Sutherland, I. W., and Wilkinson, J. F. (1963), *J. Gen. Microbiol. 30,* 105.
Tang, S. P. W., and Coleman, J. E. (1968), *J. Biol. Chem. 243,* 4286.
Teale, F. W. J., and Weber, G. (1957), *Biochem. J. 65,* 476.
Weinryb, I., and Steiner, R. F. (1968), *Biochemistry 7,* 2488.

Covalent Structure of Collagen: Amino Acid Sequence of α1-CB3 of Chick Skin Collagen†

S. N. Dixit, A. H. Kang,* and J. Gross

ABSTRACT: The amino acid sequence of α1-CB3, a peptide containing 149 residues obtained from the central portion of the α1(I) chain of chick skin collagen by cyanogen bromide cleavage, has been determined. As in the other sequences from the helical region of collagen chains, the repeating triplet Gly-X-Y extends throughout the length of the peptide. These data allow a comparison of the sequence to that of α1-CB3 from calf and rat skin collagens. As compared with calf, the chick peptide contains 14 amino acid differences, whereas it contains 17 residue differences from the rat peptide. Thus, the sequence identity level is 91 and 89%, respectively, in comparison to the calf and rat peptides. These values are significantly greater than the value of 97% observed between the peptide of the two mammalian species and reflect the greater phylogenetic distance of the species compared.

Knowledge of the primary structure of collagen is essential in elucidating the relation of structure to function. Information on the comparative biochemistry of collagens from different classes of vertebrates may, as has been shown in other proteins, significantly contribute toward our understanding of the aspects of primary structure which determine the biologic properties of the protein. In recent years, several laboratories have been committed to the determination of the complete covalent structure of several different collagens. Progress with the amino acid sequence analysis of this very large molecule, ca. 300,000 daltons, has been greatly facilitated by the application of CNBr cleavage (reviewed by Gallop et al., 1972; Traub and Piez, 1971).

As a part of our systematic effort to determine the complete amino acid sequence of chick skin collagen, we have previously reported the covalent structures of α1-CB1, α1-CB2, α2-CB1, and α2-CB2 peptides (Kang and Gross, 1970; Highberger et al., 1971). In the present communication, we describe our data delineating the amino acid sequence of α1-CB3, which contains 149 amino acids and which comprises residues 419 through 567 of the intact α1(I) chain. The covalent structure of the homologous peptides from calf skin (Fietzek et al., 1972a; Wendt et al., 1972a) and rat skin (Butler et al., 1974a) has recently been published.

Materials and Methods

Preparation of α1-CB3. The CNBr peptide, α1-CB3, was prepared from the α1 chains of purified, salt-extracted skin collagen of lathyritic chicks. The procedural details and the criteria of purity were previously described (Kang et al., 1969a,b).

Enzymatic Hydrolyses. Digestion with trypsin (TPCK[1] treated, Worthington) was performed in 0.2 *M* Tris (pH 7.6) containing 10^{-3} *M* $CaCl_2$ at 37° for 2 hr. A 1:50 molar ratio of enzyme/substrate was used. The digestion was terminated by acidification with 2 *M* acetic acid and lyophilization. Digestion with chymotrypsin (three times recrystallized, Worthington) was performed in 0.2 *M* NH_4HCO_3 (pH 8.0) at room temperature for 2 hr. An enzyme/substrate molar ratio of 1:50 was used. The reaction was terminated by lyophilization or by separating the products by gel chromatography.

Maleylation and Demaleylation. In some experiments, the peptides were maleylated prior to trypsin digestion to confine the tryptic cleavage to the arginyl residues (Butler et al., 1969). Following tryptic c hydrolysis, the treated peptide was demaleylated by incubation in 0.2*M* pyridine acetate (pH 3.0) for 6 hr at 60°.

Column Chromatography. The initial fractionation of tryptic and chymotryptic peptides was carried out by molecular sieve chromatography on a column (2 × 120 cm) of Sephadex G-50 S (Pharmacia) equilibrated with 0.04 *M* sodium acetate (pH 4.8). Samples were applied in 2 ml of the buffer and the column was eluted with the same buffer at a flow rate of 18 ml/hr.

