COLLAGEN CROSS-LINKING: MATURATION OF DIVALENT KETO-IMINES BY AN ALTERNATIVE MECHANISM TO PYRIDINOLINE FORMATION IN CARTILAGE

David R. Eyre, Mary Ann Weis, and Jiann-Jiu Wu

Orthopaedic Research Laboratories, Department of Orthopaedics and Sports Medicine, University of Washington, Seattle, Washington 98195-6500
Running head: Collagen cross-linking in cartilage revisited

Address correspondence to: David R. Eyre, Ph.D., 1959 NE Pacific St., HSB BB1052, Box 356500, Seattle, WA 98195-6500 USA. Fax: 206-685-4700; E-mail: deyre@u.washington.edu
Abstract
The tensile strength of fibrillar collagens depends on stable intermolecular cross-links formed through the lysyl oxidase mechanism. Such cross-links based on hydroxylysine aldehydes are particularly important in cartilage, bone and other skeletal tissues. In adult cartilages, the mature cross-linking structures are trivalent pyridinolines, which form spontaneously from the initial divalent keto-imines. We have examined if this is the complete story or if other keto-imine maturation products also form, since the latter are known to disappear almost completely from mature tissues. Denatured, insoluble bovine articular cartilage collagen was digested with trypsin and cross-linked peptides were isolated by copper chelation chromatography which selects for their histidine-containing sequence motifs. The results showed that in addition to the naturally fluorescent pyridinoline peptides, a second set of cross-linked peptides was recoverable in high yield from mature articular cartilage. Sequencing and mass spectral analysis identified their origin from the same molecular sites as the initial keto-imine cross-links but they did not fluoresce and were non-reducible with NaBH₄. Based on their mass spectra, they were identical to their precursor keto-imine cross-linked peptides but the cross-linking residue had an M+188 adduct. Considering the properties of an analogous adduct of identical added mass on a glycated lysine-containing peptide from type II collagen, we predict similar dihydroxyimidazolidine structures formed from their keto-imine groups by spontaneous oxidation and free arginine addition. We suggest the trivial name arginoline for the keto-imine cross-link derivative. Mature bovine articular cartilage contains about equimolar amounts of arginoline and hydroxylysyl pyridinoline (HP) based on peptide yields.

Introduction
All fibril-forming collagen types, I, II, III, V and XI, of vertebrates are cross-linked by the lysyl oxidase mechanism (1,2). Two reaction pathways can be defined, one based on lysine aldehydes, the other on hydroxylysine aldehydes, which operate tissue-dependently. A combination of both pathways is evident in some tissues, for example the unique cross-linking profile of bone collagen (3). Cartilage collagens exclusively use the hydroxylysine aldehyde route (4,5). Newly made type II collagen in hyaline cartilages and other tissues becomes cross-linked almost exclusively by divalent keto-imines, each formed by the addition of a telopeptide hydroxylysine aldehyde to a hydroxylysine at either residue 87 or 930 of the triple-helix. As fibrils mature, divalent keto-imines spontaneously interact to form a trivalent cross-link hydroxylysyl pyridinoline (HP) (6,7). The stoichiometry is such that 1 mole of HP is formed from two moles of hydroxylysino-ketonorleucine (HLKNL, the keto-imine). This was established by direct amino acid analysis after borohydride reduction of tissue to stabilize the keto-imines to acid hydrolysis (7,8) and also by radiolabeling of cartilage with 14C-lysine in vivo (9) and in vitro (10). A reaction mechanism was proposed (aldol addition followed by ring closure and oxidation) in which two keto-imines condensed to form the pyridinoline and release one of the helical domain hydroxylysine side-chains (7,8). An alternative reaction mechanism was suggested in which the reactants are a keto-imine and a free hydroxylysine aldehyde on a second telopeptide in the same collagen molecule as the first (11). This doesn’t fit the stoichiometry, relatively slow reaction kinetics of pyridinoline appearance after fibrils have formed or the low level of free aldehydes in even newly made collagen. Whatever the mechanism, the keto-imines disappear from mature tissues and pyridinolines become essentially the sole cross-linking residues (12,13). The content of HP cross-links ranges from 1 to 2 moles per mole of type II collagen in mature human articular cartilage (14) and 1.3 to 1.9 moles per mole in mature bovine articular cartilage (13,15). Cartilages contain the highest collagen HP levels of all connective tissues (2). The chemistry, content and spatial arrangement of cross-linking bonds are all thought to be important determinants of collagen fibril properties and overall tissue mechanical behavior. Since mature articular cartilage collagen turns over very slowly at best throughout life (16,17), most of it is laid down during growth and has to last a lifetime (5).
Its resistance to mechanical failure and ability to withstand other biological wear, tear and proteolytic stresses is therefore a key determinant of joint longevity. Determining what drives or enables the initial keto-imine cross-links to convert efficiently (or not) into pyridinoline residues may be important for understanding why joints fail earlier in some individuals than in others. Using a procedure we developed for selecting for the total pool of collagen type II cross-linking domains from protease digests of tissue, we aimed to answer this question in a study of fetal, calf and adult bovine articular cartilage by mass spectrometry.

