Journal of Cellular Biochemistry 92:99–103 (2004)

Heterologous CpG Island Becomes Extensively Methylated in the Genome of Transgenic Mice

Elena Ioudinkova,[1] Lilia Verbovaia,[1] Sergey Kadulin,[2] Igor Goldman,[2] and Sergey V. Razin[1]*

[1]Laboratory of Structural and Functional Organization of Chromosomes, Institute of Gene Biology RAS, 119334 Moscow, Russia
[2]Laboratory of Transgenesis, Institute of Gene Biology RAS, 119334 Moscow, Russia

Abstract It is demonstrated that a heterologous (chicken) CpG island containing five Sp1 canonical recognition sequences becomes highly methylated in the genome of transgenic mice bearing one or several copies of the transgene. Similar levels of methylation of the chicken CpG island were observed in different tissues of transgenic mice except the brain where the level of methylation of this chicken CpG-rich fragment was significantly lower than in other tissues. Analysis of susceptibility of the ''transgenic'' CpG island to Hpa II and Msp I restriction nucleases revealed an unusual methylation pattern interfering with the action of both of these enzymes. A conclusion has been drawn that heterologous CpG island *per se* does not contain all necessary signals permitting to maintain its own non-methylated status in the genome of transgenic animals. J. Cell. Biochem. 92: 99–103, 2004. © 2004 Wiley-Liss, Inc.

Key words: CpG island; DNA methylation; transgenic mice; chicken alpha-globin gene domain; SP1 transcription factor

Vertebrate genomes are characterized by significant depletion of CpG dinucleotides. The ratio of CpG to GpC dinucleotides in genomes of vertebrate animals is approximately 0.2. This is thought to be a result of instability of methylated cytosine (^{met}C) within CpG dinucleotides being the targets for eukaryotic cytosine-methylases [Holliday and Grigg, 1993]. In all genomes of vertebrates, there are, however, regions characterized by a balanced content of CpG and GpC. These regions known as CpG islands [Bird, 1986] are not methylated apparently because of the presence within them of numerous recognition sites for SP1 [Brandeis et al., 1994] and other transcription factors. CpG islands often include promoters of housekeeping genes and origins of DNA replication [Cross and Bird, 1995; Antequera and Bird, 1999]. Promoters of some tissue-specific genes are also located within CpG islands [Ariel et al., 1991]. CpG islands were thus proposed to contain signals essential for maintaining an active (open) chromatin configuration favorable for transcription initiation. In agreement with this idea, it was demonstrated that transgenes linked to CpG islands were partially protected from position effect variegation and were expressed at higher levels as compared to transgenes lacking CpG islands [Whitelaw et al., 1989; Shewchuk and Hardison, 1997; Hejnar et al., 2001]. The status of heterologous CpG islands in the genome of transgenic animals has not, however, been studied. In the present work, we decided to check whether a CpG island from the chicken genome contains signals necessary to prevent it from methylation in the genome of transgenic mice. To this end, transgenic mice bearing a CpG island from the chicken domain of alpha-globin genes were obtained. Analysis of the abovementioned CpG island for accessibility to restriction enzymes sensitive and insensitive to CpG methylation has demonstrated that in the genome of transgenic mice this chicken CpG island bearing five recognition motives for the universal transcription factor SP1 becomes heavily methylated.

Grant sponsor: Presidium of the Russian Academy of Sciences (grant for Physico-Chemical Biology).

*Correspondence to: Sergey V. Razin, Institute of Gene Biology RAS, 119334 Moscow, Russia.
E-mail: sergey.v.razin@usa.net

Received 21 October 2003; Accepted 2 December 2003
DOI 10.1002/jcb.20023

MATERIALS AND METHODS

Recombinant Clones

A recombinant clone bearing the chicken alpha globin domain upstream CpG island was obtained in our lab previously [Kalandadze et al., 1990] and the nucleotide sequence of the insertion was determined (X54965; AF098919). The CpG-rich insertion of this clone includes a replication origin [Razin et al., 1986] and a site of DNA loop anchorage to the nuclear matrix (LAR) [Farache et al., 1990; Razin et al., 1991]. A 2.9-kb DNA fragment bearing the CpG island with flanking sequences was cut out of the plasmid DNA by the *Hind*III restriction enzyme.

