Gly-Gly-Containing Triplets of Low Stability Adjacent to a Type III Collagen Epitope[†]

Naina K. Shah,[‡] Meena Sharma,[‡] Alan Kirkpatrick,[§] John A. M. Ramshaw,[§] and Barbara Brodsky*,[‡]

Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, 675 Hoes Lane, Piscataway, New Jersey 08854-5636, and CSIRO, Division of Biomolecular Engineering, 343 Royal Parade, Parkville, 3052, Australia

Received December 27, 1996; Revised Manuscript Received March 14, 1997[⊗]

ABSTRACT: Collagens, in addition to their structural role in the extracellular matrix, possess a number of functional binding domains. In this study, the binding to collagen of a monoclonal antibody is used as a model to define the molecular features involved in triple-helix interactions with other proteins. Here we report the thermal stability of an overlapping set of triple-helical peptides that includes the epitope recognized by a monoclonal antibody to type III collagen. Although the sequences of these peptides are very closely related, by a translation of a single triplet along the collagen chain, substantial variations in the melting temperatures were observed. These variations in thermal stability could not be readily explained by differences in imino acid content, or in numbers of charged or hydrophobic residues. The results indicate that Gly-Gly-Y triplets, which are adjacent to the epitope, have a strong influence in reducing the thermal stability of triple-helical peptides. Further studies, which were carried out on a set of "host-guest" triple-helical peptides containing different Gly-Gly-Y guest triplets, confirm the destabilizing effect of such tripeptides. The presence of Gly-Gly-Y triplets may play an important role in specific functions of type III collagen by modulating the local triple-helical structure or dynamics.

Collagens are triple-helix-containing structural proteins in the extracellular matrix, with at least 19 distinct genetic types now identified (Brown & Timpl, 1995; Prockop & Kivirriko, 1995). The most abundant and well-characterized collagens are members of the fibril-forming family, types I, II, and III, which are found in characteristic fibrils with an axial period of $D = 65-67$ nm in connective tissues. The different types are each found with a tissue-specific distribution. For instance, type I is the major collagen in skin, tendon, and bone, while type III collagen is found, together with type I collagen, in extensible tissues such as skin and blood vessels. Type I collagen is a heterotrimer, consisting of two distinct types of chains, while type III collagen is a homotrimer.

The distinguishing structural feature of collagens is the presence of a triple-helix domain. The triple-helix conformation consists of three polypeptide chains, each in an extended polyproline II-like helix, which are twisted about a common axis (Rich & Crick, 1961; Ramachandran, 1967; Fraser et al., 1979). The close packing of the three chains generates a constraint for glycine as every third residue, leading to amino acid sequences of the form $(Gly\text{-}X\text{-}Y)_n$. The imino acid proline has steric constraints that stabilize the extended nature of the polyproline II-like chains. When proline is incorporated into the Y position of Gly-X-Y triplets, it is posttranslationally modified to hydroxyproline, which confers additional stability by participating in a highly ordered water network (Bella et al., 1994, 1995). Triplets of the form Gly-Pro-Hyp are the most stabilizing and the most frequently found in the collagen triple-helix. Inclusion of Gly-Pro-Hyp triplets is important in the design of stable triple-helical peptides.

The collagen triple-helix is a protein motif which forms a rod-like structure and has the capacity to self-associate and to interact with other molecules (Brodsky & Shah, 1995). For example, in addition to forming D-periodic fibrils, type III collagen has a unique collagenase cleavage site and has a defined $\alpha 2\beta 1$ integrin binding site necessary for platelet binding (Fields, 1995; Brodsky & Ramshaw, 1996). Within the long triple-helix domain of type III collagen (1026 residues), these are distinct regions that are specifically recognized by other proteins. Recognition will be mediated by residues in the X and Y positions of the repeating Gly-X-Y triplets, all of which are substantially exposed to solvent in the triple-helical configuration (Jones & Miller, 1991).

The interaction explored in this investigation is the binding of a monoclonal antibody (MAb)[1] to the triple-helical region of type III collagen. The binding of collagen by a MAb is a protein−protein interaction of high specificity and presents an opportunity to define the molecular features involved in triple-helix interactions with another macromolecule. Triple-helical domains of mammalian collagens generally show little immunogenicity. However, a number of murine MAbs were raised against the triple-helical region of human type III collagen (Werkmeister & Ramshaw, 1991). These MAbs were conformation-specific, reacting only with this collagen in its triple-helical form and not with denatured, single-

[†] This work has been supported by NIH Grant AR 19626 (B.B.) and NSF U.S.−Australia International Cooperative Research Grant (B.B.) and the Australia/USA Bilateral Science Program (J.A.M.R.).

* Author to whom correspondence should be addressed. Phone: (908) 235 4048. Fax: (908) 235 4783.

[‡] University of Medicine and Dentistry of New Jersey.

[§] CSIRO, Division of Biomolecular Engineering.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1997.

