Comparative Studies on the Amino Acid Sequence of the α2-CB2 Peptides from Chick and Rat Skin Collagens*

John H. Highberger,† Andrew H. Kang,‡ and Jerome Gross

ABSTRACT: The amino acid sequence of the 30-residue cyanogen bromide peptide (α2-CB2) from the α2 chain of chick skin collagen has been determined by the Edman degradation–dansylation method applied to its tryptic digestion products. Confirmation of the structure was obtained by study of additional overlapping peptides isolated from a thermolytic digest. The primary structure consists of the triplet sequence generally characteristic of collagen, with glycine every third residue, and hydroxyproline occurring only in position 3. The sequence of the NH_2-terminal-half of the homologous peptide from rat skin collagen has also been determined. This portion of the molecule contains all of the three substitutions between the two species found in these peptides. In contrast to the well-known resistance of the arginyl–proline bond, the arginyl–hydroxyproline bond in α2-CB2 was found to be slowly cleaved by trypsin in Tris buffer at pH 7.5.

Information on the primary structure of the α chains of several collagens, primarily those of rat and chick tissues, is gradually accumulating. The general approach to this problem involving cleavage of the separated α chains at methionyl residues by cyanogen bromide treatment and isolation and sequence study on the resulting fragments, is being pursued in several laboratories. Thus, Kang *et al.* (1969a,b) have accounted for the total known amino acids of the α1 and α2 chains of chick skin collagen in terms of isolated and characterized CNBr peptides. The order in which these peptides occur along the α1 chain has been established by Piez *et al.* (1969) and Rauterberg and Kühn (1968). Igarashi *et al.* (1970) and Vuust *et al.* (1970) very recently have determined this order in the case of the α2 chain. The amino acid sequences of several of the smaller CNBr peptides from the NH_2-terminal region of the α chains of chick skin have also been determined (Kang *et al.*, 1967; Kang and Gross, 1970), and a similar body of data exists for chick bone collagen (Miller *et al.*, 1969; Lane and Miller, 1969), and for rat skin collagen (Butler *et al.*, 1967; Fietzek and Piez, 1969). All of this work is leading toward the ultimate goal of detailed knowledge of the primary structure of the collagen macromolecule, necessary for definitive understanding of its role in the tissue in normal physiologic and pathologic processes.

In the present communication we report the amino acid sequence of the 30- residue α2-CB2 peptide from chick skin collagen, and also that of the portion of the homologous peptide from rat skin collagen which contains all of the interspecies substitutions found in these peptides.

Experimental Section

Preparation of Collagen and Isolation of α2. Neutral salt-extracted chick skin collagen was prepared from 3-week-old white Leghorn lathyritic[1] chicks according to the procedure described by Kang *et al.* (1969c). The α2 chain was isolated by CM-cellulose chromatography of the denatured, solubilized collagen (Piez *et al.*, 1963; Kang *et al.*, 1969c).

Rat skin collagen was prepared as described by Kang *et al.* (1966) and the α2 chain was isolated as above.

Preparation of α2-CB2. The CNBr cleavage of the purified α2 chains, and the isolation of the resultant α2-CB2 by phosphocellulose chromatography were carried out as previously described (Bornstein and Piez, 1966; Kang *et al.*, 1969b). The α2-CB2 fractions were further purified by column chromatography on Sephadex G-50 or G-100.

Tryptic Hydrolyses. Hydrolyses with trypsin were carried out in 0.2 M NH_4HCO_3 containing 10^{-3} M $CaCl_2$, at pH 7.8, or in 0.1 M Tris containing 10^{-3} M $CaCl_2$ at the same pH. Treatments were allowed to proceed for 4 hr at 37°, or, in special cases, for a shorter time (2.5 hr), at the end of which fresh enzyme was added and the treatment repeated. The enzyme was Worthington L-(1-tosylamide-2 phenyl)ethyl chloromethyl ketone treated, and used at a molar ratio to substrate of 1:50. Reactions were terminated by addition of 2 N acetic acid to pH 4, followed by lyophilization.

Thermolytic Hydrolysis. Thermolytic hydrolysis of α2-CB2 was carried out in 0.2 M Tris buffer at pH 8.0, using an approximately 1% solution of the peptide, and 0.5% (on the weight of peptide) of enzyme, added in 0.001 M calcium acetate. Digestion was allowed to proceed for 2 hr at 37°, at which time a second aliquot of enzyme was added and the digestion continued for 1 more hr. The reaction was terminated by addition of 6 N HCl to pH 2, and the mixture was lyophilized. With the exception of the use of Tris buffer, these conditions are those of Bradshaw (1969), and of Ambler and Meadway (1968). The thermolysin was a three-times-crystallized product (Calbiochem).

Cation-Exchange Chromatography. Tryptic and thermolytic hydrolysates of α2-CB2 were fractionated by ion-exchange chromatography. The initial fractionations were carried out on a 0.9 × 23 cm column of PA-35 (Beckman) cation-ex-

* From the Developmental Biology Laboratory, Departments of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts. *Received July 27, 1970.* This is publication No. 526 of the Robert W. Lovett Memorial Group for the Study of Diseases Causing Deformities. This work was supported by an RGK Foundation grant and Public Health Service Grant No. AM 3564 from the National Institute of Arthritis and Metabolic Diseases.

† To whom to address correspondence.

