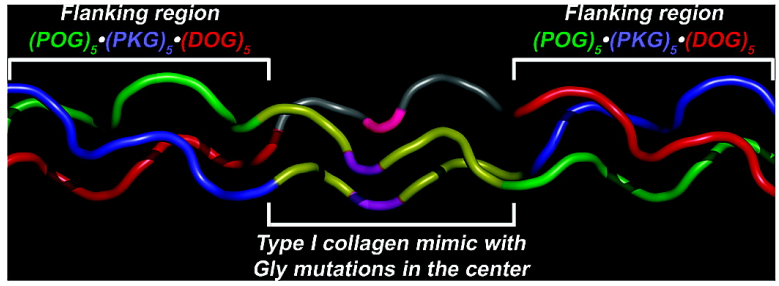
Article

Synthetic Collagen Heterotrimers: Structural Mimics of Wild-Type and Mutant Collagen Type I

Varun Gauba, and Jeffrey D. Hartgerink

J. Am. Chem. Soc., **2008**, 130 (23), 7509-7515 • DOI: 10.1021/ja801670v • Publication Date (Web): 16 May 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 05/16/2008

Synthetic Collagen Heterotrimers: Structural Mimics of Wild-Type and Mutant Collagen Type I

Varun Gauba and Jeffrey D. Hartgerink*

Department of Chemistry and Department of Bioengineering, Rice University, 6100 Main Street, Mail Stop 60, Houston, Texas 77005

Received March 5, 2008; E-mail: jdh@rice.edu

Abstract: Collagen type I is an AAB heterotrimer assembled from two α1 chains and one α2 chain. Missense mutations in either of these chains that substitute a glycine residue in the ubiquitous X-Y-Gly repeat with a bulky amino acid leads to osteogenesis imperfecta (OI) of varying severity. These mutations have been studied in the past using collagen-like peptide homotrimers as a model system. However, homotrimers, which by definition will contain glycine mutations in all the three chains, do not accurately mimic the mutations in their native form and result in an exaggerated effect on stability and folding. In this article, we report the design of a novel model system based upon collagen-like heterotrimers that can mimic the glycine mutations present in either the α1 or α2 chains of type I collagen. This design utilizes an electrostatic recognition motif in three chains that can force the interaction of any three peptides, including AAA (all same), AAB (two same and one different), or ABC (all different) triple helices. Therefore, the component peptides can be designed in such a way that glycine mutations are present in zero, one, two, or all three chains of the triple helix. With this design, we for the first time report collagen mutants containing one or two glycine substitutions with structures relevant to native forms of OI. Furthermore, we demonstrate the difference in thermal stability and refolding half-life times between triple helices that vary only in the frequency of glycine mutations at a particular position.

Introduction

Collagen is a major structural component of all connective tissues such as skin, ligament, cartilage, bone, and tendon. Collagen is characterized by a ubiquitous X-Y-Gly repeating motif, where X and Y are generally proline (P) and hydroxyproline (O), respectively, while glycine is present every third amino acid.[1–8] Three left-handed polyproline II helices coil around each other to form a right-handed triple helical structure that is stabilized by an extensive network of $CO_{(X)}-NH_{(Gly)}$ hydrogen bonds and a tightly packed triple helix. The absence of a side chain in the Gly residue make it sterically possible to fit in the interior of the crowded triple helix. Substituting Gly by any other amino acid results in the distortion of the triple helix and leads to varying degrees of destabilization depending on the substitution.[7–9] Gly mutations in natural collagens lead to a variety of connective tissue disorders. Mutations in homotrimeric AAA collagens (type II, III, VII, and X) lead to chondrodysplacia, Ehlers−Danlos syndrome, dystrophic epidermolysis bullosa, and chondrocyte hyperthrophy, respectively.[10–13] Mutations in heterotrimeric AAB collagens (type I, IV, and VIII) result in osteogenesis imperfecta, Alport syndrome, and corneal endothelial dystrophy, respectively, whereas mutations in heterotrimeric ABC collagen (type VI) result in Bethlem myopathy.[10–13] Osteogenesis imperfecta (OI), the most studied collagen disease, is primarily caused by missense mutations in either the α1 or α2 chains of type I collagen, which lead to the substitution of Gly in the ubiquitous X-Y-Gly repeat by bulky amino acids such as Arg, Asp, Glu, Cys, Ser, Ala, or Val.[10,11] OI phenotypes vary from mild to lethal forms depending on a variety of factors including the following: (1) the chain type mutated; for example, it is observed that the amino acid substitutions that are lethal when present in the α1 chain may not be lethal when present in the α2 chain;[10,11] (2) the proximity to the C-terminus; it has been observed that mutations present near the C-terminus are more lethal than the mutations present near the N-terminus, which correlates with the C- to N-terminus folding of the triple helix;[14,15] (3) the neighboring sequence of the mutated site; (4) the amino acid residue which substitutes for glycine.

(1) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483–506.
(2) Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915–916.
(3) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593–595.
(4) Ramachandran, G. N.; Kartha, G. *Nature* **1954**, *174*, 269–270.
(5) Ramachandran, G. N. *Int. Rev. Connect. Tissue Res.* **1963**, *1*, 127–182.
(6) Cowan, P. M.; Mcgavin, S.; North, A. C. T. *Nature* **1955**, *176*, 1062–1064.
(7) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75–81.
(8) Bella, J.; Brodsky, B.; Berman, H. M. *Structure* **1995**, *3*, 893–906.
(9) Bella, J.; Brodsky, B.; Berman, H. M. *Connect. Tissue Res.* **1996**, *35*, 455–460.
(10) Byers, P. H. *Philos. Trans. R. Soc. London, Ser. B* **2001**, *356*, 151–157.
(11) Kelly, J. W.; Byers, P. H.; Helenius, A.; Swoboda, B. E. P. *Philos. Trans. R. Soc. London, Ser. B* **2001**, *356*, 157–158.
(12) Myllyharju, J.; Kivirikko, K. I. *Trends Genet.* **2004**, *20*, 33–43.
(13) Ricard-Blum, S.; Ruggiero, F.; van der Rest, M. *Top. Curr. Chem.* **2005**, *247*, 35–84.
(14) Bachinger, H. P.; Morris, N. P.; Davis, J. M. *Am. J. Med. Genet.* **1993**, *45*, 152–162.
(15) Lamande, S. R.; Dahl, H. H. M.; Cole, W. G.; Bateman, J. F. *J. Biol. Chem.* **1989**, *264*, 15809–15812.

