Characterization of the Cyanogen Bromide Peptides from the α2 Chain of Chick Skin Collagen*

Andrew H. Kang,† Shogo Igarashi, and Jerome Gross

ABSTRACT: Cleavage of the α2 chain of chick skin collagen with cyanogen bromide gives rise to six peptides which have been separated and purified by a combination of ion-exchange and molecular sieve chromatography. The finding of six peptides is consistent with the known content of methionine in the α2 chain (five) of chick skin collagen. The molecular weights of these peptides range from 31,000 to 317 and total 91,000, a value which is in reasonable agreement with the known molecular weight of 95,000 for the α chain. The amino acid composition of each peptide is clearly unique and the six peptides account for all of the known amino acid content of the original α2 chain within experimental error. The six cyanogen bromide peptides from the α2 chain of chick skin collagen are clearly homologous to the corresponding peptides derived from the α2 chain of rat skin collagen and identical in every examined respect with the cyanogen bromide peptides from the α2 chain of chick bone collagen.

Since it was previously shown that cyanogen bromide cleavage of the α1 chains from chick bone and skin collagens yielded identical peptides with regard to amino acid composition and molecular weight, the differences in the properties of bone and skin collagens are unlikely to be explained on the basis of differences in the primary structure.

Definitive understanding of the complex structure and biological behavior of collagen requires complete knowledge of the amino acid sequences of the two different polypeptide chain types comprising the molecule. The approach has been to reduce these large chains (molecular weight approximately 95,000) to smaller fragments of more manageable size by chemical cleavage at the methionyl residues, to establish the order of these peptides along the length of the chain by a variety of methods and to ultimately determine the amino acid sequence for each fragment. The α1 and α2 chains of rat skin (Butler *et al.*, 1967; Fietzek and Piez, 1969) and chick bone (Miller *et al.*, 1969; Lane and Miller, 1969) collagens have been under systematic study in this manner. We have previously reported (Kang *et al.*, 1969a) the isolation, purification, and characterization of ten unique peptides accounting for all of the known amino acid content and molecular weight of the α1 chain of chick skin collagen after cleavage at the methionyl residues with CNBr. In the present paper, we report the results of a similar study on the α2 chain of chick skin collagen performed as a part of our continuing efforts to establish the primary structure of this collagen. The results of this study are particularly interesting, as the recent study of Lane and Miller (1969) on the α2 CNBr peptides of chick bone collagen provides an opportunity to compare the structure of this protein from two different tissues from the same species but with widely different biologic function.

A useful nomenclature for designating the CNBr peptides of collagen has been described elsewhere (Miller *et al.*, 1969). As in the case of the α1 CNBr peptides, the peptides from the α2 chain of chick skin collagen have been named on the basis of inferred homology to the six α2 CNBr peptides from rat skin collagen (Fietzek and Piez, 1969).

Experimental Section

Source and Preparation of α2. Neutral salt-soluble and acid-soluble collagens were prepared from the skins of normal as well as lathyritic 3-week-old chicks as previously described (Kang *et al.*, 1969a). The source of collagen did not influence the results to be reported here and, therefore, will not be specified. The α2 chain was isolated by chromatography of the mildly heat-denatured, solubilized collagen on columns of CM-cellulose (CM-32, Whatman) as described previously (Kang *et al.*, 1969b). The effluent fractions containing the α2 chain were desalted on columns of Sephadex G-25 (Pharmacia) equilibrated with pyridine acetate (pH 4.8) as described previously (Piez *et al.*, 1963).

All column chromatography was monitored continously at 230 mμ using either a Gilford or Beckman DB spectrophotometer equiped with flow cells (path length, 10 mm).

CNBr Cleavage. The α2 chain was cleaved with CNBr in 0.1 N HCl at 30° for 4 hr under an atmosphere of nitrogen (Bornstein and Piez, 1966; Kang *et al.*, 1969a). A 100-fold molar excess (relative to methionine) of CNBr was used. The reaction was terminated by lyophilization after diluting the reaction mixture with ten volumes of cold water.

