Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 1302–1306, February 1995
Genetics

Sexually dimorphic DNA demethylation in the promoter of the *Slp* (sex-limited protein) gene in mouse liver

NORIHIKO YOKOMORI, RICKIE MOORE, AND MASAHIKO NEGISHI

Pharmacogenetics Section, Laboratory of Reproductive and Developmental Toxicology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709

Communicated by Ray D. Owen, California Institute of Technology, Pasadena, CA, October 19, 1994 (received for review August 19, 1994)*

ABSTRACT **Mouse *Slp*, a duplicate of the fourth complement component (*C4*) gene, exhibits EDTA-independent complement activity with a hepatic expression that is male specific. To provide an underlying mechanism for the male-specific expression, we have analyzed the promoter activity of the various 5′-flanking sequences and CpG demethylation of the *Slp* gene. Transient transfections using HepG2 cells indicate that the element TTCCGGGC (nt −124 to −117) regulates the promoter activity. Moreover, CpG at position −121 of this regulatory element is demethylated to a much higher degree in males than in females. This sexually dimorphic DNA demethylation is consistent with the male-specific expression of the *Slp* gene in DBA/2 males. The regulatory element binds to the different TTCCGGGC-specific nuclear proteins depending on the methylation of the CpG site. In contrast, the corresponding CpG at position −119 of the *C4* gene, which is expressed in both males and females, is demethylated at equal and high levels in both sexes. We therefore propose that the DNA demethylation and methylation-sensitive transcription factors may be a part of the regulatory mechanism for the male-specific expression of the *Slp* gene.**

DNA methylation is known to regulate cellular physiology by altering gene expression. It is also programmed in the growth and developmental processes (1–4). In general, methylation is inversely associated with gene expression. This report examines the possibility that sexually differentiated expression of the *Slp* gene is also regulated by DNA methylation.

Slp is encoded by the *S* region of the major histocompatibility complex and mediates EDTA-independent complement activity (5–8), which may play a role in the mobilization of additional polymorphonuclear leukocytes to inflammation sites (see, however, ref. 9, which suggests that Slp has no biological function). Despite the high identity (95%) of the nucleotide sequences, the mouse *Slp* and *C4* genes display distinct patterns of expression. The former gene is expressed only in adult male mice with certain *H-2* haplotypes such as $H\text{-}2^d$ and $H\text{-}2^s$. In contrast, the latter gene is expressed in both females and males regardless of the *H-2* haplotype. It was once thought that androgen regulates the male-specific expression of the *Slp* gene. Moreover, a putative androgen-response element is located within a 160-bp fragment 2 kbp upstream of the *Slp* gene (10, 11). These previous studies, however, are contradicted by a recent finding that it is growth hormone that directly regulates the expression (12). Therefore, a role of this element in male-specific transcription remains to be further evaluated. The *Slp* gene, therefore, needs to be studied from other possible directions to associate its specific expression with a cis-acting element.

In a manner similar to *Slp* expression, hepatic steroid 16α-hydroxylase cytochrome P450 2d-9 is also expressed specifically in the adult male mice and is regulated by growth hormone at the transcriptional level of the gene *Cyp2d-9* (13–17). A regulatory element of the *Cyp2d-9* promoter has been characterized by a phylogenetic comparison with the other members within the gene subfamily (18). Furthermore, we have recently found that the levels of CpG methylation in this element are inversely associated with the male-specific expression of the *Cyp2d-9* gene. In this paper, we compare the promoter sequences of the two male-specific genes so as to identify the conserved regulatory element. We then determine the promoter activity and the CpG methylation of the element. DNA methylation as a general mechanism for sex-specific gene expression will be discussed.