The peptides obtained from the molecular sieve chroma-

† From the Veterans Administration Hospital and Departments of Biochemistry and Medicine, University of Tennessee Medical School, Memphis, Tennessee 38104, and the Developmental Biology Laboratory, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts *Received October 15, 1974.* This investigation was supported by grants from the U.S. Public Health Service, No. AM16506 and AM3564, and a grant from the Gebbe Research Foundation. This study was conducted under VA Research Project No. 4826-01 and 4826-02.

* Medical Investigator of Veterans Administration, to whom all correspondence should be addressed.

[1] Abbreviations used are: CNBr, cyanogen bromide; TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone; SPhNCS, 4-sulfophenyl isothiocyanate; PTH, phenylisothiohydantion; ANS, 2-amino-1,5-naphthalenedisulfonic acid; EDC, *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide.

Table I: Amino Acid Composition of the Tryptic Peptides of α1-CB3.[a]

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	Total	α-CB3[b]
Hydroxyproline	6.1 (6)	3.0 (3)	—	1.8 (2)	0.9 (1)	0.9 (1)	1.9 (2)	0.9 (1)	1.0 (1)	—	17	14
Aspartic acid	1.0 (1)	2.1 (2)	1.0 (1)	—	—	2.9 (3)	—	—	—	—	7	7
Threonine	—	—	0.9 (1)	—	—	—	—	—	—	—	1	1
Glutamic acid	6.1 (6)	2.2 (2)	3.2 (3)	—	2.1 (2)	—	2.1 (2)	—	0.9 (1)	0.9 (1)	17	16
Proline	3.6 (4)	1.1 (1)	3.8 (4)	1.4 (1)	3.0 (3)	—	—	1.1 (1)	—	—	14	14
Glycine	14.7 (15)	6.7 (7)	6.3 (6)	5.4 (5)	4.3 (4)	4.1 (4)	4.2 (4)	2.3 (2)	2.3 (2)	1.0 (1)	50	46
Alanine	7.8 (8)	3.3 (3)	3.1 (3)	4.1 (4)	—	3.0 (3)	2.1 (2)	—	—	—	23	22
Valine	1.0 (1)	—	—	0.9 (1)	1.0 (1)	—	—	—	—	—	3	3
Leucine	1.0 (1)	1.0 (1)	—	—	—	—	—	—	—	—	2	2
Phenylalanine	1.0 (1)	—	—	—	—	—	—	0.9 (1)	1.0 (1)	—	3	3
Hydroxylysine	—	—	—	—	—	—	—	0.2	—	—	0.2	0.4
Lysine	1.0 (1)	—	—	1.0 (1)	—	1.0 (1)	1.0 (1)	0.8 (1)	—	—	4.8	4.4
Arginine	1.0 (1)	—	1.0 (1)	—	1.0 (1)	—	1.0 (1)	—	1.0 (1)	1.0 (1)	6	6
Homoserine	—	1.0 (1)	—	—	—	—	—	—	—	—	1	1
Total	45	20	19	14	12	12	12	6	6	3	149	140

[a] Values are expressed as residues per peptide. A dash indicates the level was less than 0.1 residue per peptide. Numbers in parentheses indicate assumed integral values. [b] From Kang et al., 1969b.

tography were desalted on a column (2 × 120 cm) of Bio-Gel P-2 (200–400 mesh) using 0.1 *M* acetic acid as the eluent and further fractionated on phosphocellulose. A 1 × 6 cm column of phosphocellulose was equilibrated with 0.001 *M* sodium acetate at 44°, and, after application of samples, was eluted with a linear gradient of NaCl from 0 to 0.1 *M* over a total volume of 500 ml. The effluent was continuously monitored at 230 nm in a Gilford spectrophotometer.

Amino Acid Analysis. The samples were hydrolyzed in constant boiling HCl under an atmosphere of nitrogen at 108° for 24 hr. Analyses were performed on an automatic analyzer (Beckman 121) using a single column method as described previously (Kang, 1972). No correction factors were used for losses or incomplete release of amino acids.

