Jamil Serwanja, Alexander Wieland, Astrid Haubenhofer, Hans Brandstetter, Esther Schönauer

Department of Biosciences and Medical Biology, University of Salzburg, Salzburg, Austria

Title

The activator domain of bacterial collagenases drives collagen recognition, unwinding and processing

Abstract

Collagens form the resilient backbone of the extracellular matrix in mammals. Only few proteases are able to digest triple-helical collagen. Clostridial collagenases can efficiently process collagen. However, little is known about the mechanism of bacterial collagenolyis of either soluble collagen or the multi-hierarchically assembled, insoluble collagen fibers. Here we present a functional analysis of the distinct roles of the individual domains of collagenase G (ColG) from *Hathewaya histolytica.* A broad array of biochemical, biophysical, and enzymatic data consistently revealed unexpected synergistic and antagonistic interactions between the activator, peptidase and collagen-binding domains. We found the non-catalytic activator domain to act as a master regulator, coordinating the complex interactions to specifically recognize and process the diverse physiological substrates. The results presented here enable multiple applications such as the engineering of collagenase variants with selectivity for defined substrate states.

Introduction

Collagens are the most abundant proteins in mammals, constituting up to 90% of the extracellular matrix. Their supramolecular assemblies - collagen fibrils, fibers and networks – shape and organize the extracellular matrix and are essential for tissue integrity [1,2]. Of the currently known 28 types of collagen, fibrillar type I collagen is the by far most prevalent one (approx. 90%). These fibrils have dimensions of ~300 nm in length and ~1.5 nm in diameter [3,4].

Collagens are remarkable molecules due to their hallmark feature the collagen triple helix and their multi-hierarchical structure. Type I collagen consists almost exclusively of this structural motif (96%) [5]. The collagen triple helix is formed by three α-chains that intertwine into a tightly packed right-handed helix. This helix is stabilized by the repetitive triplet Gly-X-Y, in which the X and Y positions are frequently occupied by proline and hydroxyproline, respectively. Gly-Pro-Hyp is the most frequent triplet (10.5%) [6]. Triple-helical collagen is highly resistant to proteolysis (i) due to its tight helical packing, caching the peptide bonds in the interior of the helix, and (ii) due to its high content of the imino acids proline and hydroxyproline, which only few proteases can accommodate in their active sites. This is one of the reasons why natively folded triple-helical collagen is a hard-to-digest protein. Even the promiscuous proteases of the digestive system cannot process it under physiological conditions [7,8]. What is more, fibrillar collagen adopts in the body a complex multi-hierarchical structure. Soluble type I collagen molecules (*i.e.* tropocollagen) self-assemble into larger supramolecular assemblies, starting from small insoluble microfibrils with a diameter of > 5 nm, which are made up by 5 tropocollagen molecules. These microfibrils further assemble into larger insoluble fibrils (30 – 500 nm diameter) that in turn form 0.5-2 µm thick fibers. These fibers are stabilized via hydroxylysyl pyridinoline and lysyl pyridinoline crosslinks, and through interactions with glycosaminoglycans, imparting the fibers with great tensile strength and a half-life of several years. These fibers form the basic scaffold of the extracellular matrix in mammals. [2,9–11].

A true collagenase must therefore be able to (i) bind to these macroscopic supermolecular fibers, (ii) extract a single tropocollagen molecule from the composite collagen structure, (iii) unwind the collagen triple helix to excavate the scissile bonds, (iv) accommodate the imino acid-rich peptide chain in its active-site cleft, and (v) cleave pre- or post-(hydroxy)proline peptide bonds. Merely a small number of mammalian enzymes is capable of this complex task and can remodel collagen under physiological conditions. Typically, they do so with a very strict and narrow substrate specificity (*e.g.* MMP-1/2/8/13/14/18, cathepsin K, neutrophil elastase) [12,13]. However, bacteria have also evolved collagenases, most notably *Clostridium* spp., *Bacillus* spp. (M9B

subfamily) and *Vibrio* spp. (M9A subfamily) [14]. In contrast to the narrow specificities of mammalian collagenases, these bacterial zinc metalloproteases are capable of cleaving folded collagen at multiple sites and can fully digest the triple-helical collagen into small peptides [15–17]. In nature, these collagenases enable saprophytic bacterial species to consume collagen for nutrition and provide pathogenic strains with a tool for host invasion, colonization, and facilitated toxin diffusion [18–21]. However, the molecular mechanism that enables bacterial collagenases to cleave collagen 'like a meat chopper' is not fully understood.

Bacterial collagenases are calcium-dependent multi-domain proteins that belong to the gluzincin family of metalloproteases [15,22–27]. At the N-terminus, they harbor a collagenase unit of ~80 kDa consisting of an activator domain (M9N domain) and a peptidase domain (peptidase M9 domain), which are connected via a short 9 aa-long linker. The peptidase domain contains the catalytic zinc ion, which is coordinated by the two histidines of the canonical zinc-binding HEXXH motif, and by a glutamate ~ 28 amino acids downstream [23,28,29]. At the C-terminus, a varying composition of two to three accessory domains is found, typically consisting of zero to two polycystic-kidney disease-like domain(s) (PKD), and one or two collagen-binding domains (CBD) (each ~10 kDa) [23,24,26,30] (**Fig. 1a**).

The most extensively studied and biotechnologically utilized members are collagenase G (ColG) and collagenase H (ColH) from *Hathewaya histolytica* (formerly *Clostridium histolyticum*) (*e.g.* [15,21,23,24,31,32]. The first high-resolution structure of the collagenase unit of ColG in 2011, published by us, revealed a saddle-shaped architecture of the collagenase unit (**Fig. 1b**) and we identified the collagenase unit (CU) as the minimal collagenolytic entity. The CU consists of an N-terminal activator domain (AD) fused to a peptidase domain (PD). The PD alone is able to cleave small peptidic substrates, yet it cannot degrade triple-helical collagen. By contrast, the CU is able to unwind and cleave native collagen *in vitro* without the need of the accessory domains [23]. Based on these findings, we proposed a conformational two-state model of bacterial collagenolysis, *aka* chew-and-digest model, in which recognition, unwinding and processing of triple-helical collagen are driven by opening and closing of the CU. In this model, triple-helical collagen is recognized by the PD in the open CU state as seen in the crystal structure [23], thereby triggering the closing of the CU. In the proposed compact state the AD and the PD are able to interact with the triple-helical collagen, thereby facilitating the unwinding of the bound collagen molecule and enabling the successive cleavage of the entrapped α-chains in the (semi-) closed conformation of the CU. On a more speculative level, the chew-and-digest model was also extended to the processing of microfibrils. Hereby, an initially open CU binds a microfibril, which again triggers

a CU contraction movement, in which a single tropocollagen molecule is exposed from the fibrillar ensemble and then unwound and cleaved by the engulfing CU [23].

In case of the *Clostridium* collagenases, support for the existence of the closed conformation of the CU was provided by Ohbayashi *et al.* using small angle X-ray scattering (SAXS) experiments. The obtained low resolution model of ColH suggests that a closed conformation of the CU is preferred in solution [33]. In case of *Vibrio* collagenases, recent crystallographic studies identified similar saddle-shaped architectures of the CUs of the collagenase VhaC of *Vibrio harveyi VHJR7* [29] and Ghcol of *Grimontia hollisae* [28]. Molecular dynamics simulations performed with the CU of VhaC (VhaC-CU) suggest that the unbound VhaC-CU alternates between an open and closed conformation, while binding to a triple-helical substrate shifts the equilibrium to the closed state [29].

Even though these findings corroborate two pillars of the proposed chew-and-digest model for collagenolysis, namely the existence of an open and (semi-) closed conformation of the CU, other results argue for a refinement of the model. *Wang et al.* showed that the AD of VhaC (VhaC-AD) can bind in a dose-dependent manner to short triple-helical collagen mimics and insoluble type I collagen fibers, while the VhaC-PD showed only minor or no binding affinities to these substrates, respectively. The authors, therefore, suggest that the initial recognition of triple-helical collagen is mediated by the AD in bacterial collagenases [29]. In support, Ikeuchi *et al.* identified two binding sites for Gly-Pro-Hyp peptides in the AD of Ghcol (Ghcol-AD) and one binding site in the active site of the PD. Moreover, the latter identified three highly conserved residues in the VhaC-AD (Phe107, Arg153, and Tyr157), which are important for the binding of collagen fibers [29].

These recent studies have helped to shed new light onto the mechanism of bacterial collagen degradation. However, the essential questions of how bacterial collagenases manage to unfold the collagen triple helix – a feature, which sets them apart from other peptidases - and how they are able to process even higher molecular assemblies of collagen, are still not answered.

In this study, we investigate the role of the individual domains of bacterial collagenases in the recognition of the different physiological forms of fibrillar collagen, *i.e.* unwound collagen (*aka* gelatin), soluble triple-helical collagen (*aka* tropocollagen) and insoluble collagen fibers. We identify and characterize the triple-helicase entity in bacterial collagenases, and we examine the significance of the interplay between the AD and the PD in the processing of soluble and insoluble collagen using conformational trapping.

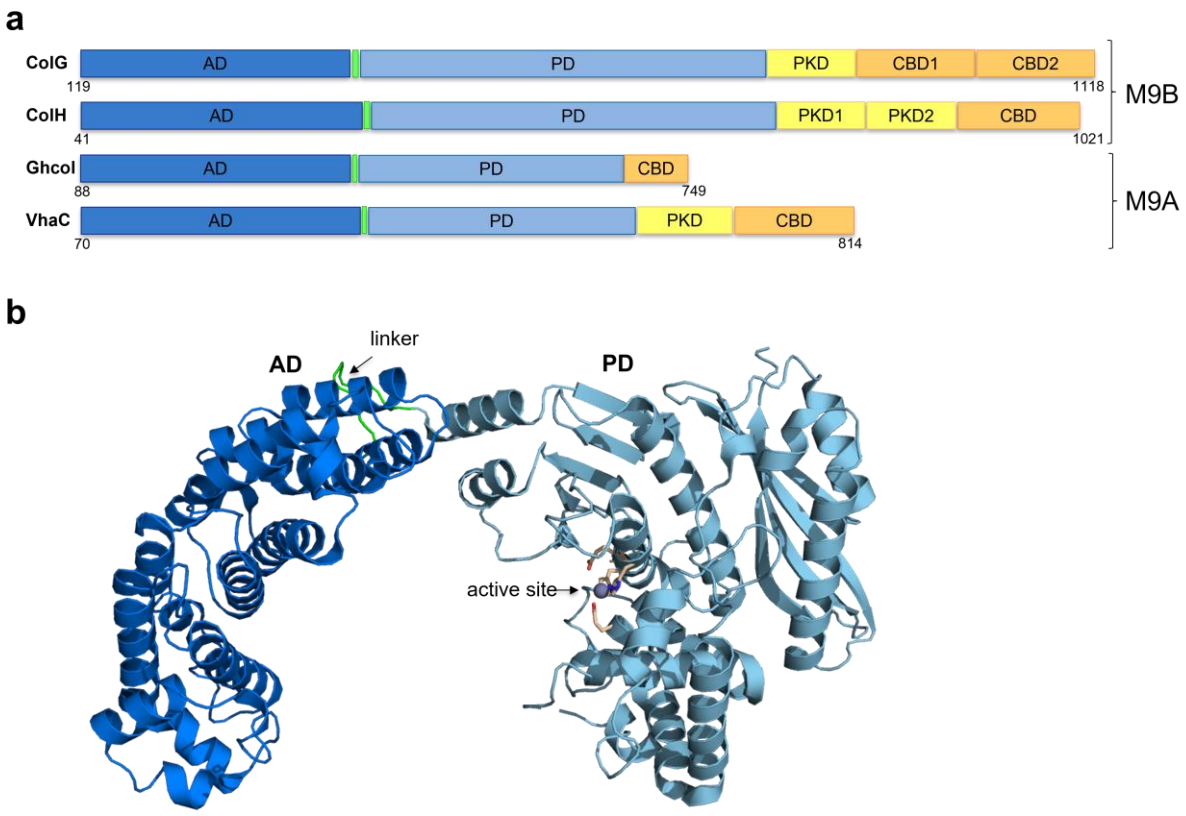


Fig. 1: **Domain composition and crystal structure of ColG-CU. a**, Schematic domain organization of ColG and ColH from *H. histolytica*, Ghcol from *Grimontia hollisae*, and VhaC from *Vibrio harveyi VHJR7* [29,34,35]. Activator domain (AD) (dark blue), linker (light green), peptidase domain (PD) (light blue), polycystic disease-like domain (PKD) (yellow), and collagen-binding domain (CBD) (orange). **b**, Ribbon representation of the collagenase unit of ColG (PDB: 2y50). The catalytic zinc ion (grey) and the catalytic residues (light orange) are shown in ball-and-stick representation. The linker is highlighted (green). Molecular figures were created with PyMOL [36].

Results

Unwound α-chains are the substrate of bacterial collagenases.

Fibrillar collagen such as type I collagen adopts a complex multi-hierarchical structure in the body, ranging from individual collagen triple helices, via (micro-)fibrils to fibers. The dimensions of the active sites of clostridial collagenases such as ColG, ColH, ColT [34], and of Ghcol[28], a *Vibrio* collagenase , but also of mammalian collagenases like MMP-1 are all around ~10 Å wide, and thus consistently exclude the access of complex, supramolecular collagen assemblies. Even individual tropocollagen molecules with an approximate diameter of ~15 Å need to be unwound to their α-chains to fit into the active-site clefts of these enzymes. This has been demonstrated for MMP-1 [37,38]. To confirm this hypothesis on a biochemical level, we investigated the temperature dependence of tropocollagen and gelatin degradation using the

collagenase unit of ColQ1 from *Bacillus cereus strain Q1* (ColQ1-CU) [39] (**Fig. 2**). At 37 °C and 25 °C, tropocollagen and gelatin were both readily degraded by ColQ1-CU. However, when the reaction temperature was decreased to 15 °C, collagen turnover nearly halted, while gelatin degradation continued, yet at a significantly slower rate compared to 25 °C and 37 °C. Upon further decrease of the reaction temperature to 4 °C, no degradation of collagen by ColQ1-CU could be observed, but cleavage of gelatin still proceeded at a very slow rate. Thus, whereas the gelatin cleavage was qualitatively governed by the Arrhenius equation[40,41], tropocollagen degradation was more sensitive to decreasing the reaction temperature, indicative for the existence of an additional distinct transition temperature for collagen as compared to gelatin cleavage. These results emphasize that the unwinding of the collagen triple helix, *i.e*. the triple-helicase activity, is an essential prerequisite for collagen cleavage by bacterial collagenases.

