Action of Proteolytic Enzymes on Tropocollagen and Insoluble Collagen*

M. P. Drake,† P. F. Davison, S. Bump, and F. O. Schmitt

ABSTRACT: The findings reported in this paper support and extend the hypothesis that the fundamental building block of collagen, the tropocollagen molecule, contains extra-helix peptide appendages; that these telopeptides, whose number remains undetermined, are susceptible to digestion by a variety of noncollagenase proteases; that such digestion changes the properties of the treated tropocollagen but does not change its main structural features; and that most of the intra- and intermolecular cross links found in collagen occur through the telopeptides. In addition, data are presented which show that most soluble tropocollagen solutions normally contain covalently linked tropocollagen polymer aggregates.

The isolation of highly purified calfskin tropocollagen and studies of its reaction with pepsin carried out in this laboratory have been previously described (Rubin *et al.*, 1963, 1965). From these studies it was concluded that peptide appendages, termed "telopeptides," external (*i.e.*, exposed to enzyme action) to the triple-helix body of the tropocollagen molecule exist, and that through some of these peptides intramolecular bonds between the three component α chains of the tropocollagen molecule are formed. Some of the telopeptides are vulnerable to pepsin action since intramolecular cross links are broken on pepsin digestion and fragments comprising a small fraction of the tropocollagen molecule (approximately 1%) become dialyzable.

Several investigators (*e.g.*, Grant and Alburn, 1960; Oneson and Zacharias, 1960; Nishihara and Miyata, 1962; Hafter and Hörmann, 1963; and Stevens, 1965) have studied the action of enzymes on soluble and insoluble collagen and it has been shown that fragments of tropocollagen may be excised by certain enzymes and that insoluble collagens may be solubilized to some extent.

Anticipating that studies on the action of enzymes of different specificities upon tropocollagen would provide additional insights into the role and importance of telopeptides and would also facilitate the elucidation of the number and sequence of amino acids in the telopeptides and the number of telopeptides on each molecule, the action of a variety of proteolytic enzymes upon soluble and insoluble calfskin collagen has been studied and is reported in this paper. An accompanying paper deals with the physical characterization of pronase-treated and helium-sonicated tropocollagen. From the investigations reported in these two papers, a mechanism for the conversion of soluble tropocollagen molecules into insoluble collagen fibrils is postulated.

Materials

Reagents. Inorganic reagents were commercial reagent grade. Water was double distilled. The dialysis tubing was precleaned as in the former experiments by boiling for 2 hr in 10% sodium carbonate solution and rinsing many times with 0.05% acetic acid solutions.

Enzymes. The following enzymes were used without further purification: α-chymotrypsin (Worthington CDI, three times crystallized), Elastase (Mann no. 183, crystallized), pepsin (Worthington PM, or Mann no. 2095, two times crystallized), pronase (California Biochemical Corp. no. 53702, grade B), trypsin (Worthington TRL, two times crystallized).

Soluble Tropocollagen. Tropocollagen solutions (acid-soluble fraction 2A from calfskin) were prepared as described previously (Rubin *et al.*, 1965) and were stored as frozen precipitates. Samples were thawed and initially dissolved as required in 0.05% acetic acid (with the precaution that no solutions were maintained in acid solvents for more than a few days and the temperature of such solutions was maintained close to 0°).

For enzyme reaction at a neutral pH, the solubilized tropocollagen was dialyzed against several changes of 0.1 M calcium acetate solution; this procedure produced a solution having little buffer capacity but a pH range of 7.0–7.3. Where a higher pH was desired, 0.1 M Tris at the desired pH was added to make the solution 0.001 M in Tris. For a lower pH (with excellent buffering

* From the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts. *Received August 11, 1965.* This investigation was supported by grant NB00024-15 from the National Institute of Neurological Diseases and Blindness, National Institutes of Health, U. S. Public Health Service.

† Supported by a U. S. Public Health Fellowship (5F3 AM-11,175) from the National Institute of Arthritis and Metabolic Diseases. Present address: Department of Biochemistry, School of Basic Medical Sciences, The University of Tennessee, Memphis 3, Tenn.

TABLE I: Amino Acid Composition[a] of Telopeptides from Dialysates of Enzyme-Treated Soluble Calfskin Tropocollagen.

	Pronase		Elastase		Chymotrypsin		Trypsin		Pepsin	
Amino Acid	A[b]	B	A[c]	B	A[c]	B	A[d]	B	A[d]	B
Hydroxyproline	0.15	6	0.10	2	0.054	1	0.23	2	0	0
Aspartic Acid	0.24	10	0.20	3	0.275	5	0.34	3	0.29	2
Threonine	0.083	3	0.022	0	0.071	1	0.021	0	0	0
Serine	0.20	8	0.084	1	0.27	5	0.23	2	0.15	1
Glutamic Acid	0.24	10	0.25	4	0.39	7	0.44	4	0.43	4
Proline	0.40	17	0.16	3	0.23	4	0.46	4	0.25	2
Glycine	0.73	30	0.33	6	0.53	9	1.26	11	0.44	4
Alanine	0.21	9	0.091	2	0.19	3	0.37	3	0.10	1
Valine	0.074	3	0.015	0	0.068	1	0.083	1	0	0
Methionine	Trace	1	Trace	0	Trace	0	Trace	0	0	0
Isoleucine	0.064	3	0.020	0	0.037	1	0.048	0	0	0
Leucine	0.15	6	0.21	4	0.23	4	0.29	2	0.39	3
Tyrosine	0.16	7	0.18	3	0.17	3	0.37	3	0.44	4
Phenylalanine	0.12	5	0.12	2	0.18	3	0.16	1	0.35	3
Hydroxylysine	0.014	1	0	0	0.018	0	0	0	0	0
Lysine	0.050	2	0.024	0	0.057	1	0.15	1	0.03	0
Histidine	0.027	1	0.037	1	0.073	1	0.071	1	0.06	1
Arginine	0.14	6	0.040	1	0.085	1	0.28	2	0.12	1
		128		32		50		40		26

[a] Values are corrected for contribution from enzyme and collagen, and for losses in dialysis and hydrolysis. Values are expressed as micromoles in column A and as residues per mole in column B (using 260,000 for the molecular weight of collagen). [b] Best values (as described in Experimental Section) for dialysate from 6 mg of tropocollagen. [c] Averaged limit digest values for dialysate from 15 mg of tropocollagen. [d] Averaged limit digest values for dialysate from 30 mg tropocollagen.

capacity around pH 5) glacial acetic acid was added dropwise.

Insoluble Collagen. The insoluble collagen used was a fraction of the residue remaining after repeated extraction with 10% NaCl, then 0.067 M Na_2HPO_4 solutions, and finally 0.15 M pH 3.7 sodium citrate buffer as described previously (Rubin *et al.*, 1965). This residue was thawed, initially suspended in 0.05% acetic acid (using a motor-driven stirrer), and centrifuged. The centrifuged material was deposited in three layers; hair predominated in the bottom layer, granular material in the second, and white finely divided insoluble collagen on the top. The top layer was carefully scraped from the tubes and is the material designated as insoluble collagen in this study. (It was not possible to remove completely small strands of hair until after enzyme treatment.) The insoluble collagen was resuspended in the desired buffer solution and was centrifuged twice more before being used.