EXPERIMENTAL PROCEDURES

Construction of Plasmids. To construct a series of pSlp-CAT plasmids, we amplified various portions of the *Slp* gene promoter using PCR kits (Perkin–Elmer/Cetus) and specific primers based on the reported sequence of the *Slp* gene (7). These primers were also designed to create the *Hin*dIII site at the 5′ end. The 3′ primer used was 5′-GGGTCTAGAGGATCCAGGAGAGGTCAACCT, in which the *Xba* I site is underlined. In addition, the following 5′ primers were used: pSlp(−248)CAT, 5′-GGGAAGCTTGTACTGGGAATAAACAAG; pSlp(−142)CAT, 5′-GGGAAGCTTCTTGACCATCACGTGGTTTCCGGG; pSlp(−124)CAT, 5′-GGGAAGCTTTTCCGGGCTCATGGGTC; pSlp(−116)CAT, 5′-GGGAAGCTTTCATGGGGTCAAAGGGAG; pSlp(−124)/M(−121)CAT, 5′-GGGAAGCTTTTCaGGGCTCATGGGGTC. The created *Hin*dIII sites are underlined, and the mutated nucleotide at position −121 is indicated by a lowercase letter. The amplified promoter DNAs were digested by *Hin*dIII and *Xba* I, ligated into the corresponding cloning sites of the pCAT-Basic vector (Promega), and transformed into *Escherichia coli* HB101 cells. The plasmids were purified by CsCl centrifugation and verified by DNA sequencing.

Cell Culture, Transfection, and Chloramphenicol Acetyltransferase (CAT) Activity. Human hepatocellular carcinoma HepG2 cells were cultured in minimum essential medium Eagle (Sigma) supplemented with 10% (vol/vol) fetal bovine serum, penicillin (100 units/ml), and streptomycin (100 μg/ml) at 37°C. At 24 hr prior to transfection, HepG2 cells were transferred to plastic plates (60-mm diameter). Using the calcium phosphate/DNA coprecipitation method provided by the CellPhect transfection kit (Pharmacia), each pSlp-CAT DNA (2 μg) was transfected into HepG2 cells at 37°C for 12 hr. The transfected cells were then washed twice with PBS solution and cultured in the medium at 37°C for 40 hr. To measure the CAT activity, the harvested cells were lysed by freezing and thawing. Then the heat-treated cell extract (5 μg of protein) was incubated with [^{14}C]chloramphenicol, according to Gorman *et al.* (19).

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.

Abbreviation: CAT, chloramphenicol acetyltransferase.
*Communication of this paper was initiated by Donald Shreffler and, after his death (August 8, 1994), completed by Ray D. Owen.

Genes	Promoter Sequences	Sex-Specificity
C4 gene:	-132 TCACGTGGTTTCCCGGCTCA	Male and Female
Slp gene:	-133 TCACGTGGTTTCCGGGCTCA	Male
Cyp 2d-9:	-109 TCCTCCCTATTCCGGGCCAA	Male
Cyp 2d-10:	-109 TCCTCCCTACAGGTGGACAA	Male and Female

FIG. 1. Nucleotide sequence comparisons. The nucleotide sequences from the 5′-flanking regions of the *Slp* and *C4* genes (7) are compared with those from the *Cyp2d-9* and *Cyp2d-10* genes (16). Potential CpG methylation sites are underlined, and the conserved sequence is shaded.

Gel-Shift Assay. Fifty adult CD-1 mice were used to obtain nuclear extracts as described by Gorski *et al.* (20). Each oligonucleotide, synthesized with an Applied Biosystems DNA/RNA synthesizer, was annealed to its complement and labeled by using [α-^{32}P]dATP (>6000 Ci/mmol; Amersham; 1 Ci = 37 GBq) and the Klenow fragment of DNA polymerase. Methylated oligonucleotides were prepared by including 5-methylcytosine during the appropriate cycle of synthesis. Each radioactive probe ($\approx$20,000 cpm) was incubated with 5 μg of male liver nuclear proteins in 10 μl of 20 mM Tris·HCl buffer (pH 7.5) containing 1 μg of poly(dI-dC), 50 mM NaCl, 0.1 mM dithiothreitol, and 10% (vol/vol) glycerol at room temperature. An $\approx$50-fold excess of unlabeled oligonucleotide was used for competition experiments.

