The Chromatographic Separation and Amino Acid Composition of the Subunits of Several Collagens*

KARL A. PIEZ, ELIZABETH A. EIGNER, AND MARC S. LEWIS

From the National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland

Received August 13, 1962

The subunit structure of collagens from the skin and tail tendon of the rat, the swim bladder of the carp, and the skin of the spiny dogfish has been studied by chromatography of the denatured proteins and amino acid analysis and ultracentrifugation of chromatographic fractions. It is demonstrated that the structure of these collagens can be represented by the formula $(\alpha 1)_2\alpha 2$ where $\alpha 1$ and $\alpha 2$ are subunits of similar size but different amino acid composition. These collagens also contain components formed by intramolecular crosslinking of $\alpha 1$ and $\alpha 2$. These are designated $\beta 1$ ($\alpha 1 - \alpha 2$) and $\beta 2$ ($\alpha 1 - \alpha 1$). The resulting structures $\alpha 1\beta 1$ and $\alpha 2\beta 2$ occur to varying degrees in different collagen preparations. Since in at least one collagen (rat skin) the amount of $\beta 1 + \beta 2$ is larger in samples which represent biologically older collagen, the structure $(\alpha 1)_2\alpha 2$ is, at least in part, in transition to the internally crosslinked structures. The amino acid compositions demonstrate a close relationship between $\alpha 1$ and $\alpha 2$. There are marked species and tissue differences, but a similar pattern is discernible. A recently identified amino acid, 3-hydroxyproline, is present in all the collagens.

It has been known for a number of years (Orekhovitch and Shpikiter, 1955) that the denaturation of soluble collagen yields fragments that fall largely into two weight classes. One of these, called the β-component, has about twice the molecular weight of the other, the α-component. A number of attempts have been made to relate these components stoichiometrically to each other and to the native collagen molecule (see Piez *et al.*, 1960). However, it has recently become clear that different collagen fractions from the same tissue may yield different proportions of the two weight classes. Thus, cold neutral sodium chloride solution extracts a collagen from calf or rat skin that consists largely of α-components when denatured, while acidic solutions extract in addition a biologically older collagen that contains a large proportion of β-components (Mazurov and Orekhovitch, 1960; Orekhovitch *et al.*, 1960; Piez *et al.*, 1961). That the two weight classes might have a monomer-dimer relationship has been considered by many investigators, but this possibility seemed to be inconsistent with the demonstration that the α- and β-components have different amino acid compositions (Piez *et al.*, 1960). However, this difficulty was overcome by the finding, presented in a preliminary report (Piez *et al.*, 1961), that rat-skin collagen consists of at least two kinds of α-components (designated $\alpha 1$ and $\alpha 2$) and two kinds of β-components (designated $\beta 1$ and $\beta 2$), which can be distinguished from one another by their chromatographic behavior and amino acid composition. $\beta 1$ consists of one $\alpha 1$ and one $\alpha 2$; $\beta 2$ consists of two $\alpha 1$. It was concluded that the primary subunits of the native collagen molecule are $\alpha 1$ (two per molecule) and $\alpha 2$ (one per molecule) and that $\beta 1$ and $\beta 2$ arise from them through the formation of intramolecular covalent bonds. Grassman *et al.* (1961) have made a similar suggestion on the basis of ultracentrifugal studies of denatured collagen and the isolation from collagen digests of peptides with two or three chains which had compositions inconsistent with identical α-components.

The present report is concerned with additional documentation for these conclusions and demonstrates the generality of the proposed structure by examining collagens from the skin and tail tendon of the rat, the swim bladder of the carp, and the skin of the spiny dogfish.

* Presented in part before the Division of Biological Chemistry at the 141st National Meeting of the American Chemical Society, Washington, D. C., March, 1962.

EXPERIMENTAL

Preparation of Collagen.—Rat skin and tail tendon were obtained from rapidly growing Sprague-Dawley rats weighing 50–100 g. All operations were conducted at 5°. Centrifugations were performed in a Servall refrigerated centrifuge with an SS-1 or SS-2 head. Neutral salt-extracted collagen was prepared by stirring the coarsely ground tissues with about ten volumes of 1 N NaCl for 48 hours. The suspension was filtered through several layers of cheesecloth and then through a Celite filter cake on a large Buchner funnel. The filtrate was acidified with acetic acid to *p*H 3–4. The precipitated collagen was removed by centrifugation at 13,000 × *g* for 15 minutes and dissolved in 0.5 M acetic acid by stirring for 24 hours. The solution was centrifuged at 40,000 × *g* for 2 hours and the collagen was precipitated by adding sodium chloride to 5% (w/v). The collagen was separated by centrifugation, dissolved in 0.5 M acetic acid, centrifuged at 40,000 × *g* for 2 hours, and dialyzed against 0.02 M disodium phosphate, the solution being changed once or twice until precipitation was complete. The precipitation by dialysis against disodium phosphate was repeated. The collagen was dissolved in 0.15 M acetic acid, dialyzed exhaustively against 0.15 M acetic acid, and lyophilized. The dry, salt-free collagen was stored over calcium chloride at 5°.

Acid-extracted rat skin collagen was prepared from the residue after salt extraction by stirring with about ten volumes of 0.5 M acetic acid for 48 hours. The collagen was purified as described above.

Acid-extracted collagens from rat tail tendon and from skin of the spiny dogfish were prepared in the same manner but without a preliminary neutral salt extraction. The spiny dogfish (*Squalus acanthias*) were obtained from the Mt. Desert Island Biological Laboratory, Mt. Desert Island, Maine.

Swim bladder collagen from the carp (ichthyocol), extracted with a citrate buffer and purified by a similar procedure, was kindly supplied by Dr. Paul Gallop (see Gallop, 1955).

Chromatography of Collagen.—The CM-cellulose used in all the experiments described here, unless otherwise indicated, was from a single batch prepared in the laboratory as described by Peterson and Sober (1956). A

jacketed column, 2.2 cm diameter, was packed with CM-cellulose (Piez *et al.*, 1960) in starting buffer to a height of 18 cm. A circle of filter paper was placed on top of the column. Water at 40° was circulated through the jacket. Polyvinyl tubing, $^1/_{32}$ in. i.d., connected to the bottom of the column, led to a flow cell (1 cm light path and 0.5 ml volume) in the sample side of a Beckman DB recording spectrophotometer. A tee, constructed of $^1/_{16}$ in. i.d. glass tubing and closed at the top with a short piece of plastic tubing and a clamp, was placed in this line to trap bubbles. Starting buffer was placed in another cell in the reference side of the spectrophotometer. The outlet of the flow cell led to a fraction collector. Eluting buffer was delivered to the top of the column through Tygon tubing, $^1/_{16}$ in. i.d., $^3/_{16}$ in. o.d., in a Sigmamotor finger pump. The tubing entered the top of the column through a ball and socket joint which was clamped to allow pressure to be maintained. The flow rate was 120 ml per hour. A Sargent SRL chart recorder with logarithmic gears and a 9.5 in. chart (0–1.0 absorbancy) was employed to record absorbancy continuously during each run. The chart speed was 3 in. per hour. Absorbancy was measured at 230 mμ.

