# Collagen breaks at weak sacrificial bonds taming its mechanoradicals

Benedikt Rennekamp[1,2,3], Christoph Karfusehr[1,3,4], Markus Kurth[1,2], Aysecan Ünal[1,2,3], Kai Riedmiller[1], Ganna Gryn'ova[1,2], David M. Hudson[5], and Frauke Gräter[1,2,3,*]

[1]Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany
[2]Interdisciplinary Center for Scientific Computing, Heidelberg University, INF 205, 69120 Heidelberg, Germany
[3]Max Planck School Matter to Life, Jahnstrasse 29, 69120 Heidelberg, Germany
[4]Physics Department and ZNN, Technical University Munich, Coulombwall 4a, 85748 Garching, Germany
[5]Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, WA, 98195, USA
[*]frauke.graeter@h-its.org

October 17, 2022

**Abstract**

Collagen is a force-bearing, hierarchical structural protein important to all connective tissue. In tendon collagen, high load even below macroscopic failure level creates mechanoradicals by homolytic bond scission, similar to polymers. The location and type of initial rupture sites critically decide on both the mechanical and chemical impact of these micro-ruptures on the tissue, but are yet to be explored. We here use scale-bridging simulations supported by gel electrophoresis and mass spectrometry to determine breakage points in collagen. We find collagen crosslinks, as opposed to the backbone, to harbor the weakest bonds, with one particular bond in trivalent crosslinks as the most dominant rupture site. We identify this bond as sacrificial, rupturing prior to other bonds while maintaining the material's integrity. Also, collagen's weak bonds funnel ruptures such that the potentially harmful mechanoradicals are readily stabilized. Our results suggest this unique failure mode of collagen to be tailored towards combatting an early onset of macroscopic failure and material ageing.

## 1 Introduction

Collagen is the most abundant protein of our body and the major material of all connective tissue, from tendon to bone to skin. It is perpetually subjected to high mechanical loads, such as up to 90 MPa in stretched achilles tendon [1]. As recently discovered, under elastic deformations, tendon collagen I forms mechanoradicals originating from covalent bond rupture [2], just as in any other synthetic polymer [3, 4]. However, in contrast to synthetic materials, collagen readily stabilizes the primary radicals on dihydroxy-phenylalanines (DOPAs), with these highly efficient radical scavenging residues built into the collagen protein [2]. Where collagen I, and potentially other protein materials, rupture under mechanical load remains to be clarified. The microscopic failure mechanisms, however, define macroscopic mechanical properties such as ultimate strain and toughness. Further, the positions of the weakest links within the structure determine where and what types of mechanoradicals are formed and how they will propagate through and potentially damage

the system until they get scavenged. A tempting hypothesis for collagen is therefore that rupture occurs in the vicinity of DOPAs to rapidly stabilize the primary radical and to thereby prevent uncontrolled radical reactions such as migration and recombination. Nailing down the rupture process of collagen, being the major substituent of tissues, is a key necessity for addressing tissue degradation and ageing, and for guiding tissue engineering.

The first determinant of rupture is the relative strength of the chemical bonds. Any protein backbone features three types of bonds: the C-N peptide bond, the $C_\alpha$-N bond and the $C_\alpha$-C bond. Due to the $\pi$-conjugation across the peptide bond, this bond is an unlikely candidate for homolytic rupture. Which of the two single bonds, $C_\alpha$-N or $C_\alpha$-C, is the weaker link is less obvious. Previous studies suggest bond dissociation energies for these two bond types to be in a similar range [2, 5], so the chemical environment matters. Collagen also features enzymatically derived crosslinks connecting the individual triple helices. A variety of crosslink chemistries has evolved, which can be categorized into divalent (or pre-mature) and trivalent (or mature) crosslinks. All of them are derived from lysines, and as a consequence feature again single bonds of the C-C and C-N type (compare Fig. 1c). The thermodynamic differences in homolytic scission of collagen's backbone and crosslink bonds are decisive for collagen mechanics but currently unknown.

The second determinant of rupture is how force distributes through the complex hierarchical structure of collagen (compare Fig. 1) and loads its chemical bonds. In collagen, $\alpha$-chains wind up to triple helices that are packed in a quasi-hexagonal shape in cross-section. They are braided into each other in a non-trivial way along the fibril axis, forming fibrils on a larger scale [6]. This three-dimensional structure, combined with the crosslinking in between the triple helices, results in a molecular force distribution network that critically decides on the rupture propensity of an individual bond. In the simplest model, the Bell-Evans model, force acting on a bond exponentially increases its rate of rupture [7–9]. Our previous work identified crosslinks as well as adjacent backbone bonds as the most strongly loaded links [2, 10]. Similarly, crosslinks have been reported as responding to stress in smaller scale molecular simulations [11] as well as in coarse-grained models [12], while recent semi-empirical calculations identified $C_\alpha$-C bonds of prolines at X-positions as rupture candidates [13]. None of the previous studies allowed quantitative conclusions on specific bond rupture sites within the collagen fibril.

In the field of polymer mechanochemistry, selective bond rupture has been achieved by synthetically incorporating weak bonds into the polymer, as sacrificial bonds increase the toughness of the material [14], or as mechanophores to report on mechanical force [15, 16]. Sacrificial bonds that so far have been observed in biological systems are non-covalent in nature, such as hydrogen bonds or ionic bonds in bone collagen [17, 18]. If protein materials harness the virtue of mechanically weak covalent bonds as sacrificial bonds or mechanophore, is to our knowledge currently unknown. In collagen, selective bond scission could critically determine mechanical stability, further radical-mediated damage, and modes of mechano-sensing.

Here, we employ comprehensive and high-level quantum mechanical calculations, scale-bridging hybrid Molecular Dynamics simulations for reactive dynamics (KIMMDY [10]), and gel electrophoresis-coupled mass spectrometry (MS) of stressed collagen to assess collagen's mechanochemical rupture. Our computations identify crosslink bonds as weakest links and major rupture sites, in particular $C_\alpha$-$C_\beta$ bonds due to the stabilization of $C_\alpha$-centered radicals by the captodative effect. We observe a large concentration of ruptures of $\sim$70 % in these bonds, the two key factors of which are $\sim$120 kJ mol$^{-1}$ lower bond dissociation energies of crosslink compared to backbone bonds, and the higher stress concentration in this region. As a result, most rupture sites are in direct vicinity to DOPA residues which can prevent uncontrolled radical reactions by rapid scavenging. However, there is a strong competition with unspecific breakages in the vast amount of backbone bonds throughout the fibril. To still direct rupture into the comparably few crosslink bonds, these bonds have to be substantially weaker. SDS-PAGE coupled to MS supports altered collagen crosslinking, as well as a widespread rupture within both $\alpha$1 and $\alpha$2-chains in collagen I, as evident from the redistribution of molecular masses in collagen upon mechanical stress. Our simulations find a particular bond in one of the

two 'arms' of the trivalent crosslink Pyridinoline (PYD) and Deoxypyridinoline (DPD) to be particularly primed for rupture, due to a stabilizing conjugation of the resulting radicals. Its rupture does not compromise collagen integrity and leads to local stress relaxation, a hallmark of a sacrificial bond.
Overall, we show both experimentally and computationally that the conjunction of topology and chemical structure of collagen by design funnels bond ruptures towards crosslinks and specific sacrificial bonds therein, very much akin weak sacrificial bonds in synthetic polymers. We propose that collagen has evolved into a high-performance biomaterial by means of highly specific mechanochemical routes that enable collagen to self-report on rupture while preventing further unspecific damage.

