Studies on the Location of Intermolecular Cross-Links in Collagen. Isolation of a CNBr Peptide Containing δ-Hydroxylysinonorleucine†

Andrew H. Kang*

ABSTRACT: Purified soluble rat-tail tendon collagen, *in vitro* reconstituted native-type fibrils, and intact tail tendon were separately cleaved with cyanogen bromide at the methionyl residues after reduction with tritiated sodium borohydride, and the resulting peptides were fractionated by ion exchange and molecular sieve chromatography. The cross-link related compounds in the CNBr peptides were identified and located by amino acid analysis and a scintillation spectrometry after 6 N HCl and 2 N NaOH hydrolyses. The soluble collagen yielded all 15 specific peptides (9 from α1 and 6 from α2) in the expected amounts. No δ-hydroxylysinonorleucine or posthistidine compound were found in any of the peptide fractions. In the digests of the reduced, reconstituted native-type fibrils and in the intact tendons, however, the NH_2-terminal peptides, α1-CB1^{Ald} and α2-CB1^{Ald}, and β_{12}-CB1 and β_{11}CB1, were present in much decreased amounts. Most of the larger CNBr peptides contained partial residues of δ-hydroxylysinonorleucine, or the posthistidine compound, or both. These data indicate that during fibrillogenesis of rat-tail tendon *in vitro* and *in vivo*, the NH_2-terminal, aldehyde-bearing peptides participate in the formation of intermolecular cross-links and further that the intermolecular cross-links form at several different sites along the neighboring molecules. Although all of the possible cyanogen bromide peptides bearing intermolecular cross-links were not isolated in pure form due to the complexity of the mixture, one such peptide of mol wt 18,000 containing one residue of δ-hydroxylysinonorleucine was isolated and purified. The amino acid compositions of the peptide was consistent, within experimental error, with the sum of that for α1-CB1^{Ald} and α1-CB6, indicating its origin as an intermolecular amino-carboxyl cross-link. The presence of β_{11}-CB1 and β_{12}-CB1, the intramolecular cross-link peptides, in the digests of soluble collagen and their disappearance in the digests of the native-type fibrils and the intact tendons indicate that the intramolecular cross-links react further to form higher order aggregates.

The nature and location of the intermolecular cross-links through which collagen forms a continuously cross-linked polymeric network have been of widespread interest since the unusual mechanical stability of the collagenous structure is largely dependent on their formation. Defects in this process result in dramatic failure in the function of connective tissue such as occurs in osteolathyrism (Levene and Gross, 1959; see Piez, 1968b; Bornstein, 1970). The time-dependent decrease in solubility of collagen fibrils both *in vitro* and *in vivo* is accompanied by concomitant increase in the cross-link content, further supporting our ideas of this important structure–function relationship (A. H. Kang and C. Franzblau, in preparation).

Previous studies have shown that in rat and chick skin collagens specific lysyl residues located near the NH_2 termini of the α chains (α1-CB1 and α2-CB1) are oxidatively deaminated to aldehydes (allysine), which then react *via* aldol condensation to form the intramolecular cross-link (Bornstein and Piez, 1966; Kang *et al.*, 1969c). Such an aldol condensation product has been isolated and characterized in its reduced form after tritiated $NaBH_4$ reduction (Kang *et al.*, 1969a; Rojkind *et al.*, 1969) and the CNBr peptides bearing the precursor aldehydes as well as the aldol cross-link have also been isolated and characterized.

The results of investigations on the intermolecular cross-link from several laboratories (Bailey and Peach, 1968; Bailey *et al.*, 1970; Tanzer and Mechanic, 1968, 1970; Kang *et al.*, 1970) indicate that the intermolecular cross-links, δ-hydroxylysinonorleucine and lysinonorleucine are derived by a Schiff base formation between the allysyl residues and hydroxylysyl or lysyl residues on adjacent molecules. In addition, another cross-link tentatively designated the posthistidine compound, is present in most soft tissue collagens (Bailey and Peach, 1968; Kang *et al.*, 1970). The structure of the last compound is not known, but evidence suggests that it is in part derived from the aldol condensate (Franzblau *et al.*, 1970; Kang and Gross, 1970a; A. H. Kang and C. Franzblau, in preparation).

We have previously reported (Kang and Gross, 1970a) that during spontaneous cross-linking *in vitro* of purified, allysine-rich collagen obtained from D-penicillamine-treated chicks, the intermolecular cross-links are generated only in native-type fibrils. In nonstriated fibrils, only the intramolecular aldol cross-links are formed. These data suggest that exact spacial alignment of functional groups involved in intermolecular cross-linking is possible only in the quarter-stagger packing of the collagen molecules such as seen in the native-type fibrils. This relationship had been suggested earlier by Tanzer (1968). Since the location of all of the CNBr peptides in rat collagen is known, we wished to localize the sites of intermolecular cross-links using CNBr cleavage *in vitro* reconstituted native-type fibrils and the intact rat-tail tendon.

† From the Developmental Biology Laboratory, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. *Received July 21, 1971.* This is publication 562 of the Robert W. Lovett Memorial Group for the study of Diseases Causing Deformities. This work has been supported by grants from the Public Health Service (AM 3564) and the National Science Foundation (GB 24481).

* Russell L. Cecil Fellow of the Arthritis Foundation.

Our results, reported here, clearly indicate that during fibrillogenesis, the NH_2-terminal allysine-containing peptides, α1-CB1^{Ald} and α2-CB1^{Ald} and the intramolecular cross-link peptides, β_{11}-CB1 and β_{12}-CB1, disappear with concomitant appearance of δ-hydroxylysinonorleucine and the posthistidine cross-links and that the intermolecular cross-linking occurs at several loci along the length of both the α1 and α2 chains in different molecules.

