Cyanogen Bromide Peptides from Insoluble Skin and Dentin Bovine Collagens†

Dino Volpin‡ and Arthur Veis*

ABSTRACT: Purified insoluble bovine corium (IC) and dentin (DC) collagens were degraded with CNBr. Neither collagen was totally soluble in the acidic digestion mixture, even after prolonged or repeated digestion. The soluble peptides were resolved by a combination of ion-exchange and gel filtration chromatographic procedures and compared with the peptides previously obtained from the isolated $\alpha 1$ and $\alpha 2$ chains from acid-soluble (SC) bovine corium or with β_{12} from the same source. The majority of peptides from IC and DC could be matched with the SC CNBr peptides. However, three distinct new classes of peptides were detected in the IC and DC digests. Both contained an acid-stable intermolecular cross-link component, $\alpha 1$-CB6′[$\alpha 1$-CB6 + $\alpha 1$-CB(0,1)]. Both also contained two additional peptides attributed to the presence of an $\alpha 1$(III) chain, $\alpha 1$(III)-CB3, and $\alpha 1$(III)-CB(4,5). While the $\alpha 1$(I) peptides of IC and DC were quite similar except for degree of hydroxylation, the $\alpha 1$(III) peptides showed more pronounced tissue specific differences. Evidence was also obtained indicating the presence of a peptide, $\alpha 2$-CB4′, consisting of $\alpha 2$-CB4 and a noncollagenous polypeptide attachment. The $\alpha 2$-CB4′ was present in both IC and DC but was not detected in SC. Each of the CNBr peptides common to both soluble and insoluble collagens was analyzed for hexose. In IC, $\alpha 1$(I)-CB5, $\alpha 2$-CB4, and $\alpha 2$-CB5 contained disaccharide units and $\alpha 1$(I)-CB6 contained a monosaccharide. In DC, only $\alpha 1$(I)-CB5 contained disaccharide, while $\alpha 1$(I)-CB3 and $\alpha 1$(I)-CB6 contained monosaccharide units. The DC $\alpha 2$-CB4 and $\alpha 2$-CB5 contained mixtures of mono- and disaccharide attachments. In addition to the hexose, DC $\alpha 1$(I)-CB6, $\alpha 2$-CB4, and $\alpha 2$-CB5 all contained covalently bound phosphate groups. The acid-insoluble CBNr digestion residues contained portions of collagen with uncleaved methionyl residues in close association, particularly in DC, with a highly acidic polypeptide. In DC, this polypeptide was rich in phosphoserine groups. These data indicate that the insoluble collagens do differ in several ways from acid-soluble collagen and it is suggested that these differences relate to the tissue-specific fibril organization as well as to the presence of intermolecular cross-linkages.

It is generally considered that the insoluble collagens differ from soluble collagens from the same tissue essentially in terms of the covalent intermolecular cross-linkages between collagen monomers. Other equally important but less well documented possibilities for differences lie in quantitative variations in degree of glycosylation or in the extent of covalent association with non-collagenous peptides or other moieties. At a still more fudamental level, the constituent α chains may be qualitatively genetically distinct (Miller *et al.*, 1971). In a preliminary report, Volpin and Veis (1971a) demonstrated that insoluble bovine skin and dentin collagens provided examples of such differences between soluble and insoluble bovine skin collagens. The present paper presents a detailed comparison of the CNBr peptides obtained from insoluble bovine skin and dentin collagens with those CNBr peptides isolated from soluble bovine skin collagens (Volpin and Veis, 1971b; Rauterberg and Kühn, 1971) as well as with each other.

Materials and Methods

Bovine Skin Collagen. The steer skin collagen preparation used by Volpin and Veis (1971b) for the isolation of soluble collagen (SC) α chains was utilized in the present study so that the results are directly comparable. Insoluble collagen (IC) was the residue left after successive extractions with neutral salt and four extractions with 0.5 N acetic acid. The β_{12} component was obtained from the same CM-cellulose chromatographic separations of denatured SC used to isolate the $\alpha 1$ and $\alpha 2$ chains (Volpin and Veis, 1971b).

Bovine Dentin Collagen. Dentin collagen was prepared from unerupted bovine teeth following the procedure of Veis and Schlueter (1964). The unerupted teeth were removed from fresh cattle jaws. The pulp and its processes were extracted and the adhering soft tissues and cementum were removed by scraping. Following several washes in 15% NaCl, the teeth were minced, then broken into small pieces and extacted repeatedly at 4° with 0.5 M EDTA, adjusted to pH 7.4 with NaOH. The teeth were shaken vigorously during the extraction. The brittle enamel layer flakes off the softer and more flexible dentin and the soluble enamel proteins are dissolved, followed by the soluble components of the dentin. Extraction was continued until the calcium content was reduced below the limit of detection. The soft, white pieces of dentin collagen (DC) were washed with distilled water to remove all EDTA, then lyophilized for storage. The dentin at this stage is completely insoluble in cold dilute organic acids or in strong solutions of lyotropic agents.

Cleavage with CNBr. Samples of SC, IC, and DC weighing 400 mg were immersed in 50 ml of 70% formic acid, flushed with nitrogen. A 200-fold molar excess (relative to

† From the Northwestern University Medical School, Chicago, Illinois 60611. *Received June 2, 1972.* This study was supported by Grants AM-13921 and DE-01374 from the National Institute of Arthritis and Metabolic Diseases and the National Institute of Dental Research.

‡ Holder of a North Atlantic Treaty Organization Postdoctoral fellowship. Present address: Institute of Histology, University of Padova, 35100 Padova, Italy.

[1] Abbreviations used are: IC, SC, and DC, insoluble, soluble, and dentin collagens.

methionine) of CNBr was added (Epstein *et al.*, 1971). The digestion was carried out at room temperature (24°) for 4 hr. The reaction was terminated by diluting the digestion mixture with water. The mixture was then lyophilized. A sample of 200 mg of β_{12} obtained by chromatography of acid-soluble collagen on CM-cellulose (Volpin and Veis, 1971b) was subjected to the same procedure.

After lyophilization, the digest was dissolved in 0.15 M acetic acid, warmed briefly at 40° and filtered while still warmed through a 1.2 μ pore size Millipore filter, or medium sintered-glass filter. In the case of dentin, a substantial residue was left on the filter. This residue was digested for a second time in CNBr under the same conditions as above for another 4 hr. After lyophilization, the fraction soluble in 0.15 M acetic acid was mixed with the filtrate from the first digestion.

Separation of the CNBr Peptides. Several steps in separation and purification were used in each case to isolate individual CNBr peptides. All peptides were obtained as homogeneous preparations before amino acid analyses were carried out. The first step was to pass the mixture of peptides over a Bio-Gel P-2 column which separated out the smallest peptides. The next stage of fractionation, which spread the remaining peptides quite well, was on CM-cellulose at pH 3.6. In most cases, each pH 3.6 peak was rechromatogramed on CM-cellulose at pH 4.8. This step was followed by preparative molecular sieve chromatography. The homogeneity of the final peptides was examined by acrylamide gel electrophoresis. The details of each procedure are given below.

CM-CELLULOSE CHROMATOGRAPHY. The soluble CNBr digests were chromatographed on a 2.5 × 12 cm column of CM-cellulose equilibrated at 41° with 0.02 M sodium citrate (pH 3.6). Elution was carried out with a linear gradient of NaCl from 0 to 0.14 M over a volume of 2000 ml. Fractions comprising a given peak were combined, desalted, and lyophilized.

Peptides isolated at pH 3.6 were rechromatographed using CM-cellulose equilibrated with 0.03 M sodium acetate (pH 4.8). Chromatography was performed using a variable linear gradient of NaCl.

CHROMATOGRAPHY OF SMALL CNBr PEPTIDES ON PHOSPHOCELLULOSE. Columns of phosphocellulose were equilibrated with 0.001 M sodium formate buffer (pH 3.6) at 41°. Chromatography was performed by applying a linear gradient of NaCl from 0 to 0.3 M ionic strength over a total volume of 2000 ml. The isolated peptides were then desalted and lyophilized.

MOLECULAR SIEVE CHROMATOGRAPHY. Further fractionation of some peptides was performed on a 2.5 × 70 cm column of Bio-Gel P-4 (200–400 mesh) equilibrated with 0.01 M acetic acid or Bio-Gel P-30 in the same solvent. Gel filtration on Bio-Gel A-1.5 (200–400 mesh) equilibrated with 1 M $CaCl_2$–0.05 M Tris-HCl (pH 7.5) using either a column of 2.5 × 75 or 2.5 × 150 cm, was used both as a final purification step for some peptides and as a means of determining the molecular weights of these peptides.

DESALTING ON BIO-GEL P-2. Lyophilized fractions containing salts were desalted at room temperature on a 4 × 40 cm column of Bio-Gel P-2 equilibrated with 0.1 M acetic acid.

ACRYLAMIDE GEL ELECTROPHORESIS. The homogeneity of some fractions was verified by acrylamide gel electrophoresis as described by Volpin and Veis (1971b) or by the method of Davis (1964).

