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.

skin collagen of lathyritic chicks. The details of these procedures and the criteria of purity were previously published (Kang et al., 1969a,b).

Hydroxylamine Cleavage. Cleavage with hydroxylamine (Eastman Organic Chemicals) of α1-CB6A into two fragments, HA1 and HA2,[1] was carried out in a manner described by Bornstein (1970).[2] The peptide was dissolved in distilled water (5 mg/ml) and denatured at 45° for 10 min, and an equal volume of freshly prepared, cold 2 *M* NH_2OH in 1 *M* K_2CO_3 (pH 10.5) was added. After thorough mixing, the reaction mixture was incubated at 35° for 90 min. The reaction was terminated by adjusting the pH of the incubation mixture to 4 with 6 *N* HCl, and the peptides were separated from the reagents and salts by chromatography on Bio-Gel P2 (100–200 mesh, Bio-Rad Laboratories) using 0.1 *M* acetic acid as the eluent.

Hydroxamate determinations were performed using the procedure described by Bornstein (1970).

Enzymatic Hydrolysis. Digestions with trypsin (L-1-tosylamido-2-phenylethyl chloromethyl ketone treated, Worthington) were performed in 0.2 *M* Tris (pH 7.6) containing 1×10^{-3} *M* $CaCl_2$. An enzyme:substrate molar ratio of 1:50 was used. The enzyme was added as a 0.5% solution in 1×10^{-3} *M* HCl. Incubations were performed at 37° for 4 hr. Digestions were terminated by acidification with 1 *M* acetic acid followed by lyophilization.

Maleylation. The peptides α1-CB6A and its hydroxylamine-produced fragment, HA1, were maleylated by the procedure of Butler et al. (1969). After tryptic digestion, maleyl groups were removed by incubation in 0.2 *M* pyridine acetate (pH 3.0) at 60° for 6 hr.

Column Chromatography. The initial fractionation of the hydroxylamine-produced fragments and the tryptic peptides was carried out on a column (2 × 130 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 same buffer and the column was eluted at a flow rate of 18 ml/hr.

Phosphocellulose chromatography was performed on a 1 × 6 cm column of phosphocellulose (Whatman) equilibrated with 0.001 *M* sodium acetate (pH 3.8) at 44°. After application of samples dissolved in 5 ml of the buffer, the column was eluted with a linear gradient of NaCl from 0 to 0.1 *M* over a total volume of 500 ml.

The column effluents were monitored continuously at 230 nm in a Gilford spectrophotometer equipped with flow cells. The peptides or peptide fractions were desalted on columns (2 × 120 cm) of Bio-Gel P2, 200–400 mesh (Bio-Rad Laboratories), using 0.1 *M* acetic acid as the eluent.

Automated peptide analyses were performed with volatile buffers on a 0.9 × 25 cm column of PA-35 resin (Beckman) using an automatic analyzer (Beckman) equipped with a stream-split device. Details of these procedures were previously described (Kang and Gross, 1970).

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

Edman Degradation. Automated Edman degradations were performed with a Beckman Sequencer, Model 890C, according to the principles described by Edman and Begg (1967). Either the Fast Protein-Quadrol (072172C) or the Slow Peptide-DMAA (071472) programs was employed to operate the instrument. Smaller peptides with COOH-terminal lysine were modified by treatment with SPhNCS by the methods of Braunitzer et al. (1970). The NH_2-terminal residue was identified by a one-step manual Edman degradation on a separate aliquot. Peptides ending in COOH-terminal arginine were treated with 2-amino-1,5-naphthalenedisulfonic acid (ANS) in the presence of *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide (EDC). The procedure used was identical with that described by Foster et al. (1973) except that only 1 mol of the reactants was used per mol of the COOH-groups in the peptide. The diminution of the reactant concentration apparently spares the β or γ COOH-groups of aspartic or glutamic acids to a large extent, as we were able to obtain these residues as their PTH derivatives in good yields (Dixit et al., 1975). These prior chemical modifications of peptides improved the retention of the peptides in the reaction cup of the automatic sequenator, allowing useful degradation through the penultimate COOH-terminus in many instances. The PTH-amino acids were identified by gas chromatography, both before and after trimethylsilylation (Pisano and Bronzert, 1969; Dixit et al., 1975). In certain instances, the amino acid derivatives were also identified by thin-layer chromatography (Inagami and Murakami, 1972) and/or by amino acid analysis after hydrolysis of PTH derivatives to their parent amino acids as described by Smithies et al. (1971). The arginyl derivative was also identified by phenanthroquinone spot test (Yamada and Itano, 1966).