Methods

The methods and equipment used in these experiments have been previously described (Rubin *et al.*, 1965). The specific optical rotation values were calculated from the concentrations determined by Kjeldahl nitrogen analyses using nitrogen content of 17.6% (Rubin *et al.*, 1963). The specific optical rotation (at 365 mμ) for a native tropocollagen solution is −1330° and for a heat-denatured solution (heated to 43° for 5 min and measured at 43°) is −460°. After the initial calibration, concentrations of the tropocollagen solutions used for other analyses were routinely determined by calculation from optical rotation measurements on the heat-denatured (43°) material.

Treatment of Collagen with Enzymes. To remove dialyzable impurities, enzymes were dialyzed as 1% solutions against the reaction buffer at 4° for 16 hr before use. Dialyzed pronase solutions were sterilized by filtration through a Millipore HA 0.45-μ filter in order to remove viable cells.

The dialyzed enzyme was added to the collagen solutions to make an enzyme:substrate weight ratio of 1:100, and the mixture was gently agitated at 20°. With pronase, when digestion periods extended over 12 hr, the enzyme–collagen mixture was sterilized by Millipore filtration after 1–3 hr of reaction time. The digestion period for the insoluble collagens was routinely 24 hr at 20° for all enzymes.

Recovery of Dialysate Fraction. After digestion, the

TABLE II: Micromoles of Amino Acids in Dialysate During Digestion of 6 mg of Soluble Calfskin Tropocollagen with 60 μg of Pronase at 20° in 0.1 M Calcium Acetate pH 7.2 Solution.

Amino Acid	Controls		Dialysate Hours of Digestion[c]						
	Pronase[a]	Collagen[b]	3	6	11	24	48	72	96
Hydroxyproline	0	0	0.10	0.14	0.17	0.19	0.27	[d]	0.34
Aspartic Acid	0.006	0.040	0.24	0.25	0.27	0.27	0.30	[d]	0.36
Threonine	0.004	0.025	0.053	0.056	0.069	0.063	0.081	[d]	0.095
Serine	0.008	0.11	0.11	0.13	0.20	0.27	0.30	[d]	0.32
Glutamic Acid	0.006	0.092	0.28	0.30	0.34	0.34	0.40	[d]	0.49
Proline	0.002	0.023	0.28	0.32	0.38	0.47	0.58	0.66	0.73
Glycine	0.009	0.098	0.72	0.80	0.96	1.00	1.21	1.36	1.53
Alanine	0.006	0.048	0.21	0.23	0.27	0.30	0.35	0.40	0.44
Valine	0.003	0.017	0.041	0.054	0.075	0.098	0.11	0.12	0.12
Methionine	Trace	Trace	Trace	Trace	Trace	Trace	Trace	Trace	Trace
Isoleucine	0.002	0.012	0.050	0.062	0.075	0.079	0.077	0.087	0.092
Leucine	0.003	0.015	0.17	0.18	0.19	0.19	0.21	0.23	0.25
Tyrosine	0.002	0.019	0.22	0.21	0.22	0.20	0.19	0.20	0.21
Phenylalanine	0.001	0.009	0.14	0.14	0.16	0.15	0.15	0.16	0.18
Hydroxylysine	0	0	0.011	0.014	0.019	0.021	0.021	0.032	0.034
Lysine	0.002	0.014	0.053	0.057	0.068	0.079	0.080	0.087	0.10
Histidine	0.002	0.024	0.031	0.031	0.037	0.046	0.049	0.056	0.060
Arginine	0.002	0.009	0.12	0.13	[d]	0.20	0.21	0.25	0.27

[a] Average (divided by 16.7) of dialysates from three 1-ml samples of pronase solution (1 mg/ml) taken during course of digestion. [b] Average of dialysates from three 2-ml samples of collagen solution (3 mg/ml) taken during course of digestion. [c] Values not corrected for hydrolysis destruction nor for contribution from collagen and pronase. [d] Peak inadequately resolved for reliable estimation.

tropocollagen and enzyme were dialyzed against a large volume of 0.05% acetic acid at 4° for 48 hr to separate the dialyzable peptides. The dialysate was concentrated on a Buchler rotating evaporator (at 50°) and the calcium ions (when present) were titrated with oxalic acid. The precipitated calcium oxalate was removed by centrifugation and washed twice with small volumes of fresh 0.05% acetic acid. The supernatant and washings were then combined and evaporated to dryness. Control experiments established that over 95% of a standard amino acid mixture could be recovered from a 0.1 M calcium acetate solution in this manner. Routinely, a second dialysis of the concentrated dialysate was made to ensure that traces of tropocollagen were not carried through the peptide isolation procedure.

Recovery of Collagen Fraction. The enzyme-treated collagen was recovered and separated from the enzyme by repeating three times the precipitation of the collagen with 15% KCl, centrifugation, and resuspension in fresh 0.05% acetic acid. In the case of pepsin, 0.02 M Na_2HPO_4 was also added to the precipitation step to denature the pepsin by raising the pH. Freed in this manner from the protease, the tropocollagen could be denatured without danger of further enzymic degradation. The presence of residual traces of enzyme in denatured solutions could be easily detected by ultracentrifuge analyses at 40° as α chains would be degraded and material sedimenting more slowly than the α chains would appear.

Sedimentation Area Measurements. The relative concentrations of the components resolved by sedimentation were determined by measuring the respective areas of the peaks by planimetry on enlarged schlieren photographs and correcting for radial dilution. The Johnson–Ogston effect was minimized by using concentrations of protein lower than 0.4% (Speakman, 1962) but was otherwise disregarded.

Results

Soluble Tropocollagen. The amino acid content of the dialyzable peptides is shown in Table I for the various enzymes employed. In every case but pronase, these are limit digests, *i.e.*, the digestion has been followed for at least 72 hr and the values shown are representative of the plateau region where no further digestion could be detected. This plateau region occurred within 24 hr for every enzyme except pronase.

With pronase it was evident from both the amino nitrogen and total amino acid content in the dialysates that a plateau region was never reached. Preliminary experiments showed evidence of microbiological contamination which dictated the subsequent use of Millipore filtration as described above. Although the 0.45-μ

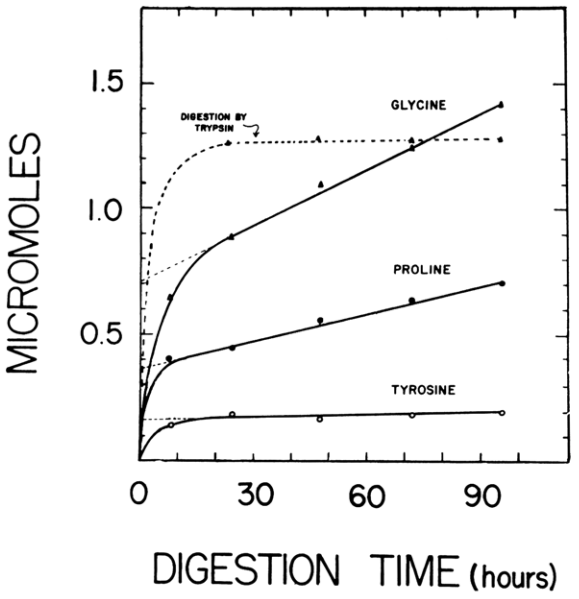

FIGURE 1: A typical plot for the micromoles of dialysate amino acids *vs.* time of digestion by pronase is shown for 6 mg of calfskin tropocollagen at 20° in 0.1 M calcium acetate pH 7.3 solution. The broken line extrapolation to zero time used to determine the values in Table I is also shown. For illustration of a limit digest, the micromoles of dialysate glycine *vs.* time for digestion of 20 mg of tropocollagen by trypsin is shown by a broken line.

