Consequences of Glycine Mutations in the Fibronectin-binding Sequence of Collagen

Panharith Chhum, Hongtao Yu, Bo An, Brian R. Doyon, Yu-Shan Lin, and Barbara Brodsky

From the Departments of Biomedical Engineering and Chemistry, Tufts University, Medford, Massachusetts 02155

Edited by Amanda Fosang

Collagen and fibronectin (Fn) are two key extracellular matrix proteins, which are known to interact and jointly shape matrix structure and function. Most proteins that interact with collagen bind only to the native triple-helical form, whereas Fn is unusual in binding strongly to denatured collagen and more weakly to native collagen. The consequences of replacing a Gly by Ser at each position in the required (Gly-Xaa-Yaa), Fn-binding sequence are probed here, using model peptides and a recombinant bacterial collagen system. Fluorescence polarization and solid-state assays indicated that Gly replacements at four sites within the Fn-binding sequence led to decreased Fn binding to denatured collagen. Molecular dynamics simulations showed these Gly replacements interfered with the interaction of a collagen β-strand with the β-sheet structure of Fn modules seen in the high resolution crystal structure. Whereas previous studies showed that Gly to Ser mutations within an integrin-binding site caused no major structural perturbations, mutations within the Fn-binding site caused the triple helix to become highly sensitive to trypsin digestion. This trypsin susceptibility is consistent with the significant local unfolding and loss of hydrogen bonding seen in molecular dynamics simulations. Protease sensitivity resulting from mutations in the Fn-binding sequence could lead to degradation of type I collagen, early embryonic lethality, and the scarcity of reported osteogenesis imperfecta mutations in this region.

Collagen is the major protein in the extracellular matrix (ECM), providing physical support to cells and appropriate mechanical properties to tissues. Collagens are distinguished by their triple helix structure, composed of three polyproline II (PPII) helices supercoiled about a common axis (1, 2). The tight packing of the three chains generates a requirement for Gly as every third residue, giving the characteristic collagen (Gly-Xaa-Yaa), repeating tripeptide sequence. The most abundant collagens form fibrils with a 67-nm axial periodicity, including the type I collagen fibrils found in bone and tendon and type II collagen fibrils present in cartilage. In addition to structural and mechanical roles, collagen mediates cell signaling by binding to integrin and other cell receptors and participates in matrix structure and remodeling through interactions with other ECM proteins (3). One important interaction is the direct binding of collagen to fibronectin (Fn), another dominant ECM protein. Collagen and Fn play critical roles in matrix structure and cell signaling, and both are known to influence ECM self-assembly and function. Fn binds very strongly to denatured collagen (gelatin) and less strongly to native collagen (4). The 6-triplet sequence within collagen that binds Fn (designated here as the fibronectin-binding sequence (FBS)) has been identified, using collagen CNBr fragments, small single-stranded (Gly-Xaa-Yaa), peptides, and recombinant collagens (5, 6), and the effect of Gly missense mutations within the FBS is reported here.

The large dimeric Fn glycoprotein is composed of a string of modular domains (FnI, FnII, and FnIII) (7), and its interaction with collagen/gelatin is mediated by the gelatin binding domain, which contains four FnI modules and two FnII molecules as follows: 6FnI–2FnII–9FnI (8). NMR spectroscopy indicated the individual Fn modules are sandwiches of two antiparallel β-sheets (9), and the high resolution co-crystal structure of a subfragment of the gelatin binding domain (8–9FnI) with a (Gly-Xaa-Yaa), collagen peptide showed the collagen Gly-Xaa-Yaa sequence adopted a β-strand conformation that extended β-sheets present in the FnI modules (5).

It has been proposed that mutations in collagen, which perturb its interactions with binding partners, may result in lethality or severe pathology (10, 11). There is evidence that a mutation within the FBS of type II collagen (R789C), observed in multiple cases of chondrodysplasias, increased collagen binding to Fn and led to intracellular aggregation and poor ECM assembly (12). Mutations within type I collagen are linked to the dominant form of osteogenesis imperfecta (OI), and the extensive database of known OI mutations within the human α1(I) collagen chain shows a scarcity of mutations at the FBS, which could suggest an embryonic lethal phenotype (10, 11). In this study, missense mutations are introduced at each Gly site within the FBS to investigate their effect on the triple helix.
structure and on the binding of Fn to native and denatured collagen, using both a recombinant collagen system and model peptides.

A recombinant bacterial collagen system with inserted sequences from the human α1 chain of type II indicated that Fn binding to both native and denatured collagen requires a minimum (Gly-Xaa-Yaa)₆ sequence as follows: α1(II) residues 775GLAGQRGIVGLPGQRGER792 (6). The 18-residue FBS in α1(II) is highly homologous to the sequence in the α1 chain of type I collagen (GJAGQRGVGGLPGQRGER), with only two very conservative changes. In this study, the consequences of Gly to Ser replacements within the FBS of type II collagen are investigated using fluorescence polarization studies of single-stranded collagen model peptides and solid-state binding assays on the recombinant bacterial protein containing the human FBS. These studies showed that Gly substitutions at four positions in the FBS decreased Fn binding to denatured collagen. Molecular dynamics simulations are consistent with the decreased binding being due to disruption of the interaction of the collagen β-strand with the Fn module β-sheet reported in the crystal structure (5). All mutations in the FBS caused the native triple helix of the recombinant collagen to become highly sensitive to trypsin, in agreement with molecular dynamics studies showing a locally unfolded triple helix with extended disruption of hydrogen bonding. The unfolding and trypsin sensitivity caused by mutations in the FBS contrasts with the trypsin resistance seen for similar missense mutations within the integrin-binding site of collagen, indicating the importance of the sequence context in determining the structural consequence of mutations.

Results

Gly Substitutions Affect Fibronectin Binding to Denatured Collagen

Fn is known to bind very tightly to denatured type I and type II collagen, and the collagen sequence and site involved in Fn binding has been determined (5, 6). The effect of Gly to Ser mutations on Fn binding to denatured collagen was explored through fluorescence polarization assays on single-stranded (Gly-Xaa-Yaa)₆ peptides and solid-state binding assays on denatured recombinant bacterial collagen.

