Hidden length lets collagen buffer mechanical and chemical stress

Benedikt Rennekamp,[1,2,3,4] Helmut Grubmüller,[3,4,*] and Frauke Gräter[1,2,3,†]
[1]*Heidelberg Institute for Theoretical Studies, 69118 Heidelberg, Germany*
[2]*Interdisciplinary Center for Scientific Computing, Heidelberg University, 69120 Heidelberg, Germany*
[3]*Max Planck School Matter to Life, 69120 Heidelberg, Germany*
[4]*Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany*

(Received 18 December 2023; revised 3 June 2024; accepted 27 August 2024; published 8 November 2024)

Collagen, the most abundant protein in the human body, must withstand high mechanical loads due to its structural role in tendons, skin, bones, and other connective tissue. It was recently found that tensed collagen creates mechanoradicals by homolytic bond scission. We here employ scale-bridging simulations to determine the influence of collagen's mesoscale fibril structure on molecular breakages, combining atomistic molecular dynamics simulations with a newly developed mesoscopic ultra-coarse-grained description of a collagen fibril. Our simulations identify a conserved structural feature, a length difference of the two helices between pairs of crosslinks, to play a critical role. The release of the extra hidden length enables collagen to buffer mechanical stress. At the same time, this topology funnels ruptures such that the potentially harmful mechanoradicals are readily stabilized, buffering the arising oxidative stress. Our results suggest collagen's hidden length to exploit a sweet spot in the trade-off between breakage specificity and strength.

DOI: 10.1103/PhysRevE.110.054408

I. INTRODUCTION

Structural proteins like collagen have to withstand high mechanical loads, reaching up to 90 MPa in stretched achilles tendon [1]. They have to buffer peak loads from exercise, or even deal with accidental overstretching, ideally without major material failure. Collagen is the main component of all connective tissues, such as tendon or ligaments. It is well known for its outstanding mechanical properties, which are attributed to the intricate structure of this protein material.

The structural organization of collagen into fibers spans multiple length scales, ranging from its protein secondary structure and molecular crosslinking to the macroscopic winding of the fibrils up to fiber bundles. The triple helices are the building blocks of collagen's fibrillar structure. They comprise three protein chains, each of about 1000 amino acids, with a length of about 300 nm. As depicted in a two-dimensional scheme in Fig. 1(a), these helices are packed in a staggered way. This organization results in the typical gap and overlap pattern of collagen, with an about 67-nm-long D-period comprising one gap and overlap region as the repeating unit along the fiber axis. Across the fiber axis, the helices form a (quasi)hexagonal pattern [2] in the overlap region. In the gap region, however, the triple helices coil around each other, thereby swaping neighbors when running along the fiber axis [Fig. 1(c)]. Furthermore, adjacent triple helices are crosslinked to one another, with up to two crosslinks at either end of a helix. These crosslinks are enzymatically derived from lysines, such that they can be formed only at certain positions along the sequence. At the next structural level, the resulting microfibrils wind up forming collagen fibrils. Deciphering the determinants of collagen stretch response and its eventual failure requires taking into account this complex structure across these scales all the way down to the molecular level.

Recent experiments have shown the scission of molecular bonds in stretched rat tail tendon collagen type I [3], even in the subfailure stress regime. The radicals generated from covalent bond ruptures have been shown to stabilize on dihydroxy-phenylalanines (DOPAs) and to cause formation of H_20_2. The arising oxidative stress can potentially trigger further reactions of the body. We have recently shown that DOPAs are enriched around the enzymatic crosslinks in collagen [4] and, intriguingly, cross-linked areas have been identified as first rupture sites based on a combination of quantum mechanical, molecular dynamics (MD), and kinetic Monte Carlo simulations [5,6]. Further, we identified sacrificial bonds on the molecular level inside the crosslinks that funnel the first ruptures to the vicinity of DOPA [5]. Hence, the picture emerges that collagen might have evolved to deal with mechanical and oxidative stress inside this regime of highly loaded yet still subfailure regime. However, the simulations of collagen rupture have so far been limited to a single D-period. How crosslinks and backbone bonds share loads and undergo bond scission across and along a whole collagen fiber has remained unknown and requires a mesoscopic collagen model able to predict covalent bond scission in systems with several layers of crosslinks, i.e., far beyond a single D-period.

Experimental studies on the influence of collagen structure onto its mechanical properties are naturally focused on larger length scales, while computational studies like molecular

*Contact author: hgrubmu@gwdg.de
†Contact author: frauke.graeter@h-its.org

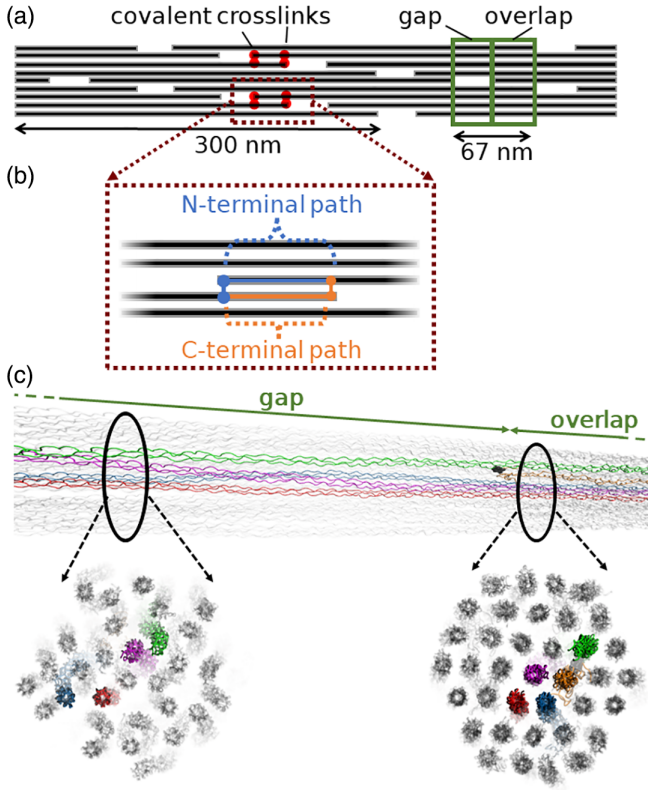

FIG. 1. Collagen (micro-)fibril structure. (a) The 300-nm-long tropocollagens are stacked in a periodically shifted pattern against each other, leading to the well-known D-periods with gap and overlap regions. A typical 2D scheme for this packing is shown in black, with two crosslinks highlighted in red. (b) In the overlap region, there are two force transmission pathways via the two crosslink sites at the N- and C-terminal regions, shown in blue and yellow, respectively. (c) Clipping of the backbone structure of the atomistic model used for the MD simulations described in the text. Exemplary molecules in one unit cell of the overlap and gap region are highlighted by color. In the gap region, the relative position of the triple helical molecules changes as they twist around each other. This shifting, which cannot easily be depicted in 2D, can be seen in more detail in the two cross-sectional views below. The atomistic model shown here is taken from ColBuilder [13], based on the 3D data by Orgel *et al*. [14].

dynamics have focused on collagen molecules or small fibril fragments, as our own previous work [5,6]. Some attempts on mesoscopic models [7,8], or also on disordered network models [9,10], have been made. Computational studies have lead to insights on the denaturation and unwinding of triple helices as a molecular failure mode in cyclic fatigue experiments [11,12]. None of these studies, to the best of our knowledge, combined rupture dynamics, as we obtained them from atomistic simulations, with larger topological fibrillar features.

Here we explore how collagen's fibril structure determines molecular breakages and influences subsequent reactions in the material. To this end, we bridge the length gap to the mesoscale by devising a super coarse model in a bottom-up parametrization from MD simulations. A key feature of our mesoscopic model is the ability to model a difference in force transmission pathways between two neighboring crosslinks [Fig. 1(b)]. According to our MD simulations at the microscopic scale, and as proposed on a smaller scale also previously [15], a difference in the peptide length between two otherwise fully equivalent crosslinks causes the N-terminal crosslink to always rupture first and never the C-terminal crosslink region. This feature, which we find to be conserved in collagen I, allows the release of hidden peptide length and thus increases the toughness of collagen.

Using our efficient mesoscopic model, we systematically explore the rupture properties of collagen across different structural parameters. We identify the hidden length between two adjacent crosslinks to strongly shift ruptures into crosslinks. This rupture specificity allows collagen to readily scavenge mechanoradicals through DOPA residues adjacent to crosslinks. The released hidden length in turn can buffer mechanical stress. Crucially, our comparisons across different hidden lengths suggest a trade-off between rupture specificity at crosslinks and the overall rupture propensity, i.e., fibril stability. Our mesoscopic model thus proved useful to assess collagen failure modes and can similarly be used to test other variations across the structure and crosslink topology of this important biomaterial.

II. RESULTS

A. Collagen breaks sequentially, releasing hidden length

To obtain first insights into the sequence of ruptures in atomistic simulations, as well as a basis for the parametrization of the coarse-grained model further below, we first resorted to a 67-nm fibril model [5,13]. It comprises an overlap region in the middle, with two half gap regions around. A clipped version can be seen in Fig. 1(c).

The breakage counts have been obtained previously with a combination of force-probe molecular dynamics simulations with kinetic Monte Carlo (KIMMDY) [5,6] which is briefly described in Sec. V and illustrated in Fig. 2(a). Note that rates sampled in the kinteic Monte Carlo step are on the order of 1/s at the force used in the MD simulations, aiming at the same regime as in macroscopic experiments (e.g., Ref. [3]). We here reanalyzed these data, which are publicly available in Ref. [16]. We observed a higher propensity of crosslink ruptures compared to backbone ruptures for both divalent HLKNL and trivalent PYD crosslinks [Fig. 2(b)]. This trend is even more pronounced for trivalent crosslinks, as those harbor a particularly weak bond. This potentially beneficial concentration of ruptures comes at the cost of higher total rupture rates.

We next asked if the two crosslink sites at the opposing ends of the overlap region [see also Fig. 1(b)] show the same or maybe different rupture propensities, despite having the same chemical structure. Figure 2(c) shows whether a crosslink in the microfibril broke on the N-terminal or C-terminal side. Overall, we see that N-terminal crosslinks are more prone to rupture than their C-terminal counterparts. Interestingly, this asymmetry increases with the PYD rupture rate. In contrast, almost all HLKNL crosslinks break at the N-terminal sites.

Next, we addressed the question of what happens after all N-terminal crosslinks have ruptured. To this end, we extended the previous data set by 18 more MD or KIMMDY

