Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine

Konrad Beck*†, Virginia C. Chan*, Nigel Shenoy*, Alan Kirkpatrick‡, John A. M. Ramshaw‡, and Barbara Brodsky*§

*Department of Biochemistry, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School, Piscataway, NJ 08854; and ‡CSIRO, Division of Molecular Science, Parkville, Victoria 3052, Australia

Communicated by Elkan R. Blout, Harvard Medical School, Cambridge, MA, February 3, 2000 (received for review July 10, 1999)

Mutations resulting in replacement of one obligate Gly residue within the repeating (Gly-Xaa-Yaa)$_n$ triplet pattern of the collagen type I triple helix are the major cause of osteogenesis imperfecta (OI). Phenotypes of OI involve fragile bones and range from mild to perinatal lethal. In this study, host–guest triple-helical peptides of the form acetyl-(Gly-Pro-Hyp)$_3$-*Zaa*-Pro-Hyp-(Gly-Pro-Hyp)$_4$-Gly-Gly-amide are used to isolate the influence of the residue replacing Gly on triple-helix stability, with Zaa = Gly, Ala, Arg, Asp, Glu, Cys, Ser, or Val. Any substitution for Zaa = Gly (melting temperature, T_m = 45°C) results in a dramatic destabilization of the triple helix. For Ala and Ser, T_m decreases to ≈10°C, and for the Arg-, Val-, Glu-, and Asp-containing peptides, T_m < 0°C. A Gly → Cys replacement results in T_m < 0°C under reducing conditions but shows a broad transition (T_m ≈ 19°C) in an oxidizing environment. Addition of trimethylamine *N*-oxide increases T_m by ≈5°C per 1 M trimethylamine *N*-oxide, resulting in stable triple-helix formation for all peptides and allowing comparison of relative stabilities. The order of disruption of different Gly replacements in these peptides can be represented as Ala ≤ Ser < CPO$_{red}$ < Arg < Val < Glu ≤ Asp. The rank of destabilization of substitutions for Gly in these Gly-Pro-Hyp-rich homotrimeric peptides shows a significant correlation with the severity of natural OI mutations in the α1 chain of type I collagen.

Osteogenesis imperfecta (OI) or "brittle bone disease" is characterized by a general decrease in bone mass resulting in multiple fractures, and its inheritance is in most cases autosomal dominant (for reviews, see refs. 1 and 2). The clinical phenotype has been subdivided into four types (3) with type I as the mildest and most common form characterized by prepubertal fractures and slight growth retardation. Type III and IV OI patients have recurrent fractures, progressive limb deformities, and severe growth deficiencies. The most severe form, OI type II, is usually perinatal lethal. Infants have short limbs, crumpled long bones, and multiple fractured ribs. In nearly all cases, OI is caused by mutations in the genes *COL1A1* and *COL1A2* coding for the two chains α1(I) and α2(I) of procollagen type I.

Collagen type I is the major structural protein of the extracellular matrix of bone, skin, and tendon (4). It consists of two α1 and one α2 chains that are synthesized in the rough endoplasmic reticulum in a procollagen form with N- and C-terminal extensions. The central region contains 338 repeating Gly-Xaa-Yaa triplets, with proline as the residue most frequently incorporated into the Xaa and Yaa position. In the Yaa position, Pro is generally enzymatically modified to 4-hydroxyproline (denoted by O or Hyp throughout), making Gly-Pro-Hyp the most common, as well as the most stabilizing, triplet in collagen (5). The presence of Gly as every third residue is considered essential to form the characteristic collagen triple helix. In this conformation, three left-hand polyproline II-like helical chains are wound around each other to form a tightly packed right-handed superhelix. Only Gly residues can be accommodated without distortion as every third residue near the center of this supercoiled helix (6–8).

Although RNA splicing variations, gene deletions, insertions, and duplications are observed, the majority of OI cases are caused by single base exchanges resulting in a replacement of Gly with another amino acid (1, 2, 9, 10). Hundreds of distinct Gly substitution mutations have been observed in different OI cases, occurring in either the α1(I) or α2(I) chain. Seven different amino acids, whose codons differ from the codon of Gly in one nucleotide, have been observed as replacement residues at many sites along both chains. Some substitutions lead to lethal cases, and others result in milder forms; however, no simple correlation has been observed relating clinical severity with the chain type, mutation site, flanking sequences, or residue substituted for Gly.

Studies on OI fibroblasts have shown excess posttranslational glycosylation and lysine hydroxylation N-terminal to the OI collagen mutation site, reflecting delayed folding and the C- to N-terminal directionality of collagen triple-helix propagation (1, 5, 11). Increased intracellular breakdown, reduced collagen secretion, and compromised helix stability are also observed for some OI collagens (1, 12–14). In a few cases of Gly → Cys or Gly → Asp substitutions resulting in lethal OI, a pronounced kink within the triple helix is observed and has been shown to interfere with proper fibril formation (15–18).

Collagen-like peptides allow characterization of the effect of Gly substitutions within a defined sequence environment in a homogeneous population of triple-helical molecules. The influence of those residues replacing Gly in OI mutations is reported herein for a host–guest peptide model. The host peptide consists of eight Gly-Pro-Hyp triplets, a model of the most stable collagenous environment (19). Near the center, one Gly residue is replaced with Ala, Arg, Asp, Cys, Glu, Ser, or Val, and the effect on stability is analyzed. As expected, the thermal stability was decreased drastically by a Gly substitution. The decrease in triple-helix stability critically depends on the nature of the residue replacing Gly. Comparison of the peptide results with the OI phenotype for mutations in the α1(I) chain indicates that the most destabilizing residues correlate with the most severe clinical symptoms, whereas the less disruptive residues show more varied phenotypes.