Collagen-like peptide homotrimers have been widely used as models to study the effects of glycine mutations.[16–24] Type I collagen, however, is an AAB heterotrimer, and the mutations which lead to OI are present in either the α1 or α2 chains, not both. Furthermore, the sequences of α1 and α2 chains are not identical. Thus, collagen-like peptide homotrimers with glycine mutations in all the three chains do not correctly mimic the mutations found in type I collagen and result in what is expected to be an exaggerated effect on stability and folding. We recently reported the synthesis and characterization of various ABC type heterotrimers, including a highly stable heterotrimer composed of three unique 30 amino acid peptides: neutral $(POG)_{10}$, positive $(PKG)_{10}$, and negative $(DOG)_{10}$. This heterotrimer was found to have a thermal stability comparable to that of the $(POG)_{10}$ homotrimer which is known to be the most stable peptide motif using natural amino acids.[25,26]

With the use of this electrostatically controlled recognition motif we show here the development of a system which can force the interaction of any three guest peptide sequences. The guest sequences can be all the same (AAA), two the same and one different (AAB), or all different from each other (ABC). These guest peptides can be used to understand the stabilities of heterotrimeric collagens which were previously synthetically inaccessible. In this article, we take advantage of this design flexibility to study mutations which give rise to OI. As long as the POG, PKG, and DOG motifs are present in the three chains to guide the formation of an ABC heterotrimer, a natural sequence can be incorporated into the peptide design so that the effect of glycine mutations can be assessed in zero, one, two, or all three chains. We demonstrate the synthesis of four ABC heterotrimers which incorporate the natural sequence from residues 242−250 from either the α1 or α2 chains of human type I collagen and the Gly to Ser mutation at position 247 in either one α2 chain,[27] two α1 chains,[28] or all three chains. For the first time, we demonstrate the difference in thermal stability and refolding times between the triple helices with zero, one, two, or three glycine mutations. We believe that using the peptide design proposed in this article, any natural mutated or normal sequence with either AAB or ABC organization can be incorporated in a synthetic system. We expect that glycine mutants will more accurately reflect natural conditions when prepared as α1·α1·α2 heterotrimers using the peptide design described in this article, and not as α1·α1·α1 or α2·α2·α2 homotrimers. This ability to prepare heterotrimeric collagen helices in a highly flexible fashion is also expected to be of use when designing synthetic biomaterials which attempt to emulate the structure of collagen and, more generally, the extracellular matrix.[29–34]

Table 1. Sequence of Peptides Studied[a]

no.	sequence
	α1 Chain Mimics (**A**)
1	$(POG)_5$ (PQGPG**G**PPG) $(POG)_5$
2	$(POG)_5$ (PQGPG**S**PPG) $(POG)_5$
3	$(PKG)_5$ (PQGPG**G**PPG) $(PKG)_5$
4	$(PKG)_5$ (PQGPG**S**PPG) $(PKG)_5$
	α2 Chain Mimics (**B**)
5	$(POG)_5$ (PVGAA**G**ATG) $(POG)_5$
6	$(POG)_5$ (PVGAA**S**ATG) $(POG)_5$
7	$(DOG)_5$ (PVGAA**G**ATG) $(DOG)_5$
8	$(DOG)_5$ (PVGAA**S**ATG) $(DOG)_5$

[a] N- and C-termini of all the peptides are acetylated and amidated, respectively.

Eight polypeptides, shown in Table 1, were synthesized for use in this study. Their synthesis and characterization are described in the Supporting Information. There are four peptides that use the sequence 242−250 of the α1 chain of human type I collagen and are numbered from **1** to **4**. Similarly, there are four peptides that use the sequence 242−250 of the α2 chain of human type I collagen and are numbered from **5** to **8**. All the peptides used in this study are N-terminally acetylated and C-terminally amidated to eliminate the charge repulsion between the termini.[35]

The heterotrimeric helices prepared in this study are abbreviated in such a way that the first, second, and third chains in the triple helix have neutral $(POG)_5$, positive $(PKG)_5$, and negative $(DOG)_5$ flanking regions, respectively. For example, a nonmutated heterotrimer abbreviated as **A·A·B** consists of polypeptides **1**, **3**, and **7**, whereas a mutated heterotrimer abbreviated as **A·A·B′** consists of polypeptides **1**, **3**, and **8**. The homotrimers are abbreviated as **3A**, **3B**, **3A′**, and **3B′** for the triple helices formed from three copies of **1**, **5**, **2**, and **6** polypeptides, respectively. The abbreviations and composition of various triple helices used in this study are shown in Table 2.