Stock buffer, *p*H 4.8 and ionic strength 0.5, was prepared from 68 g sodium acetate·3 H_2O and 19 ml acetic acid per liter of solution. This was diluted with water to an ionic strength of 0.06 for starting buffer. Limiting buffer was prepared in the same manner but with the addition of sodium chloride to give a concentration of 0.1 N (a total ionic strength of 0.16). These conditions were used for chromatography of collagens from rat tail tendon, rat skin, and carp swim bladder. For chromatography of skin collagen from the spiny dogfish, stock buffer was diluted to an ionic strength of 0.08 for starting buffer; limiting buffer also contained sodium chloride at a concentration of 0.12 M (a total ionic strength of 0.20). In all experiments, a linear gradient was obtained by eluting the column from a two-chamber, constant-level gradient device, the mixing chamber containing 400 ml of starting buffer and the reservior an equal weight of limiting buffer.

The sample, usually 25–75 mg, was dissolved in 10–20 ml of starting buffer by stirring overnight at 5°. It was denatured just before chromatography by heating to 45° for 30 minutes. If the solution was cloudy, it was filtered while warm through a fine porosity sintered glass disk. Before each experiment, starting buffer was pumped through the column and the flow cell until a stable baseline was observed on the recorder. The tee was opened to release air trapped during the previous run. The height of the solution above the CM-cellulose was reduced, if necessary, to about 2 mm by allowing the pump to deliver air for a few minutes. Without disturbing the column or stopping the pump or recorder, the inlet line to the pump was then placed in a small flask containing the sample. When the sample had been pumped onto the column, the line was moved immediately to the gradient device which had been previously prepared, the gradient was started, and the collection of fractions was begun. When the run was complete the inlet line to the pump was moved to a flask containing a solution 0.5 N in sodium chloride and 0.01 N in sodium hydroxide for about 30 minutes to remove any protein remaining on the column. Starting buffer was then pumped through the column until the baseline became stable. The column could usually be used without repacking for several runs. When air bubbles became objectionable, the column was repacked.

For rechromatography, the appropriate fractions were combined, the total ionic strength of buffer and salt was calculated from the position of the fractions in the gradient, water was added to give an ionic strength less than 0.06, and the solution was pumped onto the column and the gradient started as described above. If the volume was large, the flow rate was increased to about 300 ml per hour while the sample was being applied. In this manner, a sample in a volume of 500 ml or more could be chromatographed.

Protein was isolated from effluent fractions by gel-filtration (Porath and Flodin, 1959) at room temperature through a column of Sephadex G-25, 3.5 cm in diameter and 45 cm in height (450 ml), equilibrated with *p*H 4.4 pyridine acetate buffer (4.0 ml pyridine and 6.7 ml acetic acid per liter of solution). This size column was sufficient for volumes of protein solution up to at least 150 ml. Control experiments were performed to establish the points of emergence of protein and compounds of low molecular weight. The portion of effluent containing the protein was collected and lyophilized to give dry, salt-free material.

Absorbancy indices were measured on weighed samples of chromatographic fractions from rat-skin collagen and on the original collagen dissolved in starting buffer. All had identical indices (±2%). The values varied with the spectrophotometer employed and the instrument settings selected, since absorbancy was measured on the side of the region of "end-absorption." The absorbancy index was about 20 at 230 mμ, 10 at 234 mμ, and 5 at 240 mμ. It was assumed that absorbancy was proportional to concentration for all the collagens studied.

The chromatograms presented here were redrawn for clarity.

Amino Acid Analysis.—The purified collagens and fractions isolated from the chromatographic columns were hydrolyzed in 6 N HCl for 24 hours at 106° and analyzed on an automatic amino acid analyzer (Piez and Morris, 1960) as previously described (Piez *et al.*, 1960).

Ultracentrifugation.—Denatured collagens and collagen fractions were examined by sedimentation velocity at 40° employing Schlieren optics in a Spinco Model E ultracentrifuge. Native collagens were dissolved by stirring overnight in sodium formate buffer, *p*H 3.75 and ionic strength 0.15, and denatured by heating to 40° just before being placed in the cells (12 mm Kel-F, standard and wedge window). Collagen fractions were dissolved in the same buffer by stirring for a few minutes at 40° just before being placed in the cells.

All patterns reproduced here were taken 128 minutes after reaching a speed of 59,780 rpm with the phase plate angle at 65°. As illustrated in the figures, sedimentation is from left to right.

Results

3-Hydroxyproline.—A small but easily measurable peak of ninhydrin color absorbing maximally at 440 mμ and appearing in the effluent of the amino acid analyzer just before the methionine sulfoxides was noted in hydrolysates of rat tail collagen (Fig. 1). An apparently identical peak with characteristics of a cyclic imino acid was seen previously in hydrolysates of collagen from the sponge (Piez and Gross, 1959). The unknown compound from the sponge collagen has now been isolated and has been identified by comparison with the synthetic compound by Irreverre *et al.* (1962) as 3-hydroxyproline. When the known 3-hydroxyproline was added to a hydrolysate of rat tail collagen, the peak was unaltered except for an increase in height, indicating identity. Ogle *et al.* (1961) have

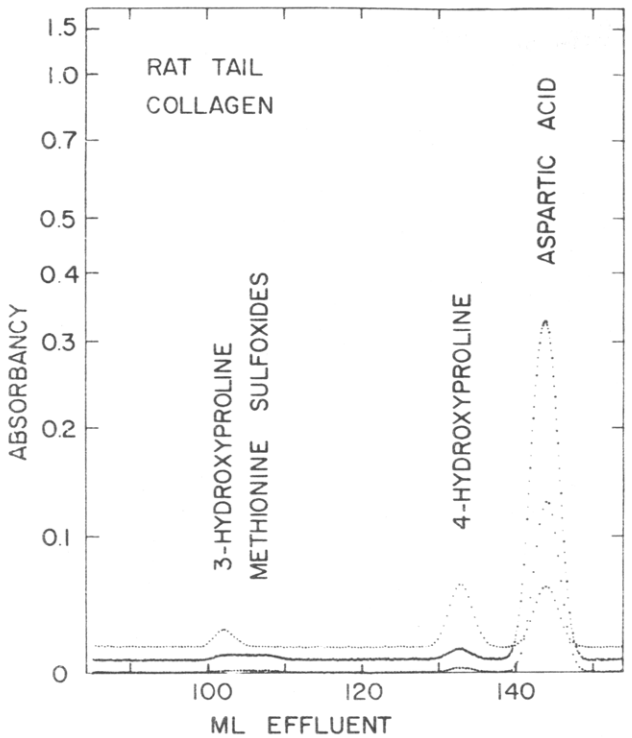

FIG. 1.—A portion of a chromatogram obtained with an automatic amino acid analyzer showing the presence of a new amino acid, 3-hydroxyproline, in a hydrolysate of rat tail collagen. The curves show absorbance by the color produced by reaction with ninhydrin at 440 mμ (upper curve in blank regions) and 570 mμ (3 mm cell, middle curve; 1 mm cell, lower curve).