Materials and Methods

Peptide Synthesis and Purification. Peptides were synthesized by using fluorenylmethoxycarbonyl chemistry, purified, and identified by mass spectrometry as described (19–22).

Abbreviations: OI, osteogenesis imperfecta; TMAO, trimethylamine *N*-oxide; O or Hyp, 4-hydroxyproline.

†Present address: Institut de Biologie et Chemie des Proteins, Centre National de la Recherche Scientifique, Unité Propre de Recherche 412, 7 Passage du Vercors, 69367 Lyon, France.

§To whom reprint requests should be addressed. E-mail: brodsky@umdnj.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.070050097. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.070050097

Sample Preparation. Lyophilized peptides were dried *in vacuo* over P_2O_5 for 2 days before weighing to prepare 10 mg/ml stock solutions in water. At least 2 days before measurements, samples were adjusted to 1 mg/ml in 150 mM NaCl/20 mM KH_2PO_4/NaOH, pH 7.2 (PBS) by dilution with concentrated buffer solutions and stored in an ice bath. For the Cys-containing peptide, measurements were performed in the absence and presence of 5 mM DTT, which was added to avoid disulfide bond formation. The monomeric state of this peptide under reducing conditions before and after melting experiments was confirmed by mass spectroscopy, whereas monomers and dimers were observed in the absence of DTT. Trimethylamine *N*-oxide (TMAO; Sigma) concentrations were varied by addition of a 4.1 M stock solution in PBS.

CD Spectroscopy. CD spectra were recorded in 1-mm quartz cuvettes on an Aviv 62DS spectropolarimeter equipped with a five-cell thermostated cell holder and a temperature controller (Aviv Associates, Lakewood, NJ). Samples were transferred into temperature-adjusted cells with precooled pipette tips. Spectra were collected from 260 to ≈210 nm in 0.5-nm intervals at 0°C, corrected for buffer baselines measured in the same cells, and normalized to mean residue ellipticities $[\Theta]_{MRW}$ with M_r derived from the sequences. Thermal transition curves were recorded at 225 nm starting at 0°C (and at −5°C for peptide DPO in 3.7 M TMAO). The temperature was increased in 0.3°C intervals with equilibration and measuring times of 90 and 20 s, respectively, resulting in average heating rates of 4–5°C/h. Spectra recorded 6 to 10 h after finishing the melt, during which time the samples were equilibrated at 0°C, showed a CD signal recovery of greater than 90% at 225 nm.

Thermodynamic Analysis. Thermal transition curves were normalized to the fraction of triple-helical peptide F with $F = [\Theta - \Theta_u(T)]/[\Theta_n(T) - \Theta_u(T)]$, where Θ_n and Θ_u represent ellipticities of the fully folded and unfolded species, respectively, corrected for their temperature dependence by linear extrapolation of the low and high temperature baselines; Θ is the observed ellipticity; and T is temperature. Assuming a two-state mechanism in which three unfolded chains u combine to a triple-helical molecule n results in an equilibrium constant $K = c_n/c_u{}^3 = F/3\,c_0{}^2\,(1 - F)^3$ where $c_0 = c_u + 3c_n$ is the total peptide concentration and $F = 3c_n/c_0$ the fraction of folded peptide (23). This assumption seems justified, because a two-state transition was verified for related peptides by analytical ultracentrifugation (24). With $\Delta G^0 = \Delta H^0 - T\Delta S^0 = -RT \ln K$, it follows that

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

where ΔG^0, ΔH^0, ΔS^0, R, T, and T_m are the standard free energy, enthalpy, entropy, gas constant, absolute temperature, and melting temperature, respectively. T_m and ΔH^0 were derived by fitting the normalized transition curves to Eq. **1** by using a nonlinear least-squares algorithm.

Results

Peptide Design. Single base changes causing substitutions of Gly within the Gly-Xaa-Yaa triplet are the most common type of mutations within the type I collagen genes leading to OI. Changes of one nucleotide within a Gly codon can result in codons for Ala, Arg, Asp, Cys, Glu, Ser, Val, or Trp or for a premature stop codon. All such mutations except Trp or a stop codon have been found in different cases of OI, with Ser and Cys replacements accounting for the most frequently observed mutations (10).

To test for the influence of different replacement residues on the stability of the collagen triple helix, a set of host–guest peptides was synthesized in which one Gly residue is substituted by all of the replacement residues seen in OI cases. A set of host–guest model peptides of the form acetyl-(Gly-Pro-Hyp)$_3$-Gly-Xaa-Yaa-(Gly-Pro-Hyp)$_4$-Gly-Gly-amide has been studied extensively in our laboratory. This design embeds a guest triplet Gly-Xaa-Yaa near the middle of a Gly-Pro-Hyp-rich environment and has an acetylated N terminus and amidated C terminus to eliminate interactions between guest triplets and ionized ends (25, 26). The host peptide acetyl-(Gly-Pro-Hyp)$_8$-Gly-Gly-amide has a T_m = 45°C, and all of the more than 40 peptides with different L-amino acids in the Xaa and Yaa positions studied thus far formed stable triple helices with melting temperatures in the 20–45°C range (19–22, 27). For this study, the Gly residue in the guest triplet is replaced by another residue, Zaa, with Xaa and Yaa being Pro and Hyp, respectively, yielding a guest triplet Zaa-Pro-Hyp. For comparison, peptides with Ser in the Xaa and Yaa position were included in this study. In the following discussion, peptides are named by their guest triplet in single-letter code, with O used to designate Hyp.

