Conformational flexibility of collagen molecules: composition and environmental effects

Naghmeh Rezaei, Aaron Lyons and Nancy R. Forde*

Department of Physics

Simon Fraser University

8888 University Drive

Burnaby, BC V5A 1S6

CANADA

*Corresponding author. 778-782-3161. nforde@sfu.ca. ORCID: 0000-0002-5479-7073

CLASSIFICATION:

BIOLOGICAL SCIENCES: Biophysics and Computational Biology

PHYSICAL SCIENCES: Physics

KEYWORDS: collagen, atomic force microscopy, single-molecule, worm-like chain, curvature, persistence length

ABSTRACT

Despite its prevalence and physical importance in biology, the mechanical properties of molecular collagen are far from established. The flexibility of the triple helix is unresolved, with descriptions from different experimental techniques ranging from flexible to semirigid. Furthermore, it is unknown how collagen type (homo- vs. heterotrimeric) and source (tissue-derived vs. recombinant) influence flexibility. Using SmarTrace, a chain tracing algorithm we devised, we performed statistical analysis of collagen conformations collected with atomic force microscopy (AFM) to determine the protein's mechanical properties. Our results show that types I, II and III collagens – the key fibrillar varieties – exhibit molecular flexibilities that are very similar. However, collagen conformations are strongly modulated by salt, transitioning from compact to extended as KCl concentration increases, in both neutral and acidic pH. While analysis with a standard worm-like chain model suggests that the persistence length of collagen can attain almost any value within the literature range, closer inspection reveals that this modulation of collagen's conformational behaviour is not due to changes in flexibility, but rather arises from the induction of curvature (either intrinsic or induced by interactions with the mica surface). By modifying standard polymer theory to include innate curvature, we show that collagen behaves as an equilibrated curved worm-like chain (cWLC) in two dimensions. Analysis within the cWLC model shows that collagen's curvature depends strongly on pH and salt, while its persistence length does not. These results show that triple-helical collagen is well described as semiflexible, irrespective of source, type, pH and salt environment.

SIGNIFICANCE STATEMENT

The predominant structural protein in vertebrates is collagen, which plays a key role in extracellular matrix and connective tissue mechanics. Previous measurements on molecular collagen have provided measures of flexibility that vary by over an order of magnitude. Thus, the mechanics of triple-helical collagen – the fundamental building block of structural tissues – are not established. Our results, obtained using single-molecule imaging analysis, find that collagen deposited on a mica surface adopts intrinsically curved structures at low salt and pH. After accounting for this curvature, we find that the flexibility of collagen (as determined by its persistence length) is not affected significantly by salt, pH or composition. Thus, collagen appears to be well described as a semiflexible polymer.

INTRODUCTION

Collagen is the predominant structural protein in vertebrates, where it represents more than one quarter of the total protein in our bodies (1). It is widely used as a biomaterial, and plays a vital physiological role in extracellular matrix and connective tissue mechanics. Collagen can assemble into many different higher-order forms, which fulfil distinct structural and mechanical roles.

Over 28 different types of human collagen have been identified (1), of which the most prevalent are the fibrillar collagens such as types I, II and III. These collagens assemble to create highly ordered fibrils, which in turn are the building blocks for the extracellular matrix and for fibres, which act as load- and tension-bearing structures in connective tissues. Not surprisingly, changes in collagen's composition are associated with a wide variety of diseases, including osteogenesis imperfecta (2) and Ehlers-Danlos syndrome (3). In addition to genetic mutations in collagen, physiological dysfunction can arise from alterations in posttranslational modifications, and aging via nonenzymatic glycation and crosslink formation (4). Such chemical changes at the protein level correlate with altered tissue structure and mechanics, and pathologically affect human health. Because of the hierarchical nature of collagen structure, identifying the mechanisms by which chemical changes modify tissue mechanics requires understanding how they impact mechanics at the molecular level.

Collagen's molecular structure is a right-handed triple helix (Figure 1A), approximately 300 nm long and 1-2 nm in diameter (1). This triple helix comprises three left-handed polyproline-II-like helices (α-chains) with a characteristic $(\text{Gly-X-Y})_n$ repeat amino acid sequence. Although frequently proline or hydroxyproline, the X and Y amino acids are variable, and provide the sequence diversity that defines individual collagen types. Different collagen types are also distinguished as homotrimeric (three identical α-chains) or heterotrimeric. Although the sequences of the different types of collagen are well established, their mechanical properties are not.

One property commonly used to describe the mechanical properties of chain-like biomolecules is the persistence length – the length over which the orientation of the chain remains correlated in the presence of thermal noise. While the persistence lengths of other biological polymers such as DNA are well established and robust to experimental technique, this is not the case for collagen (Table 1). Early solution-based studies found that collagen's triple helix should be considered semi-rigid, exhibiting a persistence length of $p \approx$ 130-180 nm (5-7). These values likely led to the textbook description of collagen as a possessing a stiff, rod-like structure (8). However, recent single-molecule experimental and simulation studies contradict this description, finding persistence lengths for collagen as short as $p \approx 10$ nm (9-15). These values imply that collagen is highly flexible, adopting compact, coiled configurations in solution. Other experimental and simulation approaches have determined persistence lengths throughout this range, with another cluster of values around $p \approx$ 40-60 nm (12, 16-18). It is remarkable that this fundamental parameter of persistence length is so poorly established for a protein of such mechanical importance and ubiquity.

Examination of Table 1 suggests that there is diversity not only in the approaches used to study collagen's flexibility but also in the types of collagen, their sources, and solution conditions. A comparison of different tissue-derived collagen types was performed in reference (16), which determined type I heterotrimer to be slightly more rigid than type III homotrimer. A similar difference between homo- and heterotrimeric collagen sequences was found in recent MD simulations (14). Collagen's remodelling by MMP1, a collagenase that unwinds the triple helix, has been suggested to

depend on the source of collagen (e.g. recombinant, from cell culture or from tissue) (19), implying that posttranslational enzymatic and age-related nonenzymatic modifications may alter its mechanics at the molecular level. A recent study found that collagen's flexibility is influenced by its solution environment, with collagen appearing flexible at low salt in acidic conditions and more rigid in high-ionic-strength, neutral pH buffers (11). This result runs counter to the expectations of the behavior of polyelectrolytes like DNA, which are expected to become more flexible as ionic strength is increased (20). Understanding collagen's response in different solution conditions is relevant not only to its mechanical role in distinct tissue environments, but also to understanding its intracellular compactness as it traverses a pH gradient during cellular processing and secretion (21, 22).

In this work, we use atomic force microscopy imaging and characterization to investigate how composition, source and chemical environment affect the flexibility of collagen's triple helix. By comparing different fibrillar collagens (types I, II and III; tissue-derived and recombinantly expressed; homo- and heterotrimeric triple helical structures), we find that molecular composition does not significantly impact collagen's overall flexibility. In contrast, we find that both ionic strength and pH independently impact the apparent flexibility of collagen, providing estimates of persistence length that span the range from flexible to semi-rigid, depending on solution environment. Careful consideration of polymer chain statistics shows, however, that treating collagen as a standard worm-like chain does not describe its properties in most of these chemical environments. Instead, we show that a curved worm-like chain model that includes inherent molecular curvature is a far more appropriate descriptor of the observed conformations of collagen on mica. With this model, collagen's bending flexibility depends much less on solution conditions; instead, salt concentration and pH modulate its global curvature. We find that curvature does not depend strongly on the type or source of collagen, nor does the persistence length. To the best of our knowledge, these results constitute the first experimental analysis of curved worm-like chains, and provide a new explanation for the conflicting reports on the flexibility of triple-helical collagen.

RESULTS AND DISCUSSION

Atomic force microscopy (AFM) is a well-established tool for imaging the conformations of flexible biopolymers (23). We have used the technique to image collagens of different types and sources, deposited from a range of solution conditions, and have characterized how these parameters influence the flexibility of collagen's triple helix.

Collagen type and source

To investigate how collagen's flexibility depends on its source, we compared four different samples of collagen. These represent three types of fibrillar collagens (types I, II and III), of either rat (type I) or human (all three) genetic origin. Although all of these samples are capable of self-assembly into highly ordered fibrils, they were imaged under sufficiently dilute conditions such that assembly did not occur. Type I collagen is a heterotrimer ($\alpha_1(I)_2\alpha_2(I)_1$), while types II and III are each homotrimeric ($\alpha_1(II/III)_3$). Additionally, to explore how collagen's source influences its flexibility, our samples encompass both tissue-derived (rat type I and human type II) and cell-derived (human types I and III) sources. The latter

were recombinantly expressed in yeast, and possess only prolyl hydroxylation as a posttranslational modification (PTM) (24). This contrasts with mammalian-produced collagens which have extensive PTMs including hydroxylation of prolines and lysines, as well as O-glycosylation of hydroxylysines (12, 21).

Representative AFM images of each of the four collagen samples are shown in Figure 1, B-E. Each image was obtained by depositing collagen from a room-temperature solution of 100 mM KCl + 1 mM HCl onto freshly cleaved mica, then rinsing with water and drying prior to imaging. For all four collagen samples, the chains appear semiflexible, with contour lengths of approximately 300 nm, as expected. There are no obvious qualitative differences in flexibility among these samples.

