cooperative Research Grant (to B. B.), and the Australia/USA Bilateral Science Program (to J. A. M. R.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: Dept. of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 675 Hoes Lane, Piscataway, NJ 08854-5636. Tel.: 732-235-4048; Fax: 732-235-4783; E-mail: brodsky@rwja.umdnj.edu.

[1] Standard three- and single-letter nomenclature is used to describe peptide sequences with Hyp and O representing hydroxyproline.

Peptides were purified by reversed phase high pressure liquid chromatography on a C-18 column eluted with a binary gradient of 0.1% trifluoroacetic acid in 0–40% (v/v) water/acetonitrile to >90% purity. Identity and purity of the peptides were confirmed by amino acid analysis using a Waters high pressure liquid chromatography system with ninhydrin detection and by laser desorption mass spectrometry.

Sample Preparation—Peptides were dried *in vacuo* over P_2O_5 for at least 48 h before weighing. Samples for circular dichroism (CD)[2] spectroscopy were prepared at a concentration of 1 mg/ml, and pH measurements were made at 20 °C. Peptides were dissolved in either 100 mM acetic acid (pH 2.7), phosphate buffer adjusted to the same ionic strength with NaCl (pH 7.0) or in 100 mM NaOH (pH 12.2).

Circular Dichroism Spectroscopy—CD spectra were recorded on an Aviv Model 62DS spectropolarimeter with a Hewlett-Packard Peltier thermoelectric temperature controller. Samples were equilibrated at 4 °C for at least 48 h before measuring, and data were recorded in 1-mm path length cells. CD spectra were collected at 2 °C from 260 to 210 nm at 1 nm intervals with a 3 s recording time at each point. For melting curves, the ellipticity was monitored at 225 nm, and the sample temperature was raised from 2 to 70 °C in increments of 0.3 °C over a time span of 15 h. For $(POG)_{10}$, renaturation curves were recorded upon cooling with the same rate in temperature change. T_m values evaluated from de- and renaturation curves varied by less than 0.5 °C, and final ellipticities after cooling differed form the starting value by less than 20% indicating nearly complete reversibility. A decrease of the heating rate by a factor of two, as measured for two host guest peptides, resulted in a decrease of 1–2 °C in T_m suggesting that the chosen conditions represent near equilibrium unfolding.[3]

Calculation of Thermodynamic Parameters—Melting curves were analyzed assuming a two-state, trimer ↔ monomer transition model. Equilibrium ultracentrifugation on similar triple-helical peptides supports a two-state model (26) and the first derivative of all melting curves showed a single gaussian-shaped minimum. All but 2 of the 16 peptides in this study fit a two-state model well (see "Results").

For the two-state model, the folding of three chains to the triple-helix conformation can be described by (27)

$$3u \leftrightarrow n \qquad \text{(Eq. 1)}$$

where u and n denote the unfolded (monomeric) and folded (trimeric) state, respectively. Thus the equilibrium constant K is

$$K = \frac{c_n}{c_u^3} = \frac{F}{3c_0^2(1-F)^3} \qquad \text{(Eq. 2)}$$

where $c_0 = c_u + 3c_n$ is the total peptide concentration and $F = 3c_n/c_0$ is the fraction of folded peptide. F was determined as $F = [\theta - \theta_U(T)]/[\theta_N(T) - \theta_U(T)]$ where θ is the observed ellipticity and θ_N and θ_U are the ellipticities for the native and unfolded forms, respectively, corrected for their temperature dependence by linear extrapolation of the baselines observed at high and low temperatures (27). The melting temperature was taken as the midpoint of the transition curve ($F = 0.5$).

The standard enthalpy, ΔH^0 (van't Hoff), was calculated by fitting the transition curve to (27)

$$K = \exp\left[\frac{\Delta H^0}{RT}\left(\frac{T}{T_m} - 1\right) - \ln(0.75c_0^2)\right] \qquad \text{(Eq. 3)}$$

The standard entropy, ΔS^0, was calculated accordingly as

$$\Delta S^0 = \frac{\Delta H^0}{T_m} - R\ln(0.75c_0^2) \qquad \text{(Eq. 4)}$$

The standard (Gibb's) free energy, ΔG^0, was calculated for $T = 298$ K as $\Delta G^0 = \Delta H^0 - T\Delta S^0$. From repeated experiments on independently prepared samples, the error in the determination of ΔH^0 and thus the other derived thermodynamic parameters is estimated to be less than 10%.

RESULTS

Host-guest peptides of the form $Ac(Gly\text{-}Pro\text{-}Hyp)_3\text{-}Gly\text{-}X\text{-}Y\text{-}(Gly\text{-}Pro\text{-}Hyp)_4\text{-}Gly\text{-}Gly\text{-}NH_2$ have been shown to provide a useful template to evaluate the contribution of individual Gly-*X*-*Y* triplets to triple-helix properties (11). The N and C termini are blocked with acetyl and amide groups, respectively, ensuring that the only ionizable groups would be those introduced in the Gly-*X*-*Y* guest triplet. In this study, the contributions of the four ionizable residues Asp, Glu, Lys, and Arg to triple-helix stability were investigated by inclusion in the guest triplet of host-guest peptides. In the first set of peptides, these residues were placed in the *X* position with Hyp in the *Y* position. In a second set, these residues were introduced in the *Y* position of the guest triplet with Pro in the *X* position. Furthermore, pairs of oppositely charged residues were introduced as *X* and *Y* in the Gly-*X*-*Y* guest triplet, generating a set of another 8 peptides.

