Precursors of Collagen Secreted by Cultured Human Fibroblasts

(isotopic labeling/chromatography/gel electrophoresis)

BURTON GOLDBERG*, ERVIN H. EPSTEIN, JR., AND CHARLES J. SHERR

Departments of Pathology and Dermatology, New York University Medical Center, New York, N.Y. 10016

Communicated by Michael Heidelberger, September 28, 1972

ABSTRACT Soluble precursors of collagen ("procollagens") in the culture medium of human diploid fibroblasts were characterized by molecular sieve chromatography on 6% agarose and by sodium dodecyl sulfate-acrylamide gel electrophoresis. Under the denaturing conditions for gels, the following molecular species were identified: a major species of three procollagen chains, a procollagen dimer, and free procollagen chains. Reduction with 2-mercaptoethanol completely dissociated the trimer and dimer, releasing procollagen $\alpha 1$ and procollagen $\alpha 2$ chains with molecular weights of about 120,000. Both pro $\alpha 1$ and pro $\alpha 2$ chains could be labeled with [35 S]cystine. In addition, free procollagen chains with molecular weights of less than 120,000, but heavier than native α chains, were identified in gels. Radioactivity in all of the above molecules could be chased into native α chains extracted from extracellular fibers of the cell layers. The data indicate that: (i) collagen is secreted as a disulfide-stabilized procollagen trimer with the composition, (pro $\alpha 1$)₂·pro $\alpha 2$; (ii) noncovalent interactions maintain the three-chain assembly during stepwise enzymatic excisions of N-terminal, cysteine-containing procollagen peptides; and (iii) after such excisions, native helical molecules, ($\alpha 1$)₂· $\alpha 2$, precipitate as fibers in the cell layers.

Collagen is secreted from fibroblasts in a soluble form and precipitates as characteristic fibers in the extracellular space. Native tropocollagen molecules isolated from such extracellular fibers generally consist of two identical $\alpha 1$ chains and one $\alpha 2$ chain in triple helical assembly. *In vitro*, correct reassembly of the molecule from separated α chains proceeds slowly and inefficiently, and assembled helical molecules tend to come out of solution to form typical collagen fibers. For these reasons it has been proposed that rapid *in vivo* assembly of collagen molecules and secretion in a soluble form are determined by extra terminal peptides that are later excised (1, 2). Supporting this view are recent reports that biosynthetic precursors of both $\alpha 1$ and $\alpha 2$ chains ("procollagens": "pro $\alpha 1$ ", and "pro $\alpha 2$ ") may be extracted from embryonic rat (3) and chick (4) calvaria, and may be found in the medium of cultured fibroblasts (5) and cultured embryonic-chick tendon cells (6, 7). The procollagens have been characterized as native $\alpha 1$ and $\alpha 2$ chains with additional terminal peptide sequences of molecular weight about 15,000-20,000.

Noncovalently assembled chains might be stabilized by formation of disulfide bonds between cysteine residues in these extra peptide sequences. In bone culture preparations pro $\alpha 1$ contains 5 or 6 half-cystine residues per chain, but little or no cysteine was detected in pro $\alpha 2$, and no evidence for disulfide-

bond formation between procollagen chains was obtained (2). In contrast, in preparations of cultured fibroblasts (5, 8) and tendon cells (7), data from molecular sieve chromatography gave evidence for association of procollagen chains. In particular, the results of Dehm *et al.* (7) raised the possibility that three procollagen chains might be assembled through disulfide bonds.

We have used molecular sieve chromatography and sodium dodecyl sulfate (SDS)-acrylamide gel electrophoresis to characterize the soluble procollagens in the medium of cultured human fibroblasts. From these studies, we propose that a trimer of pro α chains, (pro $\alpha 1$)₂·pro $\alpha 2$, is the only form of collagen secreted by these cells and suggest a detailed model for extracellular conversion of the trimer to native collagen fibers.