Peptide Conformation and Stability. Far-UV CD spectra recorded at 0°C in PBS showed maxima of varying magnitudes at 225 nm for all host–guest peptides studied (Fig. 1). The host peptide GPO (19) as well as peptides GSO and GPS showed values of $[\Theta]_{225} \approx 4{,}000$ deg·cm^2·dmol^{-1} in native conditions, indicating a fully triple-helical conformation, as also found for other related host–guest peptides (19–22, 27). The corresponding values for peptides APO (3,600 deg·cm^2·dmol^{-1}), CPO under oxidizing conditions (3,200 deg·cm^2·dmol^{-1}), and SPO (3,200 deg·cm^2·dmol^{-1}) are lower but are still in the range expected for a triple helix, whereas the CD amplitudes of peptides CPO$_{red}$ (1,600 deg·cm^2·dmol^{-1}), DPO (2,000 deg·cm^2·dmol^{-1}), EPO (1,700 deg·cm^2·dmol^{-1}), RPO (1,600 deg·cm^2·dmol^{-1}), and VPO (1,300 deg·cm^2·dmol^{-1}) are comparable with the CD amplitude of a single-chain polyproline II helix (24, 28, 29).

Thermal stability was determined by monitoring the change of ellipticity at 225 nm on heating (Fig. 2). The host peptide GPO (19) as well as peptides GSO (Fig. 3*A*) and GPS show sharp thermal transitions from the trimer (native) to monomer (denatured) state, with T_m values of 45, 38, and 35°C, respectively (Table 1). Peptides APO and SPO also show cooperative melting, but the thermal stability dramatically decreases to about 10°C.

In contrast to the cooperative melting of these peptides, a linear change of ellipticities with increasing temperature is recorded for DPO, EPO, RPO, and VPO, suggesting a denatured conformation even at low temperatures (Fig. 2*A*). This result, together with the low magnitude of the CD signal at 0°C (Fig. 1*A*), indicates a polyproline II-like structure of noninteracting peptide chains. To test whether the lack of triple-helix formation for DPO, RPO, and EPO is caused by unfavorable ionic interactions between the charged side chains, melting experiments were performed at pH 1.8 and pH 12.9 as well as in the presence of 2 M NaCl. Charge repulsion did not seem to play a major role in triple-helix destabilization, because CD amplitudes remained in the same range as was found for PBS and because melting curves still showed a linear dependence of the temperature. APO used as a control peptide showed no significant change in T_m values under these conditions.

The peptide CPO$_{red}$ shows an almost linear change in ellipticity with temperature, with just the final phase of a cooperative transition detectable near 0°C (Fig. 2*A*), indicating that a Gly → Cys$_{red}$ substitution prevents triple-helix formation. However, in the absence of DTT, CPO has a broad ($\Delta H^0 = -240$ kJ/mol) thermal transition, centered near 18°C, suggesting that formation of disulfide bonds plays a stabilizing role, although the nature of the product(s) has not been defined by this study.

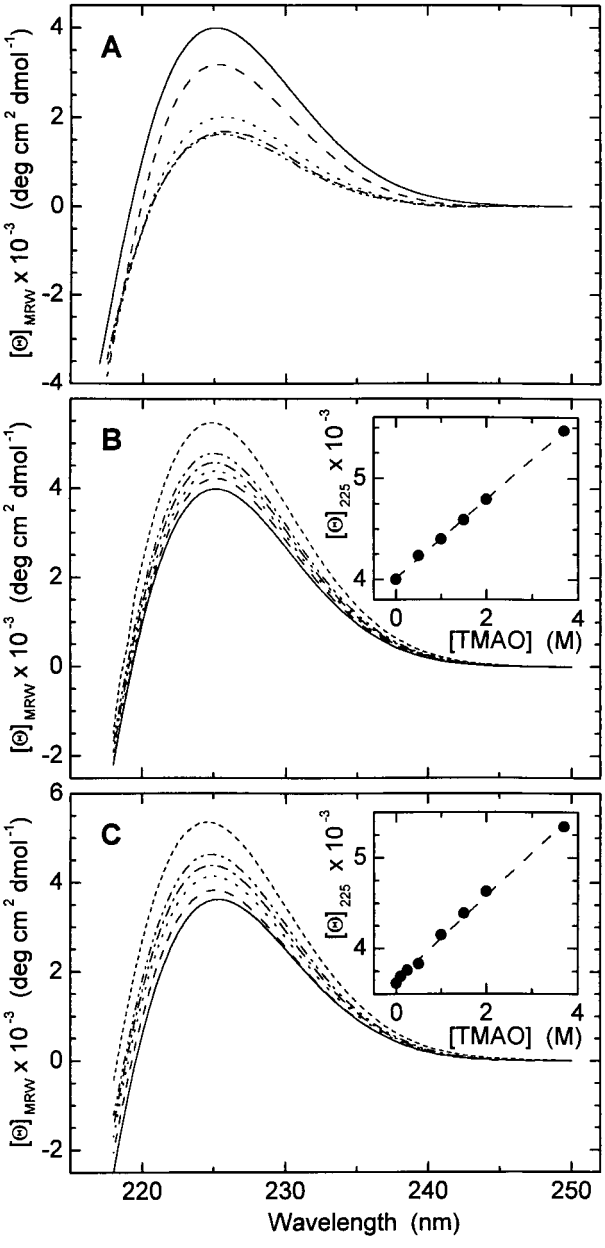

Fig. 1. Far-UV CD spectra of host–guest peptides. CPO_{red} is used to indicate the peptide with a Cys-Pro-Hyp guest triplet under reducing conditions. (*A*) Spectra are shown for peptides GPS (—), SPO (----), CPO_{red} (-·-·-), DPO (·····), and RPO (-···-) recorded at 0°C in PBS (concentration = 1 mg/ml). Corresponding spectra are shown for peptides GSO (*B*) and APO (*C*) recorded at TMAO concentrations of (from bottom to top at 225 nm) 0, 0.5, 1.0, 1.5, 2.0, and 3.7 M. (*Insets*) The increase of the molar ellipticity at 225 nm observed on increasing the concentrations of TMAO.

