Sugars and Polyols Inhibit Fibrillogenesis of Type I Collagen by Disrupting Hydrogen-Bonded Water Bridges between the Helices

N. Kuznetsova,[‡,§] S. L. Chi,[||] and S. Leikin*[,‡]

Laboratory of Physical and Structural Biology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, and Princeton University, Princeton, New Jersey 08544

Received January 12, 1998; Revised Manuscript Received April 28, 1998

ABSTRACT: To better understand the mechanism of collagen fibrillogenesis, we studied how various sugars and polyols affect the formation and stability of collagen fibers. We combined traditional fiber assembly assays with direct measurement of the interaction between collagen triple helices in fibers by osmotic stress and X-ray diffraction. We found that the effects of sugars and polyols were highly specific with respect to small structural differences between these solutes. For example, 1,2-propane diol only weakly inhibited the fiber assembly and practically did not affect the interaction between collagen helices in fibers. At the same concentration, 1,3-propane diol eliminated the attraction between collagen helices and strongly suppressed fibrillogenesis. The two diols have the same atomic composition and differ only by the position of one of their hydroxyls. Still, their ability to inhibit fiber assembly differs by more than an order of magnitude, as judged by protein solubility. We argue that this is because collagen fibrillogenesis requires formation of hydrogen-bonded water clusters bridging recognition sites on the opposing helices. The ability of various sugars and polyols to inhibit the fiber assembly and to destabilize existing fibers is determined by how efficiently these solutes can compete with water for crucial hydrogen bonds and, thus, disrupt the water bridges. The effect of a sugar or a polyol appears to be strongly dependent on the specific stereochemistry of the solute hydroxyls that defines the preferred hydrogen-bonding pattern of the solute.

Type I collagen is a triple helical protein which consists of two identical α1(I) polypeptide chains and one α2(I) chain. It is a major structural protein composing the extracellular matrix of many tissues, for example, tendon, bone, and skin. Abnormal collagen−collagen interactions are responsible for severe pathology of human development (osteogenesis imperfecta) and for severe disease complications (e.g., connective tissue failure in diabetes). A comprehensive bibliography of collagen research was recently reviewed in refs *1* and *2*.

At physiological ionic strength and neutral pH, increasing temperature induces spontaneous assembly of type I collagen into nativelike fibers. Extensive *in vitro* studies produced a detailed phenomenological picture of this process (see, e.g., refs *1* and *3* and references therein), but the underlying mechanism remained unknown. Understanding this mechanism is not only important for numerous biomedical applications but it is also likely to provide useful insights for solving the more general problems of protein folding, recognition, and assembly. Although collagen has an unusual amino acid composition and a unique structure, the interactions responsible for its assembly may be common.

One of the unknowns is why sugars and polyols inhibit collagen fibrillogenesis (*4*−*6*) and destabilize assembled fibers (*7*). This contrasts with the observation that the same solutes protect most proteins (*8*−*12*), including the collagen triple helix (*13*, *14*), against thermal denaturation. Such unusual sugar and polyol action may hold a key to understanding the molecular mechanism of fibrillogenesis.

We previously approached this problem by directly measuring forces between collagen helices in reconstituted and native fibers using osmotic stress and X-ray diffraction (*15*, *16*). The measurement in fibers solvated by pure water revealed a balance between a short-range, exponential repulsion and a longer-range attraction. The repulsion prevents the molecules from coming too close to each other, and the attraction keeps them from coming too far apart. We found that addition of some sugars, polyols, and alcohols strongly affects this balance of forces. For example, glycerol and glucose inhibit the attraction, ethylene glycol has practically no effect, and ethanol weakly enhances the attraction.

The observed specificity was difficult to explain by effects of these solutes on electrostatic and hydrophobic interactions. By elimination, we suggested that the measured attraction may be associated with either direct or water-mediated hydrogen bonding between recognition sites on opposing helices and that sugars and polyols somehow disrupt this interaction (*16*).

The goals of the present study are to verify this hypothesis, to clarify how sugars and polyols suppress the attraction, and to relate forces and solute effects observed in fibers with

* To whom correspondence should be addressed. LPSB/NICHD, National Institutes of Health, Bldg. 12A, Rm. 2041, Bethesda, MD 20892. E-mail: leikin@helix.nih.gov. Fax: 1-301-496-2172.

[‡] National Institutes of Health.

[§] On leave from Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia.

[||] Princeton University.

S0006-2960(98)00089-0 CCC: $15.00
Published on Web 08/07/1998

forces responsible for fiber formation. To achieve this, we now combine the direct measurement of forces between helices in fibers with the measurement of the effects of these compounds on collagen solubility. We also use a wider range of structurally related sugars, polyols, and alcohols compared to our previous study (*16*) and correlate structural, physical, and chemical properties of these solutes with their effects on fibrillogenesis.

METHODS

Sample Preparation. Type I collagen was extracted from rat tail tendons as described previously (*15*, *16*). It was solubilized in 0.5 M acetic acid solution (adjusted to pH 2.8 by NaOH) and digested with pepsin (≈100 mg/1 g of tendons in two doses, 24 h at 4 °C each). The pepsin treatment was used to prevent reformation of intermolecular covalent cross-links by removing terminal fragments of small, nonhelical telopeptides at the ends of each collagen molecule. The removal of a significant fraction of telopeptides was confirmed by near-UV absorption at 250−300 nm which indicated that ≤20% of the tyrosines localized in the terminal peptides remained after the treatment. The treatment was sufficient for preventing spontaneous reformation of intermolecular covalent cross-links in reconstituted fibers, as confirmed by SDS gel electrophoresis after sodium borohydride reduction. However, neither the triple helical collagen domain nor the formation and structure of the fibers was significantly affected by pepsin digestion. The protein was further purified by three cycles of salt precipitation and acetic acid solubilization.