Results

Cleavage of α1-CB6A with Hydroxylamine. The products of hydroxylamine reaction of α1-CB6A were fractionated on Sephadex G-50 S.[3] Three well-separated peaks were observed. The first peak eluting in the position of the starting material was unchanged in amino acid composition and apparently represents uncleaved material. Retreatment of this fraction with hydroxylamine under the same conditions gave rise to little or no further cleavage, indicating that the partial cleavage was not due to incomplete reaction. Similar partial cleavage of α1-CB3 and α1-CB8 by hydroxylamine was previously reported (Butler, 1969; Bornstein, 1970).

In addition to the resistant fragment, two peaks, α1-CB6A-HA1 and HA2, were consistently observed. Each of the fragments eluted as homogeneous material when subjected to phosphocellulose chromatography. The molecular weights of these peptide fragments, as determined by molecular sieve chromatography on an Agarose column calibrated with collagen chains and peptides of known molecular weight (Piez, 1968), were 5800 and 4000 for HA1 and HA2, respectively. The sum of the amino acid compositions of the two fragments is identical, within experimental error, with that of α1-CB6A (Table I), indicating that cleavage of a single bond produced HA1 and HA2. The fragment HA2

[1] Abbreviations used are: HA, hydroxylamine-produced fragment; SPhNCS, 4-sulfophenyl isothiocyanate; PTH, phenylthiohydantoin; ANS, 2-amino-1,5-naphthalenedisulfonic acid; EDC, *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride.

[2] The method has been subsequently modified (Balian et al., 1971) to improve the yield.