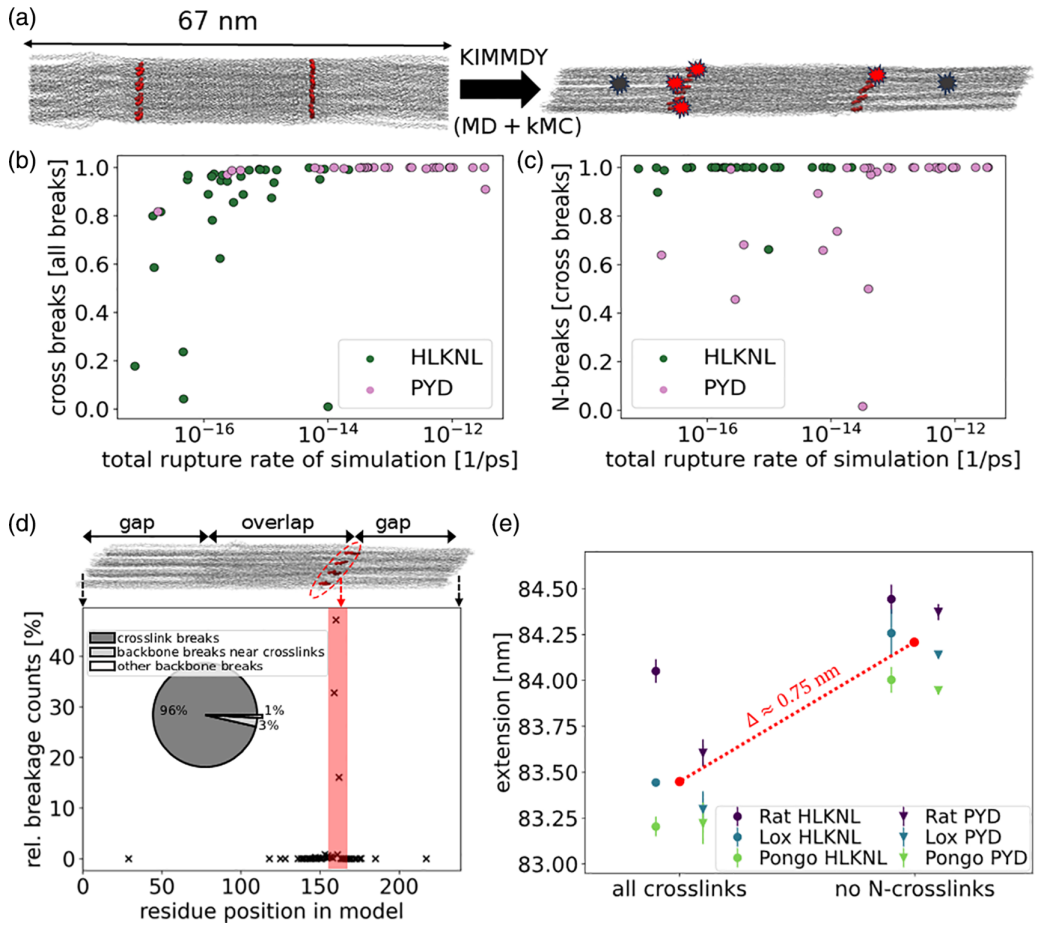

FIG. 2. Sequential rupturing: Collagen first breaks at N-terminal crosslinks and then on the C-terminal side, thus releasing hidden length between the connections. (a) Atomistic collagen fibril model before (left) and during a stretching simulation (right). We use KIMMDY, a combination of constant-force MD and kinetic Monte Carlo described previously [5,6], to obtain rupture counts. We symbolically indicated rupture points with stars (red for crosslinks and black for backbone rupture). [(b) and (c)] Rupture counts of 63 KIMMDY simulations (publicly available from pervious work in Ref. [16]) with respect to the ratio of crosslink breakages in (b) and, especially, with respect to the percentage of breakages at N-terminal crosslink sites given the crosslink breaks in (c). Color code according to crosslink type with HLKNL being divalent and PYD being trivalent crosslinks. (d) Relative breakage counts along the fibril in the second stage of a sequential rupturing after the N-terminal crosslinks have already been broken. Models are here only crosslinked at the C-terminal sites, encircled in red. Inset: Pie chart of summed-up ruptures in the crosslinks vs the backbones in the crosslinked area (up to five residues before or behind) vs elsewhere in the backbone. Data comprise a set of additional 18 pulling simulations. (e) Extension measured as end-to-end distance, averaged per model and crosslink type with three replica per data point. Error bars are standard errors of the mean, red line indicates the difference between the averaged values.

simulations of collagen models: For each of the same three species as before (*Rattus norvegicus*, *Pongo abelii*, and *Loxodonta africana*), we generated two models with ColBuilder that are crosslinked only at the C-terminal side, one model with divalent HLKNL and one with trivalent PYD crosslinks. This setup assumes exclusive rupture of N-terminal crosslinks first and substituted a series of KIMMDY simulations, which would have been computationally too demanding. With these 18 models, we conducted three replica each of force-probe MD and KIMMDY simulations.

Figure 2(d) shows the aggregated and then normalized KIMMDY rupture distribution of these 18 simulations. As expected, the concentration in the vicinity of the remaining C-terminal crosslinks is even more pronounced than for the fully crosslinked models. This finding points to a sequential rupturing mechanism, where the C-terminal crosslinks break after the N-crosslinks. Furthermore, Fig. 2(e) shows that the end-to-end distances of our new models are about 0.75 nm larger than our initially fully crosslinked models (averaged extensions are 83.46 and 84.21 nm, respectively). This additional extension was robustly observed across all three different species considered. As the models with and without N-terminal crosslinks are otherwise identical, this strongly suggests that there is a release of extra "hidden" length between the two connections that can occur only after the N-crosslinks are broken.

To identify the cause of the preferential N-terminal over C-terminal rupture and of the resulting release of hidden length, we analyzed the peptide lengths in between crosslinks in Fig. 3. Figure 3(a) shows a snapshot of an atomistic force-probe MD simulation, with two overlapping force transmission pathways highlighted in color. Already visual

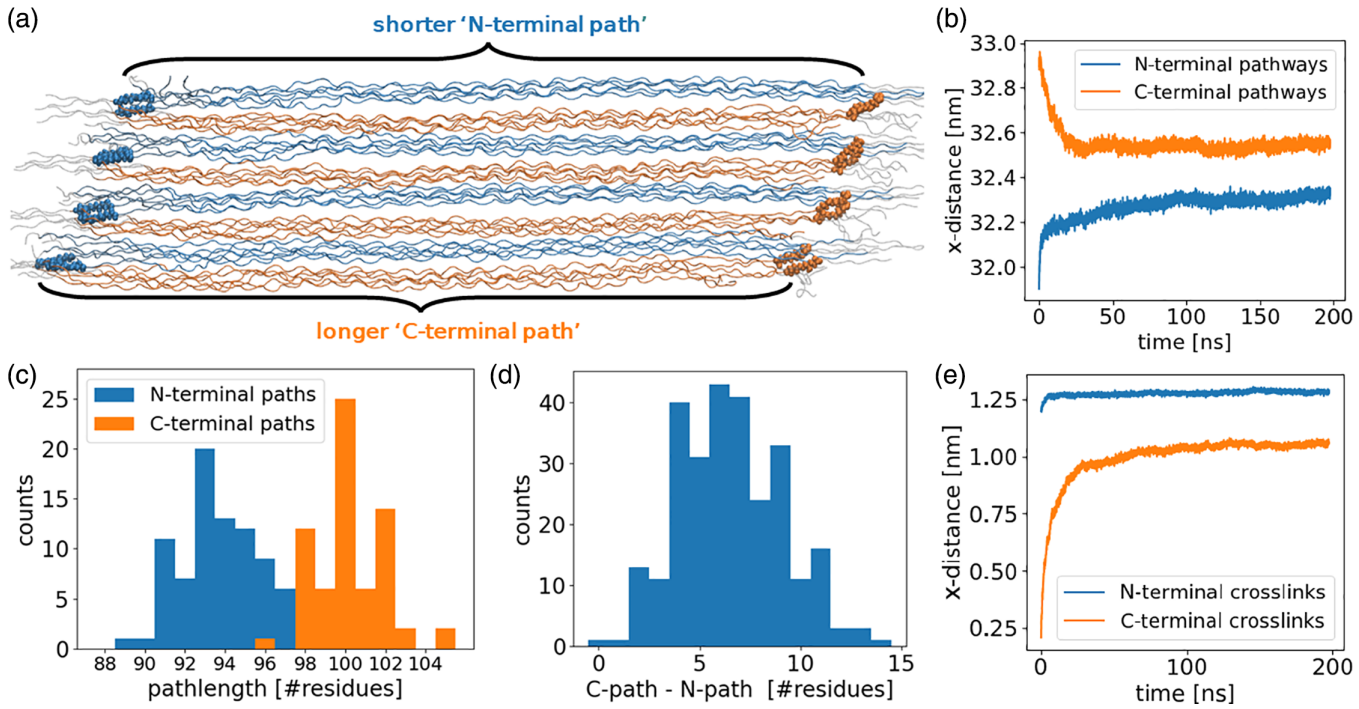

FIG. 3. Force pathways via N-terminal crosslink sites are shorter than via C-terminal crosslink sites in our collagen models. (a) Enlarged view of the overlap region, taken from a snapshot of an atomistic force-probe MD simulation. Ending strands connecting to left side are in orange, interdigitating triple helices in blue, and triple helix regions outside of the crosslink pair in gray. The crosslinks are highlighted space filling. Clearly, the N-terminal crosslinks lay down more in pulling direction, i.e., are more stressed, than their C-terminal counterparts. (b) Length along the fibril axis (x direction) of the two pathways for one simulation, averaged over the eight connections in that simulation. (c) Path lengths histograms obtained from all models from ColBuilder [13] with all possible crosslink connections available, measured in terms of the number of residues between the two crosslink sites. (d) Path length difference histogram, obtained from the differences of the path lengths in (c). (e) Crosslink elongation measured as distance between the crosslinked C_α atoms in pulling direction, averaged over the eight crosslinks per site for one simulation.

inspection suggests a clear strain difference in the respective crosslinks. The ones shown in blue, which are connected to the starting side of the new collagen molecule, are clearly more stretched in pulling direction, i.e., more stressed than the ones that connect the end of the C-terminal sites ending strands, shown in orange.

To quantify this observation, Fig. 3(b) shows the x distance (fibril axis direction) between the C_α atoms of the crosslinks for both pathways for an exemplary simulation, averaged over the eight connections in the model, thus providing a direct measure of the length of the two brackets indicated in Fig. 3(a). We see that, while the lengths approach each other under force, the C-terminal paths remain more extended than their counterparts. Similarly, Fig. 3(e) shows the z distance between the beginning and ending (C_α atoms) of the crosslinks as a measure for how much they are strained and orient themselves along the pulling direction. Clearly, the N-terminal crosslinks become almost fully stretched very early on, whereas the C-terminal crosslinks yield less length to the apparently lower load. Still, a smaller part of the total load is carried by the latter as well.

The histogram in Fig. 3(c) counts the number of amino acids between the connections at the two sites. To this end, we evaluated all models that are available in ColBuilder [13]. These involve 20 species with slightly varying sequences, each with different potential crosslink sites, i.e., lysines in regions that qualify for enzymatic crosslinks. Clearly, the connections via the N-terminal sites are shorter, with an average length of about 94 compared to the approximately 100 residues length observed for the C-terminal paths. Also, comparing individual path lengths within each of these systems, the N-terminal paths are always shorter than the C-terminal ones [Fig. 3(d)]. We find an average difference of six to seven residues, ranging up to 14 extra amino acids in some models. We also note that path length differences are less likely at every third position of this x axis, which can be explained by the Gly-X-Y sequence, which has lysines as crosslink precursors only at X and Y positions.

Overall, Figs. 3(c) and 3(d) suggest an average length difference of $6-7\%$ for the two paths of about 100 amino acids between the two crosslink sites. Note, however, that this way of counting does not necessarily define a difference in spatial distance. The N-terminal crosslink might connect to the first strand of the three chains in the triple helix, whereas the C-terminal crosslink might bind to the second or third strand of the triple helix, and both strands can be shifted relative to each other. Hence, the same residue number of one strand does not necessarily be at the same position as on another of the chains. This results in the fact that the count is not always exactly proportional to lengths. This shift is a well-known feature in collagen, and due to it the side chains of glycines in the GLY-X-Y sequence pattern can point to the inside of