Phosphocellulose Chromatography. The lyophilized CNBr digests were chromatographed on 2 × 15 cm columns of phosphocellulose in 0.001 M sodium formate buffer (pH 3.6) at 40° (Kang *et al.*, 1969b). A linear gradient of NaCl from 0 to 0.3 M over a total volume of 800 ml was used. After the elution of α2-CB2 (see Figure 1), the remaining peptides were eluted as one peak by increasing the NaCl concentration of

* From the Developmental Biology Laboratory, Departments of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. *Received April 9, 1969.* This is Publication No. 481 of the Robert W. Lovett Memorial Group for the Study of Diseases Causing Deformities. This work has been supported by grants from the U. S. Public Health Service AM-3564 and the RGK Foundation.

† Senior Investigator of the Arthritis Foundation at the Massachusetts General Hospital.

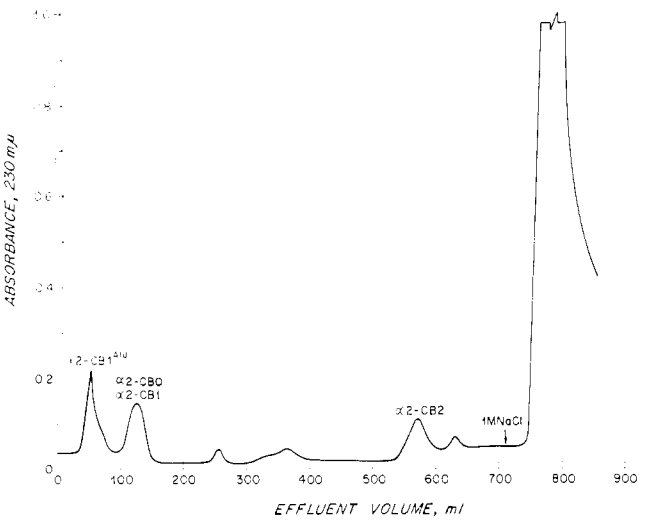

FIGURE 1: Phosphocellulose chromatography of the CNBr peptides derived from the $\alpha 2$ chain of neutral salt-soluble chick skin collagen at pH 3.6, 40°. Elution was with a linear gradient of 0.001 M sodium formate (pH 3.6) to 0.001 M sodium formate and 0.3 M NaCl (pH3.6) over a total volume of 800 ml. At the arrow NaCl concentration of elutrient was increased to 1 M.

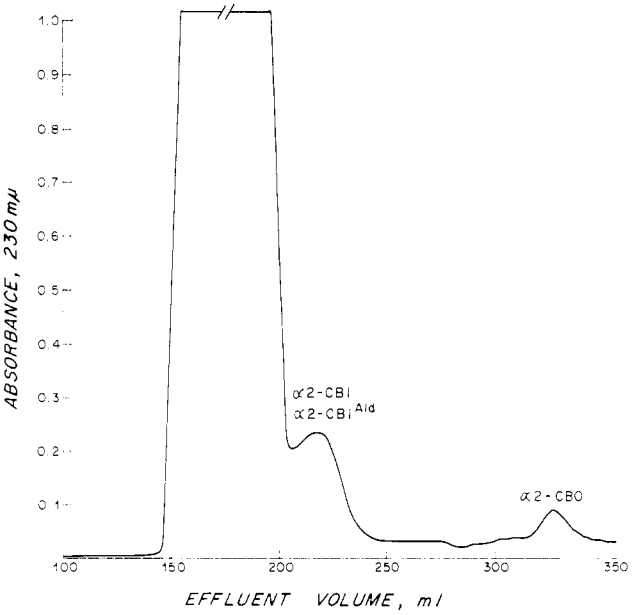

FIGURE 2: Chromatography of $\alpha 2$ CNBr digest on Bio-Gel P-4 (200–400 mesh) in 0.15 N acetic acid.

the elutrient to 1 M. The effluent fractions containing the peptides were lyophilized and desalted on columns of Bio-Gel P-2 equilibrated with 0.15 N acetic acid.