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

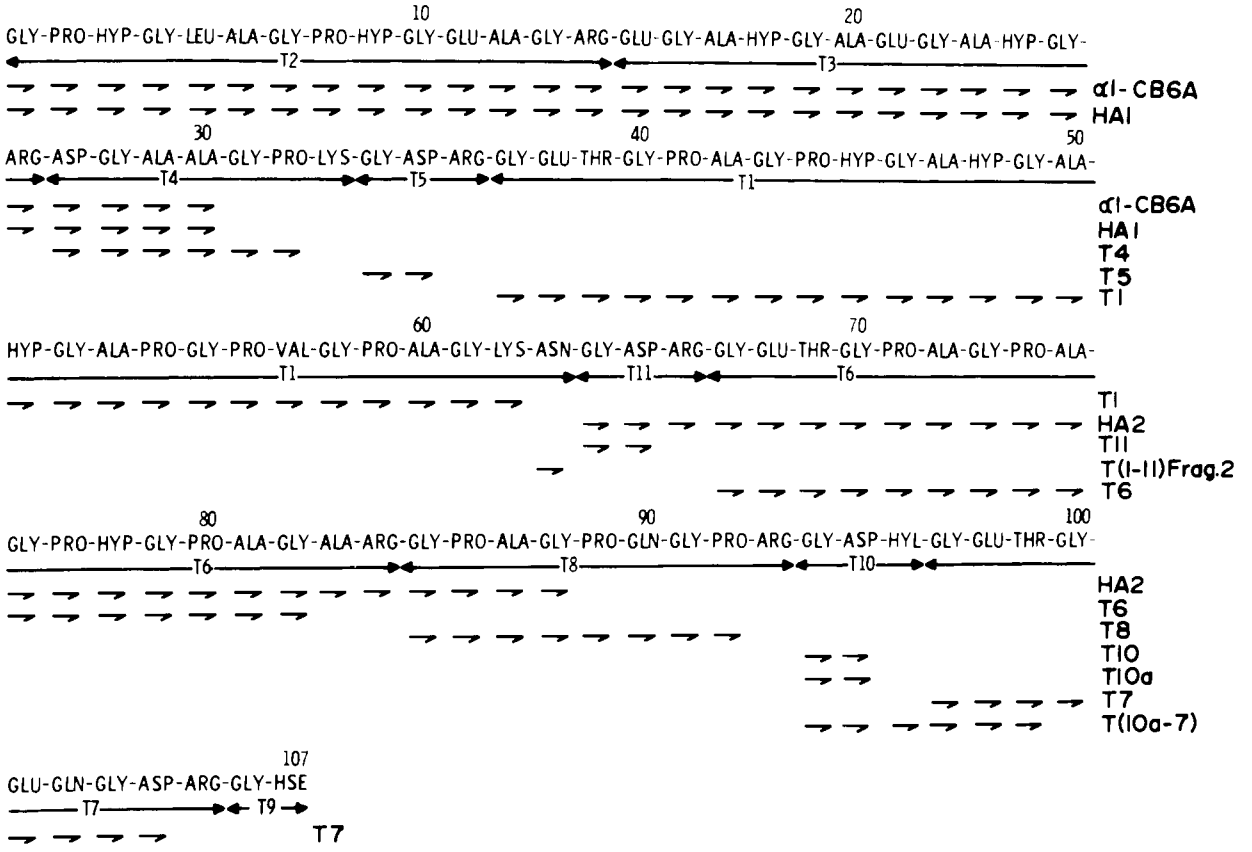

FIGURE 1: The complete amino acid sequence of α1-CB6A. The tryptic peptides are indicated by long arrows (↔). Short horizontal arrows (→) indicate the extent of Edman degradation of each peptide degraded. Hydroxylamine cleaves the bond between the residues 63 and 64.

Table I: Amino Acid Composition of α1-CB6A, Its Hydroxylamine-Produced Fragments, HA1 and HA2, and Tryptic Peptides.[a]

	α1-CB6A	HA1	HA2	T1	T2	T3	T4	T4a	T5	T6	T7	T8	T9	T10	T10a	T11	T(10a-7)	Total[b]
Hydroxyproline	8.3	7.3	1.4	3.2	1.6	1.7				0.7								8
Aspartic acid	6.1	3.2	3.0	0.9			1.0	1.0	0.9		0.9			0.9	1.0	1.0	1.7	6
Threonine	2.8	0.9	1.8	1.0						1.0	0.8						0.8	3
Glutamic acid	9.1	4.2	5.3	1.2	1.2	2.0				1.2	2.9	1.1					2.9	9
Proline	15	7.6	6.6	4.7	2.4		1.0	1.0		4.0		3.0						15
Glycine	36	21	15	9.1	5.3	4.3	1.9	1.9	1.0	6.4	3.1	2.9	1.1	1.0	1.0	1.2	3.9	36
Alanine	17	12	5.0	5.3	2.1	3.1	1.9	1.9		4.1		0.9						17
Valine	1.0	1.0		0.9														1
Leucine	1.0	1.0			1.0													1
Hydroxylysine	0.8	0.2	0.6					1.0							1.0		1.0	[c]
Lysine	2.2	1.8	0.5	0.8			1.0							1.0				3
Arginine	7.1	3.3	4.0		1.0	1.0			1.0	1.1	1.0	1.0				1.0	1.0	7
Homoserine	1.0		1.0										1.0					1
Total	107	63	44	27	14	12	7	7	3	18	9	9	2	3	3	3	12	107

[a] Expressed as residues per peptide. Values for residues present in numbers greater than 10 were rounded off to the nearest whole number. A dash indicates 0.1 residue or less. [b] Total of T1 through T11 not including T4a, T10a, or T(10a-7). [c] Present only in T4a, T10a, and T(10a-7).

contains a residue of homoserine, and, therefore, must constitute the COOH-terminal sequence. HA1 contains 1 equiv of hydroxamate per peptide[3] and therefore contains the acyl moiety of the hydroxylamine-sensitive bond (see Figure 1).

