Analyses of lysine aldehyde cross-linking in collagen reveal that the mature cross-link histidinohydroxylysinonorleucine is an artifact

David R Eyre, MaryAnn Weis and Jyoti Rai

From the Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, WA, 98195, USA

*Running title: Collagen cross-linking revisited

To whom correspondence should be addressed: David R. Eyre, Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, WA, USA, Telephone: 206-543-4700, Fax: 206-685-4700, E-mail: deyre@uw.edu

Key words: extracellular matrix, collagen, protein cross-linking, lysyl oxidase, allysine aldol, mass spectrometry (MS), skin, cornea, tendon, connective tissue

ABSTRACT

Lysyl oxidase-generated intermolecular cross-links are essential for the tensile strength of collagen fibrils. Two cross-linking pathways can be defined, one based on telopeptide lysine aldehydes and another on telopeptide hydroxylysine aldehydes. Since the 1970s it has been accepted that the mature cross-linking structures on the lysine aldehyde pathway, which dominates in skin and cornea, incorporate histidine residues. Here, using a range of MS-based methods, we reexamined this conclusion and found that telopeptide aldol dimerization is the primary mechanism for stable cross-link formation. The C-telopeptide aldol dimers formed labile addition products with glucosylgalactosyl hydroxylysine at α1(I)Lys-87 in adjacent collagen molecules that resisted borohydride reduction and after acid hydrolysis produced histidinohydroxylysinonorleucine (HHL), but only from species with a histidine in their α1(I) C-telopeptide sequence. Peptide MS analyses and the lack of HHL formation in rat and mouse skin, species that lack an α1(I) C-telopeptide histidine, revealed that HHL is a laboratory artifact rather than a natural cross-linking structure. Our experimental results also establish that histidinohydroxymerodesmosine is produced by borohydride reduction of N-telopeptide allysine aldol dimers in aldime intermolecular linkage to non-glycosylated α1(I)Lys-930. Borohydride reduction of the aldime promotes an accompanying base-catalyzed Michael addition of α1(I) His932 imidazole to the α,β-unsaturated aldol. These aldehydes are intramolecular at the N-terminus but at the C-terminus they can be both intramolecular and intermolecular according to present and earlier findings.

Collagen is the most abundant vertebrate protein. The highly evolved collagen family includes over 40 different gene products and many additional splice variants with diverse tissue-dependent roles (1,2). Fibril-forming collagens make up the bulk of an organism’s collagen mass and are responsible for the underlying strength of most tissues throughout metazoan evolution. The strength of individual collagen fibrils rests heavily on covalent cross-links being formed between individual molecules by the action of lysyl oxidase during fibril assembly and growth (3,4). The content, placement and chemistry of the resulting cross-linking amino acids is highly regulated and variable, dependent on the tissue type and in particular on the quality of the post-translational modifications of the newly synthesized collagen.
molecules (5,6). Covalent cross-links can vary among tissues even across a single genetic type of collagen such as collagen type I, the most abundant molecular type in vertebrates (7,8).

In essence, two pathways of lysyl oxidase-mediated cross-linking operate in fibril-forming collagens, one based on telopeptide lysine aldehydes and the other on telopeptide hydroxylsine aldehydes (3). The latter features in bone, cartilages and many other skeletal and tough connective tissues while the former is characteristic of skin, cornea and certain tendons notably rodent tail tendons. The aldehyde side-chains created by lysyl oxidase initially interact with specific helical domain hydroxylysines on neighboring collagen molecules to form divalent cross-links. These undergo further reactions to produce mature, chemically more stable cross-linking structures. On purely a hydroxylysine-aldehyde pathway these are the trivalent pyridinolines, hydroxylysyl pyridinoline (HP) and lysyl pyridinoline (LP). In tissues featuring purely a lysine-aldehyde pathway, the only known stable natural maturation product is histidinohydroxylysinonorleucine (HHL), which is prominent on analysis of mature human and bovine skin (9-12), but not rat or mouse skin or tail tendon type I collagen (11-14). Originally this cross-linking structure had been misidentified as hydroxyaldol-histidine (15), which later studies corrected (11-14).

Bone type I collagen in contrast has a unique cross-linking pattern which depends on the reactions between a mix of both telopeptide lysine and telopeptide hydroxylsine aldehydes and their helical domain partners. The mature structures in bone include a mix of trivalent pyridinolines, hydroxylsyl pyridinoline (HP) and lysyl pyridinoline (LP). This chemistry is largely controlled during collagen synthesis in the endoplasmic reticulum of osteoblasts by chaperone protein complexes that regulate the activity of lysyl hydroxylase isozymes against telopeptide and triple-helical domain substrate sites (5,8,18). The well recognized initial reaction products on the lysine aldehyde pathway after lysyl oxidase has converted telopeptide lysine aldehydes to aldehyde side-chains are intramolecular aldol dimers formed between α1 and α2 and α1 and α1 N-telopeptides within the individual type I collagen molecules (19). These aldols and remaining telopeptide lysine aldehyde monomers form aldime cross-links intermolecularly with adjacent helical domain hydroxylsine side-chains, which on borohydride reduction of tissue produce tetravalent histidinohydroxymerodesmosine (HHMD) and the reduced divalent aldime cross-link hydroxylsinosinorleucine (HLNL) (20,21). The reactions progress spontaneously in the fibril in vivo, catalyzed by their local protein environment within the highly ordered collagen polymer. The original methods by which the lysyl oxidase catalyzed cross-linking products were identified require borohydride reduction to stabilize labile aldamines as secondary amines. After acid hydrolysis the novel cross-linking amino acids were resolved by column chromatography and their structures determined by mass spectrometry, NMR and other analytical methods (22-29). The essential elements of lysyl oxidase mediated cross-linking in collagen and elastin were thus established in the late 1960s through the 1970s (30). There remained some controversy as to the source and origin of HHMD and the structurally related aldol histidine (31) on borohydride reduction, whether the histidine was added as an artifact of base-catalyzed Michael addition (21-28) or not (29). The controversy was never clearly resolved by methods available at the time.

Peptide mass spectrometry is a powerful tool for revealing site-specific collagen post-translational modifications including collagen cross-link analysis (32). It helped us discover effects on cross-linking in bone collagen in various human genetic variants and mouse models of osteogenesis imperfecta (OI) and in skin and tendon collagens of mice with null genes that affect collagen cross-linking (8,17,32). From these studies we discovered that the allysine aldol dimer, the first cross-link to be identified in collagen (19), which is responsible for the prominent β-chain dimers seen on SDS-PAGE of type I collagen extracts from skin and other soft tissues, was also formed between α1(I) C-telopeptide lysine aldehydes, both intramolecularly and intermolecularly (34).