Fn Binding to Model Peptides with Gly-Ser Mutations—Single-stranded peptides with type I collagen (Gly-Xaa-Yaa)₆ sequences have previously been used to study Fn binding to collagen (5, 13). Fluorescence polarization studies were carried out here on a control 18-mer containing the type II collagen (Gly-Xaa-Yaa)₆ 775GLAGQRGIVGLPGQRGER792 sequence (Fn6) (Fig. 1A). Seven homologous peptides were synthesized with Gly to Ser replacements at the 7 positions within and C-terminal to the FBS as follows: Gly-775, Gly-778, Gly-781, Gly-784, Gly-787, Gly-790, and Gly-793. When the mutations fell at the terminal Gly residue positions, the peptides were extended by two tripeptides to minimize end effects, giving (Gly-Xaa-Yaa)₆, a 24-residue peptide (Fn8) was used as a control for these longer peptides (Fig. 1A). CD spectroscopy on the control peptide and representative mutant peptides gave a PPII-like spectrum, with no cooperative thermal transition (Fig. 1B). Even though the peptides have Gly as every 3rd residue, they do not form triple-helical structures due to their short length and low imino acid content. A fluorescein label was attached to the N terminus of these peptides for fluorescence polarization studies (Fig. 1A).

Fluorescence polarization of the control Fn6 and Fn8 peptides with increasing amounts of Fn showed typical saturation profiles with Kᵣ values of 5.4 and 9.0 nm, respectively (Fig. 1C). These Kᵣ values are somewhat lower than the values of 13 nm reported for fluorescence polarization assays for the FITC-α1(I) collagen chain binding to Fn (8), suggesting the shorter peptides may bind Fn more tightly than longer single chain molecules. Fluorescence polarization was also carried out on the peptides with Gly to Ser replacements. Peptides with a Ser substitution at the positions corresponding to Gly-778, Gly-784, and Gly-790 showed a dramatic decrease in Fn binding; the fluorescence polarization did not reach saturation, and Kᵣ values could not be calculated. A small but significant decrease in binding was seen for substitutions at Gly-775 and Gly-787 (Fig. 1A). Peptides with Gly-781 and Gly-793 substitutions had binding similar to the control peptides.

Fn Binding to Denatured Recombinant Collagen with Gly-Ser Mutations—The other approach to study binding of Fn to denatured collagen utilizes a recombinant bacterial protein system based on the Scl2 protein from Streptococcus pyogenes, which includes an N-terminal trimerization domain (V domain) and a collagen-like (CL) domain consisting of 79 Gly-Xaa-Yaa triplets (designated VCL) (14, 15). The protein expressed in Escherichia coli lacks hydroxyproline (Hyp), because E. coli lacks prolyl 4-hydroxylase, but it forms a triple helix structure with stability similar to that of human fibrillar collagens (Tm, ~36–37 °C). The originally expressed bacterial VCL protein has no defined biological activity, but insertion of human type II collagen residues 775GLAGQRGIVGLPQQRGER792 (FBS) between two CL domains was shown to lead to Fn binding to native and denatured collagen (6). Because of experimental difficulties in introducing Gly mutations within the construct containing two tandem CL domains, the 6-triplet FBS was cloned here within a single triple-helical CL domain, to create the protein VCL-Fn (Fig. 2A). Homologous VCL-Fn proteins were generated with Gly to Ser mutations at seven different Gly sites as follows: G775S, G778S, G781S, G784S, G787S, G790S, and G793S.

Solid-state assays were done to determine the binding of Fn to the denatured recombinant proteins. The recombinant bacterial collagen VCL-Fn control (Fig. 2A) and homologous constructs with Gly-Ser replacements formed stable triple-helical structures. The proteins were denatured by heating to 60 °C for at least 15 min prior to coating onto the 96-well plates for the solid-state assays. Fn did not bind to the original VCL protein after denaturation, but it did bind to denatured VCL-Fn that contained the 6-triplet human type II FBS. VCL-Fn showed less binding compared with denatured type I and II collagens, consistent with previous observations (Fig. 2B) (6).

Binding of Fn to denatured recombinant collagen was measured over a range of Fn concentrations to obtain dose-response curves (Fig. 2D). At all Fn concentrations, three of the Gly to Ser mutations, G778S, G784S, and G790S, caused Fn
binding to decrease close to baseline levels, in good agreement
with the fluorescence data. The dose-response curve for the
mutation G787S did not reach saturation and showed a small
decrease in binding (Fig. 2, C and D), which was significant at all
Fn concentrations except the highest. The G781S mutation
showed a dose-response curve that is overall similar to the con-
trol VCL-Fn. A small but significant decrease in binding was
seen for several low Fn concentration points, but the solid state
observations are generally consistent with the fluorescence
results indicating G781S does not differ from the control. Fn
Effect of Gly Mutations in Collagen on Fibronectin Binding

binding appeared to be increased for mutations at the Gly residues flanking the central FBS, G775S and G793S (Fig. 2, C and D), an effect not seen by fluorescence studies on peptides with these mutations. Both methods suggest Gly-775 and Gly-793 are not essential to Fn binding. K_d values were not calculated because studies on solid-state assays suggest they are useful for qualitative internal comparisons but not for the derivation of meaningful K_d values (16).

Molecular Dynamics Simulations of Single Strand Collagen Peptide Bound to 8–9Fnl—Computational studies were carried out to clarify how Fn binding to denatured collagen might be affected by Gly to Ser replacements. The high resolution X-ray structure of a complex of two Fn modules 8–9Fnl within the gelatin binding domain bound to a 21-residue single strand collagen-like peptide showed a well defined interaction where five residues of the collagen peptide near the C terminus, 787GQRGE, form a β-sheet with hydrogen bonds in an antiparallel manner to the 8FnI β-sheet (Fig. 3) (5). Simulation modeling showed that the N-terminal portion of the FBS forms a second β-strand, at an angle to the first, and the 775GLAGQR sequence in this strand showed similar but less consistent interactions with the 8Fnl β-sheet (5).

