Vol. 118, No. 1, 1984 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 13, 1984 Pages 261-269

CHARACTERIZATION OF A SHEEP ELASTIN cDNA CLONE CONTAINING TRANSLATED SEQUENCES

Kyonggeun Yoon, Mary May, Norma Goldstein, Zena K. Indik,
*Louise Oliver, *Charles Boyd and Joel Rosenbloom

Center for Oral Health Research, School of Dental Medicine
University of Pennsylvania, Philadelphia, Pa. 19104
*MRC Unit for Molecular and Cellular Cardiology, University
of Stellenbosch Medical School, Tygerberg, 7505, South Africa

Received November 22, 1983

mRNA, isolated from the ligamentum nuchae of fetal sheep by guanidine HCl extraction and oligo(dT) cellulose chromatography, was used to synthesize blunt-ended cDNA molecules by the successive application of AMV reverse transcriptase, DNA polymerase and S1 nuclease. The cDNA was centrifuged on a 15-30% sucrose gradient and molecules greater than 700 bp were tailed with dCTP and cloned into the PstI site of pBR322 which had been tailed with dGTP. Ampicillin-sensitive and tetracycline-resistant colonies were screened by in situ hybridization with elastin-enriched mRNA that had been terminally labeled with ^{32}P. Recombinant plasmids prepared from strongly hybridizing colonies were characterized by restriction mapping and the plasmid with the largest insert (1300 bp) thought to contain elastin sequences was characterized in more detail. The nick-translated cDNA hybridized to a single 3.5 kb mRNA species upon blot hybridization, a size identical to that previously identified for chick elastin mRNA (Burnett et al. (1982) J. Biol. Chem. 259, 1569-1572). Nucleotide sequencing of the 5' end of the cDNA demonstrated a sequence which was extremely GC rich and which corresponded to an amino acid sequence partially homologous to that previously identified in porcine tropoelastin (Foster et al. (1973) J. Biol. Chem. 248, 2876-2879). This is the first report of the identification of a plasmid containing sequences complementary to a translated region of elastin mRNA.

The connective tissue protein elastin is largely responsible for maintaining the elasticity of major blood vessels, lung tissue, the dermis and certain ligaments and cartilage. Within the mature elastin fibers, the individual polypeptide chains are covalently connected by crosslinkages derived from the oxidation of lysine residues (1-5). A polypeptide, designated tropoelastin, with a molecular weight of about 72,000 has been isolated from the aortas of copper-deficient and lathyritic animals (6-8) and appears to be a soluble intermediate in the biosynthesis of the insoluble elastin fiber (9-17). A significant portion of porcine tropoelastin has been sequenced, although the isolated, sequenced peptides have not been ordered (18). Cell-free translation of elastin mRNA has demonstrated that tropoelastin is the primary translation product (19-23).

0006-291X/84 $1.50

A chick elastin cDNA clone containing 3' non-translated sequences (24) has been used to identify and to estimate the relative amounts of elastin mRNA in the developing chick aorta by blot hybridization. A single mRNA species of 3.5 kb hybridized to the cDNA probe and this species increased greatly between day 7 and day 14. When these levels were compared to functional elastin mRNA measured by translation in a rabbit reticulocyte lysate system and to the rate of elastin synthesis in freshly isolated aortas of various ages incubated in vitro, the results suggested that the changes in elastin synthesis seen during development are governed by the elastin mRNA content of the aorta (25, 26). Similar results were found when the levels of functional elastin mRNA in the developing sheep lung and ligamentum nuchae were compared to the rates of elastin synthesis in the tissues (16, 17).

Tropoelastin contains about 830 amino acid residues and therefore approximately 1000 nucleotides of the 3.5 kb elastin mRNA are untranslated. We have been unable to obtain large (greater than 1.0 kb) chick elastin cDNA clones expected to contain translated sequences and have therefore turned to the sheep ligamentum nuchae as a source of elastin mRNA. We here report the construction and characterization of a sheep elastin cDNA clone containing translated sequences.

