Purification of *BsuE* methyltransferase and its application in genome mapping

Hridayabhiranjan Shukla, Yasushi Kobayashi¹, Hartwig Arenstorf, Yukio Yasukochi¹ and Sherman M. Weissman*

Yale University School of Medicine, Department of Human Genetics, New Haven, CT, USA and

¹Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan

Received March 25, 1991; Revised and Accepted July 9, 1991

ABSTRACT

We have used a combination of *BsuE* methyltransferase (*M-BsuE*) and *NotI* restriction enzyme to cut genomic DNA at a subset of *NotI* sites. The usefulness of this system is shown in a re-examination of the restriction map of the human MHC. Combinations of methylases and restriction enzymes can be used to generate cuts at different frequencies in genomic DNA, such that they generate ends complementary to *NotI* ends, and can be used in conjunction with *NotI* linking clones in chromosome jumping experiments. These enzyme combinations have the potential to produce cutting sites in genomic DNA spaced at intervals favorable for extensive mapping, fragment enrichment, and cloning efforts.

INTRODUCTION

The average human chromosome contains about 150 million base pairs (bp) of DNA. Current gel electrophoretic techniques can resolve linear DNA molecules up to a size of about 6 million bp. *NotI* appears to cut genomic DNA less frequently than other available restriction enzymes; but even this enzyme cuts at least once every million bp on the average, and there are already several regions known where two *NotI* sites fall in close proximity to one another. A restriction system that cuts genomic DNA somewhat less frequently than *NotI* would help greatly in long-range mapping, by providing linking clones with which to rapidly scan large regions of DNA for major deletions or rearrangements. These would provide a scaffolding on which to map more frequent cutting sites or polymorphic markers, and generate large fragments for enrichment and subcloning.

A number of schemes have been proposed for creating very infrequent cutting sites, either by use of deoxyadenylic acid methylases that create cutting sites for enzymes (1), by binding proteins or nucleic acids to a predetermined sequence (2, 3), or by using methylases whose recognition sites overlap a given restriction enzyme's site and therefore block a fraction of such sites (4, 5). Gaido, Prostko and Strobl had suggested the use of

M.BsuE, which methylates the first C in the sequence CGCG, for long-range restriction mapping of eukaryotic genomic DNA (6). None of these systems had come into general use, partly because sufficiently pure enzymes were not available, and partly because of limitations such as occasional cutting of DNA at hemimethylated sites (7). While this study was in progress two methylases, *M-FnuDII* and *M-BepI*, with specificities similar to *M-BsuE*, have been shown to increase the sizes of *NotI* fragments in bacteria (8). In the present report we describe an improved purification scheme for the enzyme *M-BsuE* (6), and demonstrate that our preparation of this enzyme is sufficiently pure to be used in conjunction with *NotI* to generate very large restriction fragments.

The human major histocompatibility complex (MHC) is a cluster of genes and gene families encoding a number of products important for the function of the immune system. It is also of interest because several diseases occur more frequently in individuals carrying particular alleles of MHC genes (9). The MHC is one of the most intensively studied regions of the human chromosome and one of the first regions to which long-range mapping by pulsed field gel electrophoresis (PFGE) was applied (10–15). The complex is thought to cover three or four megabases and over one megabase of the MHC has been cloned in cosmids (16–18). Nevertheless, uncertainties exist in the published physical map due to failure to obtain definitive overlap between macro-restriction fragments in the telomeric region of the complex. We apply the *NotI* *M-BsuE* system to mapping of the MHC and demonstrate that it can be used to map *NotI* sites lying between two *NotI* sites that are resistant to inhibition by methylation.

Other restriction enzymes are available that produce overhanging ends on DNA fragments that have the same sequence as those produced by *NotI*. We discuss how various methylases might be used in combination with these enzymes to generate cutting sites that would be expected to occur at frequencies comparable to or greater than that of *NotI* sites, and therefore be useful in chromosome jumping and mapping experiments in conjunction with *NotI* linking clones.

* To whom correspondence should be addressed

METHODS

Bacterial Strains and DNA

Bacillus subtilis strain ISE15 was kindly provided by Dr. S. Ikawa (Institute of Physical and Chemical Research, Wako, Japan). Poly(dGdC)-poly(dGdC) DNA was obtained from Pharmacia LKB Biotechnology (Uppsala, Sweden).

Enzyme, Reagents, and Chromatography Supplies

Amino acids, sodium lauroylsarcosine, Tris HCl, EDTA and phenyl methyl sulfonyl fluoride were purchased from Sigma Chemical Co. (St. Louis, MO). S-[Methyl-³H] adenosyl-L-methionine (12.4 Ci/mmol) was obtained from New England Nuclear (Boston, MA). Yeast extract, nutrient agar, and glucose were from Difco (Detroit, MI). RPMI, Fetal Calf Serum, and antibiotics for tissue culture were procured from Gibco Laboratories (USA). All other chemicals were purchased from Wako (Osaka, Japan). Restriction enzymes were from New England Biolabs or Stratagene (USA). Proteinase K and different grades of Bovine Serum Albumin were obtained from Boehringer Mannheim. Various grades of agarose were purchased from FMC (USA). DEAE-Sepharose was obtained from Pharmacia LKB Biotechnology (Uppsala, Sweden). HPLC columns, Asahipak ES-502N and TSK gel Heparin-5PW were from Asahikasei (Kawasaki, Japan) and Tosoh (Tokyo, Japan), respectively.

