The Amino Acid Sequence of Chick Skin Collagen α1-CB7†

John H. Highberger,* Clare Corbett, Andrew H. Kang,‡ and Jerome Gross

ABSTRACT: The amino acid sequence of chick skin collagen α1-CB7, the 268 residue CNBr peptide from the helical portion of the α1 chain, has been determined by automatic and manual degradation of tryptic and chymotryptic peptides, and of the COOH-terminal fragment produced by cleavage with animal collagenase. The resulting sequence shows 94% identity with that of the corresponding peptide from calf skin collagen (Fietzek, P. P., Rexrodt, F. W., Hopper, K. E., and Kühn, K. (1973), *Eur. J. Biochem. 38,* 396). The bond cleaved by animal collagenase has been identified as Gly-Ile at residues 221–222 of α1-CB7.

As a result of the work of several laboratories it has recently become possible to put together a complete amino acid sequence for the collagen α1(I) chain (Hulmes et al., 1973) based on data obtained partly with rat skin and partly with calf skin collagens. It is, however, not yet possible to do this for collagen of a single species, and this remains one of the major objectives of a systematic investigation in our laboratories of the primary structure of the α chains of chick skin collagen, of which this work is a part (Kang et al., 1969a–c; Kang and Gross, 1970; Highberger et al., 1971; Dixit et al., 1975a,b). Two considerations dictated the choice of the large (268 residue) α1-CB7 fragment as the subject of the present study: it represents approximately one-quarter of the α1 chain, this being located entirely in the helical portion, and it contains the site of the cleavage of α1 by animal collagenases (Gross and Lapiere, 1962; Gross and Nagai, 1965). The highly specific nature of this cleavage had led to the hypothesis that some unique structural feature might exist at this site.

While this work was in progress Fietzek et al., (1973) (see also Rexrodt et al., 1973) published their determination of the amino acid sequence of calf skin collagen α1-CB7. The results presented in this paper therefore permit a detailed comparison of the primary structures of a considerable portion of the α1(I) chains of two widely divergent species. They also make possible conclusions concerning sequence requirements around the animal collagenase cleavage site (see also Gross et al., 1974).

Experimental Section

Preparation of α1-CB7. Chick skin collagen was prepared from 3-week-old white Leghorn lathyritic chicks by neutral salt and acid extraction, and the α1 chain was isolated from purified collagen by CM[1]-cellulose chromatography, as described by Kang et al. (1969b). The α1 chain was subjected to CNBr cleavage by treating it, after solution in 70% formic acid at a concentration of 5 mg/ml, at 37° for 4 hr under nitrogen with an approximately equal weight of CNBr. The digest was either diluted tenfold with water and lyophilized, or desalted on a Bio-Gel P2 column of appropriate size before lyophilizing. α1-CB7 was isolated from the CNBr digest by CM-cellulose chromatography in citrate buffer at pH 3.6, and purified by rechromatography on CM-cellulose in acetate buffer at pH 4.8, using the methods of Butler et al. (1967). The preparations were given a final purification by chromatography on an Agarose (Bio-Gel A-1.5m, 200–400 mesh) column (2 × 120 cm) in 1 *M* $CaCl_2$ containing 0.01 *M* Tris-HCl (pH 7.5). In the later stages of the work α1-CB7 was isolated directly from CNBr digests of whole acid-soluble collagen. This variation gave no detectable differences in results from those obtained with material isolated from the separated α1 chain.

Maleylation of α1-CB7 and Tryptic Digestion of the Product. α1-CB7 was maleylated in 0.1 *M* sodium pyrophosphate buffer (pH 9.0) with a 20-fold molar excess of resublimed maleic anhydride, by the method of Butler et al. (1969). At the end of the reaction maleic acid and salts were removed by passing the mixture through a Sephadex G-25 column in 0.1 *N* NH_4OH, and lyophilizing the eluted protein. The extent of maleylation of the product, as estimated spectrophotometrically (Butler et al., 1969), was usually at least 90%. The maleylated α1-CB7 was digested with trypsin (Worthington TPCK-treated, three times crystallized), at an enzyme/substrate weight ratio of 1:100, in 0.2 *M* NH_4HCO_3 containing 10^{-3} *M* $CaCl_2$ (pH 7.9) for 2 hr at 37°. Another similar aliquot of enzyme was then added, and digestion continued for an additional 2 hr. The reaction was then stopped by acidification with 1 *M* acetic acid. The blocking maleyl groups were removed from the digestion products by heating the solution at pH 3.6, first at 37° for 65 hr, and then at 60° for 6 hr. This slight modification of the method of Butler et al. (1969) was found necessary because when either of these sets of conditions was used alone sufficient maleylamino groups remained to complicate the subsequent chromatographic separation of the digestion products.

† From the Developmental Biology Laboratory, Departments of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, and the VA Hospital and Departments of Biochemistry and Medicine, University of Tennessee Medical School, Memphis, Tennessee. *Received December 12, 1974.* This is publication No. 666 of the Robert W. Lovett Memorial Group for the Study of Diseases Causing Deformities. This work was supported by National Institutes of Health Grants AM 3564-14 and AM 16506, and was conducted under VA Research Project No. 4826-01 and 4826-02.

‡ Medical Investigator of Veteran's Administration, VA Hospital, Memphis, Tennessee.

[1] Abbreviations used are: CM, carboxymethyl; Tris, tris (hydroxymethyl)aminomethane; PTH, phenylthiohydantion; SPITC, 4-sulfophenyl isothiocyanate; DMAA, dimethylallylamine; Dns, 5-dimethylaminonaphthalene-1-sulfonyl (dansyl); ANS, 2-amino-1,5-naphthalenedisulfonic acid.

Where unblocked lysine-containing fractions from the above tryptic digests were subjected to further tryptic digestion, this was carried out as described above.

Tryptic Digestion of α1-CB7. α1-CB7 was treated with trypsin in the same manner described above for the maleylated product, except that the buffer used was 0.1 *M* Tris-HCl containing 10^{-3} *M* $CaCl_2$. At the end of the digestion period the reaction mixtures were acidified with acetic acid to pH 3, and immediately lyophilized.

Chymotryptic Digestion of α1-CB7. α1-CB7 was prepared for chymotryptic digestion by dissolving in 0.2 *M* NH_4HCO_3 (pH 7.9) at a concentration of approximately 5 mg/ml, and heating the solution for 10 min in a water bath at 50° in order to ensure denaturation. The solution was then cooled to 37° and soybean trypsin inhibitor (Worthington) was added in a weight ratio to the substrate of 1:1000. The enzyme (α-chymotrypsin, three times crystallized, Worthington) was then added in a weight ratio to the substrate of 1:100, and the solution was incubated at 37° for the required length of time (15 min to 2 hr, in the present series of experiments). At the end of the incubation period the solution was acidified with glacial acetic acid to pH 3, and immediately lyophilized.

Animal Collagenase Treatments. Animal collagenase was prepared from the media of cultures of the back skin of bullfrog tadpoles (*Rana catesbeiana*) by the method of Nagai et al. (1966), and purified as described by Harper and Gross (1970). Native chick skin collagen was treated with the enzyme at 20° in 0.05 *M* Tris buffer (pH 7.5) containing 0.2 *M* NaCl and 0.005 *M* $CaCl_2$. The concentration of collagen was 1–2 mg/ml, and the enzyme/collagen weight ratio was 1:100. The progress of the digestion was followed by viscometry, and at its conclusion the reaction was stopped by acidification with acetic acid, and the enzyme was irreversibly inactivated by heating the solution to 60° for 10 min. It was then dialyzed against 0.5 *M* acetic acid at 4° to remove salts, and lyophilized.