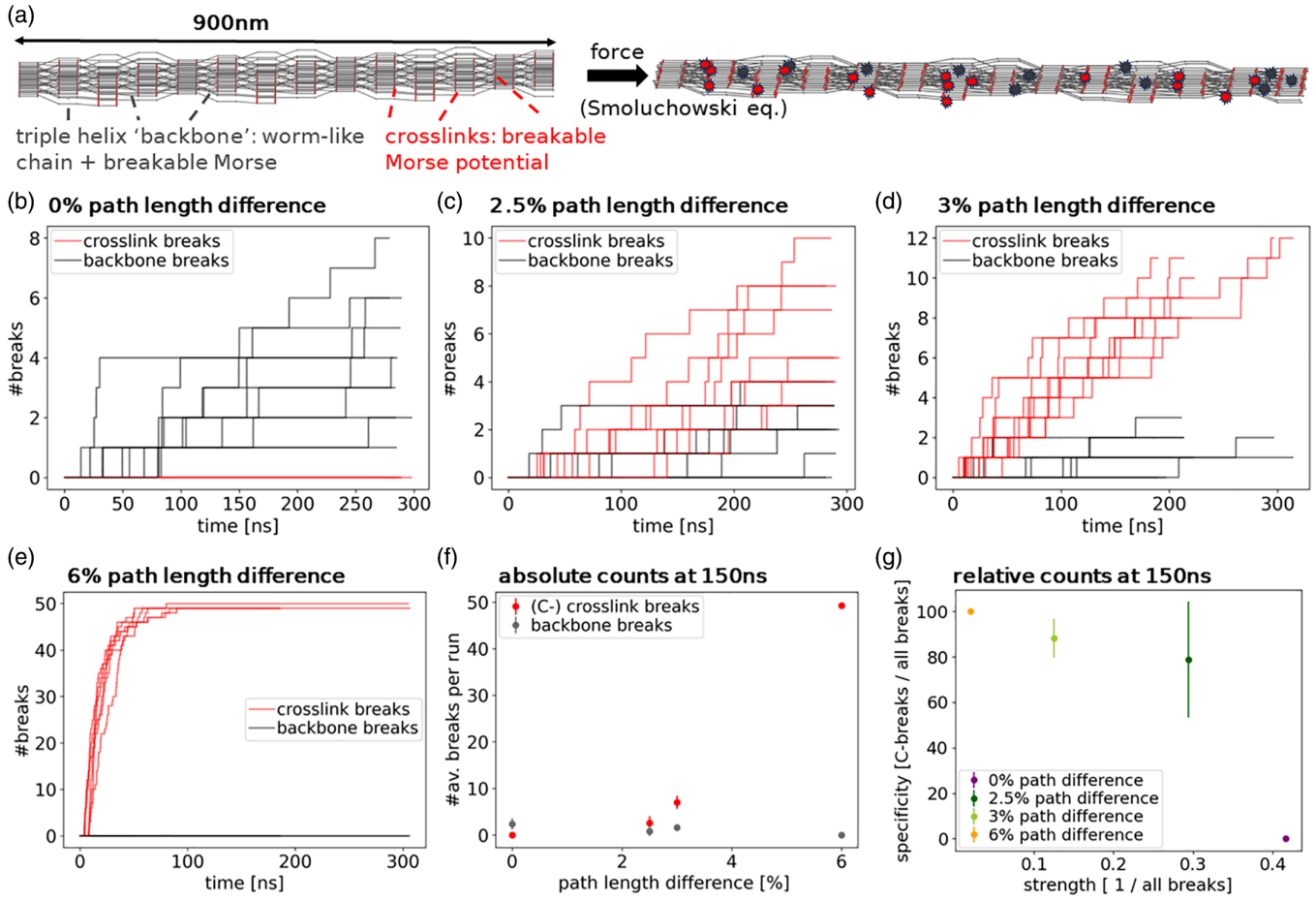

FIG. 4. Path length differences funnel ruptures into crosslinks but increase rupture counts. We show breakage events observed in the mesoscopic ColBreaker model of a 900-nm fibril; triple helical backbone ruptures are shown in black, N- or C-terminal crosslink ruptures in red. In the panels, each line represents an independent simulation, simulation time per run varied due to available computational resources. (a) Fibrillar mesoscopic model before (left) and during (right) the simulation; ruptures are indicated by stars. (b) In a model with equal path lengths between N- and C-terminal crosslinks, only backbone ruptures occur. (c) In a setup with 2.5% path difference, resembling our atomistic models, there is a competition between less backbone and more crosslink ruptures. (d) At 3% length difference, just slightly above the value actually observed in collagen fibrils, the crosslink breakages clearly outnumber occasional backbone ruptures. (e) With a 6% difference in the contour lengths, only crosslinks breakages are seen, and also markedly earlier. (f) Comparison of the number of backbone vs the number of crosslink failures at 150 ns for different path lengths, averaged over all replica. (g) Relative ratios of crosslinks vs backbone breakage as a measure of rupture specificity for different path lengths against inverse breakage counts, both again counted at 150 ns using the same data as in (f).

the helix, facilitating hydrogen bonding to side chains of other residues at the X or Y position.

Apparently, there is a difference of about 0.75 nm between the path lengths estimated from the respective amino acid counts and the direct length measurement under force in Fig. 2(e). This length corresponds to about two (stretched) amino acids or a bit more than 2% of this about 100 residue-long section. We obtain a similar result by looking at the extra strain compared to the length of the overlap region between the two crosslinks of about 31 nm:

$$\Delta\epsilon = \frac{\Delta l}{l} = \frac{0.75\,\text{nm}}{31\,\text{nm}} \approx 2.5\%. \tag{1}$$

Because the comparison of extensions under force provides a good estimate for the extra extension that is actually available, we will use these 2.5% also for the parametrization of the mesoscopic model further below. Because of this use case, the normalization with respect to the 31-nm-long overlap region is most useful here: The contour length of the overlap region in the mesoscopic model will be varied. Note, however, that this is not the strain of the full model.

B. Uneven load distribution among triple helices funnels ruptures into crosslinks

Our observation of preferencial N-terminal over C-terminal crosslink ruptures suggests a sequential rupturing of the microfibril. This scenario, however, cannot be directly monitored in costly atomistic simulations. To get more direct access to sequential rupture at the required larger length scales, and to systematically assess the function of the hidden lengths in the collagen system, we developed a mesoscopic simulation model ColBreaker, depicted in Fig. 6 in more detail and in Fig. 4(a) schematically for the simulation procedure. In

short, segments of triple helices are described by wormlike chains (WLC), with an additional breakable Morse potential to allow for backbone ruptures. These wormlike chains are interconnected via crosslinks, which, too, are described by Morse potentials. The system is assumed to be overdamped and, accordingly, propagated via the Smoluchowski equation (for further details, see methods and SI).

Although the atomistic models already comprise approximately 2.5 million atoms just to encompass one gap and one overlap region with a length of 67 nm, our coarse-grained model ColBreaker can scale up to micrometers. We have seen in Fig. 3 that the force pathways via N-terminal crosslinks are always shorter than their C-terminal counterparts in our models that follow the experimental x-ray data. Beside scaling to larger length scales, we use the model to easily explore parameters like this path difference by adjusting the contour length of the underlying WLC potential. Further, we compare the models with hidden length to, for example, homogeneous models where equal paths would equally share the force, which we do not have at hand atomistically.

To explore the influence of the connection lengths on the rupture propensities, we first conducted a set of coarse-grained simulations varying the contour factor of the triple helix sections between the two crosslinks in the overlap region, from 0% difference (i.e., both sides are of equal length) to 2.5% [similarly to the 2.5% extension after removal of N-crosslinks form MD simulations in Fig. 2(d)], via 3% right next to it, to 6% (about the average difference when counting residues as above) for a 900-nm-long collagen fibril. These models contain five triple helices on each of the six sides of the hexagonal grid, such that the system comprises a total of 1714 beads with 338 crosslinks. In contrast to the atomistic model that only contains 16 crosslinks, the coarse-grained system is therefore considered large enough to include most effects of the three-dimensional (3D) network topology. We apply mechanical strain to the system similarly as for the atomistic simulations (see Appendix B 2). Note, however, that these simulations use a higher force than those in the KIMMDY simulations (Fig. 2) in order to observe breaks directly on the simulated timescale reached by the mesoscopic model. Specifically, the virtual spring is extended to 20.5% on each side such that the force reaches an initial level of approximately 6 nN per triple helix, leading to several expected ruptures within accessible simulation time.

Figures 4(b)–4(e) shows the breakage counts for several independent runs, each of models with varying contour lengths between two adjacent crosslinks. For each simulation, the curves show the number of (N-terminal and C-terminal) crosslink ruptures versus the backbone breaks (at unspecific location) over time. When the path difference is zero and crosslinks share their load equally, as in Fig. 4(b), only backbone breakages occur. Introducing a 2.5% or 3% difference in the contour lengths [Figs. 4(c) and 4(d)], a competition occurs, and overall we see more crosslink than backbone ruptures. For even larger length differences of up to 6% [Fig. 4(e)], only crosslink ruptures occur and at a much higher rate. Notably, the intermediate path length values are those we find in the atomistic collagen structure. For this reason, we have adjusted the ColBuilder parameters to closely mimic the atomistic MD simulations. More specifically, we calibrated the relative strength of backbone versus crosslink Morse potentials such that a similar competition between ruptures in these two types of bonds in models with intermediate path lengths occurs [compare Figs. 2(b) and 7].

So far, these initial coarse-grained simulations yielded three main findings. First, the naturally occurring intermediate path length difference can now be systematically assessed in light of many scenarios ranging from both equal paths to marked differences, thus revealing the involved trade-off. Second, the ColBreaker simulations establish that also for the larger mesoscopic fiber, crosslinks rupture is preferentially driven by path length differences. Finally, with ColBreaker we assessed many consecutive ruptures in one simulation run and show this mechanism to also prevail until macroscopic failure of the fiber. In contrast, for the atomistic simulation we had to resort to several models (with and without N-terminal crosslinks) as a proxy for the multistep process, because simulating the full cycle would have been too costly.

To what extent the ruptures resulting from unequal load sharing between two adjacent triple helices is shifted towards the crosslinks is quantified and summarized in Fig. 4(f). This observation suggests that the underlying path length difference, resulting from a relative shift in crosslinked lysines in the two overlapping triple helices, can funnel the breakages away from unspecific failure somewhere in the fiber to a more localized and controllable environment in the crosslinks. We refer to this as breakage specificity in regard to location, which is critical for the radical scavenging properties of collagen, with the DOPA residues built-in nearby crosslinks.

This relative specificity ratio of crosslink to all ruptures is shown in Fig. 4(g) against the inverse of the total breakage count at a fixed time as a measure for strength. We observe a rather continuous loss of strength, emerging from many individual and stochastic ruptures. This comes along with a gain in specificity, when introducing a path length difference. This plot very clearly highlights the trade-off between specificity and strength, with the path length difference introducing specificity in rupture location at the expense of faster ruptures. Scenarios to exemplify this trade-off are compared schematically in Fig. 5(a). While it seems beneficial to put a higher load on one of the two connections, there is still a significant load share between the two, with the second connection bearing a significant minority of it. Reducing this too much, like in the case of 6% of path length difference, the system fails specifically but much more quickly. In particular the two close data sets with 2.5% and 3% difference, corresponding to the situation in actual collagen fibers as being probed in the atomistic simulations, show that the system is very sensitive to this parameter and also that the value obtained from MD seems to represent a "sweet spot." This conclusion is supported by comparison to the case of zero hidden length and thus equally shared burdens on both crosslinks [Fig. 4(b)]. Clearly, the latter scenario prevents crosslink ruptures and instead promotes unspecific backbone breakage. The latter is potentially more harmful, as radicals are generated unspecifically everywhere and can be less readily scavenged.

Inspired by these observations on the mesoscopic model, we have reanalyzed the total failure rates seen in the atomistic simulations [Fig. 5(b)], with similar conclusions. In particular, the failure rates observed for the atomistic models without

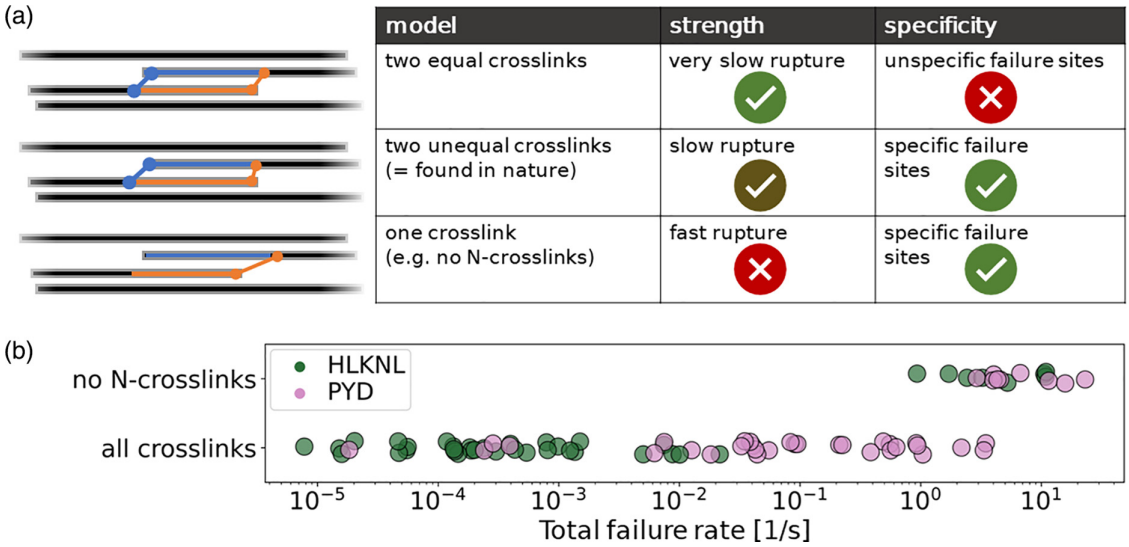

FIG. 5. Unequal load sharing is a sweet spot in the trade-off between strength and rupture specifictiy. (a) Sketch of three possible model configurations and their tabulated properties with regard to strength (measured as time to rupture) and specificity (location of ruptures clearly defined and nearby redox-active amino acids). Our atomistic models are based on experimental data [13,14] and correspond to the second case with optimal trade-off. (b) Comparison of total failure rates in atomistic simulations for the second and third case from (a). Rupture rates are taken from 63 KIMMDY simulations with fully crosslinked models and our new set of 18 models without N-terminal crosslinks, and are then summed up per simulation and colored according to crosslink type.