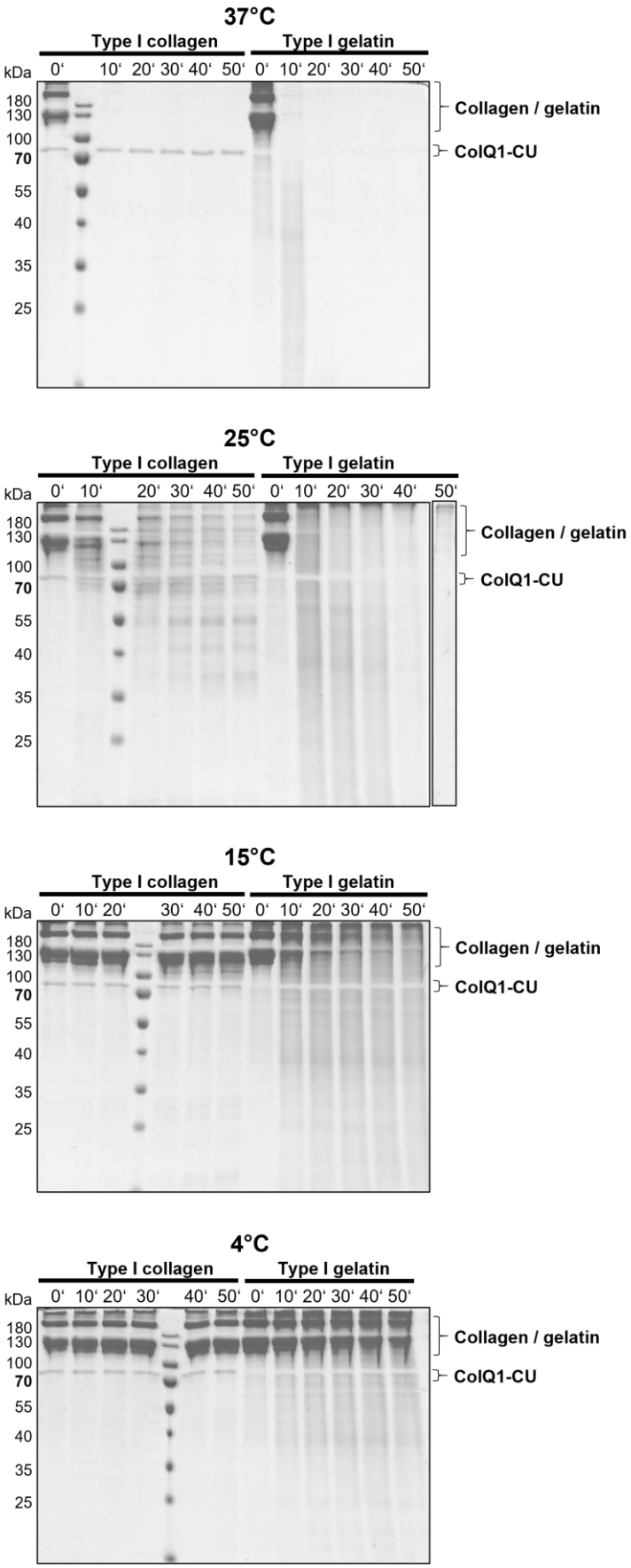


Fig. 2: Temperature dependence of type I collagen and type I gelatin degradation by ColQ1-CU. 3.33 µM type I collagen and 9.99 µM type I gelatin, respectively, were incubated with 0.2 µM and 0.02 µM ColQ1-CU, respectively, at the indicated temperatures. Reactions were terminated by addition of 35.6 mM EDTA and subjected to SDS–PAGE under non-reducing conditions.

Recognition of the hierarchical substrate collagen by clostridial collagenases

The previous finding also highlights that the processing of physiological substrates involves different recognition events. Important substrate states for binding include (i) insoluble collagen fibers, fibrils and microfibrils; (ii) soluble (tropo-)collagen; and (iii) unfolded α-chains (*aka* gelatin). Yet, a comprehensive comparative characterization of all these different binding events for bacterial collagenases is missing. Therefore, we systematically examined the affinity of full-length ColG, the CU and of its individual domains for these three different hierarchical levels of collagen. Protein variants that contained the PD carried the catalytically inactivating E524A mutation. The full-length enzyme ColG E524A (Tyr119-Lys1118) (ColG-FL E524A) and the ColG-AD (Tyr119-Ser392) contained an N- or C-terminal maltose-binding protein tag, respectively.

First, we investigated the binding of ColG-FL E524A and the individual domains to insoluble collagen fibers, the predominant physiological substrate in the mammalian body (**Fig. 3**). We observed concentration-dependent binding of ColG-FL E524A, ColG-CU E524A (Tyr119-Gly790), ColG-AD and ColG-CBD2 (Asn1004-Lys1118) to the fibers, whereas very low or hardly any binding to insoluble collagen was seen for ColG-PD E524A (Asp398-Gly790) and ColG-PKD (Ile799-Asn880). The highest affinity towards the fibers exhibited ColG-FL E524A, followed by ColG-CBD2, ColG-CU E524A, and ColG-AD. ColG-CBD2 showed the second-best binding among all ColG variants to collagen fibers. It bound better to fibers than all other ColG-domains, reflecting the eponymous ‘collagen binding/recruitment domain’ function. Interestingly, ColG-CU E524A showed a higher affinity to collagen fibers than the single AD, indicating a cooperative binding mode of the AD and PD to the fibrillar collagen substrate. Together, these results suggest that the very high binding affinity of the full-length enzyme is based on the high affinity of the CBD and the CU to collagen fibers. The binding of the latter seems to be primarily mediated by the AD, in line with findings for VhaC-AD[29].

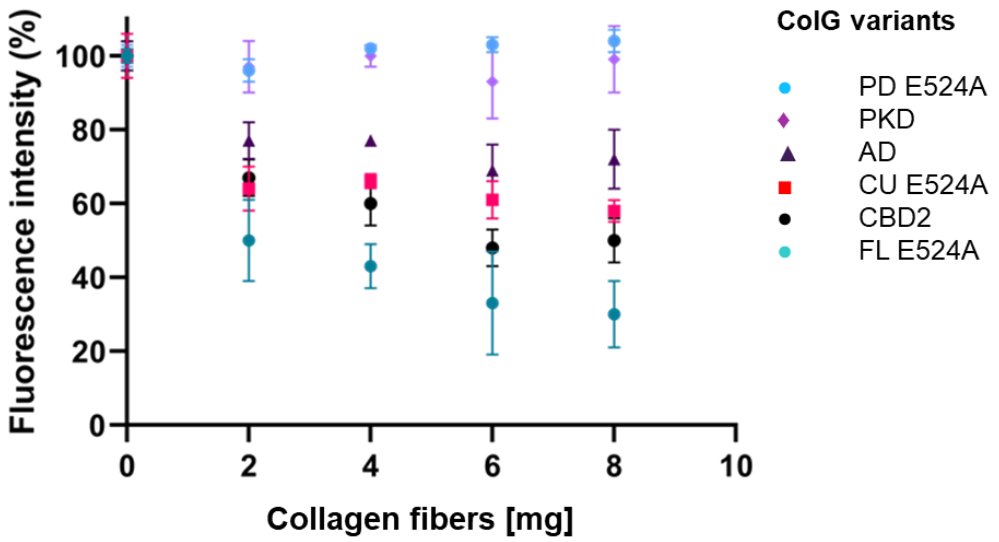


Fig. 3: Binding of ColG variants to insoluble collagen fibers. ColG-domains were fluorescently labelled and incubated at 0.2 µM (except ColG-PKD: 5 µM) for 30 min at RT with increasing concentrations of insoluble collagen fibers from bovine Achilles tendon. The fluorescence remaining in the supernatant after incubation was detected at 647/680 nm.

To test the binding affinities towards unwound collagen (produced by heat treatment of folded collagen) and soluble collagen, indirect ELISA assays were performed (**Fig. *4***). Since gelatin can partially refold into triple-helical structures depending on the local concentration upon cooling [40–42], particular care was taken to minimize the amount of triple-helical conformations by using gelatin concentrations and reaction temperatures that disfavored triple-helix formation during the coating process and the binding assays. Among all variants, ColG-FL E524A displayed, as expected, the highest affinity to both physiological substrates. Intriguingly, it displayed very similar affinities towards unfolded and triple-helical collagen (0.031 ± 0.004 µM *vs.* 0.024 ± 0.002 µM, respectively). Most of its affinity towards soluble collagen and gelatin is mediated by the CU. ColG-CU E254A bound best to gelatin with an apparent K_d of 0.08 ± 0.01 µM, and it showed an approx. 5-fold lower affinity to collagen with an apparent K_d of 0.37 ± 0.06 µM. Interestingly, the accessory domains showed only poor binding towards gelatin and collagen, except for ColG-CBD2 towards its name-giving ligand. ColG-CBD2 bound to triple-helical collagen with an apparent K_d of 14 ± 1 µM, but it displayed low binding affinity to gelatin with an apparent K_d of 253 ± 30 µM. In contrast, ColG-PKD exhibited very little binding to both substrates at all. The observed binding was so low that no saturation could be accomplished even in the presence of 1.2 mM ColG-PKD. Thus, while no K_d values could be determined, the data indicate the apparent equilibrium dissociation constants to be well over 1 mM for ColG-PKD for both substrates (**Fig. S1**).

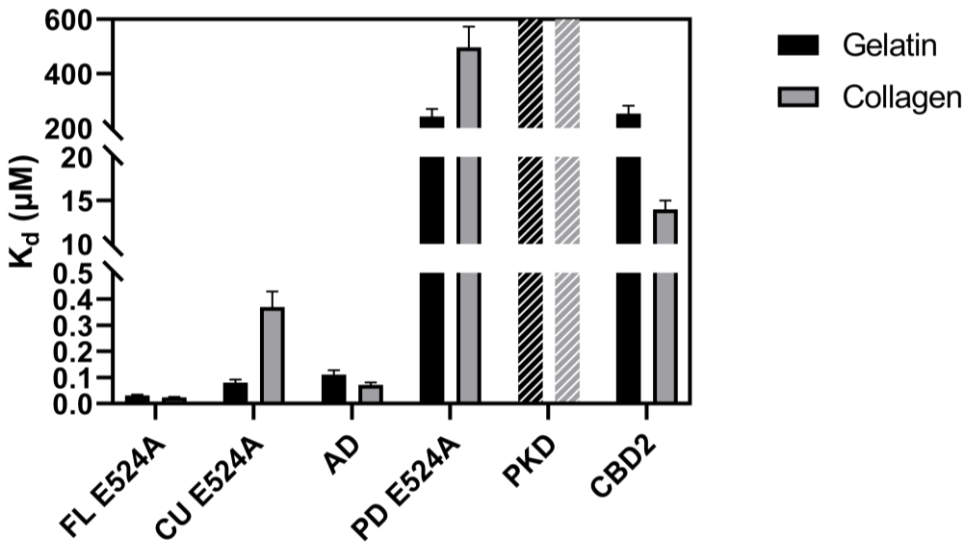


Fig. 4: Binding affinities of ColG variants to soluble type I collagen and gelatin. The binding affinities were determined via indirect ELISA. Binding to type I collagen was tested at 25 °C to preserve the triple-helical fold of the coating, while the binding affinities towards type I gelatin were tested at 37 °C. The K_d for ColG-PKD could not be determined, as no saturation up to 1.2 mM was observed. However, from the data it could be concluded that the K_d exceeds 1000 µM by far (shaded bars). The binding curves are shown in Fig. S1.

Yet, the highly affine binding of ColG-FL E524A towards gelatin and tropocollagen cannot solely be explained by the contributions of the CU. ColG-FL displayed a more than twofold higher affinity towards gelatin and a nearly 9-fold higher affinity to collagen compared to ColG-CU. These increases in affinity most likely result from the C-terminal CBD domains; in fact, the differential CBD binding affinities to collagen and gelatin consistently explain the preferential collagen binding of ColG-FL, despite a weaker collagen binding of ColG-CU (**Fig. *4***).

To dissect the binding of the CU to gelatin and collagen even further, we analyzed the individual binding affinities of the AD and PD, which together form the CU. ColG-AD exhibited an apparent K_d of 0.11 ± 0.02 µM for gelatin and an even better K_d of 0.072 ± 0.009 µM for collagen. In contrast to that, ColG-PD E524A showed only low binding to gelatin (apparent K_d of 243 ± 28 µM) and an even twofold lower binding towards collagen (apparent K_d of 497 ± 76 µM). When comparing the affinities to gelatin and collagen of the individual AD and the PD with the one of the CU, first, it is striking that the CU binds to gelatin like the AD (apparent K_d = 0.08 ± 0.01 µM *vs*. 0.11 ± 0.02 µM). This suggests that the binding to gelatin is predominantly mediated via the AD, consistent with the several hundred-fold difference in binding affinities of AD and PD. Second and counterintuitively, the binding of the CU to triple-helical collagen (apparent K_d = 0.37 ± 0.06 µM) cannot be accounted for by an additive effect of the AD (apparent K_d = 0.072 ± 0.009 µM) and the PD (apparent K_d = 497 ± 76 µM), but

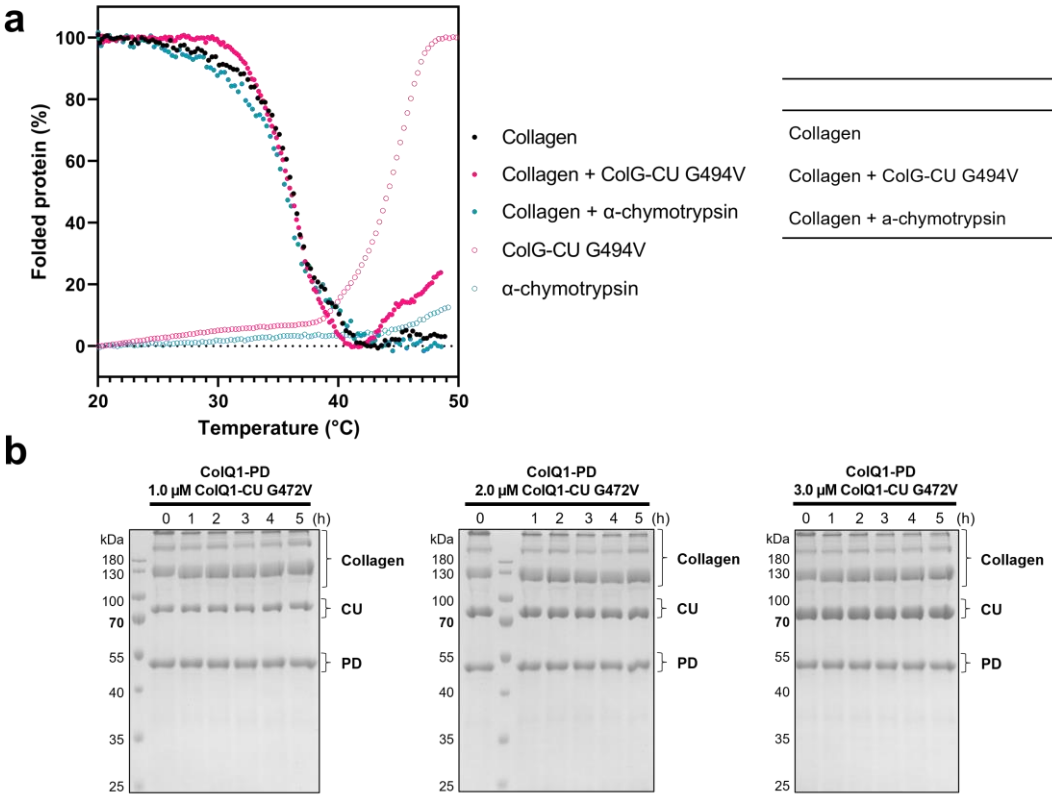
rather suggests an antagonistic interaction of the AD and PD when binding to the triple-helical substrate.