The combined COOH-terminal fragments $\alpha1^B$ and $\alpha2^B$ were separated from the other products of digestion and denaturation by chromatography of the lyophilized digest on a 4 × 120 cm column of Agarose (Bio-Gel A 1.5 m, 200–400 mesh) using 1.0 *M* $CaCl_2$ containing 0.01 *M* Tris (pH 7.5) as the eluent. The effluent was monitored continuously at 230 nm in a Gilford spectrophotometer.

It was found to be unnecessary to separate $\alpha1^B$ from $\alpha2^B$ for sequencing, as the major $\alpha1^B$ sequence was unmistakably evident, at least for the first 27 residues.

Molecular Sieve Chromatography. With a few exceptions, to be described below, enzymatic digests were first fractionated by molecular sieve chromatography on Sephadex G-50$_S$ (Pharmacia, Superfine Grade). A 2.5 × 100 cm jacketed column maintained at 35° was used. The column was calibrated for molecular weight estimation using collagenous CNBr peptides of known molecular weight and the ^3H_2O method described by Piez (1968). Portions of approximately 50 mg of α1-CB7 digestion products were applied to the column, and eluted with 0.03 *M* NaOAc (pH 4.8) at a flow rate of about 20 ml/hr. The eluate was continuously monitored at 230 nm in the equipment already described. Collected fractions were desalted on Bio-Gel P2 columns of appropriate size, in 0.1 *N* acetic acid, before lyophilizing. In the single case, described below, where an enzymatic digest was first fractionated on Sephadex G-25, the column was operated in a similar manner to that described for G-50$_S$.

Ion Exchange Chromatography. Tryptic digests of α1-CB7 were fractionated directly on a 0.9 × 23 cm column of PA-35 (Beckman) cation exchange resin at 50°, using a pyridine acetate gradient with a fraction of the effluent diverted to the ninhydrin reaction coil of an amino acid analyzer (Beckman, Model 120C) as previously described (Highberger et al., 1971). Fractions requiring further purification were rechromatographed on a 0.9 × 58 cm Dowex 1 (Bio-Rad AG-1X2) anion exchange column, using the same equipment, and eluting the column at 50° consecutively with 0.1 *M* pyridine acetate buffers of pH 9.4, 8.25, and 6.5, and finally with a linear gradient formed between equal volumes of the latter buffer and 0.5 *M* acetic acid. Further purification of fractions from anion exchange chromatography, as well as from molecular sieve chromatography, was carried out on a phosphocellulose (Whatman P 11) column at 40°, using linear NaCl gradients in 0.001 *M* sodium acetate (pH 3.8). NaCl gradients varied from 0–0.2 *M* to 0–0.6 *M*, depending upon the estimated size of the peptides being fractionated. Collected fractions were desalted on P 2 columns, in 0.1 *N* acetic acid, before lyophilizing.

In a few instances, to be described below, fractions obtained by molecular sieve chromatography of enzymatic digests were further fractionated on a 2 × 10 cm CM-cellulose column eluted at 40° with a linear gradient formed between 500 ml of 0.04 *M* sodium acetate (pH 4.8) and 500 ml of the same buffer containing 0.15 *M* NaCl. Where unblocked lysine-containing fractions from tryptic digests of maleylated α1-CB7 were subjected to further tryptic action, the digests were fractionated on a 1 × 10 cm column of QAE-Sephadex A-50 (Pharmacia), eluted at 40° in a stepwise manner with 0.05 *M* Tris-HCl buffers ranging in pH from 9.0 to 7.2. Collected fractions were desalted on P-2 columns, and lyophilized.

Peptide fractions whose amino acid analyses indicated a purity judged sufficient for sequencing were examined for homogeneity by paper chromatography and paper electrophoresis. Paper chromatography was carried out by the descending method on Whatman No. 1 paper in pyridine–H_2O, 8:2. Paper electrophoresis was performed in a Savant flat plate apparatus on Whatman 3MM paper, at 40 V/cm, in pyridine acetate at pH 6.4. Both types of chromatograms, after air evaporation of solvent, were stained by dipping in a 0.2% solution of ninhydrin in acetone.

Amino Acid Analysis. Samples for amino acid analysis were sealed under nitrogen in redistilled constant boiling HCl and hydrolyzed for 24 hr at 110°. The analyses were run on a Beckman 120C analyzer by the single column method of Miller and Piez (1966). Correction factors were used for Thr and Ser only. These were determined on collagenous samples in the course of the present work and were found to be 1.14 and 1.28, respectively.

Sequencing Methods. Automated sequential degradation of peptides was performed according to the method of Edman and Begg (1967) with a Model 890B sequencer (Beckman Instruments). In the later stages of the work the instrument was modified by the addition of a nitrogen flush during high vacuum steps. Samples were applied in either water or dimethylallylamine buffer, and dried with the Beckman peptide application subroutine. Programs used were essentially the Beckman protein and peptide programs, with the following modifications: (1) protein programs in all cases were run with single cleavage and extraction, and frequently with double vacuum times following coupling; and (2) benzene and butyl chloride washes in pep-

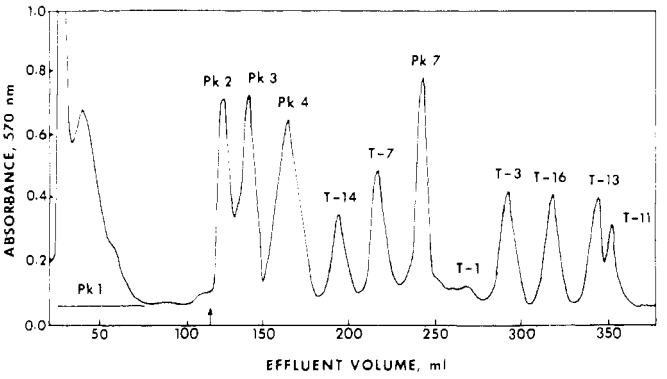

FIGURE 1: Chromatography of approximately 40 mg of tryptic digest of α1-CB7 on a 0.9 × 23 cm PA 35 column at 50°. The column was eluted at a flow rate of 72 ml/hr with 0.1 *M* pyridine acetate (pH 3.12) for the first hour, with 0.2 *M* pyridine acetate (pH 3.56) for the second hour, and thereafter with a linear gradient formed between 260 ml of the latter buffer and the same volume of 2.0 *M* pyridine acetate (pH 5.0). The horizontal bar indicates how the fractions composing the initial large, heterogeneous peak were pooled for further fractionation, and the vertical arrow shows the start of the gradient.

tide programs were each of 80 sec duration; in the COOH-terminal regions of the peptides these were each shortened to 60 sec.

Manual sequencing methods were employed on some of the smaller peptides, either according to Edman (1970), with identification of the PTH-amino acids by gas chromatography, or by the dansylation-Edman procedure as previously described (Highberger et al., 1971), with identification of the Dns-amino acids by thin-layer chromatography.

Small (i.e., 9 to approximately 30 residues) lysine-containing peptides were treated with 4-sulfophenyl isothiocyanate by the method of Braunitzer et al., (1970) in order to increase their hydrophilicity. Other selected peptides not containing lysine were coupled with 2-amino-1,5-naphthalenedisulfonic acid via 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride as described by Foster et al. (1973). In several cases both of these reagents were used to modify the same peptide, the latter reaction being carried out first, after which the solution was brought to dryness in a gentle stream of nitrogen, and then treated with the sulfophenyl isothiocyanate.

In the conversion of the phenylthiazolinones to the phenylthiohydantoins, accomplished by heating in 1 *N* HCl at 80°, glycine residues were treated for 15 min and amide residues for 3 min, instead of the usual 10 min.

