Proc. Nat. Acad. Sci. USA
Vol. 68, No. 6, pp. 1138–1142, June 1971

Evidence for Procollagen, a Biosynthetic Precursor of Collagen

(zymogen/pepsin digestion/chromatography/gel electrophoresis/isotopic labeling)

GARLAND BELLAMY AND PAUL BORNSTEIN*

Departments of Biochemistry and Medicine, University of Washington, Seattle, Wash. 98105

Communicated by Hans Neurath, March 16, 1971

ABSTRACT **Incubation of rat calvaria for short times in the presence of a labeled amino acid revealed the existence of a collagen fraction (procollagen) that functions as a biosynthetic precursor of collagen. Procollagen contains an α1-like chain (pre-α1) that elutes earlier from CM-cellulose than does rat-bone α1 and has a molecular weight, estimated by acrylamide gel electrophoresis, of 120,000. A time-dependent conversion of pre-α1 to α1 was demonstrated by incubation of calvaria for periods varying from 9 to 60 min and by a pulse–chase experiment. Limited cleavage of procollagen with pepsin resulted in a molecule with a chain resembling α1 in chromatographic properties, molecular weight, and relative hydroxyproline and proline contents. Thus, conversion of procollagen to collagen is likely to occur *in vivo* by a proteolytic mechanism. The additional peptide sequences in procollagen may serve to initiate chain association in triple-helix formation, to facilitate molecular transport, and to inhibit intracellular fibrogenesis.**

The biosynthesis of a functional collagen molecule and its subsequent utilization are critically dependent on a number of processes that follow the polyribosomal assembly of its polypeptide chains. These include the hydroxylation of prolyl and lysyl residues (1, 2), *O*-glycosylation of certain hydroxylysyl residues (3, 4), association of the three polypeptide chains to form the native triple-helical monomer(5), transport of the collagen molecule to the extracellular space (6), and the specific aggregation (7) and cross-linking (8) of collagen to form the connective tissue fibril.

The hydroxylation, glycosylation, and cross-linking of collagen have been the subject of considerable investigation, but less is known of the mechanisms by which chain association leads to the formation of a triple-helical molecule and its secretion from the cell. If we are to understand these processes, it is essential to know whether α chains obtained from extracellular collagen by heat dissociation and chromatography (9) are structurally identical with intracellular collagen polypeptides. One possibility, among others, is that additional peptide sequences facilitate triple-helix formation and transport of the collagen molecule, while at the same time inhibiting intracellular fibrogenesis. Such peptide moieties could be removed by an extracellular proteolytic mechanism, which would thereby initiate the transition of the protein from a soluble to a fibrillar form.

Abbreviations: β-APN, β-aminoproprionitrile-fumarate; SDS sodium dodecyl sulfate.

* To whom requests for reprints may be addressed; recipient of Research Career Development Award, K4-AM-42582 from the U.S. Public Health Service.

During studies designed to investigate the kinetics of collagen chain synthesis in calvarial tissue cultures, we observed that recently synthesized, chromatographically isolated, α1 chains from rat-bone collagen differed from the bulk of extracellular α1 chains in their position of elution from CM-cellulose. Several lines of evidence subsequently demonstrated that such chains are part of a precursor collagen molecule (procollagen)† and that conversion of procollagen to collagen can be simulated *in vitro* by limited proteolysis with pepsin.

METHODS

In vitro synthesis of collagen

Calvaria (frontal and parietal bones) from newborn rats were incubated at 37°C, in Dulbecco's modification of Eagle's medium, under 5% CO_2–95% O_2. The medium was supplemented with 64 μg/ml β-aminopropionitrile-fumarate (β-APN), 100 μg/ml sodium ascorbate, and 0.58 mg/ml glutamine. After an initial incubation for 60 min, the bones from 3–12 rats were placed in fresh medium containing 20 μCi/ml of [2,3-^{3}H]L-proline or 20 μCi/ml of [1-^{14}C]glycine and incubated for 9, 18, 36, or 60 min. In pulse–chase experiments, a large excess of unlabeled L-proline (12.5 mM) was added, and incubation was continued for another 60 min. After incubation, the calvaria were filtered, washed twice with cold water, and homogenized in 0.5 M acetic acid at 0°C in a glass–teflon homogenizer.