[1] Abbreviations: MAb, monoclonal antibody; CD, circular dichroism; T_m, melting temperature. Standard one-letter and three-letter abbreviations are used for the common amino acids; hydroxyproline is denoted by O (one-letter code) and Hyp (three-letter code). Ac indicates acetyl.

S0006-2960(96)03146-7 CCC: $14.00

collagen chains. Overlapping triple-helical peptides were used to define the epitope recognized by one of these MAbs and identified the sequence Gly-Leu-Ala-Gly-Ala-Hyp-Gly-Leu-Arg in a triple-helical conformation as being required for binding (Glattauer et al., 1997). Here we report studies characterizing the thermal stability of this overlapping set of triple-helical peptides that was used previously to characterize this type III collagen epitope. These stability studies indicate that the presence of Gly-Gly-Y triplets adjacent to the epitope site has a strong influence in lowering the thermal stability of the triple-helix. Further studies, carried out on a "host-guest" set of triple-helical peptides containing different Gly-Gly-Y guest triplets, confirm this destabilizing effect. The proximity of destabilizing Gly-Gly-Y triplets to the epitope may play an important role in recognition.

MATERIALS AND METHODS

Peptide Synthesis. All peptides studied were synthesized on an Applied Biosystems 430A Peptide Synthesizer using a standard FastMoc (Applied Biosystems) method. Two designs were used. The first design had a general structure of acetyl-$(Gly\text{-}X\text{-}Y)_6(Gly\text{-}Pro\text{-}Hyp)_4Gly\text{-}Val$, where $(Gly\text{-}X\text{-}Y)_6$ was a section of 6 consecutive triplets from the region 505−540 of type III collagen (Glattauer et al., 1997). The $(Gly\text{-}Pro\text{-}Hyp)_4$ sequence is included as a promoter and stabilizer of triple-helix formation, while the Gly-Val sequence is to eliminate possible diketopiperazine formation during synthesis. The peptides are designated by the residue number at the N-terminus, based on the sequence of human type III collagen (Ala-Kokko et al., 1989). For example, T3−505 includes residues 505−522 of human type III collagen. These peptides were synthesized on Fmoc-L-Val-WANG resin and were acetylated by acetic anhydride and triethylamine in dimethylformamide. Side chain protection was *tert*-butyl for Hyp, *tert*-butyloxycarbonyl for Lys, *tert*-butyl ester for Glu and Asp, and pentamethylchroman-sulfonyl for Arg.

The second peptide design was to synthesize "host-guest" peptides of the form acetyl-$(Gly\text{-}Pro\text{-}Hyp)_3Gly\text{-}Gly\text{-}Y(Gly\text{-}Pro\text{-}Hyp)_4Gly\text{-}GlyNH_2$, where Gly-Gly-Y is a guest triplet in a $(Gly\text{-}Pro\text{-}Hyp)_8$ host environment. The Y residue was either Hyp, Ala, Leu, or Phe. These peptides were synthesized on Fmoc-RINK-resin, as previously described for a related series of peptides (Shah et al., 1996), with acetylation and protection of Hyp as above. The N-terminal was acetylated and the C-terminal amidated to eliminate the charge repulsion between the three chains at both ends of the helix, which was previously seen to lower the overall thermal stability (Venugopal et al., 1994). A Gly-Gly sequence was included at the C-terminal to eliminate the likelihood of diketopiperazine formation during synthesis.

Peptides which showed an initial purity of less than 95% were further purified using reverse phase HPLC. Compositions of peptides were confirmed by amino acid analysis using a Waters HPLC system with ninhydrin detection. Mass spectrometry, using either liquid secondary positive-ion mass spectroscopy on a VG Analytical ZAB-T instrument (peptide T3−514) or laser desorption mass spectroscopy (peptides T3−511, T3−520, T3−523), also gave single peaks at expected molecular weights.

Circular Dichroism Spectroscopy. Circular dichroism (CD) spectra were recorded on an Aviv Model 62DS spectropolarimeter equipped with a Hewlett Packard Peltier thermoelectric temperature controller. Peptide solutions of concentrations 1 mg/mL were used, with peptides dried *in vacuo* over P_2O_5 for 48 h prior to weighing (Shah et al., 1996). Solutions made either in 0.1 M acetic acid, pH 2.7, or in 10 mM sodium phosphate with 0.15 M NaCl, pH 7.0 (PBS), were equilibrated for at least 48 h at 4 °C and then for 2 h at 2 °C before CD measurements. Cells of 1 mm path length were used. Spectra were recorded from 210 to 260 nm. For thermal equilibrium measurements, the wavelength was kept constant at 225 nm, while the temperature was increased in increments of 0.3 °C followed by 2 min equilibration at each temperature; data were collected for 10 s at each point.