Amino Acid Analysis. Samples were hydrolyzed in 2 ml of constant-boiling HCl at 108° for 22 hr under nitrogen in sealed tubes. The acid was removed under vacuum at about 50° on a Buchler Evapo-Mix. The dried samples were dissolved in water and a volume containing 0.2–1 mg was used for amino acid analysis on a two-column automatic Jeolco analyzer. No corrections were made for the possible partial destruction or incomplete release of individual amino acids due to hydrolysis conditions. To convert micromoles into residues per peptide, it was assumed that each peptide (with the exception of the COOH terminal) would contain one residue of homoserine. The amount of homoserine was determined in calibration runs by making the pH of the hydrolyzed sample alkaline and then changing to acid just before injecting the sample onto the column. In this way all the homoserine lactone is transformed to homoserine.

The method of Woessner (1967) was used for hydroxyproline analyses when total amino acid compositions were not determined.

Hexose Analysis. Qualitative identification of the hexoses was achieved by thin-layer chromatography (tlc) in 1-butanol–pyridine–water (6:4:3, v/v) after hydrolysis of an aliquot of material in 2 ml of 2 N HCl under N_2 for 2 hr at 110° (Butler and Cunningham, 1966; Miller, 1971). Hydrolysates representing 5 mg of material were applied at room temperature to a 1 × 20 cm column of Bio-Rad AG 50 W-X4 and two column volumes of water were used for elution. The eluent was lyophilized and redissolved in water and a part was applied along with standards to a tlc sheet. After development, the chromatograms were sprayed with 5% $AgNO_3$ for detection of hexoses.

The protein-bound hexoses were measured quantitatively by the orcinol–sulfuric acid method of Winzler (1955) and expressed as equivalents of galactose.

Phosphate Analyses. Phosphate determinations were made following the procedure of Lucena-Conde and Prat (1957).

Hydroxylysine-Linked Hexoses. PERIODATE OXIDATION. Test samples were oxidized at pH 3.8 with 0.025 M periodic acid at room temperature for 8 hr in the dark (Bulter and Cunningham, 1966). The excess periodate was removed by gel filtration on Bio-Gel P-2 and the samples were hydrolyzed under N_2 in 2 ml of 6 N HCl at 108° for 24 hr. Aliquots were then examined on the amino acid analyzer for their residual content of hydroxylsine. This periodate-resistant hydroxylsine was presumed to be hexose bound.

ALKALINE HYDROLYSIS. Samples were hydrolyzed under N_2 in alkali-resistant tubes (Corning 7280) in 2 ml of 2 N NaOH at 108° for 24 hr (Spiro, 1969) in order to hydrolyze peptide bonds but preserve O-glycosidic linkages. The samples were neutralized with HCl and aliquots were analyzed for free hydroxylysine.

Results

Neither IC, after 4 hr of treatment with CNBr, nor DC, after 8 hr of the same treatment, went entirely in solution (Volpin and Veis, 1971a). Only 85% of IC was digested and solubilized by the CNBr digestion. The double CNBr digestion of DC yielded 92% recovered as soluble peptides.

The amino acid compositions of IC, the whole CNBr extract from IC, DC, and the whole CNBr extract from DC are compared in Table I. The extracts have compositions quite close to the starting materials and could thus be taken as reasonably representative of the total insoluble material. However, as is evident below, these fractions do not completely represent the insoluble matrices from which they are derived.

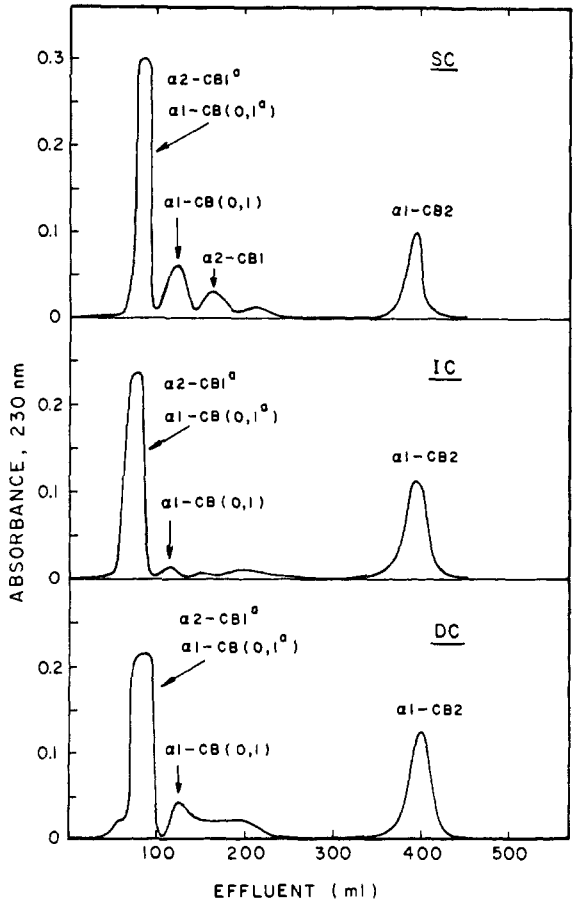

FIGURE 1: Phosphocellulose chromatograms of the CNBr peptides from acid-soluble bovine skin collagen (SC), -insoluble skin collagen (IC), and dentin collagen (DC). Elution was achieved with a linear gradient from 1000 ml of starting buffer (0.001 M sodium formate, pH 3.6) and 1000 ml of limit buffer (starting buffer containing 0.3 M NaCl). Temperature was 41°.

TABLE I: Compositions of Insoluble Bovine Skin and Dentin Collagens and the Soluble Peptides Resulting from CNBr Degradation.[a]

	Residues/1000 Total Residues			
	IC		DC	
	Intact	CNBr Peptides	Intact	CNBr Peptides
4-Hydroxyproline	94	95	102	101
Aspartic acid	46	45	46	46
Threonine	17	16	17	16
Serine	36	34	37	35
Glutamic acid	72	72	72	73
Proline	127	127	118	120
Glycine	326	331	332	337
Alanine	112	112	109	105
Valine	23	22	20	20
Isoleucine	11	10	10	10
Leucine	25	24	25	25
Tyrosine	3	3	4	3
Phenylalanine	13	13	13	14
Hydroxylysine	6	6	10	10
Lysine	27	27	23	23
Histidine	5	5	5	5
Arginine	51	51	51	51
Methionine[b]	6	–	6	–
Homoserine[c]	–[d]	6	–	6
(Phosphate)	–	–	(6.2)	(2.3)
Total	1000	1000	1000	1000

[a] Residues rounded off to nearest whole number. [b] Includes methionine sulfoxides. [c] Includes homoserine lactone. [d] Less than 0.2 residue.

CNBr Solubilized Peptides. α2-CBO. The soluble peptides obtained from the cleavage of IC and DC were first chromatographed on Bio-Gel P-2. Analysis of the effluent from the included volume of the column (Click and Bornstein, 1970) revealed traces of several amino acids but in each case, it contained about 8- to 10-fold larger amounts of glycine, leucine, and homoserine in 1:1:1 ratio. This corresponds to the peptide designated α2-CBO in SC (Volpin and Veis, 1971b).

α1-CB(0,1), α2-CB1. The peptides collected from the void volume of the Bio-Gel P-2 column were chromatographed on phosphocellulose and CM-cellulose. Figure 1 compares the phosphocellulose chromatograms for the SC, IC, and DC digests. Few qualitative differences are evident. However, the peptides α1-CB(0,1) and α2-CB1 are present in decreased amounts in the insoluble collagens. In fact, it was difficult to detect α2-CB1 in either IC or DC.

The first peak, labeled α2-CB1[a] and α1-CB(0,1[a]) in Figure 1, was not separated into its component peptides but, the peak material, in line with earlier studies (Volpin and Veis, 1971b; Click and Bornstein, 1970), had a composition typical of a mixture of α2-CB1[a] and α1-CB(0,1[a]). Very little peptide material could be recovered from this first peak in spite of its high absorbance at 230 nm. The high absorbance is apparently artifactual and is due to some uv absorbant nonpeptide material present in the digests.

CM-CELLULOSE CHROMATOGRAPHY. The peptides from β_{12}, IC, and DC were subjected to CM-cellulose chromatography with the results shown in Figure 2 (see also Volpin and Veis, 1971a,b). The chromatograms are qualitatively similar in each case except for the appearance of peaks 3 and 4 in IC and DC. Quantitative differences in absorbancies of peaks 6, 7, 10, and 11 are evident. The identifications of the peaks indicated in Figure 2 were made on the basis of the compositions of the component peptides following the isolation procedures indicated in subsequent paragraphs.

The major components of every peak except peak 1 in Figure 2 were rechromatogramed on CM-cellulose with 0.03 M sodium acetate (pH 4.8) buffer and raising the ionic strength of NaCl in the limit buffer from 0.08 to 0.20 M, depending on the molecular weights of the peptides to be separated. Further purification of each rechromatographed peptide was achieved by Agarose gel fitration chromatography.