Materials and Methods

Preparation of Collagen. Male Sprague-Dawley rats (Charles River Farm C-D strain), 7 weeks of age, were sacrificed by decapitation and tendons were stripped from the tails. All operations were performed at 5°. After washing thoroughly with distilled water, the tendons were minced into fine pieces. Soluble collagen was extracted with ten volumes of 0.5 M acetic acid and purified by repeated precipitation with 5% NaCl, followed by several more precipitations from the neutral solution with 20% NaCl as described previously (Kang *et al.*, 1969b). Purified collagen was stored in the lyophilized state in a cold desiccator over P_2O_5.

In some experiments, the tail tendons were used directly after washing with cold distilled water. The latter procedure did not solubilize significant amounts of collagen as judged by hydroxyproline assay of the supernatant.

In Vitro Reconstituted Native-Type Fibrils. Cross-striated fibrils were prepared from soluble collagen by a modification of the heat gelation method of Gross and Kirk (1958) as described previously (Kang *et al.*, 1970). The native type of fibrillar banding of each preparation was confirmed by electron microscopy.

Reduction with Tritiated Sodium Borohydride. Soluble rat-tail tendon collagen, the *in vitro* reconstituted fibrils, and tendons were reduced with tritiated $NaBH_4$ (200 mCi/mmole, New England Nuclear) at room temperature for 30 min in 0.05 M Tris (pH 7.5) containing 0.16 M NaCl (Kang *et al.*, 1970). A drop of anti-foam B (Dow Corning) was added to minimize foaming of the reaction mixture. Excess tritiated $NaBH_4$ was destroyed by lowering the pH to 4 by the careful addition of 1 M acetic acid, and salts were removed by exhaustive dialysis *vs.* 0.1 M acetic acid.

Cyanogen Bromide Cleavage. CNBr digestion of the various reduced soluble collagens, reconstituted fibrils and tendons was carried out in a manner similar to the methods of Bornstein and Piez (1966). The proteins were suspended in 70% formic acid (10 mg/ml), the vessel flushed with nitrogen, and 500-fold molar excess (relative to the methionine content) of CNBr was quickly added. The reaction was allowed to proceed for 4 hr at 30°. Any insoluble material was then removed by centrifugation at 30,000*g* and the supernatant was lyophilized. The extent of conversion of methionine to homoserine as well as solubilization of the proteins was generally above 90% and no significant differences in the extent of the reaction in the three different types of collagen preparations were observed.

Chromatography of the CNBr Peptides. The CNBr peptides were initially fractionated on 2.5 × 10 cm columns of carboxymethyl (CM)-cellulose equilibrated at 40° with 0.02 M citrate buffer (pH 3.8) containing 0.02 M NaCl. After applying the sample dissolved in 10 ml of the same buffer, the column was eluted with a linear gradient formed between 1000 ml each of 0.02 M sodium citrate–0.02 M NaCl (pH 3.8) and 0.02 M sodium citrate–0.17 M NaCl (pH 3.8) (Butler *et al.*, 1967). The effluent was continuously monitored at 230 mμ in a Gilford spectrophotometer equipped with a flow cell. Fractions of 10 ml were collected.

The small peptides eluting unretarded from the CM-cellulose column (fraction 1 of Figure 1a,b) were desalted on columns of Bio-Gel P-2, 200–400 mesh (Bio-Rad), using 0.1 M acetic acid as the eluent, and fractionated further on columns of phosphocellulose under the conditions described previously (Bornstein and Piez, 1966; Kang *et al.*, 1969b). Briefly, columns of phosphocellulose (2 × 10 cm) were equilibrated with 0.001 M sodium acetate (pH 3.8) at 40° and eluted with a linear gradient of NaCl from 0 to 0.3 M over a total volume of 800 ml.

In some instances, peptide fractions were further chromatographed on CM-cellulose (0.9 × 6 cm column) at pH 4.8 using 0.01 M sodium acetate buffer. A linear gradient of NaCl from 0 to 0.1 M was superimposed over a total volume of 250 ml.

Molecular Weight Determination. Molecular sieve chromatography on a 2 × 110 cm column of Agarose 1.5 M, 200–400 mesh (Bio-Rad), equilibrated with 0.01 M Tris (pH 7.0) and 1.0 M $CaCl_2$, was used for molecular weight determination and purification of the CNBr peptides (Piez, 1968a). The column had been calibrated using separately purified α1-CNBr peptides of known molecular weight (Kang *et al.*, 1969b).

Amino Acid Analysis. Samples were hydrolyzed in constant-boiling HCl under nitrogen at 108° for 24 hr. In some instances, hydrolysis in 2 N NaOH at 108° for 24 hr was also used. The alkaline hydrolysates were applied to the columns after tenfold dilution and adjustment of pH to 2.0 with HCl. No untoward effects on the resolution of amino acids were noted under these circumstances provided the sample volumes were kept small (less than 3 ml).

Analyses were performed on an automatic analyzer (Jeolco 5AH) using a single column method with stepwise elution. The initial buffer was 0.2 N sodium citrate (pH 2.93), containing 3% 1-propanol. After the elution of the acidic amino acids, 3-hydroxyproline, methionine sulfoxide, 4-hydroxyproline, aspartic acid, threonine, and serine in that order, the second buffer, 0.2 N sodium citrate (pH 3.29) was introduced to separate homoserine, glutamic acid, proline, glycine, alanine, cystine, and valine. The neutral amino acids, methionine, isoleucine, leucine, tyrosine, and phenylalanine, were eluted with 0.2 N sodium citrate (pH 4.10). The remaining basic amino acids were then eluted with 0.2 N sodium citrate (pH 4.61) containing 1.0 M NaCl in the order of hydroxylysine, lysine, ammonia, histidine, homoserine lactone, and arginine. The temperature of the column was maintained at 54° throughout. The instrument is equipped with a stream-split device for routine operation. A scale expander allows detection of 0.001 μM of amino acids accurately.

Correction factors for hydrolytic losses of labile amino acids (threonine, serine, methionine, and tyrosine) and incomplete release of valine were used as previously determined for 6 N HCl hydrolysates (Piez *et al.*, 1960).