Materials and Methods

Peptide Synthesis and Purification. All the peptides were synthesized on an Advanced ChemTech 396 multipeptide automated synthesizer using Fmoc solid-phase chemistry based on a 0.15 mmol scale. The methods for synthesis, purification by reversed-phase high-performance liquid chromatography (RP-HPLC), and analysis by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) are the same as described previ-

(16) Yang, W.; Battineni, M. L.; Brodsky, B. *Biochemistry* **1997**, *36*, 6930–6935.
(17) Persikov, A. V.; Pillitteri, R. J.; Amin, P.; Schwarze, U.; Byers, P. H.; Brodsky, B. *Hum. Mutat.* **2004**, *24*, 330–337.
(18) Liu, X. Y.; Kim, S.; Dai, Q. H.; Brodsky, B.; Baum, J. *Biochemistry* **1998**, *37*, 15528–15533.
(19) Hyde, T. J.; Bryan, M. A.; Brodsky, B.; Baum, J. *J. Biol. Chem.* **2006**, *281*, 36937–36943.
(20) Buevich, A. V.; Silva, T.; Brodsky, B.; Baum, J. *J. Biol. Chem.* **2004**, *279*, 46890–46895.
(21) Bhate, M.; Wang, X.; Baum, J.; Brodsky, B. *Biochemistry* **2002**, *41*, 6539–6547.
(22) Beck, K.; Chan, V. C.; Shenoy, N.; Kirkpatrick, A.; Ramshaw, J. A. M.; Brodsky, B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 4273–4278.
(23) Baum, J.; Brodsky, B. *Folding Des.* **1997**, *2*, R53–R60.
(24) Battineni, M. L.; Yang, W.; Liu, X. Y.; Baum, J.; Brodsky, B. *Matrix Biol.* **1996**, *15*, 174–175.
(25) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 15034–15041.
(26) Gauba, V.; Hartgerink, J. D. *J. Am. Chem. Soc.* **2007**, *129*, 2683–2690.
(27) Zhuang, J. P.; Tromp, G.; Kuivaniemi, H.; Castells, S.; Bugge, M.; Prockop, D. J. *Hum. Mutat.* **1996**, *7*, 89–99.
(28) Mackay, K.; Byers, P. H.; Dalgleish, R. *Hum. Mol. Genet.* **1993**, *2*, 1155–1160.
(29) Kar, K.; Amin, P.; Bryan, M. A.; Persikov, A. V.; Mohs, A.; Wang, Y.-H.; Brodsky, B. *J. Biol. Chem.* **2006**, *281*, 33283–33290.
(30) Kotch, F. W.; Raines, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3028–3033.
(31) Paramonov, S. E.; Gauba, V.; Hartgerink, J. D. *Macromolecules* **2005**, *38*, 7555–7561.
(32) Rele, S.; Song, Y. H.; Apkarian, R. P.; Qu, Z.; Conticello, V. P.; Chaikof, E. L. *J. Am. Chem. Soc.* **2007**, *129*, 14780–14787.
(33) Wang, A. Y.; Mo, X.; Chen, C. S.; Yu, S. M. *J. Am. Chem. Soc.* **2005**, *127*, 4130–4131.
(34) Cejas, M. A.; Kinney, W. A.; Chen, C.; Leo, G. C.; Tounge, B. A.; Vinter, J. G.; Joshi, P. P.; Maryanoff, B. E. *J. Am. Chem. Soc.* **2007**, *129*, 2202–2203.
(35) Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948–7956.

Table 2. Composition of Triple Helices Studied

abbreviation	peptides
Homotrimers	
3A	**1**
3A′	**2**
3B	**5**
3B′	**6**
Heterotrimers	
A·A·B	**1**, **3**, and **7**
A·A·B′	**1**, **3**, and **8**
A′·A′·B	**2**, **4**, and **7**
A′·A′·B′	**2**, **4** and **8**

ously.[26] HPLC and MALDI-TOF spectra for each peptide can be found in the Supporting Information.

Circular Dichroism Spectroscopy. CD measurements were performed with a Jasco J-810 spectropolarimeter, equipped with a Peltier temperature control system, using quartz cells with a path length of 0.1 cm. Thermal unfolding curves were obtained by monitoring the decrease in ellipticity in a desired temperature range at a wavelength where the CD spectra shows a positive maximum. The value is between 223 and 225 nm. The temperature ranged from 0 or 5 to 95 °C depending on the peptide mixture studied and at a heating rate of 10 °C/h. Fraction folded was calculated for the melting curves according to the following equation:

$$F_T = \frac{[\theta]_T - [\theta]_{\text{unfolded}}}{[\theta]_{\text{folded}} - [\theta]_{\text{unfolded}}} \quad (1)$$

where F_T is fraction folded at temperature T, $[\theta]_T$ is the mean residue ellipticity (MRE) at temperature T, $[\theta]_{\text{folded}}$ and $[\theta]_{\text{unfolded}}$ are the MRE of the maximally folded and unfolded forms measured at 5 and 95 °C, respectively. MRE is calculated as follows $[\theta] = (\theta m)/(10cln_r)$, where θ is ellipticity in mdeg, m is molecular weight in g/mol, c is concentration in mg/mL, l is path length of the cuvette in cm, and n_r is the number of amino acid residues in the peptide.

Peptide solutions were mixed in desired ratios in such a way that the final total peptide concentration was 0.2 mM, and a neutral pH was maintained by using 10 mM phosphate buffer solution. Unfolding studies were performed with and without preheating. For preheating studies, peptides were mixed in desired ratios, heated to 85 °C, and incubated for 15 min. The peptide solution was then slowly cooled to 25 °C at a rate of 1 °C/min and then incubated overnight at room temperature before performing the unfolding studies. For nonpreheating studies, peptides were mixed in the desired ratios and the unfolding studies were performed immediately. The minimum of the derivative of the fraction folded plot indicates the steepest slope of the unfolding process and is used in this article to indicate the melting temperature (T_m) under the conditions described above. This was calculated using the Jasco Spectra Manager software. All experiments were repeated once, and T_m values were found to be reproducible within ±1 °C or less.