MATERIALS AND METHODS

Construction of cDNA clones. mRNA was isolated from near-term foetal sheep nuchal ligament using 8 M guanidinium HCl as previously described (17), since this tissue has been shown to be particularly rich in elastin mRNA. Poly(A^+) mRNA was recovered by chromatography on oligo(dT) cellulose (19). The first strand of the cDNA was synthesized essentially as described by Schibler, et al. (27) in a 100 µl reaction volume containing 5 µg of Poly(A^+) mRNA, previously heated to 95° for 1 min and quenched in an ice bath, 0.5 mM of each of the 4 dNTPs, 250 µCi [α^{32}p] dCTP (3000 Ci/mMole), 5 mM DTT, 50 mM KCl, 6 mM $MgCl_2$, 50 mM Tris HCl, pH 8.3, 5 µg oligo(dT) and 50 U of AMV reverse transcriptase (obtained from J. Beard). Incubation was for 60 min at 37° and the reaction was terminated by the addition of EDTA to a final concentration of 10 mM. The RNA was degraded by the addition of NaOH to a final concentration of 0.32 M and incubation at 42° for 10 min. The solution was neutralized by the addition of Tris, pH 7.5, to a final concentration of 0.1 M and 1 M HCl until neutrality was reached using phenol red as indicator. The cDNA was desalted on a G-50 Sephadex column which was eluted with water and the fractions containing the cDNA were lyophilized. Second strand synthesis was carried out in 300 µl containing 0.5 mM dNTPs, 70 mM KCl, 30 mM Tris, pH 7.5, 4 mM $MgCl_2$, 0.5 mM 2 β-mercaptoethanol, and 25 U of the Klenow fragment of E. coli polymerase I (Bethesda Research Laboratories). Incubation was for 4.5 hr at 18°. The reaction was stopped by the addition of EDTA to a final concentration of 10 mM and the dsDNA desalted on a G-50 Sephadex column eluted with water. The fractions containing the DNA were lyophilized and dissolved in a total volume of 300 µl

containing 300 mM NaCl, 30 mM Na Acetate, pH 4.5, 3 mM $ZnCl_2$. 4 Units of S1 nuclease (Boehringer-Mannheim) were added and the digestion carried out at 37° for 45 min. The reaction solution was extracted with an equal volume of chloroform/phenol and the dsDNA was fractionated on a 15-30% sucrose gradient. Molecules greater than 700 bp were tailed with dCTP using terminal transferase and hybridized to PstI-cut pBR322 tailed with dGTP as described previously (24). The recombinant plasmids were used to transform E. coli MC 1061 and transformants were selected by growth on agar plates containing 5 μg/ml tetracycline.

Screening of transformed colonies. The tetracycline-resistant colonies were transferred to agar plates containing 20 μg/ml of ampicillin and incubated overnight at 37°. The ampicillin-sensitive, tetracycline-resistant colonies were screened by the colony hybridization method of Grunstein and Hogness (28) using ^{32}P ligamentum nuchae mRNA which had been prepared in the following way. One μg mRNA was partially hydrolyzed in 0.1 M NaOH at 4° for 30 min, and the solution then neutralized. ^{32}P was added by incubating the mRNA with 100 μCi of [γ ^{32}P] ATP in 50 mM Tris, pH 7.6, 10 mM $MgCl_2$, 1.5 mM spermine, 5 mM DTT and 5 units of T 4 polynucleotide kinase (PL Biochemicals) for 1 hr at 37°. The labeled RNA (2 x 10^6 cpm/ μg) was extracted with phenol/chloroform and separated from the ATP by Sephadex G-50 chromatography. Hybridization was carried out overnight at 42° in 70% formamide, 0.4 m NaCl. Colonies which hybridized well on duplicate filters were selected for further screening.

DNA sequencing. Transformed bacteria were grown in one liter of M-9 medium and plasmids were amplified by the addition of chloramphenicol, 170 μg/ml. Plasmid DNA was isolated by the alkaline method of Birnboim and Doly (29) followed by equilibrium centrifugation in CsCl, then restricted with PstI and the fragments isolated by electroelution after electrophoresis in 1% agarose gels. The DNA was extracted with phenol/chloroform and then cloned into the PstI site of the phage M13mp8 (30). Restriction fragments derived from the cDNA inserts now contained in the recombinant phage were sequenced by the Sanger dideoxy method (31) using a universal primer of 17 nucleotides (Collaborative Research, Cambridge, Mass).

