and Professor L. Mandelkern for discussing their results prior to publication.

References

Bovey, F. A., Ryan, J. J., and Hood, F. P. (1968), *Macromolecules 1*, 305.
Brant, D. A., Miller, W. G., and Flory, P. J. (1967), *J. Mol. Biol. 23*, 47.
Deber, C. M., Bovey, F. A., Carver, J. P., and Blout, E. R. (1970), *J. Amer. Chem. Soc. 92*, 6191.
DeTar, D. F., Gouge, M., Honsberg, W., and Honsberg, U. (1967a), *J. Amer. Chem. Soc. 89*, 988.
DeTar, D. F., Rogers, F. F., Jr., and Bach, H. (1967b), *J. Amer. Chem. Soc. 89*, 3039.
DeTar, D. F., and Vajda, T. (1967), *J. Amer. Chem. Soc. 89*, 998.
Downie, A. R., and Randall, A. A. (1959), *Trans. Faraday Soc. 55*, 2132.
Edsall, J. T., Flory, P. J., Kendrew, J. C., Liquori, A. M., Nemethy, G., Ramachandran, G. N., and Scheraga, H. A. (1966a), *Biopolymers 4*, 121.
Edsall, J. T., Flory, P. J., Kendrew, J. C., Liquori, A. M., Nemethy, G., Ramachandran, G. N., and Scheraga, H. A. (1966b), *J. Biol. Chem. 241*, 1004.
Edsall, J. T., Flory, P. J., Kendrew, J. C., Liquori, A. M., Nemethy, G., Ramachandran, G. N., and Scheraga, H. A. (1966c), *J. Mol. Biol. 15*, 399.
Eyman, D. P., and Drago, R. S. (1966), *J. Amer. Chem. Soc. 88*, 1617.
Madison, V., and Schellman, J. A. (1970), *Biopolymers 9*, 561.
Mattice, W. L., and Mandelkern, L. (1970), *J. Amer. Chem. Soc. 92*, 5285.
Mattice, W. L., and Mandelkern, L. (1971a), *Biochemistry 10*, 1926.
Mattice, W. L., and Mandelkern, L. (1971b), *Biochemistry 10*, 1934.
Schimmel, P. R., and Flory, P. J. (1968), *J. Mol. Biol. 34*, 105.
Steinberg, I., Harrington, W. F., Berger, A., Sela, M., and Katchalski, E. (1960), *J. Amer. Chem. Soc. 82*, 5263.
Torchia, D. A., and Bovey, F. A. (1971), *Macromolecules 4*, 246.
Torchia, D. A., di Corato, A., Wong, S. C. K., Deber, C. M., and Blout, E. R. (1972a), *J. Amer. Chem. Soc. 94*, 609.
Torchia, D. A., Wong, S. C. K., Deber, C. M., and Blout, E. R. (1972b), *J. Amer. Chem. Soc. 94*, 616.

Cyanogen Bromide Cleavage of Guinea Pig Skin Collagen. Isolation and Characterization of Peptides from the α1 and α2 Chains†

C. C. Clark‡ and Paul Bornstein*

ABSTRACT: After CNBr cleavage of the chromatographically purified α1 and α2 chains of guinea pig skin collagen (GPSC), eight peptides were isolated from each chain. These peptides were resolved by a combination of ion-exchange and gel filtration chromatography, and characterized by determination of their amino acid composition and molecular weight. The 16 peptides obtained in this manner adequately account for the amino acid composition and molecular weight of the respective intact chains. In contrast to the collagens of other species, the α1 and α2 chains of GPSC display a different number and distribution of methionyl residues. In the α1 chain, a methionine is missing in the sequence homologous to α1-CB7 plus α1-CB6 of other species; while in the α2 chain, two additional methionines exist in the sequence homologous to α2-CB4 of other species. Consequently, the chromatographic patterns of CNBr peptides, particularly those obtained from the α2 chain, differ considerably from those of other collagens. GPSC is, nevertheless, clearly similar to other collagens in overall amino acid composition and in molecular weight. This study provides a base line for subsequent investigations which utilize the CNBr peptides of GPSC, particularly those designed to determine the distribution of intermolecular cross-links in the insoluble protein.

Cleavage of collagen at methionyl residues with cyanogen bromide (CNBr) has proven to be a useful technique in a variety of studies aimed at characterizing this important macromolecule. Such studies include the comparative biochemistry of collagen chains from different tissues and species (see Piez *et al.* (1968) for a review), and the localization and characterization of an intramolecular covalent cross-link in the protein (Bornstein and Piez, 1966; Bornstein *et al.*, 1966; Kang *et al.*, 1969c). Recently this approach was directed to an analysis of insoluble collagen in an attempt to isolate and identify peptides containing intermolecular cross-links (Miller, 1971; Volpin and Veis, 1971a).

These experiments were initiated in an attempt to establish a good model system for the investigation of insoluble collagen. In order to interpret the complex pattern of peptides obtained

† From the Department of Biochemistry and Research Training Unit, Department of Medicine, University of Washington School of Medicine, Seattle, Washington 98195. *Received December 3, 1971.* This work was supported by National Institutes of Health Grants AM 11248 and HD 04872.

‡ Postdoctoral Fellow, supported by National Institutes of Health Grant 5-F02-AM 45106.