MATERIALS AND METHODS

Cells and Culture Conditions. All experiments were performed with a normal human diploid fibroblast strain, CRL 1121, obtained from the American Type Culture Collection. The cells were grown to confluence in plastic petri dishes in Dulbecco-Vogt medium supplemented with 10% fetal-calf serum and sodium ascorbate (75 μ g/ml). Before addition of isotope, the medium was replaced with serum-free medium supplemented with ascorbate. L-[14 C]proline (260 Ci/mol), [14 C]glycine (91 Ci/mol), L-[3,4- 3 H]proline (29.8 Ci/mmol), L-[2,3- 3 H]proline (45.7 Ci/mmol), [2- 3 H]glycine (11.1 Ci/mmol) (from New England Nuclear), and L-[35 S]cystine (69.5 Ci/mol) (Amersham/Searle) were added to the cultures in 0.1-0.3 ml of phosphate-buffered saline. After the appropriate labeling periods, the plates were chilled, and subsequent manipulations were performed at 4°. The medium was removed, made 0.5 M in acetic acid, and centrifuged at 12,000 $\times g$ for 10 min. The supernatant was dialyzed for 12 hr against 0.5 M acetic acid, and for 48 hr against distilled water before lyophilization. Cell layers were washed with water and extracted with 0.5 M acetic acid on a rotary shaker for 48 hr. The extracts were centrifuged at 27,000 $\times g$ for 15 min, and the supernatants were dialyzed for 48 hr against water before lyophilization.

Reduction and Alkylation of Medium. Medium freshly harvested from cell layers or after lyophilization was reduced in 50 mM dithiothreitol (CalBiochem)-0.5 M Tris (pH 8.2) for 60 min at room temperature. The solution was chilled, and iodoacetamide was added to 0.12 M. After the solution was stirred for 30 min, it was dialyzed against water and lyophilized.

Chromatography. Molecular sieve chromatography was performed on a 6% Biogel A-1.5 (BioRad) column (1.8 $\times$ 200 cm)

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

* Address reprint requests to Dept. of Pathology.

FIG. 1. Chromatography of medium on 6% agarose. About 2×10^8 cells in confluent layers were incubated with $100 \mu\text{Ci}$ of [^3H]proline and $100 \mu\text{Ci}$ of [^3H]glycine for 24 hr in serum-free medium. Processed medium was chromatographed directly (a), and after reduction and alkylation (b). Both samples of medium were denatured at 45° in 1 M CaCl_2 -50 mM Tris-HCl (pH 7.5) with 10 mg of added soluble human skin collagen as carrier. Continuous tracing represents absorbance of carrier collagen. Aliquots of each 3.2-ml fraction were counted for ^3H (O—O). THO (O—O—O), cpm in tritiated water in arbitrary units, added as internal standard. Brackets indicate pooling of fractions for electrophoresis in acrylamide gels.

eluted with 1 M CaCl_2 -50 mM Tris-HCl (pH 7.5). Calibration plots of log molecular weight were made from the elution positions of added tritiated water, beta components, alpha chains, and purified CNBr peptides from soluble human skin collagen, by the method of Piez (9).

Medium and cell layer extracts were also chromatographed at 45° on carboxymethyl cellulose (CMC) equilibrated with 60 mM sodium acetate (pH 4.8) (10). Chromatography of CNBr peptides was on CMC equilibrated with 20 mM sodium citrate (pH 3.6) (11).

Preparation of ^{14}C -Labeled α Chain Markers. 3×10^8 cells at confluence were labeled for 72 hr with $50 \mu\text{Ci}$ of [^{14}C]proline and $50 \mu\text{Ci}$ of [^{14}C]glycine in medium supplemented with 10% fetal-calf serum, ascorbate (75 $\mu\text{g}/\text{ml}$), and β -amino-propionitrile (50 $\mu\text{g}/\text{ml}$), and the cell layers were extracted with acetic acid as above. The extract yielded labeled $\alpha 1$ and $\alpha 2$ chains in the expected 2/1 ratio when it was run on gels and chromatographed on CMC-acetate. These chains cochromatographed with $\alpha 1$ and $\alpha 2$ chains of soluble human skin collagen used as carrier on both CMC-acetate and 6% agarose, and gave the full complement of labeled CNBr peptides after codigestion with soluble human-skin collagen and chromatography on CMC-citrate. We further verified the authenticity of the marker $\alpha 1$ and $\alpha 2$ chains by isolating them from CMC-acetate columns, desalting, hydrolyzing in 6 N HCl, and isolating the radioactive hydroxyproline and proline by ion-exchange chromatography (12). The ratio of cpm hy-

droxyproline/cpm proline was 0.8, as expected for authentic chains.