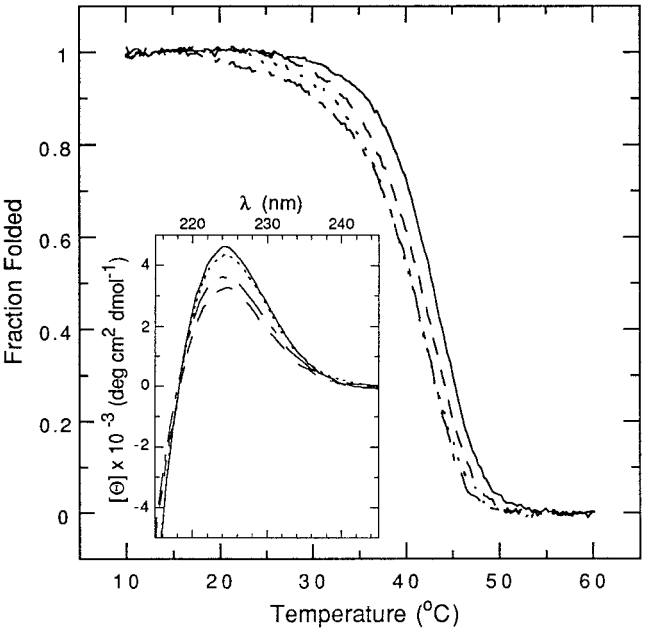

FIG. 1. **Thermal stability of Gly-*X*-Hyp-containing host-guest peptides.** CD thermal transition profiles recorded at 225 nm were normalized to the fraction folded for peptides containing GEO (——), GKO (- - -), GRO (····) and GDO (– – – – –) (pH 7.0; peptide concentration, 1 mg/ml). The *inset* shows corresponding CD spectra recorded at 2 °C that exhibit maxima at about 225 nm.

Effect of Ionizable Residues in the X Position of Gly-X-Hyp Triplets—Ionizable residues were introduced into the *X* position of the guest triplet, with a Hyp in the *Y* position, generating peptides containing Gly-Glu-Hyp, Gly-Asp-Hyp, Gly-Lys-Hyp, and Gly-Arg-Hyp. At low temperature, CD spectra showed the characteristic shape of the triple-helix conformation, with a maximum near 225 nm. Monitoring the ellipticity at 225 nm with increasing temperature results in a sharp transition in all cases (Fig. 1), which has been previously shown to represent a trimer to monomer transition (26). The melting temperatures of the 4 peptides varied slightly, ranging from T_m = 40.1 °C to 42.9 °C (Table I). The small variation in the melting temperature indicates that the identity of the ionizable residue in the *X* position has only a small influence on stability.

To examine the effect of ionization, the thermal stabilities of these peptides were determined at pH 2.7, 7.0, and 12.2 (Table I). Conclusions about ionization effects cannot be drawn for peptides containing Arg as this residue is not in a fully uncharged form even at pH 12.2. All Gly-*X*-Hyp peptides had a higher T_m at neutral, rather than acidic or basic pH. Ionization may increase stability, but additional factors must play a role since the melting temperatures for the Gly-Glu-Hyp and Gly-Asp-Hyp peptides are higher at pH 7.0 than at pH 12.2 even though Glu and Asp are ionized both at neutral and basic pH.

Thermodynamic analyses were carried out to calculate the

[2] The abbreviations used are: CD, circular dichroism; T_m, melting temperature.

[3] N. K. Shah, personal communication.

TABLE I
Thermodynamic parameters of Gly-X-Hyp and Gly-Pro-Y host-guest peptides

CD melting curves were recorded at 225 nm at peptide concentrations of 1 mg/ml in 0.1 M acetic acid (pH 2.7), 0.1 M phosphate buffered saline (pH 7.0) and 0.1 N NaOH (pH 12.2). Corresponding values for peptides containing Gly-Ala-Hyp and Gly-Pro-Ala are included for comparison (11).

