MOLECULAR AND CELLULAR BIOLOGY, Mar. 1989, p. 1289–1297
0270-7306/89/031289-09$02.00/0

Vol. 9, No. 3

Hormonal Regulation of Phosphoenolpyruvate Carboxykinase Gene Expression Is Mediated through Modulation of an Already Disrupted Chromatin Structure

Y. TONY IP, DARYL K. GRANNER, AND ROGER CHALKLEY*

Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, Tennessee 37232

Received 12 September 1988/Accepted 16 December 1988

We used indirect end labeling to identify a series of five hypersensitive (HS) sites in the phosphoenolpyruvate carboxykinase (PEPCK) gene in H4IIE rat hepatoma cells. These sites were found at −4800 base pairs (bp) (site A), at −1300 bp (site B), over a broad domain between −400 and −30 bp (site C), at +4650 bp (site D), and at +6200 bp (site E). Sites A to D were detected only in cells capable of expressing the PEPCK gene, whereas site E was present in all of the cells examined thus far. The HS sites were present in H4IIE cells even when transcriptional activity was reduced to a minimum by treatment with insulin. Stimulation of transcription by a cyclic AMP analog to a 40-fold increase over the insulin-repressed level did not affect the main features of the HS sites. Furthermore, increased transcription did not disrupt the nucleosomal arrangement of the coding region of the gene, nor did it affect the immediate 5′ region (site C), which is always nucleosome-free. In HTC cells, a rat hepatoma line that is hormonally responsive but unable to synthesize PEPCK mRNA, the four expression-specific HS sites were totally absent. Our experimental results also showed that, although there is a general correlation between lack of DNA methylation and transcriptional competence of the PEPCK gene, the role, if any, of methylation in the regulation of PEPCK gene activity is likely to be exerted at very specific sites.

Interphase eucaryotic chromatin consists of arrays of nucleosomes composed of about 160 base pairs (bp) of DNA wrapped around a core histone octamer. This primary organization is further folded into higher-order structures in the presence of histone H1 (12, 29). Within this compact chromatin organization, there must be certain special features that allow the transcription machinery to recognize and initiate RNA synthesis. Many studies have demonstrated that active genes are contained in a chromatin conformation different from that of inactive genes (for reviews, see references 12, 17, 18, and 29). This difference in packaging is revealed by increased sensitivity toward nuclease digestion, mostly with DNase I and micrococcal nuclease (MNase) (6, 7, 22).

Genes that are active or potentially active usually possess a canonical nucleosomal organization (1, 36, 40). On the other hand, it has also been argued that some particularly active genes may lose elements of normal nucleosomal arrangement (4, 10, 37). Genes that are transcriptionally active reside in a region of chromatin that is more sensitive to DNase I digestion than is bulk material (15, 16, 43). In addition, chromatin containing active or potentially active genes often possesses short stretches of DNA (50 to 400 bp) that exhibit even greater sensitivity to DNase I. While these DNase I-hypersensitive (HS) sites have been found 3′ to or even within certain genes (6, 28), they are most often detected at the 5′ ends of active genes (6, 21, 22, 26, 34). The presence of such DNase I-HS sites has been correlated with the expression of those genes. In several cases, these sites are located in or near sequences shown to be important in regulating the adjacent gene (19, 20, 25, 26). Although the precise nature of the makeup of DNase I-HS sites is unknown, increasing evidence indicates that a nonhistone protein(s) binds to the region and interferes with histone deposition, thus leading to increased sensitivity (13, 17, 31).

We used a multihormonally regulated system, the phosphoenolpyruvate carboxykinase (PEPCK) gene, as a model to study chromatin structural changes in relation to transcription. PEPCK, a key enzyme in gluconeogenesis, is regulated both positively and negatively by different hormones. In the H4IIE rat hepatoma cell line, PEPCK gene transcription is increased in response to glucocorticoids and cyclic AMP (cAMP) and is dominantly repressed by insulin (23, 38, 39). Specific DNA regions which are responsible for the hormonal regulation of PEPCK gene expression have been located in the 5′-flanking region of the gene (23, 30, 44).

We used both MNase and DNase I to probe the chromatin structure of the PEPCK gene operating at various levels of transcriptional activity as a consequence of treatment with different hormones. We found that the coding region of the gene is organized into regular nucleosomal arrays regardless of the transcriptional activity. However, in cells capable of expressing the PEPCK gene, the immediate 5′-flanking region (500 bp) lacked nucleosomes under all of the conditions tested. Using the technique of indirect end labeling, we found that this region was very sensitive to nuclease digestion. Within this generally sensitive region, we identified individual HS sites which correlate well with the regulatory elements of the gene. These nuclease-HS sites are tissue and expression specific, since they are present in liver-derived H4IIE cells but not in XC cells (a rat fibroblast cell line) or HTC cells (a rat hepatoma line that has lost the ability to transcribe the PEPCK gene). In addition, there is a tissue-specific HS site located 4.8 kilobases upstream of the transcription start site. Finally, two closely located HS sites were identified 100 bp beyond the polyadenylation site of the gene. The coding region of the PEPCK gene in H4IIE cells is highly methylated, while the 5′-flanking region of the gene in these cells and in rat liver is not methylated.

* Corresponding author.

MATERIALS AND METHODS

Materials. Radionucleotides, Zeta probe membrane, adult Sprague-Dawley rats, 8-(4-chlorophenylthio)-cAMP (abbreviated as cAMP in this report, since this is the active moiety), insulin, and dexamethasone were obtained from Amersham Corp., Bio-Rad Laboratories, Harlane, Boehringer Mannheim Biochemicals, Eli Lilly & Co., and Sigma Chemical Co., respectively. Chemicals were from Fisher Scientific Co. or Sigma. Nucleases were from Worthington Diagnostics. Restriction enzyme were obtained from Promega.

Cell culture. H4IIE and HTC cells were maintained as monolayer cultures in Swim 77 medium supplemented with fetal bovine and newborn calf sera at final concentrations of 2.5% (vol/vol) each (38). XC cells were maintained as monolayer cultures in Dulbecco modified Eagle medium supplemented with fetal bovine serum to a final concentration of 10%.

When hormone treatment was required, the cells were maintained in serum-free medium for 24 h after growth to confluence. Dexamethasone was added at a final concentration of 0.5 μM, insulin was added at 5 nM, and cAMP was added at 0.1 mM. Hormone treatments lasted 30 min.

