Volume 14 Number 24 1986 Nucleic Acids Research

The maintenance of methylation-free islands in transgenic mice

A-B.Kolsto, G.Kollias, V.Giguere, K-I.Isobe, H.Prydz and F.Grosveld

Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Received 8 October 1986; Revised and Accepted 24 November 1986

SUMMARY

The Thy-1 gene is expressed in a tissue- and stage-specific pattern and has a typical 1.6kb methylation-free island (MFI) covering about 600bp upstream and downstream of the two alternative first exons. By microinjection of a mouse Thy-1.1/human Thy-1 gene into fertilized eggs, we were able to show that the MFI is restored in the transgenic mice. The flanking sequence became methylated, but the MFI remains unmethylated in all tissues of transgenic mice at different developmental stages tested, irrespective of the site of expression of the gene. There is one exception, in extra-embryonal tissues of 14.5 day embryos a small percentage of the islands were methylated. We conclude that maintenance of the MFI is regulated by cis-acting sequences present within the gene, and indicates that the unmethylated state of the islands is consistent with a necessary but not sufficient condition for expression of the gene.

INTRODUCTION

The dinucleotide CpG is 4-5 fold under-represented in the vertebrate genome, but accounts for almost all of the methylcytosine present. Between 60% and 90% of CpGs are methylated at the 5-position on the pyrimidine ring. This modification renders vertebrate DNA less cleavable by certain restriction endonucleases that are inhibited when this doublet is methylated. The bulk of mammalian DNA is highly methylated, but Cooper et al. (1) described that about 1% of the vertebrate genome was frequently cut by HhaI and HpaII, both of which are methylation sensitive. Recently, it has been shown that these sites are clustered and that there are about 30,000 of these CpG-rich, methylation-free islands in the haploid genome (2). They were found in all tissues investigated and their sequences are, for the most part, unique. Several genes, mainly housekeeping genes, have been reported to have such methylation-free islands, covering 500-2000bp at their 5' end. They frequently include the first exon(s) as well as upstream sequences (see review (3)). At present the role of these islands in the control of gene expression is not clear, nor is it known how an MFI is maintained.

The Thy-1 gene is a member of the immunoglobulin superfamily of genes (4) and is expressed in a highly tissue specific and developmentally regulated manner (see review (5)). The mouse Thy-1 gene (6) has a CpG-rich 5' region which is a methylation-free island (7), and covers two very unusual promoters (6-9). Here we describe a study on the maintenance of the methylation pattern of the Thy-1 gene, by studying the methylation pattern of a mouse-human hybrid gene, introduced into the germ line of mice. The hybrid gene was made up from the 5' part of the murine Thy-1.1 gene and the 3' untranslated part of the human Thy-1 gene, to allow clear distinction of the introduced hybrid gene and its products from the endogenous Thy-1.2 background.

We show that in all transgenic mice and their offspring, the island was reformed in all tissues, whether the construct was highly expressed or not, and whether a few or multiple copies of the gene were inserted at any one locus. The only exception to the maintenance of the MFI was a slight degree of methylation of the island in extra-embryonic foetal tissue (placenta and yolk sac).

MATERIALS AND METHODS

(CBA/J x C57Bl/10)F1 mice were bred at NIMR, London. Restriction enzymes were obtained from Anglian Biotechnology, Colchester, Essex, U.K., Bio-Labs, Beverly, MA, USA, Bethesda Research Laboratories, Bethesda, MD, USA and Boehringer, Mannheim, Germany. Isotopes were from New England Nuclear, UK.

Construction of hybrid genes and microinjection

A mouse-human Thy-1.1 hybrid gene was constructed by replacement of the non-coding part of the murine gene by the human equivalent (10). A 9.2kb EcoRI fragment of the hybrid gene was injected into the pronuclei of fertilized eggs from (CBAxC57 B1/10)F1 mice as described previously (11).

DNA methylation

Genomic DNA from various organs of normal and transgenic embryos or adult mice was cut with 5-10-fold excess restriction enzymes. This was necessary to obtain complete digestion with the enzymes SacI, MspI, HpaII and HhaI, and was monitored by the addition of an internal λ DNA control. The fragments were analyzed by Southern blotting (23) using 0.8 or 1.0% agarose gels. The probes were prepared from the gene construct and labelled with ^{32}PdTTP and ^{32}PdATP using Klenow fragment and a

hexadeoxyribonucleotide mixture (Pharmacia, Uppsala, Sweden) as described (12).