Growth of Bacterial Cells

Bacillus subtilis ISE15 cells were grown at 37°C in modified Spizien's medium containing 0.25% yeast extract, 1% lactate, 0.02% MgSO₄, 0.5% glucose and 40 ml/l of 25×minimum salts (75 g/l KH₂PO₄, 175 g/l K₂HPO₄, 12.5 g/l sodium citrate-dihydrate and 25 g/l (NH₄)₂SO₄), supplemented with 40 mg/l leucine and 45 mg/l methionine to support the growth of strain ISE15. Methyltransferase was purified from 12 l of cells in early logarithmic growth (O.D._{600nm} = 0.5). Preliminary experiments showed that production of methylase was maximum in the early exponential phase of cell growth, as also observed by Strobl (personal communication). Therefore the bacterial cells were harvested at O.D._{600nm} = 0.5.

Methyltransferase Assay

The radiochemical assay for *Bsu*-E methyltransferase (*M-BsuE*) measures the activity of transfer of ³H-methyl groups from S-[methyl-³H] adenosyl-L-methionine to DNA. Poly(dGdC)-poly(dGdC) was used as a DNA substrate for estimating methyltransferase activity of *M-BsuE*. The reaction mixture contained 50 mM Tris buffer, pH 7.8, containing 10 mM EDTA, 5 mM 2-mercaptoethanol, 0.5 μg of poly(dGdC)-poly(dGdC), 2.2 μM S-[methyl-³H] adenosyl-L-methionine containing 1.1 μCi, and enzyme in a final volume of 40 μl.

Reactions were initiated by the addition of S-[methyl-³H] adenosyl-L-methionine and incubated at 37°C for 1 hour. Enzyme reactions were terminated by the addition of 10 μl of 4 M NaCl; the reaction mixtures were spotted on DEAE-paper (1×1 cm) and washed four times with a total volume of 4 l of 0.35 M Na₂HPO₄. The filters were rinsed with water and ethanol and air-dried. Finally, they were placed in scintillation fluid and assayed for radioactivity.

Protein Estimation

Protein concentration was estimated by the dye method described by Bradford (19) or by absorbance at 280 nm with bovine serum albumin as standard.

Gel Electrophoresis

M-BsuE protein purification was monitored by analysis on 10% SDS-polyacrylamide gel electrophoresis (20). Samples were dissolved in sample buffer, boiled for 5 minutes, and loaded onto an SDS-polyacrylamide gel with a 4% stacking gel. The gels were stained with silver staining Kit 2D-Silver Stain II 'DAIICHI' from Daiichi (Tokyo, Japan).

Enzyme and Methylation Unit

During purification of enzyme, one Unit of methyltransferase activity (U) is defined as the amount of enzyme that transfers 10⁶ cpm of ³H-methyl groups to poly(dGdC)-poly(dGdC) DNA per minute.

One methylation Unit (MU) is defined as the amount of enzyme required to completely protect 1 μg of pBluescript KS(−) DNA from digestion with AccII after one hour of methylation at 37°C.

Genomic DNA Preparation

For all experiments EBV transformed human lymphoblast cells, line JY (HLA-A2,2; B7,7 CW-,-; DR4, W6) were cultured in RPMI 1640 with 10% (V/V) Fetal Calf Serum and 1% penicillin-streptomycin. To make DNA in agarose plugs, cells were harvested and washed twice with phosphate buffered saline, containing 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄·7H₂O, and 1.4 mM KH₂PO₄ (pH 7.3) at 4°C. These cells were suspended at a density of 2×10⁷ cells/ml at 37°C, mixed with an equal volume of 1% low-melting agarose gel at 45°C and cast into 125 μl plugs. Cells were lysed within the plugs by suspension in 500 mM EDTA (pH 9.2), 1% sodium lauroylsarcosine and 1 mg/ml proteinase K (ESP) at 50°C for 48 hours. Proteinase K was quenched in 10 mM Tris HCl 1 mM EDTA (pH 7.5) (T₁₀E₁) containing 1mM phenylmethyl sulfonyl fluoride (PMSF). PMSF was removed by extensive dialysis of plugs against T₁₀E₁.