Extraction of collagen

Homogenized calvaria were extracted at 4°C with 0.5 M acetic acid for 48 hr; the extracts were clarified by centrifugation and dialyzed against 0.08 M sodium acetate, pH 4.8. No attempts were made to further purify the native collagen. Specifically, filtration through Celite was avoided since experiments indicated that a significant fraction of the biosynthetically labeled protein was adsorbed to the diatomaceous earth.

Chromatography on CM-cellulose

Calvarial extracts were denatured at 40°C for 20 min and centrifuged at 39,000 × *g*. The sample was diluted with an

† The term 'procollagen' was originally applied by Orekhovitch *et al.* (10) to a soluble precursor of the collagen fiber. For this purpose, the term has largely been replaced by the descriptive term 'soluble collagen' or by 'tropocollagen'. We suggest that the use of procollagen be reserved to describe a precursor form of the functional collagen molecule in the sense that the terms procarboxypeptidase and proinsulin are used.

equal volume of deionized 8 M urea that contained 7–15 mg of carrier lathyritic rat-skin collagen, and chromatographed on CM-cellulose at 40°C (9). Chromatography was performed in 4 M urea with a starting buffer of 0.04 M sodium acetate, pH 4.8 and a limiting buffer containing, in addition, 0.08 M NaCl. Aliquots of 0.5 ml were counted in 10 ml of toluene-based scintillant containing 10% BBS-3 (Beckman).

Acrylamide gel electrophoresis

Analytical gel electrophoresis of [^{3}H]collagen chains was performed in 5% polyacrylamide that contained 0.5% SDS (11). Heat-denatured salt-extracted guinea-pig-skin [^{14}C]collagen served as an internal standard to mark the position of migration of α chains. The positions of migration of a mixture of α chains, β components, and CNBr fragments of known molecular weight (12) were also determined by staining with Coomassie Blue. Unstained gels were sliced into 1-mm discs with a gel slicer; the slices were digested in scintillation vials with 0.5 ml of 30% H_2O_2 at 55°C for 3–6 hr and counted in 10 ml of scintillant containing 10% BBS-3 in toluene.

Limited proteolysis with pepsin

Calvaria were incubated with [^{3}H]proline for 18 min and extracted. To one aliquot of extract in 0.5 M acetic acid was added pepsin at a final concentration of 100 μg/ml. This digest, and a control containing an equal number of counts, were incubated at 15°C for 6 hr. The pH of the two samples was raised to 8.5 with NaOH to inactivate the pepsin, and the samples were dialyzed at 4°C against 0.08 M sodium acetate, pH 4.8.

Separation of proline and hydroxyproline

Samples were hydrolyzed in evacuated tubes in 6 N HCl at 108°C for 24 hr. Proline and hydroxyproline were completely separated by chromatographic elution from a cation-exchange resin (DC 1-A beads, Durrum Chemical Co.) with a complex citric acid–sodium citrate gradient (13). The radioactivity in proline and hydroxyproline was determined by counting carefully measured aliquots of column fractions. Counts were limited to the elution positions of authentic proline and hydroxyproline, and automatic external standardization ratios indicated that no corrections were necessary for differential quenching.

RESULTS

Evidence for a precursor α1-like collagen chain

Acetic acid extracts of calvaria incubated for 9, 18, and 36 min in the presence of [^{3}H]proline contained a collagen fraction with an α1 chain that eluted earlier from CM-cellulose than the α1 chain of carrier rat-skin collagen (Fig. 1). Similar results were obtained with [^{14}C]glycine. The relative amounts of labeled pre-α1‡ and α1, which coeluted with carrier α1, decreased with an increase in labeling time; this suggests conversion of pre-α1 to α1 and an accumulation of α1 in the calvaria. Although rat-bone collagen was not available as a carrier, the increase, with longer labeling periods, in radioactive α1 that coeluted with a carrier of skin α1 excluded tissue-specific differences in the chromatographic properties of the α1 chain. Chromatographic heterogeneity resulting from the oxidative deamination of specific lysyl side chains

‡ The term 'pre-α1' is used to denote an α1-like chain that precedes the α1 chain of collagen in its elution from CM-cellulose and that represents a biosynthetic precursor of the α1 chain.