To quantify the flexibility of each collagen type, the imaged chains were traced to provide backbone contours for conformational analysis. Existing chain-tracing algorithms proved problematic for many of our images, in some cases due to the background levels of noise (25), or in others from the sensitivity of the algorithm to the locations of user-input starting points (26). Hence, we developed a new chain-tracing program, dubbed SmarTrace. The algorithm incorporates pattern matching to identify the best centerline of a chain, and refines an initial guess at this centerline by using direction and width continuity constraints. Required user input is minimal: backbone contours are identified from only a few clicks near the chain (see Figure 1F). Details of SmarTrace's methodology and workflow are provided in the Supporting Information (SI).

From traced chains, we implemented tools of polymer physics to analyse the statistical properties of chain conformations and determine persistence lengths (27). To perform this analysis, chains were segmented randomly into non-overlapping pieces of different contour lengths (e.g. $s = 10, 20, \ldots, 200$ nm). This approach allows partially traced chains to be included in the analysis (25). For each segment, we calculated its squared end-to-end distance $R^2(s)$ and the change in orientation between its starting and ending tangent vectors $\hat{t}(s) \cdot \hat{t}(0) = \cos\theta\,(s)$; these quantities were then averaged over all segments of length s within the population. Length-dependent trends in mean squared end-to-end distance, $\langle R^2(s)\rangle$, and tangent vector correlation, $\langle \cos\theta\,(s)\rangle$, were compared with the predictions of the worm-like chain (WLC) model for polymers equilibrated in two dimensions (2D):

$$\langle R^2(s)\rangle = 4sp\left[1 - \frac{2p}{s}\left(1 - e^{-\frac{s}{2p}}\right)\right] \tag{1}$$

$$\langle \cos\theta\,(s)\rangle = e^{-\frac{s}{2p}}. \tag{2}$$

The validity of the SmarTrace chain-tracing algorithm and analysis approaches was established by testing their performance on AFM images of DNA and on simulated chains (see SI and Figures S1-S3).

We first investigated whether collagen molecules are equilibrated on the mica and are well described by the worm-like chain model. For type I collagen from rat deposited from 100 mM KCl + 1 mM HCl, both mean-squared end-to-end distance $\langle R^2(s)\rangle$ and tangent vector correlation $\langle \cos\theta\,(s)\rangle$ are well described by the WLC model (Figure 2A-B). The agreement between persistence lengths from the fits to equations (1) and (2), $p = 116 \pm 3$ nm and $p = 107 \pm 6$ nm respectively, further suggest that collagen deposited from this solution condition behaves as an equilibrated 2D WLC. This assumption is corroborated by Gaussian distributions of bending angles (Figure 2C). The kurtosis of the bending angle distribution is approximately 3 (the value characteristic of a normal distribution) at all contour lengths (Figure 2D). These measures all indicate that the conformations of collagen on mica represent a sample

equilibrated in two dimensions, and that rat type I collagen has a persistence length of approximately 110 nm.

The persistence lengths were similarly determined for the other collagen samples (Figure 2E). The data and fits are shown in Figure S4, while persistence lengths are presented in Figure 2F. Table 2 provides a summary of the fit parameters, representing an average of the $\langle R^2(s)\rangle$- and $\langle \cos\theta\,(s)\rangle$-derived results; the complete summary of parameters from each fit is given in Table S1, along with reduced χ^2 values used to assess goodness of fit. All four collagen samples exhibit similar persistence lengths, within the range of approximately 80-120 nm. We find no significant differences in flexibility between tissue-derived (rat I, human II) and yeast-expressed recombinant (human I, III) collagens. We observe slightly greater flexibility for homotrimeric (II, III) collagen versus heterotrimeric (I) collagen, similar to previous comparisons between tissue-derived types I and III (16). However, these differences are small in the context of the wide range of values in the literature (Table 1). Thus, differences in type, source and extent of posttranslational modifications do not appear to cause substantial differences in molecular flexibility.

Influences of salt and pH

Given a previous report of significantly different collagen flexibility in buffers with high salt concentration and neutral pH, compared with low ionic-strength, acidic solutions (11), we next investigated how pH and salt concentration independently impact the flexibility of collagen. Rat type I collagen was deposited onto mica from a range of different solution conditions in a controlled ionic environment. In one set of experiments, solutions contained only varying concentrations of potassium chloride, from 0 mM (water) to 100 mM KCl. In a second set of experiments, solutions again contained varying concentrations of KCl, but were acidified to pH ≈ 3 by 1 mM HCl.

Images of collagen at different KCl concentrations and pH are shown in Figure 3A. It is immediately apparent that salt affects the conformations of collagen: the protein transitions from more compact structures at low ionic strength to much more extended conformations at high ionic strength. A similar trend is seen in the acidic conditions. Conformations were analysed to obtain mean-squared end-to-end distances and tangent correlations as a function of contour length, and were fit to equations (1) and (2), respectively. This analysis finds persistence length to increase significantly with ionic strength, at both neutral and acidic pH (Figure 3B and Table 2). Our results corroborate and elaborate on the previous findings of increased flexibility at low pH and salt compared with neutral pH and high salt (11).

The dependence of persistence length on salt concentration is striking. In neutral solutions, the determined persistence length exhibits a four-fold increase from its value in water as the KCl concentration is increased to 100 mM. In acidic solutions, persistence lengths are shorter but still increase, by about three-fold as the concentration of KCl increases from 1 mM to 100 mM. Intriguingly, the range of persistence lengths exhibited here spans almost the entire range of values reported, and unreconciled, in the literature (Table 1).

The trend of increasing stiffness with increasing ionic strength is unexpected, as it is opposite to predictions for polyelectrolytes and to observations for biopolymers such as DNA (20). However, examination of the data and fits in some of the solution conditions, particularly at lower salt

concentrations, reveals that the experimental data are not well described by the standard worm-like chain model (e.g. Figure 4; see also Figures S5, S6). For $\langle R^2(s)\rangle$, the agreement between model and data is reasonable at shorter segment lengths (e.g. s < 100 nm), but longer segments exhibit shorter-than-expected end-to-end distances. The disagreement with the standard WLC model is even more apparent when considering the $\langle \cos\theta\,(s)\rangle$ behaviour. There appears to be an oscillatory modulation of the tangent vector correlation, whereby chains maintain directionality at small segment lengths, reorient substantially at intermediate lengths, and in many cases, exhibit anticorrelated tangent vectors at longer contour lengths. These are not features of a well equilibrated WLC. Instead, the oscillatory character of the tangent vector decorrelation suggests that curvature may play a role in the observed conformations of collagen.

Curved worm-like chain

To describe an intrinsically curved worm-like chain (cWLC), we assume that the lowest-energy conformation of the chain is curved (bent), rather than straight. The extent of angular fluctuations about this curved state is determined by the angular bending potential and the thermal energy of the system, which relate to persistence length as for the standard worm-like chain. From this cWLC model, equations for the mean squared end-to-end distance and mean tangent vector correlation can be derived (see Supporting Text 3):

$$\langle R^2(s)\rangle_c = \frac{4sp}{(1+4{\kappa_0}^2p^2)^2}\left\{1-\frac{2p}{s}(1-4{\kappa_0}^2p^2)\left[1-\cos(\kappa_0 s)\,e^{-\frac{s}{2p}}\right]+\frac{4\kappa_0 p^2}{s}\left[\kappa_0 s-2\sin(\kappa_0 s)\,e^{-\frac{s}{2p}}\right]\right\}, \tag{3}$$

$$\langle \cos\theta\,(s)\rangle_c = \cos(\kappa_0 s)\,e^{-\frac{s}{2p}}\,. \tag{4}$$

Here $\kappa_o$represents the inherent curvature of the chain, which is the inverse of its inherent radius of curvature: $\kappa_o = R_o^{-1}$. Equations (3) and (4) describe the expected behavior of an intrinsically curved, inextensible worm-like chain equilibrated in two dimensions, and are two-dimensional versions of 3D results derived previously (28). Of note, the cWLC model is distinct from approaches that assume a random local curvature (25): in the cWLC model the lowest-energy conformation of the chain is globally curved, *i.e.*, bends always in the same direction.

Figures 4A and 4B show fits of the curved WLC model to the $\langle R^2(s)\rangle$ and $\langle \cos\theta\,(s)\rangle$ data from rat type I collagen deposited from 1 mM HCl. The cWLC captures trends in both $\langle R^2(s)\rangle$ and $\langle \cos\theta\,(s)\rangle$ significantly better than the standard model (χ^2_{red} for $\langle R^2(s)\rangle$: 18 (cWLC) vs 423 (WLC); for $\langle \cos\theta\,(s)\rangle$: 4.4 (cWLC) vs 68 (WLC)). Notably, the cWLC model captures the longer-length anticorrelation of the tangent vectors ($\langle \cos\theta\rangle < 0$), which under the standard WLC model is unphysical. Additionally, the cWLC fits provide a distinct interpretation of the compact structures observed at low ionic strength (Figure 3): rather than being more flexible at lower salt concentration, collagen is instead more curved.

Salt-dependent trends in persistence length and curvature obtained from the curved WLC model are shown in Figures 4C and 4D. In contrast to the monotonic increase in persistence length with increasing KCl concentrations found when using the standard WLC model (Figure 3), the results from the cWLC fits indicate that persistence length varies only modestly over the range of concentrations used here (Figure

4C). Instead, the curvature depends strongly on ionic strength, decreasing from $\kappa_0 = 0.02$ nm^{-1} to 0 (*i.e.*, straightening) as the concentration of KCl is increased (Figure 4D). This effect is shown schematically in Figure 4E.