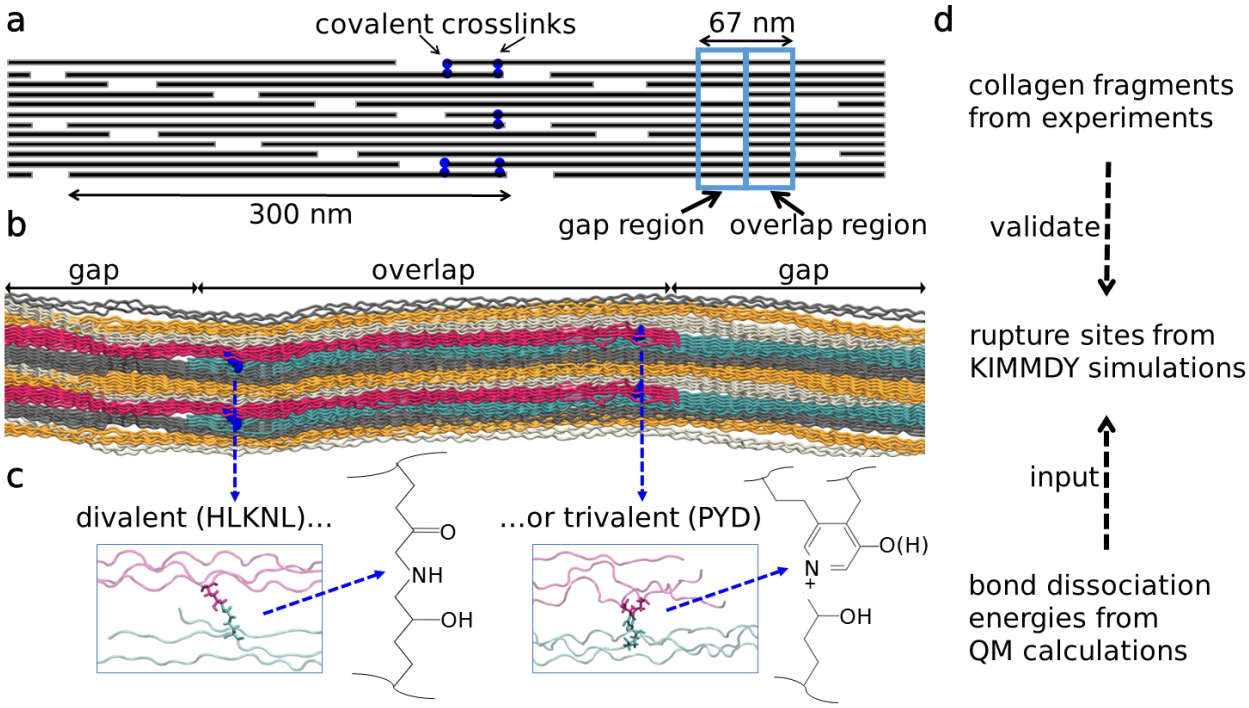


Figure 1: **Collagen structure, spanning multiple length scales, and our corresponding methods**. **a** 2D projection of the staggered arrangement of collagen triple helices, each 300 nm. This results in the typical overlap and gap regions of collagen, inlcuding the 3D braiding of triple helices. **b** Our atomistic model spanning one overlap (middle) and about one gap region (split into two parts) of collagen. **c** Zoom in on enzymatic crosslinks connecting the triple helices. Different chemistries (divalent or trivalent) are possible at these positions. **d** Our workflow, combining different methods as collagen spans multiple length scales.

# Results

## Crosslinks harbor the weakest bonds

To determine the stability of covalent bonds in collagen I and identify the thermodynamically weakest links, we performed high accuracy QM calculations on the G4(MP2)-6X level of theory, which reaches an accuracy similar to that of G4 for thermochemical properties [19]. We obtained bond dissociation energies (BDEs) of bonds from four major collagen I crosslinks (trivalent: PYD/DPD, divalent: HLKNL and deH-LNL), as well as of backbone bonds of collagen's main amino acids (glycine, proline and alanine). Further, we calculated radical stabilization energies (RSEs) of homolytic bond scission products to examine the origin of the large differences in BDEs, which can exceed 160 kJ mol$^{-1}$ across C-C bonds within a single crosslink.

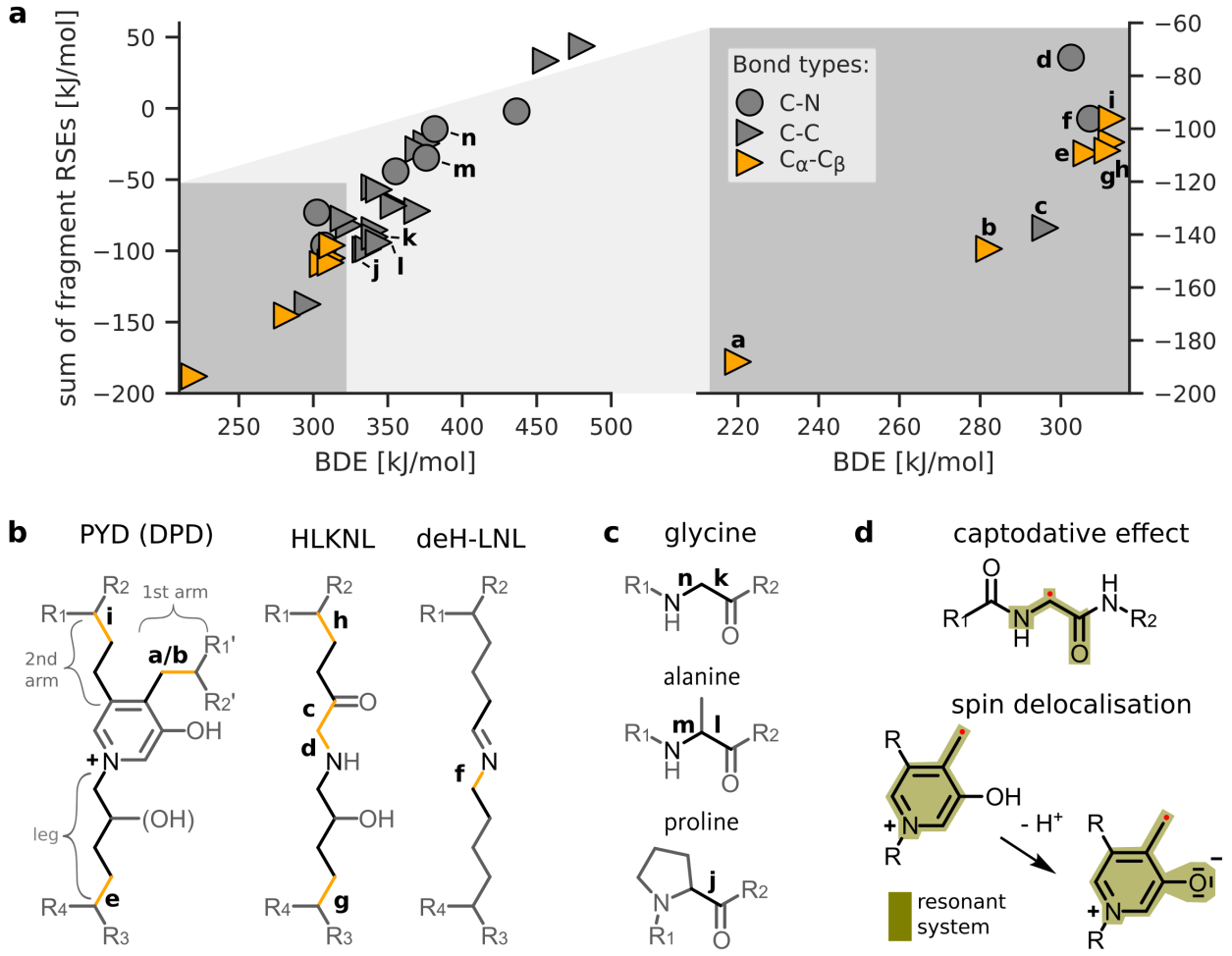


Figure 2: **Crosslinks harbor the weakest bonds in collagen type I**. **a** Correlation between the BDEs and summed RSEs of homolytically cleaved bonds and formed radicals. All molecular structures used for calculations can be found in the SI. Bonds with BDEs $< 315\,\mathrm{kJ\,mol^{-1}}$ are shown in more detail on the right. The small letters close to selected data points indicate respective bonds in the lower panels b and c. **b,c** Molecular structures of (b) investigated collagen crosslinks and (c) prevalent amino acids in collagen. BDEs of bonds colored in black and orange (BDE $< 315\,\mathrm{kJ\,mol^{-1}}$) were explicitly calculated. Residues R1, R1', R2 and R2' connect to the same collagen triple helix. The BDEs of the PYD bond denoted with letter "a" was obtained after deprotonation of the pyridine bound hydroxyl group. **d** Main electronic effects leading to the lowest found BDEs. The captodative effect describes the non-additive stabilization of the amino-carbonyl and carbonyl-amino group present at any $C_\alpha$ peptide radical. The lower part illustrates the increased radical delocalization upon deprotonation of PYD and DPD, lowering the BDE and RSE of bond "b" to that of bond "a".

Figure 2a shows the BDEs of all calculated bonds against the summed RSE of the two radicals formed by each homolytic bond scission. All values can be found in Suppl. Tables 1-10. Overall, the linear correlation between low BDEs and large negative summed RSEs shows that BDE deviations from aliphatic C-C bonds can generally be explained by stabilization of the formed radical centres by proximate chemical groups.