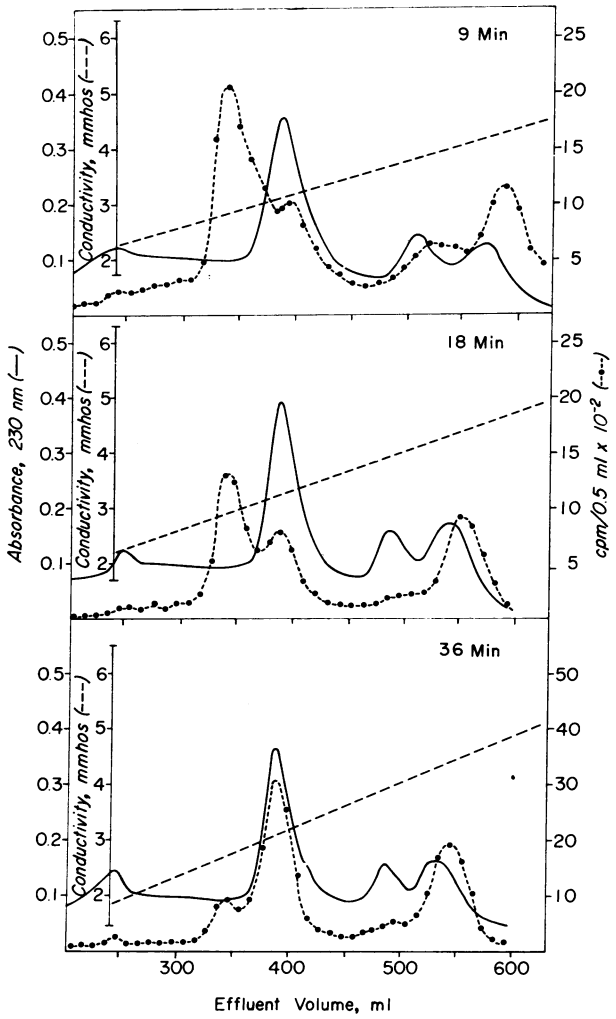

FIG. 1. CM-cellulose elution of acetic acid extracts of rat calvaria incubated with [^{3}H]proline for 9, 18, and 36 min. A constant amount of carrier lathyritic rat-skin collagen was added to each sample.

(14) was avoided by the addition of β-APN to the culture medium (15).

Both pre-α1 and α1 fractions rechromatographed reproducibly; this excludes an artifactual origin for the pre-α1 fraction. Its content of labeled hydroxyproline (see below) conclusively identified the pre-α1 peak as a collagen fraction. To provide additional evidence for the precursor nature of pre-α1, we incubated calvaria in the presence of [^{3}H]proline for 18 min. a large excess of unlabeled L-proline was added, and incubation was continued for another 60 min. The results of this pulse–chase experiment (Fig. 2) indicated that pre-α1 chains synthesized during the initial incubation period were converted to chains that coeluted with carrier α1 during the subsequent hour of incubation. Extracts of bones incubated in the presence of label for 60 min contained a large excess of α1 chains. However, the total number of counts per bone in the pre-α1 fraction was roughly the same as in 18- and 36-min incubations, which suggests a limited pool of pre-α1.