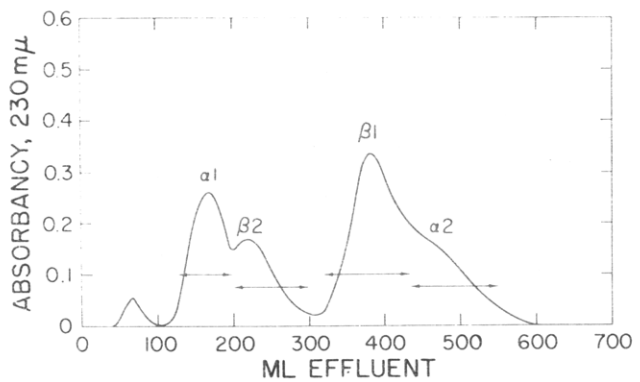

FIG. 2.—Elution pattern of approximately 40 mg of denatured rat tail collagen chromatographed on CM-cellulose at 40°. Fractions were taken as indicated by the arrows for rechromatography (Fig. 3).

also reported an unknown compound in collagen from bovine Achilles tendon that they suggested is related to the cyclic imino acids. The other collagens examined in the present study contain the same compound, which appears in each case to be identical to known 3-hydroxyproline. In the amino acid analyzer used for these studies, 3-hydroxyproline from sponge collagen has a molar absorbancy index at 440 mμ, 1.40 times the value observed for proline.

Rat Tail Collagen.—The manner in which the chromatographic peaks were first characterized is illustrated with rat tail collagen. A typical chromatogram of denatured rat tail collagen is shown in Figure 2. The four major peaks were isolated from the fractions indicated in the figure and rechromatographed separately. The results appear in a composite chromatogram in Figure 3. Since the positions of the peaks in the effluent are very sensitive to the chromatographic

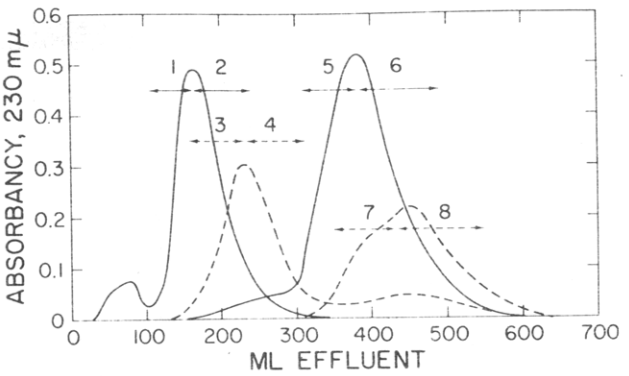

FIG. 3.—A composite chromatogram showing rechromatography of the four major components pooled from several chromatograms of denatured rat tail collagen (Fig. 2). The peaks were cut as indicated by the arrows for examination in the ultracentrifuge (Fig. 4) and amino acid analysis (Table I).

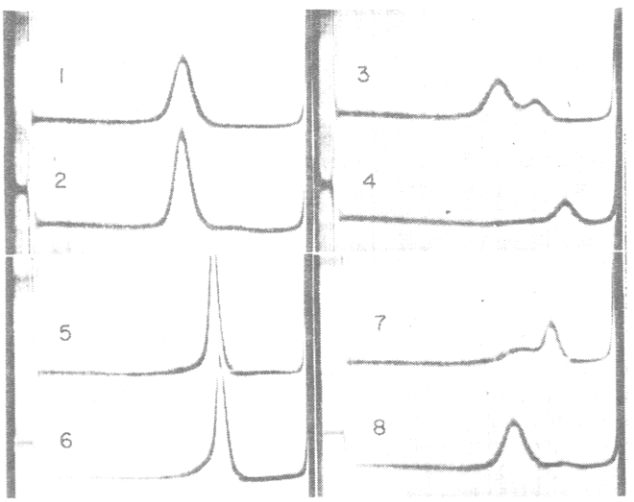

FIG. 4.—Sedimentation velocity patterns of denatured collagen fractions. The numbers refer to the fractions isolated by chromatography (see Fig. 3).

conditions, separate chromatograms may not be directly comparable. For example, fractions 7 and 8 are probably shown (Fig. 3) as appearing too early in the effluent relative to fractions 5 and 6. However, it can be seen that the major part of each fraction chromatographed very close to its original position. Each peak after rechromatography was divided in half as indicated in Figure 3 to give eight fractions. The protein from each of these fractions was isolated, examined by sedimentation velocity, and analyzed for total amino acid content. The sedimentation patterns (Fig. 4) show that fraction 1 was all α; 2 was mostly α with a trace of β; 3 contained both components; 4 was largely β; 5 and 6 were also β with some α appearing in the latter; 7 contained both; and 8 was almost entirely α. In interpreting these diagrams it must be remembered that there is a large Johnston-Ogston effect (Johnston and Ogston, 1946; see also Grassman *et al.*, 1961) which tends to exaggerate the apparent size of the more slowly moving peak at the expense of the faster moving peak. The pattern obtained from fraction 3 (Fig. 4), for example, is typical of that seen with approximately 1:1 mixtures of α and β. There is also a marked dependence of sedimentation rate on concentration. It is clear from these results that both α and β appear in two different places on the chromatograms. These have been designated $\alpha 1$, $\beta 2$, $\beta 1$, and $\alpha 2$, in the order of their appearance in the effluent (Fig. 2). The results of the amino acid

TABLE I

AMINO ACID COMPOSITION OF SUCCESSIVE FRACTIONS FROM CHROMATOGRAMS OF RAT TAIL COLLAGEN

Residues/1000 Total Residues

	Original	Fraction No.[a]							
		1	2	3	4	5	6	7	8
Glycine	331	327	331	331	334	336	336	336	335
Valine	23	19.0	19.4	19.8	19.8	25.5	25.9	27.4	30.4
Isoleucine	9.6	6.3	6.3	6.4	6.3	11.1	12.0	12.6	15.1
Leucine	24	18.0	18.1	18.1	18.8	25.2	26.2	26.7	30.9
Histidine	4.1	1.4	1.5	1.6	1.4	4.8	4.9	6.3	7.4

[a] Single analyses of the chromatographic fractions indicated in Figure 3 are shown.

analyses of the eight fractions are shown in part in Table I. Fractions 1–4 were identical; 5 and 6 showed marked changes; 8 was again very different; and 7 was intermediate between 6 and 8. Comparing these data with the ultracentrifuge patterns and the original chromatogram, it can be seen that $\alpha 1$ and $\alpha 2$ have quite different, though obviously related, amino acid compositions. $\beta 2$ is identical to $\alpha 1$, while $\beta 1$ has a composition equivalent to a 1:1 mixture of $\alpha 1$ and $\alpha 2$. The original unfractionated collagen has a composition equivalent to a mixture containing two $\alpha 1$ and one $\alpha 2$ subunits. The remainder of the amino acid data not shown in Table I was also consistent with these conclusions. The complete analysis assembled by subunit appears in Table II. The analyses of fractions 1 and 2 were averaged to obtain $\alpha 1$; 4 was taken as $\beta 2$; 5 and 6 were averaged to obtain $\beta 1$; and 8 was taken as $\alpha 2$.

The approximate distribution of the subunits estimated from the areas under the chromatographic peaks is shown in Table III.

Rat Skin Collagen.—Preliminary results have been reported for rat skin collagen (Piez *et al.*, 1961). A chromatogram of acid-extracted collagen is shown in Figure 5. Ultracentrifuge patterns of the four subunits isolated by chromatography and rechromatography and patterns of the original collagen at two concentrations are shown in Figure 6. It can be seen that the $\alpha 1$ and $\alpha 2$ preparations contained no visible traces of heavier components. $\beta 1$ showed some indication of α (presumably $\alpha 2$), but probably not enough to affect its analysis. The $\beta 2$ fraction was at best probably only half $\beta 2$, the rest representing $\alpha 1$. Some samples of $\beta 2$ were obtained which were less contaminated, but this required that only the trailing one half to one third of the peak be taken. The amino acid compositions of the subunits are shown in Table II. Although the $\beta 2$ sample contained about one half $\alpha 1$, the analysis is not affected since these two subunits have the same composition.

