Structural Heterogeneity of Type I Collagen Triple Helix and Its Role in Osteogenesis Imperfecta

Received for publication, July 13, 2007, and in revised form, November 19, 2007. Published, JBC Papers in Press, December 11, 2007. DOI 10.1074/jbc.M705773200

Elena Makareeva, Edward L. Mertz, Natalia V. Kuznetsova, Mary B. Sutter, Angela M. DeRidder, Wayne A. Cabral, Aileen M. Barnes, Daniel J. McBride, Joan C. Marini, and Sergey Leikin

From the 4th Section on Physical Biochemistry and the 5th Bone and Extracellular Matrix Branch, NICHD, National Institutes of Health, Bethesda, Maryland 20892 and the 6th University of Maryland School of Medicine, Baltimore, Maryland 21201

We investigated regions of different helical stability within human type I collagen and discussed their role in intermolecular interactions and osteogenesis imperfecta (OI). By differential scanning calorimetry and circular dichroism, we measured and mapped changes in the collagen melting temperature ($\Delta T_m$) for 41 different Gly substitutions from 47 OI patients. In contrast to peptides, we found no correlations of $\Delta T_m$ with the identity of the substituting residue. Instead, we observed regular variations in $\Delta T_m$ with the substitution location in different triple helix regions. To relate the $\Delta T_m$ map to peptide-based stability predictions, we extracted the activation energy of local helix unfolding ($AG^\ddagger$) from the reported peptide data. We constructed the $AG^\ddagger$ map and tested it by measuring the H-D exchange rate for glycine NH residues involved in interchain hydrogen bonds. Based on the $\Delta T_m$ and $AG^\ddagger$ maps, we delineated regional variations in the collagen triple helix stability. Two large, flexible regions deduced from the $\Delta T_m$ map aligned with the regions important for collagen fibril assembly and ligand binding. One of these regions also aligned with a lethal region for Gly substitutions in the $\alpha_1(I)$ chain.

The mature type I collagen molecule is a 300-nm-long triple helix formed by two $\alpha_1(I)$ and one $\alpha_2(I)$ chains, which is flanked by short terminal peptides. Based on the (Gly-Xaa-Yaa)$_{128}$ triplet repeat within each chain (1), the triple helix is commonly viewed as a single domain (Xaa and Yaa stand for variable residues). However, this picture may not accurately represent variations in the stability of different regions within the triple helix that fold and unfold cooperatively (2–4). The triple helix is only metastable or marginally stable at physiological temperature (3, 5–8). Its local structure appears to be highly dynamic and intimately related to local stability. The more labile regions may exist in a loose conformation, constantly undergoing unfolding/refolding transitions while more stable “clamp” regions prevent unfolding of the whole molecule (3, 9–12). Such structural and dynamic heterogeneity is believed to play an important role in self-assembly (9, 13) and function (3) of collagen fibers.

Significant progress in understanding regional variations in the triple helix stability in different collagens has been reported in recent years. For example, some flexible sites were localized by observing triple helix bending in electron microscopy (14) and/or increased susceptibility to proteolytic cleavage (9, 15). Genetically generated reshuffling of different triple helix regions was shown to have a significant effect on the overall stability of the molecule (11, 16). Relative local stability maps were proposed based on scoring different sequences (17) or on the denaturation temperature ($T_m$) measured for triple-helical host-guest peptides (18).

The existence of looser, less stable, and tighter, more stable structural regions within the triple helix is now commonly accepted. However, knowledge of their locations, properties, and role in type I collagen function is still incomplete. In particular, it is still not understood if and how these regions affect the osteogenesis imperfecta (OI) phenotype.

OI is an autosomal dominant bone disorder caused primarily by mutations in type I collagen (19, 20). Over 80% of lethal to moderately severe cases result from substitutions of the obligatory Gly residues that disrupt the triple helix, clearly pointing to structural defects as the underlying cause (21). However, no correlations of OI severity with the predicted local stability of the helix at the substitution site or with the change in the measured denaturation temperature of the whole helix ($\Delta T_m$) were found (17, 21–28). Regional variations in collagen structure and stability may modulate the structural and functional consequences of otherwise similar mutations, thereby contributing to the lack of simple correlations between the structural disruptions and OI phenotype.

Note that by local properties or disruptions of the triple helix we mean those involving less than 5–10 adjacent Gly-Xaa-Yaa (GXY) triplets. This is the characteristic distance scale of variations in the reported local stability maps (17, 18). We refer to disruptions or average helix properties on much larger distance scales as regional.

One approach to better understanding regional variations in the triple helix structure and stability is a systematic study of

---

*This work was supported in part by the Intramural Research Program of NICHD, National Institutes of Health, Department of Health and Human Services. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 The on-line version of this article (available at http://www.jbc.org) contains supplemental data, Figs. S1–S8, and Equations S1–S5.

2 To whom correspondence should be addressed: National Institutes of Health, Bldg. 9, Rm. 1E-127, Bethesda, MD 20892. Tel.: 301-594-8314; Fax: 301-402-0292; E-mail: leikins@mail.nih.gov.

3 The abbreviations used are: OI, osteogenesis imperfecta; EDS, Ehler-Danlos syndrome; CAPS, 3-(cyclohexylamino)propanesulfonic acid; DSC, differential scanning calorimetry; MLBR, major ligand-binding region.
molecules with Gly substitutions by the same techniques under similar conditions. For instance, such studies of α1(I) mutations at the N-terminal end of collagen allowed us to delineate a highly stable N-anchor region separated from the rest of the triple helix by a flexible microunfolding site (29). This region is formed by ~85 N-terminal residues in each chain (see “Discussion”). Mutations within this region cause structural changes in the molecule, resulting in a distinct OI/EDS phenotype with significant ligament laxity, characteristic of Ehlers-Danlos Syndrome (EDS) (30). Analysis of these structural changes allowed us to propose a molecular mechanism for the EDS-like symptoms (29, 30).

Another approach is to continue development of the local stability models. Many efforts were invested into detailed studies of collagen-like peptides with different sequences (see e.g. Refs. 18, 28, 31, 32 and references therein). We utilized the peptide-based local stability models (17, 18) as one of the tools for delineating the N-anchor region (29). However, accurate evaluation of the local stability of full-length collagens from such data is still a challenge (18). The existing models have not been experimentally tested as yet.