Sequencing of the Sodium Bisulfite-Treated Promoter. Genomic DNAs were prepared from the livers of 2-month-old CD-1 (Charles River Breeding Laboratories) and DBA/2 (The Jackson Laboratory) males and females using the SDS/proteinase K method, digested with *Pst* I, and then subjected to sequential reactions to determine CpG methylation patterns according to Frommer *et al.* (21). The oligonucleotide primers were synthesized based on the reported sequences of the *Slp* and *C4* genes (7). The top strand of promoter sequence (nt −196 to −36) of the *Slp* gene was amplified using 10 μl of the bisulfite-reacted DNA as a template and the oligonucleotides 5′-GGGTCTAGAAGATTTTGGTGTTTGGGTTGT and 5′-GGGTCTAGACTAACTATCTACTCCCTTCTA as the 5′ and 3′ primers, respectively. The underlined regions indicate that an *Xba* I site was added at each end of the amplified DNAs. The −194 to −37 region of the *C4* gene was also amplified using the same primers. Amplified DNAs were digested with *Xba* I and cloned into M13mp19 vectors for DNA sequencing. Variations of the bisulfite reaction were measured as follows: pBluescript plasmid (1 μg) was methylated *in vitro* by *Hpa* II methylase and included during the bisulfite reactions. The plasmid DNA was amplified using the oligonucleotides 5′-GGGTCTAGAAATTTAATTTTTATTAAAGGG and 5′-GGGTCTAGATATAATACAACTCACTATAA as 5′ and 3′ primers, respectively, and sequenced.

RESULTS AND DISCUSSION

Nucleotide sequences of two sets of gene pairs are compared: the *Slp* and *C4* genes, which are two different isotypes, and the *Cyp2d-9* and *Cyp2d-10* genes, which are members of the mouse cytochrome P450 2D subfamily (Fig. 1). *Cyp2d-9*, which encodes a hepatic steroid 16α-hydroxylase cytochrome P450, is selected for the comparison since it is expressed only in adult male mice (15, 16). In contrast, the *C4* and *Cyp2d-10* genes are expressed similarly in both sexes. Although *Slp* and *Cyp2d-9* are members of the different gene families and located on different chromosomes, they share the conserved sequence TTCCGGGC in their 5′-flanking regions. While this sequence is located in the middle of the strong *Cyp2d-9* promoter (18), CpG at position −97 [CpG (−97)] in the sequence is demethylated completely (93%) in adult males but only partially (48%) in adult females (unpublished results). The corresponding sequence also provides the *Slp* gene with a potential methylation site at position −121. Both sites are either eliminated by the nucleotide mutations in the *Cyp2d-10* gene or shifted downstream to position −119 in the *C4* gene. In addition to the unique CpG (−121), *Slp* contains another CpG at position −130, which lies within the region we are comparing. Since this CpG is also conserved at position −129 in the *C4* gene, it appears to be a common CpG site for the both *Slp* and *C4* genes.

It is known that the 5′-flanking regions up to positions −244 or −189 of the *Slp* gene possess the promoter activity in the transfected HepG2 cells (7, 8). Although transfection does not measure directly sex-limited promoter activity, Nonaka *et al.* (7) previously found that the *Slp* and *C4* promoters exhibited different activities. Conversely, Miyagoe *et al.* (8) have recently reported that there is no difference in promoter activity. Nevertheless, these regions contain the conserved sequence TTCCGGGC (nt −124 to −117), and the role of this sequence for the promoter activity has not been examined specifically. Therefore, successive 5′-deletion mutants of the *Slp* promoter were constructed with or without the presence of this sequence, placed in front of the CAT reporter gene, and then transfected into HepG2 cells to measure their promoter ac-