Experimental Procedures

Tissue preparation

Articular cartilage from 18-month steer knee joints (local abattoir) was extracted with 4M guanidine HCl, 0.05M Tris-HCl, pH 7.4, for 48 hr at 4°C. The washed residue was digested with pepsin and solubilized type II collagen was recovered in the 0.7M NaCl precipitate (12). The isolated type II collagen was digested with recombinant proteomics grade trypsin (Roche Diagnostics) in 0.05M Tris, 0.15M NaCl, pH 8.0, for 20 hr at 37°C with an enzyme to substrate ratio of 1:200. Cyanogen bromide (CNBr) peptides were generated by digesting pepsin-solubilized type II collagen with CNBr in 70% formic acid for 24 hr at room temperature (18).

Type II collagen was similarly extracted from fetal and 3 mo. calf epiphyseal cartilages with pepsin and salt-precipitated then digested with CNBr.

SDS-PAGE

The CNBr-digest was run on 12.5% SDS-PAGE and the gel before staining with Coomassie Blue was scanned on a Perkin-Elmer fluorescence spectrophotometer equipped with a thin layer/slab-gel scanning accessory and an E-C Apparatus densitometer (12).

Immobilized metal ion affinity chromatography (IMAC)

An IMAC column was prepared by charging HiTrap chelating HP beads (Amersham Biosciences) with CuCl2. Tryptic digests of 18 mo steer collagen were loaded on the column in 0.05M Tris, 0.15M NaCl, pH 8.0. Bound peptides were eluted in three pools: pool A with 0.1M Na acetate, pH 6.35; pool B with 0.1M Na acetate, pH 4.6; and pool C with 0.1M Na acetate, pH 4.6, containing 0.2M imidazole. Each CNBr digest of pepsin-solubilized type II collagen from fetal and calf cartilage was loaded in 0.05M Tris-HCl, 0.15M NaCl, 2M guanidine HCl, pH 8.0, and the bound peptides were eluted with 0.1M Na acetate, pH 4.6, containing 2M guanidine HCl, and 0.2M imidazole.

Reverse-Phase (RP) HPLC

Each tryptic peptide pool from the IMAC column was run on RP-HPLC (C8 RP-300, 25 cm x 4.6 mm, Brownlee Columns, Applied Biosystems) (19). Peak tryptic peptide fractions were aliquoted for protein sequencing and/or mass spectrometry.

The CNBr-peptides from fetal and calf cartilage type II collagen that bound to the Cu++ IMAC column consisted mostly of non-cross-linked and cross-linked forms of CB12 and CB9,7 that were resolved into two peaks by C8 RP-HPLC (not shown). The CB12 and CB9,7 pools were separately digested with trypsin and re-run on RP-HPLC.

Protein microsequencing

Edman amino-terminal sequencing was carried out on a Porton 2090E machine equipped with on-line HPLC analysis of phenylthiohydantoin derivatives (20).