Methylation and Other Enzyme Reactions

Methylation of DNA in agarose plugs was carried out in 50mM Tris HCl (pH 7.8), 10mM EDTA, 5mM β mercaptoethanol and 80 μM S-adenosyl methionine (SAM) in a reaction volume of 250 μl for 12 hours at 37°C with additional aliquot of 80μm SAM after 6 hours. In some cases the above cycle was repeated one more time. A half plug (about 5 μg of DNA) was used per reaction. After the methylation the enzyme was inactivated with ESP at 50°C for 3–5 hours. Proteinase K and PMSF were removed as described above. *NotI* digestion was carried out overnight at 37°C with excess of enzyme in 150mM NaCl, 10mM Tris HCl (pH 7.8), 10mM MgCl₂, and 250μg/ml bovine serum albumin.

Pulsed Field Gel Electrophoresis

Digested DNA in agarose plugs was fractionated on agarose gels using either Stratagene's RAGE or BioRad's CHEF Mapper electrophoresis system. Both of these systems basically provide a clamped homogeneous electrical field within the gel (21). Various pulse timings and total electrophoresis times were used. In all cases, a 1% agarose gel in 0.5×TBE was used at 14°C. Approximately one-fourth of a plug was loaded per lane. Specific conditions are mentioned in Figure legends.

Nucleic Acid Blotting and Hybridization

After the separation of DNA in gels it was treated with 0.1 N HCl for 20 minutes at room temperature for depurination of the

DNA to facilitate the transfer. Gels were blotted on Hybond-N (Amersham). After transfer, DNA was crosslinked to membrane by 1.6×10^5 μ J of U.V. irradiation in a U.V. crosslinker (Stratalinker).

Membranes were prehybridized in 0.5mM sodium phosphate, 7% SDS, 1% BSA and 1mM EDTA (pH 7.2) (22) containing 250 μ g/ml denatured herring sperm DNA at 65°C for 12 hours. After 12 hours, denatured probe was added to the same solution and hybridization was carried out for another 12 hours at 65°C.

DNA probes were labeled with [α^{32} P] dCTP by random hexanucleotide priming (23).

Probe non-specifically bound to the filter was removed by washing at 65°C with 0.5 or 0.2 $\times$ SSPE and 0.5% SDS.

Purification of M-BsuE

The purification of M-BsuE from 8 grams of packed *B. subtilis* strain ISE15 was performed at 4°C and is summarized in Table 1. The bacterial cells were harvested at optical density 0.5 (A_{600}) by centrifugation at 10,000g for 10 minutes. The bacterial pellet was re-suspended in 120 ml of sonication buffer (20 mM Tris-HCl pH7.8, 0.1 mM EDTA, 10 mM 2-mercaptoethanol) and sonicated on ice (10 $\times$ 1 min.) using a Branson sonifier 250 (Output 4 units, 40% Duty Cycle). The sonicate was centrifuged at 30,000g for 30 min. to yield the cell-free extract.

Crude extract was added dropwise to 30 ml of 25% streptomycin sulphate solution on ice. After 30 minutes the mixture was centrifuged at 30,000g for 30 minutes to remove nucleic acids. The 130 ml supernatant was precipitated by the addition of 65g ammonium sulfate on ice and centrifuged at 30,000g for 30 minutes. The pellet was re-suspended in 50 ml dialysis buffer (20 mM Tris-HCl pH7.8, 5 mM EDTA, 10 mM 2-mercaptoethanol, 5% glycerol, 70 mM NaCl) and dialyzed overnight against 5 l of the buffer.

The dialysate (57 ml) was loaded onto a 2.2 cm $\times$ 20 cm DEAE-sepharose column pre-equilibrated with 20 mM Tris-HCl pH7.8, 5mM EDTA, 1mM 2-mercaptoethanol, 5% glycerol buffer containing 100 mM NaCl, then eluted using a total volume of 600 ml by a linear gradient from 100 to 200 mM NaCl. M-BsuE activity eluted at 160 mM NaCl. The active fractions were pooled and dialyzed overnight against 5 l of dialysis buffer containing 20% glycerol instead of 5% glycerol.

The dialysate (117 ml) was loaded onto a Heparin-HPLC (TSKgel Heparin-5PW) column pre-equilibrated with 20 mM Tris buffer, pH 7.8, containing 190 mM NaCl, 5 mM EDTA, 1 mM 2-mercaptoethanol and 20% glycerol, and then eluted with a 79 ml 190–300 mM NaCl linear gradient. M-BsuE activity eluted at 280 mM NaCl. A part of the active fractions (4.5 ml of 12.9 ml) was loaded onto a DEAE-HPLC (Asahipak ES-502N)

column pre-equilibrated with 20 mM Tris buffer, pH 7.8, containing 70 mM NaCl, 5 mM EDTA, 1 mM 2-mercaptoethanol and 20% glycerol after dialysis against the equilibration buffer, and then eluted with a 91 ml 70–200 mM NaCl linear gradient. M-BsuE activity eluted at 170 mM NaCl.