Isolation of nuclei and nuclease digestion. Tissue culture cells were collected after the monolayers were washed with phosphate-buffered saline supplemented with 2.5 mM EDTA; the cells were washed twice with the same buffer. Nuclei were obtained by lysing the cells in ice-cold HB buffer (10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; pH 8], 50 mM NaCl, 1 mM EDTA, 0.25 mM ethylene glycol-bis(β-aminoethyl ether)-*N-N-N'-N'*-tetraacetic acid [EGTA], 0.5 mM spermidine, 0.15 mM spermine, 0.5% Triton X-100, 0.25 M sucrose). After the nuclear pellet was washed once in the same buffer and then twice in digestion buffer (10 mM Tris [pH 7.4], 50 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$, 0.25 M sucrose), the nuclear pellet was suspended in the digestion buffer at a concentration of 25 A_{260} U/ml. MNase was added at 0 to 500 U/ml (usually 0, 20, 50, 200, and 500 U/ml for a set of five or as otherwise indicated), and DNase was added at 0 to 25 U/ml (usually 0, 2, 8, 15, and 25 U/ml for a set of five). Digestion was started by incubation at 25°C. After 5 min, the reaction was stopped by adding sodium dodecyl sulfate to 1%, EDTA to 25 mM, and proteinase K to 200 μg/ml, and the DNA was purified by standard methods (24). Frozen samples of rat tissue were cut into small pieces and homogenized in 10 volumes of ice-cold HB buffer, and nuclease digestion was performed as described above.

Southern blot analysis. Purified DNA, digested with appropriate restriction enzymes when indicated, was resolved by electrophoresis in agarose gel. The concentrations of the gels varied from 1 to 1.6%, depending on the size of fragment sought. The gel was treated with alkali and set up for capillary transfer to Zeta probe membrane as described by the supplier. DNA fragments used as probes for hybridization were obtained after restriction enzyme digestion of cloned plasmids containing PEPCK gene sequences. The genomic organization of the PEPCK gene and the indirect end-labeling probes are shown in Fig. 1. The polyacrylamide gel-purified DNA fragments were used as templates for oligo-primed labeling (14). A specific activity of 10^9 cpm/μg was consistently obtained. Prehybridization and hybridization were performed at 65°C for 3 and 16 h, respectively, in a buffer containing 1.5× SSPE (1× SSPE is 180 mM NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA [pH 7.4]), 1% Carnation nonfat dry milk, 2% sodium dodecyl sulfate, 10% dextran sulfate, and 200 μg of denatured salmon sperm DNA per ml. The filters were washed as described by the supplier, and the final wash was at 65°C in 0.2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 0.1% sodium dodecyl sulfate.

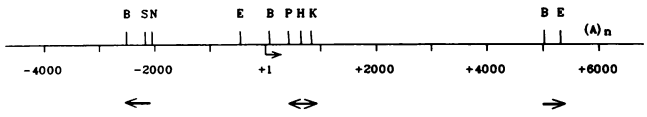

FIG. 1. Genomic organization of the PEPCK gene. The horizontal line represents both the coding and the flanking regions of the gene. The numbers below the horizontal line indicate the size of the gene in base pairs, with +1 as the transcription start site. $(A)_n$ is the polyadenylation site of the gene. The restriction sites frequently referred to in this report are indicated as follows above the line: B, *Bgl*II; S, *Sac*I; N, *Nco*I; E, *Eco*RI; P, *Pvu*II; H, *Hin*dIII; K, *Kpn*I. The arrows below the line represent the DNA fragments used for indirect end labeling in the experiments described in this communication.

RESULTS

No alteration of nucleosomal organization of the PEPCK gene coding region on hormonal activation of transcription. An unsolved problem in the field of chromatin structure is whether nucleosomes must be displaced to facilitate the passage of the polymerase II complex in the process of transcription. Some systems seem to be devoid of nucleosomes during vigorous RNA synthesis (10, 17, 37), but whether this is a cause or an effect of transcription is undetermined. In this regard, the PEPCK gene system is a good model, because the transcription rate varies 40-fold in H4IIE cells, depending on the hormonal treatment.

Nuclei isolated from control and hormone-treated H4IIE cells were digested with increasing amounts of MNase (see Materials and Methods) to assess whether canonical nucleosomes were present on the PEPCK gene. Ethidium bromide staining of purified total genomic DNA showed a typical nucleosomal ladder (Fig. 2A). When a PEPCK cDNA probe which spans the last three exons (+3000 to +5700) (2) was used, it was apparent that nucleosomal arrays are present in this region (Fig. 2B). The 6- and 15-fold-increased rates of transcription caused by dexamethasone or cAMP treatment, respectively (38), had no detectable effect on nucleosomal structure in the coding region of the PEPCK gene. The nucleosomal pattern was also assessed in H4IIE cells treated with insulin, which depresses the basal transcription level. In comparison with cells treated with cAMP or dexamethasone, no differences in nucleosomal organization were detected in the coding region of the gene (data not shown). A similar analysis performed on DNA isolated from XC cells, in which the PEPCK gene is not transcribed, also showed an identical nucleosomal pattern in this region (data not shown). Thus, within the limits of the analysis, no disruption of the nucleosomal array in the PEPCK gene was detected, an indication that there is no extensive rearrangement of nucleosomes over the coding region in response to a substantial change in the level of transcription. Obviously, the possibility that disruption and reformation of individual nucleosomes occur close to the polymerase complex during its passage along the gene cannot be excluded. The extents of MNase digestion were comparable in all of the cases tested (compare the ethidium bromide staining and the

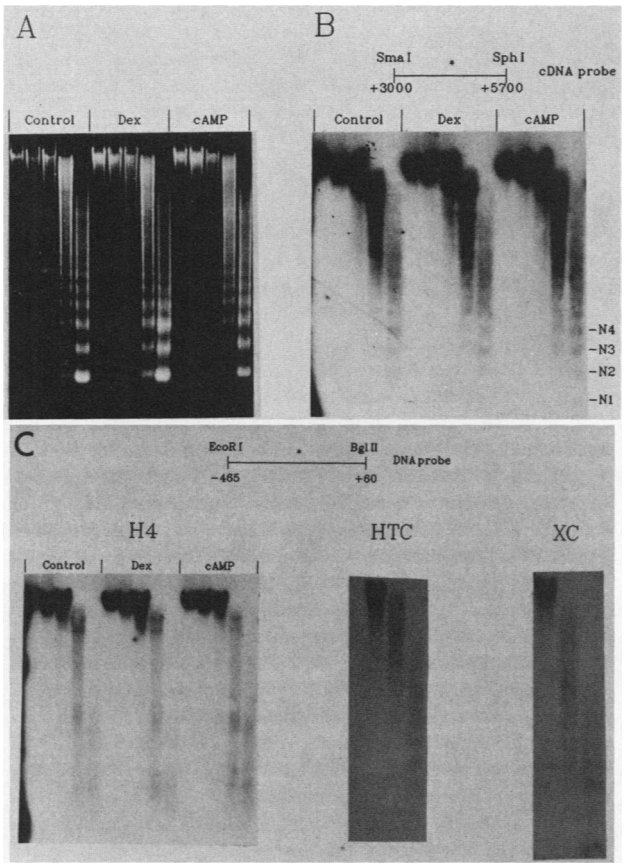

FIG. 2. Nucleosomal organization of the PEPCK gene. H4IIE hepatoma cells were grown to confluence and maintained in serum-free medium for 24 h. Nuclei were isolated from control cells or cells treated with hormones for 30 min, as indicated. Dex, Dexamethasone. Increasing amounts of MNase, as described in Materials and Methods, were used for 5 min of digestion at 25°C. DNA was purified and resolved in 1.4% agarose gel. (A) Ethidium bromide staining of the resolved total DNA. (B) Hybridization of the transferred DNA with a cDNA probe which covered the last three introns of the PEPCK gene. N1 through N4 indicate DNAs with sizes corresponding to the lengths of mono-, di-, tri-, and tetranucleosomes, respectively. (C) H4IIE cells treated as described above. HTC and XC cells were used with no hormone treatment. Nuclear isolation and nuclease digestion were done as described above, except that the amounts of MNase used in HTC and XC cells were 50, 200, and 500 U/ml, respectively. The transferred DNA was probed with a DNA fragment which spanned −465 to +60, where +1 is the transcription start site.

blots). There was no indication of more-rapid kinetics of digestion characteristic of chromatin lacking histone H1. While this does not constitute definitive proof for the presence of H1, it certainly renders it unlikely that this linker protein is completely absent from this gene, even in the most highly transcribed cases.