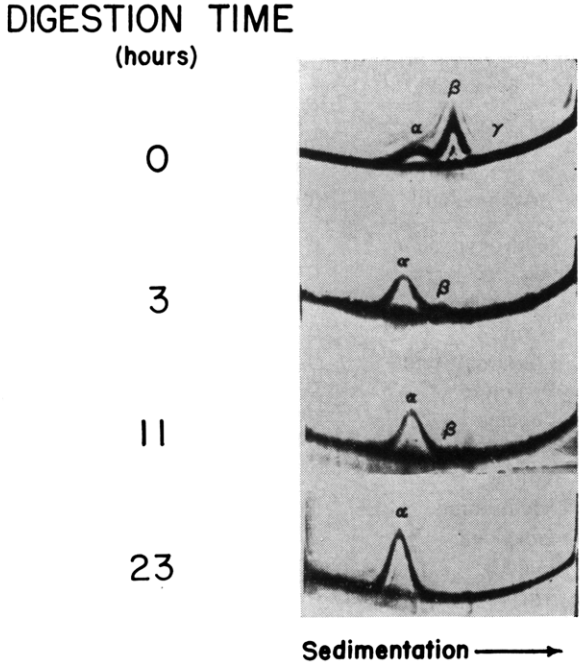

FIGURE 2: The disappearance of cross links and the conversion of soluble calfskin tropocollagen to an all α-chain molecule during pronase digestion is shown in schlieren pictures of the denatured material (after separation from the pronase) during sedimentation velocity runs at 56,100 rpm and 40°. Centrifugation time 60–90 min; protein concentration 0.3–0.4%; bar angle 60°.

filters clog very rapidly with soluble tropocollagen solutions, pronase-treated tropocollagen solutions pass quite readily through such filters after 1-hr digestion.

Data from typical analyses of dialysates obtained after varying periods of digestion by pronase are shown in Table II, and some of the values for particular amino acids are plotted in Figure 1. Similar plots of each amino acid were made, and the values shown for the pronase telopeptides content in Table I are the zero-time values for the extrapolated slopes of each amino acid.

The enzymes did not denature the tropocollagen molecule since there was only a small loss of optical rotation in any of the digested solutions. The greatest loss occurred with pronase where the specific optical rotation at 365 mμ fell from −1330 to *ca.* −1260° during the 12-hr reaction. Removal of the telopeptides by dialysis restored the specific optical rotation (calculated from nitrogen analyses) of the digested solution to −1320°, *i.e.*, almost normal. That pronase will react with denatured collagen was checked by heat denaturing the tropocollagen at 40° for 10 min and then digesting at 20° either immediately afterward or after the solution, held at 4° overnight, had regained almost 60% of its original optical rotation. In both cases, more than 50% of the tropocollagen nitrogen became dialyzable within 8 hr, and there was only a small difference in the two rates of digestion. With the other enzymes, there was much less collagen digested in the solution held overnight in the cold (*e.g.*, with pepsin only 17% of the tropocollagen nitrogen became dialyzable after 24 hr of digestion at 20°).

The course of enzymatic attack on telopeptides was also followed by ultracentrifugal analysis of the denatured tropocollagen (after removal of the enzyme). In the case of pronase, the sedimentation patterns revealed that a complete conversion to α chains occurred sometime between 11 and 24 hr of digestion (Figure 2). Although the exact time is not better defined because of the difficulty of detecting the presence of small amounts of β or γ components, this time period correlates with the marked change in the rate of peptide release shown in Figure 1. Routinely, therefore, the digestion of soluble tropocollagens by pronase was terminated after 12 hr of digestion when pronase-treated tropocollagen was being prepared for other studies. The physical characterization of such pronase-treated tropocollagen is described more fully by Davison and Drake (1966).

In Table III the relative amounts of telopeptide material released by the action of the enzymes used may be compared with the change in the ratios of the α, β, and γ components in the enzyme-treated tropocollagen. It is evident that the percentages of cross links broken

TABLE III: Comparison of Selected Parameters from the Digestion of Acid-Soluble Calfskin Tropocollagen by Various Enzymes.

Enzyme Used	Dialysate Amino Acids Number	% [a]	Composition $\alpha:\beta:\gamma$	$[\eta]$
None	0	0	32:65:3	16
Pepsin	26	0.93	74:23:3	14
Elastase	32	1.1	84:13:3	13
Chymotrypsin	50	1.8	76:21:3	14
Trypsin	40	1.4	47:50:3	14
Pronase pH 7.3	128	4.6	100:0:0	9–10

[a] Number of dialysate residues × 100/2800 amino acid residues in tropocollagen (mol wt 260,000).

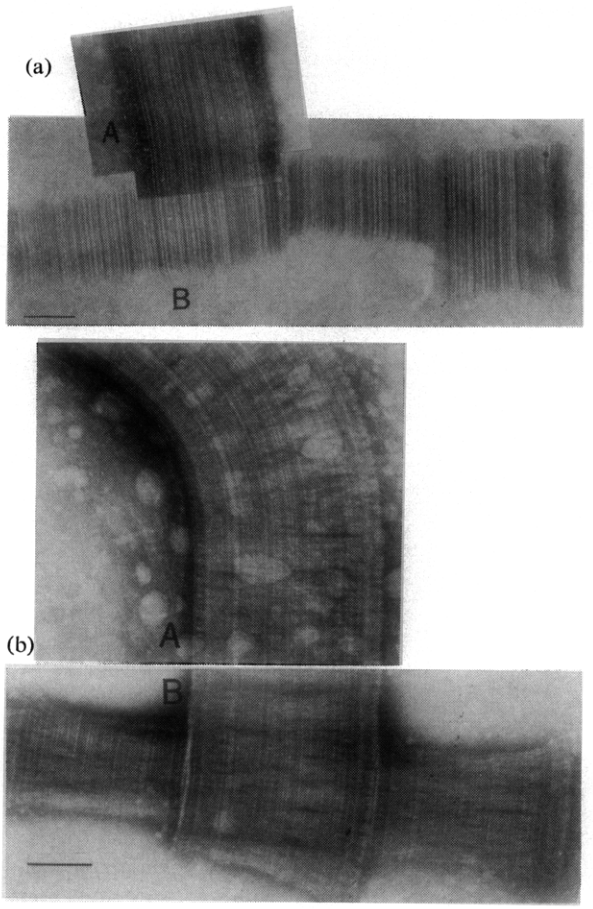

FIGURE 3: Electron micrographs of tropocollagen precipitated as ribbonlike segment-long-spacing aggregates. (a) Pronase-treated tropocollagen (A) is shown juxtaposed to several aggregates (arranged head-to-head and tail-to-tail) from untreated tropocollagen (B). Both samples were positively stained with phosphotungstic acid pH 4.2; some of the stain remains on A delineating the ends of the molecule. (b) Negatively stained (phosphotungstic acid pH 7) segments from pronase-treated (A) and untreated (B) tropocollagen are shown juxtaposed to demonstrate the integrity of the length of the enzyme-treated molecule. The bar in the lower left-hand corner indicates 1000 A.

are not directly correlated with the amount of telopeptide material split off. This disparity is most striking in the case of trypsin digests, where very few cross links appear to be broken, but much telopeptide material appears to be split off. Although some loss of intrinsic viscosity and a lowering of the concentration dependence of the viscosity was evident in each of the enzyme-treated tropocollagens, a marked reduction in viscosity occurred only when pronase was used.

