Vol. 87, No. 3, 1979 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 13, 1979 Pages 818-826

PROCOLLAGEN α2 mRNA IS SIGNIFICANTLY DIFFERENT FROM PROCOLLAGEN α1(I) mRNA IN SIZE OR SECONDARY STRUCTURE

Paul Tolstoshev, Roberta Haber and Ronald G. Crystal

Pulmonary Branch,
National Heart, Lung and Blood Institute
Bethesda, Maryland 20014

Received February 26, 1979

SUMMARY: The fractionation of fetal calf tendon messenger RNA in 85 percent formamide sucrose gradients shows a separation of the mRNAs coding for pro α1(I) and pro α2 chains of type I collagen. This difference in sedimentation in denaturing gradients suggests that pro α2 mRNA is approxmately 1000 bases shorter than pro α1(I) mRNA. However, such a size difference is significantly greater than would be predicted from consideration of the size of the polypeptide chains coded for by these mRNAs, and thus, residual secondary structure in the mRNAs may contribute to these apparent size differences.

INTRODUCTION:

Vertebrate collagens are a heterogenous set of at least five different proteins types (1-4). The most common is type I collagen which is found in a wide variety of tissues, including skin, bone, tendon and lung (1,2,5). Type I collagen is composed of three polypeptide chains, two α1(I) chains and one α2 chain. Each of these chains is synthesized as a precursor form, termed procollagen [pro α1(I) and pro α2, respectively], with extensions to the α chains at both the N- and C- terminal ends (1,2,6,7).

The size of the pro α1(I) chain is of the order of 1550 amino acids, consisting of an α chain sequence of about 1050 amino acids, an N-terminal extension of 150-200 amino acids (8-11), and a C-terminal extension of 300-350 amino acids (10-14). The pro α2 chain is remarkably similar, with an α chain of about 1050 residues, an N-terminal extension of approximately 50 amino acids, and a C-terminal extension of 300-350 amino acids, giving a total of approximately 1400 residues (10-15). Thus, from current chemical data, it is likely that the pro α1(I) chain is at most 10 percent larger than the pro α2 chain.

0006-291X/79/070818-09$01.00/0

Although the amino acid sequence of the pro α1(I) and pro α2 chains are clearly different, more than two thirds of each chain has the repetitive sequence of the form gly-X-Y, where approximately one-third of the X and Y residues are proline or hydroxyproline (1). Thus, although the pro α1(I) and pro α2 chains represent different gene products, their amino acid sequences have remarkable similarities.

Taken together, the size and sequence considerations of pro α1(I) and pro α2 chains predict that the mRNAs for these two polypeptides should be similar both in size and secondary structure. However, despite these theoretical considerations, the present study demonstrates that the mRNAs for the pro α1(I) and pro α2 chains are sufficiently different in apparent size or secondary structure to be readily separated by denaturing sucrose gradients.

MATERIALS AND METHODS

Isolation of Tendon Messenger RNA

Total RNA was isolated from 4 to 6 month old fetal calf tendons using 6M guanidine hydrochloride (16). Poly A+ containing mRNA was isolated from total RNA by affinity chromatography with oligo dT-cellulose (17).

Fractionation of mRNA with Aqueous and Denaturing Sucrose Gradients

Fractionation on aqueous sucrose gradients was performed as described by Rosen et. al. (18). Fifty μgm of fetal calf tendon poly A+ mRNA was heated to 70° for 1 minute, chilled on ice, and layered over a 10-35 percent exponential sucrose gradient. Centrifugation was performed in a Beckman SW41 rotor (246,000 g, 11 hrs., 20°), gradient fractions of 0.5 ml were collected, and the absorbance at 260 nm measured. RNA was recovered from pooled fractions by precipitation with one tenth volume of 20 percent potassium acetate, pH 6.0, and 2 volumes of ethanol.