Preparation of Transgenic Mice

Transgenic mice were generated by microinjection of a linear DNA fragment into the male pronucleus of fertilized oocytes obtained from F1 hybrid (CBA × C57BL) females. DNA constructs were diluted in 0.1 mm EDTA, 10 mM Tris-HCl (pH 7.4) buffer to the concentration of about 300–400 copies per picoliter. After microinjection, the embryos were incubated for 1–2 h at 37°C in M16 medium and then transplanted to pseudopregnant female mice. Identified transgenic mice were mated to non-transgenic ones to produce the first generation (F1) offspring. All subsequent experiments were carried out using F1 transgenic mice.

Analysis of the Methylation Status of the Chicken CpG Island in the Mouse Genome

To analyze the methylation pattern of the chicken CpG island inserted into the mouse genome, the susceptibility of the corresponding DNA fragment to digestion by *Msp*I and *Hpa*II restriction enzymes was studied. The area under study contains 17 recognition sites (CCGG) for *Msp*I and *Hpa*II restriction enzymes of which one, *Hpa*II, does not cut DNA when the internal cytosine residue in the recognition sequence is methylated. To separate the inserted chicken CpG island from the flanking mouse DNA sequences all samples were additionally digested with the Pst I restriction enzyme that releases an internal 2.2 kb fragment from the 2.9 kb fragment of chicken DNA used for making transgenic mice (Fig. 1).

RESULTS

The upstream CpG island of the chicken domain of alpha-globin genes has been described in our previous publication [Razin et al., 2000]. The restriction map and compositional characteristics of this DNA fragment are presented in Figure 1. For the present study, it is important that the fragment contains five canonical recognition motifs for the ubiquitous transcription factor Sp1. As far as the compositional characteristics are concerned, the fragment represents a typical GC-rich CpG island. The CpG:GpC ratio in the internal part of the fragment (where all Sp1 recognition sites are located) almost reaches 0.9. In chicken cells of different lineages, this part of the DNA fragment under study is never methylated while the downstream end is selectively methylated in non-erythroid cells [Razin et al., 2000]. We have obtained several transgenic mice bearing in their genomes, the 2.9 kb *Hind*III fragment of chicken DNA shown in Figure 1. Since multiple copies of transgenes are known to be frequently methylated in genomes of transgenic animals [Henikoff, 1998], we have selected a line bearing a single copy of the transgene for further studies. DNA samples were isolated from different tissues of the transgenic mice. Erythroblasts at different stages of differentiation were taken from bone marrow of mice made anemic by phenylhydrasin injection. All DNA samples were treated with *Pst*I to release the internal

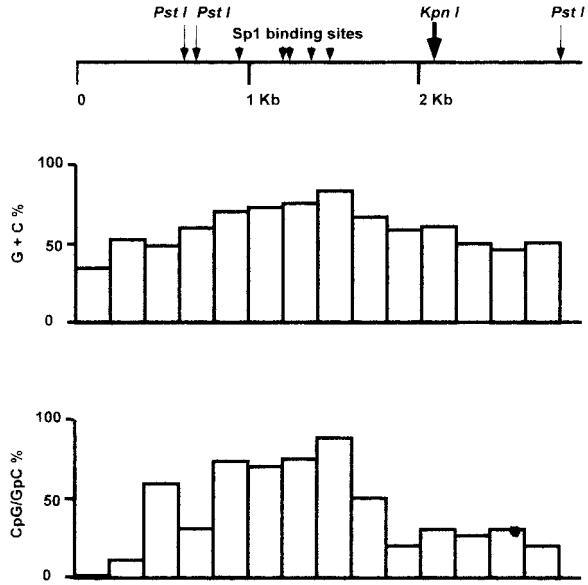

Fig. 1. A restriction map and compositional characteristics of the chicken alpha-globin domain upstream CpG island. Position 0 on the map corresponds to position 12085 on the DNA sequence deposited in Gene Bank under accession no.: AF098919.

part of the fragment under study and then with either *Hpa*II or *Msp*I. These enzymes recognize the same sequence, CCGG, but only *Msp*I can cut this sequence when the central cytosine residue is methylated by CpG methylases.

