Stability Junction at a Common Mutation Site in the Collagenous Domain of the Mannose Binding Lectin†

Angela Mohs,‡ Yingjie Li,§ Ellen Doss-Pepe,‡ Jean Baum,§ and Barbara Brodsky*,‡

Department of Biochemistry, UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, and Department of Chemistry, Rutgers University, Piscataway, New Jersey 08854

Received August 11, 2004; Revised Manuscript Received November 16, 2004

ABSTRACT: Missense mutations in the collagen triple-helix that replace one of the required Gly residues in the $(Gly-Xaa-Yaa)_n$ repeating sequence have been implicated in various disorders. Although most hereditary collagen disorders are rare, a common occurrence of a Gly replacement mutation is found in the collagenous domain of mannose binding lectin (MBL). A Gly → Asp mutation at position 54 in MBL is found at a frequency as high as 30% in certain populations and leads to increased susceptibility to infections. The structural and energetic consequences of this mutation are investigated by comparing a triple-helical peptide containing the N-terminal Gly−X−Y units of MBL with the homologous peptide containing the Gly to Asp replacement. The mutation leads to a loss of triple-helix content but only a small decrease in the stability of the triple-helix ($\Delta T_m \sim 2$ °C) and no change in the calorimetric enthalpy. NMR studies on specifically labeled residues indicate the portion of the peptide C-terminal to residue 54 is in a highly ordered triple-helix in both peptides, while residues N-terminal to the mutation site have a weak triple-helical signal in the parent peptide and are completely disordered in the mutant peptide. These results suggest that the N-terminal triplet residues are contributing little to the stability of this peptide, a hypothesis confirmed by the stability and enthalpy of shorter peptides containing only the region C-terminal to the mutation site. The Gly to Asp replacement at position 54 in MBL occurs at the boundary of a highly stable triple-helix region and a very unstable sequence. The junctional position of this mutation minimizes its destabilizing effect, in contrast with the significant destabilization seen for Gly replacements in peptides modeling collagen diseases.

A variety of hereditary disorders are due to mutations in the collagen triple-helix motif (*1*, *2*). The triple-helix is the characteristic structure in the family of extracellular matrix proteins defined as collagens and also forms a domain in various host-defense proteins (*3*). Three extended polyproline II-like chains are supercoiled around each other to form the triple-helix structure, with hydrogen bonds between the three chains (*4*−*6*). The extended chain conformation is stabilized by a high content of the imino acids, Pro and Hyp, while the close packing of the three chains requires every third residue to be glycine. These conformational requirements generate a repeating $(Gly-X-Y)_n$ sequence, where X and Y are frequently occupied by Pro and Hyp, respectively. While any residue can be accommodated in the X or Y positions (*7*), replacement of Gly by any larger amino acid appears to result in pathology. The most common mutations in triple-helices are missense mutations causing Gly to be replaced by one of eight bulkier residues generated by a single nucleotide change.

The nature of the disease associated with a mutation depends on the function and location of the triple-helix containing molecule (*1*, *2*, *8*). The first characterized collagen triple-helix molecular disease was Osteogenesis Imperfecta (OI), a hereditary bone disorder due to mutations in type I collagen, the predominant collagen in bone (*9*). A variety of hereditary disorders has been linked to mutations in other collagen types, such as dystrophic epidermolysis bullosa due to mutations in type VII collagen (*1*, *8*, *10*). Mutations have also been found in the triple-helix domains of noncollagenous proteins, such as the variants observed for C1q and mannose binding lectin (*2*, *11*). Mannose binding lectin (MBL,[1] also known as mannose binding protein or mannan binding protein) plays a critical role in host-defense against pathogenic microorganisms (*2*, *12*−*14*). It functions in the complement pathway and opsonization. The lectin pathway of complement activation requires association of MBL with MBL associated serine proteases (MASPs). It is the binding of this complex to sugar groups on the surface of microorganisms that leads to MASP autoactivation and cleavage of the complement components C4 and C2. MBL consists of

† This work was supported by Grants GM60048 (B.B.) and GM45302 (J.B.). A.M. acknowledges support of a Deutsche Forschungsgemeinschaft (DFG) postdoctoral fellowship, and B.B. was supported by a Ruth Kirchstein National Research Service Award (NIGMS).

* To whom correspondence should be addressed. Phone: (732) 235-4397. Fax: (732) 235-4783. E-mail: brodsky@umdnj.edu.

‡ UMDNJ-Robert Wood Johnson Medical School.

§ Rutgers University.

[1] Abbreviations: MBL, mannose binding lectin; CD, circular dichroism; DSC, differential scanning calorimetry; OI, osteogenesis imperfecta. The three letter code for amino acids is used, with Hyp designating hydroxyproline. Peptides are designated by the number of the first and last residues of the MBL sequence included, such as MBL 45−61. Single amino acid notation is used to indicate the sequence of peptides, with O used for hydroxyproline.

10.1021/bi0482708 CCC: $30.25
Published on Web 01/20/2005

four distinct domains: a N-terminal Cys rich domain; a collagen-like domain with a kink; a short coiled-coil α-helix; and a C-terminal globular lectin. The collagenous domain of human MBL is of the form $(Gly-X-Y)_7-Gly-Gln-(Gly-X-Y)_{12}$ and contains a kink at the Gly−Gln interruption site. The N-terminal region of the triple-helix prior to the interruption, together with the adjacent Cys-rich region, self-associates into disulfide linked biologically active oligomers of four to six molecules, which have a bouquet-like appearance (*12*−*15*). The fifth triplet of the triple-helix (Gly−Thr−Lys) is the site of a common Gly → Asp mutation at codon 54 causing a low serum level of MBL, while other less frequent variants include a Gly → Glu replacement in the sixth triplet and an Arg → Cys replacement in the fourth tripeptide unit of the collagen domain (*2*, *14*, *16*−*19*). The Gly → Asp mutation at position 54 in MBL has been implicated in childhood infections and increased susceptibility to infections and autoimmune disease (*17*, *18*). This variant has a frequency as high as 30% in sub-Saharan African populations, making it by far the most frequent Gly substitution observed in any collagenous domain (*2*, *16*). Extensive studies have been carried out to investigate the molecular mechanism that links the Gly → Asp mutation in the collagen-like domain of MBL to increased risk of infection. In particular, the mutation has been shown to interfere with normal assembly of the trimeric molecules into oligomers and to alter interactions with MASPs (*2*, *16*, *19*−*21*). Studies on recombinant MBL suggests that the triple-helix formed when the mutation is present has an enthalpy similar to the native state but a reduced stability (*19*−*21*).