* Recipient of Research Career Development Award K4-AM-42582 from the U. S. Public Health Service.

by CNBr digestion of insoluble collagen and, thereby, to derive information concerning the distribution of intermolecular cross-links, preliminary studies must establish the amino acid composition and chromatographic properties of peptides from the constituent polypeptide chains of the soluble protein. This paper describes the isolation and characterization of all the CNBr peptides from the $\alpha 1$ and $\alpha 2$ chains of soluble guinea pig skin collagen. The guinea pig was chosen because its small size makes isotopic studies feasible and because, unlike the rat, the guinea pig produces a highly insoluble skin collagen.

Unlike type I collagens[1] previously analyzed, GPSC[2] contains a different number and distribution of methionyl residues, primarily in the $\alpha 2$ chain. As a result, eight peptides were isolated from $\alpha 2$. Three of these have no homologous counterpart in other collagen $\alpha 2$ chains. In addition to the significance of these findings for comparative studies, knowledge of the structure of GPSC is useful in view of its use in immunochemical studies (Michaeli *et al.*, 1969, 1971; Timpl *et al.*, 1968) and as a substrate for mammalian collagenases (see Eisen *et al.*, 1970).

Materials and Methods

Preparation of Guinea Pig Skin Collagen. Skins from 250-g male guinea pigs were shaved, minced, washed with water, and lyophilized. The dry skin was powdered in a Wiley mill in the presence of Dry Ice, and lipids were extracted by the procedure of Nikkari and Heikkinen (1968). The dried skin powder was subjected to sequential extraction at 4° with 0.45 M NaCl–0.05 M Tris (pH 7.5), followed by 1 M NaCl–0.05 M Tris (pH 7.5), and twice with 0.5 M acetic acid. The resulting soluble collagens were purified as described previously (Piez *et al.*, 1963; Bornstein and Piez, 1966).

Preparation of Collagen Chains. Soluble collagens were denatured and chromatographed on CM-cellulose (Bornstein and Piez, 1966). By using a large column (5 × 30 cm) and a limiting buffer of slightly higher ionic strength (0.20), 2 g of collagen could be routinely resolved into $\alpha 1$ plus β_{11}, β_{12}, and $\alpha 2$. No attempt was made to separate $\alpha 1$ from β_{11}, but preparations of $\alpha 2$ were rechromatographed before use. Yields of collagen chains were approximately 60%.

CNBr Cleavage. Cleavage of α chains with CNBr was performed essentially as described by Epstein *et al.* (1971). Under these conditions, there was greater than 90% conversion of methionine to homoserine.

Ion-Exchange Chromatography. CNBr-produced peptides were chromatographed on phosphocellulose and CM-cellulose (Whatman) essentially as previously described (Bornstein and Piez, 1966; Butler *et al.*, 1967; Click and Bornstein, 1970).

Molecular Sieve Chromatography. CNBr peptides were concentrated by lyophilization and desalted on Bio-Gel P-2 (100–200 or 200–400 mesh) equilibrated with 0.03 M ammonium propionate (pH 4.5). Fractionation of the smaller CNBr peptides from $\alpha 1$ and $\alpha 2$ was accomplished on Sephadex G-75 equilibrated with 0.03 M ammonium propionate and on Bio-Gel P-6 equilibrated with 1 M $CaCl_2$–0.05 M Tris (pH 7.5).

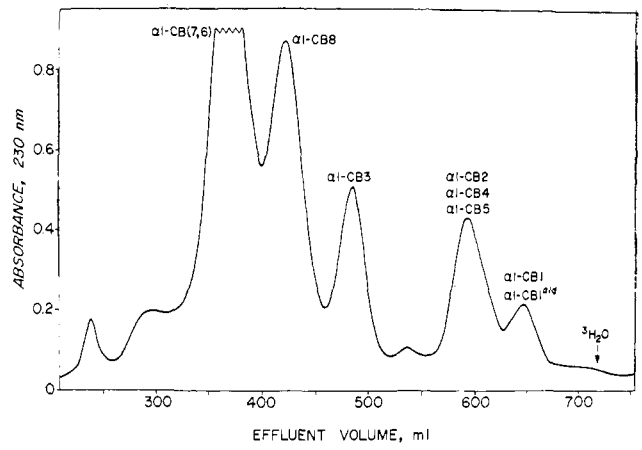

FIGURE 1: Agarose molecular sieve chromatography of the product of CNBr digestion of 70 mg of GPSC $\alpha 1$ chain. A 2.5 × 170 cm column was equilibrated with 1 M $CaCl_2$–0.05 M Tris (pH 7.5).

Gel filtration on agarose (Bio-Gel A 1.5 m) equilibrated with 1 M $CaCl_2$–0.05 M Tris (pH 7.5) was used both as a preparative and as an analytical tool. Molecular sieve columns were calibrated using appropriate CNBr peptides from rat skin collagen as standards and tritiated water to mark the included volume (Piez, 1968).

Amino Acid Analyses. Analyses were performed as described previously (Click and Bornstein, 1970). The compositions presented for each peptide represent the averages of at least two analyses.

Results

CNBr Peptides Derived from α1 (Table I)

α1-CB0 was isolated from a digest of $\alpha 1$ as described by Click and Bornstein (1970) for $\alpha 2$-CB0. A dipeptide (Glu,Hse) identical in composition with a peptide isolated from rat tendon (Bornstein, 1969) and chick (Kang *et al.*, 1969b; Miller *et al.*, 1969) collagen $\alpha 1$ chains was demonstrated in this manner (Table I).

