Isolation and Characterization of the Cyanogen Bromide Peptides from the α1(III) Chain of Human Collagen†

Endy Chung, Elizabeth M. Keele, and Edward J. Miller*

ABSTRACT: The nine peptides derived from the human α1(III) chain by cyanogen bromide cleavage have been isolated and characterized with respect to amino acid composition and molecular weight. The peptides are recovered in equimolar amounts and together account for all the amino acids and molecular weight of the chain as obtained following pepsin digestion and solubilization of the native collagen. These results corroborate previous conclusions that the α1(III) chain exists in tissues as molecules of the chain composition, $[\alpha 1(III)]_3$, and that the latter contain only a single type of chain. Two of the nine cyanogen bromide peptides from α1(III) may be tentatively identified as homologs of peptides derived from α1(I) and α1(II) chains. The remaining α1(III) peptides have no obvious counterparts in the cyanogen bromide cleavage products of the α1(I) or α1(II) indicating a much different distribution of methionyl residues and suggesting that examination of the cyanogen bromide peptide pattern would provide a precise means of distinguishing between these chains. Characterization of the α1(III) peptides has also verified the unique compositional features of the latter chain. Thus, seven of the nine peptides contain more hydroxyproline than proline, and three of the peptides exhibit a glycine content somewhat higher than would be anticipated on the basis of one glycyl residue for every three amino acids. In addition, the two cysteinyl residues of α1(III) which provide sites for interchain disulfide bonding have been shown to occur in a typical collagen-like sequence near the COOH-terminal portion of the chain.

Evidence for the existence of the genetically distinct α1(III) chain was initially obtained in studies on the cyanogen bromide (CNBr) peptides derived from the insoluble collagen of human dermis (Miller *et al.*, 1971). These studies demonstrated that the CNBr cleavage products of insoluble dermal collagen contained at least two peptides the chemical properties of which indicated derivation from a collagen chain other than the previously characterized α1(I), α2, or α1(II) chains.

More recently, collagen molecules comprised of the α1(III) chain and possessing the apparent chain composition, $[\alpha 1(III)]_3$, have been isolated and separated from $[\alpha 1(I)]_2\alpha 2$ molecules by differential salt precipitation of the collagen solubilized from several human tissues by limited cleavage with pepsin (Chung and Miller, 1974). The latter studies also showed that α1(III) chains are extensively cross-linked through cysteine-derived disulfide bonds which are presumably located in helical, pepsin-resistant portions of the native $[\alpha 1(III)]_3$ molecule. Moreover, when compared to α1(I), α2, or α1(II) chains, the α1(III) chain exhibits several other compositional differences. Most notably, these differences include a relatively high content of hydroxyproline and glycine and a relatively low content of alanine and phenylalanine (Chung and Miller, 1974).

In order to further characterize the α1(III) chain, we have isolated and characterized the expected nine CNBr peptides derived from the chain as obtained following limited cleavage and solubilization with pepsin. The nine peptides have been recovered in approximately equimolar amounts and together account for all the molecular weight and amino acids of the α1(III) chain. These results, then, provide definitive evidence for the occurrence of a single type of α chain in molecules of the chain composition, $[\alpha 1(III)]_3$. These data also confirm the unique amino acid composition of α1(III) and serve to indicate the distribution of its unique features throughout the chain.

Only two of the CNBr peptides of human α1(III) may be tentatively identified as homologs of CNBr peptides derived from human α1(I) (Click and Bornstein, 1970; Epstein *et al.*, 1971) or α1(II) (Miller and Lunde, 1973) chains. Therefore, the peptides from α1(III) have been simply designated by numbers which correspond to their order of elution from cation-exchange columns.

Materials and Methods

Source and Preparation of α1(III). Collagen molecules containing the α1(III) chain were obtained by selective precipitation at 1.5 M NaCl from neutral salt solutions of pepsin-solubilized collagen as previously described (Chung and Miller, 1974). Due to tissue availability and increased yield of pepsin-solubilized collagen, the majority of the experiments described here were performed on collagen derived from infant dermis. However, collagen from human aorta as well as uterine leiomyoma gave identical results.

Each preparation of α1(III)-containing collagen (approximately 150 mg) was dissolved at a concentration of 10 mg/ml in 0.02 M (Na^+) sodium acetate (pH 4.8) containing 1.0 M urea, denatured at 50° for 30 min, and chromatographed on a 2.5 × 10 cm column of carboxymethyl (CM-) cellulose utilizing conditions previously described for chromatography of the α1(II) chains of cartilage collagen (Miller, 1971a). The nature of the material eluted from CM-cellulose was further assessed by rechromatography of an aliquot, both before and after reduction and alkylation, on an agarose molecular sieve column (see below).