pH	Peptide guest	T_m	ΔH^0	ΔS^0	ΔG^0	Peptide guest	T_m	ΔH^0	ΔS^0	ΔG^0
		°C	kJ/mol	kJ/mol K	kJ/mol		°C	kJ/mol	kJ/mol K	kJ/mol
7.4	GAO	39.9	−420	−1.2	−60	GPA	38.3	−360	−1.00	−54.4
2.7	GDO	37.6	−540	−1.6	−61	GPD	33.1	−590	−1.8	−55
7.0		40.1	−520	−1.5	−64		30.1	−550	−2.0	−56
12.2		38.0	−560	−1.6	−62		30.1	−770	−1.7	−48
2.7	GEO	39.7	−470	−1.4	−62	GPE	41.9	−570	−1.7	−70
7.0		42.9	−590	−1.8	−73		39.7	−630	−1.9	−69
12.2		40.9	−670	−2.0	−74		38.5	−640	−1.9	−67
2.7	GKO	40.4	−510	−1.5	−64	GPK[a]	37.1	(−430)	(−1.3)	(−53)
7.0		41.5	−540	−1.6	−67		36.8	(−400)	(−1.2)	(−54)
12.2		38.3	−530	−1.6	−62		38.8	(−440)	(−1.3)	(−59)
2.7	GRO	39.4	−470	−1.4	−61	GPR[a]	45.5	(−560)	(−1.3)	(−74)
7.0		40.6	−520	−1.5	−65		44.5	(−450)	(−1.1)	(−62)
12.2		38.0	−510	−1.5	−61		43.1	(−390)	(−1.1)	(−62)

[a] Fitting of transition curves to Equation 3 shows significant deviations indicating a different unfolding process than the assumed two-state model. The values given in parentheses are estimates assuming the applicability of that model.

enthalpy change ΔH^0 assuming a two-state model. The Gly-*X*-Hyp peptides, where *X* is a charged residue, exhibit a greater enthalpic and smaller entropic contribution compared with Gly-Pro-Hyp and Gly-Ala-Hyp (11) (Table I). This is consistent with the proposal that polar residues in the *X* position of the guest triplet increase the hydrogen bonding potential and the order of the surrounding solvent network.

Effect of Ionizable Residues in the Y position of Gly-Pro-Y Triplets—Asp, Glu, Arg, and Lys residues were introduced into the Y position of the guest triplet, with Pro in the *X* position. These 4 peptides gave characteristic triple-helical CD spectra, and show melting temperatures between 30.1 °C (Gly-Pro-Asp) and 44.5 °C (Gly-Pro-Arg) at pH 7.0 (Fig. 2; Table I). Except for Gly-Pro-Arg, each Gly-Pro-*Y* peptide is less stable than the corresponding Gly-*X*-Hyp peptide. With a range of ΔT_m = 14.4 °C, the melting temperatures for the Gly-Pro-*Y* peptides show a significantly greater variation than observed for the set of Gly-*X*-Hyp peptides (ΔT_m = 1.6 °C).

Thermodynamic analyses show that peptides containing Gly-Pro-Glu and Gly-Pro-Asp have very favorable enthalpic contributions to stability with a correspondingly unfavorable entropy value (Table I). The Gly-Pro-Lys and Gly-Pro-Arg peptides do not fit the two-state model well, despite repeated purification and melts. Nevertheless, enthalpy values were estimated from the closest fitting parameters. The approximate enthalpy of Gly-Pro-Lys and Gly-Pro-Arg is much less favorable than that observed for Gly-Pro-Glu and Gly-Pro-Asp, suggesting some differences in interactions (Table I). Comparison of the thermal stability of each Gly-Pro-*Y* peptide at pH 2.7, 7.0, and 12.2 suggests that ionization results in destabilization for *Y* = Glu, Asp, and Lys, whereas the effect of ionization cannot be estimated for *Y* = Arg.

Effect of Oppositely Charged Residues in the X and Y Positions of Gly-X-Y Triplets—Peptides were synthesized with Gly-*X*-*Y* guest triplets containing oppositely charged residues in positions *X* and *Y*, providing the eight combinations Gly-Asp-Lys, Gly-Asp-Arg, Gly-Glu-Lys, Gly-Glu-Arg, Gly-Lys-Asp, Gly-Lys-Glu, Gly-Arg-Asp, and Gly-Arg-Glu. Thermal stability measurements were carried out to investigate whether the presence of oppositely charged residues in adjacent positions modulates the contribution of individual residues. At low temperature, the 8 peptides formed stable triple-helices as indicated by positive maxima at 225 nm in the CD spectra. All peptides showed sharp transition curves, and at pH 7.0, the T_m was within the range of 30.9 °C (Gly-Asp-Lys) to 40.4 °C (Gly-Glu-Arg) (Table II).

