DNA Binding of Methyl-CpG-Binding Protein MeCP2 in Human MCF7 Cells†

Christoph Koch and Wolf H. Strätling*

Institut für Biochemie und Molekularbiologie I, Universitätsklinikum Hamburg-Eppendorf, Martinistrasse 52, D-20246 Hamburg, Germany

Received October 29, 2003; Revised Manuscript Received February 2, 2004

ABSTRACT: MeCP2 has been identified as a chromatin-associated protein that recognizes MAR elements as well as methyl-CpGs. To characterize target sequences of MeCP2 in human cells, we employed two complementary methods. First, by use of a preparative chromatin immunoprecipitation protocol, we created from MCF7 cells a library enriched with sequences bound to MeCP2. A total of 154 representative clones were sequenced and analyzed. A large fraction of clones was found to be associated with retrotransposons, mostly with *Alu* repeats. A subgroup of four clones is derived from putative MARs; one clone is associated with a CpG island, and four clones contain alphoid repeats. Classical satellite DNAs II and III are not represented among clones, although they are heavily methylated. Second, using indirect immunofluorescence microscopy, we show that MeCP2 staining of human metaphase chromosomes has a dotted to knobby appearance with a reduced level of staining of centromeric regions of some chromosomes. On the other hand, an anti-5-methylcytosine antibody preferentially stained the juxtacentromeric regions of chromosomes 1, 9, and 16, which habor highly methylated, classical satellite DNAs, and methylated alphoid sequences in centromeric regions of several other chromosomes with reduced intensity. In interphase MCF7 cells, the distribution of MeCP2 exhibits a granular appearance throughout the nucleus. This distribution does not parallel that of methylated cytosine and heterochromatin. The selective binding behavior of MeCP2 revealed by these results (preference for murine major satellite DNA, *Alu* sequences, MARs, and CpG islands) is explained by its ability to recognize the sequence information (guanine bases) adjacent to CpG (TpG) as demonstrated in previous footprinting experiments.

The pattern of DNA methylation is read by a small family of proteins, the MBD[1] (methyl-CpG-binding domain) protein family (*1*). The prototype of this family is methyl-CpG-binding protein 2 (MeCP2). This protein was originally identified in rats due to its ability to recognize DNA fragments containing symmetrically methylated CpGs (*2*) and in chicken due to its preference for binding to MARs, the putative bases of chromatin loops (*3*, *4*). In mammals, the MeCP2 protein has three functional domains. First, the MBD is responsible for recognition of methylated CpGs and for binding to unmethylated sequences and MARs. Second, there is a transcriptional repression domain in the middle of the protein. According to a widely accepted model, this domain recruits a corepressor complex which contains the transcriptional corepressor protein mSin3A and histone deacetylases. This is thought to result in a localized deacetylation of core histones and gene silencing (*5*, *6*). Third, the C-terminal region of MeCP2 contains a proline-rich protein interaction surface capable of binding to group II WW domains (*7*).

In murine metaphase chromosomes, MeCP2 is enriched in pericentromeric heterochromatin, in accordance with its content of major satellite DNA, the largest fraction of methylated sequences in mice (*2*). In interphase nuclei of murine neuronal cells, MeCP2 is associated with pericentromeric and perinucleolar heterochromatin but does not bind significantly to the heavily methylated ribosomal DNA repeats (*8*). The occupancy of specific genes by MeCP2 has been analyzed using chromatin immunoprecipitation assays. MeCP2 was found to be associated with the metallothionein I gene promoter in a rat hepatoma, with the silent hypermethylated human multidrug resistance gene, and with the methylated imprinting control region of the silenced allele of two mouse imprinting genes, U2af1-rs1 and H19 (*9*−*12*). These *in vivo* binding studies have been complemented by *in vitro* band shift and DNase I and methylation protection footprinting experiments. MeCP2 can recognize a single methylated CpG. In methylated mouse major satellite DNA, MeCP2 binds with a high affinity to two sites, containing the MeCP2 protected sequences 5′-GACGAAAT-3′ (site I) and 5′-AGTGTGTATT-3′ (site II) (*4*). In nonmethylated satellite DNA, MeCP2 binds to these sites with slightly reduced affinities.

The level of interest in the involvement of MeCP2 in transcriptional control considerably increased since the discovery that mutations in the *MECP2* gene cause Rett syndrome, a neurological disorder with onset in early childhood, which occurs primarily in girls (*13*). A difficulty is that most previous DNA binding studies with MeCP2

† This work was supported by Grant SFB545B2 (to W.H.S.) from the Deutsche Forschungsgemeinschaft. C.K. acknowledges fellowship support from the University of Hamburg.

* To whom correspondence should be addressed. Phone: +49-40-42803-2392. Fax: +49-40-42803-8161. E-mail: straetli@uke.uni-hamburg.de.

[1] Abbreviations: DAPI, 4′,6′-diamidino-2-phenylindole; EST, expressed sequence tag; LINE, long interspersed nuclear element; L1, LINE-1; MAR, matrix attachment region; MBD, methyl-CpG-binding domain; 5MeC, 5-methylcytosine; MeCP2, methyl-CpG-binding protein 2; PMSF, phenylmethanesulfonyl fluoride; SINE, short interspersed nuclear element.

10.1021/bi0359271 CCC: $27.50
Published on Web 04/07/2004

utilized murine cells or murine DNA fragments, and their results cannot be applied to human cells without reservation. We therefore initiated a search for DNA fragments associated with human MeCP2 using a preparative chromatin immunoprecipitation assay complemented by indirect immunofluorescence microscopy.

EXPERIMENTAL PROCEDURES

Cell Culture, DNA Labeling, and Formaldehyde Cross-Linking. Monolayer cultures of MCF7 cells, a human breast adenocarcinoma cell line (*14*), were grown on 145 mm diameter culture dishes in RPMI 1640 medium with Glutamax-I and 25 mM HEPES (Invitrogen) supplemented with 10% fetal calf serum (Invitrogen) at 37 °C and 5% CO_2. For DNA labeling, the medium of a semiconfluent dish was replaced with 20 mL of fresh medium containing 1.3 μCi/mL [*methyl*-^{3}H]thymidine (specific activity of 2 Ci/mmol, Amersham Biosciences), and cells were incubated for an additional 24 h. To cross-link cells with formaldehyde, 20 dishes with MCF7 cells (2×10^6 cells per dish), one of which contained labeled cells, were incubated for 5 min at 37 °C in 20 mL each of fresh RPMI 1640 medium, to which 11% formaldehyde was directly added to a final formaldehyde concentration of 1%. The 11% formaldehyde solution was prepared in 0.1 M NaCl, 1 mM Na-EDTA, 0.5 mM Na-EGTA, and 50 mM HEPES (pH 8.0) from a 37% boiled and deionized formaldehyde stock solution (Merck). Pioneering time course experiments, with a fixation time from 5 min to 48 h, indicated that ~98% of the MeCP2 was cross-linked to DNA within 5 min.

Purification and Immunoprecipitation of Cross-Linked DNA−Protein Complexes. Cross-linked DNA−protein complexes were purified as described by Göhring and Fackelmayer (*15*) with some modifications. Cross-linked cells were rinsed twice with ice-cold phosphate-buffered saline [PBS; 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and 1.4 mM KH_2PO_4 (pH 7.3)], scraped off the dishes in 10 mL of PBS with a rubber policeman, and pelleted by centrifugation at 800*g* for 5 min at 4 °C. After being resuspended in 10 mL of RSB [10 mM Tris-HCl, 10 mM NaCl, 3 mM $MgCl_2$, and 0.5 mM phenylmethanesulfonyl fluoride (PMSF) (pH 8.0)], cells were homogenized by 15 strokes in a Dounce homogenizer (pestle B, Wheaton), and nuclei were collected by centrifugation at 800*g* for 8 min at 4 °C. Then nuclei were washed twice with RSB, and unbound proteins were extracted with buffer E [10 mM Tris-HCl, 10 mM $Na_2S_2O_5$, 1 M NaCl, 0.1% Nonidet P-40, 1 mM K-EDTA, and 0.5 mM PMSF (pH 8.0)]. Nuclei were pelleted as described above, resuspended in 2.7 mL of buffer E containing 0.1 M NaCl instead of 1 M NaCl, and lysed by the addition of 0.3 mL of 20% Sarcosyl. The extract was sheared by sonication for 5 min and added to 13 mL of a CsCl solution (1.55 g/mL) prepared in 20 mM Tris-HCl, 1 mM K-EDTA, and 0.5% Sarcosyl. The sample was centrifuged at 40 000 rpm in a Beckman SW60 rotor for 50 h at 20 °C. DNA−protein complexes formed a turbid, highly viscous disk near the middle of the gradient. The gradient above the disk was fractionated from the top in aliquots of 250 μL each; the disk was picked with a pair of tweezers, and then 250 μL fractions were collected from below the disk. The density of each fraction was determined by refractometry, and relative DNA concentrations were measured by liquid scintillation counting. The disk containing DNA−protein complexes was washed three times in TE buffer [50 mM Tris-HCl and 5 mM Na-EDTA (pH 7.5)], sonicated in 3 mL of TE buffer with a Branson model 450 sonifier at power setting 10 and cycle setting 50% with constant cooling until the disk was nearly dissolved (30−60 min), and clarified by centrifugation at 14 000 rpm in an Eppendorf centrifuge for 30 min at 4 °C. The clear supernatant ("cross-link fraction") was stored for up to 3 months at −25 °C.

