Site-Specific Cleavage of a Yeast Chromosome by Oligonucleotide-Directed Triple-Helix Formation

SCOTT A. STROBEL AND PETER B. DERVAN*

Oligonucleotides equipped with EDTA·Fe can bind specifically to duplex DNA by triple-helix formation and produce double-strand cleavage at binding sites greater than 12 base pairs in size. To demonstrate that oligonucleotide-directed triple-helix formation is a viable chemical approach for the site-specific cleavage of large genomic DNA, an oligonucleotide with EDTA·Fe at the 5′ and 3′ ends was targeted to a 20–base pair sequence in the 340–kilobase pair chromosome III of *Saccharomyces cerevisiae*. Double-strand cleavage products of the correct size and location were observed, indicating that the oligonucleotide bound and cleaved the target site among almost 14 megabase pairs of DNA. Because oligonucleotide-directed triple-helix formation has the potential to be a general solution for DNA recognition, this result has implications for physical mapping of chromosomes.

TECHNIQUES FOR THE SITE-SPECIFIC cleavage of double-stranded DNA are vital to chromosomal mapping, gene isolation, and DNA sequencing (*1, 2*). Restriction endonucleases with 4– to 6–base pair (bp) binding sites cleave too frequently for many chromosomal DNA manipulations (*3*). Rare-cutting restriction enzymes with 8-bp specificities have found widespread use in genetic mapping; however, these enzymes are few in number, are limited to the recognition of CpG-rich sequences, and cleave at sites that tend to be highly clustered (*4*). Combinations of methylases and restriction enzymes that require methylated sequences can produce cleavage specificities of 8 to 12 bp (*5*). Transient methylase protection can be induced by DNA binding proteins that recognize sequences with overlapping restriction-methylation sites; restriction enzyme digestion then produces specific cleavage at the protein binding site (*6*). Recently, endonucleases encoded by group I introns have been discovered that might have greater than 12 bp specificity (*7*). Unfortunately, none of these strategies can be generalized to recognize and cleave at any of the large number of unique sequences contained in human DNA.

Pyrimidine oligonucleotides bind specifically to purine sequences in duplex DNA to form a local triple-helix structure (*8–12*). The oligonucleotide binds in the major groove parallel to the Watson-Crick purine strand by Hoogsteen hydrogen bonding (*8–12*). Triple-helix specificity is derived from thymine (T) binding to adenine-thymine base pairs (T·AT base triplet) and protonated cytosine (C+) binding to guanine-cytosine base pairs (C + GC base triplet) (*8–15*). Guanine recognition of thymine-adenine base pairs (G·TA base triplet) within the pyrimidine triple-helix motif (*9*) and recognition of (purine)$_n$(pyrimidine)$_m$ type sequences by alternate strand triple-helix formation (*10*) have extended recognition of duplex DNA to a wide class of mixed purine-pyrimidine sequences (*16*). Oligonucleotides 15 to 20 bases in length equipped with an EDTA·Fe moiety produce sequence-specific double-strand breaks with efficiencies ranging from 5 to 25% at their target sites within genomes as large as that of bacteriophage λ (48.5 kbp) (*8–10*). In order to determine if this specificity can be achieved in chromosomal DNA, a triple-helix target site, 5′-$A_2GA_4GA_2GA_3GA_5$-3′, was inserted proximal to the *LEU2* gene on the short arm of the 340-kb chromosome III of *Saccharomyces cerevisiae* (*17–25*) by homologous recombination (Fig. 1). The genetic map location of the *LEU2* locus indicates that double-strand cleavage at the 20-bp target site should produce two fragments, approximately 110 ± 10 and 230 ± 10 kb in size (*26*) (Fig. 2). We report the site-specific cleavage at this genetically engineered sequence on chromosome III by an oligonucleotide-(EDTA·Fe)$_2$.