Molecular weight and homogeneity of the pre-α1 fraction

Little or no collagen containing pre-α1 was recovered from calvaria extracted with 1 M NaCl–0.05 M Tris (pH 7.5), a

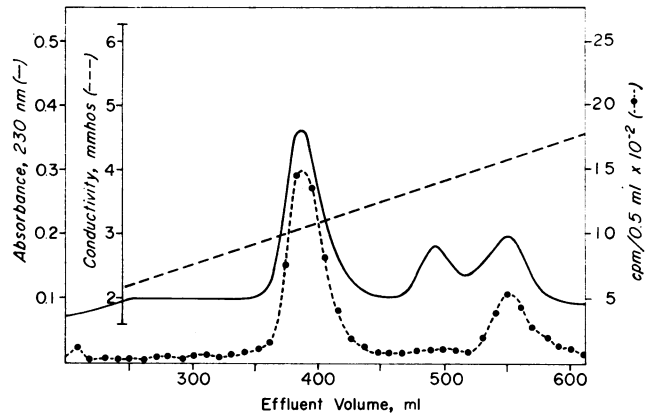

FIG. 2. CM-cellulose elution of an acetic acid extract of rat calvaria incubated with [^{3}H]proline for 18 min, chased with an excess of unlabeled L-proline, and then incubated for another 60 min. Carrier lathyritic rat-skin collagen was added to the sample.

solvent that might be expected to selectively solubilize recently synthesized extracellular collagen. On the other hand, attempts to purify the collagen in acetic acid extracts of calvaria by conventional precipitation and resolubilization procedures were limited by losses during purification. Acetic acid extracts applied to CM-cellulose columns therefore contained, in addition to collagen, noncollagenous proteins and, possibly, nascent chains stripped from ribosomes by homogenization and acetic acid extraction of the tissue. However, the molecular homogeneity of the pre-α1 fraction, suggested by its chromatographic elution profile, was confirmed by acrylamide gel electrophoresis in sodium dodecyl sulfate (SDS). The rechromatographed fraction migrated as a single band, with an R_f of 0.27 (Fig. 3). The R_f of [^{14}C]α-chains of guinea pig collagen (95,000 daltons), used as an internal standard, was 0.33. These data, together with the observed positions of migration of β components and CNBr peptides of known molecular weight, indicate that the pre-α1 chain has a molecular weight of about 120,000.

Limited proteolysis with pepsin

Extracts of calvaria, incubated with [^{3}H]proline for 18 min, were treated with pepsin under conditions that essentially

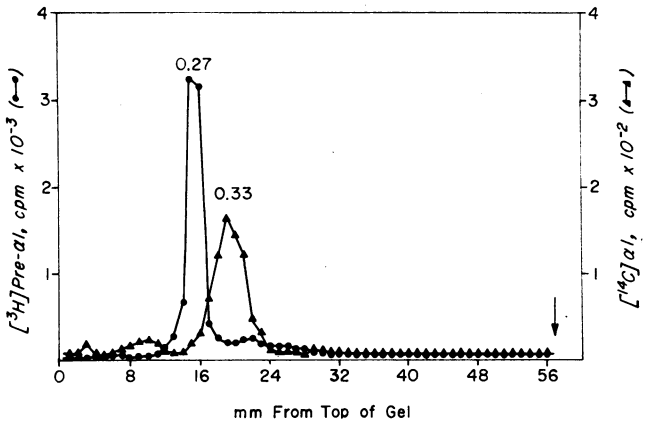

FIG 3. SDS–acrylamide gel electrophoresis of [^{3}H]pre-α1. Guina pig [^{14}C]collagen (in which the label was mainly in the α-chains) was used as an internal standard. The R_f values of [^{3}H]pre-α1 (0.27) and α1 (0.33) are indicated. The *arrow* marks the position of the tracking dye.

limit proteolysis to regions of the collagen molecule that are not triple-helical (16–18). After inactivation of pepsin and dialysis, treated and control preparations were chromatographed on CM-cellulose. The results (Fig. 4) indicate a quantitative conversion of pre-α1 to chains that coelute with α1. Since essentially the same total number of counts eluted in the regions of pre-α1 plus α1 (1.36×10^5; Fig. 4*A*) and α1 (1.28×10^5; Fig. 4*B*), the possibility that the pre-α1 fraction was digested to dialyzable peptides by pepsin, or rendered insoluble during treatment, is excluded. On SDS–acrylamide gel electrophoresis, pepsin-treated pre-α1, rat-calvaria α1, and guinea pig α1 migrated the same distance. These results indicate that the collagen molecule containing pre-α1 is largely in a triple-helical conformation, and that the conversion of pre-α1 to α1, which occurs in the tissue, can be simulated by limited proteolysis with pepsin.