Analytical Gel Electrophoresis on SDS-Acrylamide Gels. Gels were 5% in acrylamide and 0.07% in methylenebisacrylamide (Eastman), and were polymerized in the running buffer: 0.1 M phosphate (pH 7.0) containing 0.1% SDS and 0.5 M urea (13). Lyophilized medium or cell layer extracts were taken up in 10 mM phosphate buffer (pH 7.0) containing 1% SDS and 0.5 M urea, and heated for 60 min at 55 – 60° . Samples of 0.05–0.2 ml were applied to gels after they were mixed with one drop of glycerol and Bromophenol Blue tracking dye. For reduction, samples were heat-denatured as above in the presence of 1% 2-mercaptoethanol, and 5 mM 2-mercaptoethanol was added to the standard running buffer. Electrophoresis was generally performed at 15 mA/13-cm gel for 5 hr.

With ^3H -labeled samples, ^{14}C -labeled α chains were added to the gels as internal markers. Parallel gels were also calibrated with purified unlabeled β components, α chains, and CNBr peptides of known molecular weights, all obtained from human-skin collagen. These gels were stained with Coomassie Blue, and mobilities were calculated by the method of Weber and Osborne (14). The mobility of a $\gamma 111$ species was taken from calibration plots of known $\alpha 1$ components.

Unstained radioactive gels were cut into 1-mm discs with a Mickle slicer, swollen, and eluted in NCS solubilizer (Amersham/Searle) (15), and counted in a toluene-based scintillant. For double-labeled gels, ^3H radioactivity in the sample always exceeded added ^{14}C marker radioactivity; discriminator set-

tings excluded ^3H counts from the ^{14}C channel, and gave a known, low spillover of ^{14}C radioactivity into the ^3H channel.

RESULTS

Under the given culture conditions, 9% of the protein synthesized by the fibroblasts was collagen (16). After it was labeled for 24 hr with [^3H]proline, 60–80% of the radioactive collagen was present as soluble molecules in the medium. These soluble forms of collagen can be shown to be precursors of the native collagen molecules precipitating in the cell layer (manuscript in preparation, and see below). From the amounts of radioactivity in nondialyzable hydroxyproline and proline, we calculated that the soluble collagen precursors represented, at a minimum, 70% of the total radioactive protein in the medium. It is therefore likely that most, if not all, of the proteins detected in medium by our methods represent collagen intermediates.

Molecular sieve chromatography on 6% agarose

The molecular weight distribution of the soluble, nondialyzable molecules secreted into the medium is shown in Fig. 1a. After a 24-hr period of labeling, about 70% of the radioactive material in the medium is recovered in a skewed peak eluting earlier than the marker β chains (molecular weight 190,000). The remaining labeled material elutes in the region of α chains (molecular weight 95,000), although this peak may contain some slightly heavier molecules.

Calculations from calibration curves of the molecular weight of the labeled material eluting before the β peak gave values ranging between 240,000 and 330,000. Because of the limited resolution of the 6% agarose column for molecular weights greater than 300,000, we could not determine if the peak eluting before the marker β chains contained significant amounts of molecules larger than native collagen γ chains (molecular weight 285,000).

In other experiments the labeled medium was divided into two portions: half was run on agarose as in Fig. 1a, and the remainder was reduced and alkylated before chromatography. The unreduced medium gave an elution profile identical to Fig. 1a, but all the radioactivity in the reduced and alkylated sample eluted as an asymmetric peak in the region of the α marker (Fig. 1b). Thus, the material of higher molecular weight eluting ahead of marker β components must have been assembled through disulfide bonds, and reduction and alkylation appeared to release polypeptides that were somewhat heavier than α chains.

