STADLER SYMP. Vol. 14 (1982) University of Missouri, Columbia 69

CHROMATIN STRUCTURE OF EUKARYOTIC GENES: DNase I HYPERSENSITIVE SITES

Ky Lowenhaupt, Michael A. Keene, Iain L. Cartwright & Sarah C.R. Elgin

Washington University, Department of Biology, St. Louis, MO 63130

SUMMARY

We have recently learned much about the organization of the eukaryotic genome by using nucleases as probes of the protein-DNA interactions of chromatin. In this paper we review the available information from our laboratory and others concerning the DNase I hypersensitive sites of chromatin. Such sites have been found at or near the 5' ends of many genes. Frequently there is a cluster of these sites before a given gene. In addition, a number of sites have been observed which are not 5' to any known transcript; these may play roles in DNA replication or gene rearrangement. The 5' DNase I hypersensitive sites appear during development before a given gene is expressed. In several cases these events are closely temporally linked, but in some instances 5' DNase I hypersensitive sites are apparent long before a gene is due to be transcribed. For example, many sites observed in the post-blastula embryo of Drosophila are established in the pre-plastula stage. The molecular nature of DNase I hypersensitive sites is unclear. In some instances sites have been shown to be nucleosome free regions; however the pattern of association or proteins, if any, as well as the form of the DNA molecule, remain to be established.

INTRODUCTION

Much evidence has been gathered in recent years suggesting that changes in chromatin structure play an important role in the regulation of gene expression during development. In 1976 Weintraub and Groudine observed that the β-globin gene is preferentially digested by DNase I in chromatin from cells in which the gene is active (red blood cells). Similar results have been obtained subsequently for a wide variety of genes in a number of organisms and tissues (see MATHIS et al. 1980; CARTWRIGHT et al. 1982; IGO-KEMENES et al. 1982, for reviews).

In addition to genes which are actively being transcribed, genes which have been but are no longer being transcribed in a differentiated cell have been shown to be KNase I sensitive. Acctive genes are more accessible to a number of other enzymes, e.g., micrococcal nuclease, *E. coli* DNA polymerase, and restriction enzymes. It has been concluded that this sensitivity reflects the rearrangement of the chromatin around active genes into a more "open" or accessible conformation, a change which may be necessary for transcription.

It has been of interest to map the extent of this altered conformation. Studies of the region encompassing the ovalbumin gene cluster of chickens indicate that 80 kb of DNA show the same level of DNase I sensitivity as the transcribed regions in the expressing tissue (LAWSON et al. 1982). In addition, approximately 10 kb at each end of this domain are intermediate in sensitivity, less sensitive than the active genes but more so than bulk DNA, suggesting a transitional region. Similary, studies of the β-globin gene cluster indicate a large region to be involved in a DNase I sensitive domain (STADLER et al. 1980a).

In addition to these broad patterns of sensitivity to digestion, analyses of chromatin structure by means of nucleases have revealed another interesting aspect. Specifice sites along the chromatin fibre are even more sensitive to digestion with DNase I than active genes; these have come to be known as DNase I hypersensitive sites. Such sites were first observed by Wu et al. (1979a) in a study of the heat shock genes of *Drosophila melanogaster*. Nuclei were digested very lightly with DNase I, and the resultant DNA was pruified and analysed by a Southern blot. Hybridization with a specific cloned DNA revealed discrete bands (Fig. 1). Control experiments, in which pruified DNA was digested to an equivalent extent produced no such bands. Because the bands are quite sharp it was suggested that specific and limited regions in chromatin have a very high susceptibility to the enzyme. The pattern of bands differes for each of the cloned *Drosophila* DNAs used as probes, indicating that the sites are nonuniformly arranged in the chromatin. In this study, DNase I hypersensitive sites were found close to the heat shock genes, while cloned DNAs containing no known transcripts revealed only widely separated sites.

Various protocols have been used to map these DNase I hypersensitive sites. The most commonly reported has been called the "indirect end labeling" technique and was first used by Nedospasov and Georgiev, using micrococcal nuclease (1980) and by Wu, using DNase I (1980). In this technique, nuclei are lightly digested with the nuclease, such than an average of one cut is introduced within the region of interest. The DNA is purified and digested completely with a restriction endonuclease. The resulting fragments are size-separated by electrophoresis through an agarose gel and transferred to nitrocellulose by Southern blotting (SOUTHERN, 1975). The nitrocellulose is probed using 32p-labeled recombinant DNA containing a small fragment (0.2-2 kb) from the region of interest, ideally abutting the restriction site at one end of the region of interest. A diagram of the experiment is shown in Figure 2. Some possible fragments

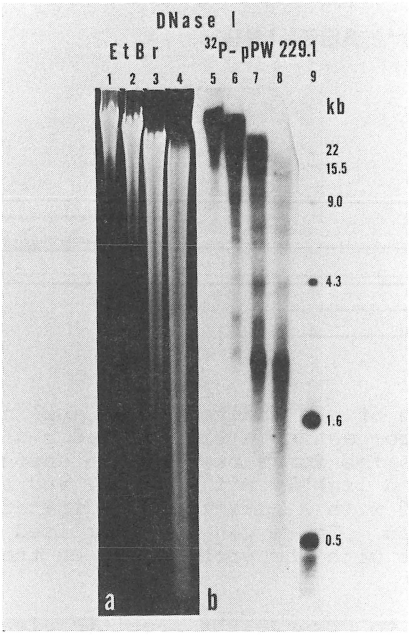

FIGURE 1. Pattern of DNA fragments from DNase I Digestion of Chromatin. Nuclei were digested with increasing amounts of DNase I and the DNA was purified and displayed on a 0.7% agarose gel as described by Wu et al. (1979a). a) The overall pattern of DNA fragments is visualized after staining with ethidium bromide. b) The specific pattern of fragments including the hsp 70 gene is shown. A Southern blot was prepared and hybridized with a nick-translated plasmid, pPW 229.1, containing part of the coding region of the hsp 70 gene. Lane 9 contains molecular weight markers. Figure from Wu et al. 1979a.

from the set that will be visualized are indicated as lines beneath the map. The largest fragment extends from one restriction site to the next, while the smaller fragments extend from the left-hand restriction site to a right-hand DNase I site. The size of the smaller fragments reflects the position of the DNase I cleavage relative to the left-hand restriction site.