A binding model of 775GLAGQRGIVGLPGQRGERGLP to 8–9Fnl was built based on the X-ray structure, extending the
Effect of Gly Mutations in Collagen on Fibronectin Binding

DECEMBER 30, 2016 • VOLUME 291 • NUMBER 53
JOURNAL OF BIOLOGICAL CHEMISTRY

known structure in the N-terminal direction by adding GLA. Molecular dynamics (MD) simulations at 400 K were performed for the wild type and for sequences with the Gly at each position replaced by Ser. The fractional times for the formation of hydrogen bonds and salt bridges between Fn and the collagen peptide were calculated for all seven Gly-Ser replacements. The data for the three substitutions that caused decreases in fractional hydrogen bonds or salt bridge are shown (G790S, G787S, and G778S). In addition, the replacement of Gly-Ser at position 784 led to increased distances for the hydrophobic interactions (\(8\text{FnI H549} \) and \(9\text{FnI F569} \) to collagen Leu-785) in two out of five simulations. The collagen peptide residues from X-ray crystal structure were colored in orange while the residues added were colored in green.

**FIGURE 3.** MD simulations of single strand collagen peptide bound to \(8\text{– 9FnI}\). The estimated fraction times for the formation of hydrogen bonds and salt bridges between Fn and the collagen peptide were calculated for all seven Gly-Ser replacements. The data for the three substitutions that caused decreases in fractional hydrogen bonds or salt bridge are shown (G790S, G787S, and G778S). In addition, the replacement of Gly-Ser at position 784 led to increased distances for the hydrophobic interactions (\(8\text{FnI H549} \) and \(9\text{FnI F569} \) to collagen Leu-785) in two out of five simulations. The collagen peptide residues from X-ray crystal structure were colored in orange while the residues added were colored in green.

Gly-784 falls in the region connecting the two \(\beta\)-strand segments. MD simulations showed that G784S did not lead to any loss of hydrogen bonding. Visual examination suggests the mutation G784S could disrupt hydrophobic interactions. In MD simulations, the distance of His-539 of \(8\text{FnI} \) and Phe-569 of \(9\text{FnI} \) to Leu-785 of the collagen peptide was monitored for G784S (Fig. 3); the disruption of this hydrophobic interaction was observed in some but not all simulations. It should be noted that the simulations were started from a configuration with the G784S collagen peptide pre-bound to the Fn modules, and this construction may render potential disruption of hydrophobic interactions difficult to observe in the simulation time scale of 50 ns.

Residue Gly-781 also falls in the region connecting the two \(\beta\)-strand segments, and no direct interactions are seen for the Gly-781 GIV triplet with the Fn module. MD simulations indicated the mutation G781S did not affect hydrogen bonding or hydrophobic interactions in the complex. This is in good agreement with the experimentally observed high level of Fn binding for peptides and denatured recombinant collagen with this mutation.

MD simulations predicted that G775S and G793S, the residues flanking the required sequence, would not affect Fn-col-
Effect of Gly Mutations in Collagen on Fibronectin Binding

Collagen interactions. This agrees reasonably well with the experimental data, which showed little change or a small increase in binding (Figs. 1, A and C, and 2, C and D).

Effect of Gly Substitutions in FBS on Native Collagen

Fn is known to bind specifically to native collagen, although the binding is less tight than to denatured collagen (4). The peptides described above are good models for binding to denatured collagen but did not form triple-helical structures. Attempts to create triple-helical model peptides containing the FBS by making longer peptides and capping the ends with stabilizing Gly-Pro-Hyp or Gly-Pro-Lys-Gly-Glu-Pro terminal sequences were unsuccessful (data not shown). Recombinant bacterial collagen with the FBS, VCL-Fn, did form a stable triple-helical structure, and recombinant constructs with mutations, as described above, were used to study how the mutations affected native collagen binding and the triple helix structure.

Fn did not bind to the original VCL protein but did show weak binding to VCL-Fn (Fig. 2B). The constructs with mutations within the internal Gly residues of the FBS showed a small decrease in binding to Fn. Dose-response curves did not yield reproducible saturation kinetics, and it was not possible to determine whether the small changes seen were significant. Although statistically significant results were not obtained from the binding studies, substantive effects of mutations on triple-helical conformation and stability were observed through biophysical studies on recombinant constructs.

Structure and Thermal Stability of Recombinant Collagens with Gly-Ser Mutations—The VCL-Fn construct, which has the insertion of the 6-triplet human type II collagen FBS within the CL domain, showed a CD spectrum similar to VCL, with a positive peak near 220 nm and a minimum just below 200 nm, indicating formation of a typical triple helix (Fig. 4A). As noted previously, the magnitude of the 220 nm maximum is smaller than seen for a standard collagen triple helix because of the opposing CD signal of the α-helical contribution of the V domain (15, 17, 18). The CD maximum near 220 nm was monitored as a function of temperature, and the first derivative of the transition was used to obtain Tm values for the mutant constructs. VCL has a sharp thermal transition with Tm = 36.1 °C. The thermal transition of VCL-Fn shows a small 1–2 °C decrease in stability (Tm = 34.6 °C), together with the appearance of a shoulder representing a small amount of a less stable species (Tm = 32 °C) (Fig. 4, B and C). The small decrease in Tm value is typical of that observed when other 6-triplet human sequences (e.g. integrin-binding sequence) have been inserted within the CL domain (19), but the appearance of a second lower stability transition has not been reported before and is related to the presence of the FBS.

The structural consequences of Gly to Ser mutations in the inserted human FBS were evaluated. CD spectra of all VCL-Fn proteins with Gly-Ser mutations showed collagen-like features, indicating the overall triple helix structure is maintained (see Fig. 4A for a representative example). All constructs with Gly-Ser mutations had biphasic melting profiles, with two distinct peaks (Fig. 4, B and C). The melting profiles of G775S, G778S, G781S, and G784S were similar to that seen for VCL-Fn, with the large majority of the triple helix melting near 34.5 °C and a small amount near 31.5 °C. There was an increased amount of the less stable species in G787S, G790S, and G793S. Estimates of the fraction of the less stable species were ~10% for VCL-Fn; 20–25% for the mutations at Gly-775, Gly-778, Gly-781, and Gly-784; and 35–40%, for sites Gly-787, Gly-790, and Gly-793. The less stable species could represent a partially unfolded species in equilibrium with the fully folded triple helix, or it could be due to a partially unfolded domain within the triple helix that melts independently. It is also possible that the lower thermal transition is due to partially degraded molecules suggested by several faster moving bands in the SDS-PAGE (0 time point in Fig. 5B), but these faster moving bands are not seen in all preparations, and the small amounts present cannot explain a 20–40% magnitude of the shoulder.