A third means of evaluating the course and the effect of the various enzyme treatments was to precipitate the treated tropocollagens as segment-long-spacing aggregates, and to stain and examine them in the electron microscope. After pronase digestion for 1 or 2 days, and after limit digests from all the other enzymes, there was no electron microscopic indication that any alteration of the main body of the molecule had occurred. This is illustrated in Figure 3 from which it is apparent that all of the normally recognizable bands along the length of the molecule are present and that the total length of the molecule appears unchanged. A better definition of the ends was obtained by application of negative staining techniques (which aided measurement of the length) and such studies verified the fact that the total length remained unchanged after enzyme action.

Electron microscopy also provided direct evidence that pronase digests the tropocollagen molecule itself, as was suggested by the continuing release of peptide material on prolonged incubation. Digestion with pronase was continued for several weeks and aliquots were withdrawn at irregular intervals for precipitation of segment-long-spacing aggregates. A characteristic feature observed in 1–3-day samples was the presence of B–B dimers (*i.e.*, two segments abutting at the B ends); very few A–A dimers were seen. After several days, however, only A–A dimers were present, which suggested that even in apparently intact B ends some changes had occurred which modified the "preferred" B–B end-to-end interaction. Subsequent examination after 12, 21, 35, and 49 days of digestion revealed a progressive attack on the molecule itself as is illustrated in Figure 4. Even after the addition of further aliquots of sterile pronase, however, segment-long-spacing aggregates from 49-day digests contained many intact segments. In all of the pictures studied, the A ends of the molecule remained (to this degree of resolution) intact and digestion proceeded, apparently stepwise, from the B end.

Digestion by trypsin appeared to change the charge profile along the body of the molecule because precipitation of trypsinized tropocollagen in segment-long-spacing form at pH 3.5 (*i.e.*, 0.05% acetic acid)

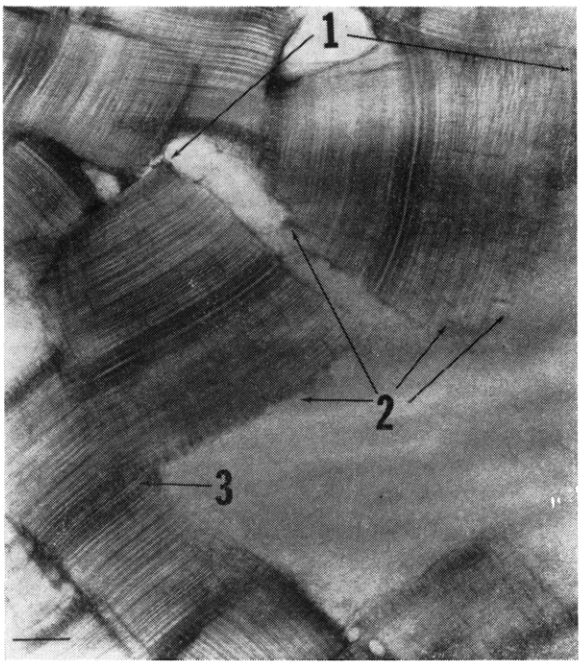

FIGURE 4: Electron micrograph of segment-long-spacing aggregates precipitated from a 12-day pronase digest of tropocollagen. Characteristically the molecules aggregate in long ribbons in which the molecules are aligned in register normal to the axis of the ribbon. These ribbons also frequently associate A end (of each molecule) to A end with a 230 A overlap giving a centrosymmetric structure with A′′–A′′ junctions. Here are shown two such overlapping segments with the B ends (Hodge and Schmitt, 1960) intact (1) and progressively etched by enzyme action (2). 3 shows an unusual B′′–B′′ junction (with some overlap of the segments). B–B junctions appear to be preferred in untreated tropocollagen, but after 12 hr of pronase attack they are rarely seen. Stain, phosphotungstic acid pH 4.2. The bar in the lower left-hand corner indicates 1000 A.

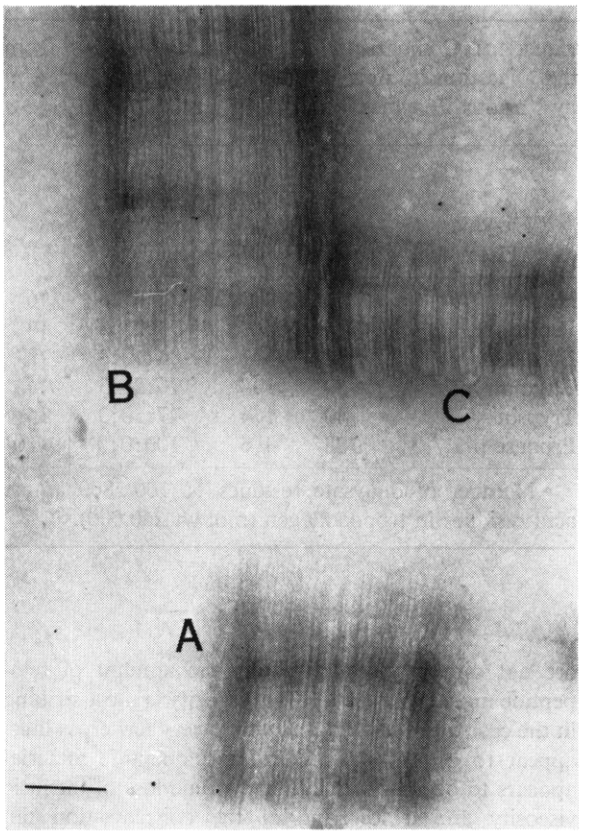

FIGURE 5: Segments precipitated from trypsin-treated tropocollagen by 0.1% adenosine triphosphate (ATP) at pH 3.5, positively stained by phosphotungstic acid pH 4. A centrosymmetric segment is seen at A. At C is seen a normal asymmetric segment. The segment at B is a mosaic of the structure in A and C. The bar in the lower left-hand corner indicates 1000 A.

produced a centrosymmetric band pattern as shown in Figure 5. A similar observation has been reported by Kühn *et al.* (1964) and by Hörmann and von Wilm (1964). The normal asymmetric segments could be precipitated from trypsin-treated tropocollagen solutions by using a pH 3.2 solutions (*i.e.*, 0.05 M acetic acid).

Insoluble Collagen. The solubilizing effect of various enzymes on acid-insoluble collagen, the $\alpha:\beta:\gamma$ composition, and the intrinsic viscosity values of the resulting solubilized collagen are shown in Table IV. Several points merit comment, particularly when the data in this table are compared with those of Table III which summarize the effects of these same enzymes upon the acid-soluble tropocollagen. Pepsin, for example, which produced less than 1% by weight of telopeptide material in the acid-soluble tropocollagens, solubilized 99% of the insoluble collagen and produced a material having a relatively high content of γ molecules [pepsin also attacks few of the cross links in γ-tropocollagen (Rubin *et al.*, 1965)]. Pronase, which produced almost 5% telopeptide material and eliminated all of the cross links in the acid-soluble tropocollagens also solubilized 99% of the insoluble collagen, but only at a slightly acid pH, and always left a residuum of cross links since β and γ chains were still present. Since these β and γ chains persisted after sonic irradiation (see below), they are evidence of pronase-resistant intramolecular bonds not detected in tropocollagen. Some of the collagens solubilized by enzymes were examined by hot-column carboxymethylcellulose chromatography (Piez *et al.*, 1961). The β component from pepsin-digested material or from the pH 7.3 pronase treatment was found to be devoid of β_{11}. A similar finding was previously reported with pepsin treatment of tropocollagen (Rubin *et al.*, 1965).