Radioactivity Determination. Radioactivity was assayed on aliquots of column effluents using Aquasol (New England Nuclear) in an automatic liquid scintillation counter (Packard).

Hydrolysis with Trypsin. Trypsin (TPCK-treated, Worthington) was dissolved in 1 × 10^{-5} M HCl and added to the samples dissolved in 0.2 M NH_4HCO_3 (pH 7.8) containing 1 × 10^{-3} M $CaCl_2$. The reaction was allowed to proceed for 2 hr at 37° and terminated by adjusting the pH of the reaction mixture to 4 and lyophilizing.

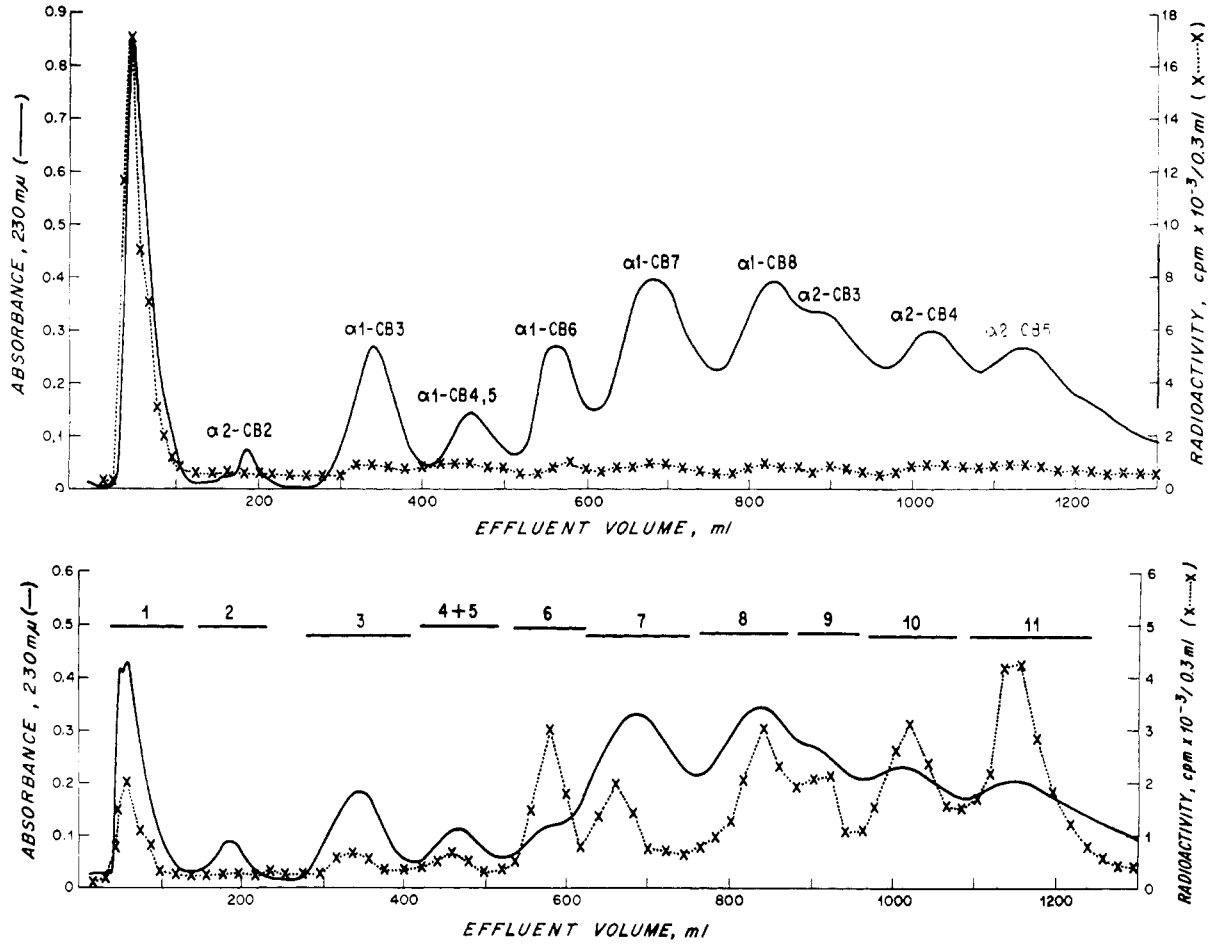

FIGURE 1: CM-cellulose chromatograms of the CNBr digests of extracted, purified rat-tail tendon collagen reduced with sodium borohydride in solution (a, top, 240 mg), and the sodium borohydride-reduced native tendon in solid state (b, bottom, 210 mg). See text for conditions. The unlabelled peak at the beginning of the chromatogram in 1a contains the small NH_2-terminal peptides. The bars in 1b indicate the fractions collected. Aliquots of 0.3 ml were used to determine radioactivity of the effluent fractions.

Results

Chromatography of the CNBr Peptides from Reduced Soluble Collagen. Lyophilized CNBr peptides were initially fractionated on CM-cellulose at pH 3.8. A typical elution pattern of the CNBr digest of reduced soluble collagen is presented in Figure 1a. Satisfactory resolution was achieved except for the mixture of the small NH_2-terminal peptides eluting unretarded at the void volume of the column, α1-CB4 from α1-CB5, and α1-CB8 from α2-CB3. The last two could be further separated by rechromatography on CM-cellulose at pH 4.8 after desalting. The yield of each peptide, estimated by amino acid analysis on known aliquots indicate that each peptide is present within experimental error (Butler *et al.*, 1967) in the expected amount assuming two α1 chains and one α2 chain per molecule. The peptides, α1-CB4 and CB5, were resolved by rechromatography on phosphocellulose (Kang *et al.*, 1969b). Only the first peak contained significant amounts of radioactivity. The intermolecular cross-links, δ-hydroxylysinonorleucine and the posthistidine compound, were not present in any of the peptides as determined by amino acid chromatography and scintillation spectrometry of 6 N HCl hydrolysates. Furthermore, analyses of 2 N NaOH hydrolysates of each peptide fraction showed that ε-hydroxynorleucine is present only in the NH_2-terminal peptides which eluted unretarded from the CM-cellulose column, and none of the other peptides derived from the main body of the molecule contained significant amounts of the compound. These data indicate that the allysyl residues are limited to the NH_2-terminal regions.