In the majority of cases, each CM-cellulose peak contained only one major peptide, and as indicated above, all of the peptides found to be present in the isolated α1, α2, and β_{12} chains from SC were also found to be present among the IC and DC CNBr peptides.

Peaks 2, 8, and 9 yielded α2-CB2, α1-CB7, and α1-CB8, identical in each preparation, after purification. Because of difficulty in cleaving the methionyl residue between α2-CB3 and α2-CB5, we did not isolate very large amounts of these

TABLE II: Compositions of α1-CB3 from CM-cellulose Peaks 4 and 5 from SC, IC, and DC Collagens.

	Residues/Peptide[a]				
	SC	Insoluble Skin Collagen		Dentin Collagen	
	α1(I) Peak 5	α1(I) Peak 5	α1(III) Peak 4	α1(I) Peak 5	α1(III) Peak 4
4-Hydroxyproline	14	14	14	15	14
Aspartic acid	6.5	6.5	6.2	6.4	6.3
Threonine	–[c]	–	2.3	–	2.1
Serine	2.8	2.8	2.8	2.8	4.3
Glutamic acid	16	16	12	15	11
Proline	16	16	19	15	18
Glycine	51	51	51	51	51
Alanine	22	22	15	23	14
Valine	3.8	3.8	2.8	3.9	3.6
Isoleucine	–	–	1.3	–	1.4
Leucine	3.0	3.0	2.7	3.4	3.8
Tyrosine	–	–	–	–	–
Phenylalanine	3.0	3.0	1.3	2.6	3.3
Hydroxylysine	0.3	0.3	0.4	1.4	1.4
Lysine	4.8	4.8	6.0	3.9	4.5
Histidine	–	–	–	–	–
Arginine	6.0	6.0	4.2	6.0	6.4
Homoserine[b]	1.0	0.9	0.9	0.9	1.0
Total	151	150	141	150	145

[a] In this and succeeding tables, the amino acids present as 10 residues or more are rounded off to the nearest whole number. Actual values are given for those with less than 10 residues. [b] Includes homoserine lactone. [c] Less than 0.2 residue.

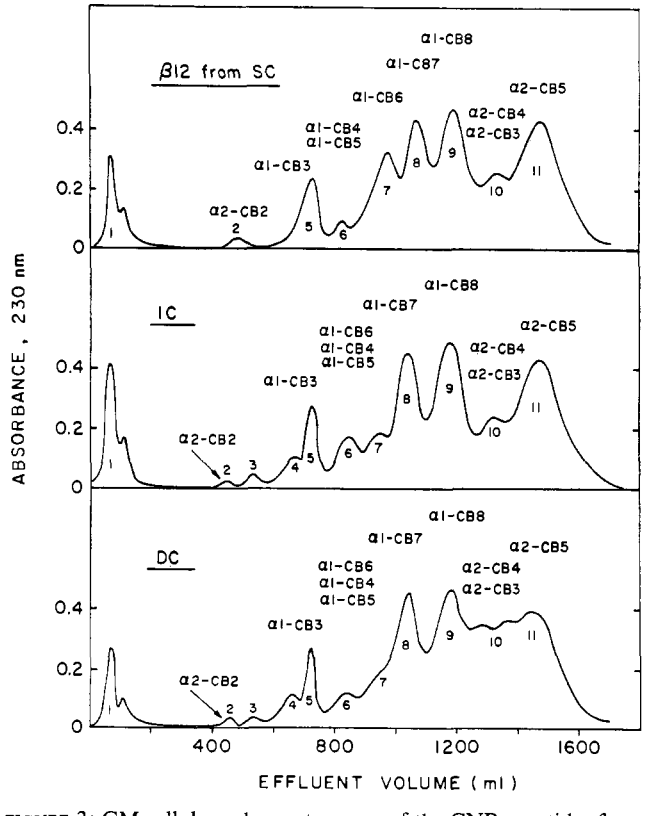

FIGURE 2: CM-cellulose chromatograms of the CNBr peptides from β_{12} of SC, from IC, and from DC. Elution was with a linear gradient from 1000 ml of starting buffer (0.02 M sodium citrate, pH 3.6) and 1000 ml of starting buffer containing 0.14 M NaCl. Temperature was 41°.

two peptides from the insoluble collagens. Instead, the uncleaved peptide α2-CB(3-5) was isolated from peak 11. Additional complications were found in the analyses of CM-cellulose peaks 4, 5, 6, 7, and 10.

α1-CB3 AND RELATED PEPTIDES. Peak 5, identified as α1-CB3 in isolated α1-chains from SC (Volpin and Veis, 1971b) proved to contain a single major homogeneous component

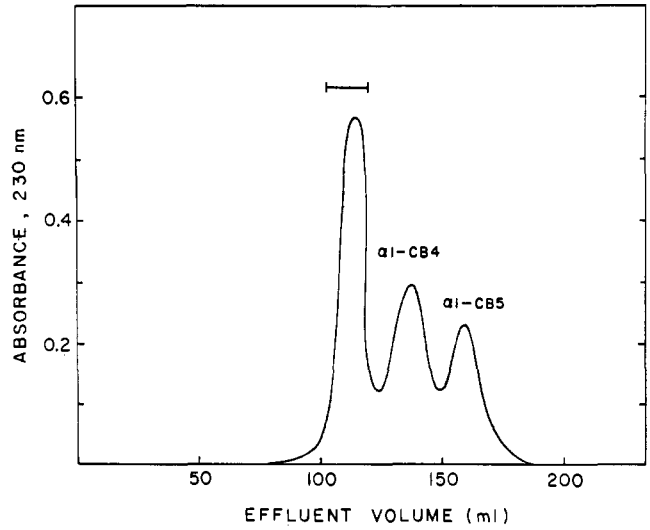

FIGURE 3: Gel filtration of peptides 6 and 7 from IC or DC on a 2.5 × 70 cm column of Bio-Gel P-4. Elution was with 0.1 M acetic acid. Flow was 30 ml/min. The fraction under the bar was collected for further chromatography.

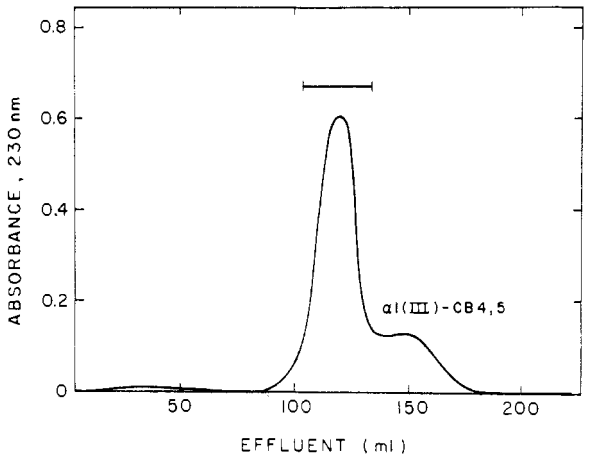

FIGURE 4: Gel filtration of the fraction under the bar of Figure 3 on a 2.5 × 70 cm column of Bio-Gel P-30. Elution with 0.1 M acetic acid. The fraction under the bar was collected for further chromatography.

upon rechromatography and gel fitration of both IC and DC peptides. Peak 4, also isolated as a homogeneous peptide upon rechromatography, had a composition very similar to that of the peak 5 α1-CB3 from IC and DC but, as presented in Table II, showed some distinct differences in composition. Miller *et al.* (1971) found that human skin contained two α1-CB3 peptides which appeared to be derived from genetically distinct α1 chains, designated α1(I) and α1(III). As indicated in the preliminary communication (Volpin and Veis, 1971a), the peak 4 peptides from the bovine collagens also correspond to an α1(III)-CB3.

Two features are immediately evident upon comparing the composition data for the purified peak 4 and 5 components shown in Table II. Except for the difference in degree of hydroxylation of both lysine and proline, the peak 5 components of SC, IC, and DC are essentially identical, containing 150–151 amino acid residues. However, the peak 4 components of IC and DC are not identical. In both cases, the peak 4 α1(III)-CB3 contains two additional residues of threonine and one residue of isoleucine as distinguishing characteristics, along with sharply lower amounts of both glutamic acid and alanine and elevated amounts of proline. The α1(III)-CB3 from DC differs from that of IC principally in having two extra residues each of serine and arginine.

α1-CB4, α1-CB5, α1-CB6 AND RELATED PEPTIDES. The peptide α1-CB6, found in CM-cellulose peak 7 (Figure 2) of β_{12} was sharply reduced in amount in both IC and DC and peaks 6 (containing α1-CB4 and α1-CB5 in isolated α1 chains) and 7 appeared to be considerably overlapped. Therefore, the material in these was combined in each case and subjected to chromatography on Bio-Gel P-4. This yielded a chromatogram (Figure 3) in which α1-CB4 and α1-CB5 were nicely separated. The major peak at the column void volume, the region under the bar in Figure 3, was put through a Bio-Gel P-30 column (Figure 4). A small amount of a new peptide was separable from the void volume peak. This new peptide was chromatogramed preparatively and analytically on the Bio-Gel A-1.5 columns. The peptide, from both IC and DC had a molecular weight of 8000.