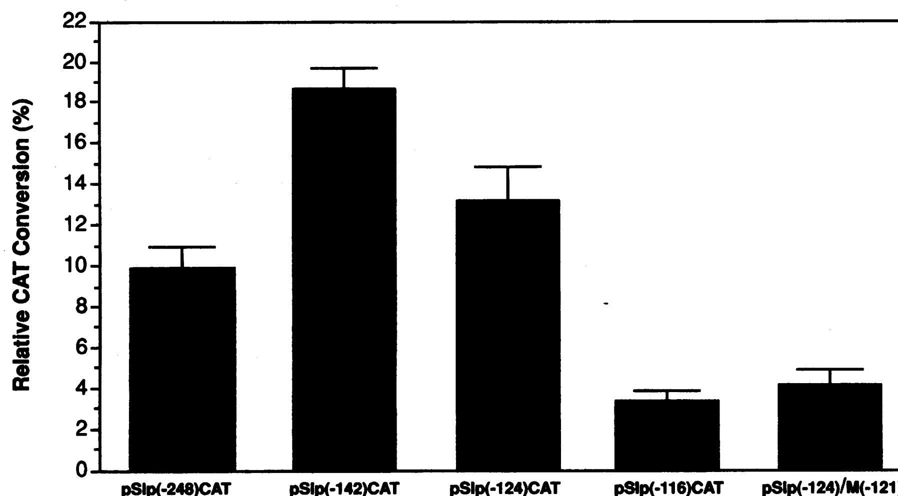

FIG. 2. Slp promoter activity in HepG2 cells. Each deletion construct was transfected into HepG2 cells, and its activity was measured as the conversion rate of chloramphenicol to its acetylated forms as described in *Experimental Procedures.* Standard errors were obtained from four independent results.

tivities (Fig. 2). pSlp(−124)CAT, which included the conserved TTCCGGGC (−124 to −117) sequence, displayed strong promoter activity. Activity was dramatically decreased by the removal of this sequence from the promoter. Moreover, the C → A mutation at position −121 decreased the activity of pSlp(−124)CAT to the low level observed with pSlp(−116)CAT. The results indicate, therefore, that TTCCGGGC (−124 to −117) is the regulatory element that plays the key role in the *Slp* promoter activity. In addition, the activity strongly depends on the type of nucleotide at position −121, the potential methylation site in the regulatory element. For technical reasons, it was not possible to directly determine whether the methylation at position −121 interfered with the promoter activity in HepG2 cells. Furthermore, HepG2 cell may not maintain an *in vivo* phenotype for sex-specific gene expression. We therefore decided to measure the *in vivo* methylation pattern of CpG (−121) in the *Slp* gene and compare it with other CpG sites in *Slp* and *C4* genes.

For that purpose, we have sequenced bisulfite-treated genomic DNA (Fig. 3) to determine the methylation levels at the unique CpG (−121) and common CpG (−130) sites of the *Slp* gene in female and male livers from DBA/2 and CD-1 mice. Slp was first reported as a variant serum protein with whose expression was associated with *H-2* locus (22). Subsequently, more extensive studies have concluded that, while the *Slp* gene is expressed only in adult male mice with certain haplotypes including $H\text{-}2^d$ and $H\text{-}2^s$, it is not expressed in either sex with other haplotypes such as $H\text{-}2^b$, $H\text{-}2^k$, and $H\text{-}2^q$ (23, 24). In contrast, the *C4* gene is expressed in both females and males regardless of the *H-2* haplotype. In DBA/2 mice carrying the $H\text{-}2^d$ haplotype, the *Slp* gene is expressed only in males (6). CpG (−121) but not CpG (−130) exhibits a demethylation