Refolding Studies. Refolding studies were performed using the Jasco J-810 spectropolarimeter, equipped with a Peltier temperature control system. Peptide mixtures were prepared as described above. The peptide mixture was heated to 85 °C in a glass vial and incubated for 15 min. It was then immediately transferred to a 0.1 cm cuvette which had been precooled to 5 °C. Refolding data in the form of ellipticity was collected immediately afterward with a data pitch of 5 s. The data was collected for 50 min. The dead time required to transfer the sample and start the data acquisition was approximately 15 s. Thus, for the data analysis, the zero time is actually 15 s after the peptide mixture is taken out of the heated glass vial. Refolding half-time, the time at which half of the polypeptides have refolded into triple helices, was calculated by monitoring when the fraction folded equated half of the maximum equilibrated value as observed after overnight incubation on a repeat sample. The fraction folded was calculated using eq 1, where $[\theta]_T$ is the MRE at time t and $[\theta]_{\text{folded}}$ is the MRE of the maximally folded form measured at 5 °C. $[\theta]_{\text{unfolded}}$ was measured by curve fitting of the refolding data by linear extrapolation to time zero.[16,36] Repeated experiments indicate that the deviation in half-life is approximately ±5 s.

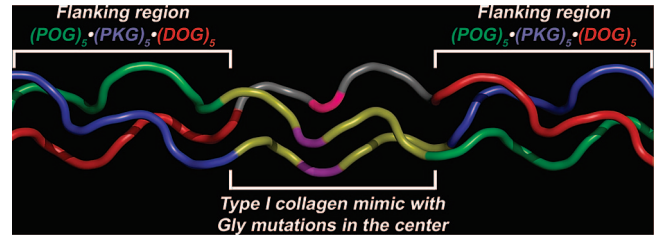


Figure 1. Schematic representation of the molecular design. The flanking regions are assembled from $(POG)_5$ (green), $(PKG)_5$ (blue), and $(DOG)_5$ (red) motifs, which are predisposed to form an ABC heterotrimer. The central region mimics the sequence 242−250 of type I collagen in the AAB register with the glycine to serine mutations in either α1 or α2 chains at position 247.

Results and Discussion

Molecular and Experimental Design. We designed peptides that utilize $(POG)_5$, $(PKG)_5$, and $(DOG)_5$ motifs to organize OI-causing glycine mutations in synthetic collagen-like heterotrimers similar to those observed in native type I collagen, as depicted schematically in Figure 1. Since the flanking regions in the designed peptides are predisposed to selectively form ABC collagen heterotrimers directed through electrostatic interactions, we can drive the assembly of designed peptides containing native sequences from type I collagen into triple helices utilizing these interactions. In this study, we assemble human collagen type I sequence 242−250 in the center of the triple helix in both its normal and mutated (Gly to Ser, at position 247) forms in an AAB type register. Two peptides with the $(POG)_5$ and $(PKG)_5$ flanking regions mimic the α1 chain (sequence, PQGPGGPPG), and one peptide with the $(DOG)_5$ flanking region mimics the α2 chain (sequence, PVGAAGATG) of type I collagen, giving rise to an α1·α1·α2 register in the central region of interest. The heterotrimer with no mutations is referred to as **A·A·B**, and the heterotrimers with mutations in either α1 chains or in the α2 chain are referred to as **A′·A′·B** or **A·A·B′**, respectively. For comparison, we also analyzed a heterotrimer with mutations in all three chains that is referred to as **A′·A′·B′**. Homotrimers of the α1 and α2 chains were also analyzed, both in normal and mutated forms, and are referred to as **3A**, **3B**, **3A′** ,and **3B′**.

Homotrimers. 3A, **3B**, **3A′**, and **3B′** homotrimers were analyzed at a concentration of 0.2 mM in 10 mM PO_4 buffer at pH 7. All the homotrimers are composed of $(POG)_5$ flanking regions, with polypeptide **A** mimicking the normal and **A′** mimicking the mutated form of the α1 chain of type I collagen. In the same fashion, polypeptide **B** mimics the normal and **B′** mimics the mutated form of the α2 chain of type I collagen. **3A** and **3B** show melting temperature of 68 and 54 °C, respectively, as shown in Figure 2. As expected, the Gly to Ser mutations show a decreased stability with melting temperatures of 45 and 33 °C for **3A′** and **3B′**, respectively. This corresponds to a destabilization of 23 °C for mutations in the **A** homotrimer and 21 °C for the **B** homotrimer. This decreased stability is not surprising as the substitution of Gly with the bulkier amino acid Ser leads to a disruption of the tight packing of the triple helix.

(36) Ackerman, M. S.; Bhate, M.; Shenoy, N.; Beck, K.; Ramshaw, J. A. M.; Brodsky, B. *J. Biol. Chem.* **1999**, *274*, 7668–7673.