Electrospray Mass Spectrometry

Mass spectrometry was performed on an LCQ Deca XP ion-trap mass spectrometer (ThermoFinnigan) with in-line liquid chromatography (LC) on a C8 capillary column 300μm i.d. x 150mm long (Grace Vydac 208MS5.315) flowing at 4.5μl/min (21). The mobile phase consisted of buffer A: 0.1% formic acid with 2% buffer B in MilliQ water; and buffer B: 0.1% formic acid in 3:1 (v/v) acetonitrile:N-propanol. The column eluent was introduced into the mass spectrometer with an atmospheric pressure electrospray ionization source (ESI) with a 3kV spray voltage and inlet capillary temperature of 160°C. The mass spectrometer was run in an automatic triple-play mode, cycling through a full scan, zoom scan and ms/ms every few milliseconds.
Results

SDS-PAGE resolution of cross-linked peptides

Three forms of the cyanogen bromide-peptide CB9,7 from CNBr-digested pepsin-solubilized type II collagen were resolved as distinct bands on SDS-PAGE (Fig. 1). By N-terminal sequencing the fastest of the bands appeared to contain a mixture of linear (i.e., non-cross-linked) CB9,7 mixed with CB12 cross-linked to either one or two C-telopeptides. The middle band is CB9,7 cross-linked to one N-telopeptide and the slowest band is CB9,7 cross-linked to two N-telopeptides. Fig. 1 shows the results of densitometric and fluorometric scans of the gel lane shown. The two fluorescent peaks correspond to the trivalent hydroxylysyl pyridinoline (HP) cross-linked forms of CB9,7 and CB12. The divalent cross-links are not fluorescent. Also as noted above, the divalent and trivalent cross-linked forms of CB12 appear to comigrate with linear CB9,7 as a single band but the divalent and trivalent cross-linked forms of CB9,7 separate as distinct bands so their relative amounts can be assessed. There are about equal amounts of the divalent and trivalent versions of CB9,7. This 1:1 ratio is not predicted from quantitative analysis for the known cross-linking residue HP and the divalent hydroxylysino-ketonorleucine (HLKNL) in borohydride-reduced and hydrolyzed 18 mo. bovine cartilage, in which the ratio of HP to HLKNL was 40/1 (13). This implies that most of the divalent cross-linking residue in CB9,7X resolved in the gel is of a form not detectable in the HLKNL assay.

Copper-affinity chromatography

The triple-helical domain of bovine α1(II) contains only two histidine residues, both in a GHR sequence immediately C-terminal to the cross-linking hydroxylysine residues at K87 and K930 (Ensembl entry: ENSBTAP00000017505). Since trypsin cleaves after arginine, all cross-linked tryptic peptides will contain histidine. A copper-loaded IMAC column was used to bind the histidine-containing peptides from a tryptic digest of pepsinized type II collagen. Figs. 2A,B and C show the RP-HPLC profiles of three successive pools eluted from the IMAC column at pH 6.35 (A), pH 4.6 (B), and 0.2M imidazole, pH 4.6 (C). Pool C contained the most tightly bound material that eluted only with the addition of imidazole to displace histidine. Fig. 2D shows the RP-HPLC profile of an aliquot of the total trypsin digest prior to running on the IMAC column. The trivalent cross-linked peptides eluted in pools A and B, whereas the divalent cross-links were more tightly bound to the IMAC column and eluted in pools B and C.

The peptide structures shown against the various peaks identified in Fig. 2 were determined by a combination of N-terminal sequence analysis and mass spectrometry. Variants owing to partial cleavage by pepsin in the telopeptide domains were also identified. For example, the series of fluorescent peaks eluting between 18-32min in Fig. 2B are all N-telopeptide to K930 pyridinoline isoforms with various missing or present terminal methionine and phenylalanine residues.

Mass spectrometric analyses

The main fluorescent peptide from pool A in Fig. 2 was shown by mass spectrometry to be the HP-bonded structure linking K87 to two C-telopeptides (Fig. 3A). The early fluorescent peptides from pool B in Fig. 2 were shown by mass spectrometry to be different pepsin-cleavage products of the HP-bonded structure linking K930 to two N-telopeptides (Fig. 3B).

The two non-fluorescent peptides that were the major components eluting from pool C on RP-HPLC as shown in Fig. 2 were identified by mass spectrometry and N-terminal sequencing to be divalent cross-linked peptides linking residue K930 to a single N-telopeptide (Fig. 4A) and residue K87 to a single C-telopeptide (Fig. 4B). The mass of the peptides (and their MS/MS fragmentation patterns) showed that in each the cross-linking residue had a mass greater by 188 Daltons than the theoretical mass of a keto-imine cross-linking residue (hydroxylysino-ketonorleucine). The same peptides were recovered in the same yield from the total digest (Fig. 2D) without IMAC chromatography and shown to have the +188 adduct still.