Cleavage with CNBr. The α1(III) chains eluted from CM-cellulose were dissolved at a concentration of 10 mg/ml in 70% formic acid and CNBr cleavage was attained as described previously (Miller *et al.*, 1971). Following a 4-hr incubation at 30°, the reaction mixture was diluted tenfold with distilled water and lyophilized.

† From the Departments of Pathology and Biochemistry, and Institute of Dental Research, University of Alabama Medical Center, Birmingham, Alabama 35294. *Received April 18, 1974.* Supported by U. S. Public Health Service Grants DE-02670 and HL-11, 310-07.

* Address correspondence to this author at the Institute of Dental Research, University of Alabama Medical Center.

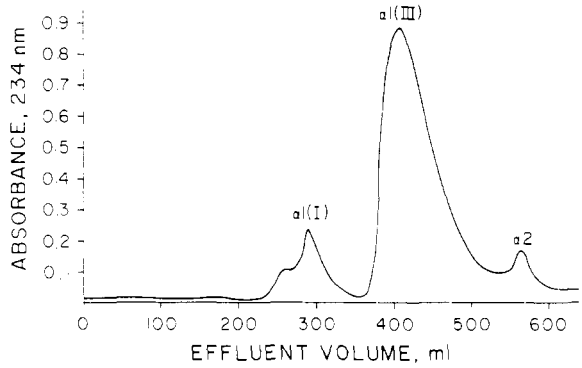

FIGURE 1: CM-cellulose elution pattern of 150 mg of collagen precipitated at 1.5 M NaCl from a neutral salt solution of pepsin-solubilized human dermal collagen. Elution was achieved in 0.02 M (Na^+) sodium acetate buffer containing 1.0 M urea at pH 4.8 using a linear gradient from 0 to 0.1 M NaCl over a total volume of 1000 ml.

Chromatography of CNBr Peptides on CM-Cellulose. The lyophilized CNBr peptides from α1(III) were redissolved in starting buffer (0.02 M sodium citrate, adjusted to pH 3.6 with citric acid and containing 0.01 M NaCl) and chromatographed initially on a 2.5 × 10 cm column of CM-cellulose employing conditions identical with those previously described for chromatography of the larger CNBr peptides from α1(II) chains (Miller, 1972; Miller and Lunde, 1973).

For further purification, the larger and more basic CNBr peptides eluted in the citrate system (above) were rechromatographed individually on a 2.5 × 10 cm column of CM-cellulose under different conditions. Peptides 3, 4, and 5 were redissolved in starting buffer (0.02 M (Na^+) sodium acetate (pH 4.8)) and applied to the column which had been equilibrated with the same buffer. Elution was then achieved at 42° and a flow rate of 200 ml/hr by means of a concave gradient. The latter was supplied by a constant-level device containing 825 ml of starting buffer in the initial chamber and 500 ml of limit buffer (starting buffer containing 0.12 M NaCl) in the second chamber. Peptide 6 was rechromatographed in the same system with the exception that both starting and limit buffers contained 0.03 M (Na^+) sodium acetate. Peptides 7, 8, and 9 were likewise rechromatographed in the same manner with the exception that both starting and limit buffers contained 0.05 M (Na^+) sodium acetate.

Chromatography of CNBr Peptides on Phosphocellulose. Peptides 1 and 2 initially resolved on CM-cellulose were further purified by rechromatography on a 1.8 × 9.0 cm column of phosphocellulose as described previously (Miller *et al.*, 1969). In the present studies, elution was achieved in 0.001 M (Na^+) sodium acetate (pH 3.8) using a linear gradient of NaCl from 0 to 0.1 M in a total volume of 400 ml.

Molecular Sieve Chromatography. For final purification and molecular weight determinations, the CNBr peptides observed in this study were rechromatographed at a flow rate of 10 ml/hr on a calibrated 1.5 × 110 cm column of agarose beads (Bio-Gel A-1.5m, 200–400 mesh, Bio-Rad Laboratories) equilibrated wuth 1.0 M $CaCl_2$ (Piez, 1968).

Aliquots of the collagen used as the starting material in these studies (see above) as well as peptide 9, the cysteine-containing peptide of α1(III), were chromatographed at a flow rate of 12.8 ml/hr on a calibrated 1.5 × 155 cm column of agarose beads (Bio-Gel A-5m, 200–400 mesh, Bio-Rad Laboratories) equilibrated and eluted with 2.0 M guanidine-HCl (absolute grade, Research Plus Laboratories) and 0.05 M Tris (pH 7.5).