Denaturing gradients were 5-22 percent (w/v) sucrose dissolved in 85 percent formamide (Eastman, Spectrograde, deionized before use), 10 mM Tris-HCl, 1 mM EDTA, pH 7.4. RNA (40 μgm) was dissolved in the 85 percent formamide buffer, heated to 70° for 1 minute, rapidly cooled and layered on the gradient. Centrifugation was at 325,000 g for 10 hours at 25° in an SW56 Beckman rotor. The absorbance profile at 260nm was measured by upward displacement of the gradients through an Isco Absorbance Monitor flow cell, and 150 μl gradient fractions were collected. The RNA in 20 μl of each fraction was precipitated with potassium acetate and ethanol in the presence of 10 μg of reticulocyte tRNA, and translated in the cell-free system. Reticulocyte tRNA at this concentration had no effect of the translation of procollagen mRNA.

Cell-Free Translation of Messenger RNA

Aliquots of RNA from gradient fractions were translated in the nuclease-treated reticulocyte lysate as described by Pelham and Jackson (19). The total reaction mixture was 30 µl, of which 10 µl was lysate. Potassium chloride was present at 40 mM, and potassium acetate at 100 mM. [^{35}S]methionine (Amersham Searle, 1245 Ci/m mole, 11 µCi/assay) was used as the tracer amino acid. Incubation was at 30^{0} for 60 minutes, and synthesis was terminated by addition of 5 µg of pancreatic RNAse followed by an additional 15 min incubation at 30^{0}.

Electrophoretic Analysis of Cell-free Products

Following incubation, 20 µl of each reaction mixture was made to 50 mM Tris-HCl, pH 6.8, 5 percent glycerol, 2 percent sodium dodecyl sulfate (SDS) and 100 mM dithiothreitol. Samples were boiled for 5 min., and them electrophoresed on 5 percent acrylamide-0.027 percent bis-acrylamide slab gels (20) for 2.5 hr at 30 ma. Gels were prepared for flurography as described by Laskey and Mills (22).

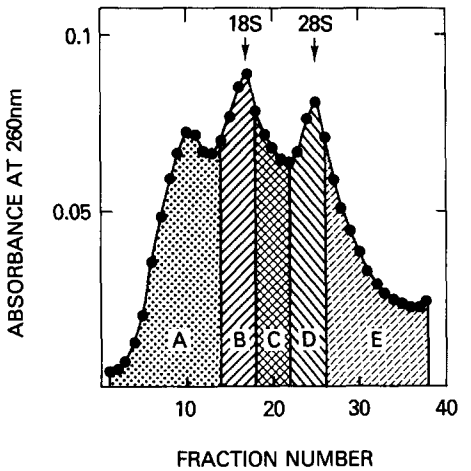
Standard [^{14}C]-labelled type I procollagen was prepared from the medium of human lung fibroblasts as preciously described (21).

Ion-Exchange Chromatography of Cell-Free Products

Following incubation, the cell free products were precipitated with 30 percent saturation of ammonium sulfate in the presence of 100 mM Tris-HCl, pH 7.4 and 5 mM unlabeled methionine, pelleted (30,000 g, 30 min, 4^{0}) and dialysed overnight against 100 mM Tris-HCl, pH 7.4, 5 mM methionine, 6M urea. To ensure that the procollagen chains in the cell-free product were not cross-linked through disulfide bonds, the dialyzed material was reduced (50 mM dithiothreitol, 4 hrs, 23^{0}) and alkylated (100 mM sodium iodoacetate, 1 hr 23^{0}). Standard type I procollagen was also reduced and alkylated in identical fashion. The samples were then chromatographed on a column of deithylaminoethyl (DEAE)-cellulose under conditions (10 mM Tris-HCl, pH 8.6 6M urea, 0 to 250 mM NaCl gradient, 23^{0}) which separate pro α2 and pro α1(I) chains (8, and Graves P, Fietzig, P., Monson, J., Prockop, D., and Olsen, B., in preparation). Fractions of cell-free products corresponding to a pro α2 standard (eluting at 35 mM NaCl) and pro α1(I) standard (eluting at 95 mM NaCl) were pooled, dialyzed into 0.5 percent acetic acid, lyophilized and electrophoresed in SDS-acrylamide slab gels as described above.