Differential scanning calorimetry (DSC) of VCL showed a sharp thermal transition at 37.0 °C; VCL-Fn showed a sharp transition at a slightly lower value, 35.5 °C, together with a small lower temperature shoulder (Fig. 4D). Two representative constructs with mutations, G781S and G790S, had slightly lower DSC Tm values, 35.0 and 35.1 °C respectively, with a more pronounced lower stability shoulder. Decreasing the DSC heating rate improved the separation and definition of the lower stability shoulder, consistent with an equilibrium between a less stable species and the standard triple helix. The calorimetric enthalpy was 3,269 kJ/mol for VCL-Fn, whereas mutations resulted in small decreases in calorimetric enthalpy (2,827 kJ/mol for G781S and 2,813 kJ/mol for G790S).

The biophysical results for recombinant constructs with the FBS are compared with a homologous recombinant collagen construct, VCL-Int, which contains six triplets from human collagen (residues 496–513 from the α1 chain of type I collagen, GARGERGFPGERGVQGGP) around the GFPGER integrin-binding site (19). The introduction of the integrin binding region had almost no effect on the single sharp thermal transition near 37 °C (19), whereas the introduction of the FBS within VCL led to a drop in stability (~2 °C) and the appearance of a lower stability species. Gly to Ser mutations within both the FBS and integrin binding domains led to small decreases in thermal stability (1–2 °C), but only the FBS mutations led to increased amounts of a second lower stability thermal transition.

Trypsin Digestion of Recombinant Collagens with Gly-Ser Mutations—The standard triple-helical conformation is known to be resistant to trypsin, so trypsin digestion was used as a probe for perturbations to the native collagen structure (20). The VCL-Fn control, with the FBS insertion, and the VCL-Int control, with the integrin binding insertion, were not susceptible to a 2-min trypsin digestion at 20 °C, the conditions typically used to test for native triple helix in animal collagens (Fig. 5, A and B) (19, 20). Introduction of Gly to Ser replacements within the inserted integrin-binding sequence in VCL-Int did not affect trypsin resistance. All VCL-Int constructs with mutations were not susceptible to trypsin digestion at 20 °C even after a 1-h digestion time (Fig. 5A). In contrast, all VCL-Fn recombinant proteins with Gly replacements within or flanking the FBS (G775S, G778S, G781S, G784S, G787S, G790S, and G793S) were highly susceptible to trypsin cleavage after a 2-min digestion. After trypsin treatment, the recombinant constructs with mutations showed two distinct bands on SDS-PAGE (Fig.
MALDI-TOF identified the trypsin-sensitive sites as one of the three Arg within the FBS, depending on the location of the mutation. For example, mass spectroscopy of the construct with the G778S mutation has a mass of 35,301 Da, and after trypsin cleavage, there are major peaks at 19,673 and 15,615 Da, supporting a cleavage at the first Arg-780 within the FBS (Fig. 5C).

Molecular Dynamics Simulations of VCL-Fn and with Gly-Ser Mutations—Observations that Gly to Ser mutations in the inserted FBS led to biphasic melting curves and to trypsin

**FIGURE 4.** Biophysical characterization of the recombinant protein VCL with no FBS, VCL-Fn with the 6-triplet FBS insertion, and a representative example of a mutation in the FBS (G790S). A, CD spectra of VCL (gray), VCL-Fn (black), and G790S (purple), showing the maximum at 220 nm and minimum near 198 nm, which are characteristic of the triple helix. B, CD melting profiles of VCL, VCL-Fn, and G790S at 220 nm. C, first derivative of the CD melting curves for VCL (top), VCL-Fn (middle), and G790S (bottom), showing a distinct minor transition of lower stability for G790S. D, DSC profiles of VCL (top), VCL-Fn (middle), and G790S (bottom).
susceptibility suggest conformational changes to the triple helix within this region. MD simulations at 300 K were performed on triple-helical structures of both VCL-Fn and VCL-Int with mutations to further explore such conformational changes (Fig. 6). The simulations were carried out initially on a control triple helix peptide modeling VCL-Fn,
which contained the 6-triplet FBS flanked by five N-terminal and three C-terminal triplets from the bacterial collagen sequence, and capped by (GPO)₃ stabilizing triplets on both ends. The triple helix showed standard interchain hydrogen bonding at 300 K (Fig. 6A, top).

To model two representative mutant constructs, MD simulations were then performed at the same temperature on a homologous sequence with a Gly to Ser replacement at position Gly-781 or Gly-787 (Fig. 6A). The fractional times for the formation of interchain hydrogen bonds within the triple helix were calculated, and a hydrogen bond is considered to be disrupted if its fractional time falls below 0.5 during the 100-ns simulation. For G781S, the mutation led to disruption of 4–5 hydrogen bonds between each pair of chains and local unfolding of 10 N-terminal residues of the human collagen insertion GLAGQRSIVG, as well as the N-terminal adjacent bacterial collagen triplet GLP (Fig. 6A, middle). The introduction of the G787S mutation caused disruption of a wider region of the tri-
ple helix involving disruption of 6–7 hydrogen bonds between chain pairs and loosening of the entire 6 triplet human FBS insertion GLAGQRGIVGLPQRSTGER, in addition to the N-terminal GLP from the bacterial collagen (Fig. 6A, bottom). Even the smaller disruption calculated for G781S appears to be sufficient to induce the experimentally observed trypsin susceptibility. The larger region of disruption for G787S compared with the G781S construct is consistent with experimental CD melting results showing a larger amount of the lower stability species in the G787S construct compared with the G781S construct. In some cases, a small amount of β-strand is seen within the unfolded triple helix, which could be related to the interaction with the Fn β-sheet.