Calculation of Melting Temperatures and Thermodynamic Parameters. Equilibrium melting curves obtained for all peptides (except T3−523) were analyzed using a two-state, trimer ↔ 3 (monomer) model. The equilibrium constant for this reaction is given by the equation:

$$K = \frac{[\text{trimer}]}{[\text{monomer}]^3} = \frac{F}{3c^2(1-F)^3}$$

where c is the total concentration of peptide expressed as monomers in moles and F is the fraction of the peptide in the folded (triple-helical trimer) state. F was calculated using the equation:

$$F = \frac{\theta_{\text{observed}} - \theta_{\text{monomer}}}{\theta_{\text{trimer}} - \theta_{\text{monomer}}}$$

where θ_{observed} is the observed ellipticity, θ_{trimer} is the ellipticity when the peptide is fully associated, and θ_{monomer} is the ellipticity of the monomer. θ_{trimer} and θ_{monomer} were obtained by linear extrapolation of the pre- and posttransition regions of the thermal equilibrium curves, respectively. The temperature at which F was 0.5 was taken as the melting temperature, T_m.

For peptide T3−523, which was not completely triple-helical at lower temperatures, θ_{trimer} could not be determined as there were not enough data points for linear extrapolation of the pretransition region of the curve. Here the T_m was taken as the temperature at which the first derivative, $d\theta/dT$, for the thermal equilibrium curve shows a minimum.

Thermodynamic parameters were calculated for the host-guest set of peptides using the method described earlier in Shah et al. (1996). The equilibrium melting transitions were fit to a two-state trimer ↔ monomer transition, and the van't Hoff enthalpy, $\Delta H°$, was determined by curve-fitting with the equation (Engel et al., 1977; Marky & Breslauer, 1987):

$$K = \exp\left[\frac{\Delta H°}{RT}\left(\frac{T}{T_m} - 1\right) - \ln(0.75c^2)\right]$$

The equation:

$$T_m = \frac{\Delta H°}{\Delta S° + R\ln(0.75c^2)}$$

was used to calculate $\Delta S°$, while $\Delta G°$ was calculated at a temperature of 298 K by $\Delta G° = \Delta H° - T\Delta S°$.

Molecular Modeling. Molecular modeling was performed using an SYBYL 6.1 molecular modeling package (Tripos Inc., St. Louis, MO). The coordinates for $(GPO)_{14}$ were

Table 1: Sequences of Seven Peptides, Including the Epitope Recognized by a Conformation-Dependent MAb, Their Reactivity to This MAb by ELISA, and Their T_m Values at pH 7.0 in PBS

Peptide	Peptide Sequence	Reactivity (ELISA)*	T_m (°C) (pH 7.0)
T3-505	Ac-**GGK**GDAGAOGERGPOGLA(GPO)$_4$GV	No	20.9
T3-508	Ac-GDAGAOGERGPOGLAGAO(GPO)$_4$GV	No	23.2
T3-511	Ac-GAOGERGPOGLAGAOGLR(GPO)$_4$GV	Yes	25.9
T3-514	Ac-GERGPOGLAGAOGLR**GGA**(GPO)$_4$GV	Yes	16.5
T3-517	Ac-GPOGLAGAOGLR**GGA**GPO(GPO)$_4$GV	Yes	15.8
T3-520	Ac-GLAGAOGLR**GGA**GPOGPE(GPO)$_4$GV	Yes	17.5
T3-523	Ac-GAOGLR**GGA**GPOGPE**GGK**(GPO)$_4$GV	No	8.9

*ELISA reactivity is taken from Glattauer et al. (1997) (see reference for original numerical data). The location of the epitope (Glattauer et al., 1997) is shown underlined while Gly-Gly-containing triplets are shown in boldface.

obtained from X-ray fiber diffraction data (Fraser et al., 1979), and then substituted to obtain the collagen sequence between residues 502 and 543. This was used without further refinement for illustration purpose only.

RESULTS

Triple-Helix Conformation and Stability of Overlapping Peptides. A set of seven overlapping peptides were synthesized, covering residues 505−540 of type III collagen, with the sequence in each peptide being offset by one triplet from the next peptide (Glattauer et al., 1997). These peptides are of the form acetyl-(Gly-X-Y)$_6$(Gly-Pro-Hyp)$_4$Gly-Val, and each contains 18 residues from the human type III collagen chain at its N-terminus, with a stabilizing (Gly-Pro-Hyp)$_4$ C-terminal (Table 1).

At low temperature, the peptides show a significant maximum near 225 nm that is characteristic of the triple-helix, and this maximum disappears as the temperature is increased (Figure 1a, inset). All peptides show a sharp thermal transition (Figure 1), which previous studies indicate corresponds to the trimer to monomer transition (Long et al., 1993). The melting temperatures of these seven peptides vary from 9 °C to 26 °C at pH 7.0, with peptide T3−511 being the most stable, while T3−523 is the least stable (Table 1). The melting temperatures of peptides were in most cases slightly higher at acid pH (pH 2.7) than at neutral pH (Table 2).

Peptide T3−523 is distinguished from the other peptides in this set by its low T_m and by the small magnitude of its maximum near 225 nm in PBS. Its low ellipticity value at 2 °C and the incomplete nature of the transition (Figure 1a) suggest it is unlikely that the peptide is fully trimeric even at 2 °C. However, the peptide was fully trimeric in 0.1 N acetic acid (pH 2.7), with trimer ellipticity values comparable to other peptides. Its T_m was 13.4 °C in 0.1 N acetic acid and approximately 8.9 °C in PBS, values considerably lower than seen for any of the other peptides.