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

[1] The tissues of lathyritic chicks were used in order to maximize the yield of extractible collagen. Lathyrism was induced by feeding 1 day old chicks a commercial diet containing 0.1% β-aminopropionitrile for 3 weeks.

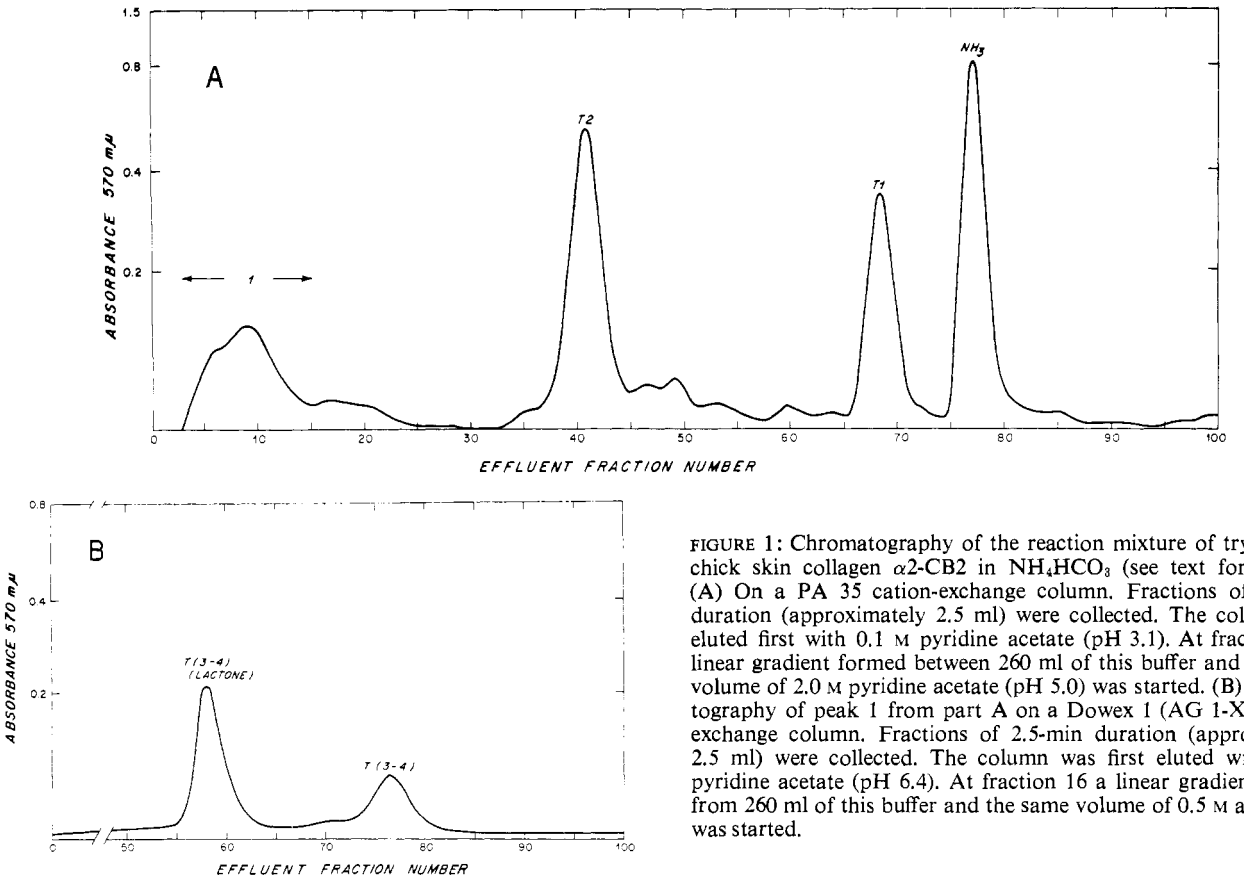

FIGURE 1: Chromatography of the reaction mixture of trypsin and chick skin collagen α2-CB2 in NH_4HCO_3 (see text for details). (A) On a PA 35 cation-exchange column. Fractions of 2.5-min duration (approximately 2.5 ml) were collected. The column was eluted first with 0.1 M pyridine acetate (pH 3.1). At fraction 12 a linear gradient formed between 260 ml of this buffer and the same volume of 2.0 M pyridine acetate (pH 5.0) was started. (B) Chromatography of peak 1 from part A on a Dowex 1 (AG 1-X2) anion-exchange column. Fractions of 2.5-min duration (approximately 2.5 ml) were collected. The column was first eluted with 0.5 M pyridine acetate (pH 6.4). At fraction 16 a linear gradient formed from 260 ml of this buffer and the same volume of 0.5 M acetic acid was started.

change resin at 50°, using a linear gradient formed between equal volumes of 0.1 M pyridine acetate (pH 3.1) and 2.0 M pyridine acetate (pH 5.0) final buffer. In most cases the column was eluted for 1 hr with the initial buffer before starting the gradient. An automatic amino acid analyzer (Beckman), modified for peptide fractionations, was used for the runs. Total buffer flow rates of approximately 70 ml/hr were used, with a split-stream pump diverting one-seventh of the column effluents through the ninhydrin reaction coil for monitoring on the analyzer recorder by measurements of absorbance at 570 mμ. The remainder of the column effluents were collected in a fraction collector as fractions of 2.5-min duration. Appropriate fractions were pooled and lyophilized, and suitable aliquots taken for amino acid analysis.