Similar studies were carried out to model the homologous peptide with the integrin-binding sequence GARGFGFRGGER. The control peptide formed a well packed triple-helical structure with standard hydrogen bonding. Mutations at both positions homologous to those studied in the FBS (G502S, G508S) led to a much more localized disruption of hydrogen bonding. For mutations G502S and G508S, 1–3 hydrogen bonds were disrupted between each pair of chains (Fig. 6B). The much larger number of hydrogen bonds seen to be disrupted in the FBS compared with the integrin-binding sequence could explain the trypsin sensitivity of the Fn-binding mutant constructs compared with the trypsin resistance of the integrin-binding mutant constructs.

Discussion

The interaction of collagen with cell receptors, extracellular matrix molecules, and enzymes is critical for its biological function. The effect of mutations on collagen interactions is used here to clarify the binding mechanism as well as to understand how altered interactions can lead to pathology. Most proteins that interact with collagen, such as integrins and SPARC, have a stringent requirement for a triple-helical collagen structure, and the structural basis for this requirement can be seen in the high resolution structure of co-crystals of the integrin I domain or SPARC with a triple helix peptide (21, 22). The ability of Fn to bind very strongly to denatured collagen and to also recognize native collagen raises questions about the mechanism and specificity of this interaction.

Fn Binding to Denatured Collagen with Mutations—The effects of Gly missense mutations on the collagen triple helix folding, stability, and secretion have been reported (11, 23–25), but this study presents the first example where mutations affect the recognition of unfolded or denatured collagen. Our experimental studies demonstrated that the change from a Gly to Ser at four internal sites within the (Gly-Xaa-Yaa)_n FBS interfered with Fn binding to single-stranded collagen. The experimentally observed decrease in Fn binding correlates well with the disruption of the β-sheet interaction between Fn and a collagen β-strand seen by MD simulations for three of the sites (Gly-778, Gly-787, and Gly-790), whereas the fourth with a decreased binding upon Ser mutation (Gly-784) may result from reduced hydrophobic interactions. Thus, the decrease in Fn binding due to Gly missense mutations within the collagen FBS is a result of interference with the recognition of the 4–5 FnI modules by the β-strand collagen peptide.

CD spectra of single strand collagen peptides containing the six triplets known to be required for Fn binding are characteristic of polyproline II-like chains, consistent with previous reports (5). There is only one Pro in the FBS sequence. Some residues in the 18-mer (Gln and Ala) are known to have a high PPII propensity, whereas several residues in the middle of the peptide (Val and Ile) have a very low propensity for PPII and a high propensity for β-strand (26–28). It is possible that contact with Fn may induce single chain PPII-like segments to undergo a transition to β-strand, because both are very extended conformations (0.31 nm rise/residue for PPII; 0.34 nm rise/residue for β).

Effect of FBS Mutations on Native Triple-helical Collagen—Erat et al. (13) propose that Fn binding to native triple-helical collagen requires local unfolding of the triple helix, so that a single Gly-Xaa-Yaa chain can form a β-strand that will interact with the β-sheet Fn modules (5). Such a mechanism is supported by the requirement for the same (Gly-Xaa-Yaa)₆ sequence for Fn binding to both native and denatured collagen (6) and is consistent with the weaker binding of Fn to the native triple helix compared with denatured collagen (8). The studies reported here suggest that introduction of the Fn binding region within the bacterial collagen domain leads to a less stable molecular species in equilibrium with the standard triple helix. The lower stability triple-helical molecule could have a locally unfolded conformation within the FBS, which could bind to Fn through the same mechanism seen for single-stranded collagen chains. The introduction of mutations within this region leads to a larger proportion of the less stable species and to trypsin sensitivity within the FBS. MD simulations support a very significant loss of hydrogen bonding in the mutant molecules, which is in agreement with the trypsin sensitivity and the lower calorimetric enthalpy observed for the mutant collagens. These observations are consistent with previous computational studies on the collagenase cleavage site, which overlaps with the Fn-binding sequence (G↓LAQQRIVGGLPQRSTGER in type II collagen), showing the typical triple helix conformation is in equilibrium with a locally loosened vulnerable conformation (29).

Overall, these studies support the presence of a loose or locally unfolded chain in the collagen triple helix that binds to Fn through the same mechanism seen for single-stranded collagen chains. Because Fn binding to collagen appears to directly involve only two Fn modules out of the six modules present in the Fn gelatin binding domain, it is possible that some of the modules play an unwinding role similar to that seen for the hemopexin domain in collagenase (30). Given that mutations within the FBS loosen the triple helix, it might be expected that such mutations would lead to a significant increase in Fn binding to native collagen. This was not observed, which is likely due to the interference of the mutations with the collagen β-strand and Fn β-sheet interactions.

These studies demonstrate the effect of Gly missense mutations depends strongly on the (Gly-Xaa-Yaa)₆ sequence context. Although mutations within the integrin binding domain (19) had little effect on the triple helix conformation and stability, a mutation within the fibronectin binding region unwound the local triple helix sufficiently to make it trypsin-sensitive.
Collagen Pathology and Mutations in the Fn Binding Region—It is well established that the replacement of even one Gly in the (Gly-Xaa-Yaa)$_n$ repeating sequence of the collagen triple helix by a larger residue leads to pathology, such as missense mutations in type I collagen that lead to the dominant form of the fragile bone disease OI (10). Although errors in processing, secretion, and degradation are considered as critical steps in OI and other collagen diseases, it has also been suggested that Gly missense mutations within ligand binding domains of collagen may cause pathology by interfering with natural collagen interactions (10, 31). The studies here report the consequences of Gly to Ser mutations within the FBS of human type II collagen, but the high homology between the human α1(I) and α1(II) chains suggests the results can also be used to consider OI mutations in type I collagen.

The recombinant triple-helical molecules studied here are well defined homotrimers, with a Gly to Ser replacement in each of the three chains. In collagen diseases, such as dominant OI with a mutation in the α1(I) chain, the collagen population includes a heterogeneous mixture of triple-helical molecules with 0 (~25%), 1 (~50%), or 2 (~25%) mutant chains. There is experimental evidence from model triple-helical peptides (32) and computational studies (33) that introduction of a Gly replacement into the first chain has the most destabilizing and perturbing effect, mutating the second chain has a much smaller effect, and introduction of the mutation in the third chain has the least influence. Consistent with this earlier work, preliminary molecular dynamics studies indicate that at certain sites within the Fn-binding sequence, a Gly to Ser mutation introduced within only one chain of the triple helix will result in a substantial disruption of interchain hydrogen bonding across 3–4 tripeptide units. Therefore, it is likely that the trypsin susceptibility seen for recombinant homotrimeric molecules with mutations are applicable to heterotrimeric molecules containing only one or two mutant chains.

