Vol. 92, No. 2, 1980 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 29, 1980 Pages 554-562

THE MOLECULAR WEIGHT OF THE CELL-FREE TRANSLATION PRODUCT OF α1(I) PROCOLLAGEN mRNA

Linda Sandell and Arthur Veis*

*Departments of Biochemistry and Oral Biology
Northwestern University Medical and Dental Schools
Chicago, Illinois 60611

Received December 3,1979

SUMMARY: Chick embryo calvaria RNA has been translated in a nuclease-treated reticulocyte lysate system. The products have been characterized by their susceptibility to collagenase digestion and precipitability by Type I procollagen chain antibodies. The largest product, pre procollagen α1(I), migrates more slowly than proto procollagen α1(I) isolated from calvaria, contains initiator methionine and is cleaved to proto procollagen α1(I) by dog pancreas microsomes. The molecular weight of pre procollagen α1(I) was estimated to be 165,000 and by a similar analysis, pre procollagen α2 was identified and its weight determined to be approximately 165,000 daltons. Bacterial collagenase digestion of pre procollagen α1(I) generated two new low molecular weight hydrophobic components.

INTRODUCTION

It has been hypothesized that proteins destined for secretion are synthesized in a precursor form containing an initial NH_2-terminal sequence which acts as a signal to direct the nascent peptide into the membrane of the endoplasmic reticulum [1,2]. Procollagen fits this model and is initially translated as a precursor which may contain the "signal" peptide [3]. The largest procollagen α1(I) chain, $P_{NC}\alpha1(I)$[+], isolated from tissue, has an estimated molecular weight of 145,000 based on sequence analyses of the NH_2-terminal and triple-helical regions [4,5] and amino acid analysis of the COOH-terminal region [6]. This investigation has been directed at determining the molecular weight of the initially translated procollagen Type I (α1) chain since it appears that the precursor portion of the collagen molecule may be considerably larger than predicted by the signal hypothesis. In this

+ $P_{NC}\alpha1(I)$: Type I procollagen α1 chain containing the NH_2 and COOH terminal propeptides. Proto $P_{NC}\alpha1(I)$: underhydroxylated $P_{NC}\alpha1(I)$. $P_C\alpha1(I)$: $P_{NC}\alpha1(I)$ without the NH_2-terminal propeptide.

0006-291X/80/020554-09$01.00/0

study, the mRNA from chick embryo calvaria has been translated in a cell-free nuclease-treated reticulocyte lysate system. In addition to determination of the molecular weight of the translation products, these components, labeled with [^{3}H]proline, [^{35}S]methionine, [^{3}H]methionine, or [^{3}H]phenylalanine, have been characterized by their susceptibility to collagenase digestion and their precipitability by Type I procollagen chain antibodies.

MATERIALS AND METHODS

RNA was extracted from calvaria of 13 day chick embryos by phenol/chloroform following incubation with proteinase K (Beckman) [3] or by ethanol precipitation of bones homogenized in 8M Gu-HCl [8]. Reticulocyte lysates were prepared by the method of Pelham and Jackson [9]. Protein synthesis was carried out in a total volume of 30 µl containing 10 µl of nuclease-treated reticulocyte lysate fortified with 0.2 mM hemin and 0.5 µg creatine phosphokinase; 140 mM potassium, 10 mM creatine phosphate, 20 mM Hepes buffer (pH 7.5), 150 µg/ml total RNA or 10-15 µg/ml mRNA. Labeled amino acids or ^{3}H-met-tRNA$_{fmet}$ (New England Nuclear) were added as desired. The cell-free product was immunoprecipitated with antiserum prepared by immunizing rabbits with dermatosparactic procollagen and collected by reaction with goat anti-rabbit IgG followed by centrifugation through sucrose [10]. Dog pancreas microsomes, containing signal peptidase activity were prepared essentially by the method of Shields and Blobel [11]. Cleavage of the putative signal sequence was accomplished by adding the pancreas microsomes to the translation system. Proteins were examined by SDS-PAGE [12] with subsequent fluorography [13,14] or cut into 1 mm slices and counted. Samples were reduced with 2-mercaptoethanol, heated to 100° and applied directly to gel.