RESULTS

To analyze complex genomes we first purified M-BsuE sufficiently so that it did not cause degradation of the genomic DNA. The purification of M-BsuE as discussed in the Methods section is summarized in Table 1. In successive column chromatographic separation on DEAE-Sephadex and heparin-HPLC, M-BsuE was well separated from the bulk of the other proteins. In the heparin-HPLC step the enzyme was extensively purified, as seen in Fig. 1. However, considerable loss of the total and specific activities was observed although the buffer contained 20% glycerol. SDS-PAGE of the heparin-HPLC purified enzyme showed one major and one minor band with apparent molecular weights of 46 K and 56 K respectively.

We tried to separate M-BsuE on a DEAE-HPLC column to determine which band is responsible for the activity. As seen in Fig. 1, only the 46 K polypeptide band was detected in the DEAE-HPLC purified enzyme on SDS-PAGE, but the specific activity was reduced to 22% of that in the heparin-HPLC purified enzyme. This might indicate that the highly purified enzyme was very labile as reported by Gaido et al. (6). Therefore we used the heparin-HPLC purified M-BsuE for analysis of the genomic DNA.

The purified enzyme did not produce detectable cleavage of human genomic DNA as judged by PFGE of DNA plugs that had been incubated in the presence of the enzyme. Preliminary tests showed that plasmid DNA containing a *NotI* site that overlapped with a M-BsuE restriction site was almost completely

Figure 1. SDS-PAGE analysis of M-BsuE at each purification step. From each purification step, an aliquot of M-BsuE was analyzed on 10% SDS-PAGE and visualized by silver staining. 1. Approximately 3 μ g of crude extract. 2. Approximately 2 μ g of streptomycin treatment of crude extract. 3. Approximately 5 μ g of ammonium sulfate precipitate. 4. Approximately 0.5 μ g of DEAE-Sephadex fraction. 5. Approximately 0.1 μ g of heparin-HPLC fraction. 6. Approximately 0.1 μ g of DEAE-HPLC fraction. An arrow indicates the 46K polypeptide band, corresponding to methylase.

Table 1. Purification of M-BsuE from *Bacillus subtilis* strain ISE15

Steps	Protein (mg)	Specific Activity (units/mg)	Total Activity (units ^a)	Yield (%)	Fold Purification
Crude extract	380	0.215	81.8	100	1
Streptomycin treatment	280	0.479	134.0	164	2.2
DEAE-Sephadex	23	1.64	37.8	46	7.6
Heparin-HPLC	0.09	29.1	2.62	3.2	135

^a During the purification, one unit was defined as the amount of enzyme which transferred 10^6 cpm of 3 H-methyl groups from S-[methyl- 3 H] adenosyl-L-methionine to poly(dGdC)-poly(dGdC) per minute.

protected from digestion with *NotI* by prior methylation (data not shown).

The JY line of human B lymphoblastoid cells was used for studies to evaluate the usefulness of the combined *M-BsuE NotI* restriction enzyme system (*Not-Bsu*) for long-range mapping. This line is heterozygous in the MHC Class II region and also in the Class I region (24), although it expresses only HLA-A2

Table 2.

	Probes from the Human Major Histocompatibility Complex Used to Analyze DNA Blots	Fragment Sizes in kb Observed With Probe On PFGE Blot
<i>HLA-E</i> (pHLAE.a)	A 240 bp <i>PvuII</i> - <i>AluI</i> fragment subcloned from the 3' untranslated region of the <i>HLA-6.2</i> gene designated as <i>HLA-E</i> . (27)	300
<i>RS 5.10</i> (pRS 5.10)	A 0.7 kb <i>EcoRI</i> , intergenic genomic fragment upstream of <i>HLA-E</i> . (28)	300
<i>HS-2</i> (pHS-2)	A 1.85 kb <i>EcoRI</i> - <i>BamHI</i> fragment upstream of <i>HLA-G</i> . (29)	680
<i>pMC27.7</i>	A second, 700 bp <i>EcoRI</i> fragment located approximately 17 kb upstream of <i>HLA-G</i> . (29)	680
<i>pMC6.7</i>	A 700 bp <i>EcoRI</i> fragment located approximately 21 kb, upstream of <i>HLA-H</i> . (29)	680
<i>pHLA 2AI</i>	A 490 bp <i>PvuII</i> <i>MspII</i> fragment derived from the 3' untranslated region of the <i>HLA-A₂</i> gene. (30)	680
<i>RS 30.4</i> (pRS 30.4)	A 3.4 kb <i>EcoRI</i> , intergenic genomic fragment upstream of <i>HLA-B</i> . (10)	1000 & 1300
<i>TNF</i>	A 238 bp long unique DNA sequence, amplified in a polymerase chain reaction by using specific primers from positions 1967 and 2205 of the <i>TNF</i> gene. (31)	300
400 bp (p Prelink-6)	A 400 bp <i>EcoRI</i> - <i>BamHI</i> genomic fragment from a linking clone which was specific to a <i>NotI</i> fragment between the <i>DR₁</i> and <i>B</i> loci of human MHC. The linking clone was isolated from cosmid T27A by Spies <i>et al.</i> (16)	1100 & 1200
1 kb	A 1 kb <i>BamHI</i> - <i>NotI</i> fragment from the telomeric end of the <i>NotI</i> fragment on which <i>DR</i> is located. This clone was isolated from cosmid T27A, described by Spies <i>et al.</i> (16)	~290
<i>DP</i> (p8MI)	A 4.1 kb <i>MspI</i> genomic fragment extending from <i>DP</i> exon 1 to <i>DPβ</i> exon 1. (32)	430
<i>DRα</i> (pp 34R13)	A 3.1 kb <i>EcoRI</i> genomic fragment, containing exon 2 to exon 5 and the 3' untranslated region of the <i>HLA DR</i> gene. (33)	1100 & 1200
<i>AH2</i> (pAH-2)	A full-length <i>B₇</i> cDNA isolated from a JY cDNA library in Charon BS. (34)	300, 680, 1000, 1300 & >1600

