Journal of Cellular Biochemistry 54:154–160 (1994)

Protected Regions in the Chicken α2(1) Procollagen Promoter in Differentiated Tissues

Sharada L. Truter and M. Iqbal Parker

MRC/UCT Research Unit for the Cell Biology of Atherosclerosis, Department of Medical Biochemistry, University of Cape Town Medical School, Observatory 7925, South Africa

Abstract The higher ordered structure of the chicken α2(I) procollagen gene was analyzed in chromatin isolated from expressing (lung) and nonexpressing (reticulocyte and erythrocyte) tissues. Digestion of DNA with methylation sensitive restriction endonucleases revealed that this gene was methylated in all tissues examined and that no differences existed in the promoter methylation patterns between expressing and nonexpressing tissues. DNAse 1 hypersensitive sites were located between 100–300 bp upstream from the transcription initiation site and within the first intron. These sites were also hypersensitive to the single-strand specific S_1 nuclease, implying that this region of the gene in the chromatin is either in an unfolded single-stranded conformation or under severe conformational stress. These differences in the α2(1) chromatin structure were confirmed by the finding that the promoter was more accessible to restriction endonuclease digestion in the expressing tissues than in the nonexpressing tissues. Digestion of chromatin with Pst I and Sma I revealed that some of these sites in the promoter were differentially protected by DNA-binding proteins in the two tissue types. These protected sites were located as far upstream as −1,600 and downstream within the first intron at +800.

Key words: DNA methylation, gene regulation, nuclease hypersensitivity, chromatin, procollagen gene

Transcriptionally active chromatin generally contains nuclease hypersensitive regions which differ from its transcriptionally inactive counterparts. This is mainly due to a difference in chromatin conformation between the different tissue or cell types. Such differences may be accounted for by torsional constraints within the DNA double helix and altered binding of regulatory proteins to the DNA [for reviews see Weintraub, 1985; Kornberg and Lorch, 1992].

DNAse 1-hypersensitivity changes in active genes is perhaps the best documented chromosomal change and such DNAse 1-hypersensitive sites have been shown to exist within the chicken and mouse α2(I) procollagen genes [McKeon et al., 1984b; Liau et al., 1986]. These hypersensitive sites have been shown to be located between 100–300 bp upstream from the transcription initiation site, and within the first intron in both chicken embryo fibroblasts and its Rous Sarcoma Virus transformed counterparts. These sites, however, are not present in brain chromatin, where the α2(1) procollagen gene is not expressed.

Although nucleosomes are associated with transcriptionally active as well as inactive genes, their particular arrangement on the DNA fiber may allow the adoption of an "open" conformation during transcription. As a result of this "open" conformation, active genes are often also hypersensitive to micrococcal nuclease and restriction endonucleases, in addition to being hypersensitive to DNAse 1 and S_1 nuclease [McGhee et al., 1981; Emerson and Felsenfeld, 1984]. Such studies are very useful in the analysis of DNA-protein interactions, as has been shown in the case of the chicken β-globin and the human dihydrofolate reductase genes [Wood and Felsenfeld, 1982; Emerson and Felsenfeld, 1984; Shimada et al., 1986]. In the chicken β-globin gene, the 5′ hypersensitive site extends over a region of 200 bp, of which 115 bp can be excised as protein-free DNA by Msp I digestion, suggesting the absence of a nucleosomal structure [McGhee et al., 1981; Emerson and Felsenfeld, 1984]. Studies by Liau et al. [1986] also suggest that the DNAse I-hypersensitive region

Received July 26, 1993; revised and accepted September 15, 1993.

Address reprint requests to M. Iqbal Parker, MRC/UCT Research Unit for the Cell Biology of Atherosclerosis, Department of Medical Biochemistry, University of Cape Town Medical School, Observatory 7925, South Africa.

of the mouse α2(I) procollagen gene is more accessible to restriction endonucleases because of a relatively "open" conformation of the coding DNA in collagen-synthesizing NIH 3T3 cells.