RESULTS

Transgenic mice

The hybrid gene construct consisted of a 6kb fragment from the 5' part of the 9.2 kb cloned Thy-1.1 (AKR) mouse gene (6). This fragment contained the first four exons (Ia and b, II, III in Fig. 1) and the 5' end of exon IV coding for the transmembrane part of the protein and about 10 bases of the 3' untranslated region. Linked to it was 3.2kb of the human Thy-1 gene (13) containing the remaining part of exon IV, i.e. the major part of the 3' untranslated region and 2kb of the 3' flanking sequence. The gene was injected as a 9.2kb EcoRI fragment, as previously described (11).

Four transgenic mice were identified in a litter of 13 newborn mice by tail DNA blots. The first mouse (Tg6) died before mating

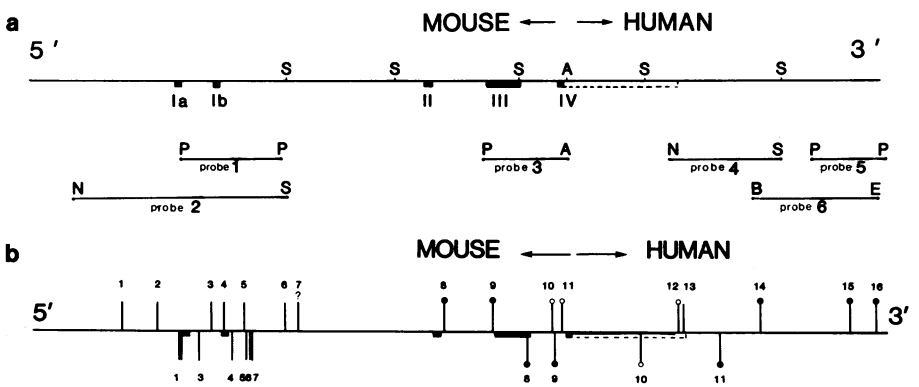

Fig. 1:

a) The mouse human gene construct of Thy-1.
The exons I-IV and the probes used for the Southern blot anlysis are indicated. A: ApaI, B: BglII, E: EcoRI, N: NcoI, P: PstI, S: SacI. The solid blocks represent mouse exon sequences, the open block human exon seqences.

b) The sites for MspI/HpaII are indicated by upwards bars and the sites for HhaI by downwards bars. Methylated sites are indicated with filled circles and partly methylated sites with open circles. Simple bars indicate unmethylated sites. The MspI sites and the HhaI sites are referred to in the text by numbers assigned from the 5' end (1-16 for MspI sites and 1-11 for HhaI sites). These were obtained by using the sequence of the mouse Thy-1.2 gene (6,8,24) and restriction digests of the Thy-1.1 construct.