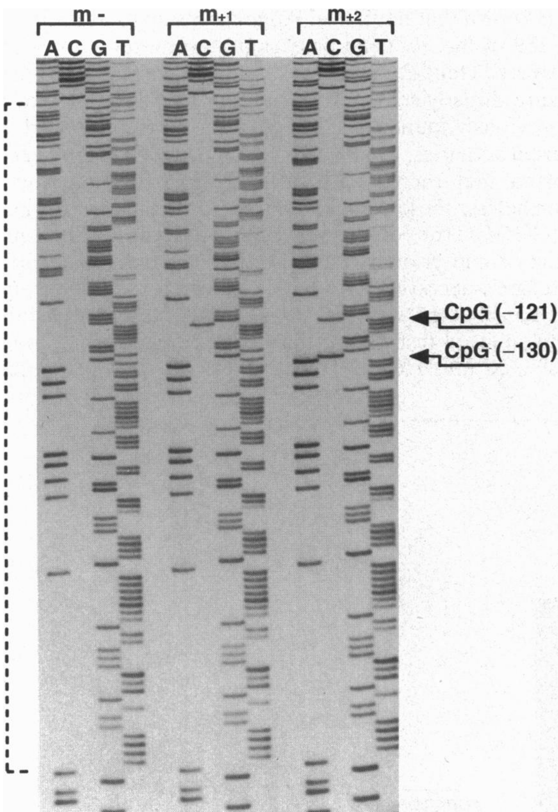

FIG. 3. Sequencing of bisulfite-treated genomic DNA. The three sequences with different methylation patterns of the *Slp* gene are shown. Genomic DNA here was prepared from DBA/2 mice. The following symbols are used: m−, not methylated; m_{+1}, methylation at position −121; m_{+2}, methylation at both positions −121 and −130. A bracket with a broken line indicates the sequences of the *Slp* gene.

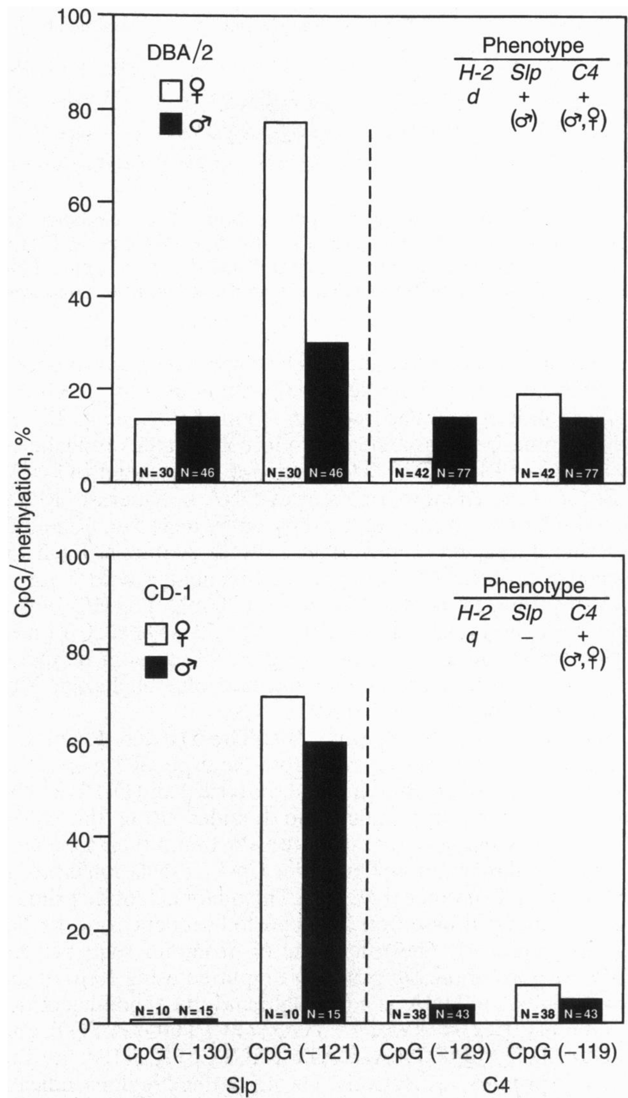

FIG. 4. Level of CpG methylation. The levels of CpG methylation are shown as percentages of the total number of sequences (indicated by N at bottom of each bar). Open and closed bars represent methylation in the female and male mice, respectively. We prepared two independent, bisulfite-treated DNAs and performed two separate amplifications from each of these two DNA samples. The values in this figure were obtained by adding together the bisulfite sequences from similar numbers from each amplified DNA template. The average error of this deamination method to measure the methylation levels was ≈10%.