Edman Degradation. Automated Edman degradations (Edman and Begg, 1967) were performed with a Beckman Sequencer Model 890C employing either the Fast Protein-Quadrol (072172C) or Slow Peptide-DMAA (071472) Programs supplied by Beckman Instruments. Small peptides with COOH-terminal arginine were modified by treatment with 2-amino-1,5-naphthalenedisulfonic acid (ANS) in the presence of *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide (EDC) prior to the application of the sample to the reaction cup. The procedure used was identical with that described by Foster et al. (1973) except that only 1 mol each of ANS and EDC was used per mol of COOH-group in the peptide rather than the threefold molar excess used by Foster et al. (1973). This diminished amount of reactant apparently spared the β- or γ-carboxyl groups of aspartic or glutamic acid as we were able to identify these amino acid residues as their PTH derivatives in good yields. Small peptides with COOH-terminal lysine were treated with SPhNCS as described by Braunitzer et al. (1970) and Foster et al. (1973). These modifications improved retention of the peptides in the reaction cup allowing automatic degradation through the penultimate COOH-terminal residues. The PTH-amino acids were analyzed by gas chromatography both before and after trimethylsilylation using the general procedure of Pisano and Bronzert (1969). Under these conditions, PTH-hydroxyproline, a residue infrequently encountered in noncollagenous proteins, is well separated from other PTH-amino acids and elutes from the column at 15.0 min between PTH-leucine (retention time, 11 min) and PTH-methionine (retention time, 17 min). In addition, selected residues were also analyzed by thin-layer chromatography (Inagami and Murakami, 1972) and/or on an amino acid analyzer after hydrolysis of PTH derivatives to their parent amino acids using the method of Smithies et al. (1971). The argininyl residues were also identified by spot test with phenanthroquinone (Yamada and Itano, 1966).

Results

Isolation of Tryptic Peptides. Ten tryptic peptides were isolated and purified from the tryptic digest of α1-CB3 by a combination of molecular sieve chromatography on Sephadex G-50 S and phosphocellulose chromatography.[2] Amino acid composition of these peptides is presented in Table I. The values of the number of residues per peptide calculated from the observed yield of each amino acid are equal or close to integers in most instances, and the sum of the number of residues of each amino acid agrees within experimental error with the composition reported previously for α1-CB3 (Kang et al., 1969b).

Since the total content of the basic residues lysine and arginine in α1-CB3 is 11 (Table I), one might have expected 12 peptides to result from tryptic hydrolysis. However, in two instances the basic residue was followed by hydroxyproline (see Figure 1, residues 15 and 78) and was not cleaved by trypsin. Similar resistance of Lys(or Arg)–Pro(or Hyp) bonds to tryptic attack has previously been observed (Butler and Ponds, 1971; Highberger et al., 1971).

The Alignment of the Tryptic Peptides. In order to gain information on the alignment of the tryptic peptides, α1-CB3 was maleylated prior to trypsin digestion, and the resulting peptides were demaleylated and fractionated on Sephadex G-50 S and phosphocellulose.[2] A total of seven peptides were obtained. Amino acid analyses[2] indicated that three of these peptides, T(8-7), T(4-3), and T(6-2), were overlap peptides and consisted of uncleaved (due to maleylation) T8 and T7, T4 and T3, and T6 and T2, respectively (see Figure 1).

In addition to the characterization of the tryptic peptides derived from maleylated α1-CB3, several experiments were performed to determine the alignment of the ten tryptic

[2] See paragraph at end of paper regarding supplementary material.