None of our cross-linking findings on skin, however, were consistent with the proposed origin of the stable maturation product, HHL, predicted to be an addition product between the α1(I) C-telopeptide-to-K87helix aldime cross-
link, dehydro-HLNL, and a third chain histidine residue at α2(1) H92 (10). In order to understand better the nature of the mature cross-links on the lysine aldehyde pathway we applied mass spectrometry and different proteolytic cleavage and detection methods to examine the structure of cross-linked peptides containing C-telopeptide and N-telopeptide cross-linking sites from various skin, cornea and tendon collagens which are known to be particularly rich in or to lack HHL cross-links (35).

The findings show that allysine aldol dimers form the major stable cross-linking bonds at both ends of the type I collagen molecule in tissues that use the lysine aldehyde pathway. Together with literature data they provide evidence for a mechanism by which HHL is produced on acid hydrolysis artificially from a C-telopeptide allysine aldol in labile linkage to glegalHyl at α1(I) K87, but only in those species that have a conserved histidine in their α1(I)-C-telopeptide sequence.

Results

Comparison of cross-linked α-chain SDS-PAGE patterns across tissue collagen extracts

Some quantitative differences in extractability of type I collagen are evident between tissues that depend on lysine aldehyde cross-linking (Fig 1), but the patterns of α, β, γ and higher oligomers are very similar. In contrast, an extract of demineralised bone was run as a contrasting type I collagen-based matrix that is cross-linked by a mix of lysine aldehydes and hydroxylsine aldehydes (17). The results show a consistent pattern of β-chain dimers being more prominent than α-chains in extracts of skin, cornea and tail tendon collagens compared with bone collagen. The skin, cornea and tendon extracts also reveal higher order oligomers of which γ112 is the most prominent in non-denaturing, acetic acid extracts. The latter γ112 trimer was previously shown to be derived from extracted native type I collagen molecules that contain an allysine aldol between the α1(I)N- and the α2(I)N-telopeptide and between the two α1(I)C- telopeptides. Under the denaturing conditions of SDS-PAGE this runs as the γ112 band.

Tissue-dependent hydroxylation and glycosylation differences between helical domain cross-linking lysines.

In-gel trypsin and Lys-C digestion were also used to measure the post-translational status of the known triple-helical domain cross-linking sites in α1(I) and α2(I) chains from the various tissues. The results are summarized in Fig 2. Type I collagen of skin and cornea differs from that of tail tendon in having K87 fully hydroxylated and glycosylated as glegal Hyl, whereas in tail tendon K87 is fully hydroxylated but is not glycosylated. Bone collagen has primarily a mix of glegal Hyl and galHyl at K87, the ratio of which varies between species (17).

Western blot analysis of CNBr-digested tissue collagens for C-telopeptide cross-linked structures

The β-chain dimers, β112 (Fig 1) were the first established products of lysyl oxidase cross-linking in collagen, the result of allysine aldol intramolecular dimerization of N-telopeptides (19). Only recently were C-telopeptide aldol dimers discovered, initially from studying the tail tendon collagen of fibromodulin-null mice (34) but subsequently in the skin and tendons of other mutant and wild-type mice (8,33). These studies showed evidence for both intramolecular and intermolecular C-telopeptide aldol cross-link formation. We therefore designed an analytical screen to profile all forms of C-telopeptide-containing structures in CNBr-digests of tissue collagen. To do this a monoclonal antibody, mAb 1G7, that recognizes an epitope specific to the α1(I)C-telopeptide was applied to SDS-PAGE resolved tissue digests. The results are shown in Fig 3.

Although immunoreactive yields vary somewhat between individual tissue sources as seen from the lane intensities, (which may depend on efficiency of C-telopeptide epitope generation by trypsin), all tissues show similar patterns except bone. The fastest positive band in each lane from skin, cornea and tendon is the linear α1CB6 sequence which ends with the C-telopeptide. Looking at the rat skin lane as an example, the next slower band is an aldol dimer of CB6 and above that is the linear CNBr partial cleavage product, α1CB7,6. The same bands repeat in the
lanes for rat tail tendon, bovine cornea and skin. In the latter two lanes, with stronger overall band intensities, the higher band is an α1CB6 x α1CB7,6 dimeric partial cleavage product. No other significant cross-linked forms of CB6 are revealed. For example, no band that would be consistent with the proposed molecular origin of HHL as a mature cross-link linking the three CB peptides, α1CB6 x α1CB5 x α2CB4 (10) was recovered from skin or cornea.

The lane from rat bone collagen gives a different band pattern that is nevertheless consistent with the known cross-linking mechanism of C-telopeptides in bone type I collagen. Stable ketoimine divalent cross-links are responsible for the α1CB6 x CB5 band and pyridinoline plus pyrrole trivalent cross-links for the α1(CB6 x CB6 x CB5) band running above the α1CB6 dimer and α1CB7,6 bands from skin, cornea and tendon.

Mass spectral identification of CB-peptides
To confirm the CB-peptide band identities, histidine-containing peptides were enriched from tissue CNBr digests by an immobilized metal ion, affinity column then fractionated on a C8 reverse-phase column (Fig 4). As shown, collected fractions were run serially across lanes of an SDS-PAGE gel. Individual bands were excised and subjected to in-gel trypsin digestion and tandem LCMS analysis, which established their indicated CB-peptide identities. These results also helped confirm the identities of the bands revealed by Western blot analysis in Fig 3.

Prominence of allysine aldol cross-linked N- and C-telopeptide dimers in bacterial collagenase-digested cornea collagen
To dissect the cross-linking sites more finely down to smaller peptides using mass spectrometric analysis, the same tissues were digested with bacterial collagenase under conditions that generated reproducible and informative cross-linked peptide structures (8,16). The digests were fractionated by reverse-phase HPLC and individual fractions analysed by LCMS and their structures determined manually from the MSMS fragmentation patterns (8,16). In Fig 5, the elution positions and the structures of the major N-telopeptide and C-telopeptide cross-linked dimers are identified.

Effect of borohydride reduction on telopeptide aldol dimeric peptide yields
Reduction of tissue with sodium borohydride failed to change the yield of C-telopeptide aldol dimeric peptides but N-telopeptide dimers were no longer found. Instead the same N-telopeptide dimeric peptides were recovered in covalent linkage to hydroxylysine at K930 in a peptide fragment from the α1(I) chain.

The masses and fragmentation profile of these three chained crosslinked structures (Fig 6) are exactly consistent with reduction of the aldimine bond formed between the α,β unsaturated allysine aldol aldehyde and the ε-amino group of K930 hydroxylysine. Apparently on tissue reduction, the N-telopeptide aldol adducted to the helical domain hydroxylysine is susceptible to borohydride reduction but the C-telopeptide aldol adduct is not.