As a preclearing step, 80 μL of a 50% slurry of salmon sperm DNA and protein A agarose beads (Upstate) was added to 1 mL of the cross-link fraction, incubated for 30 min at 4 °C, and removed by centrifugation at 4000 rpm for 5 min at 4 °C. For immunoprecipitation, 50 μL of a 50% slurry of swollen protein A and Sepharose CL4B (Pharmacia) was incubated in 400 μL of NP-40 buffer [140 mM NaCl, 50 mM K-HEPES (pH 7.5), 2 mM Na-EDTA, 10% glycerol, 0.5% Nonidet P-40, and 1 mM PMSF] (*16*) with 10 μg of anti-MeCP2 antibody (Upstate, catalog no. 07-013) for 2 h at 4 °C with continuous mixing on a rotary mixer. Sepharose beads were then collected by centrifugation at 4000 rpm for 5 min at 4 °C, and washed three times with NP-40 buffer. Loaded protein A−Sepharose beads were resuspended in 200 μL of 2× NP-40 buffer and incubated with 200 μL of precleared cross-link fraction overnight at 4 °C using rotary mixing. In control experiments, the cross-link fraction was incubated with protein A−Sepharose beads loaded with rabbit preimmune IgG (Dianova) or anti-Sp1 antibodies (Santa Cruz Biotechnology). Immunocomplexes were pelleted at 4000 rpm for 8 min at 4 °C, and pellets were washed three times (15 min each) in NP-40 buffer at 4 °C, and then eluted with 250 μL of elution buffer (0.1 M $NaHCO_3$ and 1% SDS) at room temperature for 20 min with rotation. The supernatant was collected, and the elution was repeated once. To revert the cross-links, 20 μL of 5 M NaCl was added to the pooled supernatant fractions, and the mixture was incubated for 5 h at 65 °C (*17*).

Protein and DNA Extraction (18) and Western Blotting. Protein and DNA were precipitated with ethanol overnight at −20 °C. After centrifugation at 14 000 rpm at 4 °C for 30 min, the pellet was washed with 70% ethanol, air-dried, dissolved in 100 μL of TE buffer (pH 8.0), and extracted with 100 μL of a phenol/chloroform/isoamyl alcohol mixture (25:24:1, Roth) at room temperature for 15 min. The organic phase was back-extracted once with an equal volume of TE buffer. The pooled aqueous solutions were extracted once again with a phenol/chloroform/isoamyl alcohol mixture. DNA was precipitated from the aqueous phase with ethanol and 1 μg of glycogen in the presence of 0.07 M sodium acetate (pH 5.2), washed in 70% ethanol, and air-dried. Protein was precipitated from the pooled organic phases with 10 volumes of acetone overnight at −20 °C, and pelleted by centrifugation at 14 000 rpm for 30 min at 4 °C. The pellet was washed with cold acetone, air-dried, and resolved by SDS−8% PAGE. After transfer onto nitrocellulose Protan BA85 membranes (Schleicher and Schuell), blots were incubated with anti-MeCP2 antibody (1:1700) and horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (1:5000, Southern Biotechnologies). Antigen−antibody complexes were visualized using an enhanced chemiluminescence detection system (Amersham Biosciences).

Cloning. To produce 3′-A overhangs, the DNA pellet was dissolved in a total volume of 30 μL containing PCR buffer [10 mM Tris-HCl (pH 8.8), 50 mM KCl, 0.08% Nonidet P-40, and 1.5 mM $MgCl_2$], dATP, dCTP, dGTP, and TTP (0.7 mM each), and 1 μL of Taq DNA polymerase (MBI Fermentas) and incubated for 10 min at 72 °C. The DNA was precipitated with ethanol in the presence of 0.07 M sodium acetate (pH 5.2) for 30 min at −80 °C, pelleted, washed with 70% ethanol, and air-dried. For cloning, the pGEM-T Easy System I cloning kit (Promega) was used. The DNA was incubated (total volume of 10 μL) overnight at 16 °C in the supplied ligation buffer [30 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, and 5% poly(ethylene glycol)] with 50 ng of the vector pGEM-T Easy and 3 units of T4 DNA ligase and transformed into chemically competent *Escherichia coli* strain DH10B cells. Transformed cells were plated on LB/ampicillin/isopropyl β-D-thiogalactoside/5′-bromo-4-chloro-3-indolyl β-D-galactopyranoside plates. Inserts were analyzed directly from picked white colonies by PCR using T7 forward (5′-TAATACGACTCACTATAGGG-3′) and Sp6 reverse (5′-GATTTAGGTGACACTATAG-3′) primers. Cycling conditions were as follows: 96 °C for 2 min; 40 cycles of 96 °C for 10 s, 52 °C for 30 s, and 72 °C for 25 s; and 72 °C for 5 min. Inserts were furthermore examined by restriction analysis using *Eco*RI. Inserts of 154 clones were sequenced from each end by direct automated sequencing with Dye Terminator Ready Reaction mix (Applied Biosystems) and analysis on an ABI 310 system (Applied Biosystems) (Institut für Zellbiochemie und Klinische Neurobiologie, Universitätsklinikum Hamburg-Eppendorf).

Sequence analysis was conducted using the RepeatMasker program (http://ftp.genome.washington.edu/RM/RepeatMasker.html) for repetitive elements, software programs S/MAR-Test (http://www.genomatix.de/cgi-bin/smartest_pd/smartest.pl) and MAR-Wiz1.5 (http://www.futuresoft.org/MarFinder) for putative MAR elements, and the Ensembl human database (http://www.ensembl.org) (*19*) and the NCBI blast database (http://www.ncbi.nlm.nih.gov/BLAST) (*20*) for localization of the cloned fragments within the human genome. Quantification of *Alu* and LINE-1 sequences was performed by "real-time" PCR in a LightCycler (Roche) using *Alu* primers 5′-GGCCGGGCGGTGGCTC-3′ and 5′-GAGACCGAGTCTCGCTCTGTC-3′ and 5′ LINE-1 primers 5′-AGGCATTGCCTCACCTGG-3′ and 5′-AGCGAGATTCCGTGGGCG-3′ following the manufacturer's instructions.

Immunofluorescent Staining. Chromosomal preparations from short-term human lymphocyte cultures were obtained by the standard method. The spreads were submersed in PBS and incubated for 20 min at room temperature with anti-MeCP2 antibody (1:50 dilution) in PBS containing 3% bovine serum albumin. Spreads were then washed twice with PBS, incubated for 20 min at room temperature with Alexa 488-conjugated anti-rabbit IgG, washed again with PBS, mounted, and processed for confocal image analysis using a Leica TCS 4d laser confocal scanning microscope equipped with objective PL APO 63× 1/1.40. To visualize methylated DNA, spreads were irradiated with ultraviolet light for 16 h (30 cm from a UV-C 30 W lamp) (*21*), incubated with a monoclonal anti-5-methylcytosine (5MeC) antibody (kind gift of A. Niveleau, used at a dilution of 1:30) (*22*) and FITC-conjugated anti-mouse IgG (Jackson ImmunoResearch), and analyzed by fluorescence microscopy using an Ortholux Leica fluorescence microscope with filter I2. Controls lacking the first antibody showed that fluorescent staining was specific.

MCF7 and NIH-3T3 cells were fixed with 1% formaldehyde, washed with PBS, permeabilized with −20 °C cold methanol, rehydrated with PBS for 10 min, and processed for immunostaining as described above using an anti-MeCP2 antibody and Rhodamine Red-conjugated goat anti-rabbit IgG(H+L) (Molecular Probes). After staining with 4′,6′-diamidino-2-phenylindole (DAPI), image analysis was performed by fluorescence microscopy using the Leica fluorescence microscope.