RESULTS AND DISCUSSION

Fractionation of fetal calf tendon poly A+ mRNA on an aqueous sucrose gradient demonstrated that it consisted of RNAs of a broad size distribution, with peaks corresponding to residual 18S and 28S ribosomal RNA (Figure 1A). When aliquots for RNA from five pooled gradient fractions (indicated as fractions A-E in Figure 1, upper panel) were translated in a nuclease-treated reticulocyte lysate cell-free system with [^{35}S]methionine as the tracer, analysis of the cell-free products on sodium docycel sulfate



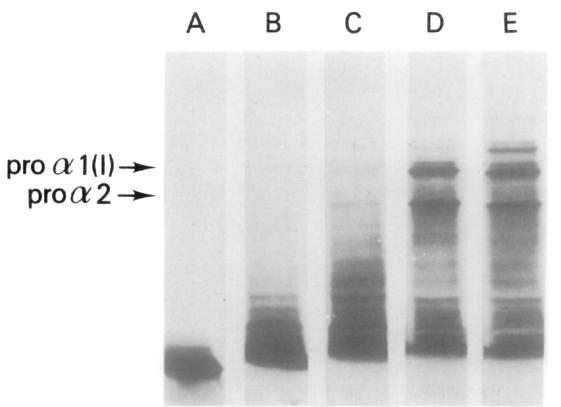
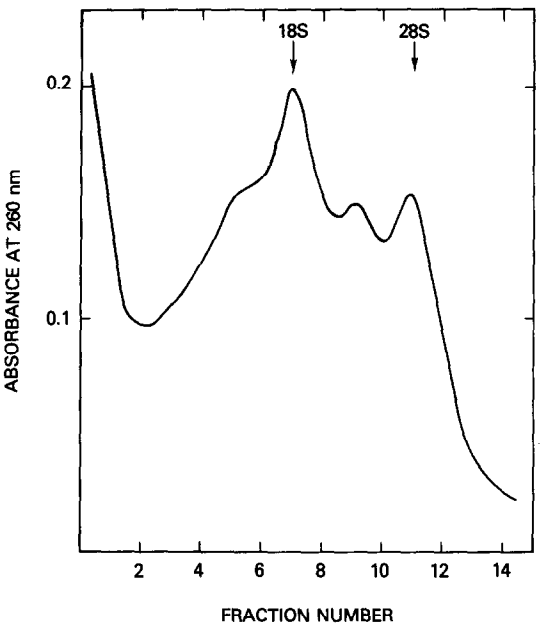


Figure 1: Fractionation of fetal calf tendon mRNA on an aqueous sucrose gradient and translation of these fractions in the reticulocyte cell-free system. Upper panel: absorbance profile (260 nm) of 50 µgm of fetal calf tendon poly A+ mRNA fractionated on an aqueous 10-35 percent exponential sucrose gradient. The gradient was divided into 5 pools, A-E, and the RNA in each was recovered by precipitation with ethanol. An equal amount (2.5 µgm) of RNA from each fraction was translated in the cell-free system. Sedimentation of 18S and 28S ribosomal RNA markers are indicated. Lower panel: fluorogram of the translation products of each of the 5 pools analyzed by SDS-acrylamide gel electrophoresis as described in Materials and Methods. Positions of pro α1(I) and pro α2 standards are indicated.

polyacrylamide gels revealed two prominent high molecular weight polypeptides encoded by fractions D and E. These polypeptides corresponded closely in size to authentic pro α1(I) and pro α2 chains (Figure 2),