The fractions corresponding to peak 1 (Figure 1a) containing the NH_2-terminal peptides were pooled, and desalted on on Bio-Gel P-2. The dipeptide, α1-CB0, and the tripeptide, α2-CB0, were lost during this procedure since they were retarded on the column and eluted unseparated from the salts. The other peptides were fractionated on phosphocellulose under conditions described above. A typical chromatogram is depicted in Figure 2a. The radioactivity profile closely follows the aldehyde-containing peptides (α2-CB1^{Ald}, α1-CB1^{Ald}, β_{12}-CB1, and β_{11}-CB1). Apparently, reduction of the aldehydic residues (allysine and aldol condensate) present in some of the peptides does not alter their chromatographic behavior, and the elution pattern resembles closely those published earlier for unreduced CNBr peptides (Bornstein and Piez, 1966; Bornstein, 1969; Kang *et al.*, 1969b,c). The peptides α1CB1^{Ald} and α2-CB1 do not separate, but the relative amount of each peptide in the peak could be estimated from the amino acid composition of the mixture assuming the presence of one residue of ε-hydroxynorleucine per α1-CB1^{Ald}. The peptides α1-CB2 and β_{11}-CB1 elute together, but are separated on rechromatography on phosphocellulose after tryptic hydrolysis of the mixture. The

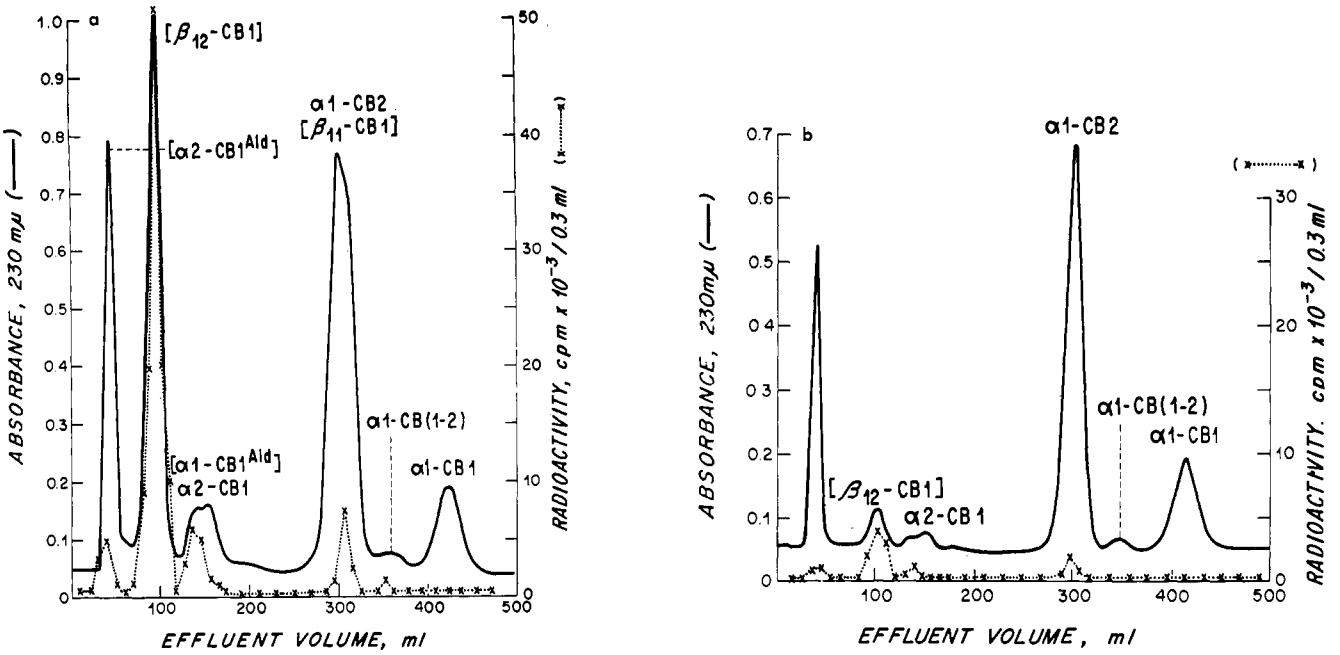

FIGURE 2: Phosphocellulose chromatograms of the fraction 1 obtained from CM-cellulose chromatography (Figure 1a,b). "2a" represents the fraction 1 obtained from 500 mg of the reduced soluble collagen and "2b" the fraction 1 obtained from 480 mg of the reduced native tendon. Brackets are used to indicate that the aldehydic group in the particular peptide has been reduced by sodium borohydride. Note the marked diminution in the amounts of the aldehyde-containing peptides, α1-CB1^{Ald}, α2-CB1^{Ald}, and β_{12}-CB1, in the reduced tendon (2b) as compared with the reduced soluble collagen (2a). Aliquots of 0.3 ml were assayed for radioactivity.

former peptide contains one arginyl residue and is cleaved by trypsin to two fragments, α1-CB2T-1 and -T-2 (Bornstein, 1969; Kang and Gross, 1970b) which chromatograph at positions different from the original whereas β_{11}-CB1 is not altered by trypsin treatment (see Figure 3a). Amino acid analysis and scintillation spectrometry confirmed the presence of the reduced aldol condensate in the intramolecular cross-link peptides, β_{11}-CB1 and β_{12}-CB1 (Kang *et al.*, 1969a; Kang and Gross, 1970a).