Anion-Exchange Chromatography. In cases where peaks isolated from the cation-exchange column were found to represent mixtures rather than pure peptides, these were rechromatographed on a 0.9 × 58 cm column of Dowex 1 (AG 1-X2, Bio-Rad). A linear gradient formed between equal volumes of an initial buffer of 0.5 M pyridine acetate (pH 6.4), and a final eluent of 0.5 M acetic acid was used. The same modified analyzer used for cation-exchange chromatography was used, and the flow rates and effluent split between reaction coil and fraction collector were the same as described above.

Amino Acid Analyses. Suitable aliquots of the isolated peptides were hydrolyzed in constant-boiling HCl at 108° for 24 hr under nitrogen. The amino acid analyses were performed on commercial analyzers (Beckman, Phoenix) by the single-column method of Miller and Piez (1966). No corrections were made for hydrolytic loss of labile amino acids.

Sequence Determination. Determinations of the amino acid sequences of the isolated peptides was carried out by the dansylation-Edman degradation procedures of Gray (1967), with slight modifications as described by Kang and Gross (1970). DNS[2]-amino acids were identified by thin-layer chromatography on silica gel sheets (Eastman Chromatogram Sheet 6061) using a sandwich-type developing apparatus (Eastman). Unequivocal results were usually obtained by one-dimensional chromatography employing the solvent systems of Deyl and Rosmus (1965), or of Morse and Horecker (1966). In some cases the two-dimensional polyamide layer system of Woods and Wang (1967) was used.

Paper Electrophoresis. The presence of asparagine or glutamine was deduced from determinations of net charge by paper electrophoresis of the DNS-peptides, or of DNS-peptides at relevant stages of Edman degradation, as described by Gray (1967). Electrophoresis was carried out on 50-cm strips of Whatman No. 3MM paper for 1 hr in pH 6.5 pyridine acetate, using a Savant flat-plate apparatus at 4000 V.

Results

Chick Skin Collagen α2-CB2

Tryptic Treatment in NH_4HCO_3. Tryptic treatment of chick skin collagen α2-CB2 was first carried out in NH_4HCO_3 buffer for 4 hr at 37° as detailed in the Experimental Section. The lyophilized reaction mixture was chromatographed on a PA 35 cation-exchange column under the conditions described above. The resulting chromatogram is shown in Figure 1A.

Peak 1 in this chromatogram was obviously a poorly re-

[2] Abbreviation used is: DNS, 1-dimethylaminonaphthalene-5-sulfonyl (dansyl).

TABLE I: Amino Acid Composition of Chick Skin Collagen α2-CB2 and Its Tryptic Digestion Products in NH_4HCO_3 Buffer.[a]

		Figure 1A		Figure 1B	
	α2-CB2[c] (1–30)	T2 (7–15)	T1 (1–6)	T(3–4) (Lactone) (16–30)	T(3–4) (16–30)
4-Hydroxyproline	2	–	–	1.7	1.9
Aspartic acid	3	–	0.9	1.9	2.0
Serine	1	0.9	–	–	–
Glutamic acid	1	–	–	1.4	1.4
Proline	3	0.9	1.2	1.2	1.0
Glycine	10	3.4	2.1	5.4	4.9
Alanine	4	2.2	0.9	1.4	1.2
Valine	1	1.0	–	–	–
Leucine	1	–	–	1.0	0.9
Lysine	1	1.0	–	–	–
Arginine	2	–	1.0	0.9	0.8
Homoserine[b]	1	–	–	1.0	1.0

[a] Composition expressed as residues per peptide. A dash indicates 0.1 residue or less. The numbers in parentheses refer to the residue numbers in Figure 4. [b] Includes homoserine lactone. [c] Kang *et al.* (1969b).

TABLE II: Amino Acid Composition of the Tryptic Digestion Products of Chick Skin Collagen α2-CB2 in Tris Buffer.[a]

	Figure 2A			Figure 2B	
	T2 (7–15)	T3 (16–23)	T1 (1–6)	T(3–4) (16–30)	T4 (24–30)
4-Hydroxyproline	–	–	–	2.1	2.0
Aspartic acid	–	2.0	1.1	2.0	–
Serine	0.9	–	–	–	–
Glutamic acid	–	–	–	1.0	1.0
Proline	1.1	1.0	0.9	0.9	–
Glycine	3.3	3.2	2.0	4.8	2.0
Alanine	2.1	0.8	1.1	1.1	–
Valine	1.0	–	–	–	–
Leucine	–	–	–	1.0	1.0
Lysine	1.0	–	–	–	–
Arginine	–	1.0	0.9	0.9	–
Homoserine[b]	–	–	–	1.0	1.0

[a] Composition expressed as residues per peptide. A dash indicates 0.1 residue or less. The numbers in parentheses refer to the residue numbers in Figure 4. [b] Includes homoserine lactone.

solved mixture, because of the marked shoulder. The amino acid analyses of the second (T2) and third (T1) peaks indicated pure nona- and hexapeptides, respectively (see Table I). The fourth peak contained only ammonia.

The total peak 1 was collected and after lyophilization was rechromatographed on Dowex 1-X2, in an attempt to resolve the components of the mixture. This chromatogram is shown in Figure 1B. Two well-resolved peaks were obtained. The amino acid analyses of these were identical within experimental error, and indicated a 15-residue peptide containing homoserine (Table I). The two peaks thus represented the homoserine and homoserine lactone forms of the COOH-terminal-half of chick skin collagen α2-CB2. This portion of the molecule furthermore contained an arginyl bond which remained uncleaved by the tryptic treatment.