Several studies have shown that either α1(1) or α2(1) procollagen gene expression may be altered in virally or chemically transformed cells, which may be accompanied by DNA rearrangements or insertions [Schnieke et al., 1983; Hartung et al., 1986], or may involve epigenetic mechanisms [Parker et al., 1982; Parker and Gevers, 1984; Marsilio et al., 1984]. The α2(1) procollagen promoter contains several regions which are involved in specific DNA-protein interactions, including the CCAAT and TATA boxes [Oikarinen et al., 1987; Hatamochi et al., 1988; Maity et al., 1992]. A clone of simian virus 40 transformed WI-38 fibroblasts does not produce any α2(1) procollagen chains [Parker et al., 1989], and it would appear that trans-acting factors which bind in the region between +54 and −120 of the α2(1) procollagen promoter may play an important role in this phenomenon [Parker et al., 1992]. It is therefore important to analyze the distribution of nucleosomes and the binding of transcription factors to this region of the gene in order to understand the regulation of collagen gene expression in normal and transformed cells.

Thus it is also important that the structural organisation of these genes in the chromatin of normal cells be elucidated. In this study, the DNase 1- and restriction endonuclease hypersensitive sites in the chicken α2(I) procollagen gene were investigated in expressing (lung) and non-expressing (reticulocytes and erythrocytes) tissues. This study differs from previous ones in that fresh tissue was used, thus eliminating the possibility of artifacts due to extended or adverse culture conditions.

MATERIALS AND METHODS

Isolation of Nuclei

Hybrid gold chickens were treated for 5 days with phenylhydrazine (6 mg/kg body weight) in order to induce anaemia. Reticulocyte, erythrocyte, and lung nuclei were prepared using a modification of the method of Bloom and Anderson [1982]. The nuclei were resuspended in 0.34 M sucrose, 10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, layered onto a cushion of 1.6 M sucrose, 10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, containing 1 mM spermidine and 0.2 mM PMSF and pelleted by centrifugation at 100,000*g* for 30 min at 4°C. The pelleted nuclei were resuspended in 0.34 M sucrose, 10 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$ containing 0.5 mM spermidine, and stored in aliquots at −70°C.

Digestion of Nuclei

The nuclei were thawed on ice, centrifuged for 30 s in a microfuge, and resuspended at a chromatin concentration of 1 mg/ml in the appropriate buffer [Shimada et al., 1986]. Nuclease concentrations ranged between 5–100 μg enzyme/mg chromatin/ml DNAse 1 or between 0.5 and 10 U/μl S_1 nuclease for the indicated time periods. The reactions were terminated by the addition of SDS to 0.1% and proteinase K to 100 μg/ml. After incubation at 50°C for 1 h, the samples were deproteinized by three extractions with phenol:chloroform:isoamylalcohol (25:24:1) and the DNA precipitated by the addition of one tenth volume 3 M sodium acetate and 2.5 volumes ethanol. The DNA was pelleted by centrifugation, dissolved in 10 mM Tris/HCl, pH 8.0, 1 mM EDTA, digested with EcoR1, fractionated on 1% agarose gels, and transferred to nylon membranes (Amersham, Arlington Heights, IL).

For digestion with restriction endonucleases, the nuclei were resuspended in the buffers recommended by the suppliers, but containing 0.34 M sucrose and 5 mM spermidine. Digestion was performed with 1 U of enzyme per μg chromatin for 1 h at either 37°C with PstI or at 25°C with SmaI. Control incubations in the absence of enzymes were also performed in order to determine whether digestion by endogenous nucleases occurred. At the end of the incubation period, the reactions were terminated by the addition of SDS/EDTA and processed as described above. Where indicated, DNA (10 μg) was digested with 2 U restriction enzyme/μg DNA for 3 h, electrophoresed on 1% agarose gels, and transferred to nylon membranes.