Reduction and alkylation of α1(III) chains as well as peptide 9 were attained by dissolving the samples in 5.0 M urea containing 0.1 M 2-mercaptoethanol and subsequent addition of iodoacetic acid (Na salt) to a concentration of 0.2 M as previously described (Chung and Miller, 1974). Products of the reduction and alkylation reactions were examined by chromatography on agarose A-5m. In those instances where [1-^{14}C]iodoacetic acid (New England Nuclear, NEC-222) was used in the alkylation reactions, 0.5-ml aliquots of the column effluent were mixed with 10 ml of Aquasol (New England Nuclear) and counted in a liquid scintillation counter (Model LS-233, Beckman Instruments, Inc.).

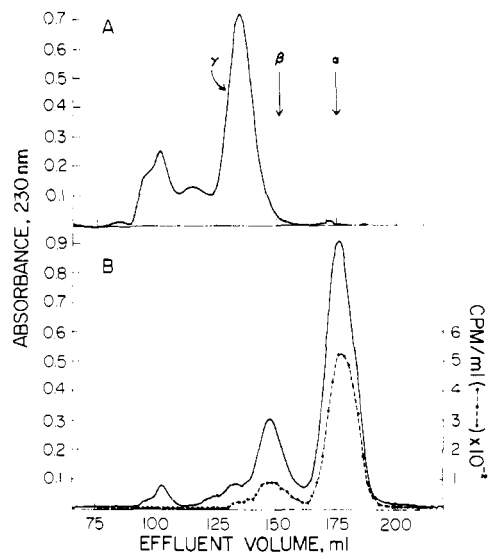

FIGURE 2: Bio-Gel A-5m chromatography of α1(III) chains eluted from CM-cellulose (Figure 1) before (A) and after (B) reduction with 2-mercaptoethanol and alkylation with iodoacetic acid (Na salt) containing [1-^{14}C]iodoacetic acid. The column was equilibrated and eluted at a flow rate of 12.8 ml/hr with 2.0 M guanidine-HCl containing 0.05 M Tris (pH 7.5).

Performic Acid Oxidation. Samples of peptide 9 (reduced with 2-mercaptoethanol) were oxidized with performic acid (Moore, 1963) for the determination of cysteine as cysteic acid.

Amino Acid Analyses. The purified peptides were hydrolyzed at 110° for 24 hr in constant-boiling 6 N HCl, and amino acid analyses were performed on an automatic amino acid analyzer as previously described (Miller, 1972).

In calculating amino acid chromatograms, correction factors were employed for losses of threonine, serine, and tyrosine, and incomplete release of valine during hydrolysis of collagen samples (Piez *et al.*, 1960).

Results

Preparation of α1(III). Figure 1 depicts a typical CM-cellulose elution pattern of the collagen precipitated at 1.5 M NaCl from neutral salt solutions of pepsin-solubilized dermal collagen. The α1(III) chains chromatograph as a single peak in a position intermediate between the more acidic α1(I) and the more basic α2 chains. The latter chains represent approximately 5% of the protein appearing in the eluent indicating that the original precipitate of $[\alpha 1(III)]_3$ molecules contained a small proportion of molecules with the chain composition, $[\alpha 1(I)]_2\alpha 2$. As judged by weight of the protein recovered, approximately 85% of the material applied to the column was recovered in the α1(III) peak.

Aliquots of the α1(III) chains appearing in CM-cellulose chromatograms were routinely examined by molecular sieve chromatography. Figure 2A is a representative elution pattern of the α1(III) peak when chromatographed on Bio-Gel A-5m.

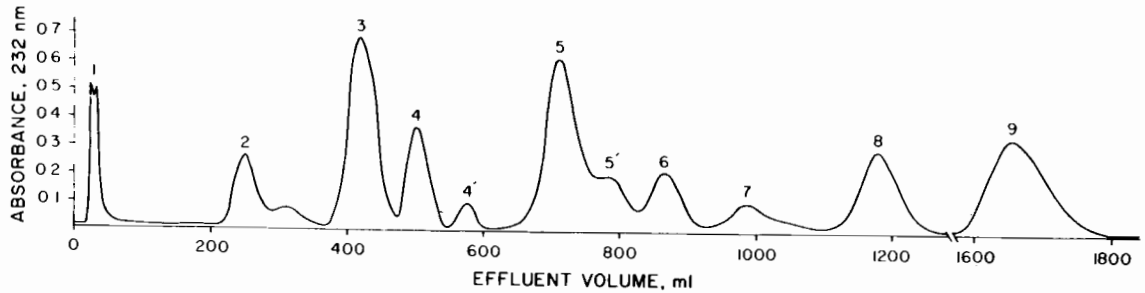

FIGURE 3: CM-cellulose elution pattern of 200 mg of CNBr peptides from α1(III). Elution was achieved in 0.02 M (Na^+) sodium citrate buffer (pH 3.6) using a linear gradient from 0.01 to 0.16 M NaCl over a total volume of 2000 ml.

TABLE I: Amino Acid Composition[a] and Molecular Weight of CNBr Peptides from Human α1(III).