Identification was mainly by gas–liquid chromatography according to Pisano and Bronzert (1969; see also Pisano et al., 1972) as modified for SP-400 column packing (Beckman). Trimethylsilylation was carried out for 5 min at 80° with *N,O*-bis(trimethylsilyl)acetamide (Pierce Chemical Co.) and was used mainly for identification of aspartic and glutamic acids, and for differentiation between leucine and isoleucine. Arginine residues were identified by the phenanthrenequinone method of Yamada and Altona (1966). Lysine, asparagine, and glutamine were identified by thin-layer chromatography, using the solvent systems of Kulbe (1971) and of Inagami and Murakami (1972). After completion of gas–liquid chromatography the remainders of the samples were hydrolyzed in hydriodic acid according to the method of Smithies et al. (1971), and residue identifications made on an automatic amino acid analyzer (Jeolco Model 5-AH).

Repetitive yields obtained by the above methods, as would be expected, were much better in the longer runs of α1-CB7 itself, than in the shorter runs on smaller peptides. For example, repetitive yields based on alanine and covering α1-CB7 from residue 6 to residue 79 ranged from 97.4 to 97.5%, while that obtained for the alanines at residues 3 and 14 of the 27-residue tryptic peptide T-10 was only 92.9%.

Results

In the following, peptides isolated from tryptic digests of α1-CB7, from tryptic digests of maleylated α1-CB7, and from chymotryptic digests of α1-CB7, are given designations starting with T, TM, and C, respectively. In each case the initial letter is followed by a number indicating the position of the peptide, relative to the other peptides isolated from similar types of digest, in the α1-CB7 sequence as finally determined.

Isolation of Peptides from Tryptic Digests of α1-CB7. Comparison of the content of basic amino acids in α1-CB7 with its total number of residues suggested that tryptic digestion would probably produce a number of relatively small peptides. It was found that when such a digest was passed through a PA 35 cation exchange column under the conditions described in the Experimental Section, the larger and more acidic peptides were eluted immediately, while further development of the column produced a number of peaks, analysis of which indicated that they were composed of small peptides.

A typical chromatogram is shown in Figure 1. Of the peaks shown, those marked T-1, T-3, T-7, T-11, T-13, T-14, and T-16 gave analyses (shown in Table I) indicating reasonably pure small peptides. Of this group, isolated directly from the PA 35 column in pure enough form for sequencing, four are tripeptides. The remaining three comprise a tetra-, a hexa-, and an octapeptide.

Materials collected from the remaining peaks of Figure 1 were rechromatographed on a Dowex 1 anion exchange column as described above. When peak 1 from Figure 1 (collected as shown in that figure) was rechromatographed under these conditions five peaks were obtained (chromatogram not shown), of which one gave the analysis shown under T-8 in Table I, indicating it to be a 24 residue tryptic peptide. The analyses of the remaining peaks indicated them to be mixtures. In other Dowex 1 runs made in a similar manner, of the remaining peaks from Figure 1 (chromatograms not shown), T-9 and T-2, two nonapeptides, were obtained from peaks 2 and 3, respectively, and T-4 and T-15, hexa- and nonapeptides respectively, were each obtained from peak 4. Analyses of all these peptides are given in Table I.

Four additional tryptic peptides were obtained by rechromatographing on phosphocellulose other peaks from the Dowex 1 chromatography as described in the Experimental Section. Figure 2 shows the chromatogram obtained when the third peak from the Dowex 1 chromatogram which had given T-8 was so treated. This gave rise, as indicated, to T-6 and T-10, of 18 and 27 residues respectively, and to T-17, a hexapeptide. The latter, since it contains no basic residues, and one residue of homoserine, clearly represents the COOH-terminal portion of α1-CB7. Similar rechromatography of the first peak from the same Dowex 1 chromatogram gave T-12, a 17-residue peptide sufficiently pure for sequencing (chromatogram not shown).

In the above chromatographic separations the yields of the tryptic peptides, based on the amount of α1-CB7 subjected to tryptic digestion, were as follows: the yields of

Table I: Amino Acid Composition of α1-CB7 Tryptic Peptides.[a]

Peptide	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8	T-9	T-10	T-11	T-12	T-13	T-14	T-15	T-16	T-17
4-Hydroxyproline	0.8	1.1		1.0		2.0	0.8	3.4		2.9		1.6			1.1		1.1
Aspartic acid			1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.0		0.2					0.1
Threonine						1.0			1.0	0.8		0.8		1.0			
Serine	0.2				0.2	0.3				0.4		1.1					0.1
Glutamic acid	1.0				0.2	0.2		2.0	1.0	0.6				1.0	0.9	1.0	
Proline				1.1		2.9		3.3		4.6		1.1	0.8	1.0			2.1
Glycine	1.2	2.6	1.3	2.3	1.3	5.9	2.2	7.6	2.9	8.8	1.3	5.8	1.2	3.0	3.0	1.0	2.0
Alanine	0.1	2.8	0.2			2.1	1.0	3.7	1.7	5.2	0.9	3.8		1.0			
Valine										0.8					1.9		
Isoleucine						1.0											
Leucine		0.9			1.0	1.0									0.9		
Phenylalanine								0.9				1.0					
Hydroxylysine										0.7							
Lysine		1.0		1.0		1.0		1.1	1.0	0.3					0.2		
Arginine	1.0		1.0		1.0		1.0				0.8	1.0	1.0	1.0	1.1	1.0	
Homoserine[b]																	0.8
Total residues	4	9	3	6	4	18	6	24	9	27	3	17	3	8	9	3	6

[a] In residues per peptide. No entry indicates that the amount found was less than 0.1 residue. [b] Includes homoserine lactone.

peptides eluted from the PA 35 column ranged from 56 to 77 mol %, and those eluted from the Dowex 1 column from 22 to 41 mol %. In both cases the amounts actually recovered were less, owing to the diversion of a portion of the column effluent to the ninhydrin reaction coil. In the case of the phosphocellulose column the peptide yields ranged from 6 to 19 mol %.

The material in peak 7 of the original PA 35 chromatography (Figure 1) had resisted all efforts at further fractionation by ion exchange chromatography, and examination by paper electrophoresis had indicated that separation of its components by such means would be extremely difficult. Resort was therefore made to paper chromatography. Two well-separated fractions were obtained by descending chromatography in pyridine–water (see Experimental Section for details). Analyses of these showed one to be the tetrapeptide T-5 (Table I); the other was apparently a still unresolved mixture.

As shown in Table I, a total of 17 tryptic peptides of sufficient purity for sequencing was thus isolated in the present work. If all basic bonds are cleavable a total of 25 would be expected. The peptides shown in Table I account for 6 of the 11[2] lysines and 10 of the 13 arginines of α1-CB7. Several additional peptides isolated by the methods described above were not obtained in sufficient purity to be useful, and are therefore not included in this report.

Isolation of Peptides from Tryptic Digests of Maleylated α1-CB7. Tryptic digests of maleylated α1-CB7, prepared as described above, were first fractionated on a calibrated Sephadex G-50$_s$ column. A typical chromatogram is shown in Figure 3. The large peak which eluted first, at a position corresponding to a molecular weight of about 6000, gave the analysis shown in Table II under TM-3. This showed it to be an overlap peptide of 66 residues, containing three lysines, one of which is the partially hydroxylated lysine of α1-CB7. Analyses of the remaining peaks of Figure 3 showed them to be mixtures requiring further fractionation.

