37. K. Mikami, A. Sakamoto, H. Takase, T. Tabata, M. Iwabuchi, *Nucleic Acids Res.* **17**, 9707 (1989).
38. G. An *et al.*, *Plant Cell* **2**, 225 (1990).
39. L. A. Chodish, S. Buratowski, P. A. Sharp, *Mol. Cell. Biol.* **9**, 820 (1989); C. S. Carr and P. A. Sharp, *ibid.* **10**, 4384 (1990).
40. D. Staiger, H. Kaulen, J. Schell, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6930 (1989).
41. C. Ettlinger and L. Lehle, *Nature* **331**, 176 (1988); S. Gallagher, T. W. Short, P. M. Ray, L. H. Pratt, W. R. Briggs, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8003 (1988); M. R. McAinsh, C. Brownlee, A. M. Hetherington, *Nature* **343**, 186 (1990).
42. G. Gonzalez and M. R. Montiminy, *Cell* **59**, 675 (1989).
43. J. Devereux, P. Haeberli, O. Smithies, *Nucleic Acids Res.* **12**, 387 (1984). Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. The Fos sequence is from T. Curran, G. Peters, C. Van Beveren, N. M. Teich, I. M. Verma, *J. Virol.* **44**, 674 (1982).
44. We thank L. Brian, R. Brosh and R. Kempainen for their excellent technical assistance and A. Baldwin and T. Thomas for helpful discussions. Thanks are also extended to P. Bedinger for use of her lab facilities during parts of this work, B. Kay for advice and providing the polymerase chain reaction primers, S. Whitfield for photographic assistance, and N. Welsch for preparation of the manuscript. Supported by the USDA Competitive Grants Program (89-37262-4456) (to R.S.Q.).

17 April 1990; accepted 21 August 1990

Cleaving Yeast and *Escherichia coli* Genomes at a Single Site

MICHAEL KOOB AND WACLAW SZYBALSKI

The 15–megabase pair *Saccharomyces cerevisiae* and the 4.7–megabase pair *Escherichia coli* genomes were completely cleaved at a single predetermined site by means of the Achilles' heel cleavage (AC) procedure. The symmetric *lac* operator ($lacO_s$) was introduced into the circular *Escherichia coli* genome and into one of the 16 yeast chromosomes. Intact chromosomes from the resulting strains were prepared in agarose microbeads and methylated with Hha I (5′-GCGC) methyltransferase (M·Hha I) in the presence of *lac* repressor (LacI). All Hae II sites (5′-$^{A}_{G}$GCGC$^{T}_{C}$) with the exception of the one in $lacO_s$, which was protected by LacI, were modified and thus no longer recognized by Hae II. After inactivation of M·Hha I and LacI, Hae II was used to completely cleave the chromosomes specifically at the inserted $lacO_s$. These experiments demonstrate the feasibility of using the AC approach to efficiently extend the specificity of naturally occurring restriction enzymes and create new tools for the mapping and precise molecular dissection of multimegabase genomes.

INTEREST IN MAPPING, MANIPULATING, and sequencing of large genomes (*1*) has inspired a search for methods of precisely and efficiently cleaving chromosomal DNA into a small number of fragments. Presently available restriction enzymes recognize sites 4 to 8 bp in size and thus cut most genomes into a very large number of fragments. Although two general approaches for generating rare cleavage sites in vitro have been known for several years, they either are limited by the size and diversity of the cleavage sites recognized (*2*) or use cleaving reagents that cut DNA with low efficiency (*3*). We have introduced an alternative approach, which we call Achilles' heel cleavage (AC), that allows both very rare and highly efficient cleavage of DNA at predetermined locations (*4*).

The key to the AC procedure is modification of the DNA substrate so as to "erase" all but a small subset of recognition sites for a restriction enzyme with a methyltransferase (MTase) that recognizes the same DNA sequence. This subset of sites is protected from methylation by a DNA-binding molecule, added just before methylation, that forms sequence-specific complexes capable of excluding the MTase. Thus, cleavable restriction sites remain only at those locations where the recognition sites for the given restriction enzyme and for the DNA-binding molecule overlap.