Amino Acid	1	2	3	4	5	6	7	8	9	Total	α1(III)[b]
4-Hydroxyproline[c]	2 (2.0)	7 (7.2)	18	19	30	11	3.4	15	23	128	127
Aspartic acid	0	1 (0.8)	2 (1.9)	7 (6.9)	9 (8.9)	4 (4.0)	4 (4.0)	5 (4.8)	11	43	43
Threonine	0	0	1 (1.2)	4 (3.9)	3 (3.1)	1 (1.0)	1 (1.0)	0	3 (2.9)	13	13
Serine	0	4 (4.0)	6 (5.9)	0	9 (8.8)	4 (3.9)	0	1 (1.0)	14	38	39
Glutamic acid	0	3 (3.0)	6 (6.1)	11	15	7 (6.9)	4 (4.2)	10	15	71	72
Proline[c]	1 (1.2)	6 (6.1)	15	20	28	7 (7.1)	1.6	9 (9.1)	31	119	109
Glycine	4 (4.0)	14	36	56	77	28	14	45	78	352	354
Alanine	2 (2.0)	0	8 (7.8)	14	20	9 (9.1)	3 (2.8)	19	19	94	97
Valine	0	1 (1.0)	1 (1.0)	0	4 (3.9)	1 (0.8)	0	2 (2.0)	4 (4.2)	13	14
Cysteine[d]	0	0	0	0	0	0	0	0	2 (1.8)	2	2
Isoleucine	1 (1.0)	0	3 (2.8)	0	3 (3.0)	1 (1.1)	0	2 (2.1)	3 (2.8)	13	13
Leucine	1 (1.0)	0	2 (2.0)	5 (5.2)	5 (4.9)	1 (1.0)	2 (2.1)	2 (2.0)	4 (3.8)	22	22
Tyrosine	0	0	2 (2.0)	0	0	0	0	0	0	2	3
Phenylalanine	0	0	1 (1.0)	1 (1.0)	2 (1.9)	1 (1.0)	1 (1.0)	1 (1.0)	1 (1.0)	8	8
Hydroxylysine[c]	0	0	0	0.5	0.8	0.5	1.8	0.5	1.6	6	5
Lysine[c]	0	1 (1.0)	2 (2.1)	7.5	8.0	1.5	1.2	4.6	4.5	30	30
Histidine	0	0	1 (0.9)	0	0	1 (1.0)	1 (1.0)	0	3 (3.1)	6	6
Arginine	0	2 (2.0)	4 (3.8)	3 (3.1)	8 (7.8)	5 (5.1)	2 (2.1)	9 (8.9)	13	46	47
Homoserine	1 (1.0)	1 (1.0)	1 (1.0)	1 (0.9)	1 (1.0)	1 (0.9)	1 (1.0)	1 (0.9)	0	8	8[e]
Total	12	40	109	149	223	84	41	126	230	1,014	1012
Mol wt by amino acid anal	1039	3666	10,058	13,100	20,106	7795	4035	11,545	21,090	92,434	
Mol wt by agarose chromatography	1039[f]	3666[f]	9,500	13,000	20,000	8000	4035[f]	12,000	22,000	93,240	

[a] Residues per peptide to the nearest whole number. Actual values are given where less than ten residues occur. [b] Calculated from α1(III) amino acid analyses (Chung and Miller, 1974) using an average residue molecular weight of 91.2 and assuming a molecular weight of 92,400 for α1(III) when obtained after solubilization with pepsin. [c] Actual values listed where evidence for partial hydroxylation of prolyl and lysyl residues is apparent. [d] Determined as cysteic acid after reduction of peptide 9. [e] Denotes methionine for α1(III). [f] Molecular weight determined by amino acid analysis only.

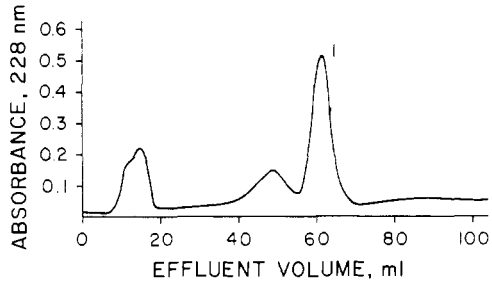

FIGURE 4: Phosphocellulose rechromatography of peptide 1. Elution was achieved in 0.001 M (Na^+) sodium acetate buffer (pH 3.8) using a linear gradient of NaCl from 0 to 0.1 M in a total volume of 400 ml.