pattern that agrees with the specific expression of the *Slp* gene in DBA/2 males (Fig. 4). CpG (−121) of the gene is demethylated ≈3 times greater in males than females (70% versus 23%). Consistent with the association of the high methylation level with the repression of the gene in DBA/2 mice is the observation that CpG (−121) is heavily methylated in the both sexes of CD-1 mice (70–60%). CD-1 mice carry the H^q haplotype and do not express the gene at a detectable level of *Slp* mRNA in either sex. CpG (−130) of this gene, on the other hand, is always highly demethylated regardless of the sex and strain of mouse. In addition, the *C4* gene, which is expressed in all of the mice tested, is always demethylated at the CpG (−119) as well as the CpG (−129) sites. Over 90% of the CpG sites in the *in vitro*-methylated pBluescript were found to always be methylated. Hence, CpG (−121) is the only site with methylation levels associated with the expression of the *Slp* gene in both DBA/2 and CD-1 mice. It remains, however, to

be determined whether any linkage exists between the demethylation and *H-2* haplotypes.

It is interesting that the basal methylation levels differ significantly in the two strains of mice ["basal level" is defined as the methylation of CpG (−130) in the *Slp* gene as well as at both CpG sites in the *C4* gene]. These sites are methylated at higher levels in DBA/2 than in CD-1 mice (Fig. 3). Assuming that the methylated gene is inactive, this evidence suggests that the number of genes that can potentially be activated is different in each strain of mice. For example, over 95% of the time, demethylation of CpG (−130) coincides with that of CpG (−121). As a result, the sex-specific expression of the *Slp* gene must be determined by not only the absolute level of demethylation at CpG (−121) but also relative differences in demethylation levels between sexes. Another intriguing finding is that CpG (−121) of the *Slp* gene is demethylated to some degree even in mice (DBA/2 males and both males and females of CD-1) that appear not to express the gene. This demethylation may result in low-level expression of the *Slp* gene in female mice. Unfortunately, experimental procedures may be involved in producing low levels of demethylation, since genomic DNA for our experiments was prepared from liver tissues containing other types of cells such as Kupffer and blood cells in addition to hepatocytes. Although only hepatocytes express the *Slp* gene, they account for approximately half of the genomic DNA prepared from liver tissue. Most of the demethylation observed in the DBA/2 female and CD-1 male and female mice may be accounted for by the nonhepatic cells.

An oligonucleotide $(TTCCGGGC)_2$ was prepared with and without CpG methylation from which a gel-shift assay was performed (Fig. 5*A*). The results indicate the presence of different nuclear proteins that specifically bind to either the unmethylated or the methylated probes. Since an Sp1 consensus does not compete, these proteins are not Sp1 or Sp1-related proteins. It remains, however, to be tested whether the proteins represent the known methylation-sensitive transcription factors and methyl-CpG-binding proteins such as MeCp1 (1). Among the organs tested, both the methylation-sensitive and the methyl-CpG-binding proteins are present in liver, brain, and heart but not in kidney, spleen, and testis (data not shown). This tissue-dependent coexpression implies that these two proteins are a pair of factors that regulate a gene through the regulatory element TTCCGGGC, most likely as a part of a large regulatory complex. Might the *C4* gene be regulated by binding of the same proteins to the corresponding sequence TTCCCGGC (−123 to −116)? A gel-shift binding is observed with the methylated $(TTCCm^5CGGC)_2$ but not with the unmethylated $(TTCCCGGC)_2$ probes (Fig. 5*B*). Additionally, the methylated probe of the *C4* gene binds to a different protein from that binding to $(TTCm^5CGGGC)_2$ of the *Slp* gene. This protein, which binds to $(TTCCm^5CGGC)_2$, is ubiquitously distributed in all organs tested, including liver, kidney, spleen, heart, brain, and testis (data not shown). Hence, the *Slp* and *C4* genes appear to be regulated by different nuclear proteins, although they are homologous genes sharing high nucleotide identity and similar G+C-rich sequences.