N-terminal crosslinks are several orders of magnitude larger than those seen for the fully crosslinked ones. Our fully crosslinked atomistic models, while more prone to rupture on one side, therefore must have some (uneven) load-sharing between the two connection sites. To put it reversely, if they did have to carry the full (or vast) amount of load alone even in the "fully crosslinked" scenario, then the failure rates would be as high as in the scenario of only single crosslinks.

Our mesoscopic model accurately replicates the rise in rupture rate of the C-terminal crosslink observed when the N-terminal crosslink is sacrificed. Here, too, rupture rates of C-terminal crosslinks are several orders of magnitude larger in the absence of N-terminal crosslinks [Fig. 8(a)]. Interestingly, we find parameter sets which ensure a significant survival time of the remaining lone crosslink. Figure 8(b) shows an example of such a rupturing mechanisms, with a pronounced lag time between the ruptures of pairs of redundant crosslinks. A reduced pulling force and a higher spring constant, that leads to a significant drop in pulling force on release of the hidden length, allowed to obtain this result.

To summarize, the seemingly small difference of a few percentages in contour length offers a powerful control mechanism, particularly for adjusting the count and specificity of covalent bond ruptures in collagen. Additionally, this hidden length reserve can contribute to enhancing collagen toughness.

III. DISCUSSION

Here we studied how the collagen type I microfibril structure affects its failure mechanisms, as well as the implications for chemical follow-up reactions in the material. To complement experiments with the necessary level of molecular detail of mechanoradical generation in tensed collagen, we employed both atomistic and mesoscale simulations to obtain a comprehensive picture on the rupture propensities in collagen.

We started out using KIMMDY, a hybrid approach of kinetic Monte Carlo and MD that can bridge the timescale gap between rare rupture events and MD accessible simulation time. Utilizing constant forces in a macroscopic subfailure regime, we obtained first rupture sites in atomistic simulation systems comprising one characteristic pattern of gap and overlap region. With about 2.5 million atoms, this simulation system is already computationally quite demanding, albeit small for collagen. We used previous [5] and additional KIMMDY simulations and analyses and observed that the two redundant and chemically identical crosslinks on either side of the overlap region differ in their rupture propensities when being stretched. Specifically, the N-terminal crosslink is connected to its adjacent helices through a shorter triple helix segment than the C-terminal crosslink and ruptures first. Remarkably, we find this notion of a sacrificial bond conserved across collagen I models of different species.

The concept of sacrificial bonds is widely known in synthetic polymer networks. In collagen, sacrificial ionic bonds have previously been proposed in bone to enable energy dissipation through the release of hidden length [17]. Recently, we already identified individual chemical bonds within trivalent collagen crosslinks to be particularly weak and sacrificial [5]. Turning our view from individual molecular bonds to the microfibril structure here, we observed the larger-scale interhelical structure, exhibiting a difference in helix length, to again funnel ruptures into the DOPA-rich regions. These radical-scavenging residues can prevent material damage, which would otherwise occur through uncontrolled radical migration [4]. In contrast to the molecular bonds, however, we here found that it is not the difference in bond dissociation

energies across covalent bonds that is the decisive adjusting screw of the arising competition, but the level of load-sharing between the two crosslinks. We see that while one side is more prone to rupture, a considerable fraction of the load is still carried by the second connection. Without the sacrificial N-terminal crosslinks, the overall failure rate would be significantly higher.

In addition to ensuring specificity of ruptures, our results also suggest that the difference in breakage propensity between the two crosslink types enables collagen to act as a mechanical buffer in that a bond scission in one connection releases additional hidden length of the other longer pathway. Previous studies of sacrificial bonds and hidden lengths in biological materials have attributed the observed mechanism to the rupture of noncovalent interactions [17]. In contrast, synthetic double network elastomers have been devised which release hidden length on rupture of scissile covalent bonds in the system [18], very much akin to the mechanism we describe here for collagen. Given that force probably distributes more heterogeneously in the real biological system than in our idealized models, this mechanism will specifically benefit very highly stressed strands for which elastic deformation is already on its limit and which otherwise might fail completely. In this regard, these strands can to a certain degree relax compared to their neighbors, thereby facilitating force redistribution and thus helping to maintain the overall integrity of the fibril. The dissipation of energy for bond rupture and fiber lengthening will not only increase strength but also toughness, as observed for synthetic materials [18].

To investigate larger length scales beyond our atomistic model, as well as to more systematically explore the effect of topological parameters on the stress response of collagen fibrils, we here developed the mesoscopic model ColBreaker (see Methods and Appendix B) and parameterized this model from our atomistic simulations. In short, this model consists of piecewise wormlike chains describing the triple helices, connected in the characteristic 3D topology via crosslink potentials which allow for both backbone and crosslink breakages. Despite a rather short integration step size which turned out to be required for technical reasons, this mesoscopic model enabled us to access substantially longer timescales as well as and much larger length scales. The obtained rupture simulations were validated by our atomistic results, where we compared fully crosslinked models to C-terminal crosslinked variants, the latter breaking much faster.

In particular, from simulations of this ColBreaker model, we obtained the following results. First, we saw that also at the fibrillar level the contour length of the connection between the two crosslinks forms a "set screw" that can adjust the ratio of backbone versus crosslink ruptures with remarkable sensitivity. Governed by the difference in force transmission pathway lengths, the collagen fibrils can adapt to uneven mechanical load in a trade-off between strength (more balanced load sharing) and specificity in breakage (more heterogeneous forces). This mechanism, second, allows for precise control of the sites of first microscopic failure, with redox-active amino acids and, in particular, the radical scavenger DOPA in close proximity. Indeed, its presence in collagen at these locations has very recently been confirmed experimentally with mass spectrometry [4]. We note that DOPAs are located both near the N- and C-terminal sites, so we did not find any particular benefit in having the N-crosslinks instead of C-crosslinks as first rupture points. Rather, the unequal load sharing between the two funnels the ruptures away from the backbone towards crosslinks that then might even both rupture (sequentially) in cases of very high load. Subsequently, in presence of water, H_2O_2 can be created, as has previously been shown in pulling experiments on collagen fibers [3]. In this way, the breakage of one crosslink offers a possibility to already start signaling high loads in collagen, triggering repair mechanisms and other reactions, even before macroscopic failure occurs.

Third, ColBreaker has enabled us to compare this scenario to a variety of other hypothetical scenarios which, too, are inaccessible to atomistic simulations. For example, models with connections of equal length only lead to unspecific backbone ruptures, whereas models with a more than realistic length differences tend to fail much earlier, and models with crosslinks on only one side cannot even sustain the same force levels. Hence, all of these scenarios have obvious functional drawbacks compared to the model setup closest to natural collagen, which suggests that the architecture and topology of collagen has evolved to achieve optimal functionality in tensed energy-storing tendons. We hypothesize that the hidden length is a structural element that can modulate collagen's mechanical response depending, for instance, on tissue type or on location, such as in low-stress positional versus high-stress energy storing tendons.

Previous experiments have suggested that crosslink rupture, followed by molecular sliding, might be a mechanism to explain force-extension curves of collagen type I fibrils [19]. It has also been speculated that even complete subsections of a fibril might be sacrificial areas or at least more susceptible to mechanical damage, with the origin of material heterogeneity to lie in different crosslink densities [20]. Our simulation results also provide mechanistic insights on the molecular level to these experimental hypotheses and support the notion that at least a major part of the sacrificial mechanism takes already place at the molecular level. Specifically, the asymmetry in the crosslink connection, i.e., the hidden length, leads to sliding and energy dissipation after initial crosslink ruptures until the C-terminal crosslink failure leads to a full disconnect of the triple helices, thus allowing for even larger displacements. This mechanism, however, is only valid in highly and thus redundantly crosslinked regions, that is, those with the maximal number of 2 mol/mol enzymatic crosslinks [21,22].

In addition, further experimental approaches to validate the computational results are desirable. While in previous work both crosslink and backbone ruptures could be seen in competition qualitatively [5], in particular the determination of (crosslink) rupture locations remains challenging. Similarly, the quantification of crosslink content by mass spectrometry might be a route to test our predictions more precisely. Another approach to circumvent handling the hard-to-control biological samples could be to use collagen-mimetic peptides that are artificially crosslinked. Such an idealized system would also allow us to quantitatively validate our computational observations, such as the force-dependent kinetics (Fig. 12) as well as the buffering through path length differences.

Clearly, by its mesoscopic nature, ColBreaker rests on simplifying assumptions and approximations. Nevertheless, its implementation is flexible enough to enable the inclusion of additional and potentially decisive structural properties of collagen—as well as other complex polymer-based materials—in future work. One possibility here would be to include "trivalent" crosslinks. Although the Morse scaling factor can be used to implement different crosslinks strengths, more changes would be required to differentiate between the behavivor of divalent vs trivalent crosslinks. Possibly, using a double-well energy barrier for trivalent crosslinks, which ruptures sequentially, could mimic such molecular sacrificial bonds [5] and might further fine-tune the observed competition between crosslink and backbone ruptures. Another option would be to more realistically capture postrupture sliding of helices, which requires inclusion of interhelical interactions, e.g., by using an adhesion term between neighboring strands. In the quite convoluted fibrillar topology, these nonbonded interactions could contribute markedly to the overall mechanical properties, for example by redistributing force to other strands, at the cost of introducing more complexity and computational cost.

Finally, our results highlight the role of crosslink topology on collagen mechanics. Both for collagen and also from a more general perspective of polymeric complex materials, it would be exciting to explore other topologies as well. For example, while the available fiber diffraction data [14] clearly favors the canonical topology studied here, it cannot rule out more heterogeneous force transmission networks with, e.g., alternating crosslink topologies, in which crosslinks from one triple helix are connected to different neighboring helices or where the three-dimensional twisting of the molecules follow a different course. ColBreaker simulations of such variants and comparison to available force or extension measurements might help to discriminate between these topologies. Similarly, advanced glycation endproduct crosslinks [23], which are more randomly located than enzymatically derived crosslinks and accumulate with ageing, could be incorporated. We speculate that the resulting heterogeneities in the force distribution could render collagen more prone to rupture, maybe also in a less controlled manner throughout the fibril than with only enzymatic crosslinks. Looking beyond collagen, it would be intriguing to see if the uncovered trade-offs and design principles are also at play in other crosslinked polymer network, both in natural ones (as actin) or synthethic ones.

IV. CONCLUSION

In summary, we here explored across several length scales the picture of a collagen topology that funnels bond ruptures into crosslinks, as a new biological counterpart to well-studied sacrificial connections in nonbiological polymers. Specifically, the topology of collagen exhibits a hidden contour length that causes such strong specificity and enhances the resilience of collagen against mechanical stress. When combined with our previous findings that there are specific sacrificial bonds on the molecular level exerting a similar function [5], and considering that—across different tissue types—particular sites of the radical scavenger DOPA are highly enriched around the crosslinked regions [3,4], our findings corroborate the emerging view that collagen fulfills not only a passive mechanical function, that of being strong and tough, but also actively controls highly specific mechanochemical routes of radical migration. The underlying and very sensitive trade-off is tuned by crosslink topology as a major "set screw."