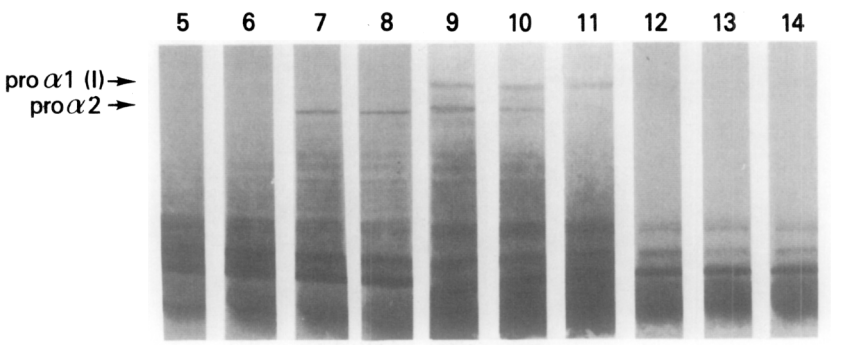


Figure 2: Fractionation of type I procollagen mRNAs on a denaturing gradient and translation of gradient fractions. Upper panel: absorbance profile (260 nm) of 40 μgm of fetal calf tendon poly A+ mRNA analyzed on a 5-22 percent exponential sucrose gradient in 85 percent formamide, as described in Materials and Methods. Sedimentation of 18S and 28S ribosomal RNA markers are indicated. Lower panel: fractions of 150 μl of the sucrose gradient were collected and the RNA in 20 μl of each gradient was translated in the cell-free system. The translation products of each fraction were analyzed by SDS-acrylamide gel electrophoresis and subsequent flurography. Position of pro α1(I) and pro α2 standards are indicated.

were completely digested by Clostridial collagenase, and were precipitated from the cell-free reaction mixture by anti-calf type I collagen antibodies (data not shown). By these criteria, these two polypeptides represent

the primary translation products of type I procollagen mRNA. Thus the mRNAs for calf tendon type I procollagens sediment in aqueous sucrose gradients in the range of 25 to 30 S, in agreement with similar analysis of chick calvaria procollagen mRNAs (23,24,25).

However, when the calf tendon type I procollagen mRNAs were denatured in 85 percent formamide (70°, 1 min) and then sedimented through a sucrose gradient in the presence of 85 percent formamide at 25°, translation of the gradient fractions demonstrated a clear separation of the mRNAs for pro α2 and pro α1(I) (Figure 2). Whereas on aqueous gradients the two mRNAs sedimented together in a range of 25 to 30 S, under the denaturing condition of the formamide gradients, the pro α2 mRNA sedimented much slower than the pro α1(I) mRNA, suggesting an apparent size difference of at least 1000 bases.

To confirm that the mRNA sedimenting slower in the formamide sucrose gradient coded for pro α2 chains, RNA from gradient fractions 7 + 8 of Figure 2 were pooled, translated in the cell-free system, and the product analyzed sequentially by diethylaminoethyl (DEAE)-cellulose chromatography and SDS-acrylamide gel electrophoresis. The cell-free collagen product directed by gradient fractions 7 + 8 eluted on the DEAE column in the region of the pro α2 standard and subsequently electrophoresed on SDS-gels with the pro α2 standard (Figure 3, gel A). In comparison, gradient fractions 7 + 8 did not direct the translation of collagen chains that eluted on the DEAE column in the region of the pro α1(I) standard (Figure 3, gel B). Thus, gradient fractions 7 + 8 of Figure 2 clearly coded for pro α2 and not pro α1(I) chains. Gradient fractions 9-11 (Figure 2B) contain the pro α1(I) mRNA activity, as well as pro α2 mRNA activity from the region of overlap. This was confirmed by sequential DEAE and SDS-acrylamide gel analysis (Figure 3, gels C + D). Thus, although there is