Chromatography of the CNBr Peptides from Reduced Reconstituted Fibrils and Tendon. The results obtained from the reduced *in vitro* reconstituted native-type fibrils and native tendon are essentially identical and therefore, the data for the native tendon only are presented here. A representative chromatogram of the CNBr digest of reduced tendon is presented in Figure 1b. As compared to similar patterns obtained for reduced soluble collagen (Figure 1a), it can be seen that the relative sizes of the peaks 1 and 6 are much smaller reflecting the decreased content of the NH_2-terminal and COOH-terminal peptides in the native tendon. In addition, several new radioactive peaks are present which were not present in the digests of reduced soluble collagen. Some broadening of the peaks also is observed. These are probably due to the fact that each peak contains, in addition to the specific CNBr peptide, the same peptides cross-linked intermolecularly to other peptides (see below).

The peptides in peak 1 (Figure 1b) were further fractionated on phosphocellulose after desalting. Figure 2b illustrates the elution pattern. It clearly shows the markedly diminished content of allysine-bearing peptides α2-CB1^{Ald} and α1CB1^{Ald}. The unlabeled peak eluting first from the column is a forepeak of nonpeptide uv-absorbing material and does not contain α2-CB1^{Ald} as determined by molecular sieve chromatography on Bio-Gel P-2. In addition the "intramolecular" cross-link peptides, β_{12}-CB1 and β_{11}-CB1 are similarly decreased. Figure 3b, a phosphocellulose chromatogram of a tryptic hydrolysate of the peak containing α1-CB2 and β_{11}-CB1 obtained from Figure 2b, indicates the virtual absence of the cross-link peptide in the CNBr digest of tendon (and the *in vitro* reconstituted native-type fibrils). This observation further supports the previous suggestion that the intramolecular cross-link, the aldol condensate, is further incorporated in the formation of higher intermolecular cross-links (the

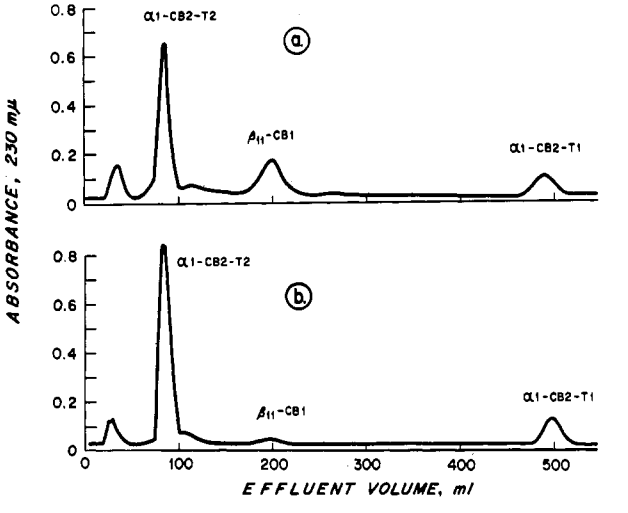

FIGURE 3: Phosphocellulose elution patterns of the tryptic hydrolysates of the mixture of α1-CB2 and β_{11}-CB1 obtained from phosphocellulose chromatography presented in Figure 2. "a" represents the material from the reduced soluble collagen (500 mg) and "b" the reduced native tendon (480 mg). See text for details. Again, note that β_{11}-CB1 is present in markedly decreased amount in the reduced native tendon (b) as compared with the reduced soluble collagen.

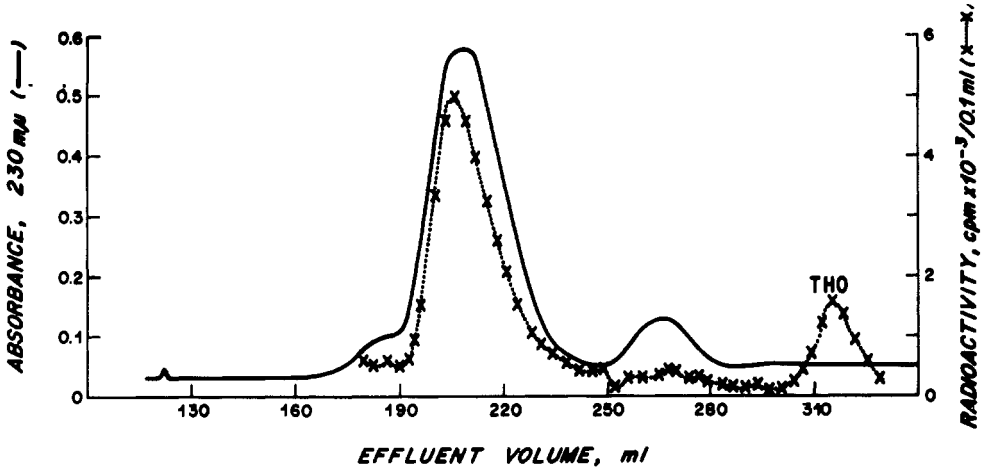

FIGURE 4: Agarose column elution pattern of fraction 6 (Figure 1b). Aliquots of 0.2 ml from each effluent fraction were assayed for radioactivity. Tritiated water (THO) was added to mark the column volume.

posthistidine compound) both *in vitro* and *in vivo* (Franzblau *et al.*, 1970; A. H. Kang and C. Franzblau, in preparation).

The marked diminution of the NH_2-terminal peptides containing the reactive aldehyde residues, allysine and aldol condensate, in the collagen fibrils suggested the possibility that these peptides may become linked to other peptide regions of the neighboring molecules during fibril formation. To examine this idea, the fractions were pooled as indicated by the bars (Figure 1b), purified by molecular sieve chromatography, and after desalting were analyzed for content of the intermolecular cross-links, δ-hydroxylysinonorleucine, and the posthistidine compound. Table I summarizes the results obtained. With the exception of fractions representing relatively small peptides, detectable amounts of the cross-link compounds are present in every fraction. These data strongly suggest that the intermolecular cross-linking involves several sites along the collagen molecule represented by these CNBr peptides. While it cannot necessarily be assumed that the intermolecular cross-link bearing peptides chromatographing under a given peak area are related to the specific CNBr peptide which normally chromatographs in that peak, the fact that the cross-link bearing peptides behave on molecular sieve chromatography in a manner which makes it difficult to separate them from the major homologous parent peptides suggests that they are closely related to each other in molecular weight. If the intermolecular cross-links are derived from linking together the small NH_2-terminal peptides to others, such results would be expected since the size contribution by the NH_2-terminal peptides is relatively small. At the very least it can be stated that the cross-link peptides chromatographing under various peaks differ from each other and must, therefore, derive from different regions.