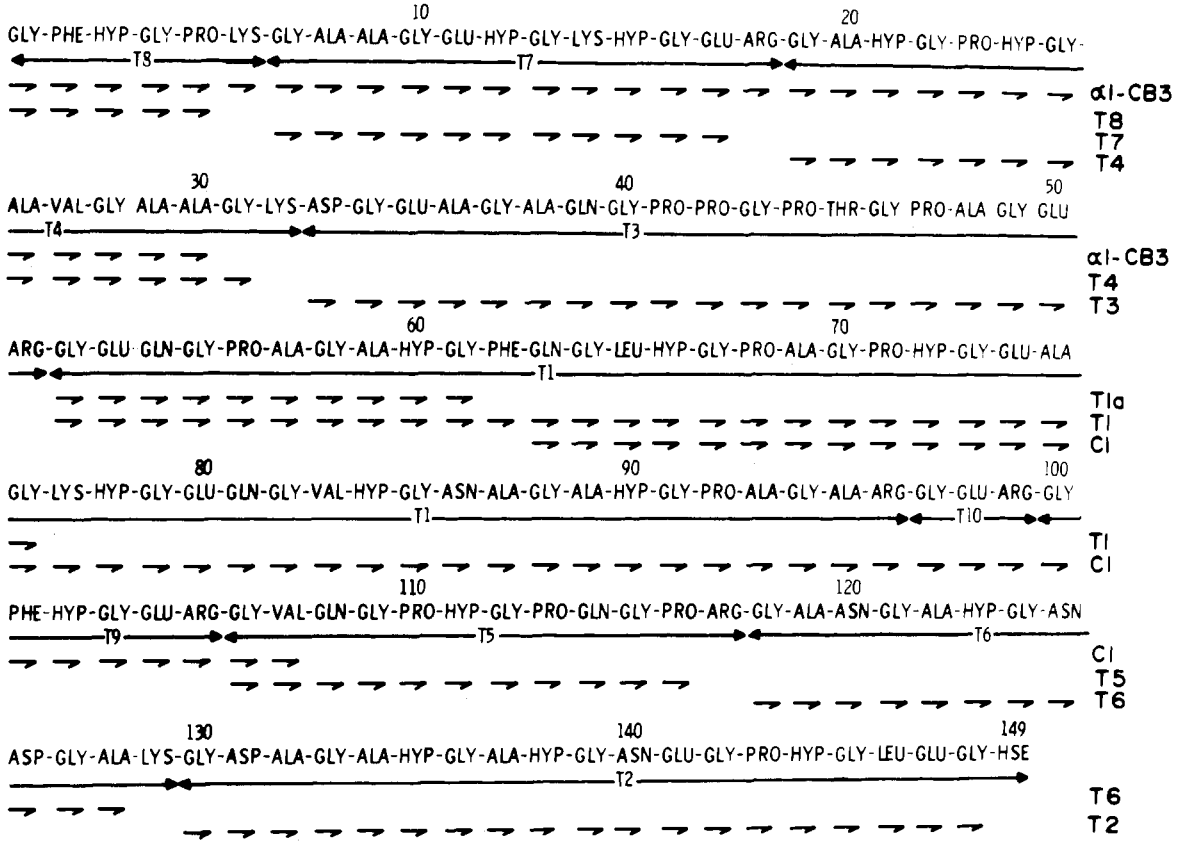

FIGURE 1: The complete amino acid sequence of α1-CB3. The tryptic peptides are indicated by long arrows (↔). Short horizontal arrows (→) indicate the extent of Edman degradation of each peptide degraded.

peptides. First, intact α1-CB3 was subjected to sequential degradation in automated sequenator, and the NH_2-terminal sequence through 30 residues was determined (see Figure 1). These data together with the amino acid compositions (Table I), further sequence studies on the individual tryptic peptides (see below), and the results obtained from maleylation experiments described above show that the order of the first four peptides is T8-T7-T4-T3.

Additional information was obtained from studies on peptide fragments derived from chymotryptic hydrolysis of α1-CB3. Two chymotryptic peptides C2 (residues 1–62) and C1 (residues 63–149) were separated from uncleaved α1-CB3 by Sephadex G-50 S chromatography.[2] The peptides C1 and C2 account for the total content of amino acids in the parent peptide.[2] The peptide C1 contains homoserine; it must be COOH-terminal, and C2 the NH_2-terminal peptide. The peptide C2 was modified by maleylation and digested with trypsin, and the resulting peptides were demaleylated and fractionated on Sephadex G-50 S; three major peptide peaks were seen.[2] Following further purification on phosphocellulose, amino acid analysis[2] indicated that peaks 1 and 2 consisted of T(4-3) and T(8-7), respectively. Peak 3 (C2-T1a) was a peptide of 11 residues containing a residue of phenylalanine but lacking a basic residue. Edman degradation of this peptide (see Figure 1) in the automatic sequenator showed that its sequence is identical with the NH_2-terminal sequence of isolated T1. These data, along with the previously deduced alignment (see above) of T8-T7-T4-T3, indicated that T1 must follow T3 in the alignment. Further supporting evidence was obtained by chymotryptic cleavage of isolated T1. Fractionation of a chymotryptic digest of isolated T1 on Sephadex G-50 S[2] resolved two peptide fragments, T1a (residues 52–62) and T1b (residues 63–96). Amino acid composition[2] of T1a is identical with the tryptic fragment obtained from C2.

The alignment of the remaining tryptic peptides were deduced as follows: T2 must be COOH-terminus since it contains a residue of homoserine. T6 must be penultimate since T(6-2) was isolated from tryptic hydrolysis of maleylated α1-CB3. The relative location of T10, T9, and T5 was determined by sequenator analysis of α1-CB3-C1, the COOH-terminal portion of the chymotryptic peptide; degradation through the first 45 residues established their relative location (see Figure 1).