The same strategy was used to examine the chromatin structure of a key regulatory domain of the PEPCK gene. The region from about −500 to the cap site includes the basal promoter elements, the cAMP response element (33, 44), and the proximal PEPCK gene glucocorticoid response elements (GREs) (30). With a probe that covers the region from −465 to +60 bp, we found a result drastically different from that observed in the analysis of the coding region (Fig. 2C). In H4IIE cells, the pattern was highly diffuse in both control and transcriptionally activated cells, with little evidence of nucleosomal structure. In contrast, XC cells, which do not express this gene and do not respond to hormones, possess a normal nucleosomal arrangement in this region. This excludes the trivial possibility that the DNA sequence in this region somehow excludes histones. Interestingly, in the HTC rat hepatoma cell line, which for some unknown reason does not express the PEPCK gene, there was also no sign of chromatin disruption in this region.

We probed other regions of the gene in H4IIE cells (including the 5′-flanking region out to −2500 bp) and found nucleosomes across the entirety of the gene. Although nuclease-HS sites were detected in other portions of the gene (as shown in sections below), no other region had an extensively disrupted nucleosomal arrangement.

Mapping of MNase-HS sites. The difference in nucleosomal organization around the 5′ regulatory region led us to assay more precisely for the details of protein arrangement in this zone by the techniques of indirect end labeling. Nuclei were isolated from H4IIE cells after treatment with different hormones and from untreated HTC and XC cells. The nuclei were briefly digested with MNase, and the DNA was isolated. Purified genomic DNA was also digested with nuclease as a control, since MNase cleaves DNA with some sequence specificity (7). The DNA samples were then digested with *Hin*dIII (sites at −599 and +637 bp) and analyzed with a small *Pvu*II-*Hin*dIII fragment (+422 to +637). The restriction fragment and the indirect end-labeling probe are indicated in Fig. 3A. In addition to the intact 1,236-bp parent band, sets of subbands were clearly visible (Fig. 3A). Four sharp bands were located at −400, −345, −265, and −150 bp. In addition, an exceptionally sensitive site was detected at −60 bp. The region from −150 to −60 bp was more resistant to nuclease attack. Treatment with the hormones individually or in combination resulted in no detectable difference in the banding pattern (Fig. 3). This pattern of HS sites was totally absent from the nontranscribed PEPCK chromatin in HTC and XC cells, and it is not likely to be the result of sequence-specific cutting of MNase, since the pattern was not detected in naked rat genomic DNA. More-extended digestion (right lane of each pair in Fig. 3A) revealed somewhat broader zones of cleavage between the *Hin*dIII site at +637 bp and the strong site at −60 bp (see the bottom of the autoradiograph). They were measured to be at +140 and +300 bp. The simplest interpretation of this finding is that the first two nucleosomes on the coding region of the gene are phased.

The organization of chromatin in this region was also analyzed by probing from the 5′ direction. The MNase-treated DNA was digested with *Bgl*II (sites at −2500 and +60 bp) and probed with a *Bgl*II-*Nco*I (−2500 to −2100 bp) fragment (Fig. 3B). The autoradiograph revealed a complicated picture. Two bands were apparent in the region immediately 5′ to the transcription start site (Fig. 3B, arrows). The decreased number of detectable sites in this region is probably due to the decreased resolution of the agarose gel at this long distance and the highly complex nature of this area. In addition, another nuclease-HS site was detected at −1240 bp (Fig. 3B, closed triangles). This site appeared only in H4IIE cells, and its presence was independent of hormone treatment. Another nearby band also appeared in this analysis (Fig. 3B, open triangles). The latter was clearly a DNA sequence-specific cleavage site of MNase, since it appeared in all cases, including the naked genomic DNA. However, it is noteworthy that this site was cut much more extensively in H4IIE nuclei than in the other

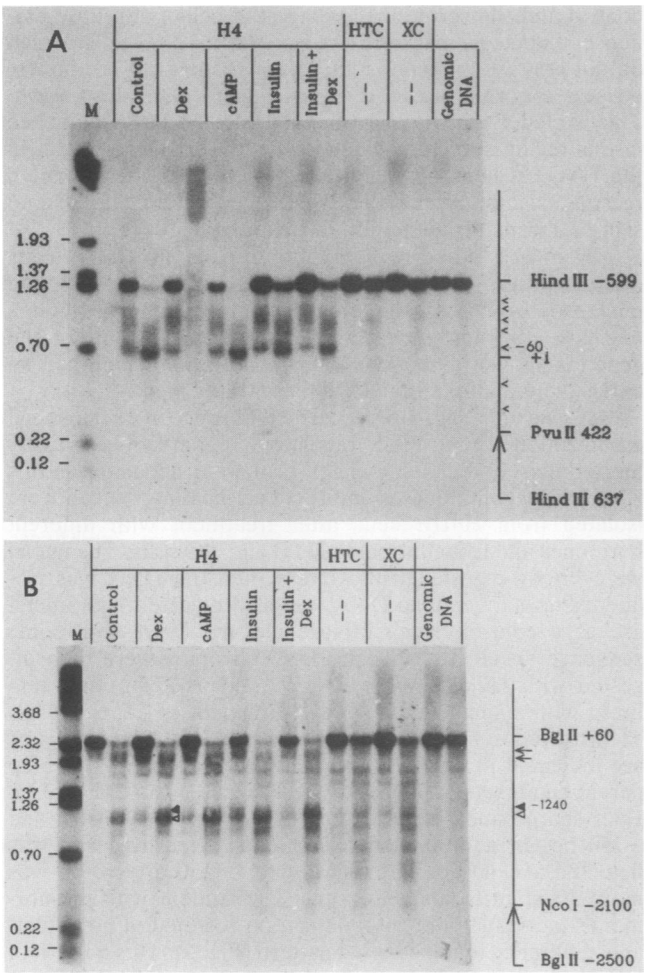

FIG. 3. MNase-HS site mapping by indirect end labeling. Cell maintenance and hormone treatment were conducted as described in the legend to Fig. 2. The amounts of MNase used were 50 and 200 U/ml for each set of experiments. The purified DNA was digested with restriction enzymes and resolved in a 1.4% agarose gel, followed by Southern blot analysis. (A) DNA was digested with *Hin*dIII and probed with a *Pvu*II-*Hin*dIII DNA fragment. Lane M contained *Bst*EII-cut λ DNA molecular weight markers; the numbers to the left represent kilobase pairs. (B) DNA was digested with *Bgl*II and probed with a *Bgl*II-*Nco*I fragment. Dex, Dexamethasone.

cell types. It appears that this region is more open for nuclease attack in H4IIE cells than in other nonexpressing cell lines. A distal GRE has been mapped between −1180 and −1140 bp (30), and a correspondance between the GRE and the new HS site described above seems highly likely. The nuclease might have detected this region because of protein binding to the regulatory domain, which leads to disruption of the chromatin structure. However, it seems unlikely that hormone-modulated binding of the glucocorticoid receptor alone contributes to this event, since we did not detect a hormone-dependent difference in the site. The bands in the lower part of the autoradiograph are probably DNA sequence-specific cleavage sites of the nuclease.