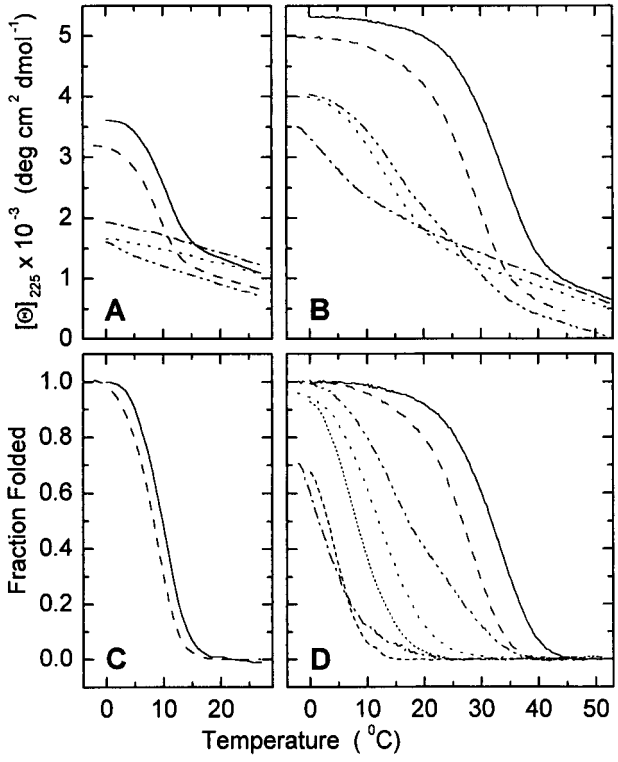

Fig. 2. Thermal denaturation of host–guest peptides. The change of ellipticity on heating was monitored at 225 nm for peptides APO (—), SPO (----), CPO_{red} (-·-·-), RPO (·····), and DPO (-··-··) (concentration = 1 mg/ml) in PBS in the absence (*A*) and presence (*B*) of 3.7 M TMAO. Transition curves were normalized to the fraction of folded peptides for profiles in the absence (*C*) and presence (*D*) of TMAO. Transition profiles for peptides VPO (short dotted line) and EPO (short dashed line) are included in *D*.

Influence of TMAO on Peptide Conformation and Stability. To test whether peptides DPO, EPO, RPO, VPO, and CPO_{red} could form triple helices, measurements were performed with different concentrations of TMAO as a cosolvent. TMAO is a naturally occurring osmolyte that counteracts the destabilizing interaction of proteins with urea in different marine elasmobranchs (30). It was also shown to increase the melting temperature of proteins *in vitro* (31, 32) and is effective in correcting temperature-sensitive protein folding abnormalities *in vivo* (33). In contrast to various polyols that stabilize collagens and collagenous peptides (23, 29, 34), TMAO has an excellent far-UV transparency.

In 3.7 M TMAO, all or part of a transition curve was observed for the entire set of Gly → Zaa peptides, allowing comparison of their relative stabilities (Fig. 2*D*; Table 1). For CPO under reducing conditions, a very broad transition was seen, with a clear biphasic nature at lower values of TMAO; this result is likely a consequence of the oxidizing effect of TMAO (35), resulting in a mixture of reduced and oxidized species. Peptides, such as APO, gave melting temperatures that linearly increased with TMAO concentration, as did host–guest peptides GSO and GPS (Fig. 3). All host–guest peptides studied showed relatively constant values of the increase in T_m per mol of TMAO (ΔT_m/[TMAO]) between 4.4 and 5.9°C (Fig. 3 *Insets*; Table 1), which are similar to the ≈4°C reported for ribonuclease (31, 32). These values could be used to extrapolate apparent T_m values in the absence of TMAO for peptides DPO, EPO, RPO, VPO, and CPO_{red} (Table 1). The increase in T_m as a result of TMAO is accompanied generally by a decrease in the absolute values of the transition enthalpies (Table 1); such a decrease is evident from the broadening of the transition curves (Fig. 2, compare *A* and *C* with *B* and *D*, respectively).

Besides its effect on the thermostability of the host–guest peptides, TMAO also increases the amplitude of the CD signal. The maximum ellipticities at 225 nm recorded at 0°C show a linear increase of ≈500 deg·cm^2·dmol^{-1} per 1 M of TMAO, depending on the peptide (Fig. 1 *B Inset* and *C Inset*; Fig. 2, compare also *A* and *B*).

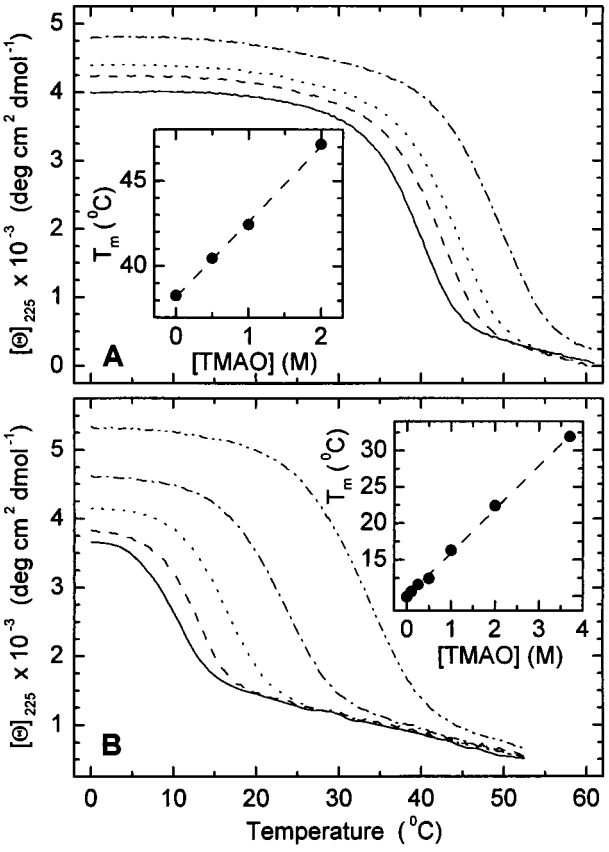

Fig. 3. Thermal denaturation of host–guest peptides at varying concentrations of TMAO. The change of ellipticity on heating was monitored at 225 nm for peptides GSO (*A*) and APO (*B*) (concentration = 1 mg/ml) in PBS (—) and after addition of 0.5 (----), 1.0 (·····), 2.0 (-·-·-), and (for APO) 3.7 (-···-) M TMAO. (*Insets*) The increase in melting temperature on increasing TMAO concentrations.