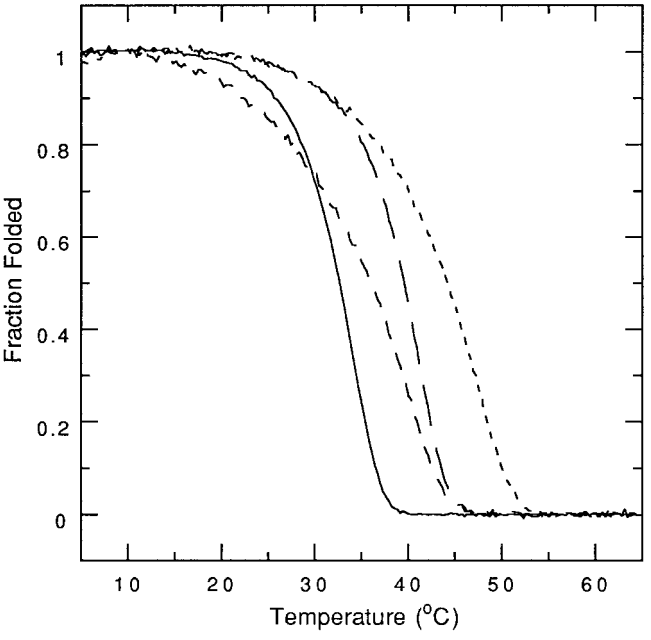

FIG. 2. **Thermal stability of Gly-Pro-*Y*-containing host-guest peptides.** CD thermal transition profiles recorded at 225 nm were normalized to the fraction folded for peptides containing (from left to right at *F* = 0.5) GPD, GPK, GPE, and GPR. An extended range in stabilities is evident and a lower degree of cooperativity observed for GPK and GPR, indicating lower transition enthalpies.

Comparison of the data indicates that a change in one residue can influence stability. For example, Gly-Glu-Lys is more stable than Gly-Asp-Lys (ΔT_m = 4 °C). Reversal of the identity of the residues in the *X* and *Y* positions also affects the melting temperature. For instance, the T_m of the peptide with the guest triplet Gly-Glu-Arg is more stable by 6 °C than that containing Gly-Arg-Glu. For two reversed pairs (Gly-Glu-Arg/Gly-Arg-Glu, Gly-Asp-Arg/Gly-Arg-Asp), the peptides with acidic residues in the *X* position and basic residue in the *Y* position were more stable, whereas Gly-Lys-Asp was found to be more stable than Gly-Asp-Lys. The stability of the Gly-Lys-Glu-containing peptide is essentially the same as that of the corresponding peptide with reversed charge distribution.

TABLE II

Thermodynamic parameters of Gly-X-Y host-guest peptides containing oppositely charged residues

CD melting curves were recorded at 225 nm at peptide concentrations of 1 mg/ml. Corresponding values for peptides containing Gly-Pro-Hyp and Gly-Ala-Ala are included for comparison (11).

pH	Peptide guest	T_m	ΔH^0	ΔS^0	ΔG^0	Peptide guest	T_m	ΔH^0	ΔS^0	ΔG^0
		°C	kJ/mol	kJ/mol K	kJ/mol		°C	kJ/mol	kJ/mol K	kJ/mol
7.4	GPO	44.5	−350	−1.0	−60	GAA	29.3	−450	−1.4	−46
2.7	GDK	26.5	−600	−1.9	−42	GKD	30.5	−770	−2.4	−53
7.0		30.9	−520	−1.6	−49		35.8	−720	−2.2	−64
12.2		29.9	−490	−1.5	−47		30.2	−720	−2.3	−52
2.7	GDR	33.4	−550	−1.7	−54	GRD	28.8	−720	−2.3	−49
7.0		37.1	−580	−1.7	−62		35.0	−720	−2.2	−63
12.2		34.4	−460	−1.3	−53		31.9	−630	−1.9	−54
2.7	GEK	29.5	−490	−1.5	−46	GKE	36.5	−620	−1.9	−62
7.0		35.0	−590	−1.8	−58		35.3	−630	−1.9	−60
12.7		33.1	−530	−1.6	−53		31.6	−680	−2.1	−54
2.7	GER	37.3	−530	−1.6	−60	GRE	35.0	−630	−1.9	−60
7.0		40.4	−520	−1.5	−64		33.8	−680	−2.1	−59
12.2		39.1	−510	−1.5	−62		32.2	−710	−2.2	−56

To investigate the effect of the ionization state, the thermal stabilities were compared at pH 2.7, 7.0, and 12.2 (Table II). All peptides except Gly-Lys-Glu and Gly-Arg-Glu were most stable at pH 7.0, the pH range in which residues in both the *X* and *Y* positions are ionized. The maximal stability at neutral pH is consistent with the formation of stabilizing ion-pairs. The thermodynamic parameters calculated for these peptides are summarized in Table II. The enthalpy was a stabilizing feature at all pH values, whereas the entropy was unfavorable compared with the host peptide containing only Gly-Pro-Hyp triplets.

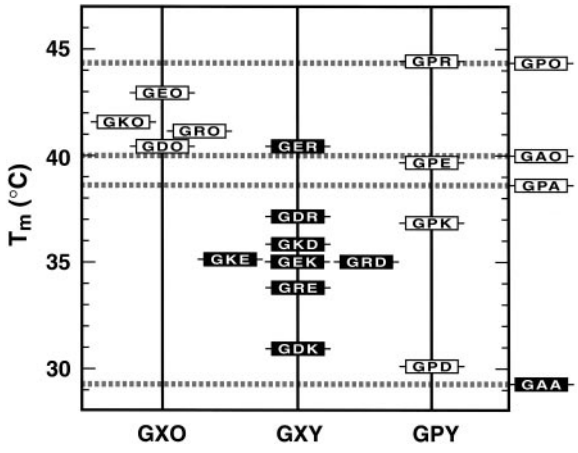

FIG. 3. **Comparison of thermal stabilities of host-guest peptides containing ionizable residues.** Melting temperatures determined at pH 7.0 are plotted in groups for peptides containing G*X*O, G*XY*, and G*PY* where *X* and *Y* denote ionizable residues. Corresponding values for peptides containing GPO, GAO, GPA and GAA are shown on the *right* axis for comparison (11). G*X*O peptides with ionizable *X* residues are more stable than the GAO-containing peptide, suggesting the stabilizing feature of charged residues in the *X* position. Two of the G*PY* peptides are less stable and two are more stable than the GPA containing peptide, indicating the strong dependence of stability on the identity of the charged residue in the *Y* position. G*XY* peptides are all more stable than the GAA peptide, suggesting a positive contribution to triple-helix stability when adjacent *X* and *Y* positions are occupied by oppositely charged residues.