To prepare nuclei, MCF7 cells were scraped with a rubber policeman, washed with PBS, and lysed by dropwise addition of prewarmed (37 °C) 0.07 M KCl. Nuclei were then pelleted by centrifugation at 800*g*, resuspended, and fixed by dropwise addition of 10 mL of a methanol/acetic acid mixture (3:1) over a period of 20 min at room temperature. Following centrifugation at 800*g*, the fixation step was repeated twice. Finally, the nuclear suspension in 300 μL of fixation solution was transferred onto slides, irradiated with UV light, incubated with anti-MeCP2 and anti-5MeC antibodies, washed, and incubated with Rhodamine Red-conjugated anti-rabbit and Cy2-conjugated anti-mouse antibodies. Confocal image analysis was performed using a Leica TCS 4d laser confocal scanning microscope.

RESULTS

A DNA Library Enriched with Sequences Bound to MeCP2. To create a library enriched with human genomic sequences bound to MeCP2, we used a preparative chromatin immunoprecipitation protocol (*15*). After MCF7 cells had been treated with formaldehyde, cross-linked chromatin was purified by CsCl isopycnic centrifugation and immunoprecipitated with anti-MeCP2 antibodies. In initial experiments, we tested the efficiency and specificity of the immunoprecipitation reaction. Proteins were purified from antibody-bound and unbound fractions and assayed for MeCP2 content by Western blotting. As shown in Figure 1, MeCP2 is substantially enriched in the chromatin fraction precipitated with anti-MeCP2 antibodies (lane 3) and only weakly represented in the unbound fraction (lane 4). As controls for the specificity of the procedure, the immunoprecipitation reaction was performed using rabbit preimmune IgG or anti-Sp1 antibodies. We found that MeCP2 was absent in the antibody-bound fractions (lanes 5 and 7) but could be detected as prominent bands in the corresponding unbound fractions (lanes 6 and 8). A parallel Western blot using anti-Sp1 antibody verified that this antibody was capable of precipitating Sp1-containing chromatin (data not shown). Collectively, these results demonstrate that the immunoprecipitation reaction using anti-MeCP2 antibodies is efficient and highly specific.

The DNA in the immunoprecipitated chromatin fraction was de-cross-linked, purified, and cloned into the vector pGEM-T Easy. The genomic inserts of a total of 154 clones were sequenced from both ends. A list of all cloned fragments along with their sequences and basic features is shown in the Supporting Information. Cloned fragments were

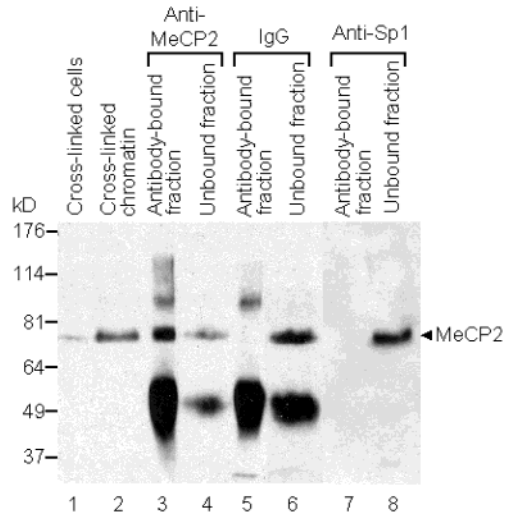

FIGURE 1: Efficiency and specificity of the immunoprecipitation procedure. MCF7 cells were cross-linked with formaldehyde (lane 1), and nuclear DNA−protein complexes were purified by CsCl isopycnic centrifugation (lane 2). Then immunoprecipitation reactions were performed using anti-MeCP2 antibodies, rabbit IgG, or anti-Sp1 antibodies. Proteins were purified from antibody-bound (lanes 3, 5, and 7) and unbound fractions (lanes 4, 6, and 8) and analyzed for MeCP2 content by Western blotting. The total protein from each pair of antibody-bound and unbound fractions was loaded onto the gel, thus allowing quantitative evaluation of the efficiency and specificity of the reaction. The apparent molecular mass of human MeCP2 (~80 kDa) is indicated at the right. The strong bands at ~50 kDa likely correspond to the light chain of the immunoprecipitating antibody.

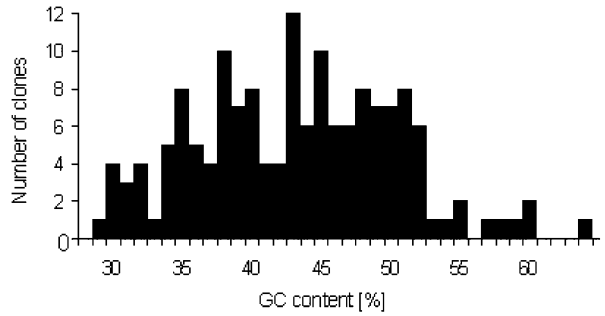

FIGURE 2: Histogram of GC content. Clones are sorted by their GC content (percentage).

first analyzed for general properties and sequence features. The size of the insert varies from 35 to 1346 bp (average of 394 bp). The GC content varies from 29.3 to 63.6% (average of 43.2%) (Figure 2). Also, the density in CpG dinucleotides varies greatly among cloned fragments. Fourteen (9.1%) clones do not contain any CpG, while in others, the CpG density approaches the value typically found in CpG islands (approximately 10 CpGs per 100 bp) (*23*). Cloned fragments were then analyzed for their location within the genome by searching the Ensembl human database and the NCBI blast database (as of June 15, 2003). The location of 116 (75.3%) clones was unequivocally determined; they proved to derive mostly from intergenic regions and introns. The failure to localize the insert of 38 clones is mostly due to their shortness or their high content of repetitive sequences. Analyzing the distribution of the localizable clones along the chromosomes, we found that in a large number of cases two to four clones are localized to identical or very close chromosomal loci

Table 1: Classes of Fragments Cloned from Immunoprecipitated Chromatin[a]

class	definition	no.	%
I	clones lacking transposable elements and alphoid repeats	61	39.6
A	GC content above average	37	24.0
B	GC content below average	24	15.6
II	clones containing transposable elements	91	59.1
A	clones containing SINEs	59	38.3
	(1) *Alu* elements	49	31.8
	AluJ	18	11.7
	AluS	22	14.3
	AluY	16	10.4
	(2) MIR elements	10	6.5
B	clones containing LINEs	28	18.2
C	clones containing LTR retrotransposons	13	8.4
D	clones containing DNA tansposons	11	7.1
III	clones containing alphoid repeats	4	2.6

[a] Classes II and III and subclasses therefrom partially overlap, since 21 clones contain more than one type of repetitive element. Furthermore, 15 clones contain more than one copy of a single type of element.

(Table S1 of the Supporting Information). For instance, locus 2p25.2−25.3 contains clones 33, 51, and 125, locus 8q22.1−22.2 contains clones 16, 58, 71, and 109, and locus 8q14 contains clones 34, 52, and 63. Overall, nearly half of the localizable clones belong to this subgroup. This suggests that MeCP2 tends to be clustered at certain chromosomal loci. Repetitive sequences are contained in 93 (60.4%) clones and account for 42.40% of all cloned sequences. Repetitive sequences include SINEs (short interspersed nuclear elements), LINEs (long interspersed nuclear elements), LTR retrotransposons, DNA transposons, and alphoid repeats. Notably, classical satellite DNAs II and III, which are components of constitutive heterochromatin, are not represented among clones, although they are heavily methylated (*21*).

Two parameters, the presence of various types of repetitive elements and the GC content, were used to divide clones into three classes and several subclasses (Table 1). Class I contains clones lacking transposable elements and alphoid repeats. On the basis of GC content, class I was further subdivided into subclasses IA and IB, members of each subclass possessing a GC content above the genome average (41%) and below, respectively (*24*). Class IA contains 28 clones with moderately elevated GC contents and four clones with a GC content between 51.6 and 63.6% (Table 2). Since CpG dinucleotide density in the latter clones varied from 2.1 CpGs per 100 bp (clone 111) to 5.8 CpGs per 100 bp (clone 76), we examined their genomic location for the association with CpG islands. Clone 111 is located on the flank of a typical CpG island on chromosome 11p11.2. This island is positioned in the center of a 134 kb intergenic region between the gene encoding the protein OASIS and a predicted gene. Since CpG islands are regarded as markers for the 5′-end of genes (*25*), we searched the Ensembl human database for expressed sequence tags (ESTs) close to the CpG island and indeed found one 0.4 kb EST adjacent to the island. On the basis of this information, we would predict that the island is located at the 5′-end of a novel gene. In band shift experiments using a labeled, methylated LINE-1 promoter fragment and bacterially expressed MeCP2 (*4*), methylated clone 111 fragments proved to be efficient competitors, while unmethylated fragments competed weakly

Table 2: Analysis of Class IA Clones and Their Relationship with a CpG Island

clone	length[a]	GC content[b]	CpG density[c]	location	relationship to a CpG island
35	620	54.7	3.4	20q11.21[d]	−
76	121	63.6	5.8	19p13.2[e]	located in a CpG-rich region downstream of a CpG island
111	516	51.6	2.1	11p11.2[f]	located on the flank of a CpG island
154	1079	52.0	4.2	−	−

[a] In base pairs. [b] Percentage. [c] CpGs per 100 bp. [d] Covers exon 5 and flanking intron sequences of the gene encoding the GDP-fucose protein *O*-fucosyltransferase 1 precursor. [e] Located in exon 6 of a predicted gene (UniGene cluster Hs.172928). [f] Located in a 134 kb intergenic region flanked by the gene encoding the protein OASIS (centromeric at a distance of 40.7 kb) and a predicted gene (ENST00000307087) (telomeric at a distance of 93.4 kb).