Synthetic collagen peptides offer an approach to characterizing the effect of mutations on the triple-helix (*22*). Peptide models composed of all Gly−Pro−Hyp triplets can successfully incorporate a Gly mutation. The high-resolution structure of a $(Pro-Hyp-Gly)_{10}$ peptide with a Gly to Ala substitution near the center indicated a local loss of direct hydrogen bonding and a slight untwisting at the mutation site (*6*). The replacement of a single Gly in a host−guest $(Gly-Pro-Hyp)_8$ system by the set of residues generated by a single base change leads to a dramatic loss of stability, with the degree of destabilization dependent on the identity of the residue replacing Gly: Ala < Ser < Cys < Arg < Val < Glu, Asp < Trp (*23*, *24*). The sequence surrounding the replacement can also affect the degree of destabilization. Peptides that model two Gly → Ser OI sites both show a significant decrease in triple-helix stability, but the change in T_m is 11 °C in one case and 22 °C for the other case (*25*).

Here, studies on peptides containing Gly replacements are extended to model the Gly to Asp mutation found at position 54 in the N-terminal triple-helical domain of MBL. This mutation is of interest because of the high frequency of its occurrence and because of the high degree of destabilization expected when an Asp replaces a Gly in the triple-helix. Circular dichroism spectroscopy (CD), NMR spectroscopy, and calorimetric studies indicate that the MBL N-terminal $(Gly-X-Y)_n$ sequence has an asymmetric nature: a highly ordered triple-helix stabilized substantially by electrostatic interactions C-terminal to residue 54 contrasts with an unstable triple-helix with weak hydrogen bonding N-terminal to residue 54. The Gly to Asp substitution occurs just at the boundary and leads to complete disordering of the N-terminus with only a very small drop in stability and no loss of enthalpy in the model peptide.

MATERIALS AND METHODS

Peptides. Peptides MBL 45−61, MBL 45−61[G10D], MBL 42−61, MBL 42−61[G13D], and MBL 45−64 were synthesized by Synpep Corporation (Dublin, CA), while peptides MBL 54−61 and MBL 57−61 were synthesized by Tufts University Core Facility (Boston, MA). Peptides were purified by a Shimadzu reversed-phase HPLC system, and the identity of the peptides was confirmed by laser desorption mass spectrometry. The concentration of MBL 42−61, MBL 42−61[G13D], and MBL 45−64 was determined by dry weight. All other MBL model peptides included a Tyr at the C-terminus, and the concentration was determined using the molar extinction coefficient 1400 (M^{-1} cm^{-1}) at 275 nm on a Beckmann model DU640 spectrophotometer.

Circular Dichroism Spectroscopy. Circular dichroism spectra were recorded on an AVIV model 62DS spectropolarimeter. Cuvettes of 1 mm path lengths were used, and the temperature of the cells was controlled using a Peltier temperature controller. Peptides were equilibrated at 0 °C in either PBS buffer, pH 7 or in 50 mM glycine/HCl buffer, pH 3 for at least 48 h prior to temperature-induced denaturation, which was monitored at 224 nm. The peptides were equilibrated for 2 min at each temperature, and the temperature was increased at an average rate of 0.1 °C/min, using our standard conditions described previously (*26*). The error in T_m values is ±0.5 °C. For equilibrium experiments, the equilibration time was 4−12 h at each temperature in the pretransition region, 2−4 h in the transition region, and 20−30 min in the posttransition region.

Calorimetry. Differential scanning calorimetry (DSC) experiments were recorded on a NANO-DSC II model 6100 (Calorimetry Sciences Corporation.). Peptides were dialyzed against PBS buffer and equilibrated at 0 °C for at least 48 h. Peptide solutions (1 mg/mL in PBS buffer, pH 7) were loaded at 0 °C into the cell and heated at a heating rate of 1 °C/min. Since this heating rate gives a scan that is far from equilibrium (*26*), the denaturation temperature T_m observed is scan-rate dependent and higher than that observed for CD studies done at slower rates. The enthalpy was calculated from the first scan since the scans were not reversible upon being cooled. Enthalpy measurements under these conditions should be valid since it was shown for similar collagenous peptides that the enthalpy is not dependent on the heating rate (*26*).

NMR Spectroscopy. The MBL 45−61 and MBL 45−61-[G10D] peptides were synthesized with ^{15}N-Gly at positions 7 and 25. All NMR samples were prepared in 10% D_2O/90% H_2O in sodium phosphate buffer at pH 7, with concentrations of 8 mg/mL for MBL 45−61 and 13 mg/mL for MBL 45−61[G10D]. NMR experiments were performed on a Varian INOVA 600 MHz spectrometer. The $\{^1H-^{15}N\}$ heteronuclear single quantum coherence (HSQC) spectra were recorded at 0 and 40 °C (*27*, *28*). Spectral widths were 6000 Hz in the proton dimension and 1500 Hz in the nitrogen dimension. All data had 2048 complex data points in the t_2 dimension and 128 increments in the t_1 dimension. Two-dimensional data were processed on a Silicon Graphics