Comparison of fetal epiphyseal with 3 mo. calf articular cartilage

Tryptic peptides from the CB12 pool, isolated by C8 RP-HPLC after Cu++ IMAC binding were compared for pepsin-solubilized type II collagen from fetal growth cartilage and 3 mo. calf articular cartilage (Fig. 5A). From fetal cartilage essentially the total pool of divalent cross-linked peptide from
the K87 site gave a mass matching that of the keto-imine form (Fig. 5B), whereas from 3 mo.
calf cartilage two forms were detected with more than half already in the +188 adduct form. The
linear peptide sequence (uncross-linked) containing K87 was also identified in the
chromatograms. Interestingly, K87 in the uncross-linked peptide was hydroxylsine glycoside but
the cross-linked peptides were largely not enzymically glycosylated.
The CB9,7 linear and cross-linked peptide fraction (see Methods) gave tryptic peptides with the same
essential findings, mostly the keto-imine cross-link at K930 from fetal cartilage and over half as the
+188 adduct from calf cartilage (results not shown).

Identification of a non-enzymically glycated linear tryptic peptide from type II collagen
Analysis of additional peptides from the Cu ++- column eluent by N-terminal sequencing and mass
spectrometry revealed a hexapeptide (Fig. 2C, fraction 16) containing K930, but the K930
sequencer cycle was blank and on mass spectrometry (Fig. 6) an additional mass of 350 Daltons on a lysine residue was observed (equals +162 +188, that matches + hexose + [O] + arginine adduct). The MS/MS fragmentation spectra shown in Fig. 6A were consistent with the structure shown in Fig. 6B, which is predicted to be the ε-amino hexosyl (presumably glucosyl) derivative of the lysine side-chain of the peptide GLKGHR, in which the keto-imine group has become oxidized to a dicarbonyl that spontaneously had added a free arginine through its guanido group. This maturation reaction is the same mechanism proposed for the +188 adduct found on the lysyl oxidase-mediated keto-imine cross-linked peptides.

Fig. 7 shows the proposed structure of the keto-imine cross-link adduct, a 3,4 dihydroxy
imidazolidine, for which we suggest the trivial name, arginoline. The unusually strong binding of the adduct to the Cu ++ IMAC column (resulting in the fraction shown chromatographed in Fig. 2C) lends support to a structure that embodies an imidazole ring.

Discussion
The results show that about half the initially formed keto-imine cross-links in type II collagen
of mature bovine articular cartilage had matured by an alternative chemical reaction to that
producing trivalent HP. The product of this alternative reaction path is a non-reducible adduct formed from the keto-imine, which functionally is still a divalent intermolecular cross-link. Quantitatively about equimolar amounts of peptides based on the arginyl keto-imine adduct (arginoline) and HP cross-linked peptides were recovered from each of the two telopeptide-to-helix cross-linking loci of bovine type II collagen. Thus, if two keto-imines effectively condense to form one HP (see below) then roughly 2/3 of the keto-imines in the cartilage collagen had matured to HP and 1/3 to arginoline. This means that cartilage collagen fibrils are even more densely cross-linked than previously appreciated from their HP content alone (Fig. 8 illustrates the molecular concept). The reported HP content of 18 mo. bovine articular cartilage ranges from 1.3 to 1.9 moles per mole of collagen (13,15), which means that 2.6 to 3.8 of the total 6 telopeptides per type II collagen molecule (three at each end) are involved in HP linkage. At an equal concentration the modified keto-imine cross-link would account for a further 1.3 to 1.9 telopeptides per molecule, i.e., both forms of cross-link can account for between 3.9 and 5.7 of the 6 telopeptides theoretically available for cross-linking. Since some telopeptide lysines will be missed by lysyl hydroxylase and/or lysyl oxidase, this can account for the remainder.
The various types of reducible cross-link in collagen disappear from most adult connective
tissues, whether their primary formation pathway is from lysine or hydroxylysine aldehydes (22,23).
Several reaction mechanisms have been proposed to account for this disappearance in addition to pyridinoline formation from keto-imines, for example oxidation of aldímnes (initial products on the lysine aldehyde pathway) to amides (24), gem-diamine formation (25) and natural reduction. Natural reduction was ruled out by amino acid analysis after acid-hydrolysis without borohydride treatment (26,27). Oxidation seems a more likely path extracellularly in biological systems and α-amino adipic acid has been identified in skin collagen after acid hydrolysis and proposed to originate from the amide formed by oxidation of the aldímne, Δ-HLNL (24).
There is also a precedent for keto-imine oxidation based on the known fate of the glycosylamines that form adventitiously through non-enzymic glycosylation (glycation) of lysine side-chains in long-lived proteins. The initial glycosylamine adducts undergo Amadori rearrangement to a keto-imine, which can oxidize to an $\alpha, \beta$ dicarboxyl (28,29). The latter show a high affinity and selectivity in forming additional compounds with the guanido group of arginine (30) or another guanido-containing molecule such as creatine (31). A complex imidazolidine, called DOGDIC (32), has been identified among the many AGE (advanced glycation end-product) structures that go on to form from the Amadori adduct between glucose and lysine. Dihydroxyimidazolidines have been identified as major products of $\alpha, \beta$ dicarboxyl addition to arginine (30). The results of our present analyses can best be explained if the keto-imine cross-links that fail to form pyridinones in collagen fibrils undergo a two-step process of slow, spontaneous oxidation followed by rapid addition of free arginine. The latter is available in tissue fluids. Indeed, concentrations of free arginine are likely to be higher in cartilage matrix than in other tissues due to the high negative fixed charge density of the sulfated proteoglycans. The results show no evidence of peptides resulting from the addition of peptidyl arginine side-chains to the keto-imines (which would be a further cross-linking reaction) since no peptides larger than the divalent structures plus an M+188 adduct, other than the HP-containing peptides, were recovered from the type II collagen. Thus, there was no evidence for trivalent cross-linking to a third peptide sequence containing an arginine residue, only the free arginine adduct.