In the present study, we combine both of these approaches: (i) We report measurements of ΔT_m for 41 Gly substitutions from OI patients, all performed at the same conditions by differential scanning calorimetry and circular dichroism. We test the effects of the identity and the position of the substitution along the triple helix. We differentiate contributions from molecules with different chain composition and map ΔT_m for molecules with one mutant chain. (ii) We generalize the peptide-based models for quantitative prediction of the local stability of type I collagen. We test the predictions and determine the corresponding model parameters by measuring the H-D exchange kinetics of glycine amides involved in interchain hydrogen bonds. (iii) Based on the measured ΔT_m map and calculated local stability map, we discuss regional variations in the triple helix stability and structure. We delineate two large, flexible regions, which may be important for collagen-collagen interactions, ligand binding, and OI phenotypes.

To better understand our findings and eliminate potential artifacts, we utilize a variety of techniques and approaches. In the main text, we focus only on the most crucial results and ideas. The auxiliary techniques, results, and analyses are presented in the supplemental material.

**EXPERIMENTAL PROCEDURES**

**Collagen Preparation**

**Human Procollagen**—Skin fibroblast cultures were established from dermal punch biopsies collected from OI patients under an IRB-approved protocol. Normal control cells CRL-2127 were purchased from American Type Culture Collection (Manassas, VA); GM04501, GM04503, GM07525, GM07753 were purchased from Coriell cell repositories (Coriell Institute for Medical Research, Camden, NJ). Fibroblasts were grown to confluence at 37 °C in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum and 2 mM glutamine in the presence of 5% CO_2_. Cultures were stimulated with Dulbecco’s modified Eagle’s medium supplemented with 0.1% fetal bovine serum, 2 mM glutamine, and 50 μg/ml ascorbate. After 24 h, medium was harvested, buffered with 100 mM Tris-HCl pH 7.4, mixed with protease inhibitors (25 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 5 mM benzamidine, 10 mM N-ethylmaleimide, and 0.2% NaN_3) and cooled to 4 °C. Procollagen from media was precipitated by 176 mg/ml ammonium sulfate.

**Pepsin-treated Human Collagen**—The ammonium sulfate precipitates were suspended in 0.5 M acetic acid and digested by pepsin (~0.1 mg/ml, ~1:10 pepsin:collagen) overnight at 4 °C. After digestion, collagen was selectively precipitated with 0.9 M NaCl (final concentration), suspended in 0.5 M acetic acid, and reprecipitated with 0.7 M NaCl. The pellet was washed with 70% ethanol and redissolved and dialyzed against 0.2 M sodium phosphate, 0.5 M glycerol buffer, pH 7.4. Purity of the preparation, triple helix resistance to pepsin at 4 °C, and the content of type III and type V collagens were evaluated by gel electrophoresis as described previously (29). Collagen concentration in each preparation was measured by circular dichroism (see below).

**Rat Tail Tendon Fibers**—Frozen tails of 6-week-old rats were purchased from Pel-Freeze Biologicals and stored at ~80 °C. Tendons were excised, washed in 3.5 M NaCl, 10 mM Tris, 20 mM EDTA, 2 mM N-ethylmaleimide, 1 mM phenylmethylsulfonyl fluoride, pH 7.5 and stored in this buffer at 4 °C. Crystalline organization of collagen in tendon fibrils was disrupted by overnight equilibration in 0.2 M sodium phosphate, 1 mM glycerol, pH 7.5 at 4 °C. Tendons were washed in 20 mM NaCl, 10 mM HEPES, pH 7.5 for several hours at 25 °C. A small fiber was separated from tendon with needlepoint tweezers, stretched, covered by a droplet of the NaCl/HEPES buffer, sandwiched between two 13 × 2 mm, round CaF_2_ infrared windows (one window with “in” and “out” ports for solution exchange) and sealed by a thin layer of ultra high-vacuum Apiezon L grease (M&I Materials, Manchester, UK) around window perimeter.

**Thermal Stability of Human Collagen**

**Differential Scanning Calorimetry**—DSC measurements were performed with 0.1–0.4 mg/ml procollagen and pepsin-treated collagen solutions in 0.2 M sodium phosphate, 0.5 M glycerol, pH 7.4. The thermograms were recorded from 10 to 60 °C in Nano II or Nano III calorimeters (Calorimetry Sciences Corporation, Lindon, UT). Except for α2(I)-G805D, all mutant collagens were scanned at 0.125 or 0.25 °C/min heating rate. We found that this 2-fold change in the heating rate results in a 0.4 ± 0.1 °C shift of the whole thermogram with negligible distortion of the shape of the denaturation peaks. Because of limited protein amount, α2(I)-G805D was scanned only at 1 °C/min, which results in a detectable distortion of the peak shape, reducing the accuracy of ΔT_m measurement to ~0.5 °C. All thermograms were corrected by subtraction of a linear baseline and normalized per unit collagen concentration (for quantitative analysis) or to give the same peak height (for visual comparison). To extract the contribution of type I collagen, whenever necessary we utilized a variety of techniques and approaches. In the main text, we focus only on the most crucial results and ideas. The auxiliary techniques, results, and analyses are presented in the supplemental material.
mutant peaks in DSC thermograms was performed with the PeakFit software (Systat, PointRichmond, CA).

Circular Dichroism—CD spectra of pepsin-treated, purified collagen in 0.2 M sodium phosphate, 0.5 M glycerol, pH 7.4 (PG) or in 2 mM HCl, pH 2.7 were measured in a quartz micro-cuvette with 1-cm optical path length in a J810 spectropolarimeter equipped with a PFD-425S thermoelectric temperature controller (Jasco, Easton, MD). Collagen concentration in the solution was evaluated from the ellipticity at 221 nm at 20–25 °C. Triple helix unfolding upon heating was evaluated from the ellipticity change at 223.8 nm (see supplemental material). The temperature in the cuvette was calibrated by a thermocouple probe inserted into the buffer solution. CD spectra were recorded after 2–5 min equilibration at the selected temperature.

Infrared Spectroscopy of Rat Tail Tendon Fibers

Infrared (IR) spectra of solvated rat tail tendon fibers were measured in a Continuum IR microscope attached to a Nexus 670 Fourier transform infrared spectrometer with a 15× Reflecchromat IR objective/condenser and a narrow band MCT/A detector (Thermo Electron Corp., Madison, WI). Tendon fibers sandwiched between CaF2 IR windows were mounted in a custom built, thermostated (±0.05 °C), flow-through IR cell with built-in passive compensation for thermal expansion (United States Patent Applications 10/926,405; 11/826,806). The cell was designed for high mechanical and optical stability, allowing measurements of highly reproducible spectra (±10−4 OD units) of solvated samples (33). The mounted sample was flattened by slight mechanical pressure to the optical path of 5–7 μm and extensively washed with 3 mM CAPS, 5 mM NaCl, pH 10.0. The IR beam was restricted by a 150 × 40 μm rectangular mask and positioned in the middle of a ~70-μm wide collagen sample. To obtain isotropic spectra, the beam was polarized before the sample with a BaF2 holographic polarizer at the “magic angle” of 55 deg with respect to the fiber direction. The spectra were collected from 900 to 7000 cm−1 at 4 cm−1 resolution by accumulation of 50–200 interferograms, depending on the desired time resolution. The H-D exchange reaction was initiated by rapid (≤10 s) replacement of the H2O buffer with a D2O buffer with matching buffer composition and pH (glass electrode reading +0.4). The disappearance of water O-H spectral bands (3000–3700 cm−1) indicated replacement of all interstitial H2O by D2O within 1 min. The spectra were collected with 1–30 min periodicity for up to a week at constant temperature. The fraction of unexchanged Gly-NH groups involved in interchain hydrogen bonds was measured at pH 10 from intensities of amide II (integrated from 1515 to 1600 cm−1) and amide A (integrated from 3150 to 3450 cm−1) bands. The intensities within these spectral regions in collagen with fully deuterated Gly amides were measured after pre-equilibration at 50 °C, pH 10 for 10 h.