It was observed in preliminary experiments that ∼10% of the protein was assembly incompetent. This fraction had a different UV absorption spectrum with a very strong tyrosine peak at 280 nm. It apparently consisted of small collagen fragments containing terminal peptides and/or residual pepsin not removed during the salt fractionation procedure and not detected by gel electrophoresis. Therefore, prior to the fibrillogenesis experiments, the protein was further purified by temperature-induced fiber formation, using the same protocol as for the fibrillogenesis measurements but without solutes. This reduced the assembly incompetent fraction to less than 1%.

D-glucose and methanol (Mallinckrodt, Paris, KY); D-fructose, α-methyl-D-glucoside (methylglucoside), and 1,3-propane diol (Fluka, Buchs, Switzerland); sorbitol (Spectrum Chemical, Gardena, CA); glycerol and ethylene glycol (J. T. Baker, Phillipsburg, NJ); 1,2-propane diol (Fisher Scientific, Fair Lawn, NJ); and ethanol (Warner-Graham, Cockaysville, MD) were all reagent grade quality and were used without further purification. Solutions of poly(ethylene glycol) (PEG, average molecular weight 8000 Da, United States Biochemical, Cleveland, OH) were used to apply osmotic stress on reconstituted collagen fibers.

Fibrillogenesis. Collagen solution (from 1.5 to 2 mg/mL) in 2 mM HCl (pH 2.7) was mixed 1:1 with a 2× initiation buffer containing 0.26 M NaCl, from 10 mM to 60 mM sodium phosphate (depending on the experiment), and different concentrations of one of the solutes listed above. The buffer pH was adjusted to give pH 7.4 after mixing with the collagen solution. The mixture was degassed for 5 min under vacuum and immediately placed into a UV-2101PC spectrophotometer (Shimadzu, Japan) where it was maintained at 32 °C. The mixing and degassing were conducted on ice to prevent assembly of collagen fibers (no assembly was observed on ice or following overnight incubation at 4 °C). Fibrillogenesis kinetics were measured by monitoring time dependence of the mixture optical density at 450 nm after the temperature jump to 32 °C. The measurement was initiated within 1−2 min after degassing and removal of the mixture from ice. The mixture temperature reached 32 °C within the first 1−2 min of the measurement.

Collagen solubility at 32 °C was measured after the end of the kinetics experiment, when no further change in the optical density was detected. After the equilibration, a white precipitate of assembled collagen fibers was spun down at 14000*g* for 5 min. The solubility was measured from protein concentration in the supernatant. Strictly speaking, the soluble protein fraction may contain dimers, trimers, or even larger aggregates together with monomers. However, this was not critical for the study since we used solubility to compare the effects of various solutes rather than to determine absolute values of the energies of interaction between collagen helices.

The concentration of collagen in the supernatant was evaluated from optical absorbency in the 215−230 nm region. A calibration curve was determined by measurement of the absorbency of a set of standard collagen solutions of different concentrations prepared gravimetrically from the dry protein. Precautions were taken to eliminate buffer components absorbing at these wavelengths. The error of the measurement did not exceed 20−30% at 5−10 μg/mL concentration of collagen, and it rapidly improved with increasing concentration of the protein. The 20−30% accuracy was not an impediment for the study since the essential solubility changes were on the scale from 10 to 100 fold, more than sufficiently exceeding the error.

Measurement of Interaction between Collagen Helices in Hydrated Fibers. Forces between collagen triple helices were measured as described previously (*15*−*17*). Briefly, collagen solution in 0.5 M acetic acid (pH 2.8, ∼1 g of total dry weight of the protein) was slowly concentrated at 4 °C and under 4 atm of pressure of nitrogen in an Amicon 8050 high-pressure ultrafiltration cell until a solid, optically transparent protein film was formed. The film was further dried by equilibrating against a 50 wt % solution of PEG 8000 for 2−3 days at 4 °C, separated from the dialysis membrane, washed in water (10 mM TrisCl, 2 mM EDTA, pH 7.5) to remove PEG, and air-dried at 4 °C. We previously established that the alignment of collagen molecules in the reconstituted film is the same as in native collagen fibers (*15*). Apparently the film consists of closely packed nativelike fibers.

Small, rectangular (∼1 × 0.5 × 0.5 mm) pieces were cut from the film and prehydrated in 10 mM TrisCl, 2 mM EDTA, pH 7.5 at 4 °C for at least one day. After this, the samples were equilibrated for at least a week in gravimetrically prepared PEG 8000 solutions at room temperature. In addition to PEG, the equilibrating solutions (adjusted to pH 7.5) contained 10 mM TrisCl, 2 mM EDTA, various concentrations of the solutes under study, and <5 mM residual NaCl from pH adjustment. Low salt conditions were important for the force measurements. Otherwise, the

osmotic action of NaCl would mask the solute-induced changes (*15*). Sample equilibration was performed in tightly sealed 1.5 mL microtubes fitted with O-rings. The solutions were refreshed in the middle of the equilibration, usually after the first 2−3 days.

After the equilibration in PEG, each sample was sealed in a specially designed cell with a small amount of the equilibrating solution and placed into an X-ray diffractometer. During X-ray exposure, the samples were maintained at 20 °C. The design and equipment used for the X-ray diffraction experiments were described (*15*−*18*). The lowest-order Bragg spacing for lateral packing of the helices was measured and an interaxial spacing, d_{int}, was calculated from it in the approximation of hexagonal packing of the helices.