When different enzymes were used sequentially, the expected specificity of the enzymes was demonstrable. With pronase at pH 7.1 as the final enzyme used, the solubilized collagen contained no intermolecular cross

TABLE IV: Comparison of Selected Parameters from the Digestion of Calfskin-Insoluble Collagen by Various Enzymes.

Enzyme Used	% Solubilized	Composition $\alpha:\beta:\gamma$	$[\eta]$
Trypsin	20	40:55:5	27
Elastase	32	49:43:8	14
Chymotrypsin	37	68:28:4	13
Pronase pH 5.2	99	80:10:10	13
Pronase pH 7.3	87	80:10:10	10
Pepsin	99	40:34:26	18
Substrate and Enzyme Used			
Pepsin solubilized; pronase pH 7.3		80:10:10	10
Insoluble after pronase pH 7.3; pepsin		59:10:26:5[a]	18
Insoluble after chymotrypsin; pepsin		52:32:16	17

[a] Extra component (X) moving faster than γ.

TABLE V: Comparison of Hexose and Hexosamine Content of Collagen after Treatment with Enzymes.

Enzyme	% Hexose	% Hexosamine
Untreated acid-soluble tropocollagen	0.5	0.1
Pepsin-treated tropocollagen	0.5	0.1
Pronase-treated tropocollagen	0.5	0.1
Pepsin-solubilized insoluble collagen	0.8	0.2
Pronase-solubilized insoluble collagen	0.6	0.2
Insoluble residue[a]	25	
Insoluble residue[b]	50	

[a] Insoluble residue from sequential digestion by pronase and pepsin (contained no hydroxyproline). [b] Insoluble residue after exhaustive acid extraction of residue from sequential digestion by chymotrypsin and pepsin (contained no hydroxyproline).

links (as judged by viscosity) and was reduced to a characteristic complement of 10% β and 10% γ intramolecularly cross linked molecules. With pepsin as the final enzyme a relatively high content of γ molecules was present, and when the insoluble material left after pronase digestion of insoluble collagen at pH 7.1 was treated with pepsin, there was present in the resulting solution (in addition to α, β, and γ components) a higher molecular weight compound (X) with an $s^{0,3^{\circ}_{\mathrm{w}}}_{20,\mathrm{w}}$ value of 6 S. This preparation proved to be particularly helpful in interpreting the data, as will be shown below. Enzyme digestion was repeated on this fraction in an attempt to determine whether the existence of the X component was due to an incomplete reaction with the enzymes, or whether it indicated that *intermolecular* cross links exist which are resistant to pronase digestion, similar to the *intramolecular* cross links already revealed. As shown in Figure 6, additional pepsin digestion had little effect, but additional 12-hr pronase digestion at pH 7.3 broke up almost half of the X component. This experiment implies that some of these intermolecular bonds at least are susceptible to pronase but are digested slowly, a fact which allows the structure X to be isolated.

The hexose and hexosamine contents of the various tropocollagen solutions are shown in Table V which also contains the analyses of the truly insoluble material, *i.e.*, the approximately 1% of material remaining after pepsin and/or pronase digestion of insoluble calfskin collagen.

Two physical properties which were consistently measured in these studies were the optical rotation at 365 mμ and intrinsic viscosity, since these physical properties radically change when the molecule denatures from the native-rod form to the random-coil form. The optical rotation results were consistent for the enzyme-treated soluble collagens from both the acid-soluble and the acid-insoluble collagens. For all enzyme treatments, the specific optical rotation for the soluble "native" material differed no more than 2% from the value of native tropocollagen (*i.e.*, $-1310°$ *vs.* $-1330°$, respectively). The specific optical rotation for the denatured solutions (43°) remained (within experimental error) the same for nontreated and enzyme-treated tropocollagen (*i.e.*, $-460°$), except for pronase-treated tropocollagen where the value was found to be $-465°$.

The intrinsic viscosities of the materials solubilized by enzyme action from insoluble collagen were significant in two respects. The first was that rather high values were obtained with trypsin-solubilized tropocollagen and with material from sequential enzyme digestions of insoluble collagen where pepsin was the last enzyme used. The second was that a rather standard value of 10 dl/g was obtained for the intrinsic viscosity of solutions where pronase at neutral pH was used as the final solubilizing enzyme (*e.g.*, pronase pH 7.3, and pepsin solubilized; pronase pH 7.3 in Table IV, and pronase pH 7.3 in Table III).

These results suggested that pronase digestion at pH 7.3 produces monomeric tropocollagen solutions, and that the higher viscosities of other enzyme-treated collagens and also the acid-soluble tropocollagens are due to the presence of firmly, probably covalently linked, polymeric structures. This point will be amplified below.

Intermolecular Cross Links. In the accompanying paper (Davison and Drake, 1966) it is shown that polymeric structures in tropocollagen solutions can be

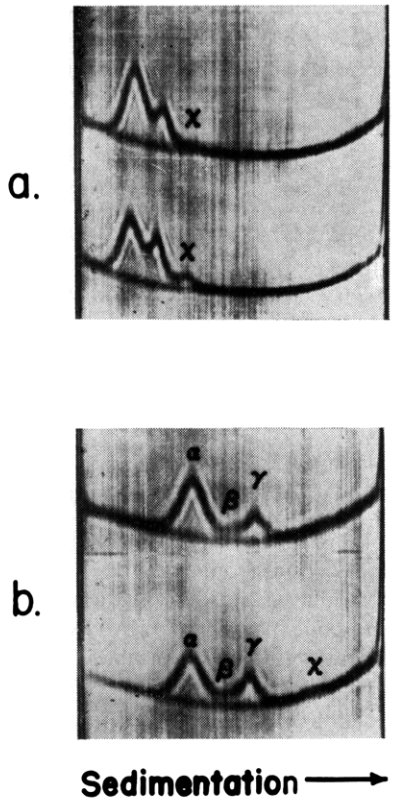

FIGURE 6: The result of an additional enzyme treatment for 24 hr to digest the X component in the soluble collagen from a multiple-enzyme treatment of insoluble collagen (see text and Table IV) as shown in schlieren pictures during sedimentation at 56,100 rpm and 40°. The positive wedge cell (upper line) contains material (concentration 0.37%) re-treated with pronase, and the normal cell (lower line) contains material (concentration 0.35%) re-treated with pepsin. (a) Sedimentation for 48 min, bar angle 60°; (b) sedimentation for 96 min, bar angle 55°.