Identification of an α1(I) C-telopeptide dimer in labile linkage to glycosylated K87 in bacterial collagenase-digested cornea collagen
In further analyses of borohydride-reduced and non-reduced cornea after bacterial collagenase digestion a large peptide was recovered spread across several fractions of the C8 HPLC chromatogram. The peptide was recovered with the same characteristic properties whether tissue had been reduced or not (Fig 7). The structure was determined manually based on fragmentation properties of the parent ions and of MS3 of the helical arm throw off Fig (7C,D). The parent ion was unusual in our experience with collagen cross-linked peptides in that the intact helical peptide arm (ions 1082.61+ and 541.82+) containing K87 hydroxylysine with sugars (glucosylgalactosyl) attached was released on MSMS as a major, intact fragment ion. This implied a labile covalent bond between the allysine aldol cross-linking residue of the C-telopeptide dimer and the glycosylated K87 hydroxylysine. The 10494+ prominent ion (Fig 7B) matches the mass of the C-telopeptide dimer being released as the α,β unsaturated allysine aldol peptide dimer.
LC/MS analysis of HHL, HHMD and total cross-linking amino acids after acid hydrolysis

The LC total ion-current profiles of the cross-linking amino acid products of acid hydrolysis of borohydride-reduced bovine tendon, skin and cornea, and mouse skin are shown in Fig 8A, together with mass spectra from scrolled regions of these profiles. The results confirm that bovine cornea and skin are rich sources of HHL and HHMD, whereas mouse skin yields only HHMD but no HHL. Bovine flexor tendon reveals products of cross-linking from both the lysine aldehyde and hydroxylysine aldehyde pathways as expected, with hydroxylysyl pyridinoline (HP), HHMD, HHL and the reduced divalent cross-links HLNL and LNL all present. From skin, desmosine (Des, from elastin) is also detected. Samples of the same tissues hydrolysed without prior borohydride reduction gave HHL but no HHMD, HLNL or LNL as expected.

Discussion

The SDS-PAGE patterns of extracted collagen chains from skin, cornea and tendon essentially are similar, with more prominent β-chain dimers than α-chains (Fig 1). The cross-linked trimer, γ112, is also consistently present. We have shown that this is the product of C-telopeptide aldol dimer formation in mouse tail tendon and skin collagen, and is especially prominent in fibromodulin-null mice due to increased telopeptide lysine oxidation by lysyl oxidase (34). Similar, abnormally high levels of γ112 and C-telopeptide aldol dimeric cross-links in skin collagen were also observed in mice with null genes encoding Sc65 or P3h3, two members of the prolyl 3-hydroxylase gene family (8,33). The latter two mouse models phenocopy a biochemical abnormality in the post-translational modification of type I collagen that results in underhydroxylation of triple-helical domain lysines at the cross-linking sites α1(I) and α2(I) K87 and α1(I) K930 and α2(I) K933 in all tissues but especially skin (8). This results in mostly lysine and little glocgalHyl at α1(I) K87 in skin collagen. The increase in γ112 and intramolecular C-telopeptide aldol dimers in skin collagen was attributed to the lack of sugar residues at K87 allowing intramolecular aldols to form when both α1(I) C-telopeptides in the same molecule had their lysines oxidised to aldehydes by lysyl oxidase. In wild type skin, intermolecular aldols at the C-terminus appeared to be the norm based on the γ and higher cross-linked chain patterns (8). For the present study, this implies that the state of glycosylation at K87 may sterically determine the reaction properties of C-telopeptide allysines in the growing fibril during spontaneous cross-link formation.

Analysis of the state of hydroxylation of lysine and glycosylation of hydroxylysine at the triple-helical domain cross-linking sites established that in skin and cornea collagens K87 was fully glycosylated, but in tail tendon type I collagen K87 was exclusively Hyl with no attached sugars (Fig 2). This lack of glycosylation of tendon type I collagen has been noted before (35). In order to determine if this had any tissue-dependent, selective effect on C-telopeptide cross-linking interactions, not evident by SDS-PAGE analysis of extracted collagen chains, we next looked at CNBr digests of whole tissue, matrix collagen. The mAb 1G7, which recognizes the C-terminus of the C-telopeptide sequence, EXAHDGGR, where K can be Lys, Hyl or a cross-link to one or more additional peptides, provides a useful screen on a Western blot for all linear and cross-linked structures that have this terminal peptide sequence. Whole tissue was treated with trypsin under non-denaturing conditions to enhance the yield of the C-terminal neoepitope that 1G7 recognizes. All tissues, except a bone control, were selected for comparison based on their collagen being cross-linked exclusively by the lysine aldehyde pathway. Bovine skin and particularly cornea were known to be rich in HHL (36), whereas this cross-link was reported absent from rat and mouse skin and tail tendon collagens (10, 12-14). The results in Fig 3 showing essentially the same patterns of linear and cross-linked forms of CB6 from skin, cornea and tendons is not consistent with the proposed intermolecular sites linked by HHL residues. The origin of HHL as a trivalent cross-link of three chains, an α1(I) C-telopeptide in aldime linkage α1(I) Hyl87 adducted to α2(I) His92 was proposed based on N-terminal Edman sequencing analyses of a partially purified tryptic peptide prepared from bovine dermis (10). No band fitting the size of such a combination of CB-peptide fragments is evident in the 1G7 Western
blot from bovine skin or cornea, the latter tissue collagen in our hands being the richest source of HHL at about 1 mole/mole of collagen (36). The only cross-linked form of CB6 on SDS-PAGE from all tissues except bone was a CB6 dimer and a larger CB7,6 partial cleavage form of it. (Fig 3). From a bone CB-digest run as a control, the main cross-linked forms of CB6 are the divalent keto-imine cross-linked peptide, CB6 x α1(I)CB5 and trivalent pyridinoline/pyrrole cross-linked CB6 x CB6 x α1(I)CB5. The latter serves as a useful size marker to compare with the smaller, faster running CB6 aldol dimer from skin, cornea and tendon CNBr digests.

To further refine the identity of the native cross-linked structures, smaller peptides were prepared by bacterial collagenase digestion. On mass spectrometry these gave prominent yields of the same C-telopeptide and N-telopeptide aldol dimeric peptides from all skin, cornea and tendon digests. (Fig 5 shows the results for bovine cornea), whether or not the tissue had been reduced with sodium borohydride. In addition, the bovine cornea digest revealed a larger peptide consisting of the CB6 dimer in labile linkage to α1(I) HyI87 contained in a short segment of helix (Fig 7). This adduct was recovered consistently from repeat digests of cornea and also from skin but not rat or mouse tail tendon (data not shown), but in relatively low and variable yields compared with the free C-telopeptide aldol dimeric peptide (Fig. 5). The mass spectral fragmentation behavior of this larger peptide with a strong throw off of the intact GMK(galglc)GHRG K87 arm is consistent with an interpretation that this was an adduct in labile linkage to the allysine aldol aldehyde group of the C-telopeptide dimer. Variable yields of this peptide together with the more abundant free C-telopeptide dimer were recovered from digests whether or not the tissue was previously treated with sodium borohydride.

