Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 51–55, January 1989
Biochemistry

Enzymatic cleavage of a bacterial genome at a 10-base-pair recognition site

(DNA methyltransferase/*Dpn* I/pulsed-field electrophoresis)

MICHAEL D. WEIL*† AND MICHAEL MCCLELLAND*

Departments of *Biochemistry and Molecular Biology and of †Surgery, University of Chicago, 920 East 58th Street, Chicago, IL 60637

Communicated by Hewson Swift, September 22, 1988

ABSTRACT **The circular genome of *Staphylococcus aureus* was cut into two fragments by a simple enzymatic method that cleaves a 10-base-pair site. The recognition sequence, A-T-C-G-mA ↓ T-C-G-mA-T, was created by the combined use of the methylase M·*Cla* I (A-T-C-G-mA-T) and the restriction endonuclease *Dpn* I (G-mA ↓ T-C). This technique is insensitive to CpG methylation and in human DNA is predicted to produce fragments that, on average, are greater than five million base pairs. The ability to create such long pieces of DNA should facilitate mapping of large, complex chromosomes.**

Cutting chromosomes into a few, long fragments would markedly simplify the task of organizing and mapping complex genomic DNA. To reduce significantly the number of pieces that result from the cleavage of large genomes, it will be necessary to cut DNA into fragments that are greater than one million base pairs (1 Mb) and perhaps as large as 10 Mb. Recent work using pulsed-field electrophoresis (PFE) has resulted in the separation of intact, chromosomal DNA as large as 12 Mb (1, 2). However, present methods of cleaving DNA do not take advantage of this separation technology because they cut too frequently. It is generally true that the longer the recognition sequence of a DNA cleavage system, the less frequently it will cut and the larger the resulting fragments. To generate consistently pieces of DNA that are longer than 1 Mb requires a cutting technique with a recognition specificity of 10 base pairs (bp). The longest recognition sequence found in the case of type II restriction endonucleases is 8 bp (3), and on most genomes, the average size of fragments produced by these rare-cutting restriction enzymes is <1 Mb (4). A number of other enzymatic and chemical techniques that cut DNA less frequently than restriction endonucleases are still in the early stages of development (e.g., refs. 5 and 6).

We previously described a method that uses adenine methylases to generate novel 8- and 10-bp recognition sequences for the methyladenine-dependent endonuclease *Dpn* I (G-mA ↓ T-C; mA, N^6-methyladenine) (7). It was possible with this technique to cleave a 10-bp recognition sequence on a specially constructed plasmid by using the adenine methylase M·*Cla* I (A-T-C-G-mA-T) with *Dpn* I. However, the technique had not been successfully applied to genomic DNA. We now demonstrate the improvement of the M·*Cla* I/*Dpn* I system and show that it can produce specific fragments of 1 Mb and greater in size by using the genome of the bacteria *Staphylococcus aureus* as a model substrate.

MATERIALS AND METHODS

DNA Preparation. Genomic DNA was prepared from *S. aureus* ISP8 that was provided by Peter Pattee (Iowa State University, Ames). The cells were grown to stationary phase, pelleted, and washed. Blocks or "inserts" of agarose containing *S. aureus* were made by first mixing SeaPlaque agarose (FMC BioProducts, Rockland, ME), final concentration of 1%, with 10^{10} cells per ml and allowing the mixture to solidify in 100-μl rectangular molds. Genomic DNA was prepared within the agarose inserts by a modification of a previously described method (8, 9). In brief, the 100-μl inserts were digested with lysozyme at 1 mg/ml and lysostaphin at 1 mg/ml (Sigma) in 100 mM EDTA at 37°C for 2 hr. This was followed by treatment with protease at 1 mg/ml (Sigma) in 1% SDS and 250 mM EDTA at 60°C for 12 hr and finally by washes in TE (10 mM Tris chloride/1 mM EDTA) at room temperature.

Bacteriophage λ *dam*$^-$ was purchased from New England Biolabs. Concatemers, formed spontaneously in the tube, were loaded directly into the gels and used as size markers. *Saccharomyces cerevisiae* chromosomes were prepared as previously described and served here as size markers (10). XP12 DNA was provided by Rick Morgan (New England Biolabs).

Enzymes. M·*Cla* I (11, 12) and *Csp* I were provided by Michael Nelson (University of Chicago). M·*Cla* I can be purchased from New England Biolabs. *Dpn* I was purchased from Boehringer Mannheim. *Taq* I and *Cla* I were purchased from New England Biolabs.

