The Limited Cleavage of Native Collagen with Chymotrypsin, Trypsin, and Cyanogen Bromide*

Paul Bornstein, Andrew H. Kang,† and Karl A. Piez

ABSTRACT: Soluble rat skin collagen was cleaved with chymotrypsin, trypsin, and CNBr under conditions which did not disrupt the helical conformation of the protein. Chymotrypsin and CNBr reduced the content of double-chain β components and converted both α chains and β components to altered α chains. The amino acid compositions of the products of chymotryptic digestion ($\alpha 1^{Chy}$ and $\alpha 2^{Chy}$) resembled that of the original $\alpha 1$ and $\alpha 2$ chains but were characterized by lower tyrosine and methionine contents. $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ each possessed a single new N-terminal glycine and each lacked a sequence from the cross-link region of $\alpha 1$ and $\alpha 2$. Previous studies utilizing complete cleavage of α chains with CNBr have shown that these sequences give rise to peptides containing tyrosine and methionine-derived homoserine. These peptides exist in two forms in digests of both chains, a lysyl-containing form ($\alpha 1$-CB-1 and $\alpha 2$-CB-1) and a lysyl-derived aldehyde-containing form ($\alpha 1$-CB-1a and $\alpha 2$-CB-1a). The aldehydes have been implicated in the formation of intramolecular cross-links. Lysine is in position five in both $\alpha 1$-CB-1, which contains 15 amino acids, and $\alpha 2$-CB-1, which contains 14 amino acids. The altered α chains produced by limited cleavage with CNBr ($\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$) resembled $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ in their amino acid composition and chromatographic and electrophoretic behavior. In addition the small peptides released by CNBr from a collagen in which lysyl-derived aldehydes were largely absent were shown to be $\alpha 1$-CB-1 and $\alpha 2$-CB-1. These data indicate that both chymotrypsin and CNBr cleave the chains of native collagen C terminal to the cross-linking site, near the N terminus of the molecule. This region is susceptible to enzymatic and chemical cleavage presumably because its amino acid composition prohibits the formation of the helical conformation typical of most of the collagen molecule. Because of its more limited specificity trypsin does not attack the N-terminal region of most native collagen preparations. However, in collagen in which the lysyl residues in $\alpha 1$-CB-1 and $\alpha 2$-CB-1 have not undergone transformation to aldehydes, cleavage does occur at these positions. Two peptides were isolated and their amino acid compositions were found to be identical with the first five amino acids in $\alpha 1$-CB-1 and $\alpha 2$-CB-1. These results confirm the findings of the chymotrypsin and CNBr experiments and demonstrate that the sequences represented by $\alpha 1$-CB-1 and $\alpha 2$-CB-1 are actually at the N-terminal ends of the $\alpha 1$ and $\alpha 2$ chains. The aldehyde–precursor lysyl residues are therefore located five amino acids from the N-terminal end of the collagen molecule.

The collagen molecule has dimensions of approximately 3000 × 15 A and consists of three α chains, each with a molecular weight of about 95,000. The chain structure of rat skin collagen may be expressed as $(\alpha 1)_2\alpha 2$ (Piez *et al.*, 1963) although the identity of the two $\alpha 1$ chains has not been established. The available evidence indicates that the three chains are parallel in direction and extend the full length of the molecule (Kang *et al.*, 1966) in a configuration which is largely helical. With time, interchain covalent links form in collagen fibrils, both between chains within the same molecule and intermolecularly. This process apparently takes place in specialized regions of the molecule.

Recent studies on peptides derived from the cross-link region of soluble rat skin collagen by cyanogen bromide (CNBr) cleavage (Bornstein *et al.*, 1966; Bornstein and Piez, 1966) have led to the conclusion that the intramolecular interchain cross-link is derived biosynthetically from the side chains of two lysyl residues on adjacent α chains. Consistent data have been obtained from studies of the chromatographic and metabolic heterogeneity of α chains utilizing the incorporation of isotopes (Piez *et al.*, 1966). Apparently only one specific lysyl residue on each α chain is available for the reaction, at least in the early stages of cross-linking. The δ-semialdehyde of α-aminoadipic acid (in peptide linkage) has been identified as an intermediate in the formation of the cross-link but the covalent bond that is formed has not been completely characterized. It may result from the condensation of two aldehyde groups to produce a bond which reacts as an unsaturated aldehyde after isolation in a CNBr-produced peptide.

Information relevant to the location of the cross-link in collagen has come from two types of studies.

* From the Laboratory of Biochemistry, National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland 20014. *Received August 25, 1966.* A portion of this work has appeared in preliminary form (Borstein *et al.*, 1966).

† Present address: Peter Bent Brigham Hospital, Boston, Mass.

It has been shown that proteolytic enzymes such as pepsin, chymotrypsin, and pronase act on native collagen to release acidic peptides constituting about 1% of its mass without a loss of helical structure (Hodge *et al.*, 1960; Rubin *et al.*, 1963, 1965; Drake *et al.*, 1966). In the process, the cross-linked β (double chain) components are converted to products very similar to α chains (Kühn *et al.*, 1963; Rubin *et al.*, 1965; Martin *et al.*, 1966). This suggests that cross-links may form in nonhelical regions of the molecules. The characterization of the two fragments produced by the digestion of native collagen with an enzyme from the tadpole has also shown that intramolecular cross-links do not form in the quarter of the molecule which includes the C-terminal ends of the α chains (Nagai *et al.*, 1965; Kang *et al.*, 1966).

The present communication is concerned primarily with more detailed studies on the location of intramolecular cross-links. This problem has been approached by the characterization of the products resulting from the proteolytic action of chymotrypsin and trypsin on native collagen and from the selective cleavage of methionyl bonds in native collagen by CNBr.

Materials and Methods

Preparation of Collagen. Rat skin collagen was obtained from the skins of 100–150-g male Sprague–Dawley rats. Acid-extracted normal collagen and 1 M NaCl-extracted collagen from lathyritic rats were prepared as described previously (Bornstein and Piez, 1964, 1966).

Chromatography on Carboxymethylcellulose and Phosphocellulose. Heat-denatured collagen was fractionated on columns of CM-cellulose and fractions were isolated by modifications (Bornstein and Piez, 1966) of previously described methods (Piez *et al.*, 1963; Bornstein and Piez, 1964). The methods for chromatography of CNBr-produced collagen fragments on phosphocellulose and their separation from buffer salts by molecular sieve chromatography on Bio-Gel P-2 have also been described (Bornstein and Piez, 1966).