CM-cellulose Chromatography. The peptides eluted from the phosphocellulose columns with 1 M NaCl were further fractionated on 2 × 15 cm columns of CM-cellulose at 40° in 0.02 M sodium formate buffer (pH 3.6; Kang *et al.*, 1969a). A linear gradient of NaCl from 0.07 to 0.16 M superimposed over a total volume of 1100 ml was used. The peptide fractions were lyophilized and desalted on Bio-Gel P-2.

In order to improve the separation of the peptides, portions of the CM-cellulose effluent (pH 3.6) were desalted on Bio-Gel P-2 and rechromatographed on CM-cellulose at pH 4.8; 2 × 15 cm columns of CM-cellulose were equilibrated with 0.04 M sodium acetate (pH 4.8) containing 0.04 M NaCl at 40°. The samples were dissolved in 10 ml of the same buffer and applied to the column. Elution was carried out by a linear gradient of NaCl from 0.04 to 0.12 M over a total volume of 1000 ml. The peptide fractions obtained in this manner were lyophilized and desalted on Bio-Gel P-2 as described above.

Molecular Sieve Chromatography on Bio-Gel P-4. The smallest peptide obtained from the $\alpha 2$ chain by CNBr cleavage, $\alpha 2$-CB0, was most conveniently separated by chromatography of the lyophilized CNBr digest on a 2.5 × 80 cm column of Bio-Gel P-4 equilibrated with 0.15 N acetic acid as described by Fietzek and Piez (1969).

Molecular Weight Determination. The molecular weights of the CNBr peptides were determined by molecular sieve chromatography on a 2 × 110 cm column of Agarose equilibrated with 0.01 M Tris (pH 7.5) containing 1 M $CaCl_2$ as described by Piez (1968) and applied for molecular weight determination of the $\alpha 1$ CNBr peptides of chick skin collagen (Kang *et al.*, 1969a).

Amino Acid Analysis. The samples were hydrolyzed with constant-boiling HCl at 108° for 24 hr in tubes sealed under nitrogen. The hydrolysates were analyzed on a single-column automatic amino acid analyzer modified for high-speed analysis (Miller and Piez, 1966). Correction factors for labile amino acids (threonine, serine, methionine, and tyrosine) and incomplete release of valine were used as determined previously (Piez *et al.*, 1960).

Results

Phosphocellulose Chromatography of CNBr Peptides. Figure 1 represents a typical phosphocellulose elution pattern of CNBr peptides derived from the $\alpha 2$ chain of neutral salt-soluble chick skin collagen. The peaks were characterized by amino acid analysis. Three peaks representing four peptides were consistently noted. The peptides $\alpha 2$-CB0 and $\alpha 2$-CB1 were not separated on phosphocellulose but could be separated by molecular sieve chromatography on Bio-Gel P-2. When the effluent fractions representing $\alpha 2$-CB0 and $\alpha 2$-CB1 were lyophilized and desalted on a Bio-Gel P-2 column, $\alpha 2$-CB0 eluted in a volume corresponding to the fluid volume of the column. After the elution of $\alpha 2$-CB2, the peptides remaining on phosphocellulose were eluted as one peak with 1 M NaCl.

The first peak consisted of the peptide $\alpha 2$-$CB1^{Ald}$ and nonprotein ultraviolet-absorbing material. It was previously shown (Kang *et al.*, 1969b) that $\alpha 2$-$CB1^{Ald}$ is the form of $\alpha 2$-CB1 whose single lysyl residue has been oxidatively deaminated in peptide linkage to allysine (α-aminoadipic δ-semialdehyde) preliminary to the formation of interchain covalent cross-links *via* an aldol condensation with a similar residue in an adjacent chain (Kang *et al.*, 1969c). Similar findings were reported earlier for rat skin collagen (Bornstein and Piez, 1966; Bornstein *et al.*, 1966) although the exact chemistry of the cross-link had not yet been established.