Fractions corresponding to $\alpha 1$, $\beta 1$, and $\alpha 2$ were also isolated from chromatograms of salt-extracted rat skin collagen (see Piez *et al.*, 1961). These could not be distinguished by chromatographic behavior or amino acid analysis from the corresponding subunits from acid-extracted collagen.

The ratio of $\alpha 1$ to $\alpha 2$ was determined by measuring the areas under the peaks. The values ranged from 1.9 to 2.2 and averaged 2.0 in chromatograms of five different samples, including both acid- and salt-extracted collagens. The ratio of $\beta 1$ to $\beta 2$ was more difficult to measure; values between 2 and 4 were obtained. The estimated proportions of the subunits are shown in Table III.

3-Hydroxyproline is present in rat skin collagen (but not in the $\alpha 2$ subunit) in very low amounts. The values given in Table II represent approximations since only a barely visible increase in absorbancy above baseline occurred on the chromatograms.

Carp Swim Bladder Collagen.—Collagen from the carp is a typical fish collagen with the same general structure and properties as mammalian collagens. A typical chromatogram obtained under the same conditions as were used for chromatography of the two rat collagens is shown in Figure 7. Four major peaks were seen in approximately the same positions as observed with rat skin collagen. Rechromatography, isolation of purified fractions, examination in the ultracentrifuge (Fig. 8), and amino acid analysis (Table II) showed that the four fractions bear the same relationship to each other and to native carp collagen as in the case of the corresponding subunits from the rat collagens.

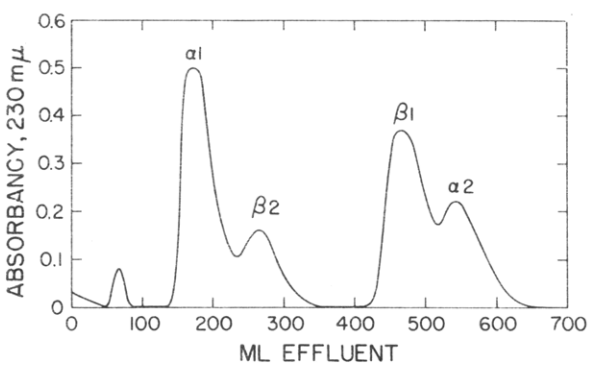

FIG. 5.—Elution pattern of approximately 50 mg of denatured rat skin collagen chromatographed on CM-cellulose at 40°.

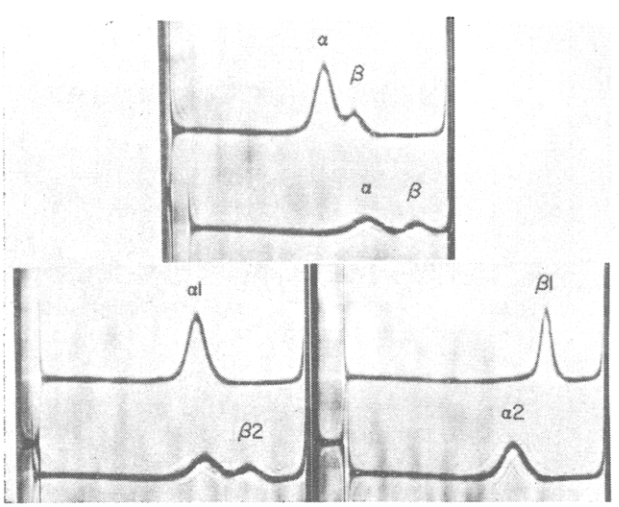

FIG. 6.—Sedimentation velocity patterns of denatured acid-extracted collagen from rat skin. Two different concentrations of unfractionated collagen (upper two patterns, 0.5 and 0.2%, approximately) and the subunits (lower patterns) isolated from the same sample of collagen by chromatography (Fig. 5) are shown.

TABLE II - AMINO ACID COMPOSITION OF SEVERAL COLLAGENS AND THEIR SUBUNITS[a]
Residues 1000 Total Residues

	Rat Skin				Rat Tail Tendon				Carp Swim Bladder				Spiny Dogfish Skin		
	Orig.	α1	β2	β1	α2	Orig.	α1	β2	β1	α2	Orig.	α1	β2	β1	α2	Orig.	α1	β1
3-Hydroxyproline	0.9	1.0	1.5	0.7	0.0	4.2	4.1	3.9	3.8	3.8	0.6	0.8	0.4	0.4	0.0	1.4	1.7	1.3
4-Hydroxyproline	92	96	97	88	86	90	93	91	87	82	76	79	75	72	70	59	65	57
Aspartic acid	46	46	45	44	44	45	47	47	46	44	47	42	43	48	53	43	44	42
Threonine	19.6	19.9	19.8	20.1	19.8	19.9	19.7	20.1	19.8	19.8	27.6	29.3	29.0	27.7	25.0	23.4	22.6	24.4
Serine	43	42	41	42	43	43	41	42	43	43	37	35	35	38	40	61	52	64
Glutamic acid	71	74	72	70	66	71	74	74	70	68	71	75	75	71	66	68	68	67
Proline	121	129	128	120	113	122	129	126	119	115	116	117	118	117	117	106	115	103
Glycine	331	330	333	338	336	331	329	334	336	335	333	335	335	336	337	338	337	339
Alanine	106	112	111	107	102	107	110	111	106	103	125	131	132	123	114	106	115	103
Valine	24.0	19.6	19.2	25.6	32.0	22.9	19.2	19.8	25.7	30.4	17.6	13.6	14.2	19.2	24.0	25.3	23.5	26.7
Methionine	7.8	8.0	6.6	7.0	6.1	8.4	9.0	9.1	7.5	6.7	16.0	17.8	17.2	14.0	11.6	18.0	17.4	16.8
Isoleucine	10.8	6.4	6.5	11.6	16.1	9.6	6.3	6.3	11.6	15.1	9.8	9.8	10.2	9.2	8.4	15.3	13.9	16.2
Leucine	23.8	18.1	18.7	25.7	32.4	23.6	18.1	18.8	25.7	30.9	20.1	14.6	15.2	22.0	28.5	24.8	19.3	27.2
Tyrosine	2.4	2.1	1.8	2.2	2.4	3.9	3.7	3.3	3.3	3.9	3.4	0.9	1.4	4.0	6.2	3.3	1.9	2.7
Phenylalanine	11.3	11.6	12.1	11.0	10.1	11.9	11.9	11.3	11.2	11.3	13.4	14.2	14.0	12.2	10.0	13.2	12.1	12.1
Hydroxylysine	5.7	4.3	3.9	6.2	8.0	6.6	4.8	4.1	7.4	9.8	7.0	5.6	5.5	7.0	8.8	5.5	5.3	6.1
Lysine	28.1	30.4	30.1	25.8	22.4	26.9	30.1	28.7	24.3	21.0	25.6	28.8	28.2	24.0	19.8	26.6	28.5	25.0
Histidine	4.9	1.9	2.0	5.3	8.5	4.1	1.5	1.4	4.9	7.4	3.6	1.7	2.0	4.0	5.8	12.5	8.3	13.7
Arginine	51	49	50	50	51	50	49	48	49	51	52	50	51	52	54	51	50	53
Amide N	(41)	(42)	(41)	(42)	(43)	(40)	(39)	(40)	(41)	(41)	(40)	(37)	(39)	(43)	(46)	(36)	(32)	(39)

[a] Averages of duplicate analyses are presented, except for rat tail tendon β2 and α2 which are single analyses.

