Biochemistry

 Volume 24, Number 13 June 18, 1985

Accelerated Publications

Structure of the 3′ Portion of the Bovine Elastin Gene†

George Cicila, Mary May, Norma Ornstein-Goldstein, Zena Indik, Scott Morrow, Helena S. Yeh, Joan Rosenbloom, Charles Boyd,‡ Joel Rosenbloom, and Kyonggeun Yoon*

Department of Anatomy and Histology and Center for Oral Health Research, School of Dental Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received March 25, 1985

ABSTRACT: A bovine genomic library constructed by partial *Sau*3A digestion and contained in λ Charon 30 was screened by in situ hybridization with a 1.3-kilobase (kb) sheep elastin cDNA clone [Yoon, K., May, M., Goldstein, N., Indik, Z., Oliver, L., Boyd, C., & Rosenbloom, J. (1984) *Biochem. Biophys. Res. Commun. 118*, 261–269]. Three clones encompassing 10 kb of the bovine elastin gene were identified and characterized by restriction mapping and DNA sequencing of the 6.2 kb of the most 3′ region of the gene. These analyses have permitted localization of eight exons in the 6.2 kb in which the translated exons vary in size from 27 to 69 base pairs, and there is an approximately 1-kb untranslated region at the 3′ end. In addition to identification of sequences homologous to those found in porcine tropoelastin, the analyses defined a 58 amino acid sequence that forms the carboxy-terminal region of tropoelastin, and this sequence, which contains two cysteine residues, was previously not observed in the protein sequence data. The analyses also suggest that functionally distinct cross-link and hydrophobic domains of the protein are encoded in separate exons.

Elastin is a unique protein in the vertebrate kingdom, imparting to the tissues that contain it the property of elastic recoil. The individual polypeptide chains in the elastic fibers are covalently connected by cross-linkages derived from the oxidation of lysine residues by a copper-requiring enzyme, peptidyl lysyl oxidase (Partridge, 1962; Franzblau et al., 1965; Miller et al., 1965; Pinnell & Martin, 1968). The extensive cross-linking results in great insolubility, and significant progress in the primary structure determination of elastin came about only after the isolation of a soluble polypeptide, designated tropoelastin (M_r ~72 000), from copper-deficient or lathyritic animals (Sandberg et al., 1969; Rucker et al., 1973; Foster et al., 1975). Although many of the tryptic peptides derived from tropoelastin have been sequenced, they have not been ordered (Foster et al., 1973; Sandberg et al., 1985). Cell-free translation of elastin mRNA has demonstrated that tropoelastin is the primary translation product (Burnett & Rosenbloom, 1979; Burnett et al., 1980; Foster et al., 1980; Karr & Foster, 1981; Davidson et al., 1982a,b), and biosynthetic labeling studies have strongly indicated that tropoelastin is a true intermediate in the formation of insoluble elastin (Murphy et al., 1972; Narayanan & Page, 1976; Bressan & Prockop, 1977; Rosenbloom & Cywinski, 1976; Smith & Carnes, 1973; Rosenbloom et al., 1980).

Measurements of elastin mRNA levels by in vitro translation and by hybridization with specific probes have suggested that the changes in elastin synthesis seen during embryonic development of the chick or sheep are largely governed by the mRNA content of the tissue (Burnett et al., 1980; Foster et al., 1980; Shibahara et al., 1981; Davidson et al., 1982a,b; Burnett et al., 1982; Davidson et al., 1984a,b). Northern hybridization analysis has demonstrated that elastin mRNA contains about 3500 nucleotides (Burnett et al., 1982; Yoon et al., 1984; Davidson et al., 1984a,b). Since only about 2550 nucleotides are necessary to code for tropoelastin, approximately 950 nucleotides are untranslated. Recent sequence analyses of a 1.3-kb sheep elastin cDNA clone and of the corresponding untranslated region of a sheep genomic clone have demonstrated that the 3′ end of the mRNA contains a large untranslated region (Yoon et al., 1984; Yoon et al., unpublished observations).

In this paper we report the isolation and characterization of genomic clones encompassing 10 kb of the 3′ portion of the

†Supported by National Institutes of Health Grant AM-20553. G.C. was supported by NIH Training Grants DE-07085 and HD-07067.

‡Present address: Department of Surgery, Academic Health Science Center, New Brunswick, NJ 08903.

0006-2960/85/0424-3075$01.50/0

bovine elastin gene. The 6.2 kb of DNA comprising the most 3′ portion of the gene have been sequenced, permitting precise determination of the exon–intron structure. In addition to identification of sequences homologous to those found in porcine tropoelastin, these analyses have elucidated the amino acid sequence of 58 amino acids ending with the carboxy terminus, which was not previously observed in the protein sequencing of tropoelastin. The analyses also suggest that functionally distinct cross-link and hydrophobic domains are encoded in separate exons. The translated exon to intron ratio was found to be extremely low (1:15).