α1-CB1 and α1-CB1^{ald}. After initial separation on agarose (Figure 1) or phosphocellulose (Figure 2), these peptides were purified by rechromatography on Bio-Gel P-6 (chromatograms not shown). The difference in compositions between

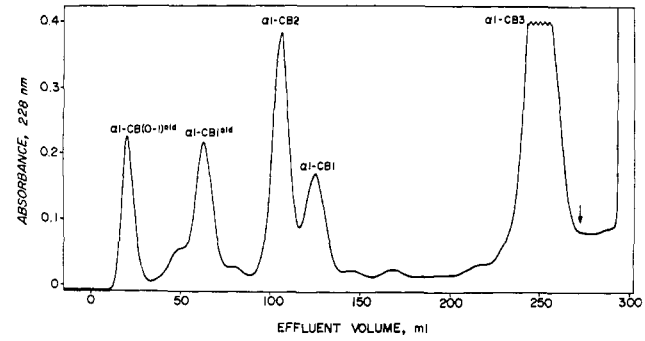

FIGURE 2: Phosphocellulose chromatography of the peptides obtained after CNBr digestion of 50 mg of GPSC $\alpha 1$ plus β_{11}. A 1.5 × 10 cm column was equilibrated in 0.001 M sodium acetate (pH 3.6) amd the peptides were eluted with a linear gradient of NaCl from 0 to 0.3 M over a volume of 350 ml. After elution of $\alpha 1$-CB3, the remaining peptides were eluted at pH 4.8 with 0.03 M sodium acetate–0.5 M NaCl (arrow).

[1] Collagens having a chain composition of $(\alpha 1)_2\alpha 2$ are designated type I to distinguish them from collagens composed of three identical $\alpha 1$ chains. The latter collagens are termed either type II (from cartilage) or type III (from skin) and have relatively different amino acid compositions and methionine contents.

[2] Abbreviation used is: GPSC, guinea pig skin collagen.

TABLE I: Amino Acid Compositions of CNBr Peptides from the α1 Chain of Guinea Pig Skin Collagen.[a]

	α1-CB0	α1-CB1	α1-CB2	α1-CB3	α1-CB4	α1-CB5	α1-CB(7,6)	α1-CB8	Total Peptides	α1[b]
3-Hydroxyproline	0	0	0	0	0	0	0.7	0	1	1.2
4-Hydroxyproline	0	0	5.5	15.9	6.4	2.9	41	27	99	99
Aspartic acid	0	1.0	0	6.2	2.8	3.0	22	9.9	45	46
Threonine	0	0	0	0	1.0	0	10	4.6	16	16
Serine	0	1.8	2.1	4.1	0	2.0	21	12	43	44
Homoserine[c]	0.9	0.9	1.0	1.0	0.8	0.9	0	1.0	7	6.5[d]
Glutamic acid	1.0	1.2	4.2	17	3.0	3.2	31	20	80	78
Proline	0	1.9	6.0	13.1	5.2	2.0	65	32	125	129
Glycine	0	3.4	12	51	16	12	155	91	340	342
Alanine	0	2.0	2.1	22	3.1	4.1	54	36	123	122
Valine	0	2.0	0	4.2	0	0	5.7	3.9	16	15
Isoleucine	0	0	0	0	0	0	5.5	2.0	8	7.9
Leucine	0	0	1.0	3.0	1.9	1.0	8.9	3.8	20	20
Tyrosine	0	1.8	0	0	0	0	1.2	0	3	3.8
Phenylalanine	0	0	0.9	2.8	0	1.0	4.5	2.7	12	12.2
Hydroxylysine	0	0	0	0.6	0	1.0	2.0	1.0	5	5.1
Lysine	0	1.0	0	4.4	1.7	1.8	12.9	8.1	30	29
Histidine	0	0	0	0	0	1.0	1.0	0	2	2.0
Arginine	0	0	1.0	6.0	3.8	1.0	23	14	49	49
Total	2	17	36	151	46	37	465	270	1024	1026

[a] Values are expressed as residues per peptide. Actual values are listed for residues present as less than 10 residues and for hydroxyproline, proline, hydroxylysine, and lysine since partial hydroxylation can give rise to nonintegral values for these amino acids. A value of zero indicates less than 0.2 residue. [b] Values calculated for a molecular weight of 93,000 and an average residue weight of 90.7. [c] Includes homoserine lactone. [d] As methionine.

α1-CB1 (Table I) and α1-CB1[ald] was limited to the presence of a single lysyl residue in the former peptide. Another peptide, α1-CB(0–1)[ald] (Figure 2), contained the uncleaved methionyl peptide bond linking α1-CB0 and α1-CB1[ald].

α1-CB2, α1-CB4, and α1-CB5 were not resolved on agarose (Figure 1), but complete separation was achieved by rechromatography on phosphocellulose (Figure 3). Alternatively, α1-CB2 was obtained by chromatography of a digest of α1 on phosphocellulose (Figure 2). α1-CB4 and α1-CB5 were also resolved by rechromatography on CM-cellulose of the material adhering to phosphocellulose. The column was equilibrated at pH 3.6 with 0.02 M sodium citrate and a linear gradient of NaCl from 0.02 to 0.14 M over a volume of 600 ml was used to elute the peptides (chromatogram not shown).

α1-CB3. After agarose chromatography (Figure 1), this peptide was purified by chromatography on phosphocellulose equilibrated with 0.001 M sodium acetate (pH 3.6). A linear gradient of NaCl from 0 to 0.4 M over a total volume of 300 ml was used (chromatogram not shown). α1-CB3 could also be prepared on phosphocellulose from a digest of α1 (Figure 2).