Digestion of Native Collagen with Chymotrypsin. Acid-extracted collagen was dissolved at a concentration of 1 mg/ml in 1 M NaCl, containing 0.005 M $CaCl_2$, and 0.05 M Tris–HCl buffer, pH 7.5, by stirring overnight at 4°. The collagen solution was then dialyzed in the cold against 20 volumes of the same buffer and centrifuged at 4° for 1 hr at 30,000*g*. The residue, if any, was discarded. Sufficient α-chymotrypsin (three times recrystallized, Worthington) was dissolved in the same solvent to give a final weight ratio of collagen to enzyme of 10:1. The collagen solution and enzyme were placed separately in a constant-temperature water bath maintained at 15°. Prior to the addition of enzyme, 2 ml of the collagen solution was removed and diluted with 2 ml of buffer. This solution served as a control for viscometry. The enzyme was then added to the collagen and after thorough mixing a small aliquot was removed and diluted with buffer to provide a collagen concentration identical with that in the control solution in total of 4 ml. This solution served to monitor changes in viscosity during the course of digestion. The flask containing collagen and enzyme was flushed with nitrogen and incubated at 15° for 24 hr. At the conclusion of the incubation, 1 N acetic acid was added to reduce the pH to 4. The solution was chilled to 4° and sufficient 25% NaCl was added to increase the salt concentration to 7.5%. All subsequent operations were performed in the cold. The collagen that had precipitated was stirred as a suspension for 1 hr and centrifuged at 10,000*g* for 30 min. The supernatant was discarded and the precipitate was dissolved by the addition of 1% acetic acid. The cycle of precipitation and redissolution was repeated four times to ensure complete removal of the enzyme. It was found convenient to perform the initial precipitation and subsequent procedures in plastic-coated, 300-ml metal centrifuge cups to avoid repeated transfer of the precipitate. At the conclusion of this purification the collagen solution was dialyzed exhaustively against 0.5% acetic acid and lyophilized.

Digestion of Native Collagen with Trypsin. Neutral salt-extracted collagen from lathyritic rats was dissolved at a concentration of 5 mg/ml in 0.5 M NH_4HCO_3, pH 7.8, by stirring overnight at 4°. The collagen solution was then dialyzed in the cold against 30 volumes of 0.3 M NH_4HCO_3, pH 7.8. Trypsin (two times crystallized, Worthington) was dissolved at a concentration of 10 mg/ml in 1×10^{-3} M HCl containing 1×10^{-3} M $CaCl_2$ and dialyzed at 4° against a large volume of the same solvent. Sufficient trypsin solution was added to the collagen solution to give a weight ratio of collagen to enzyme of 50:1 and 1 M $CaCl_2$ was added to a final concentration of 1×10^{-3} M. Digestion was carried out at 20° for 4 hr. The contents of the incubation mixture were dialyzed for 16 hr against ten volumes of water at 4°. The dialysate was lyophilized and the residue was dissolved in a small volume of water for chromatography on Bio-Gel P-2.

Viscometry. Viscosity measurements were performed in Ostwald capillary viscometers with a flow time for water of approximately 100 sec at 15°. The temperature was maintained at 15 ± 0.02°. Chymotrypsin-treated collagen was dissolved in 0.15 M, pH 4.8 potassium acetate buffer. Protein concentrations were determined by measuring the optical rotation of the denatured protein at 313 mμ and 50°. The specific rotation of gelatin under these conditions, determined on samples whose concentration was measured by micro-Kjeldahl, was −825°. A nitrogen content of 18.6% was assumed for collagen.

Limited Cleavage with CNBr. Native collagen was dissolved in cold, distilled water at a concentration of 3 mg/ml by stirring at 4° for 16–24 hr. Any undissolved material was removed by centrifugation at 30,000*g* for 1 hr. Sufficient 6 N HCl was added slowly with constant stirring to make a 0.1 N HCl solution.

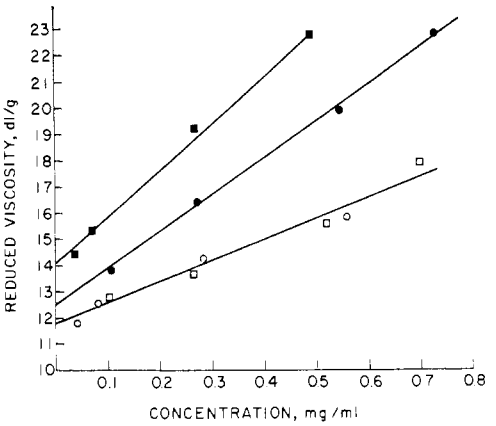

FIGURE 1: Concentration dependence of reduced viscosity of acid-extracted (■,□) and salt-extracted (●,○) collagen before (■,●) and after (□,○) limited cleavage with chymotrypsin.

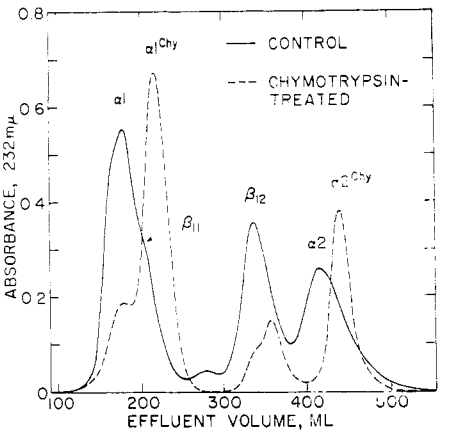

FIGURE 2: Superimposed elution patterns of control acid-extracted collagen (130 mg) and chymotrypsin-treated collagen (100 mg). The denatured collagens were dissolved in 30 ml of starting buffer, 0.06 M sodium acetate, pH 4.8, and chromatographed on CM-cellulose at 40°. A linear gradient of NaCl from 0 to 0.1 M over a total volume of 820 ml was employed.

The flask was flushed with nitrogen and solid CNBr (Eastman Organic Chemicals) was added in a 100-fold molar (relative to methionine) excess. The flask was stoppered tightly, stirred manually for a few minutes to dissolve the CNBr, and incubated at 15° for 24 hr. At the conclusion of the incubation the contents of the flask were diluted tenfold with cold water and lyophilized. The residue was redissolved in cold water by stirring overnight at 4° and lyophilized again.

Polarimetry. Optical rotation was measured in a Rudolph photoelectric spectropolarimeter equipped with a monochromator and an oscillating polarizer. Acid-extracted normal collagen was dissolved in cold water at a concentration of 1 mg/ml. After dialysis against several changes of cold water the collagen solution was clarified by centrifugation at 100,000*g* in a Spinco Model L ultracentrifuge. Limited cleavage with CNBr was then performed as described above in a 1-dm polarimeter cell jacketed at 15°. Measurements were made at the 313-mμ mercury line over a period of 24 hr.