Mutations at four sites within the FBS led to trypsin sensitivity due to local triple helix unwinding. The FBS of the human α1(I) collagen chain falls in a region known to have a low density of OI missense mutations, with an unusually high proportion of lethal mutations among the small number seen (10, 11). Examination of the OI database shows there are seven triplets devoid of known OI mutations, except for a non-lethal G784D case whose details have not been reported (10) (Fig. 7). Surrounding this largely “silent zone” are mutations at Gly sites in the three tripleptides immediately N-terminal to the FBS (G766C, G769C, and G772R) and in the three tripleptides that are one triplet C-terminal to the FBS (G796D/G796A, G799D, and G802V). The scarcity of mutations in the important FBS region is consistent with non-viability of an embryo with this mutation. As proposed previously (10, 31), mutations within the FBS could prevent normal development due to the decreased ability of the mutant collagen to interact with Fn. Our results suggest a new potential mechanism; collagens with mutations in the FBS may be susceptible to common proteases, leading to the breakdown of type I collagen. Previous studies have shown that the loss of the α1(I) chain in transgenic mice resulted in early embryonic lethality (34), and degradation of type I collagen would be expected to have a similar non-viable phenotype. We suggest that proteolytic sensitivity of native collagen could represent a novel mechanism of pathology when a Gly missense mutation occurs in the Fn binding region.

### Experimental Procedures

**Fluorescence Polarization Assay**—7.5 nM FITC-conjugated peptides were titrated with nanomolar quantities of Fn in PBS buffer with 0.05% Tween 20, pH 7.4, in a 96-well opaque non-adsorbent plate. The plate was incubated for 30 min in the dark at room temperature and read using a SpectraMax Multi-Mode Detection Platform. The excitation and emission wavelengths used were 485 and 535 nm, respectively. The data were baseline-corrected by subtracting the minimum fluorescence polarization from each data point. Each concentration of Fn was performed in triplicate. Binding affinity ($K_d$) was determined by fitting the data to a binding curve using Sigma Plot.

**Molecular Cloning**—The VCL-Fn protein sequence for the recombinant bacterial collagen was adapted from a previous

---

**FIGURE 7.** Top, known OI mutations in the Fn binding region, as reported in the OI database (53, 54). The residue replacing Gly is shown below, together with the sillecence classification of the phenotypic severity of the OI (OI-II, perinatal lethal; OI-III, severe; OI-IV, moderate; OI-I, mild) (10). In the sequence, the lethal OI II mutations are colored red, and the other non-lethal phenotypes are colored blue. Middle, experimental effect of Gly to Ser mutations on Fn binding as determined for peptides by fluorescence polarization and for denatured recombinant collagens by solid-state assays. Red bars (with +) indicate decreased binding; green bars (with +) indicate increased binding, and gray bars indicate no change from the wild-type sequence. Bottom, constructs with mutations in the FBS are highly susceptible to trypsin digestion, as indicated.
construct SCI-FnGly-775–Arg-792 that was shown to bind Fn; however, instead of cloning the FBS between two CL domains, the FBS is in the middle of a single CL domain unit for the VCL-Fn (6). The FBS GLAGQRGIVGLPGQRGER was inserted after triplet number 31 through restriction sites ApaI and Xmal. The VCL-Fn vector was cloned into a pCOLDIII vector (Takara Bio Inc.) through Ndel and Xbal restriction sites. A QuikChange II site-directed mutagenesis kit by Agilent Technologies was used to make Gly-Ser substitutions in the recombinant bacterial collagen. DNA sequencing to confirm mutants was done by Genewiz.

Protein Purification—All recombinant bacterial collagens were cloned in the pCOLDIII vector and expressed in E. coli BL21 strain, grown in 30 ml of LB medium with 100 μg/ml ampicillin overnight at 37 °C (18). The starting culture was used to inoculate 500 ml of LB/ampicillin media in a shaking flask and grown at 37 °C to an A600 = 0.8–1.0. To induce protein expression, 1 mM isopropyl β-D-thiogalactopyranoside was added to the culture, and the temperature was lowered to 22 °C. After a 16-h induction, cells were harvested by centrifugation and resuspended in His tag purification column binding buffer (20 mM sodium phosphate buffer, pH 7.4, 500 mM NaCl, 20 mM NaP, 10 mM imidazole) and frozen at −20 °C until purification.

Purification of His-tagged recombinant bacterial collagens was carried out using an AKTA Pure FPLC. Frozen cells were thawed and further lysed by sonication. The cell lysate was centrifuged, and the supernatant containing the recombinant protein was loaded onto an equilibrated HisTrap™ HP 5-ml nickel-nitrilotriacetic acid column. The column was washed with 3 column volumes of binding buffer, 3 column volumes of 50 mM imidazole, and 3 column volumes of 100 mM imidazole. The recombinant protein was eluted with 500 mM imidazole, as described previously (6).

Protein purity was checked by SDS-PAGE (NuPAGE® Bis-Tris 4–12%, Invitrogen). The concentration was determined using an extinction coefficient of ε280 = 9970 M⁻¹ cm⁻¹ after dialysis into phosphate-buffered saline (PBS, pH 7.4).

Circular Dichroism—CD spectra were obtained as described previously (35). Spectra were measured on an AVIV model 420 CD spectrometer (AVIV Biomedical, Lakewood, NJ) using glass cuvettes with a 1-mm path length. Protein solutions were equilibrated for at least 24 h at 4 °C before measurement. Wavelength scans were collected from 190 to 260 nm in 0.5-mm steps with a 4-s averaging time and a 1-nm bandwidth. Temperature scans were monitored from 0 to 60 °C with a 10-s averaging time and a 1.5-nm bandwidth. Samples were equilibrated for 2 min at each temperature, and the temperature was increased at an average rate of 0.1 °C/min, and the CD signal at 220 nm was monitored (as described previously (35)).