DNase I HYPERSENSITIVE SITES AT 5' ENDS OF GENES

When the "indirect end labeling" technique was applied to the genes for the major heat shock proteins of 70 kD and 83 kD of *Drosophila*, a number of DNase I hypersensitive sites were mapped in the region directly preceding the 5'-end of the transcripts (WU 1980). These sites are observed in nuclei from tissue culture cells or embryos which have not been subjected to heat shock as well as in nuclei from heat shocked cells. In the latter case the sites are not as readily detectable because of the increase in overall sensitivity of the genes, which makes

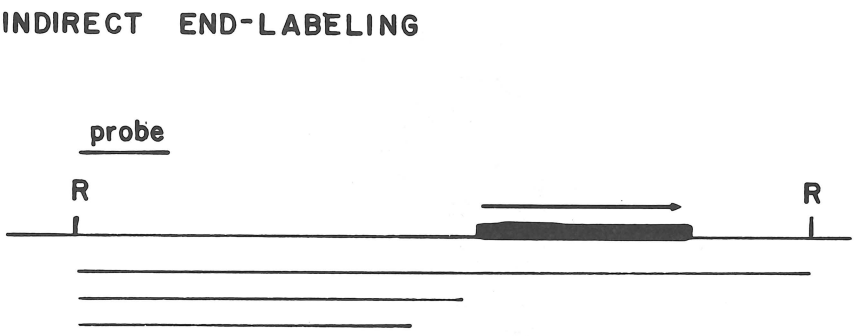

FIGURE 2. Diagram of the Indirect End Labeling Experiment. A stretch of DNA encompassing a transcribed region (heavy line) has two cleavage sites for a restriction endonuclease R. When nuclei are digested lightly with DNase I and the purified DNA completely cleaved with R, fragments indicated by the light lines are generated. These can be visualized by hybridization of a Southern blot with the probe shown in the upper left.

it more difficult to observe the specific cleavage. The gene encloding the 83 kD heat shock protein (hsp 83), a unique sequence at locus 63B, is preceded by three broad sites approximately 200-300 bp apart. In addition, there are two clusters of sites 3' to the hsp 83 gene. Since these clusters are separated by another region of transcription, one of these clusters must be 5' to this gene, and other regions of transcription in this area have not been ruled out. Because the hsp 70 coding region is not unique, the DNase I hypersensitive sites around these genes have not been mapped unambiguously. Nonetheless, the data can be interpreted to reflect sites immediately 5' to each of the five genes, as well as at other sites both up and downstream.

A unique segment of DNA mapping at locus 67B on the *Drosophila* polytene chromosomes encodes four small heat shock proteins (hsp 22, 23, 26 and 28) (CORCES et al. 1980) as well as a developmentally regulated gene (SIROTKIN & DAVIDSON 1982). Each of these five genes has a major DNase I hypersensitive site at or near the 5'-end, and a second, weaker site approximately 200 bp upstream (KEENE et al. 1981). Figure 3 shows the result of digesting nuclei with DNase I followed by digesting the purified DNa with Dam HI. The Southern blot was probled with a Bam-Sal fragment abutting the left hand end of the map, as shown in Figure 3. The position of DNase I hypersensitive sites, as determined from the length of the fragments is indicated. It is of interest to note that when chromatin is digested with micronuclease and analyzed in the same manner, the major DNase I hypersensitive site is also seen to be micrococcal nuclease sensitive. Although most preferential micrococcal nuclease sites are found to be common between DNA and chromatin (KEENE &

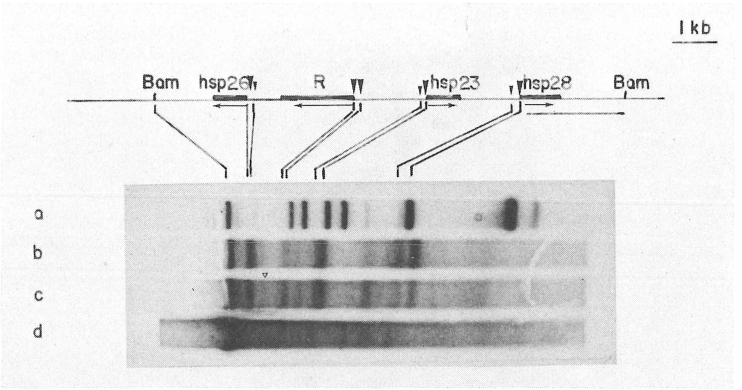

FIGURE 3. Digestion of Locus 67B by DNase I. The map shows the position of regions of transcription and the direction of transcription. (Lane d) purified DNA, (Lane c) nuclei from 0-1.5 hr embryos, (Lane b) nuclei from 6-12 hr embryos, (Lane a) molecular weight markers. The DNA from samples b, c and d was digested completely with Bam HI. After electrophoresis through a 1.2% agarose gel, a Southern blot was prepared and hybridized with a nick translated probe (lower line, upper right).

ELGIN 1981) indicating that sequence preference of the enzyme plays an important role in its choice of sites, chromatin specific sites can frequently be identified at the 5' end of genes (Figure 4).