Disruption of the mouse DNA methyltransferase gene shows that DNA methylation is programmed in the developmental processes (1–4). For instance, cell transformation results in a high degree of DNA methylation (25), whereas demethylation of the *myoD* or another regulatory gene leads to myogenesis (26). The parental imprinting of the *H-19* and *Igf-2* receptor genes is regulated by allele-specific methylation; the promoters are methylated in the inactive paternal genes and demethylated in the active maternal copies (27). Our present studies show that the regulatory element of the *Slp* gene undergoes sex-dependent DNA demethylation in accordance with the expression in male mice and binds to the specific factors depending on the methylation patterns. In addition to these previously reported regulation mechanisms,

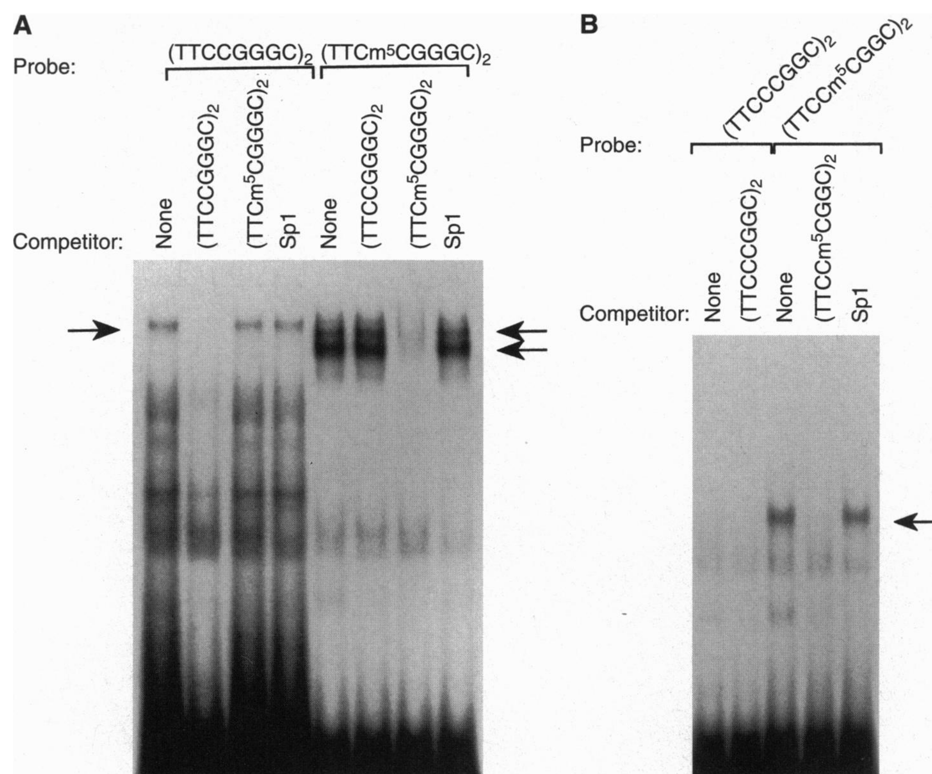

FIG. 5. Gel-shift assay. Nuclear extracts used for this experiment were prepared from CD-1 male mice. The probes used were $(TTCCGGGC)_2$ and $(TTCm^5CGGGC)_2$ from the *Slp* and *Cyp2d-9* genes (*A*) and $(TTCCCGGC)_2$ and $(TTCCm^5CGGC)_2$ from the *C4* gene (*B*). An approximately 50-fold excess of unlabeled oligonucleotide was used for the competition experiment. The binding complex formed with $(TTCCm^5CGGC)_2$ migrated faster than that with $(TTCm^5CGGGC)_2$. In addition, no cross competition was observed. Arrows indicate the binding complexes.

the sex-specific gene expression can also be regulated by DNA methylation. Until direct evidence indicating that the methylation of CpG site inhibits gene transcription, other possibilities remain to be excluded. The demethylation, for example, can indirectly affect the gene transcription by altering chromatin structure, or it may be a secondary phenomenon caused by another sex-dependent modification occurring at a topologically different site of the *Slp* gene.