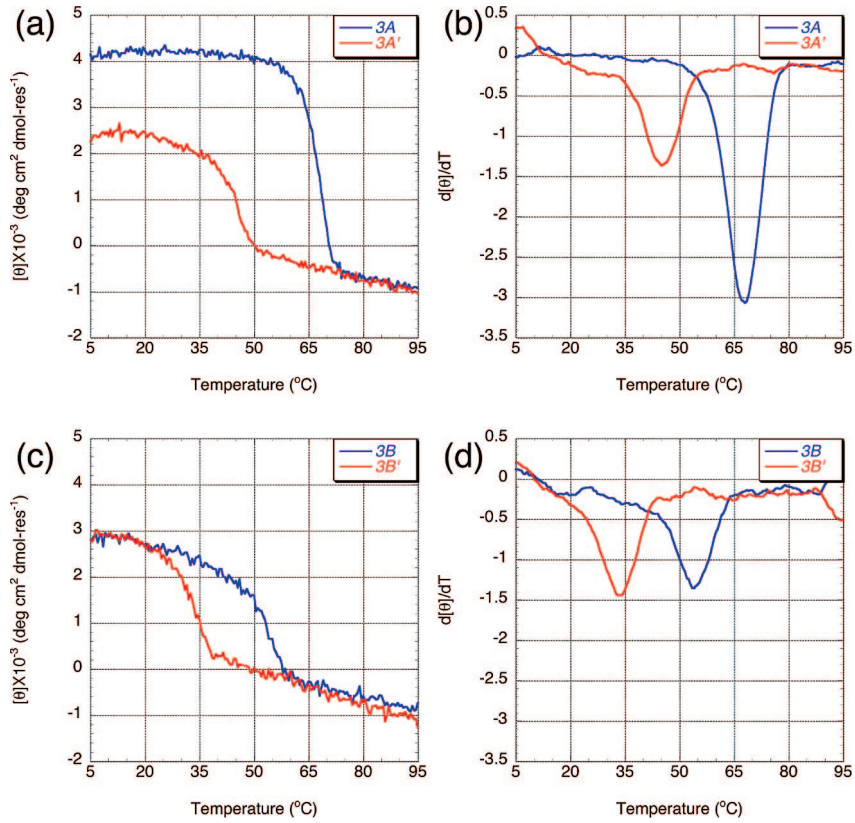


Figure 2. Circular dichroism for homotrimeric helices mimicking sequence 242−250 of the α1 and α2 chains of type I collagen, with and without glycine mutations at position 247. (a) Thermal analysis shows cooperative unfolding for both **3A** and**3A′** homotrimers mimicking the α1 chain. (b) The first derivative of unfolding vs temperature shows that the thermal stability decreases by 23 °C when the glycine to serine mutation is present in all the three chains. (c) Cooperative unfolding for both **3B** and **3B′** homotrimers mimicking the α2 chain. (d) Thermal stability decreases by 21 °C when glycine to serine mutations are present in all the three chains.

A similar drop in thermal stability of approximately 22 °C was reported for a lethal Gly to Ser mutation in homotrimers mimicking sequence from residues 904−921 from the α1 chain of human type I collagen and with the mutation present at position 913.[16] However, a drop of only 11 °C was reported for a nonlethal mutation present at position 901 in a homotrimer mimicking sequence from residues 904−921 from the α1 chain demonstrating the importance of the neighboring sequences of the mutation site. Clearly, neighboring sequences modify the extent of destabilization differently. Additionally, a lower melting temperature for **3B** and **3B′** when compared to **3A** and **3A′**, respectively, is expected on the basis on reported literature[37] where the authors report that 3α2 homotrimer is thermally less stable than the 3α1 homotrimer. **3A** and **3A′** heterotrimers were observed to be more stable by 14 and 12 °C when compared to **3B** and **3B′** heterotrimers, respectively. The thermal stability of the homotrimers is summarized in Table 3.

Type I collagen is an AAB heterotrimer, and the mutations that lead to OI are present only in either the α1 or α2 chains. Since homotrimers are composed of three identical chains, the mutation will either be absent in all the chains or present in all chains. Thus, by using synthetic homotrimers, the mutations cannot be accurately studied for type I collagen or any other heterotrimeric collagens and are limited to analyzing only naturally occurring homotrimers. The presence of mutations in all three chains, and not in just one or two chains in synthetic collagen-like peptides, exaggerates their destabilizing contribution. In contrast, synthetic heterotrimers are an ideal system for analyzing type I collagen or any other heterotrimer mutations in their native forms since the individual chains in a heterotrimer can be tailored to have the specific mutation in one, two, or all three chains. The formation of various ABC heterotrimers, which mimic the mutations in either the α1 or α2 chains of type I collagen, was assessed by combining the corresponding polypeptides.

Table 3. Thermal Stability and Refolding Half-Life of Triple Helices Studied

abbreviation	melting temp (°C)	$t_{1/2}$ (s)
	Homotrimers	
3A	68	580
3A′	45	910
3B	54	345
3B′	33	655
	Heterotrimers	
A·A·B	60	190
A·A·B′	44.5	300
A′·A′·B	42.5	325
A′·A′·B′	36.5	470

ABC Heterotrimers. Various peptides with $(DOG)_5$, $(PKG)_5$, or $(POG)_5$ flanking regions and the sequence 242−250 (Gly to Ser mutation at position 247) from type I collagen in the center were mixed together in a 1:1:1 ratio to form ABC heterotrimers with zero, one, two, or three mutations. The heterotrimer formation was confirmed by CD analysis analogous to our previous description[25,26] and is also explained in detail in the Supporting Information. The heterotrimer **A·A·B**, with no

(37) Leikina, E.; Mertts, M. V.; Kuznetsova, N.; Leikin, S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1314–1318.