Relative contents of proline and hydroxyproline in biosynthetically labeled collagen chains

The ratio of radioactivity in hydroxyproline and proline in pre-α1, α1, and pepsin-treated pre-α1 are listed in Table 1. This ratio is an accurate measure of the relative contents of the two amino acids in collagen, since both amino acids are derived from the same precursor (prolyl-tRNA) (1, 2), and hydroxylation involves the direct displacement of a hydrogen atom on carbon 4 of proline (19), which in these experiments contains no tritium. The ratio of Hyp/Pro was lower in pre-α1 than in α1, but the additional prolyl residues must be located in the pepsin-susceptible region of the pre-α1 chain since pepsin treatment of procollagen restored this ratio to a value similar to that found in α1 (Table 1). The ratio of Hyp/Pro in rat-bone collagen α1 is higher than that observed in the α1 chain of rat-skin collagen (12), but this is not unexpected in view of the known tissue-specific differences in the degree of hydroxylation of proline (20).

DISCUSSION

It is generally considered that collagen is secreted as a triple-helical molecule (21), although the possibility that individual α-chains associate extracellularly has not been excluded. Since the reconstitution *in vitro* of native collagen molecules from unfolded chains does not proceed readily under physiological conditions (22, 23), chain association *in vivo* may require coordinated biosynthesis of complementary chains on adjacent polysomes or the existence of some form of intracellular template. Another possibility, for which precedence exists in the case of pancreatic proteolytic enzymes (24) and insulin (25), is the existence of a zymogen or secretory form of the protein, characterized by the presence of additional peptide sequences (26, 27). Such peptide moieties may serve to initiate chain association, facilitate transport to the extracellular space, and prevent intracellular aggregation of collagen.

The experiments reported here provide the first clearcut evidence for the existence of a biosynthetic precursor of collagen, procollagen, containing additional peptide sequences that may serve the above functions. The ability of pepsin to convert procollagen to fractions that yield a polypeptide chain resembling α1 in chromatographic properties and molecular weight suggests that proteolytic mechanisms may perform a similar transformation *in vivo*. Since the conditions for enzymatic digestion limit cleavage by pepsin to nonhelical regions of the collagen molecule (16–18), the additional peptide se-

TABLE 1. *Ratios of total activities in hydroxyproline and proline in biosynthetically labeled rat-bone collagen chains*

Chain	Hyp/Pro
Pre-α1	0.81; 0.86
α1	0.98; 0.99
Pepsin-treated pre-α1*	1.07; 0.98
Rat-skin α1	0.83†

* Obtained from pepsin-treated procollagen.
† Determined by amino acid analysis (12).

quences in procollagen presumably exist in some conformation other than the collagen helix. However, the major portion of the procollagen molecule appears to be protected from peptic attack by a triple-helical conformation.

Although the precise region of the extracellular native collagen molecule cleaved by pepsin has not been established, the similarity of the reaction to the limited cleavage by chymotrypsin (18) suggests that the non-helical NH_2-terminal regions of collagen α-chains are the primary site of enzymatic attack. Location of the additional peptide sequences in procollagen at the NH_2-termini would permit the function of these sequences to be expressed prior to completion of the chains. This position is also favored by the finding that both α1 and α2 chains from all collagens thus far examined contain a short sequence (12–20 amino acids) at the NH_2-terminal end of the chain that differs in primary structure from the helical region of collagen (28, 29). This NH_2-terminal region in α1 may represent the COOH-terminal end of the additional sequence present in pre-α1. The finding that rat-skin collagen α1 chain lacks a tetrapeptide present at the NH_2-terminal end of the otherwise identical rat-tendon α1 chain (30), and a similar heterogeneity at the NH_2-terminal end of the chicken-skin collagen α1 chain (31), may reflect variable proteolysis during the conversion of procollagen to collagen.

