Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 7167–7171, October 1986
Biochemistry

Genomic sequencing reveals a positive correlation between the kinetics of strand-specific DNA demethylation of the overlapping estradiol/glucocorticoid–receptor binding sites and the rate of avian vitellogenin mRNA synthesis

(hemimethylation/gene regulation/5-methylcytosine/mechanism of demethylation/memory effect)

H. P. SALUZ, J. JIRICNY, AND J. P. JOST

Friedrich Miescher Institut, P.O. Box 2543, CH-4002 Basel, Switzerland

Communicated by Diter von Wettstein, June 9, 1986

ABSTRACT **Genomic sequencing was used to study the extent of cytosine methylation of four CpG sites within the regulatory region of the estradiol-inducible avian vitellogenin II gene. Three of these sites, two of which lie within the estradiol-receptor binding site and one in a short stretch of alternating purines and pyrimidines, were initially fully methylated. Analysis of DNA isolated from liver nuclei revealed that hormone treatment of immature White Leghorn roosters resulted in a demethylation of these sites, which occurred initially in only one DNA strand. This demethylation correlated well with the induction of vitellogenin mRNA synthesis. The demethylation of the complementary DNA strand lagged ≈24 hr behind. The fourth CpG, located within an overlapping glucocorticoid-receptor binding site, was already hemimethylated at the onset of the experiment. The demethylation of this site also occurred with kinetics similar to the rate of vitellogenin mRNA synthesis. All four CpGs remained demethylated even after cessation of gene transcription. A comparison of the methylation state of these four sites in DNA from different tissues demonstrated a clear dependence of the demethylation on estradiol. Our results suggest that this hormone-dependent event occurs via an active pathway through excision repair and/or enzymatic demethylation.**

The presence of methylated cytosine residues within the regulatory regions of a number of viral and eukaryotic genes has been shown to result in reduced levels of transcription of these genes (1–13). The extent of cytosine methylation was usually determined with the aid of methylation-sensitive restriction endonucleases such as *Hpa* II, *Msp* I, *Sal* I, and *Xho* I (1–13). These enzymes are, however, capable of detecting only a small fraction of all methylation sites—i.e., only those CpGs that lie within their respective recognition sequences. The enzymes are, moreover, unable to distinguish between symmetrically methylated and hemimethylated sites (14).

Direct genomic sequencing, first described by Church and Gilbert (15) and further developed in our laboratory (16), overcomes the disadvantages inherent to the use of restriction enzymes. We have used direct genomic sequencing to study the methylation pattern of the four CpGs situated 500–600 base pairs upstream of the cap sequence of the avian vitellogenin II gene. DNA binding experiments and computer analysis of this region revealed the presence of a glucocorticoid–receptor binding site that overlaps an estradiol–receptor binding site (Fig. 2).

One of these CpGs, part of a *Hpa* II recognition sequence within the estradiol–receptor binding site (17), was shown to become demethylated in the liver of immature chickens after induction of vitellogenin gene transcription with the hormone (18–20). Following this primary induction, the gene was able to respond to a secondary stimulation without the usual lag period between hormone treatment and onset of mRNA synthesis (21–25). This "memory effect" could be the direct result of a persistent demethylation of the *Hpa* II site, as well as of other CpGs that lie within the regulatory region of this gene. Here we report that the three remaining CpGs in this region were also demethylated after hormone induction. We furthermore show that this demethylation is strand-specific and that its kinetics run parallel to the rate of vitellogenin mRNA synthesis.

MATERIALS AND METHODS

All restriction enzymes were obtained from P-L Biochemicals. DNA polymerase I (Klenow fragment) was purchased from Boehringer Mannheim, and estradiol was from Serva (Heidelberg). 3′-Deoxyadenosine 5′-[α-^{32}P]triphosphate (triethylammonium salt, 3000 Ci/mmol; 1 Ci = 37 GBq), and 2′-deoxy-[5-^{3}H]uridine 5′-triphosphate (sodium salt, 10 Ci/mmol) were purchased from Amersham. GeneScreen (NEF-872) membranes were from New England Nuclear.