Amino Acid Analysis. Analyses were performed on a single-column automatic amino acid analyzer modified for high-speed analysis (Miller and Piez, 1966). Previously determined corrections were made for losses of serine, threonine, methionine, and tyrosine and incomplete release of valine (Piez *et al.*, 1960).

N-Terminal Analyses. Amino-terminal analyses were performed on isolated α1 and α2 chains of acid-extracted normal collagen and on the chains isolated by CM-cellulose chromatography of chymotrypsin-treated native collagen. Both the dinitrophenylation and the phenylisothiocyanate methods, as described for the analysis of collagen chains (Piez *et al.*, 1966), were employed.

Acrylamide Gel Electrophoresis. The method of Nagai *et al.* (1964) was used.

Results

Limited Cleavage of Collagen with Chymotrypsin

Viscometry. A small but definite decrease in relative viscosity was observed after incubation of native collagen with chymotrypsin. This effect was noted primarily in the first 6–8 hr and tended to plateau after 15–24 hr when a 1:10 enzyme:substrate weight ratio was employed at 15°. Similar findings have been reported by Martin *et al.* (1966). Although the change in relative viscosity was most marked with acid-extracted collagen, smaller reductions were consistently observed with salt-extracted collagen and with collagen from lathyritic rats. The effect of chymotrypsin on the reduced viscosity of collagen is illustrated in Figure 1 in which it can be seen that both the intrinsic viscosity and the concentration dependence of reduced viscosity is less for chymotrypsin-treated collagen than for the control acid- or salt-extracted collagens. The higher intrinsic viscosity and concentration dependence of acid-extracted as compared with salt-extracted collagen is consistent with the presence of a greater proportion of high molecular weight aggregates in the former preparation. The differences disappear after chymotrypsin treatment. These results can be accounted for by the action of chymotrypsin in converting solutions of collagen containing varying amounts of aggregates to solutions which are largely monomeric. Drake *et al.* (1966) have reached similar conclusions in their studies of the action of enzymes on soluble native collagen.

Chromatography on CM-Cellulose. The elution patterns of chymotrypsin-treated and control acid-extracted collagens chromatographed in the denatured state on CM-cellulose are illustrated in Figure 2. It is evident that a significant reduction in β components

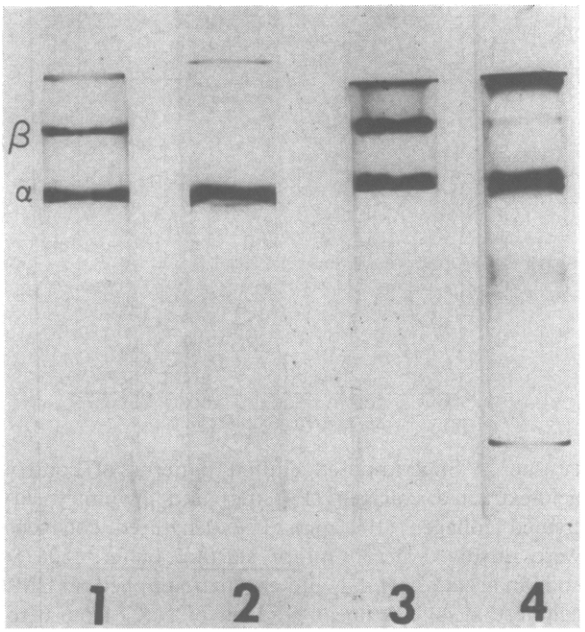

FIGURE 3: Acrylamide gel electrophoresis patterns of acid-extracted collagen before treatment (tubes 1 and 3) and after limited cleavage with chymotrypsin (tube 2) and CNBr (tube 4). Migration is toward the cathode from top to bottom.

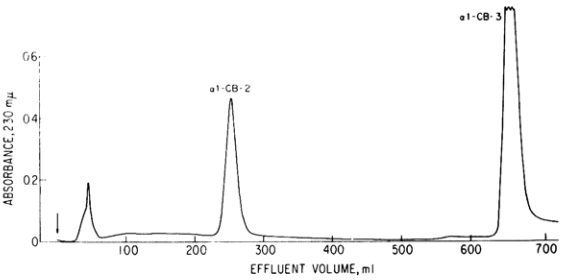

FIGURE 4: Elution pattern of the low molecular weight peptide fraction obtained by CNBr cleavage of $\alpha 1^{Chy}$. Chromatography was performed on phosphocellulose at pH 3.8, 40°. The arrow indicates placement of the sample (120 mg) dissolved in 10 ml of starting buffer, 0.001 M sodium acetate, pH 3.8. A linear gradient of NaCl from 0 to 0.3 M over a total volume of 820 ml was employed. The initial peak consists of nonpeptide material eluting at the gradient front.

occurred during chymotryptic digestion. In addition, the two peaks $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ chromatographing near the original $\alpha 1$ and $\alpha 2$ chains were eluted at a higher effluent volume and showed a narrower distribution, indicating that they are more basic and perhaps more homogeneous than the original chains. The amount of residual β_{12} as well as the shoulder preceding the $\alpha 1^{Chy}$ peak (presumably unchanged $\alpha 1$) could be reduced by using a temperature of 20° or a higher enzyme:substrate ratio but these conditions frequently resulted in poorer resolution of the products by CM-cellulose chromatography suggesting a less selective cleavage. When native salt-extracted collagen or lathyritic collagen was similarly treated with chymotrypsin and chromatographed on CM-cellulose the changes observed were essentially the same although the reduction in β components was less striking since the content of β components in the original preparations was low. These results suggest that chymotrypsin removed small, relatively acidic, sequences from both $\alpha 1$ and $\alpha 2$ (lost during purification of the enzyme-treated protein) giving rise to altered α chains ($\alpha 1^{Chy}$ and $\alpha 2^{Chy}$) which therefore were more basic. Since recoveries of chymotrypsin-treated lathyritic, salt-extracted, and acid-extracted collagens routinely ranged between 80 and 95% after purification, $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ must be derived from both α chains and β components.