RESULTS

Additions of calvaria mRNA to the reticulocyte lysate system simulated incorporation of [^{35}S]methionine into TCA precipitable protein by 35-fold, [^{3}H]phenylalanine by 20-fold, and [^{3}H]proline by 10-fold. The products of cell-free translation are shown in the fluorograph of Fig. 1. Lane A shows the translation directed by calvaria mRNA, and lane B, by globin mRNA. Lane D shows the products of translation that are immunoprecipitated by antiserum to Type I procollagen; these high molecular weight bands are removed by digestion with bacterial collagenase (lane C). Figure 2 compares the translated products with standards of known molecular weight and amino acid composition. The major labeled component of cell-free synthesis migrates more slowly than proto P$_{NC}$α1(I) on SDS-PAGE. Lane D of Fig. 2 shows [^{35}S]-

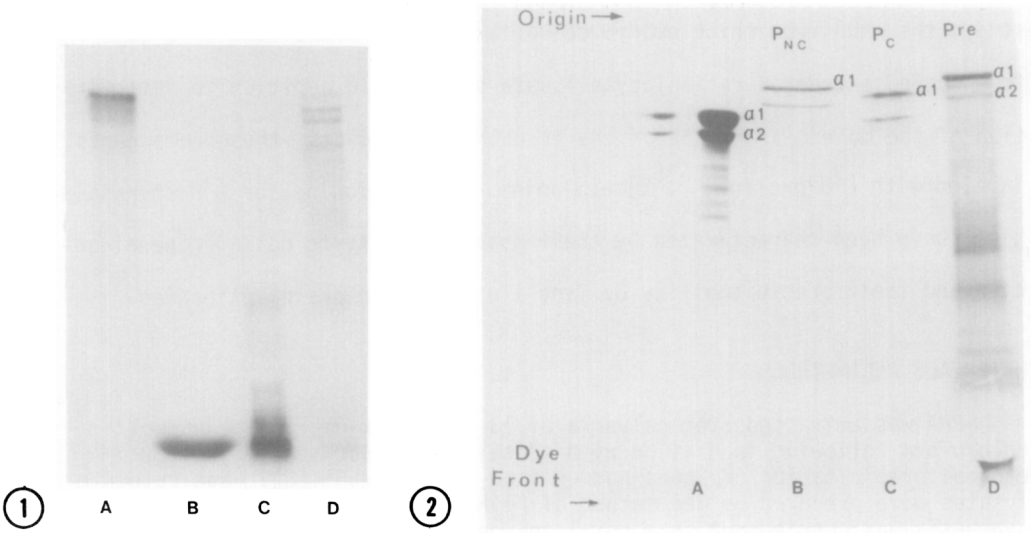

Fig. 1. Fluorograph of SDS-PAGE (6-15% gradient) of [^{3}H]proline labeled translation product of calvaria mRNA (A) and hemoglobin mRNA (B) in mRNA-dependent reticulocyte lysate. The anti procollagen immunoprecipitate of the calvaria mRNA translation product is shown in Lane D, Lane C is the result of the digestion of the material of (D) with bacterial collagenase. Bacterial collagenase (Advanced Biofactures) was added directly to the translation mix after 60 min of incubation or to the immunoprecipitate to a final concentration of 250 U/Ml collagenase, 10 mM NEM (as an inhibitor of non-specific protease activity), 2.5 mM Tris (pH 7.4), and 33 mM $Ca(Ac)_2$. Incubation was for 30 min at 30°.