Gel electrophoresis

To better define the molecular species secreted into the medium, we analyzed the radioactive peaks isolated on agarose columns by SDS-acrylamide gel electrophoresis. When the region before the β peak (Fig. 1a, fractions 56–68) was run on gels, two peaks were obtained (Fig. 2a). It is thus clear that the agarose column failed to resolve two molecular weight species: a major species (A) with a molecular weight greater than γ chains, and another species (B) with a molecular weight greater than β components. When the sample was reduced and run on gels (Fig. 2b), the radioactivity was quantitatively recovered in two major species, C and D, migrating more slowly than α chains. A leading shoulder on peak D was also reproducibly found. It was, therefore, clear that at least two different polypeptides were linked through disulfide bonds to form the molecular

FIG. 2. Acrylamide gel electrophoresis of fractions previously isolated by agarose chromatography. Fractions 56–68 of the experiment of Fig. 1a were pooled, dialyzed against water, lyophilized, and applied to a gel under standard conditions (a), and after reduction with 1% 2-mercaptoethanol (b). ^3H cpm, (●—●); [^{14}C] α -chain markers (○—○). Gel (a) was run for 75 mA-hr; gel (b) was run for 150 mA-hr to increase resolution.

species A and B of Fig. 2a. Based on their relative mobilities and their resemblance to procollagen chains demonstrated in other systems (17), peak C was designated pro α 1 and peak D pro α 2. The molecular weights of these molecules may be derived from calibration plots, allowing for the fact that α 2 components have greater mobilities in gels than α 1 components of the same molecular weight (18). This is clearly shown in our data for the ^{14}C -labeled marker α 1 and α 2 chains and for the unlabeled β markers. Based on calibration of gels with markers derived solely from α 1 chains, the presumptive pro α 1 chain has a molecular weight of 120,000. The presumptive pro α 2 is assigned a similar weight, assuming that, like the native α 2 chain, its mobility is anomalously high for its molecular weight.

If peak A (Fig. 2a) were composed of three disulfide-linked procollagen chains and peak B of two such chains, their molecular weights would be 360,000 and 240,000, respectively. Average molecular weights of 320,000 and 238,000 for peaks A and B, respectively, were obtained from experimentally observed mobilities and calibration plots derived from α 1 chain components only. By the arguments given above, the value obtained for peak A, which was lower than expected, suggests that at least one of its three procollagen chains must be pro α 2. The mobility data for peak B are in accord with a composition of two pro α 1 chains.

When medium was run directly on gels without prior agarose chromatography, the pattern shown in Fig. 3a was

FIG. 3. Acrylamide gel electrophoresis of medium. Labeling conditions are as for experiment of Fig. 1. (a) Processed medium applied directly to gel under standard conditions; (b) medium reduced and alkylated before application to gel. See text for lettering convention for radioactive peaks. ^3H cpm, (●—●); ^{14}C α -chain markers, (○—○). $\alpha 1$ -CB7 and $\alpha 1$ -CB8, CNBr peptide markers with molecular weights of 24,000.

obtained. Peaks with the same mobilities as those identified in gel runs of the pooled fractions from agarose are lettered correspondingly. Peaks A and B together contain about 60% of the incorporated radioactivity. Free pro $\alpha 2$ chains (D) are identified, but free pro $\alpha 1$ chains are absent. In addition, there is a peak (E) almost coincident with the marker $\alpha 1$ chains, and other peak (F) that runs somewhat behind marker $\alpha 2$. Also present are smaller peaks in the general regions of the β and CNBr peptide markers. As predicted from these data, pooled fractions 71–86 from the α region of the agarose column of Fig. 1a gave only peaks D, E, and F when run on gels (not shown).

Gel electrophoresis of reduced and alkylated medium is shown in Fig. 3b. As expected, species A and B have disappeared, and peaks C, D, E, and F are again identified. A similar pattern resulted when the peak obtained on agarose chromatography of reduced and alkylated medium (Fig. 1b) was subsequently run on gels (not shown). These data confirm the limited resolution of the agarose column for pro α and α chains.