Across collagen crosslinks and backbone bonds, our results suggest that $C_\alpha$-$C_\beta$ bonds are promising bond rupture candidates (having BDEs $\leq 312\,\mathrm{kJ\,mol^{-1}}$). The high stability (low RSE) of the formed $C_\alpha$ centered radical of $-87\,\mathrm{kJ\,mol^{-1}}$ explains the low $C_\alpha$-$C_\beta$ BDE, as do comparable literature values ($-74.1\,\mathrm{kJ\,mol^{-1}}$ [20] and $-75.5\,\mathrm{kJ\,mol^{-1}}$ [5]). We attribute the comparably low RSE to the stabilizing effects of carbonyl-amino and amino-carbonyl substituents present in peptides, which are further increased when combined. Such non-linear donor/acceptor interactions are known as captodative effects (Fig. 2, Panel d) [21, 22].

Interestingly, the presence of a second strong radical stabilizing system on the $C_\beta$ centered radical further lowers the BDE (Fig. 2, bond "b") in the trivalent crosslinks PYD and DPD to just $282\,\mathrm{kJ\,mol^{-1}}$. The stabilizing effect of the conjugated aromatic system becomes even more pronounced upon deprotonation of the pyridine bound hydroxyl group, extending the resonant system and giving an exceptionally low BDE of $220\,\mathrm{kJ\,mol^{-1}}$. Importantly, the pKa value of a reference compound (3-Hydroxypyridine methochloride, pKa = 4.96 [23]) suggests a pKa of about 5, implying that PYD and DPD crosslinks are mostly deprotonated at neutral pH (and even more so their respective radicals).
We note that the electronic structures of radicals conjugated with the aromatic ring in PYD and DPD showed spin contamination, indicating multi-reference natures (see further tests and discussion in the Supplementary Information, in particular Suppl. Fig. 10). As the decreased BDEs of bond "a" and bond "b" appear plausible compared to literature RSEs and chemical intuition, we proceeded with these values but also varied PYD BDEs in the following analysis to assess the resulting uncertainty of rupture propensities in the fibril. Also note that throughout this study, we do not differentiate between low BDEs of PYD and DPD, as calculation of four DPD bonds convinced us that all relevant BDEs are similar (details given in SI).
Focusing on divalent crosslinks HLKNL and deH-LNL, we found (additional to $C_\alpha$-$C_\beta$ bonds) two more weak bonds in HLKNL and one more in deH-LNL. Again, the low BDEs can be understood by the strong radical stabilization effects of neighboring groups. For example, both fragments of the weak $C_\delta$-$C_\epsilon$ bond of HLKNL have large RSEs of $-68$ and $-70\,\mathrm{kJ\,mol^{-1}}$ (literature: $-65.3\,\mathrm{kJ\,mol^{-1}}$ [24]). Furthermore, cleavage of the second weakest bond in HLKNL, $C_\epsilon$-$N_\zeta$, forms radicals stabilized captodatively by a carbonyl group in beta position, as well as an N-centred radical with a moderate RSE of $-30\,\mathrm{kJ\,mol^{-1}}$ (literature: $-25\,\mathrm{kJ\,mol^{-1}}$ [5]).
Notably, none of the backbone bonds appear as likely rupture candidates. All calculated BDEs of backbone bonds, both $C_\alpha$-C and $C_\alpha$-N bonds, are considerably higher than those of the weak crosslink bonds ($\Delta$BDE $\geq 24\,\mathrm{kJ\,mol^{-1}}$). Within the backbone, $C_\alpha$-C bonds are overall weaker than $C_\alpha$-N bonds and are the primary rupture candidates of the protein backbone, in agreement with previous semi-empirical calculations [13], here largely independent of the amino acid type.

In essence, our quantum chemical calculations identify bonds within crosslinks as weak bonds within collagen, in particular $C_\alpha$-$C_\beta$ bonds reaching values down to $220\,\mathrm{kJ\,mol^{-1}}$ in the trivalent PYD crosslink. These can be backtraced to excellent stabilization of the resulting mechanoradicals.

## Ruptures occur dominantly in and around crosslinks, but also unspecific in backbones

Rupture propensity is determined not only by the thermodynamic strength of the bonds, as calculated above, but also by the actual force that is acting upon that bond. To directly predict relative rupture counts within a stretched collagen fibril, we utilized KIMMDY [10], a hybrid simulation scheme that invokes bond rupture in kinetic Monte Carlo steps during atomistic MD simulations. Bond rupture rates are calculated based on the bond's strength given by its BDE and its mechanical weakening by the force distribution through the

protein in the MD simulation. For the simulations, we use atomistic collagen type I fibril models taken from ColBuilder [25] as described in the methods and shown in Fig. 1b. The simulated structures comprise one overlap and gap region in length and 41 triple helices in width (in the overlap region).

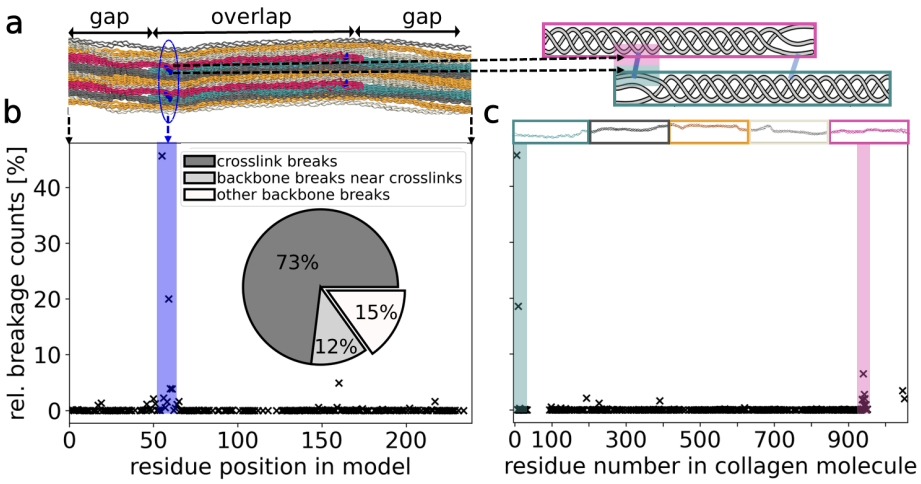


Figure 3: **Collagen ruptures primarily in crosslinks, but also unspecific ruptures occur everywhere**. Data was obtained from 63 independent MD simulations including divalent and trivalent crosslink setups, with different collagen sequences and pulling conditions, as described in the methods section. For simulations with trivalent crosslinks, we used the deprotonated BDEs of PYD. In both panels b and c the same data is presented in two different reference systems. **a** Left side: Location of residues and crosslinks in our model along the fiber axis. Triple helices with the same phase are in the same color, respectively. Upper and lower residues participating in the same crosslinks (blue) appear in the same residue position in the model, see for example the blue encircled area. This view is used in Panel b. Right side: Crosslinks along a whole 300nm collagen molecule occur at the beginning (cyan area) and at the end of a triple helix (magenta area), shown as cartoon with coloring as in the structure on the left. This reference system, counting along the collagen molecule, is used in Panel c. **b** Propensity of bond breakages in our collagen model. Most ruptures concentrate in the crosslinked area, while there are some scattered backbone ruptures. Inset: Pie chart of summed up ruptures in the crosslinks vs. the backbones in the crosslinked area (up to 5 residues before/behind) vs. elsewhere in the backbone. **c** Propensity of bond breakages within collagen fibrils. The crosslinked areas at the beginning and end of the molecules, as marked exemplary for the N-terminal crosslink in the shaded areas, are most prone to rupture. Note that both in b and c we only show data points where at least one rupture happened.

Notwithstanding the overall higher stability of the backbone bonds according to the quantum chemical calculations, KIMMDY predicts that crosslink and backbone bond ruptures can both occur. Along the fibril structure (Fig. 3a, left), ruptures, albeit at low relative counts, occurred randomly across the whole fiber axis (Fig. 3b). Most frequently, however, the crosslinks themselves break (at 73 %). Also, the backbone in its vicinity has a strongly increased rupture propensity due to local stress concentration (12 %). When mapping the rupture statistics onto the collagen I sequence (Fig. 3a, right), the prevalence of crosslink ruptures over widely distributed backbone ruptures again is evident (Fig. 3c). Only backbone regions where two crosslinked strands are overlapping and thus share the external stress, i.e. at the very beginning and end, are largely void of ruptures (areas without crosses in cyan and magenta boxes). Thus, the external force distributes such that some of the backbone bonds, in particular those in vicinity of crosslinks, are weakened sufficiently to compete in terms of bond rupture rates with the intrinsically weaker crosslink bonds. In addition, backbone bonds are statistically still prone to rupture due to their sheer number, with up to ∼1,500 aminoacids per crosslink in a fully crosslinked collagen I fibril.