DISCUSSION

This investigation uses host-guest peptides to determine the propensity of ionizable residues for the triple-helix conformation in the different positions of the Gly-*X*-*Y* repeating unit. Previous studies have used host-guest peptides as well as leucine zipper peptide models to study the effect of ionizable residues in the coiled-coil α-helix (28, 29), which, besides the triple-helix, is the other basic protein conformation of fibrous proteins consisting of supercoiled helical subunits (30–32). Interpretation of the results on the coiled-coil α-helix requires consideration of interactions of charged side chains with the intrinsic helix dipole, as well as the potential of oppositely charged ionizable residues to form internal ion-pairs if in appropriate positions (33–35). In contrast to the α-helix, in the triple-helix no net dipole results from the peptide backbone as the amide and carbonyl groups are perpendicular to the triple-helix axis and are distributed in a symmetrical fashion (36). In the absence of interactions with an intrinsic dipole, ionizable residues in the triple-helix can participate in ion-pair formation or in hydrogen bonding to unsatisfied backbone carbonyl groups (37). The polar side chains may also participate in the highly ordered hydration network that has been found for the collagen triple-helix (12).

The comparison of the four Gly-*X*-Hyp host-guest peptides with previous studies indicates that ionizable residues in the *X* position are only marginally less favorable than proline in that position and somewhat more stable than *X*=Ala (11) (Fig. 3). The favorable enthalpy values for these Gly-*X*-Hyp peptides relative to Gly-Pro-Hyp and Gly-Ala-Hyp suggest participation of the side chains in hydrogen bonding. Model building shows that Lys, Arg, and Glu side chains can form hydrogen bonds with available backbone carbonyl groups within the same chain or with neighboring chains, whereas the shorter side chain of Asp cannot (Fig. 4). The polar groups of all four side chains are fully accessible to solvent. The insensitivity of thermal stability to the identity of the residue in the *X* position suggests that either solvent exposure or backbone hydrogen bonding results in favorable enthalpic stabilization. The less favorable entropy values for the Gly-*X*-Hyp set relative to Gly-Pro-Hyp and Gly-Ala-Hyp is consistent with increased conformational constraints on the more flexible ionizable side chains in the triple-helical state or increased order of hydration. The charged forms of these peptides are more stabilizing than the uncharged forms. The lack of any destabilizing effect of having the same charges in the *X* positions of three chains in the same molecule must result from favorable side chain orientation in the triple-helix, pointing outward in near optimal directions for minimizing charge repulsion (Fig. 4*A*).

The broad range of thermal stability observed when the ionizable residues are in the *Y* position contrasts with the narrow range noticed for the *X* position (Fig. 3). The peptide containing Arg has the highest melting temperature, followed