(data not shown). This supports the thesis that the methylated clone 111 sequence is a preferred binding site of MeCP2. Clone 76 is located in exon 6 of a predicted gene characterized by an elevated CpG density (3.4 CpGs per 100 bp) in its 5′-portion, including exon 6 and a typical CpG island at its 5′-end. Clone 35 was not found to be associated with any CpG island, and clone 154 could not be localized in the genome.

In our laboratory, MeCP2 has been originally identified as a protein specifically binding to the AT-rich, curved 5′-MAR of the chicken lysozyme gene (*3*, *4*, *26*). This element has been shown to insulate transgene expression from chromosomal position effects and to act as an enhancer blocking element (*27*−*29*). Other features of MARs include their association with origins of bidirectional replication and DNA topoisomerase II binding sites (*30*, *31*). These and other characteristics have been incorporated into two software programs predicting MARs, S/MAR-Test and MAR-Wiz1.5. Using these programs, we found that four of eight clones in class IB with a conspicuously high AT content are located in intergenic regions within or at the border of sequences predicted to be MARs (Table 3).

Alu elements, the only active SINEs in the human genome, are highly methylated and sparsely transcribed *in vivo* (*32*, *33*). *Alu* sequences are contained in 49 (31.8%) clones (Table 1) and account for 14.85% of cloned sequences (Table 4). Statistical evaluation by a parametric analysis (χ^2 test) reveals that the latter value is significantly higher ($P < 0.00001$) than the fraction of *Alus* in the human draft genome sequence (10.60%) (*24*). We also compared the level of *Alu* sequences in the cross-link fraction with that in the precipitated fraction by quantitative real-time PCR. The indicative crossing point in the cross-link fraction was 8.920 ± 0.002 and that in the precipitated fraction 8.45 ± 0.14 (means ± the standard deviation of five independent experiments, data not shown), indicating a slight enrichment of *Alu* sequences through the immunoprecipitation reaction. Collectively, these results suggest that MeCP2 is preferentially associated with *Alu* sequences. *Alu* sequences fall into three evolutionary age classes: *AluJ* (60−100 million years old), *AluS* (25−60 million years old), and *AluY* (<30 million years old). *AluJ* sequences are contained in 18 clones; all of these represent truncated and mutated elements and exhibit a low CpG content, due to C>T and methylated C>T transitions. *AluS* elements, which originate from a period where *Alu* sequences have experienced an exceptional peak of activity (*24*), are contained in 22 clones, while *AluY* elements, the youngest *Alu* class, are represented in 16 clones. MIRs (mammalian-wide interspersed repeats), an inactive SINE family (*34*), were found in 10 clones.

Table 3: Analysis of Class IB Clones for Potential MAR Elements

clone	length[a]	AT content[b]	CpGs	location	MAR prediction
2	422	70.1	1	4q13.3[c]	located at the border of a MAR[k,l]
43	223	68.2	−	13q13.3[d]	−
49	226	69.9	2	12q21.33[e]	located at the border of a MAR[l]
91	592	69.9	3	5q11.2[f]	located within a MAR[l]
92	329	70.2	1	1q31.1[g]	−
97	277	65.0	3	6q16.3[h]	located on the flank of a MAR[k]
121	430	69.1	3	4q25[i]	−
158	372	68.3	5	5q14.1[j]	−

[a] In base pairs. [b] Percentage. [c] Located in a 70.5 kb intergenic region flanked by genes encoding estrogen sulfotransferase and the α-S1 casein precursor. [d] Located in the first intron of a gene similar to that for the stomatin peptide. [e] Located in a large intergenic region with dual specific protein phosphatase 6 as the next telomeric gene at a distance of 212 kb. [f] Located in a 60.2 kb intergenic region flanked by genes encoding heat-shock protein β-3 and sorting nexin 18. [g] Located in the center of an approximately one-megabase large intergenic region. [h] Located in a several-megabase large region devoid of genes. [i] Located in a 101.4 kb intergenic region flanked by genes encoding myopodin protein and calcineurin-binding protein calsarcin-1. [j] Located in the first intron of a predicted gene. [k] Predicted by S/MAR-Test. [l] Predicted by MAR-Wiz1.5.

Table 4: Comparison of the Fraction of Repetitive Sequences in Cloned Sequences with the Fraction in the Human Draft Genome Sequence

	cloned sequences		
	length (bp)	fraction (%)	fraction of the draft genome sequence (%)[a]
repetitive elements (total)	25753	42.40	44.83
SINEs	10306	16.97	13.14
Alu	9018	14.85[b]	10.60
MIR	1288	2.12	2.20
LINEs	7489	12.33	20.42
LINE-1	5042	8.30[b]	16.89
LINE-2	2171	3.57	3.22
LINE-3	276	0.45	0.31
LTR retrotransposons	2987	4.92[b]	8.29
DNA transposons	1574	2.59	2.84

[a] Taken from ref *24*. [b] Significantly different from the respective fraction of the draft genome sequence ($P < 0.00001$).

While the nonautonomous SINEs do not encode any protein, LINEs contain two open reading frames encoding proteins required for their retrotransposition. LINE sequences are contained within 28 (18.2%) clones and account for 12.33% of all cloned sequences (Tables 1 and 4). Sequences of LINE-1 (L1), the only active autonomous non-LTR retrotransposon family in the human genome, account for 8.30% of cloned sequences, which is significantly reduced compared to the fraction of L1s in the human draft genome sequence (16.89%) (Table 4) (*24*). Four clones contain L1 sequences solely occurring in primates, while the remaining ones belong to subfamilies distributed among mammals (Figure 3). We then analyzed the genomic environment and

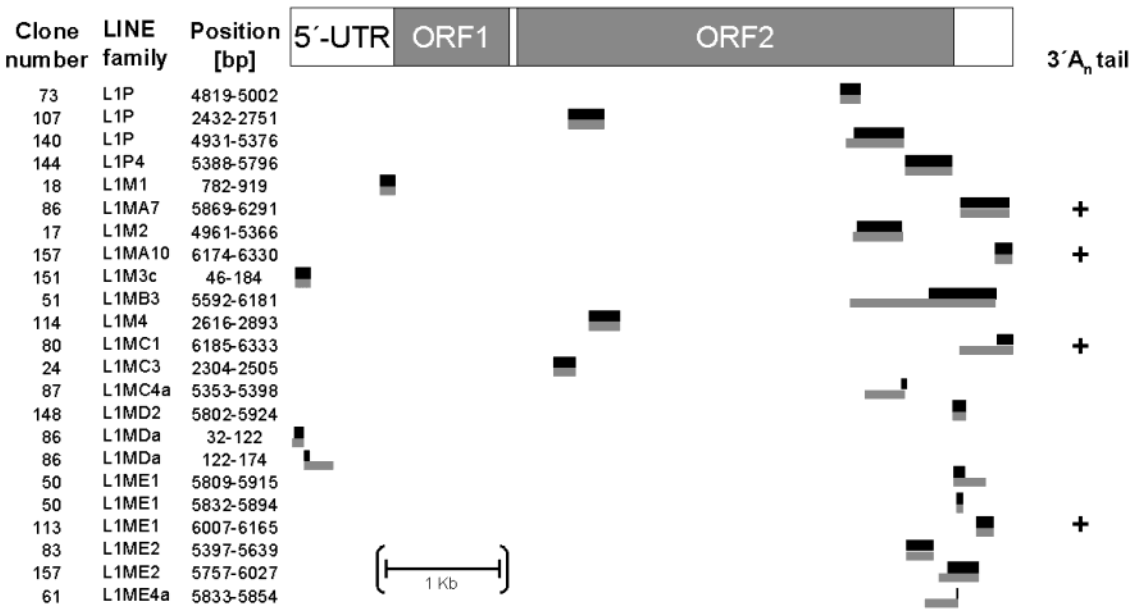

FIGURE 3: Localization of cloned L1 sequences within a functional L1 element. L1 sequences are ordered according to their evolutionary age. Positions of the sequences are shown numerically as well as graphically (black bars) below a schematic L1 map. The size of the respective L1 sequence in the genome is indicated by a gray bar. L1 sequences with putative authentic 3′-tails (A_n) are marked with a plus. 5′-UTR, 5′-untranslated region; ORF1 and ORF2, open reading frames 1 and 2, respectively.