A 20-base pyrimidine oligonucleotide, 5′-T*TCT$_4$CT$_2$CT$_3$CT$_4$T*-3′, with thymidine EDTA (T*) (*27*) at the 5′ and 3′ termini, was synthesized by automated methods beginning with 5′-*O*-DMT-thymidine-EDTA-triethylester 3′-succinyl control pore glass as the solid support (DMT, 4,4′-dimethoxytrityl). Cleavage reactions were performed on yeast transformants SEY6210 (− target site) and SEY6210B (+ target site) (*28*). Chromosomal DNA embedded in an agarose plug was equilibrated with oligonucleotide-(EDTA·Fe)$_2$ to facilitate diffusion into the agarose and triple-helix formation (pH 7.2, 22°C). The cleavage reaction was initiated by addition of dithiothreitol (DTT). To improve the cleavage yield, a second cleavage cycle was performed by disrupting the triplex at conditions of high pH (8.5), reequilibrating the plug in a triplex-compat-

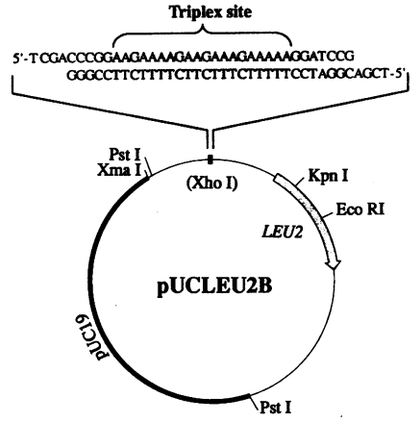

Fig. 1. Schematic diagram of pUCLEU2B constructed by insertion of the Pst I–Xma I 4.0-kb *LEU2* fragment from YEp13 into pUC19 by standard procedures (*19*). Complementary oligonucleotides containing a homopurine sequence were ligated into the unique Xho I site, upstream of the *LEU2* gene.

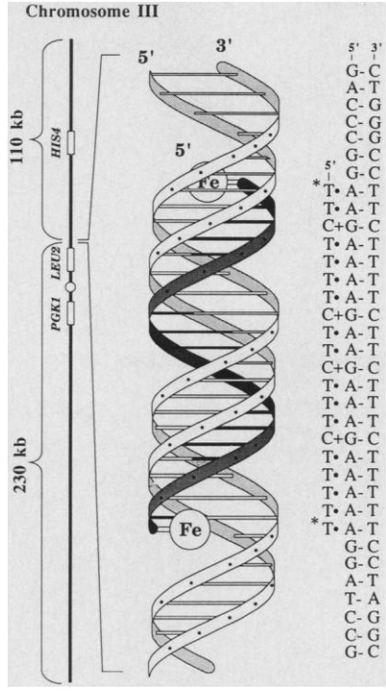

Fig. 2. (**Left**) Genetic map of *S. cerevisiae* strain SEY6210B (+ target site) chromosome III. The locations of *HIS4*, *PGK1*, *LEU2* loci (boxes), the centromere (circle), and the triple-helix target site are indicated. The sizes of the cleavage products (based upon genetic map distances and experimental results) are shown. (**Right**) Schematic diagram of the triple-helix complex. The pyrimidine oligonucleotide with EDTA·Fe at the 5′ and 3′ termini is bound in the major groove parallel to the purine strand.

Arnold and Mabel Beckman Laboratories of Chemical Synthesis, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125.

*To whom correspondence should be addressed.

ible buffer, and repeating the reaction with fresh reagents. The chromosomes were separated by pulsed-field gel electrophoresis and detected by ethidium bromide staining (Fig. 3, A and B). Cleavage products were detected by DNA blotting with chromosome III–specific probes (Fig. 3, C and D).

The *HIS4* (*29*) and *PGK1* (*30*) genes are located on the short and long arms of chromosome III, respectively (Fig. 2). DNA hybridization of the resolved cleavage products (Fig. 3A) with a radiolabeled *HIS4* probe revealed a 110 ± 10 kb fragment present only in the yeast strain containing the engineered target site (SEY6210B) (Fig. 3C, lanes 3 and 4). Hybridization with a radiolabeled *PGK1* probe revealed a second unique fragment 230 ± 10 kb in size (Fig. 3D, lanes 3 and 4). The extent of double-strand cleavage at the target site was estimated at 6% by densitometry. The observed fragment sizes are consistent with those estimated from the genetic map (*26*). Thus, after searching through almost 14 megabase pairs of yeast DNA, the oligonucleotide bound and cleaved specifically at the 20-bp target site while leaving the other chromosomes largely intact (Fig. 3B).