The force versus separation curves were determined from the following considerations. PEG 8000 does not penetrate into the collagen film because its size is much larger than the separation between collagen helices (*15*−*17*). The surface of the sample acts as a semipermeable membrane which excludes PEG but allows water and solutes to equilibrate between collagen fibers and the bathing solution. As a result, PEG osmotically draws water from collagen fibers, pushing the helices closer together. The osmotic action of PEG is conveniently described by its osmotic pressure, Π_{PEG} (i.e., the pressure which would build up in a closed compartment containing the PEG solution, separated by a semipermeable membrane from pure water). The osmotic work of PEG upon removing the volume, ΔV, of water from the collagen film is equal to $\Pi_{PEG}\Delta V$. The effective osmotic work of PEG in the presence of other solutes was corrected for the volume occupied by those solutes assuming that they did not interact with PEG. The required correction was small enough so that possible errors due to simplifying assumptions were negligible.

In equilibrium, a net repulsive force between collagen helices counteracts the osmotic action of PEG, preventing the helices from approaching each other too closely. The work of the repulsion and the osmotic work of PEG balance each other. This allows us to calculate the force, *f*, per unit length of the helices as described before (*19*)

$$f = \frac{\Pi_{PEG} d_{int}}{\sqrt{3}} \quad (1)$$

assuming that lateral packing of collagen can be approximated by a hexagonal lattice and assuming that the interactions between collagen helices are pairwise additive. For a more detailed description of this osmotic stress technique for force measurement see, for example, ref *20*.

Correction of Force Data for Osmotic Action of Excluded Solutes. To extract the direct effect of a solute on forces between collagen helices in fibers, we take into account that the concentration of the solute inside fibers may be lower than that outside. The gradient of solute concentration resulting from such solute exclusion produces the same osmotic effect as the gradient of PEG concentration. We describe this effect by an osmotic pressure, Π_{solute}, of the excluded solute. The value of Π_{solute} is determined by the dependence of the solute deficit in fibers on the spacing between collagen helices. Within a simple approximation of constant solute concentration at all locations in fibers,

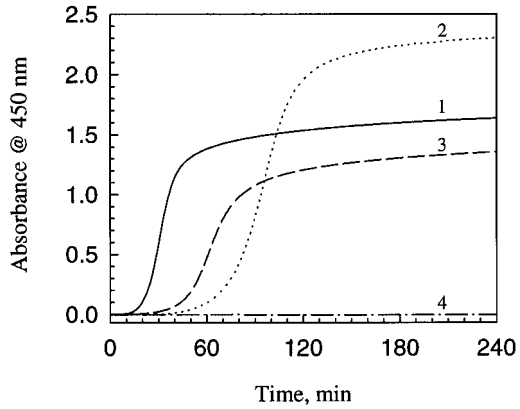

FIGURE 1: Kinetics of collagen self-assembly at 32 °C, pH 7.4, 0.13 M NaCl in different fibrillogenesis buffers: (1) 10 mM sodium phosphate, no solutes; (2) 25 mM sodium phosphate, no solutes; (3) 10 mM sodium phosphate, 1 mol of 1,2-propane diol; and (4) 10 mM sodium phosphate, 1 mol of 1,3-propane diol. All solutions contained 0.7 mg/mL collagen. No fibril assembly was observed in 1 mol of 1,3-propane diol after 24 h.

Π_{solute} is proportional to the difference in the solute concentration inside and outside the fibers. When the value of Π_{solute} is known, the measured force, *f*, can be corrected for this osmotic pressure as

$$\tilde{f} = f + \frac{\Pi_{solute} d_{int}}{\sqrt{3}} \quad (2)$$

so that the corrected force, $\tilde{f}$, reflects only the direct solute effect. This correction is based on the assumption that solute−PEG interaction is not significant. Then, the osmotic pressure of the excluded solute simply adds to that of PEG, as reflected in eq 2.

RESULTS

Fibrillogenesis: Solubility Measurement. Pepsin-treated collagen helices spontaneously assemble into nativelike fibers at an elevated temperature (*21*). Figure 1 illustrates kinetics of fibrillogenesis at 32 °C in solutions of several different compositions. Fast collagen fiber growth occurs after a lag phase. It is followed by slow fiber aggregation, as previously described (*3*). A small fraction of the protein remains dissolved in solution in equilibrium with the fibers.

Addition of various solutes affects both the fibrillogenesis kinetics (Figure 1) and the solubility (Figure 2). In this study, we concentrate on the solubility changes. It is difficult to separate kinetic effects associated with solute-induced viscosity changes from those caused by changes in collagen−collagen interactions. Furthermore, solubility can be better compared with intermolecular forces since both characterize the energy of interaction between collagen molecules at equilibrium.

It is clear from Figure 2 that solute effects on fibrillogenesis are remarkably specific. Comparison of 1,2- and 1,3-propane diols (Figures 1 and 2) is most revealing. These diols have the same atomic composition and differ only in the position of one of their hydroxyls (Figure 3). They have practically identical effects on dielectric properties of water. Nevertheless, collagen solubility in 1 molal (mol) 1,3-propane diol is ∼20 times higher than in 1 mol 1,2-propane diol. Interestingly, solute dependence of the stability of an