the 3′-ends of the cloned L1 sequences (Figure 3). Surprisingly, we found that of 22 cloned L1 sequences at most four (clones 80, 86, 113, and 157) are derived from the 3′-end of intact full-length elements. The remaining ones originate from short internal L1 fragments. This contrasts with the well-documented fact that the great majority of genomic L1 elements are variably 5′-truncated. Furthermore, with one exception, all genomically localizable L1 sequences (13 clones) reside in intergenic regions. Thus, our results suggest that the cloned L1 sequences mostly are derived from a minor subpopulation of intergenically located, internal L1 relics, possibly generated by chromosomal rearrangements, recombinations, or insertion of other repetitive sequences (see clones 17, 24, 80, and 87). Similarly, all cloned LINE-2 and LINE-3 sequences represent short internal relics located in intergenic regions (*35*). The level of full-length L1 elements in the cross-link and the precipitated fraction was assessed by real-time PCR using primers from the 5′-end (data not shown). The crossing points, 18.47 ± 0.10 for the cross-link fraction and 22.00 ± 0.71 for the precipitated fraction (means ± the standard deviation, data not shown), indicate that full-length L1 elements were significantly depleted in the precipitated fraction.

Thirteen (8.4%) clones contain LTR retrotransposon sequences. In 11 clones, these sequences are solely derived from LTRs, while only two clones (98 and 115) contain internal LTR retrotransposon sequences (ERV class I). Most LTR retrotransposons in the human genome are "fossils" that consist of only an LTR, with the internal sequence having been deleted by homologous recombination between the flanking LTRs (*24*). This situation is reflected by the high proportion of LTRs relative to internal sequences among the cloned LTR retrotransposon sequences. In total, LTR retrotransposon sequences account for 4.92% of the cloned sequences, while in the human draft genome sequence, they represent 8.29% (*24*). DNA transposons are contained in 11 (7.1%) clones and alphoid repeats in four (2.6%) clones.

Localization of MeCP2 and 5-Methylcytosine. Surprisingly, we found that classical satellite DNAs II and III are not represented among clones, although they are heavily methylated (*21*). This prompted us to compare the distribution of MeCP2 with that of 5-methylcytosine (5MeC) in metaphase chromosomes of human peripheral lymphocytes using indirect immunofluorescence microscopy with anti-MeCP2 and anti-5MeC antibodies. Laser scanning images of MeCP2-stained chromosomes showed an irregularly dotted appearance (Figure 4A, top two panels). When whole chromosomes were viewed, through the overlay of many dots, the staining gained a knobby appearance (third panel). On all brightly stained metaphase plates, we observed a lack of staining or a significantly reduced level of staining in the centromeric regions of some metacentric and submetacentric chromosomes, such as chromosomes 1, 3, 6, and 7, although in these chromosomes alphoid repeats are methylated (as an example, see the asterisks in the third panel) (*36*). Using the same protocol, in murine lymphocytes, anti-MeCP2 preferentially stained the centromers (data not shown and ref *2*), indicating that the lack of accumulation of MeCP2 in centromeric heterochromatin in human lymphocytes is unlikely to be an artifact of the fixation or staining procedure. The anti-5MeC antibody preferentially stained the juxtacentromeric regions of chromosomes 1, 9, and 16 and the heterochromatic long arm of the Y chromosome, which is consistent with published data (*21*) (Figure 4A, bottom panel). These chromosomal sites are known for the location of the classical satellite DNAs II and III (*37*, *38*). C bands of several other chromosomes were stained with intermediate intensity, confirming previous observations about the methylation of alphoid repeats (*36*). These surprising results indicate that the distribution of MeCP2 on human metaphase chromosomes is significantly different from that of 5-methylcytosine. They also indicate that MeCP2 is not concentrated on classical satellite DNAs II and III despite their high degree of methylation, providing a reasonable explanation

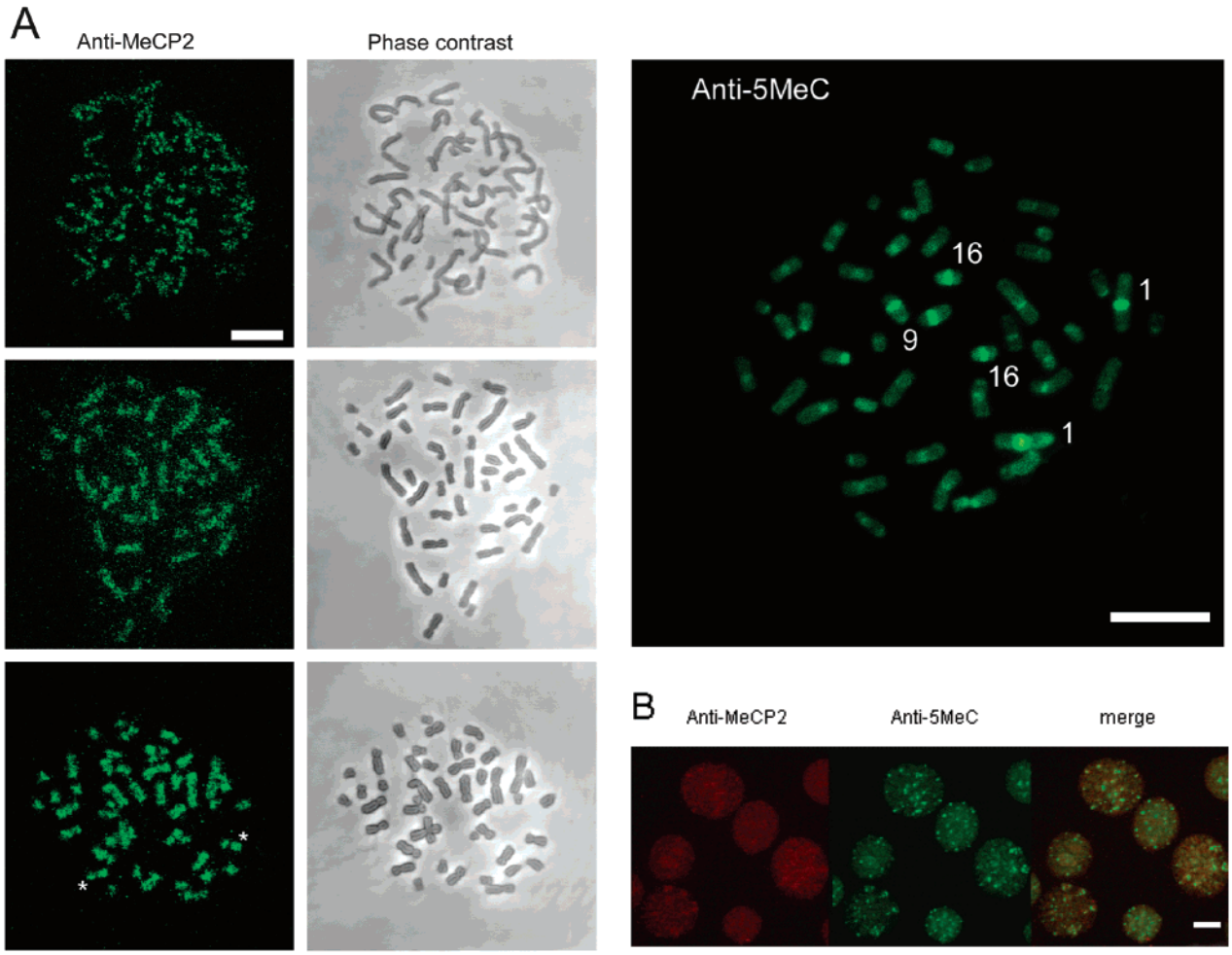

FIGURE 4: MeCP2 does not follow DNA methylation in human cells. (A) In the top three pairs of panels, metaphase chromosomes of human peripheral lymphocytes were stained with anti-MeCP2 antibodies and Alexa 488-conjugated secondary antibodies (left). Corresponding phase contrasts are shown at the right. The top two pairs of panels are confocal images, while whole chromosomes are viewed in the third pair of panels. Two examples for chromosomes showing the absence of staining in the centromeric region are marked with asterisks. In the panel at the very right, metaphase chromosomes were irradiated with UV light and stained with anti-5MeC antibodies and FITC-conjugated secondary antibodies. Scale bars are 10 μm. (B) Nuclei prepared from interphase MCF7 cells were stained with rabbit anti-MeCP2 (left) and anti-5MeC antibodies (center). Staining was detected with Rhodamin Red-conjugated anti-rabbit and Cy2-conjugated anti-mouse secondary antibodies using confocal immunofluorescence microscopy. In the right panel, both images are superimposed electronically (merge). The scale bar is 5 μm.

for our inability to find satellite DNA sequences among clones. Our results also suggest that methylated alphoid repeats are not a preferred target of MeCP2, in accordance with the presence of only four clones with alphoid repeats in our library.