The observed affinities to soluble collagen in parts mirrored the binding profile to fibrillar collagen. The negligible binding of the PD and PKD domain as well as the high affinity of the full-length protein came as no surprise for either substrate. However, there were remarkable differences. First, ColG-CBD2 showed the second-best binding among all ColG variants to fibrillar collagen. It bound better to fibers than ColG-CU E524A and ColG-AD, which both had displayed superior binding (38-fold and 194-fold tighter binding than ColG-CBD2) to soluble collagen. This difference in binding behavior suggests that additional binding motifs such as pyridinoline cross-links, which are specific to the supramolecular fibers, may modify the affinities of the individual domains to the substrates. Second, and quite unexpectedly, we found reversed affinities regarding ColG-CU E524A and ColG-AD. These two units differ by the presence or absence of the PD, which has no or only negligible affinity towards fibrillar and soluble collagen, respectively. ColG-CU E524A showed a higher affinity to collagen fibers than ColG-AD, which reflects a situation of clearly cooperative binding. One mechanistic explanation of this synergistic interaction relates to the saddle-shaped structure of the collagenase unit with dimensions matching those of collagen microfibrils, allowing for multivalent interactions of AD and PD with microfibrillar collagen, cf. fig. 1. Such multivalent interactions are not to be expected for soluble, triple-helical collagen. To the contrary and unexpectedly, we observed an apparent antagonistic interaction of AD and PD towards triple-helical collagen, which may be explained by a steric hindrance effect. Triple-helical collagen is stiffer than gelatin, explaining why the antagonistic effect is seen only with soluble collagen but not with gelatin.

Collagen is not globally unwound by bacterial collagenases.

Collagen unwinding is a necessary prerequisite for collagen degradation. To investigate the triple-helicase activity of ColG, we measured the melting temperature (T_m) of type I tropocollagen in the absence and presence of equimolar concentrations of ColG-CU G494V. α-Chymotrypsin, which cannot unwind and cleave triple-helical collagen, served as negative control. As **Fig. 5** shows, there was no significant change in the T_m of collagen in the absence or presence of ColG-CU G494V, and the melting temperature of collagen did not change in the presence of either protease (**Fig. 5**a); instead, the comparison of the melting curves seems to indicate negative cooperativity in the collagen unfolding [42]. This suggests that ColG does not

unwind triple-helical collagen globally, but that the unwinding of soluble collagen only takes place locally. To corroborate this finding, we performed a proteolytic complementation assay using ColQ1-PD as 'cutter' enzyme and the proteolytically inactive ColQ1-CU G472V as 'triple helicase' (**Fig. 5**b, **Fig. S2**). ColQ1-PD is a very efficient cutter of unfolded collagen, but it is not able to degrade triple-helical collagen (**Fig. S3**). Despite the addition of increasing concentrations of ColQ1-CU G472V to a constant high amount of ColQ1-PD, we observed no turnover of natively folded collagen, suggesting that the unwinding of collagen by the ColQ1-CU G472V only occurred very locally. This finding corroborated the previous CD results and showed that M9B collagenases do not unwind collagen globally during catalysis.



	T_m(°C)
Collagen	36.0 ± 0.5
Collagen + ColG-CU G494V	36.1 ± 0.3
Collagen + a-chymotrypsin	36 ± 1

Fig. 5: ColG does not unwind type I collagen globally. a, Thermal transition curve and melting temperatures for type I collagen with or without ColG-CU G494V and α-chymotrypsin. Pepsin-treated type I collagen (2.25 µM) was incubated with or without an equimolar concentration of ColG-CU G494V or α-chymotrypsin. The molar ellipticity was recorded at 222 nm, while the temperature increased from 20 to 50 °C at a rate of 0.15 °C/min. **b**, 3.33 µM type I collagen were incubated with equimolar concentrations of ColQ1-PD and increasing concentrations of ColQ1-CU G472V at 25 °C. Samples were taken at indicated time points and the reaction was stopped by addition of 38 mM EDTA. The integrity of the collagen fold was verified by co-incubation with 0.83 µM α-chymotrypsin (data not shown).

Unwinding of collagen-like triple helices can be monitored via CD.

As we demonstrated, the local unwinding of the collagen triple helix is not accessible to detection by circular dichroism using the physiological substrate tropocollagen. We hypothesized that the fraction of the collagen molecule that is unwound by the collagenase is

too small in comparison to the dimensions of the whole molecule (~300 nm in length, ~338 Gly-X-Y tripeptides/α-chain) to have a significant effect on the overall T_m of the molecule. Therefore, we generated a short collagen-like substrate V-$(GXY)_{26}$ to monitor the unwinding activity via CD. This construct is derived from prokaryotic collagen Scl2 (*Streptococcus pyogenes* collagen-like protein 2). It is composed of the N-terminal α-helical trimerization V domain of Scl2.3 and the 78-aa long proline-rich B segment of the collagenous CL domain of Scl2.28 [43,44]. The native B segment from Scl2.28 was shown to form a stable collagen-like triple-helix with a T_m of 29.5 °C [44]. We modified the 26 tripeptides of the B segment to increase the affinity of ColG to it by mutating 4 tripeptides to Gly-Pro-Ala (**Fig. S4**) [31]. This modified trimeric construct V-$(GXY)_{26}$ was efficiently recognized by ColG-CU. The triple-helical B-segment was degraded by ColG-CU within 120 min, leaving the V-domain trimer intact, while the catalytically inactive G494V mutant could not cleave it. α-Chymotrypsin could only partially cleave the V-domain, but left the triple-helical part intact (**Fig. 6a**).

We confirmed the triple-helical fold of the modified B segment of V-$(GXY)_{26}$ using CD spectroscopy (**Fig. 6b,e**). The net spectrum of V-$(GXY)_{26}$ is composed of the α-helical contribution of the V-domain (78 aa) and of the triple-helical contribution of the collagenous B-segment (78 aa). The spectrum of the isolated V domain showed the typical α-helical CD signature with two minima at 208 and 222 nm, respectively. The isolated modified B-segment, *i.e.* $(GXY)_{26}$, displayed the typical spectral fingerprint of triple-helical collagen with a distinct minimum at 197 nm and a positive peak at 220 nm. When monitoring the molar ellipticity at 222 nm at temperatures between 10 and 70 °C, the melting of the triple-helical $(GXY)_{26}$ part could be clearly differentiated from the melting of the V-domain (**Fig. 6c,d**). The single V-domain exhibited a T_m of 52 ± 2 °C, while the triple-helical $(GXY)_{26}$ segment within V-$(GXY)_{26}$ had a T_m of 31.82 ± 0.05 °C. The latter T_m is a bit higher than the reported T_m of the unmodified B-segment and could suggest that the newly introduced Gly-Pro-Ala triplets stabilized the triple-helical fold compared to the original sequence. Moreover, when we compared the spectra of V-$(GXY)_{26}$ recorded at 25 °C and 35 °C, *i.e.* before and after the B segment melts, and calculated the respective difference spectrum, it showed the CD transitions typical for the collagen triple helix, a negative band below 200 nm and a positive band at 222 nm [45] (**Fig. 6e**).

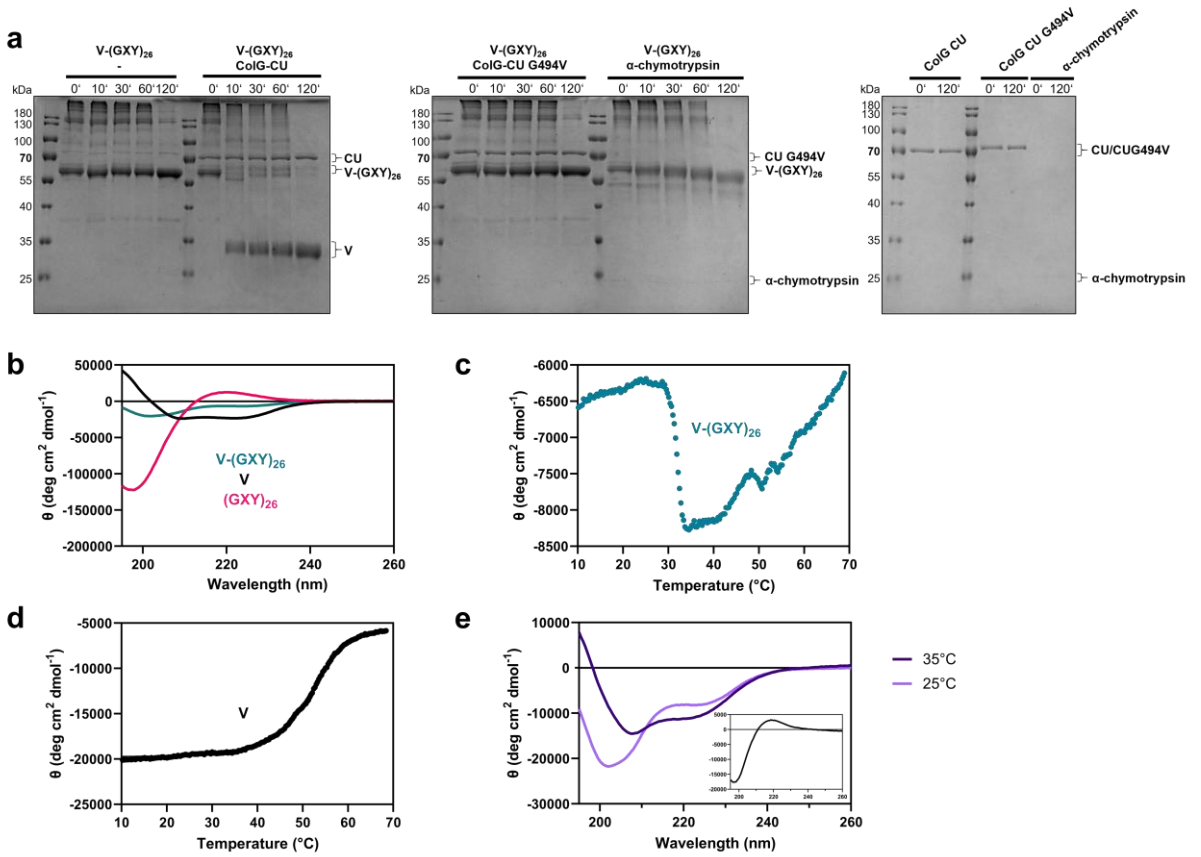


Fig. 6: V-(GXY)$_{26}$ is a short collagen-like substrate. a, Incubation of 10 µM V-(GXY)$_{26}$ with and without 0.5 µM ColG-CU WT, with 0.5 µM ColG-CU G494 and 0.5 µM α-chymotrypsin at 25 °C. The reactions were stopped by the addition of SDS-loading buffer and analyzed by SDS-PAGE. . Note that the V domain migrates as an SDS stable trimer at an apparent mass of 30 kDa; similarly, V-(GXY)$_{26}$ migrates as a trimer. **b**, CD spectrum of V-(GXY)$_{26}$ and of (GXY)$_{26}$, prepared by pepsin digest of V$_{Scl2.3}$-(GXY)$_{26}$. All CD spectra were recorded at 25 °C. **c-d**, Melting profiles of V-(GXY)$_{26}$ (**c**) and of the single V-domain (**d**). Molar ellipticities were recorded at 222 nm while the temperature increased from 10 to 70 °C at a rate of 9 °C/h. **e**, CD spectra of V-(GXY)$_{26}$ recorded at 25 °C and 35°C. The inlet shows the CD difference spectrum of the 25 °C spectrum minus the 35°C spectrum.

The activator domain is a triple helicase.

To identify the triple-helicase entity in ColG, we tested the effect of all ColG domains which were able to bind to triple-helical collagen, *i.e.* ColG-CU, ColG-AD and ColG-CBD2, on the melting temperature of V-(GXY)$_{26}$, monitoring the molar ellipticity at 222 nm between 25 and 35 °C. The interaction with ColG-CU G494V resulted in a significant decrease in the overall T_m of the triple-helical part of V-(GXY)$_{26}$. In absence of the collagenase, V-(GXY)$_{26}$ had a T_m of 31.34 ± 0.02 °C. However, in the presence of a twofold excess of ColG-CU G494V a distinct destabilization of the triple helix was observed, resulting in a decrease in T_m to 30.66 ± 0.02 °C. As a control we added a 6-fold excess of ColG-CBD2 to V-(GXY)$_{26}$, which is known to bind to triple-helical collagen, but which cannot unwind collagen. The presence of ColG-CBD2 did not alter the T_m of V-(GXY)$_{26}$ (31.35 ± 0.05 °C). This is the first time to our knowledge that the triple-helicase activity of a collagenase could be experimentally verified.

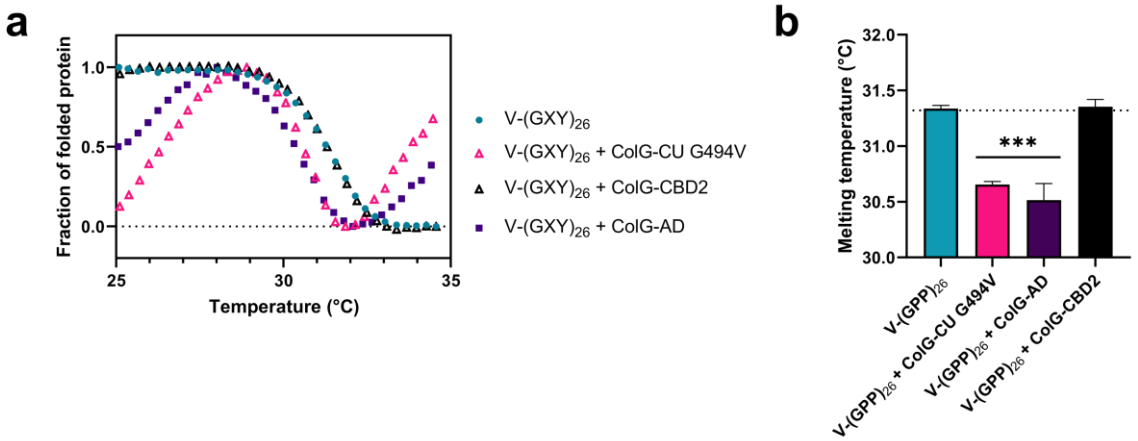


Fig. 7: Triple-helicase activity of ColG-CU is detectable via CD spectroscopy. a Melting profiles of the short collagen mimic V-$(GXY)_{26}$ in absence or presence of ColG-CU G494V, ColG-AD and ColG-CBD2. 5 µM V-$(GXY)_{26}$ were incubated with 10 µM ColG-CU G494V, 10 µM ColG-AD and 30 µM ColG-CBD2. Molar ellipticities were recorded at 222 nm, while the temperature increased from 25 to 35 °C at a rate of 4.5 °C/h. **b**, Melting temperatures of V-$(GXY)_{26}$ in absence or presence of ColG-CU G494V, ColG-AD and ColG-CBD2 (mean ± SD). Statistical significance was determined by one-way ANOVA followed by Dunnett's test for multiple comparisons (*** $P < 0.001$)

Given that triple-helical collagen preferentially binds to the AD rather than the PD and that the PD cannot unwind collagen on its own, we speculated whether the AD on its own is responsible for collagen unwinding, while the PD, which harbors the catalytic residues, is primarily responsible for hydrolysis. To test this hypothesis, we measured the melting profile of V-$(GXY)_{26}$ in presence of ColG-AD (**Fig. 7**). Upon co-incubation with ColG-AD, we observed a clear decrease in T_m of V-$(GXY)_{26}$ to 30.52 ± 0.15 °C, comparable to the temperature shift detected in presence of ColG-CU G494V (30.66 ± 0.02 °C). This finding suggests that the AD can unfold the triple-helix of V-$(GXY)_{26}$ independently of the PD.