Materials and Methods

Isolation and Restriction Mapping of Genomic Clones. A bovine genomic library, constructed by the insertion of genomic DNA partially digested with *Sau*3A into the *Bam*HI sites of λ Charon 30, was screened with a sheep elastin 1279-bp cDNA clone, pcSEL1 (the bovine library was a generous gift from Dr. Dan Capon of Genentech, Inc.). A total of 4×10^5 plaque-forming units (pfu) was screened on *Escherichia coli* LEZ grown in 150-mm Petri dishes (2×10^4 pfu/dish) by the in situ hybridization method (Benton & Davis, 1977). Only positive plaques occurring on duplicate filters were purified to homogeneity by several cycles of plating and hybridization. Phage were grown in liquid culture, and the DNA was isolated by successive steps of chloroform extraction, CsCl gradient centrifugation, proteinase K digestion, phenol extraction, and alcohol precipitation (Maniatis et al., 1982). After the DNA was digested with restriction nucleases according to the direction of the suppliers (BRL and New England Biolabs), the fragments were electrophoresed on 1% agarose gels.

DNA Sequencing. The *Eco*RI/*Bam*HI restriction fragments were isolated by electroelution after electrophoresis on 1% agarose gels, and after phenol/chloroform extraction the DNA was cloned into the phage M13mp8,9 system (Heidecker et al., 1980). The inserts were then sequenced by the Sanger dideoxy method (Sanger et al., 1977) using a universal primer of 17 nucleotides (Collaborative Research, Cambridge, MA) or oligonucleotides 15–16 bases in length synthesized in our laboratory.

Synthesis of Oligonucleotides. The oligonucleotides were synthesized by using a modification of the phosphite method of Matteucci & Caruthers (1981), employing a Genetic Design (Watertown, MA) programmable machine. Terminal, 3′-blocked nucleotides bound to control pore glass and activated, blocked nucleotides were purchased from American BioNuclear (Emeryville, CA) and used without further purification. The synthesized oligonucleotides were purified by reverse-phase chromatography on a Varian 2000 high-pressure liquid chromatograph.

Results

Restriction Endonuclease Mapping of Genomic Clones. Three independent clones, designated BEL1, BEL2, and BEL3, which hybridized strongly to the sheep elastin cDNA clone, pcSEL1, were purified by successive cycles of plating and hybridization, and *Eco*RI and *Bam*HI restriction digests of the purified DNA were subjected to agarose gel electrophoresis (gels not shown). The restriction map for the three inserts in the recombinant phage, which was deduced from these studies, is illustrated in Figure 1. The clones contained 10.3, 15.5, and 12.0 kb of insert, respectively, and BEL2 and BEL3 were seen to overlap BEL1 but to extend past it on the 3′ side. In order to determine which restriction fragments hybridized to the cDNA clone, Southern analysis was carried out. Two

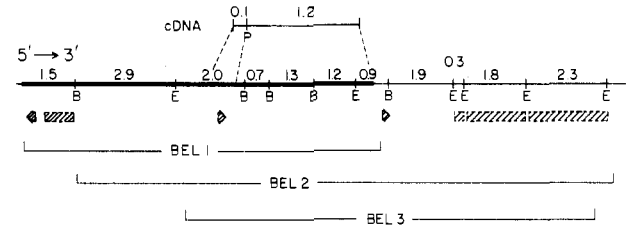

FIGURE 1: Restriction map of bovine elastin genomic clones BEL1, BEL2, and BEL3. Approximately 10 kb of the elastin gene is indicated as a heavy line. The position of the cDNA clone, pcSEL1, is shown. Symbols: E, *Eco*RI; B, *Bam*HI; cross-hatched arrows, repetitive sequences determined by sequencing; cross-hatched boxes, indication that the restriction fragment contains repetitive sequences somewhere within it.