TABLE III
THE APPROXIMATE DISTRIBUTION OF SUBUNITS IN SEVERAL DENATURED COLLAGENS AND THE CALCULATED DISTRIBUTION OF COLLAGEN STRUCTURES IN THE NATIVE SAMPLES

	Collagen Subunits[a]				Collagen Structures[b]		
	α1	α2	β1	β2	$(\alpha 1)_2\alpha 2$	α1β1	α2β2
Rat skin							
Salt-extracted	55	27	13	5	72	20	8
Acid-extracted	35	15	37	13	24	56	20
Rat tail tendon	30	10	45	15	9	68	23
Carp swim bladder	48	23	20	7	59	30	11
Dogfish skin	31	3	61	5	0	92	8

[a] Calculated from the areas under the appropriate peaks on the chromatograms. [b] Calculated from the subunit distribution. The few percent of γ-component which may be present is not included.

The amount of β1 + β2 in the unfractionated carp swim bladder collagen measured from the chromatograms was about 30% of the total weight compared to about 50% for acid-extracted rat skin collagen and 60% for tail tendon collagen (Table III). The ratio of α1 to α2 appeared to be approximately 2, although resolution was not good enough for a precise measurement. The ratio of β1 to β2 was also about the same as observed for the rat collagens.

Spiny Dogfish Skin Collagen.—The spiny dogfish is a member of the class Elasmobrachii. Although vertebrates, these animals are not true fish. They resemble and are closely related to the sharks. The skin collagen is similar to certain collagens from fish which live in cold water in that it has a very low denaturation temperature (about 16° in solution) and a low content of cyclic imino acids (Lewis and Piez, in preparation).

When a sample of this collagen was chromatographed under the conditions used for the rat and carp collagens, protein peaks were very late in appearing in the effluent. Therefore, a slightly higher ionic strength and steeper gradient were employed. Under these conditions somewhat similar chromatograms were obtained, although only two major peaks were evident (Fig. 9). The ultracentrifuge patterns reproduced in Figure 10 showed the first of these was α (α1 in Fig. 9) and the second was β (β1 in Fig. 9). Material taken from the trailing edge of the β1 peak (in the region labeled α2, Fig. 9) contained some α. (The component sedimenting ahead of β1 is undoubtedly the γ-component which is discussed below.) The shoulder following α1 on the chromatograms was not studied but presumably represents β2. Amino acid analysis (Table II) of protein from the peaks labeled α1 and β1 indicated that these identifications are correct. The amino acid composition of the original collagen is consistent with a composition equivalent to a weight ratio of α1 to β1 of 1:2. The chromatograms also support this structure; the ratio of α1 to β1 determined from the areas under the peaks was 0.50. The amounts of α2 and β2 are very uncertain, but they probably totaled less than 10% of the sample. The estimated proportions appear in Table III. It is assumed that no uncrosslinked molecules were present in the sample although a few per cent could not be detected by the distribution of subunits.

It seems from these observations that dogfish collagen has the same basic structure as the rat and carp collagens. Since this conclusion rests in part on the apparent analogy to the rat and carp collagens, it will be desirable to isolate and analyze β2 and α2.

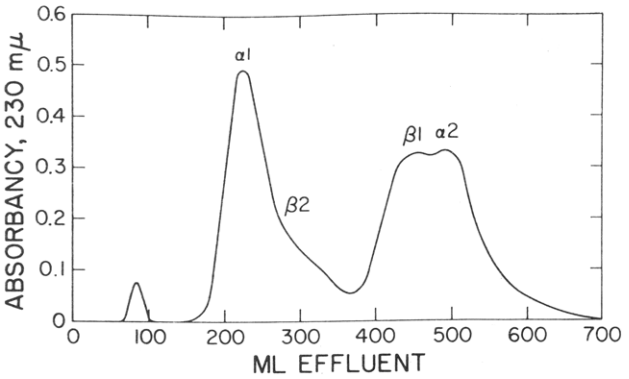

FIG. 7.—Elution pattern of approximately 50 mg of denatured collagen from carp swim bladder chromatographed on CM-cellulose at 40°.

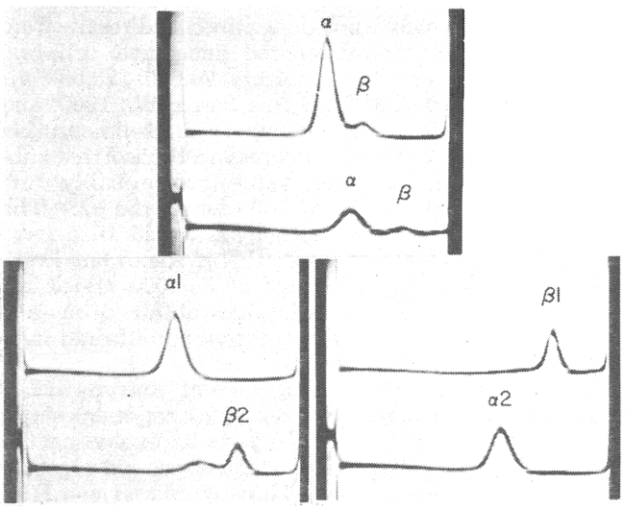

FIG. 8.—Sedimentation velocity patterns of denatured carp swim bladder collagen. Two different concentrations of unfractionated collagen (upper two patterns, 0.6 and 0.3%, approximately) and the subunits (lower patterns) isolated from the same sample of collagen by chromatography (Fig. 7) are shown.

The finding that higher ionic strength buffers are required to chromatograph denatured dogfish collagen than are required for the other collagens is consistent with the greater basicity which can be calculated from the amino acid data. The excess of basic groups over acidic groups in the unfractionated collagens is about 71 per 1000 residues for dogfish collagen and about 63 per 1000 residues for the other three. The difference is largely the result of a higher content of histidine in dogfish collagen.

3-Hydroxyproline was again found in small amounts.

General Observations.—In all the chromatograms a forepeak appeared which was unretarded by the CM-cellulose. Its size varied but it usually represented less than 2% and sometimes less than 1% of the total absorbancy. It was not present if the sample was dialyzed after denaturation. Chromatograms of samples which were judged to be impure by amino acid analysis, stored in denatured form prior to chromatography, or otherwise subjected to conditions which might bring about degradation often showed particularly large forepeaks. It is therefore likely that the forepeak represents a small amount of impurity or degraded material.

The possibility that a small but critical part of the molecule is released by denaturation and is perhaps present in the forepeak was examined by careful amino

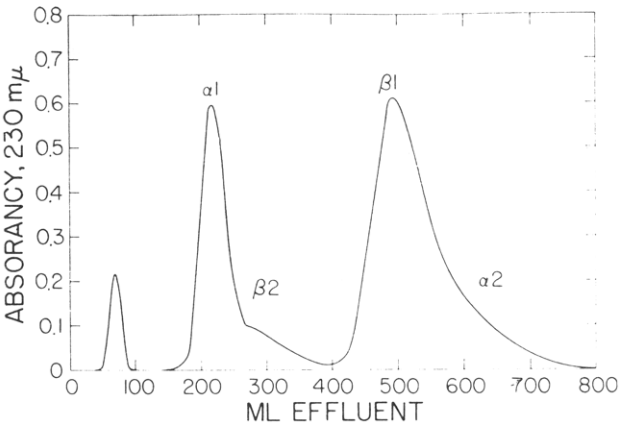

FIG. 9.—Elution pattern of approximately 70 mg of denatured collagen from the skin of the spiny dogfish chromatographed on CM-cellulose at 40°.