Model Analysis

For quantitative analysis and comparison with experiments, we describe local stability of collagen in terms of activation energy for local triple helix unfolding. We calculate this activation energy from the amino acid sequence of collagen and apparent denaturation temperatures (Tm) of host-guest peptides reported by Persikov et al. (18) within the following model.

Denaturation of Host-guest Peptides—The Tm of host-guest peptides have been investigated mostly in a kinetically limited regime, in which the transition is irreversible at the experimental time scale (34). At the same heating rate, the activation free energy of the transition G‡ is related to the apparent Tm as in Equation 1 (see supplemental material).

\[
\delta G^\text{‡} \approx \left( \frac{H^\text{m}}{T_0} \right) \delta T_m
\]

Here T0 is a reference temperature; \( \delta T_m = T_m - T_0 \), and \( \delta G^\text{‡} = G^\text{‡} (T_m) - G^\text{‡} (T_0) \) and \( G^\text{‡} \) and \( H^\text{m} \) are the activation free energy and enthalpy for a reference peptide which denatures at \( T_0 \). Hereafter we refer to thermodynamic potentials only at the reference temperature \( T_0 \) and omit the corresponding argument in their notation.

In the IR experiment at pD 10, only Gly-Gly-Gly (GXY) triplet i, which breaks interchain hydrogen bonds and exposes buried Gly-NH to water. It affects adjacent triplets \( j \) whose contribution is weighted by \( \exp[-(i-j)^2/r^2] \), where \( r \) characterizes the extent of the perturbation. (ii) The sequence dependence of the activation energies is the same in collagen and host-guest peptides. (iii) The contribution of each chain in type I collagen is independent and additive, even though type I collagen is a heterotrimer while all studied host-guest peptides were homotrimers (18).

Within this model, the difference \( \Delta G_j = G_j^\text{‡} - G^\text{‡} \) between the activation free energy for local unfolding at triplet i ([Gj]) and within a reference sequence \( G^\text{‡(ref)} \) is Equation 3,

\[
\Delta G_j^\text{‡} = \sum_i \delta G_i \exp[-(i-j)^2/r^2] \tag{Eq. 3}
\]

where \( \delta G_i = (2/3)\delta G^\text{GXY}(\alpha_1) + (1/3)\delta G^\text{GXY}(\alpha_2) \), \( \delta G^\text{GXY}(\alpha_1) \), and \( \delta G^\text{GXY}(\alpha_2) \) are given by Equation 2, and GXY(\alpha1) and GXY(\alpha2) are the sequences within \( \alpha_1 \) and \( \alpha_2 \) chains of the triplet \( j \), respectively.

H-D Exchange Kinetics—In the IR experiment at pH 10, only glycine amides contribute to the kinetics of H-D exchange. All other exchangeable protons are substituted instantaneously. The kinetics is determined entirely by local unfolding exposing glycine amides to water, after which the exchange is instantaneous as well. We approximate the number of unexchanged
protons in glycine amides at time $t$ after replacing $H_2O$ by $D_2O$ by Equation 4,

$$N(t) = \sum \exp(-k_i t)$$  

(Eq. 4)

where, as above, the index $i$ labels GXY triplets in the triple helix and Equation 5,

$$k_i = k_0 \exp(-\Delta G_i^f/RT)$$  

(Eq. 5)

is the characteristic rate constant for local triple helix unfolding at the triplet $i$, $\Delta G_i^f = G_i^f - G_i^f_{ref}$ is given by Equation 3, $R$ is the universal molar gas constant, and $k_0$ is the local H-D exchange rate constant within the same reference sequence as used for the definition of $G_i^f_{ref}$.

**RESULTS**

**Thermal Stability of Gly Substitution Mutants**—We systematically examined thermal stability of type I collagen secreted by cultured fibroblasts from 5 normal controls and 47 heterozygous patients with 41 different Gly substitutions. We measured DSC thermograms of pepsin-treated collagen at 0.125–1 °C/min heating rate (Figs. 1 and 2; Table 1) in 0.5 M glycerol, 0.2 M phosphate-glycerol buffer, as previously predicted (6) and confirmed for normal control molecules (8) and some mutants (supplemental material).

At the same heating rate, DSC thermograms of pepsin-treated type I collagen secreted by normal control cells were identical. They had a single sharp maximum (Fig. 1A) at the apparent denaturation temperature ($T_m$). The values of $T_m$ varied up to ±0.3 °C between the cell lines and even between different cultures of the same cell line (see supplemental material). We observed similar $T_m$ variation in pepsin-treated collagen purified from tail tendons of wild-type mice (35). DSC thermograms from the same collagen batch (from the same human cell culture or the same mouse tail tendon) were reproducible within 0.1 °C. Therefore, ±0.3 °C variation is not related to the measurement technique or the instrument but rather to variation in the properties of the collagen itself. One possible source is variability in posttranslational modification, e.g. due to uncontrolled changes in cell culture conditions. For instance, overmodification of collagen upon incubation of cells at 40 °C caused an almost 1 °C increase in the $T_m$ (36).

In most mutants we observed additional peaks with lower $T_m$ (Figs. 1, B–D and 2), corresponding to denaturation of molecules containing one or two mutant chains. Pepsin-treated secreted collagens with the same mutation from unrelated patients had denaturation peaks at the same temperatures. All major DSC peaks corresponded to denaturation of different molecules rather than different domains within the same molecule (supplemental material).