α1-CB(7,6). Agarose chromatography of a CNBr digest of guinea pig skin α1 revealed a distribution of peptides with

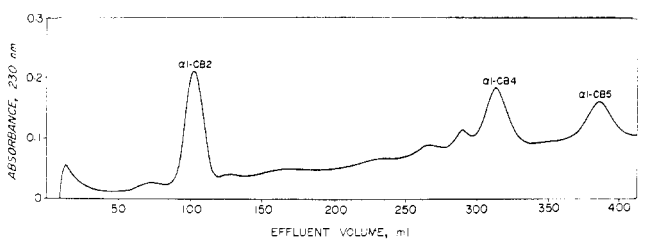

FIGURE 3: Phosphocellulose chromatography of a mixture of α1-CB2, α1-CB4, and α1-CB5 (7 mg) isolated by agarose chromatography (Figure 1). The column was equilibrated as described in the legend to Figure 2, but the gradient was extended to 0.6 M NaCl over a volume of 450 ml.

TABLE II: Molecular Weights of CNBr Peptides from the α1 Chain of Guinea Pig Skin Collagen.

Peptide	Amino Acid Anal.	Mol Sieve Chromatography
α1-CB0	181	nd[a]
α1-CB1	1,542	1,400[b]
α1-CB2	3,265	3,300[b]
α1-CB3	13,696	13,500[c]
α1-CB4	4,172	nd[a]
α1-CB5	3,356	nd[a]
α1-CB(7,6)	42,176	43,000[c]
α1-CB8	24,489	24,000[c]
Total	92,877	92,909[d]

[a] Not determined by molecular sieve chromatography. [b] Determined by Bio-Gel P-6 chromatography. [c] Determined by agarose chromatography. [d] Assuming the molecular weights determined by amino acid analysis for α1-CB0, α1-CB4, and α1-CB5.

TABLE III: Amino Acid Compositions of CNBr Peptides from the α2 Chain of Guinea Pig Skin Collagen.[a]

	α2-CB0	α2-CB1[b]	α2-CB2	α2-CB4A	α2-CB4B	α2-CB4C	α2-CB(3–5)	Total Peptides	α2[c]
4-Hydroxyproline	0	0	2.8	20.6	8.0	2.0	46	80	84
Aspartic acid	0	1.0	2.0	9.6	2.2	1.2	35	51	52
Threonine	0	0	1.8	3.6	1.8	0	13	21	21
Serine	0	0	1.0	12	1.3	1.1	20	35	36
Homoserine[d]	0.9	0.9	0.9	1.0	0.9	1.0	1.0	7	6.9[e]
Glutamic acid	0	1.0	1.1	13	7.2	1.2	42	65	68
Proline	0	1.9	3.0	19.7	9.2	2.8	80	117	117
Glycine	1.4	4.8	10	72	26	7.8	219	341	350
Alanine	0	0	2.0	25	6.4	2.1	64	99	98
Valine	0	1.0	1.0	11	2.1	0	18	33	36
Isoleucine	0	0	0	2.8	0	0	12	15	15
Leucine	0.9	1.0	1.0	8.7	1.0	1.1	16	30	30
Tyrosine	0	0.8	0	0	0	0	2.6	4	3.6
Phenylalanine	0	0	0	1.0	2.4	0	8.2	11	12
Hydroxylysine	0	0	0	2.2	0.8	0	4.2	7	8.0
Lysine	0	0.5	0	4.0	2.0	1.0	14.5	22	20
Histidine	0	0	0	0	1.7	0	6.8	9	9.7
Arginine	0	1.2	2.7	10	5.0	1.2	31	51	51
Total	3	14	30	218	77	22	633	998	1019

[a] Values are expressed as residues per peptide. Actual values are listed for residues present as less than 10 residues and for hydroxyproline, proline, hydroxylysine, and lysine since partial hydroxylation can give rise to nonintegral values for these amino acids. A value of zero indicates less than 0.2 residue. [b] Represents the average of α2-CB1 and α2-CB1^{ald}. [c] Values calculated for a molecular weight of 93,000 and an average residue weight of 91.4. [d] Includes homoserine lactone. [e] As methionine.

molecular weights different from those obtained with digests of α1 from other species (compare Figure 1 with Figure 2 of Vuust and Piez (1970)). There was no peptide in the 18,000 molecular weight range (α1-CB6), but a large peak preceded the 24,000 molecular weight range. This peptide had a molecular weight of 43,000 (Table II) and could be purified on a 1.5 × 8 cm CM-cellulose column equilibrated with 0.04 M sodium acetate (pH 4.8) using a linear gradient of NaCl from 0 to 0.4 M over a total volume of 300 ml (chromatogram not shown).

The identification of this peptide as α1-CB(7,6) was further established by amino acid analysis (Table I) showing the absence of homoserine (and methionine). This indicated that the peptide must be COOH terminal in the α1 chain and excluded the possibility that the peptide arose as a result of incomplete cleavage at a methionyl residue. These results, coupled with the knowledge (by homology) that the COOH-terminal peptide, α1-CB6, is preceded by α1-CB7 (Piez *et al.*, 1969), permit the conclusion that the guinea pig skin α1 chain is lacking a methionyl residue normally present approximately 200 residues from the COOH terminus.

α1-CB8 was isolated as described above for α1-CB(7,6).