The sequence specificity of pyrimidine oligonucleotides for local triple-helix formation on duplex DNA is dependent upon pH, temperature, and organic cosolvents (*8*). Under conditions of lower pH, lower temperature, or added ethanol, oligonucleotides have been observed to bind to sites that are in significant but not perfect match with the target-site sequence (*8*). Because the complete sequence of the yeast genome is not yet available, the location and number of secondary binding sites on chromosome III could not be predicted a priori. Interestingly, one major (300 ± 10 kb) and three minor (190, 210, and 240 ± 10 kb) secondary cleavage fragments were detected on chromosome III at pH 7.2 (Fig. 3, C and D, lanes 3 and 4) (*31*). The appearance of the three minor fragments (190, 210, and 240 kb) upon hybridization with the flanking markers *HIS4* and *PGK1* indicates that the minor secondary cleavage sites are found on the long arm of chromosome III, distal to the engineered target site. The major secondary cleavage site (300 kb) was not flanked by the markers, but must map to within 40 kb of a chromosome III telomere.

The extent of sequence similarity of the secondary sites to the target site can be estimated by examining the cleavage pattern as a function of increasing pH. The cleavage products were examined over the pH range 7.2 to 7.8 (Fig. 3E). The 190- and 210-kb bands were not observed above pH 7.4 (lanes 7 to 10), whereas raising the pH above 7.6 eliminated the 240-kb fragment (lanes 9 and 10). The 300-kb band and the fragment corresponding to the designed target site were still observed at pH 7.8 (lanes 9 and 10) though at lower cleavage efficiencies. This suggests that the order of sequence similarity of the different sites with the target site are 300 > 240 > 210, 190 kb.

A chemical approach for the site-specific cleavage of intact chromosomes at 12- to 20-bp sequences might assist the large effort being directed toward mapping genomic DNA. For an unambiguous test of site-specific cleavage on chromosomal DNA by oligonucleotide-directed triple-helix formation, a target site of known sequence and approximate physical location was chosen for this experiment. However, the ability of oligonucleotide-directed triple-helix formation to recognize a wide variety of purine and mixed purine-pyrimidine sequences (*16*) could allow the orchestrated cleavage of large genomic DNA at any genetic marker for which some sequence information is known.

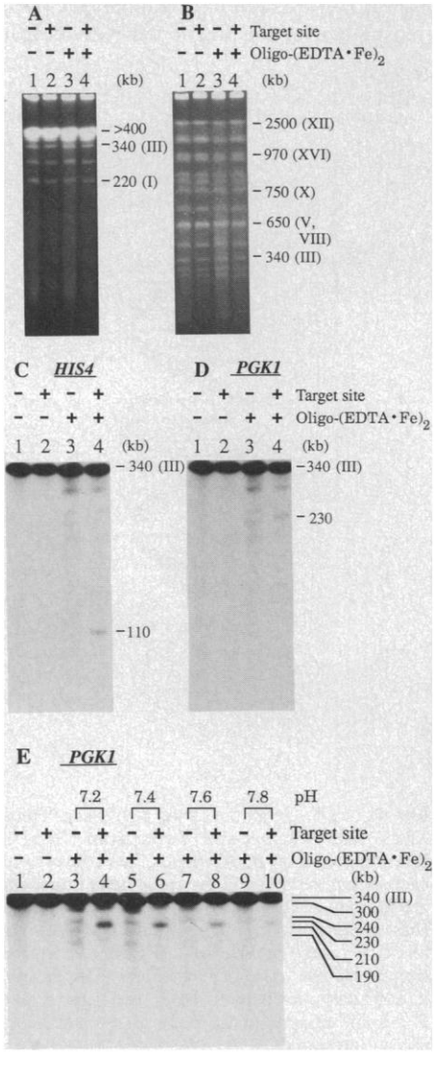

Fig. 3. Site-specific cleavage of yeast chromosomes. Lanes 1 and 2 (all gels): SEY6210 (− target site) and SEY6210B (+ target site) chromosomal DNA unreacted controls, respectively. Lanes 3 and 4 (all gels): SEY6210 and SEY6210B chromosomal DNA, respectively, after reaction with oligonucleotide-(EDTA·Fe)$_2$. (**A**) Separation of yeast chromosomes less than 400 kb in size by pulsed-field gel electrophoresis on a Bio-Rad CHEF system. Pulse times were ramped from 10 to 20 s during a 24-hour period (14°C and 200 V). Chromosomal DNA was detected by ethidium bromide staining. Fragment sizes were estimated by comparison to bacteriophage λ concatemers. (**B**) Separation of all yeast chromosomes with tentative assignments. A 60-s pulse time for 16 hours was followed by 90-s pulses for 8 hours (14°C, 200 V). Sizes were estimated by comparison to YNN295 chromosomal DNA (*24*). (**C**) DNA blot hybridization of reactions shown in (A) with a 250-bp *HIS4* fragment labeled with ^{32}P by random priming (*19*). The DNA blot transfer and hybridization were performed by standard procedures (*19*). The cleavage products were visualized by autoradiography and quantitated by laser densitometry. (**D**) DNA blot hybridization as in (C) except a 1.3-kb marker from the promoter region of *PGK1* was used for hybridization (**E**). pH profile of cleavage products hybridized with PGK1 marker.