Previous studies have demonstrated the practicality of using the AC approach to efficiently cleave plasmid DNA when the *lac* repressor (LacI) (*4*), the phage λ repressor (*4*), or a synthetic oligodeoxynucleotide capable of forming a triple-helix structure were used as blocking molecules (*5*). LacI-mediated AC has also been shown to specifically and efficiently cleave a λ genome (*6*) that contained the symmetric *lac* operator ($lacO_s$) (*7*), which is the ideal LacI-binding site and contains the recognition sequence for both Hae II (5′-AGCGCT) and Hha I (5′-GCGC). Conversion of this operator to an AC site creates a restriction recognition site of ~20 bp, which is large enough to be unique even in the human genome.

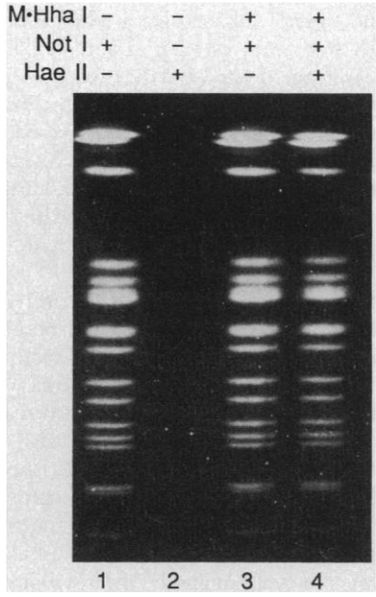

Fig. 1. Complete methylation of the *E. coli* genome. Genomic DNA from *E. coli* strain BNN103 (*9*) was prepared in agarose microbeads (*12*). The microbeads were equilibrated with methylation buffer [50 mM tris-HCl (pH 7.5), 10 mM EDTA, 10 mM dithiothreitol (DTT), 80 μM *S*-adenosyl-methionine, and bovine serum albumin (BSA) (100 μg/ml)], 12.5 U of M·Hha I (New England Biolabs) were added to the reactions indicated, and all samples were incubated for 1 hour at 37°C. The MTase and contaminating nucleases were then inactivated (20 μl of 500 mM EDTA and 1% *N*-lauroylsarcosine, 30 min at 52°C), the microbeads equilibrated with Not I buffer [150 mM NaCl, 10 mM tris-HCl (pH 8.0), 10 mM $MgCl_2$, BSA (100 μg/ml), and 0.01% Triton X-100], and 5 U of the indicated restriction enzyme were added. After incubation at 37°C for 1 hour, the samples were deproteinized as before (15 min at 52°C) and analyzed by PFGE [1% high-strength agarose (Bio-Rad), 150 V, and 25-s switch time for 19 hours at 14°C on a CHEF-DR II system (Bio-Rad)]. Treatments are specified above each lane.

Results with these relatively small DNA molecules, though encouraging, did not guarantee that AC could be successfully applied to the dissection of the large genomes for which it was designed. Direct testing of an appropriately modified AC protocol on whole chromosomes was necessary to determine whether the greater sequence complexity of these extremely large DNA molecules would result in decreased specificity and efficiency of cleavage, as was the case with another rare-cutting method (*8*).

In this report, we show the feasibility of using AC for physically mapping and precisely dissecting chromosomes. Model genomes were generated by introducing $lacO_s$ into the 4.7-Mb circular genome of *Escherichia coli* (*9, 10*) and into one of the 16 chromosomes in the 15-Mb genome of *Saccharomyces cerevisiae* (*11*). Conditions were

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison, WI 53706.

found under which isolated chromosomal DNA from the resulting strains could be efficiently methylated and digested (*12*). The efficiency and specificity of AC at the introduced operator were then determined by methylating the genome with Hha I MTase (M·Hha I) in the presence of LacI, cutting the modified DNA with Hae II, and analyzing the resulting products with pulsed-field gel electrophoresis (PFGE).