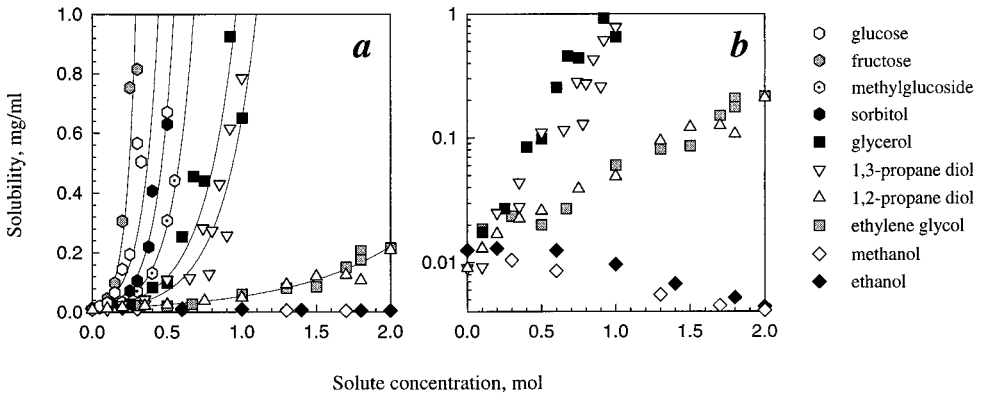

FIGURE 2: Collagen solubility at 32 °C, pH 7.4, 10 mM sodium phosphate, 0.13 M NaCl, in the presence of different solutes (as shown by different symbols) plotted in linear (a) and logarithmic (b) scales. Note that the logarithm of solubility is proportional to the energy of interaction between collagen helices if the fraction of collagen monomers is much larger than the fraction of soluble protein aggregates in the supernatant.

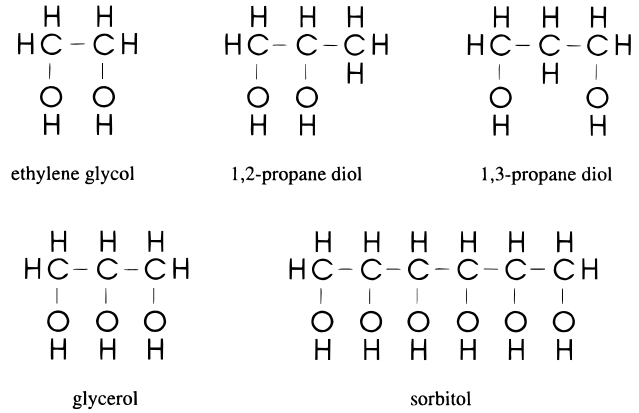

FIGURE 3: Schematic representation of diols and polyols used in the study.

isolated collagen triple helix with respect to denaturation is exactly opposite to the stability of fibers, that is, 1,3-propane diol protects the helix against denaturation while 1,2-propane diol promotes denaturation (*13*).

Note that in analyzing solute effects on collagen solubility it is important to keep in mind that fibrillogenesis buffers typically contain phosphate to produce more regular fibers (*22*). Solutes may affect the activity of phosphate ions which are very potent fibrillogenesis inhibitors (compare, for example, fibrillogenesis kinetics at 10 and 25 mM phosphate, Figure 1). Such solute−phosphate coupling should be taken into account.

Consider the effects of 1,2- and 1,3-propane diols in detail. Collagen solubility, s (μg/mL), at 32 °C and different propane diol concentrations, c_{solute} (mol), and different phosphate concentrations, $c_{phosphate}$ (M), is shown in Figure 4. From linear regression analysis we find that the measured solubility is best described by

$$\log_{10}(s) = \log_{10}(s_0) + k_{phosphate}c_{phosphate} + k_{solute}c_{solute} + \kappa c_{phosphate}c_{solute} \quad (3)$$

where s_0 is the extrapolated solubility in the absence of either phosphate or propane diols, $k_{phosphate}$ is the phosphate inhibition constant, k_{solute} is the propane diol inhibition constant, and κ is the propane diol−phosphate coupling constant. The values of the constants obtained by least-squares fitting are $k_{phosphate} \approx 13$ μg/(mL M); $k_{solute} \approx 0.5$ μg/(mL mol) and $\kappa \approx 17$ μg/(mL M mol) for 1,2-propane diol; and $k_{solute} \approx 1.5$ μg/(mL mol) and $\kappa \approx 37$ μg/(mL M mol) for 1,3-propane diol. We find from eq 3 that solute−phosphate coupling can be neglected at $c_{phosphate} \ll k_{solute}/\kappa$, where $k_{solute}/\kappa \approx 40$ mM for both propane diols. Assuming that this constant is not very different for other solutes, we compare solute effects in 10 mM phosphate buffers to emphasize direct solute effects and to minimize the solute−phosphate coupling.

Effects of other solutes on collagen solubility are highly specific as well (Figure 2). Glucose, fructose, methylglucoside, sorbitol, and glycerol strongly inhibit collagen fiber assembly. Ethylene glycol has a much weaker inhibitory effect. Methanol and ethanol promote fibrillogenesis. These solutes also induce opposite changes in the stability of fibers and in the stability of isolated collagen helices. Sugars, glycerol, and ethylene glycol protect collagen while methanol and ethanol promote denaturation (*13*).

Interaction between Collagen Helices in Hydrated Fibers. The forces measured between collagen molecules in fibers (per unit length of the helix) are plotted in Figures 5 and 6 versus separation between the central axes of helices (the helical diameter is ~12 Å). The osmotic stress technique allows us to measure forces only when they are net-repulsive. Our previous study has shown that such net force is composed of two major components, a repulsion and an attraction. These components can be separated by suppressing the attraction by several different methods (*16*). Then the repulsion is measured independently and the attraction is calculated as a difference between the net force and the repulsion.

The net force between helices in water without solutes at pH 7.5 and 20 °C is shown by a solid line on each graph. Its repulsive component is shown by a dotted line. The repulsion dominates at 12−18 Å interaxial separations (*16*). The attraction decreases with the distance slower than the repulsion and dominates at larger distances (*15*, *16*). The attraction is responsible for molecular recognition between the helices and for their spontaneous assembly (fibrillogenesis) from solution into nativelike fibers.