Three peaks, corresponding to $\alpha_1$-$\alpha_2$, $\alpha_1\alpha_1\alpha_2$, and $\alpha_1\alpha_2\alpha_2$ (the superscript labels the mutant chain) were observed in $\alpha_1\alpha_1$-G76E (29) and $\alpha_1\alpha_1$-G523C (Fig. 1B). The area under each peak (estimated by peak deconvolution) represents the fraction of the corresponding molecules in the mixture (Table 1). The differences between the fractions measured from DSC thermogram deconvolution (Table 1) and expected for heterozygous $\alpha_1$($I$) mutations (25%, $\alpha_1\alpha_2$; 50%, $\alpha_1\alpha_1\alpha_2$; and 25%, $\alpha_1\alpha_2\alpha_2$) are related to intracellular degradation of a fraction of mutant molecules. This conclusion was confirmed by studies of intracellular collagen denaturation (Table 1 and supplemental material). Note that within 2–5% experimental
secretion of molecules with one mutant chain was sufficient for
preparations from the cell culture media.

cells and become undetectable by DSC (and/or two mutant chains may be retained and degraded by
different content of mutant chains may result in fusion of the

From DSC thermograms we determined the difference
(\(\Delta T_m\)) in the denaturation temperatures of molecules with one
and no mutant chains, as illustrated in Fig. 1. In mutants with
one broad DSC peak, \(\Delta T_m\) was estimated by thermogram
deconvolution (Fig. 1D). Only \(\alpha_2(1)-G805D\) did not have a
detectable effect on \(T_m\) (although a lower resolution DSC at
1 °C/min was utilized due to limited protein availability). All
other studied substitutions reduced the triple helix \(T_m\) by 0.8 °C
to 4.6 °C (Table 1). Based on comparison of \(\Delta T_m\) for 5 different
cell lines with the same \(\alpha_1(1)-G589S\) substitution, we estimated
reproducibility of \(\Delta T_m\) as ±0.3 °C, consistent with the \(T_m\)
reproducibility in normal controls.

Mapping \(\Delta T_m\) versus substitution location in the triple helix
for pepsin-treated molecules with one mutant chain revealed
a regular pattern (Fig. 3). At the same time, we found no
apparent correlations of \(\Delta T_m\) with the substitution identity
when all mutations were pooled together (Fig. 3, inset).
Moreover, \(\Delta T_m\) were the same (±0.3 °C) within the following
groups: \(\alpha_1(1)-G121D/\alpha_1(1)-G121V/\alpha_2(1)-G121D, \alpha_1(1)-G148D/
G148R, \alpha_2(1)-G247S/\alpha_2(1)-G247C, \alpha_1(1)-G247S, \alpha_1(1)-G898S/
\alpha_2(1)-G898V, \alpha_2(1)-G922S/G925R\) (Table 1). Note that
\(\alpha_2(1)-G247S\) and -G247C had different post-translational over-
modification (37).

These observations suggested that \(\Delta T_m\) variations may be
related to heterogeneous local stability of the triple helix and
concomitant variations in the extent of structural disruptions
caused by substitutions at different locations.

### Triple Helix Disruption by Gly Substitutions

We previously
found that the whole N anchor region of ~30 N-terminal GXY
(Gly-Xaa-Yaa) triplets was disrupted by Gly substitutions
within it. DSC of procollagen and altered N-propeptide cleav-
age suggested that the disruptions propagated all the way to the
N-terminal end (30). Chymotrypsin susceptibility indicated
that the disruptions propagated all the way to \(\alpha_1(1)-F92\), where
the cleavage was monitored (29).

Similar to N-anchor mutations, the thermal stability of
secreted \(\alpha_1(1)-G997S\) procollagen was higher than the pep-
sin-treated collagen (Fig. 4), indicating that the mutation
disrupted all five GPO triplets on its C-terminal side. The
C-propeptide limited the extent of the disruption, thereby
contributing to the stability of the mutant. In \(\alpha_1(1)-G967C\)
and other mutants outside the N-anchor, the stability of
secreted procollagen was the same as pepsin-treated colla-
gen. Neither \(\alpha_1(1)-G997S\) nor \(\alpha_1(1)-G967C\) had any signifi-
cant effect on the chymotrypsin susceptibility at \(\alpha_1(1)-F935\)
(supplemental material), in contrast to the increased suscep-
tibility of N-anchor mutants at \(\alpha_1(1)-F92\) (29). Gly substitu-
tions appear to cause more limited disruptions within the
last 25 triplets at the C-terminal end compared with the first
25 triplets at the N-terminal end.

---

**FIGURE 1.** Normalized DSC thermograms of selected collagens with Gly \(\rightarrow\) Ser substitutions at different locations (measured and corrected as
described in the legend to Fig. 1). In \(\alpha_1(1)-G352S, \alpha_1(1)-G448S, \) and \(\alpha_2(1)-G706S\) the contributions from molecules with two, one, and no mutant chains are
fused into a single, broadened peak (see supplemental material). The
normal control thermogram (dashed line) is overlaid onto \(\alpha_1(1)-G352S\) and
\(\alpha_2(1)-G706S\) thermograms for better visualization of the peak broadening
and apparent \(T_m\), shift caused by the mutations.

Two peaks were observed in pepsin-treated, secreted collagens
from most other mutants. A single peak was observed in
\(\alpha_1(1)-G148D, -G148R, -G352S, -G448S, \) and -G832S as well as
in \(\alpha_2(1)-G706S, -G721S, \) and -G805D (Figs. 1 and 2). A smaller
number of peaks than possible collagen compositions may have
two explanations. (i) A small difference in \(T_m\) of molecules with
different content of mutant chains may result in fusion of the
responding peaks into one (Fig. 1D). (ii) Molecules with one
and/or two mutant chains may be retained and degraded by
cells and become undetectable by DSC (<5% of total collagen)
in preparations from the cell culture media.

In all cases with fewer than expected peaks, we found that the
secretion of molecules with one mutant chain was sufficient for
detection by DSC (supplemental material). Clearly, the secre-
tion of molecules with two mutant chains was also sufficient in
\(\alpha_1(1)-G313C\) and -G967C, where S-S dimers were quantified by
SDS-PAGE (Table 1). Previously, we reported evidence of the
secretion of such molecules in \(\alpha_1(1)-G13D, -G25V, \) and -G34R
as well (29). However, we could not say with certainty whether
the molecules with two affected chains contributed to DSC
thermograms of other \(\alpha_1(1)\) mutants.
### TABLE 1