The recovery and identification of the linear hexapeptide GLK$^*$GHR from trypsin-digested cartilage collagen, in which K$^*$ is a lysine bearing an oxidized hexosylamine that had bound to arginine, supports the proposed mechanism (oxidation to an $\alpha, \beta$ dicarboxyl, then arginine addition) in cartilage in vivo. Indeed the mass spectral properties of the modified hexapeptide (Fig. 6), that presumably originated from non-enzymic glycation in vivo (23), gave the clue to interpreting the structure of the keto-imine adduct on the divalent cross-linked peptides. The mass spectral evidence that the precursor is the lysine form of the peptide (rather than hydroxylysine) suggests that a lysine side-chain in this sequence motif is more readily glycylated than hydroxylysine. This lysine (K930) is almost fully hydroxylated in type II collagen but the equivalent glycylated derivative of a hydroxylysine-containing GLK$^*$GHR peptide was not found.

To confirm that this phenomenon wasn’t peculiar to bovine cartilage, we also analyzed pig articular cartilage (from carpal bones) with resulting very similar, high levels of the M+188 adduct of the equivalent keto-imine cross-linked type II collagen peptides (results not shown). Analyses of human articular cartilage also revealed the equivalent M+188 peptides from type II collagen, but in lower yields and apparently together with related peptides that had lower mass adducts on the keto-imine cross-linking residue. Since the human cartilage was from age >30 yr adults, we suspect that the initial dihydroxy adducts (Fig. 7) can undergo dehydrations or rearrangements of the kind noted for AGE products of non-enzymically glycosylated lysine residues (23,33). We also analyzed type I collagen from bovine meniscus fibrocartilage by similar methods (results not shown) and established that the same M+188 adduct on divalent cross-linked peptides from the N- and C-telopeptide cross-linking sites was a major feature of fibrocartilage type I collagen. So the proposed pathway of oxidation and arginine addition appears to operate in cartilages in general. Whether this is a widespread mechanism of keto-imine maturation in all tissue collagens or is cartilage-specific will require further study. Arginoline is not reductible with Na borohydride (results not shown) and is likely to be destroyed on acid hydrolysis so would be missed using the earlier techniques of cross-link analysis.