Practical implementation of the AC protocol with large DNA molecules requires that both digestion and methylation be readily performed. Complete digestion of intact chromosomal DNA, particularly when embedded in agarose microbeads (*12*), is straightforward and fast (Fig. 1, lanes 1 and 2). Complete methylation of chromosomes, however, is complicated by the presence of nucleases in commercially available MTase preparations and by the instability of the MTase or the methyl donor, or both, during incubation. We have overcome these complications by optimizing methylation conditions (*12*) and by methylating in the absence of Mg^{2+}, which is required for nuclease activity but not for methylation. The MTase and contaminating nucleases are then completely inactivated (see the legend to Fig. 1) before Mg^{2+} is added for the subsequent cleavage reaction. As shown in Fig. 1, we can achieve, within a 1-hour incubation, complete methylation and thus complete protection of chromosomal DNA from a specific endonuclease (compare lanes 2 and 4) without any noticeable degradation by nonspecific nucleases (compare lanes 1 and 3).

To assess the efficiency of the AC process, we methylated chromosomal DNA from the *E. coli* strain with a single $lacO_s$ (*9*) with M·Hha I in the presence of LacI repressor. After a 1-hour incubation, both the MTase and the repressor were inactivated, and the sample was digested with Not I and Hae II. (Not I was used to divide the 4.7-Mb genome into distinct fragments that could be easily analyzed by PFGE.) Hae II completely cleaved the 160-kb $lacO_s$-containing Not I fragment specifically at $lacO_s$ (Fig. 2, compare lane 6, arrowheads, with control lanes 1 and 2).

Next, we assessed the specificity of the AC process by digesting the modified chromosomes with Hae II alone (without Not I digestion) and analyzing the digestion products for the presence of non-$lacO_s$ cleavage sites (Fig. 3). This was done because even though Not I fragments that did not contain $lacO_s$ were not detectably digested by Hae II in the above experiment (Fig. 2, lanes 5 and 6), the presence of any unexpected AC cleavage sites could be more accurately determined in the absence of other digestion