Cleavage Reactions. Many different buffers and reaction conditions were used to optimize the 10-bp cleavage system in agarose. Using DNA embedded in agarose inserts, we found that in KGB (100 mM potassium glutamate/25 mM Tris acetate, pH 7.6/10 mM magnesium acetate/bovine serum albumin at 50 μg/ml/1 mM 2-mercaptoethanol) (13) the best results for the M·*Cla* I and *Dpn* I reaction (M·*Cla* I/*Dpn* I) were obtained with slightly lower temperatures, longer reaction times, and higher concentrations of *S*-adenosylmethionine (SAM) than those recommended by the vendor. To produce fragments below 1 Mb, the enzymes were used simultaneously in double digests: 5 units of *Csp* I, 5 units of M·*Cla* I, and 5 units of *Dpn* I were incubated with one-fourth of an agarose insert (25 μl) in 100 μl of KGB containing 160 μM SAM at 30°C for at least 6 hr.

The production of megabase fragments required sequential reactions to minimize nonspecific nuclease activity. DNA-containing inserts were treated with 5 units of M·*Cla* I in 100 μl of modified KGB (KGB with no magnesium) containing 160 μM SAM at 30°C for 6–12 hr, followed by heat treatment at 55°C for 30 min. The methylated DNA was digested for 6 hr with 5 units of *Dpn* I in the presence of 20 mM magnesium acetate.

Methylation and cleavage of XP12 DNA were performed in KGB under conditions that were otherwise recommended by the manufacturer.

Electrophoresis Apparatus. PFE was performed on a Gene-Line transverse alternating field electrophoresis system (14)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PFE, pulsed-field electrophoresis; SAM, *S*-adenosylmethionine; mA, N^6-methyladenine; Mb, million base pairs.

Table 1. Candidate methylases for the methylase/*Dpn* I strategy

Methylase	Recognition specificity*	Methylase/*Dpn* I site	
		Size, bp	Sequence
M·*Cvi*BIII†	TCG*A*	8	TCGATCGA
M·*Taq* I†	TCG*A*		TCGATCGA
M·*Tth* I†	TCG*A*		TCGATCGA
M·*Ban* III	ATCG*A*T	10	ATCGATCGAT
M·*Bsp* XI	ATCG*A*T		ATCGATCGAT
M·*Bsc* I	ATCG*A*T		ATCGATCGAT
M·*Cla* I†	ATCG*A*T		ATCGATCGAT
M·*Mbo* II†	GAAG*A*		GAAGATCTTC
M·*Ncu I*	GAAG*A*		GAAGATCTTC
M·*Mme* I	GTYGG*A*	11	GTYGGATCCRAC
M·*Acc* III	TCCGG*A*	12	TCCGGATCCGGA
M·*Ama* I	TCGCG*A*		TCGCGATCGCGA
M·*Bsp*HI	TCATG*A*		TCATGATCATGA
M·*Kpn*2I	TCCGG*A*		TCCGGATCCGGA
M·*Mro* I	TCCGG*A*		TCCGGATCCGGA
M·*Nru* I	TCGCG*A*		TCGCGATCGCGA
M·*Rsp*XI	TCATG*A*		TCATGATCATGA
M·*Sal*DI	TCGCG*A*		TCGCGATCGCGA
M·*Taq* II	GACCG*A*		GACCGATCGGTC
M·*Xba* I†	TCTAG*A*		TCTAGATCTAGA
M·*Eco*R124‡	GA*A*N$_6$RTCG	13	GAAN$_4$CGATCGN$_4$TTC
M·*Eco*R124/3‡	GA*A*N$_7$RTCG		GAAN$_5$CGATCGN$_5$TTC
M·*Eco*B‡	TG*A*N$_8$TGCT	14	AGCAN$_5$TGATCAN$_5$TGCT

*The adenine residues that are methylated are italicized.
†Confirmed methylase/*Dpn* I cleavages. The others are theoretical, based on known restriction recognition sequences. M·*Hph* I (T-mC-A-C-C), M·*Bsp*MII (T-C-C-G-G-A), M·*Sbo* 13 (T-C-G-C-G-A), and M·*Eco*DXXI (T-C-A-N$_7$-mA-T-T-C) are excluded because they cannot methylate at G-mA-T-C.
‡M·*Eco*R124 (G-A-A-N$_6$-R-T-C-G, where R = G or A), M·*Eco*R124/3 (G-A-A-N$_7$-R-T-C-G), and M·*Eco*B (T-G-mA-N$_8$-T-G-C-T) are type I methylases and might be impractical *in vitro*.

purchased from Beckman Instruments. A PC 750 pulse controller, purchased from Hoefer Scientific Instruments, was used to steadily increase (ramp) pulse times when necessary.