The results shown in Fig. 3 are compiled from a set of KIMMDY simulations varying in crosslinks, collagen sequence (different species) and pulling conditions (see Methods). Importantly, in all of the models and settings, we recover a competition between crosslink and backbone bond ruptures, i.e ,a concentration of ruptures at and around crosslinks, and the random distribution of the remaining backbone ruptures along the sequence (Suppl. Figs. 6-7). However, quantitative rupture counts can vary strongly and also depend on model choice and composition. For instance, if we assume acidic pH, that is, 100% protonated PYD, we have a lowest BDE of $282\,\mathrm{kJ\,mol^{-1}}$ instead of $220\,\mathrm{kJ\,mol^{-1}}$ (Fig. 2a, bond b versus a). In this case, the ratio shifts to 28%:41%:31% of breakages in crosslinks to those in adjacent backbone bonds to other unspecific backbone bonds, compare Suppl. Fig. 6c. Hence, as crosslinks now feature bonds strengths closer to those of the backbone, ruptures throughout the backbone can compete with those in crosslinks under the dynamic stress concentrations present in the stretched fibril. To conclude, crosslinks indeed represent weak spots, not only in terms of their BDEs but also in the context of the stress distribution across the collagen structure. The weaker the crosslinks are compared to the backbone, the more they can prevent any other unspecific position in the protein backbone.

## Trivalent crosslinks harbor sacrificial bonds

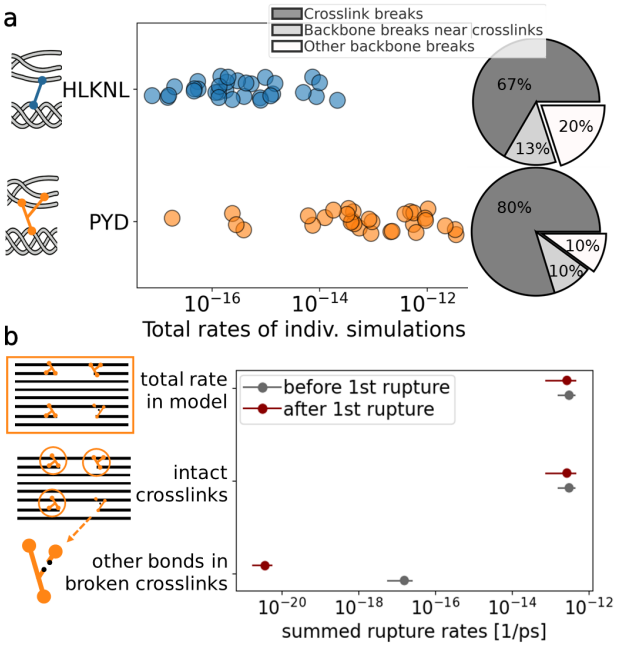


Figure 4: **Trivalent crosslinks break faster, but (at first) without loosing connection. a** Comparison of total primary bond rupture rates in simulations with models having divalent (HLKNL) vs. models having trivalent (PYD) crosslinks, which show overall higher rates. The ratio of crosslink ruptures to other breakages confirms that the increased rupture rates are indeed due to higher propensities in the crosslinks. **b** Comparison of secondary bond rupture rates before and after the first arm in one PYD crosslink broke. The rate in the second arm even decreases, being several orders of magnitude lower compared to the sum in the remaining model. Hence, other bonds will break first and the fibril stays overall intact. This data is averaged from 9 simulations with trivalent crosslinks, which we continued after a first rupture in the short PYD arm. Error bars are errors of the mean.

The trivalent crosslinks become more prevalent with age and connect two 'arms', i.e. two modified lysine

side chains of different strands from within one triple helix, to one lysine residue ('leg') of another triple helix (compare Fig. 2b). It has been speculated that due to this double connection, trivalent crosslinks provide more stability yet less elasticity. Having established crosslinks as major ruptures sites by our combined calculations, we next asked how the rupture propensity of the trivalent crosslink compares to its divalent counterpart (chemical structures shown in Fig. 2b), and how the redundant connectivity of the trivalent crosslink is compromised by rupture.

At first glance, the distribution of ruptures does not depend heavily on the crosslink type, with the exception of an even stronger concentration in the crosslinks for PYD. For the detailed rupture distributions, compare Suppl. Fig. 6b for PYD to Suppl. Fig 11b for HLKNL. In Fig. 4a, however, we summed up all bond rupture rates within one simulation and grouped these total rates per crosslink type that we used in the model. The models with trivalent (PYD) crosslink exhibit clearly higher rupture rates. While this appears counter-intuitive at the first glance, as the two arms can share the external mechanical load, the explanation is straightforward: the short arm of the trivalent crosslink has by far the lowest BDE due to the combination of the captodative effect and conjugation with the aromatic ring (compare Fig. 2d). While this leads to a higher rupture propensity for a first breakage, the overall crosslink remains intact due to the second arm. The low BDE of this bond also renders ruptures across the collagen fibril highly localized to the crosslink, with 80 % ruptures occurring therein (Fig. 4a). Unspecific backbone ruptures halved compared to collagen with divalent crosslinks, from 20 % to 10 % of the rupture counts.

Rupture of a trivalent crosslink on one of the two arms leaves the linkage of the two triple helices overall intact. We next examined if subsequent ruptures occur primarily within the same crosslink or rather within other fully intact trivalent crosslinks. To this end, we continued 9 pulling simulation of models with trivalent crosslinks after a rupture inside these arms, examined the new rupture rates and compared to the situation of a fully intact pre-rupture fibril (Fig. 4b). Surprisingly, despite now lacking one force-bearing connection, the total bond rupture rate in the damaged trivalent crosslink decreases by $\sim$3 orders of magnitude compared to the pre-rupture simulations, indicating that stress in this crosslink overall relaxed after its rupture. Indeed, we find an average distance increase of the crosslink $C_\alpha$ atoms after rupture by $\sim$0.1 nm, as the longer arm can now elongate and use the previously hidden length (Suppl. Fig. 8).

Also, the remaining intact crosslinks that still each harbor the extraordinarily weak bond, dominate the total of all rates in the simulation. Another rupture in the damaged crosslink is highly unlikely, and this crosslink stays intact. Instead, subsequent ruptures are predicted to mainly occur in the same type of weak $C_\alpha$-$C_\beta$ bonds in other triple helices. Our combined data in Fig. 4 suggests that the PYD and DPD crosslinks are less stable with respect to first ruptures, without compromising fibril integrity.

## Experiments confirm covalent backbone rupture and altered crosslinking

We next set out to experimentally validate our predictions of both crosslink and backbone covalent bond ruptures in collagen. Electron-paramagnetic resonance (EPR) spectroscopy is a highly sensitive method for detection of mechanoradicals, and has been used for this purpose successfully [2], however it does not allow spatial resolution and in addition rapid migration onto DOPA prohibits the primary rupture site identification. We instead resorted to gel electrophoresis combined with mass spectrometry. Fig. 5a shows the protein molecular weight distribution of solubilized collagen polypeptides from stressed and unstressed (control) collagen rat tail tendon resolved by an SDS-PAGE gel. We recovered the expected band pattern, including the monomeric $\alpha$-chains ($\alpha$1(I) and $\alpha$2(I)), as well as higher order oligomers (e.g. $\beta$-chains, $\gamma$-chains and $\delta$-chains). For the stressed samples, we additionally observe bands in the mass regime of fragmented chains, indicating backbone rupture of single collagen strands. To quantify the proteolytic breakdown in stressed and control tissues, we obtain the densitometry of stained collagen chains and peptide fragments across samples with ImageJ [26] in Fig. 5b. Stressing the samples gives rise to fragments of masses lower than those of full-length collagen chains, across a wide range of masses down to small fragments running at the dye front.