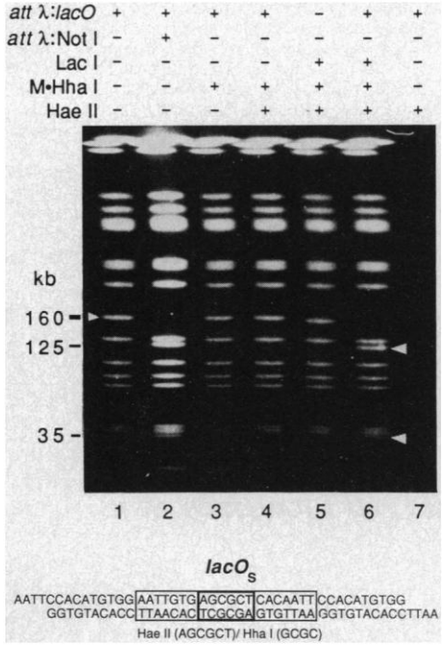

Fig. 2. Achilles' heel cleavage of the *E. coli* genome at a single site as analyzed by a change in the Not I digestion pattern. Genomes from two *E. coli* strains were tested for their susceptibility to *lac* repressor (LacI)–mediated AC: BNN103 (lane 5) (*9*), which does not contain $lacO_s$, and BNN103(*attλ*:*lacO*) (lane 6) (*9*), in which $lacO_s$ was introduced at the phage λ attachment site (*attλ*). A third strain, BNN103(*attλ*:*lacO*, Not I) (lane 2) (*9*), also had an additional Not I site introduced at *attλ*. Not I digestion of this last strain resulted in cleavage of the 160-kb Not I fragment containing *attλ* (lane 1, arrowhead) into 125- and 35-kb fragments (lane 2), mimicking the expected result from AC at *attλ*:$lacO_s$ (lane 6, arrowheads). The sequence of the $lacO_s$ fragment (with Eco RI termini) is shown below (7). Genomic DNA, which was prepared in agarose microbeads (*12*), was either incubated in methylation buffer (controls) or methylated, as indicated. LacI (8×10^{-12} mol calculated for tetramer) was added just before M·Hha I in lanes 5 and 6. The samples were deproteinized and, where indicated, digested with Hae II. All but the complete Hae II digest (lane 7) were also digested with Not I. The microbeads were again deproteinized and analyzed by PFGE (170 V and 15-s switch time for 14 hours). Methylation, deproteinization, digestion, and PFGE were performed as described in Fig. 1 with the minor modifications noted above.

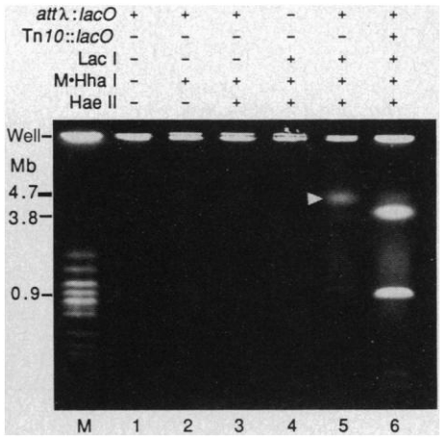

Fig. 3. Cleaving the *E. coli* genome at one or two AC sites. The intact genomes from three *E. coli* strains were tested for their susceptibility to LacI-mediated AC. The intact DNA without a $lacO_s$ (BNN103) (*9*) remained circular and was trapped in the well (lane 4). The genome with a single $lacO_s$ [BNN(*attλ*:*lacO*) (*9*)] was linearized, and thus the 4.7-Mb molecule could enter the gel (lane 5, arrowhead). *Escherichia coli* genome with two symmetric *lac* operators [BNN(*attλ*:*lacO*, Tn*10*::*lacO*) (*10*)] was cleaved into two fragments of 3.8 and 0.9 Mb (lane 6). Digestions were done in a buffer containing 20 mM tris-HCl (pH 8.0), 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM DTT, BSA (100 μg/ml), and 2 U Hae II (Not I was not used in these experiments). All other enzymatic manipulations (indicated above the gel) were performed as described in Figs. 1 and 2. Treated chromosomes were analyzed by PFGE (50 V, 0.6% agarose, and 2-min switch time for 15 hours; 6 min for 8 hours; 20 min for 19 hours; and 40 min for 26 hours). Sizes were estimated from the previously measured size of wild-type *E. coli* K-12 (*18*) and by comparison of the smallest fragment with yeast chromosomes (marker lane, M), Minor bands caused by the DNA-binding protein present as a contaminant in the M·Hha I preparations (*13, 14*) are discussed in the text.

fragments. Achilles' heel cleavage of the strain with a single $lacO_s$ linearizes the circular *E. coli* chromosome and allows the 4.7-Mb DNA molecule to enter the gel (Fig. 3, lane 5, arrowhead), whereas the chromosome without $lacO_s$ remains circular and is trapped in the well (Fig. 3, lane 4). Achilles' heel cleavage of the strain with a second $lacO_s$ introduced at the Tn*10* on the chromosome (*10*) produces two fragments of the expected sizes (Fig. 3, lane 6).

Careful inspection of Fig. 3 (lanes 5 and 6) reveals the presence of a minor band resulting from weak LacI-independent protection of a Hae II site. The degree of this protection was found to vary directly with the amount and inversely with the purity of the M·Hha I preparation used (*13*). Thus, we believe that this unexpected AC activity is due to a contaminating DNA-binding protein in commercial preparations of M·Hha I, which is overproduced from the *hhaIM* gene cloned in *E. coli* (*14*).