Electrophoresis was performed by using 1.2% agarose gels in 0.25× TAE buffer (10 mM Tris acetate/0.25 mM EDTA). The gels were cooled to 10°C during electrophoresis by circulating 0.25× TAE buffer with an RM-20 Lauda refrigerating circulator. PFE of DNA fragments produced by the *Csp* I and double digests was conducted at 10 V/cm with a pulse time of 1 sec with a 2-sec ramp for 20–24 hr. DNA fragments produced by M·*Cla* I/*Dpn* I were separated by PFE by using 5 V/cm for a total of 48 hr with pulse times of 300 or 500 sec.

Standard 6% polyacrylamide gel electrophoresis was performed in TBE buffer (89 mM Tris borate/2 mM EDTA) to separate XP12 DNA.

RESULTS

Creating a 10-bp Recognition Sequence for *Dpn* I. The 8-bp recognition sites of *Not* I and *Sfi* I are the longest of any known restriction endonuclease. Yet, by employing adenine methylases that recognize a portion of the *Dpn* I recognition sequence, *Dpn* I sites with an apparent specificity of 8 bp or more can be created (Table 1). One such *Dpn* I site results from methylation by M·*Cla* I at a 10-bp sequence that consists of two overlapping M·*Cla* I sites (Fig. 1) (7). Calculation of the expected frequency of this recognition site, 5′ A-T-C-G-mA ↓ T-C-G-mA-T 3′, on DNA of random composition indicated that fragments of 1 Mb could be expected (4, 9).

Variables Encountered when Cutting Genomic DNA into Large Pieces. Since very large pieces of DNA shear in liquid, it is necessary to prepare them embedded in agarose. However, agarose notably inhibits the activity of most restriction enzymes. To achieve satisfactory cutting in agarose, the performance of some enzymes can be optimized by altering recommended reaction buffers and conditions. We have found that a potassium glutamate buffer (13) and slightly lower temperatures than those recommended by the manufacturers yield the best conditions for complete methylation or cleavage of DNA in agarose (see *Materials and Methods*).

Commercial preparations of restriction endonucleases and methylases, including *Dpn* I and M·*Cla* I, often contain contaminating, copurified endonucleases. Though this level of nuclease contamination has little effect on the smaller fragments ordinarily produced by enzymatic digestion, pieces of DNA in the megabase range are degraded. We initially worked with pieces of 50–1000 kilobase pairs (kb) to bypass this problem. The task was accomplished by cutting a bacterial genome, *S. aureus*, with a rare-cutting enzyme, *Csp* I [C-G-G-(A/T)-C-C-G]. *S. aureus* cells were embedded in agarose, and DNA was prepared as described in *Materials and Methods*. When the DNA in the agarose inserts was cut by *Csp* I and separated by PFE, we observed 10 bands (Fig. 2), where the largest is 750 kb and the smallest is 50 kb. By comparing these pieces to the λ phage ladder (48.5 kb per concatemerized genome), run simultaneously as a size marker, we estimate the *S. aureus* ISP8 genome to be at least 2.86 Mb.

The Cleavage of Genomic DNA by *Dpn* I After Methylation with M·*Cla* I (M·*Cla* I/*Dpn* I). The small number of fragments produced by digesting the genome of *S. aureus* ISP8 with *Csp* I and the absence of a G-mA-T-C methylation system (15) permitted the detection of cutting by M·*Cla* I/*Dpn* I on genomic DNA in agarose. Fig. 3 *A* and *B*, respectively, show the simultaneous and sequential digestion of the *S. aureus* genome with *Csp* I/M·*Cla* I/*Dpn* I. Two M·*Cla* I/*Dpn* I sites are detected. First, the largest *Csp* I fragment of 750 kb disappears in the double digest, and a new band appears at 450 kb. If the 450-kb fragment is from the band at 750 kb, then one would expect to find another fragment at 300 kb. In fact, the band at 300 kb appears to be more intense than the one that results from *Csp* I digestion alone. Second, a site is located in the *Csp* I band at 275 kb. The *Csp* I 275-kb fragment is present to only a small degree after M·*Cla* I/*Dpn* I digestion. In the double digest, a new band appears between 70 and 80 kb that can be seen on the gel by applying a faster pulse time (Fig. 3*C*). If the 75-kb fragment is from the band