The compositions of the eight CNBr peptides account, within experimental error, for the composition of the α1 chain (Table I). The molecular weights of the peptides, estimated by molecular sieve chromatography, agree well with the weights calculated by amino acid analysis and with the total molecular weight of the α1 chain (Table II). In addition, analysis of the area under each peptide peak (Figure 1) when compared with the total number of residues contained in that peptide, indicated that all the peptides were present in equivalent amounts (Miller *et al.*, 1969).

CNBr Peptides Derived from α2 (Table III)

α2-CB0 was isolated from a digest of α2 as described for α1-CB0. The tripeptide (Gly,Leu,Hse) is identical in composition with α2-CB0 isolated from rat (Fietzek and Piez, 1969), chick (Kang *et al.*, 1969a; Lane and Miller, 1969), bovine (Volpin and Veis, 1971b), baboon (Epstein *et al.*, 1971), and human (Click and Bornstein, 1970; Epstein *et al.*, 1971) collagen α2 chains.

α2-CB1 and α2-CB1^{ald}. After separation of α2-CB0, the remaining digest was chromatographed on Sephadex G-75 (Figure 4). The fractions representing the low molecular weight peptides (α2-CB4C, α2-CB2, α2-CB1, and α2-CB1^{ald}) were pooled and rechromatographed on Bio-Gel P-6 (Figure 5). Amino acid analysis of the peaks in the molecular weight range of 1200 confirmed the existence of both α2-CB1 and α2-CB1^{ald}.

α2-CB2 was isolated as described for α2-CB1 (Figures 4 and 5).

α2-CB3, α2-CB5, and α2-CB(3–5) were prepared by agarose chromatography of the large peptides from α2. The resulting pattern (Figure 6) showed the presence of a large peak in the molecular weight range of 60,000 as well as a broad peak in the molecular weight range 20,000–30,000 (Table IV). Amino acid analysis of the former peak confirmed its identity as the uncleaved peptide α2-CB(3–5) (Table III). This identification is supported by the presence of one residue of homoserine, but no methionine. These findings suggest that conversion of methionine to homoserine occurred without concomitant cleavage of the methionyl peptide bond. A similar uncleaved peptide has been isolated from chick (Lane and Miller, 1969; Igarashi *et al.*, 1970), rat (Vuust *et al.*, 1970), human (Click and Bornstein, 1970;

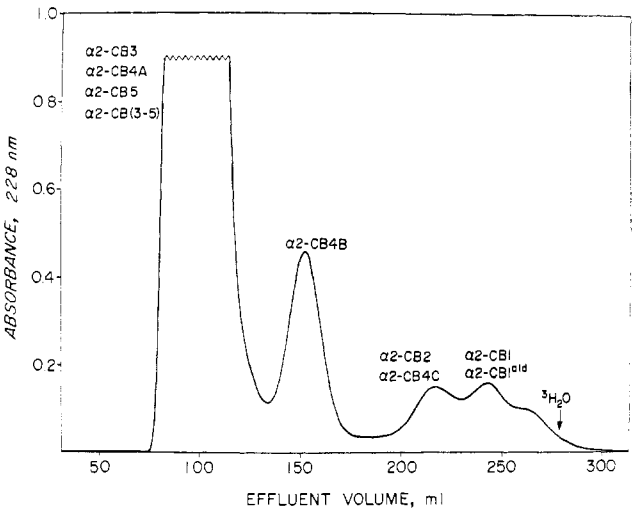

FIGURE 4: Molecular sieve chromatography (Sephadex G-75) of the product of CNBr digestion of 50 mg of GPSC $\alpha 2$ chain. A 1.8 × 90 cm column was equilibrated with 0.03 M ammonium propionate (pH 4.5).

Epstein *et al.*, 1971), and bovine (Volpin and Veis, 1971b) collagen $\alpha 2$ chains. In contrast to other studies in which $\alpha 2$-CB(3–5) was present as a minor component, in digests of guinea pig skin $\alpha 2$ chain this large peptide represents 0.6 or more of a peptide equivalent.

The peptides in the 20,000–30,000 molecular weight range (Figure 6) were rechromatographed on CM-cellulose at pH 3.6 using citrate buffers as previously described. This procedure completely separated $\alpha 2$-CB4A (see below) from $\alpha 2$-CB3 and $\alpha 2$-CB5 (chromatogram not shown). The latter peptides could be partially separated by further rechromatography on CM-cellulose at pH 4.8. However, the resolution was poor and the resulting peptides were impure.

$\alpha 2$-CB4. Resolution of CNBr digests of purified $\alpha 2$ chains as described above failed to detect the presence of a peptide homologous to $\alpha 2$-CB4 found in other species. It soon became apparent that the sequence in the GPSC $\alpha 2$ chain which corresponded to $\alpha 2$-CB4 contained two additional methionyl residues. This is consistent with the higher methionine content of this chain (Table III) relative to the $\alpha 2$ chains of other collagens. Thus, instead of a single peptide, three peptides were produced by CNBr cleavage of this region. These have been designated as $\alpha 2$-CB4A, $\alpha 2$-CB4B, and $\alpha 2$-CB4C and, on the basis of homology, account for the composition

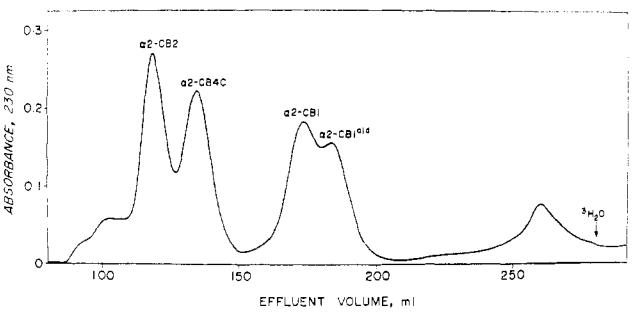

FIGURE 5: Molecular sieve chromatography on Bio-Gel P-6 of a mixture of $\alpha 2$-CB1, $\alpha 2$-CB1^{ald}, $\alpha 2$-CB2, and $\alpha 2$-CB4C (Figure 4). A 1.8 × 95 cm column was equilibrated with 1 M $CaCl_2$–0.05 M Tris (pH 7.5).