Tryptic Treatment in Tris Buffer. A second tryptic treatment of chick skin collagen α2-CB2 was carried out in Tris buffer, under the conditions described in the Experimental Section. The lyophilized reaction mixture was chromatographed on the PA 35 cation-exchange column under the conditions given above. The resulting chromatogram is shown in Figure 2A.

This chromatogram, run under the same conditions as that of the first tryptic treatment done in NH_4HCO_3 buffer shows two differences from that chromatogram (Figure 1A). First, it shows one additional peak (T3), and secondly, the poorly resolved peak 1 is somewhat better resolved.

The amino acid analyses of the second (T2), third (T3), and fourth (T1) peaks indicated pure nona-, octa-, and hexapeptides, respectively (Table II). Of these, the composition of the second and fourth peaks were identical with those of the second and third, respectively, of the chromatogram of Figure 1A. The fifth peak contained no amino acid after hydrolysis.

The two initial poorly resolved peaks were pooled as peak 1, lyophilized, and rechromatographed on Dowex 1-X2 under the same conditions described above. The resulting chromatogram is shown in Figure 2B. This chromatogram shows two peaks, of which the first is identical in composition (Table II) with two peaks of Figure 1B, and therefore represents the 15-residue COOH-terminal-half of α2-CB2. The analysis of the second peak (T4) shows it to contain seven residues, of which one is homoserine. This peptide therefore represents the fourth theoretically possible tryptide, formed by the partial cleavage, in Tris buffer, of the arginyl bond which resisted cleavage in NH_4HCO_3 buffer. The eight-residue peptide isolated from the PA 35 column (T3 of Figure 2A) must represent the NH_2-terminal portion of the 15-residue COOH-terminal-half of α2-CB2, since its amino acid composition plus that of T4 of Figure 2B accounts for the amino acid content of T(3-4) (see Table II). The four possible tryptides of α2-CB2 are therefore T2 and T1 of Figure 1A, T3 of Figure 2A and T4 of Figure 2B. The sum of the amino acid analyses of these peptides accounts for the total amino acids of chick skin collagen α2-CB2.

Order of the Tryptic Peptides. The fact that tryptic treatment of α2-CB2 in NH_4HCO_3 left one arginyl bond uncleaved provided the overlap necessary for determining the order in which the tryptic peptides occur in the α2-CB2 molecule. Since the peptide (T3-4) containing this arginyl bond also contains homoserine, and since it contains fifteen residues, it must be the COOH-terminal-half of α2-CB2, as pointed out above. The NH_2-terminal-half is therefore made up of the two peptides T1 and T2, which together with T (3–4) account for the total amino acid content of α2-CB2. Which of the two peptides T1 and T2 forms the NH_2-terminal portion of α2-CB2 was determined by subjecting the untreated CNBr peptide to four stages of Edman degradation, with identification of the NH_2-terminal residue at each stage by the methods described above. This sequence was found to be Gly-Pro-Ala-Gly-.

