Proc. Nat. Acad. Sci. USA
Vol. 70, No. 2, pp. 361–365, February 1973

Secretion and Extracellular Processing of Procollagen by Cultured Human Fibroblasts

(isotopic labeling/pulse-chase/pepsin digestion/gel electrophoresis)

BURTON GOLDBERG AND CHARLES J. SHERR

Department of Pathology, New York University Medical Center, New York, N.Y. 10016

Communicated by H. Sherwood Lawrence, November 27, 1972

ABSTRACT **Cultures of human diploid fibroblasts were labeled with radioactive proline and glycine, and the precursor of collagen (procollagen) in cells and medium was characterized by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. A covalently assembled molecule with the composition (pro $\alpha 1)_2 \cdot$ pro $\alpha 2$ (approximate molecular weight, 360,000) appeared intracellularly soon after synthesis of the constituent chains, and could be detected in the medium after 60 min of labeling. The molecule was stabilized by disulfide bonds between cysteine residues in the amino-terminal procollagen peptide sequences of the three chains. Collagenase digested the molecule to peptides of 30,000 molecular weight or less. Limited digestion with pepsin excised nonhelical procollagen peptides, yielding native, triple-helical tropocollagen. Pulse-chase experiments indicated that a peptidase in the medium sequentially excised the nonhelical peptides from the molecule, generating tropocollagen molecules that aggregated as fibers in the cell layer. The excised, nonhelical procollagen peptides contained little or no proline or glycine. Intramolecular bonds of the lysyl aldehyde type were not detected in the secreted molecule, as reduction of the medium always resulted in quantitative recovery of free pro α chains in dodecyl sulfate–urea. Lysyl-derived, covalent bonds appeared to form between tropocollagen molecules aggregating in the cell layer. We suggest the term "pro-tropocollagen" for the assembled, secreted precursor of collagen.**

Tropocollagen, the native molecule extracted from extracellular collagen fibers, is generally composed of two identical $\alpha 1$ chains and an $\alpha 2$ chain in triple helical assembly. Molecules with additional N-terminal peptides on the $\alpha 1$ and $\alpha 2$ chains ("procollagens," "pro $\alpha 1$," "pro $\alpha 2$") have been identified and characterized as biosynthetic precursors of the native chains (1–3). It has not been clearly established if the pro α chains are assembled before or after secretion from the cell. Only free procollagen chains were isolated from bone culture systems (4), but the studies of Dehm *et al.* with cultured tendon cells indicated that secreted procollagen chains were associated through disulfide bonds, and a three-chain structure was proposed (5).

We have studied the soluble forms of collagen secreted into the medium of cultured human fibroblasts by chromatography and analytical gel electrophoresis, and have presented evidence that the soluble, secreted form of collagen is a disulfide-stabilized trimer of three procollagen chains with the composition (pro $\alpha 1)_2 \cdot$ pro $\alpha 2$ (6). Other molecular species identified in the medium were interpreted as intermediates formed by the action of an extracellular, procollagen peptidase (7) on the assembled, three-chain molecule. We now present further data on the intracellular assembly, secretion, and extracellular processing of the procollagen trimer.

Abbreviations: SDS, sodium dodecyl sulfate; CMC, carboxymethyl cellulose.

MATERIALS AND METHODS

Cells and Culture Conditions. All experiments were performed with confluent cultures of a normal human diploid fibroblast strain, CRL 1121, obtained from the American Type Culture Collection. Culture conditions, isotopic labeling, and processing of medium and cell layers for gel electrophoresis were as described (6). For experiments in which intracellular forms of collagen were studied, cells were detached from washed cell layers with EDTA [0.5 mM in phosphate-buffered saline (pH 7.2)] and collected by centrifugation at 600 $\times$ *g* for 10 min. Cells were washed twice in phosphate-buffered saline, and procollagen was isolated (6).

SDS–Acrylamide Gel Electrophoresis. Gels were 5% in acrylamide and 0.07% in methylenebisacrylamide (Eastman). They were polymerized in the running buffer: 0.1 M phosphate (pH 7.0) containing 0.1% SDS and 0.5 M urea. Authentic ^{14}C-labeled α chains were added to the samples as internal markers, and sample application and electrophoresis were as described (6). Some samples were reduced with 1% 2-mercaptoethanol, heat denatured, and then applied to gels.