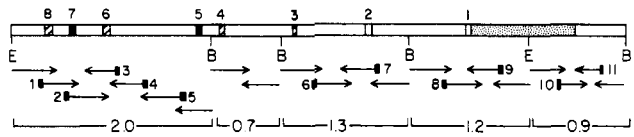

FIGURE 2: DNA sequencing strategy and exon–intron structure of the bovine elastin gene. All sequencing was performed by using the dideoxy method with fragments cloned into the M13mp8,9 system with a universal primer except where indicated by small black numbered boxes at the tail of the arrow. These boxes indicate priming with specific synthetic oligonucleotides reading 5′ → 3′ as follows: 1, GCCTAGTGACCTTGG; 2, GACTGACCTGTGGAGA; 3, TGATTCCAGAACACA; 4, GGGATCGTCTCGACCC; 5, GTGTGACCAGACTTG; 6, CCATGGGAGAATCAC; 7, GGTCCCTGGCCCACT; 8, GATTTGAAACAAGGT; 9, CCTAGAGATGGAGAG; 10, CCATCTCATGTGTCT; 11, AGGAGGCCAAGACACA. Symbols: (cross-hatched boxes) hydrophobic exon; (■) cross-link exon; (□) exon not designated as hydrophobic or cross-link; (dotted box) untranslated portion of exon 1.

nick-translated probes were utilized, the entire pcSEL1 and the 107 bp most 5′ portion of pcSEL1 that contained coding sequences exlusively. The 2.0-, 0.7-, 1.3-, 1.2-, and 0.9-kb fragments all hybridized to the complete pcSEL1, indicating that all these fragments contained exon sequences. However, only the 2.0-kb fragment hybridized to the 107-bp 5′ fragment, which permitted orientation of the genomic clones as indicated in Figure 1. None of the fragments located at the 3′ side of the 0.9-kb fragment hybridized to pcSEL1. Subsequent DNA sequencing (see below) confirmed that the 0.9-kb fragment contained two polyadenylation consensus sequences, AATAAA, demonstrating that the end of the elastin gene was contained within the 0.9-kb fragment.

DNA Sequence Analysis and Overall Structure of the Gene. All the *Eco*RI/*Bam*HI fragments shown in Figure 1 were cloned into the M13mp8,9 system in order to obtain the restriction fragments in both orientations. Sequencing was then performed by the Sanger dideoxy method using a universal primer of 17 nucleotides (Collaborative Research, Inc.) and in addition through the use of specific synthesized oligonucleotides. Figure 2 illustrates the sequencing strategy used for the 2.0-, 0.7-, 1.3-, 1.2-, and 0.9-kb fragments. The sequence of approximately 350–400 nucleotides could be determined accurately for each end of the fragments with the universal primer. We then synthesized an oligonucleotide starting at about 325 bp from the end and used this oligonucleotide to prime the same M13 clone to determine sequences further from the universal starting point. This process was repeated until the sequencing of each strand overlapped its complement strand, completing the sequencing of all the fragments. In Figure 2, each arrow indicates the direction and extent of sequence determination from a given starting point, using universal primer (unmarked tail of the arrow) or specific oligonucleotides (black numbered boxes at the tail). In this