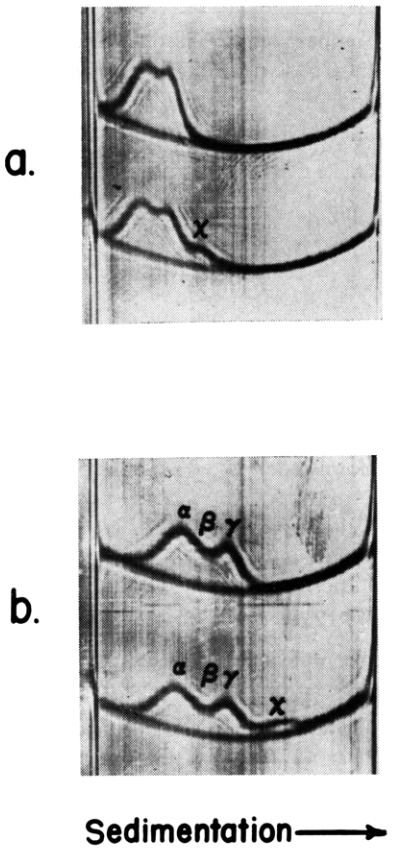

FIGURE 7: Comparison of pre- and postsonication solutions of soluble tropocollagen containing the X component shown during sedimentation at 39,500 rpm and 40°. The normal cell contains the unsonicated material, and the wedge cell contains its sonicated counterpart (concentration 0.26%). (a) Sedimentation for 80 min, bar angle 60°; (b) sedimentation for 128 min, bar angle 55°.

selectively broken by controlled sonic irradiation. If the triple-helix structure of collagen is assumed, then any material sedimenting more rapidly than the γ molecules in a solution of denatured collagen must be a compact form of γ-tropocollagen (*i.e.*, highly cross linked) or a structure containing 4 or more α chains arising from an intermolecular bonding between two or more collagen molecules. The isolation, from insoluble collagen, by sequential pronase and pepsin digestion, of a solution of collagen molecules showing a significant fraction of an "X" component sedimenting faster than γ was described above. A highly cross-linked γ molecule studied by Veis and Drake (1963) had a sedimentation coefficient lower ($s^{0.3\%}_{20,w} = 4.7$ S) than that of this X component, and it appeared probable that X in the denatured diagram arose from covalently linked collagen molecules. This material, therefore, provided an ideal opportunity to test whether intermolecular bonds give rise to linear polymers. The solutions were irradiated in a Bronson Sonifier at an intensity known to give rise to monomeric tropocollagen molecules. When the irradiated material was denatured and examined by ultracentrifugation, no heavy X component could be found (Figure 7).

With the recognition that some of the β, γ, and higher aggregates in collagen solutions can arise from intermolecular bonds and that these bonds can be selectively destroyed by sonic irradiation (if the cross links produce longitudinal rather than lateral polymers), it became possible to differentiate to some extent the intra- and intermolecular components by examination of the relative areas of components in sedimentation patterns of pre- and postsonication solutions. Such a comparison is shown in Figure 8, where the increase in α-chain content in the sonicated sample of acid-soluble calfskin tropocollagen must mean that the intermolecular cross link in many of the polymer aggregates was attached $\alpha:\alpha$ or $\alpha:\beta$, rather than $\beta:\beta$, $\beta:\gamma$, or $\gamma:\gamma$. Particularly note that in neither of the sonicated solutions (Figures

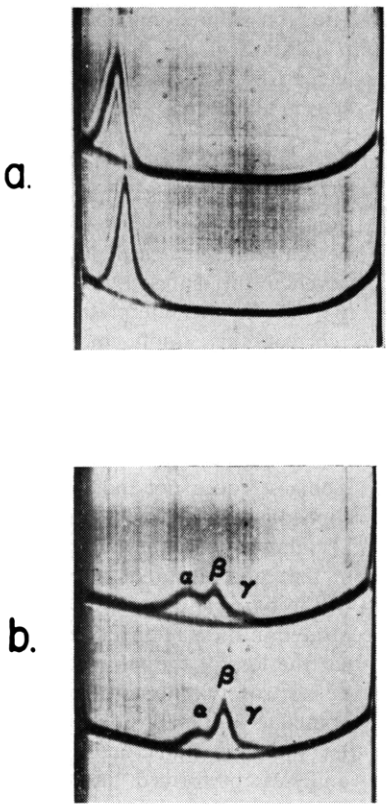

FIGURE 8: A confirmation of the presence of intermolecular cross-linked polymers in acid-soluble calfskin tropocollagen shown by the difference in subunit composition of the tropocollagen before and after sonication. The wedge cell contains the sonicated material (concentration 0.33%, $\alpha:\beta:\gamma$ ratio of 48:40:12), and the normal cell contains its nonsonicated counterpart (concentration 0.36%, $\alpha:\beta:\gamma$ ratio of 33:56:11). Centrifugation at 56,100 rpm and 40°. (a) Sedimentation 32 min, bar angle 60°; (b) sedimentation 96 min, bar angle 55°.

7 and 8) is there detectable any component of a molecular weight lower than an α chain; this is a direct demonstration that the sonic irradiation did not cause scission within the tropocollagen molecule.

Attempts were made to enrich fractions of the tropocollagen in polymeric structures, but with limited success. Ease of precipitation under selected precipitation conditions (*e.g.*, those of Wood (1964)) did give discriminable fractions, but their intrinsic viscosities were all similar (about 21 dl/g) except for the last small fraction obtained by the Wood method which had a viscosity corresponding to monomers (10 dl/g). It was concluded that precipitability is determined by chemical specificity in the molecule rather than by the presence of aggregated structures which act as nuclei for precipitation (a conclusion also reached by Wood).

Discussion

The essential observation in this and in the previous studies (Rubin *et al.*, 1963, 1965) which has shown conclusively that proteases act on collagen molecules is that cross links joining α chains intra- and intermolecularly have been broken. Thus, in the treatment of soluble tropocollagen solutions with enzymes, it is the change in the $\alpha:\beta:\gamma$ ratio which shows that reaction with tropocollagen has occurred. With the exception of pronase, the amounts of nitrogen rendered dialyzable by proteolysis are in the range of 1–3%, and in this range might be attributable to the digestion either of noncollagenous protein impurities or of traces of denatured collagen. Indeed, just such a rationale has been invoked for many years and collagen in the native state has been considered to be totally resistant to all proteases except collagenases. That discrete structural changes have been effected by proteases has now also been confirmed by the independent evidence that antigenic determinants on tropocollagen are lost under such treatment (Schmitt *et al.*, 1964). Another series of experiments by Dabbous (1965) on the formation of intermolecular cross links between tropocollagen molecules by tyrosinase action and the elimination either of the sites of cross linking or of the cross links by proteolysis offers further independent proof of this point.

The very limited nature of the attack by enzymes on the native molecule and the fact that tropocollagen *is* extensively degraded by proteases after denaturation suggest that the chain-folding characteristic of the collagen triple helix (Ramachandran, 1963) confers resistance to proteolysis by the enzymes used in this study, although steric hindrance also may be a factor since much of the peptide bond backbone for each chain lies "buried" within the molecule. Moreover, the fact that none of the enzymes used in this study (with the exception of pronase after a sufficient incubation time) breaks the α chains suggests that the sites vulnerable to enzyme action exist outside the helical main structure, *i.e.*, the vulnerable peptide bonds lie at the ends of the α chains (perhaps where the chains may not terminate in register) or on branch peptides extending laterally from the α chains. Such peptides have been termed telopeptides (Rubin *et al.*, 1963). The lack of scission within the main length of the α chains also implies that if any nonhelical regions interrupt the regular folding of the chains either the structure of the peptide chains at those sites or the local amino acid composition still confers resistance to the enzymes studied here.