In conclusion, the male-specific *Slp* gene contains the methylation site at CpG (−121) within the conserved regulatory element TTCCGGGC (−124 to −117). This CpG site is preferentially demethylated in adult male mice concordantly with their male-specific expressions in DBA/2 and CD-1 mice. In addition, the element binds to the specific nuclear proteins depending on the methylation pattern of the CpG site. We conclude, therefore, that the element and additional protein factors may be a part of the general mechanism of the sex-specific gene expression including the *Slp* gene.

We especially thank Dr. Kaoru Aida for his helpful discussion throughout this research. Our sincere thanks are also due to Drs. Joyce Goldstein, Cary Weinberger, and Lee Pedersen and colleagues in the Pharmacogenetics Section for critical reading of this manuscript.

1. Tate, P. H. & Bird, A. P. (1993) *Curr. Opin. Gene Dev.* **3**, 226–231.
2. Cedar, H. (1988) *Cell* **53**, 3–4.
3. Tilghman, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8761–8762.
4. Bird, A. (1992) *Cell* **70**, 5–8.
5. Shreffler, D. C. (1982) in *Histocompatibility Antigens: Structure and Function*, eds. Parham, P. & Strominger, J. (Chapman & Hall, London), pp. 189–219.
6. Van den Berg, C. W., Demant, P., Aerts, P. C. & Van Duk, H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10711–10715.
7. Nonaka, M., Kimura, H., Yeul, Y. D., Yokoyama, S., Nakayama, K. & Takahasi, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7883–7887.
8. Miyagoe, Y., Georgatsou, E., Varin-Black, N. & Meo, T. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5786–5790.
9. Ogata, R. T. & Zepf, N. E. (1991) *J. Immunol.* **8**, 2756–2763.
10. Stavenhagen, J. B. & Robins, D. M. (1988) *Cell* **55**, 247–254.
11. Adler, A. J., Scheller, A., Hoffman, Y. & Robins, D. M. (1991) *Mol. Endocrinol.* **5**, 1587–1596.
12. Georgatsou, E., Bourgael, P. & Meo, T. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3626–3630.
13. Negishi, M., Aida, K. & Yoshioka, H. (1993) in *Cytochrome P-450*, eds. Omura, T., Ishimura, Y. & Fujii-Kuriyama, Y. (Kodansha, Tokyo, and VCH, Weinheim, Germany), pp. 230–241.
14. Noshiro, M. & Negishi, M. (1986) *J. Biol. Chem.* **261**, 15923–15927.
15. Wong, G., Kawajiri, K. & Negishi, M. (1987) *Biochemistry* **26**, 8683–8690.
16. Wong, G., Itakura, T., Kawajiri, K., Skow, L. & Negishi, M. (1989) *J. Biol. Chem.* **264**, 2920–2927.
17. Aida, K. & Negishi, M. (1993) *J. Mol. Endocrinol.* **11**, 231–240.
18. Yoshioka, H., Lang, M., Wong, G. & Negishi, M. (1991) *J. Biol. Chem.* **265**, 14612–14617.
19. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044–1051.
20. Gorski, K., Carneiro, M. & Schibler, U. (1986) *Cell* **47**, 767–776.
21. Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L. & Paul, C. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1827–1831.
22. Shreffler, D. C. & Owen, R. D. (1962) *Genetics* **48**, 9–25.
23. Shreffler, D. C., Atkinson, J. P., Chan, A. C., Karp, D. R., Killion, C. C., Ogata, R. T. & Rosa, P. A. (1984) *Philos. Trans. R. Soc. London B* **306**, 395–403.
24. Altema, P. L. & Katz, D. D. (1933) *Inbred and Genetically Defined Strains of Laboratory Animals* (Fed. Am. Soc. Exp. Biol., Bethesda), Part 1.
25. Davis, R. L., Weintraub, H. & Lassar, A. B. (1987) *Cell* **51**, 987–1000.
26. Wu, J., Issa, J.-P., Herman, J., Bassett, D. E., Jr., Nelkin, B. & Baylin, S. B. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 8891–8895.
27. Li, E., Beard, C. & Jaenisch, R. (1993) *Nature (London)* **366**, 362–365.