Tryptic Peptides of α1-CB6A-HA1. The tryptic digest of α1-CB6A-HA1 was separated by cation-exchange chromatography on PA-35 (Beckman). Six distinct peaks (T1, T2, T3, T4, T4a, and T5) were obtained.[3] Each of the peptides was homogeneous as judged by amino acid analysis (Table I). Of these, T4 and T4a had identical composition except that the latter contained a residue of hydroxylysine instead of a residue of lysine. Similar partial hydroxylation of lysyl residues in collagen were reported previously (Butler, 1968). On the basis of the relative yields of T4 and T4a the degree of hydroxylation at this locus was estimated to be approximately 15%. Hydroxamate analyses[3] of these tryptic peptides indicated that only HA1-T1 contained 1 equiv per peptide of hydroxamate. Although one might have anticipated the hydroxamate-containing tryptic peptide not to contain a basic residue, the presence of the lysyl residue in

Table II: Automated Sequenator Analyses of Tryptic Peptides and Hydroxylamine-Produced Fragments of Chick Skin α1-CB6A.[a]

Peptide	Position in Chain	Amount Applied (μmol)	Program Used[b]	Modification	Residues Degraded
α1-CB6A	1–107	1.0	Quadrol	None	30
HA1	1–63	1.0	Quadrol	SPhNCS	30
HA2	64–107	1.0	Quadrol	SPhNCS	25
T4	27–33	0.5	DMAA	SPhNCS	6
T5	34–36	0.5	DMAA	ANS	2
T1	37–63	1.0	DMAA	SPhNCS	26
T11	64–66	0.5	DMAA	ANS	2
T(1–11) fragment 2	63–66	0.5	DMAA	None	1
T6	67–84	0.5	DMAA	ANS	16
T8	85–93	0.5	DMAA	ANS	8
T10	94–96	0.5	DMAA	SPhNCS	2
T10a	94–96	0.5	DMAA	SPhNCS	2
T7	97–105	0.5	DMAA	ANS	8
T(10a-7)[c]	94–105	0.5	DMAA	ANS	6

[a] Samples analyzed in Beckman Sequencer 890C. [b] Quadrol refers to the Fast Protein-Quadrol Program 072172C and DMAA to Slow Peptide-DMAA Program 071472 of Beckman Instrument. [c] Residue number 3 of this peptide, hydroxylysine, was not identified as PTH derivative or by amino acid analysis after back hydrolysis, but was inferred from amino acid composition and sequence analysis of T10a.

T1 was apparently due to the penultimate location of the amino acid and the consequent resistance to the tryptic attack.[4]

Tryptic Peptides of α1-CB6A-HA2. The fractionation of a tryptic digest of HA2 on PA-35 yielded eight well-resolved peaks (T6, T7, T8, T9, T10a, T(10a-7), T10, and T11). Their amino acid composition is presented in Table I. The peptides T10 and T10a have an identical composition except for the substitution of lysine with hydroxylysine in the latter. It is interesting to note that the particular hydroxylysyl bond appears partially resistant to tryptic attack under the experimental conditions used as evidence by the isolation of a peptide, T(10a-7), which has the composition consistent with it being an uncleaved peptide consisting of T10a and T7. The resistance is partial since retreatment with trypsin of the isolated T(10a-7) and chromatography on Sephadex G-50 S yielded additional T10a and T7. Based on the recovery yields of the peptides T10, T10a, and T(10a-7), it is estimated that approximately 60% of the basic residue appears hydroxylated.

Thus, by a combination of the above procedures, a total of 11 tryptic peptides were isolated from HA1 and HA2. The sum of the composition of all the tryptic peptides is, within experimental error, identical with the observed composition of α1-CB6A (see Table I).

The Alignment of the Tryptic Peptides. In order to gain information on the alignment of the tryptic peptides, α1-CB6A and HA1 were maleylated prior to tryptic hydrolysis, confining the cleavage to the arginyl residues. By a combination of molecular sieve chromatography on Sephadex G-50 S and phosphocellulose chromatography,[3] two overlap peptides were obtained: T(4-5) representing an uncleaved peptide consisting of T4 and T5, and T(1-11) consisting of T1 and T11 (see Figure 1).