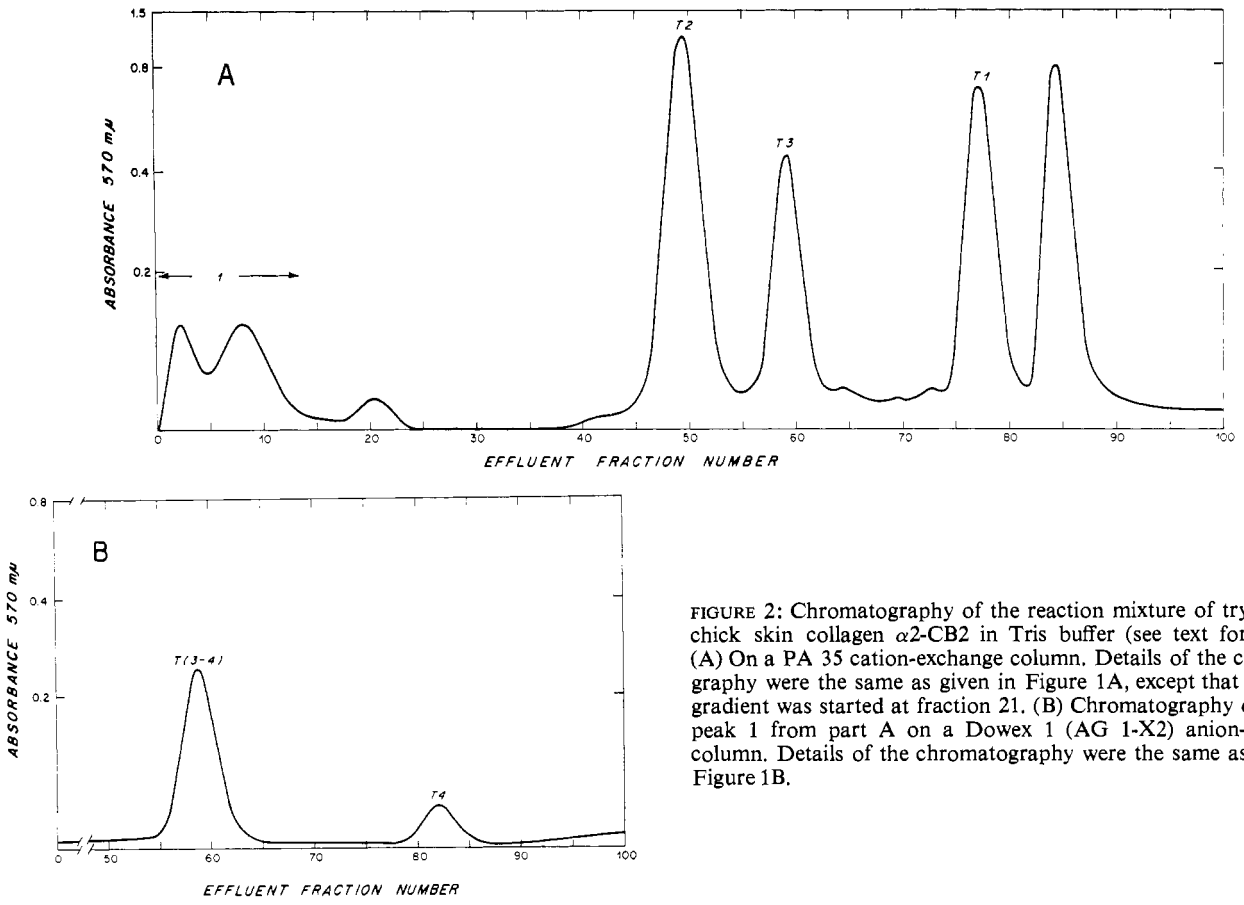

FIGURE 2: Chromatography of the reaction mixture of trypsin and chick skin collagen α2-CB2 in Tris buffer (see text for details). (A) On a PA 35 cation-exchange column. Details of the chromatography were the same as given in Figure 1A, except that the linear gradient was started at fraction 21. (B) Chromatography of pooled peak 1 from part A on a Dowex 1 (AG 1-X2) anion-exchange column. Details of the chromatography were the same as given in Figure 1B.

Since only the sequence of T1 fits this (see below), T1 must be the NH_2-terminal portion of α2-CB2, and this must be immediately followed by T2. Since, as pointed out above, T3 forms the NH_2-terminal portion of the COOH-terminal-half of the molecule, the complete order from the NH_2 terminus, of the four tryptides must be: T1, T2, T3, T4.

Amino Acid Sequence of T1 (Residues 1–6). Five stages of Edman degradation with identification of the NH_2-terminal residue at each stage, gave the sequence Gly-Pro-Ala-Gly-Asx-. The known specificity of trypsin indicated that arginine is the COOH terminus. Determination of the net charge of the DNS-peptide as zero by paper electrophoresis at pH 6.5 indicated that Asx must be Asn.

Peptide T2 (Residues 7–15). Eight stages of Edman degradation gave the sequence Gly-Ala-Ser-Gly-Pro-Ala-Gly-Val-. Lysine must be the COOH terminus from the known trypsin specificity.

Peptide T3 (Residues 16–23). Seven stages of Edman degradation gave the sequence Gly-Pro-Asx-Gly-Asx-Ala-Gly-. The known trypsin specificity indicates that arginine is the COOH terminus. Paper electrophoresis of the DNS-peptide at pH 6.5 gave a net charge of −1, showing that one Asx is Asn and one is Asp. The net charge remained unchanged through four stages of Edman degradation, indicating that the Asx at residue 18 must be Asn, and therefore that at residue 20 is Asp.

Peptide T4 (Residues 24–30). The sequence Hyp-Gly-Glx-Hyp-Gly-Leu- was deduced from six stages of Edman degradation. Homoserine must be the COOH terminus. That Glx is Glu was shown by the DNS-peptide (in the homoserine form) giving a net charge of −2 on paper electrophoresis at pH 6.5.

The complete amino acid sequence of chick skin collagen α2-CB2, deduced as described above, is shown in Figure 4.

Thermolytic Treatment. Thermolytic digestion of chick skin collagen α2-CB2 was carried out as described in the Experimental Section. The lyophilized reaction mixture was chromatographed on a PA 35 cation-exchange column under the conditions previously described. The resulting chromatogram is shown in Figure 3.

Fractions representing the seven major peaks obtained were collected, lyophilized, and analyzed. The analyses of the materials of Peaks 2 (Th2), 5 (Th3), and 6 (Th1) indicated them to be single peptides (see Table III), the sum of whose amino acid compositions accounted for the total amino acids of α2-CB2. The analyses of the materials of the remaining four peaks showed that they were peptide mixtures; since these

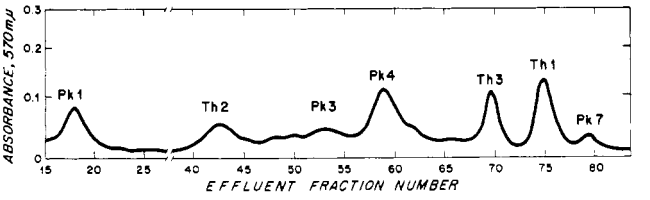

FIGURE 3: Chromatography of the reaction mixture of thermolysin and chick skin collagen α2-CB2 (see text for details) on a PA 35 cation-exchange column. Fractions of 3.0-min duration were collected. The column was first eluted with 0.1 M pyridine acetate (pH 3.1). At fraction 21 a linear gradient formed between 280 ml of this buffer and the same volume of 2.0 M pyridine acetate (pH 5.0) was started.