Second, we identified intact $\alpha$-chains and lower molecular weight polypeptides by mass spectrometry. We

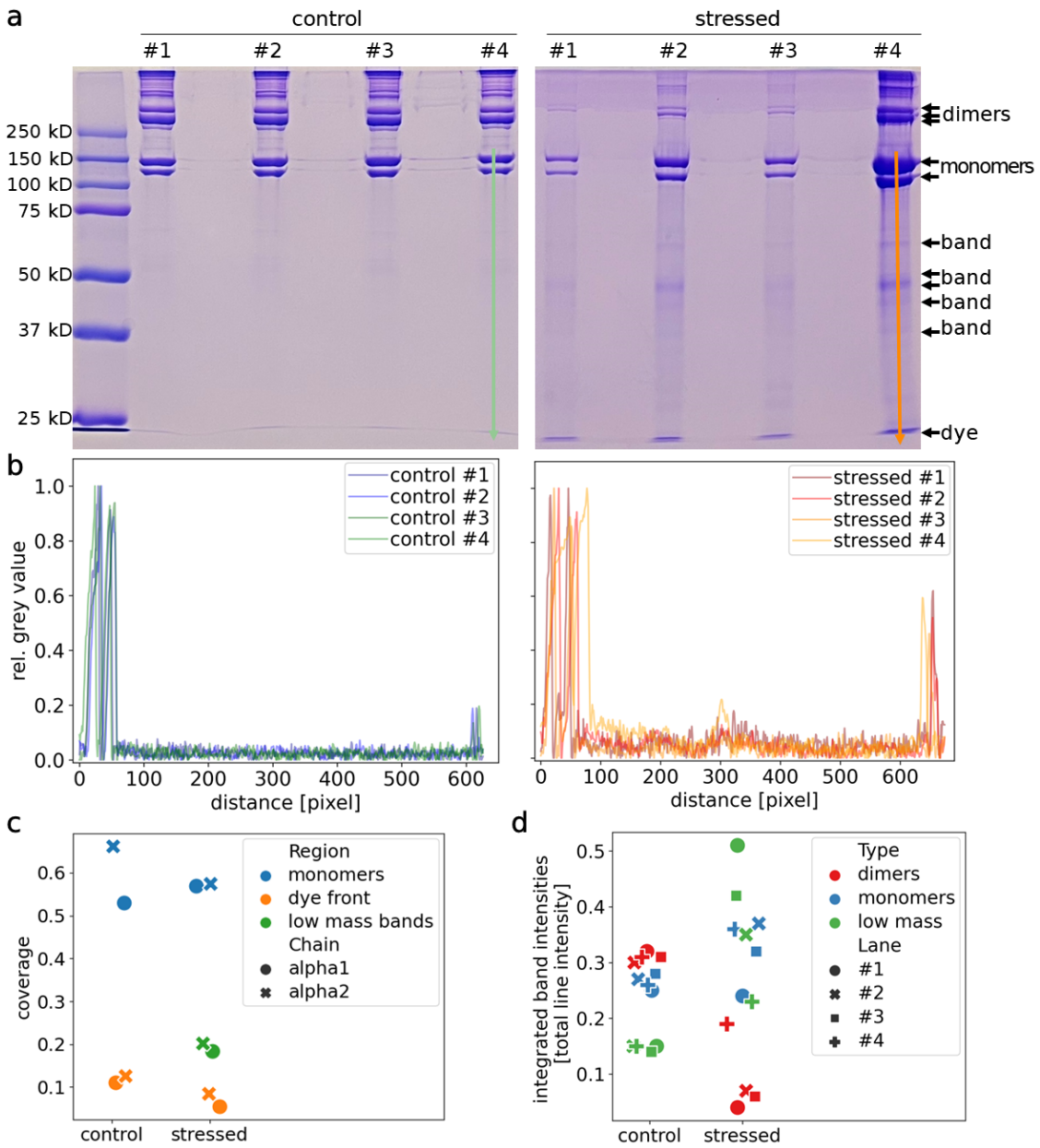


Figure 5: **SDS-PAGE and mass spectrometry analysis support the theoretical simulations of covalent backbone and crosslink rupture in stressed tendon collagen. a** SDS-PAGE analysis reveals an altered banding pattern between stressed and control tendon collagen. Sample loads were normalized to the dry weight of tendon prior to loading. We indicated a line as used for the analysis in panel b lanes #4. The bands that we cut out for the mass spectrometry in panel c are marked with arrows. **b** Densitometry to quantify stained collagen $\alpha$-chains and degradation products on SDS–PAGE. We show the inverted grey value, normalized to the maximum peak height, along the different gel lanes. We note that the stressed samples show higher intensity in the mass range of broken collagen fragments below 100 kDA. **c** Percent sequence coverage of collagen $\alpha$1(I)- and $\alpha$2(I)-chains from the bands analyzed by mass spectrometry. We find collagen peptide fragments both in stressed and control samples in the mass regime of monomers, low mass bands and even the dye front. **d** Integrated intensity in the two monomeric bands (yellow) and three dimer (red) bands, as well as the total area below the monomers (incl. dye front, green), normalized for each lane individually with the total intensity to account for overall differences in brightness. In the stressed samples, the relative amount of dimers decreases indicating that they have turned to monomers.

took them from regions as indicated in Fig. 5a from stressed samples #2 and #4, because peptide fragments from samples #1 and #3 were below the limits of detection. The sequence coverage of the collagen $\alpha$-chains, obtained with Proteome Discoverer and displayed in Fig. 5c, confirms the presence of both $\alpha$1- and $\alpha$2-chain fragments in all these areas, even in the lowest mass fractions of the dye front (see Suppl. Fig. 9 for coverage across the collagen sequence). Overall, this experimental analysis supports the unspecific breakages found in the theoretical analysis (compare Fig. 3c), as the stressed samples contained multiple $\alpha$-chains peptide fragments indicative of unspecific backbone breakages in the fiber. Given that the intensity of low molecular weight bands increases upon mechanical stress, we attribute this indeed to the mechanical loading.

Lastly, we used variations in the collagen banding pattern as observed by SDS-PAGE as a measure of collagen crosslinking quality. Here, we test our hypothesis of loss of connection between the molecules by rupture of crosslinking upon mechanical stress, as predicted by our simulations (Fig. 3b). Thus, we expect to observe less oligomers ( $\beta$-, $\gamma$- and $\delta$-chains) and instead more monomers. Indeed, there is precedent for predicting alterations in collagen crosslinking quality through changes in banding patterns, whether caused by chemical treatment or genetic disorder [27–29]. It is apparent in Fig. 5a that the concentration of oligomers is decreased in the stressed samples. We quantify the integrated intensities of dimers and monomers in each lane in Fig. 5d, after normalizing to the overall intensity to account for the different signal strengths in the lanes. In the control, dimers account for slightly more of the total intensity in their lane than monomers. In contrast, after mechanical treatment, dimer intensity decreases, and the monomeric $\alpha$-chains make up a higher fraction in the stressed samples. Hence, stress on the tissue disrupted the crosslinks typically found in these higher molecular weight collagen polypeptides. Again, this is consistent with the crosslink breakages predicted from our simulations.

Together, the experimental evidence of both backbone rupture and altered crosslinking supports the detailed simulation results. The increase in lower molecular weight fractions is indicative of unspecific backbone ruptures, and the shift from oligomeric combinations to monomers can be explained by ruptures of crosslinks that otherwise held those oligomers together.

## Discussion and Conclusions

Following the hierarchical structure of collagen, we identified collagen type I breakage sites by means of a scale-bridging approach. High-level quantum-chemical calculations returned bond dissociation energies that we incorporated into KIMMDY, yielding rupture statistics, which in turn were validated by experiments. The KIMMDY approach is more efficient than hybrid quantum mechanical/molecular mechanical (QM/MM) techniques [30, 31], especially given that the initial rupture sites are unknown and the whole molecule can potentially react. Further, having now detailed BDEs at hand, it is also more accurate than reactive force fields [32], which would also not scale well to these large atomistic models.

We observe the rupture propensity of covalent bonds in collagen to be determined by a competition between how weak a bond intrinsically is, that is, low BDEs, and how much force acts on this bond, that is, the stress distribution in the fibril. We find this interplay to depend on various conditions, most importantly the crosslink type (and BDEs therein) and the external force application, but also on computational details in KIMMDY as described in the methods section. However, notwithstanding this sensitivity of quantitative results with regard to parameters, a robust picture has emerged: Crosslinks harbor the weakest bonds. In particular the captodative effect stabilizing $C_\alpha$-centered radicals, and, in trivalent crosslinks, the additional resonance with the aromatic ring give rise to exceptionally low BDEs. Hence, they break most often. However, the stressed protein backbone still breaks randomly as well due to stress localization, but also simply due to the vast amount of bonds under force. The lower the difference of backbone and crosslink BDE, the higher the share of backbone ruptures in that competition. We could also observe both backbone ruptures and altered crosslinking in our SDS-PAGE gels.

Protein $C_\alpha$ radical stabilization by the captodative effect was found before, e.g. when inducing radical formation via UV light [33], or by mass spectrometry experiments [34]. These studies support our result, and we here extend the decisive role of the captodative effect in proteins to mechanical bond rupture and mechanoradical formation in a stretched protein material.

The stable methylpyridine radical in trivalent crosslinks extends the set of aromatic moieties known to stabilize radicals in proteins, namely histidine, tyrosine, tryptophan, and DOPA, as e.g. in ribonucleotide reductase [35–37], or collagen [2]. The benzylic methylpyridine radical is chemically remarkably similar to Vitamin B6, with well-known radical scavenging properties [38, 39]. A DOPA radical, however, is even more stable, with a BDE for hydrogen abstraction (X-H) as low as $260\,\mathrm{kJ\,mol^{-1}}$ for a DOPA anion [40]. We here determined a BDE(X-H) of around $337\,\mathrm{kJ\,mol^{-1}}$ for the deprotonated PYD, hence, in the context of the collagen structure, radical migration to DOPA seems likely from a thermodynamic viewpoint.