Origins of collagen curvature

To determine whether the induction of this curvature is ion-specific, images of rat type I collagen deposited from 20 mM acetic acid were collected and analyzed with the standard and curved WLC models (Figure 5A-D). Conformations of collagen deposited from acetic acid are better described by the cWLC model, consistent with observations at low ionic strengths of KCl. Again, naïve application of the standard WLC model results in persistence lengths smaller than those obtained from cWLC fits. To test for equilibration of the samples, we considered the distribution of bend angles. Because of chain curvature, these distributions should not be represented by a single Gaussian centered at $\theta(s) = 0$. Rather, the distributions are expected to be described by two Gaussians, centered symmetrically about zero at an angle $\theta(s) = \pm\kappa_0 s$ and with standard deviation $\sigma(s) = \sqrt{s/p}$ (equation (S20)). The experimental angular distribution is indeed well described by this bimodal function (Figure 5D). The agreement between predicted and measured angular distributions is strong evidence of collagen equilibration on the mica surface in this low-salt condition. The fit parameters in acetic acid ($p = 50$ nm, $\kappa_o = 0.018$ nm^{-1}) are similar to those from the 1 mM HCl solution (Table 2), which has comparable ionic strength and pH. Thus, there does not seem to be a strong effect of chloride versus acetate anions on collagen's mechanical properties.

It is possible that the observed curvature is an effect specific to heterotrimeric collagen. Since type I collagen has one α-chain which is different than the other two, a larger or smaller propensity for surface interactions with this chain could produce a surface-induced curvature to minimize its energy (29). To address this possibility, types I, II and III human collagens were also imaged in 20 mM acetic acid and analyzed with the curved WLC model (Figure S7). The results of this analysis are shown in Figures 5E and 5F. Although the different types vary somewhat in their persistence lengths, their curvatures are nearly identical under these solution conditions, regardless of their trimeric identity. Thus, surface interactions with a unique chain in heterotrimeric collagen are not responsible for the observed curvature.

The curvature observed at low ionic strengths may nonetheless result from molecular interactions with the surface. Intrinsically straight, chiral semiflexible chains may adopt curved conformations at an interface (29). The extent to which the chain curves at an interface will be determined by a competition between the free energy of surface-molecular interactions and the torsional twist energy of the chain. Specifically, curvature in this model requires a chiral organization of sites on the chain that interact differentially at the interface compared with the rest of the chain sites. In the current context, these could be the side-chains of the three individual α-chains of collagen, which are arranged in a right-handed helical structure about the central axis of the collagen molecule. If these sites interact strongly with the mica surface, the free energy gained by adhering to the surface could outweigh the energetic cost of altering the twist (local helical pitch) of the triple helix. In this scenario, the observed trends in curvature with salt concentration (Figure 4F) imply a salt dependence of collagen's torsional stiffness and/or its interactions with the surface. Salt-dependent interactions with mica are well established for other biopolymers (e.g. (27)) and could be expected to contribute here (30). In particular, K^+ affects the

interactions of self-assembling collagen with a mica surface (30-32). However, KCl also affects collagen stability in solution: the thermal stability decreases at lower concentrations of KCl (33). If interactions with the surface contribute to the observed low-ionic-strength curvature, destabilization of the triple helix at low salt concentration – seen also by decreased circular dichroism of the triple helix (33) – is consistent with a reduced torsional stiffness of collagen.

In addition to the ionic-strength dependence of curvature, our images and conformational analysis also demonstrate a dependence on pH: collagen appears more curved in more acidic conditions (Figure 4F). At a pH of 3, collagen carries a net positive charge. At neutral pH, collagen instead has a roughly equal balance of positively and negatively charged residues. The increased curvature at acidic pH could result from stronger electrostatic affinity to the negatively charged mica surface (30). Alternatively, electrostatic repulsion between α-chains could destabilize the triple helix and reduce its torsional stiffness, particularly at low ionic strength when shielding from co-solute anions is weak (at 1 mM ionic strength the Debye length ≈ 10 nm). At neutral pH, the potential for salt bridge formation could strengthen attractions between α-chains (33, 34), thereby increasing the torsional stiffness of the triple helix.

It is also possible that collagen itself is intrinsically bent, even in solution (35). Certainly within the fibrillar superstructure, individual collagen molecules display conformations that are locally bent (36). Studying how salt influences conformations in solution would be key to discriminating which interactions are surface-specific in our study. A change in torsional stiffness and/or alteration of preferred twist angle at lower ionic strength would be expected to alter the average shape of collagen in solution. Light-scattering or other solution-based techniques could be used to study how collagen's compactness in solution depends on salt concentration, and thereby determine how the triple helix itself responds to changes in ionic strength.

Comparison with other estimates of collagen flexibility

Taken together, our results indicate that collagen should be regarded as a semiflexible polymer at room temperature, with a persistence length in the range of 80-100 nm. This is more flexible than solution-based estimates, yet more rigid than found in almost all previous single-molecule approaches (Table 1). Inherent biases and assumptions built into data interpretation within each methodology may contribute to the range of results reported in the literature. Here, we describe possible reasons for the discrepancies.

With AFM imaging, one must play particular heed to molecule-surface interactions and to the possibility of kinetically trapping polymers in non-equilibrium configurations on the surface. In our measurements, the observed conformations appear well equilibrated in two dimensions (Figure 2C,D and Figure 5D). Furthermore, while electrostatic interactions with mica may play a role in the apparent curvature of collagen on the surface, we have found that the persistence length is not strongly affected.

In single-molecule stretching experiments, one source of bias is the underestimation of persistence length when the contour length $L < 20p$ (37), as is the case for collagen. Any structural changes in the triple helix that occur during stretching may also modify the bending energy: such changes have been

inferred from enzymatic cleavage assays of collagen under force (19, 38-40), and may contribute to softening of collagen's force-extension response (12).

Many of the simulations of collagen have used steered molecular dynamics, which may not investigate its equilibrium flexibility. Additionally, the results of molecular dynamics simulations depend on the force fields used (41), which have been developed predominantly for globular proteins rather than the unique fold of collagen.

Interpretations of light-scattering data require a model (such as an ellipsoidal scatterer (17)), and it is possible to describe decay curves equally well assuming both compact and more extended molecular geometries (42). Furthermore, the effective size of the protein, as deduced from solution-based studies such as light-scattering and rheology, is often affected by intermolecular interactions; thus, concentration-dependent studies are essential in order to determine the spatial extent of a given molecule in isolation (43). Given collagen's propensity to self-associate (44, 45), intermolecular interactions may contribute to larger estimates of collagen's size in solution. The rheological studies of reference (7) found that pH does not have a significant effect on collagen's flexibility; however, to reach this conclusion, the authors had to assume that type I collagen exhibited a non-uniform flexibility at neutral pH, with ends considerably more flexible than the central region of the triple helix.

The two studies most similar to ours, which used AFM imaging (11) and electron microscopy (EM) (16) to analyse collagen's flexibility, based their estimates of persistence length on the distance between the endpoints of the collagen chain. As we have shown (Figures 4 and 5), naïve application of equation (1) can lead to significant underestimation of persistence length for curved chains. Consistently, our estimates of collagen's persistence length are greater than theirs obtained at low ionic strength. At high ionic strength and neutral pH, our results and those of the previous AFM imaging study (11) are in reasonable agreement: in this condition, collagen is well described by the standard WLC model and curvature may be disregarded. In general, however, our findings demonstrate the importance of mapping the functional form of $\langle R^2(s)\rangle$ and $\langle \cos\theta\ (s)\rangle$, and caution against determining persistence length by using only chain end-point separations.

In our description and analysis of collagen, we have made two major assumptions. First, neither the WLC nor cWLC model takes into account the helical nature of collagen. Thus, flexibility of the chain derives only from bending deformations. If the intrinsic curvature described by our model arises due to a surface-collagen interaction inducing altered twist in the protein, then torsional stiffness of the triple helix (46) is an important parameter, and twist-bend coupling (47) may be required to fully describe collagen's flexibility.

Second, our analysis procedures treat collagen as an apolar, homogenous worm-like chain. We are not able to discriminate between N→C and C→N directionality in the imaged chains. Furthermore, persistence lengths and curvatures reported in this work represent average values for the entire collagen sequence. In adopting this approach, we have implicitly assumed that the $(\text{Gly-X-Y})_n$ triple-helix-forming motif dominates collagen's response, with local sequence variations (e.g. fraction of prolines occupying the X and Y positions) viewed as introducing only minor perturbations. The flexibility of fibrillar collagen triple helices appears approximately uniform in EM images (16), supporting this assumption. However, molecular dynamics simulations (48) and modelling of atomistic collagen structures (46) suggest that collagen possesses a sequence-dependent flexibility, commensurate with

the expected variations in pitch observed for different triple helical sequences (1, 36). It is interesting to speculate how a sequence-dependent mechanical signature may provide additional physical cues to interaction partners (40). Experimentally elucidating the detailed sequence-dependent flexibility of collagen is an exciting future prospect that will require higher-precision algorithms than applied thus far to this mechanically important protein.

CONCLUSIONS

Using atomic force microscopy, we determined the persistence length of different types and sources of collagen in the presence of different co-solutes (K^+, Cl^-, acetate). Although collagens of different type and source exhibited only minor differences in persistence length, our initial findings implied that ionic strength and, to a lesser extent, pH modulate collagen's flexibility. Closer inspection of these results, however, revealed this interpretation to be an artefact of the model used for analysis. Rather than behaving as an intrinsically straight worm-like chain, collagen deposited in low salt conditions was found to adopt preferentially curved conformations on the mica surface. Statistical analysis of chain conformations demonstrates that this curvature is a thermodynamic property, with chains appearing well equilibrated in two dimensions. The extent of curvature depends strongly on salt concentration and pH, with persistence length remaining roughly constant across all conditions.