This was accomplished in most cases by rechromatography on phosphocellulose, as described above for tryptic peptides of α1-CB7. In this way TM-5, a 48-residue peptide containing one lysine and two arginines (Table II), was obtained from peak 2 of Figure 3 (chromatogram shown in Figure 4). One of the arginines in this peptide remained uncleaved by the trypsin presumably because it occurs in an Arg-Pro or Arg-Hyp sequence. TM-5-T-2 (also shown in Table II), the 24-residue COOH-terminal portion of TM-5

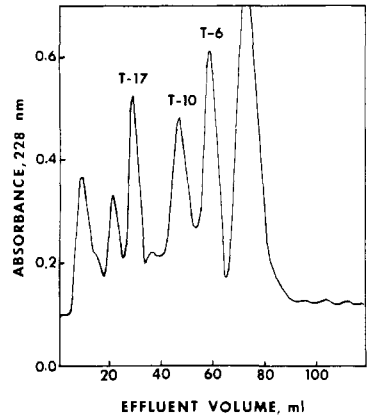

FIGURE 2: Rechromatography on a 0.9 × 10 cm phosphocellulose column of peak 3 from the rechromatography on Dowex 1 of peak 1 from Figure 1. The column was eluted, at 44° and a flow rate of 30 ml/hr, with a linear gradient formed between 200 ml of 0.001 *M* NaOAc (pH 3.8), and the same volume of the same buffer containing 0.6 *M* NaCl.

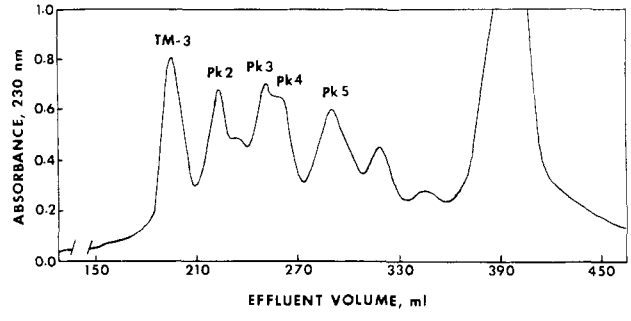

FIGURE 3: Chromatography of approximately 50 mg of a tryptic digest of maleylated α1-CB7 (unblocked after digestion) on a 2.5 × 110 cm column of Sephadex G-50$_s$ at 35°. The column was eluted with 0.03 *M* NaOAc (pH 4.8) at a flow rate of 18 ml/hr.

[2] The originally published analysis of chick skin collagen α1-CB7 (Kang et al., 1969b) showed the total of lysine and hydroxylysine residues as 10; as shown in the present work this should be 11. The present work also shows that there are 12 aspartyl plus asparaginyl residues instead of the originally published 11.

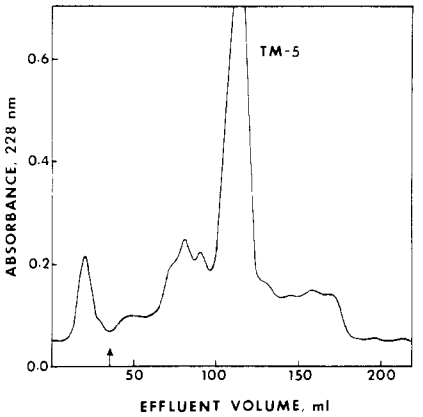

FIGURE 4: Rechromatography of peak 2 from Figure 3 on a 0.9 × 10 cm phosphocellulose column. The column was eluted at 42° and a flow rate of 40 ml/hr, initially with 0.001 *M* NaOAc (pH 3.8) and finally with a linear gradient formed between 250 ml of this buffer and 250 ml of the same buffer containing 0.6 *M* NaCl. The vertical arrow indicates the start of the gradient.

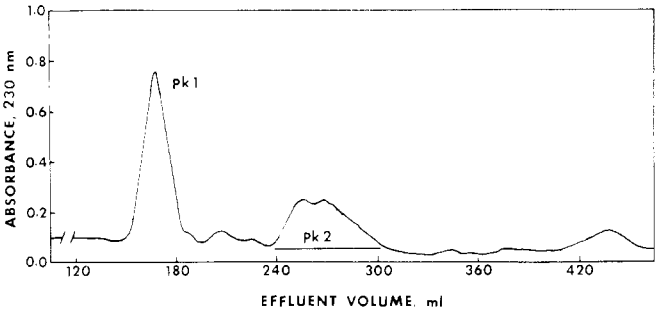

FIGURE 5: Chromatography of approximately 50 mg of a 75-min chymotryptic digest of α1-CB7, prepared as described in the text, on a 2.5 × 90 cm column of Sephadex G-50$_s$. The column was maintained at 35° and eluted with 0.03 *M* NaOAc (pH 4.8) at a flow rate of 20 ml/hr.

released by tryptic cleavage at the latter's unblocked lysyl residue, was isolated by chromatography of the TM-5 tryptic digest on QAE-Sephadex (chromatogram not shown; see Experimental Section for details). As shown in Table II, TM-5T-2 contains both isoleucines of TM-5.

TM-2, a 30-residue peptide with one lysine, and TM-4, a 29-residue peptide with two lysines, were obtained by phosphocellulose rechromatography (chromatogram not shown) of peak 3 of Figure 3. TM-6, a 24-residue peptide with one lysine, was obtained from peak 5 of Figure 3 by rechromatography on carboxymethylcellulose at pH 4.8 (chromatogram not shown).

In a search for smaller lysine-containing overlap peptides, a portion of unblocked tryptic digest of maleylated α1-CB7 was fractionated on a Sephadex G-25 column, as described in the Experimental Section. TM-1, an 18-residue peptide containing two lysines, was obtained by PA-35 rechromatography, as described above, of the first retarded peak on the Sephadex G-25 (chromatograms not shown).

In the above peptide isolations, the large peptide TM-3, isolated directly from Sephadex G-50$_s$ columns, was generally obtained in yields of 47–50 mol % based on the amount of α1-CB7 taken for maleylation. Peptides eluted from phosphocellulose columns gave yields ranging from 25 to 32 mol % on the same basis, while yields of from 35 to 45 mol % were obtained for those isolated from CM-cellulose columns.

The six lysine-containing overlap peptides isolated by the above methods, analyses of which are shown in Table II, account for 10 of the 11 lysines of α1-CB7. There thus can be no more than one additional lysine overlap. In addition to the overlap peptides, 3 of the non-lysine-containing tryptic peptides shown in Table I were also isolated from tryptic digests of maleylated α1-CB7; these are T-7, T-12, and the COOH-terminal peptide, T-17. Since these were not purified sufficiently to be of use for sequencing, they are not included here. It should be noted, however, that the occurrence of T-17 in tryptic digests of maleylated α1-CB7 indicates that it is immediately preceded in the sequence by an arginine, and that the latter is thus the last basic residue in the α1-CB7 sequence. The fact that the peptides of Table II, together with those of Table I also found in tryptic digests of maleylated α1-CB7, account for 9 of the 13 arginines, and a total of 243 residues, also indicates that the missing peptides must be of relatively small size.

Table II: Amino Acid Composition of Lysine-Containing Tryptic Peptides from Maleylated α1-CB7.[a]

Peptide	TM-1	TM-2	TM-3	TM-4	TM-5	TM-5-T-2	TM-6
4-Hydroxyproline	1.4	2.3	6.6	3.2	5.8	2.0	3.0
Aspartic acid	3.0	1.0	3.1	1.1	1.4	1.0	1.0
Threonine	0.1	2.0	1.8		1.6	0.8	
Serine	0.3	0.1		1.5	1.1	1.0	2.8
Glutamic acid	0.3	1.2	3.6	1.1	5.1	1.6	2.0
Proline	1.7	5.3	8.6	4.9	6.5	3.1	2.4
Glycine	5.6	9.6	22	10	16	7.8	7.7
Alanine	2.2	5.3	13	1.0	5.3	3.2	1.0
Valine			1.0	0.9			0.1
Isoleucine		1.0		0.8	1.6	1.5	0.2
Leucine	1.0	0.9		1.0			1.0
Phenylalanine			0.9				0.6
Hydroxylysine			0.8				
Lysine	1.9	1.1	2.2	1.8	1.1	0.3	1.2
Arginine	1.0	1.0	1.3	1.1	2.2	0.8	1.0
Total residues	18	30	66	29	48	24	24

[a] In residues per peptide. No entry indicates that the amount found was less than 0.1 residue.