Differential Scanning Calorimetry—DSC was performed as described previously (35). DSC was measured on a NANO DSC II model 6100 (Calorimetry Sciences Corp, Lindon, UT). Each sample was re-dialedyzed against PBS overnight before measurement, and the dialyzed buffer was used as reference in the experiment. Sample solutions were loaded at 0 °C into the cell and heated at a rate of 1 °C/min until 60 °C as described previously (35).

Trypsin Cleavage Assay—Trypsin digestion was carried out as follows. 10 μg (0.16 mg/ml) of protein substrate in PBS buffer was incubated to a final trypsin concentration of 10 μg/ml for 2 min at 20 °C. The reaction was stopped by adding 1 mM phenylmethylsulfonyl fluoride (PMSF). Cleavage was visualized with SDS-PAGE, and the sites of cleavage were determined by mass spectroscopy.

MALDI-Mass Spectrometry—Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra were acquired on a Microflex LT system (Bruker Corp., Billerica, MA) with 70% laser intensity using a linear positive 60-kDa method provided by the software. MALDI matrix was prepared by saturating sinapinic acid in solutions containing 50% (v/v) acetonitrile and 0.3% (v/v) trifluoroacetic acid. A 6-μl aliquot of protein substrate was mixed with 24 μl of matrix. 1 μl of solution was plated onto a 96-spot target plate and allowed to dry.

Solid-state Fn Binding Assay—Fn binding affinity for recombinant bacterial collagen constructs was measured in both native and denatured states, as described previously (6). Prior to collagen immobilization, denatured samples were heated to 60 °C for at least 15 min, and native samples were equilibrated to 4 °C overnight. 50 μl of a 100 μg/ml solution of protein substrate was then coated overnight onto Immulon 2HB plates at 4 °C. Plates were washed with PBS and subsequently blocked with 1 mg/ml BSA in PBS for 1 h at room temperature. 50 μl of 20 μg/ml human plasma Fn was then incubated at room temperature with immobilized collagen. An anti-Fn C-terminal antibody was used to show the presence of Fn in the ELISA plates. A secondary antibody anti-mouse HRP was used for detection with 3,3',5,5'-tetramethyldibenzine, followed by neutralization with 3 M HCl and measurement at A490.

Molecular Dynamics Simulations of Triple-Helical Collagen Peptides—For triple-helical collagen peptides, six model triple-helical collagen peptides were simulated in this study. The wild-type (WT) collagen peptide containing the FBS sequence, GPOGPQPOGPQGPRGEQQPGQGLAGQRGIVGLPGQERGERGLPGKDGEAGPGQPGQPOGQ, and its G781S and G787S mutants; the WT collagen containing the integrin-binding sequence, GPOGPQPOGPQGPRQEQQPGQGARGERPFPGERVGQPPGLPGKDGEAGPGQPGPQPOGQ, and its G502S and G508S mutants. The initial structure for the WT triple-helical collagens were built using the triple-helical collagen building script (THeBuScr) (36). The initial structure for the G781S/G787S or G502S/G508S mutant collagens were created using the GROMACS 4.6.7 suite (38) with the GROMOS54A7 force field (39) and SPC water model (40). For each collagen triple helix, the starting structure was first subjected to 2,000 steps energy minimization. The minimized structure was then used to perform all-atom molecular dynamics simulations. The simulations were performed using the Amber99SB force field (50, 51) and tip3p water model (52) at 300 K using NAMD (53).

It should be noted that the simulations using the Amber99SB force field (50, 51) and tip3p water model (52) at 400 K give similar results.
Effect of Gly Mutations in Collagen on Fibronectin Binding

then solvated in a rectangular periodic box of water. The dimension of the box was chosen such that the minimum distance between any atom of the collagen and the box walls is 1.5 nm. The solvated system was further energy-minimized for 5,000 steps. With all heavy atoms of the collagen harmonically restrained to their initial positions, each system was heated from 5 K to the target temperature (300 K) within 20 ps and equilibrated for an additional 30 ps to relax the solvent molecules and adjust the box size. Before production, an additional 500-ps equilibration, with only the collagen backbone restrained, was also performed to relax the collagen side chains.

The 100-ns production runs were performed in the isobaric-isothermal (NPT) ensemble at 300 K and 1.0 bar. The first and last Cα atoms of each chain were restrained with a harmonic potential of 10 kJ/mol/nm². The temperature was regulated using the Nosé-Hoover thermostat with a coupling time constant of 1.0 ps (41, 42). To avoid the “hot solvent-cold solute” problem (43, 44), the protein and solvent molecules were coupled to separate thermostats. The pressure was controlled using an isotropic Berendsen barostat with a coupling time of 2.0 ps and a compressibility of 4.5 × 10⁻⁵ bar⁻¹ (45). All bonds were constrained with the LINCS algorithm (46). A 2-fs time step was used to evolve the dynamics. The non-bonded interactions (Lennard-Jones and short-range electrostatic) were truncated at 0.8 nm. Long-range electrostatic interactions beyond the cutoff distance were calculated using the Particle Mesh Ewald (PME) summation method with a Fourier spacing of 0.12 nm and an interpolation order of four (47). A long range analytic dispersion correction was applied to both energy and pressure to account for the truncation of Lennard-Jones interactions (48).

Molecular Dynamics Simulations of Single Strand Collagen Peptide Bound to Fn—The models of single strand collagen peptide bound to 8–9FnI were built from the X-ray crystal structure of Protein Data Bank code 3EJH (chains B and F) (5). All solvent molecules were removed from the X-ray structure. The collagen peptide (chain F) was used as a template to generate the wild-type (WT) sequence of GLAGQRGIVGLPGQRGER-Glycam06 parameters (49) for the N-acetylgalacosamine (GlcNAc) residue was used in the Fn binding simulations. For each system (Fn binding with WT/G778S/G781S/G784S/G787S/G790S/G793S single strand collagen peptide), the initial structure was first solvated in a cubic water box of size 9.1 × 9.1 × 9.1 nm³ with three sodium ions added to neutralize the net charge of the system. The solvated system was further energy-minimized for 5,000 steps to remove bad contacts. With all heavy atoms of Fn and collagen harmonically restrained to their initial positions, each system was heated from 5 to 400 K within 20 ps and equilibrated for an additional 130 ps to relax the solvent molecules and adjust the box size. The final 50-ns production runs were performed using the same protocol as described under “Molecular Dynamics Simulations of Triple-Helical Collagen Peptides.”