In summary, the proposed alignment of the tryptic peptides of α1-CB3 is T8-T7-T3-T1-T10-T9-T5-T6-T2. These data are consistent with the data obtained for α1-CB3 of calf skin (Wendt et al., 1972a; Fietzek et al., 1972a) and rat skin (Butler et al., 1974a) collagens, except that in the latter two, the residue 14 (lysine) was followed by alanine, rather than hydroxyproline allowing tryptic cleavage at this site and thus giving rise to one additional tryptic peptide to total 11 rather than the 10 observed in chick α1-CB3.

Internal Sequences of Tryptic Peptides. Intact α1-CB3 and the individual tryptic peptides and some of the chymotryptic peptides were subjected to automated Edman degradation to determine the complete sequence of α1-CB3, which is depicted in Figure 1. Each peptide was analyzed at least twice. The relative position of each peptide in the chain, the number of useful degradation cycles achieved, the amount of peptide introduced into the sequencer, and the modifying agents used, if any, are summarized in Table II, together with other essential details. The use of blocking agents such as SPhNCS for lysyl side chains, and ANS in the presence of EDC for carboxyl groups, was essential to minimize the loss of sample during extraction. Where pep-

Table II: Automatic Sequenator Analysis of Tryptic and Chymotryptic Peptides from Chick Skin α1-CB3.

Peptide	Position in Chain	Amount Used (μmol)	Program Used[a]	Modification	Residues Degraded
α1-CB3	1–149	1.0	Quadrol	None	30
T8	1–6	0.5	DMAA	SPhNCS	5
T7	7–18	0.5	DMAA	ANS	11
T4	19–32	0.5	DMAA	SPhNCS	13
T3	33–51	0.5	DMAA	ANS	18
T1	52–96	1.0	DMAA	ANS	25
T1a	52–62	0.5	DMAA	ANS	10
T5	106–117	0.5	DMAA	ANS	11
T6	118–129	0.5	DMAA	ANS	11
T2	130–149	0.5	DMAA	ANS	19
C1	63–149	1.0	DMAA	SPhNCS	45

[a] Quadrol refers to the Fast Protein-Quadrol Program 072172C, and DMAA to Slow Peptide-DMAA Program 071472 of Beckman Instrument.

tides were modified with SPhNCS, the NH_2-terminal residues were identified by manual one-step Edman degradation on untreated aliquots of the peptides.

Discussion

The complete amino acid sequence of chick skin α1-CB3 as deduced from the present study is presented in Figure 1. These results, together with the results obtained for α1-CB7 (J. H. Highberger, A. H. Kang, C. Corbett, and J. Gross, manuscript in preparation), α1-CB4, α1-CB5, and α1-CB6A (S. N. Dixit, A. H. Kang, J. Gross, and C. Corbett, manuscript in preparation), and the previously published sequence for the NH_2-terminal peptides (Kang and Gross, 1970) extend the knowledge of the α1 chain of chick skin to 693 residues; the remaining sequence unreported is represented by two CNBr peptides, α1-CB8 (268) and α1-CB6B (85). Although the homologous regions of these peptides as well as several of the other CNBr peptides have been sequenced for rat skin (Kang et al., 1967; Bornstein, 1969; Butler, 1970; Butler and Ponds, 1971; Balian et al., 1971, 1972; Butler et al., 1974a) or calf skin collagens (Rauterberg et al., 1972a,b; Fietzek et al., 1972a,b, 1973; Wendt et al., 1972a,b), the complete amino acid sequence for the entire α chain from a single species has not yet been established. Such information would be useful in providing a clue to the evolution of this unusual structural protein and in delineating interactions which contribute to the stability of the triple helical molecule.

A comparison of the sequence of chick α1-CB3 with that of calf (Fietzek et al., 1972a) and rat (Butler et al., 1974a) shows that there are 14 and 17 residue differences, respectively. These substitutions are summarized in Figure 2. All substitutions between the chick as compared to the calf and rat peptide involve change of only one nucleotide in the triplet of DNA codons except two; these are Ala→Leu at residue 87 and Gln→Ala at residue 114. These substitutions involve a two-base change. Thus, the observed level of sequence identity between the chick peptide on the one hand and the calf and the rat peptides on the other is 91 and 89%, respectively.

