PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Demethylation of Genes in Animal Cells

D. Frank, M. Lichtenstein, Z. Paroush, Y. Bergman, M. Shani, A. Razin and H. Cedar

Phil. Trans. R. Soc. Lond. B 1990 **326**, 241-251
doi: 10.1098/rstb.1990.0008

Phil. Trans. R. Soc. Lond. B **326**, 241–251 (1990)
Printed in Great Britain

Demethylation of genes in animal cells

By D. Frank[1], M. Lichtenstein[2], Z. Paroush[1], Y. Bergman[2], M. Shani[3], A. Razin[1] and H. Cedar[1]

Departments of [1] *Cellular Biochemistry and* [2] *Experimental Medicine, The Hebrew University-Hadassah Medical School, Jerusalem, Israel*

[3] *Institute of Animal Sciences, ARO, The Volcani Center, Bet-Dagon, Israel*

Tissue-specific animal cell genes are usually fully methylated in the germ line and become demethylated in those cell types in which they are expressed. To investigate this process, we inserted a methylated IgG κ gene into fibroblasts and lymphocytes at various stages of development. The results show that this gene undergoes demethylation only in the mature lymphocytes and therefore suggest that the ability to demethylate a gene is developmentally regulated. These studies were supported by similar experiments using the rat *Insulin* I gene, and in this case it appears that the *cis*-acting elements that control demethylation may be different from those responsible for gene activation. The ability to demethylate the housekeeping gene *APRT* is also under developmental control, because this occurs only in embryonic cells, both in tissue culture and in transgenic mice.

1. Introduction

The role of DNA methylation in the regulation of gene expression in animal cells is slowly unveiling. Initial studies on local methylation patterns revealed that tissue-specific genes are heavily methylated in most tissues, but are unmethylated in their tissue of expression (Yisraeli & Szyf 1984). On the other hand, housekeeping genes appear to contain CpG islands (Bird *et al.* 1985), which are unmethylated in every tissue and at all stages of development (Stein *et al.* 1983). This clear-cut inverse correlation between gene expression and DNA methylation levels strongly suggested that this modification inhibits RNA transcription and thus may have a function in controlling the appearance of gene activity during development.

The major proof that methylation may have a direct effect on gene expression comes from experiments using DNA-mediated gene transfer. While unmethylated genes are generally active following transfection, enzymatically methylated constructs are strongly inhibited (Stein *et al.* 1982; Busslinger *et al.* 1983). Indirectly, this experiment suggests that DNA modification may also be the cause of inactivity of the endogenous tissue-specific genes present in the same cell type. Although the exact mechanism of transcriptional inhibition is not known, DNA modification most likely affects gene expression by interfering with DNA interactions of the gene specific (Becker *et al.* 1987) and the more general ubiquitous chromatin proteins (Keshet *et al.* 1986).

These data paint a fairly simple and attractive model for explaining the role of DNA methylation during development. All tissue-specific genes are fully methylated in the germ line (Yisraeli & Szyf 1984) and throughout development, and as such, this modification serves as a general marker, which automatically puts genes into an inactive conformation, without the need for gene-specific *cis*- and *trans*-acting factors. Gene activation and its programmed

demethylation, on the other hand, are probably carried out by specific recognition signals in the tissue of expression. In keeping with this mechanism, housekeeping genes have a fixed unmethylated pattern in all tissues, and are thus always placed in an active conformation. The best evidence in support of this hypothesis comes from transfection experiments with the muscle-specific α-actin gene (Yisraeli *et al.* 1986). When a methylated α-actin gene is inserted into fibroblasts, this modification is preserved and inhibits gene expression. In myoblasts, however, the modified α-actin gene is specifically recognized and activated and this is accompanied by a site-specific demethylation, which mimics that which occurs to the endogenous gene *in vivo*. Thus, this specific cell type, as opposed to others, has the ability to overcome α-actin DNA methylation and to put this gene in an active conformation.

These studies provide important initial information on the process of gene-specific demethylation in animal cells. By using the same gene transfer approach, it should be possible to analyse additional genes and thus derive a picture of the general aspects of this mechanism. Furthermore, our recent observation that demethylation can be observed following transient transfection should open the way to studying the mode of action of this enzymic system.

2. Results

Before we can begin to understand the role of tissue-specific gene demethylation, it is essential to characterize the molecular aspects of this process. This should include a study of the site and cell-type specificity as well as an analysis of the *cis*-acting elements which direct this event.

The Ig κ light chain has been shown to be methylated at several sites in tissues that do not express the gene, but is completely unmethylated in B cells, which actively transcribe this gene (Mather & Perry 1983). To test whether B lymphocytes retain the ability to recognize the gene and demethylate it, an *in vitro* methylated Ig κ construct of $V_\kappa 19 J_2 C_\kappa$ was prepared and introduced into these cells by DNA-mediated gene transfer.

As shown in figure 1, the κ gene has one *Hpa*II site and two other *Hha*I sites and these were all methylated *in vitro* using the appropriate enzymes. After transfection to a plasmacytoma cell line (S194), we assayed the methylation state of these same sites. When genomic DNA is cut with *Bgl*II, the expected 2.8 kilobase (kb) band is observed, but digestion with *Hpa*II, which yields a 1.4 kb band, and *Hha*I (2.2 kb) show that both of these 5′ sites are completely unmethylated. Further analysis using a *Pvu*II/*Hin* dIII digestion strategy revealed that the 3′ *Hha*I is also unmethylated, indicating that all of the sites in the exogenous κ gene underwent demethylation in this cell line. This demethylation seems to be a dominant-acting phenomenon, as it occurred in all of the κ chain molecules despite the fact that we tested pools of transformed clones, each of which has its own site of integration. It should be noted that the additional bands that appear on this blot are the result of cross-hybridization with endogenous gene sequences. Within B cells, other methylated sequences, such as those in the plasmid vector or sequences from other transfected genes remained almost fully methylated (unpublished data), indicating that this reaction is site-specific. Furthermore, the demethylation process is also cell-type specific because the exogenous κ gene remained fully methylated when introduced into mouse fibroblast cells (figure 1) or pre-B lymphocytes (unpublished data).

Another biological system that is amenable to demethylation analysis is that of the rat insulin gene, which has been shown to undergo tissue-specific demethylation in the beta- (β)-islet cells

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

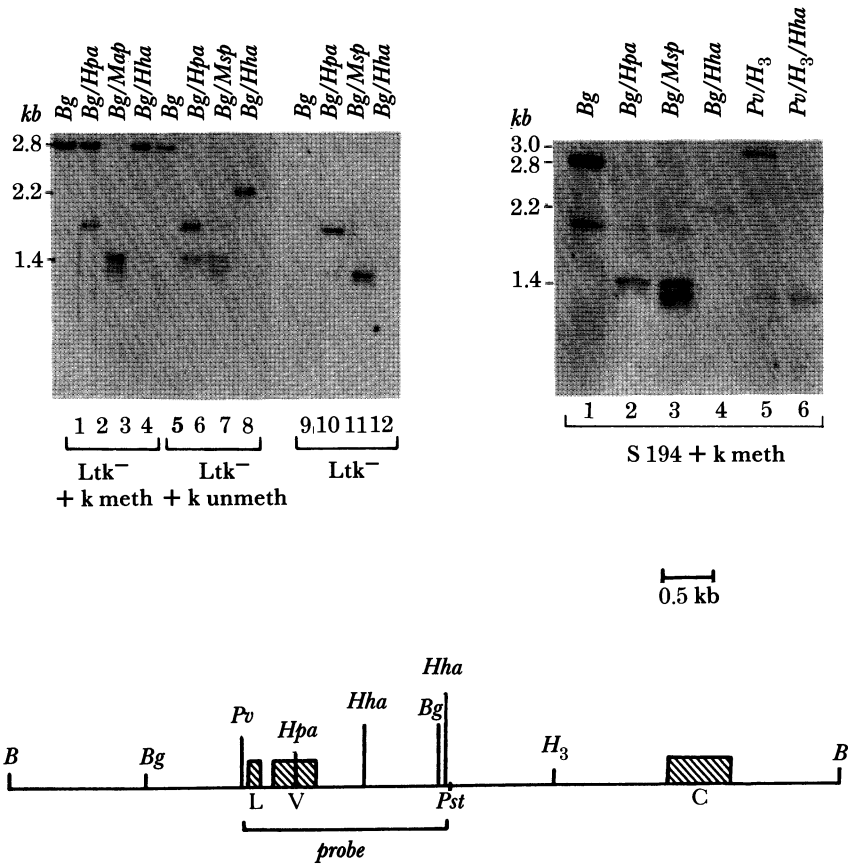


FIGURE 1. Methylation analysis of Ig κ following its transfection to cells. A plasmid (Kelley *et al.* 1988) containing the $V_{\kappa}19J_2C_{\kappa}$ was methylated *in vitro* (Yisraeli *et al.* 1986) using *Hpa*II and *Hha*I methylases and co-transfected into cell line S194 by electroporation (Kelley *et al.* 1988) or into mouse L-cells either by the calcium phosphate precipitation technique or by electroporation. In both cases, colonies containing the plasmid were isolated by selection for neomycin and these were pooled and expanded and then used to prepare DNA for Southern blot analysis. DNA was digested with *Bgl*II and in addition, with either *Hpa*II, *Msp*I or *Hha*I in order to assay methylated sites near the V region. In the case of S194, the 3′ *Hha*I site was analysed by digestion with *Pvu*II/*Hin* dIII. In all cases, 20 μg of DNA were loaded on each lane and the final filters were probed using a nick-translated *Pvu*II/*Pst*I fragment. In the case of the L-cell transfection, gel analysis is shown for both the methylated Ig κ construct as well as for a non-methylated control. In addition to the expected 2.8 kb band and its cleavage products, these autoradiograms also show hybridization to the endogenous equivalent of the Ig κ gene, as is demonstrated by the blot of the untransfected Ltk⁻ control.

of the pancreas during normal development (Cate *et al.* 1983). To analyse the demethylation that takes place on this gene, we obtained a hamster-insulinoma cell line that expresses high levels of insulin and has been shown to contain the *trans*- acting factors necessary to activate foreign insulin-gene constructs (Walker *et al.* 1983). We methylated the rat insulin gene *in vitro* using both the *Hpa*II and *Hha*I methylases and co-transfected it into both fibroblasts and insulinoma cells, selecting for neomycin resistance for isolating the colonies. DNA from these cells was then isolated and subjected to Southern blot restriction-enzyme analysis. The results shown in figure 2 demonstrate that the *Hpa*II and *Hha*I sites at the 3′ end of the gene undergo considerable demethylation in the insulinoma cell-line. Further mapping of all of the methylatable sites in the gene domain revealed that only site H1 remained methylated (unpublished data) and this corresponds to the same pattern of modification as the endogenous

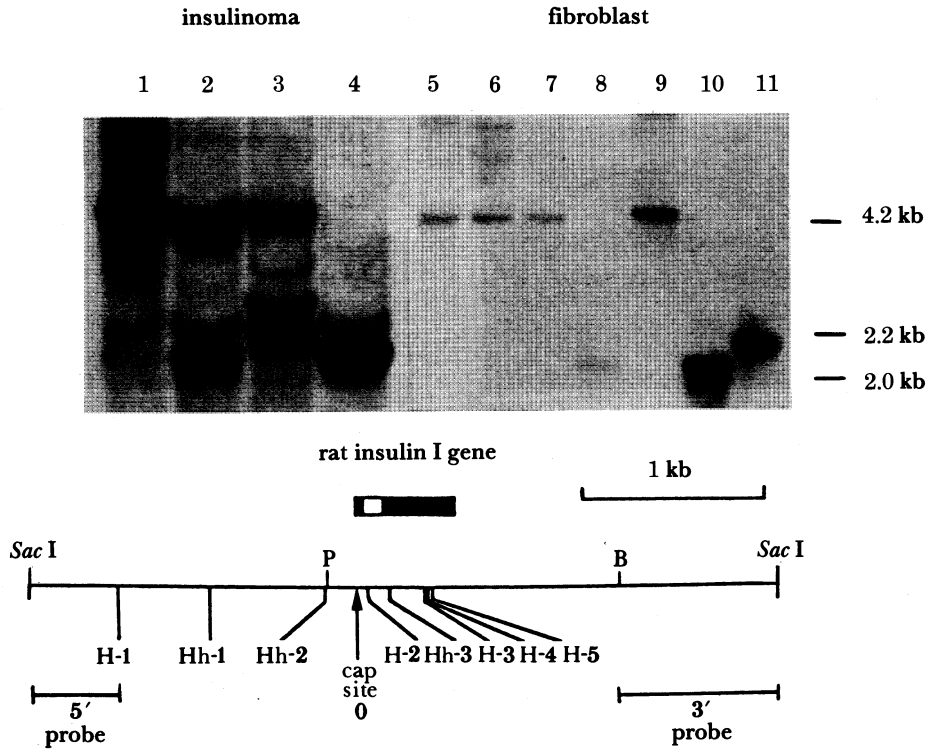


FIGURE 2. Demethylation of the rat insulin I gene. The cloned *Sac*I fragment of the insulin gene was methylated *in vitro* with *Hpa*II and *Hha*I and co-transfected using a neomycin resistance vector into hamster insulinoma cells or fibroblasts (CHO). DNA from pools of colonies were digested with *Sac*I (lanes 1 & 5) and in addition, *Hpa*II (lanes 2 & 6), *Hha*I (lanes 3 & 7) or *Msp*I (lanes 4 & 8). Marker plasmid DNA has also been cleaved with *Sac*I alone (lane 9) or together with *Hpa*II (lane 10) or *Hha*I (lane 11). The resulting blot filters were hybridized with the 3′ probe shown on the map. The accompanying map (Cate *et al.* 1983) also pictures the actual location of the insulin I gene and the positions of the *Hpa*II sites (H) and the *Hha*I sites (Hh). Note that the entire *Sac*I fragment is 4.2 kb in length. Digestion of this inserted DNA with *Hpa*II shows that in the insulinoma, the cluster of sites H3–H5 is largely unmethylated. The blot also shows that *Hha*I sites Hh-2 and Hh-3 are also unmodified. All of these sites remain totally methylated in fibroblast cells.

gene. In striking contrast, all of the sites in the gene remain resistant to both *Hpa*II and *Hha*I digestion on the exogenous DNA transfected into the fibroblast line.

From both of these transfection experiments it is clear that cells contain a tissue-specific mechanism that can recognize a methylated gene and demethylate it so that it resembles the endogenous gene counterpart. Thus the activity responsible for the developmental fixation of the methylation pattern within a given cell type remains active throughout the life of a cell grown in culture.

An overview of the methylation state in any cell type shows that while most of the genome is methylated, all active genes are undermethylated (Naveh-Many & Cedar 1981). A high percentage of these genes have a housekeeping function and are thus active in every cell type. These constitutive genes all contain CpG islands at their 5′ end that are completely unmethylated in every cell and in the germ-line DNA (Stein *et al.* 1983). It was assumed that this pattern is fixed in the genome and that the organism probably lacks the facilities for altering the methylation state of these islands. Indeed, when methylated island sequences were introduced into fibroblasts or myoblasts by DNA-mediated gene transfer, the CpG sites remained fully methylated (Stein *et al.* 1982; Yisraeli *et al.* 1986). To confirm this finding we

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

reintroduced a methylated hamster *APRT* gene to mouse L-cells and did a careful analysis of almost all of the potentially methylated restriction sites. As shown in figure 3, when this transfected DNA is cut with *Pvu*II/*Eco*RI and hybridized to a 5′ *APRT* probe, a 900 base pairs (b.p.) band is visualized. Further cleavage showed that the *Hha*I site in this fragment is almost fully methylated. In fact, all of the other assayable methylation sites also remained in the modified state (unpublished data). Only one *Hpa*II site (marked M on figure 3) underwent partial demethylation. This picture of stable methylation stands in sharp contrast to the active process of demethylation that occurs on tissue-specific sequences.

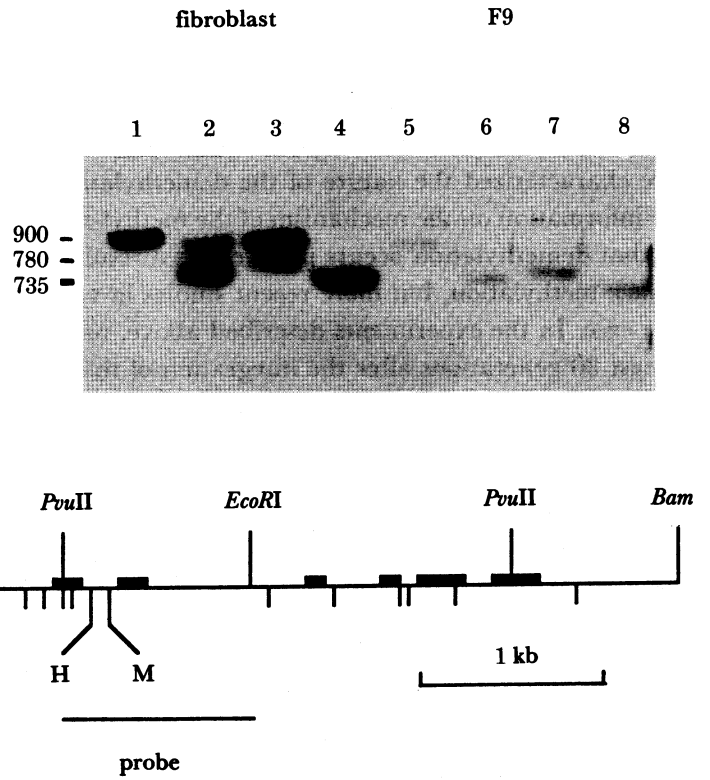


FIGURE 3. Methylation state of the *APRT* gene in transfected cells. The cloned *Bam*HI fragment containing the *APRT* gene (dark regions) was methylated *in vitro* with the *Hpa*II and *Hha*I methylases and transfected into mouse L-cells or F9 teratocarcinoma. DNA from pooled colonies was digested with *Pvu*II/*Eco*RI (lanes 1 & 5) and, in addition, with the enzymes *Hpa*II (lanes 2 & 6), *Hha*I (lanes 3 & 7) or *Msp*I (lanes 4 & 8). After electrophoresis and blotting the filters were hybridized to a nick-translated *Pvu*II/*Eco*RI probe. Shown on the map are the *Hpa*II (M) and the *Hha*I sites (H) that were analysed in this experiment. The other *Hpa*II and *Hha*I sites in this region are indicated by vertical lines, and these were assayed using different blotting strategies (data not shown).

If methylated islands are not recognized by the demethylation machinery, we reasoned that a methylated housekeeping gene introduced into a transgenic mouse would remain methylated and therefore inactive. The *APRT* gene has already been shown to be subject to methylation inhibition in fibroblast cells (Stein *et al.* 1982; Keshet *et al.* 1985). When an *APRT* construct was modified *in vitro* and injected into mouse zygotes, however, we found that the resulting transgene underwent a generalized, organism-wide demethylation at CpG sites restricted to the island region. From this experiment, it can be concluded that mice do have a mechanism for removing methyl groups from CpG islands and we assume that this activity normally occurs

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

in the early embryo. A good model for studying cells during early embryonic development is the F9 teratocarcinoma line, which behaves much like the cell types present just before the blastocyst stage. When a methylated *APRT* gene was inserted into these cells by co-transfection, we found that this island does indeed undergo active demethylation, confirming results from the transgenic animals. Figure 3 quite clearly shows that both the *Hpa*II and *Hha*I sites marked on the map are completely unmethylated, and additional experiments demonstrated that all of the sites in the 5′ end of the gene up to the *Eco*RI site underwent massive demethylation (unpublished data). Sites for *Hpa*II and *Hha*I, which are located 3′ to the *Eco*RI site, and are thus not within the CpG-island region, remained modified; this emphasizes the site-specificity of this enzymic reaction. Thus CpG islands are also subject to demethylation, but only in a stage-specific manner. Theoretically, a CpG island that has become methylated will remain fixed in this state in somatic cells, as they lack the appropriate demethylation activity.

Although we have characterized the nature of the demethylation process for several gene types, this gives little information on the mechanism of the methyl-removal reaction. For a long time it was thought that demethylation occurs passively as a result of DNA replication in the absence of maintenance methylation, but more recent studies in various systems indicate that this might not be the case. In the experiments described above, methylation was evaluated in transfected cells at least 30 generations after the integration of the foreign sequences, but this gives no information on the dynamics of demethylation. To follow the kinetics of this reaction, we introduced methylated sequences into cells in culture by transient transfection, and searched for changes in the modification pattern at 4–72 h (see figure 4). The α-actin gene has already been shown to undergo demethylation in myoblasts *in vivo* at a critical *Sma*I (CCCGGG) site within its promoter region (Yisraeli *et al.* 1986). When this site was assayed in a transient experiment, it remained resistant to both *Sma*I and *Hpa*II (CCGG) digestion, which suggested that it had not undergone demethylation. This unique sequence is also recognized by the restriction enzyme *Nci*I (CCCGG or CCGGG) and methylation at the internal CG residue protects against cleavage. It was therefore surprising that the transfected DNA from a transient experiment of as short as 4 h was cut specifically by *Nci*I. Further analysis showed that this restriction enzyme is unusual in that it digests both unmethylated and hemimethylated DNA, but not fully methylated sites. Thus the use of *Nci*I provides a sensitive assay for methylation changes and shows that demethylation of α-actin begins very quickly after the entrance of the gene into the cell, but that in the first stage of this reaction, the target site is demethylated on one strand only. Further studies with strand-specific probes suggest that this initial step can take place on either DNA strand. The removal of the second methyl group on the complementary strand does not occur within the time span of the transient experiment (72 h) and must require additional conditions that may be related to DNA replication or integration into the genome. It should be noted that there are two potential sites for *Nci*I cleavage within the α-actin promoter region. The band obtained upon *Nci*I digestion corresponds to a demethylation event exclusively at the *Sma*I site and not at the 5′ *Nci* site, once again indicating the exquisite site-specificity of this enzymatic process.

DNA inserted into cells by transient transfection also provided a good substrate for studying the relation between demethylation and replication. The plasmids used in these experiments were all grown in *Escherichia coli* and therefore harbour GATC sites methylated in the A position. Digestion of this exogenous DNA with the enzymes *Mbo*I and *Dpn*I revealed that this

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

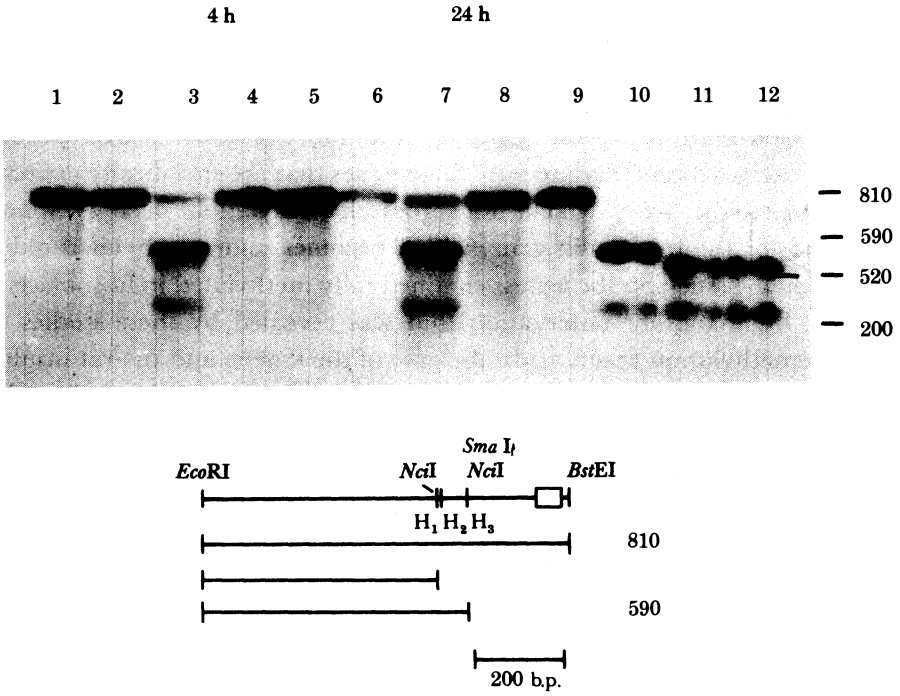


FIGURE 4. α-Actin demethylation following transient tranfection. A plasmid containing the *Eco*RI/*Bst*EII fragment from the promoter region of the rat α-actin gene was methylated *in vitro* with *Hpa*II and *Hha*I and transfected into the rat myoblast line L8 (Yisraeli *et al.* 1986). At either 4 or 24 h after the addition of the DNA, cells were harvested and DNA was prepared from nuclei and digested with *Eco*RI/*Bst*EII (lanes 1 & 5) alone or together with *Sma*I (lanes 2 & 6), *Nci*I (lanes 3 & 7) or *Hpa*II (lanes 4 & 8). Marker plasmid cleaved with the same order of enzymes is shown in lanes 9–12. After blotting, these filters were hybridized with a probe that covers the entire region shown on the map. Note that the *Eco*RI/*Bst*EII produces an 810 b.p. fragment. When this is cut with *Nci*I at site H2, approximately 600 b.p. and 200 b.p. fragments are produced. Cleavage at both H1 and H3 would yield a 540 b.p. fragment together with the same 200 b.p. piece. The map (Yisraeli *et al.* 1986) also shows the position of the first gene exon (open box).

A remains modified, suggesting that these molecules do not undergo any replication during the course of the transfection. Thus the hemidemethylation observed in α-actin must occur by an active enzymatic process, and not passively as a result of replication. This system may be useful for learning additional information on the mechanics of the demethylation process. Preliminary results using a reverse genetic approach clearly indicate that *cis*-acting sequences within the 5′ end of the gene are required to direct demethylation, but it is not yet known whether these are the same elements used for regulating tissue-specific transcription.

3. DISCUSSION

DNA modification plays an important role in regulating gene expression in animal cells. As all tissue-specific genes are fully methylated in most cells of the organism, this modification serves as a general signal at the level of the DNA that marks genes as being inactive without the necessity for a system of *cis*- and *trans*-acting elements to recognize each gene individually. Activation of these gene sequences occurs during development within each cell type and the initial tissue-specific recognition must take place on methylated molecules. It is quite clear that

these events must be mediated through specific proteins and that the process of demethylation is not the primary factor that causes gene activation.

In these studies we have attempted to learn something about this reaction by using DNA-mediated gene transfer to assay the potential of various cell types to carry out demethylation. The most obvious aspect of this process to emerge from our experiments is its specificity. Each cell type has the capacity to demethylate only those genes that are specifically expressed within it. This is true for the κ gene, which undergoes demethylation in B lymphocytes and plasmacytomas and for the rat insulin gene, which becomes undermethylated only in β-islet type cells. Both of these tissue-specific genes remained fully methylated in fibroblasts and in the F9 embryonic cell line. Another observation that was revealed by these studies is the site-specificity of the demethylation reaction. In the case of the κ gene and the rat insulin gene, all demethylation was restricted to the gene region itself and did not spread over into the flanking sequences, which in these experiments was generally plasmid DNA. Even within the insulin gene itself there was a strict specificity with only a certain domain of sites undergoing demethylation. The final pattern of methylation of the transfected sequences always mimics that seen on the endogenous gene in the same tissue, suggesting that both demethylations are carried out by the same enzymic system.

Using the experimental model system, it should now be possible to dissect out the *cis* and *trans* elements, which control the demethylation reaction. We assume that sequences at the 5′ end of the gene are required for demethylation and this can be tested by introducing gene constructs that lack these regions. Application of this reverse genetics approach to the α-actin gene has shown that there is indeed a defined DNA element that is necessary to obtain site-specific demethylation. It seems that the process of demethylation may be very similar to that of gene activation, as both events involve the interaction of tissue-specific *trans*-acting factors with *cis* elements in the gene domain. It is already quite clear that the extreme 5′ insulin enhancer is dispensable for demethylation (unpublished data), despite the fact that it is needed to provide a 40-fold boost in transcriptional levels (Karlsson *et al.* 1987). One possibility is that demethylation may result secondarily from the actual establishment of the transcription complex. In this regard it should be noted that while both transcription and demethylation occur during normal development, the order of these events has not been definitively established in most tissues. In the liver, both transcription and changes in chromosome structure precede the demethylation step (Wilks *et al.* 1982; Vedel *et al.* 1983; Benvenisty *et al.* 1985), and it is likely that a similar sequence of events occurs in other tissues as well.

In our model, methylation serves as a general signal to put DNA sequences into an inactive conformation. Gene activation involves overcoming this inhibition and appears to occur when the gene is still methylated. What then is the role of the demethylation that takes place on tissue-specific genes? One possibility is that activation occurs in two steps. In the first stage, a transcription complex is initiated through the involvement of *trans*-acting factors. In the second stage, demethylation causes a permanent change in the structure of the gene and marks it as being active and in the same class as the cells housekeeping sequences. This second consolidative step may be especially important if the original activation factors are present only transiently and do not remain for the lifetime of the cell type.

In addition to tissue-specific demethylation, we have also observed stage-specific demethylation of CpG islands in housekeeping genes. Our results not only show that these islands undergo demethylation exclusively in early embryo cells, but they also shed some light

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

on the process of island recognition. Using *APRT* as a model system, we were able to show that this gene undergoes demethylation in F9 embryonic cells, but not in other somatic cell types. This demethylation was restricted to the island portion of the gene, whereas other regions remained fully methylated and even underwent *de novo* methylation when introduced into transgenic mice (unpublished data). Thus the cell machinery must have some mechanism for recognizing and defining CpG islands. We first reasoned that transcription may play a role in the demethylation process, but this is unlikely as methyl groups are removed even in a promotorless construct. Sequences such as pBR322, which have a relatively high CpG content, also undergo specific demethylation in F9 cells and this suggests that it is the mere presence of concentrated CpG residues that is recognized by the embryonic cell.

The role of this demethylation was initially puzzling to us, as islands have been found to be uniformly unmethylated in all tissues and in the germ line, and there appeared to be no obvious need for a demethylation reaction. A close evaluation of the cells' CpG islands showed that this type of stage-specific demethylation process may play an important role for several sequence categories. The major function of this activity may be related to X-chromosome reactivation. During the early blastocyst stage of development, one X chromosome in female cells undergoes inactivation, and this is followed by the *de novo* methylation of gene sequences on this chromosome (Lock *et al.* 1987), including a large number of CpG island-containing housekeeping genes (Wolf *et al.* 1984*a*; 1984*b*; Keith *et al.* 1986). When this X is reactivated during oogenesis in the female germ line (Gartler *et al.* 1975; Kratzer & Chapman 1981), there must be a mechanism for island demethylation. The activity we have described would certainly explain the methyl metabolism of the X chromosome, as it is required in the germ line and/or in early development, but must be absent in somatic cells where X methylation maintains the inactive state of this chromosome (Liskay & Evans 1980). This demethylation activity may also play a role in the developmental changes which occur to the methylation pattern of satellite DNA. These sequences are undermethylated in the germ line and early embryo, but fully methylated in somatic cells (Chapman *et al.* 1984; Ponzetto-Zimmerman & Wolgemuth 1984). It should be noted that in the mouse these sequences, which represent 10% of the genome have a high CpG content, containing 10 residues within the 234 nucleotide repeat (Horz & Altenburger 1981) and may make up the most abundant island category in the cell. They also contain about 40% of the cell's methyl moieties (Solage & Cedar 1978). The function of satellite DNA and the role played by methylation is unknown. Other organisms have clear cut mechanisms for physically removing satellite sequences from somatic cells (John & Gabor-Miklos 1979) and DNA methylation may accomplish the functional equivalent of this in animal cells.

While demethylation is a common reaction, which takes place in various cells during development, little is known about its mechanism. As the chemical removal of the methyl moiety from the cytosine base is considered to be thermodynamically impossible (Razin *et al.* 1986), it was originally thought that functional demethylation is a passive process and takes place as a result of DNA replication in the absence of maintenance methylation. Several pieces of evidence now suggest that demethylation is an active process. Most importantly, it was shown that demethylation of the δ-crystallin gene occurs in the absence of any DNA synthesis within post-mitotic lens cells *in vivo* (Sullivan & Grainger 1987). Furthermore, experiments with mouse erythroleukaemia cells demonstrate clearly, that upon induction, the genome undergoes a massive demethylation with rapid kinetics that could not be accounted for by the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

replication model (Razin *et al* 1985). These authors suggested that active demethylation could take place through a base-replacement model similar to the mechanism used to remove unwanted methylated residues from DNA in bacteria and animal cells (Razin *et al.* 1986).

Our results for the α-actin gene in rat myoblasts provide the strongest evidence to date that demethylation occurs by an active enzymatic process. This conclusion is based on the observation that demethylation takes place within a few hours after entry of methylated DNA to the cell and from the fact that these plasmid molecules do not undergo any replication even after several days in culture. These experiments also reveal details about the stages of the demethylation mechanism. The first step after the recognition of the specific α-actin sequences appears to involve the removal of a methyl moiety from only one strand of the DNA. Although complete demethylation was not observed in transient transfection experiments, this second stage probably occurs once the DNA has integrated into the chromosome, because stably transfected molecules show total demethylation at the identical sites (Yisraeli *et al.* 1986). It is certainly possible that in this gene-specific system, the removal of the methyl group from the second strand requires replication, but we have no proof for this hypothesis. Using genomic sequencing it has been shown that the activated vitellogenin gene undergoes a similar two-step demethylation process in embryonic liver cells (Saluz *et al.* 1986). Unlike the hemi-demethylation observed for the α-actin gene, vitellogenin demethylation is programmed such that it always begins in one specific strand of the DNA.

It should be noted that all of the demethylations observed in this study are sequence-specific events. Although the *cis*-acting elements controlling demethylation have not been worked out, it is clear that these are present and that it is the recognition of such regions which provide for proper recognition and initiation of the demethylation reaction. This process may be closely coupled to transcription, and there is a good possibility that the same *cis*-acting sites are required for setting up the transcription complex and for carrying out demethylation. Since the chemical reaction of demethylation is probably identical for all genes, we favour a model in which the specificity of the demethylation is established by proteins that interact with the gene sequences, but the actual reaction is done by a non-specific enzyme complex. Using the experimental systems described in this manuscript, it should be possible to dissect the sequences involved in forming the template for demethylation and this should bring us closer to understanding the entire process.

The authors thank C. Gopin for her assistance in preparing this manuscript. This research was supported by funds from the National Institutes of Health and the Israel Cancer Research Fund (to H.C.), and the U.S.–Israel Binational Science Foundation (to H.C. and to Y.B.).

References

Becker, P. B., Ruppert, S. & Schutz, G. 1987 Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. *Cell* **51**, 435–443.

Benvenisty, N., Mencher, D., Meyuhas, O., Razin, A. & Reshef, L. 1984 Sequential changes in DNA methylation patterns of the rat phosphoenol pyruvate carboxykinase gene during development. *Proc. natn. Acad. Sci. U.S.A.* **82**, 267–271.

Bird, A., Taggart, M., Frommer, M., Miller, O. J. & Macleod, D. 1985 A fraction of the mouse genome that is derived from islands of non-methylated, CpG-rich DNA. *Cell* **40**, 91–99.

Busslinger, M., Hurst, J. & Flavell, R. A. 1983 DNA methylation and the regulation of globin gene expression. *Cell* **34**, 197–206.

Cate, R. L., Chick, W. & Gilbert, W. 1983 Comparison of methylation patterns of the two rat insulin genes. *J. biol. Chem.* **258**, 6645–6652.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Chapman, V., Forrester, L., Sanford, J., Hastie, N. & Rossant, J. 1984 Cell lineage-specific undermethylation of mouse repetative DNA. *Nature, Lond.* **307**, 284–286.

Gartler, S. M., Andina, R. J. & Grant, N. 1975 Ontogeny of X-chromosome inactivation in the female germ line. *Expl Cell Res.* **91**, 454–457.

Horz, W. & Altenburger, W. 1981 Nucleotide sequence of mouse satellite DNA. *Nucl. Acids Res.* **9**, 683–696.

John, B. & Gabor-Miklos, G. L. 1979 Functional aspects of satellite DNA and heterochromatin. *Int. Rev. Cytol.* **58**, 1–53.

Karlsson, O., Edlund, T., Barnett Moss, J., Rutter, W. J. & Walker, M. D. 1987 A mutational analysis of the insulin gene transcription control region: Expression in beta cells is dependent on two related sequences within the enhancer. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8819–8823.

Keith, D. H., Singer-Sam, J. & Riggs, A. D. 1986 Active X chromosome DNA is unmethylated at eight CCGG sites clustered in a guanine-plus-cytosine-rich island at the 5′ end of the gene for phosphoglycerate kinase. *Molec. Cell Biol.* **6**, 4122–4125.

Kelley, D. E., Pollok, B. A., Atchison, M. L. & Perry, R. P. 1988 The coupling between enhancer activity and hypomethylation of κ immunoglobulin genes is developmentally regulated. *Molec. Cell Biol.* **8**, 930–937.

Keshet, I., Lieman-Hurwitz, J. & Cedar, H. 1986 DNA methylation affects the formation of active chromatin. *Cell* **44**, 535–543.

Kratzer, P. G. & Chapman, V. M. 1981 X chromosome reactivation in oocytes of *Mus caroli*. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3093–3097.

Liskay, R. M. & Evans, R. J. 1980 Inactive X chromosome does not function in DNA-mediated cell transformation for the hypoxanthine phosphoribosyl transferase gene. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4895–4899.

Lock, L. F., Takagi, N. & Martin, G. R. 1987 Methylation of the *Hprt* gene on the inactive X occurs after chromosome inactivation. *Cell* **48**, 39–46.

Mather, L. & Perry, R. P. 1983 Methylation status and DNaseI sensitivity of immunoglobulin genes: changes associated with rearrangement. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4689–4693.

Naveh-Many, T. & Cedar, H. 1981 Active gene sequences are undermethylated. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4246–4250.

Ponzetto-Zimmerman, C. & Wolgemuth, D. J. 1984 Methylation of satellite sequences in mouse spermatogenic and somatic DNAs. *Nucl. Acids Res.* **12**, 2807–2821.

Razin, A., Feldmesser, E., Kafri, T. & Szyf, M. 1985 Cell-specific DNA methylation patterns: formation and a nucleosomal locking model for their function. In *Biochemistry and biology of DNA methylation* (ed. G. L. Cantoni & A. Razin), pp. 239–253. New York: Alan R. Liss Inc.

Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Scarpa, S., Carotti, D. & Cantoni, G. L. 1986 Replacement of 5-methylcytosine – a possible mechanism for transient demethylation during differentiation. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2827–2831.

Saluz, H. P., Jiricny, J. & Jost, J. P. 1986 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. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7167–7171.

Solage, A. & Cedar, H. 1978 The organization of 5-methylcytosine in chromosomal DNA. *Biochemistry* **17**, 2934–2938.

Stein, R., Razin, A. & Cedar, H. 1982 *In vitro* methylation of the hamster *APRT* gene inhibits its expression in mouse L-cells. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3418–3422.

Stein, R., Sciaky-Gallili, N., Razin, A. & Cedar, H. 1983 The pattern of methylation of two genes coding for housekeeping functions. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2422–2426.

Sullivan, C. H. & Grainger, R. M. 1987 δ-crystallin genes become hypomethylated in post-mitotic lens cells during chicken development. *Proc. natn. Acad. Sci. U.S.A.* **84**, 329–333.

Vedel, M., Gomez-Garcia, M., Sala, M. & Sala-Trepat, J. M. 1983 Changes in methylation pattern of albumin and α-fetoprotein genes in developing rat liver and neoplasia. *Nucl. Acids Res.* **11**, 4335–4354.

Walker, M. D., Edlund, T., Boulet, A. M. & Rutter, W. J. 1983 Cell-specific expression controlled by the 5′ flanking region of insulin and chymotrypsin genes. *Nature, Lond.* **306**, 557–561.

Wilks, A. F., Cozens, P. J., Mattaj, J. W. & Jost, J. P. 1982 Estrogen induces a demethylation at the 5′ end region of the chicken vitellogenin gene. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4252–4255.

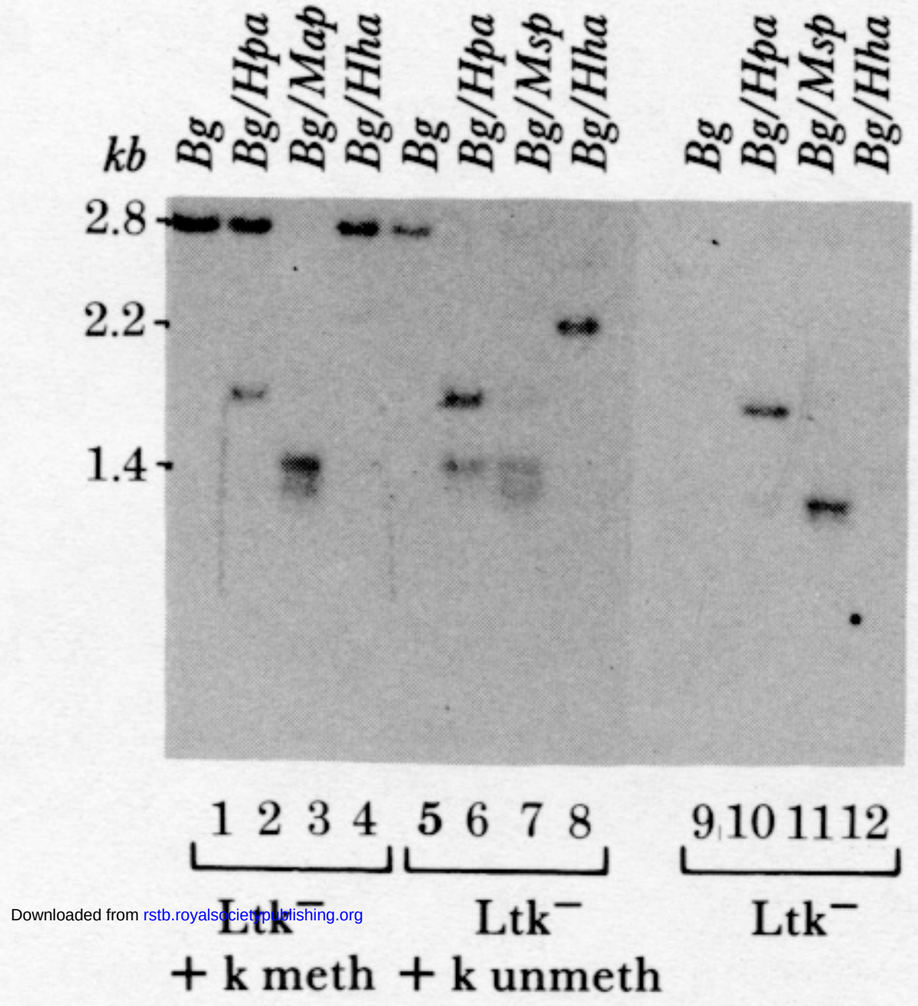
Wolf, S. F., Dintzis, S., Toniolo, D., Persico, G., Lunnen, K. D., Axelman, J. & Migeon, B. R. 1984*a* Complete concordance between glucose-6-phosphate dehydrogenase activity and hypomethylation of 3′ CpG clusters: implication for X-chromosome dosage compensation. *Nucl. Acids Res.* **12**, 9333–9348

Wolf, S.F., Jolly, D. J., Lunnen, K. D., Friedmann, T. & Migeon, B. R. 1984*b* Methylation of the hypoxanthine phosphoribosyl transferase locus on the human X chromosome: implications for X-chromosome inactivation. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2806–2811.

Yisraeli, J., Adelstein, R., Melloul, D., Nudel, U., Yaffe, D. & Cedar, H. 1986 Muscle-specific activation of a methylated chimeric actin gene. *Cell* **46**, 409–416.

Yisraeli, J. & Szyf, M. 1984 Gene methylation patterns and expression. In *DNA methylation and its biological significance* (ed. A. Razin, H. Cedar & A. D. Riggs), pp. 353–378. New York: Springer-Verlag.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES



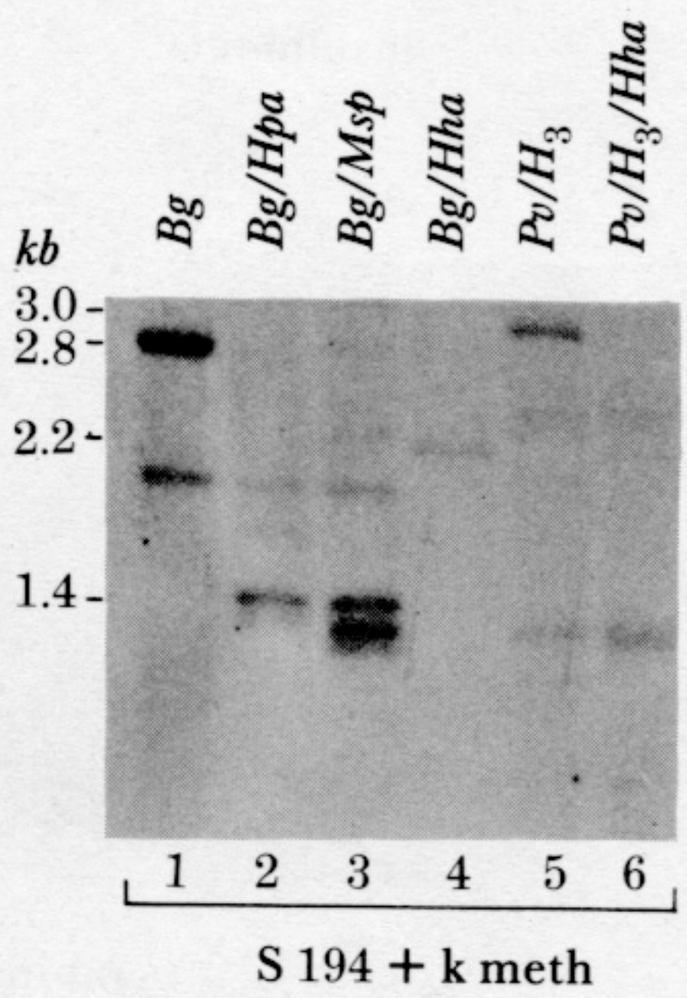
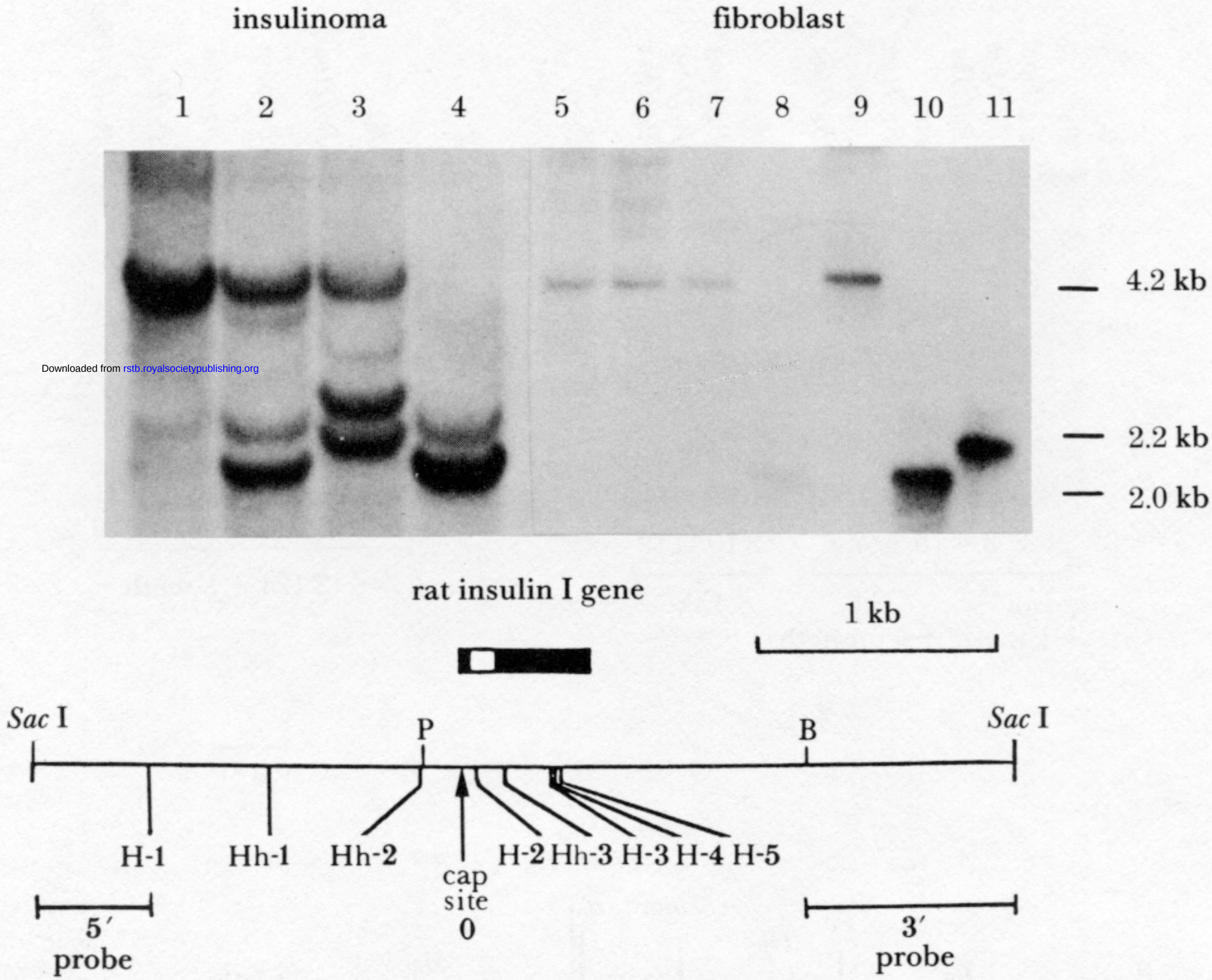
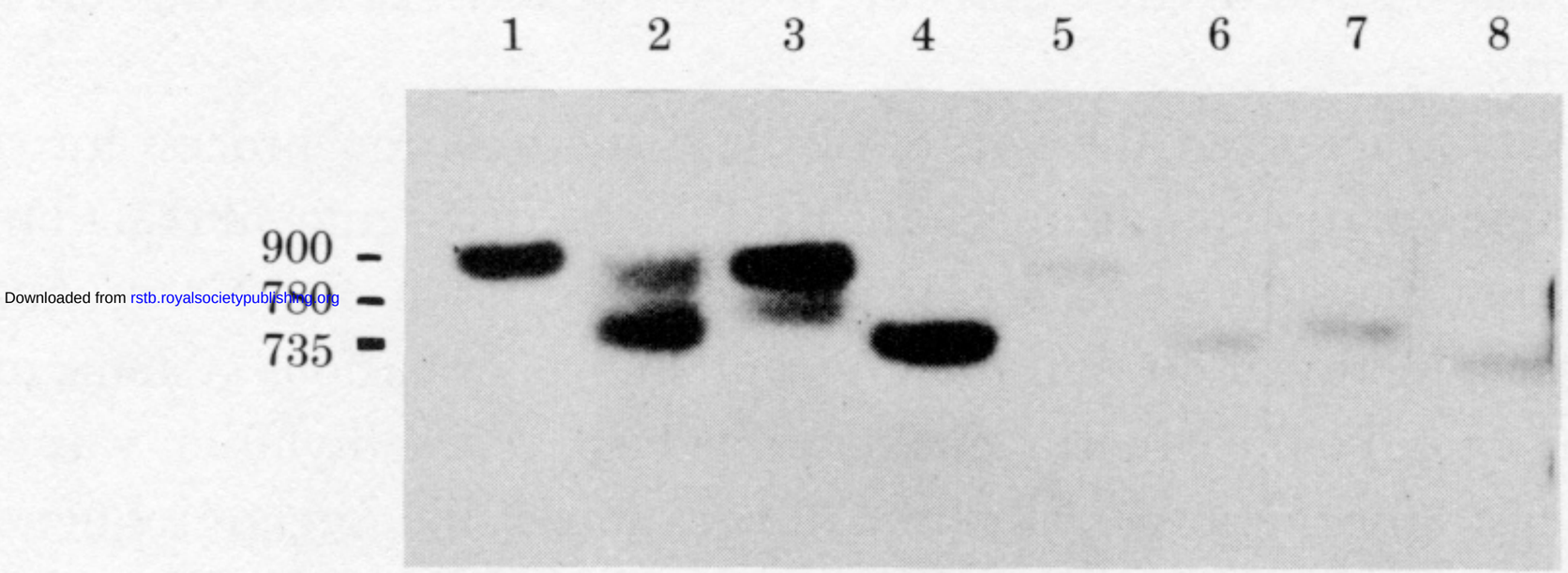
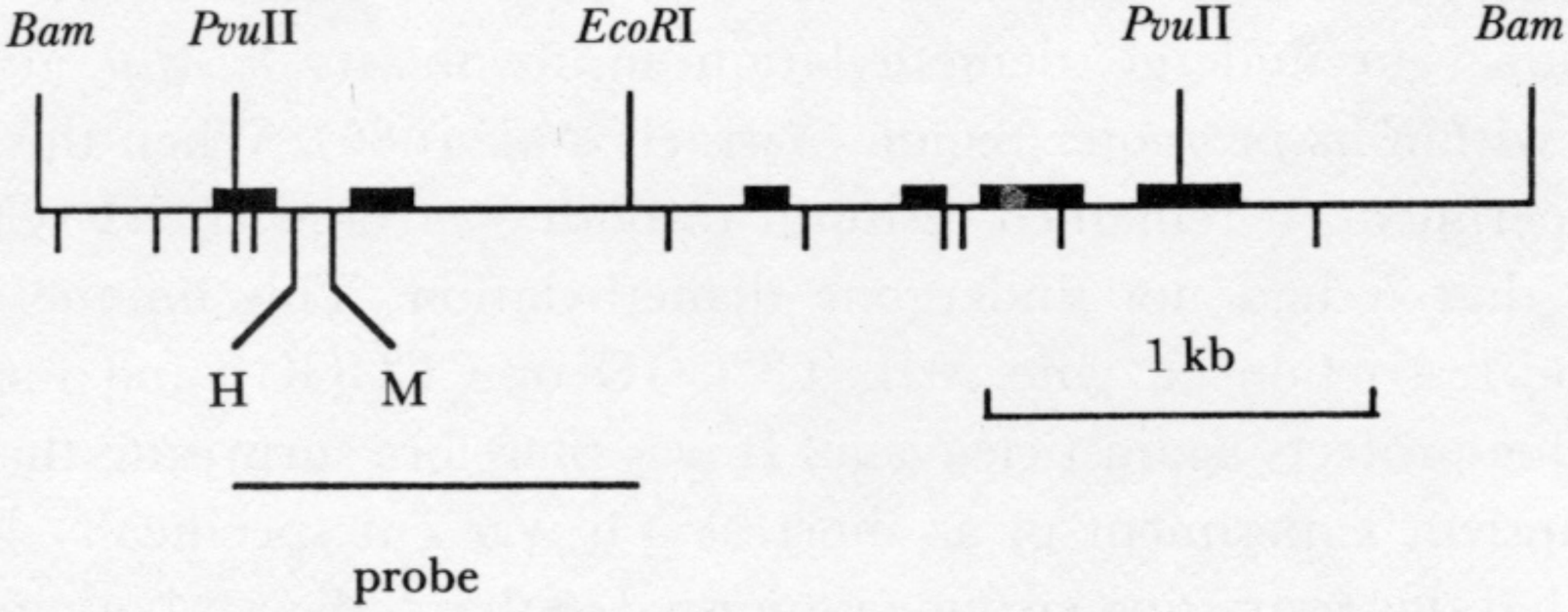


FIGURE 1. Methylation analysis of Ig κ following its transfection to cells. A plasmid (Kelley *et al.* 1988) containing the $V_{\kappa}19J_2C_{\kappa}$ was methylated *in vitro* (Yisraeli *et al.* 1986) using *Hpa*II and *Hha*I methylases and co-transfected into cell line S194 by electroporation (Kelley *et al.* 1988) or into mouse L-cells either by the calcium phosphate precipitation technique or by electroporation. In both cases, colonies containing the plasmid were isolated by selection for neomycin and these were pooled and expanded and then used to prepare DNA for Southern blot analysis. DNA was digested with *Bgl*II and in addition, with either *Hpa*II, *Msp*I or *Hha*I in order to assay methylated sites near the V region. In the case of S194, the 3′ *Hha*I site was analysed by digestion with *Pvu*II/*Hin* dIII. In all cases, 20 μg of DNA were loaded on each lane and the final filters were probed using a nick-translated *Pvu*II/*Pst*I fragment. In the case of the L-cell transfection, gel analysis is shown for both the methylated Ig κ construct as well as for a non-methylated control. In addition to the expected 2.8 kb band and its cleavage products, these autoradiograms also show hybridization to the endogenous equivalent of the Ig κ gene, as is demonstrated by the blot of the untransfected Ltk^- control.

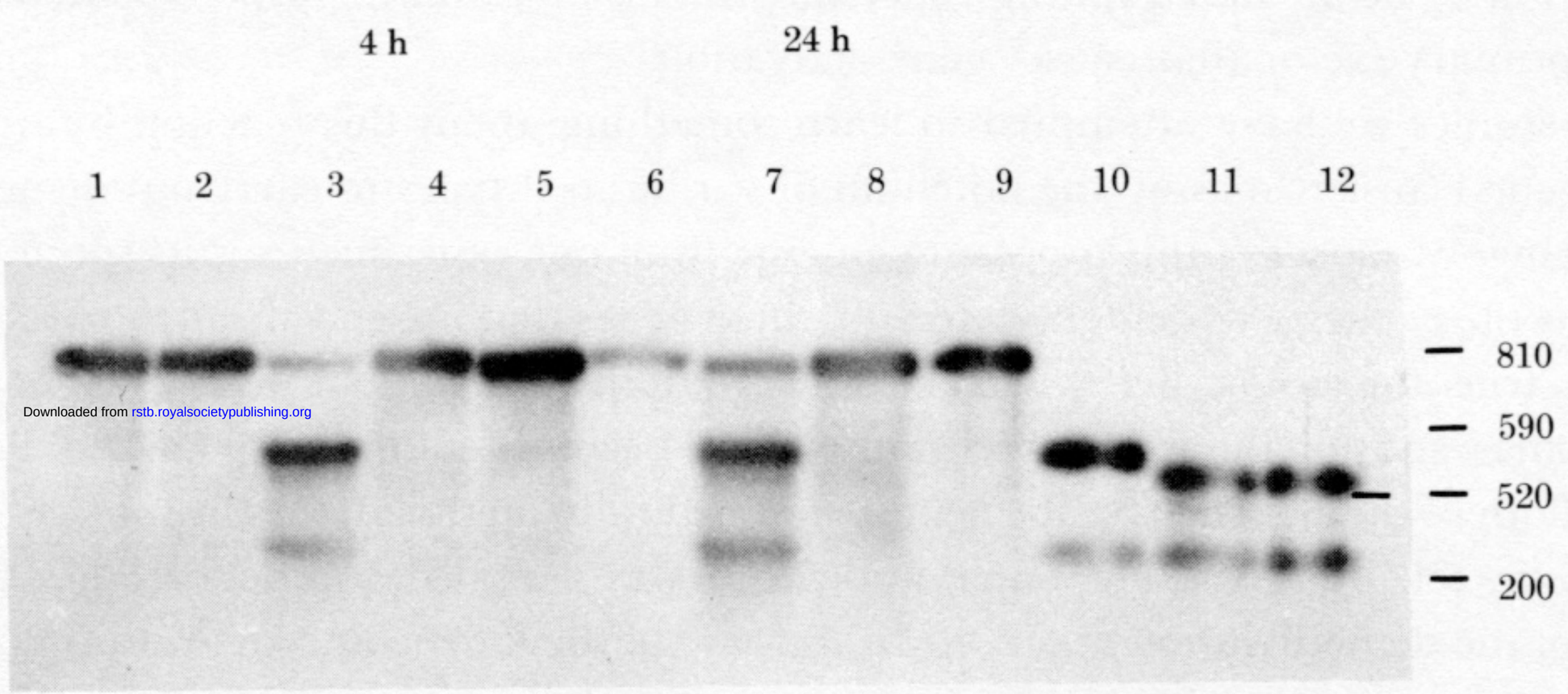


IGURE 2. Demethylation of the rat insulin I gene. The cloned *Sac*I fragment of the insulin gene was methylated *in vitro* with *Hpa*II and *Hha*I and co-transfected using a neomycin resistance vector into hamster insulinoma cells or fibroblasts (CHO). DNA from pools of colonies were digested with *Sac*I (lanes 1 & 5) and in addition, *Hpa*II (lanes 2 & 6), *Hha*I (lanes 3 & 7) or *Msp*I (lanes 4 & 8). Marker plasmid DNA has also been cleaved with *Sac*I alone (lane 9) or together with *Hpa*II (lane 10) or *Hha*I (lane 11). The resulting blot filters were hybridized with the 3′ probe shown on the map. The accompanying map (Cate *et al.* 1983) also pictures the actual location of the insulin I gene and the positions of the *Hpa*II sites (H) and the *Hha*I sites (Hh). Note that the entire *Sac*I fragment is 4.2 kb in length. Digestion of this inserted DNA with *Hpa*II shows that in the insulinoma, the cluster of sites H3–H5 is largely unmethylated. The blot also shows that *Hha*I sites Hh-2 and Hh-3 are also unmodified. All of these sites remain totally methylated in fibroblast cells.





IGURE 3. Methylation state of the *APRT* gene in transfected cells. The cloned *Bam*HI fragment containing the *APRT* gene (dark regions) was methylated *in vitro* with the *Hpa*II and *Hha*I methylases and transfected into mouse L-cells or F9 teratocarcinoma. DNA from pooled colonies was digested with *Pvu*II/*Eco*RI (lanes 1 & 5) and, in addition, with the enzymes *Hpa*II (lanes 2 & 6), *Hha*I (lanes 3 & 7) or *Msp*I (lanes 4 & 8). After electrophoresis and blotting the filters were hybridized to a nick-translated *Pvu*II/*Eco*RI probe. Shown on the map are the *Hpa*II (M) and the *Hha*I sites (H) that were analysed in this experiment. The other *Hpa*II and *Hha*I sites in this region are indicated by vertical lines, and these were assayed using different blotting strategies (data not shown).



IGURE 4. α-Actin demethylation following transient tranfection. A plasmid containing the *Eco*RI/*Bst*EII fragment from the promoter region of the rat α-actin gene was methylated *in vitro* with *Hpa*II and *Hha*I and transfected into the rat myoblast line L8 (Yisraeli *et al.* 1986). At either 4 or 24 h after the addition of the DNA, cells were harvested and DNA was prepared from nuclei and digested with *Eco*RI/*Bst*EII (lanes 1 & 5) alone or together with *Sma*I (lanes 2 & 6), *Nci*I (lanes 3 & 7) or *Hpa*II (lanes 4 & 8). Marker plasmid cleaved with the same order of enzymes is shown in lanes 9–12. After blotting, these filters were hybridized with a probe that covers the entire region shown on the map. Note that the *Eco*RI/*Bst*EII produces an 810 b.p. fragment. When this is cut with *Nci*I at site H2, approximately 600 b.p. and 200 b.p. fragments are produced. Cleavage at both H1 and H3 would yield a 540 b.p. fragment together with the same 200 b.p. piece. The map (Yisraeli *et al.* 1986) also shows the position of the first gene exon (open box).