**Structural Heterogeneity of Type I Collagen***

| Mutation | OI type | $-\Delta T_m$, °C ± 0.3 | Secreced collagen (%) | Intracellular collagen (%) |
|----------|---------|------------------------|----------------------|--------------------------|
| Normal controls | | | | |
| IV | 3.2 | 12 ± 1 | 69 ± 5 | 9 ± 6 |
| G25V, P1 | IV | 4.6 | 15 ± 2 | 62 ± 4 | 11 ± 2 | 77 ± 5 |
| G25V, P2 | IV | 4.6 | 17 ± 2 | 64 ± 4 | 13 ± 2 | 82 ± 5 |
| G34R | III/IV | 2.2 (2.4) | 14 ± 2 | 65 (14) ± 5 | 7 ± 2 | 4 ± 2 | 71 ± 5 |
| G76E | III | 3.3 | 13 ± 2 | 73 (26) ± 4 | 7 ± 2 | 78 (23) ± 4 | 68 ± 10 |
| G88E | III/IV | 2.0 | 24 ± 3 | 57 ± 5 | 7 ± 2 | 65 ± 5 |
| G121V | IV | 2.1 | 9 ± 2 | 72 ± 3 | 4 ± 2 | 71 ± 5 |
| G121D | III | 2.3 | 14 ± 2 | 70 ± 5 | |
| G136R | G121D | II/IV | 2.4 | 9 ± 2 | 63 ± 5 | |
| G148D | IV | 1.2 | 9 ± 2 | 70 ± 3 | 4 ± 2 |
| G148R | IV | 1.1 | 20 ± 1 | 64 ± 5 |
| G193S | III | 1.9 | 8 ± 2 | 70 ± 3 | 2 ± 1 | 86 ± 5 |
| G217S | III | 1.8 | 15 ± 2 | 70 ± 3 | 7 ± 2 | 66 ± 7 |
| G238S, P2 | IV | 2.4 | 13 ± 2 | 47 ± 2 | 6 ± 2 | 56 ± 5 |
| G238S, P1 | IV | 2.4 | 15 ± 1 | 50 ± 2 | 10 ± 1 | 67 ± 5 |
| G247C | III | 2.2 | 8 ± 1 | 50 ± 3 | 13 ± 2 | 64 ± 5 |
| G247S | III | 2.0 | 15 ± 1 | 71 ± 4 | 7 ± 2 | 85 ± 5 |
| G247S | II/III | 1.6 | 15 ± 1 | 75 ± 5 | 7 ± 2 | 81 ± 5 |
| G286R | III | 1.6 | 13 ± 2 | 73 (26) ± 4 | 5 ± 2 | 78 (23) ± 4 | 68 ± 10 |
| G313C | IV | 1.9 | 13 ± 2 | 38 ± 1 | 7 ± 2 | 68 ± 10 |
| G337S | III/IV | 2.6 | 7 ± 2 | 31 ± 3 | 7 ± 2 | 66 ± 7 |
| G337S | IV | 1.6 | 7 ± 2 | 13 ± 2 | 38 ± 1 | 7 ± 2 | 68 ± 10 |
| G370S | III | 1.5 | 5 ± 2 | 31 ± 3 | 7 ± 2 | 66 ± 7 |
| G421D | II | 1.0 | 11 ± 3 | 68 ± 5 |
| G436R | OP* | 1.0 | 10 ± 2 | 60 ± 10 |
| G448S | IV | 1.6 | 12 ± 2 | 48 ± 2 |
| G511S | IV | 3.5 | 12 ± 2 | 50 (20) ± 5 | 5 ± 2 |
| G523C | IV | 3.3 (3.3) | 12 ± 2 | 50 (20) ± 5 | 5 ± 2 |
| G589S, P1 | III/IV | 2.6 | 23 ± 2 | 70 ± 5 | 6 ± 2 | 80 ± 5 |
| G589S, P2 | III/IV | 2.5 | 12 ± 2 | 76 ± 5 | 4 ± 2 | 77 ± 5 |
| G589S, P3 | III/IV | 2.6 | 8 ± 1 | 70 ± 5 | 7 ± 2 | 78 ± 5 |
| G589S, P4 | IV | 2.4 | 11 ± 2 | 73 ± 5 | 6 ± 2 | 78 ± 5 |
| G589S, P5 | III | 2.0 | 6 ± 1 | 68 ± 5 |
| G610C | IV | 1.6 | 11 ± 1 | 30 ± 10 | 7 ± 2 |
| G649R | III | 1.6 | 11 ± 1 | 30 ± 10 | 7 ± 2 |
| G676V | IV | 1.0 | 12 ± 2 | 7 ± 2 |
| G706S | II | 0.9 | 13 ± 3 | 7 ± 2 |
| G721S | III | 1.2 | 9 ± 2 | 7 ± 2 |
| G805D | II | 1.0 | 9 ± 2 | 7 ± 2 |
| G832S | IV | 0.8 | 12 ± 5 | 7 ± 2 |
| G898S | III | 2.5 | 16 ± 3 | 74 ± 4 |
| G898V | III | 2.7 | 25 ± 4 | 50 ± 5 |
| G922S | IV | 2.7 | 12 ± 3 | 45 ± 4 |
| G925R | IV | 2.6 | 23 ± 2 | 44 ± 4 |
| G967C | III | 2.5 | 22 ± 3 | 59 (15) ± 5 | 8 ± 2 | 83 (35) ± 5 |
| G997S | III | 1.9 | 19 ± 3 | 55 ± 10 |

*OP, osteoporosis, vertebral compressions, no fractures.

Interestingly, we found an evidence of a ~100 triplet disruption in pepsin-treated α2(I)-G98V collagen. A shoulder in the DSC thermogram (not expected in α2(I) mutants) indicated melting of a large domain below the $T_m$ of the whole molecule (supplemental material). The contributions from molecules with and without the mutant chain were better resolved in 2 mM HCl (supplemental material). Subsequent CD measurements in 2 mM HCl revealed reversible unfolding ~30% of the triple helix below $T_m$ of the whole molecule (Fig. 5). We applied a train of temperature pulses, heating the sample from 20 to 35.5°C and cooling back to 20°C, as shown in Fig. 5B. In normal control, the first 35.5°C pulse decreased the ellipticity at 223.8 nm due to denaturation of ~5% damaged molecules present in the sample (Fig. 5A). Further pulses caused no additional changes. In α2(I)-G98V, the first 35.5°C pulse caused ~20% change in the ellipticity, of which only ~5% was due to irreversible denaturation of damaged molecules. Subsequent pulses caused ~15% oscillations due to reversible unfolding of a large triple helical region within mutant molecules (Fig. 5A), consistent with DSC observations. The upward drift of the curve was caused by irreversible complete unfolding of a small fraction of mutant molecules. Because the reversible unfolding occurred only in molecules with the mutant chain, comprising ~50% of the mixture (Table 1), we estimated that it involved ~30% of the triple helix length or ~100 triplets. The same region also appeared to be disrupted by α2(I)-G922S. Domain melting of the latter molecules was revealed by DSC in 2 mM HCl after C-terminal end truncation at or near α1(I)-F3935 by chymotrypsin (supplemental material).