Nick translation and blot hybridization. The recombinant plasmid, pcSEL1, was labeled with ^{32}P to a specific activity of 4 x 10^8 cpm/ μg DNA using the kit supplied by Bethesda Research Laboratories. Ligamentum nuchae poly(A^+) RNA was electrophoresed on 0.8% agarose gels after treatment with 1 M freshly deionized glyoxal, transferred to nitrocellulose filters and hybridized to 10^8 cpm of the probe for 24 h as described by Thomas (32). After washing, the blot was exposed to x-ray film at -70° using an intensifying screen.

RESULTS AND DISCUSSION

Seven hundred ampicillin-sensitive, tetracycline-resistant transformed colonies were screened for sequences complementary to elastin mRNA by the method of Grunstein and Hogness (28). The hybridization probe was prepared by end-labeling partially hydrolyzed ligamentum nuchae mRNA ^{32}P, as described in Materials and Methods, since such mRNA preparations have been shown to be significantly enriched in elastin mRNA (17). Plasmids containing inserts complementary to elastin mRNA would, therefore, be expected to hybridize more strongly than most other plasmids. Thirty colonies which gave a strong signal

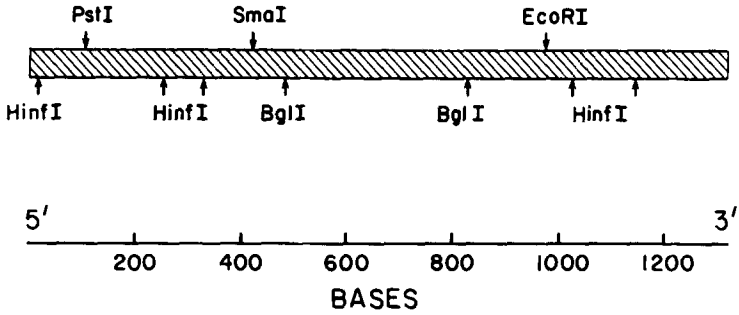

Fig. 1. Restriction enzyme map of the sheep elastin cDNA plasmid, pcSEL1. This is the largest plasmid obtained. Other smaller plasmids shared restriction sites starting at the 3' end. No Hind III, KpnI, Sal, SacI, Xba or BamHI sites were found.

on duplicate filters were chosen for further study (data not shown) and plasmids prepared from them in larger amounts. Although the cDNA used in the construction of the cDNA library had been sized, many of these plasmids contained inserts which were less than 800 bp and in others one or both PstI sites had been lost. Further analysis of these plasmids was deferred. Restriction-enzyme mapping of the remaining plasmids revealed a group which contained restriction sites in common (Fig.1). The largest of these, designated pcSEL1, which contained a 1300 bp insert, was selected for detailed characterization. First, in order to determine whether this plasmid hybridized to a mRNA comparable in size to that previously identified as elastin mRNA in chick, it was nick translated and used as a probe in a blot hybridization experiment. Fig. 2 illustrates that pcSEL1 hybridized to a single mRNA of 3.5 kb (24 S) isolated from ligamentum nuchae and that no other higher molecular weight species was detected. This result suggested that this plasmid did in fact contain elastin cDNA sequences.

The insert fragments from a PstI digestion of pcSEL1 were isolated by electroelution from a preparative agarose gel and cloned into the PstI site of the phage M13mp8. Recombinant phage were grown up and single-stranded DNA prepared for sequencing by the Sanger dideoxy method (31). We here report the nucleotide sequence of the small fragment of pcSEL1, while the sequence of the larger fragment will be reported elsewhere when completed. Both strands of the smaller fragment were sequenced since recombinant phage containing inserts in both orientations were readily obtained. Fig. 3 illustrates a sequencing gel from which the complete sequence of both strands can be read. When the non-coding