Determinants of Stability in Overlapping Peptides. The variation of about 17 °C (9−26 °C) in T_m values among peptides in this set is significant considering the similarity of the peptides (Table 1). The differences in stability must be a result of the specific N-terminal (Gly-X-Y)$_6$ type III collagen sequences included in each peptide since the

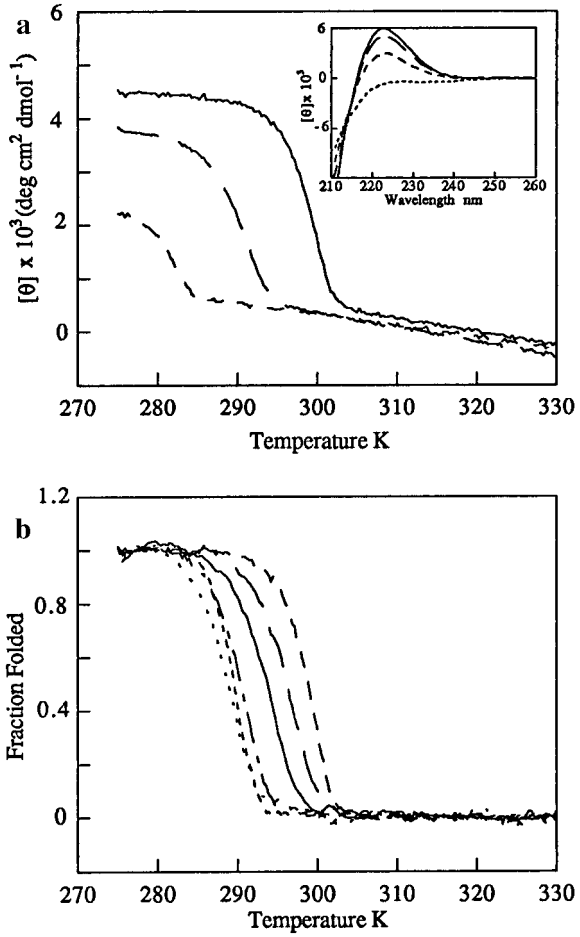

FIGURE 1: (a) Thermal transition curves for peptides in PBS (pH 7.0), showing mean residue ellipticities versus temperature for peptides T3−511 (—), T3−520 (− −), and T3−523 (- - -). Inset: CD spectra at 2 °C in PBS (pH 7.0) of the same peptides. The spectrum of peptide T3−511 at 60 °C in the monomer state is also shown (···). (b) Thermal transition curves for peptides in PBS (pH 7.0), showing fraction folded versus temperature for peptides T3−505 (—), T3−508 (− −), T3−511 (− − −), T3−514 (- - -), T3−517 (···), T3−520 (− - −). The peptide T3−523 is not included because its low T_m prevents determination of fraction folded values.

Table 2: Melting Temperatures (T_m) and Imino Acid, Hydroxyproline, Charged Pair, and Gly-Gly-Y Triplet Contents for the Set of Seven Overlapping Peptides

	T_m (°C)		total content		variable (GXY)$_6$ content			
peptide	pH 7.0	pH 2.7	imino (%)	Hyp (%)	imino (%)	Hyp (%)	potential ion pairs	GGY triplets
T3−505	20.9	20.9	34	19	17	11	2	1[a]
T3−508	23.2	24.2	38	22	22	17	1	0
T3−511	25.9	27.1	38	22	22	17	1	0
T3−514	16.5	18.5	34	19	17	11	1	1
T3−517	15.8	16.9	41	22	28	17	0	1
T3−520	17.5	20.8	38	19	22	11	0	1
T3−523	8.9	13.4	38	19	22	11	1	2

[a] N-Terminal Gly-Gly-Y triplet.

C-terminal (Gly-Pro-Hyp)$_4$Gly-Val ends are identical for the seven peptides. Attempts were made to correlate the stability of the peptides with side chain interactions which might affect the triple-helix stability, including imino acid content, electrostatic forces, and hydrophobic interactions.

Imino acids are known to be important determinants of triple-helical stability (Josse & Harrington, 1964), and Hyp plays a more stabilizing role than proline because of the

involvement of its hydroxyl group in the hydration network (Bella et al., 1994, 1995). There are relatively small variations in the total imino acid content (34−41%) and total Hyp content (19−22%) among the seven peptides in this set, and no dependence is seen between stability and either of these parameters (Table 2). Also, if one looks specifically at the imino acid content and Hyp content of the variable $(Gly\text{-}X\text{-}Y)_6$ N-terminal regions of the peptides, no correlation is evident (Table 2). Although electrostatic interactions can affect triple-helix stability (Venugopal et al., 1994; Anachi et al., 1995), the distribution of ion pairs in this peptide set does not correlate with observed stability variations. All peptides in the set contain one or two triplets with hydrophobic Leu residues, but no correlation is seen between stability and the presence of Leu (Table 2).