Molecular Sieve Chromatography on Bio-Gel P-4. $\alpha 2$-CB0 could not be obtained free of salts on phosphocellulose chromatography but could conveniently be prepared on chroma-

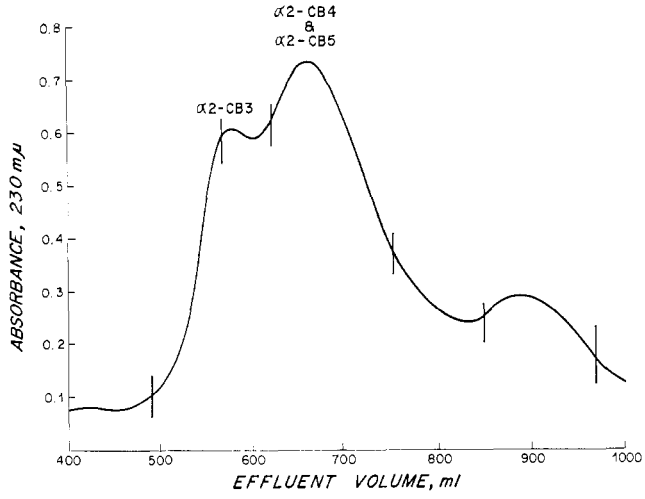

FIGURE 3: Chromatography of the CNBr peptides eluted from phosphocellulose with 1 M NaCl (see Figure 1) on CM-cellulose at pH 3.6, 40°. A linear gradient of 0.02 M sodium formate and 0.07 M NaCl (pH 3.6) to 0.02 M sodium formate and 0.16 M NaCl (pH 3.6) over a total volume of 1100 ml was used.

tography of the initial CNBr digest of the α2 chain on Bio-Gel P-4 (Figure 2). The first peak eluting at the void volume of the P-4 column is a mixture of large peptides and α2-CB2. α2-CB1 and α2-CB1^{Ald} are eluted next and α2-CB0 is eluted in a volume near the total fluid volume of the column. α2-CB0 obtained in this manner is free of salts and other peptide material.

CM-chromatography at pH 3.6. The large peptides eluted from phosphocellulose as one peak with 1 M NaCl solution were fractionated next on CM-cellulose, using 0.02 M sodium formate buffer (pH 3.6). A typical elution pattern is presented in Figure 3. The first peak contained α2-CB3 while the second peak included an incompletely separated mixture of α2-CB4 and α2-CB5. Rechromatography and amino acid analyses of

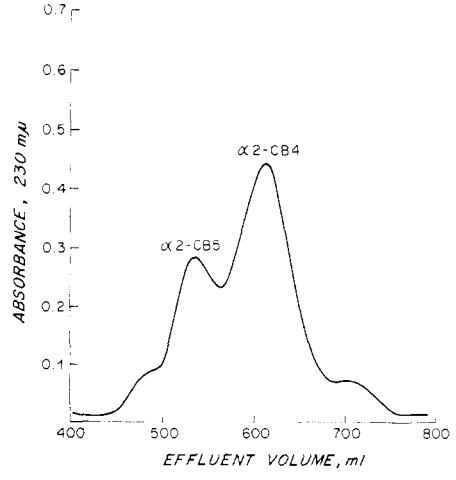

FIGURE 4: CM-cellulose rechromatography at pH 4.8 of the 620–750-ml portion of the CM-cellulose effluent (Figure 3). Chromatography was performed in 0.04 M sodium acetate (pH 4.8) with a linear gradient of NaCl from 0.04 to 0.12 M. The total volume of the gradient was 1000 ml.

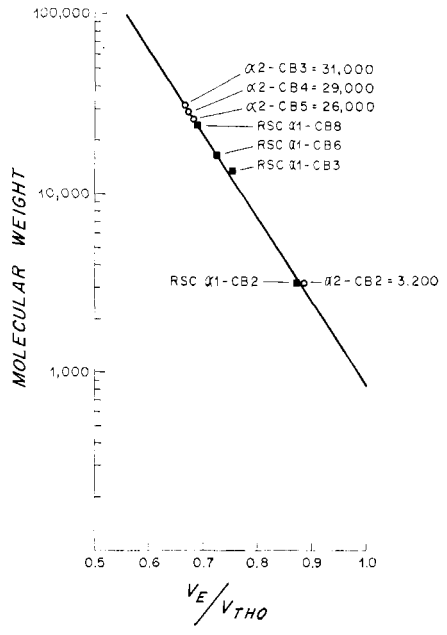

FIGURE 5: Plot of log molecular weight *vs.* elution volume, V_E/V_{THO}, from Agarose of standard peptides from rat skin collagen (RSC α1-CB8, α1-CB6, α1-CB3, and α1-CB2) and CNBr peptides from chick skin collagen α2.