Acrylamide Gel Electrophoresis. Figure 3 illustrates gel patterns of control acid-extracted collagen and the products of chymotrypsin treatment. In keeping with the results of chromatography, it can be seen that the β components largely disappeared after chymotrypsin treatment. Under these electrophoretic conditions $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ migrated together in the region of the α chains of the control sample. Since the separation was largely on the basis of molecular size this finding is consistent with the conclusions reached by amino acid analyses and CNBr cleavage of $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ (below) that only a small portion of each α chain was removed by chymotrypsin. The absence of significant bands migrating between the altered α chains and the buffer front indicates that random cleavage of the collagen molecule was minimal under the conditions of these experiments.

Cleavage of $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ by CNBr. In order to examine in greater detail the degree to which the $\alpha 1$ and $\alpha 2$ chains were altered by chymotrypsin, $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ were cleaved with CNBr and the resulting peptides were chromatographed on CM-cellulose under conditions which resolve the high molecular weight peptides (Bornstein and Piez, 1965) and on phosphocellulose using conditions suitable for the low molecular weight peptides (Bornstein and Piez, 1966). The CM-cellulose elution patterns of CNBr-cleaved $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ were identical with the patterns observed for digests of $\alpha 1$ and $\alpha 2$. However, the phosphocellulose elution patterns of the small peptides from $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ differed significantly from the elution patterns of the peptides from $\alpha 1$ and $\alpha 2$. Figure 4 illustrates the phosphocellulose elution pattern of a CNBr digest of $\alpha 1^{Chy}$. By comparison with chromatograms of peptides derived from unaltered $\alpha 1$ (see Figure 2; Bornstein and Piez, 1966), it can be seen that peptides $\alpha 1$-CB-1a and $\alpha 1$-CB-1 were not present in the digestion mixture. These peptides, which elute at 100- and 300-ml effluent volume, respectively, and contain 15 amino acids, are related forms of the same peptide. $\alpha 1$-CB-1 contains a lysyl residue while $\alpha 1$-CB-1a contains instead a lysyl-derived aldehyde. The aldehyde has been implicated

in the formation of the interchain cross-link in collagen (Bornstein *et al.*, 1966; Bornstein and Piez, 1966). The major peaks present (Figure 4) correspond in chromatographic position and amino acid composition to peptides α1-CB-2 and α1-CB-3 derived from intact α1 chains. These peptides are derived from the helical regions of the α chains and are not involved in cross-link formation. The peak eluting at the start of the ionic strength gradient was composed of ultraviolet-absorbing nonpeptide material. Figure 5 illustrates the phosphocellulose elution pattern of the CNBr digest of $\alpha 2^{Chy}$. Again, by comparison to a similar chromatogram of a digest of α2 (see Figure 4; Bornstein and Piez, 1966) the cross-link precursor peptides, α2-CB-1 and α2-CB-1a, were absent whereas α2-CB-2 remains unchanged both in chromatographic position and amino acid composition. α2-CB-1 and α2-CB-1a are analogous to α1-CB-1 and α1-CB-1a but contain 14 amino acids.

These results demonstrate that chymotrypsin cleaves the α1 and α2 chains in those regions of the chains which give rise to α1-CB-1 (or α1-CB-1a) and α2-CB-1 (or α2-CB-1a). The simplest explanation is that these peptides constitute terminal sequences. If any other changes result from the action of chymotrypsin under the conditions of these experiments they could not be detected either by changes in the chromatographic behavior of other peptides in CNBr digests or in the amino acid composition of the altered α chains (see below).

The observation that $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ when chromatographed on CM-cellulose form sharper and more symmetrical peaks than the original α1 and α2 chains is consistent with the demonstration (Piez *et al.*, 1966) that α1 and α2 from soluble rat skin collagen are chromatographically heterogeneous owing to the presence of lysine- and aldehyde-containing forms. The altered α chains are more homogeneous since the sequences containing the lysyl residue or the lysyl-derived aldehyde are removed by chymotrypsin.

The N-Terminal Residues of $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$. In agreement with other studies (Piez *et al.*, 1966), isolated α1 chains of control acid-extracted collagen possessed less than 0.1 residue of N-terminal glycine/chain (molecular weight taken to be 95,000) and no N-terminal residues were detected on α2 chains. However, after chymotrypsin treatment, analysis by both the DNP and PTH[1] methods indicated the presence of 1 equiv of N-terminal glycine/equiv of $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ (Table I). No other free N-terminal residue was present in excess of 0.1 equiv. Since only a small part of the α1 and α2 chains was removed by chymotrypsin, these results indicate that cleavage occurred near the N termini of the chains. However, the additional removal of a small portion of the chains near the C-terminal ends cannot be excluded by this evidence alone.

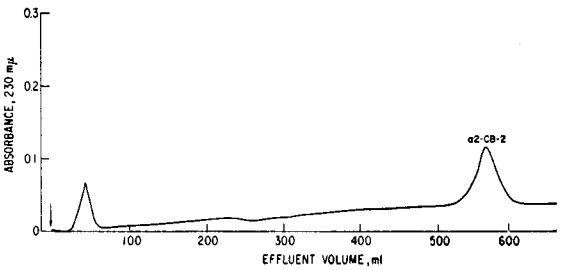

FIGURE 5: Phosphocellulose elution pattern of the low molecular weight peptide fraction obtained by CNBr cleavage of $\alpha 2^{Chy}$ (70 mg). The initial peak consists of nonpeptide material eluting at the gradient front. See legend to Figure 4 for further details.

TABLE I: Equivalents of N-Terminal Glycine[a] in $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$.

	Av	Range	Anal. (no.)
$\alpha 1^{Chy}$			
DNP[b]	1.01	0.78–1.2	4
PTH[c]	0.87	0.74–1.0	2
$\alpha 2^{Chy}$			
DNP[b]	0.93	0.88–1.1	3
PTH[c]	0.87	0.82–0.91	2

[a] No N terminals other than glycine were found in significant amounts. [b] Dinitrophenyl derivative. [c] Phenylthiohydantoin derivative.

Amino Acid Analysis of $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$. The amino acid contents of $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ are tabulated in Table II. They are almost identical with those of control α1 and α2 chains with the possible exception of a reduction of one residue each of methionine and tyrosine per chain (about 1000 residues) in both $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$. The other differences are within the analytic error of the determinations. The significance of the reduction in methionine and tyrosine would ordinarily be open to question in view of the difficulty in determining these amino acids accurately in small quantities. However, in addition to their reproducibility, support for the validity of these findings comes from the amino acid analysis of those CNBr-produced peptides which are missing after chymotryptic cleavage. These peptides, α1-CB-1 and its derivative α1-CB-1a, and α2-CB-1 and its derivative α2-CB-1a, each contain one tyrosyl and one methionine-derived homoseryl residue (Bornstein and Piez, 1966; see Table III).