Identification of DNase I-HS sites across the PEPCK gene. (i) HS sites in the 5′ portion of the gene. In many of the cases studied, sites that are MNase HS are also sensitive to DNase I (7, 22). Interpretation of DNase I sensitivity is less complicated than that of MNase sensitivity, since the former enzyme does not have a high affinity for internucleosomal regions and therefore generates a much-reduced number of cuts, thus rendering the analysis easier. Based on the initial picture indicated by MNase digestion, the indirect end-labeling technique was used to determine whether DNase I-HS sites could be detected in the PEPCK chromatin.

Nuclei from H4IIE or XC cells were digested minimally with DNase I, and the purified DNA was then cut with *Kpn*I (sites at −6540 and +878). This combination resulted in a 7,413-bp PEPCK gene fragment plus any HS site subfragments generated by the DNase I digestion. When the purified DNA was probed with the 241-bp *Hin*dIII-to-*Kpn*I fragment (sites at +637 to +878 bp), a complex array of HS zones was obtained. The zone, defined as site C, located from −367 to −30 bp, is indicated by arrowheads in Fig. 4. Close scrutiny revealed a region within this zone from −224 to −30 bp which was resistant to nuclease digestion. The region from −367 to −244 bp was somewhat more sensitive to digestion than that at −30 bp, because the former appeared early in digestion. In addition, a second distinct site (site B) was detected at −1400 bp and a third site (site A) was noted far upstream (at about −4800 bp, as defined more precisely with a probe much closer to the site [data not shown]). The PEPCK gene in XC cells was completely resistant to DNase I under the conditions used, and no evidence of HS was seen (Fig. 4).

(ii) HS sites in the 3′ portion of the PEPCK gene. The rest of the PEPCK gene, from +879 to +10470 bp, was screened for HS sites by indirect end labeling. After DNase I digestion, the purified DNA was cut with *Hin*dIII and probed with the short *Hin*dIII-*Kpn*I probe (from +637 to +878 bp) used previously. This experiment was performed with nuclei isolated from H4IIE cells treated with various hormones or from XC and HTC cells that do not express this gene. The results are shown in Fig. 5A. A weak HS site (site D) was identified at +4650 bp. As noted for the other three sites described above, this site was present only in cells that express the PEPCK gene, and the characteristics of the site seemed to be independent of hormonal treatment and the consequent level of transcriptional activity. A strongly HS site located at about +6200 bp (site E) is different from those previously reported in that it was present in all of the tissues examined so far.

The experiment shown in Fig. 5B indicated that site E was actually composed of two closely located HS sites when a probe much closer to the site was used. These sites are 120 bp apart, with the first site 100 bp 3′ to the polyadenylation signal of the PEPCK gene. These two sites were present in all of the cell lines tested. This is very unlikely to be due to sequence-specific recognition of DNase I, since the same two sites are also specifically cleaved in an HS manner by MNase (data not shown).

Methylation status of the PEPCK gene in a range of tissues. The availability of a wide variety of cell types which vary in transcriptional activity of the PEPCK gene allowed analysis of the relationship of gene activity and DNA methylation. There are 14 CpG dinucleotide sequences, potential targets for DNA methylation (11), in the broad regulatory region-complex HS site C from −600 to +48 bp (see Fig. 7, dots); three of these dinucleotides are in *Msp*I-*Hpa*II restriction sites (CCGG). The methylation status of these, as well as three additional sites external to the regulatory region, was assessed with a probe from −465 to +637. The details of the cleavage map are shown in Fig. 6. The tissues studied included liver, kidney, the liver-derived H4IIE cell line (in which the PEPCK gene is expressed), lung, sperm, XC cells,

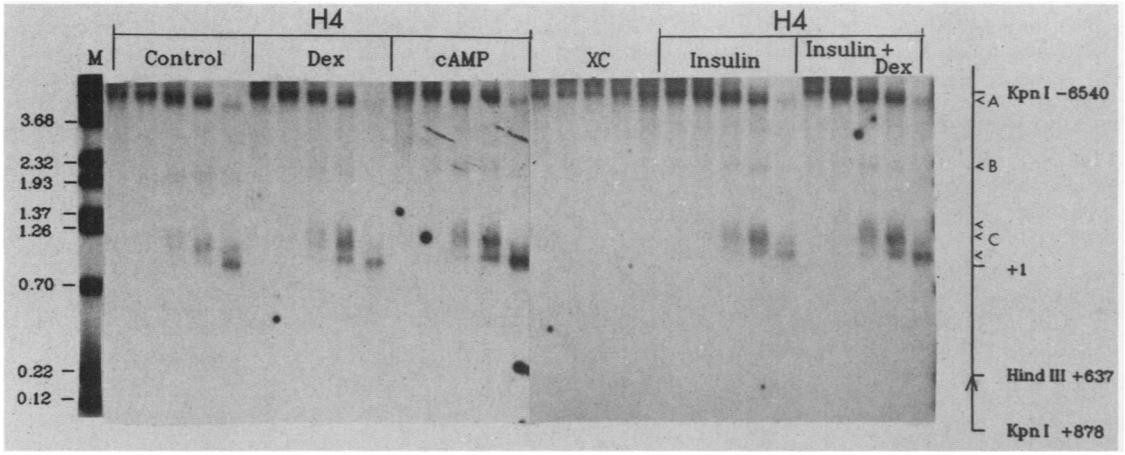

FIG. 4. DNase I-HS sites in PEPCK chromatin. Cell maintenance and hormone treatment were conducted as described in the legend to Fig. 2. The isolated nuclei were digested with increasing amounts of DNase I (see Materials and Methods), and the DNA was purified. After *Kpn*I digestion, the DNA was electrophoresed in a 1.2% agarose gel, transferred, and probed with an *Hin*dIII-*Kpn*I fragment. The lanes labeled H4 contained DNA from H4IIE cells treated with the hormones indicated. Dex, Dexamethasone. The lanes labeled XC contained DNA from XC cells with no hormone treatment. The numbers to the left indicate the molecular size markers (in kilobase pairs) in lane M.

and the liver-derived HTC cell line (in which this gene is not expressed). The results are shown in Fig. 6.