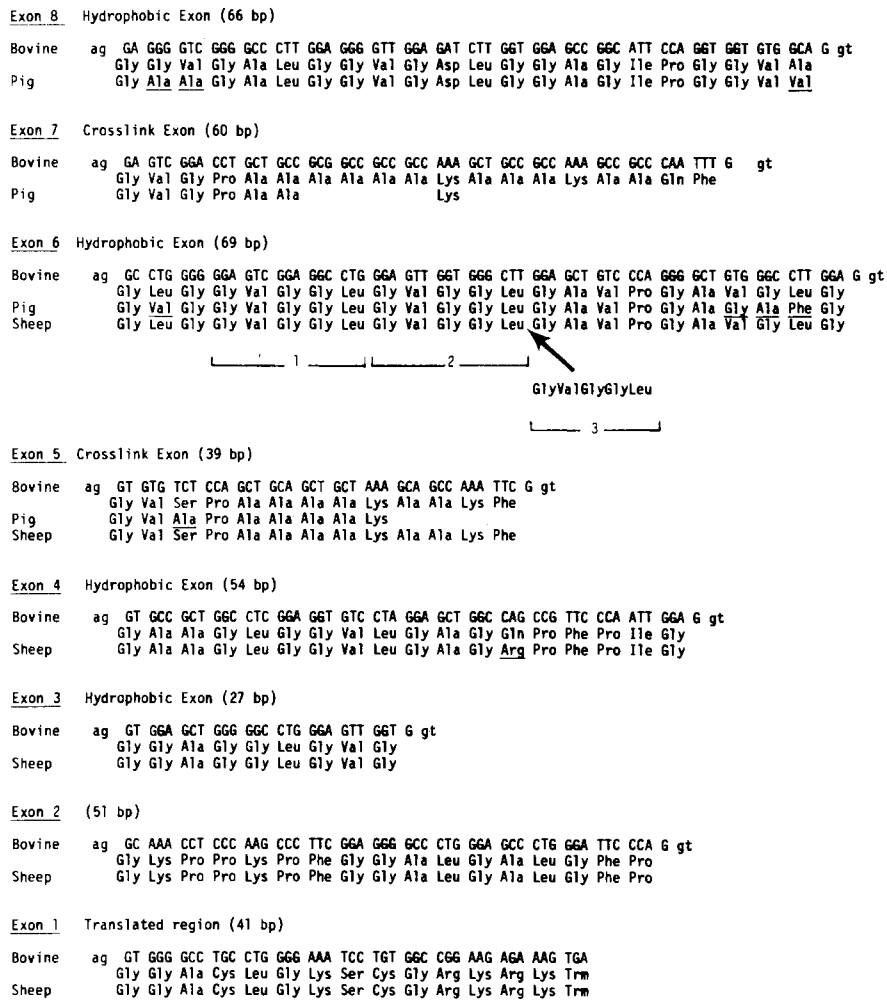

FIGURE 3: Nucleotide sequences and translated amino acid sequences for the eight exons at the 3′ portion of the bovine elastin gene. The untranslated portion of exon 1 is not shown. Corresponding sequences for pig and sheep elastin are shown where available. Differences from the bovine sequences are underlined or indicated by an insertion (arrow).

fashion, we have completed DNA sequencing of 6.2 kb of the 3′ end of the bovine elastin gene.

Exons were located mainly by homology with pcSEL1, which contains 305 bp of coding sequences (Yoon et al., 1984). The bovine genomic sequences located further 5′ to the end of pcSEL1 were translated into the three possible reading frames and compared to the available peptide sequence data for porcine tropoelastin. In this manner we were able to locate eight exons in the 6.2 kb of bovine genomic DNA. The positioning of the exons, which are numbered from the 3′ end of the gene, is indicated in Figure 2. A striking feature of this portion of the gene is the small size of the translated exons (27–69 bp), which are interspersed in large expanses of introns. The intron to exon coding ratio is about 15:1, which is very large in comparison to most other known genes. If the ratio is maintained, the gene will be over 40 kb. Although the exons are all multiples of three and glycine is found consistently at the exon–intron junctions, the exons are not regular in size, unlike some collagen genes (Boedtker et al., 1983). Another interesting feature of the elastin gene is that coding sequences corresponding to hydrophobic and cross-link domains of the protein are found in separate exons, as indicated by cross-hatched and filled boxes in Figure 2.

Detailed Analysis of Exons. It is convenient to describe the DNA sequences by beginning with the most 5′ region since the corresponding amino acid sequence will then follow in order (Figure 3). Exons 8 and 7, not contained in the cDNA, were identified by comparison to known sequences in porcine tropoelastin (Foster et al., 1975; Sandberg et al., 1985). Exon 8 is a hydrophobic exon, and there is very strong homology between the two species. It may be noted that within exon 8 there is an imperfect hexapeptide repeat of the form GVGALG. Exon 7 is a cross-link exon containing two lysyl residues separated by and flanked by alanyl residues. The exon begins with a short hydrophobic sequence terminated by a prolyl residue (GVGP) and ends with a QF sequence. The pig sequence shown for comparison is a continuation of the pig sequence illustrated in exon 8, so that the difference in the number of alanines between the bovine and pig sequences is certain.

Exon 6 contains a repeating pentapeptide (GVGGL) that is repeated 2 times in both bovine elastin and porcine elastin but 3 times in sheep elastin. It has been postulated that these repeating units can confer an unusual conformation called a β-spiral, which is created from a number of β-turns in succession in portions of the tropoelastin molecule (Venkatachalam & Urry, 1981). Differences found in the number of these repeating units among species suggest that the function of the molecule can tolerate variation in the extent of the β-spiral conformation or that elastin exists largely in a rubberlike random-coil conformation. Cross-link exon 5 contains a short leader sequence (GVSP) on the amino side but is terminated by a phenylalanyl residue rather than by QF as found in exon 7.

The remainder of the translated sequences found in exons 1–4 have not been found in porcine tropoelastin except for a

First Position	Second Position U	C	A	G	Third Position
U	F 2	S 1	Y	C 1	U
	F 4	S 1	Y	C 1	C
	L	S			A
	L	S		W	G
C	L 4	P 2	H	R	U
	L 1	P 2	H	R	C
	L 1	P 6	Q 1	R	A
	L 6	P	Q 1	R 1	G
A	I 2	T	N	S	U
	I	T	N	S	C
	I	T	K 6	R 1	A
	M	T	K 3	R	G
G	V 2	A 10	D 1	G 10	U
	V 5	A 15	D	G 9	C
	V	A 3	E	G 19	A
	V 3	A 1	E	G 10	G