REFERENCES AND NOTES

1. H. O. Smith, *Science* **205**, 455 (1979).
2. P. Modrich, *Crit. Rev. Biochem.* **13**, 287 (1982).
3. P. B. Dervan, *Science* **232**, 464 (1986); in *Nucleic Acids and Molecular Biology*, F. Eckstein and D. M. J. Lilley, Eds. (Springer-Verlag, Heidelberg, 1988), vol. 2, pp. 49–64.
4. D. P. Barlow and H. Lehrach, *Trends Genet.* **3**, 167 (1987).
5. M. McClelland, L. G. Kessler, M. Bittner, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 983 (1984); M. McClelland, M. Nelson, C. R. Cantor, *Nucleic Acids Res.* **13**, 7171 (1985); M. D. Weil and M. McClelland, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 51 (1989); Y. Patel, E. Van Cott, G. G. Wilson, M. McClelland, *Nucleic Acids Res.*, in press.
6. M. Koob, E. Grimes, W. Szybalski, *Science* **241**, 1084 (1988).
7. P. S. Perlman and R. A. Butow, *ibid.* **246**, 1106 (1989).
8. H. E. Moser and P. B. Dervan, *ibid.* **238**, 645 (1987); S. A. Strobel, H. E. Moser, P. B. Dervan, *J. Am. Chem. Soc.* **110**, 7927 (1988); T. J. Povsic and P. B. Dervan, *ibid.* **111**, 3059 (1989); L. J. Maher III, B. Wold, P. B. Dervan, *Science* **245**, 725 (1989); S. A. Strobel and P. B. Dervan *J. Am. Chem. Soc.* **111**, 7286 (1989).
9. L. C. Griffin and P. B. Dervan, *Science* **245**, 967 (1989).
10. D. A. Horne and P. B. Dervan, *J. Am. Chem. Soc.* **112**, 2435 (1990).
11. T. Le Doan *et al.*, *Nucleic Acids Res.* **15**, 7749 (1987); V. I. Lyamichev, S. M. Mirkin, M. D. Frank-Kamenetskii, C. R. Cantor, *ibid.* **16**, 2165 (1988); D. Praseuth *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 1349 (1988); J. C. Francois, T. Saison-Behmoaras, M. Chassignol, N. T. Thuong, C. Hélène, *J. Biol. Chem.* **264**, 5891 (1989). J. C. Francois, T. Saison-Behmoaras, N. T. Thuong, C. Hélène, *Biochemistry* **28**, 9617 (1989); J. S. Sun *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9198 (1989).
12. P. Rajagopal and J. Feigon, *Nature* **239**, 637 (1989); C. de los Santos, M. Rosen, D. Patel, *Biochemistry* **28**, 7282 (1989).
13. For triple-stranded polynucleotides, see: G. Felsenfeld, D. R. Davies, A. Rich, *J. Am. Chem. Soc.* **79**,

2023 (1957); K. Hoogsteen, *Acta Crystallogr.* **12**, 822 (1959); M. N. Lipsett, *Biochem. Biophys. Res. Commun.* **11**, 224 (1963); *J. Biol. Chem.* **239**, 1256 (1964); A. M. Michelson, J. Massoulié, W. Guschlbauer, *Prog. Nucleic Acids Res. Mol. Biol.* **6**, 83 (1967); G. Felsenfeld and H. T. Miles, *Annu. Rev. Biochem.* **36**, 407 (1967); F. B. Howard, J. Frazier, M. N. Lipsett, H. T. Miles, *Biochem. Biophys. Res. Commun.* **17**, 93 (1964); J. H. Miller and J. M. Sobell, *Proc. Natl. Acad. Sci. U.S.A.* **55**, 1201 (1966); A. R. Morgan and R. D. Wells, *J. Mol. Biol.* **37**, 63 (1968); J. S. Lee, D. A. Johnson, A. R. Morgan, *Nucleic Acids Res.* **6**, 3073 (1979).