When medium from cultures labeled for 24 hr with ^{35}S cysteine was analyzed on gels (not shown), peaks with mobilities of A, B, D, and F were present, but peak E was absent. Native α chains do not contain cysteine, and the result is con-

sistent with the interpretation that molecules A, B, D, and F are formed from procollagens with half-cystine residues in the peptide sequences amino terminal to the α chains. Peak F appears to represent a pro $\alpha 2$ chain with most, but not all, of the additional sequences excised.

Further confirmation that radioactive peaks A–F of the medium were collagen precursors was obtained by pulse-chase experiments (manuscript in preparation). In brief, medium from cultures labeled for 24 hr with ^3H proline and ^3H glycine were mixed with excess unlabeled proline and glycine and then placed on unlabeled cell layers. At intervals, samples of medium and acid extracts of cell layers were processed, measured for total recovered radioactivity, and run on gels. Radioactivity was principally chased from molecules A and B of the medium, and was quantitatively recovered as molecules of lower molecular weight in the cell layers and the medium. In the cell layers practically all of the radioactivity was recovered as $\alpha 1$ and $\alpha 2$ chains in the characteristic 2/1 ratio (Fig. 4). Radioactivity accumulating in the medium (not shown) during 48– to 72-hr chases was principally in two peaks that almost coincided with the marker $\alpha 1$ and $\alpha 2$ chains.

From the SDS-acrylamide gel data we conclude that a disulfide-linked trimer of procollagen chains is the major soluble collagen precursor molecule in the medium. A disulfide-linked procollagen dimer, free pro $\alpha 2$ chains, and molecules of sizes intermediate between pro α chains and α chains are also found.

DISCUSSION

The above experiments clearly indicate that SDS-acrylamide gel electrophoresis gives better resolution of molecules than agarose sieve chromatography. This can be attributed not only to the sieving characteristics of the gel, but also to the superiority of SDS-urea as a denaturing solvent. For the given conditions of denaturation, the SDS-urea solubilized essentially all of the radioactivity in the samples, whereas the 1 M CaCl_2 -Tris buffer used for agarose chromatography solubilized only 50% of the radioactivity, with disproportionate losses of the various procollagen species.

Our data are in essential agreement with the published results of Layman *et al.* (5) and Dehm *et al.* (7), and additionally, our acrylamide-gel analyses clearly establish that the

FIG. 4. Acid extract of cell layer after 48-hr chase. Labeled chase medium was placed on unlabeled cell layers for 48 hr, and then washed cell layers were extracted with 0.5 M acetic acid. Extract was applied to gel under standard conditions. ^3H cpm, (●—●); ^{14}C α -chain markers, (○—○).

principal procollagen molecule in the medium consists of three procollagen chains linked by disulfide bonds. We offer the following interpretation of our data: The only soluble secreted form of collagen is a procollagen molecule with the composition (pro $\alpha 1$)₂·pro $\alpha 2$ (peak A). Its assembly is stabilized by formation of disulfide bonds between cysteine residues in the nonhelical procollagen peptides of the three chains. The remaining portions of the chains come into register and form a triple helix by noncovalent interactions. After secretion, extracellular neutral peptidase(s) (19) progressively digest the nonhelical peptides of the procollagen trimer, but the noncovalent interactions maintain the three-chain assembly throughout the stepwise excisions. Initially, peptidase action produces a molecule in which two pro $\alpha 1$ chains are still disulfide-linked, but the pro $\alpha 2$ chain is associated only through noncovalent bonds. Denaturation of this intermediate by SDS-urea releases a pro $\alpha 1$ dimer (peak B) and free pro $\alpha 2$ chains (peak D). In accord with this model, free pro $\alpha 1$ chains (peak C) are released from peaks A and B only by reduction and denaturation. Eventually, the peptidase excises all disulfide bonds and leaves very few peptides terminal to the helical portion of the molecule. Thus, peak E has a mobility that differs minimally from $\alpha 1$ chains and is not labeled with [³⁵S]cysteine. Peak F still contains at least one cysteine residue and is also distinguished from $\alpha 2$ chains by its slower mobility. Continued proteolytic cleavage eventually generates insoluble molecules that aggregate as fibrils in the cell layers. Extraction of the cell layers thus yields native tropocollagen molecules.