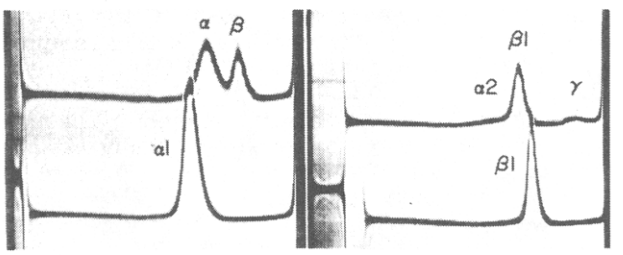

FIG. 10.—Sedimentation velocity patterns of denatured collagen from the spiny dogfish (unfractionated, upper left) and the $\alpha 1$ (lower left) and $\beta 1$ (lower right) subunits isolated by chromatography. The upper right pattern shows protein from the region labeled $\alpha 2$ (Fig. 9).

acid analysis of several samples of collagen before and after denaturation. The denatured samples were prepared in the same manner as the samples used for chromatography. They were desalted by gel-filtration through Sephadex G-25 and lyophilized. With some samples, identical analyses were obtained before and after denaturation. Other samples of undenatured collagen had slightly low glycine contents and high tyrosine contents indicating the presence of an impurity (Kühn *et al.*, 1961). In these cases, the denaturation and gel-filtration apparently constituted a purification. It is also apparent from the amino acid data in Table II that for all four collagens and each of the amino acids, the composition of the original collagen is completely accounted for by the subunit structure within the analytical error of the analysis.

The studies reported here are complicated by the fact that denatured collagens under the conditions employed for chromatography (40°, *p*H 4.8) undergo slow changes which are probably of a degradative type. This was indicated by a broadening of the chromatographic peaks and the appearance of new peaks in samples which had been heated for too long or at too high a temperature. Broadening also appeared to varying degrees on rechromatography and became pronounced if chromatography was attempted a third time. Even collagen which was denatured as described and then stored overnight in the refrigerator as a gel sometimes showed marked changes. The subunits when desalted and lyophilized appear to be reasonably stable for at least several months when stored at 5° over calcium chloride.

Native collagens usually seem to be stable nearly indefinitely when stored dry at 5°. Most of the samples used in these studies consistently gave clear to

slightly cloudy solutions when dissolved prior to chromatography. However a few became partially insoluble after several months of storage. The solubility appeared to be related to the recovery after chromatography. The preparations which gave clear solutions had recoveries near 100%, whereas samples which did not dissolve readily were recovered to an extent of less than 50% in some instances. Part of the unrecovered protein was lost on filtration before chromatography, but a large amount was not eluted from the column and was removed only by the dilute alkaline wash. Absorbancy can be followed during the wash to permit a semiquantitative measure of this portion.

Although the evidence is good that there are only four major components in the denatured collagens examined here by the criteria of sedimentation velocity and amino acid composition, the chromatograms indicate some type of heterogeneity within each major peak which cannot be easily explained. For example, when the overlapping peaks of $\alpha 1$ and $\beta 2$ were divided at the point of minimal absorbancy and rechromatographed, both peaks appeared homogeneous (Fig. 3) even though ultracentrifuge analysis showed an appreciable amount of $\alpha 1$ in the $\beta 2$ peak (Fig. 4). That is, there appears to be some part of $\alpha 1$ which chromatographs with $\beta 2$ even on rechromatography. The same relationship holds between $\beta 1$ and $\alpha 2$. In many instances it also happened that some of the peaks, particularly $\alpha 1$, showed pronounced asymmetry and even a doubling. However, this finding was not consistent. One collagen preparation would show one or more double or skewed peaks while another would not. Because of the instability of denatured collagen, it is possible that this apparent inhomogeneity results from chemical changes during preparation of the sample. It is also possible that denaturation of collagen leaves some nonrandom structure which is not necessarily the same for all the molecules in the sample. However, the possibility that $\alpha 1$ (and therefore $\beta 1$ also) represents two very closely related components cannot be completely ruled out by the present studies.

Protein isolated from the tail region of the $\alpha 2$ peak from the rat or dogfish collagens contained a minor component which sedimented faster in the ultracentrifuge than $\beta 1$ and $\beta 2$ (see Fig. 10, upper right pattern). This is undoubtedly the γ-component, which can be renatured by cooling after heat denaturation to yield a protein with many of the properties of native collagen. It probably has a structure containing all three of the primary subunits maintained in proper relationship to each other by covalent crosslinks which permit recoiling (Veis and Cohen, 1960; Rice, 1960; Grassman *et al.*, 1961; Altgelt *et al.*, 1961; Veis and Anesey, 1961).

In addition to the sample of CM-cellulose which was prepared in the laboratory and used for the studies described above, two commercial preparations of CM-cellulose have been tested Both of these gave chromatograms which were very similar to the ones reproduced here, but with one (Schleicher and Schuell Selectacel CM, Type 40) it was necessary to use higher ionic strength buffers. The other (Bio-Rad, Cellex-CM) gave poorer resolution between $\alpha 1$ and $\beta 2$ but improved resolution between $\beta 1$ and $\alpha 2$. It is not clear what properties of CM-cellulose are responsible for these variations.

Discussion

It can be concluded that most and perhaps all vertebrate collagens are initially assembled in a structure which can be represented by the formula $(\alpha 1)_2\alpha 2$. (It must be remembered, however, that it has not been demonstrated unequivocally that the two $\alpha 1$ subunits are identical.) As a function of time, intramolecular crosslinking produces the structures $\alpha 1\beta 1$ and $\alpha 2\beta 2$, in which $\beta 1$ is $\alpha 1$-$\alpha 2$ and $\beta 2$ is $\alpha 1$-$\alpha 1$. The rate at which these form and the extent to which they occur vary with the animal and the tissue over a rather wide range. That $\beta 1$ and $\beta 2$ derive largely, if not exclusively, from intra rather than intermolecular crosslinks can be concluded from the fact that collagen preparations can be made which are largely monomeric but still yield a large proportion of $\beta 1$ and $\beta 2$. Perhaps the best example of this is dogfish collagen (Lewis and Piez, in preparation). The same conclusion can be reached from other studies of molecular weight of native collagen and heterogeneity of denatured collagen (Boedtker and Doty, 1956).

As further confirmation of the proposed subunit structure, molecular weight data would obviously be useful. Several values for α and β have appeared in the literature which were determined indirectly from measurements on unfractionated denatured collagen. The weights reported range from 70,000–125,000 for α and 160,000–290,000 for β (see Piez *et al.*, 1960, and Grassman *et al.*, 1961, for a summary). Sedimentation equilibrium studies now in progress on the isolated subunits indicate that the best values are probably near 100,000 for $\alpha 1$ and $\alpha 2$ and 200,000 for $\beta 1$ and $\beta 2$. The molecular weight of native collagen would then have to be about 300,000. Though this is somewhat lower than the usually quoted value of 345,000 (Boedtker and Doty, 1956), preliminary values obtained on salt-extracted collagens by sedimentation equilibrium support the lower molecular weight.