Hormone Treatment of Animals. Immature White Leghorn roosters (200–300 g) were used in all kinetic studies. Adult roosters and egg-laying Leghorn hens were also used in organ-specific demethylation studies.

Estradiol (40 mg/ml of propylene glycol) was injected i.m. into leg muscle (1 ml/kg body weight).

Preparation of Nuclei and DNA Isolation. Liver and oviduct nuclei were prepared according to Panyim *et al.* (26) and Schweizer *et al.* (27), respectively. Erythrocyte nuclei were isolated according to McGhee *et al.* (28). High molecular weight DNA was prepared as described previously (29).

***In Vitro* Transcription.** Total RNA from purified liver nuclei was synthesized *in vitro* by the method of Jost *et al.* (30). The vitellogenin mRNA was identified by hybridization on Millipore membranes as described previously (31). Fragments A and B of plasmid pVT598 (17), containing 2 kilobases of the vitellogenin coding sequence, were used as labeled probes. The extent of DNA synthesis in the nuclei was determined *in vitro* by the incorporation of labeled dTTP (20-min pulses) into trichloracetic acid-precipitated products.

Genomic Sequencing. The modified procedure of Church and Gilbert was used (15, 16); unlabeled DNA fragments from a complete *Hin*fI restriction digest of the entire genome were subjected to a partial chemical cleavage (G, G+A, T+C, C, T) and separated by size on denaturing gels. The DNA was electrotransferred and UV-crosslinked to GeneScreen membranes. The sequence ladder was visualized by indirect end labeling (hybridization) of the sequence with a strand-specific

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

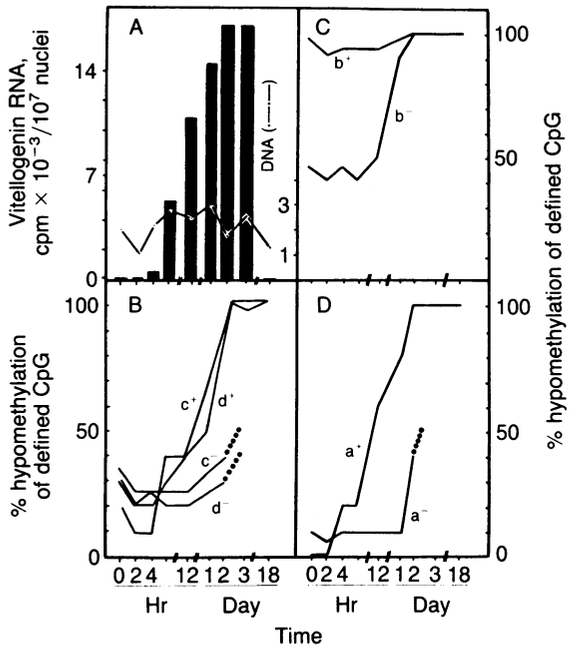

FIG. 1. Kinetics of transcription and demethylation of the avian vitellogenin II gene. (*A*) The relative rate of vitellogenin mRNA transcription in isolated liver nuclei (solid bars) and the relative rate of incorporation of ^{3}H-dTTP into DNA (also determined as cpm × $10^{-3}/10^7$ nuclei) of the same nuclei (broken line). (*B*) The kinetics of the relative amount of demethylation (%) of the CpGs c and d situated on the estradiol–receptor binding site. The + and − signs represent the upper and lower DNA strands respectively. (*C*) The kinetics of demethylation of CpG b, situated within the glucocorticoid–receptor binding site. The orientation of this binding site is opposite to that of the estradiol–receptor binding site (Fig. 2). The "upper" strand of this binding site is thus the lower DNA strand of the estradiol–receptor binding site; it therefore carries the − sign. (*D*) The kinetics of demethylation of CpG a, present on the short stretch of alternating purine–pyrimidine bases (Fig. 2).

probe; the 121-base-pair vitellogenin *Hin*fI–*Alu* I fragment cloned into the vector M13mp8 was used as a template in the synthesis of the single-stranded upper-strand probe, and the same fragment cloned into the vector M13mp9 was used as the lower-strand hybridization probe template. The specific activity of the incorporated ^{32}P in these probes was $\approx 10^9$ cpm/μg of DNA. The same filters were used for the consecutive hybridizations with these strand-specific probes. The relative extent of CpG demethylation was determined by densitometry of the autoradiograms [Shimadzu (Kyoto, Japan) Cs-930 densitometer] and by trend analysis. For calculation of the relative amount of demethylation as expressed in Fig. 1, the density of the bands above and below the tested cytosine in a given lane of Fig. 3 was taken as the 100% reference.