Isolation of Peptides from Chymotryptic Digests of α1-CB7. Chymotryptic digests of α1-CB7 were initially fractionated by chromatography on Sephadex G-50$_s$. A typical chromatogram is shown in Figure 5. The elution patterns obtained in these experiments were quite sensitive to differences in the times used for the enzymatic digestions; that shown in Figure 5 (75 min) is intermediate between the minimum of 15 min and the maximum of 2 hr. While there was a general similarity in all the chromatograms, in those obtained with the shorter digestion times the three peaks eluting first represented a larger proportion of the whole, whereas with longer digestion times these were diminished and the poorly resolved complex of later-eluting peaks was increased. This behavior is obviously compatible with the interpretation that the substrate contains few bonds cleavable at the maximum rate of the enzyme; the initial cleavage therefore results in a few fragments of relatively large size. These, however, become partially degraded with progression of the digestion, as a result of slower secondary cleavages.

For purification, the peaks obtained in gel filtration were rechromatographed on phosphocellulose as described above. Figure 6 shows the chromatogram obtained in such refractionation of peak 1 of Figure 5. The second and third peaks of this chromatogram gave identical amino acid analyses;

Table III: Amino Acid Composition of α1-CB7 Chymotryptic Peptides.[a]

Peptide	C-1	C-2
4-Hydroxyproline	5.1	21
Aspartic acid	1.9	5.3
Threonine	1.7	5.3
Serine		7.4
Glutamic acid	2.2	13
Proline	6.7	23
Glycine	14.6	62
Alanine	6.7	21
Valine		3.8
Isoleucine	0.7	2.8
Leucine		2.9
Phenylalanine	0.8	2.1
Hydroxylysine		0.9
Lysine	1.0	6.3
Arginine	1.5	8.3
Homoserine[b]		0.9
Total residues	45	185

[a] In residues per peptide. No entry indicates that the amount found was less than 0.1 residue. [b] Includes homoserine lactone.

since homoserine was present, these probably represented the free acid and lactone forms of a COOH-terminal peptide. These two, marked C-2 and C-2 (lactone) in Figure 6, were therefore pooled for sequencing. The analysis of C-2, as used in sequencing, is given in Table III. This shows it to be a large peptide extending to the COOH-terminus over approximately 70% of the length of α1-CB7.

In similar phosphocellulose rechromatography of peak 2, the heterogeneous complex pooled as shown in Figure 5, C-1, a 45-residue peptide, was obtained (chromatogram not shown). An analysis of the peak marked C-1 is given in Table III.

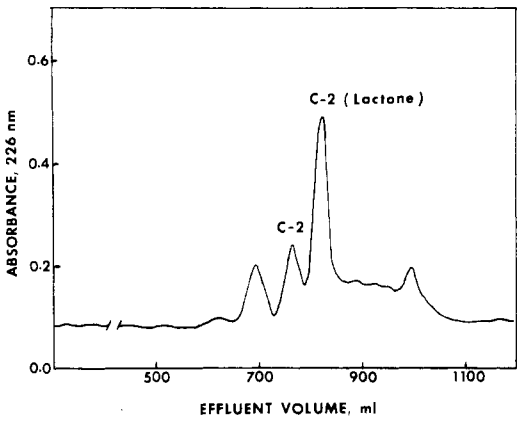

FIGURE 6: Rechromatography of peak 1 from Figure 5 on a 1.9 × 20 cm phosphocellulose column. The column was eluted at 42° and a flow rate of approximately 80 ml/hr with a linear gradient formed between 800 ml of 0.001 *M* NaOAc (pH 3.8) and an equal volume of the same buffer containing 0.6 *M* NaCl.

Based on the amount of α1-CB7 subjected to chymotryptic digestion, the peptide C-1 was obtained from the phosphocellulose column in yields of 11–24 mol %, and C-2 in yields of 13–16 mol %.

Reconstruction of the Sequence. Quantitative data on the peptides that were subjected to automatic or to manual degradation in sequencing are given in Table IV. The sequences found for all of these peptides are shown in Table V. The various peptides were aligned to give the final sequence in the manner described below.

Several automated degradations on whole α1-CB7 (of which only the last, which gave the most information, is shown in Tables IV and V) had given the NH_2-terminal se-

Table IV: Manual and Automatic Sequence Analyses of Enzymatically Produced Fragments of Chick Skin Collagen α1-CB7.

Peptide	Position in Chain Residue No.	Amount Used (μmol)	Program[a]	Modification	Cycles of Degradation
α1-CB7	1–268	0.75	Quadrol protein	None	87
T-1	1–4	0.20	Manual, Dns	None	3
T-2	5–13	0.19	Manual, GC	None	8
T-3	14–16	0.49	Manual, GC	None	2
T-4	17–22	0.22	Manual, Dns	None	5
T-5	31–34	0.20	Manual, Dns	None	3
T-6	35–52	0.17	DMAA peptide	SPITC	13
TM-2	35–64	0.42	DMAA peptide	SPITC	30
C-1	37–81	0.25	DMAA peptide	SPITC and ANS	38
T-7	65–70	0.44	Manual, GC	None	5
T-8	71–94	0.32	DMAA peptide	SPTIC	22
TM-3	71–133	0.33	DMAA peptide	SPITC	47
C-2	82–268	0.43	Quadrol protein	None	71
T-9	95–103	0.25	DMAA peptide	SPITC	8
T-10	104–130	0.13	DMAA peptide	SPITC	27
T-11	131–133	0.20	Manual, GC	None	2
T-12	134–150	0.60	DMAA peptide	SPITC	11
TM-4	151–178	0.49	DMAA peptide	SPITC	26
T-13	176–178	0.20	Manual, Dns	None	2
T-14	179–186	0.42	Manual, GC	None	7
TM-5	179–226	1.00	Quadrol protein	SPITC	40
TM-5-T-2	203–226	0.41	DMAA peptide	SPITC	24
$\alpha 1^B$	222–268	0.23	Quadrol protein	None	25
T-15	227–235	0.27	Manual, GC	None	7
T-16	236–238	0.20	Manual, Dns	None	2
TM-6	239–262	0.82	DMAA peptide	SPITC and ANS	24
T-17	262–268	0.25	Manual, Dns	None	5

[a] Dns indicates use of the Dns-Edman degradation method, with residue identification by Dns-thin-layer chromatography. GC indicates identification by gas chromatography. Identification in all other cases was as described in the text.