This was in sharp contrast to the behavior of the N-telopeptide dimers which were absent from digests of borohydride reduced tissue and quantitatively converted to the reduced aldime trimeric peptides identified in Fig 6. Whether the tissue was reduced or not, the labile C-telopeptide trimer behaved the same on mass spectrometry, with the characteristic throw off of the intact helical arm shown in Fig 7. Again in sharp contrast, the MSMS profile of the borohydride-reduced N-telopeptide trimeric peptides did not yield any intact helical arm fragmentation products (Fig 6). The low yields of the C-telopeptide aldol/K87 adduct on bacterial collagenase digestion, and of HHL on acid hydrolysis of bacterial collagenase-digested cornea (Fig 8), are consistent with the C-telopeptide aldol forming a labile adduct with glegalHyl at K87 in vivo which can break on proteolysis but produces HHL on acid hydrolysis. The mAb 1G7 Western blot results are consistent with this conclusion (Fig 3).

The present findings can explain why skin collagen from certain species yield HHL on acid hydrolysis but rat and mouse skin do not. Specifically, bovine, human and most vertebrates contain the α1(I) C-telopeptide sequence, EKAHDGGR (NCBI accession # P02453.3), with histidine (H18c) two residues C-terminal to the cross-linking lysine, whereas the rat and mouse sequence, EKSQDGGR (NCBI accession #P02454.5 and #P11087.4), does not (fig 9). Skin from other species (sheep, goat, deer as well as cow) all yield HHL on hydrolysis (37) and have the EKAHDGGR sequence conserved. From the present structural results, it is clear that it is this telopeptide histidine imidazole that is incorporated into HHL on acid hydrolysis of native fibrils. If so this raises the question, what happens to the C-telopeptide aldol dimers in rat and mouse skin on acid hydrolysis? Without the neighboring telopeptide histidine, these native aldol complexes are probably broken down on acid hydrolysis as early studies of allysine aldol complexes in collagen concluded (38). The intramolecular aldol dimers formed between N-telopeptides on the other hand clearly behave differently from the C-telopeptide aldol dimers in their chemical reaction properties in native fibrils, being quantitatively and stably reduced as aldime adducts to K930 of an α1(I) chain by sodium borohydride (Fig 6 ). Though we did not recover equivalent reduced aldol dimer peptide adducts to α2(I) K933 they are possible and we cannot rule them out.

Proposed origin and mechanism of formation of the histidine-containing adducts HHL and HHMD

The above properties support the proposed mechanisms outlined in Fig 10 whereby the C-telopeptide aldol dimer resists borohydride...
reduction as an \( \alpha_1(I) \) K87-linked enamine tautomer, whereas N-telopeptide aldol aldmine adducts are sodium borohydride-reduced producing HHMD which is released on acid hydrolysis. The speculated retro-aldol mechanism for generating HHL from the enamine tautomer on acid hydrolysis is proposed based on current knowledge of the base catalyzed mechanisms that stabilize Schiff’s base to enamine intermediates at the active sites of a range of aldolase enzymes (39-42), together with the following unique sequence features of the C-telopeptide to helix cross-linking site.

All vertebrate \( \alpha_1(I) \) C-telopeptides including rat and mouse have a conserved arginine 6 residues C-terminal (R22C) to the cross-linking lysine (see fig 9). Also in skin and cornea, \( \alpha_1(I) \) K87 is almost all glucosylgalactosyl hydroxylysine, whereas \( \alpha_1(I) \) K930 and \( \alpha_2(I) \) K933 are not glycosylated. In contrast to skin and cornea, neither K87 nor K930/933 are glycosylated in rodent tail tendon type I collagen. We believe these chemical features are responsible for the distinctive tissue dependent reactivity properties of C-telopeptide aldehydes, and their ensuing aldol dimers, compared with N-telopeptides. Specifically, we propose that R22C which will be in close proximity to the aldehyde in the hairpin loop conformation of each C-telopeptide (43), base-catalyzes both the aldol formation and the subsequent enamine tautomer structure as the dominant tautomeric form of the C-telopeptide aldol reaction product with K87. When in addition, K87 is glucalHyl, as it is in skin and cornea, the product is HHL on acid hydrolysis. The hexoses may provide the latent, local reducing agent required by the proposed reaction mechanism. In support of this mechanism, it has been established that short synthetic peptides containing the \( \alpha_1 \beta \) unsaturated aldehyde cinnamaldehyde inhibit certain classes of phosphatases by forming a reversible covalent enamine adduct with the guanido group of an arginine residue in their SH2 domain active site (44,45). As enamines, these adducts resisted borohydride reduction. We propose that R22C in \( \alpha_1(I) \) C-telopeptides can similarly interact with the aldol formed between two K16C allysines and, in native fibrils, influence the equilibrium of the addition reactions between K16C allysines and K87 helix hydroxylysines, favoring the enamine tautomer shown in Fig 10.

**Significance**

Our findings show that the long accepted collagen cross-linking pathway that involves histidine addition to produce the mature crosslinks in skin, cornea and other tissues is incorrect. In considering how tissues that use this pathway grow and remodel their fibrillar extracellular matrix the new concept has important implications.

We conclude that aldol dimers are the main, stable cross-linking bonds formed by C-telopeptides and N-telopeptides in skin, cornea and certain tendon type I collagens. Their further reaction products form relatively labile intermolecular cross-links. If the aldol-forming bonds are all intramolecular at the N-terminus and mostly intramolecular at the C-terminus, as is the case with rodent tail tendon, this can explain why rodent tail tendon collagen is highly soluble as native monomers in dilute acetic acid (8, 34) since the intermolecular cross-links will be mostly acid-labile aldimes and enamines (Fig 10). In normal skin collagen, most of the C-terminal aldol bonds appear to be intermolecular (8), which can explain the lower acid solubility of skin collagen versus tail tendon (8).

The findings are also significant when considering how collagenous tissues that depend on the lysine-aldehyde cross-linking pathway for strength, can build and remodel their extracellular fabric (46). Growing tendons, for example, may benefit from having labile intermolecular cross-links that allow collagen fibril remodelling to occur through bond breakage and reformation. In this way, new collagen molecules could be added to increase the diameter of existing fibrils, and fibrils could fuse laterally while the tissue fabric as a whole could continue its mechanical function as it grows with the skeleton and remodels as a mature tissue.