The DNA fragments were separated by electrophoresis and transferred to nylon filters. This was followed by hybridization with probes complementary to the ends of the 2.2 *Pst*I–*Pst*I fragment. The results of the hybridization experiments and their interpretation are presented in Figures 2 and 3. The 2.2-kb *Pst*I–*Pst*I

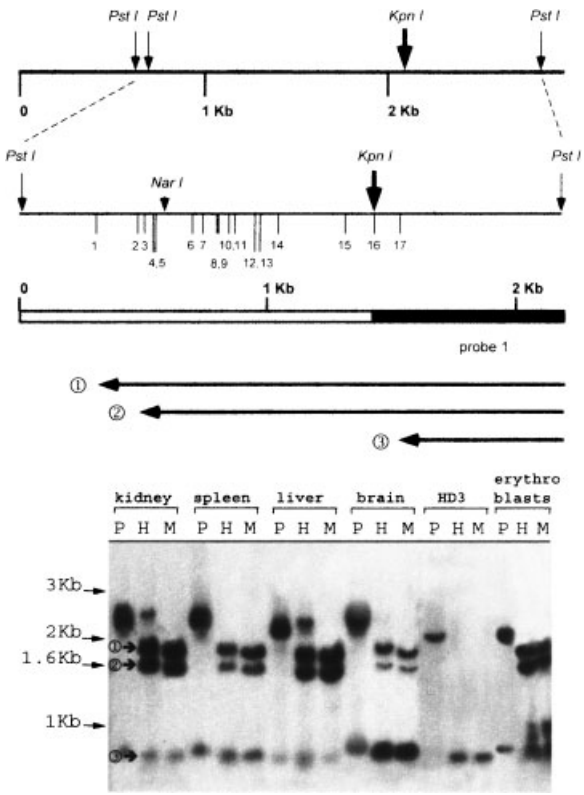

Fig. 2. Analysis of the susceptibility to *Hpa*II *and Msp*I of the internal part (*Pst*I–*Pst*I fragment) of the chicken CpG island in different tissues of transgenic mice. Hybridization with the right end-probe. The map of the 2.2 kb *Pst*I–*Pst*I fragment is shown on the top of the figure. The positions of *Hpa*II/*Msp*I recognition sites (CCGG) are indicated by numbers 1 to 17. Hybridization probe 1 is shown by a black rectangle on the distance scale. The arrows numbered (1), (2), (3) show the extensions of the visualized restriction fragments indicated on the left side of the blot by arrows with the same numbers. All samples except one named HD3 are DNAs prepared from different tissues (as indicated above the slots) of transgenic mice. HD3 are chicken erythroid cells (clone 6 of the line LSCC [Beug et al., 1982]) that were used in our previous study of the chicken alpha-globin domain upstream CpG island [Razin et al., 2000]. On the blot, the slots containing DNA digested with *Pst*I, *Pst*I and *Hpa*II, *Pst*I and *Msp*I are designated by letters P, H, and M, respectively.

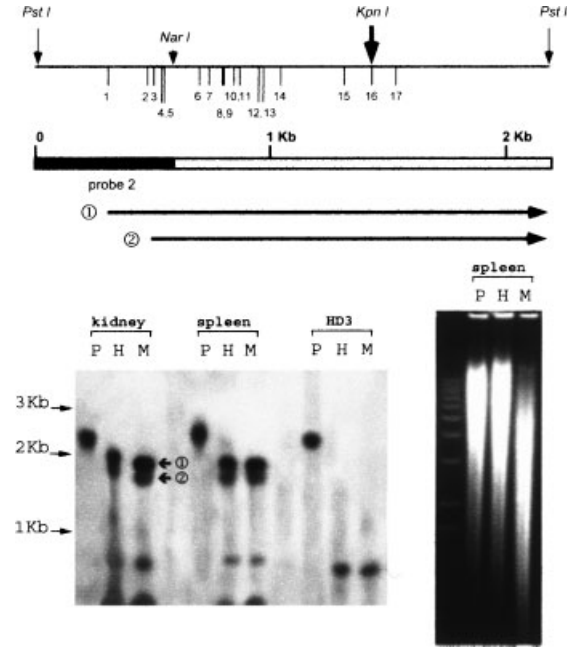

Fig. 3. Analysis of the susceptibility to *Hpa*II and *Msp*I of the internal part (*Pst*I–*Pst*I fragment) of the chicken CpG island in different tissues of transgenic mice. Hybridization with the left end-probe. All designations are the same as in Figure 2. The results of separation of DNA from kidneys of the transgenic mice treated with different restriction enzymes are shown at the bottom right side of the figure. Note that treatment with *Msp*I that insensitive to classical CpG methylation results in more profound digestion of total DNA as compared to treatment with *Hpa*II sensitive to CpG methylation.