Table 1: Thermal Stability and Calorimetry of Peptides Modeling the N-Terminal Collagenous Domain of MBL and Homologous Peptides Containing Gly54Asp Replacements

peptide	sequence[d]	T_m (°C)[a] pH 7	T_m (°C)[a] pH 3	$[\Theta_{MRW}]_{224}$ (deg cm^2 dmol^{-1})[b]	ΔH_{cal} (kJ/mol)[c]
MBL 42−61	Ac-GINGFOGKDGRD**G**TKGEKGEOGPOGPOGPOGPOGG-CONH$_2$	23.0	8.9	4820 (±680)	n.d.
MBL 42−61[G13D]	Ac-GINGFOGKDGRD**D**TKGEKGEOGPOGPOGPOGPOGG-CONH$_2$	21.3		2600 (±450)	n.d.
MBL 45−64	Ac-GFOGKDGRD**G**TKGEKGEPGQGPOGPOGPOGPOGV-CONH$_2$	<0			n.d.
MBL 45−61	NH$_2$-GFOGKDG*RD**G**TKGEKGEOGPOGPOG*POGPOGY-COOH	17.9	<0	4655 (±520)	145
MBL 45−61[G10D]	NH$_2$-GFOGKDG*RD**D**TKGEKGEOGPOGPOG*POGPOGY-COOH	15.9		2988 (±250)	141
MBL 54−61	Ac-GTKGEKGEOGPOGPOGPOGPOGY-COOH	18.5	<0	4017 (±400)	191
MBL 57−61	Ac-GEKGEOGPOGPOGPOGPOGY-COOH	11.1	<0	3564 (±350)	134

[a] T_m values are derived from temperature denaturation monitored by CD at 224 nm. [b] To obtain the mean residue ellipticity $[\Theta_{MRW}]_{224}$, peptides were equilibrated at least 48 h at 0 °C. [c] ΔH_{cal} is obtained from calorimetric experiments done at pH 7 with a heating rate of 1 °C/min. The error for ΔH_{cal} is ±10 kJ/mol. [d] Asterisk (*) is used to indicate ^{15}N-labeled residues used for NMR studies.

workstation using the FELIX 97 software package (MSI, San Diego, CA).

To accurately measure ^{15}N R_2 values, R_2 measurements were carried out separately for monomer and trimer resonances with different relaxation delays. ^{15}N R_2 measurements for monomer peaks were performed in two dimensions with seven relaxation delays (10, 50, 90, 130, 170, 210, and 250 ms), while measurements for the trimer peaks used one-dimensional experiments with seven relaxation delays (4.8, 14.4, 24.0, 33.6, 43.2, 52.8, and 62.4 ms) and were averaged over two data sets. For {^{1}H−^{15}N} NOE measurements, one-dimensional spectra were collected in the presence and absence of ^{1}H saturation (*29*). All relaxation experiments were performed at 0 °C.

RESULTS

Peptide Design. One peptide was designed to include the entire N-terminal sequence of the collagenous domain in human MBL, GINGFOGKDGRDGTKGEKGEP (MBL 42−61), while a second peptide included only residues 45−61, lacking the N-terminal GIN triplet (designated MBL 45−61). Four additional Gly−Pro−Hyp tripeptides were added at the C-terminus for stabilization. The last GEP triplet was changed to GEO (where O denotes Hyp) to increase stability, even though there is direct evidence that this Pro in the Y position remains unhydroxylated (*15*). Both peptides, with six triplets from MBL and seven triplets from MBL, formed a stable triple-helix structure. Attempts to extend the sequence in a C-terminal direction through the GQG interruption was so destabilizing that a triple-helix structure did not form, and these peptides were not studied further (Table 1).

Homologous peptides were synthesized with a Gly to Asp replacement at position 54 in the MBL sequence, corresponding to the 10th residue in the peptide MBL 45−61-[G10D]) and the 13th residue in MBL 42−61[G13D]. When evidence indicated that the peptides were not folded N-terminal to the mutation site, peptides were also designed to model only the portion of the sequence C-terminal to the mutation site: MBL 54−61, GTKGEKGEO(GPO)$_4$GY and MBL 57−61, GEKGEO(GPO)$_4$GY (Table 1).

CD Spectroscopy: *Thermal Stability and Conformation*. Circular dichroism spectroscopy was used to characterize the conformation and thermal stability of the MBL model peptides. The parent peptides MBL 45−61 and MBL 42−61 both gave a characteristic triple-helix CD spectrum at low temperature, pH 7, with a maximum near 224 nm (Figure 1A, inset). Monitoring the peak as a function of temperature using our standard conditions (see Materials and Methods) indicated a sharp thermal transition, with a T_m near 18 °C for MBL 45−61 and T_m of 23 °C for MBL 42−61 (Figure 1A; Table 1). The N- and C-terminal blocking and the increased length of MBL 42−61 are likely to be the reasons for its greater thermal stability (*30*). At pH 3, the thermal stability of both peptides was markedly decreased; for MBL 45−61, only the end of a thermal transition could be observed near low temperatures, and no T_m value could be

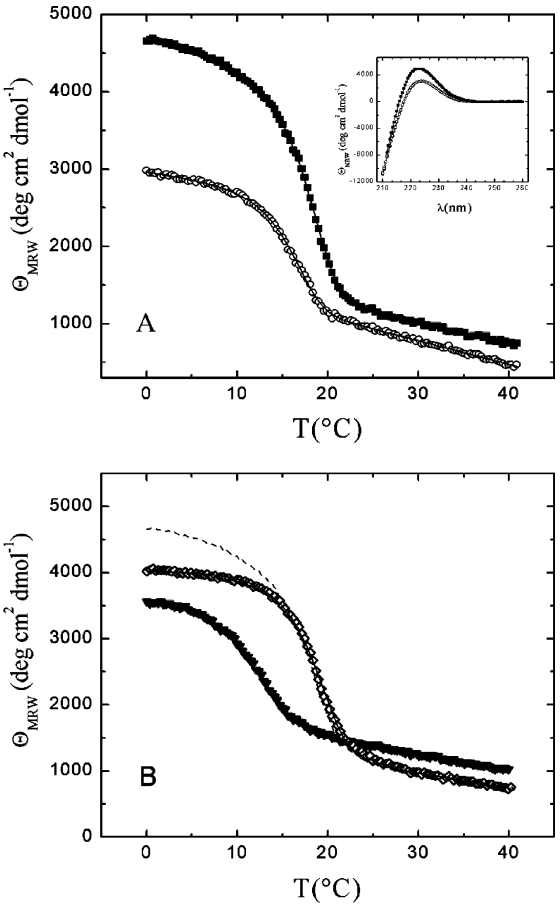

FIGURE 1: Temperature-induced denaturation monitored by circular dichroism spectroscopy at 224 nm (A): MBL 45−61 (■) and MBL 45−61[G10D] (○). The inset shows a wavelength scan monitored by CD at 0 °C. (B) MBL 54−61 (◊) and MBL 57−61 (▼). For comparison, MBL 45−61 is shown as dotted line. Solid lines represent fits to a monomer trimer model: 3M → T. All samples were run under the same standard conditions (average 0.1 °C/min) and can be compared even though these are not equilibrium conditions (*26*).

measured (Table 1). The loss of stability at low pH suggests a role for the electrostatic interactions in triple-helix formation of this highly charged sequence.