The yeast *S. cerevisiae*, with 16 linear chromosomes and a total genome size of 15 Mb, was then used as a model system for assessing the efficiency and specificity of AC of eukaryotic chromosomes. Chromosomes from M10(YIP5*lacO*) (*11*), which has $lacO_s$ inserted into chromosome V (Fig. 4, lane 2, arrowhead), were methylated with M·Hha I in the presence of LacI repressor. Subsequent digestion with Hae II resulted in complete cleavage of chromosome V at a single site (Fig. 4, lane 6, arrowheads). No other Hae II sites were detectably suscepti-

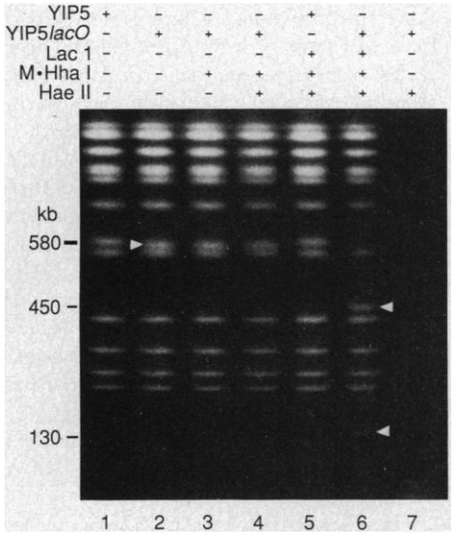

Fig. 4. Achilles' heel cleavage of a yeast chromosome. Integrative vector pYIP5*lacO*, a derivative of pYIP5 (*11*) containing a $lacO_s$ fragment, was homologously recombined with the *URA3* gene on chromosome V of *S. cerevisiae* strain M10 (*11*). pYIP5 was similarly integrated into chromosome V, and both strains were tested for their susceptibility to LacI-mediated AC. Achilles' heel cleavage of M10(YIP5*lacO*) efficiently and specifically cleaved the 580-kb chromosome (lane 2, arrowhead) into 450- and 130-kb fragments (lane 6, arrowheads). No cleavage products were evident when M10(YIP5) (which does not contain $lacO_s$) was similarly treated (lane 5). Yeast chromosomes embedded in agarose microbeads (*12*) were methylated essentially as described in Fig. 1, but with 50 mM potassium glutamate added to the buffer to enhance LacI specificity (*15, 16*). All other enzymatic manipulations indicated above the gel were performed as described in Figs. 1 and 2. Treated chromosomes were analyzed by PFGE (150 V, 1% agarose, 1-min switch time for 23 hours).

ble to digestion (Fig. 4, lanes 5 and 6).

When AC experiments with *S. cerevisiae* were performed with the MTase buffer described in Fig. 1, we observed a weak non-$lacO_s$ Hae II cleavage site, probably at the known yeast pseudooperator (*15*). However, because LacI interaction with pseudooperators is much weaker and more sharply salt-dependent than with $lacO_s$ (*16*), addition of 50 mM potassium glutamate to the methylation buffer eliminated this weak MTase protection site without detectably affecting the protection at $lacO_s$.

As seen in the examples of the yeast and *E. coli* genomes, the AC procedure permits fast, efficient, and specific cleavage of multimegabase genomes at a single predetermined site. The $lacO_s$/LacI AC protocol described here is immediately applicable to the mapping and subsequent isolation of genes in bacterial and fungal chromosomes and in yeast artificial chromosome clones (*17*), all of which can be resolved with current PFGE separate techniques. Isolation of specific regions of larger chromosomes that do not enter the gel in their intact form should also be possible when two AC sites are introduced so as to flank and thus excise the region of interest.