In addition to the heat shock genes of *Drosophila* a number of genes which are expressed in multiple tissues have been examined in several organisms. The histone genes of *Drosophila*, present at approximately 100 copies each, are each preceded by a single 5' site (SAMAL et al. 1981). Very minor sites can also be detected throughout the nontranscribed spacer regions. The region encoding a *Drosophila* ribosomal protein, rp 49, has been examined (WONG et al. 1981). The ribosomal protein gene and a region coding for a rare transcript are separated by approximately 1 kb. The two genes are transcribed divergently. A series of five DNase I hypersensitive sites was found between the genes in nuclei from embryos. In this case the site proximale to the rp 49 gene is the weakest. The α-2 collagen gene has a 5' DNase I hypersensitive site in chromatin from chick embryo fibroblasts (MC KEON et al. 1982). In yeast *S. cerevisiae*), DNase I hpersensitive sites have been mapped very near the 5' ends of two alcohol drhydrogenase genes (SLEDZIEWSKI & YOUNG 1982). One gene, ADC 1, is constitutively expressed, while the other, ADR 2, is inducible in the presence of ethanol. Although the broad DNase I sensitivity of the genes correlates with the activity state, DNase I hypersensitive sites are observed 5' to both genes under growth conditions where ADR 2 is repressed, as well as those where it is expressed. When thymidine kinase genes isolated from Herpes simplex are transfected

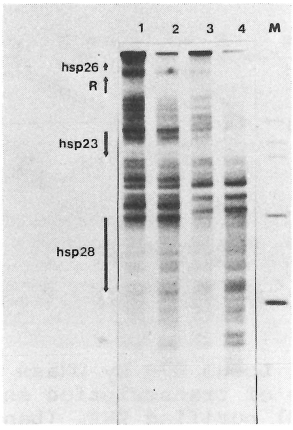

FIGURE 4. Digestion of Locus 67B by Micrococcal nuclease. The map is the same as in Figure 3 (Lane M) molecular weight markers. (Lanes 1, 2) nuclei from 6-12 hr embryos, (Lanes 3, 4) purified DNA. Protocol was as described for Figure 3.

into mouse TK$^-$fibroblasts, these genes can be expressed constitutively and allow the cells to grow under HAT selection. In such cells 180 bp upstream from the gene are sensitive to restriction endonucleases, which, like DNase I, will cleave highly preferentially in exposed regions of chromatin (SWEET et al. 1982). The extrachromosomal rDNA of Tetrahymena contains a DNase I hypersensitive site at the region of the initiation of transcription, indicating that 5' DNase I hypersensitive sites may be characteristic of genes transcribed by polymerase I as well as those transcribed by polymerase II (BORCHSENIUS et al. 1981; PALEN et al. 1982).

In several instances, the nuclear form of a viral genome is packaged as chromatin; DNase I hypersensitive sites have been observed in these cases. SV40 minichromosomes are DNase I hypersensitive for 250-400 bp over the region containing the origin of replication and the T antigen promoter (SCOTT & WIGMORE 1978; VARSHAVSKY et al. 1979). A similar site is observed in polyoma (HERBOMEL et al. 1981). Endogenous retroviruses of chickens which are transcriptionally active have DNase I hypersensitive sites in both long terminal repoeats (GROUDINE et al. 1981). These DNase I hypersensitive sites are not observed in inactive viruses. Somatically acquired Moloney murine leukemia viral sequences which are integrated into the genome show DNase I hypersnesitive sites just upstream to their 5' LTR, but no 3' site is detected (VAN DER PUTTEN et al. 1982).

The expression of the genes discussed above is not dependent on the differentiation of a specific cell type. Generally chromatin structure has been examined in dividing cells--tissue

culture cells, *Drosophila* embryos, logorithmically growing yeast, or *Tetrahymena* Genes expressed specifically in differentiated cells have been examined in expressing tissues, and again DNase I hypersensitive sites have been detected. For example, 5' DNase I hypersensitive sites have been observed by Weintraub and his colleagues around the α- and βglobin genes of chick in red blood cells (STADLER et al. 1980b; WEINTRAUB et al. 1981; LARSEN & WEINTRAUB 1982). In addition, other sites have been observed in all tissues examined. DNase I hypersensitive sites are also observed 5' to the mouse β-major- and α-globin genes in induced Friend erythroleukemia cells (SHEFFERY et al. 1982). A region of 250-300 bp upstream of the rat preproinsulin gene has been shown by Wu and Gilbert (1981) to be DNase I hypersensitive in a pancreatic β-cell tumor which secretes insulin. A surprising observation is that of a DNase I hypersensitive site within the coding region near the 3' end of the gene, found in chromatin of liver cells. This site is not observed in the tumor cells; neither site is observed in spleen, kidney or brain cells.

In contrast to most other genes examined, the ovalbumin gene of chicken does not have distinct 5' DNase I hypersensitive sites. Rather, a region of approximately 1.5 kb 5' to the gene is hypersensitive to nuclease relative to the gene itself (P. CHAMBON and M. BELLARD, personal communication). Preferential cuts are not generated, perhaps because the region is too extensive and homogeneous. Similarly, no conventional sites have been observed 5' to the conalbumen gene, although a 3' site is present in both liver and oviduct cells, nor are there sites in the region encompassing the ovalbumin X and Y genes (KUO et al. 1979). It is possible that all these genes are preceeded by DNase I hypersensitive regions rather than sites.

Thus there is evidence that 5' DNase I hypersensitive sites in chromatin are necessary, although not sufficient, for in vivo transcription of a gene by RNA polymerase II. Such sites are generally observed next to transcribing genes. In addition, recent studies of the salivary gland glue protein locus of *Drosophila melanogaster*, Sgs 4, correlate deletion of a region containing a major DNase I hypersensitive site with loss of gene expression. This gene is expressed at a high level in the salivary gland of Oregon R larvae. Two prominent DNase I hypersensitive sites are observed in chromatin from larval salivary glands, centered at 405 and 330 bp upstream from the 5' end of the gene. In addition, a minor site is seen at -480 and two more are observed near the origin of transcription. In chromatin from embryos, the 5' sites are not detectable; however, three sites are found 3' to the gene which are absent in the larval tissue (SHERMOEN & BECKENDORF, 1982). The Sgs 4 gene is transcribed at a very low level in three oriental strains, Hikone-R, Kochi-R and Seto, and no RNA transcript can be detected in the strain BER-1. BER-1 flies have a ~100 bp deletion covering the DNase I hypersensitive site at -405. The Orientals have ~50 bp deletions including the region of the -330 DNase I hypersensitive sites (MUSKAVITCH & HOGNESS, 1980, 1982). Studies of chromatin show that larval salivary glands from BER-1 flies have no 5' DNase I hypersensitive sites near the Sgs 4 gene