This alternative route of keto-imine maturation is consistent with the results of $^{14}$C-lysine labeling of collagen cross-links in vitro (10) and in vivo (9). The kinetics of keto-imine (HLKNL) and pyridinoline (HP) production from $^{14}$C-lysine were reported for explant cultures of calf cartilage plugs in vitro (10) monitored for 28 days. By 21-28 days, from a $^{14}$C-lysine pulse, HP formation had leveled off and 60% of the $[^{14}$C]-HLKNL, which peaked at 2-7 days post-pulse, had been converted to HP. But 40% of the peak $[^{14}$C]-HLKNL remained at 28 days. The present findings of a secondary maturation pathway for the fraction of
HLKNL unable to react during cartilage development to form HP is consistent with these observations and the arginoline pathway having slower kinetics than HP condensation. The present findings from fetal and calf cartilage support this concept. Fetal calf cartilage gave essentially only HLKNL in its divalent peptide pool with only trace levels of arginoline whereas 3 mo calf cartilage had both HLKNL and the arginoline adduct. This implies maturation of the keto-imine to arginoline in tissue in vivo with a t-1/2 time in the order of weeks but not years. Whether individual chondrocytes can exert some control over this reaction locally in the matrix, to close down further trivalent cross-linking for example (e.g., through a pulse of H₂O₂ or reactive oxygen species), or if it is purely adventitious would be important to understand. The material quality of articular cartilage is likely to be influenced by the quantitative balance between levels of modified divalent and trivalent cross-links and their anatomical distribution profile in a mature collagen network. It is notable that in adult human joints the content of pyridinoline cross-links in type II collagen can vary by a factor of 2 between individuals (in the range 1 to 2 moles/mole of collagen; 14).

**Conclusion**

The results reveal two maturation pathways for the initial intermolecular keto-imine cross-links in cartilage collagen, condensation to pyridinolines or oxidation and arginine addition. Both pathways involve an oxidative step. Pyridinolines presumably form to the extent permitted stereochemically by the spatial relationships set up between telopeptides and helical sites during the molecular packing of collagen molecules into fibrils (6,19). On a molar basis, about two-thirds of the divergent cross-links in adult bovine articular cartilage had converted to pyridinoline (HP) and one-third to arginoline. The latter's formation mechanism and its dihydroxyimidazolidine structure are essentially identical to that of a modified non-enzymically glycated peptide also recovered from the collagen. The findings reveal an even higher level of total cross-linking by the lysyl oxidase mechanism in cartilage collagen than previously appreciated. The oxidative status of the matrix could affect the balance between trivalent pyridinolines and divalent cross-links in mature collagen fibrils and so tissue resilience and longevity.

**References**

This work was supported in part by NIH grants AR37318 and AR36794 to DRE

Figure Legends

Figure 1
SDS-PAGE of CNBr-digested type II collagen purified from bovine articular cartilage. The collagen was solubilized by pepsin and salt precipitated before CNBr digestion and running on 12.5% SDS-PAGE. The Coomassie Blue-stained gel is shown on the left and the densitometric and fluorometric (330nm Ex, 395nm Em) scans for pyridinoline-containing peptides on the right. The CNBr-peptide map of bovine α1(II) showing positions of methionines and CB-peptides referred to by number in figure and text is shown above.

Figure 2
Reverse-phase HPLC of Cu++ IMAC affinity-column resolved tryptic peptide pools from bovine type II collagen. Equivalent aliquots of each of the three peptide pools were eluted from a C8-column with a 0-40% gradient of acetonitrile:n-propanol (3:1 v/v) in 0.1% trifluoroacetic acid over 60 min. A) Pool A; B) Pool B; C) Pool C; D) Total trypsin digest. The peptides shown in the chromatograms were identified by Edman N-terminal sequencing and ion-trap tandem mass spectrometry.

Figure 3
Tandem mass spectra identifying the main fluorescent peptides resolved from pools A and B of Fig. 2. In A) a full-scan mass spectrum from the LC-MS profile of the parent peptide identified in Fig. 2A (upper panel) and the MS/MS fragmentation spectrum of its 4+ ion (lower panel) are shown. The various b (N-terminal) and y (C-terminal) fragment ions in conjunction with N-terminal sequencing results and fluorescence properties establish the C-telopeptide to N-helix trivalent structure shown.