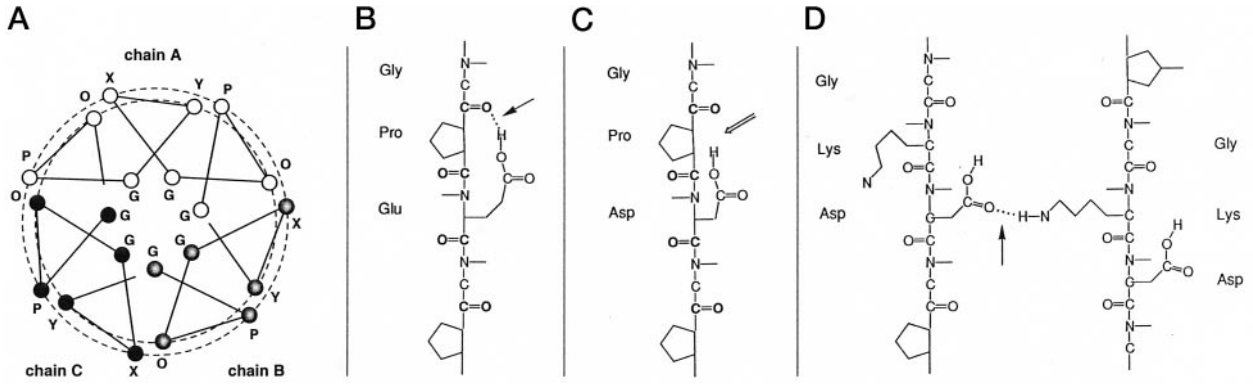

FIG. 4. **Schematics showing the steric positions of ionizable residues in the collagen triple-helix.** A, a cross-section of the 7_2 triple-helix, as found in the triple-helix crystal structure (8), illustrates the more solvent-exposed character of residues in position *X* of a G*XY* triplet when compared with position *Y*. As a result of the one-residue axial stagger of the individual chains within the triple-helix, the *Y* residue in chain A is at the same axial level and sterically close to residue *X* of chain B. *B*, a single chain of a GPE-containing peptide depicts the ability of the glutamic acid side chain to form an intrachain hydrogen bond to an unsatisfied carbonyl group (←), which may be in part the basis for the stabilizing nature of Glu in the *Y* position. *C*, the shorter side chain of the Asp residue, shown in a single chain of a GPD-containing peptide, is not able to undergo a corresponding interaction (⇐) with the peptide backbone. *D*, two staggered GKD-containing peptide chains of the triple-helix show that the basic residue in the *X* position is at the same axial level and sufficiently close to allow electrostatic interactions with the Asp residue in the *Y* position of neighboring chains (↑). This ion-pair interaction is the likely basis for the favorable contributions to triple-helix stability observed for peptides containing GRD and GKD.

in order by Glu, Lys, and Asp, and the 14 °C spread indicates a strong dependence on the identity of the residue. The most stable peptide containing Gly-Pro-Arg is as stable as the host peptide containing Gly-Pro-Hyp. A detailed investigation of a series of peptides containing Gly-Pro-Arg, Gly-Pro-Lys, Gly-Pro-*homo*Arg, and the inclusion of multiple differently spaced Gly-Pro-Arg triplets indicates that the high stability relates to the guanidinium group of Arg and its interaction with the peptide backbone (38). The least stable guest triplet, Gly-Pro-Asp, has a considerably lower T_m than Gly-Pro-Ala (Fig. 3). All polar groups in the *Y* position are substantially accessible to solvent (9), and factors that may influence the relative stability of the 4 peptides in the Gly-Pro-*Y* set include the number of hydrogen bonding sites to solvent or backbone carbonyls, relative side chain flexibility, and steric interactions of the side chain in the triple-helix. Arg has many potential hydrogen bonding sites and can hydrogen bond with backbone carbonyls (37), which may account in part for its high stability. Modeling suggests the low stability of Asp may relate to the inability of its shorter side chain to form hydrogen bonds with backbone carbonyls, in contrast with the side chains of Arg, Lys, and Glu. In addition, unfavorable steric interactions with the neighboring chain lead to restricted rotation for the shorter Asp side chain. It is difficult to correlate possible stabilizing features of the Gly-Pro-*Y* peptide group with thermodynamic parameters as the transition curves of two of its members, Gly-Pro-Lys and Gly-Pro-Arg, do not fit well to the function expected for a two-state model, suggesting a more complex unfolding process. In contrast to observations on the *X* position, the charged form of ionizable residues appears to be slightly destabilizing in the *Y* position, suggesting charge repulsion between the three residues as a result of *Y* residues being less solvent accessible (9) and sterically closer to the neighboring chain (10, 11).

Gly-*X*-*Y* triplets with oppositely charged residues in the *X* and *Y* positions comprise about 8% of the triplets in fibril forming collagens. The large majority of such triplets have the acidic residue in the *X* position and the basic residue in the *Y* position (13, 39). Thus, the host-guest peptides containing Gly-Glu-Arg, Gly-Asp-Arg, Gly-Glu-Lys, and Gly-Asp-Lys model the commonly found triplets of the form Gly-(acidic)-(basic). At neutral pH, these host-guest peptides are more stable by 3–5 °C than at acidic pH (Table I). Of the four host-guest peptides, which include the less frequently observed triplets Gly-X^+-Y^-, only the 2 peptides with a basic residue in the *X* position and Asp in the *Y* position show a substantial stability increase at neutral pH, a feature not seen when Glu is in the *Y* position.

Model building has shown that for all pairs of oppositely charged residues in Gly-*X*-*Y* positions the *X* residue of one chain is at the same axial height and sterically near the *Y* residue of a neighboring chain (14, 15) (Fig. 4). The increased stability of Gly-X^--Y^+ and Gly-X^+-Y^- peptides relative to Gly-Ala-Ala (Fig. 3) and the occurrence of maximal stability at neutral pH for six out of the eight peptides in this set could be solely due to the stabilizing nature of the individual basic and acidic residues, but an additional contribution could also result from ion-pair formation. Several analyses were carried out on the stability data of peptides with oppositely charged pairs in one triplet to evaluate whether such charge pairs stabilize the triple-helix. As a first approach, the increase in stability at neutral *versus* acidic pH of Gly-X^--Y^+ and Gly-X^+-Y^- peptides was compared with that observed for the individual charged components in separate peptides.