Note: All these probes were used to analyze the pulsed field genomic DNA blots, which confirm the map shown in Fig. 3.

and *HLA-B7* alleles. Consistent with earlier reports (15), *NotI* digests probed with a *DR* alpha probe (Class II) show that the line contains two Class II fragments of size 1100 and 1200 kilobases (kb). The larger fragment is not a partial digestion product because it does not hybridize with probes derived from the *NotI* fragments on either side adjacent to the *DR* alpha fragment. Probing with an *HLA-B* locus-specific probe showed two large restriction fragments of size 1000 and 1300kb. The larger fragment contains the *TNF* gene as well as *HLA-B* (Table 2). Therefore, it is probably a result of partial unavailability of the *NotI* site separating the 300kb *TNF* fragment from the 1000kb *HLA-B* fragment.

Hybridization with a more general Class I probe (Table 2) (either the *HLA-H* gene or an *HLA-B7* full-length cDNA clone) consistently showed six fragments. In addition to the two fragments of 1000 and 1300kb described above, there was a 300kb fragment that contained *HLA-E*, a faint 330 band that was not detected with an *HLA-E* probe, a 680kb fragment that contained *HLA-G*, *HLA-H* and *HLA-A*, and a >1600kb fragment that had not been previously reported in this or other cell lines. The largest fragment did not hybridize with either an *HLA-B* or an *HLA-E* probe but did hybridize with both *HLA-F*, *HLA-H*, and *HLA-A* probes, so that it may represent a fusion fragment although we cannot entirely rule out the possibility that it represent an additional fragment containing DNA sequences that cross hybridize with several probes from the region of the known Class I genes.

The different *NotI* restriction fragment sizes we observed were consistently somewhat higher than reported values (15). For example, the *NotI* fragment on which the heat shock protein is located has been shown to be approximately 130–150 kb by cosmid cloning (16), whereas our observed size is ~290 kb (1 kb in Table 2). Probably these variations arise because our lanes are more heavily loaded (25). In addition, they could be due to

Figure 2. a. Partial sequence of *NotI* linking clone showing that the *DR* region and heat shock gene region of the human MHC are separated by *NotI* and *M-BsuE/NotI* sites in close proximity. b. 1.6 kb *NotI* linking fragment clone in puc19 at the *BamHI* site. (1) Uncut plasmid. (2) *BamHI/NotI* cut plasmid. (3) *M-BsuE* treated plasmid. (4) *M-BsuE* treated plasmid cut with *BamHI* and *NotI*. c. JY genomic DNA blot probed with 1 kb probe. (1) Uncut JY DNA. (2) JY DNA cut with *BamHI*. (3) JY DNA cut with *BamHI/NotI*. (4) JY DNA treated with *M-BsuE* and cut with *BamHI*. (5) JY DNA treated with *M-BsuE* and cut with *BamHI/NotI*.

allelic variation, especially in the DR region, or to different extents of methylation of *NotI* sites, as in the case of the larger HLA-B containing fragments. For these reasons, the sizes given in this report should not be taken as absolute values.

Only a fraction of *NotI* sites remain unblocked after *BsuE* methylation. Unless methylated *in vivo*, these sites will be cut completely by *NotI*, and therefore can be used as end points for mapping adjacent *NotI* sites in partially methylated DNA by the Smith-Birnsteil procedure (26). To evaluate the feasibility of methylation as a mapping procedure, we first subcloned the *Bam*HI fragment spanning the telomeric end of the DR alpha fragment and the adjacent fragment containing a heat shock gene. The partial sequence of this fragment (Fig. 2a) shows that it contains two *NotI* sites 70 base pairs apart. The sequence further predicts that one of these sites should be inhibited by *M-BsuE* because it contains a CGCG sequence overlapping the GCGGCCGC site of *NotI*, while the other *NotI* site does not contain a *M-BsuE* recognition site and so should be resistant to