The nature of the subunits is not known. It is reasonable to speculate that $\alpha 1$ and $\alpha 2$ represent single chains each in a polyproline-type helix extending the full length of the collagen molecule and wound together in a larger triple-helix. (See Harrington and von Hippel, 1961, for a discussion of collagen structure.) A simple calculation shows that this is feasible. There are three residues every 2.86 A along the axis of the molecule (Rich and Crick, 1962). With three chains of similar weight and length, a molecule of molecular weight 300,000 would contain about 3300 amino acid residues (the average residue weight is 91) and extend about 3100 A, a value only slightly larger than the best value obtained from physical measurements and electron microscopy (Hall and Doty, 1958). In this proposed structure, $\beta 1$ and $\beta 2$ would be chain pairs formed by intramolecular crosslinking between groups already in proper orientation and with little or no disruption of the original structure.

Even if $\alpha 1$ and $\alpha 2$ are single continuous chains, it is not certain that they contain only α-peptide links. Blumenfeld and Gallop (1962) have reported the presence of labile bonds which are probably α- and β-aspartyl esters; Franzblau *et al.* (1962) have found γ-glutamyl peptide bonds; and Mechanic and Levy (1959) have isolated a peptide containing a peptide bond to the ϵ-amino group of lysine. Any of these bonds could occur in the backbone chain of $\alpha 1$ and $\alpha 2$ with probably a local modification of the polyproline type helix.

These bonds are also possibilities for the crosslink or crosslinks which form $\beta 1$ and $\beta 2$. No evidence is yet available concerning the nature of the bond other than that it is probably covalent since it will withstand the denaturing action of 5 M guanidine and 2 M potassium thiocyanate. Disulfide bridges are not involved, since collagen contains no cysteine or cystine.

Since bonds can be formed between any two of the three subunits, it is not unexpected that there would

be a small proportion of material in which all three subunits are joined. The γ-component found in the chromatograms and identified in the ultracentrifuge has the necessary characteristics of this species. It should also be remembered that with many tissues a large part of the collagen cannot be extracted without degradation, indicating the presence of higher aggregates (see Veis and Anesey, 1961). Therefore, the $\alpha \rightarrow \beta$ transformation may be only the first step in a more extensive crosslinking process.

It is difficult to speculate on the function of the intramolecular crosslinks, but it would seem to be logical that they would contribute to the stability of the collagen structure to a small but perhaps critical degree. Evidence that this is so is found in studies of the toxic condition known as lathyrism. Animals which have been given certain chemical agents such as β-aminopropionitrile show a connective tissue disorder in which the gross manifestations are preceded by an interference with the crosslinking mechanism. Very little β1 and β2 is formed in the collagen of these animals. It has been suggested that the stability of the collagen is affected, resulting in the gross symptoms of lathyrism (Martin *et al.*, 1961; Martin *et al.*, in press).

The collagens studied here show characteristic differences in the extent to which β1 and β2 are formed. Collagen from rat skin obtained by salt extraction contains largely α1 and α2 (Orekhovitch *et al.*, 1960; Piez *et al.*, 1961). The salt-extractable fraction is a relatively small part of the total which has been shown by labeling studies to be in transition to a less easily extractable and, it is now evident, more highly intramolecularly crosslinked material (see Gross, 1959; Jackson and Bentley, 1960). The bulk of the collagen, represented by acid-extractable collagen, contains about 50% β1 + β2. The collagen extracted by acid from rat tail tendon contains somewhat larger amounts of β1 and β2, about 60%, and therefore has less collagen containing only α1 and α2. This may be the result of a more rapid transition from uncrosslinked to crosslinked collagen. Collagen from carp swim bladder represents a type which is easily extractable by acidic solutions but is crosslinked to a much smaller extent. In contrast to this observation, collagen from the spiny dogfish consists entirely, or nearly so, of crosslinked components. A very large part of the collagen in the skin of the dogfish can be dissolved in 0.5 M acetic acid and approximate measurements indicate that more than 90% has the structure α1β1. It can be concluded that, in this collagen, α2 carries the function that is involved in crosslinking in a highly reactive state. α1 apparently cannot easily crosslink with itself. A similar but not as pronounced preference is also found in the other collagens. If α1 and α2 could crosslink to form β1 and β2 with equal ease, the expected ratio of β1 to β2 would be 2:1. However, the average measured value for rat skin collagen is approximately 3:1. Fewer measurements have been made on collagens from rat tail tendon and carp swim bladder, but the same result is indicated.

The relative amounts of the three main structures [$(\alpha 1)_2\alpha 2$, α1β1, and α2β2] present in the native collagen samples have been calculated from the distribution of the subunits as measured from the chromatograms of denatured preparations. The results appear in Table III. These values are only approximate but demonstrate the wide range of distributions which occur in extractable collagens. The small amount of γ-component present was not calculated. Grassman *et al.* (1961) found only a few per cent in collagen extracted from calf skin.

The data in Table II constitute a detailed comparative study of collagen composition. The two collagens from the rat are unlike in several respects. The differences in lysine and hydroxylysine in these and other collagens have already been reported (Piez and Likins, 1960). However, the largest dissimilarity is in the content of 3-hydroxyproline. The tyrosine values probably differ significantly, but other small differences have doubtful significance. The differences may be expressed at higher levels of structure and may be critically related to the different functions which tail tendon and skin perform.

The carp collagen shows many obvious similarities to the rat collagens. There is a similar pattern of amino acids which are present in different amounts in α1 and α2, and these differences are all in the same direction except for isoleucine and possibly aspartic acid. However, there are also some pronounced differences. Threonine, serine, tyrosine, and phenylalanine are present in different amounts in the subunits of the carp collagen but not the rat collagens. The dogfish collagen is also closely related. Most of the differences found in the other collagens occur here but to different degrees.

In spite of these differences, there is one similarity in the chemistry which is outstanding. In all the subunits of each collagen, the glycine content is always very close to one third of the total residues. This is significant since it is a requirement of the structure of collagen derived from x-ray diffraction studies that the packing of the three polyproline-type helices into the collagen structure is possible only if every third residue is glycine (Rich and Crick, 1962).

In an earlier chromatographic investigation (Piez *et al.*, 1960) of denatured calf skin collagen, employing methods that did not give as good resolution as is now possible, only two major components were observed. In the light of the present studies, it is apparent that these were α1 and β1. The amino acid analyses indicated a 1:1 weight ratio in native collagen (of the type α1β1) which according to the structures proposed here should have been 1:2. This incorrect result probably arose from contamination of β1 with the unrecognized α2 together with experimental error.

References

Altgelt, K., Hodge, A. J., and Schmitt, F. O. (1961), *Proc. Nat. Acad. Sci. U. S. 47*, 1914.

Blumenfeld, O. O., and Gallop, P. M. (1962), *Fed. Proc. 21*, 409.

Boedtker, H., and Doty, P. (1956), *J. Am. Chem. Soc. 78*, 4267.

Franzblau, C., Gallop, P., and Seifter, S. (1962), *Fed. Proc. 21*, 409.

Gallop, P. M. (1955), *Arch. Biochem. Biophys. 54*, 486.

Grassman, W., Hannig, K., and Engel, J. (1961), *Z. physiol. Chem.*, Hoppe-Seyler's *324*, 284.