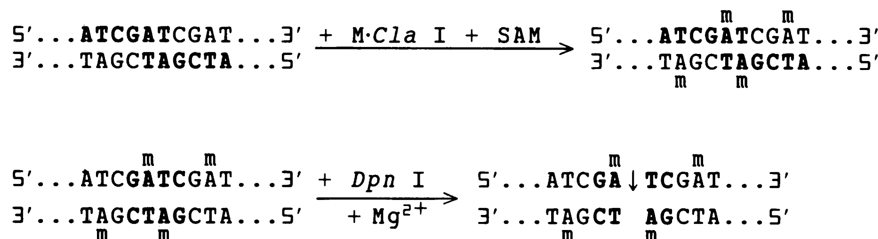

FIG. 1. Cleavage at a 10-bp sequence with M·*Cla* I and *Dpn* I. Double-strand methylation by M·*Cla* I at two overlapping recognition sequences (5′ A-T-C-G-mA-T 3′) generates a *Dpn* I cleavage site (5′ G-mA ↓ T-C 3′) at a novel 10-bp sequence, 5′ A-T-C-G-mA ↓ T-C-G-mA-T 3′.

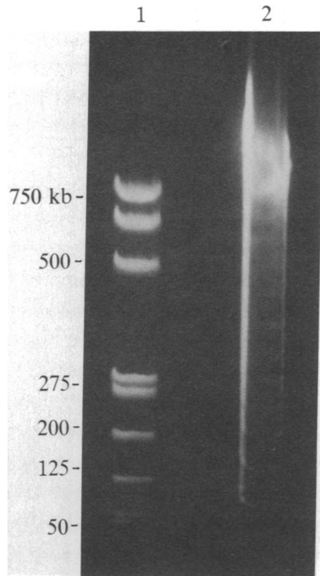

FIG. 2. PFE of *S. aureus* DNA fragments after cleavage by *Csp* I. Digested *S. aureus* DNA, in agarose inserts, was loaded into a 1.2% agarose gel. To show separation up to 800 kb, the samples underwent electrophoresis at 10 V/cm with a pulse time of 1 sec and a 2-sec ramp for 21 hr. Cutting with *Csp* I (lane 1) results in 10 bands whose sizes are approximately 750, 600, 500, 300, 275, 200, 125, 80, 70, and 50 kb. Concatemers of bacteriophage λ DNA (lane 2) were run as size markers. Each band of λ DNA represents an increment of 48.5 kb.

at 275 kb, one would expect to find another fragment at 200 kb. In fact, the band at 200 kb appears more intense than the one that results from *Csp* I digestion alone. This indicates that the *Csp* I band at 275 kb is cut to yield pieces of 75 kb and 200 kb (visualized as a doublet).

The Cleavage of the *S. aureus* Genome into Two Pieces. Two very large fragments result after *S. aureus* DNA is treated with M·*Cla* I followed by *Dpn* I (Fig. 4). This is the first demonstration of a 10-bp cleavage system producing fragments in the million base-pair range and represents the largest such fragments to be separated by PFE. The smaller fragment, as determined from the yeast chromosomal marker, is approximately 1 Mb in size. The DNA fragment represented by the larger band is estimated to be 1.86 Mb or greater. This value was derived as the difference between the minimum genome size of 2.86 Mb (determined by *Csp* I digests) and the other fragment of 1 Mb. However, the size of the larger *S. cerevisiae* chromosomal bands, present in the marker lane, has not been published.

The greater susceptibility of large fragments of DNA to the low levels of nuclease contaminating restriction enzyme preparations is seen with the 10-bp cleavage system. Digestion with M·*Cla* I/*Dpn* I alone produces more detectable degradation, seen as greater background, compared to cleavage by *Csp* I/M·*Cla* I/*Dpn* I. Most of this degradation takes place during *Dpn* I digestion in the presence of magnesium ions, although M·*Cla* I in the presence of magnesium also causes damage to large DNA fragments (data not shown). The sequential reaction, with the methylation done in the absence of magnesium, limits the effects of contaminants that are copurified with the M·*Cla* I.