Limited Pepsin Digestion. Lyophilized medium from cultures labeled for 24 hr with [^{3}H]proline and [^{3}H]glycine was dissolved in 0.5 M acetic acid, and two equal aliquots were taken. One sample received 100 μg of pepsin (Worthington, twice crystallized); both samples (final volumes 1.0 ml) were incubated for 5 hr at 15°. The pH was brought to 8.5 with sodium hydroxide to inactivate the pepsin and the samples were dialyzed against distilled water for 18 hr and lyophilized. Samples were heat denatured in sodium dodecyl sulfate–urea–phosphate, and the total radioactivity recovered was measured before application to gels.

Collagenase Digestion. Processed medium from cultures labeled for 24 hr with radioactive proline and glycine was dissolved in 50 mM Tris·HCl (pH 7.5) containing 140 mM NaCl and 1 mM $CaCl_2$. The sample was incubated with purified collagenase (Worthington CLSPA, 100 μg/ml) for 16 hr at 37°. SDS was added to 1%. The sample was heated at 60° for 30 min and dialyzed against the running buffer for gel electrophoresis. Radioactivity in the retentate was measured and applied to a gel.

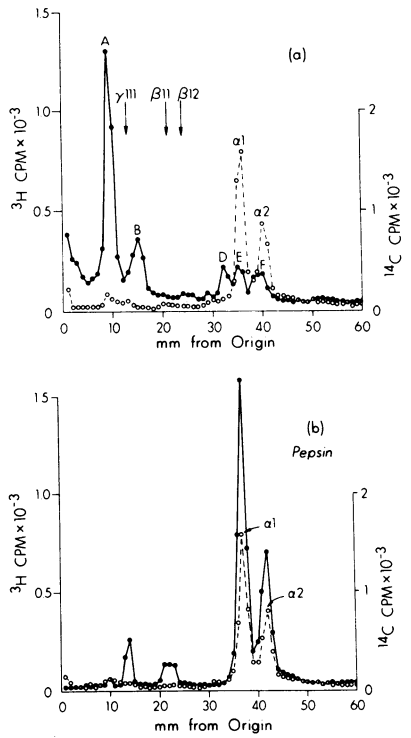

FIG. 1. Limited digestion of medium with pepsin. Gel electrophoresis of (*a*) control sample, and (*b*) pepsin-digested sample. Equal amounts of ^{3}H radioactivity were applied to gels. ^{3}H cpm, (●——●); ^{14}C cpm, α chain markers (O- - -O).

Pulse-Chase Experiments. Medium from cultures labeled for 24 hr with [^{3}H]proline and [^{3}H]glycine was supplemented with sodium ascorbate (75 μg/ml) and a 20,000-fold excess of unlabeled proline and glycine, and placed on unlabeled, replicate cell layers. At intervals, the medium and cell layers were processed as described (6). Total radioactivity recovered was measured, and the samples were applied to gels. Addition of fetal-calf serum to the chase medium did not affect the results.

RESULTS

Enzymatic digestion of procollagen

We have shown that the procollagen trimer has the composition (pro α1)$_2$·pro α2 and that disulfide bonds between cysteine residues in pro α1 and pro α2 chains help stabilize the molecule (6). Extracellular processing of the assembled molecule presumably proceeds through the action of a secreted peptidase, which excises the nonhelical procollagen peptides, generating native, triple-helical tropocollagen. To demonstrate helical and nonhelical regions in the procollagen molecule and to test their susceptibility to enzymatic attack, we performed an incubation with the endopeptidase, pepsin, under conditions in which native helical tropocollagen is not digested (8). The percentage of radioactivity recovered after incubation and dialysis was about 75% for both control and enzyme-treated samples, indicating that products of digestion were either mostly nondialyzable, or if dialyzable, contained relatively little radioactive proline and glycine. Equal amounts of radioactivity were applied to gels *a* and *b* of Fig. 1. The gel pattern of the control medium (Fig. 1*a*) is identical to that previously published for cultures labeled for 24 hr (6). Peak *A* is the (pro α1)$_2$·pro α2 trimer of molecular weight 360,000, and peaks *B*, *D*, *E*, and *F* are identified as molecular species released by SDS–urea from digestion intermediates of the assembled molecule. After pepsin digestion (Fig. 1*b*), virtually all the radioactivity was recovered in two peaks with the same mobilities as the marker α1 and α2 chains, and in the same 2/1 ratio. This result indicates that: (*i*) All the molecular species detected in the medium after 24 hr of labeling are forms of procollagen, and together they contain twice as many pro α1 chains as pro α2 chains. (*ii*) The procollagen peptides are sensitive to pepsin digestion, but the remaining portion of the molecule is resistant, probably because of triple helical assembly.