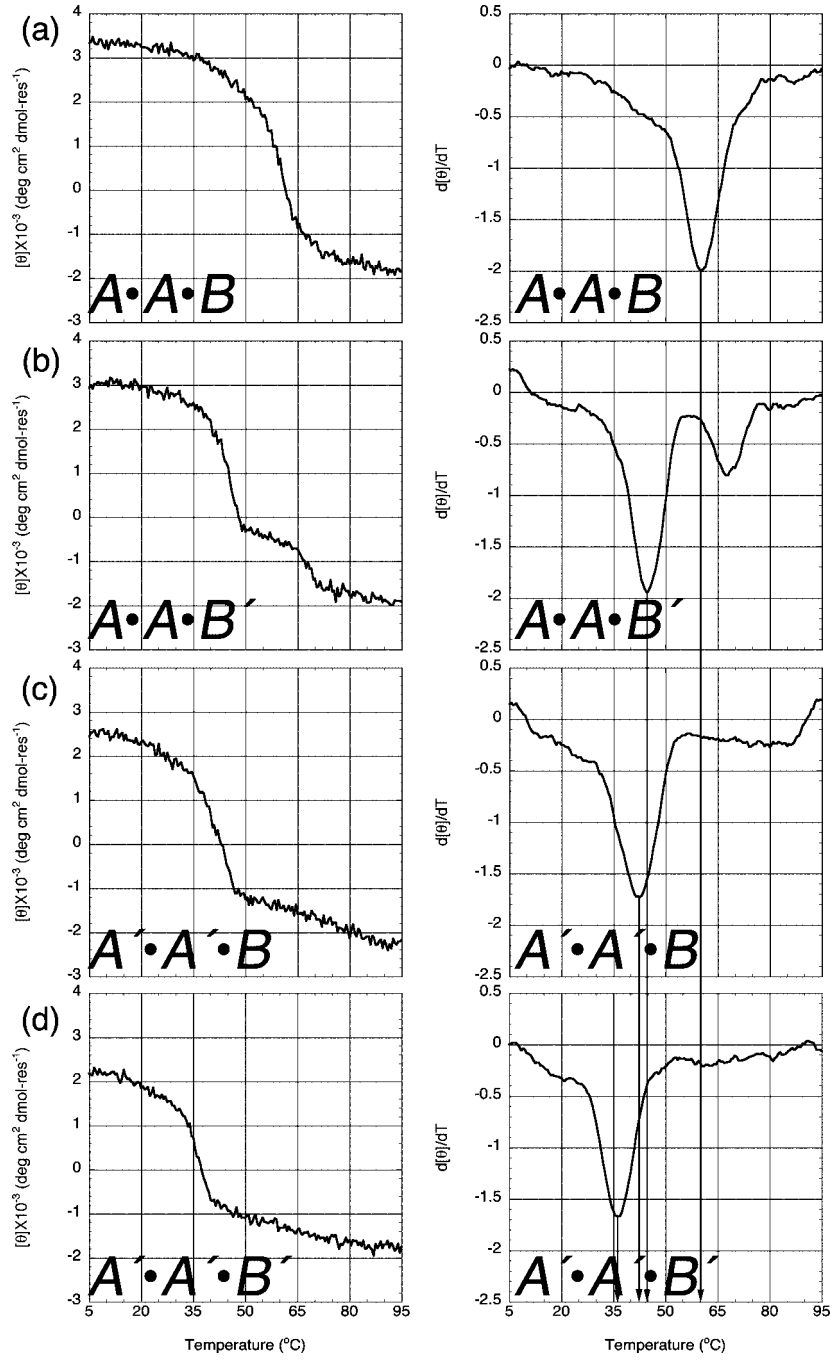


Figure 3. Circular dichroism for heterotrimeric helices mimicking sequence 242−250 of α1 and α2 chains of type I collagen and glycine to serine mutation at position 247 in either of the two chains. (a) Thermal unfolding for **A·A·B** heterotrimer shows a single transition with a melting temperature of 60 °C as observed in the first derivative of unfolding vs temperature. (b) Thermal unfolding for **A·A·B′** heterotrimer shows a melting temperature of 44.5 °C as observed in the first derivative of unfolding vs temperature. (c) Thermal unfolding for **A′·A′·B** heterotrimer shows a melting temperature of 42.5 °C as observed in the first derivative of unfolding vs temperature. (d) Thermal unfolding for **A′·A′·B′** heterotrimer shows a melting temperature of 36.5 °C as observed in the first derivative of unfolding vs temperature. The thermal stability decreases as the number of mutations increase.

mutation and an α1·α1·α2 register in the central region showed the highest stability with a melting temperature of 60 °C, as shown in Figure 3. The thermal stability dropped by 15.5 °C when the single mutation was introduced in the α2 chain with the heterotrimer **A·A·B′** showing a melting temperature of 44.5 °C. The unfolding curve for **A·A·B′** heterotrimer also showed a second transition at 67.5 °C which can be attributed to **3A** homotrimer. However, on the basis of the intensity of the transition, the **3A** homotrimer is a minority component and the bulk of the peptides are present as **A·A·B′** heterotrimers. Importantly, the formation of a small fraction of homotrimer in this peptide mixture does not affect the interpretation of the results in any way as the homotrimer transition is well-separated from the transition corresponding to the **A·A·B′** heterotrimer. The results from this particular peptide mixture are explained in a greater detail in the Supporting Information. The heterotrimer **A′·A′·B** with two mutations, both in the α1 chain, showed only a small additional decrease in stability, with a melting temperature of 42.5 °C. Thermal stability of the **A′·A′·B′** heterotrimer, with mutations in all three chains, decreases further to 36.5 °C. A continuous decrease in MRE values was also observed for all the heterotrimers when the number of mutations

increased from zero to three. This decrease in MRE value suggests that an increasing percentage of polypeptides fail to assemble into triple helices as the number of mutations increase. The thermal stability of all the heterotrimers discussed in this study is summarized in Table 3.

Thermal stability decreases nonlinearly as the number of mutations in the triple helix increases. Incorporation of the Gly to Ser mutation in the α2 chain has a severe effect on the thermal stability and leads to a 15.5 °C drop in the melting temperature when compared to the nonmutated **A·A·B** heterotrimer. The incorporation of two mutations in the α1 chain leads to a small additional decrease in thermal stability of about 2 °C when compared to the mutation in the α2 chain. Although **A·A·B′** heterotrimer contains only one and **A′·A′·B** heterotrimer contains two mutations, respectively, the difference in thermal stability between the two heterotrimers does not correctly represent the destabilization caused by the incorporation of second mutation. The first mutation is present in a different sequence environment (α2 mimic peptide) when compared to the second mutation (α1 mimic peptide), and thus the difference in thermal stability is a combined effect of different environment and an additional mutation, thereby making a direct comparison impossible. To correctly quantify the decrease in thermal stability by individual Gly to Ser mutations, we synthesized **A′·A·B** and **A·A′·B** heterotrimers where the first mutation is incorporated in either of the α1 mimetic peptides. Both **A′·A·B** and **A·A′·B** heterotrimers showed a thermal stability of 47 °C, as shown in Supporting Information Figure S6, a drop of 13 °C when compared to that of the nonmutated **A·A·B** heterotrimer. The incorporation of second mutation in the α1 chain leads to a further decrease of 4.5 °C. When the mutations are present in all the three chains, a further destabilization of 6 °C is observed. This result indicates that it is comparatively less damaging to incorporate the second and third mutations in a triple helix when compared to the first mutation. We believe that the incorporation of the first mutation in triple helix is most disruptive and leads to a correspondingly large drop in stability, whereas the addition of a second and third mutations to a already disrupted packing structure leads to only a small further decrease in stability.