Our findings provide a possible explanation for the range of persistence lengths reported for collagen in the literature: rather than a direct modulation of its flexibility, the presence of certain co-solutes may induce curvature along the collagen backbone. Whether or not the curvature seen in our results is intrinsic to collagen or is caused by co-solute-dependent interactions with the imaging surface remains to be determined. If intrinsic, then the modulation of collagen's conformation by salt and pH has implications for its cellular processing and for its assembly into higher-order structures.

The strong agreement between measures of chain flexibility and the predictions of the curved worm-like chain model suggest that this formalism may be useful for interpreting the mechanical properties of other biological filaments. Examples such as amyloid fibrils (29) and coiled-coil proteins (49) have been found to exhibit preferentially curved conformations at interfaces; applying the cWLC model to describe these systems and others may provide further insight into their equilibrium properties and the energies governing their mechanical properties.

MATERIALS AND METHODS

Collagen Sources

Recombinant human type I (RhC1-003) and recombinant human type III (RhC3-012) collagen were expressed in yeast and obtained from FibroGen (generous gifts of Alexander Dunn, Stanford University), rat tail tendon-derived type I collagen (3440-100-01) was purchased from Cultrex, and human cartilage-derived type II collagen (CC052) was purchased from EMD Millipore. All stocks are between 2 and 5 mg/ml in 20 mM acetic acid.

Sample Preparation

The desired solution conditions were obtained by solution exchange using Millipore Amicon Ultra-0.5 spin filters (NMWL 50 kDa, UFC505096), then dilution to approximately 1 µg/ml collagen prior to deposition. 50 µl was deposited onto freshly cleaved mica (Highest Grade V1 AFM Mica Discs, 10 mm, Ted Pella) for 20 seconds. After deposition, samples were rinsed five times with 1 ml ultrapure water to remove unbound proteins, and the mica was dried under a flow of filtered compressed air. It is important to note that all collagen molecules were imaged in these dry conditions. Thus, solution conditions quoted refer to the condition under which collagen was deposited onto mica, at room temperature.

Atomic Force Microscopy Imaging

Images of collagen adsorbed to mica were collected with an Asylum Research MFP-3D atomic force microscope using AC tapping mode in air. AFM tips with a 325 kHz resonance frequency and 40 N/m force constant (MikroMasch, HQ:NSC15/AL BS) were used for image collection, and were changed as necessary to preserve image quality.

Chain Tracing

The SmarTrace algorithm used to trace collagen chains from AFM images was developed in MATLAB (50) and uses a graphical user interface adapted from (26). It is available from the authors upon request. A detailed description of the SmarTrace workflow and validation can be found in the Supporting Information. Briefly, the algorithm uses cross-correlation analysis of a template cross-section with the imaged chain in order to identify its center line and width. Continuity constraints on the local directionality and width enable the tracing of chains in noisy environments.

Data Analysis

Traced chains were sampled using bootstrapping to extract statistics of the chain for flexibility determination. Following the method of Faas *et al.* (25), each chain was randomly divided into non-overlapping segments of lengths drawn from a set of input values (here, $s = 10$ nm, 20 nm, 30 nm, ... , 200 nm). The maximum segment length does not need to be the same as the chain contour length, so it allows the use of partially traced chains. This is particularly useful when ends of a molecule are not clear or chains intersect, as it allows subsections of the chain to be included in the analysis. $\langle R^2(s)\rangle$ and $\langle \cos\theta(s)\rangle$ were determined from all segments of each length s.

Resultant $\langle R^2(s)\rangle$ and $\langle \cos\theta(s)\rangle$ were fit to equations derived from the inextensible worm-like chain (WLC) model with (equations 3-4) or without (equations 1-2) intrinsic curvature. Derivations for equations 3 and 4 are presented in the Supporting Information.

The kurtosis of each angular distribution was calculated at each segment length s as $n[\sum_i(\theta_i - \langle\theta\rangle)^4]/[\sum_i(\theta_i - \langle\theta\rangle)^2]^2$, for a distribution across n angle bins θ_i. The standard error of the kurtosis was determined using Eq. S19.

Values for p and κ_o presented in Table 2 are an average of the results from independent fits to $\langle R^2(s)\rangle$ and to $\langle\cos\theta(s)\rangle$. Reported errors, Δ, represent the propagated error of the 95% confidence intervals of the respective fit parameters or half of the difference between the $\langle R^2(s)\rangle$ and $\langle\cos\theta(s)\rangle$ fit parameters, whichever is larger. A full list of fitting parameters from all models and samples is provided in Table S1. The total length of collagen chains traced in each sample condition is also included in Table S1. All data sets are available from the authors upon request.

Validation Tests

The performance of the SmarTrace algorithm and data analysis code was evaluated by tests on simulated polymers, and on images of DNA. Details are provided in the supporting information.

ACKNOWLEDGMENTS

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC). We gratefully acknowledge Patricia Mooney's gift of the atomic force microscope and ChangMin Kim for technical assistance. We thank Alex Dunn for collagen samples, Alaa Al-Shaer for testing robustness of the SmarTrace algorithm, Peter Olmsted for introducing to us the concept of surface-induced fiber curvature, David Sivak for a critical review of the manuscript, the Forde lab for many useful discussions, and Guillaume Lamour for the EasyWorm code, DNA sample images and an introduction to AFM sample preparation.

AUTHOR CONTRIBUTIONS

NR and NRF designed the project. NR and AL performed experiments. NR developed the SmarTrace algorithm and software. AL performed molecular simulations and derived the cWLC fitting equations. NR and AL analysed the data. All authors contributed to the writing of the manuscript.

REFERENCES

1. Shoulders MD & Raines RT (2009) Collagen Structure and Stability. *Annual Review of Biochemistry* 78(1):929-958.
2. Marini JC*, et al.* (2007) Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: Regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Human Mutation* 28(3):209-221.
3. Byers PH & Murray ML (2012) Heritable Collagen Disorders: The Paradigm of the Ehlers—Danlos Syndrome. *Journal of Investigative Dermatology* 132:E6-E11.
4. Myllyharju J & Kivirikko KI (2001) Collagens and collagen-related diseases. *Annals of Medicine* 33(1):7-21.
5. Utiyama H, Sakato K, Ikehara K, Setsuiye T, & Kurata M (1973) Flexibility Of Tropocollagen From Sedimentation And Viscosity. *Biopolymers* 12(1):53-64.
6. Nestler FHM, Hvidt S, Ferry JD, & Veis A (1983) Flexibility Of Collagen Determined From Dilute-Solution Viscoelastic Measurements. *Biopolymers* 22(7):1747-1758.
7. Amis EJ, Carriere CJ, Ferry JD, & Veis A (1985) Effect of pH on collagen flexibility determined from dilute solution viscoelastic measurements. *International Journal of Biological Macromolecules* 7(3):130-134.
8. Alberts B*, et al.* (2007) *Molecular Biology of the Cell* (Garland Science, New York, NY) 5th Ed.
9. Sun YL, Luo ZP, Fertala A, & An KN (2002) Direct quantification of the flexibility of type I collagen monomer. *Biochemical and Biophysical Research Communications* 295(2):382-386.
10. Sun YL, Luo ZP, Fertala A, & An KN (2004) Stretching type II collagen with optical tweezers. *Journal of Biomechanics* 37:1665-1669.
11. Lovelady HH, Shashidhara S, & Matthews WG (2014) Solvent specific persistence length of molecular type I collagen. *Biopolymers* 101(4):329-335.
12. Wieczorek A*, et al.* (2015) Development and characterization of a eukaryotic expression system for human type II procollagen. *BMC Biotechnology* 15(1):112.
13. Buehler MJ & Wong SY (2007) Entropic Elasticity Controls Nanomechanics of Single Tropocollagen Molecules. *Biophys. J.* 93(1):37-43.
14. Chang S-W, Shefelbine Sandra J, & Buehler Markus J (2012) Structural and Mechanical Differences between Collagen Homo- and Heterotrimers: Relevance for the Molecular Origin of Brittle Bone Disease. *Biophysical Journal* 102(3):640-648.
15. Varma S, Orgel Joseph PRO, & Schieber Jay D (2016) Nanomechanics of Type I Collagen. *Biophysical Journal* 111(1):50-56.
16. Hofmann H, Voss T, Kühn K, & Engel J (1984) Localization of Flexible Sites in Thread-like Molecules from Electron Micrographs - Comparison of Interstitial, Basement Membrane and Intima Collagens. *Journal Of Molecular Biology* 172(3):325-343.
17. Claire K & Pecora R (1997) Translational and rotational dynamics of collagen in dilute solution. *Journal of Physical Chemistry B* 101(5):746-753.
18. Vesentini S, Redaelli A, & Gautieri A (2013) Nanomechanics of collagen microfibrils. *Muscles Ligaments Tendons J* 3(1):23-34.
19. Adhikari AS, Glassey E, & Dunn AR (2012) Conformational Dynamics Accompanying the Proteolytic Degradation of Trimeric Collagen I by Collagenases. *J. Am. Chem. Soc.* 134(32):13259-13265.
20. Baumann CG, Smith SB, Bloomfield VA, & Bustamante C (1997) Ionic effects on the elasticity of single DNA molecules. *Proceedings Of The National Academy Of Sciences Of The United States Of America* 94(12):6185-6190.