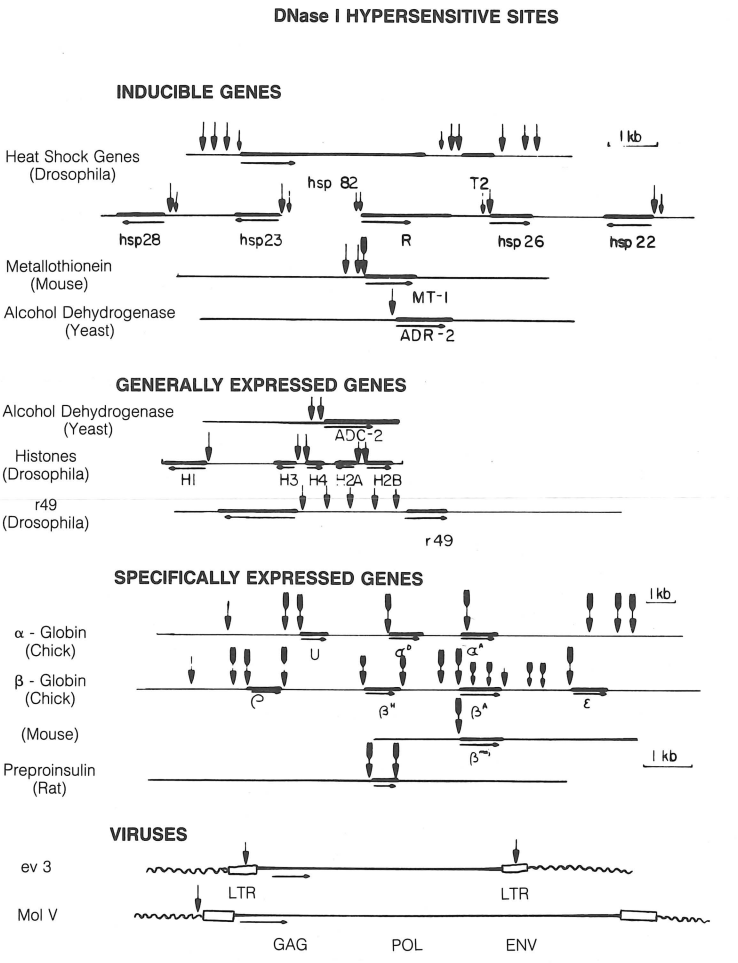

FIGURE 5. Diagram of DNase I Hypersensitive Sites Mapping Near Regions of Transcription. Transcribed regions are shown as heavy lines, with the direction of transcription as indicated (fine arrows). Large vertical arrows represent prominant DNase I hypersensitive sites while smaller arrows represent less preferred sites. DNase I hypersensitive sites which are found only in certain tissues are indicated by fledged ar ɔws. References for the original data are as follows: hsp 8?, WU (1980); hsp 22, 23, 26 and 28, KEENE et al. (1981); metallothionein, SENEAR & PALMITER 1982); ADH genes, SLEDGEWSKI & YOUNG (1982); r49, WONG et al. (1981); histones, SAMAL et al. (1981); chick globin, WEINTRAUB et al. (1981) and LARSEN & WEINTRAUB (1982); mouse β globin, SHEFFERY et al. (1982); preproinsulin, WU & GILBERT (1981); ev3, GROUDINE et al. (1981); Moloney virus, VAN DER PUTTEN et al. (1982).

while the Oriental flies are missing the site at -330, but retain the site at -405 (SHERMOEN & BECKENDORF, 1982). The correlation between the presence of these chromatin structures and expression of the gene is very striking.

Further support for the idea that DNase I hypersensitive sites are important for gene expression can be found in studies of the mouse metallothionein gene, MT-1 (SENEAR & PALMITER 1982). In this case two sites at -30 and -225 are detected. These sites encompass the regions which are necessary for cadmium-induced gene expression as determined in gene transfection and oocyte injection experiments. The glucocorticoid receptor binding site is encompassed by the distal DNase I hypersensitive site as well. An additional site is found only in cadmium-induced cells, and correlates with high level expression. In a mouse thymoma cell line which cannot express metallothionein, no hypersensitive sites can be detected. Likewise, the DNase I hypersensitive sites are not observed in tissues from two transgenic mice containing only inactive copies of the MT-1 gene. Once again one observes a correlation between the presence of nuclease sensitive sites in chromatin and gene expression. However, it should be noted that several instances of low level transcription with no concomitant hypersensitive sites have been recorded, as discussed below.

The positions of a number of DNase I hypersensitive sites relative to regions of transcription are summarized diagramatically in Figure 5.

DNASE I HYPERSENSITIVE SITES AT OTHER POSITIONS

DNase I hypersensitive sites have been observed at a number of positions that appear not to fall at or near the 5' ends of genes. The origin of replication of SV40 (VARSHAVSKY et al. 1979), polyoma (WALDECK et al. 1978), and the extrachromosomal rDNA of Tetrahymena (BORCHSENIUS et al. 1981; PALEN et al. 1982) are all DNase I hypersensitive. In addition, sites involved in DNA rearrangement events may be DNase I hypersensitive. DNase I hypersensitive sites have been observed in the region involved in DNA rearrangement events between J_H and $C\mu$ genes in the chromatin of T lymphocytes, but not in that of liver cells (STORB et al. 1981). A DNase I hypersensitive site may also be involved in the switching at the mating type locus of yeast (NASMYTH, 1982). One can speculate that DNase I hypersensitive sites exist at those positions in chromatin where accessibility of DNA is required for interaction with some macromolecule. Indeed, a given site may be important for more than one function. Alternatively, there may be different classes of sites with different molecular structures.