FIGURE 4: Codon usage in the bovine elastin gene. The numbers beside the single-letter designations for the amino acids are the values for the use of the particular codon in the translated DNA sequences determined to date. Note the usage of purines in the third position for glycine in contrast to the case of collagen (Boedtker et al., 1983).

short segment (IGGKPPKPG), which has recently been identified in a thermolysin digest (Sandberg et al., 1985). Interesting features of this carboxy-terminal region of the protein that have been deduced from the DNA sequence are (1) a sequence (VGGKPPKPF) that contains paired lysyl residues separated by prolyl residues rather than alanine but also terminated by phenylalanine, suggesting that the lysyl residues may participate in cross-link formation, (2) the presence of two cysteine residues located at positions 6 and 11 from the putative carboxy terminus, and (3) the exceedingly basic nature of the terminus since the four terminal residues are basic and another lysine is located at position 8 from the terminus.

Codon Usage. One of the major features of this portion of the elastin gene is its high content of guanine and cytosine, amounting to 42% guanine and 25% cytosine for the coding strand when all the translated sequences determined so far are analyzed. When the hydrophobic exons are analyzed separately, the remarkably high guanine content is 53% and the cytosine content is 14% for the coding strand. This increased guanine content is a reflection of the high content of amino acids such as glycine, alanine, and valine whose codons use guanine to an increased extent. Figure 4 gives the complete codon usage for the sequences determined to date and demonstrates that the use of guanine at the wobble position for glycine is significantly greater in elastin (GGG; 10/48) than in other mammalian genes including collagen (Boedtker et al., 1983; Hastings & Emerson, 1983).

Intron Structure. The intron–exon splice junctions conform to the consensus sequences for eukaryotic genes in which introns begin with GT and terminate in AG. As demonstrated in Figure 5, the junctions in the elastin gene conform to the sequence CAG-exon-GTPuPuGT. Furthermore, in four of seven introns, the first six bases at the 5′ end of the introns are identical. These sequences in the splice junctions are complementary to the 5′ end of U1 small nuclear RNA, supporting the concept that U1 may be involved in mRNA splicing (Rogers & Wall, 1980). The splice sites divide glycine codons, unlike the case of fibrillar collagen in which the glycine codons remain intact (Boedtker et al., 1983).

Remarkably, the introns are GC rich to nearly the same extent (60% GC) as the exons, and they contain several

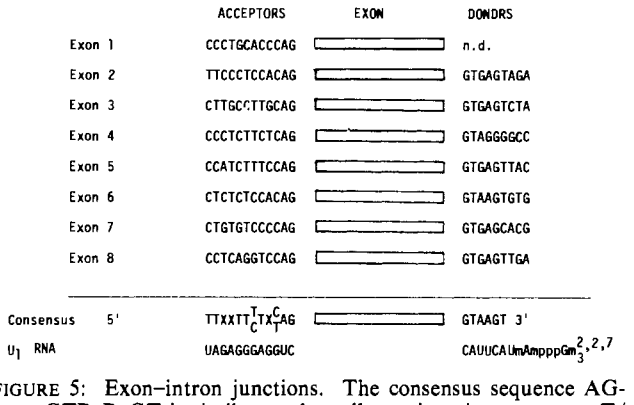

FIGURE 5: Exon–intron junctions. The consensus sequence AG-exon-GTPuPuGT is similar to the collagen junction sequence, T/CAG-exon-GTAAGT (Boedtker et al., 1983).

elastin-like sequences such as GPGGVGALGG or GGAGG. However, these sequences do not contain the consensus exon–intron splice junctions and are not present in the sheep elastin cDNA, which was used to locate the majority of the exons. They may represent the remnants of unspliced exons during evolution.

In order to determine whether repetitive sequences were found in the isolated genomic clones, the clones were restricted with *Eco*RI and *Bam*HI and the restriction fragments subjected to Southern analysis using nick-translated total bovine genomic DNA as probe (blots not shown). Of all the fragments shown in Figure 1, six fragments, 1.5, 2.0, 0.9, 0.3, 1.8, and 2.3 kb, hybridized to the genomic DNA, as indicated by cross-hatched boxes. The repetitive sequences found inside the elastin gene in the 1.5- and 2.0-kb fragments have been sequenced, which demonstrated that they were strongly homologous to Alu-like sequences found in the bovine corticotropin–β-lipotropin precursor gene and bovine fetal and adult globin genes (Watanabe et al., 1982; Schimenti & Duncan, 1984). These Alu-like sequences, varying somewhat in size, are approximately 120 bp long and are represented about 100000 times in the bovine genome. Thus, the occurrence of three such sequences within the 10 kb of the elastin gene suggests that the gene may contain more repetitive sequences than other genes. A more detailed analysis of the repetitive sequences is in preparation.