Different solutes have qualitatively different effects on the interaction between helices (Figures 5 and 6). As a result, glucose, fructose, glycerol, and 1,3-propane diol cause stronger swelling of fibers; ethylene glycol has virtually no effect; and methylglucoside, sorbitol, 1,2-propane diol,

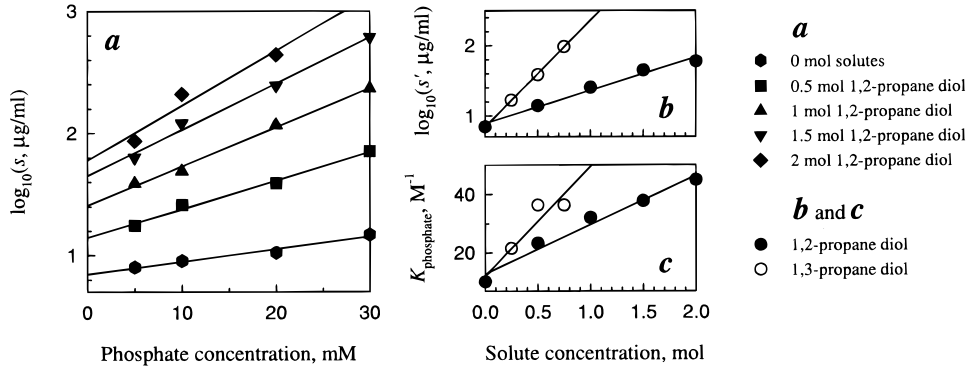

FIGURE 4: Coupling between the effects of propane diols and phosphate on collagen fibrillogenesis at 32 °C, pH 7.4, 0.13 M NaCl: (a) Collagen solubility at different concentrations of phosphate and 1,2-propane diol. Straight lines represent least-squares fits by the equation $\log_{10}(s) = \log_{10}(s') + K_{phosphate}c_{phosphate}$. (b) Collagen solubility in 1,2-propane diol and 1,3-propane diol solutions extrapolated to 0 phosphate concentration ($\log_{10}(s')$) obtained from fitting the data in (a) and similar data for 1,3-propane diol). (c) Propane diol−phosphate coupling, as characterized by the dependence of $K_{phosphate}$ on 1,2-propane diol and 1,3-propane diol concentrations.

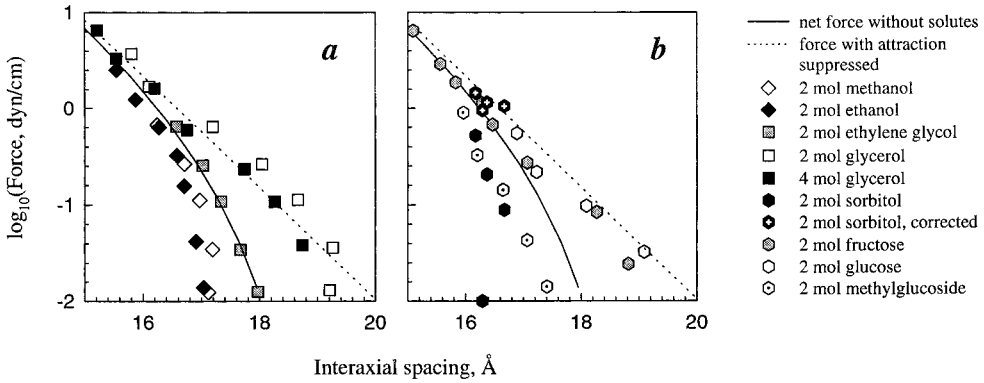

FIGURE 5: Force per unit of molecular length between collagen triple helices measured at 20 °C, pH 7.5, 10 mM Tris, 2 mM EDTA without (solid line) and with various solutes (shown by different symbols). Except for sorbitol solutions, net forces are shown without correction for osmotic action of solutes excluded from collagen fibers (see Methods). Forces between helices in sorbitol solutions are shown without and after the correction. The dotted line shows the repulsive component of the net interaction, measured by suppressing the attraction with a variety of different methods (*16*).

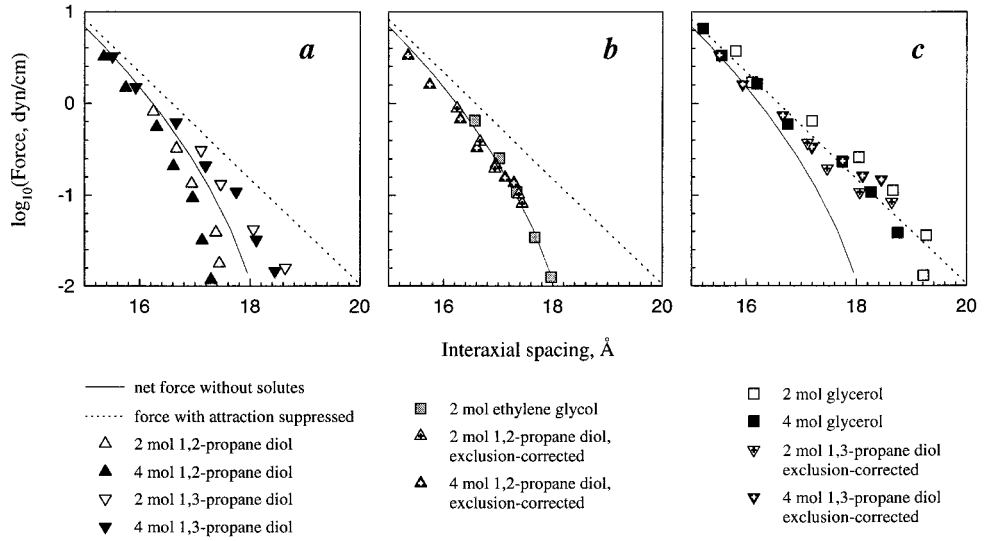

FIGURE 6: Force per unit of molecular length between collagen triple helices measured at 20 °C, pH 7.5, 10 mM Tris, 2 mM EDTA without (solid line) and with 1,2-propane diol and 1,3-propane diol. The forces are shown before (a) and after (b, c) the correction for the osmotic effect associated with 1.3% exclusion of propane diols from collagen fibers. The forces between helices in ethylene glycol (b) and glycerol (c), same as in Figure 5, are shown for comparison.

methanol, and ethanol reduce the swelling. Separation of repulsive and attractive components of the measured net forces has shown that glycerol reduces the attraction, without affecting the repulsion (*16*). Below we argue that glucose, fructose, and 1,3-propane diol act in a similar way.