Insensitivity of M·*Cla* I to CpG Methylation. Almost all restriction systems with CpG in their recognition sequence are inhibited by CpG methylation (16), a modification that is found extensively in human DNA. Although *Dpn* I is insen-

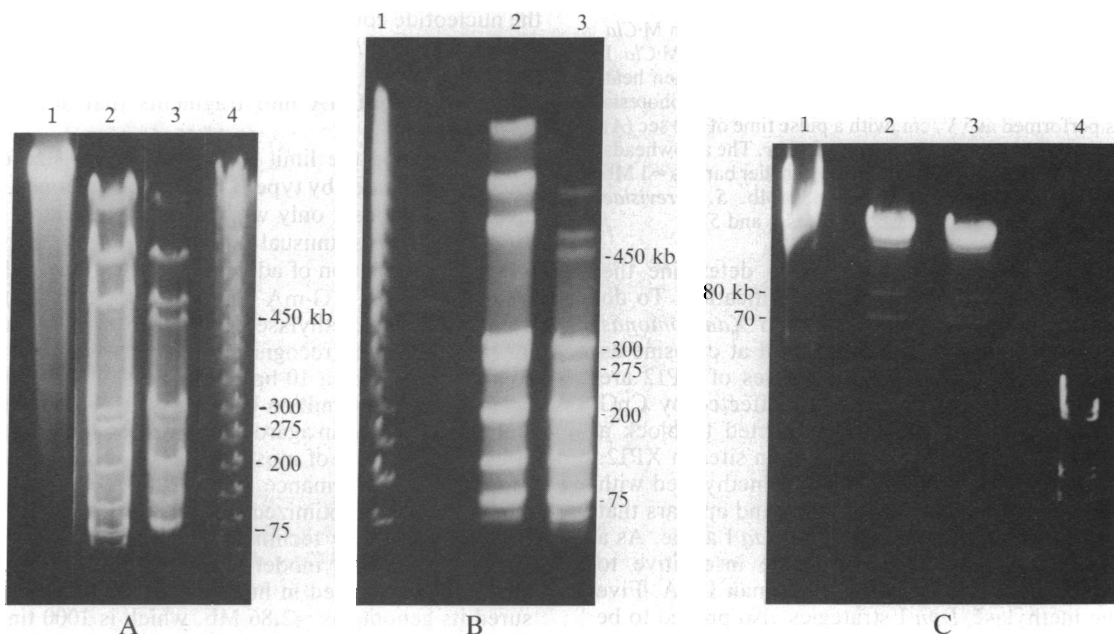

FIG. 3. PFE of double digest products of *S. aureus* DNA by *Csp* I/M·*Cla* I/*Dpn* I. (*A*) Simultaneous technique. *S. aureus* DNA, in agarose inserts, was digested with *Csp* I alone (lane 2) or *simultaneously* with *Csp* I and M·*Cla* I/*Dpn* I (lane 3). The samples were then loaded into 1.2% agarose gels. To show separation up to 800 kb, the samples underwent electrophoresis at 10 V/cm with a pulse time of 1 sec and a 2-sec ramp for 21 hr. Two M·*Cla* I/*Dpn* I (lane 3) sites are detected. The largest *Csp* I (lane 2) fragment is cut to give a new band at 450 kb and a doublet at 300 kb. The second site results in cleavage of the *Csp* I band at 275 kb to produce doublets at 200 kb and 75 kb. Uncut genomic *S. aureus* DNA (lane 1) and bacteriophage λ DNA (lane 4) were used as standards. The photograph was assembled from nonadjacent lanes of the same gel. (*B*) Sequential technique. *S. aureus* DNA was first methylated with M·*Cla* I, in the absence of Mg^{2+} and then heat treated. *Csp* I alone (lane 2) or *Csp* I and *Dpn* I (lane 3) were then added with Mg^{2+}. Electrophoresis of the DNA was carried out as described in *A*. As a size marker, λ phage DNA was used (lane 1). (*C*) To increase separation of bands below 100 kb, electrophoresis was done at a shorter pulse time and ramp. The gels were pulsed at 10 V/cm for 1 sec with a 0.2-sec ramp for 18 hr. The bottom four bands of the *Csp* I digest (lane 2) are seen. A new band appears between 70 and 80 kb on the *Csp* I/M·*Cla* I/*Dpn* I double digest (lane 3). Concatemerized λ phage DNA ladders (lane 1) and *Hin*dIII digest of λ DNA (lane 4) were used as size markers.