RESULTS

Correlation Between the Kinetics of Vitellogenin mRNA Synthesis and the Strand-Specific DNA Demethylation of the Estradiol- and Glucocorticoid–Receptor Binding Sites. Vitellogenin mRNA synthesis, as measured by an *in vitro* "run on" transcription assay (30), began ≈4 hr after the injection of immature roosters with estradiol, reaching a maximum 2 days later. No transcription was detected at day 18. A very slight but reproducible increase in DNA synthesis was also observed during the experiment (Fig. 1*A*).

The same batch of nuclei was used to study the extent of methylation of the four CpGs within the regulatory region of this gene. Genomic sequencing revealed that CpGs a, c, and d (Figs. 1 and 2) became demethylated in the upper strand with kinetics parallel to those of the vitellogenin mRNA synthesis. The demethylation of the same CpGs on the lower strand began approximately 24 hr later.

Prior to estradiol treatment, CpG b (Figs. 1 and 2) within the glucocorticoid–receptor binding site was found to be fully demethylated in the upper and approximately 50% methylated in the lower strand (Fig. 1*C*; Fig. 3, *Lower*). Both strands were fully demethylated 24 hr after the hormone treatment.

The Demethylation of the Four CpGs in the DNA of Different Hen and Rooster Organs Correlates with the Presence of Steroid–Receptor Complexes. Using *Hpa* II restriction analysis Wilks *et al.* (18, 19) demonstrated that in adult rooster liver and erythrocyte cells, CpG c (Fig. 2) was fully methylated. Our results (Fig. 4) show that CpGs a and d, as well as c, are also fully methylated in these tissues. By contrast, the same CpGs in the DNA of liver and oviduct of egg-laying hens were fully demethylated in both strands. This state of methylation correlates well with the presence of estradiol–receptor complexes in the latter tissues.

CpG b which lies in the glucocorticoid–receptor binding site exhibited a much weaker tissue-specific effect (Fig. 4). This result was not unexpected as all the tissues tested carry glucocorticoid receptors.

Thymine-Specific Sequencing Reactions. In DNA sequencing by the chemical modification method (15), 5-methylcy-

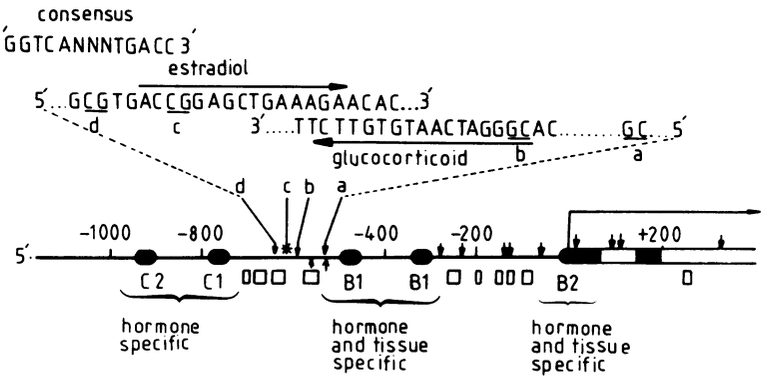

FIG. 2. Summary of the organization of the 5′ end of the chicken vitellogenin II gene. The CpGs a, b, c, and d are located in the alternating purine–pyrimidine stretch $^{-531}$(ATGTG*CG*TTGGTGCACATATG)$^{-511}$ and within the overlapping glucocorticoid and estradiol–receptor binding sites, respectively. The dyad symmetry of the consensus sequence (32) is part of the estradiol–receptor binding site. Open boxes situated between the DNase I-hypersensitive sites *B*1–*B*2 and *B*1–*C*1 (35) represent (A+T)-rich sequences (19–34 nucleotides long with 78–91% A+T). All CpGs on the sequence are marked by vertical arrows pointing downwards, and the asterisk is at the *Msp* I restriction site. Two small arrows pointing upwards indicate the two stretches of alternating purine–pyrimidine bases (33). Solid bars represent the first and second exon, and the adjacent open bars represent the introns of the vitellogenin II gene.