various areas under the peak suggested that the first part of the second peak was relatively enriched in α2-CB5, and the latter part was enriched in α2-CB4. The third peak contained peptide material which had a molecular weight of 60,000 and methionine on amino acid analysis and must, therefore, represent incompletely cleaved material. This finding is not surprising since it is known that under the conditions of CNBr digestion used in the present study, cleavage is only about 85–90% complete (Bornstein and Piez, 1966). Since the resolution of the peptides is incomplete on CM-cellulose at pH 3.6, the column effluent was divided into four fractions and rechromatographed on CM-cellulose at pH 4.8.

CM-cellulose Rechromatography at pH 4.8. The effluent fractions obtained from CM-cellulose chromatography at pH 3.6 were rechromatographed on CM-cellulose at pH 4.8 using sodium acetate buffer. The 490–570-ml fraction (Figure 3) contained a single peptide, α2-CB3. The 620–750-ml fraction was resolved into two peptides, α2-CB5 and α2-CB4, by CM-cellulose chromatography at pH 4.8 as shown in Figure 4. Pure α2-CB4 and α2-CB5 could be obtained by rechromatography of each on CM-cellulose under the same conditions. Separation of the 570–620-ml fraction (Figure 3) on CM-cellulose at pH 4.8 showed it to be a mixture of α2-CB3 and α2-CB5.

By a combination of these procedures, then, a total of six unique CNBr peptides has been obtained. This is consistent with the known content of methionyl residues (five) in the α2 chain of chick skin collagen and agrees with data obtained by Lane and Miller (1969) on the α2 CNBr peptides of chick bone collagen. The six α2 CNBr peptides of chick skin collagen are clearly homologous to the α2 CNBr peptides of rat skin collagen (Fietzek and Piez, 1969) and indentical with the α2 CNBr peptides of chick bone collagen with regard to amino acid composition and molecular weight (see below).

TABLE I: Amino Acid Composition of the CNBr Peptides from the α2 Chain of Chick Skin Collagen.[a]

Amino Acid	α2-CB0	α2-CB1	α2-CB2	α2-CB3	α2-CB4	α2-CB5	Total	α2[b]
3-Hydroxyproline	0	0	0	0	0	1 (0.5)	1	1 (1.0)
4-Hydroxyproline	0	0	2 (2.1)	27	42	31	102	100
Aspartic acid	0	2	3 (2.8)	14	15	17	51	52
Threonine	0	0	0	7 (6.8)	7 (6.8)	5 (4.9)	19	21
Serine	0	1	1 (0.9)	8 (7.8)	12	9 (9.2)	31	32
Glutamic acid	0	1	1 (1.1)	23	20	17	62	69
Proline	0	3	3 (3.1)	39	36	39	120	121
Glycine	1 (1.1)	2	10	117	115	112	357	339
Alanine	0	2	4 (4.1)	38	28	25	97	108
Valine	0	0	1 (1.2)	12	9 (9.1)	10 (9.9)	32	29
Methionine	0	0	0	0	0	0	0	5 (4.9)
Isoleucine	0	0	0	7 (7.1)	6 (5.9)	6 (5.9)	19	18
Leucine	1 (0.9)	0	1 (1.1)	7 (7.0)	10	12 (11.9)	31	33
Tyrosine	0	1	0	0	0	1 (0.7)	2	2 (1.9)
Phenylalanine	0	1	0	5 (4.9)	4 (4.0)	4 (4.1)	14	14
Hydroxylysine	0	0	0	3.8	4.1	2.1	10.0	10.3
Lysine	0	1	0.9	10.0	6.7	5.1	23.7	22
Histidine	0	0	0	1 (1.2)	3 (3.0)	4 (4.0)	8	8 (7.9)
Arginine	0	0	2 (2.0)	17	17	16	52	52
Homoserine[c]	1 (0.9)	1	1 (0.9)	1 (1.1)	1 (1.0)	0	5	0
Total	3	15	30	337	336	316	1037	1036