Amino acid analyses of the Agarose-isolated peptides showed these new peptides to have a close homology to a combined uncleaved α1-CB(4,5) but with one methionine missing (Table III) and no extra homoserine. In accordance

TABLE III: Amino Acid Composition of α1-CB4 and α1-CB5 Peptides in IC and DC (Peaks 6 and 7).

	Residues/Peptide			
	Insoluble Collagen		Dentin Collagen	
	α1(I)-CB4 + α1(I)-CB5	α1(III)-CB(4,5)	α1(I)-CB4 + α1(I)-CB5	α1(III)-CB(4,5)
4-Hydroxyproline	8.5	8.7	8.6	8.6
Aspartic acid	5.7	3.1	5.6	3.6
Threonine	0.9	0.9	1.0	1.4
Serine	1.7	3.1	1.7	2.9
Glutamic acid	6.0	5.7	5.6	6.4
Proline	8.0	9.2	8.1	9.2
Glycine	28	28	28	28
Alanine	7.2	8.9	6.8	9.0
Valine	–[a]	2.1	–	2.0
Isoleucine	–	0.9	–	0.9
Leucine	2.8	2.4	3.0	2.4
Tyrosine	–	–	–	–
Phenylalanine	0.9	0.8	0.8	0.9
Hydroxylysine	1.2	0.3	1.6	0.6
Lysine	3.4	2.1	3.3	2.1
Histidine	0.9	0.3	0.8	0.3
Arginine	4.9	4.0	5.0	4.0
Homoserine[b]	2.0	1.0	2.0	1.0
Methionine	–	–	–	–
Total	83	81	83	83
Molecular weight by Agarose chromatography		~8000		~8000

[a] Less than 0.2 residue. [b] Includes homoserine lactone.

with our finding of α1(III)-CB3 in each case, and Miller *et al.*'s (1971) finding of an α1(III)-CB(4,5) in human skin collagen, we can conclude that this new peptide also derives from the same genetically distinct α1(III) and identify it as α1(III)-CB(4,5). In addition to the methionine difference, the distinctive features of α1(III)-CB(4,5) are the appearance of two residues of valine and one of isoleucine, an extra residue of serine and a reduction of two residues of aspartic acid. There are two residues fewer of lysine in α1(III)-CB(4,5) than in α1(I)-CB(4,5) and one of these is in a position which in α1(I) is hydroxylated to hydroxylysine.

The remaining material under the bar in Figure 4, which should have been α1-CB6, was not homogeneous. Hence, it was rechromatogramed on a CM-cellulose column, equilibrated with 0.03 M sodium acetate buffer (pH 4.8) at 41° (Figure 5). With both DC and IC, two major distinct peaks were apparent, designated α1-CB6 and α1-CB6′. Material was taken from the central regions of the α1-CB6 and α1-CB6′ peaks and chromatogramed on the Bio-Gel A-1.5 columns. The α1-CB6′ emerged slightly ahead of the position of the α1-CB6 peak in each case, indicating a higher molecular weight. The α1-CB6 peaks yielded a molecular weight of ~17,000 for both IC and DC peptides.

TABLE IV: Analyses of Peptides Related to α1-CB6, Residues/Peptide.

	Insoluble Skin Collagen					Dentin Collagen				
	α1-CB6	α1-CB6′	Δ[a]	α1-CB(0,1)	Δ′[b]	α1-CB6	α1-CB6′	Δ[a]	α1-CB(0,1)	Δ′[b]
3-Hydroxyproline	0.7	0.7	–	–	–	1.1	1.1	–	–	–
4-Hydroxyproline	20	18	(2)[c]	–	–	19	19	–	–	–
Aspartic acid	8.7	13	4.3	1	3.3	8.6	11	2.4	1.0	1.4
Threonine	3.8	5.2	1.4	0.9	0.5	4.0	4.8	0.8	0.8	–
Serine	7.4	10	2.6	3.0	–	7.1	10	2.9	2.7	–
Glutamic acid	13	16	3	2	1	13	16	3	1.9	1.1
Proline	26	27	1	2	–	27	28	1	1.9	(0.9)[c]
Glycine	64	68	4	3.6	–	64	68	4	3.7	–
Alanine	22	23	1	–	1	22	22	–	–	–
Valine	3.2	4.0	0.8	1.1	–	3.3	3.8	0.5	1.2	(0.7)[c]
Isoleucine	2.1	3.0	0.9	1.0	–	1.9	2.8	0.9	0.8	–
Leucine	4.0	8.1	4.1	1.0	3	4.0	5.9	1.9	1.0	0.9
Tyrosine	–[e]	0.8	0.8	1.6	(0.8)[c]	–	2.0	2.0	1.6	0.4
Phenylalanine	1.9	2.6	0.7	–	0.7	1.8	2.7	0.9	–	0.9
Hydroxylysine	0.7	1.0	0.3	0.2	–	1.4	2.2	0.8	0.5	0.3
Lysine	5.5	6.7	1.2	0.7	0.5	4.9	4.4	(0.5)[c]	0.3	(0.8)[c]
Histidine	0.7	1.6	0.9	–	0.9	0.8	1.9	1.1	–	–
Arginine	11	12	1.0	–	1.0	11	11	–	–	–
Homoserine[d]	–	1.0	1.0	1.0	–	–	1.0	1.0	1.0	–
Total	195	222	29	19.1	12	195	218	23	18.4	5.0

[a] Δ = (α1-CB6′) − (α1-CB6). [b] Δ′ = Δ − [α1-CB(0,1)]. [c] Parentheses indicates negative quantity. [d] Includes homoserine lactone. [e] Less than 0.2 residue.

The amino acid analyses of the components designated as α1-CB6 and α1-CB6′, as isolated from the final preparative Agarose rechromatography, are shown in Table IV. The α1-CB6 from IC and DC, and the α1-CB6 isolated from the separated α1 chains of SC (Volpin and Veis, 1971b) are very similar in composition, in each case no homoserine or homoserine lactone could be detected substantiating the COOH-terminal position of this peptide. However, one residue of homoserine was detected per mole of α1-CB6′ in the peptides from both IC and DC. Since, according to the elution position on Bio-Gel A-1.5 columns, α1-CB6′ is larger than α1-CB6 in each case by about 10%, and since there is no small peptide of about 20 residues contiguous with α1-CB6 in α1 chains, it appears that α1-CB6′ is a cross-link containing peptide. The most likely partner for α1-CB6 in this cross-link is α1-CB(0,1) from another end-overlapped α1 chain. The compositions of α1-CB(0,1) from IC and DC are shown in Table IV along with the compositions of α1-CB6 and α1-CB6′. As indicated, there is a good correspondence between the sum of α1-CB6 and α1-CB(0,1) and the total composition of α1-CB6′ in both cases. However, there is a discrepancy of 5–12 residues which remains in each case which leads to the conclusion that either there is an α1-(III)-CB6 which has been isolated, or that since there is no extra glycine, there is a COOH-terminal extension on α1-CB6 which is preserved in the cross-linked peptide α1-CB6′. Rauterberg *et al.* (1972) have shown that the COOH-terminal region is an intermolecular cross-linkage site in calf skin collagen extracted from skin with 8.0 M urea. They also showed that this α1-CB6 has a small peptide extension in which the repeating Gly-X-Y sequence is not maintained. The cross-linking lysine derived aldehyde is present in this nonhelical segment. The nonhelical region is present to a variable extent in different preparations. According to Rauterberg *et al.* (1972), this region is usually missing in acid-soluble collagen but is present to some extent in urea-soluble and -insoluble collagen.

Most of the CNBr peptides were recovered in yields proportional to their expected values, *e.g.*, α1-CB3 represents 8–9% of the initial collagen. The combined amount of α1-CB6 and α1-CB6′ recovered amounted to only 5–6% of the CNBr digest as compared with a theoretical 13–14% value. It is evident in Figure 5 that additional components other than those identified here as α1-CB6 and α1-CB6′ are present in combined peaks 6 and 7. These may be related

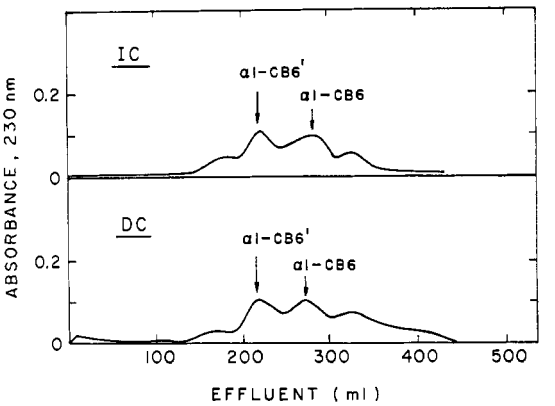

FIGURE 5: CM-cellulose chromatograms of the fraction under the bar of Figure 4. Elution was achieved with a linear gradient from 400 ml of starting buffer, 0.03 M sodium acetate (pH 4.8) and 400 ml of starting buffer containing 0.12 M NaCl. Temperature was 41°.