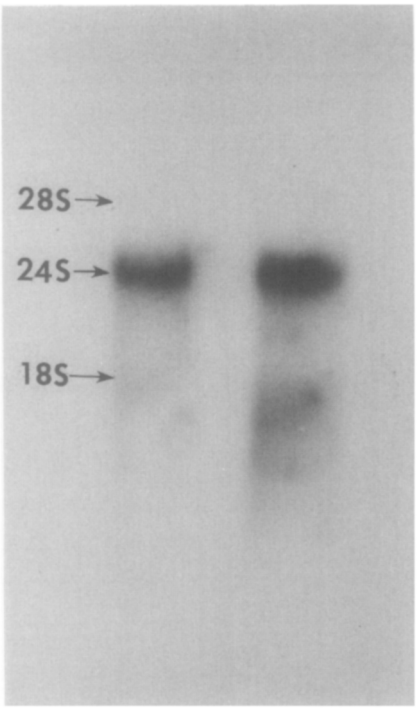

Fig. 2. Blot hybridization of ligamentum nuchae poly(A+) RNA. 10 μg of two different batches of RNA were subject to electrophoresis and transfer to a nitrocullulose filter as described in Materials and Methods (32). The filter was then hybridized to ^{32}P-labeled pcSEL1, and the filter washed and autoradiographed. 18 S and 28 S ribosomal RNA isolated from reticulocytes was run in adjacent lanes and stained with acridine orange.

strand was used as template so that the sequence determined was 5' to 3' with respect to the coding strand, little difficulties were encountered. However, when the coding strand served as template there was nonspecific termination at multiple positions resulting in ambiguities (marked by arrows in Fig. 3). Elastin is largely composed of amino acids (glycine, alanine, proline) whose codons are very GC rich which may lead to secondary structure resulting in termination when the coding strand is used as template for DNA polymerase. This tendancy to terminate may also explain the substantial difficulty there has been in constructing large elastin cDNA clones using reverse transcriptase and DNA polymerase.

The sequence of the 107 bp fragment is given in Fig. 4 along with the corresponding correct-frame translated amino acid sequence. Also given is the sequence of one of the tryptic peptides obtained from porcine tropoelastin