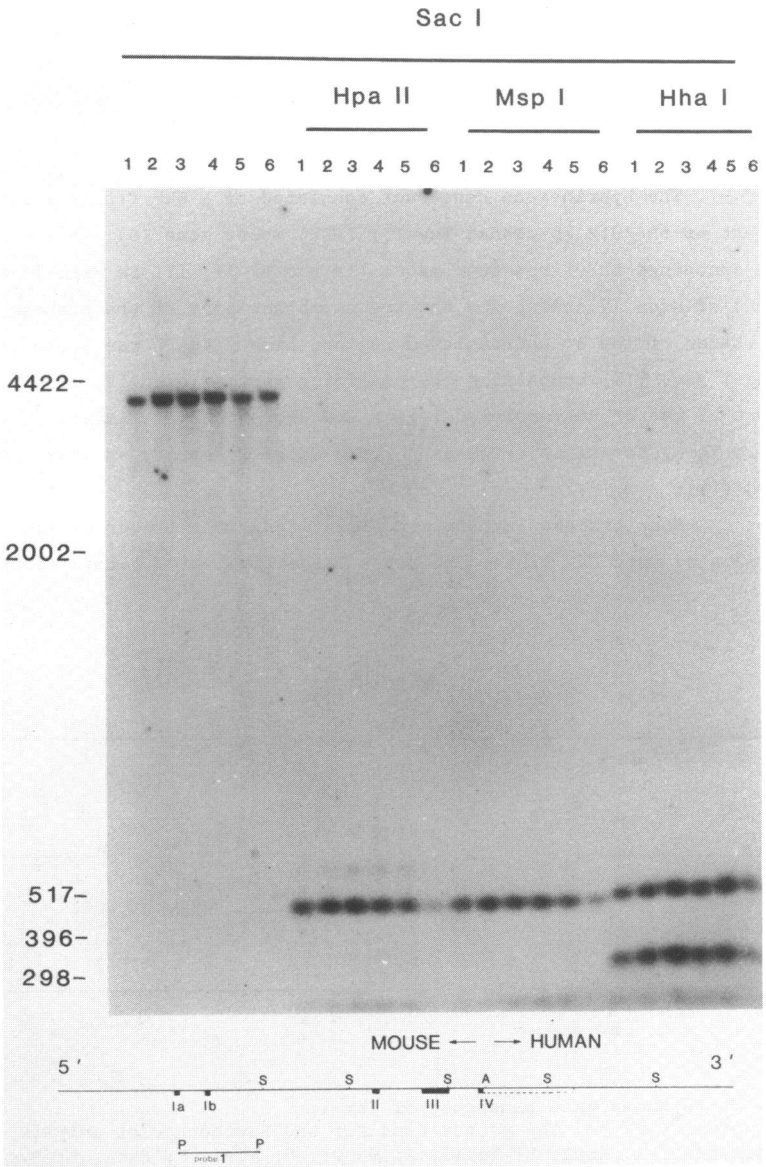

Fig. 2:
Southern blot analysis (23) of DNA from transgenic mouse Tg12 tissues.

The DNA was cut with SacI and with MspI, HpaII or HhaI as indicated. 5µg DNA was used for each lane. Lane 1: heart, lane 2: thymus, lane 3: brain, lane 4: kidney, lane 5: spleen. lane 6: liver. Probe 1 (see Fig. 1) was used for hybridization. λ DNA (digested with HindIII) and pBR322 DNA (digested with HinfI) were used as markers.

and was only partially analyzed, but showed essentially the same results as the other mice. Founder mice Tg9, Tg11 and Tg12 contained 20, 20 and 60 copies of the Thy-1 gene, respectively. Southern blotting showed that the genes were integrated in a head to tail concatamer (see below and Fig. 4). Very little DNA was lost during the integration of the exogenous DNA, since 50% of the inserted copies still contained the terminal EcoRI sites, while all copies retained all other restriction sites (data not shown).

Mapping of potential methylation sites in the gene construct

Most of the sequence of the mouse Thy-1.2 and the human Thy-1 genes is known (6,8,14,24). The murine Thy-1.1 and Thy-1.2 sequences show several differences in their non-coding regions (6) and the HpaII and HhaI sites in the constructed gene were therefore mapped by restriction fragment analysis. The construct contained a total of 16 Msp sites and 11 HhaI sites, of which 14 (7 for MspI and 7 for HhaI) are clustered in a 5' 1.6 kb fragment surrounding exons Ia and Ib (Fig. 1b). This region contains the MFI and has a high CG dinucleotide frequency (6,7). The remainder of the gene has a low CG frequency as described for the bulk of DNA (2). We have numbered the MspI (1-16) and HhaI (1-11) sites in the hybrid gene, starting from the 5' end.

Methylation pattern

Sac I-digested genomic DNA from transgenic mice was redigested with the methylation sensitive enzymes HpaII and HhaI or the methylation insensitive enzyme MspI as a control and Southern blotted. The blots were hybridized with different radioactively labelled probes as indicated in Fig.1. When probes 1 or 2 were used, a SacI fragment of 4.3kb was detected indicating a head-to-tail arrangement of the tandemly inserted genes in Tg12 offspring (Fig. 2). Double digestion with HpaII or HhaI showed the presence of a methylation-free island in the 5' part of the construct gene. The HpaII digest showed an almost identical pattern to the MspI digests. The largest fragment from the promoter region hybridizing to probe 1 was 460 bp, similarly, small fragments were seen in the HhaI digests (Fig. 2). Only a very faint band corresponding to a fragment of about 560bp was noticed after HpaII digestion, which shows a slight degree (<10%) of methylation of MspI site no.6. This site might therefore represent the 3' border of the MFI in the transgenic mice.