To confirm the triple-helicase activity of the AD on a biochemical level, we performed two substrate degradation assays (**Fig. 8**). First, we performed an assay monitoring the cleavage of soluble type I collagen at 25 °C via non-reducing SDS-PAGE. For this assay, we used ColQ1-CU, which cleaves natively folded collagen six times more efficiently than ColG[39]. Our control experiments showed that 0.2 µM ColQ1-CU readily degraded 3.33 µM collagen within 50 min (**Fig. S5a**), while 3-fold more ColQ1-PD did not result in any visible degradation of collagen even after a 150-min incubation (**Fig. S5b**). In contrast to that, half the amount of ColQ1-PD completely degraded 3.33 µM gelatin within 30 min (**Fig. S5c**), showing that the triple-helical fold of collagen, not its primary sequence, prevented ColQ1-PD from processing it. However, co-incubation of soluble collagen with increasing concentrations of ColQ1-AD in the presence of a constant amount of ColQ1-PD resulted only in minute degradation fragments after 120 and 150 min in the presence of high ColQ1-AD concentrations (**Fig. 8a**).

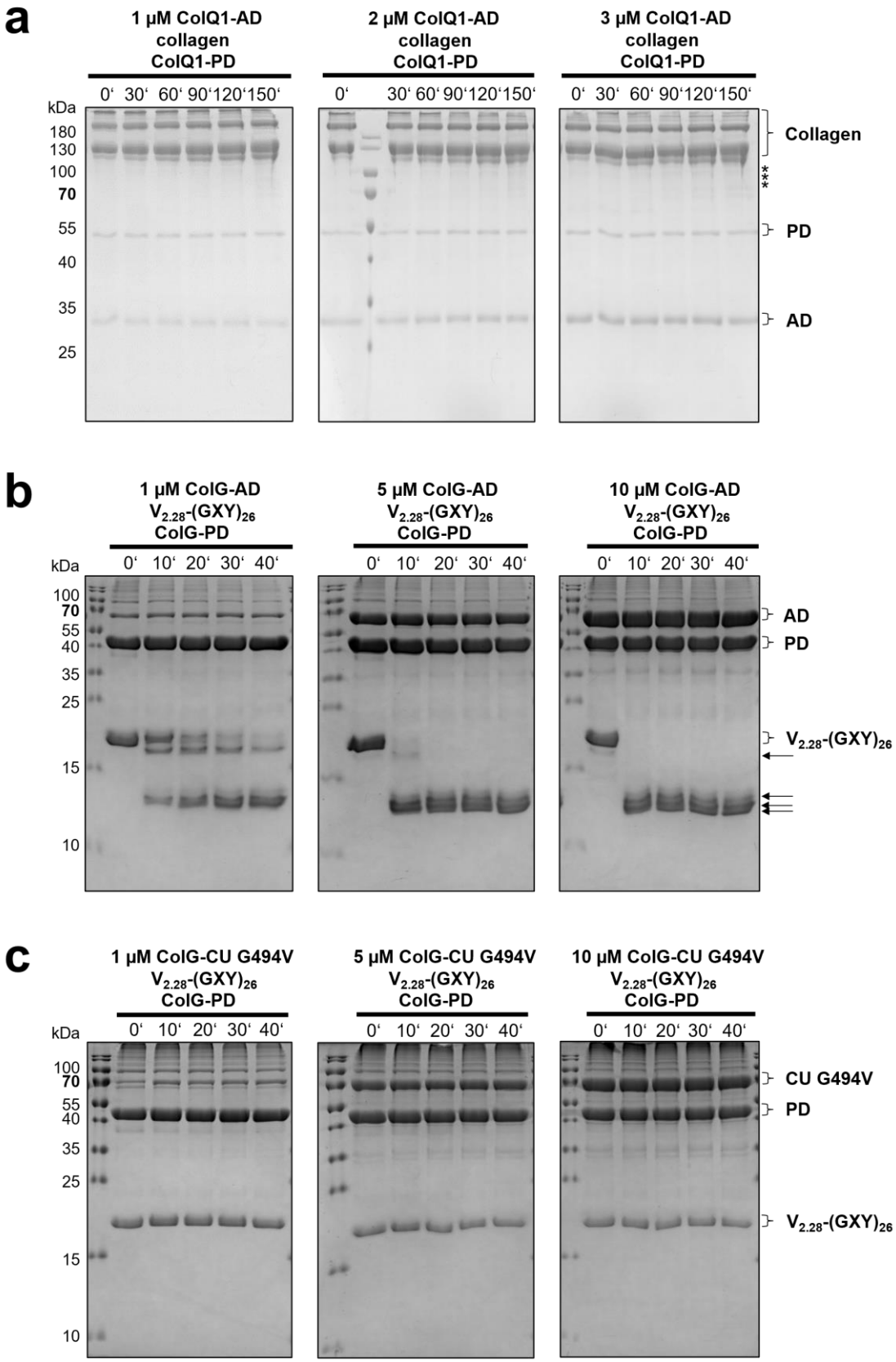


Fig. 8: Triple-helicase activity of AD is not easily detected with collagen as substrate, but can be uncovered with a short collagen mimic. a, Co-incubation of 3.33 µM type I collagen with increasing concentrations of ColQ1-AD (1-3 µM) in the presence of 0.6 µM ColQ1-PD at 25 °C resulted in minor degradation fragments (*) after 120 and 150 min at higher ColQ1-AD concentrations and not in efficient degradation of collagen. **b**, Co-incubation of 10 µM $V_{2.28}$-$(GXY)_{26}$ with 1, 5 or 10 µM ColG-AD-MBP in the presence of 10 µM ColG-PD**. c,** Triple helicase ColG-CU G494V inhibits triple-helix degradation by

ColG-PD. Co-incubation of 10 µM $V_{2.28}$-$(GXY)_{26}$ with 1, 5 or 10 µM ColG-CU G494V in the presence of 10 µM ColG-PD. Control reactions can be found in **Fig. S7**. Reactions were terminated by addition of 35.6 mM EDTA and subjected to SDS–PAGE under non-reducing conditions.

We hypothesized that the apparent failure of the two singular domains to complement each other functionally for collagen degradation was caused by its large dimensions, resulting in a vast binding surface compared to the relatively small number of enzyme molecules added. Therefore, we repeated the complementation assay using a smaller triple-helical substrate, *i.e.* $V_{2.28}$-$(GXY)_{26}$, a modified version of V-$(GXY)_{26}$ containing the V-domain of Scl2.28 [44] instead of the V-domain of Scl2.3 (**Fig. 8b**). By using this smaller substrate, we enhanced the local concentrations of the AD and the PD on the substrate in the assay. Compared to V-$(GXY)_{26}$, $V_{2.28}$-$(GXY)_{26}$ is not SDS-PAGE stable and, thus, it could be clearly differentiated from ColG-AD and ColG-PD on the gel level. Co-incubation of increasing amounts of ColG-AD with 10 µM $V_{2.28}$-$(GXY)_{26}$ in the presence of equimolar concentrations of ColG-PD resulted in a clear dose-dependent degradation of $V_{2.28}$-$(GXY)_{26}$ over time. In the absence of the cutter ColG-PD, the triple-helical substrate was not degraded (**Fig. S6a**). In the absence of ColG-AD, ColG-PD was not able to cleave trimeric $V_{2.28}$-$(GXY)_{26}$ (**Fig. S6b**). Thus, the higher local concentrations of the single ColG-AD and ColG-PD on the surface of the substrate $V_{2.28}$-$(GXY)_{26}$ indeed allowed for the productive interaction of both domains for the degradation of the triple-helical substrate. These results confirmed our previous spectroscopic findings that the AD functions as the triple helicase in bacterial collagenases.

In sum, these results show that the activator domain on its own is able to locally unwind triple-helical collagen, *i.e.* the AD is the minimal entity capable of collagen triple-helicase activity.

AD and the PD need to be in close proximity for efficient catalysis of triple-helical collagen.

Having established the role of the AD as triple helicase and main binding interface to triple-helical collagen, as supported by *Wang et al.*, but also to unfolded collagen, these finding together argued for a major refinement, if not dismissal, of the proposed chew-and-digest model for the processing of triple-helical collagen, as the binding and unwinding of triple-helical collagen was found to be compatible with the open conformation. Therefore, we questioned whether the (semi-)closed conformation of the CU, whose existence was supported by SAXS studies on ColH[33] and molecular dynamics simulations of VhAC[29], was necessary for hydrolysis. We wondered whether efficient collagen turnover required the simultaneous presence of the AD and PD at the binding site and therefore, also a distinct spatial interplay between the AD and the PD, or whether an isolated gelatinase could easily fill in as ‘cutter’ in

the presence of an 'unwinder'. To test this, we repeated the degradation assay with $V_{2.28}$-$(GXY)_{26}$ as substrate, shown in **Fig. 8b**, but replaced ColG-AD as unwinder by ColG-CU G494V (**Fig. 8c**). Interestingly, in the presence of ColG-CU G494V the isolated ColG-PD could no longer process the collagen mimic, indicating that the unwound peptide chains were not accessible to the single PD when ColG-CU G494V was bound to $V_{2.28}$-$(GXY)_{26}$. This suggests that the presentation of unwound substrate chains by the AD happens in close proximity to the PD (**Fig. 1b**), supporting the idea of a closed conformation of the CU for efficient substrate cleavage.

AD residues critical for collagen unwinding

To identify and characterize the residues crucial for substrate recognition and triple-helicase activity in the AD, we performed a bioinformatical analysis and compared the AD sequences of six bacterial collagenases (ColA from *Clostridium perfringens* (P43153), ColG from *C. botulinum* (B2TJU5), ColT from *C. tetani* (Q899Y1), ColG and ColH from *H. histolyticum* (Q9X721 & Q46085), ColQ1 from *Bacillus cereus* (B9J3S4), followed by an alanine scan of strictly or highly conserved, surface-exposed residues on the inner surface of the AD, which faces the PD, based on the crystal structure of ColG as reference model (PDB: 2y6i). Seven promising candidate residues were identified (ColG: F148, E191, R194, Y198, Y201, N251, and F295). We generated the homologous single-point mutants of ColQ1-CU and its catalytically inactive variant ColQ1-CU G472V (ColQ1: F123, E166, R169, Y173, F176, N226, and Y270) to investigate their effects on the hydrolytic activities, and binding affinities to type I gelatin and collagen, and triple-helix unwinding. The structural integrity of the single-point mutants was confirmed for the catalytically active variants using a peptidolytic assay (**Fig. S8**), showing to significant changes in the K_M, and for the inactive variants by monitoring their thermal stability using nanoDSF, as compared to the WT (**Fig. S9**). In addition, the proper folding of the inactive variants was confirmed via the CD spectra taken before melting experiments.

Intriguingly, when performing degradation studies with soluble type I collagen, all seven mutations displayed impaired tropocollagen turnover *in vitro*, while they were able to turnover a peptidic substrate like the WT (**Fig. 9, Fig. S8**). In contrast, the three control mutants (ColQ1: Y251A, N317A, Y321A, also surface-exposed residues in the AD) did not show any aberrant collagen turnover compared to ColQ1-CU WT (**Fig. S10**).