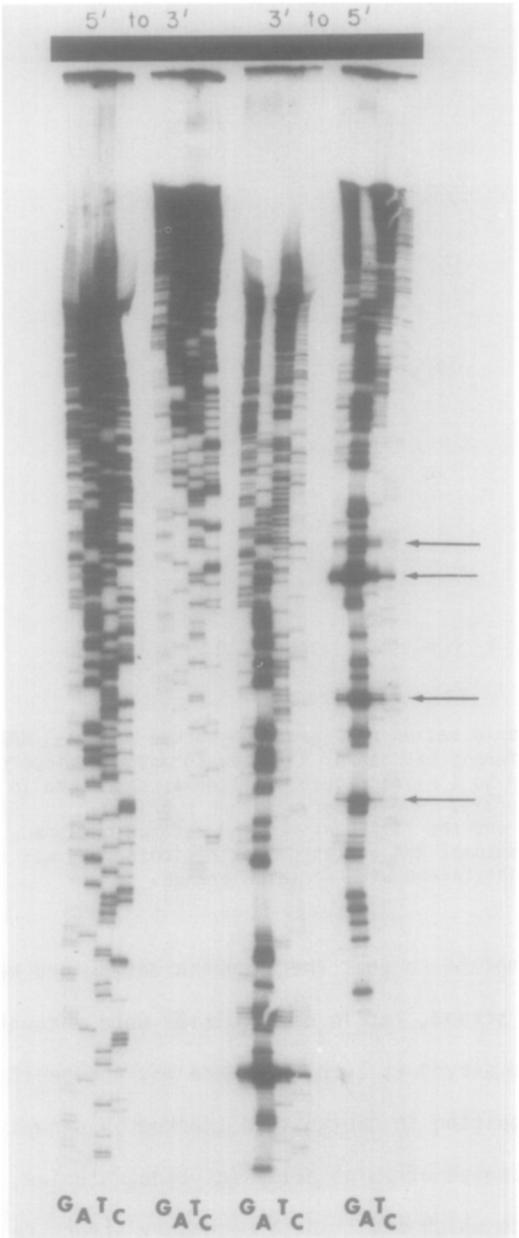

Fig. 3. Sequencing gel of the 107 bp fragment of pcSEL1. The small fragment of pcSEL1 was cloned into the PstI site of the phage M13mp8 (30) and single-stranded DNA prepared from recombinant phage. Sequencing of both DNA strands by the dideoxy method of Sanger (31) is illustrated. The 5' to 3' and 3' to 5' designations refer to the directions with respect to the messenger strand. Arrows indicate positions of non-specific termination.

(18), since no sheep elastin sequences are available. An identical sequence of sixteen amino acids is found in the translated sheep and porcine sequences using one of the two possible amino acids where ambiguities were located in the

	1			10			20			
	AA	TTT	GGC	CTG	GGG	GGA	GTC	GGT	GGG	CTC
SHEEP		PHE	GLY	LEU	GLY	GLY	VAL	GLY	GLY	LEU
PIG				ALA	ALA	GLU	PHE	GLY	VAL	GLY
	30				40			50		
		GGA	GTT	GGA	GGA	CTG	GGA	GTT	GGT	GGG
SHEEP		GLY	VAL	GLY	GLY	LEU	GLY	VAL	GLY	GLY
PIG		GLY	VAL	GLY	GLY	LEU	GLY	VAL	GLY	GLY
		60				70			80	
		CTC	GGA	GCT	GTC	CCA	GGG	GCT	GTG	GGC
SHEEP		LEU	GLY	ALA	VAL	PRO	GLY	ALA	VAL	GLY
PIG		LEU	GLY	ALA	VAL	PRO	GLY	ALA	GLY	
				90			100			
		CTT	GGA	GGT	GTG	TCT	CCA	GCT	GCA	
SHEEP		LEU	GLY	GLY	VAL	SER	PRO	ALA	ALA	

Fig. 4. Nucleotide sequence and corresponding translated amino acid sequence of the 107 bp fragment of pcSEL1. Also shown for comparison is a portion of a comparable amino acid sequence found in porcine tropoelastin (18), as determined by sequencing of tryptic peptides. Residues identical between species are underlined.

published porcine sequence. This amino acid sequence, therefore, unequivocally identifies the cDNA clone as being derived from elastin. Several aspects of this nucleotide sequence are of interest. Remarkably, 50 of the 107 bases are guanine and another 20 are cytosine, a result which serves to emphasize the possible problem of secondary structure. The total GC content is comparable to that found in the mRNA for the triple helical portion of the pro α chains of Type I collagen. However, in contrast to the glycine codon usage found in collagen in which there is a strong preference for a pyrimidine in the third position and very infrequent use of guanine (for review, see 33), in the sheep elastin sequence a purine is found in 11 of the 16 glycine codons and of these, guanine is found 4 times. Thus, 25% of the observed glycine codons are GGG including two adjacent ones. These observations suggest that at least short runs of guanine can be tolerated in mRNA and do not lead to intractable secondary structures.

This is the first report of cloned elastin cDNA containing translated sequences. The present results suggest that there is strong conservation of amino acid sequences between mammalian elastins. Since the elastin mRNA contains about 1000 untranslated nucleotides and since it is likely that most of these are at the 3' end, large cDNA clones are required to obtain cloned translated sequences using oligo(dT) as primer. In order to obtain cDNA clones which contain sequences which are further 5', we plan to use restriction frag-

ments isolated from the present plasmid, pcSEL1, in primer extension experiments. Hopefully this should facilitate the ordering of the rather large amount of sequence data available for tryptic peptides of porcine tropoelastin.

Acknowledgements. This investigation was supported by the National Institutes of Health Research Grants AM-20553, DE-02623 and HL-29702, and grants from the South African Medical Research Council and the National Cancer Association.