TABLE III: Amino Acid Composition of the Thermolytic Digestion Products of Chick Skin Collagen α2-CB2.[a]

Chromatogram Thermolytic Peptide	Figure 3 Th2 (8–28)	Th3 (29–30)	Th1 (1–7)
4-Hydroxyproline	1.7	–	–
Aspartic acid	2.1	–	1.0
Serine	1.0	–	–
Glutamic acid	1.4	–	–
Proline	2.1	–	0.9
Glycine	6.6	0.3	2.7
Alanine	2.7	–	1.0
Valine	1.0	–	–
Leucine	–	1.0	–
Lysine	1.0	–	–
Arginine	1.0	–	1.0
Homoserine[b]	–	1.0	–

[a] Composition expressed as residues per peptide. A dash indicates 0.1 residue or less. The numbers in parentheses refer to the residue numbers in Figure 4. [b] Includes homoserine lactone.

TABLE IV: Amino Acid Composition of Rat Skin Collagen α2-CB2 and Its Tryptic Digestion Products.[a]

	α2-CB2[c] (1–30)	T2 (7–15)	T1 (1–6)
4-Hydroxyproline	3	–	0.9
Aspartic acid	3	–	1.1
Threonine	1	1.0	–
Serine	1	0.9	–
Glutamic acid	1	–	–
Proline	3	0.9	0.9
Glycine	10	2.8	2.2
Alanine	2	0.9	–
Valine	1	0.9	–
Leucine	1	–	–
Arginine	3	1.3	1.0
Homoserine[b]	1	–	–

[a] Composition expressed as residues per peptide. A dash indicates 0.1 residue or less. The numbers in parentheses refer to the residue numbers in Figure 4. [b] Includes homoserine lactone. [c] Bornstein and Piez (1966).

contributed no additional information, they were not further examined.

The order of the thermolytic peptides was confirmed by determination of the NH_2-terminal residue of each, and of the following two residues in the cases of Th1 and Th2.

Peptide Th1 (Residues 1–7). Analysis showed this peptide to contain seven residues. Dansylation and two stages of Edman degradation showed that its NH_2-terminal sequence is Gly-Pro-Ala-. The peptide therefore is the seven-residue NH_2-terminal peptide formed by cleavage of the Gly–Ala bond at residues 7–8.

Peptide Th2 (Residues 8–28). This 21-residue peptide overlaps the one lysyl bond, and the arginyl–hydroxyprolyl bond of α2-CB2. Dansylation and two stages of Edman degradation showed its NH_2-terminal sequence to be Ala-Ser-Gly-. It is therefore the second of the three thermolytic peptides isolated, immediately adjacent to Th1. It is to be noted that this peptide contains an uncleaved Gly–Val bond (residues 13–14), cleavage of which would have been expected, since the adjacent lysyl ε-NH_2 group apparently does not affect the susceptibility of the bond (Ambler and Meadway, 1968). Peptides arising from cleavage of this bond would probably be found in the fractions containing mixtures, not further investigated.

Peptide Th3 (residues 29–30). The analysis indicates that this is the COOH-terminal dipeptide, leucyl-homoserine, expected to arise from cleavage of the Gly–Leu bond at residues 28–29. That leucine is NH_2 terminal was confirmed by the observation that dansylation and hydrolysis produced DNS-Leu.

The three thermolytic peptides isolated, providing new overlaps, thus provide confirmation of the sequence of chick skin collagen α2-CB2 derived from study of the tryptic digestion products. The proposed structure is shown in Figure 4.

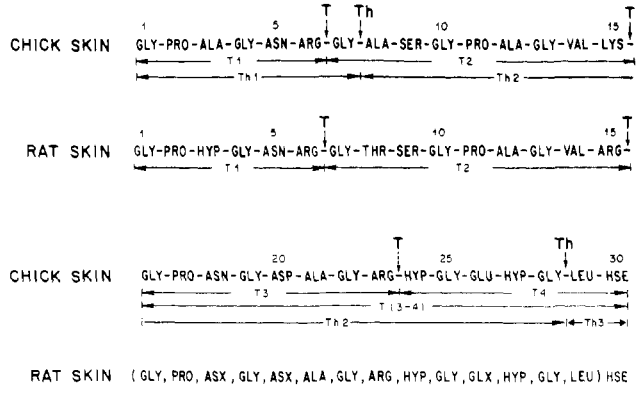

FIGURE 4: Proposed amino acid sequences of the α2-CB2 peptides from chick skin and rat skin collagens. Bonds cleaved in the enzymatic treatments are indicated by vertical arrows. The Arg–Hyp bond partially cleaved by trypsin in Tris buffer is indicated by the dashed arrow. T, trypsin; Th, thermolysin. The thermolytic peptides shown are those isolated from a reaction mixture containing other cleavage products.

Rat Skin Collagen α2-CB2

The amino acid analysis of the rat skin collagen α2-CB2 used in the present work was identical with that of Bornstein and Piez (1966). On comparing it (Table IV) to that of the chick skin collagen peptide (Table I) it is seen that the following differences exist between the two. With the same total number of residues, the rat skin peptide contains two less alanines, and one more each of hydroxyproline and threonine, then the chick skin peptide. In addition, it contains no lysine, but one more arginine, to make the same total number of basic residues.

On the basis of homology it is immediately suspected that in the rat skin peptide two of the alanine residues of the chick skin peptide have been replaced by hydroxyproline and threonine, respectively, and that one lysine of the chick skin has been replaced in the rat skin by arginine. The following results demonstrate that this is indeed the case.