Discussion

It became apparent from the hybridization experiments with pcSEL1 and preliminary DNA sequence data that the exon coding to intron ratio was very low and that many of the individual exons were quite small. Electron microscopic heteroduplex analysis of a segment of the sheep elastin gene, which probably corresponds to at least a portion of the bovine gene analyzed here, also indicated a small exon to intron ratio (Davidson et al., 1984b), but it was not possible to estimate the size of individual exons or locate their position precisely from that study. This discouraged us from attempting heteroduplex mapping, and we elected to characterize this portion of the bovine elastin gene by complete DNA sequence analysis. These sequence data in conjunction with the previously determined sequence of the sheep elastin clone, pcSEL1, enabled us to define a significant portion of the 3′ portion of the gene. In the completely sequenced region, the ratio of translated nucleotides to intron nucleotides is 1 to 15. Although there is extremely strong homology between the sheep and cow at both the nucleotide and amino acid levels, as might be expected from the relatively close evolutionary positions of the two species, differences do exist. Although most substitutions are

of a minor nature, such as substitution of hydrophobic amino acids, there is one charge change in exon 4 in which glutamine in the cow is replaced by arginine in the sheep. Also, a significant difference exists in the number of repeat units that may be found in a hydrophobic region (exon 6). This suggests that the particular number of amino acid residues in a given hydrophobic region is not crucial to the adequate functioning of the molecule. Similarly, in the cross-link region of exon 7, six alanine residues are found before the first lysine while only two alanines occur in the corresponding porcine tropoelastin peptide. This alteration may lead to variation in the particular type of cross-link formed.

A most important tentative conclusion that may be drawn from the available exon–intron structure is that hydrophobic and cross-link domains of tropoelastin appear to be encoded in separate exons, which may alternate with one another except in the extreme carboxy-terminal region. Moreover, there appears to be considerable homology between cross-link exons:

exon 7	GVGP	AAAAAAK	AAAK	AAQF
exon 5	GVSP	AAAAK	AAK	F

The sequence data also demonstrate that in the cross-link regions the lysine residues that will go to form the desmosine cross-links are found in pairs, as was suspected from the tryptic peptide data (Foster et al., 1975). Thus a given desmosine serves to cross-link only two individual polypeptide chains. Hydrophobic exons also appear to share several common features including the presence of repetitive sequences. Although these repetitive sequences are not identical between hydrophobic exons, these repetitive sequences may have been generated in part from tripeptides such as GVG, GLG, and GAG, which are shared among the hydrophobic exons. Significantly, the codons for the individual amino acids in these tripeptides can differ by a single base, raising the possibility of interconversion of one tripeptide into another and the generation of larger repeating units. Thus, although more analysis is required, particularly of the exons encoding hydrohobic regions, the present data suggest that the elastin gene may have arisen by duplication and recombination of prototypic exons encoding cross-link and hydrophobic domains.

Novel sequences, not previously observed in the porcine tropoelastin tryptic peptides, form the carboxy-terminal region of the molecule. The distinctive features of this region, including the presence of cysteine residues and its highly basic nature, suggest that this portion of tropoelastin may interact strongly, possibly through disulfide bonds, with other matrix proteins such as those composing the microfibrillar component. Although the precise nature of the proteins composing the microfibrillar component is not known, it is extremely likely that these proteins are acidic and rich in cysteine [for review, see Cleary & Gibson (1984)]. The most 3′-terminal region of the gene contains a relatively long untranslated sequence ending with one AATAAA polyadenylation signal found 971 bp after the termination codon. A second AATAAA is found 220 bp beyond the first. In the sheep cDNA clone, pcSEL1, the first polyadenylation signal was used. Although we and others have looked by Northern hybridization experiments for multiple mRNA species, only a single species of 3500 bases has been identified, in contrast to the case of many collagens in which multiple mRNA species have been identified corresponding to multiple polyadenylation signals (Aho et al., 1983; Myers et al., 1983). Since tropoelastin contains approximately 850 amino acids, roughly 950 bases of the mRNA are untranslated and these untranslated bases can be accounted for by those present at the 3′ end. This untranslated region is strongly conserved among cow, sheep, and human (Emanuel et al., 1985), suggesting a possible regulatory function, but further analyses and other experiments directly testing this hypothesis are necessary to support it. In addition, because of the imprecision in the estimates of both the number of amino acids in tropoelastin and the size of the mRNA, we have not eliminated the possibility of untranslated sequences being found at the 5′ end.

References

Aho, S., Tate, V., & Boedtker, H. (1983) *Nucleic Acids Res. 11*, 5443–5450.