Limited Cleavage of Collagen with CNBr

In view of the presence of methionine in a region of the native collagen molecule particularly susceptible to digestion by chymotrypsin and other enzymes, it seemed possible that CNBr might selectively cleave

[1] Abbreviations used: PTH, phenylthiohydantoin derivative, SLS, segment long spacing.

TABLE II: Amino Acid Composition of $\alpha 1$ and $\alpha 2$ and the Altered Chains Resulting from Limited Cleavage with Chymotrypsin ($\alpha 1^{Chy}$ and $\alpha 2^{Chy}$) and CNBr ($\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$).[a]

	Residues/1000 Residues					
	$\alpha 1$	$\alpha 1^{Chy}$	$\alpha 1^{CNBr}$	$\alpha 2$	$\alpha 2^{Chy}$	$\alpha 2^{CNBr}$
3-Hydroxyproline	0.8	0.6	0.6	0	0	0
4-Hydroxyproline	97	100	102	89	88	84
Aspartic acid	45	46	45	45	44	43
Threonine	19.4	20.0	19.4	20.1	20.1	19.0
Serine	40	40	42	44	42	40
Homoserine	—	—	0	—	—	0
Glutamic acid	74	74	74	68	69	71
Proline	129	129	128	112	111	110
Glycine	330	331	333	330	336	334
Alanine	114	114	116	103	107	107
Valine	18.4	15.3	13.6	31	31	34
Methionine	7.6	6.7	6.0	5.4	4.2	3.6
Isoleucine	6.2	5.7	4.8	15.7	16.4	18.0
Leucine	18.7	18.8	19.5	34	34	34
Tyrosine	1.8	1.1	1.2	3.1	2.0	1.0
Phenylalanine	12.0	11.5	12.1	9.9	9.5	10.6
Hydroxylysine	5.1	5.1	4.5	9.7	9.9	8.7
Amide N	(40)	(39)	(42)	(41)	(42)	(44)
Lysine	30	32	30	21	22	23
Histidine	1.9	2.2	2.1	8.5	9.2	7.9
Arginine	51	50	48	50	51	53

[a] Corrections were made for hydrolytic loss of threonine, serine, methionine, and tyrosine and incomplete release of valine. Methionine includes methionine sulfoxide. Homoserine includes homoserine lactone.

TABLE III: Amino Acid Composition of Peptides Resulting from the Limited Cleavage of Lathyritic Collagen with CNBr and Trypsin.[a]

	CNBr		Trypsin	
	$\alpha 1$-CB-1	$\alpha 2$-CB-1	$\alpha 1$-T-1	$\alpha 2$-T-1
Aspartic acid	1.0	0.9	1.0	1.0
Serine	2.0	2.0	(0.1)	1.1
Homoserine	0.9	1.0	0	0
Glutamic acid	1.1	1.0	1.0	1.0
Proline	1.9	2.0	0	0
Glycine	3.1	3.1	1.1	(0.3)
Alanine	1.0	1.0	(0.1)	(0.1)
Valine	2.0	1.2	0	0
Leucine	0	0	0	(0.2)
Tyrosine	1.0	1.1	1.0	1.0
Lysine	1.0	0.7	0.9	0.8

[a] The values are given as residues per peptide. No amino acid other than those listed was present in amount greater than 0.1 residue. Residues in parentheses are fractional residues present as impurities. Corrections were made for hydrolytic loss of serine and tyrosine and incomplete release of valine. Homoserine includes homoserine lactone.

these methionyl bonds while leaving intact the potentially susceptible bonds in helical regions. Incubation of native collagen with CNBr at 15° resulted in changes which qualitatively resembled those produced by chymotrypsin. A reduction in viscosity similar to that observed with enzymatic digestion was noted during the first 12–24 hr of the reaction. However, thereafter a very gradual fall in viscosity persisted for as long as the reaction was followed (72 hr). No change in optical rotation of a solution of native collagen treated with CNBr at 15° was noted after 24 hr.

Chromatography on CM-Cellulose. Acid-extracted collagen, treated with CNBr at 15° for 24 hr, was denatured by warming at 40–45° for 20 min and then chromatographed on CM-cellulose. The resulting chromatogram was compared with a sample of untreated collagen chromatographed under identical conditions (Figure 6). The significant differences are the presence in the chromatogram of the treated collagen of a larger amount of material eluting at the start of the gradient and the reduction in β components with a corresponding increase in the material chromatographing in the region of the α chains ($\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$). The recoveries from the CM-cellulose column of the CNBr-treated material and the control collagen were estimated by placing a weighed quantity of protein on the column and measuring the areas under

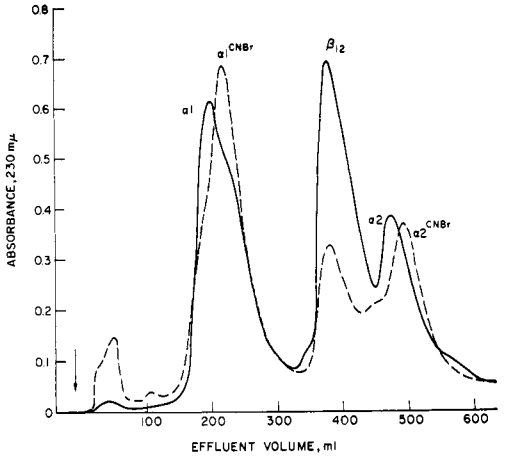

FIGURE 6: Superimposed elution patterns of control acid-extracted collagen (110 mg, solid line) and acid-extracted collagen cleaved in the native state with CNBr (90 mg, dashed line). The collagens were chromatographed on CM-cellulose in the denatured state. See legend to Figure 2 for further details.

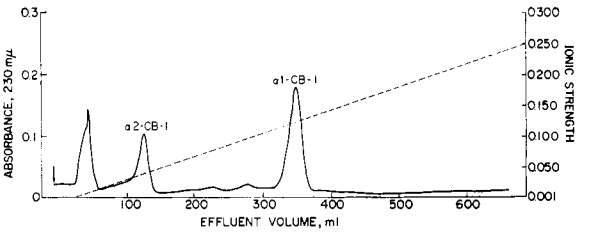

FIGURE 7: Phosphocellulose elution pattern of the peptides obtained by limited CNBr cleavage of 250 mg of salt-extracted collagen from lathyritic rats. The initial peak consists largely of nonpeptide material eluting at the gradient front. See legend to Figure 4 for further details.

the curves in the elution patterns. Recoveries of control and CNBr-treated collagen agreed within 5%. The increase in the material chromatographing in the region of the $\alpha 1$ and $\alpha 2$ chains must therefore have been at the expense of the β components indicating, as in the case of chymotrypsin cleavage, a conversion of both α chains and β components to chains resembling α chains in their chromatographic behavior. As in the case of $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$, $\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$ eluted at slightly higher effluent volumes than the control α chains consistent with an increase in their net positive charge. When native collagen was incubated in 0.1 N HCl at 15° for 24 hr in the absence of CNBr and subsequently chromatographed on CM-cellulose, the elution pattern did not differ significantly from that of the control sample in Figure 6. The changes observed in the treated samples must therefore have resulted specifically from the action of CNBr.