Many of the properties of the pre-α1 chain, such as its tendency to adsorb to chromatographic media and glass surfaces, and the difficulties encountered in purification of procollagen, resemble those observed with underhydroxylated (protocollagen) chains (32). It seems likely that such properties of protocollagen, attributed to underhydroxylation, result from the fact that collagen synthesized in the presence of inhibitors of collagen proline hydroxylase contains underhydroxylated pre-α1 chains. The existence of the pre-α1 fraction studied in these experiments cannot, however, be accounted for by such underhydroxylation since chains obtained from pepsin-treated procollagen contained a Hyp/Pro ratio that was the same as, or slightly higher than, that determined for α1 chains in rat calvaria (Table 1). The possible relation of pre-α1 to the α1(II) chain of cartilage collagen (33) was also considered. Although the $[\alpha 1(II)]_3$ molecule may conceivably function in tissues other than cartilage, the time-dependent conversion of pre-α1 to an α1-like chain effectively excludes its identity with α1(II).

In a recent study, Vuust and Piez (34) used pulse-labeling of rat calvaria in culture to study the biosynthesis of collagen chains. The failure of these investigators to detect a procollagen fraction may be due to their method of extraction and purification of the recently synthesized collagen. In analyses of the specific activities of CNBr-produced peptides of pulse-labeled [^{14}C]α1 chains, Vuust and Piez attributed the anomalously high relative specific activity of the NH_2-terminal

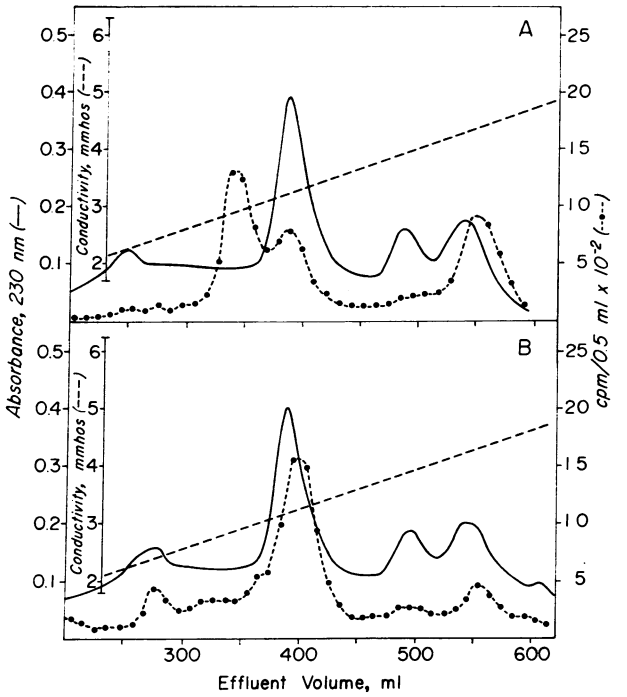

FIG. 4. CM-cellulose elution of control (*A*) and pepsin-treated (*B*) acetic acid extracts of rat calvaria incubated with [^{3}H]proline for 18 min. A constant amount of carrier lathyritic collagen was added to each sample.

peptide, α1-CB1, to preferential destruction of the ^{3}H-labeled peptide used as an internal standard. An alternative explanation is that the pepsin-susceptible portion of pre-α1 (of which α1-CB1 may represent the COOH-terminal end) is synthesized as a separate polypeptide and that linkage of this chain with the α1 chain occurs after its synthesis.

NOTE ADDED IN PROOF

D. L. Layman, E. B. McGoodwin and G. R. Martin (*Proc. Nat. Acad. Sci. USA*, **68**, 454–8, 1971) have recently reported that human fibroblasts in culture secrete an unusual form of collagen into the medium. This protein could also be converted to a collagen-like molecule by limited cleavage with pepsin. The relationship of this protein to procollagen will require further study.

This work was supported in part by National Institutes of Health Grant AM 11248. G. B. was a Predoctoral Fellow of the National Institutes of Health.