H4IIE cells clearly contain all of the expected CCGG sequences (predicted from the DNA sequence of the cloned liver PEPCK gene [2]), as indicated by the sizes of the restriction fragments obtained by *Msp*I digestion. Digestion with *Hpa*II indicated that the sites at −343, −101, and +48 bp are all unmethylated, whereas the other nearby sites (at −1748, +242, and +832 bp) are extensively modified. Clearly, methylation at these latter sites does not interfere

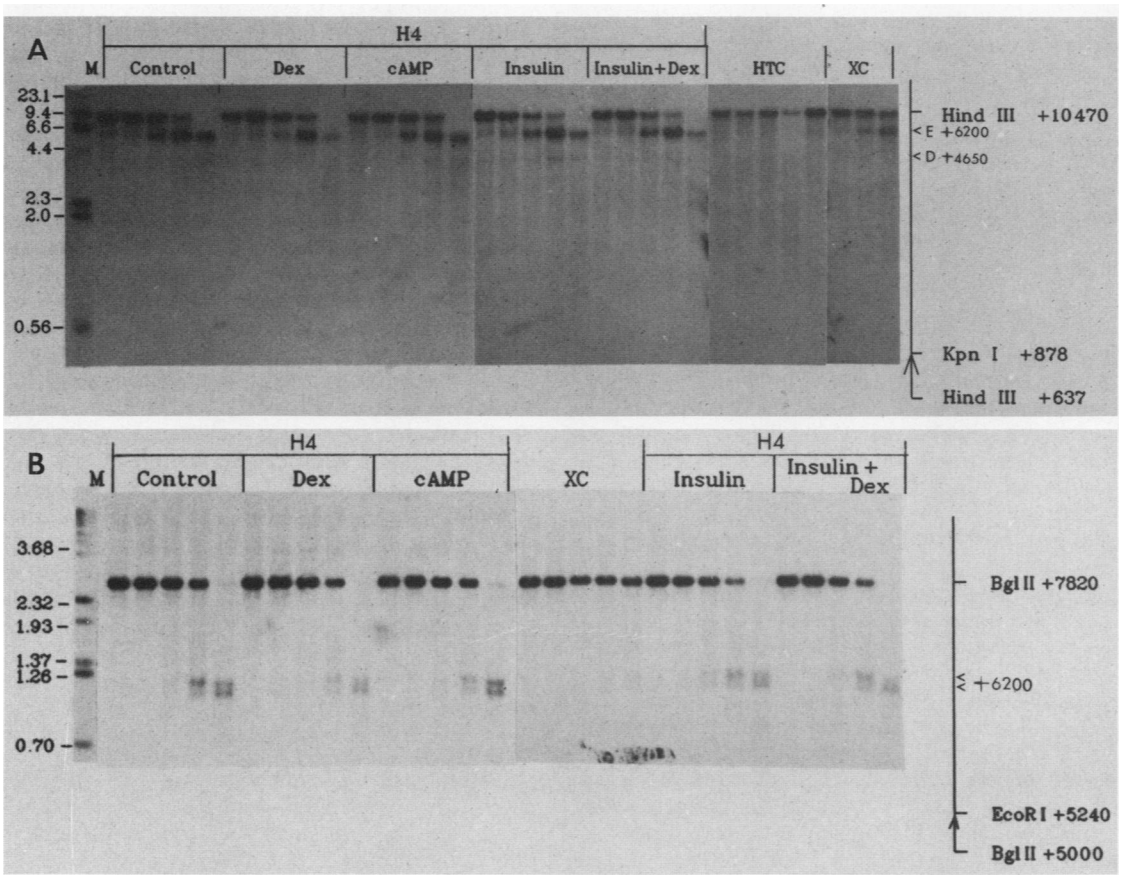

FIG. 5. DNase I-HS sites in the 3′ end of PEPCK chromatin. The experiment was performed as described in the legend to Fig. 4. Purified DNA was digested with *Hin*dIII and probed with a *Hin*dIII-*Kpn*I fragment (A) or digested with *Bgl*II and probed with a *Bgl*II-*Eco*RI fragment (B). The numbers to the left indicate the molecular size markers (in kilobase pairs) in lanes M.

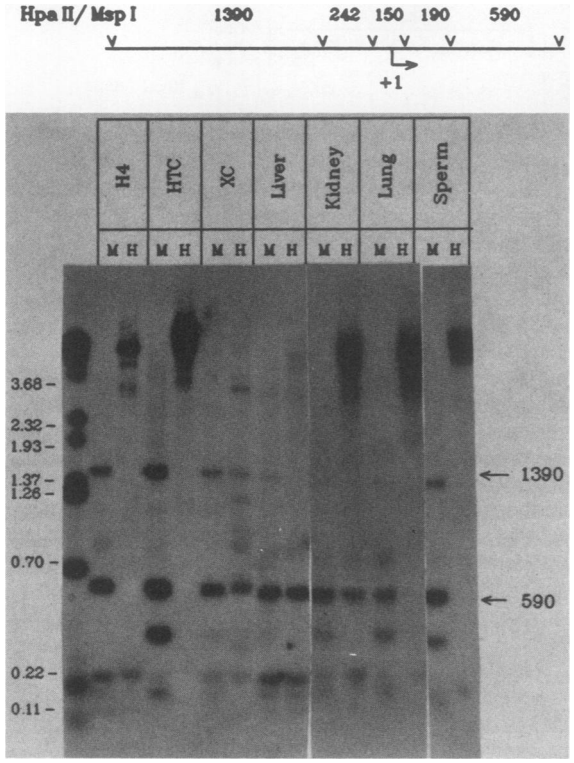

FIG. 6. DNA methylation status of the regulatory domain of the PEPCK gene. Genomic DNA purified from various cell lines or rat tissues was digested with either *Msp*I (M) or *Hpa*II (H). The DNA was then analyzed by the Southern blot technique with a PEPCK DNA fragment from −465 to +637 as a hybridization probe. The arrowheads in the diagram at the top show the *Msp*I cutting sites, and the numbers between the arrowheads indicate the expected sizes of the fragments. The band that appeared in the blot in many tissues between 590 and 242 bp was not expected on the basis of sequence analysis and may be an inaccessible site or point mutation in the *Msp*I site at −101 bp. The numbers to the left indicate the molecular size markers (in kilobase pairs) in the unlabeled lane.

with transcriptional activity. Liver and kidney DNAs showed a pattern similar to that found in H4IIE cell DNA, except that the neighboring CCGG sites were less methylated in liver and somewhat more methylated in kidney. Thus, for all of these tissues in which the PEPCK gene is expressed, although there may be variations in overall CCGG modification, the sites at −343, −101, and +68 bp were invariably unmethylated. In contrast, in nonexpressing lung and HTC cells these sequences are part of an extensive domain which is heavily methylated, although lung tissue does show a small amount of cleavage at −343 and −101 bp. Sperm DNA is also extensively methylated in the region of the PEPCK gene. There appears to be an excellent correlation among methylation of certain promoter-proximal CCGG sequences, the absence of HS sites, and the attendant transcriptional inactivity. However, when we analyzed DNA obtained from XC cells (incapable of transcribing the PEPCK gene and no HS sites upstream of the gene), the detectable methylation in this region was very low. The major differences between XC and H4IIE cells lie in the presence of methylation at the *Msp*I-*Hpa*II site at +68 bp in XC DNA and the curious partial-methylation pattern throughout the rest of the regulatory region.