Discussion

More than 250 Gly → Zaa OI mutations in type I collagen have been determined for different individuals whose phenotypes vary from lethal to mild forms (1, 2, 9, 10, 12). The mutation sites range over the 1,014-residue triple helix of both the α1(I) and α2(I) chains, with Gly → Ser and Gly → Cys substitutions being the most common replacements (Fig. 4). Although it is not yet possible to predict the severity of an OI phenotype for a given Gly substitution at a specific site, factors that are thought to be important include the nature of the residue substituted for Gly, the chain in which the mutation occurs [α1(I) or α2(I)], the location of the mutation with respect to the C terminus of the triple helix, and the amino acid sequence environment adjacent to the mutation site (1, 36, 37).

Different models have been proposed that focus on the location of the mutation and its environment. The gradient model emphasizes that changes within the C-terminal three-quarters of the molecule are usually more severe than those within the N-terminal region and suggests a relation to the C to N directionality of collagen triple-helix folding (1, 14, 38). With an increase in the number of OI cases, this general correlation has held up for Gly → Arg and Gly → Val mutations in the α1(I) chain, whereas all Gly → Asp mutations reported in the αl(I) chain are lethal (Fig. 4). Mutations resulting in Ser or Cys show both lethal and nonlethal effects all along the α1(I) chain, except for a region with only nonlethal Cys mutations near the N terminus. The α2(I) chain does not have the polarity seen for α1(I) but seems to have clusters of lethal mutation sites alternating with clusters of milder OI sites along the chain. This pattern led to the suggestion that regional factors related to local sequence stability or binding sites must be important in determining phenotype (37, 39). Local sequences N-terminal to the mutation site could also influence the propensity for renucleation, if the Gly substitution has interrupted the C- to N-terminal propagation of the triple helix (40, 41). Calculations of relative stabilities, assigned on the basis of the frequency of different tripeptides in collagen, suggest that mutations within a very stable local environment result in a more severe clinical effect (36, 42), and studies on peptides have supported the influence of local triplets surrounding a mutation on its disruptive effect (43, 44). Cooperative blocks that undergo local breathing or microunfolding could also be destabilized selectively by Gly substitutions and prevent folding of the molecule (2, 12).

Although studies on OI collagens have indicated that the residue substituted for Gly impacts on its consequences, the influence of other factors, such as the location of the mutation site with respect to the C terminus and the local surrounding sequences, and the presence of a heterogeneous population with respect to the mutated chains, has obscured any simple corre-

Table 1. Melting temperatures T_m and transition enthalpies ΔH^0 of host–guest peptides as determined by CD spectroscopy

	0 M TMAO		3.7 M TMAO		ΔT_m/[TMAO],
Peptide	T_m (°C)*	ΔH^0 (kJ/mol)	T_m (°C)	ΔH^0 (kJ/mol)	K/mol/liter
GPO†	44.5	−350	n.d.	n.d.	n.d.
GSO	38.3	−510	47.1‡	−410‡	4.4
GPS	34.8	−440	n.d.	n.d.	n.d.
SPO	8.6	−590	26.5	−350	4.8
APO	9.9	−550	31.9	−380	5.9
CPO_{red}	[−3]	—	18.1	−190	5.6
RPO	[−7]	—	12.2	−320	5.1
VPO	[−10]	—	8.2	−360	n.d.
EPO	[−16]	—	3	−380	n.d.
DPO	[−17]	—	2	−300	n.d.

Data are derived from CD melting curves recorded at 225 nm at peptide concentrations of 1 mg/ml in PBS in the absence and presence of TMAO. Values for the change of T_m with increasing amounts of TMAO are based on measurements performed at varying concentrations by assuming a linear relationship. n.d., not determined.
*Numbers in brackets are obtained from extrapolation of TMAO data to 0 M TMAO.
†Values from ref. 22 are included for comparison.
‡Values are for 2 M TMAO, because peptide GSO precipitates at higher osmolyte concentrations.