Benton, W. D., & Davis, R. W. (1977) *Science (Washington, D.C.) 196*, 180–182.

Boedtker, H., Fuller, F., & Tate, V. (1983) *Int. Rev. Connect. Tissue Res. 10*, 1–63.

Bressan, G. M., & Prockop, D. J. (1977) *Biochemistry 16*, 1406–1412.

Burnett, W., & Rosenbloom, J. (1979) *Biochem. Biophys. Res. Commun. 86*, 478–484.

Burnett, W., Eichner, R., & Rosenbloom, J. (1980) *Biochemistry 19*, 1106–1111.

Burnett, W., Finnigan-Bunick, A., Yoon, K., & Rosenbloom, J. (1982) *J. Biol. Chem. 257*, 1569–1572.

Cleary, E. G., & Gibson, M. A. (1984) *Int. Rev. Connect. Tissue Res. 10*, 97–209.

Davidson, J. M., Leslie, B., Wolt, T., Crystal, R. G., & Sandberg, L. B. (1982a) *Arch. Biochem. Biophys. 218*, 31–37.

Davidson, J. M., Smith, K., Shibahara, S., Tolstosheve, P., & Crystal, R. G. (1982b) *J. Biol. Chem. 257*, 747–754.

Davidson, J. M., Shibahara, S., Boyd, C., Mason, M. L., Tolstosehv, P., & Crystal, R. G. (1984a) *Biochem. J. 220*, 653–663.

Davidson, J. M., Shibahara, S., Schafter, M. P., Harrison, M., Leach, C., Tolstoshev, P., & Crysta, R. G. (1984b) *Biochem. J. 220*, 643–652.

Emanuel, B. S., Cannizzaro, L., Goldstein-Ornstein, N., Indik, Z. K., Yoon, K., May, M., Oliver, L., Boyd, C., & Rosenbloom, J. (1985) *Am. J. Hum. Genet.* (in press).

Foster, J. A., Burenger, E., Gray, W. R., & Sandberg, L. B. (1973) *J. Biol. Chem. 248*, 2876–2879.

Foster, J. A., Shapiro, R., Voynow, R. C., Crombie, G., Faris, B., & Franzblau, C. (1975) *Biochemistry 14*, 5343–5347.

Foster, J. A., Rich, C. B., Fletcher, S., Karr, S. R., & Przybyla, A. (1980) *Biochemistry 19*, 857–864.

Franzblau, C., Sinex, F. M., Faris, B., & Lampidis, R. (1965) *Biochem. Biophys. Res. Commun. 21*, 525–531.

Hastings, K. E. M., & Emerson, C. P., Jr. (1983) *J. Mol. Evol. 19*, 214–218.

Heidecker, G., Messing, J., & Gronenborn, B. (1980) *Gene 10*, 69–73.

Karr, S. R., & Foster, J. A. (1981) *J. Biol. Chem. 256*, 5946–5949.

Maniatis, T., Fritsch, E. F., & Sanbrook, J. (1982) *Molecular Cloning*, pp 75–85, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Miller, E. J., Martin, G. R., Mecca, C. H., & Iez, K. A. (1965) *J. Biol. Chem. 240*, 3623–3627.

Murphy, L., Harsch, M., Mori, T., & Rosenbloom, J. (1972) *FEBS Lett. 21*, 113–117.

Myers, J. C., Dickson, L. A., deWet, W. J., Bernard, M. P., Chu, M.-L., DiLiberto, M., Pepe, G., Sangiogi, F. O., & Ramirez, F. (1983) *J. Biol. Chem. 258*, 10128–10135.

Narayanan, A. S., & Page, R. C. (1976) *J. Biol. Chem. 251*, 1125–1130.

Partridge, S. M. (1962) *Adv. Protein Chem. 17*, 227–297.