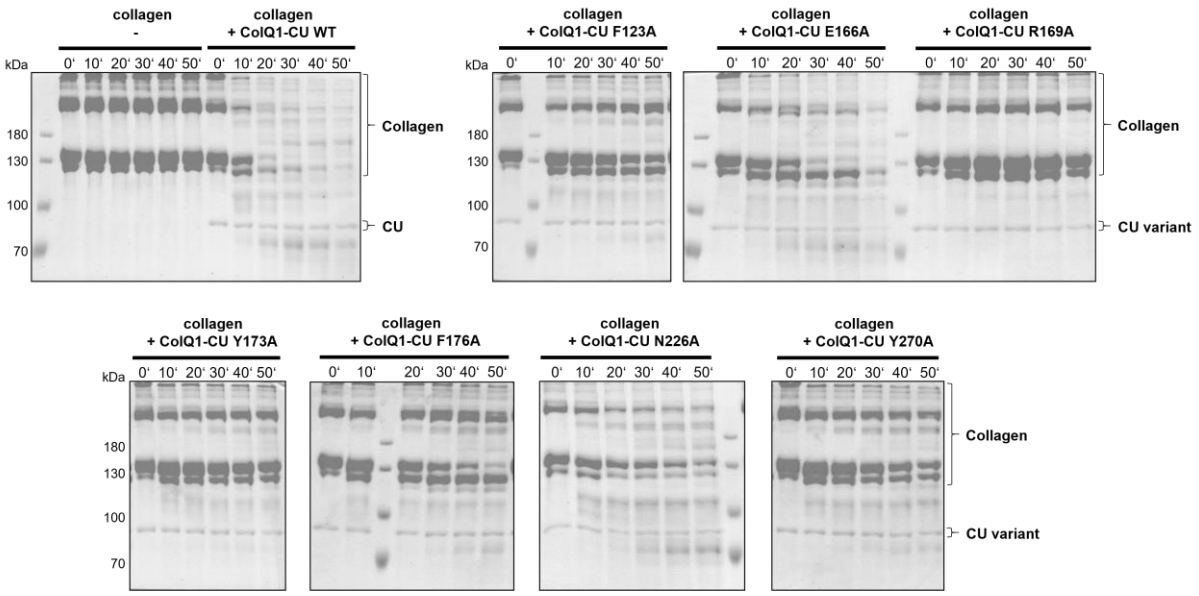


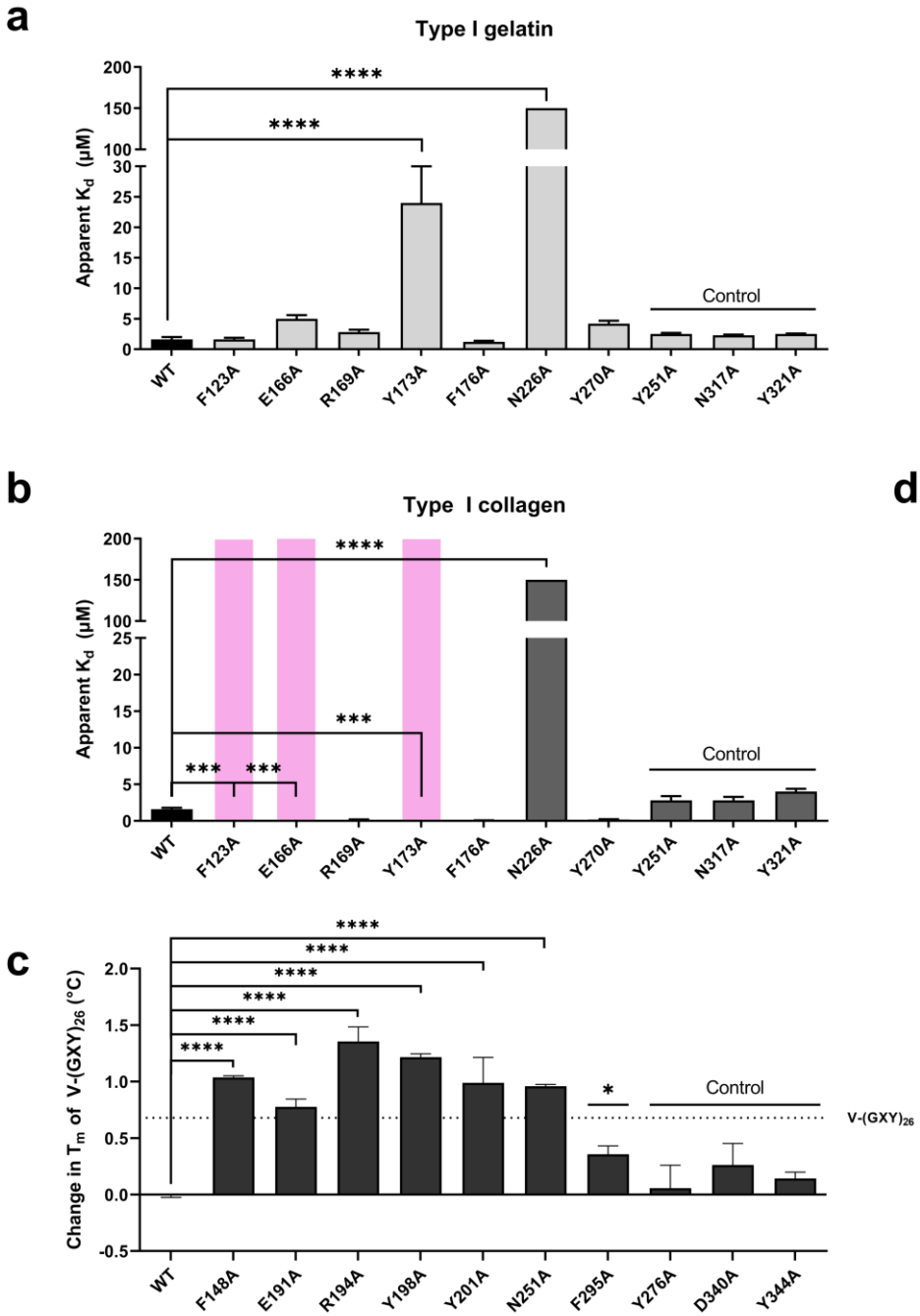
Fig. 9: Activity of ColQ1-CU WT and its mutants towards soluble type I collagen. 3.33 µM type I collagen were incubated with 0.2 µM ColQ1-CU variants at 25 °C. The reactions were stopped by the addition 38 mM EDTA and SDS-loading buffer.

To determine whether this inhibition in collagen degradation was the result of binding and/or unwinding defects, we investigated (i) the binding affinities of the ColQ1-CU mutants towards type I gelatin and type I collagen using ELISA assays, and (ii) the effect of the mutations on the triple-helicase activity by determining the melting temperature of V-$(GXY)_{26}$ co-incubated with the homologous ColG-CU G494V mutants (**Fig. 10 & Table S1**).

In the ELISA binding assays, we investigated the binding of the seven AD mutants compared to the WT, using up to 100 µM of the target proteins. When no binding up to this concentration was detected, this was classified as 'loss in binding affinity'. ColQ1-CU G472V (WT) displayed similar binding affinities to gelatin and type I collagen with an apparent K_d of 1.6 ± 0.4 µM and 1.6 ± 0.2 µM, respectively. The binding assays revealed two crucial residues for gelatin binding in the AD of ColQ1, Y173 and N226. The binding affinities to gelatin dropped drastically – 15-fold and more than 90-fold, respectively, – when these residues were replaced by an alanine (apparent K_d (Y173A) = 24 ± 6 µM and K_d (N226A) > 150 µM). These two residues also proved to be important for collagen binding. In total, four mutations resulted in significant binding defects to triple-helical collagen, *i.e.* F123A, E166A, Y173A and N226A. Similarly to gelatin binding, the N226A mutation resulted in a more than 90-fold reduction in binding affinity to collagen (apparent K_d (N226A) > 150 µM). However, more crucially, the mutations F123A, E166A and Y173A resulted in a complete loss of binding affinity of ColQ1-CU G472V in the assay. In contrast to that and interestingly, the mutations R169A, F176A and Y270A resulted in a marked increase in affinity of the CU to collagen of 9-fold, 20-fold, and 8-fold, respectively.

Next, we examined the effect of the AD mutations on the triple-helicase activity. These were performed using the homologous mutations in ColG-CU G494V. To ease the comparison, we use the following double nomenclature for each mutant, *e.g*. ColG-F148A/ColQ1-F123A. The four mutations in the AD that had resulted in binding defects to collagen in ColQ1-CU (ColG-F148A/ColQ1-F123A, ColG-E191A/ColQ1-E166A, ColG-Y198A/ColQ1-Y173A, and ColG-N251A/ColQ1-N226A), as expected, also failed to unwind V-$(GXY)_{26}$ in the triple-helicase assay. This confirms that triple-helix binding is a necessary prerequisite for unwinding. Interestingly, the two mutants ColG-R194A/ColQ1-R169A and ColG-Y201A/ColQ1-F176A, which displayed a higher binding affinity to soluble collagen than the WT, were also not able to lower the melting temperature of V-$(GXY)_{26}$ upon co-incubation. This suggests that these two residues are critical for the unwinding step. To a lesser degree, this was also found to be the case for the ColG-F295A/ColQ1-Y270A mutation. This finding suggests that triple-helicase unwinding is not only dependent on integral binding properties, but rely also on specific interactions interfere directly or indirectly with the triple-helical structure, e.g. by disturbing its local hydration shell.

In sum, we identified four residues in the AD that are crucial for the recognition of triple-helical collagen (ColG-F148A/ColQ1-F123A, ColG-E191A/ColQ1-E166A, ColG-Y198A/ColQ1-Y173A, and ColG-N251A/ColQ1-N226A), of which ColG-Y198A/ColQ1-Y173A and ColG-N251A/ColQ1-N226A are also key residues for binding to unfolded collagen. In addition, we identified three residues in the AD vital for the unwinding of collagen (ColG-R194A/ColQ1-R169A, ColG-Y201A/ColQ1-F176A, and ColG-F295A/ColQ1-Y270A).



d

Homologous residues

ColQ1	ColG
F123	F148
E166	E191
R169	R194
Y173	Y198
F176	Y201
N226	N251
Y270	F295

Fig. 10: Effect of single-point mutations in AD on binding affinities towards type I gelatin and type I collagen, and on triple-helicase activity. a, Apparent K_d values of ColQ1-CU G472V (=WT) and mutants were determined by indirect ELISA (**a**, **b**). The binding affinities towards collagen were tested at 25 °C to preserve the triple-helical fold, while the binding affinities towards gelatin were tested at 37 °C. Loss of affinity was marked with a magenta bar. **c**, Melting temperatures of V-$(GXY)_{26}$ in presence of ColG-CU G494V (= WT) or its mutants. Molar ellipticities were recorded at 222 nm, while the temperature increased from 25 to 35 °C at a rate of 4.5 °C/h. Homologous variants of ColQ1 and ColG are arranged on top of each other (**a-c**) in the figure. Melting temperature of V-$(GXY)_{26}$ alone is indicated with dotted line. **d**, Table of the homologous residues in ColQ1-AD and ColG-AD that were investigated. Statistical significance was determined by one-way ANOVA followed by Dunnett's test for multiple comparisons (* $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$).

Conformational trapping of the CU through intramolecular crosslinking

To investigate the effect of the spatial geometry and of the dynamic interplay between the AD and the PD on collagen binding and processing in more detail, we generated conformationally

trapped variants of ColG-CU. Conformational control or 'conformational trapping' of multi-domain proteins is a powerful method to elucidate the function of interdomain motion[46–48]. In this study, conformational trapping of ColG-CU was achieved by engineering intramolecular disulfides that crosslink the AD and the PD at various positions. The positions of the crosslinks were spread along the axis of ColG-CU with the aim to trap the CU in different conformational states, from semi-closed to closed (**Fig. 11a,b**). As a model for the quasi-open conformation, we used the ColG-CU Δ(Gly389–Val397) mutant, in which the linker between the AD and PD had been deleted. The activity towards short peptides in unaffected by this linker deletion [23].

Rational design and production of crosslinked ColG-CU variants.

To ensure efficient disulfide-bridge formation, we were looking for non-conserved residues on the inner-facing surfaces of the AD and PD. We generated a model of (semi)-closed conformations of ColG-CU based on PDB entry 2y50 using PYMOL software [36] and identified two residue pairs Y280/Q512 (mutant CL2), and E294/T483 (mutant CL3) for the introduction of cysteines in the upper half of ColG-CU, located at varying distances from the linker region (**Fig. 11a,b**). For mutant CL4, cysteines were introduced in a loop of ColG-PD and just before the N-terminus, in order to crosslink the ColG-CU at the tips of the AD and PD domains, locking the CU in a closed conformation. The mutants CL2-CL4 were generated on the basis of a cysteine-free ColG-CU (C218S/C262S) (mutant CF).

All ColG-CU constructs yielded over 20 mg of homogenous monodisperse protein after purification and oxidation from two liters of *E. coli* cell culture. They migrated with an apparent molecular mass of 79 kDa on a denaturing non-reducing SDS-PAGE gel and were estimated to be approximately 95% pure (**Fig. 11c**). SDS-PAGE analysis revealed that there were negligible amounts of oligomeric forms of the crosslinked mutants, indicating the robustness of the crosslinking approach, and we confirmed that the oxidation process did not negatively affect the collagenolytic activity *per se* (**Fig. S11c**).

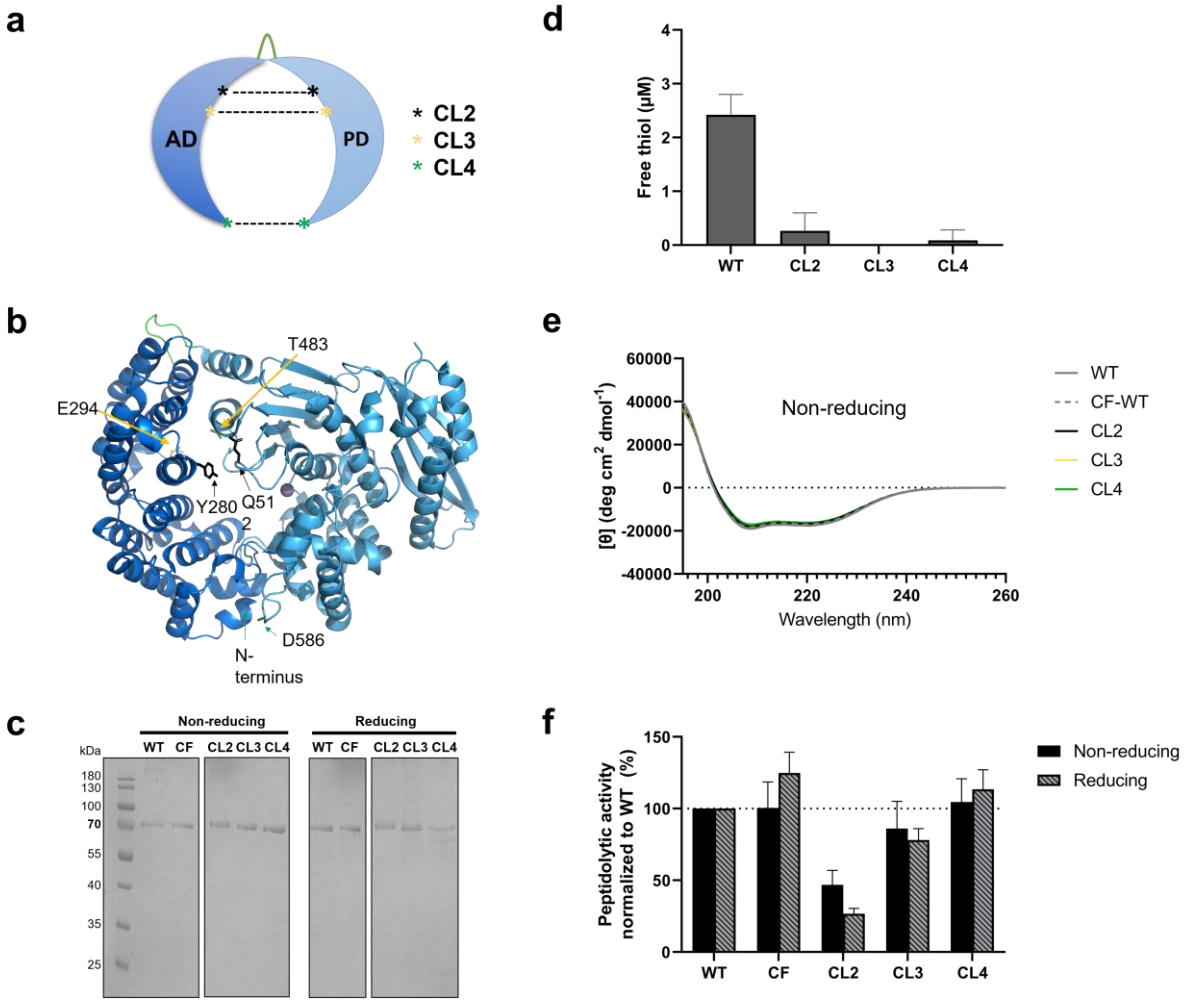


Fig. 11: Design and quality control of crosslinked variants CL1-CL4 of ColG-CU. a, Scheme of crosslinked mutants. **b**, Model of (semi)-closed conformation of ColG-CU. Mutation sites for the introduction of cysteines are shown in sticks. **c**, SDS-PAGE analysis of ColG-CU WT, CF and the mutants CL1-CL4 under non-reducing and reducing conditions on a 12% polyacrylamide gel. **d**, The presence of free thiols was detected using the thiol-specific fluorochrome 7-diethylamino-3-(4-maleinimidophenyl)-4-methyl coumarin after thermal denaturation of the variants at 60 °C. **e**, CD spectra of ColG-CU WT and the mutants CL1-CL4 in the absence of reducing agent. The data shown are a representative of triplicate experiments. **f**, Peptidolytic activity of ColG-CU WT compared to CF and the crosslinked mutants under reducing and non-reducing conditions. 16 nM ColG-CU variants were co-incubated with 2 µM quenched-fluorescent peptide FS1-1 in reaction buffer containing ± 0.5 mM TCEP.

The presence of the disulfide linkage was confirmed by a thiol quantitation assay (**Fig. 11d**). All crosslinked ColG-CU variants were tested at 1 µM concentration. ColG-CU WT which harbors two buried cysteines in the AD was used as positive control and we could confirm the presence of the disulfide bonds in CL2 to CL4. Non-reducing CD spectroscopy analysis showed that all mutants had a secondary structure similar to ColG-CU WT, suggesting that the formation of the disulfide bond did not compromise the overall fold (**Fig. 11e**). Finally, we compared the activity of the mutants towards a small quenched-fluorescence peptide substrate to the activity of ColG-CU WT to confirm the proper folding of the PD in the crosslinked variants (**Fig. 11f**). The removal of the two native cysteines in the AD in CF did not compromise its peptidolytic activity and collagenolytic activity (**Fig. S12**). In CL3 and CL4, the additional

introduction of the cysteines for crosslinking also did not inhibit peptide hydrolysis in the reduced state (86 ± 19%, 105 ± 16%, respectively) and in the oxidized state, when the crosslink was established (78 ± 8%, and 113 ± 14%, respectively). However, mutant CL2 showed a notably reduced substrate turnover in both states (47 ± 10% and 27 ± 4% residual activity under non-reducing and reducing conditions, respectively).