Table V: Sequence of Peptides Used in Reconstructing the Sequence of α1-CB7.[a]

Peptide	
	1 5 10 15 20
α1-CB7	Hyp-Gly-Glu-Arg-Gly-Ala-Ala-Gly-Leu-Hyp-Gly-Ala-Lys-Gly-Asp-Arg-Gly-Asp-Hyp-Gly-
T-1	Hyp-Gly-Glu,Arg,
T-2	Gly-Ala-Ala-Gly-Leu-Hyp-Gly-Ala,Lys,
T-3	Gly-Asp,Arg,
T-4	Gly-Asp-Hyp-Gly-
	21 25 30 35 40
α1-CB7 (contd)	Pro-Lys-Gly-Ala-Asp-Gly-Ala-Pro-Gly-Lys-Asp-Gly-Leu-Arg-Gly-Leu-Thr-Gly-Pro-Ile-
T-4 (contd)	Pro,Lys,
T-5	Asp-Gly-Leu,Arg,
T-6	Gly-Leu-Thr-Gly-Pro-Ile-
TM-2	Gly-Leu-Thr-Gly-Pro-Ile-
C-1	Thr Gly-Pro-Ile-
	41 45 50 55 60
α1-CB7 (contd)	Gly-Pro-Hyp-Gly-Pro-Ala-Gly-Ala-Hyp-Gly Asp Lys-Gly-Glu Ala-Gly-Pro-Ala-Gly Pro
T-6 (contd)	Gly-Pro-Hyp-Gly Pro Ala Gly Ala Hyp Gly Asp,Lys,
TM-2 (contd)	Gly-Pro-Hyp-Gly-Pro-Ala-Gly-Ala-Hyp-Gly-Asp-Lys-Gly-Glu-Ala-Gly-Pro,Ala,Gly-Pro
C-1 (contd)	Gly-Pro-Hyp-Gly-Pro-Ala-Gly-Ala-Hyp-Gly-Asp Lys Gly-Glu-Ala Gly Pro-Ala-Gly Pro
	61 65 70 75 80
α1-CB7 (contd)	Ala-Gly-Thr-Arg-Gly-Ala-Hyp-Gly-Asp Arg Gly-Glu-Hyp-Gly-Pro-Hyp-Gly-Pro-Ala-Gly-
TM-2 (contd)	Ala Gly Thr,Arg,
C-1 (contd)	Ala Gly Thr Arg-Gly-Ala-Hyp-Gly-Asp-Arg-Gly-Glu-Hyp-Gly-Pro Hyp Gly Pro Ala Gly
T-7	Gly-Ala-Hyp-Gly-Asp,Arg,
T-8	Gly Glu-Hyp-Gly-Pro-Hyp-Gly-Pro-Ala-Gly-
TM-3	Gly Glu-Hyp-Gly-Pro-Hyp-Gly-Pro-Ala-Gly-
	81 85 90 95 100
α1-CB7 (contd)	Phe Ala-Gly-Pro-Hyp-Gly-Ala
C-1 (contd)	,Phe,
T-8 (contd)	Phe-Ala-Gly-Pro-Hyp-Gly-Ala Asp Gly Gln Hyp Gly Ala,Lys,
TM-3 (contd)	Phe-Ala-Gly-Pro Hyp Gly-Ala-Asp-Gly Gln Hyp-Gly-Ala Lys Gly-Glu-Thr-Gly-Asp-Ala-
C-2	Ala-Gly-Pro-Hyp-Gly-Ala-Asp-Gly-Gln-Hyp-Gly-Ala Lys Gly-Glu-Thr-Gly-Asp-Ala-
T-9	Gly Glu-Thr-Gly-Asp-Ala-
	101 105 110 115 120
TM-3 (contd)	Gly Ala Lys Gly Asp Ala-Gly-Pro Hyp Gly Pro-Ala Gly Pro Thr-Gly-Ala Hyp Gly Pro
C-2 (contd)	Gly-Ala-Lys-Gly-Asp-Ala-Gly-Pro-Hyp-Gly-Pro-Ala-Gly-Pro-Thr-Gly-Ala-Hyp-Gly-Pro-
T-9 (contd)	Gly-Ala,Lys,
T-10	Gly-Asp-Ala-Gly-Pro-Hyp-Gly-Pro-Ala-Gly-Pro-Thr-Gly-Ala-Hyp-Gly-Pro-
	121 125 130 135 140
TM-3 (contd)	Ala Gly Glx Val Gly Ala Hyp Gly Pro Hyl Gly Ala,Arg,
C-2 (contd)	Ala-Gly-Glx-Val-Gly-Ala-Hyp-Gly-Pro Hyl Gly-Ala-Arg-Gly-Ser-Ala-Gly-Pro-Hyp-Gly-
T-10 (contd)	Ala-Gly Glx Val Gly Ala Hyp Gly Pro,Hyl,
T-11	Gly-Ala,Arg,
T-12	Gly-Ser-Ala-Gly-Pro Hyp Gly
	141 145 150 155 160
C-2 (contd)	Ala-Thr-Gly-Phe-Hyp-Gly-Ala-Ala-Gly-Arg-Val-Gly
T-12 (contd)	Ala Thr Gly Phe Hyp Gly Ala Ala Gly,Arg,
TM-4	Val Gly-Pro-Hyp-Gly-Pro-Ser-Gly-Asn-Ile-

Table V (*continued*)

Peptide	
	161 165 170 175 180
TM-4 (contd)	Gly-Leu-Hyp-Gly-Pro-Hyp-Gly-Pro-Ala-Gly,Lys,Glx,Gly,Ser,Lys,Gly,Pro,Arg,
T-13	Gly-Pro,Arg,
T-14	Gly-Glu-
TM-5	Gly,Glu-
	181 185 190 195 200
T-14 (contd)	Thr-Gly-Pro-Ala-Gly,Arg,
TM-5 (contd)	Thr-Gly-Pro-Ala-Gly-Arg-Hyp-Gly-Glu-Hyp-Gly-Pro-Ala-Gly-Pro-Hyp-Gly-Pro-Hyp-Gly-
	201 205 210 215 220
TM-5 (contd)	Glu-Lys-Gly-(Ala)-Hyp-Gly (Ala) Asp Gly Pro Ile Gly (Ala) Hyp Gly Thr Pro Gly Pro Gln
TM-5-T-2	Gly (Ala)-Hyp-Gly-(Ala)-Asp-Gly-Pro-Ile-Gly-(Ala)-Hyp-Gly-Thr-Pro-Gly-Pro-Gln-
	221 225 230 235 240
TM-5 (contd)	Gly Ile Ala Gly Gln,Arg,
TM-5-T-2 (contd)	Gly-Ile-Ala Gly Gln-Arg
$\alpha 1^B$	Ile-Ala-Gly-Gln-Arg-Gly-Val-Val-Gly-Leu-Hyp-Gly-Gln-Arg-Gly-Glu Arg Gly-Phe-
T-15	Gly-Val-Val-Gly-Leu-Hyp-Gly,Gln,Arg,
T-16	Gly-Glu,Arg,
TM-6	Gly Phe-
	241 245 250 255 260
$\alpha 1^B$ (contd)	Hyp-Gly-Leu-Hyp
TM-6 (contd)	Hyp-Gly-Leu-Hyp-Gly-Pro-Ser-Gly-Glu-Hyp-Gly-Lys-Gln-Gly-Pro-Ser-Gly-Ala-Ser-Gly-
	261 265 268
TM-6 (contd)	Glu-Arg
T-17	Gly-Pro-Hyp-Gly-Pro,Met,

[a] Horizontal half-arrows indicate completion of one cycle of Edman degradation, with identification made by methods described in the text. Residues are connected by hyphens where the sequence was so established. Residues identified by compositional data only are set off by commas. Unidentified residues are shown as unconnected symbols in order to facilitate recognition of the extent of the peptide. To avoid undue complication of the table, residues in α1-CB7, C-2, and the fragment $\alpha 1^B$ are shown only as far as Edman degradation proceeded; all other peptides are shown in their entirety. The parentheses at residues 204, 207, and 213 indicate that one of these, although shown as Ala, is probably a Ser. See text for details.

quence, with relatively few gaps, as far as residue 87. This sequence clearly included the tryptic peptides T-1, T-2, T-3, T-4, T-5, T-6, and T-7, the lysine containing overlap peptides TM-1 and TM-2, and the chymotryptic peptide C-1. Inspection of the sequences found for the remaining overlap peptides then showed that the NH_2-terminal sequence of the largest of these, TM-3, could be aligned with the α1-CB7 sequence starting at residue 71, thereby extending the reconstructed α1-CB7 sequence through residue 90. The NH_2-terminal sequence of the largest of the chymotryptic peptides, C-2, which must extend to the COOH-terminal of α1-CB7 since it contains homoserine, could then be aligned with the α1-CB7 sequence starting at residue 82, and extending the reconstructed sequence through residue 152. As seen from the analysis of C-2 given in Table III, it contains a total of seven residues of lysine plus hydroxylysine, and these added to the four occurring in the sequence up to residue 82 make a total of 11 in α1-CB7, as mentioned above.