TABLE V: Compositions of the α2-CB4 Peptides from Insoluble Corium Collagen.

	Residues/Peptide		
	α2-CB4	α2-CB4′	Δ[a]
4-Hydroxyproline	30	30	–[b]
Aspartic acid	14	15	1
Threonine	6.2	6.4	–
Serine	10	13	3
Glutamic acid	22	23	1
Proline	40	41	1
Glycine	109	110	1
Alanine	38	39	1
Valine	11	11	–
Isoleucine	4.2	4.3	–
Leucine	10	11	1
Tyrosine	–	0.5	0.5
Phenylalanine	4.1	5.1	1
Hydroxylysine	3.6	3.6	–
Lysine	6.2	8.8	2.6
Histidine	2.3	2.3	–
Arginine	18	20	2
Homoserine[c]	1.0	1.1	–
Total	329	345	15
Molecular weight by Agarose chromatography	~29,000	~30,000	

[a] Δ = [α2-CB4′] − [α2-CB4]. [b] Less than 0.2 residue. [c] Includes homoserine lactone.

TABLE VI: Correlation of Hexose Content and *O*-Glycoside-Linked Hydroxylysine in CNBr Peptides from IC.

	Hydroxylysine[a]				Predominant Saccharide Unit
	Acid	Alkaline	Δ	Hexose[b,c]	
α1-CB(0,1)	–[e]	–	–	0	
α1-CB2	–	–	–	0	
α1-CB3	0.3	0.3	0	Trace	
α1-CB4	–	–	–	0	
α1-CB5	1.1	0.1	1.0	1.8–1.9	Di
α1-CB6	1.1	0.8	0.3	0.4	Mono
α1-CB7	0.9	0.9	0	0	
α1-CB8	0.8	0.7	0	0	
α2-CB0	–	–	–	0	
α2-CB1	–	–	–	0	
α2-CB2	–	–	–	0	
α2-CB(3–5)[d]	6.6	6.0	0.6	1.5	Di
α2-CB4	3.6	2.4	1.2	1.7–1.9	Di

[a] Free hydroxylysine, residues per peptide. [b] Hexose expressed in moles per peptide. [c] Hexose determined by the Winzler procedure and reported as galactose equivalents. [d] The alkaline degradation and hexose determinations were carried out on the uncleaved peptide. Qualitative hexose determinations on the small amounts of separated peptides available showed the hexose to be in the α2-CB5 portion. [e] Less than 0.2 residue.

to the α1-CB6 peptides but no further examination of these substances was made. We have no other explanation for this particular discrepancy.

α2-CB3, α2-CB4, AND α2-CB5. Peak 10, as shown in the chromatogram of Figure 2, contains a mixture of α2-CB4 and α2-CB3 in SC. In the insoluble collagens, because of the difficulty in cleaving the α2-CB3–α2-CB5 bond, most of the α2-CB3 is found in peak 11, along with the α2-CB5. The remaining peak 10 material in IC and DC still represents a mixture of components which elute on the Agarose columns at a position corresponding to a molecular weight of 29,000–30,000. Rechromatography of this Agarose peak from the IC digest on CM-cellulose at pH 4.8 (Figure 6) showed the peak to split into two major components with a lesser amount of a third component.

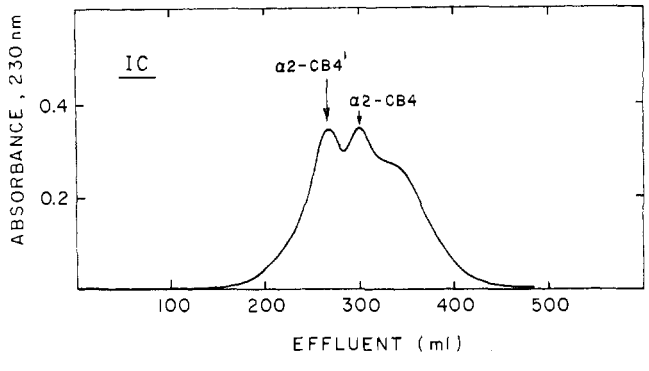

FIGURE 6: CM-cellulose chromatography of the fraction of peak 10 which emerges in Agarose columns with a molecular weight of 29,000. Elution was achieved with a linear gradient from 400 ml of starting buffer, 0.03 M sodium acetate (pH 4.8) and 400 ml of starting buffer containing 0.2 M NaCl. Temperature was 41°.

Material was collected from the central regions of the two peaks labeled α2-CB4 and α2-CB4′ in Figure 6. These materials were rechromatogramed on the Bio-Gel A-1.5 columns for final preparative and analytical purposes. Table V shows the amino acid compositions of the final purified fractions and their molecular weights for the α2-CB4 and α2-CB4′ from IC. The peptide identified as α2-CB4 is very close in composition to the α2-CB4 isolated from the α2- chain of SC (Volpin and Veis, 1971b). The difference of 15 residues between α2-CB4 and α2-CB4′ is consonant with the difference in molecular weight. The small peptides adjacent to α2-CB4 (α2-CB1, α2-CB0, and α2-CB2) all contain one-third glycyl residues. Hence, any uncleaved peptide sequence of 15 residues at either end of α2-CB4 should have contained five glycyl residues. However, the extra residues in α2-CB4′ are largely of polar character and do not contain the requisite additional glycine. The α2-CB4′ was homogeneous when examined by acrylamide gel electrophoresis and it seems unlikely from the final isolation of the peptide by Bio-Gel A-1.5 chromatography, that the difference in composition and size could have been due to contamination with a low molecular weight peptide unless that peptide was linked covalently. We are forced to conclude that α2-CB4′ represents α2-CB4 with a peptide attachment at a side-chain position. A similar situation exists with respect to the α2-CB4′ from DC.

TABLE VII: Correlation of Phosphate, Hexose, and *O*-Glycoside-Linked Hydroxylysine Residues in CNBr Peptides from DC.

	Hydroxylysine[a]					Predominant Saccharide Unit
	Phosphate[b]	Acid	Alkaline	Δ	Hexose[c,d]	
α1-CB(0,1)	–[e]	–	–	–	–	
α1-CB2	–	–	–	–	–	
α1-CB3	–	1.6	1.1	0.5	0.5	Mono
α1-CB4	–	–	–	–	–	
α1-CB5	–	1.5	0.6	0.9	1.7–1.9	Di
α1-CB6	0.4	1.4	0.8	0.6	0.4	Mono
α1-CB7	–	1.8	1.6	0.2	–	
α1-CB8	–	1.6	1.4	0.2	–	
α2-CB0	–	–	–	–	–	
α2-CB1	–	–	–	–	–	
α2-CB2	–	–	–	–	–	
α2-CB(3–5)[e]	1.2	6.8	5.7	1.1	1.4	Mono
α2-CB4	1.4	4.1	2.9	1.2	1.5–1.7	Mono

[a] Free hydroxylysine, residues per peptide. [b] Phosphate in groups per peptide, determined by method of Lucena-Conde and Prat (1957). [c] Hexose expressed in moles per peptide. [d] Hexose determined by the Winzler procedure and reported as galactose equivalents. [e] The alkaline degradation and hexose determinations were carried out on the uncleaved peptide. Qualitative hexose determinations on the small amounts of separated peptides available showed the hexose to be in the α2-CB5 portion. [f] Less than 0.2 residue.

The α2-CB4 peptide region, in the main body of the α2 chain, seems to be the only one of the α2-chain CNBr peptides exhibiting such a marked heterogeneity.

Hexose and Phosphate Distribution among the Soluble CNBr Peptides. Periodate degradation studies showed that both IC and DC contained from 1.7 to 2.5 mol of protected hydroxylysine per 1000 total amino acid residues, and this correlated well with alkaline hydrolysis data on O-glycosidically linked hydroxylysine and the total hexose content. Hexose analyses were made on every isolated CNBr peptide from IC and DC, and alkaline hydrolysis analyses on each peptide known to contain hydroxylysine. The digestion mixture from DC contained PO_4, whereas none was found in the IC digest. Phosphate determinations were carried out on each DC CNBr peptide. These data are summarized in Tables VI and VII. Hydroxylysine, as expected, appears in more peptides in DC than IC. In IC, α1-CB5, α1-CB6, α2-CB(3–5), and α2-CB4 all contain both hexose and O-glycosidically linked hydroxylysine. The IC α1-CB5, α2-CB4, and α2-CB-(3–5) contain disaccharides, while α1-CB6 is bound to a monosaccharide moiety. In DC (Table VII), only α1-CB5 bears a disaccharide, while α1-CB6, α2-CB(3–5), and α2-CB4 contain only monosaccharide moieties. In addition, α1-CB3 contains an O-glycosidically linked monosaccharide moiety. Neither α1-CB7 nor α1-CB8 showed complete correspondence of hydroxylysine when comparing acid and alkali hydrolysis procedures; however, no hexoses were detected by the Winzler (1955) procedure.