Approximately 75% of the material chromatographs in the position corresponding to γ components (300,000) and the remainder as higher molecular weight aggregates. If, however, the samples were reduced and alkylated (with [1-^{14}C]iodoacetic acid) prior to agarose chromatography, the majority of the protein is recovered as radioactive α chains (Figure 2B). The protein eluting in the position of β components (Figure 2B) was shown to have an amino acid composition essentially the same as that of the α chains, α1(III) (Chung and Miller, 1974). This, then, indicates that the β component appearing only after reduction and alkylation originates as the result of incomplete reduction of the α1(III) aggregates.

Since the results described above indicated that the α1(III) peak eluted from CM-cellulose (Figure 1) was essentially pure and contained only disulfide-linked α1(III) chains, CNBr cleavage and all subsequent experiments were performed on the protein as recovered following CM-cellulose chromatography.

Initial CM-Cellulose Chromatography of CNBr Peptides. The CNBr peptides derived from α1(III) were resolved by CM-cellulose chromatography as illustrated in Figure 3. Each of the expected 9 peptides of the chain were adequately resolved by this procedure. The peptides eluting in the peaks designated 4′ and 5′ were determined to be identical in amino acid composition and molecular weight to the parent peptides, 4 and 5, respectively. At present, there is no explanation for the apparent tendency of peptides 4 and 5 to chromatograph in multiple forms under the conditions employed.

Phosphocellulose Rechromatography of CNBr Peptides. Peptides 1 and 2 eluted from CM-cellulose (Figure 3) were further purified by rechromatography on phosphocellulose. An example of the results of this procedure is illustrated in Figure 4 for peptide 1. The initial peak in the chromatogram was comprised of nonprotein absorbing material, and peptide 1 was recovered as the major retained peak eluting at approximately 60 ml.

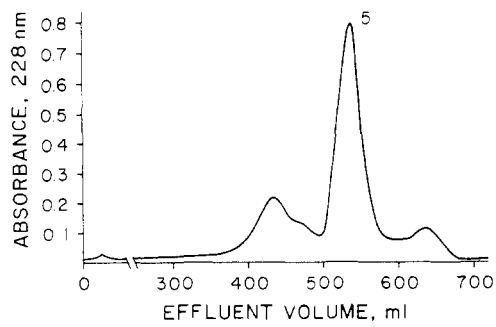

FIGURE 5: CM-cellulose rechromatography of peptide 5. Elution was achieved in 0.02 M (Na^+) sodium acetate buffer (pH 4.8) by means of a concave gradient established by a constant-level device with 825 ml of starting buffer in the initial chamber and 500 ml of limit buffer (starting buffer containing 0.12 M NaCl) in the second chamber.

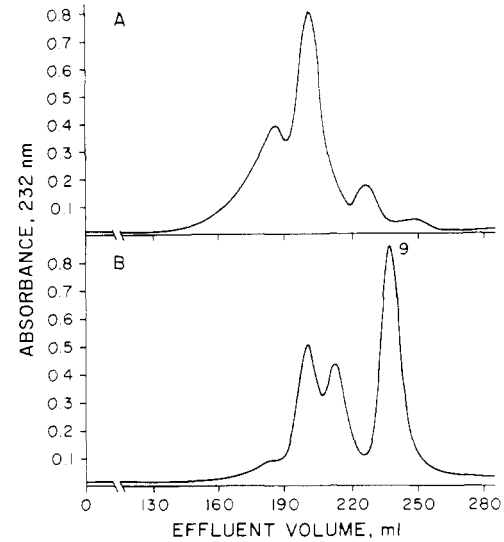

FIGURE 6: Bio-Gel A-5m rechromatography of peptide 9 before (A) and after (B) reduction with 2-mercapthanol. The column was eluted as described in the legend for Figure 2.

CM-Cellulose Rechromatography of CNBr Peptides. The remaining peptides eluted initially from CM-cellulose (Figure 3) were further purified by rechromatography on CM-cellulose at pH 4.8 and utilizing a concave gradient for elution. A representative chromatogram illustrating the results of rechromatography is presented in Figure 5 for peptide 5. The latter peptide was recovered as the major peak eluting at approximately 540 ml. The minor peaks in this chromatogram occurring both before and after peptide 5 were identified on the basis of amino acid composition and molecular weight to be identical with peptide 5 appearing in the major peak. These results are consistent with the pronounced tendency of individual collagen CNBr peptides to exhibit multiple peaks when chromatographed on CM-cellulose at pH 4.8 (Butler *et al.*, 1967).

Similar results to those depicted in Figure 5 were obtained on rechromatography of peptides 3, 4, 6, 7, and 8.