DNA Hybridization

The DNA samples were electrophoresed on either 1 or 2% horizontal agarose gels in 40 mM Tris-acetate, pH 7.8, 50 mM Na-acetate, 10 mM EDTA, and transferred to nitrocellulose filters (SS, BA85, 0.4 μM pore size) or nylon membranes (Hybond-N, Amersham). The filters were prehybridized for 2–4 h at 68°C in 6 × SSC (1 × SSC = 0.15 NaCl, 0.015 M Na citrate), 10 × Denhardt's solution, (1 × Denhardt's = 0.02% BSA, 0.02% polyvinylpyrrolidone, and 0.02% ficoll), 0.1% SDS, and 50 μg/ml Herring sperm

DNA. Hybridization was performed in the above solution containing 2 × 10^6 dpm/ml of ^{32}P-labelled DNA (1–2 × 10^8 dpm/μg DNA) at 68°C for 24–36 h. After hybridization, the filters were washed twice in 2 × SSC, 0.1% SDS for 30 min at 65°C, followed by one wash in 0.1 × SSC, 0.1% SDS for 15 min and exposed to Du Pont Cronex 4 X-ray film.

RESULTS

DNAse 1- and S_1 Nuclease-Hypersensitive Sites

The DNAse 1-hypersensitive sites within the 5′ flanking region of the α2(I) procollagen gene was determined using a 5.7 kb genomic clone containing an EcoRI fragment which contains sequences 1.7 kb upstream and 4.0 kb downstream from the transcription initiation site [Tate et al., 1983; Wozney et al., 1981].

Digestion of reticulocyte and lung chromatin with DNAse 1, resulted in specific cleavage of the 5.7 kb fragment to a smaller fragment of 4.3 kb and a barely visible fragment of 3.5 kb (Fig. 1A). The relative intensities of the 4.3 and 3.5 kb bands increased with increasing DNAse 1 concentration, but with these sites being more sensitive in the lung than in the reticulocyte chromatin. These results confirmed the existence of two DNAse 1-hypersensitive sites within the α2(1) procollagen gene in these tissues as previously described for cultured chicken embryo fibroblasts [McKeon et al., 1984a].

Digestion of nuclei with S_1 nuclease generated similar sized fragments as those obtained in the DNAse 1 experiment (Fig. 1B), indicating that these regions are probably in a single-stranded conformation. This could be due to conformational stress or the formation of H-DNA structures. It is interesting that these same structures are also sensitive to S_1 nuclease in supercoiled plasmid DNA [Finer et al., 1984; McKeon et al., 1984b].

In order to control for endogenous nuclease activity, nuclei were digested in their respective buffers in the absence of any added nuclease. In both these cases neither the 4.3 or 3.5 kb fragments were generated (data not shown).

Restriction Endonuclease Sites

Certain restriction endonuclease sites in chromatin may be protected by DNA-binding proteins and may, therefore, render these sites inaccessible to digestion by restriction endonucleases. In order to further analyse the differences in