When labeled medium was incubated with collagenase, 80% of the radioactivity became dialyzable, and the remainder was identified in gels as peptides with molecular weights of less than 30,000. This result confirms that all the radioactive molecules are collagen proteins.

Intracellular assembly and secretion of procollagen

For substantiation that the procollagen trimer was the principal secreted form of collagen, replicate cultures were labeled with [^{3}H]proline and [^{3}H]glycine for up to 2 hr. At 20-min intervals, the medium and extracts of the cells were analyzed for amounts and molecular forms of nondialyzable radioactivity. Although incorporation of radioactivity into the cells was linear for the first 60 min, labeled

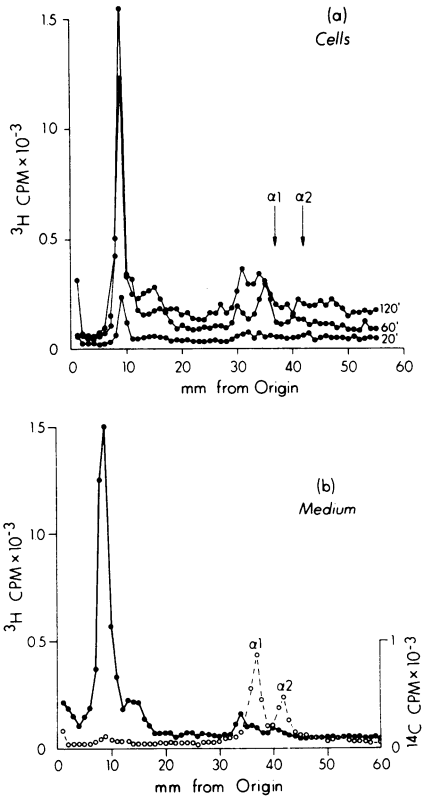

FIG. 2. Intracellular and secreted forms of collagen. Replicate cultures (4 × 10^6 cells per plate) were labeled with 50 μCi each of [^{3}H]proline and [^{3}H]glycine. Two plates were harvested every 20 min. (*a*) Cells were detached with EDTA, extracted with acid, and processed. (*b*) Processed medium after 2 hr of labeling. ^{3}H cpm, (●——●); ^{14}C cpm, α chain markers (O- - -O).

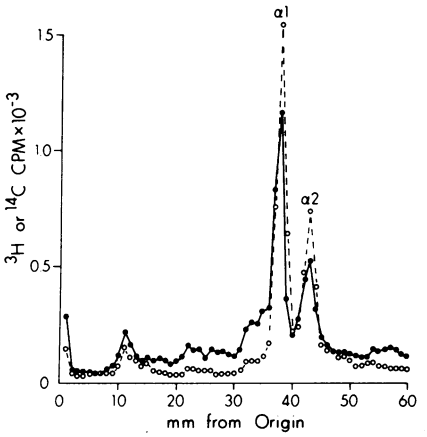

FIG. 3. Acid extract of cell layers after 24 hr of chase. See Fig. 2 for labeling conditions. After 24 hr of labeling, supplemented chase medium was placed on unlabeled cell layers. 24 Hr later, washed cell layers were extracted with acid and processed for gel electrophoresis. ^{3}H cpm, (●——●); ^{14}C cpm, α chain markers (O- - -O).

proteins did not appear in significant amounts in the medium until after 1 hr. Fig. 2*a* shows the gel pattern of protein extracted from the cells at 20, 60, and 120 min. Qualitatively similar patterns were obtained at intervening time points. Most of the radioactivity that accumulates intracellularly is in the procollagen trimer, and much smaller amounts of radioactivity are in molecular species running just behind the α chain markers. Medium was analyzed after 60 min of labeling, and the gel pattern of Fig. 2*b* is representative. Again, most of the radioactivity is in the procollagen trimer. When samples from both cells and medium were reduced and run on gels, the trimer disappeared and the radioactivity was quantitatively recovered in free pro α chains (not shown). We conclude that the disulfide-stabilized (pro α1)$_2$·pro α2 molecule is assembled intracellularly soon after synthesis of the procollagen chains, and it is the major form of collagen secreted into the medium.