Peptides α1-CB6, α2-CB(3–5), and α2-CB4 from DC showed the presence of phosphate moieties in an amount approximately equivalent to the hexose present.

Insoluble Residues from CNBr Digestion. Although the explicit assumption in the foregoing presentation of the data was that the soluble CNBr peptides represented the total insoluble collagen composition, the presence of the insoluble residues in the digestion mixtures calls this assumption into serious question on a qualitative level. Clearly, the amount of residue is only on the order of 5% and the compositions of the formic acid insoluble digestion residues, after washing free of soluble peptides (Table VIII) indicate the presence of unreacted methionyl residues and, hence, incomplete cleav-

TABLE VIII: Amino Acid Compositions of the Formic Acid Insoluble Residues from CNBr Digestion of IC and DC.

	Residues/1000 Residues	
	IC	DC
4-Hydroxyproline	86	74
Aspartic acid	51	85
Threonine	20	21
Serine	39	83
Glutamic acid	82	74
Proline	118	114
Glycine	326	278
Alanine	108	83
Valine	30	22
Methionine[a]	3	3
Isoleucine	14	15
Leucine	30	35
Tyrosine	3	9
Phenylalanine	14	16
Hydroxylsine	3	7
Lysine	22	25
Histidine	3	6
Arginine	48	50
(Phosphate)	–	(45)
Total	1000	1000

[a] Methionine and methionine sulfoxide.

age. However, the digestion residues do differ in more significant ways from the soluble CNBr peptides. They have a lower content of proline and hydroxyproline than the total insoluble collagens from which they derive. The DC residue is the most interesting in that the marked reduction in proline and hydroxyproline is accompanied by a similar decrease in glycine, and increase in serine, aspartic acid, and a large accumulation of organic phosphate. The phosphate was identified as serine phosphate on the amino acid analyzer following both standard and limited acid hydrolysis (Veis *et al.*, 1972). The DC CNBr residue thus appears to be a fragment of collagen containing an anionic phosphoprotein in close association. This phosphoprotein corresponds in composition to the phosphoprotein previously isolated from periodate digests of collagen (Veis and Perry, 1967; Carmichael *et al.*, 1971) and from EDTA extracts of dentin (Veis *et al.*, 1972). The anionic nature of the non-collagen component appears to prevent the CNBr degradation of some methionyl residues, either in the non-collagen component or in the collagen backbone itself. The identification of the region of the collagen backbone involved with the anionic phosphoprotein is under investigation and will be described elsewhere (S. Ananthanarayanan and A. Veis, in preparation).

The corresponding appearance of the insoluble IC residue with uncleaved methionyl residues and somewhat elevated acidic peptide residues suggests the possibility that an anionic component similar but not identical with that associated with DC is also present in IC in small amounts. In the skin collagen residues, the glutamic acid content is enriched, in contrast to the elevated aspartic acid in the dentin residue. Both IC and DC residues contain hydroxylysine and hexose. The DC residue contains one equivalent of galactose per hydroxylysine, whereas the IC residue contains two hexoses per hydroxylysine, in keeping with the observation of the higher disaccharide content of the corresponding IC CNBr peptides.

Discussion

Tables IX and X summarize the results for the compositions of the CNBr peptides from insoluble bovine IC and DC, respectively, which have the greatest homology with the CNBr peptides obtained from the iolated α1(I) and α2 chains from acid-soluble bovine collagen (Volpin and Veis, 1971b) and with the corresponding peptides from rat (Butler *et al.*, 1967; Fietzek and Piez, 1969), chick (Kang *et al.*, 1969; Miller *et al.*, 1969; Lane and Miller, 1969) and human collagens (Click and Bornstein, 1970). Except for the variations in hydroxylation of proline and lysine, there are few differences of significance between the α1(I) and (α2) peptides of IC and DC and the peptides from soluble bovine collagen. A complete description of the CNBr peptides of soluble calf skin collagen has also been presented by Fietzek *et al.* (1970) and Rauterberg and Kühn (1971), in which a heterogeneity of α1-CB6 was noted. Subsequently (Rauterberg *et al.*, 1972), an α1-CB6 peptide was isolated from a urea-solubilized calf skin preparation which differs from the skin α1-CB6 presented in Table X by virtue of a glycine poor COOH-terminal sequence of 21 residues. This peptide does not contain a homoserine, as does α1-CB6′ (Table IV).

In addition to the expected peptides, the CNBr digests of IC and DC show the presence of several new components which, in themselves, are not identical with either the CNBr peptides from the soluble collagen or to each other in a skin–dentin tissue comparison. These differences are summarized schematically in Figure 7.

Cross-Link Peptides. Under the conditions used and without prior reduction to stabilize aldehyde-mediated cross-links which are acid labile and would be destroyed under the CNBr digestion conditions, only one peptide, α1-CB6′, could be interpreted as representing a cross-linked component in either IC or DC [α1-CB6 + α1-CB(0,1)]. Kang (1972) has found an [α1-CB6 + α1-CB1] cross-link peptide in borohydride reduced rattail tendon soluble collagen but not in unstabilized soluble tail tendon collagen. Thus, it appears likely that this intermolecular linkage is already present in a reduced, stabilized state in the insoluble bovine collagens, or that its location and adjacent residues provide an environment protecting it from acid hydrolysis. The intermolecular [α1-CB6-α1-CB(0,1) cross-linkage appears to be the main stabilizing covalent bond in acid-treated-insoluble collagen. The cross-link peptides in IC and DC do not match exactly with the expected compositions indicating that some additional COOH-terminal or NH_2-terminal segment is present. These small extensions of 5 and 12 residues do not contain glycine or hydroxyproline and are different in IC and DC preparations. The 12 residues in IC α1-CB6′ attributed to a chain extension do not correspond with the residues in the sequence determined by Rauterberg *et al.* (1972) for the –COOH-terminal extension in urea-solubilized calf skin collagen. Further work on these cross-link containing peptides is obviously necessary.

Genetically Distinct Collagens. Both IC and DC contain peptides indicating the presence of an α1(III) type chain in bovine skin and dentin. These appear only in the insoluble collagens and were not detected in the acid-soluble collagen. Tissue specificity also appears here with different compositions for α1(III) in the skin and dentin preparations; however, the differences are more marked in α1(III)-CB3 than in the smaller peptide α1(III)-CB(4,5).

No differences in backbone composition have been detected among the α2 peptides from the insoluble collagens. This is consistent with the comparative studies of Pikkarainen (1968) in which he indicated a relatively lower degree of mutation in the α2 chain as compared with the α1 chains.

Prosthetic Group Modifications. The presence of carbohydrate moieties O-glycosidically linked to hydroxylysine has been known since the work of Butler and Cunningham (1966) and the fact that both mono- and disaccharide units may be present was demonstrated by Spiro (1969). Tables VI and VII and the schematic presentation in Figure 7 show that the insoluble skin and dentin contain hexose attachments on the same set of four CNBr peptides. In both cases, α1-CB5 bears a galactosyl-glucose disaccharide moiety and α1-CB6, a monosaccharide. However, the α2-CB(3-5) and α2-CB4 peptides of skin are linked to disaccharide units while the corresponding peptides in dentin bear mainly monosaccharide attachments. The discrepancy between alkali-stable O-linked hydroxylysine and the hexose content in the α2 peptides from DC indicates that some disaccharide is also present in each case. The α1-CB3 from DC is also partially glycosylated with a monosaccharide unit. In addition to the hexose, α1-CB6, α2-CB(3-5), and α2-CB4 from DC all contain a phosphate group in an amount equivalent to the hexose content. The location of the phosphate has not been deduced, but it seems likely that the hexose is present as a hexose phosphate rather than the phosphate attachment being *via* a serine hydroxyl (Carmichael *et al.*, 1971).

The α2-CB4 region is of special interest because it contains hexose and because of its apparent association, in α2-CB4′, with a non-collagen peptide. The α2-CB4 peptide is

TABLE IX: Amino Acid Composition of CNBr Peptides from IC.