Note that interpretation of the measured force curves in terms of direct interaction between collagen helices in fibers is based on the assumption that none of the solutes acts osmotically. This assumption holds for some but not all of the solutes. To extract the true change in the interaction

between collagen helices, a correction for solute osmotic effect should be applied, as described in Methods and discussed below.

DISCUSSION

Comparison of Force and Solubility Measurements. Solutes may affect fibrillogenesis (a) directly, they may modify interaction between collagen helices, for example, by changing dielectric properties of the solution or by disrupting hydrogen-bonded water clusters bridging opposing helices; (b) osmotically, they may not have any direct influence on the interaction and may affect fibrillogenesis because solute-depleted (or enriched) layers at collagen surfaces overlap at close separations between helices. To better distinguish these effects and to understand their underlying mechanisms, we combine traditional measurement of collagen solubility with direct measurement of forces between helices in fibers.

Our previous study suggested that glycerol has virtually no osmotic effect and that it inhibits fibrillogenesis by directly suppressing an attractive interaction between helices which is responsible for the self-assembly of collagen fibers (*16*). Comparison of the force and solubility data (Figures 2 and 5) suggests that glucose and fructose act in a very similar manner to that of glycerol, except the monosaccharides suppress the attraction more efficiently.

Surprisingly, sorbitol and methylglucoside appear to act differently (Figure 5). Structurally, sorbitol is a di-glycerol (Figure 3), and it inhibits fibrillogenesis in solution with nearly the same efficacy as a glycerol concentration two times higher (Figure 2). Methylglucoside is a glucose derivative with a single proton replaced by methyl. It has almost the same effect on collagen solubility as glucose. One would expect sorbitol and methylglucoside to affect interaction between collagen helices in fibers similarly to the way glycerol and glucose affect the interaction. However, their observed action is opposite to this expectation (Figure 5).

The most likely explanation is that sorbitol and methylglucoside are excluded from collagen fibers and act osmotically, in addition to directly suppressing the attraction between collagen helices. Indeed, excluded solutes compress fibers by drawing water from them (this osmotic effect is similar to the action of PEG, see Methods for more details). The osmotic compression may counteract suppression of the attraction between helices. The outcome of the competition depends on the extent of solute exclusion. Apparently, strong exclusion of sorbitol and methylglucoside completely masks the attraction loss so that uncorrected force curves create an impression that these solutes reduce net repulsion between collagen molecules (Figure 5b).[1]

Consider the effect of sorbitol in more detail. If our assumptions are correct, then the direct effect of sorbitol remaining inside fibers is similar to that of glucose while the osmotic effect of excluded sorbitol is similar to that of PEG. The net effect of sorbitol is a sum of the two. To explain the observed spacings between helices in fibers in 2 mol of sorbitol (Figure 5b) quantitatively, we assume that the excluded sorbitol applies an osmotic pressure of ~10 atm (0.5 osmol), that is, that the sorbitol concentration inside fibers is ~25% lower than that outside (which is quite realistic). Then, after correction for the osmotic action (see Methods), the force data in sorbitol solution overlap with those for glycerol and glucose (Figure 5b). This hypothesis also explains why the spacing in sorbitol is not sensitive to the addition of PEG at low PEG concentration. This is because excluded sorbitol itself acts as PEG, that is, PEG may affect the spacing only when its osmotic pressure exceeds 10 atm.

Similar osmotic effects appear to contribute to the action of 1,2- and 1,3-propane diols. Indeed, 1,2-propane diol affects collagen solubility exactly as ethylene glycol does (Figure 2), but it has a different effect on the measured forces between helices in fibers (compare Figure 5a and Figure 6a). Similarly, the effects of 1,3-propane diol and glycerol on the solubility are identical but their effects on the forces are different (Figures 5a and 6a). A correction of the force curves for the osmotic pressure of propane diols excluded from fibers (see Methods) completely removes this discrepancy (Figure 6, parts b and c), if we assume that the concentration of either diol inside fibers is ~1.3% lower than that outside. Therefore, the direct action of 1,3-propane diol on the interaction between collagen helices is exactly the same as that of glycerol while the direct effect of 1,2-propane diol is similar to that of ethylene glycol. Although we achieve this agreement by fitting the solute partition coefficient, this single fitting parameter allows us to describe the distance and concentration dependence of the forces measured in both diols. This is unlikely to be a mere coincidence.

It is somewhat surprising that sorbitol, methylglucoside, and 1,2- and 1,3-propane diols are excluded from fibers while closely related glucose and glycerol are not. We believe that this can be explained as follows. The exclusion of sorbitol from collagen fibers is probably associated with its much greater conformational flexibility compared to glucose and glycerol. The conformational freedom of sorbitol may be restricted near the collagen surface resulting in sorbitol depletion. Methylglucoside and propane diols appear to be excluded because of their extra hydrophobicity. (Methylglucoside is a glucose derivative with a methyl group replacing one proton; both propane diols are glycerol derivatives with a proton replacing one of hydroxyls.) This is consistent with our previous observation that collagen fibers swell less in more hydrophobic nonaqueous solvents (*17*). Apparently collagen has a solvation pattern such that, on average, hydrophobic compounds are weakly depleted near the protein surface.