Molecular Sieve Rechromatography of Peptide 9. Elution patterns obtained when peptide 9 (Figure 3) was rechromatographed on CM-cellulose at pH 4.8 usually exhibited 4–5 poorly resolved peaks suggesting an unusual degree of complexity in this particular fraction. This was verified by rechromatography of peptide 9 on Bio-Gel A-5m (Figure 6A). Although the majority of the protein eluted in a single peak at 200 ml (64,000), a significant proportion of the material was broadly distributed throughout much higher molecular weight ranges. Following reduction with 2-mercaptoethanol, however, there is a pronounced shift in molecular weight distribution and peptide 9 is eluted as three distinct peaks with molecular weights estimated at 64,000, 42,000, and 22,000 (Figure 6B). The smallest component is the monomer form of peptide 9, the COOH-terminal, cysteine-containing CNBr peptide from the α1(III) chain (see Table I). In monomer form, peptide 9 rechromatographs on CM-cellulose at pH 4.8 in a manner similar to that described above for peptides 3, 4, 5, 6, 7, and 8. Amino acid analyses of the higher molecular weight peptides (Figure 6B) indicated that they represent di- and trimer aggregates of peptide 9 which result either from incomplete reduction or re-formation of disulfide bonds during chromatography. The latter possibility is considered likely since if peptide 9 were both reduced and alkylated the observed conversion to monomer form was considerably more quantitative than in the ex-

ample shown in Figure 6B where the protein had been reduced only.

Amino Acid Composition and Molecular Weight of CNBr Peptides from α1(III). The results of amino acid analyses and molecular weight determinations of the nine CNBr peptides derived from α1(III) are presented in Table I. These nine peptides represent the expected number since α1(III), when isolated after pepsin solubilization of the native collagen, contains eight methionyl residues (Chung and Miller, 1974). As noted in the last two columns of Table I, the nine peptides account for all the amino acids of the chain. In addition, the molecular weight determined for each peptide by molecular sieve chromatography agrees quite well with values calculated from amino acid analyses.

Stoichiometry of Peptides Derived from α1(III). Relative recoveries of the CNBr peptides from α1(III) were estimated essentially as described previously for the CNBr peptides from α1(II) chains (Miller, 1971b). Approximately 250 mg of α1(III) was cleaved with CNBr and the peptides were isolated and purified as described above. A known aliquot of each peptide was then submitted to amino acid analysis in order to assess purity and determine homoserine content. The results were then calculated as the amount of homoserine associated with each peptide. In the case of peptide 9, which does not contain homoserine, the calculations were based on phenylalanine, which occurs only once in the latter peptide (see Table I).

The results of these studies are presented in Table II and clearly indicate that each of the peptides identified as CNBr cleavage products of α1(III) is recovered in equivalent amounts. The results further indicate that the extraordinarily high absorbance associated with the relatively small peptide 3 (Figure 3) may be ascribed to the presence of tyrosyl residues in this sequence.

TABLE II: Stoichiometry of the CNBr Peptides from α1(III).

Peptide	Homoserine Content (μmol)	Relative Homoserine Content[a]
1	2.089	0.9
2	2.179	1.0
3	2.221	1.0
4	2.163	1.0
5	2.227	1.0
6	2.004	0.9
7	2.111	0.9
8	2.313	1.0
9[b]	1.990	0.9

[a] The data in this column are expressed relative to the quantity of homoserine associated with peptide 5 which is assigned a value of 1.0. [b] For peptide 9, the data are calculated from phenylalanine content.

Discussion

These studies have shown that cleavage of human α1(III) with CNBr gives rise to nine unique peptide fragments. The latter have been characterized with respect to molecular weight and amino acid composition demonstrating that they account for all the molecular weight and amino acids present in the chain as obtained following solubilization of the native molecules with pepsin. In addition, the nine peptides observed in this study correspond to the expected number and are recovered in approximately equimolar amounts. The results, then, substantiate the previous conclusions (Chung and Miller, 1974) that the α1(III) chain exists in tissues in the form of molecules with the chain composition, $[\alpha 1(III)]_3$, and that the latter molecules are comprised of only a single type of chain.

One may assume, by anology with other collagen types, that treatment and solubilization of the collagen used in these studies had resulted in removal of short nonhelical sequences located at the extremities of the $[\alpha 1(III)]_3$ molecules. Thus, the CNBr peptides and associated amino acids recovered in this study must be regarded as conservative in number and lacking approximately 30–40 amino acid residues of the intact α1(III) chain.