The one feature that correlates with stability is the presence of internal Gly-Gly-Y triplets. The introduction of Gly-Gly-Ala, going from T3−511 to T3−514, leads to a 9 °C decrease in T_m. The introduction of a second triplet of this form, Gly-Gly-Lys, in going from T3−520 to T3−523, results in a dramatic decrease in the T_m of the trimer that is formed, from 18 °C to 9 °C. In addition to these instances with internal Gly-Gly-Y triplets, peptide T3−505 has a Gly-Gly-Lys at its N-terminus, which may explain its lower stability (2−5 °C) compared with peptides T3−508 and T3−511. The decrease in T_m resulting from a Gly-Gly-Y at the N-terminus of a peptide may be less than that for an internal location, since previous X-ray and NMR studies suggest that the peptide ends are "frayed" and not fully triple-helical (Li et al., 1993; Bella et al., 1994).

Host-Guest Peptides with Gly-Gly-Y Guest Triplets. The overlapping peptide studies suggested that Gly-Gly-Y triplets had a destabilizing effect on the triple-helix conformation. A set of peptides was constructed to test this hypothesis. Triple-helical peptides of the form acetyl-$(Gly\text{-}Pro\text{-}Hyp)_3$Gly-Gly-Y$(Gly\text{-}Pro\text{-}Hyp)_4$Gly-Gly$NH_2$ allow evaluation of an isolated Gly-Gly-Y triplet in a defined triple-helical environment. The validity of this host-guest approach has been demonstrated (Shah et al., 1996). Host-guest peptides with Gly-Gly-Hyp, Gly-Gly-Ala, Gly-Gly-Leu, and Gly-Gly-Phe as the guest triplets were examined. All four peptides formed stable triple-helical molecules, as shown by their characteristic CD spectra and their sharp thermal transitions (Figure 2). The T_m values of these peptides varied depending upon the residue in the Y position. The most stable was Gly-Gly-Hyp, $T_m = 33$ °C, while the least stable was Gly-Gly-Phe, $T_m = 20$ °C (Table 3).

The presence of Gly in the X position of Gly-X-Y was found to be very destabilizing compared with other residues. Previously, host-guest peptide studies have shown Pro and Hyp to be the most stabilizing residues in the X and Y positions, respectively, while Phe was the most destabilizing residue of those studied (Shah et al., 1996). The guest triplets with Gly in the X position (Gly-Gly-Y) were substantially less stable than Pro, Ala, and Leu in the X position, but were comparable to X = Phe.

Thermodynamic parameters were calculated for the host-guest set peptides containing Gly-Gly-Y triplets (Table 3). Peptides of the host-guest design appear to fit a two-state monomer to trimer model which is amenable to thermodynamic analysis (Shah et al., 1996). The van't Hoff enthalpy is determined from the thermal equilibrium curve and used to calculate the entropy and free energy assuming a two state

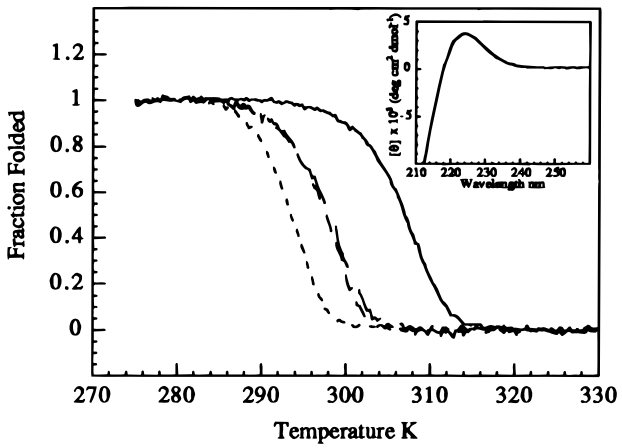

FIGURE 2: Thermal transition curves for host-guest peptides in PBS (pH 7.0), showing fraction folded versus temperature for peptides GGO (—), GGA (− −), GGL (- - -), and GGF (···). The inset shows the CD spectrum of peptide GGO (in the triple-helical state) at 2 °C in PBS.

Table 3: Thermal Stability (T_m) of Five Host-Guest Peptides and Their Thermodynamic Parameters, $\Delta H°$, $-T\Delta S°$ (where $T = 298$ K), and $\Delta G°$, in PBS

guest triplet	T_m (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ [kcal/(mol·K)]	$-T\Delta S°$ (kcal/mol)	$\Delta G°$ (kcal/mol)
GGF	19.7	−154.6	−0.50	148.1	−6.5
GGL	23.9	−138.1	−0.43	129.3	−8.8
GGA	25.0	−133.5	−0.42	124.2	−9.3
GGO	33.2	−137.4	−0.42	124.4	−13.0
GPO[a]	44.5	−84.0	−0.23	68.5	−14.5

[a] From Shah et al. (1996).

model. The presence of Gly in the X position is destabilizing relative to Pro and Ala (Shah et al., 1996) because of the unfavorable entropic term, while the enthalpy is more favorable. The presence of Gly-Gly sequences may lead to a more flexible monomer chain with a higher entropy value than seen for Gly-Ala or Gly-Pro sequences. The adoption of a similar triple-helical conformation would then require a greater loss in entropy.