The amount of telopeptide material external to the collagen helix and susceptible to proteolytic attack is best estimated from the results with pronase. Approximately 5% of the tropocollagen molecule is digested very quickly; thereafter, a slower attack continues with a progressive nibbling away of the molecules from the B end. Molecules isolated after the first rapid hydrolysis show no indication of attack on the body of the molecule itself, the segment-long-spacing aggregates

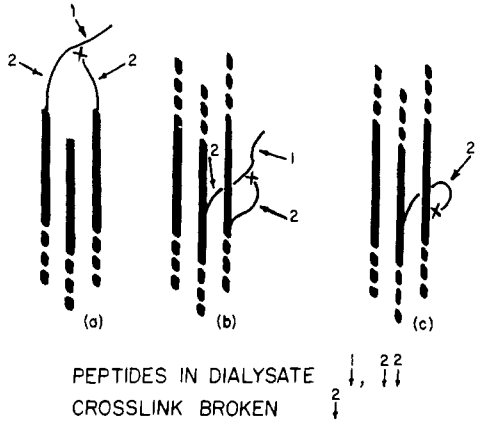

FIGURE 9: Diagram to illustrate how proteases might act to break cross links and produce dialyzable peptides in a situation where the cross link joins two telopeptides along the main chain (a) or at the terminal ends (b), and where the cross link joins a telopeptide to an amino acid of the main peptide chain (c). Potential scissions are labeled 1 when inter-α-chain cross links are not affected, and are labeled 2 when a scission would functionally "break" such a cross link.

appear to be of unchanged length and band pattern, and indeed, some of the bands in the phosphotungstic acid stained electron micrographs may be more sharply defined than usual. Moreover, at this stage the segments still show B–B end-to-end dimers and the measurements across these junctions reveal no change in dimensions.

These findings indicate that one or a number of predominantly peptide chains containing amino acids amounting in molecular weight to *ca.* 15,000/tropocollagen molecule of molecular weight 260,000 (Davison and Drake, 1966) must be arranged somewhere on the molecule in positions which have not been resolvable by electron microscopy but where they are labile to pronase attack. These chains bear apparently none of the carbohydrate in the molecule since approximately the same content of hexose and hexosamine is detected in the acid-soluble tropocollagens and in the protease-solubilized collagen after pepsin and pronase digestion. This suggests that the carbohydrates which have been proposed as cross-linking sites, are either primarily intrachain (as suggested in the subunit hypothesis of Gallop, 1964) or that, if they are interchain cross-linking sites, protease treatment does not remove them from their point of attachment.

Of the enzymes tested, only pronase broke an appreciable fraction of the intramolecular cross links in γ-tropocollagen (see Rubin *et al.*, 1965). The fact that pepsin-solubilized acid-insoluble collagen contained only 25% γ components indicates that such collagen does not differ greatly from soluble tropocollagen except for presumably covalent cross links between the molecules; it does not consist predominantly of γ molecules. The common question of the relevance of studies on soluble collagen to the structure of the preponderant insoluble collagen fibrils may therefore be answered positively.

In experiments treating tropocollagen with pepsin and in similar digestions of insoluble collagen, it was consistently found that no hydroxyproline was detectable in the dialyzable material. Hydroxyproline *is* detected, however, in the dialysate following treatment with other enzymes. If the interpretation of our observations is correct, this simply implies that lengths of the α chains exist which are not included in the protease-resistant helical structure or that the residues are situated on the end of an α chain beyond a protease-vulnerable site. The occurrence of hydroxyproline in these protease digests does not indicate that a small number of tropocollagen molecules are digested since the amino acid composition of the dialysates is markedly different from that of the parent tropocollagen. The greater part of the tyrosine in the tropocollagen molecule is concentrated in these telopeptide structures.

The fact that the loss of the telopeptides is not discernable in the electron micrograph might suggest that they are distributed as a number of small peptides. The possibility that they are numerous is supported by N-terminal analyses performed in this laboratory (J. Rosmus and Z. Deyl, 1965, unpublished observations) on tropocollagen before and after enzyme treatment. Pronase treatment appears to liberate fourteen or more N-terminal groups—a number similar to the number of acetyl groups recently found in tropocollagen by Hörmann and Joseph (1965). These experiments lead to a putative picture of a bewhiskered native tropocollagen molecule, rather than a bare, rodlike structure.

The pepsin-released peptides, in contrast to those released by pronase, tend to adsorb to most chromatographic materials (*e.g.*, paper, cellulose acetate), a fact which may point to their being large peptides. The question of the size and number of the telopeptides is under active study.

Although the presence of subunits in each α chain has been suggested (Gallop, 1964; Petruska and Hodge, 1964) but not proven, the possibility does suggest a mode of origin of numerous telopeptide structures. If the subunits of the α chains are terminated by acetyl groups, polymerization must proceed by subterminal covalent bond formation (*e.g.*, peptide bonds to the ϵ-amino group of a lysine residue or a γ-carboxyl of a glutamyl residue) and the peptide beyond the subterminal link would be a telopeptide.

The nature of the actual intramolecular cross links is still unknown, and there may very well be more than one type of cross link. It is not necessary that the proteases break the cross link itself, however, if the cross link occurs through two telopeptides in adjacent α chains or from a telopeptide to an amino acid of an adjacent α chain. The possibilities are illustrated in Figure 9. If the cross link occurs between two telopeptides, whether terminal or along the length of the molecule, a peptide bond rupture labeled (1) would produce

a dialysate peptide without breaking a cross link, while a rupture labeled (2) would appear as a broken cross link. In general, however, a minimum of two breaks would be required for dialyzable peptides to appear. Similarly, if the cross link is between a telopeptide and a molecule on the backbone peptide chain, or if the enzyme breaks the cross link itself, two breaks would be required for peptides to appear. Furthermore, we would expect that enzymes having different specificities would produce varying amounts of dialyzable peptides or cross-link breaks.

The data from treatment of both soluble and insoluble collagens are in accord with the possibilities outlined above. Trypsin, for example, breaks few cross links and produces 1.5% dialyzable material, while pepsin breaks many cross links but produces a smaller amount of dialyzable peptides. This may indicate that trypsin primarily produces type 1 breaks, but more probably it means that telopeptides digestible by trypsin are not usually involved in cross linking since trypsin also solubilizes only a small amount of the insoluble collagen. The efficiency of pepsin in solubilizing insoluble collagen may be indicative that pepsin attacks the cross links selectively. It is equally likely, however, that the pepsin-sensitive telopeptides are the sites of both the intra- and intermolecular cross-linking reaction. Since the β_{11} structure is completely absent after pepsin treatment of either soluble or insoluble collagen, this would appear to add support to the suggestion of Veis and Anesey (1965) that β_{11} (or, if the three α chains are different, $\alpha_1{:}\alpha_1$, $\alpha_1{:}\alpha_3$, or $\alpha_3{:}\alpha_3$) intermolecular cross links are predominant in insoluble collagen.

The lowered solubilizing efficiency of pronase reaction upon insoluble collagen at a neutral pH appears to involve both a bond specificity and a failure of the enzyme to penetrate the polymer network present in the neutral calcium acetate solutions employed. Since pronase is a mixture of enzymes, it could also indicate the pH dependence of certain of the component proteases. Addition of pronase to pepsin-solubilized "insoluble" collagen reduced the viscosity and interchain cross links to the level seen in the pronase-solubilized tropocollagen solutions, but the latter enzyme alone produces only a limited solubilization. By using the greater dispersive action of a more acid solution (pH 5), solubilization of insoluble collagen with pronase was increased, but the tropocollagen solution produced then contained polymer aggregates as shown by a value greater than 10 dl/g for the intrinsic viscosity. The fact that the X component in the material solubilized from insoluble collagen by pronase and pepsin treatment was reduced only to a limited extent by redigestion with pronase also shows that some resistant intermolecular cross-link areas exist; such pronase-resistant bonds have not been observed in acid-soluble tropocollagen. Presumably the X component was also present in pronase-treated pepsin-solubilized "insoluble" collagen, but was so diluted by monomers to be undetectable in the sedimentation diagram. Apparently this fraction is enriched in the "core" of the collagen fibers less readily attacked by pronase, and becomes manifest only when the core is dissolved by pepsin. Preliminary denaturation–renaturation studies suggest that X is a cross-linked dimer of γ molecules (*i.e.*, γ–γ).