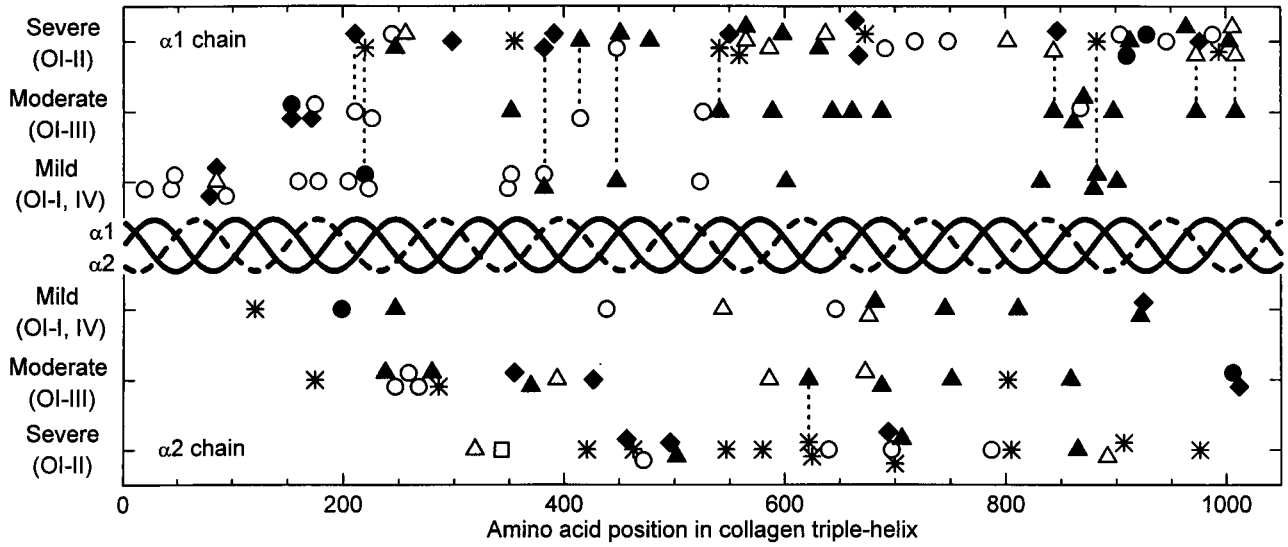

Fig. 4. Distribution of Gly → Xaa mutations within the collagen type I triple-helical region causing OI. Residues replacing Gly within the Gly-Xaa-Yaa triplets are indicated (●, Ala; ♦, Arg; *, Asp; ○, Cys; □, Glu; ▲, Ser; △, Val) and classified as mild, moderate, or severe. Mutations are those collected in the collagen mutation database (www.le.ac.uk/genetics/collagen/; ref. 10), and further unpublished ones are included with permission of the group of A. De Paepe (personal communication). Dotted lines indicate positions for which different mutations are known where the severity of the disease depends on the identity of the encoded amino acid residue.

lation. The host–guest peptide study reported herein isolates the influence of the particular residue replacing Gly on triple-helix disruption by providing a constant environment and a homogenous population of molecules. Our results indicate a dramatic destabilization when a Gly near the middle of a (Gly-Pro-Hyp)-rich triple-helix model peptide is replaced by any of the seven residues found in OI mutations. The destabilization was so great that five of the seven peptides in the Gly → Zaa set did not form triple helices in aqueous solution even at 0°C. To compare the degree of destabilization of different Gly replacements, TMAO was used as a stabilizing cosolvent, which allowed the observation of thermal transitions for all peptides (Fig. 2*A*; Table 1). Its relatively uniform stabilizing influence (Fig. 3; Table 1) for all Gly → Zaa peptides as well as other uninterrupted host–guest triple-helical peptides is consistent with its protective influence originating from an unfavorable interaction of the protein backbone with TMAO (45) and with its effect being nearly independent of the residue side chains (46). In view of previous studies supporting a compaction of protein structure in the presence of TMAO, the increase in the magnitude of the positive Gaussian CD band at 225 nm in these peptides is likely to arise from a tighter packing of the triple helix (28, 29, 47, 48). Because of the high surface-to-volume ratio of the triple-helical conformation, such forces could have a drastic effect on the packing of the collagenous peptides and result in the observed large increase of the CD amplitude.

Comparison of the T_m values of the seven Gly → Zaa peptides in 3.7 M TMAO or of their values extrapolated to 0 M TMAO indicates that the order of disruption can be represented as Ala ≤ Ser < Cys_{red} < Arg < Val < Glu ≤ Asp. At high TMAO concentrations, the magnitude of the CD maximum at 225 nm is lowered for all substituted peptides relative to the host peptide, suggesting that the conformation deviates from a standard triple-helical structure. Structural alterations could relate to the breaking of direct NH···CO peptide bonds and local untwisting at the substitution site, as seen in the crystal structure of a triple-helical peptide containing a Gly → Ala replacement (8). Analysis of the thermodynamic parameters obtained from the melting curves suggests that the drop in melting temperature results from entropic destabilization. Factors that could be involved in the relative destabilizing effect of different residues substituted for Gly include charge and size. Although Asp, Glu, and Arg are among the most destabilizing residues, our studies on the effect of varying salt concentration and pH indicate that charge repulsion is not a dominant factor in the destabilization. The smallest residues, Ala and Ser, are the least disruptive, whereas the bulky residues Val and Arg are among the most disruptive. But size is not the only determinant, because Asp and Glu are more destabilizing than Arg. It is notable that Asp is one of the most destabilizing residues in the Yaa position of Gly-Xaa-Yaa triplets (22), as well as being the most destabilizing residue in the Gly position.

The strongest evidence for an effect of the identity of the residue substituted for Gly in OI collagen on clinical phenotype is best shown when more than one substitution has been observed at a given site in the α1(I) chain (see dotted lines in Fig. 4). For instance, a Gly220Ala mutation in the α1 chain leads to moderate OI type IV (49), whereas Gly220Asp is lethal (50). Both Ser (51) and Cys (52) exchanges for Gly at position 382 result in nonlethal forms in contrast to a substitution by Arg, which leads to a lethal phenotype (53). Also, near the C terminus, a Gly973Ser mutation has a less severe effect[¶] than the lethal Gly973Val exchange (54). These examples indicate that the identity of the residue replacing Gly is important in determining the phenotype at sites all along the α1(I) chain. The relative clinical severity observed for different residues is consistent with the order of disruption of different residues in the host–guest peptides, showing that the more destabilizing Asp, Val, and Arg residues lead to more severe phenotypes than the less disruptive Ser, Cys, and Ala.