[a] Residues per peptide. Values are rounded off to the nearest whole number except for lysine and hydroxylysine when both are present since they may not be integers (Miller *et al.*, 1969). Where less than ten residues were found, actual values are shown in parentheses. A value of zero indicates less than 0.2 residue. Detailed data on a α2-CB1 have been given previously (Kang *et al.*, 1969b). [b] Values are calculated for a molecular weight of 95,000 with an average residue weight of 91.8. [c] Includes homoserine lactone.

Amino Acid Compositions of the CNBr Peptides. Table I shows the amino acid compositions of the six CNBr peptides. Residues per peptide were calculated by using the values of amino acids present in small amounts and assuming one residue per peptide of homoserine (plus homoserine lactone).

α2-CB0 is the smallest peptide containing one residue each of glycine, leucine, and homoserine. α2-CB1 (and α2-CB1^{Ald}) is the NH_2-terminal peptide which contains the lysyl residue involved in the formation of the intramolecular cross-link (Kang *et al.*, 1969b). It has an amino acid composition identical with that of α2-CB1 from chick bone collagen (Lane and Miller, 1969) except that the lysyl residue is not hydroxylated at all in chick skin α2-CB1. α2-CB2 is a small peptide containing 30 residues of which one-third is glycine. Each of the three remaining peptides is large, containing more than 300 residues of amino acids. In each instance, glycine is one-third of the total. Although the amino acid compositions of the three large peptides are similar, each has some unique features which distinguish it from the others. Thus α2-CB3 has one residue of histidine, α2-CB4 has three residues, while α2-CB5 has four residues. α2-CB5 does not contain homoserine and thus must contain the COOH terminus of the α2 chain. It also has the one residue of 3-hydroxyproline and one of the two tyrosyl residues. The compositions of these peptides are identical with those of chick bone α2 CNBr peptides (Lane and Miller, 1969) and clearly homologous to the corresponding CNBr peptides of rat skin α2.

Table I also compares the measured amino acid composition of the whole α2 chain and the total amino acid content accounted for by the six CNBr peptides. It is clear from the comparison that the six isolated peptides account for, within experimental error, all of the amino acids in the original α2 chain.

Molecular Weights of CNBr Peptides. Figure 5 presents the plot of log molecular weight *vs.* V_E/V_{THO} of the rat skin collagen α1 CNBr peptides used to standardize the Agarose column and shows the positions at which the α2 CNBr peptides were eluted. The molecular weights of the rat skin collagen α1 CNBr peptides used here had been determined previously by the sedimentation equilibrium method of Yphantis (Butler *et al.*, 1967). Table II compares the molecular weights of the chick skin collagen α2 CNBr peptides as determined by amino acid analysis and by molecular sieve chromatography. The values obtained by the two methods agree well in every instance. The sum of molecular weights of the six peptides also agrees well with the reported molecular weight of about 95,000 for the whole α chain (Lewis and Piez, 1964; Piez, 1965; Kang *et al.*, 1966).

Discussion

After cleavage of the α2 chain of chick skin collagen with CNBr we have isolated and characterized six unique peptides. The finding of six CNBr peptides is consistent with the known

TABLE II: Molecular Weights of the CNBr Peptides of the $\alpha 2$ Chain of Chick Skin Collagen.

Peptide	Amino Acid Anal.	Molecular Sieve Chromatography
$\alpha 2$-CB0	317	(317)[a]
$\alpha 2$-CB1	1,577	(1,577)[a]
$\alpha 2$-CB2	2,745	3,200
$\alpha 2$-CB3	30,585	31,000
$\alpha 2$-CB4	30,920	29,000
$\alpha 2$-CB5	29,040	26,000
Total	94,867	91,094

[a] These values were calculated from the amino acid composition only.

methionine content (five) of the $\alpha 2$ chain of chick skin collagen. The six peptides account for all of the known amino acid content and the molecular weight of the $\alpha 2$ chain of chick skin collagen. Together with a similar study on the $\alpha 1$ chain of chick skin collagen previously reported (Kang *et al.*, 1969a), this present study completes the characterization of the CNBr peptides of this particular collagen. The amino acid sequence of some of the $\alpha 1$ CNBr peptides as well as $\alpha 2$ CNBr peptides from chick skin collagen will be separately reported (A. H. Kang and J. Gross, manuscript in preparation).