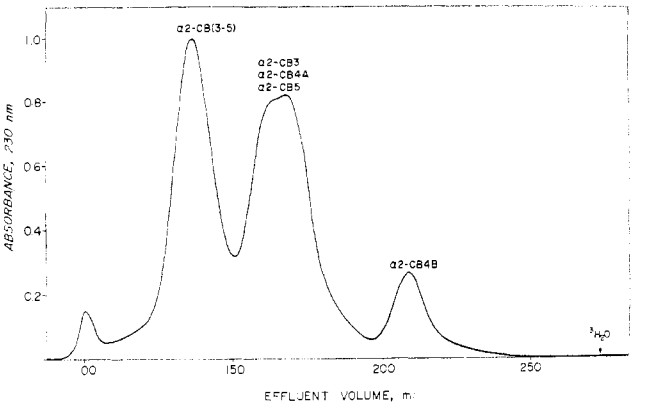

FIGURE 6: Agarose molecular sieve chromatography of the large peptides from a CNBr digest of GPSC $\alpha 2$ chain (23 mg). A 1.5 × 140 cm column was equilibrated with 1 M $CaCl_2$–0.05 M Tris (pH 7.5).

(Table III) and molecular weight (Table IV) of $\alpha 2$-CB4 from other species.

$\alpha 2$-CB4A was purified on CM-cellulose at pH 3.6 as described in the previous section. Its amino acid composition (Table III) and molecular weight (Table IV) suggest that it corresponds to more than two-thirds of the total residues in $\alpha 2$-CB4 from other species.

$\alpha 2$-CB4B was isolated in a relatively pure form from digests of $\alpha 2$ by gel filtration chromatography on either Sephadex G-75 (Figure 4) or agarose (Figure 6).

$\alpha 2$-CB4C was obtained as previously described for $\alpha 2$-CB1 (Figures 4 and 5).

The composition of the eight CNBr peptides adequately account for the composition of the $\alpha 2$ chain (Table III). Table IV shows that the molecular weights of the peptides, estimated by molecular sieve chromatography, agree well with those calculated by amino acid analysis. Together, the molecular weights of the peptides total the molecular weight of the $\alpha 2$ chain. As described for $\alpha 1$ CNBr peptides, analysis of areas under peaks from representative chromatograms indi-

TABLE IV: Molecular Weights of CNBr Peptides from the $\alpha 2$ Chain of Guinea Pig Skin Collagen.

Peptide	Amino Acid Anal.	Mol Sieve Chromatography
$\alpha 2$-CB0	274	nd[a]
$\alpha 2$-CB1	1,280	1,200[b]
$\alpha 2$-CB2	2,742	2,800[b]
$\alpha 2$-CB4A	19,925	20,000[c]
$\alpha 2$-CB4B	7,038	6,800[c,d]
$\alpha 2$-CB4C	2,011	2,200[b]
$\alpha 2$-CB(3–5)	57,856	60,000[c]
Total	91,126	93,274[e]

[a] Not determined by molecular sieve chromatography. [b] Determined by Bio-Gel P-6 chromatography. [c] Determined by agarose chromatography. [d] Determined by Sephadex G-75 chromatography. [e] Assuming the molecular weight determined by amino acid analysis for $\alpha 2$-CB0.

cated that all the peptides from α2 were present in equivalent amounts (Miller *et al.*, 1969).

Discussion

These studies characterize the CNBr peptides from soluble GPSC and provide data essential to subsequent investigations which utilize either the collagen molecule or the polymeric protein. The sixteen peptides isolated in stoichiometric amounts account for the total amino acid composition and molecular weight of the α1 and α2 chains which constitute the collagen molecule. In view of the homology among the CNBr peptides from rat, chick, bovine, baboon, and human collagens, it can be assumed that the order of the peptides from GPSC is identical with that previously established for these type I collagens; *viz.*, 0-1-2-4-5-8-3-7,6 for α1 (Piez *et al.*, 1969) and 1-0-(4A+4B+4C)-2-3-5 for α2 (Igarashi *et al.*, 1970; Vuust *et al.*, 1970).

In contrast to the species listed above, a single methionyl residue approximately 200 residues from the COOH terminus of the α1 chain is missing in GPSC. As a result, the COOH-terminal CNBr peptide from α1 is quite large (mol wt 43,000). The absence of this methionyl residue should be noted since GPSC is widely used as a substrate for mammalian collagenases. The site of initial cleavage of these enzymes in rat skin collagen was localized in part to the α1-CB7 region of the α1 chain by subsequent CNBr digestion (Piez *et al.*, 1968). Characterization of the products released from GPSC by mammalian collagenases must therefore take into account that subsequent CNBr digestion will have no effect on $\alpha 1^{B}$ (Kang *et al.*, 1966).