Pinnell, S. R., & Martin, G. R. (1968) *Proc. Natl. Acad. Sci. U.S.A. 61*, 708–716.
Rogers, J., & Wall, R. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1877–1879.
Rosenbloom, J., & Cywinski, A. (1976) *Biochem. Biophys. Res. Commun. 69*, 613–620.
Rosenbloom, J., Harsch, M., & Cywinski, A. (1980) *J. Biol. Chem. 255*, 100–106.
Rucker, R. B., Gottlich-Rieman, W., & Tom, K. (1973) *Biochim. Biophys. Acta 317*, 193–201.
Sandberg, L. B., Weissman, N., & Smith, D. W. (1969) *Biochemistry 8*, 2940–2945.
Sandberg, L. B., Leslie, J. G., Leach, C. T., Alvares, V. A., Torres, A. R., & Smith, D. W. (1985) *Pathol. Biol.* (in press).
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schimenti, J., & Duncan, C. H. (1984) *Nucleic Acids Res. 12*, 1641–1655.
Shibahara, S., Davidson, J. M., Smith, K., & Crystal, R. G. (1981) *Biochemistry 20*, 6577–6584.
Smith, D. W., & Carnes, W. H. (1973) *J. Biol. Chem. 248*, 8157–8161.
Venkatachalam, C. M., & Urry, P. W. (1981) *Macromolecules 14*, 1225–1229.
Watanabe, Y., Tsukada, T., Notake, M., Nakamishi, S., & Numa, S. (1982) *Nucleic Acids Res. 10*, 1459–1469.
Yoon, K., May, M., Goldstein, N., Indik, Z. K., Oliver, K., Boyd, C., & Rosenbloom, J. (1984) *Biochem. Biophys. Res. Commun. 118*, 261–269.

Articles

Exchange of Histones H1, H2A, and H2B in Vivo†

Larry Louters and Roger Chalkley*
Department of Biochemistry, College of Medicine, The University of Iowa, Iowa City, Iowa 52242
Received August 27, 1984

ABSTRACT: We have asked whether histones synthesized in the absence of DNA synthesis can exchange into nucleosomal structures. DNA synthesis was inhibited by incubating hepatoma tissue culture cells in medium containing 5.0 mM hydroxyurea for 40 min. During the final 20 min, the cells were pulsed with [^{3}H]lysine to radiolabel the histones (all five histones are substantially labeled under these conditions). By two electrophoretic techniques, we demonstrate that histones H1, H2A, and H2B synthesized in the presence of hydroxyurea do not merely associate with the surface of the chromatin but instead exchange with preexisting histones so that for the latter two histones there is incorporation into nucleosome structures. On the other hand, H3 and H4 synthesized during this same time period appear to be only weakly bound, if at all, to chromatin. These two histones have been isolated from postnuclear washes and purified. Some possible implications of in vivo exchange are discussed.

The basic repeating unit of chromatin structure is the nucleosome. As a consequence of the structure of the nucleosome, it would seem to be unavoidable that the nucleosome must be displaced or modified in some manner to permit the passage of enzymes involved in transcription and replication. This modification may be reflected in some identifiable dynamic characteristic of nucleosomes. In this regard, recent experiments have demonstrated that several of the histones (H1, H2A, and H2B) are capable of exchanging in and out of nucleosome structures (Carine & Thomas, 1981; Cremsi & Yaniv, 1980; Louters & Chalkley, 1984). While these studies were done in vitro, there are also several indications that histones also exchange in vivo. First, data concerning the deposition of new histones indicate the likelihood of some exchange of at least histones H1, H2A, and H2B prior to deposition (Jackson, 1978; Jackson & Chalkley, 1981a,b; Jackson et al., 1981). Second, all histones turn over faster than DNA. Different types of histones show different turnover rates: H1 turns over more rapidly than the H2A/H2B pair which turns over faster than H3/H4 (Djondjarov et al., 1983; Commerford & Cronkite, 1982; Gurley et al., 1972). This implies a continual but differential histone replacement process. Finally, histones are synthesized throughout the cell cycle (albeit to a lower degree during non-S-phase periods), and these histones synthesized in the absence of replication have been shown to associate in some way with chromatin (Groppi & Coffino, 1980; Sheinin & Lewis, 1980; Tarnowka et al., 1978; Waithe et al., 1983; Wu & Bonner, 1981; Gurley & Hardin, 1969; Nadeau et al., 1978; Russev & Hancock, 1981; Russev et al., 1980; Wu et al., 1983).

The purpose of this study was to test directly whether histone exchange also occurs in vivo. To do this, histones were radiolabeled in the presence of an inhibitor of DNA synthesis (hydroxyurea). Thus, deposition of new histones can only be by exchange since previous studies have shown that in the presence of hydroxyurea, DNA synthesis is drastically reduced (greater than 95%) while histone synthesis continues at a substantial rate (20–40% of normal) for at least an hour (Nadeau et al., 1978; Russev & Hancock, 1981). We then isolated either nucleosomal fragments or cross-linked, internal core octamers (Louters & Chalkley, 1984) and showed that radiolabeled histones H1, H2A, and H2B are isolated with these nucleosomal structures. We conclude that histones do

† This work was supported by grants from the National Institutes of Health to R.C. (Grant GM 28817) and to the Diabetes and Endocrinology Research Center (Grant AM 25295). L.L. was supported by Training Grant GM 07045-05 from the NIH.

0006-2960/85/0424-3080$01.50/0