Conventional Interpretations of Sugar and Polyol Effects on Fibrillogenesis. The dependence of collagen solubility on glycerol concentration has been analyzed before (*5*) in the traditional terms of preferential solvation (see, e.g. ref *23* and references therein). Such analysis is very useful to the extent that it allows one to determine thermodynamic relationships between the dependence of collagen solubility

[1] Our data suggest that the osmotic effects of the solutes on collagen solubility are much weaker than the osmotic effects on forces between helices in fibers. The reason for this may be understood as follows. The logarithm of the solubility is proportional to the force integrated from infinity to the equilibrium separation between helices in fibers. Thus, the osmotically induced change in the solubility is proportional not only to the osmotic force but also to the depth of overlap between the solute-depleted layers at collagen surfaces. Such overlap may be small because collagen fibers are strongly hydrated. If the solutes are depleted mostly from the primary solvation shell of the protein, such shells barely touch (if at all) at the equilibrium separation in fibers.

on solute concentration and the change in the net surface excess (deficit) of the solute upon fiber formation. However, such an approach is not appropriate for answering the central questions of the present study. Indeed, as with any purely thermodynamic approach, it does not address the physical mechanisms of solute action and solute specificity. (For example, both direct and osmotic solute effects are accompanied by a change in the surface excess of the solute.)

The disruptive effect of glucose on collagen fiber stability and fibrillogenesis has been analyzed in numerous studies (see refs *6*, *24*, and *25* and references therein). The prevalent point of view is that this is related to chemical modification of collagen by glucose which may exist in the open-chain, aldehyde form. It has been shown that such nonenzymatic glycosylation plays an important role in fibrillogenesis and in the destabilization of collagen fibers *in vivo* (*6*, *24*, *25*). However, by itself, collagen glycosylation cannot explain the action of glucose observed in our experiments. Indeed, sorbitol and methylglucoside are closely related to glucose. They inhibit collagen fibrillogenesis with nearly the same efficacy as glucose, but they do not have the same ability to modify the protein chemically. Furthermore glycerol and 1,3-propane diol appear to act in exactly the same way as glucose acts, but they also cannot react with the protein. The mechanism of fibrillogenesis inhibition and fiber destabilization in our experiments must be different.

Molecular Mechanisms of Collagen Self-Assembly and Solute Action. To summarize, previous studies have shown that important interactions between collagen helices in fibers involve a short-range exponential repulsion and a longer-range attraction (*16*). Comparison of the force and solubility data now indicates that it is this attraction that is responsible for the recognition between collagen helices and fibrillogenesis. Solutes which have the same direct effect on the attraction in fibers also have the same effect on fibrillogenesis. Glucose, fructose, methylglucoside, sorbitol, glycerol, and 1,3-propane diol inhibit fibrillogenesis by directly suppressing the attraction. In contrast, 1,2-propane diol, ethylene glycol, methanol, and ethanol have minimal, if any, effects on it. This amazing specificity provides us a key to understanding the mechanism of the attraction and how the strong inhibitors suppress it.

Previous comparison of the effects of methanol, ethanol, ethylene glycol, glycerol, and glucose on the attraction suggested that electrostatic, van der Waals, and hydrophobic interactions do not contribute significantly to it (*16*). We believe that comparison of the very different effects of 1,2- and 1,3-propane diols proves that this is the case. Indeed, the two propane diols have identical effects on dielectric properties of aqueous solutions (both static and optical) which determine the strength of electrostatic and van der Waals interactions. It is 1,2-propane diol rather than 1,3-propane diol that is capable of disrupting hydrophobic attraction (e.g. 1,2-propane diol promotes denaturation of collagen and other proteins while 1,3-propane diol inhibits it (*13*)). In contrast our experiments show that 1,3-propane diol strongly suppresses the attraction between collagen helices while 1,2-propane diol does not.

By excluding electrostatic, van der Waals, and hydrophobic mechanisms of attraction we arrive at the conclusion that collagen self-assembly is caused either by direct or by water-mediated hydrogen bonding between some recognition sites on the molecules. (The recognition sites are still not known, but it has been argued that Gly-X-Hyl-Gly-His-Arg sequences, which are highly conserved in all fibril-forming collagens (*26*), are likely to be involved (*3*, *16*).)

High-resolution X-ray structure of a hydrated crystal formed by a collagenlike, triple helical peptide suggests that hydrogen-bonded water bridges are likely sources of the attraction (*27*, *28*). This structure revealed malleable hydrogen-bonded water clusters bridging and holding peptide triple helices together. Although this peptide does not contain sequences similar to the recognition sites on collagen, the X-ray structure gives a good visual example illustrating the likelihood and power of such interaction. The interaction is sufficiently strong to hold together the triple helices in a crystal.

The present study indicates that sugars and polyols inhibit fibrillogenesis by incorporating into and disrupting hydrogen-bonded water clusters bridging collagen helices. In particular, the observed linear dependence of the logarithm of collagen solubility and, therefore, of the energy of interaction between the helices on the concentration of the solutes (Figure 2b) is more consistent with such a model. One would expect a nonlinear dependence with saturation if the solutes directly bind to and block the recognition sites. The drastic difference in the action of 1,2- and 1,3-propane diols supports this model as well. These two diols may have different effects on water structure because of different distances between their hydroxyls.