The peptides containing the Gly → Asp replacement (MBL 45−61[G10D] and MBL 42−61[G13D]) also showed a collagen-like CD spectrum, but the mean residue ellipticity of the maximum dropped by about 30% as compared to the wild-type peptide (Figure 1A, inset). The T_m values for the mutant peptides in PBS, pH 7, decreased only 2 °C as compared with the parent peptides (Figure 1A; Table 1). At pH 3, no triple-helix signal was observed. The small drop in the T_m value as a consequence of introducing the Gly → Asp replacement was surprising, given the high destabilization ranking of Asp as a Gly replacement (*23*) and the large destabilization previously reported for Gly substitutions in OI peptides of similar design (*25*).

Since both MBL 45−61 and MBL 42−61 peptides showed similar effects of the Gly54Asp replacement, further CD, NMR, and calorimetry studies focused on MBL 45−61, which has a length and design homologous to peptides used to characterize Gly replacements in collagens (*22*, *25*). Equilibrium thermal transitions for MBL 45−61 yielded lower values than our standard melting conditions, $T_m = 11$ °C for the parent peptide and $T_m = 9$ °C for the mutant peptide (data not shown). This confirmed that the small drop in thermal stability due to the Gly replacement in MBL peptides is not a result of nonequilibrium conditions.

The thermal stabilities at our standard conditions (pH 7) for two shorter peptides containing only the sequence C-terminal to the mutation site, including or excluding residues 54−56, were also studied. Both MBL 54−61 (23-mer), which contained the GTK triplet where the mutation is located, and MBL 57−61 (20-mer), which lacks the GTK triplet, formed stable triple-helices, with T_m values of 18.5 and 11.1 °C, respectively (Figure 1B; Table 1). These peptides were unblocked at the C-terminus, like MBL45−61, but blocked at the N-terminus, to model the fact that there are extra sequences rather than the charged terminal amino group in the MBL 45−61 model peptides. The parent $(Gly-X-Y)_{10}$ peptide MBL 45−61 has a stability very similar to that of the shorter $(Gly-X-Y)_7$ peptide MBL 54−61, suggesting that the N-terminal triplets are not contributing to the stability. The full-length mutant peptide, MBL 45−61[G10D], has a T_m of 15.9 °C, between the values seen for the MBL 54−61 and MBL 57−61 peptides (Table 1). These results both suggest the sequences N-terminal to residue 54 are not contributing to the stability in the parent peptide or the homologous mutant peptide.

NMR Spectroscopy: *Conformation, Dynamics, and Aggregation.* NMR studies were carried out on peptides MBL 45−61 and MBL 45−61[G10D] with ^{15}N-labeled residues at positions Gly7 and Gly25. The $\{^1H-^{15}N\}$ HSQC spectrum of MBL 45−61 gave Gly7 and Gly25 trimer peaks that disappeared at temperatures higher than the T_m. Monomer peaks for Gly7 and Gly25 were present at both high and low temperatures (Figure 2). The monomer and trimer resonance peaks of Gly25 had chemical shifts nearly identical to those reported previously for Gly25 in a well-characterized peptide of similar design (peptide T1-892) (*31*, *32*). A broad trimer peak with indications of multiple components was assigned to the Gly7 residue. Trimer peaks of both Gly25 and Gly7 have high R_2 values and high NOE values,

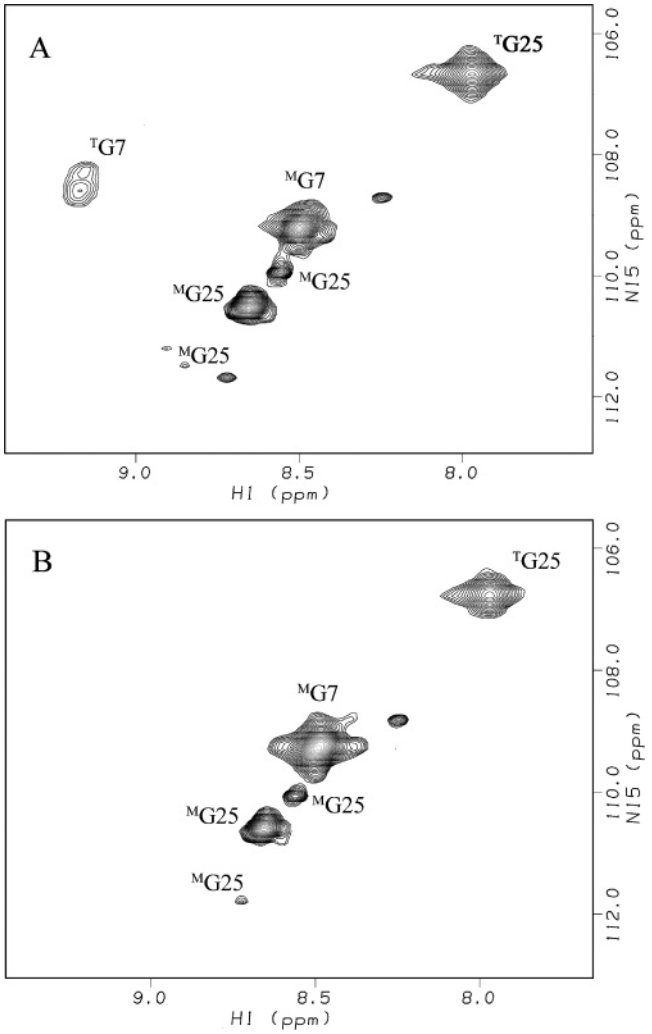

FIGURE 2: $\{^1H-^{15}N\}$ HSQC spectra obtained at 0 °C for (A) MBL 45−61 and (B) MBL 45−61[G10D]. The peaks corresponding to the monomer or trimer state are denoted with a superscript M or T, respectively. The two peptides were synthesized with ^{15}N-labeled positions at Gly7 and Gly25.