Cornea is a unique tissue that depends on the lysine aldehyde pathway for the natural strength of its highly organized type I collagen fibrillar array (47). The new insight that aldol dimers are the principle source of stable cross-linking bonds in cornea is potentially significant with respect to the recognized clinical benefits from artificially cross-linking keratoconus corneas with riboflavin then UV-A light (48). The
underlying molecular mechanism by which this treatment strengthens corneal tissue and prevents bulging is unclear, but there is experimental evidence that endogenous carbonyl groups are the reactive components that form new cross-links (49). Since cornea type I collagen contains roughly two moles of allysine aldol per mole, this is a strong candidates for being the proposed reactive species when corneas are treated with riboflavin and UV-A light.

**Experimental procedures**

**Tissue sources and preparation**

Adult bovine tissues (skin, eyes and superficial digital flexor tendon) were purchased from Sierra for Medical Science (Whittier, CA). Adult normal Sprague-Dawley laboratory rats and adult normal (C57B6/j) mice were obtained as byproducts from approved and completed animal studies. Tissues were scraped clean and washed in saline (0.15 M NaCl, 0.05 M Tris) pH 7.4 with protease inhibitors (2 mM EDTA, 5 mM Benzamidine, 10 mM 1,10-phenanthroline and 2 mM phenylmethylsulfonyl fluoride) for 24 h at 4°C. Tendon and skin were defatted with chloroform/methanol (1:1 v/v) for 24 h at room temperature. Cornea was extracted in 4 M guanidine HCl, 0.05M Tris pH 7.4 with protease inhibitors overnight at 4°C to remove proteoglycans before collagen extraction. Bone was demineralized in 0.1M HCl, 24 h at 4°C.

**Collagen extraction and proteolysis**

Type I collagen was solubilized from 2-3 mg samples of prepared bone, tendon, skin and cornea by heat denaturation for 3 min at 90°C in Laemmli buffer (denatured extract), or 3% acetic acid at 4 °C for 24 h, or cyanogen bromide digestion in 70% formic acid for 24 h at room temperature. Cornea was extracted in 4 M guanidine HCl, 0.05M Tris pH 7.4 with protease inhibitors overnight at 4°C to remove proteoglycans before collagen extraction. Bone was demineralized in 0.1M HCl, 24 h at 4°C.

**Cornea CNBr peptide chromatography**

An IMAC (immobilized metal ion affinity chromatography) column was prepared by charging HiTrap chelating HP beads (Amersham Biosciences) with CuCl2. CNBr digested cornea was loaded onto the IMAC column in 0.05 M Tris-HCl, 0.15 M NaCl, 2 M guanidine HCl, pH 8.0, and the bound peptides were eluted with 0.1 M sodium acetate, pH 4.6, containing 2 M guanidine HCl and 0.2 M imidazole. Eluted CNBr peptides were run on reversed-phase HPLC and individual fractions run on 12.5% SDS-PAGE.

**SDS-PAGE**

The method of Laemmli (50) was used with 5 or 12.5% gels for extracts of tissue collagen. Type I α-chains and Cyanogen Bromide peptides were cut from gels and digested with trypsin or Endo Lys-C (Mass Spec Grade, Promega, Madison, WI, USA) in-gel, for mass spectral analyses.

**Immunoblotting with mAb 1G7**

A mouse monoclonal antibody (1G7 isotype IgG1,k) was generated against the synthetic peptide EKAHDGGR that recognizes a C-terminal neo-epitope of the α1(I) collagen C-telopeptide. This can be generated by trypsin digestion of native collagen. Peptide conjugation, immunization and hybridoma selection were done as in (51) using RBF/DnJ mice. The synthetic peptide EKAHDGGR was used for screening and characterization. Using variable length C-telopeptide synthetic peptides, we determined that the C-terminal arginine is necessary for 1G7 antibody recognition. Tissue samples were activated with 0.1mg/mL trypsin (Trypsin, Proteomics grade, Roche Diagnostics) in 0.1 M NH4HCO3 for 24h then digested with CNBr in 70% formic acid both at room temperature for 24 h. CNBr peptides where separated by 12.5% SDS-PAGE and transblotted to PVDF membrane. Membranes were blocked in 5% nonfat-milk, probed with 1G7 monoclonal antibody (1:5,000 dilution), and detected with an HRP-conjugated secondary antibody (1:10,000, anti-mouse IgG). Blots were developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Scientific).

**LCMS peptide analysis**

Electrospray LC-MS/MS was performed on tryptic and Lys-C peptides generated from gel bands and individual collagenase HPLC fractions using an LTQ XL ion-trap mass spectrometer (Thermo Scientific) equipped with in-line liquid chromatography using a C4 5um capillary column (300um x 150mm; Higgins Analytical RS-15M3-
Collagen Cross-linking Revisited

W045) eluted at 4.5ul min. The LC mobile phase consisted of buffer A (0.1% formic acid in MilliQ water) and buffer B (0.1% formic acid in 3:1 acetonitrile:n-propanol v/v). The LC sample stream was introduced into the mass spectrometer by electrospray ionization (ESI) with a spray voltage of 4kV. Proteome Discoverer search software (Thermo Scientific) was used for peptide identification using the NCBI protein database. Proline and lysine modifications were examined manually by scrolling or averaging the full scan over several minutes so that all of the post-translational variations of a given peptide appeared together in the full scan. Cross-linked peptides were identified manually by calculating theoretical parent ion masses and possible ms/ms ions and matching these to the actual parent mass and ms/ms spectrum.

**LCMS of cross-linking amino acids**

Samples of freeze-dried tissues both untreated and reduced with sodium borohydride (36), bacterial digested tissue and subsequent chromatographic peptide pools were hydrolyzed in 6N HCl at 108 C for 24 hours. Cross-linking amino acids were enriched from hydrolysates by partition on a hydrated cellulose column (37). Briefly, a column was prepared by pouring a 5% cellulose (CF1, Whatman) slurry in butanol-water-acetic acid (4:1:1 v/v/v) into a 10 mL syringe plugged at the bottom with glass wool. The column was washed three times with butanol-water-acetic acid (4:1:1 v/v/v) before the sample dissolved in a 5% cellulose/ butanol-water-acetic acid slurry was applied. Free amino acids were removed by washing the column with butanol-water-acetic acid solution. The multi-charged bound cross-linking amino acids were eluted with water and freeze-dried for mass spectral analysis. Electrospray LC-MS/MS was performed on the LTQ XL using a Cogent 4 diamond hydride column (15cm x 1mm; Microsolv technology, 70000-15P-1) eluted at 50ul min. The LC mobile phase consisted of buffer A (0.1% formic acid in MilliQ water) and buffer B (0.1% formic acid in 80% acetonitrile).
Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

This work was supported in whole or in part by grants from the National Institutes of Health (NIAMS Awards AR037318, AR036794 (DE) and NICHD Award HD070394 (DE).