The rank of destabilization of Gly substitutions in host–guest peptides also correlates with clinical phenotype of natural OI mutations in the α1(I) chain. The most disruptive substitution in the host–guest peptide set is Asp, which correlates with consistently lethal phenotypes observed for Gly → Asp substitutions in

[¶]Gomez-Lira, M., Mottes, M., Zolezzi, F., Cohen-Solal, L., Valli, M. & Piquatti, P. F., 5th International Conference on Osteogenesis Imperfecta, Sept. 27–30, 1993 Oxford, U.K., p. 120 (abstr.).

MEDICAL SCIENCES

α1(I) chains. The next most destabilizing replacement residues, Arg, Val, and Glu, result in lethal OI phenotypes in α1(I) chains, unless the site is near the N terminus, suggesting a small modulation by location. The less destabilizing residues Ser and Cys appear all along the chain in lethal and nonlethal cases. The modulation of lethal/nonlethal phenotype may now depend on more subtle factors related to sequences surrounding the mutation site (40, 41). Our results on nonreduced peptides with a Gly → Cys substitution raise the possibility that disulfide bonds, and their likely heterogeneity, might also modulate the clinical phenotype. The rare occurrence of Ala replacements in OI, even though it is the least disruptive residue in peptides, might reflect an ascertainment problem, because Ala mutations could result in a milder phenotype, with symptoms different from those expected for OI (38).

Collagens with different Gly substitutions extracted from OI fibroblasts show no change or only a small decrease in thermal stability compared with normal collagen, except for a few unusual cases (1, 2). The lack of destabilization of OI collagens compared with peptides may be caused in part by the stabilizing Gly-Pro-Hyp environment and the short length of the host–guest peptides, which magnify the destabilizing effects of Gly substitutions. In addition, direct stability measurements on OI collagen molecules may be complicated by heterogeneity of the molecular population and the use of enzyme digestion to measure thermal stability, problems not encountered in peptide studies.

Although a correlation is seen between the phenotype for mutations in the α1(I) chain and the destabilization of host–guest peptides with different Gly substitutions, such a correlation is not seen for mutations in the α2(I) chain. Lethal and nonlethal Gly → Asp, Gly → Val, and Gly → Arg mutations are present all along the α2(I) chain, even though the predictions from our data and α1(I) chain results suggest that these should be all lethal, except perhaps at the N terminus. The different "rules" for the α2(I) and α1(I) chains may arise from the nature of the α2(I) chain as a less imino acid-rich, more flexible chain. These rules could also arise from the different effects on molecular composition, because 50% of the collagens would have a mutant α2(I) chain and 50% would be normal; however, α1(I) mutations lead to 50% of the collagens having one mutant α1(I) chain and 25% having two mutant α1(I) chains (1).

To define a molecular basis for the severity of OI cases, it is necessary to determine the effect on the triple helix of each factor independently. This study on host–guest peptides has isolated the effect of one factor, the residue substituted for Gly. This study has also shown that mutations in the α1(I) chain have a dominant effect for the most disruptive residues, whereas other factors become determining for the less disruptive substitutions. These data provide a basis for evaluating the role of other factors, such as the surrounding sequence and location within the collagen chain, in determining the clinical phenotype of OI mutations.

We thank Mr. N. Bartone for assistance with peptide characterization and Dr. A. De Paepe (University Hospital Ghent, Belgium) for her permission to include unpublished mutations in Fig. 4. We appreciate the helpful discussions with Dr. Jean Baum and Dr. Peter Byers. This work was supported by grants from the National Institutes of Health (to B.B.), the Children's Brittle Bone Foundation (to B.B.), the National Science Foundation U.S.–Australia International Cooperative Research (to B.B.), and the Australia/U.S.A. Bilateral Science Program (to J.A.M.R.). The current work of K.B. is supported by the Fondation pour la Recherche Médicale.