Achilles' heel cleavage is a general, simple, and efficient means of combining the specificities of restriction enzymes with those of other DNA-binding molecules. This approach will eventually be used in conjunction with a wide range of DNA-binding proteins and synthetic DNA-binding agents to create a new class of tools for the physical mapping and precise molecular dissection of multimegabase genomes.

REFERENCES AND NOTES

1. J. D. Watson, *Science* **248**, 44 (1990); C. R. Cantor, *ibid.*, p. 49.
2. M. McClelland, L. G. Kessler, M. Bittner, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 983 (1984).
3. H. E. Moser and P. B. Dervan, *Science* **238**, 645 (1987); J. P. Sluka, S. J. Horvath, M. F. Bruist, M. I. Simon, P. B. Dervan, *ibid.*, p. 1129.
4. M. Koob, E. Grimes, W. Szybalski, *ibid.* **241**, 1084 (1988); *Gene* **74**, 165 (1988).
5. L. J. Maher III, B. Wold, P. B. Dervan, *Science* **245**, 725 (1989).
6. E. Grimes, M. Koob, W. Szybalski, *Gene* **90**, 1 (1990).
7. J. R. Sadler, H. Sasmor, J. L. Betz, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6785 (1983).
8. S. A. Strobel and P. B. Dervan, *Science* **249**, 73 (1990).
9. DNA manipulations involved in plasmid constructions were performed as described [T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982)]. *Escherichia coli* strain BNN103 [R. A. Young and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1194 (1983)], was the parental strain for all *E. coli* strain construction. Operator $lacO_s$ was introduced at the *att*λ site on the chromosome by cloning the Ssp I to Nhe I $lacO_s$ fragment from pOE310 (7) between the unique Ssp I and Xba I sites of plasmid pANT123 (to be described elsewhere), which contains *att*P (27480 to 27873 on λ map) and an origin of replication flanked by Not I sites, to yield pANT*lacO*310. This plasmid was digested with Not I, and the fragment containing $lacO_s$ and *att*P was isolated. This fragment was then either self-ligated [in order to preserve the Not I site for the construction of the positive control strain BNN103(*att*λ::*lacO*, Not I), Fig. 2, lane 2] or treated with the Klenow fragment of DNA polymerase I and all four deoxynucleotides, ligated, and then again digested with Not I [in order to eliminate the Not I site for the construction of strain BNN103(*att*λ::*lacO*); Fig. 2, lane 1]. These ligation mixes were transformed into a derivative of BNN103 in which the λ Int protein was supplied in trans by a plasmid containing the λ *int* gene (27758 to 29019 on λ map) under the control of the *lac* promoter. Colonies resistant to ampicillin were selected and cured of the λ Int–producing plasmid with acridine orange [see J. H. Miller, *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972)].
10. Operator $lacO_s$ was integrated at the Tn*10* (near *hflA*) of BNN103 by homologous recombination. The Xba I to Eco RV *tet* promoter region of a plasmid containing the 3.9-kb Bam HI to Hind III fragment of Tn*10* was replaced with the Ssp I to Nhe I $lacO_s$-containing fragment from pOE310 (7). A kanamycin-resistance gene was then introduced into the resulting plasmid by cloning the Hinc II fragment from pUC4K [J. Vieira and J. Messing, *Gene* **19**, 259 (1982)] into the unique Eco RV site. Gene replacement was then performed [see J. A. K. W. Kiel, J. P. M. J. Vossen, G. Venema, *Mol. Gen. Genet.* **207**, 294 (1987)] to yield BNN103 (Tn*10*::*lacO*). Another $lacO_s$ operator was introduced at the *att*λ site of this strain as described in (*9*).