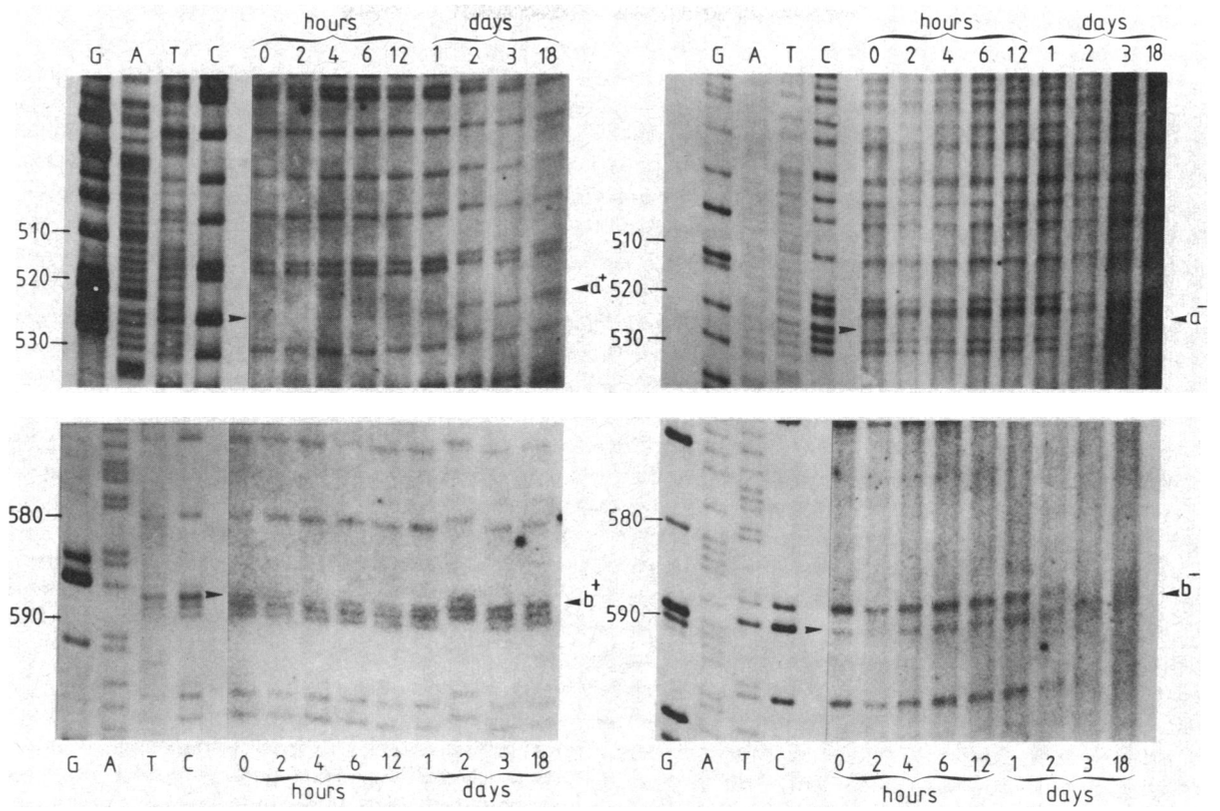

FIG. 3. Examples of demethylation kinetics of CpGs a and b on upper and lower DNA strands. CpG a (*Top*) is demethylated with kinetics similar to that of c and d (Fig. 1). The CpG b (*Bottom*) is already hemimethylated in the control untreated animals. The + and − signs denote upper and lower DNA strand, respectively. Autoradiograms were scanned as described in *Materials and Methods*, and percentage of demethylation is plotted in Fig. 1. All lanes in these kinetic experiments represent the "C," cytosine-specific, reaction made from total liver DNA.