	Residues/1000 Amino Acid Residues								
	α1-CB(0,1)	α1-CB2	α1-CB3	α1-CB4	α1-CB5	α1-CB6	α1-CB7	α1-CB8	Total of CNBr Peptides[a]
	A. Peptides Derived from α1(I) Chains								
3-Hydroxyproline						0.7			1.0
4-Hydroxyproline		5.6	14	5.6	2.9	20	23	31	102
Aspartic acid	1.0		6.5	2.7	3.0	8.7	12	10	44
Threonine	0.9			0.9		3.8	5.0	5.2	16
Serine	3.0	1.8	2.8		1.7	7.4	7.4	8.2	32
Glutamic acid	2.0	3.8	16	3.0	3.0	13	17	19	77
Proline	2.0	6.0	16	5.8	2.2	26	37	28	123
Glycine	3.6	12	51	16	12	64	90	90	339
Alanine		2.3	22	3.0	4.2	22	34	35	123
Valine	1.1		3.8			3.2	4.7	4.6	17
Isoleucine	1.0					2.1	2.7	1.7	8
Leucine	1.0	1.0	3.0	1.8	1.0	4.0	3.9	3.0	19
Tyrosine	1.6								2
Phenylalanine		1.0	3.0		0.9	1.9	2.9	2.7	12
Hydroxylysine	0.2		0.3		1.2	1.1	0.9	0.8	4
Lysine	0.7		4.8	1.8	1.6	5.5	8.8	8.8	32
Histidine					0.9	0.7			2
Arginine		1.0	6.0	3.9	1.0	11	13	15	51
Homoserine	1.0	1.0	0.9	0.8	1.0		1.0	1.0	7
Total	19	36	150	45	37	195	263	264	1011

	Residues/1000 Amino Acid Residues					
	α2-CB0	α2-CB1[a]	α2-CB2	α2-CB4	α2-CB(3–5)	Total of CNBr Peptides[a]
	B. Peptides Derived from the α2 Chain					
3-Hydroxyproline					0.9	1
4-Hydroxyproline			2.3	30	50	82
Aspartic acid		1.0	2.1	14	31	48
Threonine			1.0	6.2	11	18
Serine		1.0	1.7	10	21	34
Glutamic acid		2.0	1.2	22	44	69
Proline		2.0	2.8	40	70	115
Glycine	1.0	3.0	10	109	212	335
Alanine		1.0	3.1	38	64	106
Valine			1.0	11	17	29
Isoleucine				4.2	10	14
Leucine	1.0		1.0	10	19	31
Tyrosine		0.7			1.6	2
Phenylalanine		0.8		4.1	8.6	14
Hydroxylysine		0.2		3.6	6.6	10
Lysine		0.4[b]		6.2	12	19
Histidine				2.3	5.7	8
Arginine			3.0	18	34	55
Homoserine	1.0	1.0	1.0	1.0	1.0[c]	5
Total	3	13	30	330	619	995

[a] Rounded off to nearest whole numbers. [b] As aldehyde. [c] Methionine and methionine sulfoxide.

located in the interior of the helical region of the α2 chains and, hence, the non-collagen peptide could only be added on as a side-chain, attachment. This would be akin to the very early suggestion of Rubin *et al.* (1963) that collagen contained non-collagen "telopeptides" not continuous with the main collagen peptide backbone. A detailed study of α2-CB4′ from DC is in progress, as is a related study of the insoluble residue from the CNBr digest of DC.

TABLE X: Amino Acid Composition of CNBr Peptides from DC.

	Residues/1000 Amino Acid Residues								
	α1-CB(0,1)	α1-CB2	α1-CB3	α1-CB4	α1-CB5	α1-CB6	α1-CB7	α1-CB8	Total of CNBr Peptides[a]
A. Peptides Derived from the α1(I) Chains									
3-Hydroxyproline						1.1			1
4-Hydroxyproline		5.6	15	5.8	2.8	19	28	31	107
Aspartic acid	1.0		6.4	2.8	2.8	8.6	12	9.8	43
Threonine	0.8			1.0		4.0	5.0	5.3	16
Serine	2.7	1.8	2.8		1.7	7.1	7.6	8.4	32
Glutamic acid	1.9	3.8	15	2.8	2.8	13	17	18	74
Proline	1.9	6.0	15	5.7	2.4	27	33	28	119
Glycine	3.7	12	51	16	12	64	90	89	338
Alanine		2.2	23	3.1	3.7	22	35	33	122
Valine	1.2		3.9			3.3	5.1	4.8	18
Isoleucine	0.8					1.9	2.6	1.8	7
Leucine	1.0	1.0	3.4	2.1	0.9	4.0	4.4	4.4	21
Tyrosine	1.6								2
Phenylalanine		1.0	2.6		0.8	1.8	3.1	2.8	12
Hydroxylysine	0.5		1.4		1.6	1.4	1.8	1.6	8
Lysine	0.3		3.9	1.9	1.4	4.9	8.6	8.2	29
Histidine					0.8	0.8			2
Arginine		1.0	6.0	3.8	1.2	11	13	15	51
Homoserine	1.0	1.0	0.9	0.8	0.9		0.9	0.9	6
Total	18	35	150	46	36	195	267	262	1008
Molecular weight by Agarose chromatography				4200	3400	16,500	23,500	23,000	

	Residues/1000 Amino Acid Residues					
	α2-CB0	α2-CB1[a]	α2-CB2	α2-CB4	α2-CB(3-5)	Total of CNBr Peptides[a]
B. Peptides Derived from the α2 Chain						
3-Hydroxyproline					1.2	1
4-Hydroxyproline			2.4	30	48	80
Aspartic acid		1.0	1.9	14	33	50
Threonine			1.0	6.1	11	18
Serine		0.8	1.7	9.3	22	34
Glutamic acid		1.9	1.2	22	44	69
Proline		2.0	2.7	39	71	115
Glycine	1.1	3.0	10	109	210	333
Alanine		1.1	3.3	37	64	105
Valine			1.0	10	17	28
Isoleucine				3.7	8.8	13
Leucine	1.0		1.0	9.4	18	29
Tyrosine		0.7			2.2	3
Phenylalanine		0.8		3.6	6.9	11
Hydroxylysine		0.4		4.1	6.9	11
Lysine		0.4[b]		6.4	12	19
Histidine				2.2	6.0	8
Arginine			3.0	17	34	54
Homoserine	0.9	1.0	1.0	1.0	1.0[c]	4
Total	3	13	30	322	617	985
Molecular weight by Agarose chromatography			2800	29,500	58,000	

[a] Rounded off to nearest whole number. [b] As aldehyde. [c] Methionine and methionine sulfoxide.

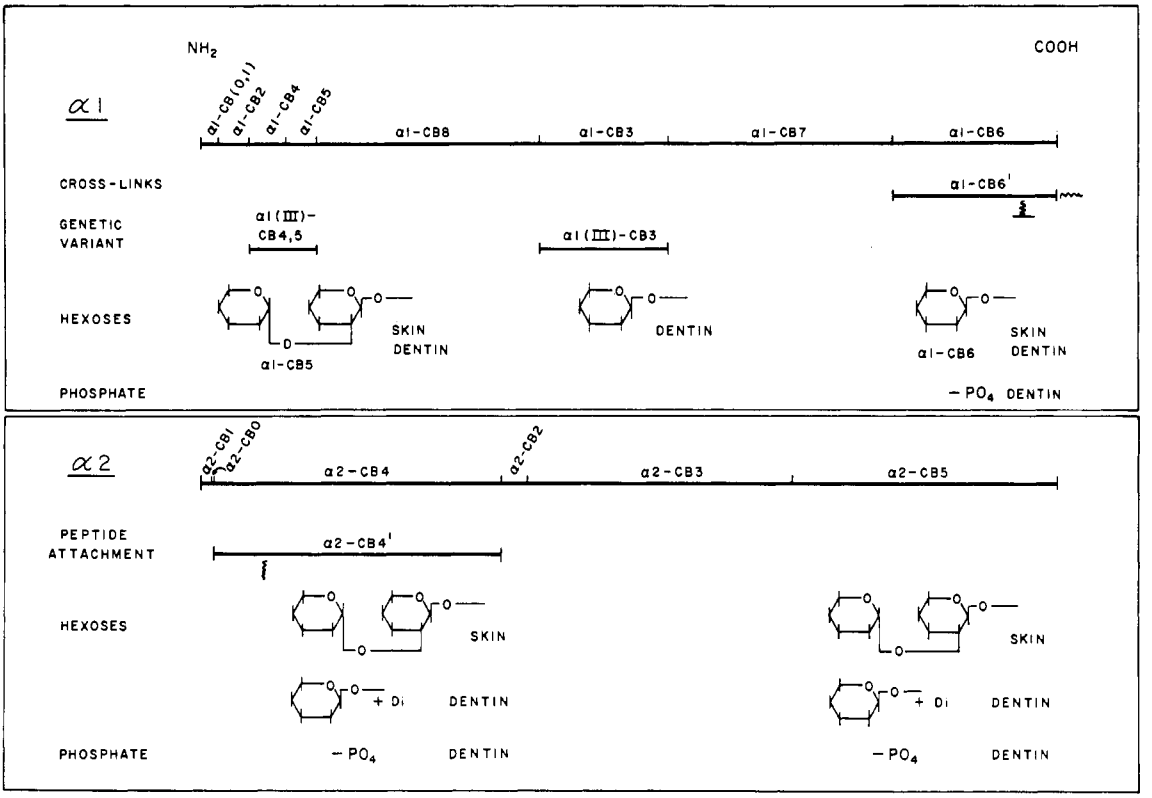

FIGURE 7: Schematic outline of locations and nature of tissue specific variations found in insoluble collagens.