Triple-helical collagen turnover, but not binding is significantly reduced in the open conformation of the CU.

The collagenase unit from ColG spans from top of the linker region to the tips of the AD and PD approx. 70 Å. Assuming a hinge-like architecture of the CU in bacterial collagenases, the oxidized mutants CL3 and CL4 are expected to constrain the conformational freedom of the ColG-CU to varying degrees. In CL3 the engineered crosslink is located about one-third down the length of ColG-AD (approx. 27 Å below the linker), but still around 8 Å above the level of the active-site cleft. It is designed to trap the CU in a semi-closed conformation. In CL4, where the AD and PD are linked at their tips, this will maintain the CU in a closed conformation (**Fig. 11b**); while the deletion of the linker in ColG-CU Δ(Gly389–Val397) (hereafter: Δlinker) enforces a quasi-open conformation of the CU.

When looking at the collagenolytic activity towards soluble type I collagen of the Δlinker mutant compared to the WT, we observed a drastic inhibitory effect. Compared to the WT, the linker deletion nearly abolished collagen degradation (**Fig. 12a**). This result suggests that in the open conformation cleavage of the triple-helical substrate is severely hampered, most likely because the presentation of the substrate to the PD is compromised. This conclusion is supported by the observation that the binding affinity towards tropocollagen was unchanged compared to the WT (**Fig. 12d**).

When we compared the collagenolytic activities of CL3 and CL4 towards soluble type I collagen in the crosslinked *vs*. open state (*i.e.* under non-reducing and reducing conditions), respectively, the results were intriguing (**Fig. 12b,c**). The crosslinkage in CL3 did not compromise tropocollagen turnover. CL3 could degrade soluble collagen similarly to the WT, and no significant difference between crosslinked and non-crosslinked state was observed, suggesting that the semi-closed conformation did not interfere with binding and processing of the triple-helical substrate. Only the conformationally most restricted variant CL4 showed a 6-fold slower collagen turnover than the WT. This reduction in collagen turnover could be completely reversed by opening of the crosslink via the addition of ß-mercaptoethanol (**Fig. 12b,c**).

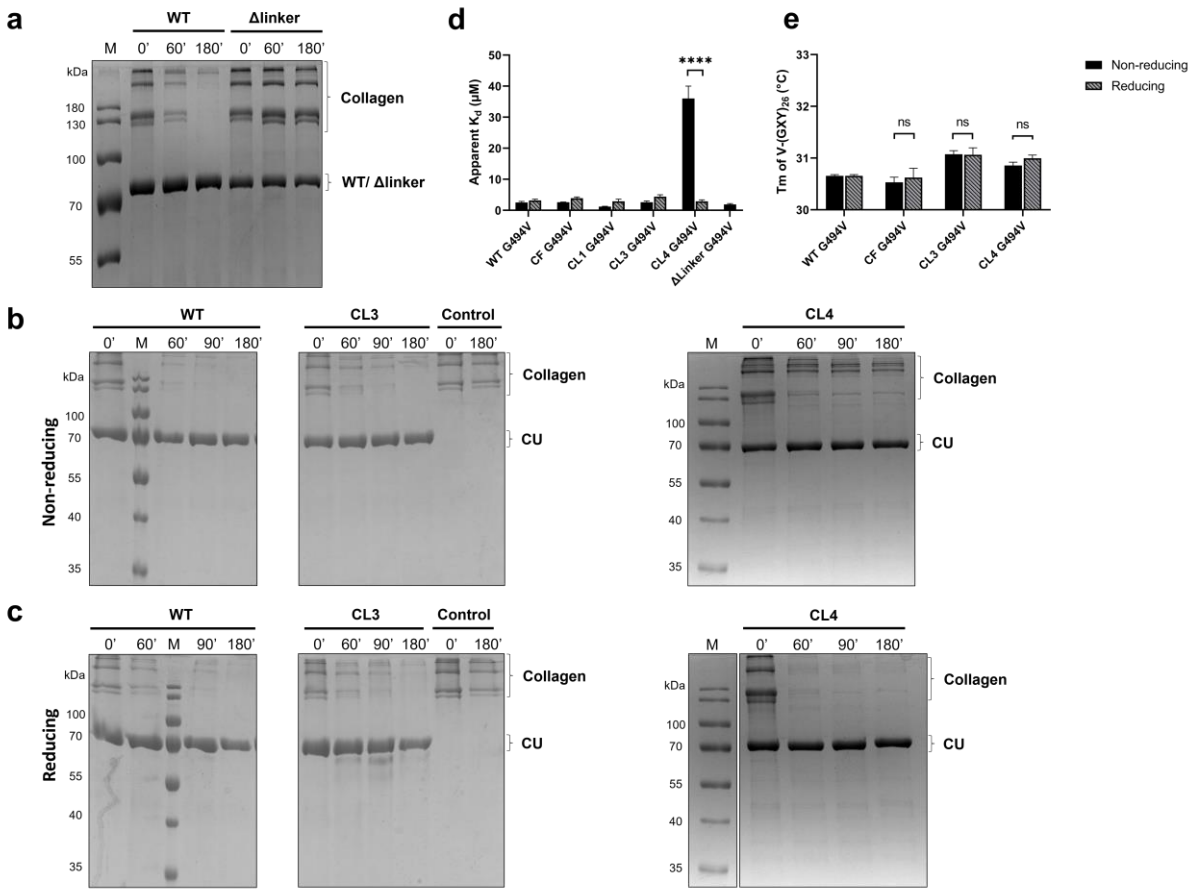


Fig. 12: Conformational trapping and its effect on triple-helix degradation and binding. a-c, Degradation of soluble collagen. 3.33 µM type I tropocollagen were incubated at 25 °C with 4.54 µM ColG-CU variant for up to 3 h in reaction buffer ± 10 mM ß-mercaptoethanol and stopped by the addition of 38 mM EDTA. The degradation was monitored on 12% non-reducing SDS-PAGE gels. **d,** Effect of conformational trapping on binding to triple-helical V-$(GXY)_{26}$. Apparent K_d values of ColG-CU G494V (=WT G494V) and of the mutants were determined by microscale thermophoresis at 22 °C using fluorescently labelled variants. Statistical significance was determined by one-way ANOVA followed by Dunnett's test for multiple comparisons (**** $P < 0.0001$). The binding curves can be found in **Fig. S14**. **e,** Melting profiles of the short collagen mimic V-$(GXY)_{26}$ in presence of WT G494V and the mutants. 5 µM V-$(GXY)_{26}$ were incubated with 10 µM ColG-variant. Molar ellipticities were recorded at 222 nm, while the temperature increased from 25 to 35 °C at a rate of 4.5 °C/h. Statistical significance was determined by one-way ANOVA followed by Šídák's test for multiple comparisons ($P < 0.05$). Experiments (**d**, **e**) were performed in the presence or absence of 1 mM ß-mercaptoethanol.

To see whether this decrease in collagen cleavage was the result of impaired triple-helix binding and/or compromised triple-helix unwinding, we examined the binding and unwinding behavior of the crosslinked mutants (CF, CL1-CL4) compared to ColG-CU (WT) using catalytically inactivated variants harboring the G494V mutation. The binding affinities towards the collagen mimic V-$(GXY)_{26}$ in the absence or presence of 1 mM ß-mercaptoethanol were determined using microscale thermophoresis. We found no significant change in binding affinity to V-$(GXY)_{26}$ in CF G494V and CL3 G494V compared to WT G494V, neither under non-reducing conditions (apparent K_d = 2.5 ± 0.4 µM), nor under reducing conditions (apparent K_d = 3.2 ± 0.4 µM) (**Fig. 12d**). Remarkably, however, we observed a ~12-fold drop in binding affinity of CL4 G494V in the crosslinked state compared to the non-crosslinked state (apparent K_d = 36 ±

4 µM *vs*. 2.9 ± 0.4 µM, respectively). The 'closed' conformation led to a sharp decrease in binding affinity, explaining the observed decrease in tropocollagen turnover by CL4 in the non-reduced state.

To examine whether also the triple-helicase activity of the CU was affected by the conformational constraints, we performed, as described above, the CD-based denaturation assay which monitors the unfolding of triple-helical segment in V-$(GXY)_{26}$ in the presence of potential 'unwinding' ColG variants. However, we found no ß-mercaptoethanol-dependent effect on the melting temperatures of V-$(GXY)_{26}$ in the presence of CL3 G494V and CL4 G494V (**Fig. 12e**).

Conformational trapping impairs cleavage of collagen fibers, but not their binding.

Finally, we investigated the effect of conformational trapping of the ColG-CU on the recognition and processing of insoluble collagen fibers (**Fig. 13**). We wondered whether these processes are more sensitive to conformational CU restraints imposed by the disulfide crosslinks, given the complex architecture of the fibers, which are multi-hierarchically organized from tropocollagen molecules (1.5 nm diameter) into microfibrils (> 5 nm diameter)[9] and then further into fibrils (20 – 500 nm diameter)[49].

Importantly, we observed no significant binding defects in all mutants compared to the WT G494V. Substrate recognition was neither substantially impaired by the linker deletion, nor by the conformational restraints imposed by the crosslinks. Thus, in contrast to binding to soluble collagen, the closed conformation did not result in a major decrease in binding affinity. Apparently, in the disulfide-restrained CU synergistic AD-PD interactions with collagen fibrils were still able to compensate for effects of impaired cooperativity between the AD and the PD in the context of triple-helical collagen (cf. **Fig. 13**, **Fig. 3**).

However, when looking at the processing of collagen fibers by the crosslinked mutants, the findings were puzzling. The Δlinker mutant with its locked-open conformation showed a drastically diminished rate of collagen fiber cleavage, a reduction of 77 ±4% compared to the WT. CL3, trapped in a semi-closed conformation, showed also a notable 47 ± 7% decrease in fiber processing compared to the WT, which could be completely reversed by the addition of a reducing agent. These findings are in line with the chew-and-digest model for the processing of microfibrils, in which domain movements of the CU are required for efficient provisioning of individual tropocollagen molecules from microfibrils for unwinding and cleavage. One might expect an even sharper decrease in fiber cleavage in case of the crosslinked mutant CL4.

However, CL4, the closed CU, exhibited only a 19 ± 9% reduction in activity compared to the WT in the crosslinked state. At first glance, this finding appears contradictory, as CL4-disulfide will restrain the opening of the CU more than that in CL3. We wish to emphasize, however, that the CL4 variant, which is locked "at the bottom" tip, can benefit from inter-domain flexibility by the flexible linker "at the top" connecting the AD and PD, possibly more than the variant CL3, which is restrained by a central disulfide bond.

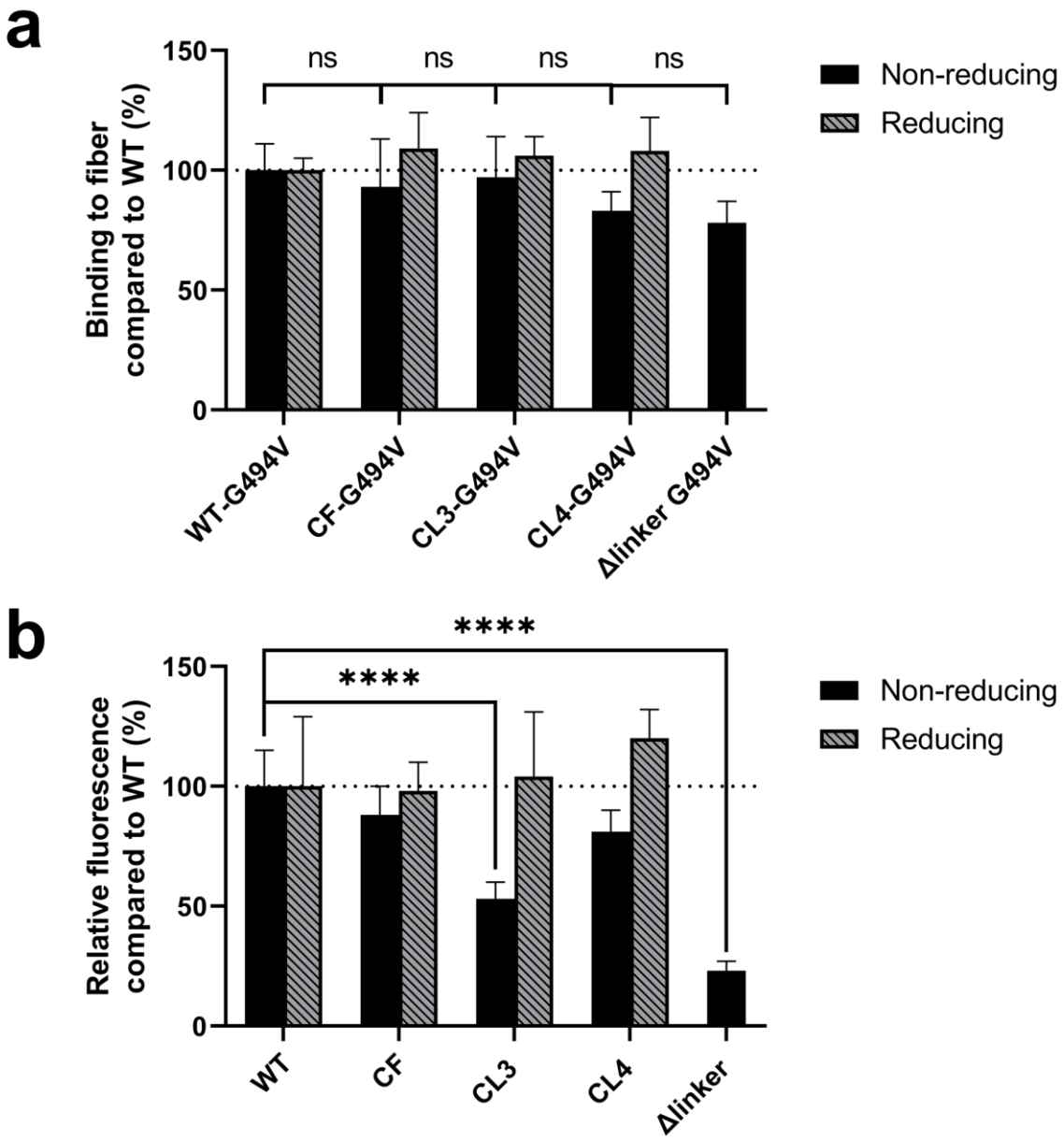


Fig. 13: Binding and processing of insoluble collagen fibers by ColG-CU (WT) and its crosslinked variants. a, Binding to insoluble collagen fibers. The catalytically inactive ColG-CU variants were fluorescently labelled and incubated at 0.2 µM for 30 min at RT with 6 mg insoluble collagen fibers from bovine Achilles tendon in the presence or absence of 10 mM ß-mercaptoethanol. The fluorescence remaining in the supernatant after incubation was detected at 647/680 nm. **b,** Degradation of insoluble collagen fibers. 2 mg insoluble fibrillar collagen from bovine Achilles tendon were incubated with 1.0 µM ColG-CU variants and incubated for 2 h at room temperature in the presence or absence of 10 mM ß-mercaptoethanol. The reaction filtrate was collected and supplemented with 38 mM EDTA to stop the reaction. The amount of hydrolysis was quantified by a fluorescamine-citrate assay (excitation: 390 nm, emission: 475 nm). All experiments were performed at least in triplicates and results are presented as mean ± SD.