**Analysis of Local Stability by H-D Exchange**—The heterogeneity in local stability of type I collagen triple helix can be experimentally studied using hydrogen to deuterium (H-D) exchange.

---

*OP, osteoporosis, vertebral compressions, no fractures.
Structural Heterogeneity of Type I Collagen

....

FIGURE 3. \( \Delta T_m \) map for collagens with different Gly substitutions in one polypeptide chain labeled by the corresponding black (\( \alpha(1)I \)) or gray (\( \alpha(2)I \)) one-letter amino acid symbols. The values of \( \Delta T_m \) were measured as shown in Fig. 1. The inset summarizes \( \Delta T_m \) dependence on the substitution identity. Analysis of the inset suggests no statistically significant correlations of \( \Delta T_m \) with the substitution identity. Note that small differences in the apparent average \( \Delta T_m \) for different residues may be related to special locations of few outlying mutations. For instance, the average \( \Delta T_m \) is the largest for V (\(-2.6^\circ C\)) and the smallest for D (\(-1.6^\circ C\)). However, upon exclusion of just \( \alpha(1)I\)-G25V and \( \alpha(2)I\)-G805D from these sets, the average \( \Delta T_m \) for V (\(-1.9^\circ C\)) and D (\(-2.0^\circ C\)) become the same and consistent with the overall average \( \Delta T_m \) (\(-2.1^\circ C\)). Both excluded mutations are located within special triple helix regions, see “Discussion.” Moreover substitutions of the same Gly with these residues, \( \alpha(1)I\)-G121V and -G121D, result in the same \( \Delta T_m \), within the 0.3 °C reproducibility of the measurements (Table 1).

kinetics of amide NH groups. At physiological conditions, the H-D exchange is catalyzed by OD\(^{-} \) ions in surrounding water (38, 39). Because Gly-NH groups in the GXY triplets are buried in the core of the helix, their H-D exchange rate is limited. Local triplet helix unfolding is required to expose them to surrounding water and OD\(^{-} \) (2, 38, 40). Different rates of local unfolding within different regions of the helix may contribute to the observed wide range of slow exchange rates (from minutes to days). However, the H-D exchange measurements in collagen reported previously (2, 40–42) did not differentiate Gly-NH within GXY triplets from other amides and did not distinguish the contributions of the unfolding and catalytic exchange processes into the overall reaction rate.

To capture the kinetics of local triple helix unfolding, we studied slow (>1 min) time evolution of amide II and amide A bands in infrared (IR) spectra of rat tail tendon fibrils at pD 10. The high pD decreased the time of OD\(^{-} \) catalyzed H-D exchange in partially exposed NHs to less than a second. It eliminated the contributions of non-Gly amides and of the exchange process itself. The observed frequency of the amide A band of the slowly exchanging groups (>3331 cm\(^{-1} \)) was consistent with that expected (43–46) for Gly-NH involved in interchain hydrogen bonds within the triple helix and much higher than observed for other amides (3280–3310 cm\(^{-1} \)).

The relative amplitude of the slow exchange was consistent with the fraction of Gly-NH calculated from the amide acid sequence. The exchange rate was independent of pD, confirming that it was determined entirely by the rate of local unfolding.

The measured time evolution of the fraction of unexchanged amides (Fig. 6) was, therefore, determined only by the kinetics of local triple helix unfolding. At the same time, high pD/pH did not cause appreciable changes in high definition IR spectra of the triple helix at or below 50 °C, indicating that the normal collagen structure and dynamics were fully preserved.

The measured time evolution of the normalized fraction of unexchanged Gly-NH (Fig. 6) indicates the presence of a broad variation in stabilities of different triple helical regions. The data exhibit a broad distribution of helix unfolding rates \( k \) from \( k_{\text{max}} \sim 0.5 \text{ min}^{-1} \) to \( k_{\text{min}} \sim 4 \times 10^{-3} \text{ min}^{-1} \). The range of variation in the activation energy of local unfolding estimated from the unfolding rates was \( \Delta G^\ddagger \sim 4.5 \text{ kcal/mol} \) (\( \Delta G^\ddagger \sim RT \ln(k_{\text{max}}/k_{\text{min}}) \)).

Mapping of Local Stability from Peptide Data—Mapping of the local triple helix stability was previously based on relative stability scores (17) or \( T_m \) of peptides with different guest GXY triplets inserted in the middle of a (GXY)\(_n\) host (18). Both models, however, expressed the stability in relative units appropri-
Structural Heterogeneity of Type I Collagen

FIGURE 5. Reversible unfolding of a large region in molecules with the mutant α2(I)-G898V chain. The fraction (%) of denatured protein (A) in response to a train of heating pulses (B) was measured by monitoring ellipticity of collagen solution in 2 mM HCl (pH 2.7) at 223.8 nm.

FIGURE 6. Kinetics of local triple helix unfolding in rat tail tendon measured by exchange of glycine amide protons to deuterium. The inset shows the same graph with logarithmic ordinate. Squares and circles represent the results of two independent measurements of the amount of unexchanged Gly-NH normalized by extrapolation to zero time. The solid line was calculated from the local stability model described under “Experimental Procedures” (Equations 4 and 5). Five triplets from each end of the triple helix were excluded from the calculation since the model may fail within these regions. The number of triplets from each triplet excluded from the calculation affected only the extreme tail of the kinetic curve at characteristic times >1000 min. The best agreement of the theoretical curve with the observed H-D exchange was achieved at $H_0^0 = 45 \pm 5$ kcal/mol and $n = 5 \pm 1$.

DISCUSSION

Δ$T_m$, Map of Human Type I Collagen

Substitution Identity—At the onset of this study, we expected systematic DSC measurements to reveal larger Δ$T_m$ for Gly substitutions with Val, Arg, Asp, or Glu, which are more destabilizing in peptides (24) and tend to be more clinically severe (21). Instead, we found no overall correlations of Δ$T_m$ with the substitution identity (Fig. 3, inset). Moreover, substitutions of the same Gly and most substitutions within adjacent GXY triplets resulted in the same Δ$T_m$ (±0.3 °C), regardless of their identity or the chain in which they occurred (Table 1 and Fig. 3).

At first glance, these findings seem counterintuitive because different substitutions cause different local disruptions of the triple helix (24). However, Δ$T_m$ measures the global stability of the molecule, which may be the same at different local disruptions. For instance, two mutant α1(I) chains result in stronger local disruptions than one chain but the same Δ$T_m$ in G13D, G25V, and G34R (29). One mutant chain disrupts a large region surrounding the mutation, apparently eliminating the contribution of the region into $T_m$. Two mutant chains disrupt this region at a lower temperature, but Δ$T_m$ remains the same, being determined by the helix stability outside the disrupted region. Similar Δ$T_m$ in most molecules with two and one mutant chain may explain why we observed three DSC peaks only in 2 of 23 studied α1(I) mutations.