CORRELATIONS WITH DEVELOPMENT

If DNase I hypersensitive sites are necessary for transcription of a given gene in vivo, it is possible that the establishment of these sites is part of the commitment of a cell to expression of a specific gene. This hypothesis is supported by the observation that, for genes expressed specifically in one

differentiated cell type, 5' DNase I hypersensitive sites are found only in the expressing tissue and not in other differentiated cell types. For example, a 5' DNase I hypersensitive site is found next to the preproinsulin gene in rat β-cell tumor nuclei, but not in chromatin from liver, spleen, kidney or brain (WU & GILBERT, 1981). The chick α- and β-globin genes are preceded by DNase I hypersensitive sites in those erythroid cells in which these genes are transcribed, but not in brain cells where no globin mRNA can be detected (STALDER et al. 1980b; WEINTRAUB et al. 1981). The finding of 5' DNase I hypersensitive sites only in expressing tissues could be explained by suggesting that the sites are the result of transcription rather than required for it. The data do not agree with this interpretation. Sites are found next to *Drosophila* heat shock genes in cells which have been maintained at 25° and in which no heat shock mRNA can be detected (WU 1980; MILLER & ELGIN 1981; KEENE et al. 1981). The regulated yeast alcohol dehydrogenase gene described previously is associated with a hypersensitive site even when transcription is repressed (SLEDZIEWSKI & YOUNG 1982). Weintraub and his colleagues (1982) have been able to separate the establishment of hypersensitive sites from the initiation of transcription of chick globin genes using a temperature sensitive transformed erythroid cell line as described below. As previously described, in *Drosophila* a group of 5' DNase I hypersensitive sites precede the glue protein gene, Sgs 4, in the chromatin from larval salivary gland cells. These sites are not observed in chromatin from embryos or tissue culture cells. This makes it unlikely that differentiation involves only the sequential masking of DNase I hypersensitive sites. Rather these sites must come into existence before the gene becomes available for transcription.

It is of interest to determine when during differentiation the DNase I hypersensitive sites are established. Weintraub and his colleagues have analyzed the α- and β-globin gene clusters in developing chick embryos (STALDER et al. 1980a; WEINTRAUB et al. 1981). They found that in embryonic red blood cells, sites were present only preceding the embryonic β and U (α-like) genes. Later in development, sites appear preceeding the adult α- and β-globin genes and the sites before the embryonic genes are lost. This pattern corresponds to the pattern of gene expression. This is in contrast to the pattern of broad DNase I sensitivity as assayed by solution hybridization: globin genes retain their overall sensitivity after transcription has ceased (WEINTRAUB & GROUDINE 1976), as does the ovalbumin gene in the differentiated tissue (PALMITER et al. 1977). The relationship between the formation of DNase I hypersensitive sites and transcription of the globin genes was analyzed in greater detail using chicken bone marrow cells infected in vitro with a temperature sensitive avian erythroblastosis virus (WEINTRAUB et al. 1982). These cells grow in culture at 36°C, producing no hemoglobin. At 42° the product of the transforming gene (erb) is inactivated and hemoglobin is synthesized. Analysis of various clonal lines revealed cells in which both DNase I hypersensitive sites and mRNA accumulation were seen after increasing the temperature. Other cells were observed which contained DNase I hypersensitive site 5' to the globin genes at both low and high temperatures,

although globin mRNA was found only at high temperatures. This was interpreted to be because these latter cells had been "frozen" in their program of differentiation by transformation at a developmental time between the establishment of the DNase I hypersensitive site and the onset of globin mRNA synthesis.

Very early *Drosophila* embryos (from 0-2 hr after laying) synthesize little or no mRNA (ANDERSON & LENGYEL 1981). Although protein synthesis is taking place, it appears to result from translation of maternal mRNAs. We were interested in determining whether or not this apparent transcriptional imcompetence reflects an absence of appropriate chromatin structure, in particular of DNase I hypersensitive sites. Therefore we examined a number of genes, comparing the results of "indirect end labeling" experiments using 0-1.5 hr embryos (preblastula) with 6-18 hr embryos (postblastula). We examined the sites near a gene which is expressed in neither of these samples, Sgs 4, and of a gene which is expressed constitutively in the postblastula embryos, a ribosomal protein gene rp49. Further, we analyzed a gene cluster mapping at locus 67B which includes four small heat shock inducible genes, hsp 22, 23, 26 and 28, as well as another gene, R, whose expression varies during development. In every case, DNase I hypersensitive sites are found in the same position near the transcribed region(s) in both pre- and postblastula embryos. Therefore these sites are established very early in development. For generally expressed loci, the lack of mRNA synthesis in preblastula embryos appears not to be the result of the inaccessibility of these sites (LOWENHAUPT et al. 1982).

In the case of the "R" gene no transcript is detected in postblastula embryos or first or second instar larvae (SIROTKIN & DAVIDSON 1982, R. FREUND, personal communication). Only in late third instar larvae and early pupae is a transcript observed; then it is very abundant mRNA. However, there are a pair of DNase I hypersensitive sites at the 5' end of the gene in pre- and postblastula embryos similar to the sites found next to the inducible heat shock genes at the same locus and not unlike the array of sites abutting the actively transcribed ribosomal protein. Transcription from the 5' end of the R gene has been reported in postblastula embryos (SIROTKIN 1982), but this transcription has not been carefully quantitated and no mature product can be observed until the late larval stage. In contrast, the glue protein gene, which is expressed in third instar larvae like the R gene, does not show DNase I hypersensitive sites in embryo chromatin (SHERMOEN & BECKENDORF 1982). Similarly, for chick globin genes, the establishment of the DNase I hypersensitive site and the initiation of transcription are closely linked temporally (WEINTRAUB et al. 1982). In Friend erythroleukemia cells, an increase in the rate of transcription appears to be accompanied by the establishment of a DNase I hypersensitive site 5' to the β and α_1-globin genes (SHEFFERY et al. 1982), suggesting that perhaps some major sites are necessary for high level but not low level transcription. This is supported by the finding that certain mouse cell lines which produce metallothionein at very low levels do not have a 5' hypersensitive site (SENEAR & PALMITER 1982).