Conversion of procollagen in medium to collagen fibers in the cell layer

To follow the processing of secreted procollagen to native tropocollagen molecules, we performed chase experiments. Medium from cultures labeled for 24 hr was diluted with excess unlabeled proline and glycine and transferred to unlabeled replicate plates. At 5, 24, 48, and 72 hr, the radioactivity in the medium and extractable from the cell layer was measured, and samples were taken for gel electrophoresis. Figs. 3 and 4 are typical of gel patterns obtained from cell layer extracts and medium in three independent experiments. For the experiment shown, 42% of the original radioactivity in the medium was chased into the cell layer in 72 hr. This relatively slow processing might result from dilution of intermediates in the large extracellular volume. Fig. 3 shows that the radioactivity chased into the cell layer was recovered in SDS–urea as α chains in the characteristic 2/1 ratio, indicating that native tropocollagen molecules were aggregating

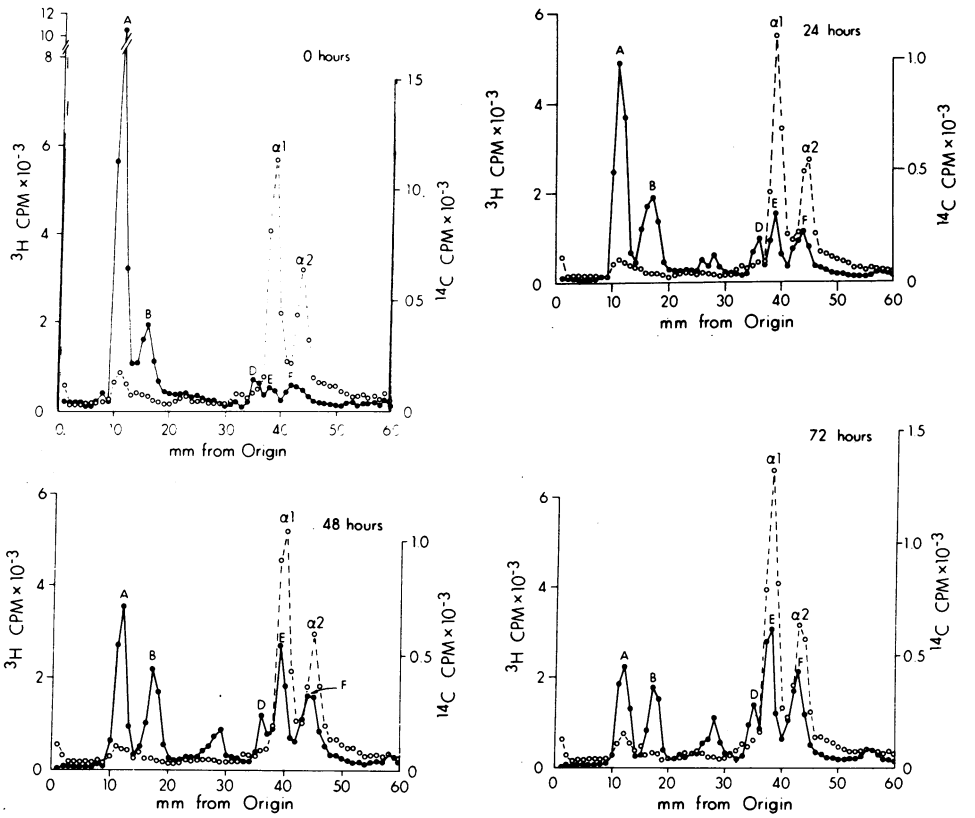

FIG. 4. Gel patterns of medium during 72-hr chase. Labeling conditions were as for Fig. 2. After 24 hr of labeling, medium from one plate was processed directly (0 hr). Medium from three plates was supplemented with unlabeled proline and glycine and placed on three unlabeled cell layers. Medium was removed at indicated times and processed. Equal amounts of ^{3}H radioactivity were applied to gels. ^{3}H cpm, (●——●); ^{14}C cpm, α chain markers (O- - -O).