Gross, J. (1959), *in* Connective Tissue, Thrombosis and Arteriosclerosis, Page, I. H., editor, New York, Academic Press, Inc., p. 77.

Hall, C. E., and Doty, P. (1958), *J. Am. Chem. Soc. 80*, 1269.

Harrington, W. F., and von Hippel, P. H. (1961), *Adv. Prot. Chem. 16*, 1.

Irreverre, F., Morita, K., Robertson, A. V., and Witkop, B. (1962), *Biochem. Biophys. Research Communs. 8*, 453.

Jackson, D. S., and Bentley, J. P. (1960), *J. Biophys. Biochem. Cytol. 7*, 37.

Johnston, J. P., and Ogston, A. G. (1946), *Trans. Faraday Soc. 42*, 789.

Kühn, K., Kühn, J., and Hannig, K. (1961), *Z. physiol. Chem.*, Hoppe-Seyler's *326*, 50.

Martin, G. R., Gross, J., Piez, K. A., and Lewis, M. S. (1961), *Biochim. et Biophys. Acta 53*, 599.

Martin, G. R., Piez, K. A., and Lewis, M. S. (in press), *Biochim. et Biophys. Acta.*

Mazurov, V. I., and Orekhovitch, V. N. (1960), *Biokhimiya 25*, 814.
Mechanic, G. L., and Levy, M. (1959), *J. Am. Chem. Soc. 81*, 1889.
Ogle, J. D., Arlinghaus, R. B., and Logan, M. A. (1961), *Arch. Biochem. Biophys. 94*, 85.
Orekhovitch, V. N., and Shpikiter, V. O. (1955), *Doklady Akad. Nauk SSSR 101*, 529.
Orekhovitch, V. N., Shpikiter, V. O., Mazurov, V. I., and Kunina, O. V. (1960), *Bull. Soc. Chim. Biol. 42*, 505.
Peterson, E. A., and Sober, H. A. (1956), *J. Am. Chem. Soc. 78*, 751.
Piez, K. A., and Gross, J. (1959), *Biochim. et Biophys. Acta 34*, 24.
Piez, K. A., Lewis, M. S., Martin, G. R., and Gross, J. (1961), *Biochim. et Biophys. Acta 53*, 596.
Piez, K. A., and Likins, R. C. (1960), *in* Calcification in Biological Systems, Sognnaes, R., editor, Washington, D. C., Am. Assoc. for the Advancement of Sci., p. 411.
Piez, K. A., and Morris, L. (1960), *Anal. Biochem. 1*, 187.
Piez, K. A., Weiss, E., and Lewis, M. S. (1960), *J. Biol. Chem. 235*, 1987.
Porath, J., and Flodin, P. (1959), *Nature 183*, 1657.
Rice, R. V. (1960), *Proc. Nat. Acad. Sci. U. S. 46*, 1186.
Rich, A., and Crick, F. H. C. (1962), *J. Mol. Biol. 3*, 483.
Veis, A., and Cohen, J. (1960), *Nature 186*, 720.
Veis, A., and Anesey, J. (1961), *Arch. Biochem. Biophys. 94*, 20.

Properties of Chromatographically Purified Trypsin Inhibitors from Lima Beans

GLYN JONES,* STANFORD MOORE, AND WILLIAM H. STEIN

From The Rockefeller Institute, New York 21, New York

Received September 13, 1962

The trypsin inhibitor fraction prepared from lima beans by the procedure of Fraenkel-Conrat *et al.* (1952) has been resolved by chromatography on DEAE-cellulose into four similar, chromatographically homogeneous components. The same substances may be obtained by direct extraction of the beans. Amino acid analyses of each component give integral molar ratios of residues. Assuming a molecular weight of about 10,000, as shown by Fraenkel-Conrat *et al.*, the substances contain by analysis 76, 77, 86, and 93 residues. The corresponding calculated molecular weights are 8291, 8408, 9423, and 9892. The inhibitors lack methionine and tryptophan, but are characterized by an unusually high content of cystine, ranging from 17.1 to 20.2 g of cystine per 100 g of protein. Cleavage of the disulfide bonds by oxidation with performic acid or by reduction with mercaptoethanol abolishes inhibitor activity. The oxidized and the reduced-carboxamidomethylated derivatives are still very resistant to the action of trypsin, however; an understanding of the basis for this resistance will depend upon further structural investigation of the inhibitors.

After the discovery by Kunitz (1947) of the presence in soybeans of a highly specific inhibitor of trypsin, substances possessing similar inhibitory properties were isolated from a number of sources. The material first obtained from lima beans by Fraenkel-Conrat *et al.* (1952) forms the subject of the present communication. The most active preparations, which Fraenkel-Conrat *et al.* were careful to point out might not be pure, were shown to have molecular weights of about 10,000 and to retain activity after exposure to extremes of *p*H and temperature or after treatment with proteolytic enzymes such as pepsin or papain. Amino acid analyses showed that these inhibitor preparations contained 16% cystine and were thus similar to keratin in degree of —S—S— cross-linking. Jirgensons *et al.* (1960) have examined such preparations chromatographically and reported further data on their amino acid composition and physical properties.

The present report describes the isolation and some of the properties of four similar inhibitors which may be prepared chromatographically either by direct extraction of lima beans or from the inhibitor fraction obtained by the procedure of Fraenkel-Conrat *et al.*

Materials and Methods

The lima beans (var. Fordhook, certified seed) were purchased from Burnett Bros., New York. The commerically prepared samples of lima bean trypsin inhibitor were obtained from Worthington Biochemical Corporation. Two lots, No. 539/40 and No. 541, were examined.

DEAE-Cellulose (exchange capacity 0.96 meq/g), Selectacel Type 40, Lot No. 1305, was obtained from Carl Schleicher and Schuell Co. The cross-linked dextrans, Sephadex G-75 and Sephadex G-25, were purchased from AB Pharmacia, Uppsala, Sweden.

The crystalline proteolytic enzymes were obtained from Worthington Biochemical Corporation: trypsin Lot No. TL 550 and Lot No. SF 926 and pepsin Lot No. P 636.

Extraction of Inhibitor from Lima Beans.—Four aqueous extractants were tested—water, 0.05 N HCl, 0.1 M ammonium formate at *p*H 3.20, and 0.25 N H_2SO_4. The extracts had *p*H values of 6.20, 4.25, 3.50, and 1.55, respectively, when measured at 25°.

All operations were performed in a cold room at 4–5°. The standard procedure was as follows: Finely ground beans (100 g) were extracted for 30 minutes with 80% ethanol (500 ml). The mixture was filtered at the water pump. The semidry meal was then stirred with 500 ml of the appropriate extractant for 1 hour and the mixture was centrifuged for 30 minutes at 1200 × *g*. The supernatant solution was removed, and the sediment was extracted for 1 hour with another 500-ml portion of extractant.

The extracts were adjusted to *p*H 5.0–5.1 with 5 N NH_4OH and combined, and solid ammonium sulfate was added with stirring to 50% saturation. After 2 hours, the precipitated protein was removed by centrifugation and taken up in 50 ml of 0.1 M ammonium formate–formic acid (*p*H 3.20). After removal of insoluble matter, the solution, which contained about

* Sophie Fricke Royal Society Research Fellow in The Rockefeller Institute 1960–62. On leave from the Rowett Research Institute, Aberdeen, Scotland.