Our data propose several advantages of the unique trivalent crosslinks of collagen. First, its rupture results in exceptionally stable radical species, reducing unspecific reactions such as uncontrolled migration. The stabilizing effect is enhanced by the localization of breakages, as low crosslink BDEs funnel ruptures to sites where a rapid transfer to nearby DOPAs is likely. Thus, further damage to the material is mitigated. Secondly, we find the mechanical strength of the trivalent crosslink to be pH-sensitive, i.e. BDEs depend on the protonation state of PYD (Fig. 2). We speculate that collagen's rupture propensity could adapt to variations in physiological pH, e.g. in muscle after exercising [41, 42] or under inflammatory conditions [43]. Thirdly, we find the weak bonds within the trivalent crosslinks to act as sacrificial covalent bonds. Stress relaxes in the trivalent crosslink after initial rupture, and subsequent ruptures occur in other crosslinks, according to our computational analyses. Mechanically, the sacrificial arm in the trivalent crosslink can increase the toughness of the material due to energy dissipation and relaxation afterwards, similar to what is known from sacrificial bonds in polymers [17]. Finally, radicals formed in crosslinks can migrate rapidly to nearby DOPAs and then in presence of water react with superoxide radicals to create $H_20_2$ as has previously been shown in pulling experiments on collagen fibers [2]. In this way, the breakage of one of the two arms offers a possibility to already start signaling high loads in collagen, triggering repair mechanisms and other reactions, even before macroscopic failure occurs. Speculating further, a self-healing effect similar to what can be observed in polymers [44] or biological fibers [45] could also come about by either direct DOPA recombination or by the formation of DOPA-Lys and DOPA-His crosslinks [46, 47].

Collagen's remarkable trivalent crosslink chemistry appears perfectly adapted towards handling mechanical stress and mechanoradicals, and as such can inspire the design of synthetic (biocompatible) material systems with high toughness, self-healing, or self-reporting strategies. If this specific mechanochemistry is similarly at play in other biomaterials such as elastin, and to what extent it plays a role in tissue ageing, remains to be elucidated.

# Methods

## Quantum-chemical calculations

Molecular structures were built in GaussView. Initial phi and psi values of amino acids were set to common values for the respective amino acid in a Ramachandran plot, based on a stretched collagen MD simulation. All radical structures were built from already geometry optimized molecular structures. H-capped radicals, used for radical stabilization energy calculations, were built by hydrogen addition to geometry optimized radicals. All structures were geometry optimized at the BMK/6-31+G(2df,p) level of theory in Gaussian09, following a slightly modified G4(MP2)-6X protocol described elsewhere [19]. Before calculation of G4(MP2)-6X energies, structures were assured to have converged and be free of imaginary frequencies. More detailed information can be found in the Supplementary Information.

## BDE and RSE calculations

Bond dissociation energies (BDEs) were calculated as described in Equation 1 based on G4(MP2)-6X enthalpies ($\Delta H_{298\mathrm{K}}$) of a given molecule ($\mathrm{R_1}$-$\mathrm{R_2}$) and its fragments formed after homolytic bond scission ($\mathrm{R_1^*}$, $\mathrm{R_2^*}$).

$$\mathrm{BDE} = \Delta H_{298\mathrm{K}}(\mathrm{R_1^*}) + \Delta H_{298\mathrm{K}}(\mathrm{R_2^*}) - \Delta H_{298\mathrm{K}}(\mathrm{R_1\text{-}R_2}) \tag{1}$$

Radical stabilization energies (RSEs) show how well a radical is stabilized relative to the reference methyl or ammonia radical. They were calculated with BDEs of $CH_3$-H bonds for C-centred and $NH_2$-H bonds for N-centred radicals. G4(MP2)-6X enthalpies of radicals ($\Delta H_{298\mathrm{K}}(R_1^*)$, $\Delta H_{298\mathrm{K}}(\mathrm{Ref}^*)$) and their H-capped counterparts ($\Delta H_{298\mathrm{K}}(R_1\text{-}H)$, $\Delta H_{298\mathrm{K}}(\mathrm{Ref}\text{-}H)$) were used to obtain RSEs as given in Equation (2).

$$\begin{aligned}\mathrm{RSE}(\mathrm{R_1^*}) &= \mathrm{BDE}(\mathrm{R_1}\text{-}H) - \mathrm{BDE}(\mathrm{Ref}\text{-}H) \\ &= \Delta H_{298\mathrm{K}}(\mathrm{R_1^*}) - \Delta H_{298\mathrm{K}}(\mathrm{R_1}\text{-}H) - \Delta H_{298\mathrm{K}}(\mathrm{Ref}^*) + \Delta H_{298\mathrm{K}}(\mathrm{Ref}\text{-}H)\end{aligned} \tag{2}$$

## Molecular Dynamics simulations

All MD simulations were carried out with the software package GROMACS, version 2018 or 2020 [48], using the amber99sb-star-ildnp force field [49, 50]. We employed 2 fs time steps with LINCS [51] constraints on h-bonds. In order to account for the asymmetric stretch of bonds under these high mechanical loads and also in order to enable bond scissions, we used Morse potentials for the bond interactions. Simulations were conducted at 300 K and 1 bar, which were maintained by a v-rescale thermostat [52] and a Parrinello-Rahman pressure coupling [53], respectively. We cut off both Lennard-Jonas and Coulomb interactions at 1.0 nm. The different collagen models were solvated in TIP3P water leading to system sizes of approx. 2.5 - 2.9 million atoms, for which periodic boundary conditions were chosen. After energy minimizations, we conducted equilibrations in NVT and NPT ensemble for 10 ns each.

For production, we subjected the collagen fibrils to a constant force of 1 nN per chain in four different schemes: One is that we distributed the forces equally leading to 3 nN on each triple helix, pulling from both sides. In schemes two and three we also pulled equally on all strands, but just from one side and restrained the other side of the molecule. We used this as a control that our results are independent of the time the force needs to propagate through the system and of internal friction.

In the last scheme, we inhomogenously distributed forces on the inner triple helices by drawing forces from a Gaussian distribution with the same average of $F_{\mathrm{av}} = 1\,\mathrm{nN}$ and a width of $\sigma = F_{\mathrm{av}}/3$. The ring of outer helices still was subject to 1 nN to avoid a sliding of the chains against each other. We refer to this scheme as shear pulling, it mimics more imperfect biological samples. In all cases, we additionally employed torque restraints to prevent unwinding of the capped collagen triple helices [54].

Overall, we used collagen fibril models from 3 different species Rattus Norvegicus, Pongo Abelii and Loxodonta Africana (as published on the webserver ColBuilder [25]), which yet again varied in their crosslink position and also crosslink type (divalent HLKNL or trivalent PYD). Including replica, we in total conducted 63 simulations. In each instance, the collagen fibrils were pulled for 100 ns before executing the KIMMDY scheme of bond rupture as discussed below. For 9 simulations with PYD crosslinks, we thereafter conducted a second cycle of 100 ns pulling to investigate secondary breakages. All models and their replica taken together yield a combined total simulation time of more than 7.2 $\mu$s in this manuscript.

## Bond breakages in KIMMDY

We use our previously developed simulation scheme **Ki**netic **M**onte Carlo / **M**olecular **Dy**namics (KIMMDY) [10] to enable bond scissions on the molecular level. In short, it takes the bond elongations from the MD simulations as input for a Bell-type model to calculate rupture rates. The rates are used in a Kinetic Monte Carlo step determining the scission site and the associated propagation of the system.

After the aforementioned 100 ns pulling MD simulations, 5,000 bond elongations, taken from another 1 ns of simulation now using a PLUMED-patched [55] Gromacs version, per breakage candidate were averaged to generate the bond rupture rates as described previously [10]. For each of these ensembles of rates, the Kinetic Monte Carlo step was invoked 10,000 times to sample the rupture propensities properly. In the cases where a second rupture was simulated, which we did for 9 simulations with PYD crosslink, we adjusted the topology and relaxed the system with a small 0,2 fs time step before pulling for another 100 ns and repeating the procedure.