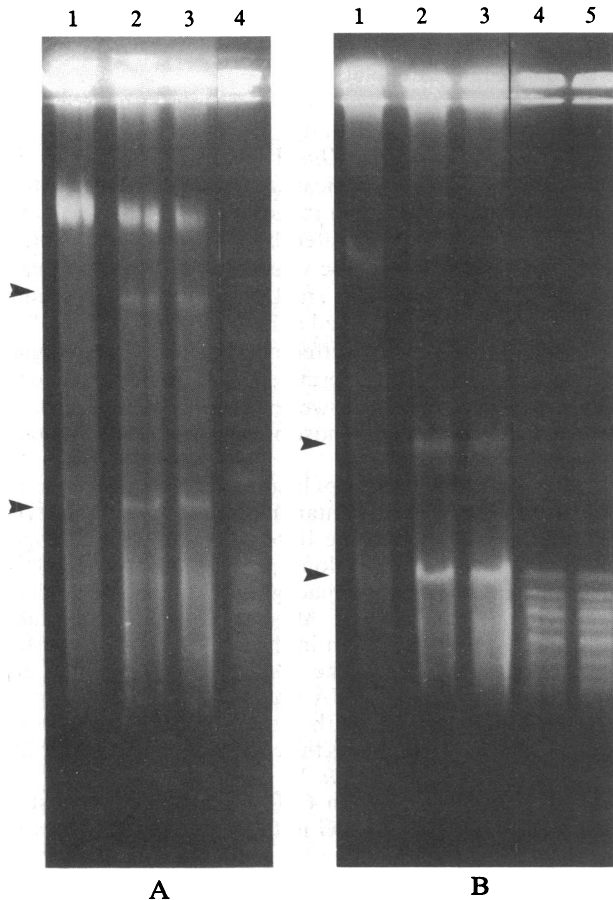

FIG. 4. PFE of sequential digest of *S. aureus* DNA with M·*Cla* I/*Dpn* I. *S. aureus* DNA inserts were methylated with M·*Cla* I in magnesium-free buffer (lanes 1–3). The samples were then heat treated and digested with *Dpn* I only (lanes 2 and 3). Electrophoresis of the gels was performed at 5 V/cm, with a pulse time of 300 sec (*A*) or with a pulse time of 500 sec (*B*) for a total of 48 hr. The arrowheads point to two bands in the megabase range. The smaller band is ≈1 Mb and the larger band is estimated to be 1.86 Mb. *S. cerevisiae* chromosomes were used as size markers (lanes 4 and 5).

sitive to CpG methylation, we wanted to determine the sensitivity of M·*Cla* I to this common modification. To do this, we used XP12, a bacteriophage from *Xanthomonas oryzae*, the DNA of which is fully modified at cytosine as 5-methylcytosine (17). All CpG dinucleotides of XP12 are therefore methylated. If M·*Cla* I were unaffected by CpG methylation, then M·*Cla* I would be expected to block a subset of *Taq* I (5′ T ↓ C-G-A 3′) recognition sites in XP12. Fig. 5 shows that when XP12 DNA is first methylated with M·*Cla* I and is then cut with *Taq* I a new band appears that is not seen when the DNA is treated with *Taq* I alone. As a result, M·*Cla* I/*Dpn* I is expected to be insensitive to methylation of the dinucleotide CpG in human DNA. Five other adenine methylase/*Dpn* I strategies also proved to be insensitive to CpG methylation (data not shown and ref. 16).

DISCUSSION

Physical mapping of large genomes would be facilitated by cleavage strategies that result in fragments averaging over 1 Mb. DNA fragments of more than 1 Mb are not ordinarily produced by type II restriction endonucleases. The frequency of occurrence of a restriction enzyme's recognition site determines the average size of the fragments it produces. A recognition sequence's frequency is partly a function of its length. It is also dependent on the relationship between the

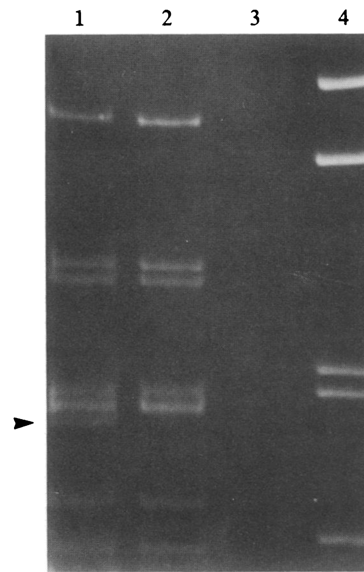

FIG. 5. Cross-protection of XP12 DNA by M·*Cla* I. Two micrograms of XP12 DNA was treated with 5 units of M·*Cla* I at 37°C for 1 hr prior to digestion with 15 units of *Taq* I at 65°C for 4 hr (lane 1) or was simply digested with *Taq* I alone (lane 2). The resulting fragments were separated by standard electrophoresis in a 6% polyacrylamide gel. The arrowhead points to a band that results from the protection by M·*Cla* I of a *Taq* I site (lane 1). As controls, λ dam$^-$ DNA was methylated with M·*Cla* I (lane 3) or was untreated (lane 4) prior to digestion with *Cla* I. The two extra *Taq* I fragments in lane 2 and the uncut λ genomic band in lane 3 were clearly visible but are not seen in this close-up photograph of a portion of the gel.