Acrylamide Gel Electrophoresis. The electrophoretic patterns on polyacrylamide gel of CNBr-treated collagen are shown in Figure 3. The reduction in β components is again apparent. $\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$ migrated with $\alpha 1$ and $\alpha 2$ suggesting that CNBr, like chymotrypsin, removed only a small portion of each chain under the conditions of these experiments. The more intense band at the buffer front represents low molecular weight material resulting from CNBr cleavage. In the case of the chymotrypsin-treated collagen, peptide material would have been lost during purification by precipitation. The protein migrating more slowly than β components in the CNBr-treated collagen may represent aggregates that were too large to enter the running gel in the control preparation and formed a sharp band at the interface, but were partially degraded by the reagent and formed a broad band extending into the running gel. The faint bands intermediate between the altered α chains and the buffer front may have resulted from a small amount of more extensive cleavage to fragments smaller than α chains.

Chromatography on Phosphocellulose. If the action of CNBr on native collagen resembles that of chymotrypsin, it should be possible to demonstrate the formation by limited cleavage with CNBr of precisely those peptides representing the missing sequences in $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$. Collagen from lathyritic rats was chosen for this experiment since the small percentage of β components and α chains with lysyl-derived aldehyde residues in this preparation (Bornstein and Piez, 1966) would simplify the expected peptide mixture. That is, $\alpha 1$-CB-1a, $\alpha 2$-CB-1a, and β_{12}-CB-1 (the peptide from β_{12} containing the cross-link) would be largely absent. After removal of CNBr the partially cleaved protein was denatured and chromatographed on phosphocellulose. The elution pattern (Figure 7) revealed the presence of two peaks chromatographing in the position expected for peptides $\alpha 1$-CB-1 and $\alpha 2$-CB-1. This identification was confirmed by amino acid analysis (Table III, compare with Table I; Bornstein and Piez, 1966). The peptides were present in a ratio very close to 2:1 consistent with the 2:1 ratio of $\alpha 1$ and $\alpha 2$ in rat skin collagen (Piez *et al.*, 1963). The presence of somewhat more than one residue of valine and less than one residue of lysine in $\alpha 2$-CB-1 (Table III) suggested a small amount of contamination of this peptide with $\alpha 1$-CB-1a. $\alpha 1$-CB-1a is present to a small extent in the preparation used and does not separate from $\alpha 2$-CB-1 during elution from phosphocellulose. Its composition differs from $\alpha 2$-CB-1 only in that it lacks lysine and contains 2 equiv of valine. The initial peak at the start of the ionic strength gradient consisted largely of nonpeptide material but amino acid analysis indicated that it also contained a small amount of $\alpha 2$-CB-1a. There was no evidence for the presence of any other low molecular weight products.

Amino Acid Analysis of $\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$. Amino acid analysis of $\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$ (Table II) demonstrated that these altered α chains resemble $\alpha 1^{Chy}$ and $\alpha 2^{Chy}$ in their reduced content of methionine

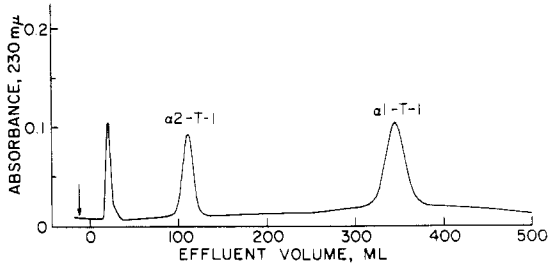

FIGURE 8: Phosphocellulose elution pattern of the dialyzable peptides obtained by limited tryptic cleavage of 300 mg of salt-extracted collagen from lathyritic rats. The initial peak consists of nonpeptide material eluting at the gradient front. See legend to Figure 4 for further details.

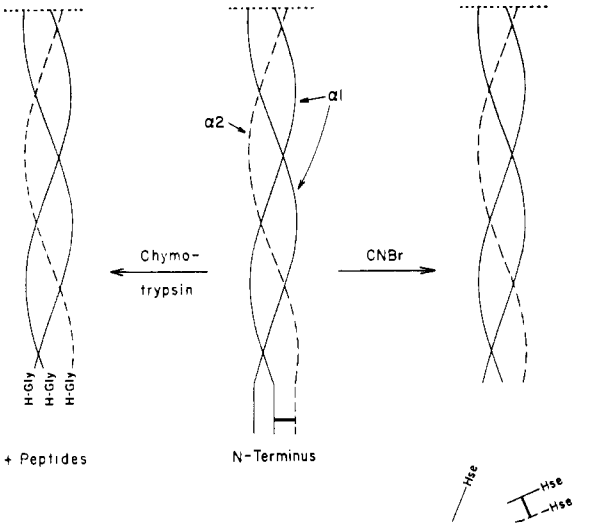

FIGURE 9: Diagrammatic representation of the N-terminal region of the collagen molecule and the action of chymotrypsin and CNBr on the native protein. The minor helix of the individual chains is not shown. The postulated absence of helical structure at the N terminus of the molecule is indicated by the straight line portions of the α chains. The bar represents an intramolecular cross-link between $\alpha 1$ and $\alpha 2$ forming β_{12}. C-terminal homoserine (or its lactone) is indicated by Hse.

and tyrosine, consistent with the loss of the sequences which constitute the cross-link precursor peptides $\alpha 1$-CB-1 and $\alpha 2$-CB-1. In other respects the analyses of $\alpha 1^{CNBr}$ and α^{CNBr} were identical, within experimental error, with those of $\alpha 1$ and $\alpha 2$. Since CNBr converts methionyl residues to C-terminal homoserine lactone (Gross and Witkop, 1961), the fragments produced will each contain 1 equiv of homoserine lactone with the exception of the products resulting from sequences C-terminal to the last methionyl bonds cleaved. The complete absence of homoserine in $\alpha 1^{CNBr}$ and $\alpha 2^{CNBr}$ and the presence of 1 equiv of $\alpha 1$-CB-1 and $\alpha 2$-CB-1 permit the conclusion that limited cleavage with CNBr occurred near the N-terminal ends of the α chains. However, these data alone do not establish that $\alpha 1$-CB-1 and $\alpha 2$-CB-1 are N terminal since sequences containing methionyl residues may have preceded these peptides and given rise to peptides that were missed. This possibility was examined in the following experiment.