In B) similarly, the parent peptide mass spectrum (upper panel) of the fluorescent peptide eluting at 29 min in Fig. 2B and its MS/MS fragmentation pattern (lower panel) establish the N-telopeptide to C-helix trivalent structure shown.
Figure 4
Tandem mass spectra identifying the two main non-fluorescent divalent cross-linked peptides resolved from pool C of Fig. 2. In A) a full-scan mass spectrum from the LC-MS profile of the parent peptide eluting at 24 min in Fig. 2C (upper panel) and the MS/MS fragmentation pattern of its 3+ ion (lower panel) are shown. In B) similarly, the parent peptide mass spectrum of the peptide eluting at 36 min in Fig. 2C (upper panel) and its MS/MS fragmentation pattern (lower panel) establish the peptide structures shown. The molecular mass of each peptide is 188 Daltons more than the theoretical mass of the equivalent keto-imine structure (see also Fig. 5 for a direct mass spectral comparison of an adduct versus keto-imine pair of otherwise identical tryptic peptides).

Figure 5
Reverse-phase HPLC of tryptic peptides from the CB12 pool comparing fetal and 3-mo. calf cartilages. The C8 RP-HPLC elution profiles (0-30% gradient) of the isolated CB12 pools are shown (A). Tandem mass spectra compare keto-imine (B) and adduct (C) versions of the same divalent cross-linked peptide from fetal and 3 mo. calf, respectively. It should be noted that this latter peptide differs in lacking an N-terminal arginine on its C-terminal telopeptide sequence compared with the variant that is the major form in the peptide preparation resolved in Fig. 4. Such partial cleavage variants were consistently observed and reflect variable cleavages by pepsin in different digests (if pepsin cuts immediately before arginine then trypsin fails to cleave after it). The other structures shown in (A) were all identified by their mass spectra together with N-terminal sequencing results.

Figure 6
Tandem mass spectra of a non-enzymically glycosylated linear peptide from type II collagen. Trypsin digests of mature (18 mo. steer) type II collagen on RP-HPLC gave a peptide (Fig. 2C, 16 min elution) of N-terminal sequence GLXGHR where X was a blank cycle. (A) Tandem mass spectra: top panel, full scan mass spectrum across the LC-MS peak, consecutive next three panels: MS/MS fragmentation spectra of the parent peptide 3+, 2+ and 1+ ions respectively. The boxed dihydroxyimidazolone structure shown in A is predicted from the parent mass, its neutral losses and fragment ions. In B) a proposed mechanism of formation by non-enzymic glucose addition, keto-imine (Amadori) rearrangement, oxidation and addition of arginine through its guanido group.

Figure 7
Proposed mechanism of formation, structure and trivial name for the keto-imine cross-link adduct, arginoline. The mechanism parallels that shown in Fig. 6 for the keto-imine resulting from non-enzymic glycosylation of a lysine-containing collagen peptide. Oxidation produces a reactive dicarbonyl that adds free arginine through its guanido group. The proposed structure of arginoline fits exactly the mass of the cross-linking residue in the N-telopeptide-to-helix and C-telopeptide-to-helix structures shown in Fig. 4.

Figure 8
Illustration of a cross-linking site between registered pairs of collagen molecules staggered by 4D (where D=234 amino acid residues, the Hodge-Petruska stagger). In the type II collagen molecule all three telopeptides at both N- and C- ends can form a cross-linking aldehyde from hydroxylysine (hydroxyallysine, see Fig. 7). Just one end is shown. The pyridinol ring illustrates how linkages of two telopeptides to a helical site can form trivalent pyridinoline cross-links and the vertical dumbbells illustrate divalent (keto-imine) bonds. Such an arrangement at maximum theoretical cross-link occupancy would use all 6 telopeptides/molecule and result in equimolar trivalent and divalent cross-links each at a stoichiometry of 2 moles per mole of collagen.
Fig. 1

Collagen II CNBr peptide map

Absorbance

Migration Distance (cm)

Fluorescence

0 4 8 12

CB10

CB11

CB8

trivalent CB9,7X

divalent CB9,7X

CB9,7 and CB12X

CB12
Fig. 2
Fig. 3
Fig. 4
Fig. 5
Fig. 6
A 3,4 dihydroxy imidazolidine formed by condensation of free arginine with the oxidized keto-imine cross-link.

Fig. 7
Collagen cross-linking: Maturation of divalent keto-imines by an alternative mechanism to pyridinoline formation in cartilage
David R. Eyre, Mary Ann Weis and Jiann-Jiu J. Wu

J. Biol. Chem. published online April 2, 2010

Access the most updated version of this article at doi: 10.1074/jbc.M110.111534

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