Abbreviations
HHL, Histidinohydroxylysinonorleucine; HP, hydroxylysyl pyridinoline; LP, lysyl pyridinoline; HHMD, histidinohydroxymerodesmosine; HLNL, hydroxylysinonorleucine; LNL, lysinonorleucine
References

1. Ricard-Blum, S. (2011) The collagen family. Cold Spring Harb Perspect Biol. 3, 1-19
2. Bella, J., Hulmes DJ. Fibrillar Collagens. (2017) Subcell Biochem. 82,457-490.
3. Eyre, D. R., Paz, M. A., and Gallop, P. M. (1984) Cross-linking in collagen and elastin. Annu.Rev.Biochem. 53, 717–748
4. Rodriguez-Pascual, F., Slatter D.A. (2016) Collagen cross-linking: insights on the evolution of metazoan extracellular matrix. Sci Rep. 6:37374. doi:10.1038/srep37374.
5. Gjaltema, R.A., Bank, R.A. (2017) Molecular insights into prolyl and lysyl hydroxylation of fibrillar collagens in health and disease. Crit Rev Biochem Mol Biol. 52, 74-95.
6. Yamauchi, M., and Sricholpech, M. (2012) Lysine post-translational modifications of collagen. Essays In Biochemistry. 52, 113–133
7. Hudson, D.M. and Eyre, D.R. (2013) Collagen prolyl 3-hydroxylation: a major role for a minor post-translational modification? Connect. Tissue Res. 54, 245–251
8. Hudson, D.M., Weis, M., Rai, J., Joeng, K.S., Dimori, M., Lee, B.H., Morello, R, Eyre, D.R. (2017) 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.
9. Yamauchi, M., London, R.E., Guenat, C., Hashimoto, F., Mechanic, G.L. (1987) Structure and formation of a stable histidine-based trifunctional cross-link in skin collagen. J Biol Chem. 262, 11428-34.
10. Mechanic, G.L., Katz, E.P., Henmi, M., Noyes, C., Yamachi, M. (1987) Locus of a histidine-based, stable trifunctional, helix to helix collagen cross-link: stereospecific collagen structure of type I skin fibrils. Biochemistry. 26, 3500-9.
11. Yamauchi, M., Woodley, D.T., Mechanic, G.L. (1988) Aging and cross-linking of skin collagen. Biochem Biophys Res Commun. 152, 898-903.
12. Woodley, D.T., Yamauchi, M., Wynn, K.C., Mechanic, G., Briggaman, R.A. (1991) Collagen telopeptides (cross-linking sites) play a role in collagen gel lattice contraction. J Invest Dermatol. 97: 580-5.
13. Reiser, K.M. Influence of age and long-term dietary restriction on enzymatically mediated crosslinks and nonenzymatic glycation of collagen in mice. (1994) J Gerontol. 149: B71-9.
14. Paul, R.G., Tarlton, J.F., Purslow, P.P., Sims, T.J., Watkins, P., Marshall, F., Ferguson, M.J. Bailey AJ. (1997) Biomechanical and biochemical study of a standardized wound healing model. Int J Biochem Cell Biol. 29: 211-20.
15. Housley, T., Tanzer, M.L., Henson, E., Gallop, P.M. (1975) Collagen crosslinking: isolation of hydroxyldol-histidine, a naturally-occurring crosslink. Biochem Biophys Res Commun. 67, 824-30.
16. Hanson, D.A., Eyre, D.R. (1996) Molecular site specificity of pyridinoline and pyrrole cross-links in type I collagen of human bone. J Biol Chem. 271: 26508-16.
17. Eyre, D.R., Weis, M.A. (2013) Bone collagen: new clues to its mineralization mechanism from recessive osteogenesis imperfecta. Calcif. Tissue Int. 93: 338-47.
18. Yamauchi M, Taga Y, Hattori S, Shiiba M, Terajima M. (2018) Analysis of collagen and elastin cross-links. Methods Cell Biol. 143, 115-132.
19. Bornstein P, Kang AH, Piez KA. (1966) The nature and location of intramolecular cross-links in collagen. Proc Natl Acad Sci USA. 55, 417-24.
20. Tanzer ML, Housley T, Berube L, Fairweather R, Franzblau C, Gallop PM. (1973) Structure of two histidine-containing crosslinks from collagen. J Biol Chem. 248, 393-402.
21. Robins SP, Bailey AJ. (1973) The chemistry of the collagen cross-links. The characterization of fraction C, a possible artifact produced during the reduction of collagen fibres with borohydride. *Biochem J.* 135, 657-65.

22. Rojkind M, Blumenfeld OO, Gallop PM. (1966) Localization and partial characterization of an aldehydic component in tropocollagen. *J Biol Chem.* 241:1530-6.

23. Bailey AJ, Peach CM. (1969) Identification of a labile intermolecular cross-link in collagen. *Biochem J.* 111:12P-13P.

24. Bailey AJ, Peach CM, Fowler LJ. (1970) Chemistry of the collagen cross-links. Isolation and characterization of two intermediate intermolecular cross-links in collagen. *Biochem J.* 117:819-31.

25. Kang AH, Gross J. Relationship between the intra and intermolecular cross-links of collagen. (1970) *Proc Natl Acad Sci USA.* 67:1307-14.

26. Tanzer ML. Cross-linking of collagen. (1973) *Science.* 180:561-6.

27. Bailey AJ, Sims TJ. Chemistry of the collagen cross-links. (1976) Nature of the cross-links in the polymorphic forms of dermal collagen during development. *Biochem J.* 153:211-5.

28. Bernstein PH, Mechanic GL. (1980) A natural histidine-based imminium cross-link in collagen and its location. *J Biol Chem.* 255:10414-22.

29. Sims TJ, Bailey AJ. (1992) Quantitative analysis of collagen and elastin cross-links using a single-column system. *J. Chromatogr.* 582: 49-55.

30. Fairweather RB, Tanzer ML, Gallop PM. (1972) Aldol-histidine, a new trifunctional collagen cross-link. *Biochem Biophys Res Commun.* 48, 1311-5.

31. Eyre DR, Weis MA, Wu JJ. (2008) Advances in collagen cross-link analysis. *Methods* 45: 65-74.

32. Heard ME, Besio R, Weis M, Rai J, Hudson DM, Dimori M, Zimmerman SM, Kamykowski JA, Hogue WR, Swain FL, Burdine MS, Mackintosh SG, Tackett AJ, Suva LJ, Eyre DR, Morello R (2016) Sc65-Null Mice Provide Evidence for a Novel Endoplasmic Reticulum Complex Regulating Collagen Lysyl Hydroxylation. *PloS Genet.* 12 :e1006002.