Although the **A′·A′·B** heterotrimer showed only a slightly lower thermal stability (by approximately 2 °C) than the **A·A·B′** heterotrimer, the melting temperature observed for both the mutated heterotrimers is in agreement with the known clinical severity of the OI types caused by mutations in the α1[28] or α2[27] chains. Mutations in the α1 chains lead to lethal OI type II, whereas the mutation in the α2 chain leads to the relatively mild OI type IV.[38] As there are two α1 chains and one α2 chain in type I collagen, the lethal phenotype for the mutation in the α1 chain is expected since there are two mutated chains instead of a single mutation present in the α2 chain. However, since the thermal stability difference between the two mutated heterotrimers in only 2 °C, the known severity of the disease may not be because of a drop in thermal stability but may be a result of abnormal collagen packing during fiber formation. Additionally, the comparatively longer flanking regions in our system (five triplets on both the boundaries when compared to three guest triplets in the center) may be too stabilizing, thereby shielding the destabilizing effect caused by an increased number of mutations. To accurately identify the factors responsible for the different severity of the α1 or α2 chain mutations at position 247, we are currently exploring the simultaneous effects of shortened boundary regions and longer guest regions. When the mutation is present in all three chains, the thermal stability and the MRE values are the lowest showing that the triple helix formed is very weak. Apart from the nonlinear decrease in thermal stability, for the first time, we show that the effect of just one glycine to serine mutation in a particular chain can be measured. Thus, we are able to differentiate between various triple helices that differ from each other in only one amino acid residue. Additionally, we have designed a series of collagen-like ABC heterotrimers that can mimic the mutations in two different chains in type I collagen, either present individually in the α1 and α2 chains (**A′·A′·B** and **A·A·B′** heterotrimers) or present in both the α1 and α2 chains (**A′·A′·B′** heterotrimer). To further understand the effects of mutations, refolding studies were performed on the heterotrimers and refolding half-times ($t_{1/2}$) were compared for all triple helices studied. Additionally, the refolding studies were performed on both normal and mutated homotrimers, and the results were compared with both the heterotrimers reported in this article and the homotrimers reported elsewhere in the literature.[16,19,21,36] Refolding half-time, the time at which half of the polypeptides have recovered to form a triple helix, was calculated from the plots by monitoring when the fraction folded equaled half of the maximum equilibrium value as observed after overnight incubation on a repeated samples.

Refolding studies showed that all the heterotrimers recovered to approximately 85% of their respective fraction folded values in fewer than 50 min. The respective MRE values observed after 50 min and the refolding half-times followed the same trend that was observed in the unfolding studies, with **A·A·B** showing the highest recovered MRE and best refolding rate followed by **A·A·B′**, **A′·A′·B**, and **A′·A′·B′**. A refolding half-time of 190, 300, and 325 s was observed for **A·A·B**, **A·A·B′**, and **A′·A′·B** heterotrimers, respectively, as shown in Figure 4 and Table 3. The **A′·A′·B′** heterotrimer, with mutation in all three chains, refolded much more slowly with a refolding half-time of 470 s. All the homotrimers folded slower than the heterotrimers, and refolding half-times of 580 and 910 s were observed for **3A** and **3A′**, respectively. **3B** and **3B′** showed refolding half-times of 345 and 655 s, respectively. Although **3B** and **3B′** show a diminished thermal stability when compared to **3A** and **3A′**, respectively, they show a much faster recovery to triple helices. Even though the refolding half-times increase with the numbers of mutations for both heterotrimers and homotrimers, they are still comparatively faster than the reported values for mutated homotrimers.[19,21] Refolding half-times of 1620 s were reported for nonmutated homotrimers that increased to 2280 s upon Gly to Ala mutation at position 901 in a peptide mimicking sequence from residues 892−909 from the α1 chain of human type I collagen.[21] These peptides were 30 amino acids long and had a $(POG)_4$ flanking region only on the C-terminal. Homotrimers reported in our study are 39 amino acids long and have $(POG)_5$ flanking regions on both the terminals. Therefore, a comparatively faster folding is expected for homotrimers when compared to the reported literature. It has been reported that $(POG)_{10}$ homotrimer has a refolding half-time of 360 s.[36] Refolding half-time for both **3A** and **3B** fall in a similar range. Heterotrimers, however, folded much faster than both the homotrimers reported in this study and in the literature. We believe that the $(DOG)_5$, $(PKG)_5$, and $(POG)_5$ flanking regions on both the sides in our systems help lead to a comparatively faster refolding. We believe that the electrostatic driving force

(38) DeVos, A.; Sermon, K.; VandeVelde, H.; Joris, H.; Vandervorst, M.; Lissens, W.; DePaepe, A.; Liebaers, I.; VanSteirteghem, A. *Hum. Genet.* **2000**, *106*, 605–613.