confirming the assignments and supporting the presence of a triple-helical conformation at both ends of the peptide. The intensity of the Gly7 trimer peak is weak relative to that of the Gly25 trimer peak. The Gly7 trimer peak also has an abnormally downfield chemical shift as compared with trimer peaks in other triple-helical peptides (*31*), which could relate to its highly charged sequence environment. At acid pH, the intensity of the Gly25 trimer peaks drops by a factor of 2, while the trimer peak of Gly7 is not observed. Both Gly7 and Gly25 monomer peaks have multiple resonances because of cis/trans isomerization and have low R_2 and NOE values as expected for disordered chains.

In the HSQC spectrum of peptide MBL 45−61[G10D] at low temperature, no trimer peak was observed for Gly7. In contrast, the trimer peak seen for Gly25 has chemical shift, ^{15}N R_2, and NOE values very close to those in MBL 45−61. The results suggest that the introduction of a Gly to Asp mutation at position 10 leads to a partially triple-helical conformation, with a rigid triple-helix retained at the C-terminal Gly−Pro−Hyp region and a monomer-like state at the N-terminal sequence.

The trimer ^{15}N R_2 values for both MBL 45−61 (29.51 ± 1.49 for Gly7 and 25.89 ± 0.47 for Gly25) and MBL 45−

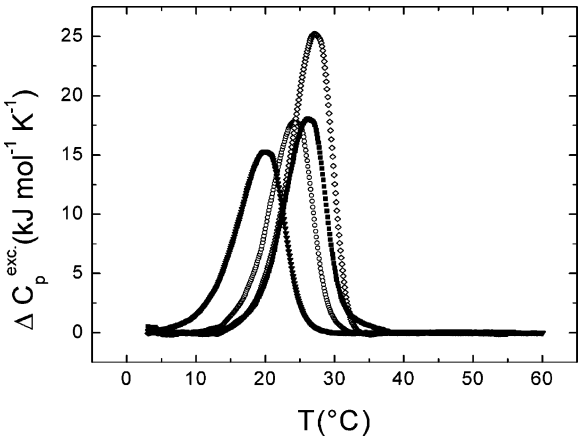

FIGURE 3: Temperature dependence of the excess partial molar heat capacity (heating rate 1 °C/min), showing MBL 45−61 (■), MBL 45−61[G10D] (○), MBL 54−61 (◊), and MBL 57−61 (▼). Although these are not equilibrium conditions, the enthalpy has been shown for a similar peptide to be heating rate independent (*26*).

61[G10D] (25.26 ± 0.60 for Gly25) are significantly higher than the trimer values in peptide T1−892 (19.29 ± 0.42 for Gly 25), which has the same length and is characterized to be triple-helical (*32*, *33*). Since ^{15}N R_2 values are governed by size, shape, and local flexibility, a higher value suggests the possibility of aggregation, which is supported by the decreased R_2 values seen for Gly25 and Gly7 trimers after the dilution of MBL 45−61.

Calorimetric Studies. Differential scanning calorimetry was carried out on peptides MBL 45−61, MBL 45−61[G10D], MBL 54−61, and MBL 57−61. The apparent T_m values are higher than those obtained by CD because of the higher scan rate (1 °C/min) utilized in DSC and the nonequilibrium state. However, the pattern of the apparent T_m values is the same as that seen by CD (Figure 3). The calorimetric enthalpy is the same within experimental error for the wild-type MBL 45−61 peptide (ΔH = 145 kJ/mol) and the mutant peptide MBL 45−61[G10D] (ΔH = 141 kJ/mol). Since the NMR data show that the residues N-terminal to the G10D replacement site are not triple-helical in the mutant peptide, the similarity in enthalpy values suggests that in the wild-type peptide, little contribution is made to the enthalpy by the N-terminal residues of the peptide (Table 1). The enthalpy of the peptide with the C-terminal 7 triplets, MBL 54−61, is ΔH = 191 kJ/mol, higher than that seen for the full length peptide MBL 45−61. This is consistent with the low contribution of the missing N-terminal sequence and even suggests that a more favorable hydrogen bonding network is established when this unstable region is completely absent. The enthalpy for the shortest peptide, with the C-terminal 6 triplets, MBL 57−61, is ΔH = 134 kJ/mol, showing the high enthalpic effect of excluding the GTK tripeptide sequence (a drop from ΔH = 191 kJ/mol to ΔH = 134 kJ/ mol).

DISCUSSION

Studies on MBL model peptides have increased our understanding of the factors leading to the loss in triple-helix stability as a consequence of a Gly substitution. In the MBL peptide, the introduction of the Gly to Asp substitution leads to loss of triple-helix N-terminal to the substitution site, but with only a very small drop in stability (2 °C). CD