Author Contributions—P. C., H. Y., Y. S. L., and B. B. designed the study and wrote the paper. P. C., H. Y., B. A., Y. S. L., and B. B. analyzed and critically evaluated the results. P. C. performed experiments with recombinant bacterial collagen, peptide fluorescence polarization, and biophysical experiments. H. Y. performed the MD simulations. B. A. participated in experimental design and supervised execution of experiments. B. R. D. developed the fluorescence polarization assay and did initial fluorescence experiments and early circular dichroism studies. All authors read and approved the final version of the manuscript.

Acknowledgments—We thank Dr. David Wilbur from Tufts University Chemistry for the use of the MALDI-TOF mass spectrometer. We thank Sezin Yigit for helpful discussions and generous experimental support.

References
1. Brodsky, B., and Persikov, A. V. (2005) Molecular structure of the collagen triple helix. Adv. Protein Chem. 70, 301–339
2. Shoulders, M. D., and Raines, R. T. (2009) Collagen structure and stability. Annu. Rev. Biochem. 78, 929–958
3. An, B., Lin, Y.-S., and Brodsky, B. (2016) Collagen interactions: Drug design and delivery. Adv. Drug Deliv. Rev. 97, 69–84
4. Dessau, W., Adelmann, B. C., and Timpl, R. (1978) Identification of the sites in collagen α-chains that bind serum anti-gelatin factor (cold-insoluble globulin). Biochem. J. 169, 55–59
5. Erat, M. C., Slater, D. A., Lowe, E. D., Millard, C. J., Farndale, R. W., Campbell, I. D., and Vakonakis, I. (2009) Identification and structural analysis of type I collagen sites in complex with fibronectin fragments. Proc. Natl. Acad. Sci. U.S.A. 106, 4195–4200
6. An, B., Abbonante, V., Yigit, S., Balduini, A., Kaplan, D. L., and Brodsky, B. (2014) Definition of the native and denatured type II collagen binding site for fibronectin using a recombinant collagen system. J. Biol. Chem. 289, 4941–4951
7. Schwarzbauer, J. E., and DeSimone, D. W. (2011) Fibronectins, their fibrillogenesis, and in vivo functions. Cold Spring Harb. Perspect. Biol. 3, a005041
8. Ingham, K. C., Brew, S. A., and Isaacs, B. S. (1988) Interaction of fibronectin and its gelatin-binding domains with fluorescent-labeled chains of type I collagen. J. Biol. Chem. 263, 4624–4628
9. Pickford, A. R., and Campbell, I. D. (2004) NMR studies of modular protein structures and their interactions. Chem. Rev. 104, 3557–3566
10. Marini, J. C., Forlino, A., Cabral, W. A., Barnes, A. M., San Antonio, J. D., Milgrom, S., and P., Symoens, S., Glorieux, F. H., Roughley, P. J., Lund, A. M., et al. (2007) Consortium for osteogenesis imperfecta mutations in the helical domain
Effect of Gly Mutations in Collagen on Fibronectin Binding

Candidate cell and matrix interaction domains on the collagen fibril, the predominant protein of vertebrates. *J. Biol. Chem.* **283**, 21187–21197

Gauva, V., and Hartgerink, J. D. (2008) Synthetic collagen heterotrimers: structural mimics of wild-type and mutant collagen type I. *J. Am. Chem. Soc.* **130**, 7509–7515

Mooney, S. D., Huang, C. C., Kollman, P. A., and Klein, T. E. (2001) Computed free energy differences between point mutations in a collagen-like peptide. *Biopolymers* **58**, 347–353

Schneike, A., Harbers, K., and Jaenisch, R. (1983) Embryonic lethal mutation in mice induced by retrovirus insertion into the α(I) collagen gene. *Nature* **304**, 315–320

Persikov, A. V., Xu, Y., and Brodsky, B. (2004) Equilibrium thermal transitions of collagen model peptides. *Proteins Sci.** **13**, 893–902

Rainey, J. K., and Goh, M. C. (2004) An interactive triple-helical collagen builder. *Bioinformatics* **20**, 2458–2459

Petersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612

Hess, B., Kutzner, C., van derSpoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435–447

Schmid, N., Eichenberger, A. P., Choutko, A., Riniker, S., Winger, M., Mark, A. E., and van Gunsteren, W. F. (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur. Biophys. J.* **40**, 843–856

Berendsen, H. J., Postma, J. P., van Gunsteren, W. F., and Hermans, J. (1981) in *Intermolecular Forces* (Pullman, B., ed) pp. 331–342, Springer, Dordrecht, Netherlands

Nosé, S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268

Hoover, W. G. (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697

Cheng, A., and Merz, J., and Kenneth M. (1996), Application of the Nosé-Hoover chain algorithm to the study of protein dynamics. *J. Phys. Chem.* **100**, 1927–1937

Lingenheil, M., Denschlag, R., Reichold, R., and Tavan, P. (2008) The “Hot-Solvent/Cold-Solute” problem revisited. *J. Chem. Theory Comput.* **4**, 1293–1306

Berendsen, H. J., Postma, J. P., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690

Hess, B., Bekker, H., Berendsen, H. J., and Fraaije, J. G. (1997) LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593

Allen, M. P., and Tildesley, D. J. (1987) *Computer Simulation of Liquids*, pp. 64–65, Clarendon Press, Oxford, UK

Kirschner, K. N., Yongye, A. B., Tschemp, S. M., González-Outeiriño, I., Danilo, C. R., Foley, B. L., and Woods, R. I. (2008) GLYCAM06: A generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* **29**, 622–655

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725

Park, S., Radmer, R. J., Klein, T. E., and Pande, V. S. (2005) A new set of molecular mechanics parameters for hydroxyproline and its use in molecular dynamics simulations of collagen-like peptides. *J. Comput. Chem.* **26**, 1612–1616

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935

Dalgleish, R. (1997) The human type I collagen mutation database. *Nucleic Acids Res.* **25**, 181–187

Dalgleish, R. (1998) The human collagen mutation database 1998. *Nucleic Acids Res.* **26**, 253–255