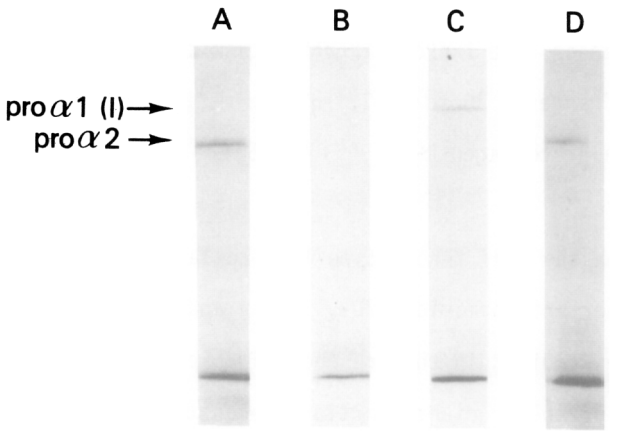


Figure 3: Confirmation of the identity of pro α2 mRNA. RNA from fractions 7 + 8, and fractions 9-11 of the formamide gradient (Figure 2) was pooled respectively, and 3.75 μgm of each RNA were translated in a 150 μl cell-free incubation. The cell-free products were analyzed by sequential DEAE-cellulose chromatography and SDS-acrylamide gel electrophoresis, as described in Materials and Methods. (A) SDS-acrylamide gel analysis of the pro α2 region of the DEAE chromatogram of gradient fractions 7 + 8 translation products. (B) the pro α1(I) region of the same DEAE chromatogram as in (A). (C) SDS-acrylamide gel analysis of the pro α2 region of the DEAE chromatogram of gradient fractions 9-11 translation products. (D) The pro α1(I) region of is the same DEAE chromatogram as in (C). The position of pro α1(I) and pro α2 markers are indicated.

some overlap, the mRNA coding for pro α2 collagen chains clearly sedimented slower than the bulk of the pro α1 mRNA under the conditions of the formamide sucrose gradients.

If the gradient separation is due to an actual size difference, the difference must come in the untranslated regions at the 5' and 3' ends of the mRNAs. Such variation in untranslated regions has been found in a variety of eukaryotic mRNAs. For example, the 3' untranslated region of human β-globin mRNA is 21 percent larger than the comparable region of human α-globin mRNA and 43 percent larger than the 3' untranslated region of rabbit β-globin mRNA (26-28). Pro α2 mRNA may differ significantly in size from pro α1(I) mRNA, with the size difference being in untranslated regions.

The observed difference in the sedimentation of the two mRNAs may also result from differences in secondary structure. Although 85 percent

formamide is sufficient to denature most RNAs, including ribosomal RNA (29), the collagen mRNAs might retain secondary structure even under these conditions, due to their high (G + C) base composition.

Finally, it is possible that the difference in sedimentation of the pro α1(I) and pro α2 mRNA is due to their binding to other macromolecules. However, this is unlikely, since such interactions would need to be sufficiently stable to survive the extraction (6M guanidine-HCl) and fractionation (85 percent formamide) conditions, yet not interfere with the translation of the mRNAs in the cell-free system.

The observation that pro α2 mRNA is significantly different from pro α1(I) mRNA in size or secondary structure or both is of interest for several reasons. First, a difference in length of the untranslated regions of the mRNAs may imply differences in regulatory sequences related to the processing or translation of the two mRNAs. Second, differences in secondary structure might result in less efficient translation of pro α2 mRNA. Since Type I collagen consists of two α1(I) chains and one α2 chain, it is conceivable that pro α1(I) and pro α2 mRNAs are present in equal quantities, but that differences in the efficiency of translation result in the final α1(I)/α2 ratio of 2:1 that is found in most cells producing Type I collagen. Third, the separation of these mRNAs will be useful for the preparation of chain-specific complementary DNA probes for molecular hybridization studies of the regulation of expression of these two genes.