1. Byers, P. H. (1993) in *Connective Tissue and Its Heritable Disorders*: *Molecular, Genetic and Medical Aspects*, eds. Royce, P. M. & Steinmann, B. (Wiley–Liss, New York), pp. 317–350.
2. Kuivaniemi, H., Tromp, G. & Prockop, D. J. (1991) *FASEB J.* **5,** 2052–2060.
3. Sillence, D. O., Senn, A. & Danks, D. M. (1979) *J. Med. Genet.* **16,** 101–116.
4. Kielty, C. M., Hopkins, I. & Grant, M. E. (1993) in *Connective Tissue and Its Heritable Disorders*: *Molecular, Genetic and Medical Aspects* eds. Royce, P. M. & Steinmann, B. (Wiley–Liss, New York), pp. 103–147.
5. McLaughlin, S. H. & Bulleid, N. J. (1997) *Matrix Biol.* **16,** 369–377.
6. Rich, A. & Crick, F. H. C. (1961) *J. Mol. Biol.* **3,** 483–506.
7. Fraser, R. D. B., MacRae, T. P. & Suzuki, E. (1979) *J. Mol. Biol.* **129,** 463–481.
8. Bella, J., Eaton, M., Brodsky, B. & Berman, H. M. (1994) *Science* **266,** 75–81.
9. Kuivaniemi, H., Tromp, G. & Prockop, D. J. (1997) *Hum. Mutat.* **9,** 300–315.
10. Dalgleish, R. (1998) *Nucleic Acids Res.* **26,** 253–255.
11. Bächinger, H.-P., Bruckner, P., Timpl, R., Prockop, D. J. & Engel, J. (1980) *Eur. J. Biochem.* **106,** 619–632.
12. Engel, J. & Prockop, D. J. (1991) *Annu. Rev. Biophys. Biophys. Chem.* **20,** 137–152.
13. Bateman, J. F., Mascara, T., Chan, D. & Cole, W. G. (1984) *Biochem. J.* **217,** 103–115.
14. Bonadio, J. & Byers, P. H. (1985) *Nature (London)* **316,** 363–366.
15. Vogel, B. E., Doelz, R., Kadler, K. E., Hojima, Y., Engel, J. & Prockop, D. J. (1988) *J. Biol. Chem.* **263,** 19249–19255.
16. Lightfoot, S. J., Holmes, D. F., Brass, D., Grant, M. E. & Byers, P. H. (1992) *J. Biol. Chem.* **267,** 25521–25528.
17. Forlino, A., Keene, D. R., Schmidt, K. & Marini, J. C. (1998) *Matrix Biol.* **17,** 575–584.
18. Kadler, K. E., Torre-Blanco, A., Adachi, E., Vogel, B. E., Hojima, Y. & Prockop, D. J. (1991) *Biochemistry* **30,** 5081–5088.
19. Shah, N. K., Ramshaw, J. A. M., Kirkpatrick, A., Shah, C. & Brodsky, B. (1996) *Biochemistry* **35,** 10262–10268.
20. Shah, N. K., Sharma, M., Kirkpatrick, A., Ramshaw, J. A. M. & Brodsky, B. (1997) *Biochemistry* **36,** 5878–5883.
21. Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A. M. & Brodsky, B. (1997) *J. Biol. Chem.* **272,** 28837–28840.
22. Chan, V. C., Ramshaw, J. A. M., Kirkpatrick, A., Beck, K. & Brodsky, B. (1997) *J. Biol. Chem.* **272,** 31441–31446.
23. Engel, J., Chen, H.-T., Prockop, D. J. & Klump, H. (1977) *Biopolymers* **16,** 601–622.
24. Long, C. G., Braswell, E., Zhu, D., Apigo, J., Baum, J. & Brodsky, B. (1993) *Biochemistry* **32,** 11688–11695.
25. Berg, R. A., Olsen, B. R. & Prockop, D. J. (1970) *J. Biol. Chem.* **245,** 5759–5763.
26. Venugopal, M. G., Ramshaw, J. A. M., Braswell, E., Zhu, D. & Brodsky, B. (1994) *Biochemistry* **33,** 7948–7956.
27. Ramshaw, J. A. M., Shah, N. K. & Brodsky, B. (1998) *J. Struct. Biol.* **122,** 86–91.
28. Mattice, W. L. & Mandelkern, L. (1970) *Biochemistry* **9,** 1049–1058.
29. Brown, F. R., III, Carver, J. P. & Blout, E. R. (1969) *J. Mol. Biol.* **39,** 307–313.
30. Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D. & Somero, G. N. (1982) *Science* **217,** 1214–1222.
31. Yancey, P. H. & Somero, G. N. (1979) *Biochem. J.* **183,** 317–323.
32. Lin, T.-Y. & Timasheff, S. N. (1994) *Biochemistry* **33,** 12695–12701.
33. Brown, C. R., Hong-Brown, L. Q. & Welch, W. J. (1997) *J. Clin. Invest.* **99,** 1432–1444.
34. Gekko, K. & Koga, S. (1983) *J. Biochem. (Tokyo)* **94,** 199–205.
35. Brzezinski, B. & Zundel, G. (1993) *FEBS Lett.* **333,** 331–333.
36. Bächinger, H.-P., Morris, N. P. & Davis, J. M. (1993) *Am. J. Med. Genet.* **45,** 152–162.
37. Wang, Q., Orrison, B. M. & Marini, J. C. (1993) *J. Biol. Chem.* **268,** 25162–25167.
38. Byers, P. H., Wallis, G. A. & Willing, M. C. (1991) *J. Med. Genet.* **28,** 433–442.
39. Marini, J. C., Lewis, M. B., Wang, Q., Chen, K. J. & Orrison, B. M. (1993) *J. Biol. Chem.* **268,** 2667–2673.
40. Wenstrup, R. J., Shrago-Howe, A. W., Lever, L. W., Phillips, C. L., Byers, P. H. & Cohn, D. H. (1991) *J. Biol. Chem.* **266,** 2590–2594.
41. Baum, J. & Brodsky, B. (1999) *Curr. Opin. Struct. Biol.* **9,** 122–128.
42. Bächinger, H.-P. & Davis, J. M. (1991) *Int. J. Biol. Macromol.* **13,** 152–156.
43. Yang, W., Battineni, M. L. & Brodsky, B. (1997) *Biochemistry* **36,** 6930–6935.
44. Liu, X., Kim, S., Dai, Q.-H., Brodsky, B. & Baum, J. (1998) *Biochemistry* **37,** 15528–15533.
45. Wang, A. & Bolen, D. W. (1997) *Biochemistry* **36,** 9101–9108.
46. Qu, Y., Bolen, C. L. & Bolen, D. W. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9268–9273.
47. Mashino, T. & Fridovich, I. (1987) *Arch. Biochem. Biophys.* **258,** 356–360.
48. Baskakov, I. & Bolen, D. W. (1998) *J. Biol. Chem.* **273,** 4831–4834.
49. Sarafova, A. P., Choi, H., Forlino, A., Gajko, A., Cabral, W. A., Tosi, L., Reing, C. M. & Marini, J. C. (1998) *Hum. Mutat.* **11,** 395–403.
50. Culbert, A. A., Lowe, M. P., Atkinson, M., Byers, P. H., Wallis, G. A. & Kadler, K. E. (1995) *Biochem. J.* **311,** 815–820.
51. Mackay, K., Byers, P. H. & Dalgleish, R. (1993) *Hum. Mol. Genet.* **2,** 1155–1160.
52. Byers, P. H. (1990) *Trends Genet.* **6,** 293–300.
53. Cohen-Solal, L., Zolezzi, F., Pignatti, P. F. & Mottes, M. (1996) *Mol. Cell. Probes* **10,** 219–225.
54. Lamande, S. R., Dahl, H. H., Cole, W. G. & Bateman, J. F. (1989) *J. Biol. Chem.* **264,** 15809–15812.