The temporal relationship between DNase I hypersensitive sites and broad DNase I sensitivity remains to be determined. In many cases, DNase I hypersensitive sites appear to be established before broad sensitivity of the domain. This is apparently the case for chick globin genes (WEINTRAUB et al. 1982), and is clearly established for *Drosophila* heat shock genes, where broad sensitivity is observed only in the induced gene (WU et al. 1979b). In contrast, the globin genes of uninduced Friend erythroleukemia cells are DNase I sensitive; however, no DNase I hypersensitive site is present in these cells until after induction (SHEFFERY et al. 1982).

DNASE I HYPERSENSITIVE SITES AND GENE ACTIVITE: POSSIBLE MODELS

The physical or chemical nature of the DNase I hypersensitive site remains a matter for speculation almost entirely uninhibited by data. It seems likely that the hypersensitivity results from a region of the chromatin structure which is less ordered and more accessible. However, an alternative is that DNase I is responding to the boundaries between two structures rather than a discrete structural entity. The idea of DNase I hypersensitive sites as boundaries is supported by the possibility that nucleosomes are ordered starting from these points (a possibility which has by no means been established, unfortunately). On the other hand, the idea of the DNase I hypersensitive site as a structure is supported by the breadth of the sites (as indicated by the fuzziness of bands generated by cleavage in them), and by evidence that they may include single stranded DNA and/or be nucleosome or protein free (SARAGOSTI et al. 1980; MCGHEE et al. 1981; LARSEN & WEINTRAUB 1982). It is, of course, possible that DNase I hypersensitive sites are structures which act as boundaries.

MCGHEE et al. (1981) have released a fragment from chicken erythrocyte chromatin, by means of the restriction endonuclease Msp I, from a DNase I hypersensitive site. This region is hypersensitive to micrococcal nuclease, DNase II and Msp I as well as DNase I, and maps from 60-260 base pairs from the 5' end of an adult β globin gene. The region is extremely GC-rich (70%). The excized fragment behaves as if 35% of the peices are protein-free DNA. Whether this reflects the situation in vivo, or whether this is the result of aggregation of protein to DNA during the isolation process or of disassociation of loosely bound proteins is unknown. Since they did not find the region to be particularly sensitive to S_1 nuclease, and it was readily cleaved with restriction enzymes which recognize only double-stranded substrates, the authors concluded that this DNase I hypersensitive site has little single-stranded or unwound character.

In contrast, Larsen and Weintraub (1982) found the 5' DNase I hypersensitive site to be sensitive to S-1 nuclease and to chemical reagents which bind covalently to single, but not double-stranded DNA. Small hairpin loops or an altered DNA structure rather than regular B-form might account for this sensitivity. The S-1 sensitive site is also present in supercoiled plasmids containing the same DNA sequences. The region which

McGhee et al. (1981) identified is hypomethylated in the expressing tissue. However, the relationship between hypomethylation and gene expression remains tenuous and confusing (for review, see RAZIN & RIGGS 1980), and there is no evidence that hypomethylation correlates with the DNase I hypersensitive sites in general.

The possibility that these sites are nucleosome free or entirely protein free remains, although the idea of naked DNA within a eukaryotic nucleus runs counter to most prejudices. In addition, this possibility only causes the question to be rephrased from "Why is this region DNase I hypersensitive?" to unprotected and obviously available DNA?" An interesting alternative form of DNA is present at DNase I hypersensitive sites. There is some evidence that Z-DNA will not bind to histone octomers to form normal stable nucleosomes (NICKOL et al. 1982) and therefore it is well suited to serve as a boundary. Since it has not been determined whether DNase I, DNase II or micrococcal nuclease are able to recognize and cleave Z-form DNA, it cannot yet be determined whether this possibility is worthy of serious consideration.

The role of DNase I hypersensitive sites in the initiation of transcription in vivo is not as yet understood. It is not uncommon to see a DNase I hypersensitive site covering the TATA box, but neither is this ubiquitous. In many cases, the proximal site is weaker than others in the array. Another interesting region has been reported to be required for enhancers, have been found in a variety of systems; the 72 base pair repeat of SV40 has enhancer activity (MOREAU et al. 1981; BANERJI et al. 1981), as does part of the long terminal repeat of Moloney sarcoma virus (LEVINSON et al. 1982) and a segment of polyoma virus (DE VILLIERS & SCHAFFNER 1981). These sequences are cis enhancers of transcription when placed in either orientation, proximal or distal to the gene in a circular DNA molecule. No sequence or tertiary structure homology has as yet been recognized among the regions which have been identified. Whether or not DNase I sensitive sites will in general include DNA sequences with enhancer function remains to be determined.

It seems reasonable to state that one or more DNase I hypersensitive sites are necessary but not sufficient for efficient expression of genes transcribed by RNA polymerase II. This might be accomplished by "tagging" regions to identify them to regulatory molecules or to polymerase, or orienting nucleosomes and important sequences, e.g., promoters, with respect to each other or by holding the chromatin in a more accessible conformation. No explanations have been suggested for the varying number and strength of the sites next to different genes. In addition there are DNase I hypersensitive sites not associated with the 5' ends of genes. It is possible that hypersensitivity to DNase I reflects a number of different chromatin structures and that many of the pending questions will have more than one answer. The specificity of these sites does provide a focus for the organization of chromatin structure, and a starting point for further investigation.

ACKNOWLEDGEMENTS

We wish to thank the many scientists who have shared their work with us prior to publication and the MIT Press for permission to reprint Figure 1. Work from this lab discussed here was supported by a grant GM 30273 from NIH to SCRE; KL was an NIH postdoctoral fellow and MAK an NIH predoctoral trainee.