DISCUSSION

We analyzed the chromatin organization of the PEPCK gene in a range of cell types in which the gene is either capable (H4IIE cells and liver) or incapable (XC and HTC cells, lung, and brain) of being expressed and regulated. The PEPCK gene in H4IIE cells contains five DNase I-HS sites (Fig. 7). These are situated at sites A (−4800 bp), B (−1400 to −1240 bp), C (−400 to −30 bp), D (+4650 bp), and E (+6200 bp). Sites A to D are found exclusively in cells that express the gene, whereas the 3′ site E is found in all of the rat cells examined thus far, regardless of the activity of the gene. Sites A to D are presumably involved in some way with expression of the gene in specific cell types. A DNA sequence close to site B has been identified as the distal GRE (30), and sequences within site C contain the proximal GRE (30), the cAMP response element, and various elements involved in basal expression of the gene (23, 33, 44). In contrast, site E is likely to be involved in some function which is common to all of the cell types examined. Possible functions for this zone include action as a replication origin, action as a site of attachment to the nuclear matrix, and action as a topoisomerase II attachment site (27, 41, 42). Furthermore, the 3′ end of PEPCK mRNA has been shown to be important in the stability of the molecule (D. Petersen and D. Granner, manuscript in preparation). This raises the possibility that site E has a special DNA structure that is retained in the mRNA molecule to confer additional stability. This region of the gene has not been sequenced, and these possibilities are currently under investigation.

Sites A to D are present in H4IIE chromatin over a wide range of transcriptional activities under the influence of a variety of hormones. We do not believe that these findings contradict the results of Richard-Foy and Hager and Zaret and Yamamoto (35, 45), who reported that HS sites were induced upon treatment of constructs containing the mouse mammary tumor virus enhancer with dexamethasone. These workers were studying the shift of a toally inactive gene to the active state following hormone administration, so they may have observed a different level of control. Nonetheless, our findings appear to exclude a role for the hormone-bound glucocorticoid receptor in specifying any of these sites in the PEPCK chromatin as an immediate response to hormone treatment, since in the absence of any hormone (the control cells are exposed to serum-free medium which is depleted of hormones) we still detected all of the sites. In this regard, the results obtained with HTC cells were particularly revealing. HTC cells were (like H4IIE cells) derived from rat liver. They possess the capacity to respond to glucocorticoids, as evidenced by the induction of tyrosine aminotransferase (32). However, these cells have lost the capacity to synthesize PEPCK mRNA, even though the PEPCK gene sequence is still present and, as far as can be determined by restriction mapping, is intact (data not shown). Curiously, HTC cells have lost the ability to form sites A to D. This implies that there is some sort of hierarchy in establishing and maintaining the HS sites and that if a key dependent stage is interfered with in some way, the whole hierarchy falls.

The most-detailed dissection of HS site C comes from indirect end-labeling analysis after controlled MNase digestion. Site C covers some 400 bp of DNA immediately 5′ to the transcriptional start site and is divided into two HS zones, with a region of resistance to nuclease digestion between the two zones from −150 to −60 bp. The more 5′ of the two HS zones shows four points of nuclease cutting

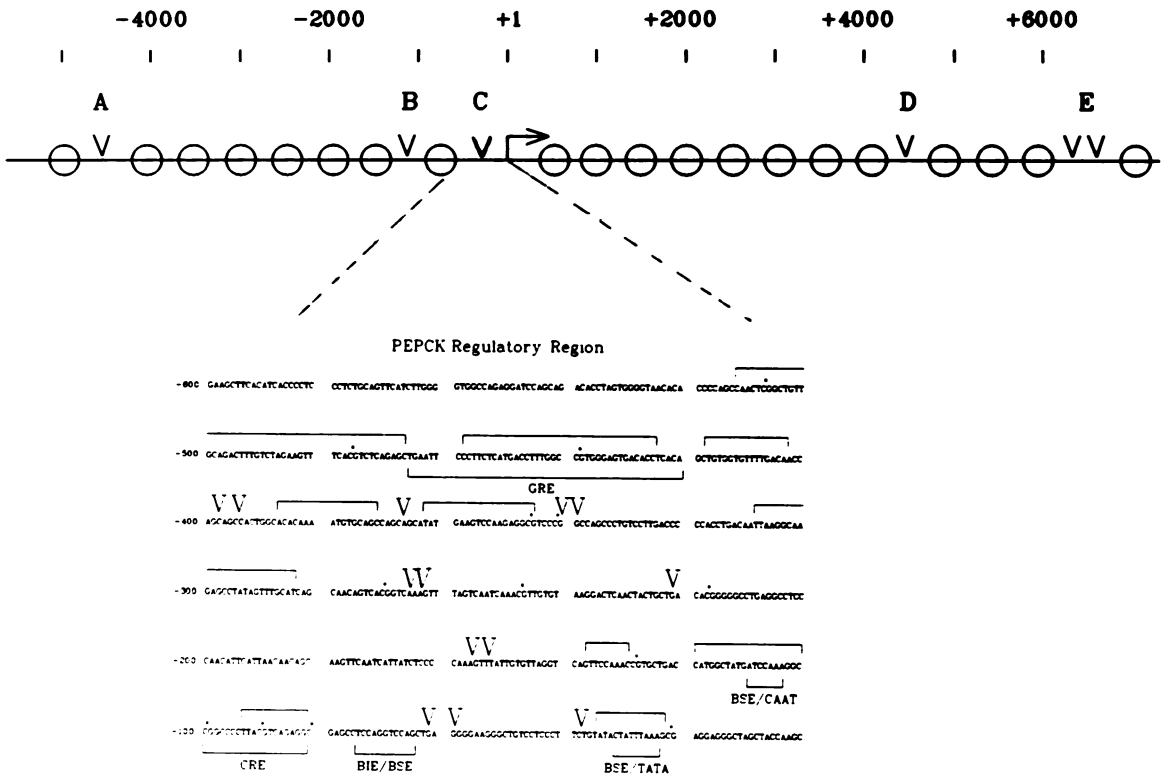

FIG. 7. Organization of PEPCK chromatin. The line represents DNA, and the circles represent nucleosomal structures. Nuclease-HS sites are indicated by V's and are designated as sites A to E. The lower portion is an enlargement of site C, which is the promoter–hormone-responsive region of the gene (2, 29, 32). The brackets above the sequence show the in vitro footprints obtained with rat liver extracts (Wong and Granner, unpublished data). The brackets below the sequence indicate the locations of functional elements defined by transfection studies. V's indicate DNase I-HS sites, and VV's indicate MNase-HS sites. The dots indicate the CpG sequences in this region.

spaced some 50 to 100 bp apart (Fig. 3A). The 3′ zone, from −60 to +1 bp, exists as a continuous region of sensitivity. These probably reflect protein-DNA interactions between −150 and −60 bp. The basal enhancer-cAMP response element was mapped from −82 to −100 bp. Furthermore, sequence analysis revealed that this region is close to CAAT and nuclear factor 1 consensus sequences. We suspect that the overall protected region between HS domains within site C is due to an extended array of nonhistone proteins presumably involved in the regulation of transcription from the PEPCK gene promoter. Indeed, nuclear extracts from rat liver protect a range of specific sequences in this region (T. W. Wong and D. K. Granner, unpublished data), indicating that many different proteins are able to bind, at least in vitro.