tosine can be distinguished from cytosine by its lack of reaction with hydrazine. This results in the disappearance of a band in the C-specific sequencing lane. To unambiguously ascertain that the absence of a band in the sequence actually represents 5-methylcytosine and not thymine (which could arise as a result of 5-methylcytosine deamination), we carried out thymine-specific sequencing reactions as described by Rubin and Schmid (34). The results (Fig. 5) indicate that none of the methylation sites under investigation contained thymine instead of cytosine or 5-methylcytosine.

DISCUSSION

Previous results from our laboratory established a definite connection between hormone induction of gene transcription and DNA methylation. A *Hpa* II site, upstream of the avian vitellogenin II gene, became demethylated following the treatment of immature chickens with estradiol (18, 19). Genomic sequencing has enabled us to study this phenomenon in unprecedented detail. As this method is not dependent on the availability of suitable restriction sites (i.e., *Hpa* II), it is possible to analyze the state of methylation at all four CpGs lying within the regulatory region of this gene. Furthermore, we were able to follow the progress of the demethylation in the two strands separately. Our kinetic studies with hormone-induced roosters demonstrated a very close correlation between the synthesis of the vitellogenin mRNA and the demethylation of all four CpGs in only one of the DNA strands in the liver. This hemimethylated state, undetectable by *Hpa* II restriction analysis of CpG c (18, 19), persisted up to the time of maximal mRNA synthesis. The demethylation of the complementary strand commenced only at this time. Interestingly, the kinetics of demethylation of CpGs a^+, b^-, c^+, and d^+ correlate also with the appearance of DNase I-hypersensitive sites B1 and C1 (Fig. 2) (35, 36).

Recently Keshel *et al.* (37) described experiments in which *in vitro* methylated genes transfected into mouse L cells contained no DNase I-hypersensitive sites, while the same unmethylated genes transfected into L cells did. These findings strongly suggest that a change in the methylation pattern may alter the chromatin structure and thus affect gene expression. It has already been shown that methyl groups can influence the affinity of specific proteins at their recognition sites as in the case of *Escherichia coli lac* repressor (38) and some restriction enzymes (*Hpa* II, *Msp* I, etc.).

The demethylation kinetics, shown in Fig. 1, are the results of studies carried out on rooster liver cells that possess estradiol receptors but normally are not exposed to the hormone. The fact that three of the four CpGs in the region under study were found to be fully methylated before estradiol treatment could mean that *both*, estradiol receptor and estradiol, are required for this demethylation to take place. We therefore extended the experiment to also include rooster and hen erythrocytes and hen oviduct and liver tissues. The CpGs under investigation in both hen and rooster erythrocytes were fully methylated. This would be expected, as this tissue lacks estradiol receptors. The liver and oviduct of hens, both of which have estradiol receptors and hormone exposure, contained only unmethylated CpGs in this region. It is therefore conceivable that the binding of an activated estradiol-receptor complex to DNA is one of the prerequisites for demethylation. This proposal is substantiated by the fact that CpG b, which lies within a glucocorticoid–receptor binding site, was found to be undermethylated in all the tissues studied. An explanation of this effect would be that both glucocorticoid receptors and the hormone are present in all these tissues, suggesting that the binding of these complexes not only initiates the demethylation but also helps maintain it.