LITERATURE CITED

ANDERSON, K.V. and J.A. LENGYEL 1981 Changing rates of DNA and RNA synthesis in *Drosophila* embryos. Dev. Biol. 82:127-138.

BANERJI, J., S. RUSCONI and W. SCHAFFNER 1981 Expression of a β globin gene is enhanced by remote SV40 DNA sequences. Cell 27:299-308.

BELLARD, M., M.T. KUO, G. DRETGEN and P. CHAMBON 1980 Differential nuclease sensitivity of the ovalbumin and β-globin chromatin regions in erythrocytes and oviduct cells of laying hen. Nuc. Acids Res. 8:2737-50.

BELLARD, M., G. DRETGEN, F. BELLARD, P. OUDET and P. CHAMBON 1982 Disruption of the typical chromatin structure in a 2500 base-pair region at the 5' end of the actively transcribed ovalbumen gene. The EMBO Journal 1:223-230.

BORCHSENIUS, S., B. BONVEN, J.C. LEER and O. WESTERGAARD 1981 Nuclease-sensitive regions on the extra chromosomal r-chromatin from Tetrahymena pyriformis. Eur. J. Biochem. 117:245-250.

CARTWRIGHT, I.L., M.A. KEENE, G.C. HOWARD, S.M. ABMAYR, G. FLEISCHMANN, K. LOWENHAUPT and S.C.R. ELGIN 1982 Chromatin structure and gene activity: The role of nonhistone chromosomal proteins. CRC Critical Reviews of Biochemistry (in press).

CORCES, V., R. HOLMGREN, R. FREUND, R. MORIMOTO and M. MESELSON 1980 Four heat shock proteins of *Drosophila melanogaster* coded within a 12-kilobase region in chromosomal subdivision 67B. Proc. Natl. Acad. Sci. U.S.A. 77:5390-93.

DE VILLIERS, J. and W. SCHAFFNER 1981 A small segment of polyoma virus DNA enhances the expression of a cloned β-globin gene over a distance of 1400 base pairs. Nucl. Acids Res. 9:6251-6264.

GROUDINE, M., R. EISENMAN, and H. WEINTRAUB 1981 Chromatin structure of endogenous retrovirus genes and activation by an inhibitor of DNA methylation. Nature 292:311-317.

HERBOMEL, P., S. SARAGOSTI, D. BLANGY and M. YANIV 1981 Fine structure of the origin-proximal DNase I-hypersensitive region in wild-type and EC mutant polyoma. Cell 23:651-658.

IGO-KEMENES, T., W. HORZ and H.G. ZACHAU 1982 Chromatin. Annual Review of Biochem. 51:89-122.

KEENE, M.A., V. CORCES, K. LOWENHAUPT and S.C.R. ELGIN 1981 DNASE I hyper-sensitive sites in *Drosophila* chromatin occur at the 5' ends of regions of transcription. Proc. Natl. Acad. Sci. U.S.A. 78:143-146.

KEENE, M.A. and S.C.R. ELGIN 1981 Micrococcal nuclease as a probe of DNA sequence organization and chromatin structure Cell 27:57-64.

KUO, M.T., J.L. MANDEL and P. CHAMBON 1979 DNA methylation: Correlation with DNase I sensitivity of chicken ovalbumin and

conalbumin chromatin. Nucl. Acids Res. 7:2105-2114.
LARSEN, A. and H. WEINTRAUB 1982 An altered DNA conformation detected by S-1 nuclease occurs at specific regions in active chick globin chromatin. Cell 29:609-622.
LAWSON, G.M., B.J. KNOLL, C.J. MARCH, S.L.C. WOO, M.-J. TSAI and B.W. O'MALLEY 1982 Definition of the 5' and 3' structural boundaries of the chromatin domain containing the ovalbumin maltigene family. J. Biol. Chem. 257:1501-1507.
LEVINSON, B., G. KHOURY, G. VAN DE WOUDE and P. GRUSS 1982 Activation of SV 40 genome by 72-base pair tandem repeats of Moloney sarcoma virus. Nature 295:568-572.
LOWENHAUPT, K., I.L. CARTWRIGHT, I.L. ZIMMERMAN, M.A. KEENE, and S.C.R. ELGIN 1982 Chromatin structure in pre- and post-blastula embryos of *Drosophila*. In preparation.
MATHIS, D., P. OUDET and P. CHAMBON 1980 Structure of transcribing chromatin. Prog. Nucl. Acid Res. Mol. Biol. 24:2-49.
MCGHEE, J.D., W. WOOD, M. DOLAN, J.D. ENGEL and G. FELSENFELD 1981 A 200 bp region at the 5' end of the chicken adult β-globin gene is accessible to nuclease digestion. Cell 27:45-55.
MCKEON, C., I. PASTAN and B. DECROMBRUGGHE 1982 Methylation status of the α 2 (1) collagen gene does not alter its chromatin structure as determined by DNase I sensitivity. J. of Cell. Biochem. Supplement 6:327.
MILLER, D.W. and S.C.R. ELGIN 1981 Transcription of heat shock loci of *Drosophila* in a Nuclear System. Biochem. 20:5033-42.
MOREAU, P., R. HEN, B. WASYLYK, R. EVERETT, M.P. GAUB and P. CHAMBON 1981 The SV40 70 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nuc. Acids Res. 9:6047-6068.
MUSKAVITCH, M.A., and D.S. HOGNESS 1980 Molecular analysis of a gene in a developmentally regulated puff of *Drosophila* melanogaster. Proc. Natl. Acad. Sci. U.S.A. 77:7362-7366.
MUSKAVITCH, M.A.T. and D.S. HOGNESS 1982 An expendable gene that encodes a *Drosophila* glue protein is not expressed in variants lacking remote upstream sequences. Cell 29:1041-1051.
NASMYTH, K.A. 1982 The regulation of yeast mating type chromatin structure by SIR: An action at a distance affecting both transcription and transposition. Cell, in press.
NEDOSPASOV, S. and G. GEORGIEV 1980 Non-random cleavage of SV40 DNA in the compact minichromosome and free in solution by micrococcal nuclease. Biochem. Biophys. Res. Comm. 92:532-539.
NICKOL, J., M. BEHE and G. GEORGIEV 1982 The effect of the B-Z transition in poly(dG-m^{5}dC)·poly(dG-m^{5}dC) on nucleosome formation. Proc. Natl. Acad. Sci. U.S.A. 79:1771-1775.
PALEN, T., D.S. GOTTSCHLING and T. CECH 1982 Transcribed and non-transcribed regions of the ribosomal RNA gene of Tetrahymena exhibit different chromatin structures. J. Cell. Biochem. Supplement 6:336.
PALMITER, R., E. MULVIHILL, S. MCKNIGHT and A. SENEAR 1977 Regulation of Gene Expression in the Chick Oviduct by Steroid Hormones. Cold Spring Harbor Symp. Quant. Biol. 42:773-778.
RAZIN, A. and A. RIGGS 1980 DNA methylation and gene fucntion. Science 210:604-610.
SAMAL, B., A. WORCEL, C. LOUIS and P. SCHEDL 1981 Chromatin structure of the histone genes of *D. mleanogaster*. Cell 23:

401-409.
SARAGOSTI, S., G. MOYNE and M. YANIV 1980 Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. Cell 20:65-75.
SCOTT, W.A. and D.J. WIGMORE 1978 Sites in Simian virus 40 chromatin which are preferentially cleaved by endonucleases. Cell 15:1511-1518.
SENEAR, A. and R. PALMITER 1982 5' regulatory sites of the mouse metallothionein-1 gene are packaged in a special chromatin structure. Cold Spring Harbor Symp. Quant. Bio. 47 (in press).
SHEFFERY, M., R. RIFKIND, and P.A. MARKS 1982 Murine erythroleukemia cell differentiation: DNase I hypersensitivity and DNA methylation near the globin genes. Proc. Natl. Acad. Sci. U.S.A. 79:1180-1184.
SHERMOEN, A. and S. BECKENDORF 1982 A complex of interacting DNase I hypersensitive sites near the *Drosophila* glue protein gene Sgs-4. Cell 29:601-607.
SIROTKIN, K. 1982 Developmentally regulated transcription at the 67B heat shock cluster. In "Induction of Heat Shock Proteins", M.J. Schlesinger, A. Tissieres and M. Ashburner, eds., Cold Spring Harbor Press N.Y. (in press).
SIROTKIN, K. and N. DAVIDSON 1982 Developmentally regulated transcription from *Drosophila melanogaster* chromosomal site 67B. Develop. Biology 89:196-210.
SLEDZIEWSKI, A., and E.J. YOUNG 1982 Chromatin conformational changes accompany transcriptional activation of a glucose-repressed gene in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. U.S.A. 79:253-256.
SOUTHERN, E.M. 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.
STALDER, J., M. GROUDINE, J.B. DODGSON, J.D. ENGEL and H. WEINTRAUB 1980a Hb switching in chickens. Cell 19:973-980.
STALDER, J., A. LARSEN, J.D. ENGEL, M. DOLAN, M. GROUDINE and H. WEINTRAUB 1980b Tissue specific DNA cleavages in the globin chromatin domain introduced by DNase I. Cell 20:451-460.
STORB, U., B. ARP and R. WILSON 1981 The switch region associated with immunoglobulin genes is DNase I hypersensitive in T. lymphocytes. Nature 294:90-92.
SWEET, R.W., M.U. CHAO, and R. AXEL 1982 The structure of the TK gene promoter: Nuclease Hypersensitivity correlates with expression. Cell (in press).
VAN DER PUTTEN, H., W. QUINT, I.M. VERMA and A. BERNS 1982 Moloney murine leukemia virus-induced tumors: Recombinant proviruses in activ chromatin regions. Nuc. Acids Res. 10: 577-592.
VARSHAVSKY, A.S., O.H. SUNDIN and M. BOHN 1979 A stretch of "late" SV40 viral DNA about 400 bp long which includes the origin of replication is specifically exposed in SV40 minichromosomes. Cell 16:453-466.
WALDECK, W., B. FOHBRING, K. CHOWDHURY, P. GRUSS and G. SAUER 1978 Origin of DNA replication in papovavirus chromatin is recognized by endogenous nuclease. Proc. Natl. Acad. Sci. U.S.A. 75:5964-5968.
WEINTRAUB, H. and M. GROUDINE 1976 Chromosomal subunits in

active genes have an altered conformation. Science 193:848-856.

WEINTRAUB, H., A. LARSEN, and M. GROUDINE 1981 α-globin gene switching during development of chicken embryos: Expression and chromosome structure. Cell 24:333-344.

WEINTRAUB, A., H. BEUG, M. GROUDIN and T. GRAF 1982 Temperature-sensitive changes in the structure of globin chromatin in lines of red cell precursors transformed by ts-AEV. Cell 28: 931-940.

WONG, Y.C., P. O'CONNELL, M. ROSBASH and S.C.R. ELGIN 1981 DNase I hypersensitive sites of the chromatin for *Drosophila malanogaster* ribosomal protein 49 gene. Nucl. Acids Res. 9:6749-6762.

WU, C. 1980 The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. Nature 286:854-860.

WU, C. and W. GILBERT 1981 Tissue-specific exposure of chromatin structure at the 5' terminus of the rat preproinsulin II gene. Proc. Natl. Acad. Sci. U.S.A. 78:1577-1580.

WU, C., P.M. BINGHAM, K.S. LIVAK, R. HOLMGREN and S.C.R. ELGIN 1979a The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence. Cell 16: 797-806.

WU, C., Y.C. WONG and S.C.R. ELGIN 1979b The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. Cell 16:807-814.