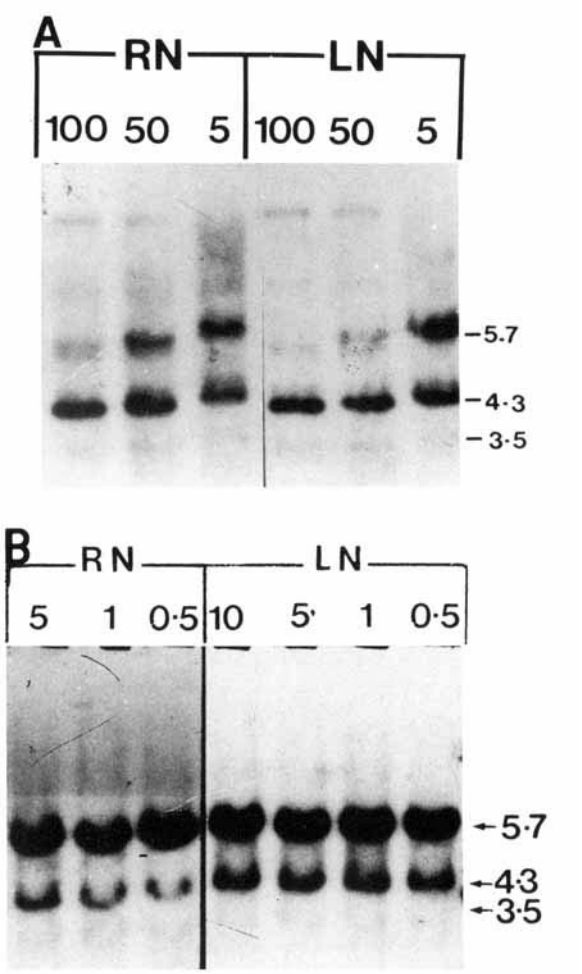

Fig. 1. DNAse 1- and S_1 nuclease hypersensitivity of the α2(1) procollagen gene. Nuclei were digested with the indicated concentrations of **(A)** DNAse 1 (μg/mg chromatin) and **(B)** S_1 nuclease (units/μg chromatin) as described in Materials and Methods, followed by digestion of the purified DNA with EcoRI. DNA was transferred to nylon membranes and hybridized to an α2(1) procollagen genomic clone. The molecular weight markers are indicated in kilobases. RN and LN indicate reticulocyte and lung nuclei, respectively.

α2(1) procollagen chromatin between active and inactive tissues, the accessibility of several PstI sites in the promoter was investigated.

Control digestion of DNA with PstI and EcoRI resulted in several fragments upon hybridization to the 5.7 kb EcoRI genomic clone (Fig. 2). Only the 1.5 and 2.0 kb fragment adjacent to the transcription initiation site are shown, while the 140 bp fragment is too small to visualize on this gel. The digestion pattern of purified DNA was identical in all three tissues investigated in this study. Digestion of nuclei with PstI (and subsequent digestion of the purified DNA with EcoRI), however, resulted in only two fragments of 5.3

and 5.7 kb. The persistence of the 5.7 kb fragment in all three tissues is probably due to protection of most of the PstI sites in this region by chromosomal proteins. Cellular heterogeneiety can be ruled out since the reticulocytes and erythrocytes were fairly homogeneous preparations of cells. The presence of the 5.3 kb fragment on the other hand, indicated that the PstI site at "a" (Fig. 2C) was digested by PstI whereas the other sites were not. This site is located 2 kb from the transcription initiation site, indicating that a conformational change of some sort has occurred at a significant distance upstream in this promoter. A time course experiment revealed that this site was more sensitive in the expressing lung tissue when compared with the non-expressing tissues (Fig. 2A,B). More important, the 5.3 kb band appeared after 5 min of PstI digestion in lung nuclei, whereas a much weaker signal appeared only after 30 min in reticulocyte nuclei and an even weaker signal in the erythrocyte nuclei. Longer digestion times resulted in decreased intensity of the 5.3 kb band in the lung tissue, and was accompanied by a concomitant increase in the lower molecular weight fragments (not shown in Fig. 2). In order to confirm these findings, a time-course Southern blot was hybridized to the 0.4 kb SmaI fragment containing the TATA box at −30 and an inverted CCAAT box at −80. After a 30 min digestion period, both erythrocyte and reticulocyte nuclei yielded only the 5.7 kb PstI resistant fragment, whereas lung nuclei yielded both the 5.3 and 5.7 kb fragments (Fig. 2B).