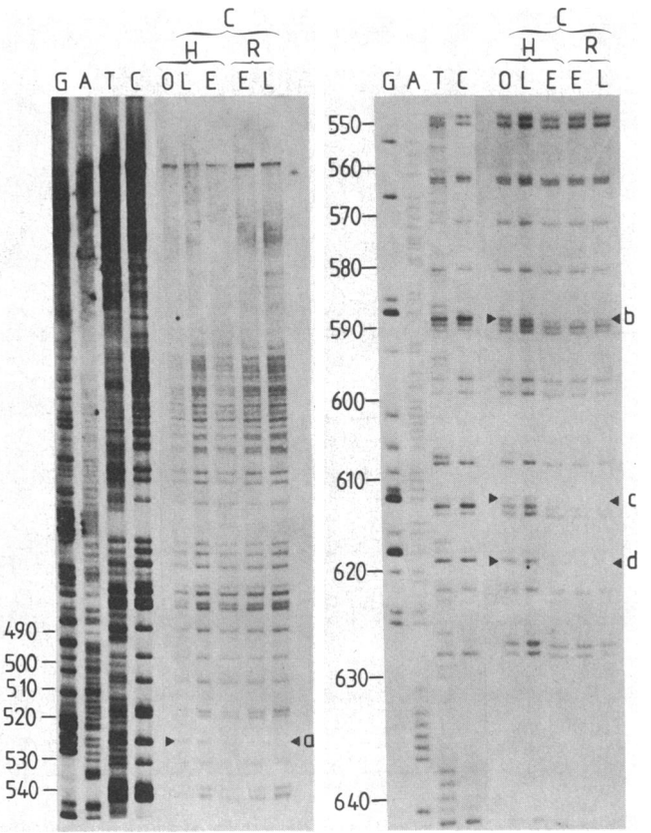

FIG. 4. Hormone dependence of the demethylation of CpGs a–d. This figure shows the cytosine (C) methylation state of CpGs a, b, c, and d (Fig. 2) present in the upper DNA strand of oviduct (O), liver (L), and erythrocytes (E) of egg-laying hens (H); and the liver (L) and erythrocytes (E) of mature roosters (R).

To discount the possibility that the lack of bands in the C-reactions of DNA of different tissues originated from developmentally fixed sequence changes, some of which could arise by the deamination of 5-methylcytosine to thymine, T-specific reactions were also done. In all cases, two of which are shown in Fig. 5, no such base changes were detected.

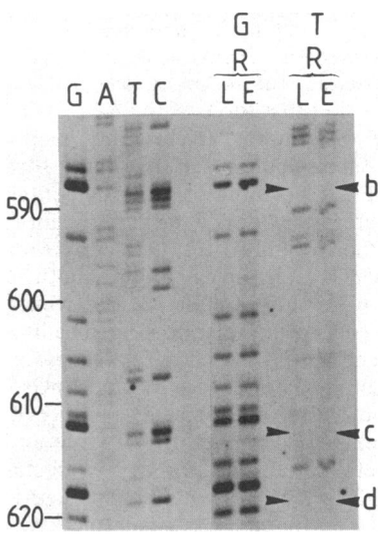

FIG. 5. Comparison of G- and T-specific reactions (G, T) of CpGs, b, c, and d on the upper DNA strand in liver (L) and erythrocytes (E) of roosters (R). No base transitions such as deamination of 5-methylcytosine to thymidine could be detected.

Our experiments provide us with several valuable insights into the mechanism of hormone-dependent demethylation. Should the hormone–receptor complex bind to both (Fig. 6*A*) or only one DNA strand (Fig. 6*B*), inhibiting thus the repair methylase activity, 4 and 7 cycles of DNA replication, respectively, would be required to yield 94% demethylated duplex DNA in both cases. In the former case, an excess of the hormone–receptor complex would also be required. Since the total DNA mass increases by no more than 17% during this period (39), this pathway may be excluded. Work from this laboratory has also shown that inhibition of total DNA synthesis by hydroxyurea and cytosine arabinoside impaired neither vitellogenin mRNA synthesis nor the demethylation of the *Hpa* II site (19). Limited DNA replication, such as gene amplification that might account for this demethylation, could also be excluded by gene titration experiments before and after estradiol treatment (ref. 40; H. P. Saluz, unpublished results).

Two plausible demethylation mechanisms remain: active demethylation by excision repair and/or by the activity of a demethylase (41). An enzymatic removal of a methyl group from 5-methylcytosine in DNA is mechanistically improbable, because the transfer of the methyl group of *S*-adenosyl-L-methionine to a carbon atom is irreversible at neutral pH (42). However, Razin *et al.* (43) have recently demonstrated that in differentiating Friend erythroleukemia cells a replacement of 5-methylcytosine with cytosine occurs, presumably via an enzymatic pathway. In our case, preliminary results (H. P. Saluz, unpublished observations) reveal the presence of strand-specific DNA cuts in the region of the CpGs studied. This evidence may lend support to the excision repair pathway.