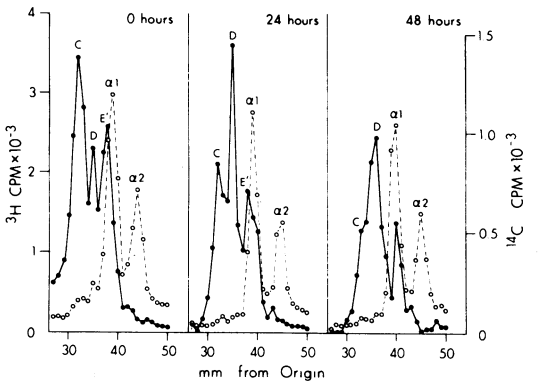

FIG. 5. Difference plots of reduced chase medium. Chase media from experiment of Fig. 4 were reduced with 2-mercaptoethanol, and the same amount of ^3H radioactivity as in Fig. 4 was applied to each gel. ^3H radioactivities in peaks *D*, *E*, and *F* of unreduced samples were subtracted from radioactivities in the same peaks in the respective reduced samples. α Chain markers were coincident for the respective unreduced and reduced samples. ^3H cpm by difference, (●——●); ^{14}C cpm, α chain markers (O- - -O).

as extracellular collagen fibers in the cell layer (9). Similar gel patterns were obtained at 48 and 72 hr. At 24 hr only 40% of the radioactivity in the cell layers was acid-extractable, and by 72 hr the value fell to 20%. We ascribe this decreasing solubility to the formation of covalent, lysyl-derived aldehydic crosslinks between the tropocollagen molecules in the cell layer (10, 11), for after addition of β-aminopropionitrile to the cultures, 80% of the radioactivity was acid-extractable at all time points.

The gel patterns of the medium (Fig. 4) showed that about 70% of the radioactivity in the trimer (peak *A*) moved into other molecular species in 72 hr. Throughout the chase period, the ratio of radioactivity in peak *A* to radioactivity in peak *B* diminished, and radioactivity in peaks *E* and *F* increased. By our interpretation (see Fig. 6), the cells secrete a peptidase that makes stepwise excisions of the procollagen segment of the assembled three-chain molecule. Initially, the disulfide linkage between pro α2 and pro α1 chains is excised, and SDS–urea releases a pro α1 dimer (peak *B*) and a free, partially digested pro α2 chain (peak *D*). Further excisions along the pro α1 chains remove remaining disulfide bonds, and then only noncovalent bonds maintain chain assembly. Eventually only a few aminoacid residues are left N-terminal to the native helical assembly, and, thus, peaks *E* and *F* are interpreted as almost completely processed pro α1 and pro α2 chains, respectively.

The ratio (cpm in all pro α1 species)/(cpm in all pro α2 species) in the medium should approximate 2.0 if we have correctly deduced the chain composition of each gel peak, and if the procollagen peptides contain little radioactive proline and glycine relative to the rest of the chains. Ratios for each time point in three chase experiments were calculated as follows: cpm in peak *A* = $^2/_3$ pro α1, $^1/_3$ pro α2; cpm in peaks *B* and *E* = all pro α1; cpm in peaks *D* and *F* = all pro α2. The calculated average ratio was 1.9 ± 0.1 (SD), consistent with the above assumptions, the results from pepsin digestion of the medium, and with our model for procollagen processing.