In order to analyse sites closer to the CCAAT and TATA boxes, nuclei were digested with SmaI, and the purified DNA was subsequently digested with EcoRI and analysed using the 0.8 kb SmaI fragment as a probe (see Fig. 3B). SmaI digested DNA hybridized to fragments of 0.8 and 4.0 kb (Fig. 3A). Since SmaI is sensitive to methylation of the internal cytosine in its recog-

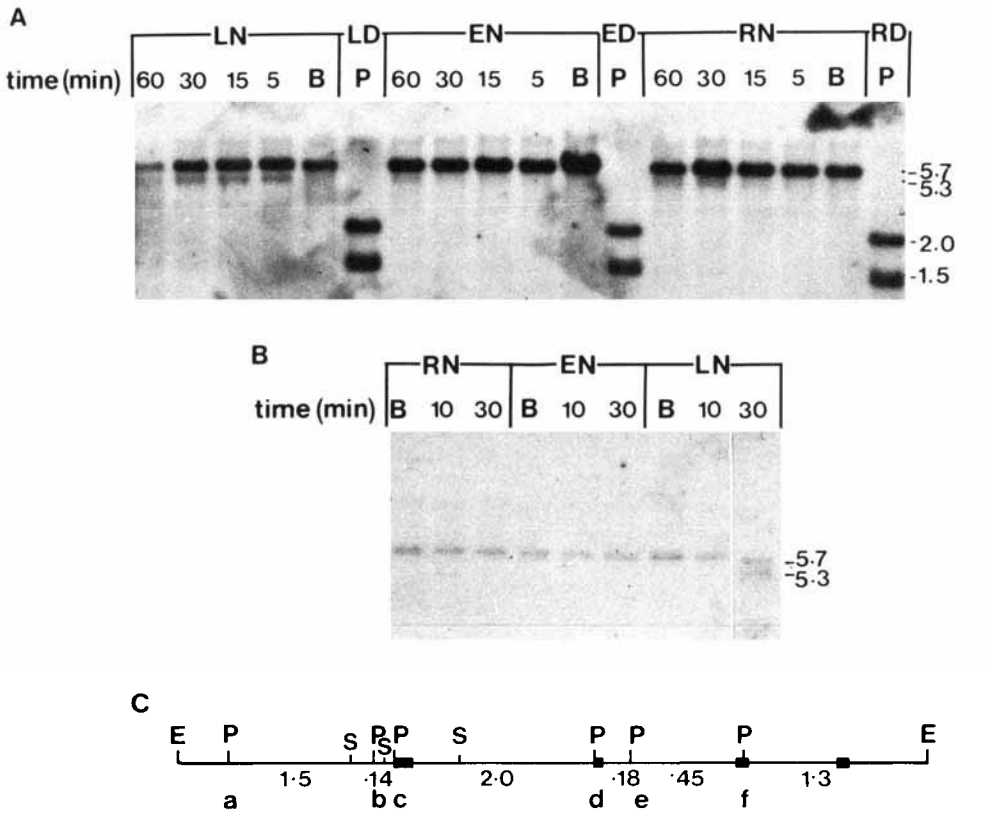

Fig. 2. Restriction endonuclease PstI-sensitivity of the α2(1) procollagen chromatin. Nuclei or DNA were digested with PstI for the indicated time periods, followed by digestion of the purified DNA with EcoRI. The Southern blots were hybridized to either the 5.7 kb EcoRI genomic clone **(A)**, or the 400 bp SmaI fragment **(B)**. The restriction map of the procollagen α2(1) promoter is presented in **C**. RN, EN, and LN indicate reticulocyte, erythrocyte, and lung nuclei, while RD, ED, and LD indicate reticulocyte, erythrocyte, and lung DNA, respectively. Lanes marked B are enzyme blanks, i.e., nuclei incubated at the appropriate temperature in the absence of added enzyme. The length of the PstI fragments in C are indicated in kilobases, with the solid boxes indicating the location of exons.