DISCUSSION

The availability of a set of overlapping, triple-helical peptides which react with a conformation-sensitive MAb to type III collagen presents an opportunity for investigation of features required for a highly specific interaction of a triple-helix with another protein. Because of its rod-like triple-helical nature, collagen has binding sites consisting of linear amino acid sequences, so it is possible to successfully model a recognition site, such as that of an epitope, using relatively short peptides. Of the seven overlapping peptides, four (T3−511, T3−514, T3−517, T3−520) reacted with the MAb (Glattauer et al., 1997). All four peptides that reacted with the MAb contained the sequence Gly-Leu-Ala-Gly-Ala-Hyp-Gly-Leu-Arg, leading to the conclusion that this nine residue sequence defines the epitope necessary for reactivity (Glattauer et al., 1997). This reactive sequence is found only at a single site in human type III collagen. It is also found at one site in mouse and rat type II collagens, but at a different location in the sequence, residues 421−429 (Metsäranta et al., 1991; Glattauer et al., 1997). Although both collagens contain the same nine residue sequence, the Mab binding to the mouse type II collagen is much weaker (Glattauer et al., 1997).

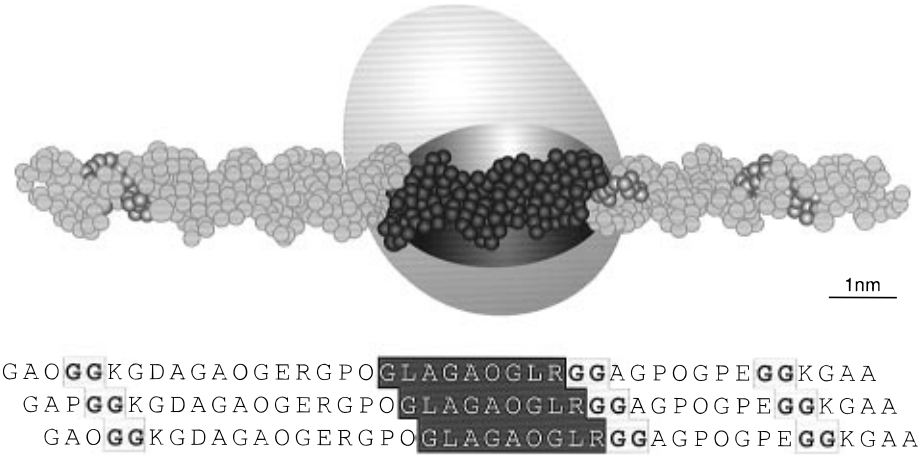

FIGURE 3: Schematic model of the α1(III) collagen triple-helical segment (residues 502−543), indicating the location of the epitope and Gly-Gly sites. The amino acid sequence of the triple helix is shown underneath the model with a one-residue stagger of the three chains. The F_v domain of an antibody has been schematically represented as an ellipse, to illustrate that the minor axis of the ellipse of the antibody binding domain is similar to the length of the triple-helical epitope (2.5−3.0 nm).

The results from this study and the previous work (Glattauer et al., 1997) allow consideration of features that mediate the recognition of the triple-helix Gly-Leu-Ala-Gly-Ala-Hyp-Gly-Leu-Arg sequence by a MAb. For the homotrimeric type III collagen, this three-triplet sequence represents approximately one pitch of the triple helix, and this length is comparable with the expected maximum size of a potential antibody binding site (Figure 3). Given an epitope size of three Gly-X-Y triplets, the number of variable sites for antibody recognition is limited to the three X and three Y positions. In the present case, the residues identified in the X and Y positions are among the most commonly observed in collagen. Further, none of the three particular triplet sequences is unusual. Consideration of these features makes it difficult to see how the amino acid sequence of the epitope would alone confer sufficient specificity. Therefore, it is possible that the conformation of the region including and surrounding the epitope is important for recognition. Thus, the weaker binding of the MAb to the same nine residue sequence found in mouse type II collagen (Glattauer et al., 1997) may be due to the different flanking regions. The mouse collagen (Metsäranta et al., 1991) lacks the destabilizing Gly-Gly sequences next to this nine residue sequence and instead has more stable triplet, GLO (Shah et al., 1996), that would lessen the chain flexibility.