$$\Delta G^0(\mathrm{GXY}_{\mathrm{pH7}}) - \Delta G^0(\mathrm{GXY}_{\mathrm{pH3}}) = [\Delta G^0(\mathrm{GXO}_{\mathrm{pH7}}) - \Delta G^0(\mathrm{GXO}_{\mathrm{pH3}})] + [\Delta G^0(\mathrm{GPY}_{\mathrm{pH7}}) - \Delta G^0(\mathrm{GPY}_{\mathrm{pH3}})] \quad \text{(Eq. 5)}$$

and

$$T_m(\mathrm{GXY}_{\mathrm{pH7}}) - T_m(\mathrm{GXY}_{\mathrm{pH3}}) = [T_m(\mathrm{GXO}_{\mathrm{pH7}}) - T_m(\mathrm{GXO}_{\mathrm{pH3}})] + [T_m(\mathrm{GPY}_{\mathrm{pH7}}) - T_m(\mathrm{GPY}_{\mathrm{pH3}})]. \quad \text{(Eq. 6)}$$

Both T_m and ΔG^0 as measures of stability are closely correlated for this set of peptides, but due to the uncertainties in the evaluation of ΔG^0 for Gly-Pro-Arg and Gly-Pro-Lys (see above), the application of T_m values leads to more consistent results. The stability of the Gly-Glu-Lys-containing peptide, for example, indicates that the pH related changes ($\Delta\Delta G^0 = -11.8$ kJ/mol, $\Delta T_m = 5.5$ °C) can be explained almost completely by the favorable effects of neutral pH relative to pH 3 on the individual Glu residue in the *X* position of Gly-Glu-Hyp ($\Delta\Delta G^0 = -11.3$ kJ/mol; $\Delta T_m = 3.2$ °C) and on the individual Lys in the *Y* position of Gly-Pro-Lys ($\Delta\Delta G^0 = -1.3$ kJ/mol, $\Delta T_m = -0.3$ °C) (Tables I and II). Such analyses for all 8 peptides indicates nearly no ion-pair stabilization except for the 2 peptides containing Gly-Lys-Asp and Gly-Arg-Asp.

A second approach is to compare the effect of the reversal of

charged residues in Gly-*X*-*Y* pairs on stability with that expected on the basis of the relative stabilities of individual residue components in G*X*O and GP*Y* triplets. One illustration is shown in Equation 7.

$$T_m(\mathrm{GER}) - T_m(\mathrm{GRE}) = [T_m(\mathrm{GEO}) - T_m(\mathrm{GRO})] + [T_m(\mathrm{GPR}) - T_m(\mathrm{GPE})] \quad \text{(Eq. 7)}$$

Such calculations show that the relative stabilities of Gly-Glu-Arg *versus* Gly-Arg-Glu and Gly-Glu-Lys *versus* Gly-Lys-Glu were close to those expected on the basis of individual residue stabilities. In contrast, Gly-Lys-Asp was much more energetically favorable than expected compared with Gly-Asp-Lys, as was Gly-Arg-Asp over Gly-Asp-Arg.

Consideration of the pH dependence of individual components, and analysis of the additivity of individual residues in the *X* and *Y* positions both indicate that ion-pair contribution to stability is rather weak for all Gly-*X*-*Y* charged pairs except Gly-Arg-Asp and Gly-Lys-Asp where the contribution is substantial. Recent studies suggest that the contribution of salt bridges on the surface of proteins to stability is small. The high dielectric constant of the solvent, the energetic cost of desolvating charges, and the unfavorable entropic result of constraining side chain positions to form a salt bridge have been suggested as possible reasons (29, 40). Both the *X* and *Y* positions where charged residues may occur in the triple-helix are highly accessible to solvent (Fig. 4*A*), so that ion-pairs in the triple-helix would be located on the protein surface. The data on 6 of the 8 triple-helix host-guest peptides suggest weak energetic contributions when oppositely charged residues are adjacent in Gly-*X*-*Y* triplets and sterically close in neighboring chains. These include the double-charged triplets most commonly found in collagens, namely Gly-Glu-Arg and Gly-Glu-Lys. This relatively frequent occurrence is unlikely to be related to triple-helix stability but probably relates to higher level interactions, such as fibril formation and ligand binding. However, our data indicate significant stabilization for Gly-(basic)-Asp triplets. The inability of Asp to form hydrogen bonds to carbonyl groups and its steric hindrance in the *Y* position may create a situation where ion-pair formation has a greater stabilizing effect than in the other cases (Fig. 4). Although triplets with acidic residues in the *Y* position are relatively rare, in most of them this position is occupied by Asp.

The highly soluble, stable host-guest triple-helical peptides are a valuable means for evaluating basic principles for the sequence dependent nature of triple-helix stability in collagen and other proteins. The position-dependent differences observed for the ionizable residues highlight the inherent non-equivalence of the *X* and *Y* positions in the triple-helix and suggest that changes in the *X* position have a relatively minor effect on stability in contrast to changes in the *Y* position, as it was also observed for hydrophobic residues (11). The surprising observation that Gly-(basic)-Asp triplets form stabilizing ion-pairs, whereas Gly-(basic)-Glu and Gly-(basic)-(acidic) triplets do not, indicates the need to look for interactions relating to *X* and *Y* residues within one triplet as well as extending host-guest peptide studies to include adjacent triplets.