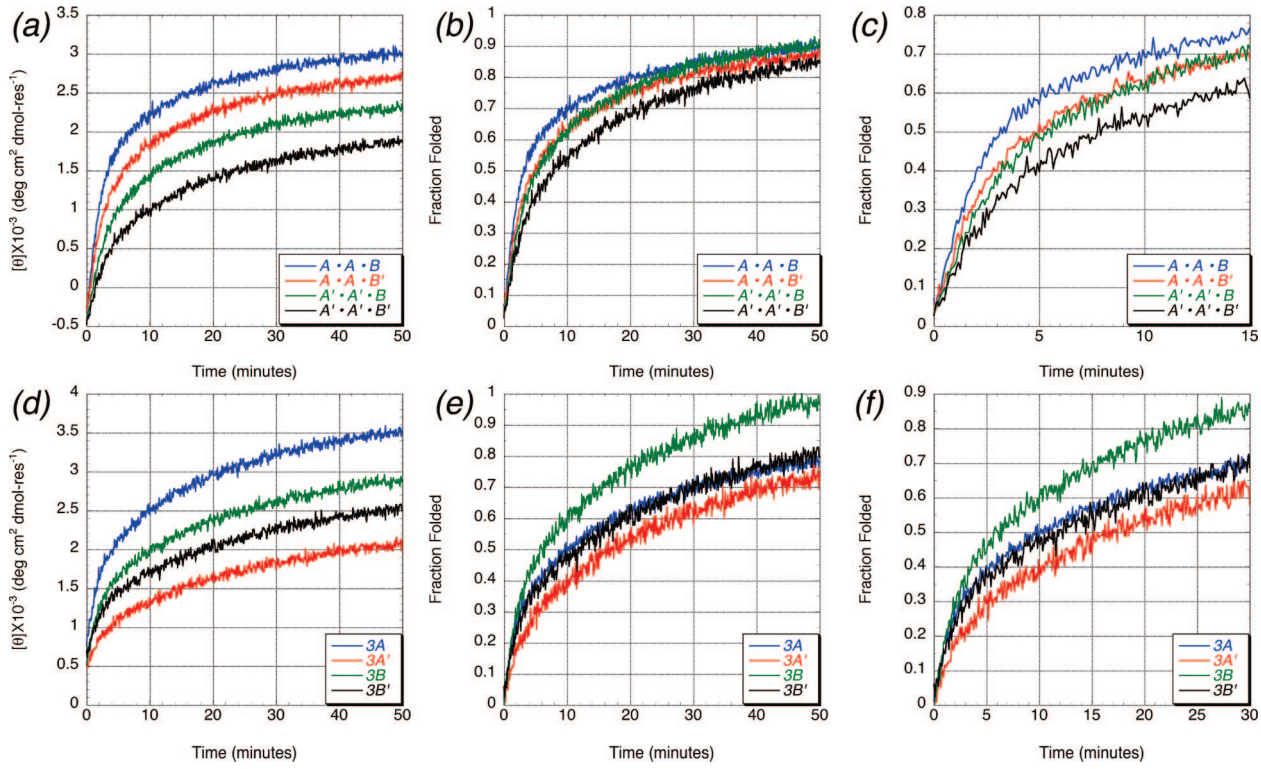


Figure 4. Refolding analysis for various heterotrimers and homotrimers showed that the refolding half-times increase as the number of mutations increase. Refolding spectra for the (a) heterotrimers and (d) homotrimers plotted as MRE vs time. Refolding spectra for the (b) heterotrimers and (e) homotrimers plotted as fraction folded vs time. (c) Refolding spectra in (b) zoomed in to clearly show $t_{1/2}$ for all the heterotrimers. (f) Refolding spectra in (e) zoomed in to clearly show $t_{1/2}$ for all the homotrimers.

for assembly in our system is responsible for this dramatic acceleration in folding, perhaps because of the long-range nature of these interactions as compared to hydrogen bonding. An additional factor that may be responsible for this difference is the higher percentage of amino rather than imino acids in our systems. In summary, all heterotrimers were found to reach 85% folding in approximately 50 min. The nonmutated heterotrimer refolded fastest, followed by single mutation, double mutation, and triple mutation heterotrimers. Combining the results from the unfolding and refolding experiments, we can conclude that for our system the heterotrimer which folds the slowest is the least thermally stable and the heterotrimer which folds the fastest is the most thermally stable. Both the refolding rates and thermal stability are affected by substituting glycine with a bulkier amino acid, and the magnitude of these effects is entirely dependent upon the site and frequency of the glycine mutation. For future studies, we plan to explore the effects of shorter flanking regions and longer guest regions so as to accurately estimate the factors responsible for the known phenotypes of the mutations found in type I collagen, at position 247 in particular and any other position in general using the same design principle as described here.

Conclusions

Substitution of glycine by bulky amino acid residues results in a multitude of disorders in natural homotrimeric and heterotrimeric collagens and leads to diseases such as OI, Alport syndrome, and Ehlers−Danlos syndrome to name a few.[10–13] Designing systems that can accurately mimic these mutations in synthetic analogs can provide us with a means to understand the structural and biological changes observed in mutant collagens. Various parameters including mutation site, chain type, amino acid residue replacing glycine, and neighboring sequence can affect the phenotype of a disorder,[10,11,14,15] and these parameters need to be studied in their native form in order to estimate the true effect of these mutations. Synthetic homotrimers can have the mutation in either none or all three chains and thus are limited to the study of only naturally occurring homotrimers in their native forms. ABC heterotrimers can be designed to assemble three completely different sequences in the center of the triple helix and give us the ability to study naturally occurring heterotrimers in their native forms. Three different sequences can mimic a natural ABC heterotrimer, whereas two similar sequences and a third different sequence can mimic a natural AAB heterotrimer. In the study reported here, we mimic the sequence 242−250 of type I collagen, with a glycine to serine mutation at position 247 in either of the α1 or α2 chains. We are able to differentiate between four triple helices that vary only in the frequency of glycine mutations at a particular position. The ease of preparation of heterotrimers and this degree of resolution in terms of separating single mutations can have major implications in our understanding of the mutations in natural collagens that lead to various connective tissue disorders in general and OI in particular.

Acknowledgment. This work was funded by the Robert A. Welch Foundation Research Grant No. C1557 and NSF CAREER Award DMR-0645474. We thank Siao Zhang for assistance with HPLC and CD.

Supporting Information Available: MALDI MS data, HPLC, and additional CD analysis. This material is available free of charge via the Internet at http://pubs.acs.org

JA801670V