Most likely, the remarkable specificity of fibrillogenesis with respect to minute differences in solute structure is related to differences in the hydrogen-bonding stereochemistry specific to each solute. For example, 1,2-propane diol and ethylene glycol have stereochemically identical hydroxyls (Figure 3) and the same effect on fibrillogenesis, regardless of the fact that 1,2-propane diol is more hydrophobic than ethylene glycol and that these solutes differ in other physical properties as well. At the same time, 1,2- and 1,3-propane diols have drastically different effects on fibrillogenesis, although they have more similar physical properties. Apparently, the difference in the position of one hydroxyl group plays the dominant role. We believe that the stereochemistry of 1,3-propane diol hydroxyls allows this solute to incorporate into water clusters bridging recognition sites on collagen helices and to terminate essential hydrogen-bond chains. The same applies to other polyols and sugars which are strong fibrillogenesis inhibitors.

CONCLUSIONS

Our most important findings can be summarized as follows. (1) Formation of salt bridges and hydrophobic interactions between opposing nonpolar amino acid side chains is not a major driving force of collagen self-assembly, in contrast with what is frequently assumed. (2) Collagen fibrillogenesis is driven primarily by the formation of hydrogen-bonded water clusters bridging recognition sites on opposing helices. The exact location and structure of the recognition sites are still unknown, but Gly-X-Hyl-Gly-His-Arg sequences near each end of the molecule are likely candidates. (3) Glucose, fructose, sorbitol, methylglucoside, glycerol, and 1,3-propane diol strongly inhibit fibrillogenesis by competing with water for crucial hydrogen bonds and,

thus, disrupting water clusters bridging the recognition sites. The ability of a solute to incorporate and disrupt these water clusters appears to be determined by the stereochemistry of the solute hydrogen-bonding pattern. (4) Solutes may also hinder fibrillogenesis by enhancing the inhibitory action of phosphate ions. This effect becomes pronounced above ~10 mM phosphate, the concentration which is frequently used in fibrillogenesis buffers.

ACKNOWLEDGMENT

We are grateful to Adrian Parsegian and Donald Rau for stimulating discussions.

REFERENCES

1. Kadler, K. (1994) *Protein Profile 1*, 519−638.
2. Prockop, D. J., and Kivirikko, K. I. (1995) *Annu. Rev. Biochem. 64*, 403−434.
3. Veis, A., and Payne, K. (1988) in *Collagen* (Nimni, M. E., Ed.) Vol. I, pp 113−137, CRC Press, Boca Raton, FL.
4. Hayashi, T., and Nagai, Y. (1972) *J. Biochem. 72*, 749−758.
5. Na, G. C., Butz, L. J., Bailey, D. G., and Carroll, R. J. (1986) *Biochemistry 25*, 958−966.
6. Rathi, A., Padmavathi, S., and Chandrakasan, G. (1989) *Biochem. Med. Metab. Biol. 42*, 209−215.
7. Leonardi, L., Ruggeri, A., Roveri, N., Bigi, A., and Reale, E. (1983) *J. Ultrastruct. Res. 85*, 228−237.
8. Gerlsma, S. Y., and Stuur, E. R. (1972) *Int. J. Pept. Protein Res. 4*, 377−383.
9. Back, J. F., Oakenfull, D., and Smith, M. B. (1979) *Biochemistry 18*, 5191−5196.
10. Gekko, K., and Timasheff, S. N. (1981) *Biochemistry 20*, 4667−4676.
11. Arakawa, T., and Timasheff, S. N. (1982) *Biochemistry 21*, 6536−6544.
12. Gekko, K. (1983) in *Ions and Molecules in Solution* (Tanaka, N., Ohtaki, H., and Tamamushi, R., Eds.) pp 339−358, Elsevier, Amsterdam, The Netherlands.
13. Harrap, B. S. (1969) *Int. J. Protein Res. 1*, 245−252.
14. Gekko, K., and Koga, S. (1983) *J. Biochem. 94*, 199−205.
15. Leikin, S., Rau, D. C., and Parsegian, V. A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 276−80.
16. Leikin, S., Rau, D. C., and Parsegian, V. A. (1995) *Nat. Struct. Biol. 2*, 205−10.
17. Kuznetsova, N., Rau, D. C., Parsegian, V. A., and Leikin, S. (1997) *Biophys. J. 72*, 353−362.
18. Mudd, C. P., Tipton, H., Parsegian, V. A., and Rau, D. C. (1987) *Rev. Sci. Instrum. 58*, 2110−2114.
19. Rau, D. C., Lee, B., and Parsegian, V. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2621−5.
20. Parsegian, V. A., Rand, R. P., Fuller, N. L., and Rau, D. C. (1986) *Methods Enzymol. 127*, 400−16.
21. Comper, W. D., and Veis, A. (1977) *Biopolymers 16*, 2113-2131.
22. Williams, B. R., Gelman, R. A., Poppke, D. C., and Piez, K. A. (1978) *J. Biol. Chem. 253*, 6578−6585.
23. Timasheff, S. N. (1993) *Annu. Rev. Biophys. Biomol. Struct. 22*, 67−97.
24. Lien, Y. H., Stern, R., Fu, J. C. C., and Siegel, R. C. (1984) *Science 225*, 1489−1491.
25. Tanaka, S., Avigad, G., Brodsky, B., and Eikenberry, E. (1988) *J. Mol. Biol. 203*, 495−505.
26. Ramirez, F. (1989) in *Collagen* (Olsen, B. R., and Nimni, M. E., Eds.) Vol. IV, pp 21−30, CRC Press, Boca Raton, FL.
27. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) *Science 266*, 75−81.
28. Bella, J., Brodsky, B., and Berman, H. (1995) *Structure 3*, 893−906.

BI980089+