It remains now to establish the cause–effect relationship between these observations and the mechanism of the hormone-dependent CpG demethylation. The CpG b, situated within the glucocorticoid–receptor binding site, is hemimethylated in all organs studied (Figs. 3 and 4), and it becomes fully demethylated in the liver following estradiol treatment. The existence of a hemimethylated site in DNA is surprising, as it has generally been accepted that all hemimethylated CpGs are efficiently methylated by a repair methylase (7, 12). In the present case it is possible that the glucocorticoid–receptor complex has a higher affinity for its binding site than the repair methylase, thus preventing the enzymatic methylation. Since it is known that glucocorticoids play an important

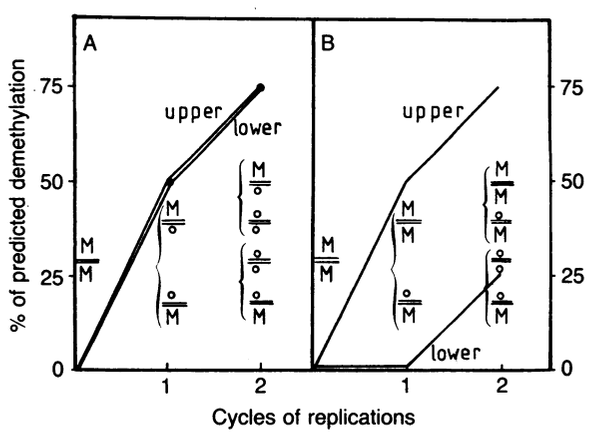

FIG. 6. Predicted pattern of demethylation of the upper and lower DNA strand based on the hypothesis of repair methylase inhibition following DNA replication. Protein inhibition of the repair methylase by (*A*) binding to both DNA strands at the same time and (*B*) binding only the upper strand. M represents 5-methylcytosine, ○ unmethylated cytosine.

role in eukaryotic cell metabolism in general and in liver in particular (44), a gradual demethylation of the CpG b may take place during development. A similar effect of glucocorticoids on the demethylation of mouse mammary tumor virus DNA has also been reported in some mutants of myeloma cells (45). Whether glucocorticoids and estradiol trigger the same mechanism of demethylation remains unknown.

We thank Drs. K. Wiebauer, P. Kuenzler, and B. Rudkin for critical reading of the manuscript and Mr. M. Ackerman for his expert photographic assistance.