In view of the association of extensive protein binding to the C site in liver cells and the array of internal nuclease-sensitive sites, we need to draw attention to a remarkable aspect of these findings. In HTC cells, which presumably contain at least most of the proteins that bind to the C domain (particularly since at least some of these proteins are likely to be common, ubiquitous transcription factors), there is no evidence of any protein binding to the C region. Since these proteins bind DNA in vitro with high affinity and are able to bind in transcriptionally competent H4IIE cells, we conclude that they are inhibited from binding to the appropriate zones in transcriptionally inert cells (such as HTC cells), whereas a mechanism exists to facilitate their binding in cells able to transcribe the PEPCK gene. An understanding of such a mechanism should be of profound significance for our understanding of differential gene control in differentiated cells.

Identification of the limits of the HS sites enabled us to study the nature of the remaining chromatin over a wide range of transcriptional activities. We found that, regardless of the amount of transcription, the PEPCK chromatin was largely organized in canonical nucleosomal arrays. This includes the sequences both far distal to the gene and within the coding regions. The rate of production of nucleosomal arrays by MNase digestion was constant throughout, indicating that the chromatin substrate did not, in all likelihood, significantly lack H1, the loss of which would have been expected to lead to a much greater rate of digestion. The finding that apparently normal nucleosomes were present, even on chromatin that was being transcribed vigorously, was interesting, since contradictory results have been reported in this regard (4, 10, 36, 37, 40). Clearly, gross nucleosomal disruption is not a sine qua non for transcription, although extremely vigorous transcription might be expected to lead to more-diffuse nucleosomal patterns. The maximal rate of transcription of the PEPCK gene, however, is not small (the transcription rate is 1,300 ppm when induced with cAMP [38]), and we estimate that up to 30 polymerase molecules (from a total of about 24,000 polymerase molecules in the cell [9]) could be associated with the coding region (6,000 bp and 30 nucleosomes) at any time. Another possibility which might account for an apparently diffuse nucleosome pattern is that the hybridization probe encompasses part of a HS zone (Fig. 2C and 3A). This may

be of considerable significance if the coding domain contains internal HS sites.

DNA methylation has also been correlated with regulation of gene expression (8, 11). Bevenisty and co-workers (5) have shown that during development, the rat PEPCK gene goes through sequential changes in DNA methylation. Our analysis confirms this idea in general. Thus, early in development in the sperm the gene is almost completely methylated. In contrast, in the adult liver the gene is largely unmethylated, while in other nonexpressing tissues these sequences are methylated. An intriguing finding concerns the three cell lines tested. H4IIE cells show very extensive methylation in the coding region, while at least three of the CpG sites in the regulatory region are unmethylated. In contrast, in liver-derived HTC cells (which do not express the PEPCK gene) both regulatory and coding sequences are fully methylated. This implies that methylation at the coding region does not affect the ongoing transcription process, while methylation at critical sites may lead to inhibition of gene expression, probably through interference of DNA-protein interactions (3). However, somewhat surprising are the results obtained with the XC fibroblast cell line, which also does not express the PEPCK gene. In these cells, the extent of methylation is quite low, although one site within the regulatory-promoter region is methylated. Thus, extensive methylation cannot explain the total absence of HS sites in XC cells. Since the fibroblast is never programmed to express the PEPCK gene, it may use means other than methylation to keep the gene in a quiescent state. Nonetheless, we can conclude that, in general, tissues which do not express this gene show a more extensively methylated regulatory region than cells which can express the gene, in which case the DNA of the regulatory region is fully unmethylated at the sites available for assay. Further analysis must clearly await genomic sequencing and direct determination of all methylated sites throughout this region of the DNA in these various cell types.

ACKNOWLEDGMENTS

We thank our many colleagues within our laboratories for their interest and support. In particular, we thank D. Petersen and M. Magnuson for a generous supply of plasmids. We thank P. Quinn for making us aware of his unpublished data.

This work was supported by Public Health Service grants from the National Institutes of Health and by the Vanderbilt Diabetes Research and Training Center.