21. Canty EG & Kadler KE (2005) Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci* 118(7):1341-1353.
22. Demaurex N (2002) pH Homeostasis of Cellular Organelles. *Physiology* 17(1):1-5.
23. Usov I & Mezzenga R (2015) FiberApp: An Open-Source Software for Tracking and Analyzing Polymers, Filaments, Biomacromolecules, and Fibrous Objects. *Macromolecules* 48(5):1269-1280.
24. Báez J, Olsen D, & Polarek JW (2005) Recombinant microbial systems for the production of human collagen and gelatin. *Appl Microbiol Biotechnol* 69(3):245-252.
25. Faas FGA, Rieger B, van Vliet LJ, & Cherny DI (2009) DNA Deformations near Charged Surfaces: Electron and Atomic Force Microscopy Views. *Biophysical Journal* 97(4):1148-1157.
26. Lamour G, Kirkegaard J, Li H, Knowles T, & Gsponer J (2014) Easyworm: an open-source software tool to determine the mechanical properties of worm-like chains. *Source Code for Biology and Medicine* 9(1):16.
27. Rivetti C, Guthold M, & Bustamante C (1996) Scanning Force Microscopy of DNA Deposited onto Mica: Equilibration versus Kinetic Trapping Studied by Statistical Polymer Chain Analysis. *Journal of Molecular Biology* 264(5):919-932.
28. Craig A & Terentjev EM (2006) Auxiliary Field Theory of Polymers with Intrinsic Curvature. *Macromolecules* 39(13):4557-4565.
29. Jordens S*, et al.* (2014) Adsorption at Liquid Interfaces Induces Amyloid Fibril Bending and Ring Formation. *ACS Nano* 8(11):11071-11079.
30. Narayanan B, Gilmer GH, Tao J, De Yoreo JJ, & Ciobanu CV (2014) Self-Assembly of Collagen on Flat Surfaces: The Interplay of Collagen–Collagen and Collagen–Substrate Interactions. *Langmuir* 30(5):1343-1350.
31. Jiang FZ, Horber H, Howard J, & Muller DJ (2004) Assembly of collagen into microribbons: effects of pH and electrolytes. *J. Struct. Biol.* 148(3):268-278.
32. Leow WW & Hwang W (2011) Epitaxially Guided Assembly of Collagen Layers on Mica Surfaces. *Langmuir* 27(17):10907-10913.
33. Freudenberg U*, et al.* (2007) Electrostatic Interactions Modulate the Conformation of Collagen I. *Biophysical Journal* 92(6):2108-2119.
34. Keshwani N, Banerjee S, Brodsky B, & Makhatadze George I (2013) The Role of Cross-Chain Ionic Interactions for the Stability of Collagen Model Peptides. *Biophysical Journal* 105(7):1681-1688.
35. Walker KT*, et al.* (2017) Non-linearity of the collagen triple helix in solution and implications for collagen function. *Biochemical Journal* 474(13):2203-2217.
36. Orgel JPRO, Irving TC, Miller A, & Wess TJ (2006) Microfibrillar structure of type I collagen in situ. *Proceedings of the National Academy of Sciences* 103(24):9001-9005.
37. Seol Y, Li J, Nelson PC, Perkins TT, & Betterton MD (2007) Elasticity of Short DNA Molecules: Theory and Experiment for Contour Lengths of 0.6 7 {micro}m. *Biophys. J.* 93(12):4360-4373.
38. Adhikari AS, Chai J, & Dunn AR (2011) Mechanical Load Induces a 100-Fold Increase in the Rate of Collagen Proteolysis by MMP-1. *J Am Chem Soc* 133(6):1686.
39. Camp RJ*, et al.* (2011) Molecular Mechanochemistry: Low Force Switch Slows Enzymatic Cleavage of Human Type I Collagen Monomer. *J Am Chem Soc* 133(11):4073.
40. Kirkness MWH & Forde NR (2018) Single-Molecule Assay for Proteolytic Susceptibility: Force-Induced Collagen Destabilization. *Biophysical Journal* 114(3):570-576.
41. Varma S*, et al.* (2015) Effect of intrinsic and extrinsic factors on the simulated D-band length of type I collagen. *Proteins* 83(10):1800-1812.
42. Shayegan M (2014) Determining local viscoelastic properties of collagen systems using optical tweezers. PhD (Simon Fraser University, Burnaby, Canada).

43. Muschol M & Rosenberger F (1995) Interactions in undersaturated and supersaturated lysozyme solutions: Static and dynamic light scattering results. *The Journal of Chemical Physics* 103(24):10424-10432.
44. Shayegan M & Forde NR (2013) Microrheological Characterization of Collagen Systems: From Molecular Solutions to Fibrillar Gels. *PLoS ONE* 8(8):e70590.
45. Shayegan M, Altindal T, Kiefl E, & Forde NR (2016) Intact Telopeptides Enhance Interactions between Collagens. *Biophysical Journal* 111(11):2404-2416.
46. Xu F*, et al.* (2017) Parallels between DNA and collagen – comparing elastic models of the double and triple helix. *Sci Rep* 7(1):12802.
47. Nomidis SK, Kriegel F, Vanderlinden W, Lipfert J, & Carlon E (2017) Twist-Bend Coupling and the Torsional Response of Double-Stranded DNA. *Physical Review Letters* 118(21):217801.
48. Ravikumar KM & Hwang W (2008) Region-specific role of water in collagen unwinding and assembly. *Proteins: Structure, Function, and Bioinformatics* 72(4):1320-1332.
49. Murray DH*, et al.* (2016) An endosomal tether undergoes an entropic collapse to bring vesicles together. *Nature* 537(7618):107-111.
50. MATLAB and Statistics Toolbox Release 2011b, The MathWorks, Inc., Natick, Massachusetts, United States.

Table 1 – Literature estimates of persistence length for fibrillar, molecular collagens.

Persistence length, p (nm)	Collagen type and source	Solution conditions	Method	Reference
10	57-residue segment of homotrimeric mouse type I	water, neutral pH	Molecular dynamics	(14)
12	Bovine dermis type I	water, pH~5	AFM imaging	(11)
13	45-residue segment of rat type I	10 mM NaCl, neutral pH	Molecular dynamics	(15)
11-15	Cell-derived human type I procollagen; recombinant human type II procollagen	>10 mM buffer, neutral pH	Optical tweezers stretching	(9, 10)
16	30-residue homotrimeric peptide	water, neutral pH	Steered molecular dynamics	(13)
22	57-residue segment of heterotrimeric mouse type I	water, neutral pH	Molecular dynamics	(14)
15-65	Recombinant human type II procollagen	>10 mM buffer, neutral pH	Optical tweezers stretching	(12)
40	Fetal bovine type III	50 mM acetic acid + equal volume glycerol	Electron microscopy	(16)
51	60-nm segment of human type I collagen	water, neutral pH	Coarse-grained molecular dynamics	(18)
57	Calf dermis type I	50 mM acetic acid + equal volume glycerol	Electron microscopy	(16)
130	Rat skin type I	>10 mM buffer, pH ~5	Viscometry	(5)
135-165	Bovine dermis type I	>10 mM buffer, neutral pH	AFM imaging	(11)
161	Rat skin type I	0.3 M acetate+ 6 mM NaCl, pH 4	Rheology	(6, 7)
160-165	Bovine dermis type I	1 mM HCl	Dynamic light scattering	(17)
167	Rat skin type I	>10 mM buffer, neutral pH	Rheology	(7)

Table 2 – Summary of fitting parameters determined in this work.

Collagen Type	Solvent Condition	Standard Model Fits		Curved Model Fits			
		p (nm)	Δp (nm)	p (nm)	Δp (nm)	κ_o (nm^{-1})	$\Delta\kappa_o$(nm^{-1})
Rat Type-I	Water	35	8	108	22	0.015	0.002
Rat Type-I	100 µM KCl	30	7	76	18	0.016	0.002
Rat Type-I	1 mM KCl	48	7	102	18	0.0114	0.0009
Rat Type-I	10 mM KCl	93	6	101	14	0.002	0.005
Rat Type-I	100 mM KCl	145	13	144	41	0	N/A
Rat Type-I	1 mM HCl	28	4	49	7	0.015	0.001
Rat Type-I	10 mM KCl + 1 mM HCl	54	7	131	33	0.011	0.001
Rat Type-I	100 mM KCl + 1 mM HCl	112	4	117	14	0.002	0.002
Human Type-I	100 mM KCl + 1 mM HCl	91	4	98	11	0.002	0.002
Human Type-II	100 mM KCl + 1 mM HCl	94	4	96	14	0.0008	N/A
Human Type-III	100 mM KCl + 1 mM HCl	85	5	90	15	0.0018	N/A
Rat Type-I	20 mM Acetic Acid	23	4	51	7	0.018	0.002
Human Type-I	20 mM Acetic Acid	26	4	51	4	0.0166	0.0006
Human Type-II	20 mM Acetic Acid	31	7	76	13	0.0158	0.0008
Human Type-III	20 mM Acetic Acid	28	5	60	4	0.0163	0.0005

Fit parameters are averages of the values from independent $\langle R^2(s)\rangle$ and $\langle \cos\theta\ (s)\rangle$ fits of the data; errors, Δ, represent the error in these estimates. The total contour lengths of collagen traced in each condition is provided along with a comprehensive list of fitting parameters and χ^2 values in Table S1.