In the species thus far examined, the distribution of methionines, particularly in α2, has been relatively invariant leading to the isolation of homologous peptides with identical (within experimental error) molecular weights and similar amino acid compositions. However, the GPSC α2 chain possesses two additional methionines (Table III), and hence three unique peptides not found in rat, chick, bovine, baboon, and human collagen α2 chains. The additional methionines in the GPSC α2 chain were localized to the region occupied by α2-CB4 since peptides homologous to α2-CB0, α2-CB1, α2-CB2, and α2-CB(3–5) of other species were identified. In addition, the sums of the molecular weights (Table IV) and amino acid compositions (Table III) of α2-CB4A, α2-CB4B, and α2-CB4C are similar to those reported for α2-CB4 from other species. The nomenclature used for these peptides is consistent with that suggested by Miller *et al.* (1969) to designate peptides resulting from the occurrence of additional methionyl residues in a sequence corresponding to a single peptide from rat skin collagen. The precise order of these peptides in the α2-CB4 region has not yet been established.

Comparison of the CNBr peptides of GPSC with those from other type I collagens shows that the compositions of GPSC α1-CB2, α1-CB4, and α1-CB5 (Table I) are identical with those of the corresponding peptides from human and baboon skin collagens (Click and Bornstein, 1970; Epstein *et al.*, 1971) except for an apparent increase in the hydroxylation of proline in GPSC α1-CB4. In addition, the compositions of the remaining CNBr peptides from the helical region of the collagen molecule (Tables I and III) are very similar to the compositions of the corresponding peptides from other collagens. Thus, despite the difference in the number and distribution of methionyl residues which leads to an altered chromatographic pattern of CNBr peptides, the homology of GPSC with the collagens of other species is evident.

It has been observed that the NH_2-terminal sequence of the chains, particularly that of the α2 chain, exhibits a somewhat greater interspecies variability than does the main body of the collagen molecule (Bornstein, 1968; Bornstein and Kang, 1970). However, GPSC α2-CB1 is identical in composition with the corresponding human and baboon peptides. This is of some interest insofar as a major antigenic determinant of both human and GPSC is at the NH_2 terminus of the α2 chain (Michaeli *et al.*, 1969, 1971; Michaeli and Epstein, 1971). If the sequence of α2-CB1 from these two collagens is also identical, it may be possible to prepare relatively specific antibodies to human collagen using the more readily available GPSC. In the α1 chain of GPSC, the composition of the NH_2-terminal sequence (α1-CB0 plus α1-CB1) is identical with that of the corresponding region from chick skin and bone collagen (Kang *et al.*, 1969b; Miller *et al.*, 1969), but differs slightly from that of the other species examined.

Acknowledgment

The authors gratefully acknowledge the skillful technical assistance of David Wright and Felicia Arguelles.

References

Bornstein, P. (1968), *Science 161*, 592.
Bornstein, P. (1969), *Biochemistry 8*, 63.
Bornstein, P., and Kang, A. H. (1970), *in* Chemistry and Molecular Biology of the Intercellular Matrix, Vol. I, Balazs, E. A., Ed., New York, N. Y., Academic, p 99.
Bornstein, P., Kang, A. H., and Piez, K. A. (1966), *Proc. Nat. Acad. Sci. U. S. 55*, 417.
Bornstein, P., and Piez, K. A. (1966), *Biochemistry 5*, 3460.
Butler, W. T., Piez, K. A., and Bornstein, P. (1967), *Biochemistry 6*, 3771.
Click, E. M., and Bornstein, P. (1970), *Biochemistry 9*, 4699.
Eisen, A. Z., Bauer, E. A., and Jeffrey, J. J. (1970), *J. Invest. Derm. 55*, 359.
Epstein, E. H., Jr., Scott, R. D., Miller, E. J., and Piez, K. A. (1971), *J. Biol. Chem. 246*, 1718.
Fietzek, P. P., and Piez, K. A. (1969), *Biochemistry 8*, 2129.
Igarashi, S., Kang, A. H., and Gross, J. (1970), *Biochem. Biophys. Res. Commun. 38*, 697.
Kang, A. H., Faris, B., and Franzblau, C. (1969c), *Biochem. Biophys. Res. Commun. 36*, 345.
Kang, A. H., Igarashi, S., and Gross, J. (1969a), *Biochemistry 8*, 3200.
Kang, A. H., Nagai, Y., Piez, K. A., and Gross, J. (1966), *Biochemistry 5*, 509.
Kang, A. H., Piez, K. A., and Gross, J. (1969b), *Biochemistry 8*, 1506.
Lane, J. M., and Miller, E. J. (1969), *Biochemistry 8*, 2134.
Michaeli, D., Benjamini, E., Leung, D. Y. K., and Martin, G. R. (1971), *Immunochemistry 8*, 1.
Michaeli, D., and Epstein, E. H., Jr. (1971), *Israel J. Med. Sci. 7*, 462.
Michaeli, D., Martin, G. R., Kettman, J., Benjamini, E., Leung, D. Y. K., and Blatt, B. A. (1969), *Science 166*, 1522.
Miller, E. J. (1971), *Biochem. Biophys. Res. Commun. 45*, 444.
Miller, E. J., Lane, J. M., and Piez, K. A. (1969), *Biochemistry 8*, 30.
Nikkari, T., and Heikkinen, E. (1968), *Acta Chem. Scand. 22*, 3047.
Piez, K. A. (1968), *Anal. Biochem. 26*, 305.
Piez, K. A., Bladen, H. A., Lane, J. M., Miller, E. J., Born-