Discussion

Collagen is hierarchically organized, with soluble triple-helical collagen assembling into insoluble microfibrils and fibers. The data presented here show that bacterial collagenases, by virtue of their multi-domain structures, recognize, disassemble and process these different collagen conformers in distinct and previously underappreciated ways.

Our data reveal that the collagen-binding domain CBD2 of ColG binds fibrillar collagen more tightly than the activator domain AD, while the binding affinities are reversed for soluble triple-helical collagen; this indicates that CBD2 employs multiple binding sites, benefiting from multivalent interaction in collagen fibers. The AD dominates the binding interaction toward triple-helical collagen. Unexpectedly, however, the protease domain PD significantly impacts the binding to collagen in a conformer-dependent manner. While within the collagenase unit CU, consisting of AD and PD, PD interacts synergistically to increase the binding affinity to insoluble collagen fibers (**Fig. 3**), the opposite is true for soluble collagen (**Fig. *4***), where PD antagonizes AD binding. These remarkable and counterintuitive observations can be rationalized by general principles of conformational selection and multivalency: the structure of the CU adopts a saddle-like shape that matches the dimension of collagen microfibrils and allows for multivalent binding interactions (**Fig. 1b**). By contrast, in case of soluble triple-helical collagen, such attractive multivalent interactions are sterically not possible, while the PD still blocks access to the AD, explaining the observed antagonistic effect towards conformationally rigid triple-helices.

AD's prominent role is not limited to collagen binding. We show here that the AD can unwind collagen triple helices. Monitoring AD-induced unfolding of collagen is challenging because the unfolding is locally restricted, transient and reversible (**Fig. 5**, **Fig. 6, Fig. 7**). These results could be further corroborated by identifying and characterizing individual AD sites, which inactivated its unwinding activity either by loss of collagen binding or by gain of collagen binding and stabilization (**Fig. 8**, **Fig. 9, Fig. 10**).

Finally, we tested whether the discovered synergistic and antagonistic domain interactions were consistent with the previously proposed chew & digest model [23]. To this end, we used conformationally restrained CU variants, which either locked the CU in an open state by minimizing the linker between AD and PD ("ΔLinker") or by disulfide linking the AD with the PD, thus locking in more closed CU conformations (**Fig. 11**). Again, we found that the

mechanism of collagen processing depends on the collagen conformer serving as the substrate. For soluble collagen, the AD-PD flexibility does not seem to play a major role (**Fig. 12**). By contrast, the insoluble, fibrous collagen indeed requires the AD-PD flexibility, not for binding, but for efficient processing (**Fig. 13**). Consequently, bacterial collagenases cleave insoluble collagen fibers by a chew & digest mechanism.

Materials and methods

Materials. Acid-soluble type I atelocollagen from bovine hides was purchased from Cell Guidance Systems (UK). Type I gelatin was produced by heating the acid-soluble type I collagen for 5 min to 95°C. Type I collagen fibers from bovine achilles tendon were purchased from Merck (Germany).

***In silico* analysis of the ColG linker region.** Multiple-sequence alignment of the linker region in ColG (Q9X721) and ColH (Q46085) from *H. histolytica*, ColA (A0A6L7H295) from *B. anthracis*, and ColA (Q81BJ6) from *B. cereus ATCC 14579*, Ghcol (F7IZI6) from *G. hollisae*, VhaC (A0A8B3DGT3) from *V. harveyi VHJR7*, and ColA (Q9KRJ0) from *V. cholerae serotype O1* were performed using Clustal Omega (Sievers et al. 2011). The linker conservation in subfamily M9B was visualized using the WebLogo program (Crooks et al. 2004) based on the sequences of 11 clostridial (Q899Y1, Q9X721, Q46085, P43153, Q84IM1, Q84IL7, Q84IN0, Q84IL2, B1L1W8, Q84IM4, and Q84IM7) and of 7 bacillary collagenases (B9J3S4, A0A6L7H295, Q81BJ6, EEK67530.1, WP_011983597.1, A0A2J9BDH1, WP_098635255.1).

Construction of V-$(GXY)_{26}$. The coding sequences of V-B from Scl2.28 incorporating the 4 mutated tripeptides and the coding sequence of the V domain from Scl2.3 were purchased from Genscript (Germany). The modified coding sequence of V-B was cloned into a modified pET15b vector. The endogenous V domain was replaced by the V domain from Scl2.3 via Gibson assembly. All constructs were verified by sequencing at Eurofins Genomics (Germany).

Construction of ColG-CU variants. Based on a pET15b expression plasmid of ColG-CU WT (Tyr119–Gly790) [23], a plasmid encoding a cysteine-free ColG-CU variant (CF) (C218S, C262S) was generated by Gibson assembly. Using site-directed mutagenesis via inverse PCR [50], the plasmids for the ColG variants CL1, CL2, CL3, and CL4 were obtained from the

cysteine-free template, while the linker variants were generated from the plasmid encoding ColG-CU WT. All constructs were verified by sequencing at Eurofins Genomics (Germany).

Expression and purification of ColG variants, ColQ variants and V-(GXY)$_{26}$. All protein variants were expressed in *E. coli* Nico21 (DE3) cells and purified as described previously [39,51]. Monodisperse protein fractions were collected and stored at -80 °C, and the sample purity was confirmed by SDS-PAGE analysis under denaturing conditions.

Preparation of (GXY)$_{26}$. 11 µM V-(GXY)$_{26}$ were co-incubated in a molar ratio of 1:20 with pepsin in 50 mM acetic acid pH 3.0 at 4 °C for 72 h. The reaction was stopped by pH adjustment to pH 7.4 and the addition of 10 mM ß-mercaptoethanol. The structural integrity of (GXY)$_{26}$ was confirmed by co-incubation with α-chymotrypsin in a molar ratio 1:10 for up to 3 h (data not shown).

Disulfide formation and thiol quantification assay. For the crosslinkage, purified collagenases (0.1 mg/ml final concentration) were suspended in 50 mM Tris-HCl pH 8.5, 1 mM β-mercaptoethanol, 300 mM NaCl, 5% glycerol, 1 mM $CaCl_2$ and 3 mM NaN_3. The oxidation reactions were kept at 4 °C for 10 days followed by centrifugation (16,500 g for 20 min). The non-oxidized molecules were removed by Activated Thiol Sepharose™ 4B chromatography (Sigma Aldrich) performed according to the manufacturer's recommendations. In short, the clarified samples were loaded onto a self-packed activated thiol sepharose 4B column equilibrated in 50 mM Tris-HCl pH 8.5, 300 mM NaCl, 5% glycerol, 1 mM $CaCl_2$ and 3 mM NaN_3 and incubated overnight at 4 °C. The crosslinked variants were collected in the flowthrough fraction. The crosslinked monomers were separated from misoxidized aggregates by size exclusion chromatography using a Superdex 200 10/300 GL (Cytiva, Germany) and 10 mM HEPES pH 7.5, 100 mM NaCl, 5% glycerol and 3 mM NaN_3 as buffer. The extent of disulfide-bridge formation in the samples CL1-CL4 was examined using the thiol-specific fluorochrome 7-diethylamino-3-(4-maleinimidophenyl)-4-methyl coumarin (CPM) [52] and ColG-CU WT (contains 2 cysteines) as positive control. CPM (Sigma-Aldrich, Germany) was dissolved at 4 mg/ml in DMSO and stored at -80 °C. Prior to use, the stock was diluted in the reaction buffer (10 mM HEPES pH 7.5 and 100 mM NaCl supplemented with 20% DMSO). The assay was performed in a total volume of 120 µL. The protein samples were diluted in the reaction buffer to 1.0 µM. 10 µL diluted dye and 16 µL DMSO were added to 94 µL protein solution. After 3 min of incubation at 60 °C for protein denaturation, the fluorescence was measured in an Infinite M200 plate reader (Tecan, Austria) at an excitation and emission wavelength of 387 nm and 463 nm, respectively.

Circular dichroism spectroscopy. Far UV circular dichroism (CD) spectra in the wavelength range from 195 to 260 nm were recorded using a Chirascan Plus CD Spectrophotometer (Applied Photophysics, Leatherhead, UK) at 25 °C. The instrument was flushed with nitrogen, the pathlength was 0.5 mm, spectral bandwidth was set to 1 nm and the scan time per point 1 s. The samples (2.25 µM type I collagen, 5.0 µM V-$(GXY)_{26}$, 5.0 µM V, 2.25 µM and 10.0 µM ColG-CU G494V, 30.0 µM ColG-CBD2, 2.25 µM α-chymotrypsin, 10 µM ColG-CU G494V mutants, 10.0 µM CF, and 10.0 µM CL1-CL4) were measured in 15 mM Tris-SO_4 pH 7.5, 100 mM NaF, 1 mM $CaCl_2$ and ± 1 mM ß-mercaptoethanol. Melting profiles were measured at similar concentrations at a wavelength of 222 nm from 10 °C to 70 °C or 20 °C to 50 °C (setting time: 120 s, scan time per point: 1 s, step: 0.3 °C) and from 25 °C to 35 °C (setting time: 240 s, scan time per point: 10 s). To verify the proper folding of the samples, CD spectra from 195 to 260 nm were recorded at the start temperature prior to initiating the temperature ramp. All measurements were done in triplicates. In case of the melting temperatures of V-$(GXY)_{26}$ in absence or presence of ColG-CU G494V, ColG-AD and ColG-CBD2, the statistical significance was determined by one-way ANOVA followed by Dunnett's test for multiple comparisons (***$P < 0.001$). When comparing the melting temperatures of the V-$(GXY)_{26}$ in the presence of the different crosslinked variants, the statistical significance was determined by one-way ANOVA followed by Holm-Šídák's multiple comparisons test, comparing the reduced variant to its non-reduced control.

Determination of thermal stability. Thermal denaturation assays were performed using Tycho NT. 6 (NanoTemper Technologies, Germany). The measurements were performed at 0.1 mg/ml protein concentration in 15 mM Tris-SO_4 pH 7.5, 100 mM NaF and 1 mM $CaCl_2$ in triplicates. Intrinsic fluorescence was recorded at 330 and 350 nm while heating the sample from 35 to 95 °C at a rate of 30 °C/min. Fluorescence ratio (350/330 nm) and inflection temperature were calculated by Tycho NT. 6 software.

Peptidolytic assay. The peptide-degradation assay was performed as described previously [53]. In short, all ColG-CU variants were tested at a final concentration of 16 nM and co-incubated with 2 µM of the quenched-fluorescent peptidic substrate Mca-Ala-Gly-Pro-Pro-Gly-Pro-Dpa-Gly-Arg-NH2 (FS1-1) (Mca = (7-Methoxycoumarin-4-yl)acetyl; Dpa = N-3-(2,4-dinitrophenyl)-L-2,3-diaminopropionyl). The final reaction buffer contained 250 mM HEPES pH 7.5, 400 mM NaCl, 10 mM $CaCl_2$, 10 µM $ZnCl_2$, 2% DMSO and 3 mM NaN_3. Reactions were performed in the presence and absence of 0.5 mM TCEP. Cleavage of the substrate was monitored for 2 min at 25 °C (excitation: 328 nm, emission: 392 nm) in an Infinite M200 plate

reader (Tecan, Austria) and the initial velocity (v_0) was determined from the progress curves (<10% substrate conversion).

Steady state measurements were used to determine K_M for FS1-1. The final concentration of ColQ1-CU WT and its variants was 1 nM and the concentration of the substrate was screened from 0 – 175 µM. The initial velocity was determined from the progress curves (<10% substrate conversion) using linear regression and inner filter effect-correction [54]. The ε_{ex328} and ε_{em392} were estimated in the buffer used for the kinetic assay to be 21794 M^{-1} cm^{-1} and 9561 M^{-1} cm^{-1}. K_M was calculated by non-linear regression from the resulting Michaelis − Menten plot using GraphPad Prism 9.1.2 (Graph Pad Software, San Diego, CA, USA).

Degradation of V-(GXY)$_{26}$ monitored via SDS-PAGE. 10 µM V-(GXY)$_{26}$ were digested at 25 °C by 0.5 µM ColG-CU or 0.5 µM α-chymotrypsin in 250 mM HEPES pH 7.5, 400 mM NaCl, 10 mM $CaCl_2$, and 10 µM $ZnCl_2$ for up to 2 h at 25 °C. Samples were taken at indicated time points and the reaction was stopped by addition of SDS-PAGE loading buffer on ice. The degradation was monitored using 12% non-reducing SDS-PAGE gels.

Degradation of V$_{2.28}$-(GXY)$_{26}$ monitored via SDS-PAGE. 10 µM V$_{2.28}$-(GXY)$_{26}$ were co-incubated with 1, 5 or 10 µM ColG-AD-MBP in the presence or absence of 10 µM ColG-PD in 250 mM HEPES pH 7.5, 400 mM NaCl, 10 mM $CaCl_2$, and 10 µM $ZnCl_2$ for up to 40 min at 25 °C. Samples were taken at indicated time points and the reaction was stopped by addition of SDS-PAGE loading buffer on ice. The degradation was monitored using 16% non-reducing SDS-PAGE gels.

Degradation of soluble collagen monitored via SDS-PAGE. 1 mg/ml acid-soluble type I atelocollagen from bovine hides (Cell Guidance Systems, Cambridge, UK) was digested at 25 °C by 4.54 µM collagenase in 250 mM HEPES pH 7.5, 400 mM NaCl, 10 mM $CaCl_2$, 10 µM $ZnCl_2$, and ± 10 mM ß-mercaptoethanol for up to 4 h. Samples were taken at indicated time points and the reaction was stopped by addition of 38 mM EDTA. The integrity of the triple-helical collagen fold was verified by co-incubation with 0.83 µM α-chymotrypsin (FLUKA, Switzerland) (data not shown). The degradation was monitored on 12% non-reducing SDS-PAGE gels. All experiments were performed at least in triplicates. Densitometric analysis were performed using GelAnalyzer 19.1 (www.gelanalyzer.com).

Collagen-fibril degradation monitored via fluorescamine-citrate assay. Two milligrams of insoluble fibrillar collagen from bovine Achilles tendon (Merck, Germany) were added to an Nanosep microcentrifugal device (0.2 µm pore size) with a low-binding Bio-Inert membrane

(Pall, Germany). 500 µL reaction buffer (250 mM HEPES pH 7.5, 400 mM NaCl, 10 mM $CaCl_2$, and 10 µM $ZnCl_2$, ± 10 mM ß-mercaptoethanol) were added for 15 min at room temperature to swell the fibers and then removed via centrifugation (13,000 g for 2 min). Then 200 µL 1.0 µM ColG-CU variants were added and incubated for 2 h at room temperature. The filtrate was collected and supplemented with 38 mM EDTA to stop the reaction. The amount of hydrolysis was quantified in comparison to the control reaction with ColG-CU WT exploiting the N-terminal-specific adduction of fluorescamine to peptides proteins at mildly acidic pH [55]. In short, 5 µL of the stopped reaction were diluted 1:10 with reaction buffer and then mixed 1:1 with 1 M citrate pH 5.6. 100 µL of this mixture were added to 10 µL 2.5 mg/ml fluorescamine in acetone and incubated for 5 min at room temperature, before the fluorescence was measured at 25 °C (excitation: 390 nm, emission: 475 nm) in an Infinite M200 plate reader (Tecan, Austria). All experiments were performed at least in triplicates.