We furthermore compared the distribution of MeCP2 with that of 5-methylcytosine in interphase MCF7 cell nuclei using indirect confocal immunofluorescence microscopy. The 5MeC staining exhibited a dotted appearance, with ~5−12 bright dots and a number of weaker dots (Figure 4B, center panel). The bright dots likely represent heavily methylated juxtacentromeric regions. In contrast, the MeCP2 staining exhibited a granular pattern distributed evenly throughout the nucleus (Figure 4B, left panel). When the electronically captured images were superimposed, distinctive overall distributions of 5-methylcytosine and MeCP2 became apparent (Figure 4B, right panel). At only very few spots is a colocalization visible. These results indicate that MeCP2 is not concentrated at highly methylated sites in interphase MCF7 cells.

Thus, our data suggest that in human cells the distribution of MeCP2 differs significantly from that in murine cells. While MeCP2 is concentrated at the highly methylated major satellite DNA in pericentromeric regions of murine cells, the distribution of MeCP2 does not parallel that of methylated CpGs in human cells. To further explore this fundamental difference, we compared the distribution of MeCP2 with that of DAPI-stained heterochromatin in human and murine cells. In MCF7 cells, nucleoli are surrounded by "shells" of heterochromatin which are strongly stained with DAPI (Figure 5C). Immunostaining of MeCP2 again exhibited a granular pattern thoughout the nucleus, except for the nucleoli (Figure 5A). Sparing of the nucleoli from MeCP2 distribution, which is consistent with previous observations (*8*), was also verified by double immunostaining with MeCP2 and nucleolin antibodies (data not shown). Furthermore, MeCP2 did not preferentially localize to DAPI-stained shells of heterochromatin. In contrast, in murine 3T3 cells, preferential MeCP2 staining was colocalized with DAPI-stained heterochromatin (panels B and D of Figure 5).

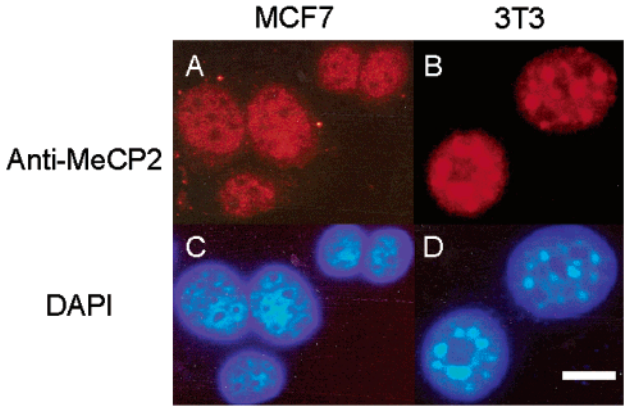

FIGURE 5: In human MCF7 cells, MeCP2 is not preferentially colocalized with DAPI-stained heterochromatin. Human MCF7 cells (A) and murine 3T3 cells (B) were stained with anti-MeCP2 antibodies and Rhodamin Red-conjugated secondary antibodies. After staining with DAPI (C and D), the distribution of MeCP2 and DAPI staining was photographed. The scale bar is 5 μm.

DISCUSSION

Using a preparative chromatin immunoprecipitation protocol, we created a library enriched with human DNA sequences to which MeCP2 is bound. By analyzing this library, we reach several interesting conclusions about the DNA binding features of MeCP2. First, the localizable clones are unevenly distributed along the chromosomes, suggesting that MeCP2 tends to cluster at certain chromosomal loci. A recent study using chromatin immunoprecipitation coupled with genomic hybridization reached a similar conclusion that a total of four MBD proteins, including MeCP2, cluster at certain chromosomal regions (*39*). Though the resolution of this analysis is much lower than that in our approach, it seems that some of these regions coincide with the loci we identified as preferred binding sites, for instance, 1q21, 3p14, 6q22, 12q12, and 19p13. Second, in general, features such as GC content, CpG density, and repeat content clones proved to vary greatly. We used a combination of these parameters to divide clones into three classes and several subclasses (Table 1). CpG islands are usually unmethylated (*23*) and hence not expected to bind MeCP2. However, during carcinogenesis, the DNA methylation pattern undergoes dramatic changes, including aberrant methylation of CpG islands (*40*). Since the source of the library was a human breast cancer cell line (MCF7 cells), we inspected class IA clones, which lack transposable elements and alphoid repeats and have a GC content above the genome average (41%), for potential CpG islands. One clone (111) proved to derive from a CpG island located in the center of a 134 kb intergenic region. Since CpG islands are mostly associated with the 5′-end of genes (*25*) and since we found an EST close to the CpG island, we predict that this island marks the 5′-end of a novel gene.

Kries and co-workers originally identified MeCP2 as a protein specifically binding to the 5′-MAR of the chicken lysozyme gene (*3*, *4*). MARs from several different organisms, including plants, have been shown to elevate the levels of and insulate transgene expression from chromatin position effects (*27*, *28*, *41*, *42*, and references therein). A MAR located close to the intronic κ chain enhancer plays a particularly important role. Together with the enhancer, it is responsible for somatic hypermutation and opening of the chromatin domain (*43*, *44*). MARs are thought to function as chromatin domain boundaries and to partition chromatin into independently regulated units. A role in higher chromatin organization also seems to be supported by the ability of MeCP2 to mediate *in vitro* packaging of nucleosomal filaments into defined higher-order structures and to promote assembly into oligomeric suprastructures (*45*). However, only a few MARs have been shown to be effective in an enhancer blocking assay (*29*). Using two computer programs which predict MARs, we found that four of our class IB clones are derived from genomic regions predicted to be MARs (Table 3). Binding of MeCP2 to MARs is remarkable since MARs, including the putative ones found here, have few CpGs. This can be best explained by the ability of MeCP2 to bind not only to methyl-CpGs but also to the related dinucleotide TpG. In fact, MeCP2 has been shown to bind to methylated mouse major satellite DNA at two sites with high affinity (K_D = 2−6 × 10^{-10} M) (*4*). As demonstrated by DNAse I footprinting, these sites contain the DNA sequences 5′-GACG-3′ (site I) and 5′-GTGTGT-3′ (site II). Since the cloned mouse satellite DNA unit used in these studies contains a total of seven CpGs, it is surprising that only one of these (methylated) CpGs (site I) is a high-affinity binding site, in addition to a site (site II) that contains two TpGs but no CpG. Thus, MeCP2 binds to a TpG-containing site with *higher* affinity than to the remaining six (methylated) CpGs.

The majority of methylated CpGs occur in repeats derived from transposable elements. Thus, it is not surprising that 59.1% of clones of the library harbor transposable elements (Table 1). The *Alu* consensus contains a strikingly high CpG content (25 in 283 bp), and there is evidence that most *Alu* sequences are heavily methylated in somatic cells (*46*). Thus, it is likewise not surprising that a large fraction (31.8%) of clones contain *Alu* sequences (Table 1). Furthermore, the fraction of *Alu* sequences (14.90%) in our library is elevated compared to the fraction of *Alu* sequences (10.60%) in the draft genome sequence (*24*). Also, the level of *Alu* sequences in the precipitated fraction is slightly increased relative to that in the cross-link fraction, as determined by real-time PCR. These results agree well with data about the composition of a library created with DNA fragments from human adenocarcinomas bound to the MBD of MeCP2 immobilized on a solid support (*47*). A minute fraction (0.9%) of clones was associated with CpG islands, yet a much larger fraction (72% of the subgroup having short inserts) contained *Alu* sequences (*48*). These published results and the presence of *Alu* sequences in nearly one-third of clones in our own library seem to argue for the involvement of MeCP2 in transcriptional repression of *Alu* sequences. However, in transient transfection experiments using an *AluSx* reporter tagged with 7SL upstream sequences, the transcriptional repression domain of MeCP2 when targeted to the reporter was unable to repress transcription (*49*). This paradox may be explained by the requirement for additional factors for repressing *Alu* transcription. *Alu* sequences contain an internal RNA polymerase III promoter that consists of an A box located just downstream of the transcription start site, and a B box located approximately 50 bp downstream of the A box. Methylation of a single CpG in the A box is sufficient to bind an as yet uncharacterized repressor and cause transcriptional repression (*33*). Furthermore, CpG methylation of the B box interferes

with binding of protein factors required for the function of the B box control region (*32*). We thus entertain the view that binding of MeCP2 to methylated *Alu* sequences plays, if any, only an auxiliary role in Alu repression. This notion is supported by the fact that *Alu* promoters can be activated by stress of various kinds without altering DNA methylation, although artifical demethylation also stimulates expression (*50*).