TABLE I: Content of δ-Hydroxylysinonorleucine and the Posthistidine Compound in Various CNBr Peptide Fractions from the Reduced Native Rat-Tail Tendon.[a]

Fraction	δ-Hydroxylysinonorleucine	Posthistidine Compound
1	0	0
2	0	0
3	0	0
4, 5	0.1	0.2
6	0.6	0
7	0.3	0.1
8	0.3	0.3
9	0.3	0.4
10	0.5	1.0
11	0.4	1.6

[a] Expressed as leucine equivalent per 1000 total residues. The fraction numbers refer to the peptides as shown in Figure 1b.

Quantitative considerations, however, are illuminating. The soluble tail tendon collagen of 7-week-old rats contains 0.35/1000 residues of allysine determined as ε-hydroxynorleucine after borohydride reduction and 2 N NaOH hydrolysis. Native tendon from animals of the same age contains 0.75 leucine equiv of δ-hydroxylysinonorleucine per 1000 residues. Since a residue of the latter contains two amino groups capable of yielding chromaphores absorbing at 570 mμ with ninhydrin, it would indicate approximately half that amount of the cross-link. These findings indicate that essentially all of the δ-hydroxylysinonorleucine present in the native fibril could be accounted for by the allysyl residues located in the NH_2-terminal peptides. The case of the posthistidine compound is more difficult to calculate quantitatively because its chemical structure is unknown. Recent studies, however, suggest that it is derived in part from the aldol condensate and further that it appears to contain four ninhydrin-reactive groups per molecule (Franzblau *et al.*, 1970; A. H. Kang and C. Franzblau, in preparation). Since it has been shown that all of the aldol content of collagen can be accounted for by the peptides β_{11}-CB1 and β_{12}-CB1 (Kang *et al.*, 1970), it follows that the cross-link peptides which contain the posthistidine compound must include these NH_2-terminal peptides. If these indications are correct, tendon collagen of 7-week-old rats, which contains 0.25/1000 residues of aldol condensate or 0.5 leucine equiv/1000 residues, should give rise to approximately 1 leucine equiv/1000 residues of the posthistidine compound. The observed value of the posthistidine compound in 7-week-old rats is 1.1 leucine equiv/1000 residues (A. H. Kang and C. Franzblau, in prepara-

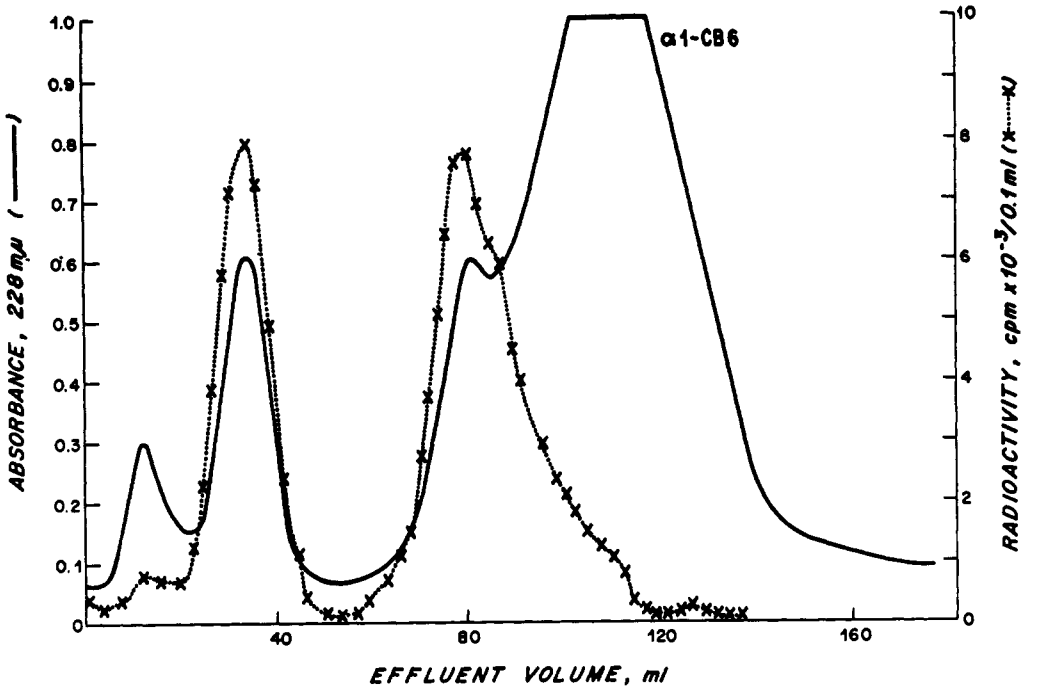

FIGURE 5: CM-cellulose chromatogram of fraction 6 after purification on Agarose. Radioactivity of effluent fractions was determined on aliquots of 0.2 ml.

tion). This type of consideration strongly suggests that indeed most, if not all, of the intermolecular cross-linking involves the NH_2-terminal peptides.