The high degree of sequence identity of α1 chains displayed by phylogenetically widely spread species belonging to separate classes of vertebrates is in contrast to the relatively greater degree of variability observed in the sequences of different α chains within the same species. Butler et al. (1974b) have reported a sequence difference of 20% between α1(I)-CB3 (skin) and α1(II)-CB8 (cartilage) from the same species. Similarly, relatively greater differences were noted between the homologous regions of α1(I) and α2 chains from calf (Fietzek et al., 1972c) and rat (Piez et al., 1972).

	8	15	20	29	41	42	45	59	86	87
Chick	ALA	HYP	ALA	ALA	PRO	PRO	THR	ALA	ASN	ALA
Calf	ALA	ALA	VAL	PRO	PRO	HYP	ALA	SER	ASP	LEU
Rat	THR	ALA	VAL	PRO	ALA	HYP	ALA	SER	ASP	LEU

	93	108	114	119	132	140	141	143	147
Chick	ALA	GLN	GLN	ALA	ALA	ASN	GLU	PRO	GLU
Calf	SER	GLU	ALA	ALA	ALA	SER	GLN	ALA	GLN
Rat	SER	GLN	ALA	ASN	THR	SER	GLN	ALA	GLX

FIGURE 2: Residue differences among α1-CB3 of chick, calf, and rat skin collagens. Calf data from Fietzek et al. (1972a); rat data from Butler et al. (1974a).

The level of sequence homology maintained by collagen α1 chains from the same tissue between the avian and the mammalian species is also significantly closer than that of other proteins from these species. For example, sequence identity between the chick and bovine species for cytochrome b_5 is 68%, for hemoglobin α chain, 73%, and for homoglobin β chain, 71% not including the large segment missing in the chick chain (Dayhoff et al., 1972). These figures are significantly smaller than that for the α1-CB3 peptide of collagen even when correction is made for the invariant glycine in collagen.

Acknowledgment

The authors are grateful to Ms. C. Corbett, Ms. K. Lawley Rizzi, Mr. N. Flynn, and Mr. Gordon Randall for their excellent technical help.

Supplementary Material Available

Tables and figures containing additional data will appear following these pages in the microfilm edition of this volume of the journal. Photocopies of the supplementary material from this paper only or microfiche (105 × 148 mm, 24× reduction, negatives) containing all of the supplementary material for the papers in this issue may be obtained from the Journals Department, American Chemical Society, 1155 16th st., N. W., Washington, D. C. 20036. Remit check or money order for $4.50 for photocopy or $2.50 for microfiche, referring to code number BIO-75-1929.

References

Balian, G., Click, E. M., and Bornstein, P. (1971), *Biochemistry 10,* 4470.

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

Bornstein, P. (1969), *Biochemistry 8,* 63.

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

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

Butler, W. T. (1970), *Biochemistry 9,* 44.

Butler, W. T., Miller, E. J., Finch, J. E., and Inagami, T. (1974b), *Biochem. Biophys. Res. Commun. 57,* 190.

Butler, W. T., and Ponds, S. L. (1971), *Biochemistry 10,* 2076.

Butler, W. T., Underwood, S. P., and Finch, J. E. (1974a), *Biochemistry 13,* 2946.