More importantly, we found no effect of the substitution identity or chain on Δ$T_m$ at the same site in the tested cases. These cases included three pairs of substitutions of the same Gly in the same chain (G121D/G121V and G148D/G148R in α1(I) and G247S/G247C in α2(I)), two pairs of the same substitution in different chains (G121D/G121D and G247S/G247S), three pairs of different substitutions in different chains, (G121D/G121V, G247S/G247C, and G898S/G898V), and one pair of adjoining substitutions (G922S/G925R in α2(I)). More testing would be interesting, but considering insufficient number of available cell lines, in the present study we focused on the role of the mutation location instead.

Substitution Location—Mapping of Δ$T_m$ versus mutation location revealed a regular pattern (Figs. 3 and 7). Moreover, the pattern for just Gly → Ser substitutions at 16 different locations appeared to be the same as for various other substitutions pooled together. Evidently, Δ$T_m$ depends on which region is disrupted by the mutation (17). Larger Δ$T_m$ are caused by disruptions of stronger regions, contributing more to the stability of the whole helix. Smaller Δ$T_m$ are caused by disruptions of weaker regions, contributing less. The Δ$T_m$ map then reveals regional contributions into the global triple helix stability. We observed average Δ$T_m$ $\approx -2.1$ °C (±2σ, where σ = 0.3 °C is the Δ$T_m$ reproducibility) for most mutations, except for three distinct regions marked on the Δ$T_m$ map in Fig. 7.
N-anchor is a highly stable region extending from Gly-1 to Gly-85 (29). In all Gly substitutions within this region we observed large $\Delta T_m$ from 3.2 to 4.6 °C in molecules with one and/or two mutant chains (Table 1). These substitutions disrupted the whole region (29, 30). Even $\alpha$1(I)-G88E, located at the region boundary, appeared to have an effect on N-propeptide cleavage on the other side of the region (30).

Mid-flex is a flexible region in the middle of the molecule, in which triple helix weakness results in less than average contribution to the overall stability. This region includes substitutions of four Gly residues from Gly-352 through Gly-436 with $\Delta T_m$, from 0.8 to 1.5 °C (Fig. 7 and Table 1). The exact boundaries of this region are presently difficult to pinpoint. For instance, we excluded adjacent $\alpha$2(I)-G337S and -G448S, which have $\Delta T_m = -1.6$ °C. Given the 0.3 °C uncertainty in the $\Delta T_m$, such exclusion is somewhat arbitrary.

C-flex is a flexible region near the C-terminal end of the molecule, which has even lower stability. It includes substitutions of five residues from Gly-676 through Gly-832 with $\Delta T_m$ from 0 to $-1.2$ °C. As with the Mid-flex region, the exclusion of $\alpha$2(I)-G610C and -G649R (both with $\Delta T_m = -1.6$ °C) from this region is somewhat arbitrary. The exact C-terminal boundary of this region is not defined yet as well, because the closest studied mutations are located at Gly-898.

Presently, we do not know whether $\alpha$1(I)-G148D/G148R ($\Delta T_m \approx -1.2$ °C) and $\alpha$2(I)-G511S/α1(I)-G523C ($\Delta T_m \approx -3.4$ °C) are just “outliers” or they are located within other distinct structural regions, as the mutations discussed above. Our
Structural Heterogeneity of Type I Collagen

current collection of mutant cells is not sufficient for answering this question.

Local Stability (ΔG‡) Map

To test whether DTm are correlated with the local stability of the triple helix at the substitution site (17), we extended the model proposed in (18). From host-guest peptide stabilities reported in (18) we calculated sequence dependent activation free energy ΔG‡ for local helix opening. We measured H-D exchange upon concomitant exposure of normally-buried amide groups of Gly to water. We found the exchange kinetics to be in good agreement with the one predicted from the ΔG‡ calculations (Fig. 6). This agreement suggests that the ΔG‡ map (Fig. 7) is a reasonable working model. However, it does not guarantee every detail of the map to be accurate, because the exchange kinetics does not resolve such details (Equations 4 and 5).

When comparing the ΔTm and ΔG‡ maps (Fig. 7), it is important to keep in mind that Gly substitutions may disrupt large regions of the helix upon ΔTm measurement. Disruptions localized to just a few adjacent triplets were reported only in peptide crystals (47) or at 4 °C (48). At relevant temperatures in solution, Gly substitutions cause complete unfolding of the same crystals (47) or at 4 °C (48). At relevant temperatures in solution, Gly substitutions cause complete unfolding of the same triple helix (47), as suggested by α1(I)-G148D/R mutations with low DTm in the middle of MLBR1. We hope to test this hypothesis as soon as we have enough mutant cell lines within the region.

Potential Implications

It is instructive to overlay the ΔTm and ΔG‡ maps with the triple helix regions important for collagen function and pathology (Fig. 7).

Fibrillogenesis—From the D-period map of type I collagen packing in fibers (49–52), we see that the Mid-flex region aligns with the gap region of the D2 segment. This is exactly where the most significant triple helix bending occurs in fibrils (53), which is probably not a coincidence.

Ligand Binding—The C-flex region aligns with MLBR2, which is one of the three Major Ligand Binding Regions proposed in Ref. 54. An increased helix flexibility and less tight packing of the chains within this region are important for binding and triple helix cleavage by collagenases (55). Because tight triple helix is likely to reduce sequence recognition, the increased flexibility may be important for binding of other ligands as well.

Interestingly, MLBR1 may also align with a flexible region in the triple helix, as suggested by α1(I)-G148D/R mutations with low DTm in the middle of MLBR1. We hope to test this hypothesis as soon as we have enough mutant cell lines within the region.

OI Phenotypes of α1(I) Mutations—To date, close to 400 cases of glycine substitutions in the α1(I) chain have been described, but only two large (>20 triplets) regions with distinct OI phenotypes have been discovered (21).

Mutations within the N-anchor result in an OI/EDS phenotype (29, 30). Except for one G76R case, no lethal Gly substitutions were found within this region (21). In a previous study, we argued that these mutations cause unfolding of the whole N-anchor in procollagen secreted from cells, preventing proper N-propeptide cleavage (29). Incorporation of the uncleaved pN-collagen molecules into fibers leads to symptoms reminiscent of EDS VII A and B, caused by mutations in the N-propeptide cleavage site (30).