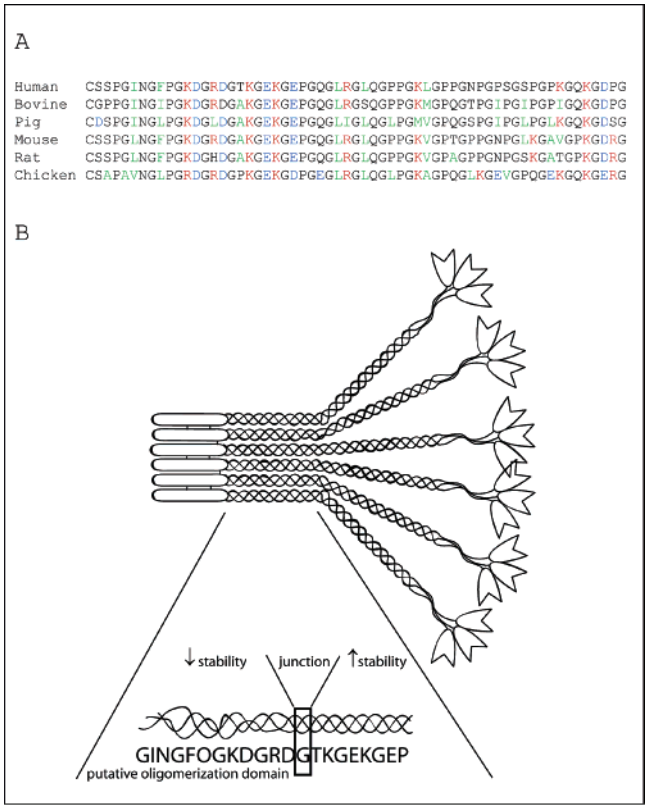

FIGURE 4: (A) Multiple sequence alignment of the MBL collagenous domain, showing acidic residues in blue, basic residues in red, and hydrophobic residues in green. Higher conservation is clearly seen N-terminal to the GQG interruption. (B) Schematic diagram of MBL and the N-terminal region of the triple-helix modeled in these studies, illustrating the discontinuity in triple-helix stability at the location of the common Gly54Asp mutation site.

and calorimetric data suggest that the small drop in T_m is a consequence of the highly stable triple-helix C-terminal to the mutation site and the very unstable triple-helix N-terminal to the mutation site. On the basis of host−guest peptide T_m values, one can assign a relative triple-helix propensity value for each Gly−X−Y tripeptide (*7*, *34*) and values for interactions between adjacent Gly−X−Y units (*35*), which can be used to predict the stability of a given sequence. Looking within the sequence, the MBL region flanking the mutation site on the N-terminal side (GINGFOGKDGRD) consists of residues that give a large net loss of stability as compared to Gly−Pro−Hyp (GIN, −24.5 °C; GFO, −14 °C; GKD, −12 °C; GRD, −13 °C). For the region C-terminal to the mutation site, GTKGEKGEO, the individual tripeptide sequences are still unfavorable (GTK, −21.6 °C; GEK, −12.3 °C; GEO, −4.4 °C as compared to GPO), but the presence of KGE sequences is highly favorable, adding a stability of 15.4 °C in the context of a GPKGEO sequence (*35*). If the stabilizing effect of KGE sequences is ignored, the MBL 45−61 peptide is not predicted to form a stable triple-helix. The importance of electrostatic interactions is confirmed by the destabilization seen at acid pH and by the conservation of KGEKGE/D sequences in MBL across species (Figure 4A). However, the consideration of the KGEKGE sequence as being equal to two isolated KGE sequences leads to an overestimate of peptide stability (Table 2), suggesting strings of such sequences are not simply additive. Calculations using only the sequence C-terminal

Table 2: Experimental and Predicted T_m Values for Different Collagen-Like Peptides with and without Gly Replacements

peptide	sequence	obs. T_m (°C)[a]	parent peptide pred. T_m (°C)[b]	peptide with mutation pred. T_m C-terminal (°C)[c]
MBL 45−61	NH_2-GFOGKDGRD**G**TKGEKG EOGPOGPOGPOGPOGY-COO H	17.9	27.0[d]	
MBL 45−61[G10D]	NH_2-GFOGKDGRD**D**TKGEKG EOGPOGPOGPOGPOGY-COO H	15.9		11.9−16.7
T7-2031[e]	Ac-GLAGEOGKOGIO**G**LOGRA GPOGPOGPOGPOGV-$CONH_2$	25.4	23.3	
T7-2031[G1 3R][e]	Ac-GLAGEOGKOGIO**R**LOGRA GPOGPOGPOGPOGV-$CONH_2$	<0		<0
T1-892[f]	Ac-GPAGPAGPV**G**PA**G**ARGPAG POGPOGPOGPO-$CONH_2$	27.5	28.2	
T1-892[G10 S][f]	Ac-GPAGPAGPV**S**PAGARGPAG POGPOGPOGPO-$CONH_2$	16		6.7−15.1
T1-904[f]	Ac-GARGPAGPQ**G**PRGDKGET GPOGPOGPOGPO-$CONH_2$	30.8	38.9	
T1-904[G10 S][f]	Ac-GARGPAGPQ**S**PRGDKGET GPOGPOGPOGPO-$CONH_2$	8.9		6.4−20.1

[a] Experimental T_m values were obtained from temperature-induced denaturation monitored by CD at either 224 or 225 nm. Peptides were equilibrated for at least 48 h at low temperatures in PBS buffer, pH 7. Thermal transitions were recorded under our standard conditions. [b] Predictions are based on a host−guest system (*7*, *34*) taking into account electrostatic interactions (*35*). [c] Predicted T_m C-terminal applies the prediction methods for the host−guest system to only those triplets that are C-terminal to the mutation site. Calculations are done both excluding and including the actual triplet which contains the mutation with the value always coming out higher when the Gly−X−Y triplet including the mutation is considered. Both values are given for each mutant peptide, as a range: T_m value excluding triplet with mutation − T_m value of C-terminal sequence including triplet with mutation. [d] Using only the stability contributions of individual Gly−X−Y tripeptide units, the calculated thermal stability is −3.8 °C, predicting peptide MBL 45-62 would not form a stable triple-helix. Adding 15.4 °C for each KGE sequence (Persikov et al., 2004), the predicted T_m value is 27.0 °C, which is higher than the observed 17.9 °C value, suggesting that the contribution of a KGEKGE sequence is less than the sum of two independent KGE units. [e] Mohs, A., and Brodsky, B., unpublished data. The sequence was designed to model the region in the type VII collagen surrounding a Gly to Arg mutation site at position 2043 that leads to the dystrophic form of epidermolysis bullosa (*10*). [f] Ref *25*.

to the mutation site give a predicted T_m consistent with that observed for the mutant peptide MBL[G10D] (Table 2). The T_m value of 15.9 °C observed for the MBL 45−61 mutant peptide is between the observed T_m value of 11.9 °C for the 6 C-terminal Gly−X−Y units (MBL 57−61), and the observed T_m value of 16.7 °C for the 7 C-terminal tripeptide Gly−X−Y units (MBL 54−61), suggesting that part of the GTK tripeptide unit may still be in triple-helical form and supporting the involvement of a KGE interaction for the lysine in this triplet.