Dayhoff, M. D., Eck, R. V., and Park, C. M. (1972), in

Atlas of Protein Sequence and Structure, Vol. 5, Dayhoff, Margaret O., Ed., Silver Spring, Md., National Biomedical Research Foundation, p 89.
Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1,* 80.
Fietzek, P. P., Kell, I., and Kuhn, K. (1972c), *FEBS Lett. 26,* 66.
Fietzek, P. P., Rexrodt, F. W., Hopper, K. E., and Kuhn, K. (1973), *Eur. J. Biochem. 38,* 396.
Fietzek, P. P., Rexrodt, F. W., Wendt, P., Stark, M., and Kuhn, K. (1972b), *Eur. J. Biochem. 30,* 163.
Fietzek, P. P., Wendt, P., Kell, I., and Kuhn, K. (1972a), *FEBS Lett. 26,* 74.
Foster, J. A., Bruenger, C. L., Hu, C. L., Albertson, K., and Franzblau, C. (1973), *Biochem. Biophys. Res. Commun. 53,* 70.
Gallop, P. M., Blumenfeld, O. O., and Seifter, S., (1972), *Annu. Rev. Biochem. 41,* 617.
Highberger, J. H., Kang, A. H., and Gross, J. (1971), *Biochemistry 10,* 610.
Inagami, T., and Murakami, K. (1972), *Anal. Biochem. 47,* 501–504.
Kang, A. H. (1972), *Biochemistry 11,* 1828.
Kang, A. H. Bornstein, P., and Piez, K. A. (1967), *Biochemistry 6,* 788.
Kang, A. H., and Gross, J. (1970), *Biochemistry 9,* 796.
Kang, A. H., Piez, K. A., and Gross, J. (1969a), *Biochemistry 8,* 3648.
Kang, A. H., Piez, K. A., and Gross, J. (1969b), *Biochemistry 8,* 1506.
Piez, K. A., Balian, G., Click, E. M., and Bornstein, P. (1972), *Biochem. Biophys. Res. Commun. 48,* 990.
Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 244,* 5597.
Rauterberg, J., Fietzek, P. P., Rexrodt, F. W., Becker, U., Stark, M., and Kuhn, K. (1972b), *FEBS Lett. 21,* 75.
Rauterberg, J., Timpl, R., and Furthmayr, H. (1972a), *Eur. J. Biochem. 27,* 231.
Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., and Ballantyne, D. L., (1971), *Biochemistry 10,* 4912.
Traub, W., and Piez, K. A. (1971), *Adv. Protein Chem. 25,* 243.
Wendt, P., Fietzek, P. P., and Kuhn, K. (1972a), *FEBS Lett. 26,* 69.
Wendt, P., von der Mark, K., Rexrodt, F., and Kuhn, K. (1972b), *Eur. J. Biochem. 30,* 169.
Yamada, S., and Itano, H. A. (1966), *Biochim. Biophys. Acta 130,* 538.

Covalent Structure of Collagen: Amino Acid Sequence of α1-CB6A of Chick Skin Collagen†

S. N. Dixit, J. M. Seyer, A. O. Oronsky,‡ C. Corbett, A. H. Kang,* and J. Gross

ABSTRACT: The amino acid sequence of chick skin collagen α1-CB6A, a peptide containing 107 residues obtained from the helical region near the carboxy-terminus of the α1(I) chain by cyanogen bromide cleavage, has been determined. This was accomplished by automated Edman degradation of the hydroxylamine-produced fragments and of the tryptic peptides prepared with and without prior maleylation. The data show that this portion of the α1(I) chain from chick skin is identical in 90% of the residues to the corresponding peptide region of calf skin collagen reported previously.

The primary structure of collagens from several species, in particular rat, calf, and chick, has been the subject of intensive investigation by several laboratories in recent years. The successful application of CNBr cleavage, and isolation and characterization of the resulting peptides representing segments from known regions of this very large molecule, ca. 300,000 daltons, have greatly facilitated progress on the problem (reviewed by Gallop et al., 1972; Traub and Piez, 1971).

Our laboratories have been committed to determining the complete amino acid sequence of chick skin collagen type I. As a part of this effort, we have previously reported the covalent structure of α1-CB1, α1-CB2, α1-CB3, α2-CB1, and α2-CB2 peptides (Kang and Gross, 1970; Highberger et al., 1971; Dixit et al., 1975). In the present paper, we present our data delineating the amino acid sequence of α1-CB6A, which contains 107 amino acids and which comprises residues 836–942 of the α1(I) chain. The covalent structure of α1-CB6 from calf skin, whose NH_2-terminal 107 residues is homologous to chick α1-CB6A, was recently reported (Mark et al., 1970; Fietzek et al., 1972; Wendt et al., 1972).

Materials and Methods

Preparation of α1-CB6A. The CNBr peptide, α1-CB6a, was prepared from the α1 chain of purified, salt-extracted

† From the Veterans Administration Hospital and Departments of Biochemistry and Medicine, University of Tennessee Medical School, Memphis, Tennessee 38104, and the Developmental Biology Laboratory, Departments of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114. *Received November 11, 1974.* This investigation was supported by grants from the U.S. Public Health Service (AM 16506 and AM 3564) and grants from the Gebbe Research Foundation and Tennessee Heart Association. This study was conducted under VA Research Project No. 4826-01 and 4826-02.

‡ Present address: Ciba-Geigy Pharmaceuticals, Ardsley, N. Y. 10502.

*Medical Investigator of Veterans Administration, to whom all correspondence should be addressed.