In addition, several experiments were performed to determine the alignment of the 11 tryptic peptides. First, intact α1-CB6A and HA1 were subjected to automated Edman degradation. The results were identical. Starting with 1.0 μmol of the peptides, the following NH_2-terminal sequence was obtained: Gly-Pro-Hyp-Gly-Leu-Ala-Gly-Pro-Hyp-Gly-Glu-Ala-Gly-Arg-Glu-Gly-Ala-Hyp-Gly-Ala-Glu-Gly-Ala-Hyp-Gly-Arg-Asp-Gly-Ala-Ala. These results, along with amino acid composition (Table I) and further sequence analyses on the individual tryptic peptides (see below), and the results obtained from maleylation experiments described above show that the first four peptides in the alignment are T2-T3-T4-T5. Since HA1 contains only one additional tryptic peptide, T1, it must follow T5. Furthermore, the isolation of T(1-11) from the tryptic hydrolysate of maleylated α1-CB6A indicates that T11 must follow T1.

The alignment of the tryptic peptides in α1-CB6A-HA2 was deduced from the data obtained from a sequential degradation of HA2, modified by treating with SPhNCS, in the automated sequencer. The following NH_2-terminal sequence was observed: Gly-Asp-Arg-Gly-Glu-Thr-Gly-Pro-Ala-Gly-Pro-Ala-Gly-Pro-Hyp-Gly-Pro-Ala-Gly-Ala-Arg-Gly-Pro-Ala-Gly. These results together with amino acid composition (see Table I) and sequence studies on the individual peptides (see below) show that the order of the first three peptides is T11-T6-T8. The peptide T9 contains a residue of homoserine, and therefore must be the COOH-terminus. The remaining two peptides T10 and T7 were located between T8 and T9 by exclusion. Further, the order of these two peptides must be T10-T7 as indicated by the isolation of the peptide T(10a-7) (see Table I) and the sequence studies on the peptide (see below).

In summary, the alignment of the 11 tryptic peptides of α1-CB6A was deduced to be T2-T3-T4-T5-T1-T11-T6-T10-T7-T9. The proposed alignment is consistent with the data obtained for α1-CB6 of calf skin collagen (Mark et al., 1970; Fietzek et al., 1972; Wendt et al., 1972).

Internal Sequences of Tryptic Peptides. Intact α1-CB6A, HA1, HA2, and several tryptic peptides were subjected to automated Edman degradation to determine the complete amino acid sequence of α1-CB6A, which is given in Figure 1. Each peptide was analyzed at least twice. Table II summarizes the location of the peptides analyzed, the number of useful cycles degraded, the amounts of each peptide applied to the sequencer reaction cup, and the use of blocking agents, if any. Where SPhNCS was used, the NH_2-terminal residue was identified by a one-step manual degradation on an untreated aliquot of the peptide. The use

[4] In the present paper, the nomenclature of the tryptic peptides is based on the tryptic peptides of the hydroxylamine fragments, HA1 and HA2. Note that HA1-T1 includes a COOH-terminal Asn at residue 63 (see Figure 1).

	11	23	29	30	32	44	47	63	75	78
Chick	ALA	ALA	ALA	ALA	PRO	PRO	ALA	ASN	ALA	HYP
Calf	SER	SER	SER	HYP	ALA	ALA	PRO	SER	ILE	VAL

	95	98	101	102	104	107
Chick	ASP	GLU	GLU	GLN	ASP	MET
Calf	ASX	GLX	GLX	GLX	ASX	ILE

FIGURE 2: Residue differences between α1-CB6A of chick and the corresponding segment from calf skin collagens. Calf data from Wendt et al. (1972).

of blocking agents, such as SPhNCS for lysyl side chains, and ANS in the presence of EDC for peptides not containing lysine was most helpful in minimizing the loss of the peptide from the reaction cup during extraction, allowing in many instances the degradation to proceed to the penultimate residues.

The location of Asn at residue 63, however, deserves some clarification. Tryptic hydrolysis of T(1-11) and subsequent fractionation of the products on Bio-Gel P2 (200–400 mesh) yielded two peptides: T(1-11) fragment 1 (residues 37–62) and fragment 2 (residues 63–66). Amino acid analyses of these two fragments indicated that fragment 1 had a composition identical with that of T1 except for the absence of a residue of aspartic acid, and that fragment 2 was identical with T11 except for an additional aspartic acid residue. One-step Edman degradation of T(1-11) fragment 2 indicated it to be Asn. These data together with the sequence data on T1 and T11 located Asn at residue 63.