Figure 3. The sizes of the markers shown on the side of each blot are in kilobases. Partial methylation of JY genomic DNA cut with *NotI* after successive methylation by the larger amounts of *M-BsuE*. In Fig. 3a, lane 1 is uncut DNA, lane 2 is DNA cut with *NotI*; lanes 3 to 7 are DNA cut with *NotI* after 1MU, 2MU, 10 MU and 20 MU of *M-BsuE* treatment successively. In Fig. 3b, lanes 1 and 2 are the same as above; in lanes 3 through 9, DNA is cut with *NotI* after 2.5MU, 5MU, 10MU, 15MU, 20MU, 25MU and 30MU of *M-BsuE* treatment, respectively. The blot in Fig. 3a is probed with B₇ full-length cDNA probe AH₂. The blot in Fig. 3b is probed with HLA2A1. The position of the arrows indicates the methylated product. Electrophoresis was performed on Biorad's CHEF mapper in 0.5×TBE buffer at 14°C in 1% agarose gel. Pulse time was increased from 29.19 seconds to 3 min. 49 seconds at 5 V/cm field strength with a field angle of 120°, over a period of 33 hours.

Figure 4. Partial Restriction Map of Human Major Histocompatibility Complex. a. Megabase Scale. b. Genes: DP, DR are the Class II genes encoding alpha and beta chains. C2 is the gene for complement component 2. TNF is the tumor necrosis factor gene. The remainder of the letters refer to Class I genes. c. *NotI* sites. The arrow and heavy vertical line indicate the position of two *NotI* sites within 70 base pairs, as described in the text. The fragments in parentheses correspond to lighter bands on the PFGE blot. Their position and relation to the other fragments is presently uncertain. The hatched line indicates the position where there is potentially another *NotI* site between HLA-A and HLA-F. The centromere of chromosome 6 is to the left.

methylase inhibition. These predictions have been confirmed both in the cloned *Bam*HI fragment and in the corresponding fragment from genomic DNA (Fig. 2 b,c). Thus the human MHC was cut at this site by *NotI* even after *M-BsuE* treatment with a 10-fold excess of the enzyme. JY DNA in agarose plugs was methylated with *M-BsuE* digested with *NotI*, run under a variety of PFGE programs to display DNA fragments of various sizes, and probed with DNA fragments representing each of the known *NotI* fragments of the MHC. Partial methylation was performed by incubating DNA embedded with agarose plugs in various dilutions of the *M-BsuE* but with other conditions kept the same as in the complete methylation experiments.

After methylation, the site centromeric to the TNF fragment was not detectably protected so that the linking clone continued to detect a ~290kb fragment. Other *NotI* sites in the Class I region were quite sensitive to the methylase. For example a general Class I probe consistently detected a new ~945kb fragment after treatment even with low levels of the methylase (Fig. 3a). This fragment hybridized with an HLA-A probe. One interpretation is that it represents a fusion fragment containing the HLA-E fragment (300kb) and the HLA-A fragment (680kb), placing these fragments very near one another (Fig. 4). Similarly the >1600kb fragment was quantitatively removed, presumably shifting into the compression zone (Fig. 3b).

In addition to *NotI* the restriction enzymes *EagI* and *EaeI* produce the 5' overhang sequence GGCC when they digest DNA. Fragments derived from digestion of genomic DNA with *NotI* on one end and with one of these other enzymes on the other will therefore circularize and could be used in chromosome jumping experiments. These enzymes tend to cut genomic DNA more frequently than is useful for long-range mapping and jumping experiments. However, they may be used in combination with methylases to generate systems that cut less frequently (Table 3). With the exception of *Fnu4H* methylase, we have confirmed the specificity of the combinations of methylase and restriction enzymes on a plasmid containing the HLA-H gene (data not shown). Methylation of DNA with *HpaII* methylase prior to cleavage with *EagI* will prevent cutting at a significant percentage of the *EagI* sites, because most occur in HTF islands. Cutting sites that occur at a frequency intermediate between that of *EagI* and that of *EaeI* can be generated by treating DNA with a general CG methylase such as the *Mss* methylase prior to treatment with *EaeI*.