FIGURE LEGENDS

Figure 1. Imaging individual collagen proteins. (A) Schematic of triple-helical collagen. (B-E) Representative AFM images of each of the four collagen samples investigated in this work. For each sample, collagen was deposited from a solution of 100 mM KCl + 1 mM HCl. Scale bars = 250 nm. (B) Rat tail-derived type I collagen. (C) Recombinant human type I collagen. (D) Cartilage-derived human type II collagen. (E) Recombinant human type III collagen. (F) SmarTrace starts with user input points along a collagen chain (left image). Initial splines connecting the three input points (dashed lines) do not follow the chain. The SmarTrace algorithm identifies the both centreline of the chain (red) and its width (blue – used for chain tracing but no further analysis), as shown in the right image.

Figure 2. Determination of collagen's persistence length using the standard worm-like chain (WLC) model. Parts (A-D) present analysis for rat type I collagen deposited from a solution of 100 mM KCl + 1 mM HCl. (A) Mean squared end-to-end distance as a function of segment length, $\langle R^2(s)\rangle$. Points represent mean values determined at 10 nm segment-length intervals, with error bars representing standard errors of the mean. The red line is a fit with equation 1, yielding a persistence length of $p =$ 116 ± 3 nm (error represents 95% confidence interval of the fit). (B) Tangent vector correlation as a function of segment length, $\langle \cos\theta(s)\rangle$. Points and errors are similarly represented as in (A). The blue line is a fit with equation 2, yielding $p =$ 107 ± 6 nm. Fits to the WLC model in (A) and (B) capture the trends in the data and yield comparable persistence lengths. (C) Angular histogram for 50 nm segment lengths (blue bars). Error bars represent $\sqrt{N}$ counting error. This distribution agrees well with that expected for a worm-like chain of persistence length $p =$ 112 nm (the average of the results from (A) and (B)), which is a normal distribution with a mean of zero and a variance of s/p (red line). (D) Kurtosis of the angular distributions extracted from the traced collagens at different segment lengths. Error bars represent the standard error in the kurtosis (Eq. S19). The kurtosis of the angular distributions is close to 3 (red line) for all segment lengths, indicating that the collagens behave as equilibrated Gaussian chains on the mica surface. (E) Venn diagram illustrating the similarities and differences between the different collagen samples tested, including trimeric identity and source. (F) Persistence length for each collagen sample deposited from a solution of 100 mM KCl + 1 mM HCl, obtained using $\langle R^2(s)\rangle$ and $\langle \cos\theta(s)\rangle$ analyses. The similarity among these samples suggests that collagen type and source have little impact on the mechanical properties of collagen at the molecular level.

Figure 3. Effects of ionic strength on persistence length. (A) AFM images of rat type I collagen deposited from solutions covering a range of KCl concentration and pH. All scale bars represent 250 nm. As salt concentration increases, the collagen chains appear to straighten. (B) Persistence lengths in these conditions. The reported values are the average of p extracted from fits to $\langle R^2(s)\rangle$ and to $\langle \cos\theta(s)\rangle$, and errors are the larger of the propagated error in this mean or half the separation between the two values of p. For the purposes of graphical representation, water was assumed to have an ionic strength of 10^{-7} M. Increasing the ionic strength causes a large increase in the apparent persistence length, an effect that is reduced slightly in the acidic condition for the same ionic strengths.

Figure 4. At low ionic strength, collagen is better described as a curved worm-like chain. (A) Fits of the mean squared end-to-end distance using both the standard (Eq. 1, dashed red line) and curved (Eq. 3, solid red line) WLC models, to data extracted from rat tail type I collagen deposited from 1 mM HCl. (B) Corresponding fits of the standard (Eq. 2, dashed blue line) and curved (Eq. 4, solid blue line) WLC tangent vector correlation functions. The data are more accurately described by the curved model

equations, as shown statistically by a significant improvement in reduced χ^2 values (see text). The standard WLC underestimates the persistence length of the collagens in this condition, presumably interpreting the induced curvature as additional fluctuation. (C) Persistence lengths for the different co-solute conditions presented in Fig. 3. Rather than the trend with ionic strength seen with the standard WLC fits, there is a less significant, and less obviously monotonic, variation in persistence length as a function of ionic strength. Instead, the observed systematic conformational changes are attributed to a variation in innate curvature, which drops as the ionic strength of the solution is increased (D). For the purposes of graphical representation, water was assumed to have an ionic strength of 10^{-7} M. Error bars on $\langle R^2(s)\rangle$ and $\langle \cos\theta(s)\rangle$ values are as described for Figure 2, and error bars on p and κ_o are obtained as described for Figure 3. (E) Schematic illustrating the transition of collagen from a molecule with a curved backbone at low ionic strength (left) to one with a straight backbone at higher ionic strength (right). The persistence length characterizes fluctuations about this lowest-energy conformation.

Figure 5. Behavior of different collagen types in acetic acid. (A) Representative image of rat tail type I collagen deposited from 20 mM acetic acid. (B) Standard (Eq. 1, dashed red line) and curved (Eq. 3, solid red line) WLC model fit to the mean squared end-to-end distance, $\langle R^2(s)\rangle$. (C) Standard (Eq. 2, dashed blue line) and curved (Eq. 4, solid blue line) WLC model fit to the tangent vector correlation, $\langle \cos\theta(s)\rangle$. Data in both (A) and (B) are described more accurately by the curved WLC model, as seen also by the reduced χ^2 values (Table S1). (D) Absolute-valued angular distribution of the same collagen molecules at a segment length of $s = 50$ nm. The red line represents the expected distribution (Eq. S3) for a curved WLC using a persistence length p = 51 nm and intrinsic curvature κ_o= 0.0182 nm^{-1}, the average values of each parameter extracted from the fits shown in (B) and (C). The dashed red lines are the two Gaussian components which make up this distribution, as discussed in the SI. (E) and (F) Persistence lengths and curvatures, respectively, of different collagen types deposited from 20 mM acetic acid. As with the different collagen types deposited from 100 mM KCl + 1 mM HCl (Figure 2F), there is only modest variation amongst the different samples, demonstrating that, also at lower ionic strength and different co-solute conditions, only minor mechanical differences exist between different collagen types and sources.

Figure 1.

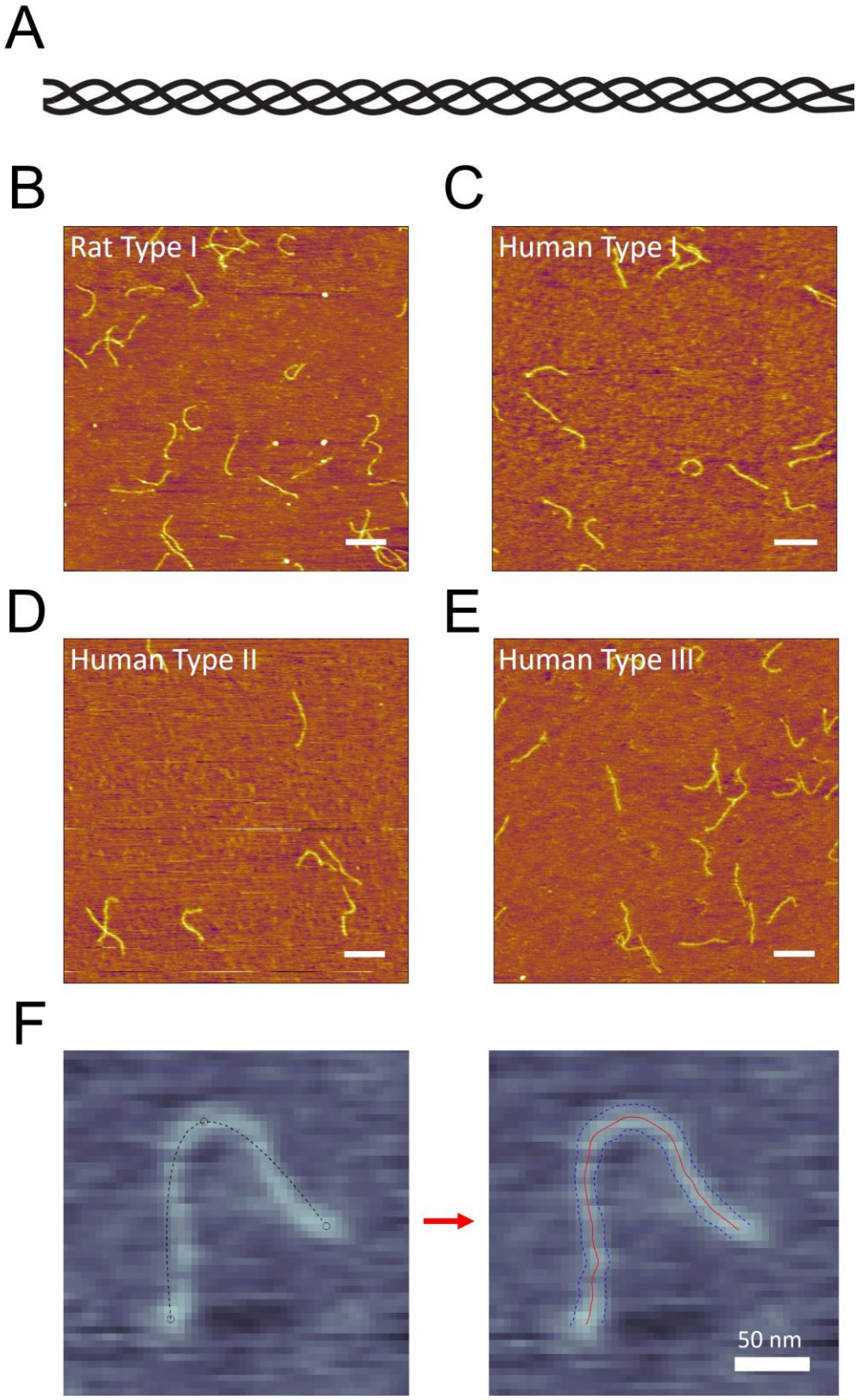

A
B
Rat Type I
C
Human Type I
D
Human Type II
E
Human Type III
F
50 nm

Figure 2.