There was no evidence for significant formation of covalent crosslinks of the lysyl aldehyde type in the soluble procollagen molecule, as reduction with 2-mercaptoethanol at all time points of the chase caused the complete disappearance of peaks *A* and *B*, with recovery of radioactivity in free pro α chains. The gel patterns of the reduced samples did not, however, unambiguously show free pro α1 and pro α2 chains in a 2/1 ratio. Reduction and SDS–urea released pro α chains of about 120,000 molecular weight from the trimer and lower molecular weight chains from digestion intermediates. Since pro α2 chains move with faster mobilities in SDS gels than pro α1 chains of the same molecular weight (6), any peak obtained by reduction could contain both chain classes, with the pro α1 component being of lower molecular weight. This result can be shown by subtraction of the radioactivity in peaks *D*, *E*, and *F* of the samples from the chase experiment (Fig. 4) from the radioactivity appearing in the same regions of the gels after reduction of the samples. Such "difference plots" (Fig. 5) only show the radioactivity in pro α chains released by reduction of the trimer (peak *A*) and the pro α1 dimer (peak *B*), and reflect partial digestion of disulfide-assembled chains throughout the chase period. With time, radioactivity in peak *C* (complete pro α1 chains) is chased into the region (peak *D*) where pro α2 chains released from the trimer are usually identified. This result shows that as pro α1 chains are shortened, they acquire mobilities similar to higher molecular weight pro α2 chains. An additional peak *E'* with a mobility close to marker α1 chains is also seen, and is interpreted as an almost completely processed pro α1 chain. These data make it clear that reduction of the medium can yield peaks that are heterogeneous for the two chain classes, thus obscuring the true stoichiometric ratio of the pro α components in the procollagen molecule.

DISCUSSION

Fig. 6 schematically summarizes our results. A procollagen molecule is assembled intracellularly soon after synthesis of the constituent pro α chains. It has the composition (pro α1)$_2$· pro α2, the individual chains having molecular weights of about 120,000. The trimer is stabilized by disulfide bonds between nonhelical pro α peptides, and a triple helix is formed

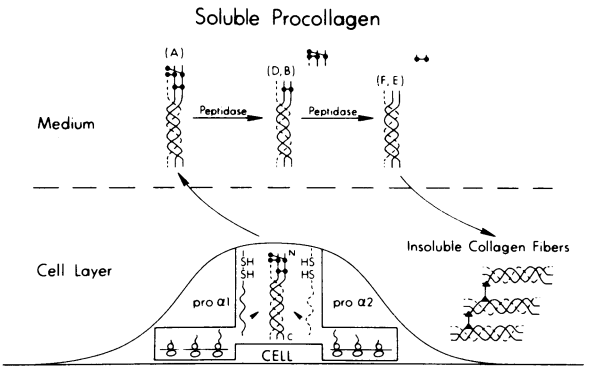

FIG. 6. Pro-tropocollagen assembly, secretion, and extracellular processing. The chains are assembled after synthesis on polyribosomes. *N* and *C* denote amino- and carboxy-terminal ends of the assembled molecule. (●——●), Indicates disulfide bonds between cysteine residues in the chains. Letters *A*, *D*, *B*, *F*, and *E* above the depicted molecules correspond to peaks generated in sodium dodecyl sulfate gels. (▲——▲), Indicates lysyl-derived, covalent crosslinks between tropocollagen molecules.

by noncovalent interactions between other portions of the chains. The molecule is secreted about 1 hr after synthesis and is the only secreted precursor form of collagen. The cells presumably secrete a peptidase that makes stepwise excisions of the nonhelical peptides until tropocollagen molecules are generated. The latter aggregate in the cell layer as fibers, which are stabilized by lysyl-derived, covalent crosslinks. We propose the name "pro-tropocollagen" for the secreted form of collagen, signifying that it is an assembled precursor of native tropocollagen.

Fig. 6 also indicates how the various gel peaks are generated from the medium and cell layer. The chain compositions of the different gel peaks were deduced from the relative mobilities and radioactivities of the molecular species before and after reduction with 2-mercaptoethanol, but the analysis is complicated by the anomalous mobility of pro $\alpha 2$ components with respect to pro $\alpha 1$ chains (12) and the overlapping mobilities of the two chain classes as they undergo enzymatic excisions. However, the model is supported by other kinds of data. For example, the model predicts a ratio of 2/1 for pro $\alpha 1$ components to pro $\alpha 2$ components in the medium, independent of the duration of isotope incorporation or chase. An average ratio of 1.9 was calculated from the data of several such experiments, confirming the prediction, as well as our deductions as to chain compositions of the different gel peaks. The results of digestion with pepsin and collagenase confirmed that virtually all the molecular species identified in the medium were procollagen components. Limited pepsin digestion also confirmed that the molecule had the composition (pro $\alpha 1)_2 \cdot$ pro $\alpha 2$ and that only nonhelical procollagen peptides were excised. Based on the recovery of radioactivity after digestion, the excised sequences appeared to contain relatively little proline and glycine. We have identified low molecular weight peptides that are cysteine-rich and glycine- and proline-poor, in the medium (6), and the chase and pepsin data now support our assumption that these are the excised N-terminal peptide sequences. Independent corroboration of this point is given by the aminoacid analysis of the pro $\alpha 1$ chain of chick calvaria published by Bornstein *et al.* (4).