sequence's G+C content plus rare di- or trinucleotides and the nucleotide content of the DNA to be cut (for reviews, see refs. 4 and 18). *Not* I (G-C ↓ G-G-C-C-G-C), which recognizes the rarest known restriction site for human DNA, cleaves human DNA into fragments that are <1 Mb, on average.

To overcome the limit to the maximum size of a DNA fragment produced by type II restriction enzymes, a strategy that had previously only worked on a plasmid in liquid was used (7). *Dpn* I is unusual in that in order to cleave DNA it requires methylation of adenine in both strands of its recognition sequence, 5′ G-mA ↓ T-C 3′ (Fig. 1) (19, 20). We have used an adenine methylase, M·*Cla* I, to create a cleavage site for *Dpn* I at a 10-bp recognition sequence. In DNA of random composition, such a 10-bp sequence would be predicted to occur once every million base pairs. However, the need to embed large DNA in agarose (to prevent shearing) adversely affects the activity of most restriction enzymes. To achieve satisfactory performance of M·*Cla* I/*Dpn* I, the reaction conditions were optimized (e.g., ref. 13).

Due to its many technical advantages, we have used *S. aureus* DNA as a model system for developing cleavage strategies to be used in human genome mapping. We measured its genome as ≈2.86 Mb, which is 1000 times smaller than the human genome. As a result, rare-cutting restriction enzymes will produce few and discrete fragments. These fragments can be evaluated by ethidium bromide staining instead of by hybridization with radioactive probes, a procedure that would be necessary for more complex genomes. Furthermore, the G+C content of *S. aureus* DNA (≈34%) and its di- and trinucleotide frequencies are similar to those of human DNA (4, 9). Thus, restriction endonuclease recognition sequences that are rare in the human genome are also rare in the *S. aureus* genome. Calculation of the expected frequency of the M·*Cla* I/*Dpn* I recognition site, 5′ A-T-C-G-mA ↓ T-C-G-mA-T 3′, in *S. aureus* indicated that two

or three fragments of 1 Mb could be expected. The *S. aureus* ISP8 genome cut with M·*Cla* I/*Dpn* I results in only two fragments when the digest products are run by PFE. This represents the longest recognition sequence for efficient cleavage of DNA yet produced on a large genome.

In mammalian DNA, the dinucleotide sequence CpG is not only rare, but it is also distributed in a nonrandom fashion (21, 22) and is extensively methylated at cytosine (m^5CpG) in a tissue-specific and developmentally regulated manner (23). Restriction endonucleases that cut infrequently in the human genome, such as *Not* I (G-C ↓ G-G-C-C-G-C), do so primarily because their recognition sequences contain CpG (4). Unfortunately, these enzymes are blocked by m^5CpG. As a result, some restriction sites are not cut or are only partially cleaved, and the cleavage pattern varies between tissues and cell lines. Although the unique M·*Cla* I/*Dpn* I recognition sequence contains two CpG dinucleotides, methylation and cleavage are not blocked by m^5CpG (Fig. 5).

M·*Cla* I/*Dpn* I or other adenine methylase/*Dpn* I strategies (Table 1) may be useful in constructing physical maps of complex genomes. M·*Cla* I/*Dpn* I is predicted to produce fragments averaging over 5 Mb in human DNA (4). Recently, a 12.6-Mb chromosome of *Neurospora crassa* has been separated by PFE (2). Thus, the major obstacle to producing and then separating fragments of this size from human DNA is the need for extremely pure adenine methylases and *Dpn* I. The problem will likely be solved by the availability of cloned enzymes. It is to be expected that the recent cloning of *Dpn* I (24) will allow *Dpn* I of high purity to be prepared.