The studies on the seven overlapping peptides indicated the marked destabilizing influence of Gly-Gly-Y triplets in the vicinity of the epitope site. A 9 °C drop in T_m was seen as a Gly-Gly-Ala sequence is introduced, and an additional large drop is seen when a Gly-Gly-Lys sequence is introduced. This destabilizing effect of Gly-Gly-containing triplets was confirmed by studies on model host-guest peptides. The results draw attention to the location of three Gly-Gly-Y triplets in relation to the epitope: one is adjacent to the C-terminal, one is three triplets from the C-terminal, and the other is five triplets from the N-terminal (Figure 3). The presence of these destabilizing triplets on either side of the epitope sequence may lead to a perturbation in the dynamics, energetics, or conformation of the helix that would "mark" the otherwise uniform helix and allow for the necessary discrimination. Previous studies have suggested that flexibility in the triple-helical conformation plays a role in recognition. For example, the fibril-forming collagens have a region of low imino acid content C-terminal to the highly specific cleavage site for mammalian collagenase, "marking" this site for enzyme binding (Fields, 1991). Binding of specific molecules has also been demonstrated to sites containing interruptions in the Gly-X-Y repeating sequence in type IV, which may represent regions of greater flexibility (Fields, 1995).

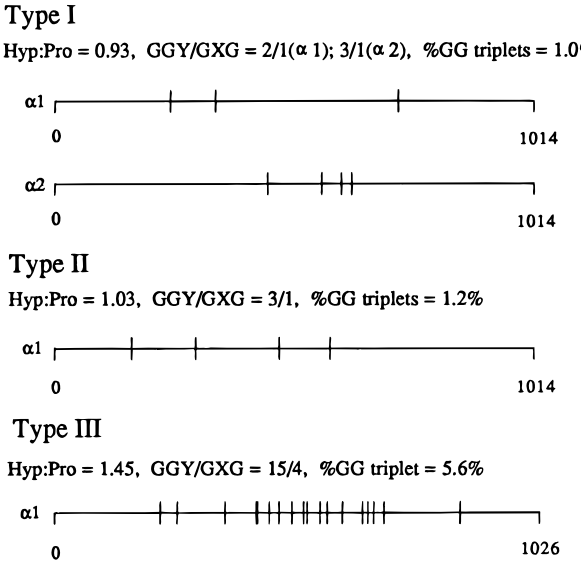

FIGURE 4: Plots of the locations of Gly-Gly-Y and Gly-X-Gly triplets in the α1[I], α2[I], α1[II], and α1[III] collagen chains, showing the differences in frequency and distribution of these triplets. The Hyp:Pro ratio, the numbers of Gly-Gly-Y and Gly-X-Gly triplets, and the total percentage of Gly-Gly triplets are given for each collagen type.

Type III collagen is the only collagen which contains significant amounts of triplets with Gly-Gly sequences (Figure 4). There are 15 Gly-Gly-Y triplets and 4 Gly-X-Gly triplets in the α-chain of type III collagen, compared with 3 Gly-Gly-containing sequences in the α1[I] chain and 4 in the α1[II] chain. The 19 Gly-Gly sites in the αl[III] chain are not distributed randomly, but are clustered near the center of the molecule, while the ends of the triple-helix contain few or no sites (Figure 4). The studies presented here indicate that Gly-Gly-containing triplets have a destabilizing effect in peptides, yet the global thermal stability of type III collagen is similar to that seen for other fibril-forming collagens (Davis & Bächinger, 1993). The de-

stabilizing influence of the high Gly-Gly content in type III collagen may be offset in part by its overall high hydroxyproline content. Even though the imino acid contents of the major interstitial collagens, types I, II, and III, are all very similar (ranging from 21.9 to 23.3%), the Hyp:Pro ratio is markedly greater for α1[III] (1.49) than for type I (0.93) or type II (1.03) collagens (Figure 4). It has been shown previously that Hyp contributes a greater stability than Pro to the triple-helix (Sakakibara et al., 1973; Berg & Prockop, 1973). The folding behavior of type III collagen may also be affected by its high content of Gly-Gly-containing tripeptide units. Given the destabilizing influence of these triplets, there may be "micro-unfolding", a breathing motion involving one or two tripeptides, at Gly-Gly sites (Privalov, 1982; Ryhänen et al., 1983). However, the size of the cooperative folding unit is unaffected, since this unit remains at 80−100 tripeptide units for type I, II, and III collagens (Davis & Bächinger, 1993). The Gly-Gly triplets could play a role in the distinctive hysteresis effect seen for type III collagen, which suggests a folding intermediate (Davis & Bächinger, 1993).

The presence of Gly-Gly triplets may confer unique properties on type III collagen molecules. The significant number of conformation-dependent MAbs that have been generated to the human type III collagen triple-helix, in contrast to other human interstitial collagens, may be related to the presence of Gly-Gly sites (Keene et al., 1987; Hori et al., 1992; Werkmeister & Ramshaw, 1991). It is interesting to note that another epitope identified in type III collagen is also in a region that contains Gly-Gly-Y triplets (Hori et al., 1992). The unusual destabilization at Gly-Gly sites could relate to the function of type III collagen in fibrils of tissues with some degree of extensibility, such as blood vessels and skin. The pathological conditions that result from the absence of type III collagen in Ehlers−Danlos Syndrome type IV or hereditary aortic aneurysms (Prockop & Kivirriko, 1995) makes it clear that the type I collagen alone cannot satisfy the tissue extensibility requirements.