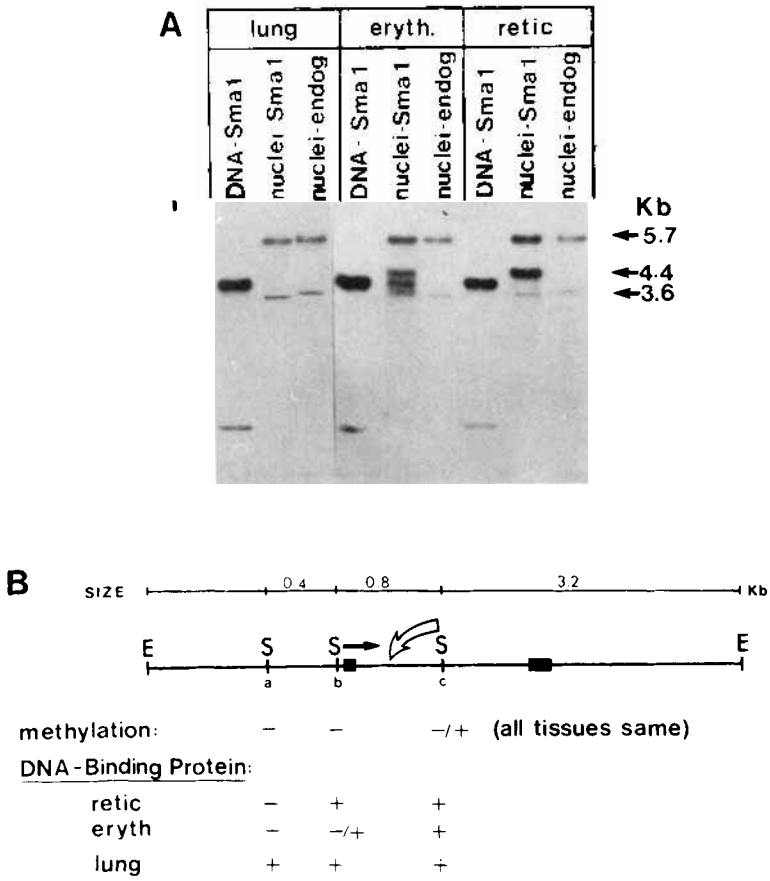

Fig. 3. SmaI sensitivity of the α2(1) procollagen gene (**A**). Nuclei or DNA were digested with SmaI followed by digestion of the purified DNA with EcoR1 as described in the legend to Figure 2. After transfer of the DNA to nylon membranes (Amersham Hybond-N), the Southern blot was hybridized to the 400 bp SmaI fragment shown in the restriction map in the methylation and protein-binding data are summarized in (**B**). S and E indicate SmaI and EcoRI, respectively. The methylation status and the binding of nuclear proteins to SmaI sites "a," "b," and "c" are indicated below the restriction map.

nition sequence, partial methylation of the Sma 1 site at "c" (Fig. 3B) will give rise to the 4.0 kb band which would still hybridize to the 800 bp Sma I probe. This site is clearly partially methylated in the DNA of all three tissues examined. Digestion of reticulocyte nuclei produced a major band of 4.4 kb indicating that SmaI site at "b" was protected from cleavage in these nuclei. The generation of this 4.4 kb fragment is probably due to inhibition of SmaI cleavage as a result of cytosine methylation at site "c" and protection of site "b" by DNA-binding proteins. Since the SmaI site "c" is only partially methylated, but no digestion occurs at this site in nuclei, this site must also be involved in DNA-protein interactions. Digestion of erythrocyte nuclei, on the other hand, generated bands of 4.0 and 4.4 kb, which indicated that the SmaI site at "b" was protected from SmaI cleavage in either only some of the chromosomes, or in only a certain population of reticulocytes. In lung nuclei, however, the presence of only the 5.7 kb fragment was compatible with protection of the SmaI sites at both "a" and "b" by DNA-binding proteins. These data indicated that the SmaI site at "a" was protected only in the expressing lung tissue and not in the non-expressing tissues. SmaI site "c," on the other hand, was protected in all three tissues since the 0.8 kb fragment was not obtained upon SmaI digestion. This result was not due to incomplete or partial digestion by SmaI since other SmaI sites on the same DNA molecule were completely digested. Also, the addition of Lambda DNA to the nuclear preparation resulted in complete digestion of the exogenous DNA. The third band of 3.6 kb which was seen in all the nuclear digests is due to cleavage by an endogenous