fragment contains 17 recognition sites for *Hpa*II/*Msp*I (see the scheme at the top of Fig. 2). Assuming all these sites are not methylated, the largest DNA fragment recognized by probe 1 is an approximately 0.5 kb fragment delimited by *Hpa*II/*Msp*I recognition site 17 and the right *Pst*I recognition site (Fig. 2). This fragment is indeed seen as a single band in the samples from chicken HD3 erythroid cells and as the major band in the DNA samples from the brain of transgenic mice (fragment (3) in Fig. 2). In all other samples, two major bands of 1.6 kb and ~1.9 kb are clearly seen (fragments (2) and (1) in Fig. 2). These fragments represent the products of cleavage at sites 2/3 and 1 (see the map in Fig. 2). Thus, all CCGG palindromes in the central part of the fragment (*Hpa*II/*Msp*I recognition sites 4–16) are heavily methylated. Most surprisingly, these recognition sites are not susceptible to cleavage even with *Msp*I which is not sensitive to methylation of the central cytosine residue within the CCGG palindrome ($C^{met}CGG$), but

can not cut this sequence when both cytosine residues are methylated ($^{met}C^{met}CGG$) [Haigh et al., 1982]. In fact, the patterns of the transgene cleavage by *Hpa*II and *Msp*I were very similar if not identical in all samples studied. It was in drastic contrast with the overall degree of cleavage of DNA from different samples by *Hpa*II and *Msp*I. As expected, the cleavage by *Msp*I was much more profound (see the photo of agarose gel stained with ethidium bromide, Fig. 3). The conclusions drawn after inspection of the results of hybridization with probe 1 were fully confirmed when the results of hybridization with a probe complementary to the left end of the DNA fragment under study (probe 2) were analyzed. This probe again stained large protected fragments identical to fragments (2) and (1) visualized with probe 1 (Fig. 3). Similar results were obtained when several other transgenic mice bearing up to ten copies of the chicken alpha globin domain upstream CpG island were analyzed (not shown).

DISCUSSION

The results of the present study corraborate two important conclusions. First, CpG islands per se do not necessarily contain all signals preventing their methylation in an ectopic position. It might, thus, happen that CpG islands exist because of a complex interaction of different regulatory elements. They do not simply provide a favorable genomic context for the activity of promoters but are possibly maintained themselves (at least to a certain extend) because of the presence of active promoters. One may propose that interaction of general transcription factors (Sp1) bound to CpG islands with the tissue-specific ones bound to promoters/enhancers is necessary to prevent methylation of CpG islands. The question deserves further investigation.

Our results show that heterologous CpG islands should be used with caution in constructions designed for delivering transgenes. Being methylated, CpG islands (possessing compositional characteristis similar to these of prokaryotic DNA) may even suppress transcription from the neighboring promoter. Of course, individual CpG islands may contain different combinations of recognition sequences for transcription factors and we cannot exclude a possibility that some of them can stabilize the expression of transgenes especially in combination with strong authentic promoters. It is, however, worth mentioning that in our case the CpG island from the chicken erythroid-specific gene domain became heavily methylated even in erythroid cells of transgenic mice (in contrast to the fully non-methylated status of this CpG island in chicken erythroid cells). Extensive methylation of a heterologous CpG island observed in our study did not depend on the transgene copy number and was observed even when a single copy of the chicken DNA fragment was integrated into the mouse genome. Hence, the mechanisms triggering methylation in this particular case could hardly be similar to those involved in inactivation of tandem copies of a transgene [Henikoff, 1998].