Tryptic Treatment in NH_4HCO_3 Buffer. The rat skin collagen α2-CB2 was treated with trypsin in NH_4HCO_3 buffer

under conditions identical with those of the chick skin collagen α2-CB2 treatment. The lyophilized reaction mixture was chromatographed on a PA 35 cation-exchange column under the conditions previously described.

The chromatogram obtained was very similar to that of the corresponding tryptic digest of chick skin collagen α2-CB2 (Figure 1A). The same poorly resolved initial mixture (peak 1) was obtained, followed by two well-resolved peaks, and a fourth peak, which contained only ammonia. The amino acid analyses of the second and third peaks are given in Table IV and these are seen to indicate peptides of nine and six residues, respectively. These peptides are obviously homologous with T2 and T1 of Figure 1A in the case of the chick skin collagen α2-CB2. The second and third peaks were collected, lyophilized, and subjected to the Edman degradation–dansylation procedures described above. The results are given below, basing the ordering of the peptides on homology with those of chick skin collagen α2-CB2.

Amino Acid Sequence of Peptide T1 (Residues 1–6). Five stages of Edman degradation gave the sequence Gly-Pro-Hyp-Gly-Asx. Arginine must be COOH terminal from the known specificity of trypsin. Paper electrophoresis at pH 6.5 showed the DNS-peptide to have zero net charge; therefore Asx at residue 5 is Asn, as in the case of chick skin collagen peptide.

Peptide T2 (Residues 7–15). Eight stages of Edman degradation gave the sequence Gly-Thr-Ser-Gly-Pro-Ala-Gly-Val-. Arginine must be the COOH-terminal residue.

The two peptides above account for the NH_2-terminal-half of the rat skin collagen α2-CB2 molecule, and contain all of the substitutions noted above in comparison to chick skin collagen α2-CB2. Therefore the COOH-terminal-half of rat skin collagen α2-CB2 has the same composition as the corresponding half of chick skin collagen α2-CB2, and the amino acid sequence of this part of the molecule is assumed to be identical with that determined for the homologous peptide. The proposed sequence for rat skin collagen α2-CB2 is shown in Figure 4, where, however, the COOH-terminal-half is shown only compositionally, since the sequence of this part of the molecule was not actually determined in this work.

Discussion

According to Igarashi *et al.* (1970) and Vuust *et al.* (1970) α2-CB2 follows α2-CB4, and is in turn followed by α2-CB3 proceeding from the NH_2 terminus, in the sequence of CNBr peptides of the α2 chain. It therefore represents a segment of that chain about 85 Å in length and located about one-third in from the NH_2 terminus. The primary structure of this segment, in the cases of both chick skin and rat skin collagens, as shown in Figure 4, consists of the typically collagenous triplet sequence with glycine every third residue, and hydroxyproline occurring only in position 3.

Comparison of the sequences shown in Figure 4 shows that within this 30-residue segment of the α2 chain only three substitutions have occurred between chick skin and rat skin collagens. These are at residues 3, 8, and 15 (hydroxyproline for alanine, threonine for alanine, and arginine for lysine, respectively). This is in contrast to the six substitutions between the same two species in the 15-residue α2-CB1 from the NH_2-terminal region, observed by Kang and Gross (1970). These findings appear to support the concept discussed by several authors that the NH_2-terminal region of the α2 chain may be the site of evolutionary changes associated with the development of species-specific properties in collagen (Bornstein, 1968; Bornstein and Kang, 1970; Kang and Gross, 1970).

It is well known that where a basic residue is followed by proline the bond is quite resistant to tryptic cleavage. It is of interest that, as shown in the present work, hydroxylation of the proline increases the susceptibility of the bond to cleavage. The extent of cleavage of the arginyl–hydroxyproline bond during the 4-hr tryptic treatment in Tris buffer was estimated at about 60% from the amounts of peptides obtained. The question arises as to why cleavage to this extent was obtained in Tris buffer, but not in NH_4HCO_3 under the same conditions of pH, time, and temperature. It is believed that the difference arises from the fact that in the NH_4HCO_3 system a large part of the Ca^{2+} necessary for stabilization of the enzyme is precipitated. The trypsin therefore only remains active long enough to cleave such susceptible bonds as arginyl–glycine, but not long enough to permit significant cleavage of the arginyl–hydroxyproline bond. In Tris buffer, on the other hand, the full complement of Ca^{2+} is present in solution, and the trypsin remains active for much longer periods, permitting attack on the more resistant arginyl–hydroxyproline bond to the extent noted above.

Acknowledgment

We thank Dr. William T. Butler for a generous gift of rat skin collagen α2-CB2. We are grateful to Mrs. J. Hazard and Miss K. Lawley for their skillful technical assistance.

References

Ambler, R. P., and Meadway, R. J. (1968), *Biochem. J. 108*, 893.

Bornstein, P. (1968), *Science 161*, 592.

Bornstein, P., and Kang, A. H. (1970), Chemistry and Molecular Biology of the Intercellular Matrix, New York, N. Y., Academic Press, pp 99–107.

Bornstein, P., and Piez, K. A. (1966), *Biochemistry 5*, 3460.

Bradshaw, R. A. (1969), *Biochemistry 8*, 3871.

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

Deyl, Z., and Rosmus, J. (1965), *J. Chromatog. 20*, 514.