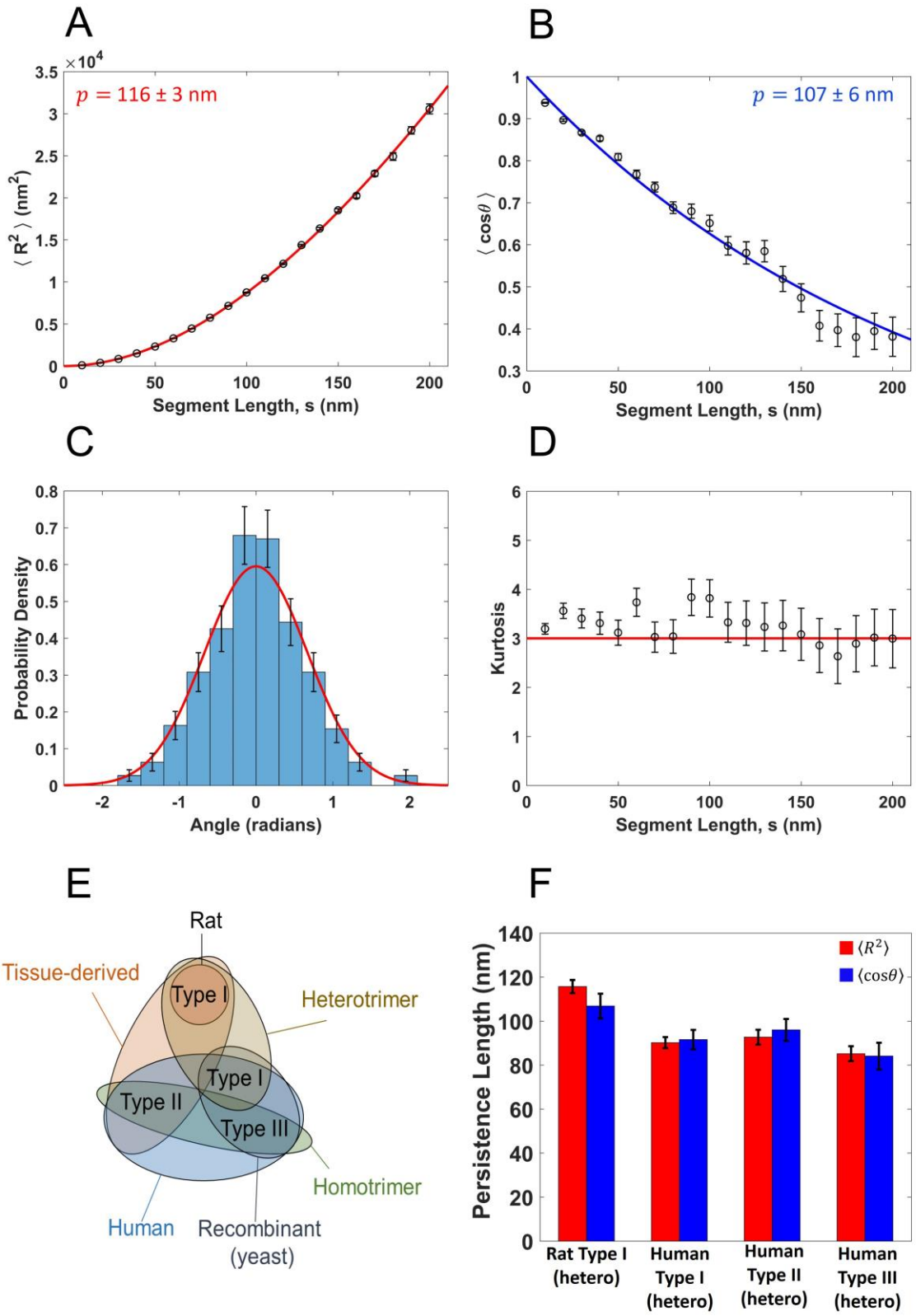

A
$p = 116 \pm 3$ nm
$\langle R^2 \rangle$ (nm^2)
Segment Length, s (nm)
B
$p = 107 \pm 6$ nm
$\langle \cos\theta \rangle$
Segment Length, s (nm)
C
Probability Density
Angle (radians)
D
Kurtosis
Segment Length, s (nm)
E
Rat
Tissue-derived
Type I
Heterotrimer
Type I
Type II
Type III
Homotrimer
Human
Recombinant
(yeast)
F
$\langle R^2 \rangle$
$\langle \cos\theta \rangle$
Persistence Length (nm)
Rat Type I (hetero)
Human Type I (hetero)
Human Type II (hetero)
Human Type III (hetero)

Figure 3.

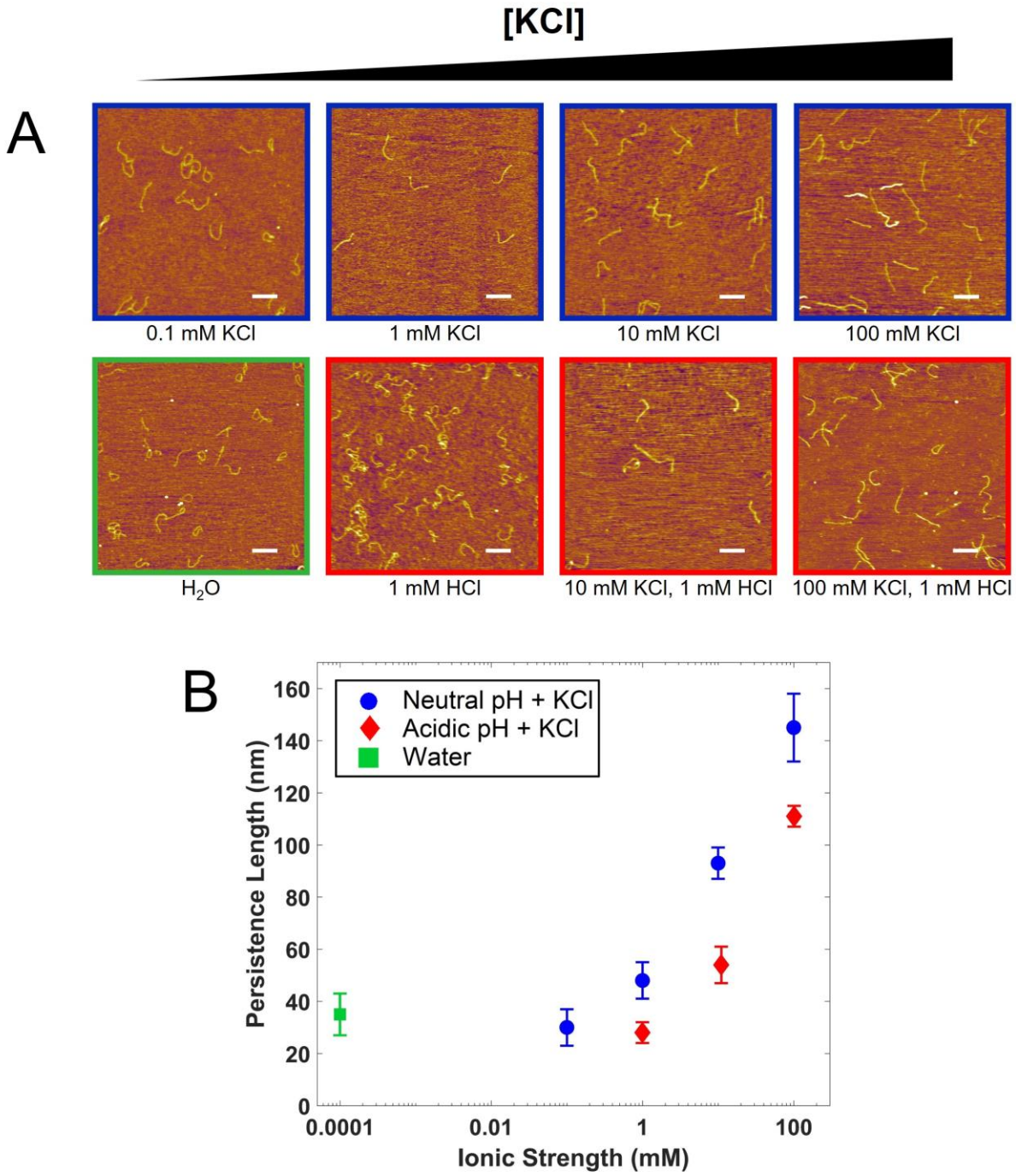

[KCl]
A
0.1 mM KCl
1 mM KCl
10 mM KCl
100 mM KCl
H2O
1 mM HCl
10 mM KCl, 1 mM HCl
100 mM KCl, 1 mM HCl
B
Neutral pH + KCl
Acidic pH + KCl
Water
Persistence Length (nm)
Ionic Strength (mM)
0
20
40
60
80
100
120
140
160
0.0001
0.01
1
100

Figure 4.