Isolation of the Amino-Carboxyl-Terminal Cross-Link Peptide. Since analyses of the CM-cellulose fractions of the CNBr digest of the reduced native tendon indicated the presence of the cross-link compounds in most fractions, attempts were made to isolate the cross-linked peptides in pure forms. However, due to the complexity of the mixture, it was possible only in the case of fraction 6 (Figure 1b). This was accomplished by molecular sieve chromatography (Figure 4) of the fraction on a calibrated column of Agarose followed by CM-cellulose chromatography of the Agarose fraction at pH 4.8 (Figure 5). The molecular weight as determined by Agarose chromatography, is 18,000 (Figure 6). Two radioactive peaks were observed on CM-cellulose chromatography, along with unaltered α1-CB6. The amino acid composition of these two radioactive peaks was indistinguishable from each other and is presented in Table II along with that of α1-CB6 of rat-tail tendon. α1-$CB1^{Ald}$ is included for comparison. It can be seen that the composition agrees with the sum of that for α1-CB6 and α1-$CB1^{Ald}$. In addition, it contains a residue of δ-hydroxylysinonorleucine. The loss of one residue of hydroxylysine is consistent with formation of δ-hydroxylysinonorleucine. These data together with the molecular weight indicate that it is a peptide intermolecularly cross-linked between the NH_2-terminal α1-$CB1^{Ald}$ and the COOH-terminal α1-CB6. The fortunate circumstance that α1-CB6 contains the one residue of 3-hydroxyproline, one of two histidines and one of three tyrosines, and that α1-$CB1^{Ald}$ contains two of three tyrosines present in the α1 chain strongly suggests the correctness of this interpretation.

Discussion

Past attempts to localize the position of intermolecular cross-links along the polypeptide chains of collagen utilizing proteolytic digestion yielded only limited information because of the difficulty of determining the site of the origin of the relatively small peptides containing the cross-links. de Luque *et al.* (1970), for example, obtained three peptides from

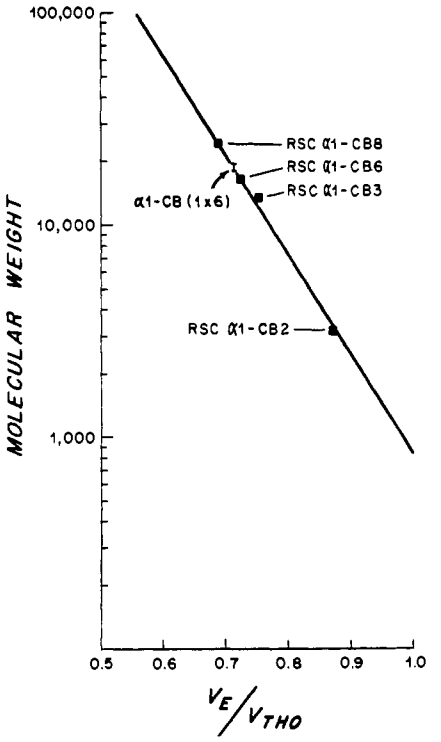

FIGURE 6: Elution data from Agarose 1.5m plotted as a function of log molecular weight. V_E/V_{THO} is the ratio of the elution volume of the peptides to the elution volume of exogenously added tritiated water (THO). The molecular weights of RSC (rat skin collagen) α1-CB8, α1-CB6, α1-CB3, and α1-CB2 had been previously determined by different methods.

TABLE II: Amino Acid Composition of the Purified α1-CB(1 × 6) Intermolecularly Cross-Linked Peptide, α1-CB6, and α1-CB1^{Ald} from Rat-Tail Tendon.[a]

	α1-CB6	α1-CB1^{Ald}	Amino-Carboxyl Cross-Link Peptide
3-Hydroxyproline	1.0		1.0
4-Hydroxyproline	20		19
Aspartic acid	12	1	11
Threonine	4.9		4.5
Serine	12	3	14
Glutamic acid	15	1	17
Proline	33	2	35
Glycine	70	3	75
Alanine	21	1	22
Valine	2.0	2	4.0
Isoleucine	3.2		2.6
Leucine	4.4		4.1
Tyrosine	1.0	2	2.5
Phenylalanine	1.8		2.4
Hydroxylysine	1.9		0.8
Lysine	4.3		4.3
Histidine	1.1		0.9
Arginine	11		11
Homoserine[b]		1	0.9
Hydroxylysinonorleucine			0.9

[a] Residues per peptide. Values are rounded off to the nearest whole number. Where less than ten residues were found, actual values are shown. A dash indicates 0.2 residue or less. Detailed data on α1-CB1^{Ald} has been published previously (Bornstein, 1969). [b] Includes homoserine lactone.

proteolytic hydrolysates of sodium borohydride reduced, reconstituted, calf skin collagen ranging in size from tetrapeptide to septapeptide, but were unable to determine their origin.

In order to circumvent some of these problems, we have employed CNBr cleavage in view of the low methionine content of most mammalian collagens and its successful application to structural studies of soluble collagens and their component polypeptide chains (see Piez *et al.*, 1969). It was reasoned that if insoluble collagen, presumably cross-linked to a greater degree than soluble collagen, and in which the cross-links were stabilized with sodium borohydride reduction, could be degraded in a limited but reproducible manner, a comparison of the products with those obtained from soluble collagen might be expected to reveal differences which could then be attributed to the cross-links. Specifically, the cross-linked CNBr peptides should contain all of the amino acids present in the constituent CNBr peptides as well as a residue of the cross-link compounds.

The results obtained in this study indicate that intermolecular cross-linking occurs at several sites along the length of the molecule, as evidenced by the presence of δ-hydroxylysinonorleucine and/or the posthistidine compound in most of the larger peptide regions on CM-cellulose chromatography of CNBr digests of the reduced native tendon. Furthermore, marked diminution of the aldehyde bearing NH_2-terminal peptides during *in vitro* fibrillogenesis and their absence in the digests of the reduced native tissue suggest that intermolecular cross-linking involves the NH_2-terminal peptides. In addition, quantitative considerations of the amounts of δ-hydroxylysinonorleucine and the posthistidine compound in relation to the amounts of their precursors, allysine and the aldol condenstate present, indicate that all of the intermolecular cross-links formed can be accounted for by the functional groups located in the NH_2-terminal peptides, and does not require the assumption that allysyl residues located at loci other than the NH_2 termini must also contribute towards the formation of δ-hydroxylysinonorleucine. The case for the posthistidine cross-link is equally convincing since its precursor, the aldol condensation product of two residues of allysine, has been localized exclusively to the NH_2 termini (Bornstein and Piez, 1966; Kang *et al.*, 1969c, 1970).