V. METHODS

A. Molecular dynamics simulations

All MD simulations in this study were conducted with the software package GROMACS 2020 or GROMACS 2021 [24] using the amber99sb-star-ildnp force field [25,26]. For the crosslinks, Amber force field parameters were taken from our websever ColBuilder [13]. We constrained h-bonds with LINCS [27] enabling a 2-fs time stepping. As under these high mechanical loads the asymmetric bond stretch becomes relevant and as we consider bond scissions, we used Morse potentials for the bond interactions as previously described [6]. We cut off both Lennard-Jonas and Coulomb interactions at 1.0 nm and used periodic boundary conditions. The different collagen models were neutralized with ions and solvated in TIP3P water [28] leading to system sizes of approximately 2.5–2.9 million atoms. In order to provide enough space for extension of the initially 67-nm-long fibrils under force, a simulation box of 95 nm length was used. We first minimized the energy and then equilibrated the systems in NVT and subsequently in NPT ensembles for at least 5 and 10 ns, respectively. All simulations were conducted at 300 K and 1 bar, maintained by a v-rescale thermostat [29] and a Parrinello-Rahman pressure coupling [30] with coupling constants of 0.1 and 2 ps, respectively. For the production runs, we subjected the caps on both sides of the collagen fibril to a constant force of 3 nN per triple helix (or 1 nN per chain). In each instance, the collagen fibrils were subjected to force for 100 ns, with average bond elongation monitored at the last 10 ns as input in the bond rupture simulations of KIMMDY. Additional torque restraints were applied to prevent unwinding of the collagen triple helices [31].

In total, we used collagen fibril models from three different species, *R. norvegicus* (rat), *P. abelii* (orangutan), and *L. africana* (elephant) [13], which yet again varied in their crosslink position having either divalent HLKNL or trivalent PYD crosslinks. For each of these 2×3 models, we conducted three replica simulations, such that there is a total of 18 independent trajectories.

For the calibration and test of the ColBreaker mesoscopic model, we simulated one of the fibril models and an individual 67-nm triple helix extracted thereof at varying force levels, as described in Appendix B.

B. KIMMDY

To quantitatively describe bond rupture at low rates, KIMMDY was used, as described previously [6]. Briefly, this stochastic method uses the average bond elongations from the previous MD simulation as input, utilizing the dynamic force distribution in the stretched protein. The second input are the thermodynamic strengths, also known as bond dissociation energies (BDEs), of the covalent bonds that

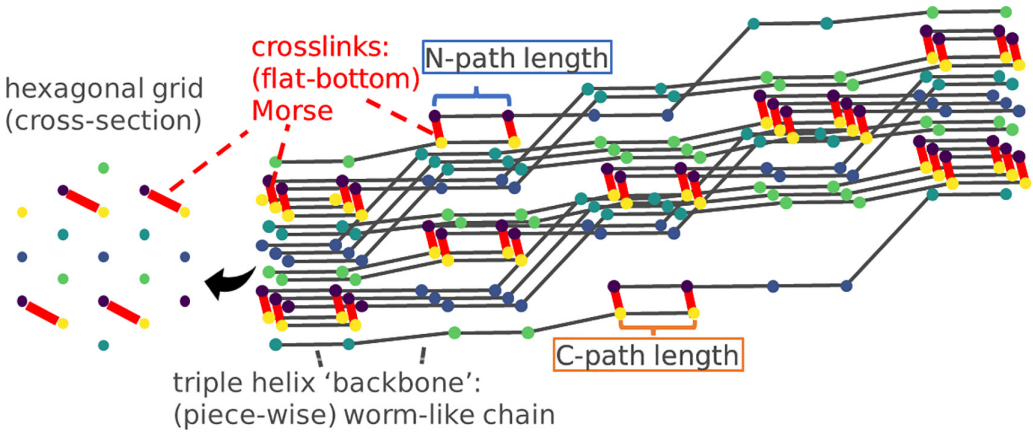

FIG. 6. The 300-nm section of the mesoscopic fibril model discussed in the text, with main potentials and parameters choices highlighted. This portion of the model comprises a fully packed hexagonal grid in the first as well as in the last layer. The grid of the first layer is also shown in a cross-sectional view on the left. Note that the fibril axis is horizontally condensed, in order to more clearly show the topology of the otherwise too long fibril. The starting beads (dark purple) of each molecule are connected to ending beads (yellow) of adjacent molecules via crosslink potentials (red). Variants of wormlike chain potentials (black lines) connect the beads along the collagen polymer.

are considered breakage candidates. The BDEs have been thoroughly explored and calibrated in previous work with quantum-chemical calculations, especially with respect to the different crosslink chemistries that have been shown to harbor particularly weak bonds [5]. Combining both the initial strength (BDEs) and the weakening of the stretched bonds (the force obtained from the MD), a Bell-type approach of an effective Morse potential yields bond rupture rates for backbone and crosslink bonds in the system. These failure events are then sampled from the pool of all events with a kinetic Monte Carlo algorithm. For each bond rupture, the respective bonded force-field terms are subsequently removed and the topology of the system adjusted accordingly. This procedure has been applied to a large set of 63 collagen systems in previous work, with the computational results being supported by experimental findings [5]. We here extend this set of breakage simulations with the 18 simulations of models without N-terminal crosslinks as described above.

C. ColBreaker: Simulations of mesoscale collagen failure

In this section, we summarize our mesoscopic collagen model used to explore larger, micrometer sized fibrils; for details we refer to the Appendix B. In this model, the triple helix "backbones" are described as WLC polymers, with interaction centers ("beads") at the ends or branching points of the crosslinks (see Fig. 6).

The motion of these particles and WLCs is described via Brownian dynamics,

$$\gamma \dot{x} = F(x) + \sigma \xi(t), \tag{2}$$

using a discretized version of the Smoluchowski equation,

$$\Delta x = \sqrt{2D\Delta t} \cdot \sigma \xi(t) + D\beta F(x)\Delta t. \tag{3}$$

Here $F(x) = -\nabla V(x)$ are forces derived from the interaction potential $V(x)$ described in Appendix B, which act along the collagen fibril axis; $\sigma\xi(t)$ are random forces described by Gaussian white noise with amplitude $\sigma = \sqrt{2\gamma k_B T}$ obeying the fluctuation dissipation theorem, and the diffusion constant $D = k_B T/\gamma$ is related to the friction γ via the Einstein relation.

Collagen molecules are segmented along the gap and overlap pattern shown in Fig. 1. In Fig. 6 the WLC polymer segments connecting these "beads" are indicated as black lines, with a total of 10 beads per triple helix. Beads at the termini of the triple helices (dark purple) were connected via crosslink potentials (red) to adjacent triple helix termini (yellow), such that the crosslinked network shown in Fig. 6 is obtained. Both the crosslinks and the "backbone" wormlike chain segments rupture under force. To this end, Morselike potentials at their maximal extensions are used. MD simulations were used to bottom-up parametrize properties such as the force-extension curves, and the Morse parameters were adapted such as to match the atomistic KIMMDY simulations. Further, the relative strength of "backbone" wormlike chain connections was adjusted versus the strength of the crosslink potentials in order to recover the competition between many crosslink breaks and fewer backbone ruptures observed previously [5]. In analogy to our webserver ColBuilder [13], we termed this model *ColBreaker*.

To define the proper connection topology shown in Fig. 6, we used the 3D structure determined by Orgel *et al.* [14], which also suggests how the triple helices are intertwined in the different regions of a collagen molecule. The fibril is aligned along the x axis, shown here slightly tilted, such that the hexagonal grid in the cross section and the resulting network can be seen. To get a better visualization of the gap and overlap features, the figure shows the fibril compressed along the x axis relative to the yz plane. On the left, we also show a cross section of the hexagonal arrangement of the individual layers. In this relatively small model used here for illustration, we choose a width of three triple helices per side of the hexagon that is known for collagen as the typical packing structure [2]. Beside the two potential types marked in the figure, different contour length between the overlapping segments were used in order to test the effect of varying distances between the crosslinks, marked in the

figure as N-path or C-path length, respectively, causing differing force transmission modes.

1. Implementation and Performance

The first version of ColBreaker has been implemented in Python and can be found on GitHub [32]. We used just-in-time compilation libraries that convert certain parts of the code to a C-based execution but are still limited in speed by this choice of programming language. We employed a 2-fs integration time step, as in the atomistic simulations, because we were limited by the fasted movements due to the steep Morse potential required to properly model bond rupture dynamics. For a "small" ColBreaker system of a 300-nm fibril consisting of 374 beads with 72 crosslinks, we still reached an average single core performance of about 70 ns/day on an Intel i5 10th Gen processor. This system corresponds in size to a 35 million all-atom collagen system. For comparison, using GROMACS 2020.3 on 32 SuperMUC-NG fat nodes (Intel Skylake with 48 cores each and 768 RAM) with a total of 1535 physical cores, the MD performance is about 5 ns/day. This corresponds to a speed up of about 21 500 by ColBreaker over atomistic simulations.

ACKNOWLEDGMENTS

This work was supported by the Klaus Tschira Foundation (F.G.). This research was conducted within the Max Planck School Matter to Life supported by the German Federal Ministry of Education and Research (BMBF) in collaboration with the Max Planck Society (B.R., H.G., and F.G.). This project has also received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 101002812) (F.G.). The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. [33] for funding this project by providing computing time on the GCS Supercomputer SuperMUC-NG at the Leibniz Supercomputing Centre (pn34ci) [34] (B.R. and F.G.).

APPENDIX A: FIGURES

1. Relative strength of backbone and crosslinks is a set screw for rupture specificity

As we discuss in Appendix B 6 below, we scaled the strength of the backbone and crosslink Morse potentials in ColBreaker to fit the relative breakage propensities that we obtained from our KIMMDY simulations. In Fig. 7, we explored the influence of this scaling parameter on the relative breakage counts, around the default setting of 1.45 for 900-nm-long models that contain a 3% path length distance, with the same simulation settings as in Fig. 4. As expected, a weaker backbone leads to more backbone ruptures and vice versa. This parameter is a sensitive set screw, as a change of only a few percentages strongly influences the narrow balance. In natural collagen, however, twisting the backbone strength is not likely (as the backbone is identical in all proteins), but the same effect can be achieved by having a different relative strength of crosslinks compared to that. We found in previous work that crosslinks harbor weak bonds with lower bond dissociation energies [5], and with ColBreaker we can now validate that there is a similar principle behind this trade-off as it is the case for the relative crosslinks force transmission pathways: Weakening the crosslink bonds slightly compared to the backbone, the rupture locations become much more controlled (that is, they are concentrated at the crosslinks).

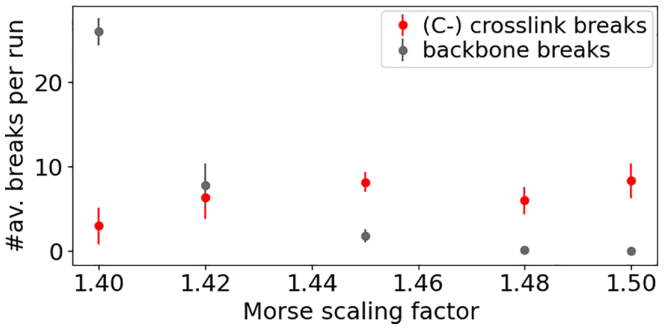

FIG. 7. Relative strength of backbone and crosslinks is, expectedly, another set screw for rupture specificity. For a set of at least three ColBreaker simulations per setting, in which we varied the strength of the backbone Morse potential, we show the backbone and crosslink rupture counts (at C-crosslinks, i.e., when the strand fully ruptured) per run in a 900-nm fibril pulled to 20.5% extension, at a simulation time of 200 ns.

2. Regaining atomistic behavior of a very fast or a sequential rupturing with ColBreaker in special parameter regimes

In Fig. 8(a), we can see that models with N- or C-terminal crosslinks fail very fast (within 2 to 3 ns after the ramp-up phase), confirming that ColBreaker recovers the same behavior as the atomistic simulations. In both, a lone crosslink leads to much faster failure, showing that a load sharing is beneficial, even if it is unequal between the two.

Figure 8(b), shows an example of how a sequential rupturing of N- before C-terminal crosslinks can be regained. Using a more extreme path length difference of 10% between the two connections, and pulling with a harder spring constant k as in the main article, the release of hidden length leads to a larger drop in the average pulling force (black curve, note that they are averaged over all strands, so the drop in the individual strands is much higher per rupture), with $\Delta F \sim k \cdot \Delta x$. Subsequently, the remaining sole C-crosslink in that strand can still sustain the now lower pulling force. While this regime is somewhat more extreme than our best-guess bottom-up parametrization below suggests, we note that there is a large variety in collagen types and crosslinking, rendering this case still possible.