The method of predicting the observed T_m of the native MBL peptide and the MBL peptide with a mutation suggests that it may be possible to use a similar approach to other triple-helical peptides with Gly replacements. It has proven difficult to obtain a consistent interpretation for the effect of Gly replacements in peptides modeling OI and other collagen diseases, even though these peptides follow a common design $(GXY)_6(GPO)_4$, with the Gly mutation introduced in the fourth Gly from the N-terminus (*25*, *31*). For instance, the introduction of a Gly to Ser replacement in peptide T1−892 results in an 11 °C drop in the T_m value, while a Gly to Ser replacement in peptide T1−904 of similar design leads to a decrease in the T_m value of 22 °C (*25*). At one extreme, there are cases where a peptide cannot form a triple-helix following the introduction of a Gly replacement (e.g., peptide T7−2031, Table 2), and on the other end of the spectrum is the case of MBL 45−61 reported here, where the Gly to Asp replacement causes only a very small decrease in stability. Applying a sequence-stability analysis described previously for MBL peptides to peptides modeling OI mutations in type I collagen and DEB mutation in type VII collagen helps explain the drop in stability caused by a Gly replacement (Table 2). This supports the concept that in these peptides with a Gly mutation and a disordered region N-terminal to the mutation, it is the sequence and length of residues C-terminal to the mutation site that determine the triple-helix stability.

The current peptide studies complement other investigations on the oligomerization of MBL and its interactions with MASPs (*2*, *16*, *19*−*21*). The studies reported here suggest a previously unrecognized feature of the N-terminal triple-helical region of MBL: the presence of a highly stable KGEKGE triple-helix adjacent to a much less stable triple-helix, which may have imperfect hydrogen bonding making little enthalpic contributions. This triple-helix region N-terminal to the kink, together with the adjacent Cys-rich domain, is known to have a critical role in determining the self-association of MBL trimers into higher order oligomers (*12*, *13*). In the native MBL, it is likely that constraints from the Cys-rich N-terminus support the triple-helix nature of the N-terminal domain, but the inherent low stability of the tripeptide units on the N-terminus of the triple-helix domain may play a role in the oligomerization process. Studies on the collagen triple-helix have indicated that microunfolding is essential for self-assembly of collagen molecules in fibril formation (*36*, *37*), and low triple-helix stability seems to be a requirement for ligand binding by triple-helix domains in many cases (*38*−*41*). The low triple-helix propensity of the GINGFOGKDGRD sequence, together with its highly conserved hydrophobic and charged nature, make this sequence a strong candidate for a role in MBL oligomerization (Figure 4).

The studies presented here suggest that the expected highly disruptive effect of a Gly → Asp substitution on a triple-helix, as observed for simple model peptides (*23*), may be modulated by its occurrence at the boundary between a highly stable and a weak triple-helix region (Figure 4B). The introduction of a Gly → Asp mutation at this discontinuity could result in further disordering of the already unstable triple-helix at the N-terminus, impairing its ability to self-associate but with little effect on the overall stability of the molecule. This hypothesis is consistent with the observation that recombinant mutant MBL appears to form a triple-helix with an enthalpy similar to that of the native state, but a somewhat reduced stability, and that oligomerization is impaired (*16*, *19*−*21*). Other stability junctions along the triple-helix, such as that seen at the collagenase cleavage site in fibrillar collagens (*41*), could have implications for biological function.

ACKNOWLEDGMENT

Dr. Yujia Xu collected preliminary NMR data on mannose binding protein model peptides. We thank Dr. Anton Persikov for helpful discussions and Dr. Seho Kim for assistance in NMR interpretation.