In contrast to the original implementation which used default backbone values for BDEs and did not differentiate between the various crosslink chemistries, we here utilize the BDEs that we calculated in this manuscript as barrier heights. Concerning peptide backbone bonds, we found that $C_\alpha$-N BDEs of alanine and glycine are identical within methodological errors. The same was the case for $C_\alpha$-C bonds of alanine, proline and glycine, even if we would expect some difference due to different chemical environments given by an amino acid's side chain. For subsequent simulations, we used the resulting averages of 379 and 341 $\mathrm{kJ\,mol^{-1}}$ for all $C_\alpha$-N and $C_\alpha$-C bonds, respectively. Note that this is different from our previous work, where we assumed that $C_\alpha$-N and $C_\alpha$-C bonds have similar BDEs. Hence, beside the collagen-specific crosslinks, our method KIMMDY also improves generally for other protein use cases with these new backbone values.

Secondly, we also revised the way KIMMDY calculates forces. We noticed that rupture probabilities varied depending on the amino acid type. While these differences did not qualitatively change our results, we carried out a detailed investigation on this as can be found in the supplement, see in particular Suppl. Fig. 11. In summary, we attribute this effect partially to influences of the force field and partially to actual differences in the systems, which can hardly be separated due to the nature of the KIMMDY protocol.

## Experiments

Rat (Rattus Norvegicus) tails were sampled earlier from female rats, 5-6 months old, snap-frozen and stored at -80 °C until the day of preparation. Tendon samples from the tail were excised by pulling out single collagen fibers. The fibers were collected into ice-cold Dulbecco's phosphate-buffered saline (DPBS) w/o calcium and magnesium. Collected samples were then briefly washed by over-head agitation at 4 °C and collected by centrifugation (500 rcf, 2 min, 4 °C). This step was repeated twice. After washing, the fibers were equilibrated at atmosphere for 1 h. Dried fibers were crushed in a porcelain mortar, cooled and filled with liquid nitrogen, for 5 min. Control samples were left untreated. Both crushed and control samples were collected on dry ice and stored at -80 °C until shipment.

Intact type I collagen was solubilized from the lyophilized rat tail tendon by SDS-extraction in SDS sample buffer for 5 min at 100 °C. Sample loads were normalized to the dry weight of tendon. SDS-extracted collagen $\alpha$-chains and collagen peptides were resolved by SDS-PAGE and stained with Coomassie Blue R-250. Densitometry of stained collagen chains ($\alpha$, $\beta$, $\gamma$) from SDS-PAGE was determined using NIH ImageJ software as previously described [27]. Collagen peptides and $\alpha$-chains were cut from SDS-PAGE gels and subjected to in-gel trypsin digestion. Electrospray mass spectrometry was carried out on the trypsin-digested peptides using an LTQ XL linear quadrapole ion-trap mass spectrometer equipped with in-line Accela 1250 liquid chromatography and automated sample injection [56, 57]. Thermo Xcalibur software and Proteome Discoverer software (ThermoFisher Scientific) were used for peptide identification and sequence coverage analysis. Tryptic peptides were also identified manually by calculating the possible MS/MS ions and matching these to the actual MS/MS spectrum. Protein sequences used for MS analysis were obtained from the Ensembl genome database.
For the analysis, we used ImageJ [26] and converted the pictures of the SDS-PAGE gel to a 32bit grey image and substracted the average background, measured as mean intensity of a line between lanes #1 and #2.

## Acknowledgements

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. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC-NG at Leibniz Supercomputing Centre (www.lrz.de).
We thank Marilyn Archer for technical assistance with the collagen extraction and SDS–PAGE. We thank Helmut Grubmüller and the respective department at the Max-Planck-Institute for Multidisciplinary Sciences for useful discussions and suggestions about the research in this manuscript.

## Author contributions

B.R. conducted together with A.Ü. MD simulations, together with C.K. KIMMDY simulations, and analyzed the simulation data and the experimental results. C.K. conducted and analyzed QM calculations. M.K. prepared samples for experiments and supported their interpretation. D.M.H. conducted SDS-page gels and mass spectrometry experiments and supported their interpretation. K.R. and G.G. supervised QM calculations and supported their interpretation. F.G. supervised the project. B.R., C.K. and F.G. wrote the manuscript. All authors commented on the manuscript and contributed to it.

## Competing Interests

The authors declare no competing interests.

## References

1. Komi, P. V. Relevance of in vivo force measurements to human biomechanics. *Journal of biomechanics* **23,** 27–34 (1990).
2. Zapp, C. *et al.* Mechanoradicals in tensed tendon collagen as a source of oxidative stress. *Nature Communications* **11.** ISSN: 2041-1723. https://dx.doi.org/10.1038/s41467-020-15567-4%20https://www.nature.com/articles/s41467-020-15567-4.pdf (2020).
3. Staudinger H; Leupold, E. Über Isopren und Kautschuk, 18. Mitteil.: Viscosittäts-Untersuchungen an Balata. *Ber Dtsch Chem Ges* **63,** 730–733 (1930).
4. Caruso, M. M. *et al.* Mechanically-induced chemical changes in polymeric materials. *Chem Rev* **109,** 5755–98. ISSN: 1520-6890 (Electronic) 0009-2665 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/19827748 (2009).
5. Luo, Y.-R. *Comprehensive handbook of chemical bond energies* (CRC press, 2007).
6. Orgel, J. P., Irving, T. C., Miller, A. & Wess, T. J. Microfibrillar structure of type I collagen in situ. *Proc Natl Acad Sci U S A* **103,** 9001–5. ISSN: 0027-8424 (Print) 0027-8424 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/16751282 (2006).
7. Bell, G. I. *Models for the specific adhesion of cells to cells* 1978. arXiv: arXiv:1011.1669v3.
8. Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophysical Journal.* ISSN: 00063495 (1997).
9. Dudko, O. K., Hummer, G. & Szabo, A. Intrinsic rates and activation free energies from single-molecule pulling experiments. *Physical Review Letters.* ISSN: 00319007 (2006).

10. Rennekamp, B., Kutzki, F., Obarska-Kosinska, A., Zapp, C. & Gräter, F. Hybrid Kinetic Monte Carlo/Molecular Dynamics Simulations of Bond Scissions in Proteins. *Journal of Chemical Theory and Computation* **16,** 553–563. ISSN: 1549-9618. https://dx.doi.org/10.1021/acs.jctc.9b00786%20https://pubs.acs.org/doi/pdf/10.1021/acs.jctc.9b00786 (2020).

11. Kwansa, A. L., De Vita, R. & Freeman, J. W. Tensile mechanical properties of collagen type I and its enzymatic crosslinks. *Biophys Chem* **214-215,** 1–10. ISSN: 1873-4200 (Electronic) 0301-4622 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/27160969 (2016).

12. Depalle, B., Qin, Z., Shefelbine, S. J. & Bühler, M. J. Influence of cross-link structure, density and mechanical properties in the mesoscale deformation mechanisms of collagen fibrils. *J Mech Behav Biomed Mater* **52,** 1–13. ISSN: 1878-0180 (Electronic) 1878-0180 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/25153614 (2015).

13. Rowe, J. & Röder, K. Chemical bonds in collagen rupture selectively under tensile stress. https://doi.org/10.1101/2022.09.23.509192 (Sept. 2022).

14. Ducrot, E., Chen, Y., Bulters, M., Sijbesma, R. P. & Creton, C. Toughening elastomers with sacrificial bonds and watching them break. *Science* **344,** 186–9. ISSN: 1095-9203 (Electronic) 0036-8075 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/24723609 (2014).

15. Willis-Fox, N., Rognin, E., Aljohani, T. A. & Daly, R. Polymer Mechanochemistry: Manufacturing Is Now a Force to Be Reckoned With. *Chem* **4,** 2499–2537. https://doi.org/10.1016%2Fj.chempr.2018.08.001 (Nov. 2018).

16. Shaikhrazieva, Y. E. & https://panor.ru/articles/vliyanie-levonorgestrela-i-folievoy-kisloty-na-sostoyanie-endometriya-i-podgotovku-k-beremennosti/85202.html. Effect of levonorgestrel and folic acid o n the state of the endometrium and preparation for pregnancy. *Pharmacy and Pharmaceutical Technology,* 58–59 (July 2022).

17. Fantner, G. E. *et al.* Sacrificial Bonds and Hidden Length: Unraveling Molecular Mesostructures in Tough Materials. *Biophysical Journal* **90,** 1411–1418. https://doi.org/10.1529%2Fbiophysj.105.069344 (Feb. 2006).

18. Fantner, G. E. *et al.* Sacrificial bonds and hidden length dissipate energy as mineralized fibrils separate during bone fracture. *Nature Materials* **4,** 612–616. https://doi.org/10.1038%2Fnmat1428 (July 2005).

19. Chan, B., Deng, J. & Radom, L. G4(MP2)-6X: A Cost-Effective Improvement to G4(MP2). *Journal of Chemical Theory and Computation* **7,** 112–120. ISSN: 1549-9618. https://dx.doi.org/10.1021/ct100542x (2011).