LITERATURE CITED

1. **Allegra, P., R. Sterner, D. F. Clayton, and V. G. Allfrey.** 1987. Affinity chromatographic purification of nucleosomes containing transcriptionally active DNA sequences. J. Mol. Biol. **196**:379–388.
2. **Beale, E. G., N. B. Chrapkiewicz, H. A. Scoble, R. J. Metz, D. P. Quick, R. L. Noble, J. E. Donelson, K. Biemann, and D. K. Granner.** 1985. Rat hepatic cytosolic phosphoenolpyruvate carboxykinase: structures of the protein, messenger RNA and gene. J. Biol. Chem. **260**:10748–10760.
3. **Becker, P. B., S. Ruppert, and G. Schutz.** 1987. Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. Cell **51**:435–443.
4. **Bellard, M., G. Dretzen, F. Bellard, P. Oudet, and P. Chambon.** 1982. Disruption of the typical chromatin structure in a 2500 basepair region at the 5′ end of the actively transcribed ovalbumin gene. EMBO J. **1**:223–230.
5. **Bevenisty, N., D. Mencher, O. Meyuhas, A. Razin, and L. Reshef.** 1985. Sequential changes in DNA methylation patterns of the rat phosphoenolpyruvate carboxykinase gene during development. Proc. Natl. Acad. Sci. USA **82**:267–271.
6. **Burch, J. B. E., and H. Weintraub.** 1983. Temporal order of chromatin structural changes associated with activation of the chicken vitellogenin gene. Cell **33**:65–76.
7. **Cartwright, I. L., and S. C. R. Elgin.** 1986. Nucleosomal instability and induction of new upstream protein-DNA association accompany activation of four small heat shock protein genes in *Drosophila melanogaster*. Mol. Cell. Biol. **6**:779–791.
8. **Cedar, H.** 1988. DNA methylation and gene activity. Cell **53**:3–4.
9. **Cochet-Meilhac, M., P. Nuret, J. C. Courvalin, and P. Chambon.** 1974. Animal DNA-dependent RNA polymerases: determination of the cellular number of RNA polymerase B molecules. Biochim. Biophys. Acta **353**:185–192.
10. **Cohen, R. B., and M. Sheffery.** 1985. Nucleosome disruption precedes transcription and is largely limited to the transcribed domain of globin genes in murine erythroleukemia cells. J. Mol. Biol. **182**:109–129.
11. **Doerfler, W.** 1983. DNA methylation and gene activity. Annu. Rev. Biochem. **52**:93–124.
12. **Eissenberg, J. C., I. L. Cartwright, G. H. Thomas, and S. C. R. Elgin.** 1985. Selected topics in chromatin structure. Annu. Rev. Genet. **19**:485–536.
13. **Emerson, B. M., L. D. Lewis, and G. Felsenfeld.** 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult β-globin gene: the nature of the binding domain. Cell **41**:21–30.
14. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
15. **Garel, A., and R. Axel.** 1976. Selective digestion of transcriptionally active ovalbumin genes from oviduct nuclei. Proc. Natl. Acad. Sci. USA **73**:3966–3970.
16. **Garel, A., M. Zolan, and R. Axel.** 1977. Genes transcribed at diverse rates have a similar conformation in chromatin. Proc. Natl. Acad. Sci. USA **74**:4867–4871.
17. **Gross, D. S., and W. T. Garrard.** 1988. Nuclease hypersensitive sites in chromatin. Annu. Rev. Biochem. **57**:159–197.
18. **Hebbes, T. R., A. W. Thorne, and C. Crane-Robinson.** 1988. A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J. **7**:1395–1402.
19. **Hemenway, C., and D. M. Robins.** 1987. DNase I-hypersensitive sites associated with expression and hormonal regulation of mouse C4 and S1P gene. Proc. Natl. Acad. Sci. USA **84**: 4816–4820.
20. **Jackson, P. D., and G. Felsenfeld.** 1985. A method for intranuclear footprinting and its application to a nuclease hypersensitive site in chromatin. Proc. Natl. Acad. Sci. USA **82**:2296–2300.
21. **Keene, M. A., V. Corces, K. Lowenhaupt, and S. C. R. Elgin.** 1981. DNase I hypersensitive sites in *Drosophila* chromatin at the 5′ end of regions of transcription. Proc. Natl. Acad. Sci. USA **78**:143–146.
22. **Levy-Wilson, B., C. Fortier, B. D. Blackhart, and B. J McCarthy.** 1988. DNase I and micrococcal nuclease-hypersensitive sites in the human apolipoprotein B gene are tissue specific. Mol. Cell. Biol. **8**:71–80.
23. **Magnuson, M. A., P. G. Quinn, and D. K. Granner.** 1987. Multihormonal regulation of phosphoenolpyruvate carboxykinase-chloramphenicol acetyltransferase fusion gene: insulin's effects oppose those of cAMP and dexamethasone. J. Biol. Chem. **262**:14917–14920.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 458–462. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **McGinnis, W., A. W. Shermoen, J. Heemskerk, and S. K. Beckendorf.** 1983. DNA sequence changes in an upstream DNase I-hypersensitive region are correlated with reduced gene expression. Proc. Natl. Acad. Sci. USA **80**:1063–1067.
26. **Nahon, J.-L., A. Venetianer, and J. M. Sala-Trepat.** 1987. Specific sets of DNase I-hypersensitive sites are associated with the potential and overt expression of rat albumin and α-fetal-protein genes. Proc. Natl. Acad. Sci. USA **84**:2135–2139.
27. **Palen, T. E., and T. R. Cech.** 1984. Chromatin structure at the

replication origin and transcription initiation regions of the ribosomal genes of *Tetrahymena*. Cell **36**:933–942.

28. **Parslow, T. G., and D. K. Granner.** 1982. Chromatin changes accompany immunoglobulin κ gene activation: a potential control region within the gene. Nature (London) **299**:449–451.
29. **Pederson, D. S., F. Thoma, and R. T. Simpson.** 1986. Core particle, fiber and transcriptionally active chromatin structure. Annu. Rev. Cell Biol. **2**:117–147.
30. **Petersen, D. D., M. A. Magnuson, and D. K. Granner.** 1988. Location and characterization of two widely separated glucocorticoid response elements in the phosphoenolpyruvate carboxykinase gene. Mol. Cell. Biol. **8**:96–104.
31. **Piette, I., M.-H. Kryszke, and M. Yanif.** 1985. Specific interaction of cellular factors with the B enhancer of polyoma virus. EMBO J. **4**:2675–2685.
32. **Plesko, M. M., J. L. Hargrove, D. K. Granner, and R. Chalkley.** 1983. Inhibition by sodium butyrate of enzyme induction by glucocorticoids and dibutyryl cyclic AMP: a role for the rapid form of histone acetylation. J. Biol. Chem. **258**:13738–13744.
33. **Quinn, P. G., T. W. Wong, M. A. Magnuson, J. B. Shabb, and D. K. Granner.** 1988. Identification of basal and cyclic AMP regulatory elements in the promoter of the phosphoenolpyruvate carboxykinase gene. Mol. Cell. Biol. **8**:3467–3475.
34. **Rasmussen, J. A., and E. Gilboa.** 1987. Significance of DNase I-hypersensitive sites in the long terminal repeats of a Moloney murine leukemia virus vector. J. Virol. **61**:1368–1374.
35. **Richard-Foy, H., and G. L. Hager.** 1987. Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. EMBO J. **6**:2321–2328.
36. **Rocha, E., J. R. Davie, K. E. van Holde, and H. Weintraub.** 1984. Differential salt fractionation of active and inactive genomic domains in chicken erythrocyte. J. Biol. Chem. **259**: 8558–8563.
37. **Rose, S. M., and W. Garrard.** 1984. Differentiation-dependent chromatin alterations precede transcription of immunoglobulin light chain genes. J. Biol. Chem. **259**:8534–8544.
38. **Sasaki, K., T. P. Cripe, S. R. Koch, T. L. Andreone, D. D. Petersen, E. G. Beale, and D. K. Granner.** 1984. Multihormonal regulation of phosphoenolpyruvate carboxykinase gene transcription: the dominant role of insulin. J. Biol. Chem. **259**: 15242–15251.
39. **Sasaki, K., and D. K. Granner.** 1988. Regulation of phosphoenolpyruvate carboxykinase gene transcription by insulin and cAMP: reciprocal actions on initiation and elongation. Proc. Natl. Acad. Sci. USA **85**:2954–2958.
40. **Sun, Y. L., Y. Z. Xu, M. Bellard, and P. Chambon.** 1986. Digestion of the chicken β-globin gene chromatin with micrococcal nuclease reveals the presence of an altered nucleosomal array characterized by an atypical ladder of DNA fragments. EMBO J. **5**:293–300.
41. **Thoma, F.** 1986. Protein-DNA interaction and nuclease sensitive regions determine nucleosome positions on yeast plasmid chromatin. J. Mol. Biol. **190**:177–190.
42. **Villeponteau, B., M. Lundell, and H. Martinson.** 1984. Torsional stress promotes the DNase I sensitivity of active genes. Cell **39**:469–478.
43. **Weintraub, H., and M. Groudine.** 1976. Chromosomal subunits in active genes have an altered conformation. Science **193**: 848–858.
44. **Wynshaw-Boris, A., T. G. Lugo, J. M. Short, R. E. K. Fournier, and R. W. Hanson.** 1984. Identification of a cAMP regulatory region in the gene for rat cytosolic phosphoenolpyruvate carboxykinase: use of chimeric genes transfected into hepatoma cells. J. Biol. Chem. **259**:12161–12169.
45. **Zaret, K. S., and K. R. Yamamoto.** 1984. Reversible and persistent changes in chromatin structure accompany activation of a glucocorticoid-dependent enhancer element. Cell **38**:29–38.