DISCUSSION

Gaido *et al.* (6) showed that *M-BsuE* methylates the external cytidine in the sequence CGCG, which suggested that *M-BsuE* should be useful for blocking a subset of *NotI* sites. We have shown that *M-BsuE* does indeed block cleavage at overlapping

Table 3. List of Combinations of Restriction Enzyme and Methylase That Generate Cutting Sites Compatible With *NotI* Ends

Enzyme Combination	Cleavage Site*
<i>BsuE</i> methylase- <i>NotI</i>	DGCGGCCGCH
<i>NotI</i>	GCGGCCGC
<i>HpaII</i> methylase- <i>EagI</i>	DGCGCCGH
<i>Fnu4H</i> methylase- <i>EagI</i>	HCGGCCGD
<i>EagI</i>	CGGCCG
<i>Mss</i> -methylase- <i>EaeI</i>	TGGCCA

* D = G, A, or T
H = C, A, or T

NotI sites, but it remains uncertain whether M-*BsuE* is a (cytosine-N⁴)methyltransferase or a 5-methyltransferase. Although both Gaido *et al.* (6) and Oda and Marmur (35), who earlier purified M-*BsuE*, assume that the methylation is at position 5 of cytidine, the experimental evidence that they presented does not rule out the possibility of N⁴ methylation. (That the methylase purified by Oda and Marmur (35) is indeed M-*BsuE* follows from the demonstration by Ikawa *et al.* (36) that *B. subtilis* ATCC6633 has only the *BsuE* (*Bsu1231I*) modification and restriction system.)

Our present results show that the combination of M-*BsuE* and *NotI* restriction endonuclease provides a system that will cut genomic DNA specifically at relatively infrequent sites without producing any non-specific nicking or cutting at a level that interferes with PFG analysis. In the data presented above as well as other unpublished results this combination has frequently generated genomic fragments in excess of 2 megabases in size that are difficult to resolve by conventional PFGE. This suggests that a substantial number of the *NotI* sites in DNA from cell lines that are not methylated by endogenous enzymes are susceptible to methylation by the *BsuE* enzyme. Relaxed specificity of the enzyme was not seen *in vitro* and a test *NotI* site in genomic DNA that did not overlap an M-*BsuE* site was resistant to methylation. The number of methylase-sensitive sites that occur in genomic DNA could not be accurately estimated because of the bias in base composition in GC rich islands in genomic DNA and the possibility that there may be further bias in the fraction of those islands that are completely unmethylated in passaged cell lines. In the MHC three of ten sites evaluated were blocked by methylation.

The use of partial methylation and complete *NotI* digestion together with probes derived from *NotI* linking clones resistant to the methylase should in principle provide an excellent opportunity for mapping the *NotI* sites in the vicinity of a linking clone by standard methods. The results with the MHC show that this approach can give useful information, but is limited by the resolution of large fragments by PFGE. The use of various combinations of methylase and restriction enzymes (Table 3) generates a number of sites that occur with desired frequencies in the genome, and that produce overhanging DNA ends compatible with *NotI* ends. These sites are of considerable potential use in mapping regions of DNA that contain few available *NotI* sites, and for application of chromosome jumping (37) to proceed out from *NotI* sites. In addition, because there are a variety of combinations available, there is an increased likelihood that a system can be found for generating DNA fragments in any region of interest that are of a size suitable for taking maximal advantage of affinity capture (38) or yeast artificial chromosome cloning techniques (39).

Although the *Fnu4H* methylase is not available, it would be expected to protect *EagI* sites that are preceded by a G or followed by a C. If used together with *HpaII* methylase it would then generate the equivalent of a seven base pair recognition sequence that contained two CG dinucleotides as well as an A and a T residue. This site would be expected to be of the same order of frequency as is the *NotI* site and could prove useful in long-range mapping.

The results from partial methylation mapping of the human MHC Class I region are consistent with earlier maps. In particular, they suggest that there is no large fragment between the HLA-E and HLA-A containing *NotI* fragments, and that earlier estimates that the size of the MHC from DP to HLA-A

(and HLA-G and -H) was about 3.2 to 3.4 megabases were not markedly low. Other studies from this lab suggest that HLA-F lies in a separate *NotI* fragment telomeric to, but within 400 kb of, the HLA-A fragment. Thus the Class I genes occupy a region of more than one but probably less than two megabases. In view of the high gene density found in the most extensively examined parts of the MHC we might anticipate that a number of additional new genes will be found also in this region of the MHC (40).

ACKNOWLEDGEMENTS

We want to thank Dr. Thomas Spies for providing us with the cosmid clones spanning the *NotI* site telomeric to the DR locus, and Dr. Rajendra P. Kandpal for lively discussion and suggestions. This work was supported by National Institutes of Health Grant R35 CA42556-05, by Department of Energy Grant DE-FG02-88ER60687, and by grants from the Ministry of Education, Science and Culture of Japan.