In striking contrast to *Alu* elements, the fraction of L1 sequences (8.27%) in the library is decreased in comparison to that in the draft genome sequence (16.89%). Furthermore, the cloned L1 sequences seem to be derived from the small subpopulation of intergenically located L1 relics consisting of short internal L1 fragments. This may be explained by assuming that these relics are methylated at a higher frequency than typical 5′-truncated L1 copies. We also found that the level of full-length L1 elements in the precipitated fraction is significantly decreased relative to that in the cross-link fraction. Either MeCP2 disfavors the highly methylated CpG island of full-length L1 elements (*51*), or most full-length L1 elements in MCF7 cells are not methylated and expressed (*52*).

The library did not contain a single clone harboring classical satellite DNA II or III. This is consistent with our immunohistochemical studies showing that the pericentromeric regions of human metaphase chromosomes 1, 9, and 16 do not stain brighter with MeCP2 antibodies than the chromosomal arms. Taken together, these results contrast significantly with the heavy methylation status of classical human satellite DNAs (*21*). In mouse metaphase chromosomes, on the other hand, the centromeric regions harboring heavily methylated major satellite DNA stain much brighter with MeCP2 antibodies than the chromosome arms (*2*). These apparently contradicting sets of data can be reconciled by assuming that MeCP2 also recognizes sequence and structural information adjacent to methyl-CpGs. We have mapped two high-affinity binding sites in the chicken lysozyme 5′-MAR, both containing the sequence 5′-G<u>G</u>**TG**-3′, and two high-affinity binding sites in mouse major satellite DNA, containing the core sequences 5′-<u>G</u>A**C**<u>**G**</u>A-3′ and 5′-G**T**<u>**G**</u>TGT-3′ (*4*, *53*). As determined in previous footprinting studies, the underlined guanine bases are protected by MeCP2 from methylation with dimethyl sulfate (*4*). It is evident that not only the guanine of the central CpG or TpG (bold) but also guanine bases in the −1 or −2 position are protected. First, this indicates that MeCP2 also recognizes sequence information adjacent to the central CpG (TpG) and seems to prefer guanine bases on one side. The C-terminal portion (helix α3) of the helical coil of the MBD may be involved in this recognition process, as indicated by recent NMR studies showing significant chemical shift changes in this region after formation of a complex with DNA (*54*). Second, the consensus sequences of the AT-rich human satellite DNAs II and III, $[CC(G)ATT]_n$ (satellite II) and $(CCATT)_2(CGATT)_2$-$[ATT(G)]_{1-2}$ (satellite III), lack any guanines upstream of CpGs. This easily explains MeCP2's aversion to human classical satellite sequences and its weak propensity for associating with AT-rich alphoid repeats. In contrast, GC-rich mouse major satellite DNA is a highly preferred target of MeCP2. This leads to the question of the role MeCP2 plays in mouse pericentromeric heterochromatin, if it is not found there in human chromosomes. During the embryonic development of the murine central nervous system, MeCP2 is redistributed from a diffuse pattern to pericentromeric loci (*55*). A very similar relocalization was observed after depolarization-induced activation of cultured embryonic neurons (*56*). It may be hypothesized that, during the neuronal maturation process, MeCP2 is shuttled from sites of active function to storage sites in pericentromeric heterochromatin. In human neurons, the identity of storage sites for MeCP2 is at present elusive. MeCP2 has often been characterized as a constituent of heterochromatin, since in murine cells it is enriched in pericentromeric heterochromatin (Figure 5 and ref *2*). However, in human MCF7 cells, MeCP2 does not preferentially associate with the heterochromatic shells surrounding the nucleoli, in accordance with our finding that in MCF7 cells the distribution of MeCP2 does not parallel that of methylated CpGs.

Summarizing, we conclude that MeCP2 exhibits surprisingly selective binding behavior. It is preferentially associated with murine major satellite DNA, *Alu* sequences, MARs, and CpG islands. On the other hand, classical human satellite DNAs II and III and L1 elements are not preferred target sites of MeCP2, and the level of binding to alphoid repeats is reduced, although these sequences are heavily methylated. These sequence preferences are explained by the ability of MeCP2 to recognize sequence information (guanine bases) adjacent to CpG (TpG).

ACKNOWLEDGMENT

We are grateful for use of the microscopic facilities of the Zentrum für Molekulare Neurobiologie Hamburg (ZMNH). In particular, we gratefully acknowledge the engaged help of M. Schweizer in the staining of metaphase spreads with anti-MeCP2 and data processing. We also thank A. Niveleau for the kind gift of the 5-methylcytosine antibody, J. Weitzel for help with real-time PCR and band shift experiments, and S. Singh for metaphase spreads.

SUPPORTING INFORMATION AVAILABLE

List of cloned fragments containing their nucleotide sequences and general features such as length, GC content, CpG density, genomic location, and information about repetitive elements, associated genes, and MARs and a listing of those chromosomal loci at which two to four clones cluster (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hendrich, B., and Bird, A. (1998) Identification and chracterization of a family of mammalian methyl-CpG binding proteins, *Mol. Cell. Biol. 18*, 6538−6547.
2. Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA, *Cell 69*, 905−914.
3. von Kries, J. P., Buhrmester, H., and Strätling, W. H. (1991) A matrix/scaffold attachment region binding protein: identification, purification and mode of binding, *Cell 64*, 123−135.
4. Weitzel, J. M., Buhrmester, H., and Strätling, W. H. (1997) Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2, *Mol. Cell. Biol. 17*, 5656−5666.