Our data complement the observations of Layman *et al.* (11), who gave evidence for a molecule with a molecular weight greater than α chains, containing intramolecular crosslinks not derived from lysyl aldehydes, and containing triple-helical collagen structure. However, these authors reported that when heat-denatured medium from fibroblast cultures was chromatographed on carboxymethyl cellulose (CMC) at pH 4.8 with a 0.1 M NaCl gradient, only 20% of the labeled protein was eluted; the rest could be released from the resin only by addition of alkali and 1 M NaCl. The material that eluted at low salt concentration was identified as free α chains and, based on the recoveries, the authors postulated that $\alpha 2$ chains were synthesized in excess, and that free α chains accumulated in small amounts in the medium. By first separating procollagens on agarose and then applying them to gels and CMC, we have determined that only partially digested procollagen chains released from the trimer by denaturation (our gel peaks *D*, *E*, and *F*) are eluted from CMC at acid pH with the standard salt gradients. The covalently assembled trimer and dimer are eluted only when alkali and 1 M NaCl are added. We interpret all the data to mean that: (*i*) The majority of procollagen in the medium is in a disulfide stabilized form; (*ii*) any free chains recovered from the medium have been released by denaturation of an assembled, three-chain procollagen molecule; (*iii*) the chains released by denaturation have some or most of their procollagen sequences excised; and (*iv*) the cell secretes the two chain classes in the expected 2/1 ratio.

Disulfide-assembled procollagen chains have not been identified to date in acid extracts of bone organ cultures, but CMC chromatography was again used for purification. It was reported (4) that only 50% of the applied material could be eluted from the resin, and while free pro $\alpha 1$ and pro $\alpha 2$ chains were eluted, they might have been enzymatically processed chains lacking some portion of the procollagen peptide sequence. Thus, the calvarial pro $\alpha 2$ chains eluted from CMC contained little or no [^{35}S]cysteine; in contrast, the pro $\alpha 2$ chains released from our gel peak *A* by 2-mercaptoethanol contained radioactive cysteine residues (6). In addition, the reported analyses from bone cultures were performed on extracts of the bone, but our data suggest that a disulfide-assembled procollagen molecule is more likely to be recovered from the medium. For these reasons, we believe that the data from the two culture systems are not necessarily in conflict, and the manner of assembly and processing of procollagen might prove to be similar for most tissues.

The expert assistance of Sheila Heitner and Josa Williamson is gratefully acknowledged. This work was supported by NIH Grant AM 15656.

1. Bellamy, G. & Bornstein, P. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 1138–1142.
2. Müller, P. K., McGoodwin, E. & Martin, G. R. (1971) *Biochem. Biophys. Res. Commun.* **44**, 110–117.
3. Ehrlich, H. P. & Bornstein, P. (1972) *Biochem. Biophys. Res. Commun.* **46**, 1750–1756.
4. Bornstein, P., Von Der Mark, K., Wyke, A. W., Ehrlich, H. P. & Monson, J. M. (1972) *J. Biol. Chem.* **247**, 2808–2813.
5. Dehm, P., Jimenez, S. A., Olsen, B. R. & Prockop, D. J. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 60–64.
6. Goldberg, B., Epstein, E. H., Jr. & Sherr, C. J. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 3655–3659.
7. Bornstein, P., Ehrlich, H. P. & Wyke, A. W. (1972) *Science* **175**, 544–546.
8. Rubin, A. L., Pfahl, D., Speakman, P. T., Davison, P. F. & Schmitt, F. O. (1963) *Science* **139**, 37–39.
9. Goldberg, B. & Green, H. (1964) *J. Cell Biol.* **22**, 227–258.
10. Pinnell, S. R. & Martin, G. R. (1968) *Proc. Nat. Acad. Sci. USA* **61**, 708–716.
11. Layman, D. L., McGoodwin, E. B. & Martin, G. R. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 454–458.
12. Furthmayr, H. & Timpl, R. (1971) *Anal. Biochem.* **41**, 510–516.