It is noted from Table I that all of the hydroxylysine of α1-CB7 is contained in the 27 residue tryptic peptide T-10, which aligns with the α1-CB7 sequence starting at residue 104. Although, as shown in Table V, this peptide was only partially sequenced, the 70% hydroxylated lysine must form its COOH-terminus, and this places this residue at 130 of α1-CB7. It was not possible to directly identify the hydroxylysine in sequencing C-2, by any of the procedures described in the Experimental Section, but the above placement of it agrees with that found by Fietzek et al. (1973), in their calf skin sequence.

The α1-CB7 sequence thus reconstructed includes, in addition to those already mentioned, the tryptic peptides T-8, T-9, T-11, and T-12. This sequence contains two valines, at residues 124 and 151, respectively. Since α1-CB7 contains only four valines, and since the sequence obtained for the $\alpha 1^B$ fragment produced by animal collagenase cleavage contains two adjacent valines, it follows that the NH_2-terminal valine of the lysine-containing overlap peptide TM-4 must be identical with the valine at residue 151. This permits the alignment of TM-4, the analysis of which indicates it to be a 28 residue peptide containing two lysines. These two lysines, as well as the arginine, glycine, proline, serine, and glutamic acid, are left unaccounted for in the sequence obtained for TM-4, as shown in Table V. Since arginine must be COOH-terminal in TM-4, the tripeptide, T-13, clearly forms the three terminal residues of TM-4, and therefore must be immediately preceded by one of the two

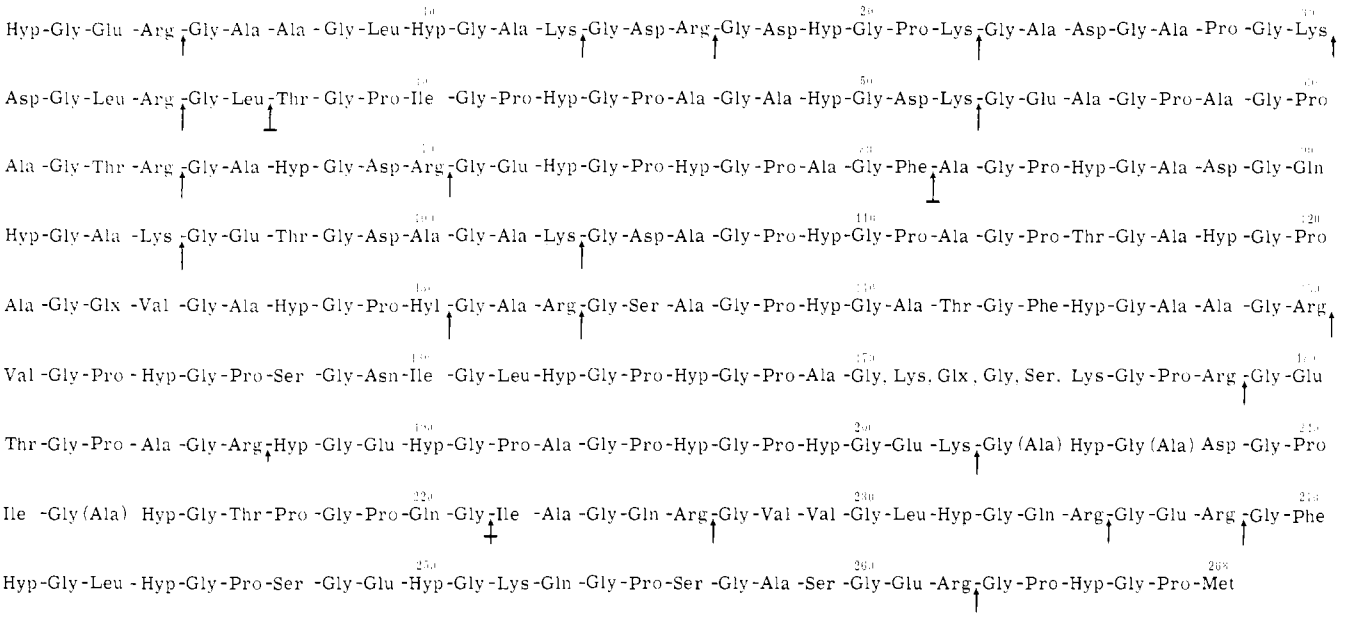

FIGURE 7: Amino acid sequence of chick skin collagen of α1-CB7. Vertical arrows indicate the observed points of enzymatic cleavage, giving rise to the peptides shown in Tables IV and V. Plain arrow, trypsin; short plain arrow, partial cleavage by trypsin; arrow with bar at distal end, chymotrypsin; arrow with midpoint bar, tadpole collagenase. See Table IV for designations of peptides produced by these cleavages.

lysines. This leaves one lysine and the glutamic acid and serine still unplaced, with blanks at 171, 172, and 174 into which they must fit. They have therefore been placed in these positions in that respective order by homology with the calf skin sequence of Fietzek et al. (1973), although in the present case the glutamyl residue must be shown as Glx.

It is known from electron microscopic studies that the site in α1 at which cleavage by animal collagenase occurs is close to residue 226 of α1-CB7 (Bruns and Gross, 1973). Turning now to the sequence obtained for the $\alpha 1^B$ fragment derived from tadpole collagenase cleavage, it is noted that the NH_2-terminal isoleucine of this must be one of the two isoleucines of TM-5, since of the four isoleucines of α1-CB7, two are already accounted for as residues 40 and 160 of the reconstructed sequence. Furthermore, as shown above, both isoleucines of TM-5 occur in TM-5-T-2, the 24 residue COOH-terminal peptide produced from TM-5 by tryptic cleavage of the latter's unblocked lysine, and only that at residue 20 of TM-5-T-2 is followed by an alanine. The NH_2-terminal isoleucine of $\alpha 1^B$ therefore must be identical with residue 20 of TM-5-T-2, and it becomes evident that the TM-5 overlap peptide immediately follows the arginine at residue 178 of α1-CB7, with the tryptide T-14 forming its first eight residues. The bond cleaved by animal collagenase is then Gly-Ile at residue 221–222, and the reconstructed sequence is extended to residue 244.

It will be noted from Tables II and V that the analysis of TM-5, and also that of its COOH-terminal component tryptide, TM-5-T-2, shows a serine that is not accounted for in the relevant portion of our sequence (179–226). Since these analyses also show one less alanine than contained in the homologous portion of the calf skin α1-CB7 sequence (Fietzek et al., 1973), it seems probable that one of the alanines at 204, 207, 213, and 223 may be a serine in chick skin. Because of scarcity of material, and the well-known tendency of PTH-serine to appear as alanine on back hydrolysis with hydriodic acid, as well as its tendency to cochromatograph with alanine in gas chromatography, we have been unable to resolve this situation further. The residue at 223 seems the least likely candidate for such a substitution, since it has been satisfactorily identified several times in $\alpha 1^B$. We have therefore chosen to show the other three residues mentioned above as alanines within parentheses, to indicate that one of them may in fact be a serine in chick skin.