REFERENCES:

1. Ramachandran, G.N. and Reddi, A. (1976) in Biochemistry of Collagen, Plenum Publishing Corporation, New York.
2. Miller, E.J. (1976) Molec. Cell. Biochem. 13, 165-192.
3. Burgeson, R.E., El Adli, F.A., Kaitile, I.I. and Hollister, D.W. (1976) Proc. Natl. Acad. Sci. USA, 73, 2579-2583.
4. Chung, E., Rhodes, R.K. and Miller, E.J. (1976) Biochem. Biophys. Res. Commun. 71, 1167-1174.
5. Hance, A.J. and Crystal, R.C. (1975) Am. Rev. Resp. Disease, 112, 657-711.
6. Bornstein, P. (1974) Ann. Rev. Biochem. 43, 567-603.
7. Martin, G.R., Byers, P.H. and Piez, K.A. (1975) Adv. Enzymol. 42, 167-191.
8. Monson, J.M., Click, E.M. and Bornstein, P. (1975) Biochemistry, 14, 4088-4092
9. Becker, U., Timpl, R., Helle, O., and Prockop, D.J. (1976) Biochemistry, 15, 2853-2862.

10. Bornstein, P., and Traub, W. (1979) The Chemistry and Biology of Collagen. The Proteins, Academic Press, New York, in press.
11. Prockop, D.J., Kivirikko, K.I., Tuderman, L., and Guzman, N.A. (1979) New Engl. J. Med. in press.
12. Byers, P.H., Click, E.M., Harper, E. and Bornstein, P. (1976). Proc. Natl. Acad. Sci. USA. 72, 3009-3013.
13. Davidson, J.M., McEneany, L.S.G. and Bornstein, P. (1975) Biochemistry, 14, 5188-5194.
14. Olsen, B.R., Guzman, N.A., Engel, J., Condit, C. and Aase, S. (1977) Biochemistry, 16, 3030-3036.
15. Becker, U., Helle, O. and Timpl, R. (1977) FEBS Letts. 73, 197-200.
16. Strohman, R.C., Moss, P.S., Micou-Eastwood, J., Spector, D., Przybyla, A. and Paterson, B. (1977) Cell, 10, 265-273.
17. Aviv, J. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA, 69, 1408-1412.
18. Rosen, J.M., Woo, S.L.C. and Comstock, J.P. (1975) Biochemistry, 14, 2895-2903.
19. Pelham, H.R. and Jackson, R.J. (1976) Eur. J. Biochem. 76, 247-256.
20. Davis, B.J. (1964) Annals N.Y. Acad. Sci. 121, 464-427.
21. Hance, A.J., Bradley, K. and Crystal, R.G. (1976) J. Clin. Invest. 57, 102-111.
22. Laskey, R.A. and Mills, A.D. (1975) Eur. J. Biochem. 56, 335-341.
23. Boetker, H., Frischauf, A.M. and Lerach, H. (1976) Biochemistry, 15, 4765-4770.
24. Rowe, D.W., Moen, R.C., Davidson, J.M., Byers, P.H., Bornstein, P. and Palmiter, R.D. (1978) Biochemistry, 17, 1581-1590.
25. Howard, B.J., Adams, S.L., Sobel, M.E. Pastan, I. and de Crombrugghe, B. (1978) J. Biol. Chem. 253, 5869-5874.
26. Marotta, C.A., Wilson, J.T., Forget, B.G. and Weissman, S.M. (1977) J. Biol. Chem. 252, 5040-5053.
27. Proudfoot, N.J., Gillam, S., Smith, M. and Longley, J.I. (1977) Cell, 11, 807-818.
28. Efstratiadis, A., Kafatos, F.C. and Maniatis, T. (1977) Cell, 10, 571-586.
29. MacNaughton, M., Freeman, K.B. and Bishop, J.O. (1974) Cell, 1, 117-125.