The Limited Cleavage of Collagen with Trypsin

The presence of lysyl residues in both $\alpha 1$-CB-1 and $\alpha 2$-CB-1 suggested that trypsin might remove a small segment at the N terminus of each α chain in native collagen. Since trypsin would not be expected to cleave chains in which these lysyl residues were converted to their aldehyde derivatives, or in which these aldehydes had condensed to form cross-links, collagen from lathyritic rats was chosen for these experiments. The dialyzable material resulting from tryptic digestion of lathyritic collagen was desalted on a column of Bio-Gel P-2 equilibrated with pH 4.5 ammonium propionate buffer (Bornstein and Piez, 1966) in order to separate peptides from $CaCl_2$ and residual Na_2CO_3. The column effluent was monitored at 230 mμ. A major peak appeared in the effluent just before the retarded salt in the position expected for low molecular weight peptides. The fractions comprising this peak were lyophilized and chromatographed on phosphocellulose. The elution pattern (Figure 8) revealed the presence of two peaks. The amino acid compositions of the isolated peptides, $\alpha 1$-T-1 and $\alpha 2$-T-1 (Table III) are identical with the sequences containing five amino acids from the N-terminal ends of $\alpha 1$-CB-1 and $\alpha 2$-CB-1 as deduced by tryptic and chymotryptic digestion of these peptides (Bornstein and Piez, 1966; A. H. Kang, P. Bornstein, and K. A. Piez, in preparation). It can be concluded from these data that $\alpha 1$-CB-1 and $\alpha 2$-CB-1 contain the sequences at the N-terminal ends of $\alpha 1$ and $\alpha 2$.

Discussion

In contrast to bacterial collagenase which digests solutions of native collagen to low molecular weight products (see Mandl, 1961), the action of other proteinases is extremely limited if conditions are chosen which do not disrupt the triple-helical conformation of the native protein. The possibility that trypsin and pepsin possess the ability to cleave a limited number of peptide bonds in regions of the collagen molecule which differ in structure from that of the collagen helix was first considered by Hodge *et al.* (1960). These authors noted that certain properties of the enzyme-treated protein had been altered and suggested that the enzymes release peptides from chains which extend beyond the helical body of the collagen molecule and that these peptides might be involved in the polymerization of collagen molecules during fibril formation. Subsequently a number of workers (Rubin *et al.*, 1963, 1965; Kühn *et al.*, 1963; Drake *et al.*, 1966; Martin *et al.*, 1966)

have shown that the action of proteolytic enzymes on native collagen in solution is associated with a conversion of cross-linked components to products that resemble α chains in their ultracentrifugal and chromatographic behavior.

The nature of the peptides released from native collagen by pepsin (Rubin *et al.*, 1963, 1965; Steven, 1965) as well as by a number of other proteolytic enzymes (Drake *et al.*, 1966) has been investigated. In contrast to the composition of collagen, these peptides, termed telopeptides (Schmitt, 1964), were characterized by a high content of tyrosine, a predominance of acidic over basic amino acids, and a low content or absence of hydroxyproline. These investigators have suggested that these peptides are multiple, representing either branches from the main collagen chain or possibly the N-terminal ends of overlapping intrachain subunits which have been proposed for the α chains of collagen (Gallop, 1964; Petruska and Hodge, 1964). Among a variety of functions attributed to the regions of the molcule from which these peptides are derived, the formation of intra- and intermolecular cross-links is well established by the observations that β components are converted to α-like chains (already cited) and that insoluble collagen can be solubilized (Kühn *et al.*, 1963; Drake *et al.*, 1966) by the action of proteolytic enzymes.

The findings presented here permit a more detailed view of the region of the native molecule accessible to proleolytic enzymes and the action of CNBr. Our conclusions are illustrated schematically in Figure 9. The characteristic triple-helical structure of collagen does not include the N-terminal ends of the α chains. The latter region, schematically represented by straight rather than helical lines, assumes some other configuration which renders it susceptible to these agents. The intramolecular interchain cross-link is formed in the N-terminal region of the molecule, perhaps in part because the absence of a helical structure permits the selective *in vivo* oxidative deamination of its lysyl residues (Bornstein and Piez, 1966). The cross-link is represented in Figure 9 as a bar linking an $\alpha 1$ and $\alpha 2$ chain to form β_{12}, but two $\alpha 1$ chains might be linked to form β_{11}, or in a small percentage of cases all three chains may be linked to form γ_{112}. Chymotrypsin and CNBr cleave near the beginning of the collagen helix, liberating small peptides. During the separation of the chymotrypsin-cleaved protein from the enzyme these peptides are lost, but presumably similar to those released by chymotrypsin can be isolated from the products of the limited cleavage of native collagen with CNBr. The chymotrypsin- or CNBr-cleaved protein after denaturation consists largely of altered α chains which differ from the original α chains only in that they lack a short segment of their N-terminal ends. Cross-linked components yield the same shortened α chains since the cross-link is in the N-terminal region.

Our data indicate that only a single, interchain covalent cross-link is responsible for the formation of the β component. However, it should be emphasized that the present studies deal only with intramolecular cross-links in soluble rat skin collagen. It is possible that older collagen or collagen from other sources may present a more complex picture.

The observation that trypsin is relatively ineffective in reducing the content of cross-linked components in native collagen (Drake *et al.*, 1966; Martin *et al.*, 1966) can be explained by the finding that the only trypsin-sensitive bonds in the N-terminal region of the rat skin collagen molecule are the aldehyde-precursor lysyl residues which are no longer available after conversion to aldehydes and the subsequent formation of the intramolecular cross-link. At higher temperatures at which trypsin is effective (Martin *et al.*, 1966) the collagen helix begins to unfold, perhaps permitting cleavage of near-by trypsin-sensitive bonds.