The NH_2-terminal sequence of the 24-residue overlap peptide TM-6 can now be aligned with the α1-CB7 sequence starting at residue 239, residues 2 of TM-6 and 240 of α1-CB7 being identical as the third and last phenylalaline of α1-CB7. The reconstructed sequence is thus extended to residue 262. At this point all of the basic residues of α1-CB7 are accounted for, as well as all of its phenylalanines, valines, isoleucines, and leucines, and it is therefore clear that the COOH-terminal hexapeptide T-17 must follow the arginine at residue 262, and terminate α1-CB7 at residue 268. The complete sequence is shown in Figure 7.

Discussion

The chick skin collagen α1-CB7 sequence obtained by the methods described above is homologous to a high degree to that published by Fietzek et al. (1973) (see also Rexrodt et al., 1973) for the corresponding peptide from calf skin collagen. The completion of the present work permits a detailed comparison of the primary structures of the α chains of two widely divergent species over the longest segments yet possible—almost exactly one-quarter of the total length. The substitutions observed between the two species are shown in Table VI. There are 15 in all, establishing a level of sequence identity between the two species of 94%. This is in good agreement with that found in other segments of the two α chains by Dixit et al. (1975a,b).

Most of the observed substitutions may be regarded as chemically conservative, at least within a not too stringent definition of the term; possible exceptions are the substitutions at 121, 160, 162, and 211, where interchanges between smaller, hydrophilic, and larger and hydrophobic residues have occurred. Another possible exception is, of course, that at 123, where, because in the present work the identification was only by analysis after back hydrolysis, and paucity of relevant material precluded further exami-

Table VI: Amino Acid Substitutions in α1-CB7 between Chick and Calf-Skin Collagens.

Residue No.	Chick Skin	Calf Skin[a]
12	Ala	Pro
19	Hyp	Ala
33	Leu	Val
58	Ala	Ser
97	Thr	Hyp
115	Thr	Ala
117	Ala	Pro
121	Ala	Ile
123	Glx	Asn
160	Ile	Ala
162	Leu	Pro
190	Hyp	Val
193	Ala	Hyp
199	Hyp	Ala
211	Ile	Ala

[a] From Fietzek et al. (1973).

nation, the designation has to be Glx.

The isolation of the tryptic peptide T-14 in the present work duplicates a phenomenon first observed in α2-CB2—partial cleavage of an Arg–Hyp bond by tryptic digestion in Tris buffer, but not in NH_4HCO_3 buffer (Highberger et al., 1971). In the present case cleavage of Arg-Hyp at 186–187 of α1-CB7 under the former conditions led to the formation of T-14, while TM-5, the overlap peptide produced by digestion of maleylated α1-CB7 under the latter conditions, contains the uncleaved bond.

As mentioned in the introduction, one of our chief interests in α1-CB7 has been that it contains the site of the cleavage of α1 by animal collagenase. This is a highly specific cleavage, and for this reason, and because of the well-known resistance of the native collagen structure to the action of proteases in general, it had been thought that some unique structural feature might exist at this site. Examination of the amino acid sequence of α1-CB7, however, reveals no immediately obvious feature of this nature, the Gly-X-Y triplet sequence is unbroken, and nothing is apparent that is in any way unusual. It is clear the Gly–Ile bond itself cannot be the sole factor, since this occurs, in different species, in at least two other places in α1 which are apparently not cleaved. These are bond 103–104 of rat skin α1-CB8 (Balian et al., 1972) and bond 106–107 of calf skin α1-CB6 (Wendt et al., 1972). The latter is not present in chick skin α1, since the isoleucine at 107 is substituted by methionine in the chick, giving rise to the extra CNBr peptide found in this species (Dixit et al., 1974b). It is, however, probably not without significance that the chick and calf α1-CB7 sequences are identical from residue 211 to the COOH-terminus, the longest segment of α1-CB7 containing no substitutions, which also contains the site of cleavage by animal collagenase. Some minimal specific sequence around the Gly–Ile bond is probably required for cleavage, but sufficient data are not yet available to permit its definite characterization. These matters are discussed in more detail in another place (Gross et al., 1974).

Acknowledgments

The authors thank Drs. Peter Fietzek and Klaus Kühn for showing them their sequence of calf skin α1-CB7 prior to publication, and Ms. Cynthia A. Walter for carrying out the tadpole collagenase cleavage. They are grateful to Ms. Joanne T. Hazard, Ms. Karen L. Rizzie, and Mr. Wing Yu for excellent technical assistance.

References

Balian, G., Click, E. M., Hermodson, M. A., and Bornstein, P. (1972), *Biochemistry 11,* 3798.

Braunitzer, G., Schrank, B., and Ruhfus, A. (1970), *Hoppe-Seyler's Z. Physiol. Chem. 351,* 1589.

Bruns, R. R., and Gross, J. (1973), *Biochemistry 12,* 808.

Butler, P. J. G., Harris, J. I., Hartley, B. S., and Leberman, R. (1969), *Biochem. J. 112,* 679.

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

Dixit, S. N., Kang, A. H., and Gross, J. (1975a), *Biochemistry 14,* 1929.

Dixit, S. N., Seyer, J. M., Oronsky, A. O., Corbett, C., Kang, A. H., and Gross, J. (1975b), *Biochemistry 14,* 1933.

Edman, P. (1970), in Protein Sequence Determination, Needleman, S. B., Ed. New York, N.Y., Springer-Verlag, p 211.

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1,* 80.

Fietzek, P. P., Rexrodt, F. W., Hopper, K. E., and Kuhn, K. (1973), *Eur. J. Biochem. 38,* 396.

Foster, J. A., Bruenger, E., Hu, C. L., Albertson, K., and Franzblau, C. (1973), *Biochem. Biophys. Res. Commun. 56,* 70.

Gross, J., Harper, E., Harris, E. D., McCroskery, P. A., Highberger, J. H., Corbett, C., and Kang, A. H. (1974), *Biochem. Biophys. Res. Commun. 61,* 605.

Gross, J., and Lapiere, C. M. (1962), *Proc. Natl. Acad. Sci. U.S.A. 48,* 1014.

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

Harper, E., and Gross, J. (1970), *Biochim. Biophys. Acta 198,* 286.

Highberger, J. H., Kang, A. H., and Gross, J. (1971), *Biochemistry 10,* 610.

Hulmes, D. J. S., Miller, A., Parry, D. A. D., Piez, K. A., and Woodhead-Galloway, J. (1973), *J. Mol. Biol. 79,* 137.

Inagami, T., and Murakami, K. (1972), *Anal. Biochem. 47,* 501.

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

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

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

Kang, A. H., Piez, K. A., and Gross, J. (1969c), *Biochemistry 8,* 3648.

Kulbe, K. D. (1971), *Anal. Biochem. 44,* 548.

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

Nagai, Y., Lapiere, C. M., and Gross, J. (1966), *Biochemistry 5,* 3123.

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

Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 224,* 5597.

Pisano, J. J., Bronzert, T. J., and Brewer, H. B., Jr. (1972), *Anal. Biochem. 45,* 43.

Rexrodt, F. W., Hopper, K. E., Fietzek, P. P., and Kühn, K. (1973), *Eur. J. Biochem. 38,* 384.

Smithies, O., Gibson, D., Fanning, E. M., Goodfliesch, R. M., and Gilman, J. G. (1971), *Biochemistry 10,* 4912.

Wendt, P., Mark, K. v. der, Rexrodt, F., and Kühn, K. (1972), *Eur. J. Biochem. 30,* 169.

Yamada, S., and Altano, H. (1966), *Biochim. Biophys. Acta 130,* 538.