Discussion

The complete amino acid sequence of chick skin α1-CB6A as determined from the present study is presented in Figure 1. The results allow a comparison with the sequence of the homologous portion of calf skin α1-CB6 (Mark et al., 1970; Wendt et al., 1972). A total of at least 11 differences are present; these are summarized in Figure 2. Determination of the state of amidation of two Asx and three Glx residues in calf skin may increase the number to 16. All 11 interspecies substitutions involve change of only one nucleotide in the triplet DNA condons except two; these are Ala->Ile at residue 75 and Pro(Hyp)->Val at residue 78. The latter involve two base changes. Thus, the observed level of sequence identity between the peptides of two species is 90%, a figure similar to that obtained for α1-CB3 (Dixit et al., 1975). The high degree of interspecies sequence identity of the α1 chains was also reported previously for α1-CB3 (Butler et al., 1974; Dixit et al., 1975) of rat skin collagen.

The subintegral number of hydroxylysine in α1-CB6A indicates the occurrence of partial hydroxylation of lysine in this sequence. On the basis of the relative yields of the tryptic peptides T4 and T4a obtained in this study, the lysyl residue at position 33 was estimated to be hydroxylated approximately 15%. Similarly, the lysyl residue at position 96 is about 60% hydroxylated. At neither position is hydroxylysine glycosylated, as 2 *N* NaOH hydrolysis and subsequent amino acid analysis of T4a and T10a yielded no detectable galactosyl- or glucosylgalactosylhydroxylysine. As noted previously (Fietzek et al., 1973), both of the lysine hydroxylations occurred at Y positions in the Gly-X-Y triplet sequence.

The amino acid sequence of α1-CB6A presented here displays several features which were noted previously for the other known helical regions of the collagen chains. This includes not only the regular occurrence of glycine at every third position, but also the preferential distribution of certain of the other amino acids at positions X or Y in the triplet sequence (Balian et al., 1972; Fietzek et al., 1973). Thus, five of seven glutamic acid residues are found in position X, whereas five of seven arginine residues are located at position Y. All three residues of threonine are found in position Y and the single residue of leucine occurs at position X. Presumably, this type of nonrandom distribution of some amino acids in X and Y positions may play a role in interactions between side chains in the triple-stranded molecule or between molecules during aggregation and fibril formation, but the precise basis for the selective distribution is not presently understood.

Hydroxylamine cleaves α1-CB6A between residues 63 and 64 (asparaginyl-glycyl) producing two fragments, HA1 and HA2. It was previously shown to cleave asparaginyl-glycyl bonds in α1-CB3 (Butler, 1969) and α1-CB8 (Bornstein, 1970) of rat skin collagen. The proposed mechanism involves cyclization of the asparaginyl-glycyl bond to give rise to the cyclic imide, anhydroaspartylglycine, which then is cleaved by nucleophilic attack (Bornstein, 1970). Of the five aspartyl residues in α1-CB6A, four (residues 35, 65, 95, and 104) precede either arginine or lysine, making cyclic imide formation at these sites unlikely; one precedes the glycyl residue at position 27. The absence of cleavage at this site is consistent with the suggestion that amidation of the aspartyl side chain enhances the likelihood of imide formation (Bornstein, 1970). Hydroxylamine cleavage therefore is a useful supplement to other nonenzymatic cleavage in sequence determination of collagen chains.

Acknowledgment

The authors are grateful to Ms. K. L. Rizzi, Ms. C. Dickey, and Mr. G. Randall for their expert 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 BI0-75-1933.

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. (1970), *Biochemistry 9,* 2408.

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. (1968), *Science 161,* 796.

Butler, W. T. (1969), *J. Biol. Chem. 244,* 3415.

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

Dixit, S. N., Kang, A. H., and Gross, J. (1975), *Biochemistry,* preceding paper in this issue.