REFERENCES

1. Myllyharju, J., and Kivirikko, K. I. (2001) Collagens and collagen-related diseases, *Ann. Med. 33*, 7−21.
2. Turner, M. W. (2003) The role of mannose-binding lectin in health and disease, *Mol. Immunol. 40*, 423−429.
3. Kielty, C. M., and Grant, M. E. (2002) The Collagen Family: Structure, Assembly, and Organization in the Extracellular Matrix, in *Connective tissue and its heritable disorders molecular, genetic, and medical aspect* (Royce, P. M., Steinmann, B., Eds.) pp 159−222, New York, Wiley-Liss, Inc.
4. Rich, A., and Crick, F. H. (1961) The molecular structure of collagen, *J. Mol. Biol. 3*, 483−506.
5. Ramachandran, G. N. (1964) in *Treatise on Collagen* (Ramachandran, G. N., Ed.) Vol. 1, pp 103−183, Academic Press, New York.
6. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution, *Science 266*, 75−81.
7. Persikov, A. V., Ramshaw, J. A., and Brodsky, B. (2000) Collagen model peptides: Sequence dependence of triple-helix stability, *Biopolymers 55*, 436−50.
8. Myllyharju, J., and Kivirikko, K. I. (2004) Collagens, modifying enzymes, and their mutations in humans, flies, and worms, *Trends Genet. 20*, 33−43.
9. Byers, P. H., and Cole, W. G. (2002) Osteogenesis Imperfecta, in *Connective tissue and its heritable disorders molecular, genetic, and medical aspect* (Royce, P. M., Steinmann, B., Eds.) pp 385−430, New York, Wiley-Liss, Inc.
10. Pulkkinen, L., and Uitto, J. (1999) Mutation analysis and molecular genetics of epidermolysis bullosa, *Matrix Biol. 18*, 29−42.
11. Petry, F., Hauptmann, G., Goetz, J., Grosshans, E., and Loos, M. (1997) Molecular basis of a new type of C1q-deficiency associated with a nonfunctional low molecular weight (LMW) C1q: parallels and differences to other known genetic C1q defects, *Immunopharmacology 38*, 189−201.
12. Presanis, J. S., Kojima, M., and Sim, R. B. (2003) Biochemistry and genetics of mannan-binding lectin (MBL), *Biochem. Soc. Trans. 31*, 748−752.
13. Hoppe, H. J., and Reid, K. B. (1994) Collectins—soluble proteins containing collagenous regions and lectin domains—and their roles in innate immunity, *Protein Sci. 3*, 1143−1158.
14. Wallis, R. (2002) Structural and functional aspects of complement activation by mannose-binding protein, *Immunobiology 205*, 433−445.
15. Wallis, R., and Drickamer, K. (1999) Molecular determinants of oligomer formation and complement fixation in mannose-binding proteins, *J. Biol. Chem. 274*, 3580−3589.
16. Larsen, F., Madsen, H. O., Sim, R. B., Koch, C., and Garred, P. (2004) Disease-associated mutations in human mannose-binding lectin compromise oligomerization and activity of the final protein, *J. Biol. Chem. 279*, 21302−21311.
17. Super, M., Thiel, S., Lu, J., Levinsky, R. J., and Turner, M. W. (1989) Association of low levels of mannan-binding protein with a common defect of opsonization, *Lancet. 334*, 1236−1239.
18. Summerfield, J. A., Sumiya, M., Levin, M., and Turner, M. W. (1997) Association of mutations in mannose binding protein gene with childhood infection in consecutive hospital series, *Br. Med. J. 314*, 1229−1232.
19. Wallis, R., and Cheng, J. Y. (1999) Molecular defects in variant forms of mannose-binding protein associated with immunodeficiency, *J. Immunol. 163*, 4953−4959.
20. Wallis, R., and Dodd, R. B. (2000) Interaction of mannose-binding protein with associated serine proteases. Effects of naturally occurring mutations, *J. Biol. Chem. 275*, 30962−30969.
21. Wallis, R., Shaw, J. M., Uitdehaag, J., Chen, C. B., Torgersen, D., and Drickamer, K. (2004) Localization of the serine protease-binding sites in the collagen-like domain of mannose-binding protein: indirect effects of naturally occurring mutations on protease binding and activation, *J. Biol. Chem. 279*, 14065−14073.
22. Baum, J., and Brodsky, B. (1999) Folding of peptide models of collagen and misfolding in disease, *Curr. Opin. Struct. Biol. 9*, 122−128.
23. Beck, K., Chan, V. C., Shenoy, N., Kirkpatrick, A., Ramshaw, J. A., and Brodsky, B. (2000) Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine, *Proc. Natl. Acad. Sci. U.S.A. 97*, 4273−4278.
24. Persikov, A. V., Pillitteri, R. J., Amin, P., Schwarze, U., Byers, P. H., and Brodsky, B. (2004) Stability related bias in residues replacing glycines within the collagen triple-helix (Gly−Xaa−Yaa) in inherited connective tissue disorders, *Hum. Mutat. 24*, 330−337.
25. Yang, W., Battineni, M. L., and Brodsky, B. (1997) Amino acid sequence environment modulates the disruption by osteogenesis imperfecta glycine substitutions in collagen-like peptides, *Biochemistry 36*, 6930−6935.
26. Persikov, A. V., Xu, Y., and Brodsky, B. (2004) Equilibrium thermal transitions of collagen model peptides, *Protein Sci. 13*, 893−902.
27. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc. 114*, 10663−10665.
28. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of DRK in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques, *J. Biomol. NMR 4*, 845−858.
29. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry 33*, 5984−6003.
30. Venugopal, M. G., Ramshaw, J. A., Braswell, E., Zhu, D., and Brodsky, B. (1994) Electrostatic interactions in collagen-like triple-helical peptides, *Biochemistry 33*, 7948−7956.
31. Liu, X., Kim, S., Dai, Q. H., Brodsky, B., and Baum, J. (1998) Nuclear magnetic resonance shows asymmetric loss of triple-helix in peptides modeling a collagen mutation in brittle bone disease, *Biochemistry 37*, 15528−15533.
32. Buevich, A., and Baum, J. (2001) Nuclear magnetic resonance characterization of peptide models of collagen-folding diseases, *Philos. Trans. R. Soc. London, Ser. B. 356*, 159−168.
33. Buevich, A. V., and Baum, J. (2002) Residue-specific real-time NMR diffusion experiments define the association states of proteins during folding, *J. Am. Chem. Soc. 124*, 7156−7162.
34. Persikov, A. V., Kirkpatrick, A., and Brodsky, B. (2002) Peptide investigations of pairwise interactions in the collagen triple-helix, *J. Mol. Biol. 316*, 385−394.
35. Persikov, A. V., Shamshaw, J. A. M., Kirkpatrick, A., and Brodsky, B. (2005) Electrostatic interactions involving lysine make major contributions to collagen triple-helix stability, *Biochemistry*, *44*, in press.
36. Kadler, K. E., Hojima, Y., and Prockop, D. J. (1988) Assembly of type I collagen fibrils de novo. Between 37 and 41 °C the process is limited by micro-unfolding of monomers, *J. Biol. Chem. 263*, 10517−10523.
37. Leikina, E., Mertts, M. V., Kuznetsova, N., and Leikin, S. (2002) Type I collagen is thermally unstable at body temperature, *Proc. Natl. Acad. Sci. U.S.A. 99*, 1314−1318.
38. Shah, N. K., Sharma, M., Kirkpatrick, A., Ramshaw, J. A., and Brodsky, B. (1997) Gly−Gly containing triplets of low stability adjacent to a type III collagen epitope, *Biochemistry 36*, 5878−5883.
39. Deprez, P., Doss-Pepe, E., Brodsky, B., and Inestrosa, N. C. (2000) Interaction of the collagen-like tail of asymmetric acetylcholinesterase with heparin depends on triple-helical conformation, sequence, and stability, *Biochem. J. 350*, 283−290.
40. Sacca, B., Fiori, S., and Moroder, L. (2003). Studies of the local conformational properties of the cell-adhesion domain of collagen type IV in synthetic heterotrimeric peptides, *Biochemistry 42*, 3429−3436.
41. Fields, G. (1991). A model for interstitial collagen catabolism by mammalian collagenases, *J. Theor. Biol. 153*, 585−602.

BI0482708