Further downstream in the mouse part of the construct, very few sites for MspI/HpaII and HhaI are available and these sites were fully

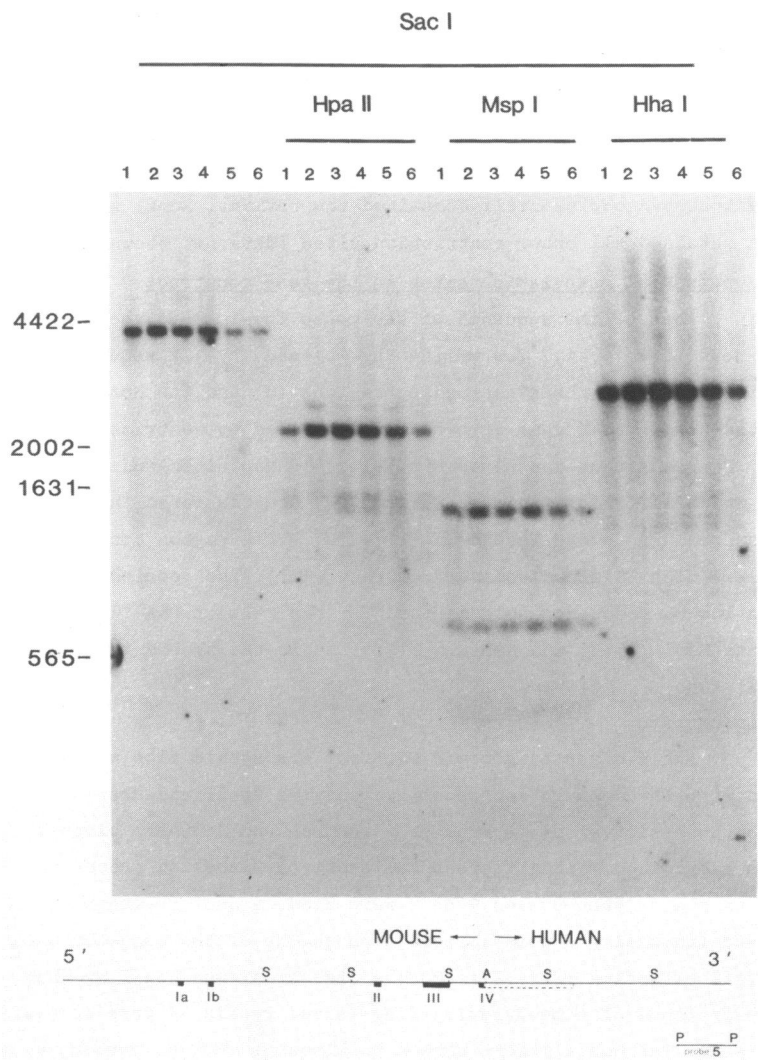

Fig. 3:
Southern blot analysis of DNA from transgenic mouse (Tg12) tissues. The lanes were as in Fig. 2. Probe no. 5 was used for hybridization.

or partially methylated using probes 2, 3, 4, 5 and 6 (Msp/HpaII sites 8-11 and HhaI sites 8 and 9, see Fig. 1b).

For example, Fig. 3 shows a blot of the very 3' end of the constructed gene with probe 5. Since multiple copies of the gene are inserted in a head to tail configuration, the digest gives rise to the

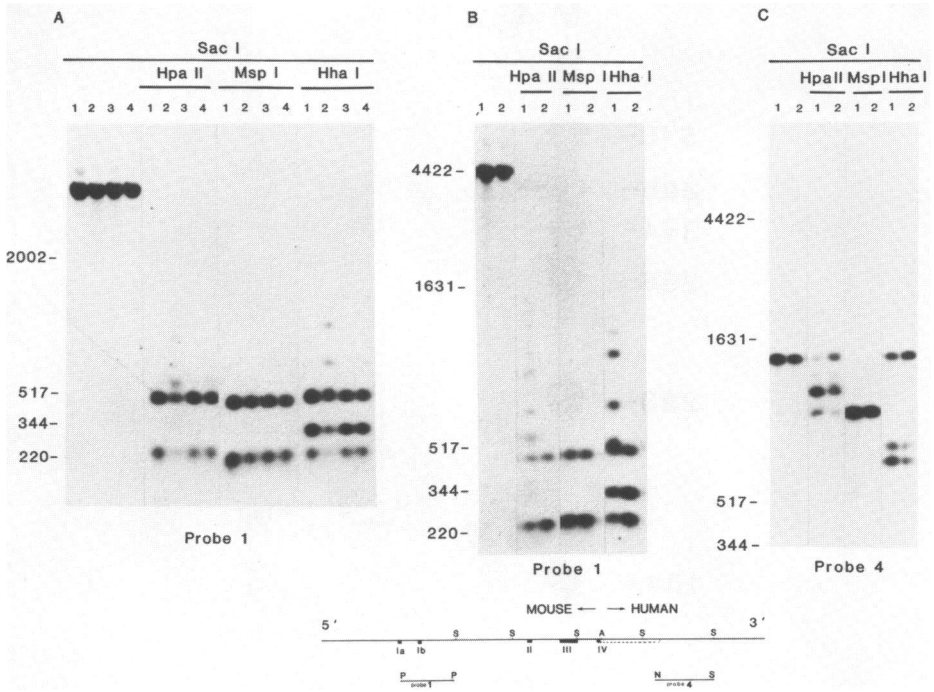

Fig. 4:

Southern blot analysis of DNA from 14.5 day embryos of transgenic mouse (Tg12). The DNA was cut with SacI and with MspI, HpaII or HhaI as indicated. 5µg DNA was used for each lane.

a) Lane 1: brain, lane 2: yolk sac, lane 3: muscle, lane 4: liver.

b) Lane 1: yolk sac, lane 2: brain.

c) Reprobing of blot b with probe 4.

Blots a and b show the same methylation in the yolk sac. Reprobing of blot b excludes that the methylation pattern was due to partial cutting with SacI.

4.3kb Sac I fragment containing the 3' end of one copy linked to the 5' end of the next. Methylation of the sites in the 3' end of the construct, and non-methylation of the sites in the 5' end (MFI part) of the construct would be expected to yield a 2.7kb fragment with HhaI and a 2.3 kb fragment with HpaII. These fragments were indeed observed (Fig. 3), indicating that the two MspI sites (nos. 15 and 16, Fig. 1b) were fully methylated and confirming that MspI site no. 1 and the HhaI sites nos. 1 and 2 are the first measurable sites at the 5' border of the MFI which are unmethylated.

The island, including the faint methylation of MspI site no.6 appeared to be fully conserved in all the gene copies in all the tissues of

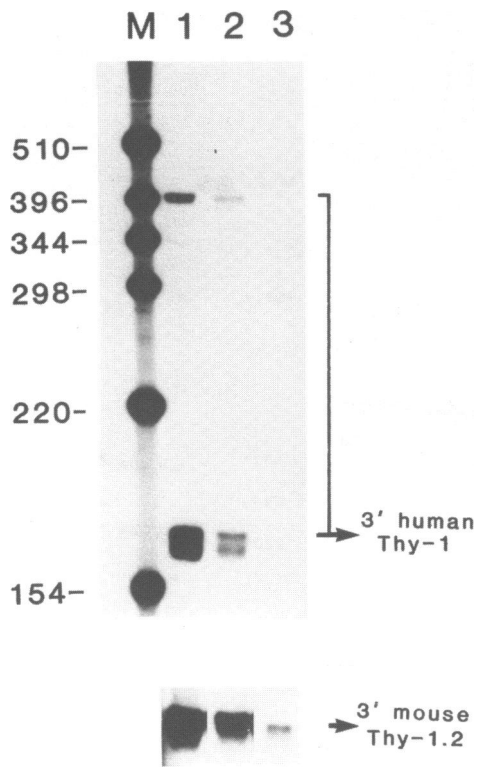

Fig. 5:
Expression of the Thy-1 transgene in adult mice.
Total RNA was extracted from 8 week old transgenic brain (lane 1), thymus (lane 2) and spleen (lane 3). The 3' termini of the hybrid mouse human Thy-1 transgenic RNA and the endogenous Thy-1.2 RNA were mapped by S1 nuclease using either a 950bp 3' labelled NcoI/BglII human probe or a 680bp 3' labelled BamH1/BamH1 mouse Thy-1.2 probe. The sizes of the protected fragments are as expected (6,14). In the case of the human Thy-1 RNA termination is shown to occur in two sites. Sizes were estimated using 5' end labelled fragments of pBR322 x HinfI digest.

all the transgenic mice investigated, with one exception. The extra-embryonal foetal tissues (yolk sac and placenta) of 14.5 day embryos (but not other tissues from the same embryos), showed slight methylation of several sites (nos. 3-5 of the MspI and nos. 3-7 of the HhaI sites) (Fig. 4). The endogenous murine Thy-1.2 gene showed an unmethylated island both in foetal, including extra-embryonal and in postnatal tissues (not shown). Thus, the MFI of the endogenous gene was also fully unmethylated in the extra-embryonal tissues (yolk sac and placenta) in contrast to the result observed for the introduced gene construct.

Correlation of expression of Thy-1 antigen and methylation of the gene

Normally brain and thymocytes of the adult mouse and only brain of man are the major tissues of Thy-1 expression (5).

Using S1 nuclease protection (25) assays with probes from the ApaI site in the fourth exon, we measured the RNA levels originating from the endogenous and the transgenic Thy-1 gene. Fig. 5 shows that high levels of transgenic transcript in Tg12 mice are present in brain and thymus, as represented by the 165nt (major polyA site) and the 400nt (minor polyA site) protected fragments. The transgene is, however, not expressed in spleen (Fig. 5, lane 3) or in other Thy-1 negative tissues (not shown). The endogenous gene is represented by a 115nt protected fragment (only one major polyA site) which is present in brain, thymus and spleen (Fig. 5 inset), but not in other Thy-1 negative tissues (not shown). The other transgenic mice show the same pattern of expression, although the levels of Thy-1 transcripts are different (15). The spleen result is somewhat surprising and a detailed analysis indicates that the absence/presence of a positive/negative regulatory element might be responsible for the difference observed between the endogenous and the transgene (15). On the other hand, the tissue-specific expression in brain and thymus indicates that most of the control elements for specific expression are present in the construct. The expression pattern of the transgene in the different tissues (including spleen) indicates that although the presence of the MFI might be a necessary requirement for expression, the maintenance or formation of an island is not related to active expression.

DISCUSSION

The murine and human Thy-1 genes both have methylation-free islands in their 5' ends as evidenced by the unhindered cleavage by HpaII and HhaI of all potentially susceptible sites in this region. The island extends over about 1.6kb in the mouse, covering exons Ia and Ib and about 600 bases upstream and downstream, and is well within the typical size range for such islands (2).

The mouse-human hybrid gene used to establish the transgenic mice was injected into eggs in a non-methylated state. It was subsequently methylated to a pattern that is essentially the same as that of the endogenous mouse Thy-1 gene, i.e. methylated sequences flank an MFI containing promoter sequences. We do not know at what stage this methylation took place. This result is highly reproducible in different

mice, i.e. the maintenance of the methylation-free island, as well as the methylation pattern of the human part of the gene construct is independent of the (presumably random) site(s) of insertion in the murine genome, and the number of genes in the integrated multimer. Moreover, although we do not know its importance, MspI/HpaII site 13 is never methylated. This implies that the information necessary to maintain these patterns lies within the injected gene itself and might even be an intrinsic property of CG-rich DNA in eukaryotic cells. It is perhaps interesting to note in this respect that the H-2K gene which normally contains an MFI is expressed and remains unmethylated in F9 teratocarcinoma cells, while the pBR vector with a normal CG frequency is not expressed and becomes at least partially methylated (16 and unpublished).

It has been suggested (3) that methylation-free islands distinguish regions of the genome available for interaction with nuclear factors in all cells, a property that would relate, in particular, to house-keeping genes. Genes that are expressed in a highly tissue-specific manner are thought not to be regularly associated with such CpG-rich, non-methylated sequences and be part of the (methylated) bulk of the DNA. However, the Thy-1 antigen is expressed in a highly tissue-specific manner (5), but the methylation-free island was found in all tissues, irrespective of whether the antigen was expressed or not. Other examples are collagen II and α-globin (3), which clearly shows that MFI are not necessarily related to "house-keeping" genes. A second possibility is that MFI's are related to constitutively transcribed genes, which might be regulated in different tissues by transcriptional enhancement and RNA stability. However, Thy-1 RNA run-off studies on nuclei from various mouse tissues exclude the possibility that transcription of a highly unstable RNA takes place in non-expressing tissues (15). This leaves the possibility that an MFI could constitutively bind common transcription factors, which might initiate transcription from constitutive promoters (such as the HPRT gene), but not from tissue-specific promoters such as Thy-1 promoters. Such promoters would need additional tissue specific factor(s) or chromatin alterations to activate (measurable) transcription in a tissue specific manner. DNA methylation in the bulk of the DNA might simply serve to prevent the binding of common factors, allowing the cell to use a limited amount of factors. On the other hand, tissue specific factors would not be prevented from binding and could activate methylated genes (such as globin genes) to transcribe and become unmethylated as a consequence. The opposite situation

i.e. methylation of an MFI clearly might play a role in the inactivation of a gene, as shown for the HPRT gene on the inactive X chromosome (17). Similar observations have been made for *in vitro* methylated genes in transfected cells (18-22). It is unknown whether methylation of the Thy-1 gene would inhibit transcription in the appropriate tissues, i.e. our results are consistent with a role for the methylation free island as a permissive and possibly necessary, but not sufficient, condition for transcription of the gene.

ACKNOWLEDGEMENTS

A.-B.K. and H.P. are on leave from the University of Oslo, Norway, and are supported by the Norwegian Research Council for Sciences and the Humanities and the Norwegian Cancer Union. GK was supported by a long-term EMBO fellowship, VG was supported by a Canadian MRC fellowship and KI was supported by the ICRF. This work was supported by the British Medical Research Council.

REFERENCES

1. Cooper, D. N., Taggart, M. H. and Bird, A. P. (1983) Nucl. Acids Res. 11, 647-658.
2. Bird, A., Taggart, M., Frommer, M., Miller, O. J. and McLeod, D. (1985) Cell, 40, 91-99.
3. Bird, A. P. (1986) Nature, 321, 209-213.
4. Williams, A. F. and Gagnon, J. (1982) Science, 216, 696-703.
5. Morris, R. (1985) Dev. Neurosci. 7, 135-160.
6. Giguere, V., Isobe, K-I. and Grosveld, F. G. (1985) EMBO J. 4, 2017-2024.
7. Giguere, V., Spanopoulou, E. and Grosveld, F. (1986) submitted.
8. Ingraham, H., Lawless, G. and Evans, G. (1985) J. Immunol. 136, 1482-1489.
9. Ingraham, H. and Evans, G. (1986) Molec. Cell. Biol. 6, 2923-2931.
10. Isobe, K-I., Kollias, G., Kolsto, A-B. and Grosveld, F. (1986) J. Immunol., in press.
11. Kollias, G., Wrighton, N., Hurst, J. and Grosveld. F. (1986) Cell, 4, 89-94.
12. Feinberg, A. P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266-267.
13. van Rijs, J., Giguere, V., Hurst, J., van Agthoven, T., van Kessel, A. G., Goyert, S. and Grosveld, F. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 82, 5832-5835.
14. Seki, T., Spurr, N., Obata, F., Goyert, S., Goodfellow, P. and Silver, J. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 6657-6661.
15. Kollias, Spanopoulou, E., Grosveld, F., Ritter, M., Beech, J. and Morris, R. (1986) submitted.
16. Rosenthal, A., Wright, S., Quade, K., Gallimore, P., Cedar H. and Grosveld, F. G. (1985) Nature, 315, 579-581.
17. Wolf, S. F., Jolly, D. J., Lunmen, K. D., Friedman, T. and Migeon, B. R. (1984) Proc. Natl. Acad. Sci. U.S.A., 81, 2806-2810.

18. Busslinger, M., Hurst, J. and Flavell, R. A. (1983) Cell, 34, 197-206.
19. Chapman, V. M., Kratzer, P. G., Siracusa, L, O. O., Quaratillo, B. A., Evans, R. and Liskay, R. M. (1983) Proc. Natl. Acad. Sci. U.S.A., 79, 5357-5371.
20. Venolia, L. and Gartler, S. M. (1983) Nature, 302, 82-83.
21. Keshet, I., Yisraeli, J. and Cedar, H. (1985) Proc. Natl. Acad. Sci. U.S.A., 82, 2560-2564.
22. Murray, E. and Grosveld, F. G. (1985) In Razin, A. and Cantoni, G. L. (eds.) Biochemistry and Biology of DNA Methylation. Progress in Clinical and Biological Research. Alan R. Liss, Inc., New York, Vol. 198, 157-176.
23. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
24. Chang, H. C., Seki, T., Moriuchi, T. and Silver, J. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 3819-3823.
25. Weaver, R. and Weissmann, C. (1979) Nucl. Acids Res., 6, 1175.