5. Nan, X., Ng, H.-H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex, *Nature 393*, 386−389.
6. Jones, P. L., Veenstra, G. J. C., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription, *Nat. Genet. 19*, 187−191.
7. Buschdorf, J., and Strätling, W. H. (2003) A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome, *J. Mol. Med. 82*, 135−143.
8. Akhmanova, A., Verkerk, T., Langeveld, A., Grosveld, F., and Galjart, N. (2000) Characterisation of transcriptionally active and inactive chromatin domains in neurons, *J. Cell Sci. 113*, 4463−4474.
9. Majumder, S., Ghoshal, K., Datta, J., Bai, S., Dong, X., Quan, N., Plass, C., and Jacob, S. T. (2002) Role of *de novo* DNA methyltransferases and methyl CpG-binding proteins in gene silencing in a rat hepatoma, *J. Biol. Chem. 277*, 16048−16058.
10. El-Osta, A., Kantharidis, P., Zalcberg, J. R., and Wolffe, A. P. (2002) Precipitous release of methyl-CpG binding protein 2 and histone deacetylase 1 from the methylated human multidrug resistance gene (*MDR1*) on activation, *Mol. Cell. Biol. 22*, 1844−1857.
11. Gregory, R. I., Randall, T. E., Johnson, C. A., Khosla, S., Hatada, I., O'Neill, L. P., Turner, B. M., and Feil, R. (2001) DNA methylation is linked to deacetylation of histone H3, but not H4, on the imprinted genes Snrpn and U2af1-rs1, *Mol. Cell. Biol. 21*, 5426−5436.
12. Drewell, R. A., Goddard, C. J., Tomas, J. O., and Surani, M. A. (2002) Methylation-dependent silencing at the H19 imprinting control region by MeCP2, *Nucleic Acids Res. 30*, 1139−1144.
13. Amir, R. E., Van den Veyver, L. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2, *Nat. Genet. 23*, 185−188.
14. Soule, H. D., Vazguez, J., Long, A., Albert, S., and Brennan, M. J. (1973) A human cell line from a pleural effusion derived from a breast carcimoma, *J. Natl. Cancer Inst. 51*, 1409−1416.
15. Göhring, F., and Fackelmayer, F. O. (1997) The scaffold/matrix attachment region binding protein hnRNP-U (SAF-A) is directly bound to chromosomal DNA *in vivo*: a crosslinking study, *Biochemistry 36*, 8276−8283.
16. Hecht, A., Strahl-Bolsinger, S., and Grunstein, M. (1996) Spreading of transcriptional repressor SIR3 from telomeric heterochromatin, *Nature 383*, 92−96.
17. Solomon, M. J., Strauss, F., and Varshavsky, A. (1986) A mammalian high mobility group protein recognizes any stretch of six A•T base pairs duplex DNA, *Proc. Natl. Acad. Sci. U.S.A. 83*, 1276−1280.
18. Kuo, M.-H., and Allis, C. D. (1999) *In vivo* cross-linking and immunoprecipitation for studying dynamic protein:DNA associations in a chromatin environment, *Methods 19*, 425−433.
19. Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T., et al. (2002) The Ensembl genome database project, *Nucleic Acids Res. 30*, 38−41.
20. Madden, T. L., Tatusov, R. L., and Zhang, J. (1996) Applications of network BLAST server, *Methods Enzymol. 266*, 131−141.
21. Miller, O. J., Schnedl, W., Allen, J., and Erlanger, B. F. (1974) 5-Methylcytosine localised in mammalian constitutive heterochromatin, *Nature 251*, 636−637.
22. Reynaud, C., Bruno, C., Boullanger, P., Grange, J., Barbesti, S., and Niveleau, A. (1992) Monitoring of urinary excretion of modified nucleosides in cancer patients using a set of six monoclonal antibodies, *Cancer Lett. 61*, 255−262.
23. Bird, A. P. (1986) CpG rich islands and the function of DNA methylation, *Nature 321*, 209−213.
24. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001) Initial sequencing and analysis of the human genome, *Nature 409*, 860−921.
25. Cross, S. H., and Bird, A. P. (1995) CpG islands and genes, *Curr. Opin. Genet. Dev. 5*, 309−314.
26. von Kries, J. P., Phi-Van, L., Diekmann, S., and Strätling, W. H. (1990) A non-curved chicken lysozyme 5′ matrix attachment site is 3′ followed by a strongly curved DNA sequence, *Nucleic Acids Res. 18*, 3881−3885.
27. Stief, A., Winter, D. M., Strätling, W. H., and Sippel, A. E. (1989) A nuclear DNA attachment element mediates elevated and position-independent gene activity, *Nature 341*, 343−345.
28. Phi-Van, L., von Kries, J. P., Ostertag, W., and Strätling, W. H. (1990) The chicken lysozyme 5′ matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes, *Mol. Cell. Biol. 10*, 2302−2307.
29. Bell, A. C., West, A. G., and Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators, *Cell 98*, 387−396.
30. Gilbert, D. M., Miyazawa, H., and DePamphilis, M. L. (1995) Site-specific initiation of DNA replication in *Xenopus* egg extract requires nuclear structure, *Mol. Cell. Biol. 15*, 2942−2954.
31. Adachi, Y., Käs, E., and Laemmli, U. K. (1989) Preferential, cooperative binding of DNA topoisomerase II to scaffold-associated regions, *EMBO J. 8*, 3997−4006.
32. Kochanek, S., Renz, D., and Doerfler, W. (1993) DNA methylation in the Alu sequences of diploid and haploid human cells, *EMBO J. 12*, 1141−1151.
33. Liu, W.-M., and Schmid, C. W. (1993) Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation, *Nucleic Acids Res. 21*, 1351−1359.
34. Smit, A. F. A., and Riggs, A. D. (1995) MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation, *Nucleic Acids Res. 23*, 98−102.
35. Koch, C. (2003) DNA target sequences of the human repressor methyl-CpG-binding protein 2 (MeCP2), Ph.D. Thesis, University of Hamburg, Hamburg, Germany.
36. Miniou, P., Jeanpierre, M., Bourc'his, D., Coutinho Barbosa, A. C., Blanquet, V., and Viegas-Péquignot, E. (1997) α-Satellite DNA methylation in normal individuals and in ICF patients: heterogeneous methylation of constitutive heterochromatin in adult and fetal tissues, *Hum. Genet. 99*, 738−745.
37. Jones, K. W., and Corneo, G. (1971) Location of satellite and homogeneous DNA sequences on human chromosomes, *Nat. New Biol. 233*, 268−271.
38. Jones, K. W., Prosser, J., Corneo, G., and Ginelli, E. (1973) The chromosomal location of human satellite DNA 3, *Chromosoma 42*, 445−451.
39. Ballestar, E., Paz, M. F., Valle, L., Wei, S., Fraga, M. F., Espada, J., Cigudosa, J. C., Huang, T. H.-M., and Esteller, M. (2003) Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer, *EMBO J. 22*, 6335−6345.
40. Costello, J. F., Frühwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomäki, P., Lang, J. C., et al. (2000) Aberrant CpG-island methylation has non-random and tumour-type-specific patterns, *Nat. Genet. 24*, 132−138.
41. Poljak, L., Seum, C., Mattioni, T., and Laemmli, U. K. (1994) SARs stimulate but do not confer position independent gene expression, *Nucleic Acids Res. 22*, 4386−4394.
42. Namciu, S. J., Blochlinger, K. B., and Fournier, R. E. K. (1998) Human matrix attachment regions insulate transgene expression from chromosomal position effects in *Drosophila melanogaster*, *Mol. Cell. Biol. 18*, 2382−2391.
43. Betz, A. G., Milstein, C., Gonzáles-Fernández, A., Pannell, R., Larson, T., and Neuberger, M. S. (1994) Elements regulating somatic hypermutation of an immunoglobulin κ gene: critical role for the intron enhancer/matrix attachment region, *Cell 77*, 239−248.
44. Jenuwein, T., Forrester, W. C., Fernández-Herrero, L. A., Laible, G., Dull, M., and Grosschedl, R. (1997) Extension of chromatin accessibility by nuclear matrix attachment regions, *Nature 385*, 269−272.
45. Georgel, P. T., Horowitz-Scherer, R. A., Adkins, N., Woodcock, C. L., Wade, P. A., and Hansen, J. C. (2003) Chromatin compaction by human MeCP2, *J. Biol. Chem. 278*, 32181−32188.
46. Kochanek, S., Renz, D., and Doerfler, W. (1995) Transcriptional silencing of human Alu sequences and inhibition of protein binding in the box B regulatory elements by 5′-CG-3′ methylation, *FEBS Lett. 360*, 115−120.

47. Shiraishi, M., Chuu, Y. H., and Sekiya, T. (1999) Isolation of DNA fragments associated with methylated CpG islands in human adenocarcinomas of the lung using a methylated DNA binding column and denaturing gradient gel electrophoresis, *Proc. Natl. Acad. Sci. U.S.A. 96*, 2913−2918.
48. Shiraishi, M., Sekiguchi, A., Chuu, Y. H., and Sekiya, T. (1999) Tight interaction between densely methylated DNA fragments and the methyl-CpG binding domain of the rat MeCP2 protein attached to a solid support, *Biol. Chem. 380*, 1127−1131.
49. Yu, F., Zingler, N., Schumann, G., and Strätling, W. H. (2001) Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription, *Nucleic Acids Res. 29*, 4493−4501.
50. Liu, W. M., Chu, W. M., Choudary, P. V., and Schmid, C. W. (1995) Cell stress and translational inhibitors transiently increase the abundance of mammalian SINE transcripts, *Nucleic Acids Res. 23*, 1758−1765.
51. Woodcock, D. M., Lawler, C. B., Linsenmeyer, M. E., Doherty, J. P., and Warren, W. D. (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon, *J. Biol. Chem. 272*, 7810−7816.
52. Bratthauer, G. L., Cardiff, R. D., and Fanning, T. G. (1994) Expression of LINE-1 retrotransposons in human breast cancer, *Cancer 73*, 2333−2336.
53. Buhrmester, H., von Kries, J. P., and Strätling, W. H. (1995) Nuclear matrix protein ARBP recognizes a novel DNA sequence motif with high affinity, *Biochemistry 34*, 4108−4117.
54. Heitmann, B., Maurer, T., Weitzel, J. M., Strätling, W. H., Kalbitzer, H. R., and Brunner, E. (2003) Solution structure of the matrix attachment region-binding domain of chicken MeCP2, *Eur. J. Biochem. 270*, 3263−3270.
55. Shahbazian, M. D., Antalffy, B., Armstrong, D. L., and Zoghbi, H. Y. (2002) Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation, *Hum. Mol. Genet. 11*, 115−124.
56. Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., Fan, G., and Sun, Y. E. (2003) DNA methylation-related chromatin remodeling in activity-dependent *Bdnf* gene regulation, *Science 302*, 890−893.

BI0359271