Essentially all (14 of 15 cases) α1(I) Gly substitutions from Gly-691 through Gly-823 were found to be lethal (21). The clinical outcome of the only case reported as non-lethal is unknown since the pregnancy was terminated after early detection of severe deformities by ultrasound. Furthermore, the abnormally low density of the substitutions discovered within this region (0.34/triplet versus average 1.16/triplet in the entire α1(I) chain) may also point to an even more severe, embryonic lethal phenotype in addition to the detected cases (21). This lethal phenotype region aligns almost perfectly with the C-flex region and MLBR2 (Fig. 7). Presently, we do not know whether such an alignment is coincidental. We can only hypothesize that weaker destabilization of the triple helix (smaller ΔTm) within this region may allow more complete folding and secretion of molecules containing mutant chains. Secretion of molecules with an altered MLBR2 structure may be particularly detrimental for bone.

OI Phenotypes of α2(I) Mutations—In about 300 reported cases of Gly substitutions in the α2(I) chain, clustering of lethal mutations was delineated (21). Most lethal cases were found within eight clusters that do not exhibit any obvious correlations with either ΔTm or ΔG‡ maps (Fig. 7). Evidently, changes in helical stability are not directly involved in the lethal outcome of heterozygous α2(I) mutations, perhaps because of the lower fraction of mutant molecules. Instead, the pattern of lethal α2(I) mutations may be determined, e.g., by altered ligand binding at disrupted mutation sites (54, 56).

Conclusions

In summary, we found: 1) no significant overall correlations of ΔTm with Gly substitution identity; 2) no straightforward correlations of ΔTm with the local helix stability at the substitution site; 3) a regular ΔTm variation with the substitution location, revealing regional variations in the stability and dynamic structure of the triple helix; and 4)
some correlations of these regions with those important for fiber formation, ligand binding, and OI phenotypes of α1(I) mutations.

We argue that OI severity may be related to the extent of structural disruptions caused by Gly substitutions and the resulting $\Delta T_m$, but this relationship is neither simple nor straightforward. For instance, the tendency of Val, Arg, Asp, and Glu to cause more severe OI compared with Ser or Ala substitutions of the same Gly (21) may be related to more extensive local helix disruptions (24). However, $\Delta T_m$ may simply not reveal the difference in the extent of local disruptions for the reasons discussed above.

It is also useful to keep in mind that the disruption of collagen structure may be the initial cause of the disease and yet only one of many factors affecting the disease outcome. Significant OI phenotype variations were reported in patients with the same mutation (21). Moreover, lethal to moderate phenotype range observed in the Brtl IV mouse model with mutation (21). Moreover, lethal to moderate phenotype range observed in the Brtl IV mouse model with mutation (21). Moreover, lethal to moderate phenotype range observed in the Brtl IV mouse model with mutation (21).

Thus, we focus on large regions of distinct structural disruptions and distinct OI phenotypes rather than on overall correlations between $\Delta T_m$ and OI severity. We hope that this approach may better reveal any underlying structure-function relationships. We believe that our progress is encouraging, but many challenges remain. The map of known OI mutations is rapidly saturating possible glycine substitution sites, but systematic studies of properties and interactions of collagens with these mutations are only beginning.

REFERENCES

1. Chu, M.-I., and Prockop, D. J. (2002) in *Connective Tissue and Its Heritable Disorders* (Royce, P. M., and Steinmann, B., eds) pp. 223–248, Wiley-Liss, New York.

2. Privalov, P. L., Tiktopulo, E. I., and Tischenko, V. M. (1979) *J. Mol. Biol.* 127, 203–216.

3. Privalov, P. L. (1982) *Adv. Protein Chem.* 36, 1–104.

4. Westerhausen, A., Kishi, J., and Prockop, D. J. (1990) *J. Biol. Chem.* 265, 13995–14000.

5. Burjanadze, T. V. (2000) *Biopolymers* 53, 523–528.

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

7. Piez, K. A. (1984) in *Extracellular Matrix Biochemistry* (Piez, K. A., and Reddi, A. H., eds) pp. 1–40, Elsevier, New York.

8. Makareeva, E., and Leikin, S. (2007) *PLoS ONE* 2, e1029.

9. Ryhanen, L., Zaragoza, E. J., and Uitto, J. (1983) *Arch. Biochem. Biophys.* 223, 562–571.

10. Burjanadze, T. V., and Bezhitadze, M. O. (1992) *Biopolymers* 32, 951–956.

11. Arnold, W. V., Fertala, A., Sieron, A. L., Hattori, H., Mechling, D., Bachinger, H. P., and Prockop, D. J. (1998) *J. Biol. Chem.* 273, 31822–31828.

12. Miles, C. A., and Bailey, A. J. (2001) *Micron* 32, 325–332.

13. Kadler, K. E., Højima, Y., and Prockop, D. J. (1988) *J. Biol. Chem.* 263, 10517–10523.

14. Hofmann, H., Voss, T., Kuhn, K., and Engel, J. (1984) *J. Mol. Biol.* 172, 325–343.

15. Kuznetsova, N. V., McBride, D. J., and Leikin, S. (2003) *J. Mol. Biol.* 331, 191–200.

16. Zafarullah, K., Sieron, A. L., Fertala, A., Tromp, G., Kuivaniemi, H., and Prockop, D. J. (1997) *Matrix Biol.* 16, 245–253.

17. Bachinger, H. P., and Davis, J. M. (1991) *Int. J. Biol. Macromol.* 13, 152–156.

18. Persikov, A. V., Ramshaw, J. A., and Brodsky, B. (2005) *J. Biol. Chem.* 280, 19343–19349.
52. Wess, T. J., Hammersley, A. P., Wess, L., and Miller, A. (1998) J. Struct. Biol. 122, 92–100
53. Orgel, J. P., Irving, T. C., Miller, A., and Wess, T. J. (2006) Proc. Natl. Acad. Sci. U. S. A. 103, 9001–9005
54. Di Lullo, G. A., Sweeney, S. M., Korkko, J., Ala-Kokko, L., and San Antonio, J. D. (2002) J. Biol. Chem. 277, 4223–4231
55. Chung, L., Dinakarpandian, D., Yoshida, N., Lauer-Fields, J. L., Fields, G. B., Visse, R., and Nagase, H. (2004) EMBO J. 23, 3020–3030
56. Forlino, A., and Marini, J. C. (2000) Mol. Genet. Metab. 71, 225–232
57. Forlino, A., Porter, F. D., Lee, E. J., Westphal, H., and Marini, J. C. (1999) J. Biol. Chem. 274, 37923–37931
58. Forlino, A., Kuznetsova, N. V., Marini, J. C., and Leikin, S. (2007) Matrix Biol. 26, 604–614
59. Forlino, A., Tani, C., Rossi, A., Lupi, A., Campari, E., Gualeni, B., Bianchi, L., Armini, A., Cetta, G., Bini, L., and Marini, J. C. (2007) Proteomics 7, 1877–1891