nuclease, since this fragment was present in nuclei incubated at 37°C in the absence of any added nuclease (Fig. 3A). This cleavage site was localized to the first intron (position +400) within the vicinity of the minor DNAse 1 hypersensitive site. It is interesting that this endogenous nuclease activity was detected only in the restriction endonuclease buffer system and not in the DNAse $1/S_1$ buffer system. The presence of the 5.7 kb band in all nuclear digests in this experiment is probably due to inefficient digestion by SmaI due to interference by nuclear proteins, since it could be eliminated by digestion with higher concentrations of restriction endonuclease.

DISCUSSION

The chicken α2(1) procollagen promoter in reticulocyte and lung chromatin was found to contain two DNAse 1 hypersensitive sites: a major site located between −100 and −300, and a minor site located within the first intron. The location of the 5′ DNAse 1-hypersensitive sites in the chicken lung chromatin was similar to that observed by McKeon et al. [1984a] in the case of cultured chicken embryo fibroblasts. In addition, the fact that these same sites were also hypersensitive to S_1 nuclease indicated that these regions were in a single-stranded form, or under considerable conformational stress. These sites have been shown to be S_1 nuclease hypersensitive in supercoiled plasmids and are located in a pyrimidine-rich sequence which can adopt a staggered loop in order to relieve the stress introduced by DNA supercoiling. This region of the α2(1) procollagen gene in the form of chromatin must, therefore, be under considerable conformational stress. The significance and function of the endogenous nuclease which also cleaves at the minor DNase 1 hypersensitive site is not clear.

Restriction enzyme analysis of the chromatin structure of the α2(I) procollagen promoter indicated that the gene in active lung nuclei was in a different conformation when compared with their inactive counterparts in reticulocyte and erythrocyte nuclei. The data presented here indicate that the CCAAT box region was in a different conformation in expressing tissues. These differences in restriction endonuclease sensitivity could be due to a more compact chromatin structure, nucleosome phasing, or other DNA-binding proteins. No difference in micrococcal nuclease sensitivity was observed between the three tissues, thus ruling out the possibility of nucleosome phasing (data not shown). It is possible that the CCAAT box could be located within a nucleosome-free region in order to facilitate interaction with other DNA binding or regulatory proteins. A similar nucleosome-free region of approximately 700 bp of DNA has been proposed for the mouse α2(I) procollagen promoter [Liau et al., 1986]. Previous studies have shown differences in DNA-binding proteins between collagen producing and non-producing cells in culture and these proteins could easily play a role in such altered chromatin conformation [Parker et al., 1992].

This study shows that certain tissue specific factors may be involved in binding to specific sites on the α2(1) procollagen promoter. In addition to the tissue specificity, it is possible that these factors may also be able to distinguish between methylated and non-methylated sites, as has been reported for the tyrosine amino transferase gene in non-expressing cells [Becker et al., 1987]. This study extends the previously demonstrated differences in chromatin structure in cultured cells to an in-vitro situation and shows that the data, at least for the α2(1) procollagen gene, are comparable.

ACKNOWLEDGMENTS

This work was supported by grants from the Medical Research Council, the National Cancer Association, and the University of Cape Town.

REFERENCES

Becker PB, Rupert S, Schutz G (1987): Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. Cell 51:835–843.

Bloom KS, Anderson JN (1982): Hormonal regulation of the conformation of the ovalbumin gene in chick oviduct chromatin. J Biol Chem 257:13018–13027.

Emerson BM, Felsenfeld G (1984): Specific factor confering nuclease hypersensitivity at the 5′ end of the chicken adult β-globin gene. Proc Natl Acad Sci USA 81:95–99.

Finer MH, Fodor EJB, Boedtker H, Doty P (1984): Endonuclease S_1-sensitive site in the chicken pro α2(1) collagen 5′-flanking gene region. Proc Natl Acad Sci USA 81:1659–1663.