The peptide sequence studied here represents one of the first cases where a region of collagen which binds to another molecule has been studied in some detail to clarify how such an apparently uniform molecule can undergo specific recognition and binding. The observation of unstable Gly-Gly-Y triplets adjacent to the recognition region suggests involvement of some flexibility or instability near the actual binding site. Since Gly-Gly sequences are found in significant numbers only in type III collagen, they are unlikely to play a role in the binding of molecules to other collagen types. However, it is possible that the interruptions of type IV collagen or the presence of other less stable triplets may play a similar role in modulating conformation such that recognition may occur for other collagens.

ACKNOWLEDGMENT

We thank Dr. Prem Yadav and Dr. Jose Varghese for assistance with molecular modeling and Veronica Glattauer and Dr. Jerome Werkmeister for helpful discussions. We appreciate the useful suggestions of Dr. Konrad Beck in improving Figure 3.

REFERENCES

Ala-Kokko, L., Kontusaari, S., Baldwin, C. T., Kuivaniemi, H., & Prockop, D. J. (1989) *Biochem. J. 260*, 509−516.

Anachi, R. B., Siegel, D. L., Baum, J., & Brodsky, B. (1995) *FEBS Lett. 368*, 551−555.

Bella, J., Eaton, M., Brodsky, B., & Berman, H. M. (1994) *Science 266*, 75−81.

Bella, J., Brodsky, B., & Berman, H. M. (1995) *Structure 3*, 893−906.

Berg, R. A., Prockop, D. J. (1973) *Biochem. Biophys. Res. Commun. 52*, 115−120.

Brodsky, B., & Shah, N. K. (1995) *FASEB J. 9*, 1537−1546.

Brodsky, B., & Ramshaw, J. A. M. (1997) *Matrix Biol.* (in press).

Brown, J. C., & Timpl, R. (1995) *Int. Arch. Allergy Immunol. 107*, 484−490.

Davis, J. M., & Bächinger, H. P. (1993) *J. Biol. Chem. 268*, 25965−25972.

Engel, J., Chen, H. T., Prockop, D. J., & Klump, H. (1977) *Biopolymers 16*, 601−622.

Fields, G. B. (1991) *J. Theor. Biol. 153*, 585−602.

Fields, G. B. (1995) *Connect. Tissue Res. 31*, 235−243.

Fraser, R. D. B., MacRae, T. P., & Suzuki, E. (1979) *J. Mol. Biol. 129*, 463−481.

Glattauer, V., Werkmeister, J. A., Kirkpatrick, A., & Ramshaw, J. A. M. (1997) *Biochem. J. 323*, 45−49.

Hori, H., Keene, D. R., Sakai, L. Y., Wirtz, M. K., Bächinger, H. P., Godfrey, M., & Hollister, D. W. (1992) *Mol. Immunol. 29*, 759−770.

Jones, E. Y., & Miller, A. (1991) *J. Mol. Biol. 218*, 209−219.

Josse, J., & Harrington, W. F. (1964) *J. Mol. Biol. 9*, 269−287.

Keene, D. R., Sakai, L. Y., Burgeson, R. E., & Bächinger, H. P. (1987) *J. Histochem. Cytochem. 35*, 311−318.

Li, M. H., Fan, P., Brodsky, B., & Baum, J. (1993) *Biochemistry 32*, 7377−7387.

Long, C. G., Braswell, E., Zhu, D., Apigo, J., Baum, J., & Brodsky, B. (1993) *Biochemistry 32*, 11688−11695.

Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.

Metsäranta, M., Toman, D., de Crombrugghe, B., & Vuorio, E. (1991) *J. Biol. Chem. 266*, 16882−16869.

Privalov, P. L. (1982) *Adv. Protein Chem. 35*, 1−103.

Prockop, D. J., & Kivirriko, K. I. (1995) *Annu. Rev. Biochem. 64*, 403−434.

Ramachandran, G. N. (1967) in *Treatise on Collagen* (Ramachandran, G. N., Ed.) Vol. 1, pp 103−183, Academic Press, New York.

Rich, A., & Crick, F. H. C. (1961) *J. Mol. Biol. 3*, 483−506.

Ryhänen, L., Zaragoza, E. J., & Vitto, J. (1983) *Arch. Biochem. Biophys. 223*, 562−571.

Sakakibara, S., Inouye, K., Shudo, K., Yasua, K., Kobayashi, Y., & Prockop, D. J. (1973) *Biochim. Biophys. Acta 303*, 198−202.

Shah, N. K., Ramshaw, J. A. M., Kirkpatrick, A., Shah, C., & Brodsky, B. (1996) *Biochemistry 35*, 10262−10268.

Venugopal, M. G., Ramshaw, J. A. M., Braswell, E., Zhu, D., & Brodsky, B. (1994) *Biochemistry 33*, 7948−7956.

Werkmeister, J. A., & Ramshaw, J. A. M. (1991) *Biochem. J. 274*, 895−898.

BI963146C