This model proposes that an assembled three-chain procollagen molecule is the natural substrate for procollagen peptidase(s), and that free collagen chains are not normally present in the medium. If we have correctly deduced the chain composition of the various gel peaks, the ratio of radioactivity in all pro $\alpha 1$ chains to radioactivity in all pro $\alpha 2$ chains should approximate 2.0. This is the case for the experiment of Fig. 3a, and it has held for many experiments in which different labeling times and pulse-chase techniques have been used.

Stepwise cleavage of the procollagen chains by peptidase should release peptide fragments of low molecular weight into the medium. The gels gave little evidence of such peptides when [³H]proline and [³H]glycine were used for labeling (see Fig. 3a). However, after labeling with [³⁵S]cysteine, peaks corresponding to peptides of molecular weight 18,000–25,000 contained about 30% of the radioactivity recovered in the gel, and these peptides yielded smaller fragments on reduction. These peaks might well correspond to cysteine-rich, N-terminal peptides excised from the procollagen chains.

Pulse-chase experiments of the type shown in Fig. 4 indicated that efficient conversion of the procollagen trimer of the medium to a native molecule in the cell layer required many hours. This slow processing might result from the large medium to cell volume ratio in tissue culture, as intermediates would be enormously diluted, and the net rate of diffusion between medium and cell layer might be slow. The same intermediates might be processed more rapidly in the smaller extracellular space of the cell layer. Accordingly, in bone organ cultures where the extracellular volume is relatively small, efficient conversion of procollagen chains to native chains occurs within minutes (3). Perhaps for this reason it has not been possible to demonstrate disulfide-linked procollagen intermediates in the calvarial system. The possibility remains however, that the manner of assembly and processing of procollagen chains might differ for different tissues.

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

1. Speakman, P. T. (1971) *Nature* **229**, 241–243.
2. Bornstein, P., Von Der Mark, K., Wyke, A. W., Ehrlich, H. P. & Monson, J. M. (1972) *J. Biol. Chem.* **247**, 2808–2813.
3. Bellamy, G. & Bornstein, P. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 1138–1142.
4. Müller, P. K., McGoodwin, E. & Martin, G. R. (1971) *Biochem. Biophys. Res. Commun.* **44**, 110–116.
5. Layman, D. L., McGoodwin, E. B. & Martin, G. R. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 454–458.
6. Jimenez, S. A., Dehm, P. & Prockop, D. J. (1971) *FEBS Lett.* **17**, 245–248.
7. Dehm, P., Jimenez, S. A., Olsen, B. R. & Prockop, D. J. (1972) *Proc. Nat. Acad. Sci. USA* **69**, 60–64.
8. Burgeson, R. E., Wyke, A. W. & Fessler, J. H. (1972) *Biochem. Biophys. Res. Commun.* **48**, 892–897.
9. Piez, K. A. (1968) *Anal. Biochem.* **26**, 305–312.
10. Piez, K. A., Eigner, E. A. & Lewis, M. (1963) *Biochemistry* **2**, 58–66.
11. Miller, E. J., Epstein, E. H., Jr. & Piez, K. (1971) *Biochem. Biophys. Res. Commun.* **42**, 1024–1029.
12. Goldberg, B. & Green, H. (1967) *J. Mol. Biol.* **26**, 1–18.
13. Shapiro, A. L., Vinuela, E. & Maizel, J. V., Jr., (1967) *Biochem. Biophys. Res. Commun.* **28**, 815–820.
14. Weber, K. & Osborn, M. (1969) *J. Biol. Chem.* **244**, 4406–4412.
15. Basch, R. D. (1968) *Anal. Biochem.* **26**, 184–188.
16. Green, H. & Goldberg, B. (1964) *Nature* **204**, 347–349.
17. Tsai, R. L. & Green, H. (1972) *Nature New Biol.* **237**, 171–173.
18. Furthmayr, H. & Timpl, R. (1971) *Anal. Biochem.* **41**, 510–516.
19. Bornstein, P., Ehrlich, H. P. & Wyke, A. W. (1972) *Science* **175**, 544–546.