Hartung S, Jaenisch R, Breindl M (1986): Retrovirus insertion inactivates mouse α1(1) collagen gene by blocking initiation of transcription. Nature 320:365–367.

Hatamochi A, Golumbek PT, van Schaftingen E, de Crombrugghe B (1988): A CCAAT DNA binding factor consisting of two different components that are both required for DNA binding. J Biol Chem 263:5940–5947.

Kornberg RD, Lorch Y (1992): Chromatin structure and transcription. Ann Rev Cell Biol 8:563–587.

Liau G, Szapary D, Setoyama C, de Crombrugghe B (1986): Restriction enzyme digestion identify discrete domains in

the chromatin around the promoter of the mouse α2(1) collagen gene. J Biol Chem 261:11362–11368.

Maity SN, Sinha S, Ruteshouser EC, de Crombrugghe B (1992): Three different polypeptides are necessary for DNA binding of the mammalian heteromeric CCAAT binding factor. J Biol Chem 276:16574–16580.

Marsilio E, Sobel M, Smith BDR (1984): Absence of procollagen mRNA in chemically transformed rat liver epithelial cells. J Biol Chem 259:1401–1404.

McGhee JD, Wood WI, Dolan M, Engel JD, Felsenfeld G (1981): A 200 basepair region at the 5′-end of the chicken adult β-globin gene is accessible to nuclease digestion. Cell 27:45–55.

McKeon C, Pastan I, de Crombrugghe B (1984a): DNase 1 sensitivity of the α2(1) collagen gene: Correlation with its expression but not with its methylation pattern. Nucleic Acids Res. 12:3491–3502.

McKeon C, Schmidt A, de Crombrugghe B (1984b): A sequence conserved in both the chicken and mouse α2(1) collagen promoter contains sites sensitive to S_1 nuclease. J Biol Chem 259:6636–6640.

Oikarinen J, Hatamochi A, de Crombrugghe B (1987): Separate binding sites for nuclear factor 1 and a CCAAT binding factor in the mouse α2(1) collagen promoter. J Biol Chem 262:11064–11070.

Parker MI, Gevers W (1984): Demethylation of the type I procollagen genes in transformed fibroblasts treated with 5-azacytidine. Biochem Biophys Res Commun 124:236–243.

Parker MI, Judge K, Gevers W (1982): Loss of type I procollagen gene expression in SV40-transformed human fibroblasts is accompanied by hypermethylation of these genes. Nucleic Acids Res 10:5879–5891.

Parker MI, Smith AA, Gevers W (1989): Absence of α2(1) procollagen synthesis in a clone of SV40-transformed WI-38 human fibroblasts. J Biol Chem 264:7147–7152.

Parker MI, Smith AA, Mundell K, Collins M, Boast S, Ramirez F (1992): The abolition of collagen gene expression in SV40-transformed fibroblasts is associated with trans-acting factor switching. Nucleic Acids Res 20:5825–5830.

Schnieke A, Harbers K, Jaenisch R (1983): Embryonic lethal mutation in mice induced by retrovirus insertion into the α1(1) collagen gene. Nature 304:315–320.

Shimada T, Inokuchi K, Nienhuis AW (1986): Chromatin structure of the human dihydrofolate reductase gene promoter. J Biol Chem 261:1445–1452.

Tate VE, Finer M, Boedtker H, Doty P (1983): Chick pro α2(1) collagen gene: Exon location and coding potential for the prepropeptide. Nucleic Acids Res 11:91–104.

Weintraub H (1985): Assembly and propagation of repressed and derepressed chromosomal states. Cell 42:705–711.

Wood WI, Felsenfeld G (1982): Chromatin structure of the chicken β-globin gene region. Sensitivity to DNase I, micrococcal nuclease and DNase II. J Biol Chem 257:7730–7736.

Wozney J, Hanahan D, Morimoto R, Boedtker H, Doty P (1981): Fine structural analysis of the chicken pro α2(1) collagen gene. Proc Natl Acad Sci USA 78:712–716.