Acknowledgments—We thank Drs. Naina K. Shah and Wei Yang for experimental guidance in the early stages of this project and for helpful discussions. We also appreciate the productive meetings and talks with Dr. Jean Baum.

REFERENCES

1. Brodsky, B., and Shah, N. K. (1995) *FASEB J.* **9,** 1537–1546
2. Brodsky, B., and Ramshaw, J. A. M. (1997) *Matrix Biol.* **15,** 545–554
3. Hoppe, H. J., and Reid, K. B. M. (1994) *Protein Sci.* **3,** 1143–1158
4. Kadler, K. (1995) *Protein Profile* **2,** 491–619
5. Ramachandran, G. N., and Kartha, G. (1955) *Nature* **176,** 593–595
6. Rich, A., and Crick, F. H. C. (1961) *J. Mol. Biol.* **3,** 483–506
7. Fraser, R. D. B., MacRae, T. P., and Suzuki, E. (1979) *J. Mol. Biol.* **129,** 463–481
8. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) *Science* **266,** 75–81
9. Jones, E. Y., and Miller, A. (1991) *J. Mol. Biol.* **218,** 209–219
10. Bansal, M., and Ramachandran, G. N. (1978) *Int. J. Pept. Protein Res.* **11,** 73–81
11. Shah, N. K., Ramshaw, J. A. M., Kirkpatrick, A., Shah, C., and Brodsky, B. (1996) *Biochemistry* **35,** 10262–10268
12. Bella, J., Brodsky, B., and Berman, H. M. (1995) *Structure* (*Lond.*) **3,** 893–906
13. Hofmann, H., Fietzek, P. P., and Kühn, K. (1980) *J. Mol. Biol.* **141,** 293–314
14. Salem, G., and Traub, W. (1975) *FEBS Lett.* **51,** 94–99
15. Traub, W., and Fietzek, P. P. (1976) *FEBS Lett.* **68,** 245–249
16. Katz, E. P., and David, C. W. (1990) *Biopolymers* **29,** 791–798
17. Katz, E. P., and David, C. W. (1992) *J. Mol. Biol.* **228,** 963–969
18. Venugopal, M. G., Ramshaw, J. A. M., Braswell, E., Zhu, D., and Brodsky, B. (1994) *Biochemistry* **33,** 7948–7956
19. Hulmes, D. J. S., Miller, A., Parry, D. A. D., Piez, K. A., and Woodhead-Galloway, J. (1973) *J. Mol. Biol.* **79,** 137–148
20. Li, S. T., Golub, E., and Katz, E. P. (1975) *J. Mol. Biol.* **98,** 835–839
21. Wallace, D. G. (1990) *Biopolymers* **29,** 1015–1026
22. Brown, E. M., Chen, J. M., and Feairheller, S. H. (1993) *J. Am. Leather Chem. Assoc.* **88,** 2–11
23. Sakakibara, S., Inouye, K., Shudo, K., Kishida, Y., Kobayashi, Y., and Prockop, D. J. (1973) *Biochim. Biophys. Acta* **303,** 198–202
24. Heidemann, E., and Roth, W. (1982) *Adv. Polym. Sci.* **43,** 143–203
25. Fields, G. B., and Prockop, D. J. (1996) *Biopolymers* **40,** 345–357
26. Long, C. G., Braswell, E., Zhu, D., Apigo, J., Baum, J., and Brodsky, B. (1993) *Biochemistry* **32,** 11688–11695
27. Engel, J., Chen, H. T., Prockop, D. J., and Klump, H. (1977) *Biopolymers* **16,** 601–622
28. O'Neil, K. T., and DeGrado, W. D. (1990) *Science* **250,** 646–651
29. Lumb, K. J., and Kim, P. S. (1995) *Science* **268,** 436–439
30. Brodsky, B. (1990) in *Protein Folding* (Gierasch, L. M., and King, J., eds), pp. 55–62, American Association for the Advancement of Science, Washington
31. Cohen, C., and Parry, D. A. D. (1990) *Proteins* **7,** 1–15
32. Lupas, A. (1996) *Trends Biochem. Sci.* **21,** 375–382
33. Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., and Baldwin, R. L. (1987) *Nature* **326,** 563–567
34. Merutka, G., and Stellwagen, E. (1991) *Biochemistry* **30,** 1591–1594
35. Scholtz, J. M., Qian, H., Robbins, V. H., and Baldwin, R. L. (1993) *Biochemistry* **32,** 9668–9676
36. Wada, A. (1976) *Adv. Biophys.* **9,** 1–63
37. Vitagliano, L., Nemethy, G., Zagari, A., and Scheraga, H. A. (1993) *Biochemistry* **32,** 7354–7359
38. Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A. M., and Brodsky, B. (1997) *J. Biol. Chem.* **272,** 28837–28840
39. Doyle, B. B., Hukins, D. W. L., Hulmes, D. J. S., Miller, A., Rattew, C. J., and Woodhead-Galloway, J. (1974) *Biochem. Biophys. Res. Commun.* **60,** 858–864
40. Waldburger, C. D., Schildbach, J. F., and Sauer, R. T. (1996) *Nat. Struct. Biol.* **2,** 122–128