The second important conclusion following from the results of the present study is that the system of methylation used in vertebrates for inactivation of heterologous genetic material may differ from classical CpG methylation carried out by all known mammalian DNA methyltransferases [Robertson, 2002], as both cytosine residues within the CCGG palindromes were found to be methylated in a CpG-rich fragment of chicken DNA integrated into the mouse genome. That is why they cannot be cut by *Msp*I which is not sensitive to the methylation of the central cytosines, i.e., to the CpG metilation per se within the CCGG [Haigh et al., 1982]. A similar pattern of transgenes methylation, although at a much lower frequency, was observed in transgenic mice previously along with the classical CpG methylation [Snibson et al., 1995]. In our case, almost all CCGG palindromes were not susceptible to *Msp*I cleavage apparently because of methylation of both cytosine residues.

REFERENCES

Antequera F, Bird A. 1999. CpG islands as genomic footprints of promoters that are associated with replication origins. Curr Biol 9:R661–R667.

Ariel M, McCarrey J, Cedar H. 1991. Methylation patterns of testis-specific genes. Proc Natl Acad Sci USA 88:2317–2321.

Beug H, Doederlein G, Freudenstrein C, Graf T. 1982. Erythroblast cell lines transformed by a temperature sensitive mutant of avian erythroblastosis virus. A model system to study erythroid differentiation in vitro. J Cell Physiol 1:195–207.

Bird AP. 1986. CpG-rich islands and the function of DNA methylation. Nature 321:209–213.

Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H. 1994. Sp1

elements protect a CpG island from de novo methylation. Nature 371:435–438.

Cross SH, Bird AP. 1995. CpG islands and genes. Curr Opin Genet Dev 5:309–314.

Farache G, Razin SV, Rzeszowska-Wolny J, Moreau J, Recillas Targa F, Scherrer K. 1990. Mapping of structural and transcription-related matrix attachment sites in the α-globin gene domain of avian erythroblasts and erythrocytes. Mol Cell Biol 10:5349–5358.

Haigh LS, Owens BB, Hellewel OS, Ingram VM. 1982. DNA methylation in chicken alpha-globin gene expression. Proc Natl Acad Sci USA 79:5332–5336.

Hejnar J, Hajkova P, Plachy J, Elleder D, Stepanets V, Svoboda J. 2001. CpG island protects Rous sarcoma virus-derived vectors integrated into nonpermissive cells from DNA methylation and transcriptional suppression. Proc Natl Acad Sci USA 98:565–569.

Henikoff S. 1998. Conspiracy of silence among repeated transgenes. Boiessays 20:532–535.

Holliday R, Grigg GW. 1993. DNA methylation and mutation. Mutat Res 285:61–67.

Kalandadze AG, Bushara SA, Vassetzky JYS, Razin SV. 1990. Characterization of DNA pattern in the site of permenent attachment to the nuclear matrix located in the vicinity of replication origin. Biochem Biophy Res Commun 168:9–15.

Razin SV, Kekelidze MG, Lukanidin EM, Scherrer K, Georgiev GP. 1986. Replication origins are attached to the nuclear skeleton. Nucl Acids Res 14:8189–8207.

Razin SV, Petrov P, Hancock R. 1991. Precise localization of the α-globin gene cluster within one of the 20- to 300-kilobase DNA fragment released by cleavage of chicken chromosomal DNA at topoisomerase II site in vivo: Evidence that the fragment are DNA loops or domains. Proc Natl Acad Sci USA 88:8515–8519.

Razin SV, Ioudinkova ES, Scherrer K. 2000. Extensive methylation of a part of the CpG island located 3.0–4.5 kbp upstream to the chicken alpha-globin gene cluster may contribute to silencing the globin genes in non-erythroid cells. J Mol Biol 209:845–852.

Robertson KD. 2002. DNA methylation and chromatin-unraveling the tangled web. Oncogene 21:5361–5379.

Shewchuk BM, Hardison RC. 1997. CpG islands from the alpha-globin gene cluster increase gene expression in an integration-dependent manner. Mol Cell Biol 17:5856–5866.

Snibson KJ, Woodcock D, Orian JM, Brandon MR, Adams TE. 1995. Methylation and expression of a metallothionein promoter ovine growth hormone fusion gene (*MToGH1*) in transgenic mice. Transgenic Res 4:114–122.

Whitelaw E, Hogben P, Hanscombe O, Proudfoot NJ. 1989. Transcriptional promiscuity of the human alpha-globin gene. Mol Cell Biol 9:241–251.