3. Stress-strain curves of ColBreaker

In Fig. 9, we depict the stress-strain relation of our ColBreaker model, with the same settings as used in the main paper. All cases exhibit first a fast wormlike chain part (up to about 7–10% strain), then an almost linear section due to the "extensible" part in the eWLC (for more, see Appendix Methods below) and last a stiffening switching back to the molecular Morse potentials. Overall, the rupture forces are quite similar (close to 50 MPa), with the equal crosslinked

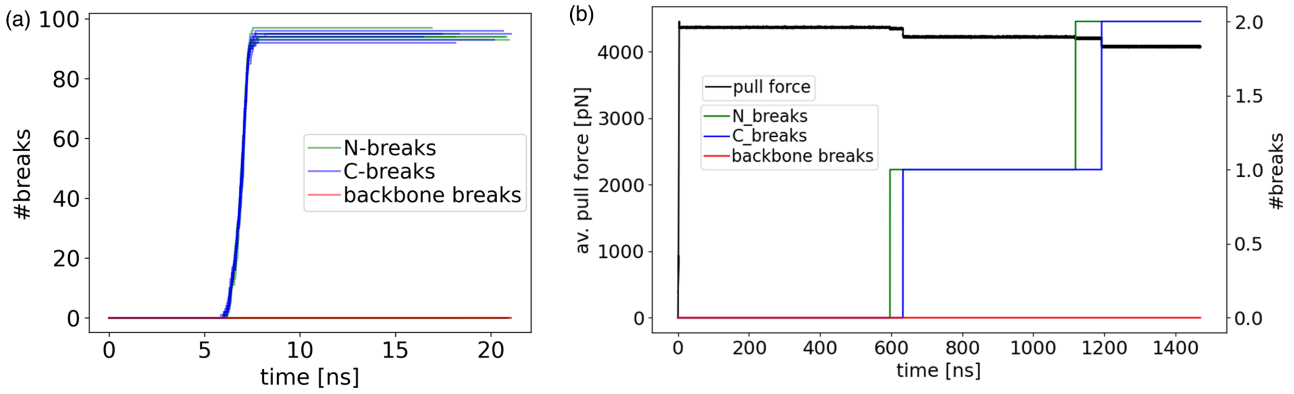

FIG. 8. Very fast or sequential rupturing in ColBreaker can be observed at more extreme parameter regimes. **a** Rupture counts in a 900-nm ColBreaker model that has crosslinks only at either N- or C-terminal sites (same setup as in the main Fig. 4), pulled up to 20.5% extension, with a linear ramp-up phase of 5 ns. (b) In a 300-nm fibril pulled at constant extension with a 10 times harder spring constant and having a path length difference of 10% between the crosslinks, we monitor the average pulling force (black) and the crosslink ruptures, separated by N- and C-terminal side. A sequential rupturing is observed.

model (0% path difference) being the strongest. A slower loading rate in the ramp-up phase (dashed yellow lines) can lead to lower initial stresses (only in the first few nanoseconds of the simulations or low strain), but ultimately catches up in the more stiff part (largest part of the simulation time). Overall, this is in line with experimental results [35–37].

APPENDIX B: METHODS (COLBREAKER)

ColBreaker is a mesoscopic, particle-based model, in which the most important interactions are defined between the connected beads. We here discuss the underlying physical potentials, their parameters, and our implementation of them into ColBreaker in more detail. Of course, other settings like length (in units of gap and overlap regions) and width (in units of beads in the hexagonal cross section) can also be adjusted and are listed here for completeness, both for the topology of the collagen fibril (Table I) and the more general and physics-related settings of the simulation (Table II).

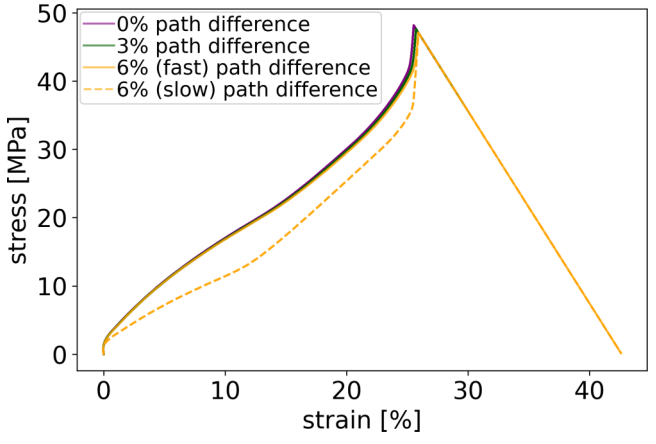

FIG. 9. Stress-strain curves for a 900-nm ColBreaker model at different path lengths (color coded). For the 6% case, two different loading speeds of the inital force ramp-up have been tested.

1. Topology of the ColBreaker fiber

A decisive feature of collagen is the twisting of collagen molecules. Much like a rope, the strength of a fibril is determined by how the individual triple helices wind around each other around each other with respect to the position on the quasihexagonal grid that collagen is known to exhibit in its cross-section [2]. This grid is the building block of ColBreaker, as in each gap and each overlap region we place exactly one bead per triple helix there. The connections between the beads, analogous to bonds, then define the course of the triple helices through theses "phases" of an individual collagen molecule. We refer to this here as the topology of ColBreaker.

The 3D topology of ColBreaker follows the x-ray fiber diffraction data of Orgel *et al.* [14], as do our atomistic models. An example of a 900-nm ColBreaker configuration can be seen in Fig. 10. To enable a representation that fits a page, we had to shrink the fibril axis compared to the cross-sectional plane such that this view is distorted and exaggerates the characteristic features along the fibril. The immense size difference between this model and our 67-nm atomistic microfibril underlines the need for such a coarse-grained description. Although we hardly incorporate the twisting of triple helices around each other in our atomistic models, we here can easily reach scales that include the 3D network topological effects up to the micrometer scale.

Figure 10(a), shows the course of the fibril, starting with the beads on the quasihexagonal shape in the overlap region. Once placed, they can move freely in the x axis (fibril direction) in the simulation, with drift and diffusional forces acting on them. We have colored the beads according to the "phases" in the braiding, starting from dark violet over shades of blue and green to yellow—in this way, it is easier to follow the course of one connected triple helix in this convoluted situation. As a result, it is always a yellow bead at the end that is connected to a dark bead at a beginning triple helix via the red crosslinks.

In the overlap, the triple helix "backbone" connections (black lines) are parallel to the fibril axis. In the gap region, the aforementioned twisting leads to connections to the next

TABLE I. Options in ColBreaker: Part I (topological parameters).

parameter	values [unit] (default)	description
layers	integer (5)	number of layers of gap and overlap regions, so length of fibril. Use 1+4*x to get x full triple helices
TH_per_side	integer (4)	number of triple helices per side in hexagonal building grid, so thickness of fibril
gap_ratio	float (0.54)	length ratio between gap and overlap region
periodicity	float [m] (6.7×10^{-8})	length of periodic d-phase (one gap plus one overlap region)
spacing	float [m] (1×10^{-8})	lateral spacing between triple helices on hexagonal grid
allow_switches	boolean (True)	if crosslink would go outside the model, switch direction to retrieve fully crosslinked network
connectedness	float [0,1] (1.0)	percentage of crosslinks. If lower than 1.0, then randomly chosen sites will not be crosslinked
double_crosslink	boolean (False)	use this option to double the amount of crosslinks: Connect both up and downwards
low_up_ratio	float [0,1] (1.0)	fraction of crosslinks that take the standard path or switch direction
crosslink_sites	[N, C, BOTH] (BOTH)	side of the gap region where crosslinks will be made
N_path_difference_factor	float (1.0)	relative length of N-terminal path in crosslinked phase compared to standard contour length
C_path_difference_factor	float (1.0)	relative length of C-terminal path in crosslinked phase compared to standard contour length

layer at different positions—see the zoom-in in Fig. 10(b) (note that it is turned a few degrees to enable a view into the cross-section). The twisting is also the reason for the perceived widening of the fibril diameter, as the triple helix does not stay within the same position of the hexagonal cross-sectional grid during its course through the fibril. Note that the resulting edge effect of a not fully packed configuration on the outside of the fibril is not relevant here, as we do not include nonbonded interactions between the strands at this point.

TABLE II. Options in ColBreaker: Part II (simulation and physics parameters).

parameter	values [unit] (default)	description
time_total	float [s] (1×10^{-7})	simulation time
dt	float [s] (2×10^{-15})	integration time step
diffusion_constant	float [m^2/s] (4×10^{-10})	diffusion constant used in Smoluchowski Eq. that will influence the internal time of the system.
pulltype	[VELOCITY, FORCE, STRAIN]	pulling setup
force_constant_in_kT	float [multiple of kT] (200×10^{17})	force constant for velocity or strain pulling
max_extension	float [initial length] (0.23)	in strain pulling, maximum extension of the virtual spring
v_pull	float [m/s] (100)	speed of virtual spring in constant velocity pulling, or for build-up phase in constant force and constant strain
constant_force	float [N] (6×10^{-9})	force level for constant pulling scheme
force_constant_pwWLC	float [N/m] (2.05×10^{-8})	force constant for enthalpic stretching in piecewise WLC
k_f	float [N/m] (440.5)	width of Morse potentials, 440.5 by harmonic bondtype C-CT from amber99sb* force field
persistence_length	float [m] (0.47×10^{-9})	persistence length of entropic part in (piecewise) WLC
contour_factor	float (1.17)	contour factor of entropic part in (piecewise) WLC
r_fb	float [m] (1×10^{-9})	flat-bottom extension of crosslinks before Morse potential starts
Edis_cross	float [multiple of kT] (119)	depth of crosslink Morse potential
Edis_bb	float [multiple of kT] (137)	depth of backbone Morse potential

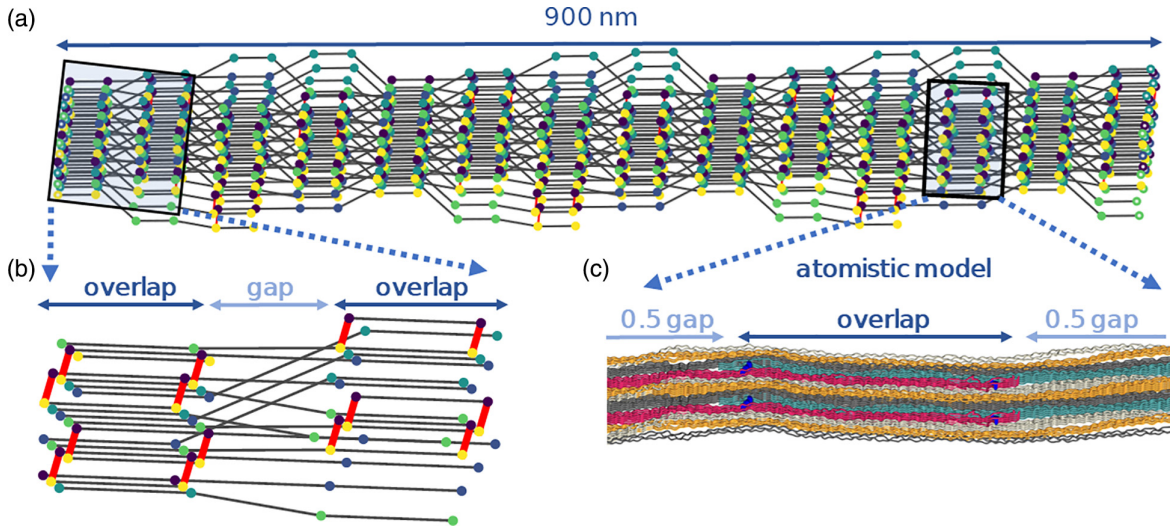

FIG. 10. Comparison between the topology of a ColBreaker fibril and our atomistic model. (a) We depict a ColBreaker model for a 900-nm-long fibril, which is about the length of three full individual collagen molecules. Note that the fibril axis is compressed here compared to the *y*/*z* plane for better visualization. In each gap and overlap region, there is a layer of beads connected via potentials that mimic the collagen triple helix. In the more dense overlap region, they proceed parallel to the fibril axis. In contrast, the twisting and change of relative positions happens in the less dense gap region. The beads are arranged on a hexagonal grid in the cross section. (b) More detailed and enlarged (and, for clearer visualization, thinned out) zoom-in onto the first two overlap regions. The red crosslinks connect beads in the last phase (yellow) of a molecule to ones that have just started (violet). (c) For comparison, we show our 67-nm-long atomistic model that has one overlap region in the middle and one half gap regions on each side.