1. Udenfriend, S., *Science*, **152**, 1335 (1966).
2. Juva, K., and D. J. Prockop, *J. Biol. Chem.*, **241**, 4419 (1969).
3. Butler, W. T., and L. W. Cunningham, *J. Biol. Chem.*, **241**, 3882 (1966).
4. Spiro, R. G., *J. Biol. Chem.*, **244**, 602 (1969).
5. von Hippel, P. H., in *Treatise on Collagen*, ed. G. N. Ramachandran (Academic Press, 1967), Vol. 1, p. 253.
6. Ross, R., in *Treatise on Collagen*, ed. B. S. Gould (Academic Press, 1968), Vol. 2A, p. 1.
7. Hodge, A. J., and F. O. Schmitt, *Proc. Nat. Acad. Sci. USA*, **46**, 186 (1960).
8. Piez, K. A., *Annu. Rev. Biochem.*, **37**, 547 (1968).
9. Piez, K. A., E. A. Eigner, and M. S. Lewis, *Biochemistry*, **2**, 58 (1963).
10. Orekhovitch, V. N. Tustanowski, A. A. Orekhovitch, K. D. and Plotnikuva, N. E., *Biokhimiya*, **13**, 55 (1948).

11. Weber, K., and M. Osborn, *J. Biol. Chem.*, **244**, 4406 (1969).
12. Butler, W. T., K. A. Piez, and P. Bornstein, *Biochemistry*, **6**, 3771 (1967).
13. Miller, E. J., and K. A. Piez, *Anal. Biochem.*, **16**, 320 (1966).
14. Piez, K. A., G. R. Martin, A. H. Kang, and P. Bornstein, *Biochemistry*, **5**, 3813 (1966).
15. Siegel, R. C., S. R. Pinnell, and G. R. Martin, *Biochemistry*, **9**, 4486 (1970).
16. Rubin, A. L., M. P. Drake, P. F. Davison, D. Pfahl, P. T. Speakman, and F. O. Schmitt, *Biochemistry*, **4**, 181 (1965).
17. Kuhn, K., P. Fietzek, and J. Kuhn, *Biochem. Z.*, **344**, 418 (1966).
18. Bornstein, P., A. H. Kang, and K. A. Piez, *Biochemistry*, **5**, 3803 (1966).
19. Fujita, Y., A. Gottlieb, B. Peterkofsky, S. Udenfriend, and B. Witkop, *J. Amer. Chem. Soc.*, **86**, 4709 (1964).
20. Bornstein, P., *Biochemistry*, **6**, 3082 (1967).
21. Gould, B. S., in *Treatise on Collagen*, ed. B. S. Gould (Academic Press, 1968), Vol. 2A, p. 139.
22. Beier, G., and J. Engel, *Biochemistry*, **5**, 2744 (1966).
23. Harrington, W. F., and N. V. Rao, *Biochemistry*, **9**, 3714 (1970).
24. Neurath, H., K. A. Walsh, and W. P. Winter, *Science*, **158**, 1638 (1967).
25. Steiner, D. F., D. Cunningham, L. Spigelman, and B. Aten, *Science*, **157**, 697 (1967).
26. Schmitt, F. O., *Bull. N.Y. Acad. Med.*, **36**, 725 (1960).
27. Speakman, P. T., *Nature*, **229**, 241 (1971).
28. Piez, K. A., H. A. Bladen, J. M. Lane, E. J. Miller, P. Bornstein, W. T. Butler, and A. H. Kang, *Brookhaven Symp. Biol.*, **21**, 345 (1968).
29. Bornstein, P., and A. H. Kang, in *Chemistry and Biology of the Intercellular Matrix*, ed. E. A. Balazs (Academic Press, 1970), Vol. 1, p. 99.
30. Bornstein, P., *Biochemistry*, **8**, 63 (1969).
31. Kang, A. H., K. A. Piez, and J. Gross, *Biochemistry*, **8**, 3648 (1969).
32. Lukens, L. N., *J. Biol. Chem.*, **245**, 453 (1970).
33. Miller, E. J., and V. J. Matukas, *Proc. Nat. Acad. Sci. USA*, **64**, 1264 (1969).
34. Vuust, J., and K. A. Piez, *J. Biol. Chem.*, **245**, 6201 (1970).