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1,* 80.
Fietzek, P., Rexrodt, F. W., Hopper, K. E., and Kühn, K. (1973), *Eur. J. Biochem. 38,* 396.
Fietzek, P. P., Rexrodt, F. W., Wendt, P., Stark, M., and Kühn, K. (1972), *Eur. J. Biochem. 30,* 163.
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., 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.
Mark, K. v.d., Wendt, P., Rexrodt, F., and Kühn, K. (1970), *FEBS Lett.* 11,105.
Piez, K. A. (1968), *Anal. Biochem. 26,* 305.
Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 244,* 5597.
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., Mark, K. v.d., Rexrodt, F., and Kuhn, K. (1972), *Eur. J. Biochem. 30,* 168.
Yamada, S., and Itano, H. A. (1966), *Biochim. Biophys. Acta 130,* 538.

The Amino Acid Sequence of Ferredoxin II from *Chlorobium limicola,* a Photosynthetic Green Bacterium†

Masaru Tanaka, Mitsuru Haniu, Kerry T. Yasunobu,* M. C. W. Evans, and Krishna K. Rao

ABSTRACT: The amino acid sequence of ferredoxin II from the photosynthetic green sulfur-reducing bacterium, *Chlorobium limicola,* was deduced to be: Ala-His-Arg-Ile-Thr-Glu-Glu-Cys-Thr-Tyr-Cys-Ala-Ala-Cys-Glu-Pro-Glu-Cys-Pro-Val-Asn-Ala-Ile-Ser-Ala-Gly-Asp-Glu-Ile-Tyr-Ile-Val-Asp-Glu-Ser-Val-Cys-Thr-Asp-Cys-Glu-Gly-Tyr-Tyr-Asp-Glu-Pro-Ala-Cys-Val-Ala-Val-Cys-Pro-Val-Asp-Cys-Ile-Ile-Lys-Val. The ferredoxin was shown to consist of 61 amino acids in a single polypeptide chain. The presence of 8 g-atoms of Fe and 8 mol of sulfide led to a calculated molecular weight of 7289. In contrast to the ferredoxin I from *C. limicola,* ferredoxin II contains basic amino acids in positions 2 and 3 and 60 from the NH_2-terminal end of the protein. The sequences of all the various ferredoxins from photosynthetic bacteria reported to date are compared with one another.

The amino acid sequence of ferredoxin I, one of the two types of ferredoxin present in the photosynthetic green sulfur-reducing bacterium, *Chlorobium limicola,* has been determined (Tanaka et al., 1974). The amino acid sequence of ferredoxin from a photosynthetic, but purple sulfur-reducing bacterium, *Chromatium,* has also been reported by Matsubara et al. (1970). From the comparison of the amino acid sequences of the *Chlorobium limicola* ferredoxin I and the *Chromatium* ferredoxin, it was pointed out that both ferredoxins show great sequence homology although the latter contains 21 additional amino acids at the C-terminal end of the protein.

Both of these ferredoxins are of the 8 Fe, 8 S^{2-}-type in which the iron is probably chelated to 8 cysteine residues to form two iron–sulfur clusters which contain 4 Fe and 4 S^{2-} each (Adman et al., 1973).

Now the amino acid sequence of ferredoxin II, the other ferredoxin in the photosynthetic green bacterium, *Chlorobium limicola,* has been determined. The present report gives the details of the research which has led to the elucidation of the primary structure of the *Chlorobium limicola* ferredoxin II.

Materials and Methods

Chlorobium limicola Ferredoxin II. The preparation of ferredoxin from a pure culture of *C. limicola* and the procedures for the isolation and purification of the CM[1] derivative of *C. limicola* ferredoxin II have already been described (Tanaka et al., 1974). The *C. limicola* CM-ferredoxin II was then further purified by column rechromatography on Dowex 1-X2 and was used in the present sequence studies.

Chymotrypsin was obtained from the Worthington Biochemical Corp. as a three-times crystallized preparation. L-(1-Tosylamido-2-lysyl)ethyl chloromethyl ketone was purchased from Cyclo Chemical.

† From the Department of Biochemistry-Biophysics, University of Hawaii, Honolulu, Hawaii 96822, and the Botany Department, University of London King's College, London, England. *Received October 7, 1974.* This project was supported by Grants GM 16784 and GM 16228 from the National Institutes of Health, The National Science Foundation (Grants GB 18739 and GB 43448), and the Science Research Council of Great Britain.

[1] Abbreviations used are: CM, carboxymethyl; PTH, phenylthiohydantoin.