Fig. 2. Fluorograph of SDS-PAGE (6-15% gradient, 9.5 cm in length). The labeled components are: collagen α1 and β_{11} chains (Lane A), proto P_{NC}α1 chains (Lane B); P_Cα1 chains (Lane C); cell-free product of chick calvaria mRNA (Lane D). Lanes A-C are labeled with [^{3}H]proline; Lane D is labeled with [^{35}S]methionine. Standards were prepared by labeling 17 day chick embryo calvaria with [^{3}H]proline for 18 min and extracting procollagen [19]. Proto-procollagen was prepared by inhibiting prolyl and lysyl hydroxylase activity with 0.02 mg/ml 2,2-dipyridyl.

methionine incorporation and should be compared directly to lane A of Fig. 1 depicting proline incorporation.

Collagens migrate anomalously in SDS-PAGE when compared with non-collagenous proteins. However, log M_r is a linear function of R_f in each case (Fig. 3), and the slopes of the two plots are virtually parallel to each other. It appears that the anomalous migration of the collagens is proportional to their relative content of collagen helix region sequence. For intact chains of the procollagen α1 series, each chain has an identical length of helical region peptide. Thus when the R_f's and M_r's of these collagens are plotted, the points fall intermediate between the "pure" colla-

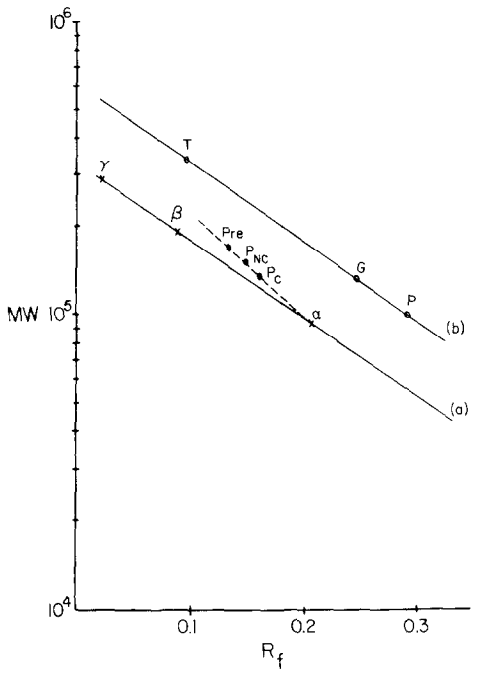

Fig. 3. Molecular weight determination of translated product. All points were taken from the electrophoresis represented in the fluorograph of Fig. 2, including the non-collagen and collagen standards which do not appear on the fluorograph but were stained with Coomassie Blue. Line (a) is the standard curve for collagenous proteins. Line (b) is the standard curve for non-collagenous proteins; phosphorylase a (P), β-galactosidase (G) and thyroglobulin subunit (T). The dashed line is the procollagen standard curve (see text): α1, M_r = 95,000, 100% collagenous; P_Cα1, M_r = 135,000, 71% collagenous; proto P_{NC}α1, M_r = 145,000, 64% collagenous, preprocollagen, placed on this procollagen curve according to its determined R_f value, has a molecular weight of 165,000 and is presumably 58% collagenous.

gen and "pure" globular standard curves. Each point would lie on a line which is one of a set of lines parallel to the two "pure" standard plots, a and b of Fig. 3, which represent 100% and 0% collagenous composition, respectively. The line joining the "precursor" collagen points thus represent the correct calibration line for an unknown collagen precursor with the major portion of triplet sequence chain intact. On this basis the molecular weight of the pre P_{NC}α1(I) is estimated to be 165,000 ± 5000. By similar analysis, the pre P_{NC}α2 chain is identified (Fig. 2) and also has a molecular weight of 165,000 ± 5000.