In another strategy, one can take advantage of human DNA markers that are polymorphic when cut with restriction endonucleases (RFLPs) and have been genetically linked to each other and to specific chromosomal sites (25). Currently, the average physical distance between adjacent RFLPs is estimated to be 5–10 centimorgans (5–10 Mb) (26). However, genetic distances only loosely correlate with the physical distance between markers (27). The physical mapping of these genetically linked sites could be accomplished with very large DNA fragments (generated by the methylase/*Dpn* I cutting systems) that are distinguishable by PFE. In human DNA, these big fragments could be linked to form a complete physical/recombination map.

We thank Peter Pattee, Michael Nelson, and Rick Morgan for generously supplying materials. M.D.W. was self-supported during most of this project and now is supported by a National Institutes of Health Surgical Oncologic Fellowship. M.M. is a Lucille P. Markey Biomedical Scholar. This work was funded by The Lucille P. Markey Charitable Trust, The National Science Foundation, and The Department of Energy.

1. Turmel, C. & Lalande, M. (1988) *Nucleic Acids Res.* **16,** 4727.
2. Orbach, M. J., Vollrath, D., Davis, R. W. & Yanofsky, C. (1988) *Mol. Cell. Biol.* **8,** 1469–1473.
3. Roberts, R. J. (1988) *Nucleic Acids Res.* **16,** r271–r313.
4. McClelland, M. & Nelson, M. (1987) *Gene Amplif. Anal.* **5,** 257–282.
5. Moser, H. E. & Dervan, P. B. (1987) *Science* **238,** 645–650.
6. Chen, C. H. & Sigman, D. S. (1987) *Science* **237,** 1197–1201.
7. McClelland, M., Kessler, L. G. & Bittner, M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 983–987.
8. McClelland, M. (1987) *Methods Enzymol.* **155,** 22–32.
9. McClelland, M., Jones, R., Patel, Y. & Nelson, M. (1987) *Nucleic Acids Res.* **15,** 5985–6005.
10. Schwartz, D. C. & Cantor, C. (1984) *Cell* **37,** 67–75.
11. McClelland, M. (1981) *Nucleic Acids Res.* **9,** 6795–6804.
12. Nelson, M. & McClelland, M. (1987) *Methods Enzymol.* **155,** 32–41.
13. McClelland, M., Hanish, J., Nelson, M. & Patel, Y. (1988) *Nucleic Acids Res.* **16,** 364.
14. Gardiner, K., Laas, W. & Patterson, D. (1986) *Somatic Cell Mol. Genet.* **12,** 185–195.
15. Brooks, J. E., Blumenthal, R. M. & Gingeras, T. R. (1983) *Nucleic Acids Res.* **11,** 837–851.
16. McClelland, M. & Nelson, M. (1988) *Gene*, in press.
17. Kuo, T.-T., Huang, T.-C. & Teng, M.-H. (1968) *J. Mol. Biol.* **34,** 373–375.
18. McClelland, M. (1988) *Nucleic Acids Res.* **16,** 2283–2294.
19. Lacks, S. & Greenberg, B. (1977) *J. Mol. Biol.* **114,** 153–168.
20. Lacks, S. & Greenberg, B. (1975) *J. Biol. Chem.* **250,** 4060–4066.
21. McClelland, M. & Ivarie, R. (1982) *Nucleic Acids Res.* **10,** 7865–7877.
22. Bird, A. P. (1986) *Nature (London)* **321,** 209–213.
23. Waalwijk, C. & Flavell, R. A. (1978) *Nucleic Acids Res.* **5,** 4631–4641.
24. de la Campa, A. G., Springhorn, S. S., Purushottam, K. & Lacks, S. A. (1988) *J. Biol. Chem.* **263,** 14696–14702.
25. Botstein, D., White, R., Skolnick, M. & Davis, R. W. (1980) *Am. J. Hum. Genet.* **32,** 314–331.
26. Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T. P., Bowden, D. W., Smith, D. R., Lander, E. S., Botstein, D., Akots, G., Rediker, K. S., Gravius, T., Brown, V. A., Rising, M. B., Parker, C., Powers, J. A., Watt, D. E., Kauffman, E. R., Bicker, A., Phipps, P., Muller-Kahle, H., Fulton, T. R., Ng, S., Schumm, J. W., Braman, J. C., Knowlton, R. G., Barker, D. F., Crooks, S. M., Lincoln, S. E., Daly, M. J. & Abrahamson, J. (1987) *Cell* **51,** 319–337.
27. Hartley, D., Davis, K. E., Drayna, D., White, R. L. & Williamson, R. (1984) *Nucleic Acids Res.* **12,** 5277–5285.