Inspection of the CNBr peptides derived from α1(III) (Table I) allows several interesting observations with respect to the relationship of this chain to other collagen α chains. Of the nine CNBr peptides of human α1(III), only two exhibit a molecular weight and amino acid composition consistent with their identification as homologs of CNBr peptides from α1(I) and α1(II) chains of several species. Peptide 6 from α1(III) is clearly homologous to the sequences represented by α1(I)-CB4, 5 (Traub and Piez, 1971) and α1(II)-peptide 12 (Miller, 1971b; Miller and Lunde, 1973) as previously indicated (Miller *et al.*, 1971). In addition, peptide 4 from α1(III) bears a striking resemblance to the sequences represented by α1(I)-CB3 (Traub and Piez, 1971) and α1(II)-peptide 8 (Miller, 1971b; Miller and Lunde, 1973). It should be noted that peptide 4 as isolated in the present study (Table I) has an amino acid composition somewhat different from that reported when the peptide was isolated from the total pool of insoluble collagen (Miller *et al.*, 1971). The reason for this discrepancy is due to the tendency of these peptides to exhibit multiple chromatographic forms making it virtually impossible to clearly separate α1(III)-peptide 4 from α1(I)-CB3 when mixtures of α1(III) and α1(I) are cleaved with CNBr.

The remainder of the peptides derived from α1(III) (1, 2, 3, 5, 7, 8, and 9) have no obvious counterparts in α1(I) (Traub and Piez, 1971) or α1(II) (Miller, 1971b; Miller and Lunde, 1973) chains, and, with the exception of peptide 9, cannot be tentatively assigned a location along the α1(III) chain. The latter peptide is, therefore, derived from the COOH-terminal portion of the chain. In view of its size peptide 9 would represent that portion of the α1(III) chain homologous to α1(I)-CB6 and the COOH-terminal portion of the α1(I)-CB7 (Traub and Piez, 1971) as well as α1(II)-peptide (9,7) and the COOH-terminal portion of α1(II)-peptide (10,5) (Miller *et al.*, 1973; Miller and Lunde, 1973).

Although the nature of the CNBr peptides derived from α1(III) is such that assigning homologies to CNBr peptides from α1(I) or α1(II) is difficult, this circumstance will allow clear and distinct differentiation of these chains when examined after CNBr cleavage.

Examination of the CNBr peptides derived from α1(III) is of further interest with respect to information concerning the α1(III) chain itself. Previous results obtained during studies on the whole chain (Chung and Miller, 1974) indicated that the most prominent features of this unique chain are the presence of relatively high contents of hydroxyproline and glycine, low amounts of alanine and phenylalanine, and the presence of cysteine which occurred in pepsin-resistant and presumably helical portions of the native molecule. In essence, each of these findings has been verified in the studies reported here, and several deserve additional comment.

Each peptide, with the exception of peptides 4 and 9, contains more hydroxyproline than proline. This result indicates that the tendency to hydroxylate relatively high numbers of

prolyl residues is not confined to any particular portion of the chain and suggests the occurrence of increased numbers of prolyl residues in sequences recognized by prolyl hydroxylase or increases in activity of the latter enzyme in certain cells. Alternatively, this observation may signify a relatively slow rate of helix formation during synthesis of $[\alpha 1(III)]_3$ molecules allowing more time for optimum prolyl hydroxylase activity on the newly synthesized nonhelical substrate as was recently postulated to account for various levels of hydroxylysine formed through the activity of lysyl hydroxylase (Kivirikko *et al.*, 1973). At the moment, there are insufficient data to discern whether any or all of the above mechanisms are operative. In this regard, however, it is of interest to note that human $\alpha 1(III)$ contains approximately the same amount of hydroxylysine as human $\alpha 1(I)$ (Bornstein and Piez, 1964) as well as $\alpha 1(I)$ chains from several other species (Piez *et al.*, 1963; Miller *et al.*, 1967; Rauterberg and Kühn, 1971).

The relatively high glycine content of $\alpha 1(III)$ (Chung and Miller, 1974) suggesting that somewhat more than one-third of the amino acids in this chain are glycine has been more precisely established by the present data. As noted in Table I, essentially all of the expected glycyl residues were recovered in the CNBr peptides and the amount recovered is 14 residues in excess of that required for one glycyl residue for every three amino acids in 1014 total residues. The present data further indicate that the additional glycyl residues occur chiefly in the sequences represented by peptides 4, 5, and 8 (Table I).

The present studies also demonstrate that the two cysteinyl residues of $\alpha 1(III)$ occur in peptide 9, and must therefore be located in the COOH-terminal portion of the chain. It is also clear that interchain disulfide bonding at this location is responsible for the relatively high molecular weight components observed when $[\alpha 1(III)]_3$ molecules are denatured and chromatographed on Bio-Gel A-5m (Figure 2A). The latter figure shows that about 75% of the denatured protein is recovered as γ components and the remainder as higher molecular weight aggregates. As shown in Figure 6A, peptide 9 is recovered in a similar state of aggregation with the trimer and higher molecular weight aggregates occurring in approximately the same proportions as γ components and higher molecular weight aggregates of the starting material.