Although it has been reported that the allysyl residues are also present in the CNBr peptides derived from the main body of the molecule in cysteamine-extracted rat skin collagen (Deshmukh and Nimni, 1971), the present study does not confirm this observation. The discrepancy might be related to the differences in the tissue source (skin *vs.* tendon) and the methods of extraction of collagen (cysteamine *vs.* acetic acid) used in the two investigations.

Efforts to purify each cross-linked peptide from the un-cross-linked material were unsuccessful except in the case of fraction 6 (Figure 1 b) due largely to the complexity of the mixtures, and the close similarities in chromatographic properties and molecular weight of the cross-linked and un-cross-linked peptides. Incidentally, these observations are consistent with the above suggestion that intermolecular cross-linking involves the smaller NH_2-terminal peptides. Addition of these small peptides would not be expected to increase the molecular weight significantly enough to separate on molecular sieve chromatography. If cross-linking had involved any combination of two or more of the larger peptides, such as α1-CB7 or α1-CB8 to α2-CB3, -CB4, or -CB5, for example, such cross-linked peptides would have been sufficiently larger than any other peptides to become separable on molecular sieve chromatography. No such peptides were observed. However, it is possible that such very large peptides, especially the ones containing the posthistidine, might have remained uneluted from the CM-cellulose column with the gradient used in the present investigation. This possibility seems less likely since no significant amounts of hydroxyproline-containing material can be eluted from the CM-cellulose on further application of a 0.02 M citrate buffer (pH 3.8) containing 0.5 M NaCl. Perhaps a definitive information on the structure of the posthistidine compound would be helpful in clarifying this point. Another factor operating against the effort to obtain the cross-linked peptides in pure form was the fact that the cross-linked peptide chromatographing under any given region of the chromatogram (Figure 1b), is a very small fraction of the total peptide material present in the peak. In any case, a clear result was obtained in the case of fraction 6. As shown in Figure 5, two radioactive peptides were eluted from CM-cellulose. Both peptides have an identical molecular weight of 18,000 as determined by Agarose chromatography. The amino acid compositions are also indistinguishable. The reason for their separation on ion-exchange chromatography is unexplained. Similar chromatographic heterogeneity of the CNBr peptides were previously observed by several investigators (Butler *et al.*, 1967; Kang *et al.*, 1969b) and were partially ascribed

to the existence of homoserine residues as free acids as well as lactone forms under the conditions of chromatography. On the basis of amino acid composition as shown in Table II, and molecular weight, we conclude that these peptides are the cross-linked peptides between α1-CB1^{Ald} and α1-CB6.

Recently, Miller (1971) also reported isolation from the tritiated $NaBH_4$-reduced chick cartilage a double-chain CNBr peptide containing the CNBr peptides 4 and 9 and a residue of hydroxylysine-derived cross-link. However, since the linear order of the CNBr peptides of cartilage collagen is not known, the location of the cross-linking site could not be determined.

Acknowledgments

The author is grateful to Dr. J. Gross for his support as well as for his critical discussions during this investigation, and to Mrs. J. Hazard and Miss K. Lawley for their excellent technical assistance.

References

Bailey, A. J., and Peach, C. M. (1968), *Biochem. Biophys. Res. Commun. 33*, 812.

Bailey, A. J., Peach, C. M., and Fowler, L. J. (1970), *Biochem. J. 117*, 819.

Bornstein, P. (1969), *Biochemistry 8*, 63.

Bornstein, P. (1970), *Amer. J. Med. 49*, 429.

Bornstein, P., and Piez, K. A. (1966), *Biochemistry 5*, 340.

Butler, W. T., Piez, K. A., and Bornstein, P. (1967), *Biochemistry 6*, 3771.

de Luque, O., Mechanic, G., and Tanzer, M. L. (1970), *Biochemistry 9*, 4987.

Deshmukh, K., and Nimni, M. E. (1971), *Biochemistry 10*, 1640.

Franzblau, C., Kang, A. H., and Faris, B. (1970), *Biochem. Biophys. Res. Commun. 40*, 437.

Gross, J., and Kirk, D. (1958), *J. Biol. Chem. 233*, 355.

Kang, A. H., Faris, B., and Franzblau, C. (1969a), *Biochem. Biophys. Res. Commun. 36*, 345.

Kang, A. H., Faris, B., and Franzblau, C. (1970), *Biochem. Biophys. Res. Commun. 39*, 175.

Kang, A. H., and Gross, J. (1970a), *Proc. Nat. Acad. Sci. U. S. 67*, 1307.

Kang, A. H., and Gross, J. (1970b), *Biochemistry 9*, 796.

Kang, A. H., Piez, K. A., and Gross, J. (1969b), *Biochemistry 8*, 1506.

Kang, A. H., Piez, K. A., and Gross, J. (1969c), *Biochemistry 8*, 3648.

Levene, C. I., and Gross, J. (1959), *J. Exp. Med. 110*, 771.

Miller, E. J. (1971), *Biochem. Biophys. Res. Commun. 45*, 444.

Piez, K. A. (1968a), *Anal. Biochem. 26*, 305.

Piez, K. A. (1968b), *Annu. Rev. Biochem. 37*, 547.

Piez, K. A., Bladen, H. A., Bornstein, P., Butler, W. T., Kang, A. H., Lane, J. M., and Miller, E. J. (1969), *Brookhaven Symp. Quant. Biol. 21*, 345.

Piez, K. A., Weiss, E., and Lewis, M. S. (1960), *J. Biol. Chem. 235*, 1897.

Rojkind, M., Gutiennez, A. M., Zeichner, M., and Lent, R. W. (1969), *Biochem. Biophys. Res. Commun. 36*, 350.

Tanzer, M. L. (1968), *J. Biol. Chem. 243*, 4045.

Tanzer, M. L., and Mechanic, G. (1968), *Biochem. Biophys. Res. Commun. 32*, 885.

Tanzer, M. L., and Mechanic, G. (1970), *Biochem. Biophys. Res. Commun. 39*, 183.