REFERENCES

1. Partridge, S. M. (1962) Adv. Protein Chem. 17, 227-297.
2. Franzblau, C., Sinex, F. M., Faris, B. and Lampidis, R. (1965) Biochem. Biophys. Res. Commun. 21, 525-581.
3. Partridge, S. M., Elsden, D. F., Thomas, J., Dorfman, A., Telser, A. and Ho, P. (1966) Nature 209, 399-400.
4. Miller, E. J., Martin, G. R., Mecca, C. H. and Piez, K. A. (1965) J. Biol. Chem. 240, 3623-3627.
5. Pinnell, S. R. and Martin G. R. (1968) Proc. Natl. Acad. Sci. U. S. A. 61, 708-716.
6. Sandberg, L. B., Weissman, N. and Smith, D. W. (1969) Biochemistry 8, 2940-2945.
7. Rucker, R. B., Gottlich-Reimann, W. and Tom, K. (1973) Biochem. Biophys. Acta 317, 193-201.
8. Foster, J. A., Shapiro, R., Voynow, D, Crombie, G., Faris, B. and Franzblau, C. (1975) Biochemistry 14, 5343-5347.
9. Murphy, L., Harsch, M., Mori, T. and Rosenbloom, J. (1972) FEBS Lett. 21, 113-117.
10. Narayanan, A. S. and Page, R. C. (1976) J. Biol. Chem. 251, 1125-1130.
11. Bressan, G. M. and Prockop, D. J. (1977) Biochemistry 16, 1406-1412.
12. Rosenbloom, J. and Cywinski, A. (1976) Biochem. Biophys. Res. Commun. 69, 613-620.
13. Uitto, J., Hoffman, H-P. and Prockop, D. J. (1976) Arch. Biochem. Biophys. 173, 187-200.
14. Smith, D. W. and Carnes, W. H. (1973) J. Biol. Chem. 248, 8157-8161.
15. Rosenbloom, J., Harsch, M. and Cywinski, A. (1980) J. Biol. Chem. 255, 100-106.
16. Shibahara, S., Davidson, J. M., Smith, K. and Crystal, R. G. (1981) Biochemistry 20, 6577-6584.
17. Davidson, J. M., Smith, K. Shibahara, S., Tolstoshev, P. and Crystal, R. G. (1982) J. Biol. Chem. 257, 747-754.
18. Foster, J. A., Bruenger, E., Gray, W. R. and Sandberg, L. B. (1973) J. Biol. Chem. 248, 2876-2879.
19. Burnett, W. and Rosenbloom, J. (1979) Biochem. Biophys. Res. Commun. 86, 478-484.
20. Burnett, W., Eichner, R. and Rosenbloom, J. (1980) Biochemistry 19, 1106-1111.
21. Foster, J. A., Rich, C. B., Fletcher, S., Karr, S. R. and Przybyla, A. (1980) Biochemistry 19, 857-864.
22. Karr, S. R. and Foster, J. A. (1981) J. Biol. Chem. 256, 5946-5949.
23. Davidson, J. M., Leslie, B., Wolt, T., Crystal, R. G. and Sandberg, L. B. (1982) Arch. Biochem. Biophys. 218, 31-37.
24. Burnett, W., Yoon, K. and Rosenbloom, J. (1981) Biochem. Biophys. Res. Commun. 99, 364-372.
25. Burnett, W., Finnigan-Bunick, A., Yoon, K. and Rosenbloom, J. (1982) J. Biol. Chem. 257, 1569-1572.

26. Eichner, R. and Rosenbloom, J. (1979). Arch. Biochem. Biophys. 198, 414-423.
27. Schibler, U., Marcu, K. B. and Perry R. P. (1978) Cell 15, 1495-1509.
28. Grunstein, M. and Hogness, D. S. (1975) Proc. Natl. Acad. Sci., U.S.A. 72, 3961-3968.
29. Birnboim, H. C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523.
30. Heidecker, G., Messing, J. and Gronenborn, B. (1980) Gene 10, 69-73.
31. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
32. Thomas, P. S. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 5201-5205.
33. Tate, V., Finer, M., Boedtker, H. and Doty, P. (1983) Cold Spring Harbor Symposium Quant. Biol. 47, 1039-1049.