Indirect ELISA. For the coating, 1 mg/ml bovine type I atelocollagen stock solutions were prepared in 0.1 M HCl and diluted to 5 µg/ml in 20 mM phosphate buffer pH 7.4, 150 mM NaCl. Gelatin solutions were prepared by heating type I collagen stock for 5 min at 95 °C. 96-well high-binding microplates (Greiner Bio-One, Germany) were incubated with 100 µL coating solution per well overnight at 4 °C for collagen plates, and at 37 °C for gelatin plates. After incubation, the plates were washed four times with PBST. Coated wells were blocked with 1x PBS supplemented with 10% skim-milk for 90 min and then washed four times with PBST. Collagen plates were stabilized with PBST supplemented with 1% BSA fraction V and 5% sucrose. Plates were dried and stored at 4 °C. For the binding assay, the hexahistidine-tagged ColG variants were prepared in PBST supplemented with 1% BSA fraction V. For K_D determination, the samples were serially diluted 1:3. 75 µL sample per well were incubated for 140 min with collagen at room temperature, while gelatin plates were incubated at 37 °C for the same time period. The plates were washed four times with PBST and were then incubated with 1:15,000 rabbit polyclonal 6x His-tag antibody conjugated to HRP (Abcam, Austria) for 1 h at room temperature, followed by four washes with PBST and a final wash with 1x PBS. As substrate 75 µL 3,3′,5,5′-tetramethylbenzidine were added per well. The peroxidase activity was followed by measuring the absorption at 650 nm every 15 s for 225 s in an Infinite M200 plate reader (Tecan, Austria) at 25 °C. The initial velocity was determined by regression analysis. For K_d determination, the data was fitted to the Hill equation using GraphPad Prism 9 (Graph Pad Software, USA). The apparent dissociation constant K_d is given as mean values of three independent experiments ± standard deviation.

Binding assay to insoluble collagen fibers. ColG-variants were labeled using the Monolith NT™ Protein Labeling Kit RED-NHS 2nd Generation Amine reactive (NanoTemper, Germany) and the yield of labelled protein and the degree of labelling were determined according to the manufacturer's manual. Insoluble type I collagen from bovine Achilles tendon (Sigma, Germany) (0, 2, 4, 6, and 8 mg or 6 mg) was prewetted with 225 µL reaction buffer (50 mM Hepes pH 7.5, 100 mM NaCl, 10 mM $CaCl_2$, 0.1% Tween-20, 1% fraction V of bovine serum albumin, 3 mM NaN_3, ± 10 mM ß-mercaptoethanol) for 15 min at RT, before 100 µL labelled ColG-variants solubilized in reaction buffer were added. The final concentration of labelled protein was 0.2 µM, except for ColG-PKD, which was added at 5 µM final concentration to ensure a proper signal-to-noise ratio. The reactions were incubated at 25 °C for 30 min with stirring in the dark and then centrifuged at 13,000 g for 5 min at RT. Labelled ColG variants in reaction buffer without any substrate were used as controls. The collagen-binding ability of labelled proteins was determined monitoring the free fluorescence intensity in the supernatant after incubation using a Tecan M200 Infinite plate reader (Tecan, Austria) (647 nm excitation/680 nm emission).

Binding assay to V-$(GXY)_{26}$ monitored via microscale thermophoresis. Microscale thermophoresis experiments were performed on a NanoTemper Monolith NT.115 instrument (NanoTemper Technologies). ColG variants were labeled using the Monolith NT™ Protein Labeling Kit RED-NHS 2nd Generation Amine reactive (NanoTemper Technologies, Germany). After labeling, the ColG variants were eluted into 250 mM Hepes pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 10% glycerol, 3 mM NaN_3 and stored at -80 °C. The yield of labelled protein and the degree of labelling were determined according to the manufacturer's manual. For the assay, the labelled proteins and the ligand were diluted into 50 mM Hepes pH 7.5, 150 mM NaCl, 10 mM $CaCl_2$, 10 µM $ZnCl_2$, 0.05% Tween-20, 3 mM NaN_3, ± 1 mM ß-mercaptoethanol. The ColG variants were used at a concentration of final concentration of 25 nM (except for 9xAla-Linker where 40 nM were used because of the low degree of fluorescent labelling), while the ligand was titrated in a 1:2 dilution series. After that ColG variants and ligand were mixed 1:2 and the samples were centrifuged for 10 min at 17,000 g at RT. The solutions were immediately transferred into Monolith NT.115 standard capillaries and measured using 60% excitation power at 22 °C. The binding of the ligand caused a reduction in the initial fluorescence signal, confirmed via specificity test performed according to the manufacturer's guidelines. The experiments were performed in triplicates. The change in the initial fluorescence signal was used to calculate the apparent binding constant K_d using non-linear regression analysis in GraphPad Prism 9 (Graph Pad Software, USA).

Acknowledgments

The authors are grateful for funding by the Austrian Science Fund (FWF), P31843.

Declaration of interests

The authors declare no competing interests.

References:

1. Burgeson, R. E. & Nimni, M. E. Collagen types. Molecular structure and tissue distribution. *Clin. Orthop.* **282**, 250–72 (1992).
2. Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **78**, 929–58 (2009).
3. Gelse, K., Pöschl, E. & Aigner, T. Collagens—structure, function, and biosynthesis. *Adv. Drug Deliv. Rev.* **55**, 1531–1546 (2003).
4. Kadler, K. E., Holmes, D. F., Trotter, J. A. & Chapman, J. A. Collagen fibril formation. *Biochem. J.* **316 (Pt 1)**, 1–11 (1996).
5. Ricard-Blum, S. The collagen family. *Cold Spring Harb. Perspect. Biol.* **3**, a004978 (2011).
6. Ramshaw, J. A., Shah, N. K. & Brodsky, B. Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides. *J. Struct. Biol.* **122**, 86–91 (1998).
7. Bruckner, P. & Prockop, D. J. Proteolytic enzymes as probes for the triple-helical conformation of procollagen. *Anal. Biochem.* **110**, 360–8 (1981).
8. Bächinger, H. P., Bruckner, P., Timpl, R., Prockop, D. J. & Engel, J. Folding mechanism of the triple helix in type-III collagen and type-III pN-collagen. Role of disulfide bridges and peptide bond isomerization. *Eur. J. Biochem. FEBS* **106**, 619–32 (1980).
9. Orgel, J. P. R. O., Irving, T. C., Miller, A. & Wess, T. J. Microfibrillar structure of type I collagen in situ. *Proc. Natl. Acad. Sci.* **103**, 9001–9005 (2006).

10. Scott, J. E. Structure and function in extracellular matrices depend on interactions between anionic glycosaminoglycans. *Pathol. Biol. (Paris)* **49**, 284–289 (2001).
11. Raspanti, M. *et al.* Glycosaminoglycans show a specific periodic interaction with type I collagen fibrils. *J. Struct. Biol.* **164**, 134–139 (2008).
12. Nagase, H., Visse, R. & Murphy, G. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc. Res.* **69**, 562–73 (2006).
13. Fields, G. B. Interstitial collagen catabolism. *J. Biol. Chem.* **288**, 8785–93 (2013).
14. Schönauer, E. & Brandstetter, H. Inhibition and Activity Regulation of Bacterial Collagenases. *Top. Med. Chem.* 1–26 (2016) doi:10.1007/7355_2016_9.
15. Bond, M. D. & Van Wart, H. E. Characterization of the individual collagenases from Clostridium histolyticum. *Biochemistry* **23**, 3085–91 (1984).
16. Seifter, S. & Harper, E. The Enzymes. in (ed. Boyer, P.) 649–697 (Academic Press, 1971).
17. Wang, Y. *et al.* Mechanistic Insight into the Fragmentation of Type I Collagen Fibers into Peptides and Amino Acids by a Vibrio Collagenase. *Appl. Environ. Microbiol.* e0167721 (2022) doi:10.1128/aem.01677-21.
18. Hatheway, C. L. Toxigenic clostridia. *Clin. Microbiol. Rev.* **3**, 66–98 (1990).
19. Singh, B., Fleury, C., Jalalvand, F. & Riesbeck, K. Human pathogens utilize host extracellular matrix proteins laminin and collagen for adhesion and invasion of the host. *FEMS Microbiol. Rev.* **36**, 1122–80 (2012).
20. Matsushita, O. & Okabe, A. Clostridial hydrolytic enzymes degrading extracellular components. *Toxicon* **39**, 1769–1780 (2001).
21. Watanabe, K. Collagenolytic proteases from bacteria. *Appl. Microbiol. Biotechnol.* **63**, 520–6 (2004).
22. Bruggemann, H. *et al.* The genome sequence of Clostridium tetani, the causative agent of tetanus disease. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1316–21 (2003).

23. Eckhard, U., Schönauer, E., Nüss, D. & Brandstetter, H. Structure of collagenase G reveals a chew-and-digest mechanism of bacterial collagenolysis. *Nat. Struct. Mol. Biol.* **18**, 1109–14 (2011).
24. Jung, C. M. *et al.* Identification of metal ligands in the Clostridium histolyticum ColH collagenase. *J. Bacteriol.* **181**, 2816–22 (1999).
25. Lee, J.-H. *et al.* Characterization of the enzyme activity of an extracellular metalloprotease (VMC) from Vibrio mimicus and its C-terminal deletions. *FEMS Microbiol. Lett.* **223**, 293–300 (2003).
26. Matsushita, O. *et al.* Gene duplication and multiplicity of collagenases in Clostridium histolyticum. *J. Bacteriol.* **181**, 923–33 (1999).
27. Yu, M. S. & Lee, C. Y. Expression and characterization of the prtV gene encoding a collagenase from Vibrio parahaemolyticus in Escherichia coli. *Microbiol. Read. Engl.* **145 (Pt 1**, 143–50 (1999).
28. Ikeuchi, T. *et al.* Crystal structure of Grimontia hollisae collagenase provides insights into its novel substrate specificity toward collagen. *J. Biol. Chem.* **298**, (2022).
29. Wang, Y. *et al.* Structure of Vibrio collagenase VhaC provides insight into the mechanism of bacterial collagenolysis. *Nat. Commun.* **13**, (2022).
30. Matsushita, O. *et al.* A study of the collagen-binding domain of a 116-kDa Clostridium histolyticum collagenase. *J. Biol. Chem.* **273**, 3643–3648 (1998).
31. Eckhard, U., Huesgen, P. F., Brandstetter, H. & Overall, C. M. Proteomic protease specificity profiling of clostridial collagenases reveals their intrinsic nature as dedicated degraders of collagen. *J. Proteomics* **100**, 102–14 (2014).
32. Matsushita, O. *et al.* A study of the collagen-binding domain of a 116-kDa Clostridium histolyticum collagenase. *J. Biol. Chem.* **273**, 3643–3648 (1998).
33. Ohbayashi, N. *et al.* Solution structure of clostridial collagenase H and its calcium-dependent global conformation change. *Biophys. J.* **104**, 1538–45 (2013).
34. Eckhard, U., Schönauer, E. & Brandstetter, H. Structural basis for activity regulation and substrate preference of clostridial collagenases g, h, and T. *J. Biol. Chem.* **288**, 20184–94 (2013).

35. Tanaka, K. *et al.* Recombinant collagenase from Grimontia hollisae as a tissue dissociation enzyme for isolating primary cells. *Sci. Rep.* **10**, 3927 (2020).

36. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. (2010).

37. Chung, L. *et al.* Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *EMBO J.* **23**, 3020–30 (2004).

38. Manka, S. W. *et al.* Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12461–6 (2012).

39. Hoppe, I. J., Brandstetter, H. & Schönauer, E. Biochemical characterisation of a collagenase from Bacillus cereus strain Q1. *Sci. Rep.* **11**, 4187 (2021).

40. Arrhenius, S. Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte. *Z. Für Phys. Chem.* **4U**, 96–116 (1889).

41. Arrhenius, S. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Z. Für Phys. Chem.* **4U**, 226–248 (1889).

42. Lei, Q., Ren, C., Su, X. & Ma, Y. Crowding-induced Cooperativity in DNA Surface Hybridization. *Sci. Rep.* **5**, 9217 (2015).

43. Squeglia, F., Bachert, B., De Simone, A., Lukomski, S. & Berisio, R. The crystal structure of the streptococcal collagen-like protein 2 globular domain from invasive M3-type group A Streptococcus shows significant similarity to immunomodulatory HIV protein gp41. *J. Biol. Chem.* **289**, 5122–5133 (2014).

44. Yu, Z., Brodsky, B. & Inouye, M. Dissecting a bacterial collagen domain from Streptococcus pyogenes: sequence and length-dependent variations in triple helix stability and folding. *J. Biol. Chem.* **286**, 18960–8 (2011).

45. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2006).

46. Arslan, S., Khafizov, R., Thomas, C. D., Chemla, Y. R. & Ha, T. Engineering of a superhelicase through conformational control. *Science* **348**, 344–347 (2015).

47. Esmenjaud, J.-B. *et al.* An inter-dimer allosteric switch controls NMDA receptor activity. *EMBO J.* **38**, e99894 (2019).

48. Shi, Y. *et al.* cAMP-dependent protein kinase phosphorylation produces interdomain movement in SUR2B leading to activation of the vascular KATP channel. *J. Biol. Chem.* **283**, 7523–7530 (2008).

49. Siadat, S. M., Silverman, A. A., DiMarzio, C. A. & Ruberti, J. W. Measuring collagen fibril diameter with differential interference contrast microscopy. *J. Struct. Biol.* **213**, 107697 (2021).

50. Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G. & Galas, D. J. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res.* **17**, 6545–51 (1989).

51. Eckhard, U., Huesgen, P. F., Brandstetter, H. & Overall, C. M. Proteomic protease specificity profiling of clostridial collagenases reveals their intrinsic nature as dedicated degraders of collagen. *J. Proteomics* (2014) doi:10.1016/j.jprot.2013.10.004.

52. Sippel, T. O. Microfluorometric analysis of protein thiol groups with a coumarinylphenylmaleimide. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* **29**, 1377–1381 (1981).

53. Schönauer, E. *et al.* Discovery of a Potent Inhibitor Class with High Selectivity toward Clostridial Collagenases. *J. Am. Chem. Soc.* **139**, 12696–12703 (2017).

54. Liu, Y. *et al.* Use of a Fluorescence Plate Reader for Measuring Kinetic Parameters with Inner Filter Effect Correction. *Anal. Biochem.* **267**, 331–335 (1999).

55. Santra, M. & Luthra-Guptasarma, M. Assaying Collagenase Activity by Specific Labeling of Freshly Generated N-Termini with Fluorescamine at Mildly Acidic pH. *Int. J. Pept. Res. Ther.* **26**, 775–781 (2020).