33. Kalamajski S, Liu C, Tillgren V, Rubin K, Oldberg Ö, Rai J, Weis M, Eyre DR. (2014) Increased C-telopeptide cross-linking of tendon type I collagen in fibromodulin-deficient mice. *J Biol Chem.* 289:18873-9.

34. Henkel W, Rauterberg J and Stirtz T. (1976) Isolation of a cross-linked cyanogen bromide peptide from insoluble rabbit collagen. Tissue differences in glycosylation and hydroxylation of the cross-link. *Eur J Biochem.* 69, 223-31

35. Eyre D. Collagen cross-linking amino acids. Methods Enzymol. 1987;144:115-39.

36. Naffa R, Holmes G, Ahn M, Harding D, Norris G. (2016) Liquid chromatography-electrospray ionization mass spectrometry for the simultaneous quantitation of collagen and elastin crosslinks. *J Chromatogr A.* 23, 1478, 60-67.

37. Eyre DR, Glimcher MJ. (1973) Evidence for intramolecular crosslinks in chicken bone collagen: the isolation of peptides containing allysine aldol. *Biochim Biophys Acta.* 295: 301-7.

38. Heine A, DeSantis G, Luz JG, Mitchell M, Wong CH, Wilson IA. (2001) Observation of covalent intermediates in an enzyme mechanism at atomic resolution. *Science.* 294: 369-74. Erratum in Science (2001) 294:2096.

39. Giger L, Caner S, Obexer R, Kast P, Baker D, Ban N, Hilvert D. (2013) Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nat Chem Biol.* 209 :94-8.

40. Amar,A, Meghezzi,H, Boucekkine,A, Kaoua, R Kolli, B, (2010) How to drive imine-ename tautomerism of pyronic derivatives of biological interest - A theoretical and experimental study of substituent and solvent effects. *Comptes Rendus Chimie.* 13 :553-560.

41. Inani H, Jha AK, Easwar S. (2017) An arginine-mediated protocol for the aldol addition of methyl vinyl ketone in water. *ChemistrySelect.* 2: 11666-11672.
43. Malone JP, Veis A. (2004) Heterotrimeric type I collagen C-telopeptide conformation as docked to its helix receptor. *Biochemistry*. 43:15358-66.

44. Park J, Fu H, Pei D. (2003) Peptidyl aldehydes as reversible covalent inhibitors of Src homology 2 domains. *Biochemistry*. 42: 5159-5167.

45. Park J, Fu H, Pei D. (2004) Peptidyl aldehydes as slow-binding inhibitors of dual-specificity phosphatases. *Biorg. Med. Chem. Lett.* 14: 685-687.

46. Holmes DF, Lu Y, Starborg T and Kadler KE (2018) Collagen fibril assembly and function. *Curr. Top. Dev. Biol.* 130, 107-142.

47. Zhang G, Chen S, Goldoni S, Calder BW, Simpson HC, Owens RT, McQuillan DJ, Young MF, Iozzo RV, Birk DE. (2009) Genetic evidence for the coordinated regulation of collagen fibrillogenesis in the cornea by decorin and biglecan. *J Biol Chem.* 284: 8888-97.

48. Belin MW, Lim L, Rajpal RK, Hafezi F, Gomes JAP, Cochen B. (2018) Corneal cross-linking: current USA status report from The Cornea Society. *Cornea*. 37, 1218-1225.

49. McCall AS, Kraft S, Edelhauser HF, Kidder GW, Lundquist RR, Bradshaw HE, Dedeic Z, Dionne MJC, Clement EM, Conrad GW. (2010) Mechanisms of corneal tissue cross-linking in response to treatment with topical riboflavin and long-wavelength ultraviolet radiation (UVA). *Invest. Ophthalmo. Vis. Sci.* 51, 129-138.

50. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227, 680-685.

51. Hanson DA, Weis MA, Bollen AM, Maslan SL, Singer FR, Eyre DR. (1992) A specific immunoassay for monitoring human bone resorption: quantification of type I collagen cross-linking N-telopeptides in urine. *J Bone Miner Res.* 7, 1251-1258.
Figure 1. SDS-PAGE showing similar α, β, γ and oligomeric cross-linked chain profiles for collagen extracts from skin, cornea and tendon. Native collagen molecules extracted by 3% acetic acid were run in the left five lanes and heat denatured collagen from a separate tissue sample in the right five lanes. Loads based on the following original tissue weights extracted were: bovine cornea, 60ug; bovine skin, 30ug; rat skin, 30ug; tail tendon, 15ug; demineralized bone, 300ug.
Figure 2. Summary of tissue-dependent differences in hydroxylation and glycosylation of helical domain cross-linking in type I collagen. The four known lysine sites of helical domain cross-linking were interrogated for post-translational status by in-gel trypsin or endo Lys-C digestion followed by LC/MS. (A) Summary of the results for bovine cornea and skin, rat skin, tail tendon and bone. (B) Example of the mass spectra from cornea for the trypsin-digested α1(I) K87 peptide and (D) its fragmentation pattern. (C) Example of the mass spectrum from cornea of the endo Lys-C digested α1(I) K933 peptide and (E) its fragmentation pattern. Abbreviations: Lys; lysine, Hyl: hydroxylysine, glegal; glucosylgalactosyl, gal; galactosyl.
Figure 3. Western blot detection of α1(I) C-telopeptide cross-linked structures from CNBr-digested skin, cornea and tendon gives a similar, characteristic pattern distinct to that from bone. (A) Illustration of the cross-linking interaction sites between telopeptides and helical domains of 4D-staggered neighboring collagen molecules packed in fibrils. Underneath, a line drawing shows the position of CNBr peptides defined by methionine residues (marked by vertical lines) in rat and bovine α1(I) and α2(I) chains. The mAb 1G7 recognizes an epitope at the C-terminus of the α1(I) C-telopeptide whether or not its cross-linking lysine is hydroxylated, oxidised to aldehyde by lysyl oxidase or cross-linked in oligomeric peptides. (B) Aliquots of CNBr-digested samples of bovine cornea and skin, rat skin, tail tendon and demineralized bone were run on 12.5% SDS-PAGE, stained with Coomassie blue (left panel) or transblotted to PVDF membrane and α1(I) C-telopeptide-containing bands detected by probing with mAb 1G7 and chemiluminescence development (right panel). Lane loads as follows were adjusted to give approximately equal stained band intensities across the left gel (bovine cornea, 10 µg; bovine skin, 9 µg; rat skin 8 µg; tail tendon, 5 µg; bone, 7 µg) and the right immunoblot (bovine cornea, 10 µg; bovine skin, 2.5 µg; rat skin 8 µg; tail tendon, 6 µg; bone, 2 µg).
**Figure 4.** Mass spectral identification of $\alpha_1$CB6 (C-telopeptide containing) linear, cross-linked and partial-cleavage forms in a CNBr-digest of bovine cornea fractionated by C8-HPLC/serial fraction SDS-PAGE. An IMAC column was used to enrich histidine-containing peptides from a CNBr-digest of bovine cornea. This pool was run on C8-HPLC as shown in the lower profile, and serial collected fractions were run on 12.5% SDS PAGE and stained with Coomassie blue (insert panel). All bands were excised, digested with trypsin in-gel and the digests run on LC/MS to identify them. Identities of all individual and partial-cleavage linear peptides are shown, as well as a peptide identified as the cross-linked dimer of $\alpha_1$CB6. No $\alpha_1$CB6-containing peptides other than the monomer, dimer and their 7,6 partial-cleavage forms were found.
Figure 5. Allysine aldol cross-linked dimers of N-telopeptides and C-telopeptides are major products in a bacterial collagenase digest of corneal collagen if unreduced by sodium borohydride. Bovine corneal tissue was completely digested with bacterial collagenase under conditions that preserved reproducible telopeptide and helical domain cross-linking sequences. A) Elution profile of resulting peptides on C8 reverse phase HPLC. The late-eluting peaks labeled C-telo aldol and N-telo aldol were shown by MS and MSMS (B, C, D, E) to have the structures shown.
Figure 6. Mass spectral identification of N-telopeptide aldol dimers linked to a peptide containing α1(I)K930 from bacterial collagenase-digested, sodium borohydride-reduced mouse skin and tail tendon collagens. C8 RP-HPLC chromatography of tissue digests prepared as in Fig 5 gave the same prominent α1(I)N x α2(I)N x α1(I) helix (A, B) and α1(I)N x α1(I)N x α1(I) helix (C, D) peptides from mouse skin and tail tendon, which were identified by their mass spectra. E) MSMS fragmentation pattern of the 1200.58+ parent ion of α1(I)N x α1(I)N x α1(I) helix from mouse skin, shown as an example.
Figure 7. Mass spectral identification of a C-telopeptide aldol dimer in labile linkage to a peptide containing glycosylated α1(I) K87 from bacterial collagenase-digested corneal collagen. Further mass spectral analysis of fractions from the C8 HPLC chromatogram shown in Fig 5 revealed a larger peptide spread through fractions 35-40 underlying the aldol dimeric peptides. (A) Mass spectrum and determined structure of the trimeric peptide. (B) MSMS (ms2) fragmentation of the parent 880.26+ ion followed by ms3 fragmentation patterns from the helical arm fragment GMKGHR with (C) and without (D) the originally attached glucosylgalactosyl modification. In the helical domain sequence shown, K is hydroxylysine, as confirmed by parent and fragment ion masses.
Collagen Cross-linking Revisited