The identification of $\alpha 1$-CB-1 and $\alpha 2$-CB-1 as the N-terminal sequences of $\alpha 1$ and $\alpha 2$, established by the present studies, is further supported by the finding that N-terminal analyses of $\alpha 1$ and $\alpha 2$ (Piez *et al.*, 1966) are consistent with the amino acid sequence and chemistry of $\alpha 1$-CB-1 and $\alpha 2$-CB-1 (A. H. Kang, P. Bornstein, and K. A. Piez, in preparation). It therefore appears that under the conditions of the experiments described, chymotrypsin and CNBr shorten the $\alpha 1$ chain by 15 amino acids and the $\alpha 2$ chain by 14 amino acids. Since lysine is the fifth amino acid in $\alpha 1$-CB-1 and $\alpha 2$-CB-1 (Bornstein and Piez, 1966), it can also be concluded that the intramolecular cross-link is formed five residues from the N-terminal ends of the $\alpha 1$ and $\alpha 2$ chains.

The conclusion that the intramolecular cross-link in soluble collagen is located near the N-terminal end of the molecule is consistent with recent studies on the nature of an aldehyde-containing peptide isolated from collagenase digests of carp swim bladder collagen (Rojkind *et al.*, 1966). The aldehydic component, present as a 2,4-dinitrophenylhydrazone derivative, was further localized to a seven or eight amino acid-containing peptide which lacked a free amino-terminal group and therefore probably arose from a blocked N-terminal region. Although the aldehydic group was not completely characterized, its reactivity suggested a role in the cross-linking of collagen. The N-terminal location of the intramolecular cross-link is also in accord with recent observations on the limited cleavage of native mammalian collagen with an enzyme isolated from the tadpole (Gross and Nagai, 1965; Kang *et al.*, 1966). These experiments indicated that only two fragments were formed from native soluble collagen by the tadpole enzyme and that the band patterns of segment-long-spacing (SLS) aggregates of these fragments, observed in the electron microscope, accounted for the total band pattern of SLS aggregates of uncleaved collagen. The smaller piece, designated TC^B, included the B end of the collagen molecule on the basis of its electron microscopic appearance. No interchain cross-links could be demonstrated in TC^B; however the larger fragment, designated TC^A, contained a proportion of cross-linked chains equal to that of the original preparation. N-terminal studies (Nagai *et al.*, 1965) have shown that newly formed

N terminals are limited to TC^B indicating that TC^A must include the N-terminal end of the collagen molecule.

In an earlier study, Nordwig and Dick (1965) were unable to detect a loss of methionine or a change in the ultracentrifugal or electrophoretic behavior of collagen after treatment of the native protein with CNBr. These authors concluded that denaturation of the protein was essential for cleavage to occur. However, it is likely that limited cleavage of the sort reported here would not have been detected by the techniques employed.

It is possible to identify provisionally the peptide bond split in $\alpha 1$ which results in its conversion to $\alpha 1^{Chy}$. Peptide $\alpha 1$-CB-2 has been shown to be located adjacent to the N-terminal peptide $\alpha 1$-CB-1 (Bornstein and Piez, 1966). Since $\alpha 1$-CB-2, which possesses an N-terminal glycine, is recovered intact after chymotryptic digestion, cleavage must have occurred N terminal to this peptide. However $\alpha 1$-CB-1 itself contains no chymotrypsin-sensitive bond (A. H. Kang, P. Bornstein, and K. A. Piez, in preparation); therefore the methionylglycyl bond joining $\alpha 1$-CB-1 and $\alpha 1$-CB-2 would appear to be the bond cleaved by chymotrypsin. This supposition is supported by the finding that after CNBr digestion of $\alpha 1^{Chy}$ no homoserine-containing peptide other than peptides previously identified in digests of $\alpha 1$ could be found. Although chymotryptic cleavage of methionyl bonds is unusual, a number of examples of such action have been reported, including the hydrolysis of a methionylglycyl bond in the γ chain of human fetal hemoglobin (Schroeder *et al.*, 1963).

Acknowledgment

The authors are indebted to Suzanne Roniger and Robert Katzen for expert technical assistance.

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. (1964), *J. Clin. Invest. 43*, 1813.

Bornstein, P., and Piez, K. A. (1965), *Science 148*, 1353.

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

Drake, M. P., Davison, P. F., Bump, S., and Schmitt, F. O. (1966), *Biochemistry 5*, 301.

Gallop, P. M. (1964), *Biophys. J. 4, Suppl. 1*, 79.

Gross, J., and Nagai, Y. (1965), *Proc. Natl. Acad. Sci. U. S. 54*, 1197.

Gross, E., and Witkop, B. (1961), *J. Am. Chem. Soc. 83*, 510.

Hodge, A. J., Highberger, J. H., Deffner, G. G. J., and Schmitt, F. O. (1960), *Proc. Natl. Acad. Sci. U. S. 46*, 197.

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

Kühn, K., Fietzek, P., and Kühn, J. (1963), *Naturwissenschaften 50*, 444.

Mandl, I. (1961), *Advan. Enzymol. 23*, 163.

Martin, G. R., Mecca, C. E., and Piez, K. A. (1966), *in* Environmental Variables in Oral Disease, Washington, D. C., American Association for Advancement of Science, pp 155–170.

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

Nagai, Y., Gross, J., and Piez, K. A. (1964), *Ann. N. Y. Acad. Sci. 121*, 494.

Nagai, Y., Piez, K. A., and Gross, J. (1965), Proceedings of the 16th Conference on the Structures of Proteins, Japan, Fukuoka, p 94.

Nordwig, A., and Dick, Y. P. (1965), *Biochem. Biophys. Acta 97*, 179.

Petruska, J. A., and Hodge, A. J. (1964), *Proc. Natl. Acad. Sci. U. S. 51*, 871.

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

Piez, K. A., Martin, G. R., Kang, A. H., and Bornstein, P. (1966), *Biochemistry 5*, 3813 (this issue; following paper.

Piez, K. A., Weiss, E., and Lewis, M. S. (1960), *J. Biol. Chem. 235*, 1987.

Rojkind, M., Blumenfeld, O. O., and Gallop, P. M. (1966), *J. Biol. Chem. 241*, 1530.

Rubin, A. L., Drake, M. P., Davison, P. F., Pfahl, D., Speakman, P. T., and Schmitt, F. O. (1965), *Biochemistry 4*, 181.

Rubin, A. C., Pfahl, D., Speakman, P. T., Davison. P. F., and Schmitt, F. O. (1963), *Science 139*, 37.

Schmitt, F. O. (1964), *Federation Proc. 23*, 618.

Schroeder, W. A., Shelton, J. R., Shelton, J. B., Cormick, J., and Jones, R. T. (1963), *Biochemistry 2*, 992.

Steven, F. S. (1965), *Biochim. Biophys. Acta 97*, 465.