The presence of a series of linear polymers in most extracted tropocollagen solutions has been intuitively recognized by most investigators for many years, although perusal of the literature suggests that the aggregates were not generally considered to be covalently bonded (*e.g.*, Harding (1965) states that intermolecular bonding does not occur in soluble collagen solutions). Boedtker and Doty (1956) used long periods of ultracentrifugation to reduce the content of aggregates, and these same solutions permitted Hall and Doty (1958) to make electron micrographs of single molecules with much less variation in length than had been possible before centrifugation. More recent indications that intermolecularly cross-linked polymers exist in tropocollagen solution have come from disc and zone electrophoresis experiments by several investigators (*e.g.*, Nagai *et al.*, 1964; Veis and Anesey, 1965) which have shown the presence of material with molecular weights greater than that of a γ molecule. The quantitative chromatography studies of Tristram *et al.* (1965) may also be interpreted to show the presence of intermolecularly cross-linked polymers in acid-soluble calfskin tropocollagen.

With the recognition that some of the α, β, and higher aggregates in collagen solutions can arise from intermolecular bonds, it is apparent that the present nomenclature for identifying the α-chain structure of molecular components needs modification. The apparent nonidentity of the three α-chains in a tropocollagen molecule (Piez, 1964) is a welcome discovery since, if this fact applies to all collagens, it becomes apparent that any β_{11}, β_{22}, or β_{33} components arise from intermolecular bonding. There exists no way at present, however, to easily identify a β_{12} or a γ_{123} component whose constituent α chains come from more than one molecule. We propose, therefore, that different molecular origins of α chains be designated by adding a single or double prime as appropriate. The components above, then, would be $\beta_{12'}$ and $\gamma_{123'}$, or $\gamma_{12'3''}$. This solution would preserve the superscript areas for the designation of subunits of α chains as was originally proposed (Kohn *et al.*, 1964) and prefixes could be reserved to designate structures involved in the cross link. From these findings and those reported in our earlier papers we would propose the following model for collagen fibrogenesis.

After synthesis the monomer tropocollagen molecule is secreted from the fibroblast. The molecule may carry upon it a number of protruding peptide chains, and these may exist at the sites where the subunits, if any, making up the α chains, are connected.

Since a large fraction of the tyrosine in the tropocollagen molecule is removed by enzyme treatment it would appear to be concentrated in the accessible telopeptides. However, the fraction removed by pepsin ($^1/_3$) or pronase ($^3/_4$) is not readily reconcilable with any current subunit hypothesis and it is possible that extra peptides are attached to the nascent tropocollagen

protropocollagen molecule by an intra- or extracellular process.

The exposed telopeptide chains are vulnerable to protease attack, and if homeostatic control of collagen deposition within an organism is mediated by enzymes, it is to be anticipated that this enzyme activity is directed to these vulnerable sites. We postulate, therefore, that in the course of the maturation of the collagen fibril, enzymes attack protropocollagen to produce the α_3-tropocollagen molecule which interacts specifically with its fellows to form a semicrystalline quarter-staggered aggregate. Subsequently enzymes are selectively activated to produce intra- and intermolecular cross links (perhaps by transpeptidation) between the telopeptide structures. In this way linear polymers (protofibrils) of tropocollagen are formed, and the creation of β and γ chains and of polymers of the tropocollagen molecules may be assumed to be contemporaneous and perhaps competitive processes (see also Veis and Anesey, 1965). A subsequent lateral bonding of these protofibrils in a quarter-staggered array by the synthesis of lateral cross links (again through telopeptides) between the molecules results in the appearance of an insoluble cross-linked collagen fiber. Subsequent attack by other proteases on the telopeptides might result in the reversal of this process and the solubilization of the insoluble fibril. Thus it may be conjectured that homeostatic control over the deposition and removal of collagen fibers is achieved by enzymatic activation of precursors, by enzyme-mediated intramolecular and intermolecular cross linking and by selective enzymatic cleavage of cross links already laid down.

Acknowledgment

The authors gratefully acknowledge the excellent technical assistance of Mr. J. W. Jacques, Mrs. A. Gould, Mrs. J. Moor, Mrs. E. Myers, and Miss A. Holzer.

References

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

Dabbous, M. (1965), M.S. Thesis, Massachusetts Institute of Technology.

Davison, P. F., and Drake, M. P. (1966), *Biochemistry 5*, 313 (this issue; following paper).

Gallop, P. M. (1964), *Biophys. J. 4*, 79.

Grant, N. H., and Alburn, H. E. (1960), *Arch. Biochem. Biophys. 89*, 262.

Hafter, R., and Hörmann, H. (1963), *Z. Physiol. Chem. 330*, 169.

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

Harding, J. J. (1965), *Advan. Protein Chem. 20*, 109.

Hodge, A. J., and Schmitt, F. O. (1960), *Proc. Natl. Acad. Sci. U. S. 46*, 186.

Hörmann, H., and Joseph, K. T. (1965), *Biochim. Biophys. Acta 100*, 598.

Hörmann, H., and von Wilm, M. (1964), *Naturwiss. 51*, 464.

Kohn, R. R., Bensusan, H. B., and Klein, L. (1964), *Science 145*, 186.

Kühn, K., Schuppler, G., and Kühn, J. (1964), *Z. Physiol. Chem. 338*, 10.

Nagai, Y., Gross, J., and Piez, K. A. (1964), *Ann. N. Y. Acad. Sci. 121*, 495.

Nishihara, T., and Miyata, T. (1962), *Trans. 3rd Collagen Sym. Collagen Res. Soc. Japan*, 84.

Oneson, L., and Zacharias, J. (1960), *Arch. Biochem. Biophys. 89*, 271.

Petruska, J. A., and Hodge, A. J. (1964), *Proc. Natl. Acad. Sci. U. S. 51*, 871.

Piez, K. A. (1964), *J. Biol. Chem. 239*, PC4315.

Piez, K. A., Lewis, M. S., Martin, G. R., and Gross, J. (1961), *Biochim. Biophys. Acta 53*, 596.

Ramachandran, G. N. (1963), *Intern. Rev. Connective Tissue Res. 1*, 127.

Rubin, A. L., Drake, M. P., Davison, P. F., Pfahl, D., Speakman, P. T., and Schmitt, F. O. (1965), *Biochemistry 4*, 181.

Rubin, A. L., Pfahl, D., Speakman, P. T., Davison, P. F., and Schmitt, F. O. (1963), *Science 139*, 37.

Schmitt, F. O., Levine, L., Drake, M. P., Rubin, A. L., Pfahl, D., and Davison, P. F. (1964), *Proc. Natl. Acad. Sci. U. S. 51*, 493.

Speakman, P. T. (1962), *Biochim. Biophys. Acta 69*, 480.

Stevens, F. S. (1965), *Biochim. Biophys. Acta 97*, 465.

Tristram, G. R., Worrall, J., and Steer, D. C. (1965), *Biochem. J. 95*, 350.

Veis, A., and Anesey, J. (1965), *J. Biol. Chem.* (in press).

Veis, A., and Drake, M. P. (1963), *J. Biol. Chem. 238*, 2003.

Wood, G. C. (1964), *Intern. Rev. Connective Tissue Res. 2*, 1.