The precise location and nature of the sequence(s) in which the cysteinyl residues of $\alpha 1(III)$ occur cannot be inferred at this time. It is clear, however, that they occur near the COOH-terminal end of the chain and in sequences which are not susceptible to proteolysis with pepsin. Presumably, these sequences are capable of forming the collagen helix. Moreover, the amino acid composition of peptide 9, which typically contains one-third of its amino acid residues as glycine and relatively large amounts of proline and hydroxyproline, does not suggest an appreciable degree of nonhelicity in this portion of the molecule. In this regard, the cysteinyl residues of $\alpha 1(III)$, described here, may be clearly differentiated from those in pro-$\alpha 1(I)$ and pro-$\alpha 2$ chains. In the latter chains cysteine occurs in noncollagen-like sequences (Bornstein *et al.*, 1972a; Furthmayr *et al.*, 1972) at the NH_2-terminal portions of the chains (Stark *et al.*, 1971; Dehm *et al.*, 1972) which are readily removed from procollagen molecules by nonspecific proteolytic enzymes (Layman *et al.*, 1971; Bellamy and Bornstein, 1971; Bornstein *et al.*, 1972b).

In contrast to collagen molecules of the chain composition, $[\alpha 1(II)]_3$, which are apparently restricted to hyaline cartilages (Miller, 1973) $[\alpha 1(III)]_3$ molecules have a more ubiquitous distribution. The latter molecules are found in dermis (Miller *et al.*, 1971; Chung and Miller, 1974), major vessels of the cardiovascular system, and uterine leiomyoma (Chung and Miller, 1974), and preliminary evidence suggesting their presence in dentin has also been presented (Volpin and Veis, 1973). This information indicating that several cell types are capable of synthesizing $\alpha 1(III)$ suggests that some $[\alpha 1(III)]_3$ molecules might occur in virtually every major connective tissue. In any event, in those tissues where $[\alpha 1(III)]_3$ molecules are known to occur, they coexist in varying proportions with molecules of the chain composition, $[\alpha 1(I)]_2\alpha 2$. Further physical and chemical studies on $\alpha 1(III)$ and its molecules should provide considerable insight into the respective roles of $[\alpha 1(I)]_2\alpha 2$ and $[\alpha 1(III)]_3$ molecules in maintaining the functional integrity of the tissues in which they occur.

References

Bellamy, G., and Bornstein, P. (1971), *Proc. Nat. Acad. Sci. U. S. 68,* 1138.

Bornstein, P., Ehrlich, H. P., and Wyke, A. W. (1972b), *Science 175,* 544.

Bornstein, P., and Piez, K. A. (1964), *J. Clin. Invest. 43,* 1813.

Bornstein, P., von der Mark, K., Wyke, A. W., Ehrlich, H. P., and Monson, J. M. (1972a), *J. Biol. Chem. 247,* 2808.

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

Chung, E., and Miller, E. J. (1974), *Science 183,* 1200.

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

Dehm, P., Jimenez, S. A., Olsen, B. R., and Prockop, D. J. (1972), *Proc. Nat. Acad. Sci. U. S. 69,* 60.

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

Furthmayr, H., Timpl, R., Stark, M., Lapiere, C. M., and Kühn, K. (1972), *FEBS (Fed. Eur. Biol. Soc.) Lett. 28,* 247.

Kivirikko, K. I., Ryhanen, L., Anttinen, H., Bornstein, P., and Prockop, D. J. (1973), *Biochemistry 12,* 4966.

Layman, D. L., McGoodwin, E. G., and Martin, G. R. (1971), *Proc. Nat. Acad. Sci. U. S. 68,* 454.

Miller, E. J. (1971a), *Biochemistry 10,* 1652.

Miller, E. J. (1971b), *Biochemistry 10,* 3030.

Miller, E. J. (1972), *Biochemistry 11,* 4903.

Miller, E. J. (1973), *Clin. Orthop. 92,* 260.

Miller, E. J., Epstein, E. H., Jr., and Piez, K. A. (1971), *Biochem. Biophys. Res. Commun. 42,* 1024.

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

Miller, E. J., and Lunde, L. G. (1973), *Biochemistry 12,* 3153.

Miller, E. J., Martin, G. R. Piez, K. A., and Powers, M. J. (1967), *J. Biol. Chem. 242,* 5481.

Miller, E. J., Woodall, D. L., and Vail, M. S. (1973), *J. Biol. Chem. 248,* 1666.

Moore, S. (1963), *J. Biol. Chem. 238,* 235.

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

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,* 1987.

Rauterberg, J., and Kühn, K. (1971), *Eur. J. Biochem. 19,* 398.

Stark, M., Lenaers, A., Lapiere, C., and Kuhn, K. (1971), *FEBS (Fed. Eur. Biol. Soc.) Lett. 18,* 225.

Traub, W., and Piez, K. A. (1971), *Advan. Protein Chem. 25,* 243.

Volpin, D., and Veis, A. (1973), *Biochemistry 12,* 1452.