2. External pulling forces

In ColBreaker, external forces are applied by pulling at the outer beads of those triple helices that would continue if not capped due to the finite size of the model. We have implemented three protocols of steering simulations in ColBreaker. Analogous to what can be done in the atomistic simulations with GROMACS, the first two options are "constant force" or "constant velocity" pulling. For constant force, one simply adds a constant external force to each bead (with ColBreaker also having an optional gradual build-up phase to increase the force level more smoothly in the beginning). In the second case of constant velocity pulling, a virtual spring with spring constant k is extended at a constant speed v_{pull} and drags on the beads.

In addition, a third protocol is implemented that is not available in our MD engine. We term it "constant extension" or "constant strain." It mimics a situation in which tendons get stretched up to a certain extension. This protocol is also implemented as a virtual spring, which gets pulled until it reaches a certain position x_{max}:

$$x_{\text{spring}} = x_0 + \max(v_{\text{pull}} \cdot t, x_{\text{max}}). \tag{B1}$$

This scheme combines several advantages, rendering it better suited to mimick the biological situation. In particular, we can keep up a certain force $F_{\text{max}} = k \cdot (x_{\text{spring,max}} - x_{\text{bead}})$ to obtain meaningful rupture rates and then have a relaxation after rupture. After a breakage inside one of the two crosslinks, the broken strand can slightly extend due to the release of extra length between the connections, and the resulting force decreases (as the distance to x_{max} is lower). In contrast, the other setups keep up the forces leading to higher strains after these micro-ruptures and do not lead to the expected relaxation.

3. Crosslinks as flat-bottom extended Morse potentials

For the crosslinks, which are (beside the backbone) one of the two modelled key interactions, a Morse potential is used, similar to the individual bonds in the atomistic simulations,

$$V_{\text{morse}}(x) = E_{\text{dis}}[1 - \exp(-\beta(x - x_0)]^2, \tag{B2}$$

with E_{dis}, the bond dissociation energy, as height of the potential and x_0 as equilibrium bond length. The width β of the potential can be expressed in terms of the harmonic force constant k_{bond} and, hence, obtained from the default force field parameters,

$$\beta = \sqrt{\frac{k_{\text{bond}}}{2E_{\text{dis}}}}. \tag{B3}$$

Under force, the crosslinks first lay down in the MD simulations. To mimic this, we added a flat bottom part of length $x_{\text{fb}} = 1\,\text{nm}$ before the Morse potential. If the actual bond is stretched more than another $x_{\text{cut}} = 0.25\,\text{nm}$, then the crosslink breaks. This distance is sufficiently larger than the inflection point of the Morse potential, after which the restoring force monotonically decreases to zero. Overall, the resulting force as a function of the distance between the two beads then reads:

$$f(x) = \begin{cases} 0 & x \leqslant x_{\text{fb}} \\ 2\beta E_{\text{dis}} e^{-\beta(x-x_{\text{fb}})}(1 - e^{-\beta(x-x_{\text{fb}})}) & x_{\text{fb}} < x \leqslant (x_{\text{fb}} + x_{\text{cut}}) \\ 0 \text{ (crosslink broken)} & x > (x_{\text{fb}} + x_{\text{cut}}) \end{cases}. \tag{B4}$$

In essence, we coarse-grain the several atomistic crosslink bonds together into one effective potential. For the barrier, we chose the weakest link with the lowest bond dissociation energy E_{dis}. We saw previously that typically the weakest bonds in the crosslinks have a much lower barrier than the others [5] such that we can neglect the influence of other bonds breaking here. We validate the crosslink breakage behavior together with the backbone against KIMMDY data in Appendix B 6.

4. Collagen backbone as extensible wormlike chains with Morse breakage barrier

A key question for building a mesoscopic collagen model is how to approximate the triple helices. Beside the formation of mechanoradicals, another common ground of collagen and polymers is that due to its building block nature, collagen can be viewed as a biopolymer. In the literature, there is a wide range of both experimental and computational studies [38,39] that compare collagens to WLC or similar models. Particularly, we are interested in the force-extension curves of stretched collagen molecules.

A standard wormlike chain model can be defined by the fact that the next segment points in a similar direction as the previous one. This leads to the fact that the directions of neighboring segments are correlated up to a characteristic length scale, the persistence length l_p.

Another possibility to calculate the persistence length l_p is to consider the WLC under force when stretching a polymer of contour length L_c in (without loss of generality) the x direction. When the chain extends, the number of possible configurations and, hence, the entropy of the system decreases. The resulting counteracting entropic force can be calculated from the total (bending plus pulling) energy. There is no exact solution for the force-extension relation, but several interpolation formulas exist. Probably most known is the version from Marko and Siggia [40,41]:

$$F(x) = \frac{k_B T}{l_p}\left[\frac{1}{4}\left(1 - \frac{x}{L_c}\right)^{-2} + \frac{x}{L_c} - \frac{1}{4}\right]. \tag{B5}$$

For collagen, the WLC model has also been applied to different experimental data, obtained by various techniques from rheology to viscometry to Atomic Force Microscopy (AFM) to optical tweezers, as reviewed in Refs. [38,39]. These reviews point out that the measured persistence lengths range from 5 to 167 nm and vary depending on the utilized method, as well as on other factors like ion concentration.

In addition to these uncertainties for the experimental values, the force-extension formula in Eq. (B5) by definition only covers entropic contributions, but in the high-force regime collagen also exhibits structural changes like (un-)twisting and enthalpic stretching of internal degrees of freedom, like molecular bonds [42]. As we are mostly interested in this high-force regime where occasional covalent bond rupture is possible, more sophisticated models that add an enthalpic contribution (usually an extra harmonic term with spring constant K_0) are more applicable. One early approximation formula for the extension including the high force regime was derived by Odijk [43],

$$x = L_c\left[1 - \frac{1}{2}\left(\frac{k_B T}{F l_p}\right)^{0.5} + \frac{F}{K_0}\right]. \tag{B6}$$

We refer to this as extensible Wormlike chain (eWLC) model. As noted in an article reviewing this and other more advanced WLC models, the eWLC description is comparatively accurate in bridging between the entropic lower force regime and the enthalpic stretching but has the disadvantage that it cannot be inverted to a form that could be used in simulations to obtain forces at given extensions [44]. To circumvent this, the authors of that study suggest a piecewise wormlike chain (pwWLC), defining the two regimes separately:

$$F(x) = \begin{cases} F_s(x) = \frac{k_B T}{4 l_p}\left(1 - \frac{x}{L_c}\right)^{-2}, & \text{if } x \leqslant x^* \\ F_e(x) = \frac{K_0}{L_0}(x - x^*) + F_s(x^*), & \text{if } x > x^*. \end{cases} \tag{B7}$$

In there, the transition happens at the inflection point given by

$$x^* = L_0\left[1 - \frac{1}{2}\left(\frac{k_B T}{l_p F^*}\right)^{0.5}\right] \quad \text{with} \tag{B8}$$

$$F^* = \frac{1}{4}\left(\frac{k_B T K_0^2}{l_p}\right)^{1/3}. \tag{B9}$$

This pwWLC provides a computationally accessible formulation at the cost of a kink at the transition point and a worse approximation in the intermediate regime. We test the various models against our own all-atom MD data here.

5. Bottom-up validation of WLC models with MD data

For a "bottom-up" comparison, we conducted constant force MD simulations, this time of both our full fibril model (67-nm-long, containing 41 triple helices) and of an individual 67-nm-long triple helix, at various force levels for at least 100 ns each. For low pulling forces, where fluctuations and bending of the elastic rod play a lager role compared to the stretching, longer simulation times were necessary to obtain converged end-to-end distances in pulling direction. Thus, for the single triple helix, we were able to explore a larger range of forces due to the smaller system size.

The obtained MD force-extension values for the fibril and triple helix are displayed in Fig. 11 as the orange and dark red data points, respectively. We show a linear scaling on the upper panel and a logarithmic depiction on the lower one, respectively, to resolve both the more relevant high force regime and the typical WLC characteristics in the lower force regime. A direct comparison of the two MD sets shows good agreement, with the full fibril yielding slightly larger extensions at identical force levels; as expected due to extra length that can be gained by crosslink extension and by a relative shift of helices with respect to each other.

Next, we fit three variants of the discussed WLC-based models to our MD data. First, we use the eWLC by Odiijk [43] that is most accurate but does not have an explicit force-extension formulation and can, hence, not be used directly in our code. It reproduces the MD behavior well over the full force range, as can be seen in both the linear and the logarithmic depiction (yellow line). With this result at hand, we obtain a persistence length l_p of 4.39 nm and a contour

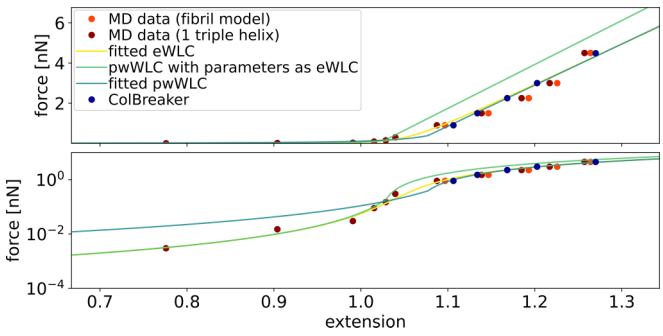

FIG. 11. Comparing force-extension curves of wormlike chain models to MD data of collagen under force. Average extensions, projected in pulling direction, at a given constant force for different simulation setups. Dark red and orange data points are obtained from constant force MD simulations of our fibril model and of an individual 67-nm-long triple helix, respectively. To these, we fit the eWLC (yellow), calculate the pwWLC with the parameters obtained from the eWLC (light green), and again fit the pwWLC focusing on the upper force range (green-blue). The parameter set obtained from the latter is utilized for ColBreaker. Thereafter, we compare these to the extensions of ColBreaker at the same force levels (dark blue). The lower panel has a logarithmic scale for the force, showing the same data.

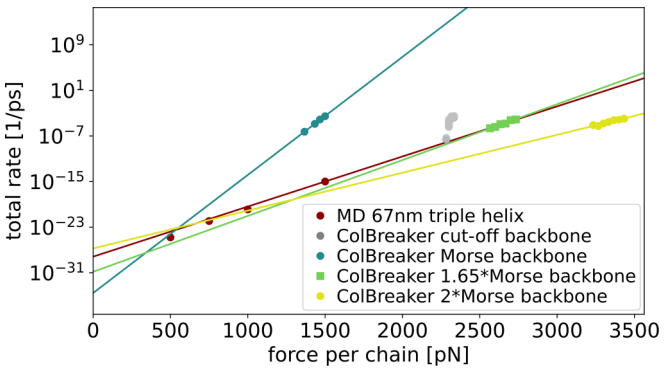

FIG. 12. The force dependency of backbone rupture rates of ColBreaker can be fitted to MD data. We use the summed-up KIMMDY rupture rates for a 67-nm triple helix, with a log-linear relation fitted to it (dark red line) as reference value. To this, we compare different rupture implementations of ColBuilder: A simple cutoff (gray), a Morse potential at the end of the WLC with the strength of an individual covalent bond (blue), with twice that strength (yellow) and, as best fit, with a increased strength of factor 1.65 (green). The ColBreaker rates are obtained as the inverted average failure times of five simulations per data point.

factor c_f of 1.07 for the wormlike chain regime, as well as a spring constant K_0 of 21.88 N/m. This is consistent, though on the very low end, with collagen persistence lengths reported elsewhere [39].

Second, using these values as input for the piecewise defined model, (light green line), good agreement in the lower force regime is regained, but with an offset for high forces. The two curves overlap up to an extension of about 1.07, where the kink due to the piecewise formulation is visible, especially in the logarithmic scaling.

Finally, fitting the pwWLC (to the nonlogarithmic values) with a least-squares algorithm naturally focuses more on high forces, which is the most relevant regime for our purposes. With this, we obtain the third parameter set (green-blue line). Although these parameters do not accurately reproduce extensions at low forces (and, hence, the fit yields meaningless values of 0.47 nm and 1.16 for persistence length and contour factor, respectively, for that regime), it is especially the upper linear part after the kink that is decisive in our model. This section is now more accurate, without the previous offset that the second parameter set would cause. This procedure yields, therefore, our final parameter set, in particular the effective spring constant K_0 of 20.45 N/m that we use for ColBreaker.

As a consistency check, we then employed ColBreaker and obtained a force-extension curve (dark blue points) matching the pwWLC. On the one hand, as an end-to-end test of the code, this validates that our implementation reproduces the input. On the other hand, this validation shows that, despite this very simplified description of collagen, good overall agreement with the MD data is achieved.