1. Holliday, R. & Pugh, J. E. (1975) *Science* **187,** 226–232.
2. Sager, R. & Kitchin, R. (1975) *Science* **189,** 426–433.
3. Razin, A. & Riggs, A. D. (1980) *Science* **210,** 604–610.
4. Razin, A. & Friedman, J. (1981) *Prog. Nucleic Acid Res. Mol. Biol.* **25,** 33–52.
5. Ehrlich, M. & Wang, R. Y. H. (1981) *Science* **212,** 1350–1357.
6. Felsenfeld, G. & McGhee, J. (1982) *Nature (London)* **296,** 602–603.
7. Adams, R. L. P. & Burdon, R. H. (1985) in *Molecular Biology of DNA Methylation,* ed. Rich, A. (Springer, New York).
8. Doerfler, W. (1983) *Annu. Rev. Biochem.* **52,** 93–124.
9. Bird, A. P. (1984) *Curr. Top. Microbiol. Immunol.* **108,** 129–141.
10. Jaenisch, R. & Jaehner, D. (1984) *Biochim. Biophys. Acta* **782,** 1–9.
11. Taylor, J. H. (1984) *DNA Methylation and Cellular Differentiation* (Springer, New York).
12. Razin, A., Cedar, H. & Riggs, A. D. (1984) *DNA Methylation: Biochemistry and Biological Significance* (Springer, New York).
13. Meselson, M. & Yuan, R. (1968) *Nature (London)* **217,** 1110–1114.
14. Gruenbaum, Y., Cedar, H. & Razin, A. (1981) *Nucleic Acids Res.* **9,** 2509–2515.
15. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
16. Saluz, H. P. & Jost, J. P. (1986) *Gene* **42,** 151–157.
17. Jost, J. P., Seldran, M. & Geiser, M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 429–433.
18. Wilks, A. F., Cozens, P. J., Mattaj, I. W. & Jost, J. P. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 4252–4255.
19. Wilks, A. F., Seldran, M. & Jost, J. P. (1984) *Nucleic Acids Res.* **12,** 1163–1177.
20. Meijlink, F. C. P. W., Philipsen, N. J., Gruber, M. & Ab, G. (1983) *Nucleic Acids Res.* **11,** 1361–1373.
21. Jost, J. P., Ohno, T., Panyim, S. & Schuerch, A. R. (1978) *Eur. J. Biochem.* **84,** 355–361.
22. Deeley, R. G., Udell, D. S., Burns, A. T. H., Gordon, J. I. & Goldberger, R. F. (1977) *J. Biol. Chem.* **252,** 7913–7915.
23. Ryffel, G. U., Wahli, W. & Weber, R. (1977) *Cell* **11,** 213–221.
24. Searle, P. F. & Tata, J. R. (1981) *Cell* **23,** 741–746.
25. Hayward, M. A., Brock, M. L. & Shapiro, D. J. (1982) *Nucleic Acids Res.* **10,** 8273–8284.
26. Panyim, S., Ohno, T. & Jost, J. P. (1978) *Nucleic Acids Res.* **5,** 1353–1370.
27. Schweizer, G., Cadepond-Vincent, F. & Baulieu, E. E. (1985) *Biochemistry* **24,** 1742–1749.
28. McGhee, J. D., Wood, W. I., Dolan, M., Engel, J. D. & Felsenfeld, G. (1981) *Cell* **27,** 45–55.
29. Wilks, A., Cato, A. C. B., Cozens, P. J., Mattaj, I. W. & Jost, J. P. (1981) *Gene* **16,** 249–259.
30. Jost, J. P., Moncharmont, B., Jiricny, J., Saluz, H. P. & Hertner, T. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 43–47.
31. Jost, J. P., Geiser, M. & Seldran, M. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 988–991.
32. Walker, P., Germond, J. E., Brown-Luedi, M., Givel, F. & Wahli, W. (1984) *Nucleic Acids Res.* **12,** 8611–8626.
33. Geiser, M., Mattaj, I. W., Wilks, A. F., Seldran, M. & Jost, J. P. (1983) *J. Biol. Chem.* **258,** 9024–9030.
34. Rubin, C. M. & Schmid, C. W. (1980) *Nucleic Acids Res.* **8,** 4613–4619.
35. Burch, J. B. E. & Weintraub, H. (1983) *Cell* **33,** 65–76.
36. Burch, J. B. E. (1984) *Nucleic Acids Res.* **12,** 1117–1135.
37. Keshel, I., Lieman-Hurwitz, J. & Cedar, H. (1986) *Cell* **44,** 535–543.
38. Fisher, E. F. & Caruthers, M. H. (1979) *Nucleic Acids Res.* **7,** 401–416.
39. Jost, J. P., Keller, R. & Dierks-Ventling, C. (1973) *J. Biol. Chem.* **248,** 5262–5266.
40. Jost, J. P., Schuerch, A. R. & Walz, A. (1977) *FEBS Lett.* **75,** 133–137.
41. Gjerset, R. A. & Martin, D. W., Jr. (1982) *J. Biol. Chem.* **257,** 8581–8583.
42. Cantoni, G. L. (1960) *Comp. Biochem.* **1,** 172–241.
43. Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Scarpa, S., Carotti, D. & Cantoni, G. L. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2827–2831.
44. Hers, H. G. & Hue, L. (1983) *Annu. Rev. Biochem.* **52,** 617–653.
45. Mermod, J. J., Bourgeois, S., Defer, N. & Crepin, N. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 110–114.