Figure 8. LC/MS analysis of collagen and elastin cross-link derivatives in acid hydrolysed cornea, skin and tendon. Cross-linking amino acid products in acid hydrolysates of whole tissues were separated from bulk amino acids by organic solvent partition on a hydrated cellulose column, then resolved and analyzed by electrospray LC/MS on the LTQ XL using a Cogent 4 diamond hydride column essentially according to the method of Naffa et al (2016). (A) Total ion current elution profile from the LC column for hydrolysates of sodium borohydride-reduced samples of bovine tendon, bovine skin, mouse skin and bovine cornea. The elution positions of the various major collagen cross-linking amino acids HP, hydroxylysyl pyridinoline; HHMD, histidinohydroxymerodesmosine; HHL, histidinohydroxylysinonorleucine; HLN, hydroxylysino-norleucine; LNL, lysinonorleucine and Des, desmosine (from elastin) are shown. (B-E) Mass spectra scrolled across the regions of LC eluent for each tissue sample showing the masses of the various cross-linking entities. Both HHL and HHMD are prominent in bovine skin and cornea but absent from mouse skin. (F) Acid hydrolyzed pooled fractions from the C8 RP-HPLC chromatogram of bacterial collagenase-digested cornea containing the C-telopeptide aldol dimer to α1(I) K87 adduct (Fig 6) reveal a low but significant yield of HHL, as did acid hydrolysates of whole bacterial collagenase-digested cornea samples.
Figure 9. Protein sequence alignment of the α1(I) C-telopeptide. The cross-linking lysine (K16C) is shown in red. The histidine (H18C), which is incorporated on acid hydrolysis into HHL, is shown in blue, and the arginine (R22C) that helps catalyze the cross-link formation is shown in green. It is notable that both mouse and rat lack the histidine residue and do not form HHL. Sequences are from Ensembl databases.

|            | α1(I) C-telopeptide |
|------------|---------------------|
| Human      | GPPSAGFDFSFLPQPPQEKAHDGGRYY |
| Bovine     | GPPSGGYDLSFPLPQPPQEKAHDGGRYY |
| Goat       | GPPSGGYDLSFPLPQPPQEKAHDGGRYY |
| Horse      | GPPSGGFDFSFLPQPPQEKSQDGGGRYY |
| Chicken    | GPPSGGFDFSFLPQPPQEKAHDGGRYY |
| Dog        | GPPSGGFDFSFLPQPPQEKAHDGGRYY |
| Mouse      | GPPSGGYDGSFLPQPPQEKSQDGGGRYY |
| Rat        | GPPSGGYDGSFLPQPPQEKSQDGGGRYY |
Figure 10. Proposed mechanism of formation of HHL from C-telopeptide aldols compared with HHMD from N-telopeptide aldols. (A) HHL formation on acid hydrolysis is restricted to skin and cornea from species that have a conserved α1(I) C-telopeptide histidine residue. Both this histidine and glucosylgalactosylhydroxylysine at α1(I) K87 appear to be necessary for HHL formation from tissue on acid hydrolysis. The proposed reaction mechanism involves a retro-aldol cleavage and covalent addition of the telopeptide histidine imidazole to give the accepted structure for HHL (10, 47). (B) HHMD is a product of borohydride reduction of the aldimine addition product between N-telopeptide \( \alpha,\beta \)-unsaturated aldol cross-links and K930 hydroxylysine together with the Michael adduction of the histidine at H932 in the same helical sequence through its imidazole group. The peptide structures obtained from borohydride-reduced, bacterial collagenase-digested skin and tail tendon collagen show that the histidine has to be from this same chain, not another α-chain. If the latter occurred, more complex peptides should have been found. The accepted structure of HHMD with the ring nitrogen of histidine imidazole added to the \( \beta \)-carbon of the aldol is shown (10, 37).
Analyses of lysine aldehyde cross-linking in collagen reveal that the mature cross-link histidinohydroxylysinonorleucine is an artifact
David R. Eyre, MaryAnn Weis and Jyoti Rai

J. Biol. Chem. published online February 7, 2019

Access the most updated version of this article at doi: 10.1074/jbc.RA118.007202

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts