Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA

Richard R.Meehan*, Joe D.Lewis+ and Adrian P.Bird
Institute of Cell and Molecular Biology, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JR, UK

ABSTRACT

Methylated DNA in vertebrates is associated with transcriptional repression and inactive chromatin. Two activities have been identified, MeCP1 and MeCP2, which bind specifically to DNA containing methyl-CpG pairs. In this report we characterize MeCP2. We show that it is more abundant than MeCP1, is more tightly bound in the nucleus, and is distinguishable chromatographically. The two proteins share widespread expression in somatic mammalian cells, and barely detectable expression in early embryonic cells. DNAs containing thymidine which has a methyl group at position 5 are not ligands for the MeCPs. The possible role of MeCP2 in methylation-associated gene inactivation was tested in *in vitro* transcription extracts. Purified MeCP2 inhibited transcription from both methylated and nonmethylated DNA templates *in vitro*, probably due to the presence of nonspecific DNA binding domains within the protein. We hypothesise that MeCP2 normally binds methylated DNA in the context of chromatin, contributing to the long-term repression and nuclease-resistance of methyl-CpGs.

INTRODUCTION

The major modification of vertebrate DNA occurs at the dinucleotide CpG. Approximately 60%–90% of CpGs are methylated at the 5 position of cytosine while the rest are unmodified. A general consensus view derived from analyses of tissue specific genes and transfection studies is that methylated DNA affects the formation of nuclease-resistant chromatin and leads to repression of gene activity (1,2,3,4). In native chromatin it has been demonstrated that m^5C is refractory to digestion by micrococcal nuclease (MNase) and to nucleases that can cleave at CpG (5,6). One model which can accomodate both gene repression and altered chromatin proposes that there are factors in the nucleus which bind differently to methylated and to non-methylated DNA. These factors might bind to a methylated gene leading to an altered chromatin structure which would in turn deny access to the transcription machinery (6).

In support of this model we have identified two DNA binding proteins which have no specific sequence requirements for binding DNA other than the presence of a number of symmetrically methylated CpGs in the test molecule (7,8). These proteins are the Methyl CpG binding Proteins, MeCP1 and MeCP2. Neither hemi-methylated DNA molecules nor DNA molecules containing m^5C in a non-CpG context are good substrates for MeCP1 and MeCP2. Two other proteins that bind to methylated DNA have also been identified, but they differ from MeCP1 and -2 in that methylcytosine must be part of a specific recognition sequence for binding to occur (9,10). In this paper we demonstrate that MeCP1 and MeCP2 are two distinct activities, and show biochemical evidence that MeCP2 is associated with chromatin. We show that, unlike MeCP1, MeCP2 cannot preferentially repress transcription from methylated DNA *in vitro*. We also report on the steady state levels of MeCP2 in different tissues and species.

MATERIALS AND METHODS

Preparation of nuclei

Nuclei from animal tissues were isolated essentially as described by Meehan et al (7). Nuclei from chick blood and cell lines were prepared as follows. Samples were spun down and washed twice in cold PBS. Cells were resuspended in lysis buffer, 1ml packed cells/6ml lysis buffer (10mM Tris.HCl [pH7.5], 10mM NaCl and 2mM MgCl$_2$). Cells were lysed by the addition of 1% NP-40 at about 25–100μl per ml, the amount being calibrated for each sample. The samples were left on ice for 3–5 minutes and monitored for cell lysis by phase-contrast microscopy. Nuclei were recovered by layering 2–4ml of the homogenate onto a 5ml cushion of 1.2M sucrose in lysis buffer in a 15 ml corex tube. This was spun for 2 minutes at 5,000 rpm in a JS 13.1 rotor. The nuclear pellet was resuspended in nuclear freezing buffer (40% glycerol, 100mM Tris.HCL [pH8.3], 10mM MgCl$_2$ and 0.2mM EDTA). The nuclei were stored at −80°C.

Bandshift and southwestern assays

The bandshift assay was performed as described (7). The DNA fragment used was CG11 in either its methylated or non-methylated form. CG11 is 135 bp long and has 27 methylatable CpGs (7). Assays used either 8μg of protein from a 0.5M rat brain nuclear extract or 6μl of each fraction from the CM

sepharose column (see below) in the presence of 1μg of sonicated *E.coli* competitor DNA. Protein/DNA complexes were detected by running the samples in a 1.5% agarose 0.5× TAE gel as described (7).

The detection by southwestern of methylation dependent DNA binding proteins was also as described (8). The probes used are described in figure legends.

Preparation of 0.55M NaCl extract from rat brain nuclei

Rat brain nuclei equivalent to 15mg of DNA were spun down and resuspended in 15 ml of 0.25M NaCl extraction buffer (0.25M NaCl, 20mM HEPES [pH7.9], 5mM EDTA, 20% glycerol and 10mM β-mercaptoethanol). The volume was then increased to 20ml and at the same time the NaCl salt concentration was increased to 0.55 molar. Protein was extracted from the nuclei by stirring for 30 minutes at 4°C. The extract was then decanted into two ultracentrifuge tubes and spun at 24,000 rpm for 40 minutes in an SW40 rotor. The supernatant was aspirated off and the protein yield determined to be 12.5mg. Histone H1 was removed by selective ammonium sulphate precipitation. Saturation to 50% with ammonium sulphate leaves histone H1 in solution. After neutralization with NaOH (40 μl of 1N NaOH per gramme of ammonium sulphate added) the protein precipitate was allowed to form overnight with gentle stirring and was collected by centrifugation at 16,000 rpm in an SS34 rotor at 0°C for 30 minutes. The pellet was resuspended in 1 ml of resupension buffer. It was then dialyzed against two changes of resuspension buffer (50mM NaCl, 20mM Tris.HCl [pH7.8], 0.1mM EDTA, 10% glycerol, 0.1% Triton X-100 and 10mM β-mercaptoethanol). The protein yield was determined to be 5.2mg at this step by the Bradford method (11).

CM sepharose chromatography

Approximately 4mg of the 0.55M NaCl extract from rat brain nuclei was loaded onto a 0.5ml CM sepharose column in buffer A (50mM Tris.HCl [pH7.8], 0.1mM EDTA, 10% glycerol, 0.1% Triton X-100 and 10mM β-mercaptoethanol) in the presence of 50mM NaCl. Bound proteins were eluted in 2ml steps of buffer A at the following NaCl salt concentrations in buffer A; 100mM, 300mM, 600mM and 1M. 0.5ml fractions were collected and assayed for MeCP1 and MeCP2 activities.

Salt extraction of MeCP2 from rat brain nuclei

Salt extracts were made from rat brain nuclei by extracting equal aliquots of nuclei (equivalent to 150μg of DNA) with different concentrations of NaCl ranging from 0.2M to 0.45M in 0.05M steps. Extraction buffer was 20mM HEPES [pH7.9], 0.1mM EDTA, 10% glycerol and 10mM β-mercaptoethanol. Extracts were left on ice for 1 hour and then spun at 14,000 rpm in an Eppendorf centrifuge for 10 minutes at 4°C. The supernatant was removed and the protein was precipitated by the addition of TCA to 10%. After washing in acetone/0.1 N HCl and acetone the protein was resupended in 200 μl of protein gel loading buffer. The nuclear pellets were resupended in 0.25M NaCl extraction buffer in the presence of 10mM CaCl₂ and 10mM MgCl₂. 15 units of MNase were added and the pellets incubated at room temperature for 10 minutes. The proteins were precipitated by the addition of TCA as before and resuspended in 200 μl of protein loading buffer. Equal volumes of the nuclear salt extract and its corresponding pellet were then analysed for MeCP2 activity by southwestern.

MNase treatment of nuclei

Rat brain nuclei equivalent to 40μg of DNA were suspended in 60μl of solution C (300mM sucrose, 50mM TEA, 25mM KCl, 4mM MgCl₂ and 1mM CaCl₂). Different aliquots were either treated with 40μg of RNase A or increasing amounts of MNase at the following concentrations 0.22 units, 0.44 units, 0.88 units, 4.4 units and 44 units of MNase per 60μl at room temperature for 15 minutes. Then 5μl was taken for DNA analysis and 5μl was taken for southwestern analysis. DNA was extracted, run on a 1.5% TAE agarose gel and stained with ethidium bromide. After MNase treatment, EDTA was added to 20mM and the nuclei put on ice for 15 minutes. The nuclei were then spun at 14,000 rpm in an Eppendorf centrifuge for 5 minutes at 4°C. The supernatant was removed and the protein precipitated with 10% TCA as above. The protein from each aliquot was resuspended in 100 μl of protein loading buffer and analysed by southwestern blotting for MeCP2 release.

Purification of rat MeCP2

MeCP2 was purified from either rat kidney or brain nuclei according to previously published methods (8).

Transcription extract preparation and assay

Rat liver nuclear extracts were prepared according to the method of Dignam (12) except that after ammonium sulphate precipitation the protein pellets were resuspended and dialysed in buffers containing 0.1% Triton X-100. *In vitro* transcription reactions were carried out as described (4), the total protein concentration was 30μg and contained 200ng of the human α-globin 2 template (4), 50 ng of the control adenovirus major late promoter and 2μg of *E.coli* competitor DNA. The reactions were incubated at 30°C for 1 hour and then terminated by the addition of RNase-free DNase for 15 minutes followed by proteinase K treatment, phenol-chloroform extraction and ethanol precipitation. The RNA products were mapped by quantitative S1 analysis as described (4) using the −89 to +155 region of the human α-globin 2 gene and the −253 to +197 region of the adenovirus major late promoter. The human α-globin 2 construct was methylated at all CpGs using the Sss I methyltransferase (Biolabs) as recommended by the supplier.

Tissue, species and embryo cell line southwestern blot

The mouse embryonic cell line, D2, was kindly provided by Erwin Wagner (IMP, Vienna). Nuclei equivalent to 50μg of DNA were extracted with 0.25M NaCl extraction buffer as above and this extract was used for testing for MeCP1 activity by the bandshift assay. The nuclear pellet was resuspended in 200μl of 2M NaCl/5M urea/0.1% Triton X-100/10mM phosphate buffer pH 6.8 and the DNA pelleted by spinning at 14,000 rpm in an Eppendorf centrifuge for 15 minutes at 4°C. 50μl of the supernatant was then used for southwestern analysis.

RESULTS

Tissue and species distribution of MeCP2

MeCP2 was first detected by the southwestern assay (8). For this assay, nuclear proteins were fractionated by SDS polyacrylamide gel electrophoresis, transfered to nitrocellulose membranes and challenged with radioactively labeled methylated and non-methylated DNA in the presence of non-methylated competitor DNA. A rat protein, MeCP2, with a relative
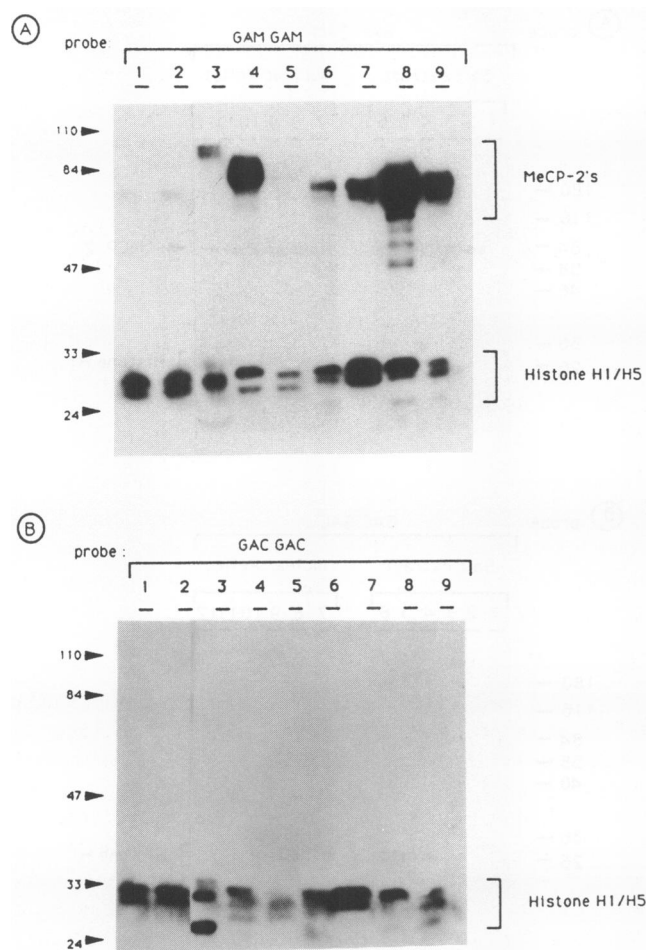
Figure 1. Detection of MeCP2 by Southwestern in rodent and avian extracts. Nuclei were isolated from different sources and subjected to a low salt extraction (0.25M NaCl) followed by a high salt extraction (2M NaCl/5M urea). The high salt extract was used in the southwestern analysis. Panel A shows the methylated DNA probe poly(GAM) and panel B the non-methylated DNA probe poly(GAC). MeCP2 and histone H1/H5 are indicated by brackets. Protein size markers are indicated by the numbers on the left and are expressed in kiloDaltons. The nuclei used were from the following sources; lane 1, PC13 a mouse embryonal carcinoma cell line; lane 2, D2 an SV40 transformed mouse embryo stem cell line; lane 3, chick erythrocytes; lane 4, rat brain; lane 5, rat testes; lane 6, mouse liver; lane 7, mouse spleen; lane 8 mouse brain; lane 9, mouse kidney.

molecular weight of 84kD was able to bind a DNA molecule with as few as one methyl-CpG pair. Figure 1A demonstrates that MeCP2-like proteins and histone H1 variants can be detected in rat and mouse tissues (figure 1A, lanes 4−9) with poly(GAM), whereas poly(GAC) only detects histone H1 like proteins (see later). However, there is a relative size difference between rat MeCP2 with an Mr of 84kD (lanes 4 and 5) and mouse MeCP2 (lanes 6−9) with an Mr of 80kd. There is also significant variation in the amount of MeCP2 in different tissues. For example brain nuclear extracts from rat and mouse have the highest levels (figure 1A, lanes 4 and 8) whereas rat testis has the lowest level (figure 1A, lane 5). Other somatic tissues have intermediate amounts of MeCP2 in their nuclei, there being a decreasing order of expression levels from brain through kidney, spleen, liver and testis of mouse. This pattern of expression is also seen in rat tissues (figure 1A, lanes 4 and 5; and data not

shown). The mouse embryonal stem cell line D2, and embryonal carcinoma cell lines PC13 and F9, have very reduced levels of MeCP2 (figure 1A, lanes 1 and 2, and data not shown). Interestingly these cell lines are also deficient in MeCP1 activity (7).

We assayed nuclei from other species for MeCP2-like proteins. Figure 1A (lane 3) shows that chicken erythrocytes contain MeCP2-like proteins which appear as a doublet with an approximate Mr of 95kD. There was no difference by southwestern blotting between MeCP2 from chicken erythrocyte, liver or brain nuclei (data not shown). MeCP2 was also detected in southwestern blots of nuclear extracts from pig brain, cow kidney and rabbit liver (data not shown). We also detected MeCP2 like activity in a number of permanent cell lines from rat (rat 1A cell line), mouse (NIH3T3 and L cell lines) and human (HeLa, K652 and PES; data not shown). No MeCP2 activity was observed in a Drosophila cell line (data not shown), in line with the absence of detectable CpG methylation in this organism.

MeCP2 can be distinguished from MeCP1

MeCP1 was originally identified as a nuclear factor that can discriminate between methylated and non-methylated DNA using the band shift assay (7). Radioactive DNA molecules that contained 12 or more symmetrically methylated CpG pairs specifically complexed with MeCP1 leading to a shift in electrophoretic mobility. Cross-linking experiments using BrdU-substituted probes showed that a protein of Mr120kD was bound to the methylated DNA. Although different in size from MeCP1 (84kD compared to 120kD), it was unclear whether MeCP2 was a component of the MeCP1 band shift or an entirely different protein. To address this question we first fractionated a rat brain extract containing both the MeCP1 and MeCP2 activities on a CM sepharose column. Eluted fractions were analysed for MeCP1 and MeCP2 using their respective assays. Figure 2B shows that MeCP1 is found in the flow-through fraction from the column (figure 2B, lanes 1−3). In contrast MeCP2 activity is retained on the column and elutes between 100 and 300mM NaCl (figure 2A, lanes 10−11). This shows that MeCP1 and MeCP2 are chromatographically separable and may therefore be two distinct activities.

We noted that the eluted fractions containing MeCP2 gave rise to new complexes in the band shift assay whose mobility differed from that of the MeCP1 complex (figure 2B, lanes 10−11). This result suggests an explanation for our initial failure to detect MeCP2 by the band shift assay: MeCP1 may have a greater affinity for the methylated probe than MeCP2, and may therefore mask the MeCP2 bandshift unless the two proteins are separated. While MeCP2 can be detected by both assays, MeCP1 has not been detected using the southwestern assay. It may be that the 120 kDa protein is unable to renature on the filter, or that the activity depends on the interaction of several different proteins that become separated on a denaturing gel.

Salt extraction of MeCP2 from nuclei

Increasing salt concentration can be used to selectively solubilise nuclear proteins. For example 0.5M NaCl completely removes histone H1 from mammalian chromatin whilst leaving the core histones in place (13). In order to determine how tightly bound MeCP2 was in rat brain nuclei, equal aliquots of rat brain nuclei were incubated with concentrations of NaCl ranging from 0.2M to 0.45M. After incubation, the nuclei were pelleted and the salt
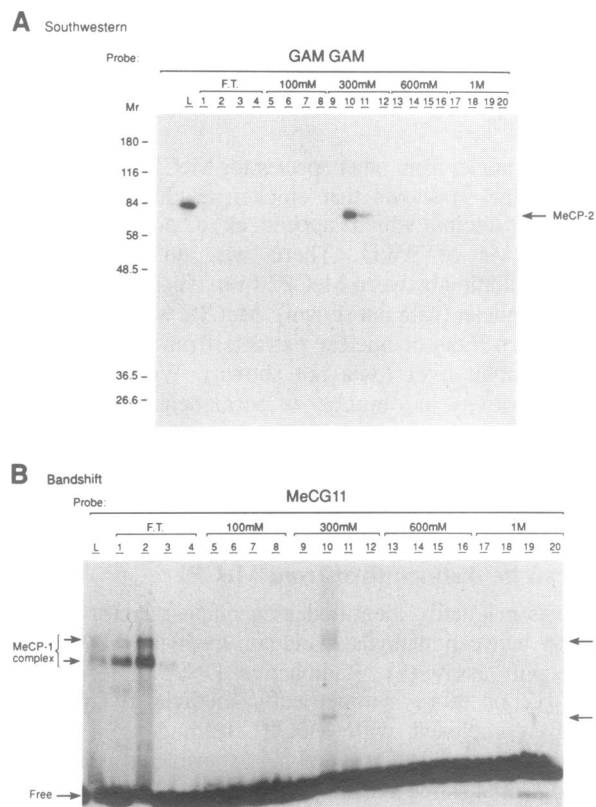
A Southwestern

B Bandshift

Figure 2. MeCP2 and MeCP1 show distinct chromatographic properties. A high salt extract (0.55*M*) from rat brain nuclei was chromatographed over a CM sepharose column (see Materials and Methods) and bound proteins were eluted using stepped increases in NaCl salt concentration in the range of 100mM to 1M. Eluted fractions were assayed by southwestern (panel A) and by bandshift (panel B). In this figure GAM GAM is poly(GAM) and MeCG11 is the methylated form of CG11 (7). Only the methylated DNA experiments are shown as the non-methylated probes did not form any protein/DNA complexes under the conditions used. In panel A the size, in kiloDaltons, of prestained molecular weight markers is indicated on the left. Lane L corresponds to the loaded extract and F.T. is the flow through from the column.

extract and its corresponding nuclear pellet were tested for MeCP2 activity using the southwestern assay. An example of the results obtained are shown in figure 3A and 3B and in graphic form in figure 3C. Both methylated and non-methylated DNA probes detect a doublet of proteins with an approximate Mr of 30 kD in the high salt extracts (see for example figure 3A and 3B, lanes 5−11). Subsequent purification and characterization established that these proteins correspond to forms of histone H1 (data not shown). MeCP2 is removed from the nucleus over a wide range of salt concentrations. More than 90% is extracted by 0.45M salt, and additional experiments have shown that 0.5M NaCl is sufficient to remove all the MeCP2 from rat brain, kidney and liver nuclei (data not shown), as is the case for histone H1. However the pattern of release of MeCP2 from the nucleus is different from that of histone H1. This becomes more apparent when the autoradiographs are densitometrically scanned and the percentage release of histone H1 and MeCP2 plotted as a function of NaCl salt concentration (figure 3C). It can be seen that it requires a salt concentration of at least 0.35*M* to release histone H1, whereas over 50% of MeCP2 is eluted with 0.3M NaCl. In fact the pattern of release of MeCP2 appears to be biphasic in that a significant proportion of the protein is retained in the
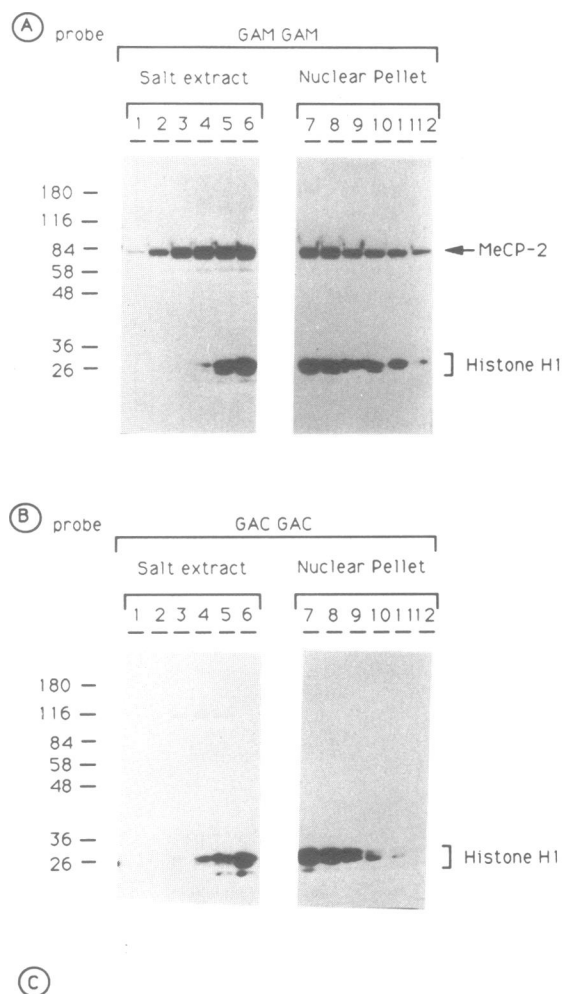

Figure 3. Salt extraction of MeCP2 from rat brain nuclei. Equal amounts of rat brain nuclei were extracted with various NaCl concentrations ranging from 0.2M to 0.45M in 0.05M steps. Each extract and its corresponding nuclear pellet were then assayed for MeCP2 activity by southwestern. The salt concentrations were as follows; lanes 1,7 0.2M ; lanes 2,8 0.25M ; lanes 3,9 0.3M ; lanes 4,10 0.35M ; lanes 5,11 0.4M ; lanes 6,12 0.45M. Panel A shows a southwestern with a methylated DNA probe GAM GAM (poly(GAM)) and panel B shows a southwestern with a non-methylated DNA probe GAC GAC (poly(GAC)). MeCP2 and histone H1 are indicated by an arrow and a bracket respectively. The size, in kiloDaltons, of the prestained molecular weight markers is indicated on the left. Panel C shows a graph of the percentage release of histone H1 and MeCP2 from rat brain nuclei with increasing salt concentration. This was determined by densitometric scanning of the autoradiographs.
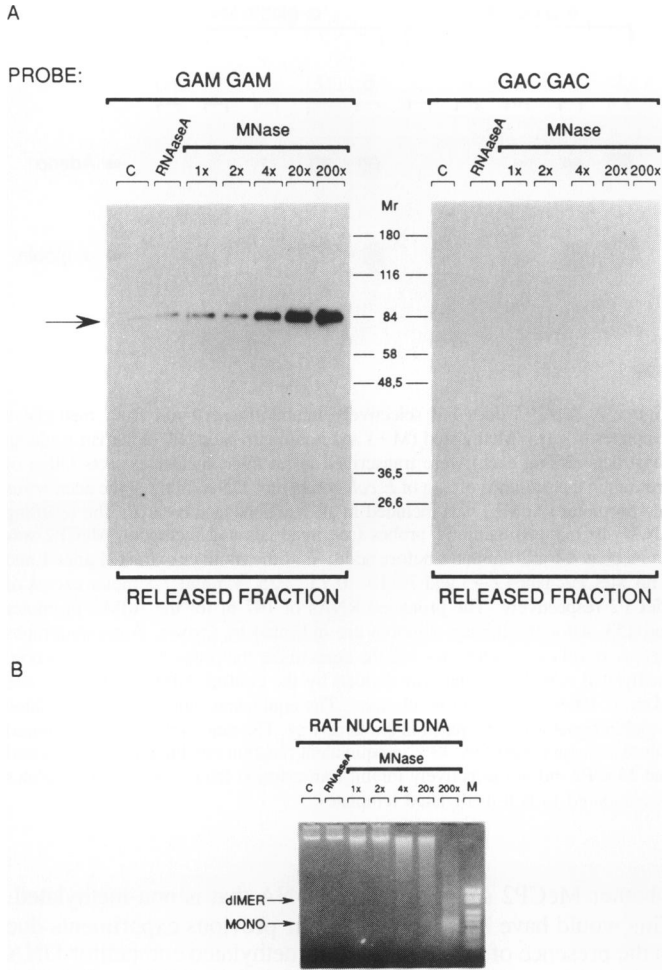
A

PROBE:

RELEASED FRACTION   RELEASED FRACTION

B

Figure 4. MNase release of MeCP2 from rat brain nuclei. Equal amounts of rat brain nuclei were incubated either alone, with RNase A (40μg total) and increasing amounts of MNase for 15 minutes at room temperature. The soluble fraction was isolated after EDTA treatment (see Materials and Methods) and tested for MeCP2 activity. Panel A shows the southwestern with a methylated DNA probe GAM GAM [poly(GAM)] and a non-methylated DNA probe GAC GAC [poly(GAC)]. An arrow indicates MeCP2. The numbers in the middle refer to the sizes, in kiloDaltons, of the prestained molecular weight markers. In the MNase treated nuclei 1× is equivalent to 0.22 units of MNase. Panel B shows the effect of the various treatments on the rat brain genomic DNA. The nucleosome monomer and dimer are indicated. Lane M is *Hpa*II digested pUC19, the largest band being 501 bp.

nucleus after 0.35M salt extraction (approximately 30%). Whether this is indicative that a high proportion of MeCP2 is unbound, or that there are high and low affinity sites for MeCP2 in the brain nucleus is unknown at present. In contrast to this, all of the measurable MeCP1 activity was released by 0.25M NaCl extraction of rat and mouse tissue nuclei (data not shown).

MeCP2 is associated with chromatin

The salt extraction experiments suggest that MeCP2 is complexed in the nucleus. Its affinity for methylated DNA *in vitro* would suggest that it might be complexed with chromatin. This was tested indirectly by examining the effect of nuclease treatment of nuclei on MeCP2 release. Rat brain nuclei were treated with increasing amounts of MNase and the resulting soluble fraction assayed for MeCP2 activity. As a control, nuclei were treated with an excess of RNAase A, which should not disrupt chromatin.
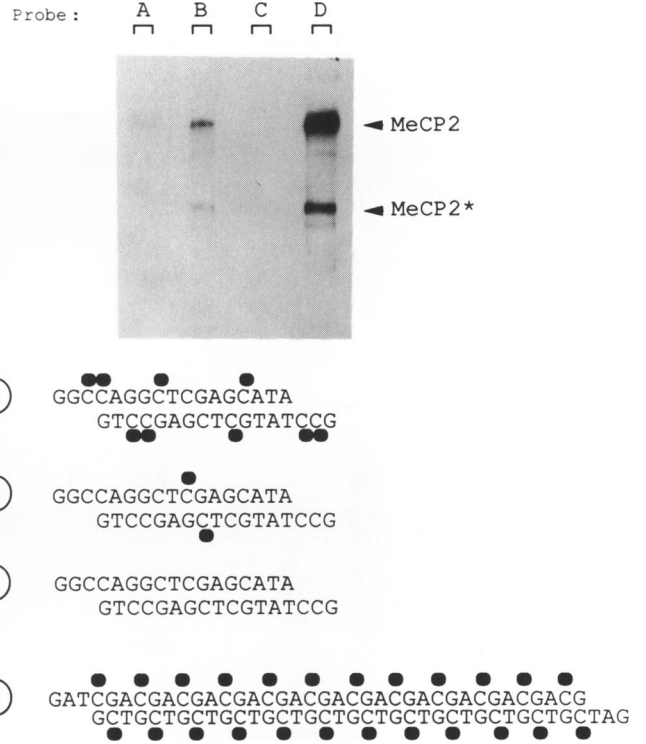

Figure 5. Binding specificity of MeCP2. Purified rat MeCP2 (20ng) was assayed by southwestern in the presence of competitor DNA with methylated or non-methylated ³²P labelled probes (panel A–D). Sequences below the figure (A–D) correspond to the probes used in each panel. Black dots indicate the presence of m⁵C in the chemically synthesized DNA strands. The appropiate DNA strands were annealed, labeled with ³²P γ-ATP and ligated into high molecular weight forms before use as southwestern probes. The filled arrows indicate the full lenth rat MeCP2 and its breakdown product, MeCP2* present in this particular preparation.

In figure 4A and B, it can be seen that increasing amounts of MNase, but not RNase A, release MeCP2 into the soluble chromatin fraction. MeCP2 is quantitatively released from the nuclei even under conditions where nuclear DNA remains at a high average molecular weight (compare lane 20× in figures 4A and B). Thus MeCP2 is not able to remain bound to mononucleosomes. The small amounts of MeCP2 released in the control and RNAase A treated nuclei (50 times less than the nuclei treated with the highest concentration of MNase) may be due to nicking of DNA in chromatin by endogenous nucleases. In addition to MNase, treatment of nuclei with restriction enzymes released MeCP2 in proportion to their ability to digest nuclear chromatin. In particular, fractionating *Hae*III digested chromatin on sucrose gradients suggested that MeCP2 co-sediments with chromatin with a length of greater than 6–7 nucleosomes (data not shown).

MeCP2 has a preference for DNA methylated at CpG

DNA containing a single methyl-CpG pair can bind MeCP2 whereas hemi-methylated DNA is a poor substrate (8). This implies that symmetrically methylated CpG is important for MeCP2 binding. To learn more about the binding requirements of MeCP2, we investigated the affinity of ligated forms of the AB17 series of oligonucleotide duplexes for purified rat MeCP2. The probe AB17 was synthesized in three forms: methylated at all cytosines *except* CpG (figure 5A); methylated *only* at the single
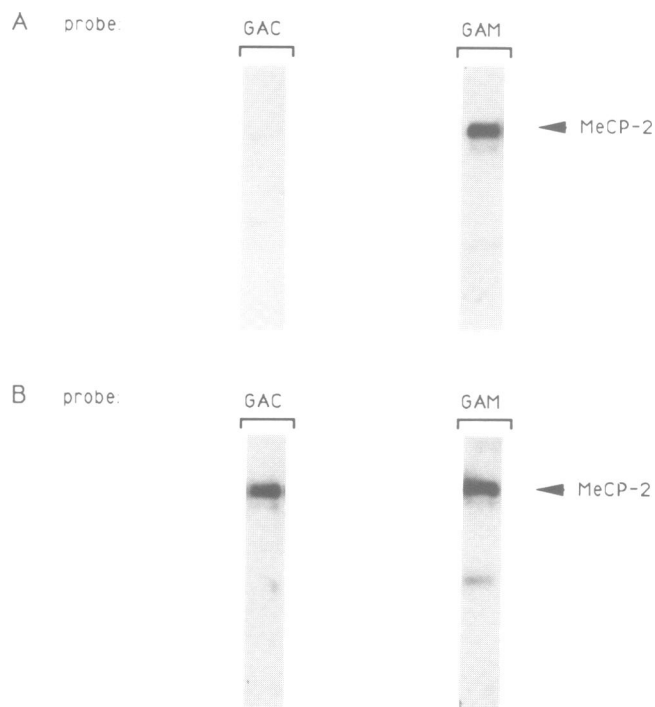
Figure 6. MeCP2 has more then one DNA binding activity. MeCP2 was purified from rat brain nuclei (see Materials and Methods) and 100ng was used in each panel for southwestern analysis with poly(GAM) and poly(GAC) probes in the presence (panel A) and absence (panel B) of competitor DNA. Rat MeCP2 is indicated with an arrowhead.

Figure 7. MeCP2 does not selectively inhibit transcription from methylated templates in vitro. Methylated (M+) and non-methylated (M−) human α-globin constructs (200ng each) were transcribed in rat liver nuclear extracts (30μg of protein) in the presence of 2μg of E.coli competitor DNA. 50ng of the adenovirus late promoter (AdML) was included in all reactions as a control. The resulting RNA was mapped using S1 probes (see materials and methods). MeCP2 was preincubated for 20 minutes before addition of the rat liver extract. Lanes 1 and 6 no MeCP2; lanes 2−5 and 7−10 , 0.75, 2.25, 4.5 and 7.5 molar excess of MeCP2 respectively. The protected RNAs of 197 nt for the AdML promoter and 153 nt for the human α-globin are indicated by arrows. Autoradiographs were scanned by densitometry and the areas under the peaks measured. The non-methylated α-globin signal was divided by the control AdML signal and was taken as 100% expression in all cases. The equivalent ratio for the methylated α-globin signal was determined in the same way. The ratio of these two normalized values was approximately 30% in equivalent reaction conditions. This indicated that MeCP2 did not selectively inhibit transcription from methylated templates as compared to non-methylated templates.

CpG (figure 5B); and non-methylated (figure 5C). We compared the binding of ligated versions of AB17 with each other and with that of the poly(GAM) sequence (figure 5D). Lane A of figure 5 shows that multimers of AB17 in which all non-CpG cytosines are methylated binds MeCP2 very weakly. In comparison multimers of AB17 in which only the CpG is methylated (1 CpG pair per 17 bp) shows stronger binding (figure 5B) even though the density of m⁵C residues is over four-fold lower than in probe A. As expected, the control sequence, poly(GAM), which had the highest density of methyl CpG (12 pairs per 40 bp), showed the strongest binding to MeCP2.

Thymidine also contains a methyl group at position 5. Thus TpG might also serve as substrate for MeCP2. The AB17 duplex contains 2 TpGs per 17 bp but the non-methylated form of AB17 (figure 5, oligo C) showed no binding to MeCP2 under the conditions used. TpG cannot substitute for m⁵C when placed opposite m⁵CpA (figure 5, oligo A), as the affinity of MeCP2 for this oligo is much reduced compared to multimers of AB17 which contain 1 MeCpG pair per 17 bp (figure 5, oligo B). The AB 17 series of oligos also contains the TpA motif (1 TpA per 17 bp) but the methyl binding domain of MeCP2 does not recognise this dinucleotide under the conditions used (figure 5, oligo C). We conclude that the methyl binding domain of MeCP2 is specific for DNA molecules containing symmetrically methylated CpG and that the methyl group of thymidine cannot substitute for m⁵C to facilitate MeCP2 binding.

MeCP2 has at least two DNA binding domains

DNA symmetrically methylated at CpG is the prefered ligand for MeCP2 *in vitro* in the presence of competitor (see ref. 8). Wi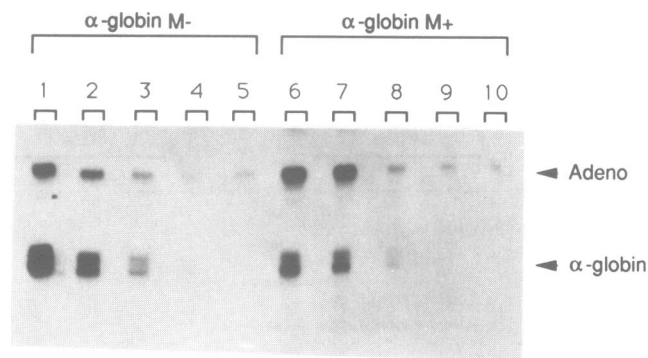th the purified protein available, it was of interest to ask whether MeCP2 can also bind to DNA that is non-methylated. This would have been masked in our previous experiments due to the presence of an excess of non-methylated competitor DNA in the southwestern probe mixture. To test this, we performed a southwestern assay in the absense of competitor DNA. In figure 6B it can be seen that both poly(GAM) and poly(GAC) bind MeCP2, though the signal with the non-methylated poly(GAC) probe was reduced, suggesting that the affinity for the methylated probe is higher. The result implies that MeCP2 has more than one DNA binding site. Since the specificity for methylated DNA was undiminished in the presence of a vast (20,000 fold) excess of non-methylated DNA, we deduce that the two domains are not overlapping.

MeCP2 does not preferentially repress transcription from a methylated template *in vitro*

Although the primary role of methylated CpGs in mammalian DNA is uncertain, it has been demonstrated that methylation near the promoters of genes is associated with transcriptional repression *in vitro* and *in vivo* (1, 2, 3, 4, 14, 15, 16). There is increasing evidence that this repression is indirect and is mediated by the binding of methyl-CpG binding proteins, such as MeCP1, which prevent access by transcription factors to the methylated DNA (4, 14, 15). We wished to know if highly purified MeCP2 could selectively repress transcription of methylated DNA. It has been shown that when high amounts of template are used in HeLa cell nuclear extracts both the methylated and non-methylated forms of the human α-globin (1 CpG per 10 bp) gene can be transcribed *in vitro* (4), as there is insufficient MeCP1 to repress all copies of the methylated template. We repeated this experiment using rat liver nuclear extracts and found that at a template concentration of 20μg/ml both methylated and non-methylated forms of the human α-globin

Table 1. Properies of MeCP1 and MeCP2 from adult rat nuclei.

	MeCP1	MeCP2
Methylated CpG pairs required for binding.[a]	12	1
Size (gel filtration)[b]	400, 800 kDa	100kDa
Size (SDS/PAGE).[b]	120 kDa	84 kDa
Nonspecific DNA binding domain.[c]	no	yes
Salt elution from nucleus.[d]	> 0.25 M	> 0.45 M
Abundance (molecules/somatic adult nucleus).[e]	5×10^3	$1-5 \times 10^5$
Repression of methylated genes in vitro.[f]	yes	no

(a) The number of methyl groups needed for binding and (b) the native and the relative molecular weights came from Meehan et al (7) and Lewis et al (8). (c) The detection of non-methylated DNA binding domains in MeCP2 is described in this work (figure 6) and reference 8. The absense of a non-methylated DNA binding domain in MeCP1 was was determined on semi-purified preparations of MeCP1 by bandshift in the absense of competitor DNA (J.L., unpublished data). (d) Salt elution was determined for MeCP2 as described in figure 3, the salt elution for MeCP1 was determined in a similar manner but assayed by bandshift (R.M. unpublished). (e) The abundance of MeCP1 was determined by titration experiments and is an approximate value. The abundance of MeCP2 was calculated from the predicted molecular weight (53 kDa) from translation of the rat cDNA (8) and the initial weight of protein present at the start of purification from kidney and brain nuclei. (f)The transcriptional analysis is from this work and references 4 and 14.

gene were transcribed. Transcription of the methylated human α-globin gene was 30% of the non-methylated form when the signals were normalized relative to the adenovirus major late promoter (AdML) control signal (figure 7, lanes 1 and 6). At lower template concentrations, transcription of the methylated human α-globin gene was completely abolished whereas the non-methylated template was transcribed (data not shown). To test whether MeCP2 could preferentially inhibit trancription of methylated DNA templates, we added purified MeCP2 to transcription extracts (figure 7, lanes 2−5 and 7−10). Increasing amounts of MeCP-2 repressed transcription from all tested templates, irrespective of their methylation status, including the AdML control gene. The relative signal from the methylated template remained at 30% of the non-methylated template under all conditions tested. Indiscriminate repression is probably due to the non-specific DNA binding activity of MeCP2 (figure 6). The absense of selective repression of the methylated templates by MeCP2 was also seen using promoter constructs (methylated and non methylated) that were driven by purified T7 RNA polymerase (data not shown). Thus the generalised inhibition of transcription by MeCP2 is also seen in a transcription system made up of purified components. This implies that MeCP2 does not inhibit transcription by interaction with components of the complex mammalian nuclear transcription extract.

DISCUSSION

MeCP1 is different from MeCP2

Two methylated DNA binding proteins have been identified which bind DNA with relatively relaxed sequence specificity (7, 8). Their only sequence requirement for binding is the presence of symmetrically methylated CpGs. The present study shows that these proteins differ in a number of important respects demonstrating that they are distinct activities. Table 1 lists some the properties of MeCP1 and MeCP2. They differ in their chromatographic affinity and exhibit size differences on gel filtration and on SDS protein gels. Polyclonal antibodies raised

to the rat MeCP2 (Mr 84kD) detect only a single band on western blots of rat nuclear proteins (8), supporting the view that the 120kD MeCP1 protein is not closely related. MeCP1 has been estimated to be at least an order of magnitude less abundant then MeCP2 (7). They also differ in the number of methyl groups required for binding in their respective assays. MeCP1 requires 12 methyl-CpG pairs and MeCP2 can bind a DNA substrate with only one methyl-CpG pair. MeCP2 has an additional non-specfic DNA binding domain which has not been detected in MeCP1. There is good evidence that MeCP1 can act as a transcriptional repressor of methylated DNA in vitro and in vivo (4, 14). MeCP2, in contrast, cannot preferentially inhibit transcription from methylated genes in vitro (figure 7 and see below). Finally there are differences in their steady state levels, with MeCP2 being much reduced in adult rodent testis (figure 1) compared to other adult tissues, whereas MeCP1 exhibits comparable levels in all adult rodent tissues tested, including testis.

The nature of MeCP2 binding sites

Although the two proteins are distinct they share some overlap in their methylated DNA binding specificities. Hemi-methylated DNAs are poor substrates for both proteins (7, 8). In addition the methyl group and the identity of the base is critical for binding by MeCP1 and MeCP2. Thymidine, which also contains a methyl group at the 5 position, is not a substrate for the methyl binding domains of either protein when tested in a variety of configurations. For example oligos containing TpG opposite m⁵CpA, TpG only or TpA (which is self complementary) failed to bind either MeCP1 or MeCP2 efficently (figure 5 and reference 7).

A consistent feature of the MeCP2 assay was the need to include non-methylated competitor DNA in addition to the probe (even when using highly purified MeCP2 protein). In the absence of competitor DNA it could be demonstrated that MeCP2 has a non-specific DNA binding activity that probably does not overlap with the methylation-specific domain (figure 6 and unpublished data). In this connection it is of interest that the predicted protein sequence of rat MeCP2 contains a number of amino acid motifs that have been found in proteins which have a preference for AT-rich DNA (8). Indeed, cloned non-methylated mouse satellite-I DNA (63% A+T, 8 CpGs) can bind rat MeCP2 weakly even in the presence of non-methylated competitor DNA, although binding is considerably enhanced upon addition of the methyl groups (8 methyl-CpGs per 234bp) (8). This observation, which has been confirmed for mouse and purified pig MeCP2 (R.M. unpublished data), suggests that the prefered substrate for MeCP2 is AT-rich DNA containing methyl-CpG. The satellite DNAs of pig and rat are also AT-rich and contain a number of methylatable CpGs (17, 18) which makes these sequences potential binding sites for MeCP2. Immunological localization shows that MeCP2 is in fact concentrated in satellite-containing chromatin in mouse, although it is also present at lower levels in the euchromatic arms of mouse and rat chromosomes (8).

Function of MeCP2?

The two phenomena associated with methylated DNA in vertebrates are gene repression and altered chromatin (19). There is an accumulating body of evidence that MeCP1 can play a direct role in methylation-associated gene repression especially for genes whose promoters are rich in methylatable CpGs (4, 14, 15). Our initial attempt to test the effects of purified MeCP2 on transcription by adding it to transcription extracts gave negative

results (figure 7), as the nonspecific DNA binding activity of MeCP2 repressed both the methylated and non-methylated human α-globin gene. It would be premature to conclude, however, that MeCP2 is not involved in transcriptional repression. Nuclease treatment of chromatin indicates that MeCP2 is bound to chromatin, but the transcription experiments were carried out in histone-free extracts in which chromatin cannot form. If the natural ligand for MeCP2 is chromatin (as is the case for histone H1), this could explain our failure to observe specific effects.

MeCP2 may play a role in the genome-wide protection of methyl-CpGs against nucleases, as it is much more abundant than MeCP1 (which is present at about 5,000 molecules per nucleus), ranging from an estimated 100,000 molecules per rat liver nucleus to 500,000 molecules per rat brain nucleus (8). Brain nuclei, which have the highest levels of MeCP2, show particularly striking protection of methyl-CpGs against nucleases (6). Conversely PC13 cells, which have very reduced levels of MeCP1 and MeCP2, show markedly reduced levels of protection (6). A role for MeCP2 in the formation of specific chromatin structures is supported by our biochemical evidence indicating that MeCP2 is associated with chromatin (figures 3 and 4) and by the reported immunological localization of MeCP2 to both centromeric heterochromatin and euchromatic arms in rodent metaphase chromosomes (8). In order to test the idea more critically, it would be useful to know the proportion of methyl-CpGs that are available for binding to MeCPs in the nucleus. Can methyl-CpGs on the nucleosome surface bind, or is it just methyl-CpGs in the internucleosome linkers that are available for binding? We also need to know how many methyl-CpG pairs are actually bound by MeCP1 and MeCP2 *in vivo*.

The similarities and differences between MeCP1 and 2 suggest a working model for their roles in methylation-mediated repression. MeCP1 is of relatively low abundance (about 5000 molecules per nucleus), and is loosely bound, whereas MeCP2 is comparatively abundant (about 200,000 molecules per nucleus) and is only released from chromatin by high salt concentrations. MeCP1 may thus compete with transcription factors in the nucleoplasm for binding to methylated DNA, the outcome depending on the density of methylation and the affinity of the factors (14). We hypothesise that binding to MeCP1 then guides the DNA into a heterochromatic structure involving stable association with MeCP2. How this might happen is unknown, but we suggest that it occurs at DNA replication, since the resistance of methylated DNA to nucleases is known to increase dramatically at this time (20).

ACKNOWLEDGEMENTS

REFERENCES

1. Cedar, H. (1988) Cell **53**, 3−4.
2. Bird, A.P. (1992) Cell **70**, 5−8.
3. Keshet, I., Lieman-Hurwitz, J. and Cedar, H. (1986) Cell **44**, 535−545.
4. Boyes, J. and Bird, A.P. (1991) Cell **64**, 1123−1134.
5. Solage, A. and Cedar, H (1978) Biochemistry **17**, 2934−2938.
6. Antequera, F., MacLeod, D. and Bird, A.P. (1989) Cell **58**, 509−517.
7. Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E. and Bird, A.P. (1989) Cell **58**, 499−507.
8. Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A.P. (1992) Cell **69**, 905−914.
9. Wang, R., Zhang, X-Y. and Ehrlich, M. (1986) Nucleic Acids Res. **14**, 1599−1614.
10. Pawlak, A., Bryans, M. and Jost, J.-P., (1991) Nucleic Acids Res. **19**, 1029−1034.
11. Bradford, M. (1976) Anal. Biochem. **72**, 248−254.
12. Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) Nucleic Acids Res. **11**, 1475−1489.
13. Burton, D.R., Butler, M.J., Hyde, J.E., Philips, D., Skidmore, C.J. and Walker, I.O. (1975) Nucleic Acids Research **5**, 3643−3663.
14. Boyes, J. and Bird, A.P. (1991) EMBO J. **11**, 327−333
15. Levine, A, Cantoni, G.L. and Razin, A (1991), Proc. Nat'l. Acad. Sci. (USA) **88**, 6515−6518.
16. Busslinger, M., Hurst, J. and Flavell, R.A. (1983) Cell **34**, 197−206.
17. Pech, M., Igo-Kemenes,T. and Zachau, H.G. (1979) **7**, 417−432.
18. Jantsch, M.,Hamilton, B.H., Mayr, B and Schweizer, D. (1990) **99**, 330−335.
19. Lewis, J.D. and Bird, A.P. (1991) FEBS Lett. **285**, 155−159.
20. Hsieh, C-L. and Lieber, M.R. (1992) EMBO J. **11**, 315−325.