Conclusion

Aside from the presence of the intermolecular cross-linkage region, the insoluble bovine collagens from skin and dentin differ from acid-soluble bovine skin collagen in two important ways. First, the insoluble collagens appear to be a mixture of two distinct collagen forms, one with the typical composition $[\alpha 1(I)]_2\alpha 2$, and the other with the presumed composition of $[\alpha 1(III)]_3$. In skin, the acid-soluble collagen is exclusively of the type $[\alpha 1(I)]_2\alpha 2$. In comparing IC and DC it is also of note that small tissue specific differences are observed in the $\alpha 1$(III) peptides. Second, either as an artifact of degradation during extraction or as a true reflection of the *in vivo* situation, the insoluble collagens contain small non-helical peptide extensions in the region of the *in vivo* stabilized intermolecular cross-link involving the $\alpha 1$-CB6 peptide, whereas these extensions are not present in the acid-soluble collagen. Similarly, it appears that some modification of the $\alpha 2$-CB4 peptide is possible and that this modified peptide region, $\alpha 2$-CB4′, is stable in the insoluble collagens but is either absent or removed during extraction of the soluble collagen. These observations, along with the recent study of artifactual cross-linking in acid-soluble collagen (Davison *et al.*, 1972), indicate that caution should be employed in extrapolating data on the soluble collagens to predict the behavior and properties of insoluble collagens, particularly with regard to intermolecular interactions.

In the schematic diagram of Figure 7, it is also of interest to note that the same portions of the $\alpha 1$ and $\alpha 2$ chains are subject to tissue-specific modification. Thus, a region near the NH_2-terminal end in each chain ($\alpha 1$-CB4, $\alpha 1$-CB5, and $\alpha 2$-CB4) may be modified by glycosylation, phosphorylation, or backbone sequence changes and possibly by the addition of a non-collagen peptide. Likewise, the COOH-terminal peptide ($\alpha 1$-CB6) may have variable degrees of glycosylation and phosphorylation and variable contents of non-helical peptide-chain extensions. Clark and Bornstein (1972) have recently shown that the $\alpha 2$-CB4 peptide region of guinea pig skin collagen is different from that of most other species studies in containing two extra methionyl residues. One may speculate that it is these regions of the collagen molecule which carry the tissue-specific information which modulate the processes of fibril organization and polymerization into the insoluble fibrous network.

References

Butler, W. T., and Cunningham, L. W. (1966), *J. Biol. Chem. 241*, 3882.

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

Carmichael, D. J., Veis, A., and Wang, E. T. (1971), *Calc. Tiss. Res. 7*, 331.

Clark, C. C., and Bornstein, P. (1972), *Biochemistry 11*, 1468.

Click, E. M., and Bornstein, P. (1970), *Biochemistry 9*, 4699.

Davis, B. J. (1964), *Ann. N. Y. Acad. Sci. 121*, 404.

Davison, P. F., Cannon, D. J., and Anderson, L. P. (1972), *Conn. Tiss. Res. 1*, 205.

Epstein, E. H., Jr., Scott, R. D., Miller, E. J., and Piez, K. A. (1971), *J. Biol. Chem. 246*, 1718.

Fietzek, P. P., Münch, M., Breitkrentz, D., and Kühn, K. (1970), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 9*, 229.

Fietzek, P. P., and Piez, K. A. (1969), *Biochemistry 8*, 2129.

Kang, A. H. (1972), *Biochemistry 11*, 1828.

Kang, A. H., Igarashi, S., and Gross, J. (1969), *Biochemistry 8*, 3200.

Lane, J. M., and Miller, E. J. (1969), *Biochemistry 8*, 2134.

Lucena-Conde, F., and Prat, L. (1957), *Anal. Chim. Acta*

16, 473.
Miller, E. J. (1971), *Biochemistry 10*, 9.
Miller, E. J., Epstein, E. H., and Piez, K. A. (1971), *Biochem. Biophys. Res. Commun. 42*, 1024.
Miller, E. J., Lane, J. M., and Piez, K. A. (1969), *Biochemistry 8*, 30.
Pikkarainen, J. (1968), *Acta Physiol. Scand. Suppl.*, 309.
Rauterberg, J., Fietzek, P., Rexrodt, F., Becker, U., Stark, M., and Kühn, K. (1972), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 21*, 75.
Rauterberg, J., and Kühn, K. (1971), *Eur. J. Biochem. 19*, 398.
Rubin, A. L., Pfahl, D., Speakman, P. T., Davison, P. F., and Schmitt, F. O. (1963), *Science 139*, 37.
Spiro, R. G. (1969), *J. Biol. Chem. 244*, 602.
Veis, A., and Perry, A. (1967), *Biochemistry 6*, 2409.
Veis, A., and Schlueter, R. J. (1964), *Biochemistry 3*, 1650.
Veis, A., Spector, A. R., and Zamoscianyk, H. (1972), *Biochim. Biophys. Acta 257*, 404.
Volpin, D., and Veis, A. (1971a), *Biochem. Biophys. Res. Commun. 44*, 804.
Volpin, D., and Veis, A. (1971b), *Biochemistry 10*, 1751.
Winzler, R. J. (1955), *Methods Biochem. Anal. 2*, 279.
Woessner, J. F. (1967), *Arch. Biochem. Biophys. 93*, 440.

Isolation and Properties of a Low Molecular Weight Protein (Apovitellenin I) from the High-Lipid Lipoprotein of Emu Egg Yolk†

R. W. Burley

ABSTRACT: As part of a study of protein–lipid interactions in avian egg yolk, the high-lipid lipoprotein (density 0.96 g/ml) of the egg yolk of the emu (*Dromaeus novaehollandiae*) has been examined. The major fraction of this lipoprotein has a particle weight of 3×10^6 with 13% of apoprotein. This apoprotein is more soluble than the corresponding apoprotein from the egg yolk of the hen (*Gallus domesticus*). Consequently several proteins (the emu "apovitellenins") that range in molecular weight from 10^4 to more than 10^5 have been recognized in the emu lipoprotein. The protein of lowest molecular weight ("apovitellenin I"), which accounts for nearly half the total apoprotein, has been isolated by chromatography in urea solution. Purified apovitellenin I contains no histidine, cystine, sulfhydryl groups, or phosphate. It contains a small amount of amino sugar. Lysine is the N-terminal residue. In disaggregating solvents (6 M guanidine hydrochloride and 6 M urea) apovitellenin I is present as a randomly coiled monomer at low concentrations. In water, in methanol, and in aqueous methanol, the protein has a high viscosity and a large proportion of α-helical structure according to optical rotatory dispersion, the maximum (nearly 80% helix) being in 50% aqueous methanol. In aqueous solutions above about pH 4 large aggregates were present and the protein was precipitated by low concentrations of salt. From its physical properties it is suggested that apovitellenin I has a structural role in the lipoprotein.

The arrangement of lipid and protein molecules in the soluble high-lipid low-density lipoproteins of avian egg yolk is not known. Lipoproteins of domestic hen's eggs have been studied intensively, but little is known about their apoproteins mainly because they are extremely insoluble in the usual nondegradative solvents for proteins. Consequently, reliable methods for their isolation and purification have not been available (Martin, 1961; Steer *et al.*, 1968; Evans *et al.*, 1968; Cook and Martin, 1969). From measurements in formic acid Martin (1961) suggested that the apoproteins of hen's egg lipoprotein have a low monomeric molecular weight. Recently, by using a mild procedure, it has been possible to isolate from the hen's lipoprotein a low molecular weight apoprotein fraction that is soluble in a water–methanol–chloroform mixture but not in solely aqueous media (Burley, 1968; Burley and Sleigh, 1971). In an extension of this work, it has been found that the apoproteins from the high-lipid lipoprotein of emu's egg yolk are all more soluble in aqueous solutions and in other solvents than those of the hen, although the lipoproteins are similar. There is apparently a series of emu apoproteins for which the term "emu apovitellenins" is proposed. The emu apovitellenin of lowest molecular weight, referred to here as "apovitellenin I," has been studied in detail in order to try to explain its function in the lipoprotein. Experiments on the location of this protein in the yolk, and on its isolation, purification, composition, and properties are described here. Its effect on lipid will be reported later.

Materials and Methods

Egg Yolk. Eight emu's eggs were studied, six of them from birds fed on a commercial laying-hen diet. When necessary, the eggs were stored at 1° for up to 9 months. Results of experiments on an egg less than 1 day old suggested that storage had no effect on apovitellenin I. The lipid composition of the yolks varied slightly with diet, and there may have been slight variations in the proportions of the apovitellenins in different eggs. Most of the results given here were from four eggs from birds fed on hen's food. The emu's eggs contained 225–243 g of yolk (about 31% of the weight of the egg) of which 56% by weight was nonvolatile in fresh yolk.

Acidic 6 M Urea. This was prepared by adding 10 N hydro-

† From the Commonwealth Scientific and Industrial Research Organization, Division of Food Research, North Ryde, N.S.W., 2113, Australia. *Received November 17, 1972.*