stein, P., Butler, W. T., and Kang, A. H. (1968), *Brookhaven Symp. Biol. 21*, 345.
Piez, K. A., Eigner, E. A., and Lewis, M. S. (1963), *Biochemistry 2*, 58.
Piez, K. A., Miller, E. J., Lane, J. M., and Butler, W. T. (1969), *Biochem. Biophys. Res. Commun. 37*, 801.
Timpl, R., Wolff, I., Wick, G., Furthmayr, H., and Steffen, C. (1968), *J. Immunol. 101*, 725.
Volpin, D., and Veis, A. (1971a), *Biochem. Biophys. Res. Commun. 44*, 804.
Volpin, D., and Veis, A. (1971b), *Biochemistry 10*, 1751.
Vuust, J., Lane, J. M., Fietzek, P. P., Miller, E. J., and Piez, K. A. (1970), *Biochem. Biophys. Res. Commun. 38*, 703.
Vuust, J., and Piez, K. A. (1970), *J. Biol. Chem. 25*, 6201.

Spectrophotometric Measurement of Binding of S-Peptide Analogs to S-Protein†

Frances M. Finn

ABSTRACT: The binding to S-protein of S-peptide and eleven synthetic analogs was determined by a spectrophotometric method in the absence of substrate. Dissociation constants were in the range of 10^{-4}–10^{-5} M. Structural modifications of S-peptide which have a marked influence on S-protein activation or RNase S inhibition in the presence of substrate are not necessarily detectable by this technique. Modifications which disrupt the noncovalent interaction between Glu^2 and Arg^{10} in the S-peptide–S-protein complex and replacement of methionine by the sulfoxide are instances where a difference in the strength of the peptide–protein interaction is detectable spectrophotometrically.

The discovery of the S-peptide–S-protein system[1] (Richards, 1958) provided a unique opportunity to measure a number of parameters of peptide–protein interactions. The fortuitous location of part of the active site in the S-peptide portion of ribonuclease S permitted synthetic manipulations to be made leading eventually to the discovery of a number of potent inhibitors (Finn and Hofmann, 1967; Hofmann *et al.*, 1970, 1971). Furthermore a systematic replacement of residues along the peptide chain has allowed definition of amino acids contributing to the strong, highly specific, noncovalent binding between peptide and protein. Through this approach it has been possible to show that His^{12} is the only residue in the peptide essential for regeneration of enzymic activity in RNase S and that a number of hydrophilic (Glu^2,Arg^{10},Gln^{11},Asp^{14}) and hydrophobic (Phe^8,Met^{13}) residues are involved in binding. Such studies have led to the discovery that a peptide containing only the first 14 amino acid residues is equivalent to S-peptide as concerns activation of and binding to S-protein.

Initially the importance of amino acids in the S-peptide chain was assessed by comparing the ability of analogs to regenerate ribonuclease activity with S-protein using the potency of S-peptide as the 100% standard. However, the sharp break that occurs in the curve for activation of S-protein by S-peptide and a number of analogs indicated that the peptides bound quite firmly and that with such strong binding it seemed reasonable to expect that subtle differences in binding ability might not be discernible.

The finding that Pyr^{12}S-peptide$_{1-14}$ (Table I) is capable of competing as efficiently as S-peptide$_{1-20}$ for S-protein (50% inhibition at a ratio of 1:1, Pyr^{12}S-peptide$_{1-14}$:S-peptide$_{1-20}$) provided an alternate, more sensitive method for comparing peptide–protein interactions (Finn and Hofmann, 1967). Since both S-peptide and the various inhibitors were exposed simultaneously to the effects of substrate and since they were competing for the same binding sites, the ability to inhibit would be directly related to the relative binding efficiency and thus provide more information than activation studies. A number of such competitive inhibition studies were performed and the results of these will be discussed later.

Recently, we have become interested in determining directly the strength of peptide–protein interactions in general and the effect of amino acid substitutions on them. The S-peptide–S-protein system is obviously ideally suited to such an investigation owing to the availability of a large number of analogs. From the point of view of the techniques employed for the determination of binding constants, however, the system has several drawbacks. Equilibrium dialysis methods, as they are commonly used, are effective only when a large difference in molecular weight exists between the protein and the ligand. The combination of a ligand of molecular weight 2×10^3 and a protein of molecular weight 12×10^3 virtually precludes the use of conventional dialysis techniques. Even the Craig (Craig and Konigsberg, 1961) modification proved difficult insofar as quantitation was concerned. It should be noted that there are no tryptophan or tyrosine residues in S-peptide so that trials with dialysis techniques necessitated preparation of radioactively labeled peptides.

The generation of a difference spectrum upon addition of S-peptide to S-protein was first noted by Richards and Logue

† From the Protein Research Laboratory, University of Pittsburgh, School of Medicine, Pittsburgh, Pennsylvania 15213. *Received October 28, 1971.* Supported by Grant No. GB 13647 from the National Science Foundation.

[1] The following abbreviations will be used: RNase S, Subtilisin-modified beef pancreatic ribonuclease A; S-peptide, the peptide obtained from RNase S; S-protein, the protein component obtained from RNase S; RNase complex, a general term denoting the reconstituted enzyme obtained by mixing S-peptide or S-peptide analogs with S-protein; Pyr, β-(pyrazolyl-3)-alanine; 3-CM-His, 3-carboxymethylhistidine; RNA refers to yeast ribonucleic acid.