To determine whether the pre P_{NC}α1(I), the major protein, had an intact NH_2-terminal precursor region, calvaria mRNA was translated in the presence

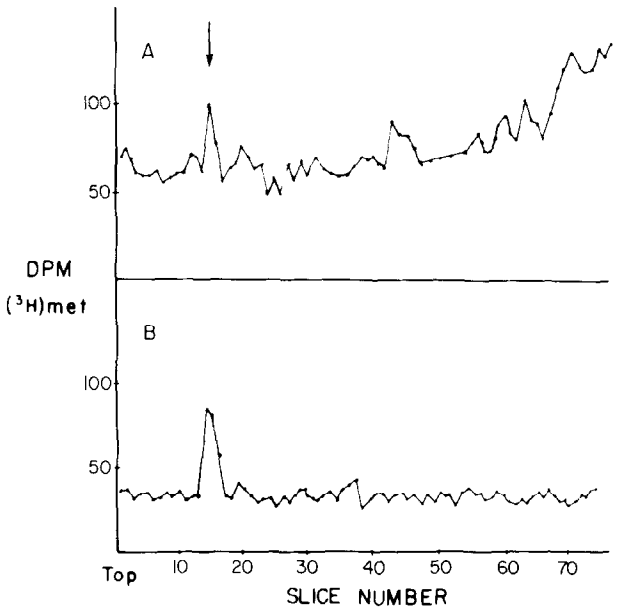

Fig. 4. Fluorograph of SDS-PAGE (6%) of [^{3}H]met labeled cell-free translation product. (A) Gels were sliced into 1 mm slices and counted by liquid scintillation. (B) Comparable SDS-PAGE (6%) of anti procollagen immunoprecipitate of material from A. Gels were impregnated with PPO and then sliced and counted.

of [^{3}H]methionyl-tRNA$_{fmet}$. This tRNA will deliver methionine only in the initiator position and the methionine will be retained in the in vitro translation product [15]. Figure 4 shows the results of this experiment. The products of translation were electrophoresed immediately (Fig. 4a) or first immunoprecipitated by antiprocollagen serum and then electrophoresed (Fig. 4b). Initiator methionine is incorporated into the band which is the dominant cell-free product labeled by [^{3}H]proline in the control translation (arrow 1). This band is more clearly seen when the cell-free product is immunoprecipitated (Fig. 4b) demonstrating that it contains the intact α1(I) chain NH_2-terminus.

In separate experiments RNA was translated in the presence of [^{3}H]-phenylalanine, or [^{3}H]proline and the products analyzed by SDS-PAGE. Phenylalanine was chosen because in calf skin procollagen there are no phenylalanine residues among residues 1-98 of $P_{NC}\alpha$1(I), where it is likely to occur in

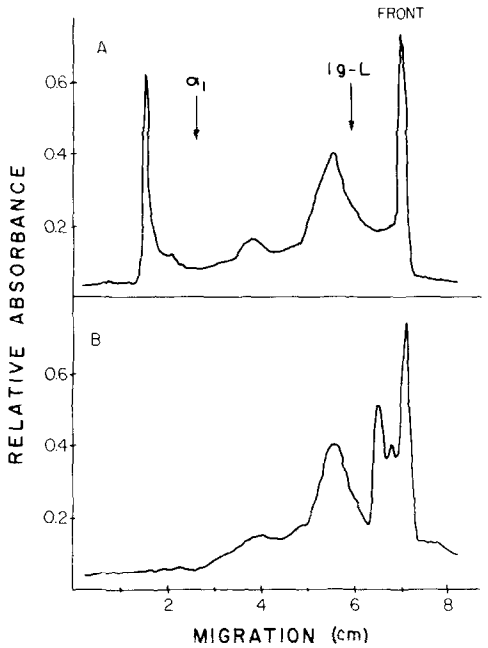

Fig. 5. Scan of fluorograph of SDS-PAGE (6-15%) of cell-free product labeled with [^{3}H]phenylalanine without digestion with collagenase (A), with collagenase (B). Markers are α1 chains and IgG light chain. Samples were reduced in 2-mercaptoethanol. With collagenase digestion under the conditions in the legend of Fig. 1, the high molecular weight peak in (A) disappears while two new low molecular weight peaks appear. Hemoglobin runs with the dye front.