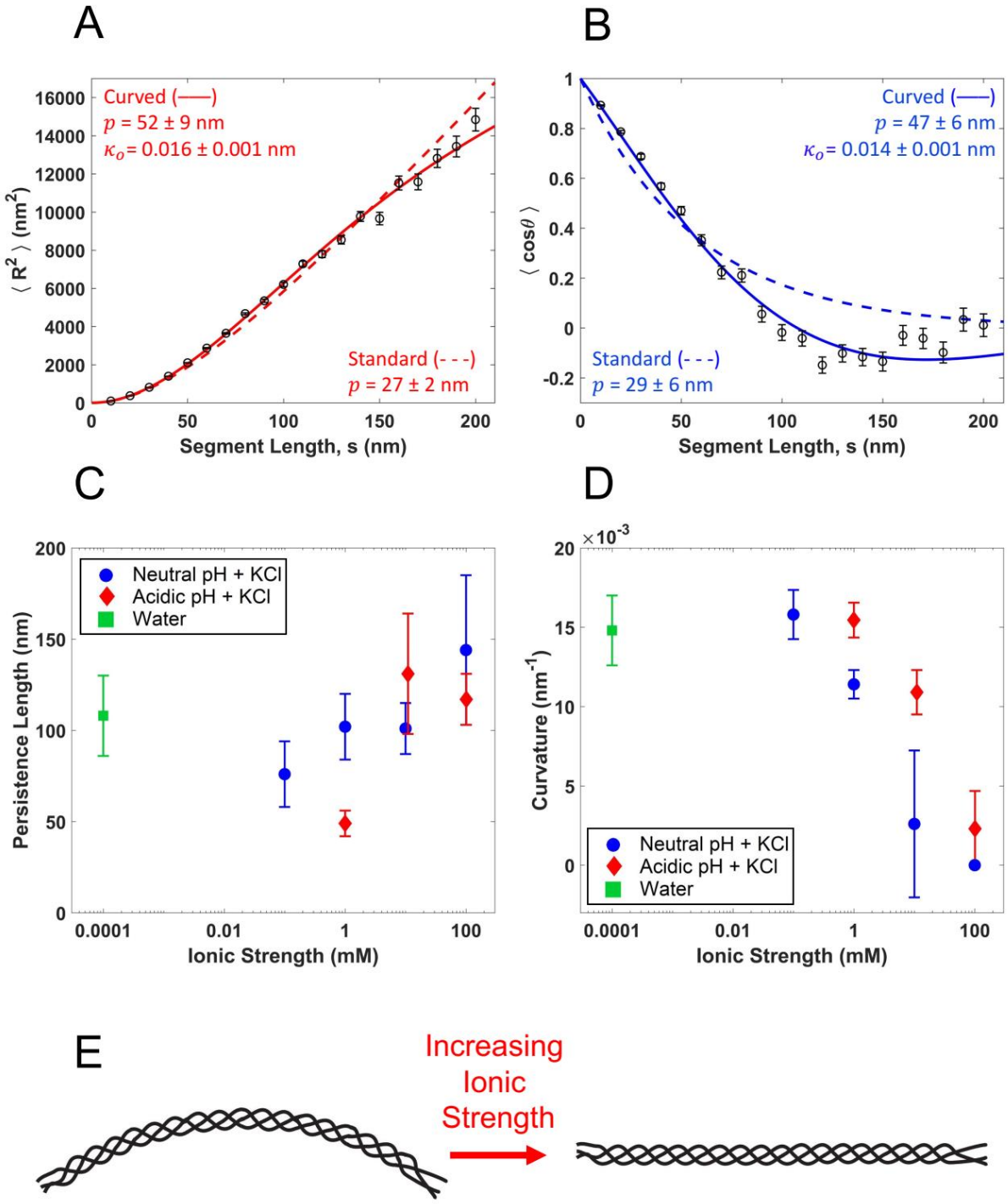
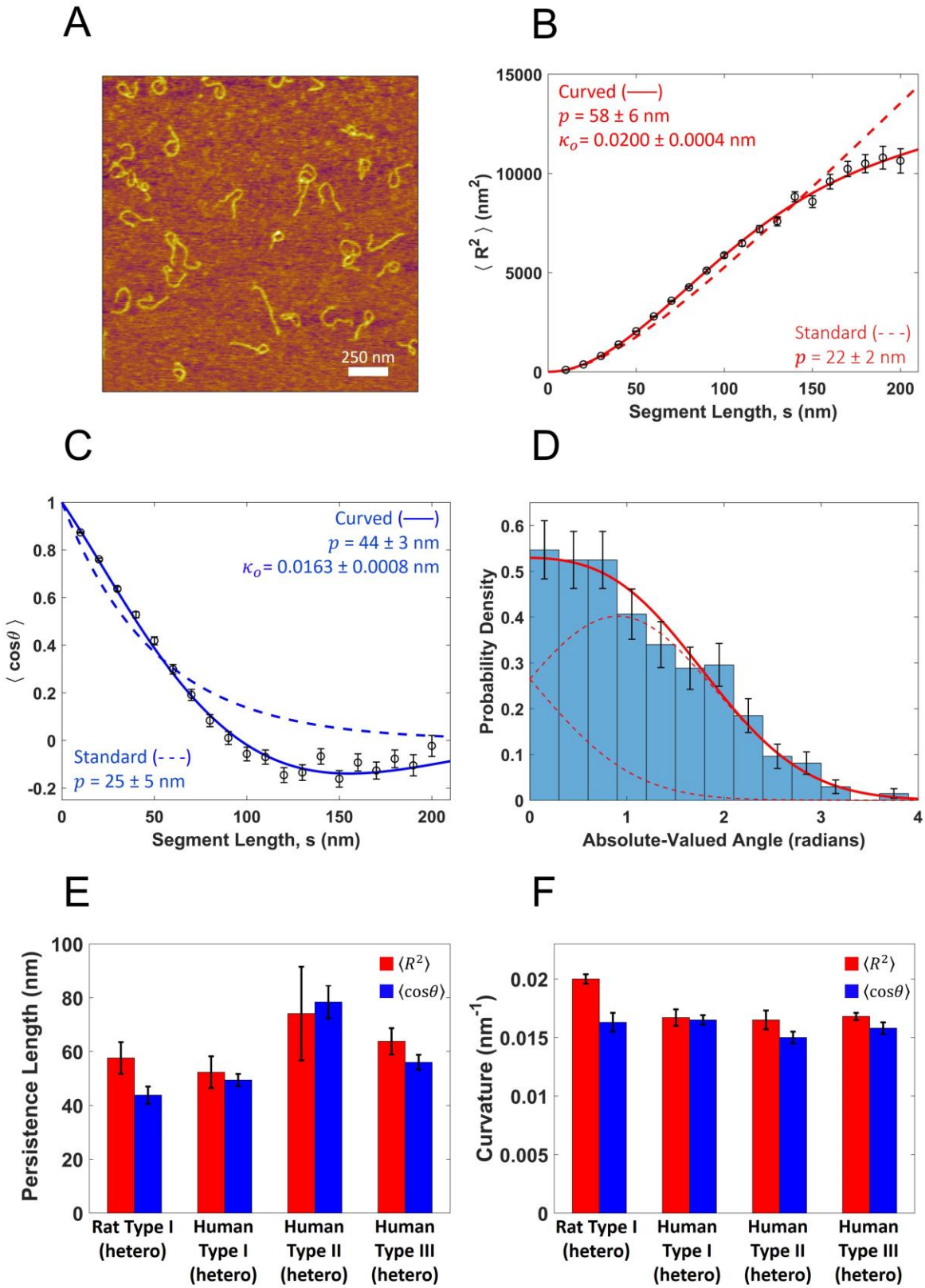

A
Curved (——)
p = 52 ± 9 nm
κ_o = 0.016 ± 0.001 nm
Standard (- - -)
p = 27 ± 2 nm
⟨ R² ⟩ (nm²)
Segment Length, s (nm)
B
Curved (——)
p = 47 ± 6 nm
κ_o = 0.014 ± 0.001 nm
Standard (- - -)
p = 29 ± 6 nm
⟨ cosθ ⟩
Segment Length, s (nm)
C
Neutral pH + KCl
Acidic pH + KCl
Water
Persistence Length (nm)
Ionic Strength (mM)
D
×10⁻³
Curvature (nm⁻¹)
Neutral pH + KCl
Acidic pH + KCl
Water
Ionic Strength (mM)
E
Increasing
Ionic
Strength

Figure 5.


A
250 nm
B
Curved (——)
p = 58 ± 6 nm
κ_o= 0.0200 ± 0.0004 nm
Standard (- - -)
p = 22 ± 2 nm
⟨ R² ⟩ (nm²)
Segment Length, s (nm)
C
Curved (——)
p = 44 ± 3 nm
κ_o= 0.0163 ± 0.0008 nm
Standard (- - -)
p = 25 ± 5 nm
⟨ cosθ ⟩
Segment Length, s (nm)
D
Probability Density
Absolute-Valued Angle (radians)
E
⟨R^2⟩
⟨cosθ⟩
Persistence Length (nm)
Rat Type I (hetero)
Human Type I (hetero)
Human Type II (hetero)
Human Type III (hetero)
F
⟨R^2⟩
⟨cosθ⟩
Curvature (nm⁻¹)
Rat Type I (hetero)
Human Type I (hetero)
Human Type II (hetero)
Human Type III (hetero)