REFERENCES

- McClelland, J., Kessler, L.G., and Bittner, M. (1984) *Proc. Natl. Acad. Sci. (USA)* **81**, 983–87.
- Koob, M., and Szybalski, W. (1990) *Science* **250**, 271–73.
- Maher, L.J.3d., Wold, B., and Dervan, P.B. (1989) *Science* **245**, 725–30.
- Nelson, M., Christ, C., and Schildkraut, I. (1984) *Nucleic Acids Res.* **12**, 5165–73.
- Nelson, M., and McClelland, M. (1989) *Nucleic Acids Res.* **17** Suppl., 389–415.
- Gaido, M.L., Prostko, C.R., and Strobl, J.S. (1988) *J. Biol. Chem.* **263**, 4832–4836.
- Davis, T.B., Morgan, R.D., and Robinson, D.P. (1990) Abstract #30, Human Genome II.
- Qiang, B.Q., McClelland, M., Poddar, S., Spokauskas, A., and Nelson, M. (1990) *Gene* **88**, 101–105.
- Svejgaard, A., Platz, P., and Ryder, L.P. (1983) *Immunol. Rev.* **70**, 193–218.
- Lawrance, S.K., Smith, C.L., Srivastava, R., Cantor, C.R., and Weissman, S.M. (1987) *Science* **235**, 1387–90.
- Dunham, I., Sargent, C.A., Trowsdale, J., and Campbell, R.D. (1987) *Proc. Natl. Acad. Sci. (USA)* **84**, 7237–41.
- Hardy, D.A., Bell, J.I., Long, E.O., Lindsten, T., and McDevitt, H.O. (1986) *Nature (London)* **323**, 453–55.
- Carroll, M.C., Katzman, P., Alicot, E.M., Koller, B.H., Geraghty, D.E., Orr, H.T., Strominger, J.L., and Spies, T. (1987) *Proc. Natl. Acad. Sci. (USA)* **84**, 8535–39.
- Ragousis, J., Bloemer, K., Phola, H., Messer, G., Weiss, E.M., and Ziegler, A. (1989) *Genomics* **4**, 301–308.
- Lawrance, S.K. and Smith, C.L. (1990) *Genomics* **8**, 394–99.
- Spies, T., Bresnahan, M., and Strominger, J.L. (1989) *Proc. Natl. Acad. Sci. (USA)* **86**, 8955–8958.
- Kappes, D. and Strominger, J.L. (1988) *Ann. Rev. Biochem.* **57**, 991–1028.
- Koller, B.H., Geraghty, D.E., Orr, H.T., Shimizu, Y., DeMars, R. (1987) *Immunol. Res.* **6**, 1–10.
- Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248–254.
- Laemmli, U.K. (1970) *Nature* **227**, 680–685.
- Chu, G., Vollrath, D., and Davis, R.W. (1986) *Science* **234**, 1582–85.
- Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. (USA)* **81**, 1991–95.
- Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
- Van Severter, G.A., Spits, H., Yssel, H., Melief, C.J.M., and Ivanyi, P. (1988) *J. of Immunol.* **141**, 417–22.
- Smith, C.L., Klee, S.R., and Cantor, C.R. (1988) In Davies, K. (ed.), *Genome Analysis*. IRL Press, Oxford, pp. 41–72.
- Smith, H.O. and Birnstiel, M.L. (1976) *Nucleic Acids Res.* **3**, 2387–98.
- Koller, B.H., Geraghty, D.E., Shimizu, Y., DeMars, R., and Off, H.T. (1988) *Journal of Immunology*, **141**, 897–904.
- Srivastava, R., Chorney, M.J., Lawrance, S.K., Pan, J., Smith, Z., Smith, C.L., Weissman, S.M. (1987) *Proc. Natl. Acad. Sci. (USA)* **84**, 4224–28.
- Shukla, H., Gillespie, G.A., Srivastava, R., Collins, F., and Chorney, M.J. (Communicated, Genomics.)

30. Koller, B.H., Sidwell, B., DeMars, R., and Orr, H.T. (1984) *Proc. Natl. Acad. Sci. (USA)* 81, 5175–78.
31. Nedwin, G.E., Naylor, S.L., Sakaguchi, A.Y., Smith, D., Nedwin, J.J., Pennica, D., Goeddel, D.V., and Gray, P.W. (1985) *Nucleic Acids Res.* 13, 6361–6373.
32. Lawrance, S.K., Das, H.K., Pan, J., and Weissman, S.M. (1985) *Nucleic Acids Res.* 13, 7515.
33. Das, H.K., Lawrance, S.K., and Weissman, S.M. (1983) *Proc. Natl. Acad. Sci. (USA)* 80, 3543–47.
34. Swaroop, A. and Weissman, S.M. (1988) *Nucleic Acids Res.* 16, 8739.
35. Oda, K., and Marmur, J. (1966) *Biochemistry* 5, 761–773.
36. Ikawa, S., Shibata, T., Ando, T., and Saito, H. (1980) *Mol. Gen. Genet.* 177, 359–368.
37. Collins, F.S. and Weissman, S.M. (1984) *Proc. Natl. Acad. Sci. (USA)* 81, 6812–6816.
38. Kandpal, R.P., Ward, D.C. and Weissman, S.M. (1990) *Nucleic Acids Res.* 18, 1789–1795.
39. Burke, D.T., Carle, G.F., Olson, M.V. (1987) *Science* 235, 806–812.
40. Abe, K., Wei, J-F., Wei, F-S., Hsu, Y-C., Uehara, H., Artzt, K. and Bennett, D. (1988) *EMBO J.* 7, 3441–3449.