The order in which the CNBr peptides occur in the $\alpha 1$ chain of chick skin collagen was tentatively established (Piez *et al.*, 1969; Kang *et al.*, 1969a). Similar data on the order of the $\alpha 2$ CNBr peptides are not yet available. Evidence was presented that $\alpha 2$-CB1 ($\alpha 2$-CB1^{Ald}) is the NH_2 terminus (Kang *et al.*, 1969b). $\alpha 2$-CB5 clearly must be the COOH-terminal peptide as it contains no homoserine. The relative positions of the other four peptides are unknown, but presently under investigation.

We have previously demonstrated that the CNBr peptides of the $\alpha 1$ chain of chick skin collagen are identical with the CNBr peptides of the $\alpha 1$ chain of chick bone collagen (Kang *et al.*, 1969a; Miller *et al.*, 1969). The present study allows a similar comparison of the two collagens to be extended to the $\alpha 2$ chain and reveals that the $\alpha 2$ CNBr peptides of the two collagens are identical with regard to amino acid composition and molecular weight except for the degree of hydroxylation of the lysyl residues. It can be most clearly seen in the case of $\alpha 2$-CB1 where the single lysyl residue of bone collagen is hydroxylated about 50% of the time (Lane and Miller, 1969) while the same residue in skin collagen is never hydroxylated. Similar differences in the degree of hydroxylation of the lysyl residues between the two collagens have been shown previously for the $\alpha 1$ chains (Kang *et al.*, 1969a; Miller *et al.*, 1969). Whether similar differences in the degree of hydroxylation of prolyl residues also exist cannot be decided at present.

The data now available indicate the collagens from two tissues of the same animal, with widely different biologic function and properties, have identical primary structures as far as can be determined from studies of the CNBr peptides, although it is not yet possible to exclude completely the existence of small but functionally significant differences in the primary sequences of such collagens. Nevertheless, the possibility is suggested that alterations which occur after completion of synthesis of the polypeptide chains may be responsible for differences in properties of collagens from different tissues. Some of these modifications known to occur in collagen include formation of aldehydic groups and subsequently of covalent cross-links, hydroxylation of lysyl and prolyl residues, and addition of carbohydrates; occurrence of some unusual compounds such as reported by Gallop *et al.* (1968) may also belong to this category. It may be anticipated that extension of the work described herein may lead to further information concerning this possibility.

Acknowledgments

We are grateful to Dr. Karl A. Piez, Dr. Peter Fietzek, Dr. Joseph M. Lane, and Dr. Edward H. Miller of the National Institutes of Health, Bethesda, Md., for allowing us to read their manuscripts before publication and for helpful suggestions during our investigation. Skillful technical assistance of Mrs. Marilyn Bunnell is gratefully acknowledged.

References

Bornstein, P., Kang, A. H., and Piez, K. A. (1966), *Proc. Natl. Acad. Sci. U. S. 55*, 417.

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

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

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

Gallop, P. M., Blumenfeld, O. O., Henson, E., and Schneider, A. L. (1968) *Biochemistry 7*, 2409.

Kang, A. H., Faris, B., and Franzblau, C. (1969c), *Biochem. Biophys. Res. Commun.* (in press).

Kang, A. H., Nagai, Y., Piez, K. A., and Gross, J. (1966), *Biochemistry 5*, 509.

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

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

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

Lewis, M. S., and Piez, K. A. (1964), *J. Biol. Chem. 239*, 3336.

Miller, E. J., Lane, J. M., and Piez, K. A. (1969), *Biochemistry 8*, 30.

Miller, E. J., and Piez, K. A. (1966), *Anal. Biochem. 16*, 320.

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

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., Eigner, E. A., and Lewis, M. S. (1963), *Biochemistry 2*, 58.

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