6. Modelling the breakage behavior

Above, we have already introduced the flat-bottom Morse potentials for the crosslinks in ColBreaker, which break once the potential energy barrier is crossed. So the next question arises on how to implement the failure of a wormlike chain mimicking the rupture of a backbone bond in the collagen triple helix.

Figure 12 compares different backbone breakage implementations in ColBuilder with reference KIMMDY data. For the KIMMDY part, we used single 67-nm (one gap and overlap region) spanning parts of collagen triple helices under varying levels of external force; continuing with KIMMDY the MD simulations that we used for the MD triple helix data in the force-extension curve in Fig. 11. We added up all the individual bond rupture rates to obtain a total failure rate of that segment. We compare this to the average breakage times of a ColBreaker set-up with the same length (three beads connected by two WLCs, so one overlap and gap as well), conducting five simulations until failure for each force level. According to the Bell-model [45], a linear dependency on the logarithmic scale of this curve is expected or, considering more refined modes such as the one by Dudko, Hummer, and Szabo [46], a slight bending.

Initially, we tested a simple cutoff of the potential such that the WLC can sustain a maximal force level before a breakage occurs immediately; the data are shown in gray in Fig. 12, for one exemplary cutoff value. As can be seen, the force-dependency does not fit the slope at all, and it is not possible to cover a range of force values due to strong dependency of the rates on the force. The reason is that effective energy landscape is lowered proportional to $F \cdot \Delta x$, and the reaction distance Δx is much larger for a soft wormlike chain potential than for an individual covalent bond with a steeper Morse potential.

To describe the actual breakage process more accurately, we instead define another switching point x_{switch} at a certain force F_{switch} (per default at 4 nN). At this force, the functional form switches Morse potential after the pwWLC. This switching force should be lower than the maximal force given

at the inflection point of the Morse potential, and we shift the potentials relative to each other such that the transition is continuous in the force level,

$$F_{\mathrm{pwWLC}}(x_{\mathrm{switch}}) = F_{\mathrm{Morse}}(x_{\mathrm{switch}} + x_{\mathrm{shift}}), \tag{B10}$$

$$\stackrel{!}{=} F_{\mathrm{switch}}. \tag{B11}$$

In this way, the beads have to cross a steep bond potential equivalent to what happens in the crosslinks and the MD simulation instead of reaching the cutoff via a soft ramp. Using the parameters of a standard C_α-C backbone bond for this last part, we obtain the blue data points in Fig. 12. Now an overall more sensible force vs rupture rate relationship appears but at comparatively high rates. This can be explained by the fact that we used the equivalent of a single bond for the whole triple helix, whereas in the MD simulation three strands would share the force. Analogously to having three springs in parallel, we expect the force to be shared and the effective potential to be stronger. Another difference between the atomistic and mesoscopic simulations is that we also have to consider that in our atomistic system we have many more bonds in a row per backbone that could all break (but have to withstand about the same force level). This second effect works in the opposite direction, though. Third, the idea of springs is idealized and, in practice, the total rupture rate is dominated by individual rates of bonds where force concentrates more due to heterogeneities or simple differences between the amino acids.

To effectively take these difference into account, we added an empirical strength scaling factor of the Morse potential. From the above considerations, it should be somewhere in the range between 1 and 3. For comparison, we show in Fig. 12 in yellow a Morse potential with twice the strength and, finally, the best fit in green. For this, we enhanced the barrier by a factor of 1.65. We still see a slight deviation in the fitted slope to the MD data but also note that this is within the given accuracy, as we only cover a small fraction of that graph. Second, the comparison of breakage data in the next paragraph will yield another iteration over this parameter set, which is why we abstained from more detail here.

Last, we compare the relative occurrences of both backbone and crosslink rupture propensities. For the crosslinks, we set a BDE of 296 kJ/mol for the barrier height in the Morse potential, which is the lowest value in the divalent HLKNL crosslink obtained from the QM calculations [5]. Using the backbones enhanced by the above factor of 1.65, however, we noticed that now the crosslinks are comparatively weak against the backbones, breaking much faster even without a crosslink path difference. The above fitting of the backbone accounted empirically for heterogeneities and imperfections in the backbone, but the crosslinks still have the parameters of an idealized bond. In the MD, in contrast, they would only lay down up to a certain angle, force might distribute into other degrees of freedom such as angles, and also other more complex effects of the fibrillar network can lead to a reduction in force uptake. We did not consider these points in the above fit to a simple collagen triple helix for the backbone. In order to take these factors into account, we iterated once more over the empirical strength of the backbone parameter, to come back to the competitive rupture regime between backbones and crosslinks that we observed previously with KIMMDY in for a wide range of parameters [5]. We could have adjusted the crosslink potential instead, but opted to work again with the backbone strength in order to keep the number of empirically fitted parameters as low as possible. In the end, an enhancement of the backbone strength by a factor of about 1.45 (instead of the above 1.65) turns out to be most reasonable. This value was obtained by comparing the relative rupture ratios of the mesoscopic model to the KIMMDY data in Fig. 2(b) in the main article, with a typical share of crosslink breakages in the range of 65% to 80%, or even 80–90% if we include backbones in the vicinity. We have also explored the influence of this parameter in Fig. 7 in more detail.

[1] P. V. Komi, Relevance of *in vivo* force measurements to human biomechanics, J. Biomech. **23**, 23 (1990).

[2] D. J. Hulmes and A. Miller, Quasihexagonal molecular packing in collagen fibrils, Nature (Lond.) **282**, 878 (1979).

[3] C. Zapp, Mechanoradicals in tensed tendon collagen as a source of oxidative stress, Nat. Commun. **11**, 2315 (2020).

[4] M. Kurth, DOPA residues endow collagen with radical scavenging capacity, Angew. Chem. Int. Ed. Engl. **62**, e202216610 (2023).

[5] B. Rennekamp, Collagen breaks at weak sacrificial bonds taming its mechanoradicals, Nat. Commun. **14**, 2075 (2023).

[6] B. Rennekamp, F. Kutzki, A. Obarska-Kosinska, C. Zapp, and F. Gräter, Hybrid kinetic Monte Carlo/molecular dynamics simulations of bond scissions in proteins, J. Chem. Theory Comput. **16**, 553 (2020).

[7] M. J. Buehler, Nanomechanics of collagen fibrils under varying cross-link densities: Atomistic and continuum studies, J. Mech. Behav. Biomed. Mater. **1**, 59 (2008).

[8] B. Depalle, Z. Qin, S. J. Shefelbine, and M. J. Bühler, Influence of cross-link structure, density and mechanical properties in the mesoscale deformation mechanisms of collagen fibrils, J. Mech. Behav. Biomed. Mater. **52**, 1 (2015).

[9] E. Ban, Strong triaxial coupling and anomalous poisson effect in collagen networks, Proc. Natl. Acad. Sci. USA **116**, 6790 (2019).

[10] F. Burla, Connectivity and plasticity determine collagen network fracture, Proc. Natl. Acad. Sci. USA **117**, 8326 (2020).

[11] J. L. Zitnay, Molecular level detection and localization of mechanical damage in collagen enabled by collagen hybridizing peptides, Nat. Commun. **8**, 14913 (2017).

[12] J. L. Zitnay, Accumulation of collagen molecular unfolding is the mechanism of cyclic fatigue damage and failure in collagenous tissues, Sci. Adv. **6**, eaba2795 (2020).

[13] A. Obarska-Kosinska, B. Rennekamp, A. Ünal, and F. Grater, Colbuilder: A server to build collagen fibril models, Biophys. J. **120**, 3544 (2021).

[14] J. P. Orgel, T. C. Irving, A. Miller, and T. J. Wess, Microfibrillar structure of type I collagen *in situ*, Proc. Natl. Acad. Sci. USA **103**, 9001 (2006).

[15] A. L. Kwansa, R. De Vita, and J. W. Freeman, Tensile mechanical properties of collagen type I and its enzymatic crosslinks, Biophys. Chem. **214-215**, 1 (2016).
[16] B. Rennekamp and F. Gräter, Collagen breaks at weak sacrificial bonds taming its mechanoradicals [Data] (2023), doi:10.11588/data/HJ6SVM, heiDATA, V1.
[17] G. E. Fantner, Sacrificial bonds and hidden length dissipate energy as mineralized fibrils separate during bone fracture, Nat. Mater. **4**, 612 (2005).
[18] E. Ducrot, Y. Chen, M. Bulters, R. P. Sijbesma, and C. Creton, Toughening elastomers with sacrificial bonds and watching them break, Science **344**, 186 (2014).
[19] E. Gachon and P. Mesquida, Stretching single collagen fibrils reveals nonlinear mechanical behavior, Biophys. J. **118**, 1401 (2020).
[20] S. Baldwin, A new longitudinal variation in the structure of collagen fibrils and its relationship to locations of mechanical damage susceptibility, J. Mech. Behav. Biomed. Mater. **110**, 103849 (2020).
[21] D. R. Eyre, M. A. Weis, and J. J. Wu, Advances in collagen cross-link analysis, Methods **45**, 65 (2008).
[22] A. Ellingson, N. Pancheri, and N. Schiele, Regulators of collagen crosslinking in developing and adult tendons, Eur. Cells Mater. **43**, 130 (2022).
[23] M. Stammers, Age-related changes in the physical properties, cross-linking, and glycation of collagen from mouse tail tendon, J. Biol. Chem. **295**, 10562 (2020).
[24] M. J. Abraham, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX **1-2**, 19 (2015).
[25] R. B. Best and G. Hummer, Optimized molecular dynamics force fields applied to the helix- coil transition of polypeptides, J. Phys. Chem. B **113**, 9004 (2009).
[26] K. Lindorff-Larsen, Improved side-chain torsion potentials for the amber ff99SB protein force field, Proteins **78**, 1950 (2010).
[27] B. Hess, P-lincs: A parallel linear constraint solver for molecular simulation, J. Chem. Theory Comput. **4**, 116 (2008).
[28] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. **79**, 926 (1983).
[29] G. Bussi, D. Donadio, and M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. **126**, 014101 (2007).
[30] M. Parrinello and A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. **52**, 7182 (1981).
[31] C. Kutzner, J. Czub, and H. Grubmüller, Keep it flexible: Driving macromolecular rotary motions in atomistic simulations with gromacs, J. Chem. Theory Comput. **7**, 1381 (2011).
[32] https://github.com/HITS-MBM/col_breaker.
[33] www.gauss-centre.eu.
[34] www.lrz.de.
[35] B. Sun, The mechanics of fibrillar collagen extracellular matrix, Cell Rep. Phys. Sci. **2**, 100515 (2021).
[36] G. A. Lichtwark and A. M. Wilson, In vivo mechanical properties of the human Achilles tendon during one-legged hopping, J. Exp. Biol. **208**, 4715 (2005).
[37] N. C. Chambers, T. W. Herod, and S. P. Veres, Ultrastructure of tendon rupture depends on strain rate and tendon type, J. Orthop. Res. **36**, 2842 (2018).
[38] N. Rezaei, A. Lyons, and N. R. Forde, Environmentally controlled curvature of single collagen proteins, Biophys. J. **115**, 1457 (2018).
[39] M. W. Kirkness, K. Lehmann, and N. R. Forde, Mechanics and structural stability of the collagen triple helix, Curr. Opin. Chem. Biol. **53**, 98 (2019).
[40] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, Entropic Elasticity of lambda-Phage DNA, Science **265**, 1599 (1994).
[41] J. F. Marko and E. D. Siggia, Statistical mechanics of supercoiled DNA, Phys. Rev. E **52**, 2912 (1995).
[42] M. J. Buehler and S. Y. Wong, Entropic elasticity controls nanomechanics of single tropocollagen molecules, Biophys. J. **93**, 37 (2007).
[43] T. Odijk, Stiff chains and filaments under tension, Macromolecules **28**, 7016 (1995).
[44] M. Hillgärtner, K. Linka, and M. Itskov, Wormlike chain model extensions for highly stretched tropocollagen molecules, J. Biomech. **80**, 129 (2018).
[45] L. Vinet and A. Zhedanov, A "missing" family of classical orthogonal polynomials, J. Phys. A: Math. Theor. **44**, 085201 (2011).
[46] O. K. Dudko, G. Hummer, and A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments, Phys. Rev. Lett. **96**, 108101 (2006).