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

Gray, W. R. (1967), *Methods Enzymol. 11*, 139, 469.

Igarashi, S., Kang, A. H., and Gross, J. (1970), *Biochem. Biophys. Res. Commun. 38*, 697.

Kang, A. H., Bornstein, P., and Piez, K. A. (1967), *Biochemistry 6*, 788.

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

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

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. (1969c), *Biochemistry 8*, 3648.

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

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.

Morse, D., and Horecker, B. L. (1966), *Anal. Biochem. 14*, 429.

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

Piez, K. A., Miller, E. J., Lane, J. M., and Butler, W. T. (1969), *Biochem. Biophys. Res. Commun. 37*, 801.
Rauterberg, J., and Kühn, K. (1968), *FEBS (Fed. Eur. Biochem. Soc.) Lett. 1*, 230.
Vuust, J., Lane, J. M., Fietzek, P. P., Miller, E. J., and Piez, K. A. (1970), *Biochem. Biophys. Res. Commun. 38*, 703.
Woods, K. R., and Wang, K. T. (1967), *Biochim. Biophys. Acta 133*, 369.

Studies of the Metal Sites of Copper Proteins. Ligands of Copper in Hemocuprein*

Giuseppe Rotilio, Alessandro Finazzi Agrò, Lilia Calabrese, Francesco Bossa, Pietro Guerrieri, and Bruno Mondovì

ABSTRACT: The electron paramagnetic resonance, optical absorption, and circular dichroism spectra of bovine erythrocuprein (hemocuprein) were studied under various conditions to obtain information on the copper binding site. When the pH of the protein solution was raised to pH 11.5, a superhyperfine pattern of nine lines with spacing of about 14 G was observed in the electron paramagnetic resonance spectrum of the protein-bound copper. This spectrum can be accounted for by the presence of three to four nitrogen atoms as ligands of copper. The original spectrum was fully recovered by lowering the pH. Also the parallel changes noticed in the absorption and circular dichroism spectra were reversible. The circular dichroism spectrum at pH 11.5 displayed the same multiplicity of ellipticity bands as at neutral pH, only the relative intensity of the peaks being changed; this strongly supports the idea that the copper site is still in its native state at the pH where the nitrogen hyperfine pattern is observed. On the other hand, raising the pH above pH 12 brought about another type of nitrogen hyperfine pattern in the electron paramagnetic spectrum of the protein and drastic changes in optical spectra, which are related to an irreversible denaturation of the protein.

Addition of cyanide was also able to induce a reversible appearance of nitrogen hyperfine pattern, but the modification of the copper site seemed more profound in this case, as indicated by the circular dichroism spectrum in the presence of cyanide. The effect of exchanging deuterium oxide for water in the electron paramagnetic resonance spectrum of hemocuprein suggests that the copper site is exposed to the solvent. Tryptophanyl, tyrosyl, and sulfhydryl residues seem not to be involved in the direct binding of the copper.

In a search for direct information on ligands of copper in copper proteins, we have extended our studies (Finazzi Agrò *et al.*, 1970) to hemocuprein (bovine erythrocuprein), the copper protein of ox red blood cells, for which an enzymatic function, *i.e.*, superoxide dismutase activity, has recently been claimed (McCord and Fridovich, 1969). This protein has a molecular weight of approximately 34,000 and contains 2 Cu(II)/molecule. Copper appears to be essential to the catalytic action, and cannot be replaced by other ions in this role (McCord and Fridovich, 1969). Thus hemocuprein is a new copper enzyme which lacks a detailed physicochemical characterization as yet. In spite of that, it seemed to us particularly suited for a study of the mode of binding of the copper ion, since its metal site appears to be intermediate, as far as can be argued from spectroscopic data, between the unusual copper coordination of "blue" copper proteins and the copper complexes of model peptides. In fact the electron paramagnetic resonance parameters of human erythrocuprein (Malmström and Vänngård, 1960), which is most probably very similar to the bovine protein, can be accounted for by a symmetry only slightly distorted from square planar. Moreover the protein is blue-green (maximum around 680 nm) with an extinction coefficient ($\epsilon_{680} \simeq 300$ cm^{-1} M^{-1}) which is significantly higher than that of low molecular weight copper complexes in this region (Brill *et al.*, 1964) but very much lower than in the case of blue copper proteins. In the present paper we report the results of spectroscopic studies, namely electron paramagnetic resonance, circular dichroism, and optical absorption spectra in different conditions, of hemocuprein as a first approach to the description of its copper site in terms of ligand atoms and in comparison with what is known about the far more deeply analyzed metal site of the blue copper proteins.

Material and Methods

All chemicals were reagent grade, and were used without further purification. Hemocuprein was purified from ox blood according to McCord and Fridovich (1969). Protein concentration was measured by a biuret procedure (Goa, 1953) or by optical absorption at 258 and 680 nm (McCord and Fridovich, 1969); the values obtained by different methods were in good agreement with each other.

Electron paramagnetic resonance spectra were recorded on a Varian 4502-14 spectrometer, equipped with 100-KHz field modulation within a Varian multipurpose cavity, and variable-temperature accessory. Frequencies were measured

* From the Institute of Biological Chemistry, University of Rome, and the Centre for Molecular Biology, C. N. R., Rome, Italy. *Received August 21, 1970.* This is the second paper of a series on the metal sites of copper proteins. The preceding paper has been published (Finazzi Agró *et al.*, 1970).