11. Growth and manipulation of yeast strains were performed as described [R. L. Rodriguez and R. C. Tait, *Recombinant DNA Techniques: An Introduction* (Addison-Wesley, Reading, MA, 1983)]. Plasmid pYIP5*lacO*, a derivative of pYIP5 [K. Struhl, D. T. Stinchcomb, S. Scherer, R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1035 (1979)] with the 44-bp Eco RI $lacO_s$ fragment from pOE310 inserted at the unique Eco RI site, was linearized within the *URA3* gene with Apa I and transformed into lithium acetate–treated *S. cerevisiae* strain M10, an a derivative of strain M11 [M. G. Sandbaken and M. R. Culbertson, *Genetics* **120**, 923 (1988)]. Ura3$^+$ transformants were selected to yield M10(YIP5*lacO*). M10(YIP5), which does not contain the $lacO_s$ operator and thus served as a negative control, was derived in the same manner from pYIP5 and M10. PFGE and Southern (DNA) analysis confirmed the integration of both plasmids in the expected chromosome. Chromosome V of strain M10(YIP5) is slightly larger than the same chromosome of M10(YIP5*lacO*), due to the integration of more than one plasmid at the *URA3* gene. The sequence content of these strains, however, differ in only the presence or absence of the $lacO_s$ insert.
12. We used agarose microbeads, which were originally developed to study the chromatin structure of human cell lines [P. R. Cook, *EMBO J.* **3**, 1837 (1984); D. A. Jackson and P. R. Cook, *ibid.* **4**, 913 (1985); see also adaptations for yeast and *E. coli* chromosomes by G. F. Carle and M. V. Olson [*Methods Enzymol.* **155**, 468 (1987)] and M. McClelland (*ibid.*, p. 22). Critical modifications are as follows. *Escherichia coli* cells were embedded at ~4 μg of DNA per milliliter in agarose microbeads, and yeast cells were embedded at 2.5 times the density of the overnight culture. To methylate the released but embedded DNA, 20 μl of microbeads was washed once with 200 μl of TE buffer [see Maniatis *et al.* in (*9*)] and 0.01% Triton X-100 and twice with 200 μl of methylation buffer (see the legends to Figs. 1 and 4). All buffer above the beads was then removed and the indicated amounts of BSA, M·Hha I, *S*-adenosylmethionine, and LacI (when used) were added. After incubation and inactivation of the MTase (see the legend to Fig. 1), the samples were equilibrated in digestion buffer (in the manner described for the MTase buffer), and BSA and restriction enzyme were added in the amounts indicated in the figure legends. After incubation, the digestions were stopped as before, and ~10 μl of treated beads was loaded and fixed on the comb before the agarose gel was poured for PFGE. A detailed point-by-point description of how we prepare and enzymatically manipulate microbeads will be published elsewhere (M. Koob and W. Szybalski, in preparation).
13. M. Koob, unpublished data.
14. M. Caserta, W. Zacharias, D. Nwankwo, G. G. Wilson, R. D. Wells, *J. Biol. Chem.* **262**, 4770 (1987).
15. A. Simons, D. Tils, B. von Wilcken-Bergmann, B. Müller-Hill, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1624 (1984).
16. M. T. Record, Jr., P. L. deHaseth, T. M. Lohman, *Biochemistry* **16**, 4791 (1977); M. C. Mossing and M. T. Record, Jr., *J. Mol. Biol.* **186**, 295 (1985).
17. D. T. Burke, G. F. Carle, M. V. Olson, *Science* **236**, 806 (1987).
18. C. L. Smith, J. G. Econome, A. Schutt, S. Klco, C. R. Cantor, *ibid.*, p. 1448.

19. We thank D. Comb and I. Schildkraut of New England Biolabs for methyltransferases and restriction enzymes; J. L. Betz, G. P. Roberts, M. R. Culbertson, and N. Hasan for plasmids and strains; and M. T. Record, Jr., for *lac* repressor. We also thank N. Hasan and S. L. Sturley for helpful discussions and guidance regarding strain construction. Supported by NIH grant GM 39715 (to W.S.), National Cancer Institute core grant 5-P30-CA-07175, and by an NSF graduate fellowship (to M.K.).

26 April 1990; accepted 5 September 1990