20. Johnny Hioe, H. Z. *Radical Stability - Thermochemical Aspects* ISBN: 978-0-470-97125-3 (2012).

21. Sustmann, R. & Korth, H.-G. in *Advances in Physical Organic Chemistry* 131–178 (Elsevier, 1990). https://doi.org/10.1016/s0065-3160(08)60045-3.

22. Peterson, J. P. & Winter, A. H. Solvent Effects on the Stability and Delocalization of Aryl Dicyanomethyl Radicals: The Captodative Effect Revisited. *Journal of the American Chemical Society* **141,** 12901–12906. https://doi.org/10.1021/jacs.9b06576 (July 2019).

23. Albert, A. & Phillips, J. N. 264. Ionization constants of heterocyclic substances. Part II. Hydroxy-derivatives of nitrogenous six-membered ring-compounds. *J. Chem. Soc.,* 1294–1304. http://dx.doi.org/10.1039/JR9560001294 (0 1956).

24. Feng, Y., Liu, L., Wang, J.-T., Huang, H. & Guo, Q.-X. Assessment of Experimental Bond Dissociation Energies Using Composite ab Initio Methods and Evaluation of the Performances of Density Functional Methods in the Calculation of Bond Dissociation Energies. *Journal of Chemical Information and Computer Sciences* **43.** doi: 10.1021/ci034033k, 2005–2013. ISSN: 0095-2338. https://doi.org/10.1021/ci034033k (2003).

25. Obarska-Kosinska, A., Rennekamp, B., Ünal, A. & Grater, F. ColBuilder: a server to build collagen fibril models. *Biophys J* **120,** 3544–3549. ISSN: 1542-0086 (Electronic) 0006-3495 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/34265261 (2021).

26. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9,** 671–675. https://doi.org/10.1038/nmeth.2089 (June 2012).

27. Kalamajski, S. *et al.* Increased C-telopeptide cross-linking of tendon type I collagen in fibromodulin-deficient mice. en. *J. Biol. Chem.* **289,** 18873–18879 (July 2014).

28. Hudson, D. M. *et al.* P3h3-null and Sc65-null Mice Phenocopy the Collagen Lysine Under-hydroxylation and Cross-linking Abnormality of Ehlers-Danlos Syndrome Type VIA. *J Biol Chem* **292,** 3877–3887. ISSN: 1083-351X (Electronic) 0021-9258 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/28115524 (2017).

29. Hudson, D. M., Archer, M., King, K. B. & Eyre, D. R. Glycation of type I collagen selectively targets the same helical domain lysine sites as lysyl oxidase–mediated cross-linking. *Journal of Biological Chemistry* **293,** 15620–15627 (2018).

30. Senn, H. M. & Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **48,** 1198–1229. ISSN: 14337851 (2009).

31. Cui, Q. Perspective: Quantum mechanical methods in biochemistry and biophysics. *J. Chem. Phys.* **145.** ISSN: 00219606 (2016).

32. Senftle, T. P. *et al.* The ReaxFF reactive force-field: development, applications and future directions. *npj Computational Materials* **2.** https://doi.org/10.1038/npjcompumats.2015.11 (Mar. 2016).

33. Metreveli, N. *et al.* Identification of free radicals induced by UV irradiation in collagen water solutions. *Journal of Photochemistry and Photobiology B: Biology* **93,** 61–65. https://doi.org/10.1016/j.jphotobiol.2008.06.003 (Nov. 2008).

34. Kjeldsen, F. *et al.* C??C Backbone Fragmentation Dominates in Electron Detachment Dissociation of Gas-Phase Polypeptide Polyanions. *Chemistry - A European Journal* **11,** 1803–1812. https://doi.org/10.1002/chem.200400806 (Mar. 2005).

35. Bleifuss, G. *et al.* Tryptophan and Tyrosine Radicals in Ribonucleotide Reductase: A Comparative High-Field EPR Study at 94 GHz. *Biochemistry* **40,** 15362–15368. https://doi.org/10.1021/bi010707d (Nov. 2001).

36. Srinivas, V. *et al.* Metal-free ribonucleotide reduction powered by a DOPA radical in Mycoplasma pathogens. *Nature* **563,** 416–420. ISSN: 1476-4687. https://doi.org/10.1038/s41586-018-0653-6 (2018).

37. Blaesi, E. J. *et al.* Metal-free class Ie ribonucleotide reductase from pathogens initiates catalysis with a tyrosine-derived dihydroxyphenylalanine radical. *Proceedings of the National Academy of Sciences* **115,** 10022–10027. ISSN: 0027-8424. eprint: https://www.pnas.org/content/115/40/10022.full.pdf. https://www.pnas.org/content/115/40/10022 (2018).

38. Matxain, J. M., Padro, D., Ristilä, M., Strid, Å. & Eriksson, L. A. Evidence of High •OH Radical Quenching Efficiency by Vitamin iB/isub6/sub. *The Journal of Physical Chemistry B* **113,** 9629–9632. https://doi.org/10.1021/jp903023c (June 2009).

39. Wondrak, G. T. & Jacobson, E. L. in *Subcellular Biochemistry* 291–300 (Springer Netherlands, Aug. 2011). https://doi.org/10.1007/978-94-007-2199-9_15.

40. Treyde, W., Riedmiller, K. & Gräter, F. Bond dissociation energies of X-H bonds in proteins. https://doi.org/10.26434/chemrxiv-2022-d7lcs (June 2022).

41. Street, D., Bangsbo, J. & Juel, C. Interstitial pH in human skeletal muscle during and after dynamic graded exercise. *The Journal of Physiology* **537,** 993–998. https://doi.org/10.1113/jphysiol.2001.012954 (Dec. 2001).

42. Kemp, G., Böning, D., Beneke, R. & Maassen, N. Explaining pH Change in Exercising Muscle: Lactic acid, Proton Consumption, and Buffering vs. Strong Ion Difference. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **291,** R235–R237. https://doi.org/10.1152/ajpregu.00662.2005 (July 2006).

43. Dıaz, F. E., Dantas, E. & Geffner, J. Unravelling the Interplay between Extracellular Acidosis and Immune Cells. *Mediators of Inflammation* **2018,** 1–11. https://doi.org/10.1155/2018/1218297 (Dec. 2018).

44. Imato, K. & Otsuka, H. in *Dynamic Covalent Chemistry* 359–387 (John Wiley & Sons, Ltd, Sept. 2017). https://doi.org/10.1002/9781119075738.ch9.

45. Krauss, S., Metzger, T. H., Fratzl, P. & Harrington, M. J. Self-Repair of a Biological Fiber Guided by an Ordered Elastic Framework. *Biomacromolecules* **14,** 1520–1528. https://doi.org/10.1021/bm4001712 (Apr. 2013).

46. Usha, R., Rajaram, A. & Ramasami, T. Stability of collagen in the presence of 3, 4-dihydroxyphenylalanine (DOPA). *Journal of Photochemistry and Photobiology B: Biology* **97,** 34–39. https://doi.org/10.1016/j.jphotobiol.2009.07.009 (Oct. 2009).

47. Fuentes-Lemus, E., Hägglund, P., López-Alarcón, C. & Davies, M. J. Oxidative Crosslinking of Peptides and Proteins: Mechanisms of Formation, Detection, Characterization and Quantification. *Molecules* **27,** 15. https://doi.org/10.3390/molecules27010015 (Dec. 2021).

48. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1,** 19–25 (2015).

49. Best, R. B. & Hummer, G. Optimized molecular dynamics force fields applied to the helix- coil transition of polypeptides. *The journal of physical chemistry B* **113,** 9004–9015 (2009).

50. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics* **78,** 1950–1958 (2010).

51. Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation* **4,** 116–122 (2008).

52. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of chemical physics* **126,** 014101 (2007).

53. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics* **52,** 7182–7190 (1981).

54. Kutzner, C., Czub, J. & Grubmüller, H. Keep it flexible: Driving macromolecular rotary motions in atomistic simulations with GROMACS. *Journal of chemical theory and computation* **7,** 1381–1393 (2011).

55. Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **185,** 604–613. ISSN: 0010-4655. http://www.sciencedirect.com/science/article/pii/S0010465513003196 (2014).

56. Eyre, D. R., Weis, M. A. & Wu, J. J. Advances in collagen cross-link analysis. *Methods* **45,** 65–74. ISSN: 1046-2023 (Print) 1046-2023 (Linking). https://www.ncbi.nlm.nih.gov/pubmed/18442706 (2008).

57. Eyre, D. in *Methods in Enzymology* 115–139 (Elsevier, 1987). https://doi.org/10.1016/0076-6879(87)44176-1.