the signal peptide region. Figure 5a is a densitometer scan of the fluorograph of the phenylalanine labeled translation, showing marked phenylalanine incorporation. Two [^{3}H]Phe-labeled bands are generated by bacterial collagenase digestion (Fig. 5b), representing collagenase resistant propeptides. The integrated total counts of the bacterial collagenase-generated phenylalanine-labeled band are almost equal to the counts of the major translated product. In the comparison experiment, the two proline-labeled bands generated in the same way have many fewer counts (data not shown). Thus, most of the proline residues must have been released by collagenase digestion whereas more of the phenylalanine residues are in the portion of collagen which is not susceptible to bacterial collagenase, the pre or propeptide regions. Pre $P_{NC}\alpha 1(I)$ is the major product of the translation system and as shown above, does contain initiator methionine. Preliminary experiments in

which translation was conducted in the presence of dog pancreas microsomes indicate that the pre $P_{NC}\alpha 1(I)$ is reduced in size to a component migrating identically to proto $P_{NC}\alpha 1(I)$ (data not shown).

DISCUSSION

The major proline-labeled translation product of calvaria mRNA, referred to as pre $P_{NC}\alpha 1(I)$, is a collagen chain which migrates more slowly on SDS-PAGE than the proto $P_{NC}\alpha 1(I)$ chain. Unhydroxylated $P_{NC}\alpha 1(I)$ migrates more rapidly and resolves clearly from hydroxylated $P_{NC}\alpha 1(I)$ chains in this system. For molecular weight calibration unhydroxylation $P_{NC}\alpha 1(I)$ was used as a standard since the translation products are not hydroxylated. The pre $P_{NC}\alpha 1(I)$ chain has an apparent molecular weight of 165,000 based on the calibration of Fig. 3. Although it is difficult to compare molecules of this size precisely, this is the most direct determination of its molecular weight. Proto $P_{NC}\alpha 1(I)$ has a molecular weight of 154,000, thus the additional pre-sequence would have a molecular weight of $\sim$20,000. The typical signal sequence of 15-30 residues would contribute only $\sim$3000 to the chain weight so it is evident that the precursor contains a large additional sequence, 15,000-18,000 in weight.

The NH_2-terminal piece cleaved by bacterial collagenase from calf skin $P_{NC}\alpha 1(I)$, called col 1 [16], has a molecular weight of about 10,000 [4]. Consequently, if pre $P_{NC}\alpha 1(I)$ is approximately 20,000 daltons larger than $P_{NC}\alpha 1(I)$, pre col 1 should have a molecular weight around 30,000. A 30,000 dalton fragment was not present but there were at least two small collagenase-resistant phenylalanine-containing moieties generated by collagenase digestion. The smaller pieces could have resulted from a second cleavage by collagenase or another uninhibited protease at a region between $P_{NC}\alpha 1(I)$ and the pre NH_2-terminal residues. At this time we cannot rule out the possibility that one of them could be derived from the COOH-terminus. These collagenase-resistant pieces contain substantial amounts of phenylalanine and very little proline. Palmiter et al. [3] found a collagenase-released fragment from the translated prepropeptide to be 5000 daltons greater than

the mass of a peptide from similarly treated $P_{NC}\alpha 1(I)$. As the estimation of the size of this propeptide has varied [3,17-18], and, as indicated above, all the sites of collagenase cleavage are not known, this method may not be reliable. The pre $P_{NC}\alpha 2$ chain was identified and sized by a calibration based on $P_{NC}\alpha 2$ and $\alpha 2$ chains and coincides with the SDS-PAGE component suggested by others [19-22]. It is interesting to note that pre $P_{NC}\alpha 1(I)$ and pre $P_{NC}\alpha 2$ chains appear to have identical molecular weights.

Many secretory proteins have been translated in vitro, and, thus far, all except ovalbumin contain a hydrophobic signal sequence of 15-30 amino acids at the amino-terminus (see ref. 23). Our results indicate that the pro $\alpha 1(I)$ chain collagen also contains a phenylalanine-rich precursor sequence but that this may be much longer than in the other secretory proteins. It is possible that the signal sequence may be comparable in length to that of other secretory proteins, but that procollagen contains still another precursor sequence section immediately adjacent which may function in some other way. Since our preliminary experiments have indicated that the signal peptidase activity of dog pancreas microsomes cleaves pre $P_{NC}\alpha 1(I)$ to the size of proto $P_{NC}\alpha 1(I)$, the signal sequence may be internal rather than amino-terminal.

ACKNOWLEDGEMENTS

Supported by NIH Grant AM13921 from the National Institute of Arthritis, Metabolic and Digestive Diseases.

REFERENCES

1. Milstein, C., Brownlee, G.G., Harrison, T.M. and Mathews, M.D. (1972) Nature (London) New Biol. 239, 117-120.
2. Blobel, G. and Dobberstein, B. (1975) J. Cell Biol. 67, 835-851.
3. Palmiter, R.D., Davidson, J.M., Gagnon, J., Rowe, D.W. and Bornstein, P. (1979) J. Biol. Chem. 254, 1433-1436.
4. Fietzek, P.P. Personal communication.
5. Fietzek, P.P. and Kuhn, K. (1976) Int. Rev. Connect. Tissue Res. 7, 1-60.
6. Olsen, B.R., Guzman, N.H., Engel, J., Condit, C. and Hase, S. (1977) Biochemistry 16, 3030-3036.
7. Hilz, H., Wiegers, U. & Adamietz, P. (1975) Eur. J. Biochem. 56, 103-108.
8. Strohman, R.C., Moss, P.S., Micou-Eastwood, J., Spector, D., Przybyla, A. and Paterson, B. (1977) Cell 10, 265-273.
9. Pelham, H.R.B. and Jackson, R.J. (1976) Eur. J. Biochem. 67, 247-256.

10. Kiely, M.L., McKnight, G.S. and Schimke, R.T. (1976) J. Biol. Chem. 251, 5490-5495.
11. Shields, D. and Blobel, G. (1978) J. Biol. Chem. 254, 3753-3756.
12. Ugel, A.R., Charmbach, A. and Rodbard, D. (1971) Anal. Biochem. 43, 410-426.
13. Bonner, W.M. and Laskey, A.B. (1974) Eur. J. Biochem. 46, 83-88.
14. Laskey, A.B. and Mills, A.D. (1975) Eur. J. Biochem. 56, 335-341.
15. Housman, D., Jacobs-Lorena, M., Rajbhandary, U.L. and Lodish, H.R. (1970) Nature 227, 913-918.
16. Furthmayer, H., Timpl, R., Stark, M., Lapiere, C.M. and Kuhn, K. (1972) FEBS Lett. 28, 247-251.
17. Smith, B.C., McKenney, K.H. and Lustberg, T.S. (1977) Biochemistry 16, 2980-2985.
18. Harwood, R., Merry, A.H., Woolley, D.E., Grant, M.E. and Jackson, D.S. (1977) Biochem. J. 161, 405-418.
19. Rowe, D.W., Moen, R., Davidson, J.M., Byers, P.H., Bornstein, P. and Palmiter, R.D. (1979) Biochemistry 17, 1501-1590.
20. Monson, J.M. and Goodman, H.M. (1978) Biochemistry 17, 5122-5128.
21. Howard, B.H., Adams, S.L., Sobel, M.E., Pastan, I., and de Crombrugghe, B. (1978) J. Biol. Chem. 253, 5869-5874.
22. Tolstochev, P., Haber, R. and Crystal, R.G. (1979) Biochem. Biophys. Res. Commun. 87, 818-826.
23. Kemper, A., Habener, J.F., Mulligan, R.C., Potts, J.T., Jr. and Rich, A. (1974) Proc. Natl. Acad. Sci. U.S.A. 71, 3731-3735.