14. For intramolecular triplexes (H-form DNA), see: S. M. Mirkin *et al.*, *Nature* **330**, 495 (1987); O. N. Voloshin, S. M. Mirkin, V. I. Lyamichev, B. P. Belotserkovskii, M. D. Frank-Kamenetskii, *ibid.* **333**, 475 (1988); H. Htun and J. E. Dahlberg, *Science* **241**, 1791 (1988); B. H. Johnston, *ibid.*, p. 1800; Y. Kohwi and T. Kohwi-Shigematsu, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3781 (1988); R. D. Wells, D. A. Collier, J. C. Hanvey, M. Shimizu, F. Wohlrab, *FASEB J.* **2**, 2939 (1988).
15. For oligonucleotide recognition extended to purine motifs (G·GC base triplets), see: M. Cooney, G. Czernuszewicz, E. H. Postel, S. J. Flint, M. E. Hogan, *Science* **241**, 456 (1988).
16. Recognition sites 16 bp in size have 2,147,516,416 unique sequences which is the size range of the human genome (3×10^9 bp) (*3*). In a formal sense, pyrimidine oligonucleotide-directed triple-helix formation utilizing T·AT and C + GC base triplets could recognize 65,536 homopurine sites. The ability to recognize one T·A base pair by use of a single G·TA base triplet in a 16-base sequence (*9*) yields an additional 524,288 binding sites. Alternate strand triple-helix formation (*10*) of the type $(\text{purine})_n$-$\text{NN}(\text{pyrimidine})_m$ ($n = 1$ to 7 and $n + m = 14$) affords 967,044 additional sequences. Thus the total number of 16-bp sites potentially recognized by current pyrimidine triple-helix motifs is 1,556,868. This is approximately one site per 2000 bp in the human genome. Sequence composition effect should be heeded when considering these values.
17. An oligonucleotide duplex containing the 20-bp target site was ligated 650-bp upstream of the *LEU2* gene at the unique Xho I site of the yeast shuttle vector YEp13 (*18*). The orientation of the insert was determined by sequencing from a *LEU2*-specific primer. The 4.0-kb Pst I–Xma I fragment containing *LEU2* and flanking sequences was subcloned (*19*) into pUC19 by a three-piece ligation to generate pUCLEU2B (Fig. 1). Competent haploid *S. cerevisiae* cells (SEY6210 *leu2*) (*20*) were transformed (*21*) with Pst I–linearized pUCLEU2B DNA and recombinants were selected on leucine-deficient minimal media. Chromosomal DNA from recombinant colonies was prepared from a log phase yeast culture by spheroplasting with Zymolyase in 0.9 M sorbitol at 37°C (pH 5.6) followed by sarcosyl and Proteinase K treatment in 0.5% low melting point agarose and 0.5 M EDTA at 50°C (*22*). Insertion of the Pst I *LEU2* fragment into chromosome III was confirmed by pulsed-field gel electrophoretic separation of the yeast chromosomes (*23*) followed by DNA blotting (*19, 24*) with the random primer ^{32}P-labeled Kpn I–EcoR I fragment from the *LEU2* gene (Fig. 1). The presence of the triple-helix target site in the yeast construct was verified by the polymerase chain reaction (*25*) with a *LEU2*-specific oligonucleotide and copies of the inserted oligonucleotides as primers for amplification.
18. J. R. Broach *et al.*, *Gene* **8**, 121 (1979); A. Andreadis, Y. P. Hsu, G. B. Kohlhaw, P. Schimmel, *Cell* **31**, 319 (1982); D. A. Fischhoff, R. H. Waterston, M. V. Olson, *Gene* **27**, 239 (1984).
19. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989).
20. For complete genotype, see J. S. Robinson *et al.*, *Mol. Cell. Biol.* **8**, 4936 (1988).
21. A. Hinnen, J. B. Hicks, G. R. Fink, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 1929 (1978); H. Ito, Y. Fukuda, K. Murata, A. Kimura, *J. Bacteriol.* **153**, 163 (1983).
22. C. L. Smith, S. R. Klco, C. R. Cantor, in *Genome Analysis: A Practical Approach*, K. Davies, Ed. (IRL Press, Oxford, 1988), pp. 41–72.
23. D. C. Schwartz and C. R. Cantor, *Cell* **37**, 67 (1984); G. F. Carle and M. V. Olson, *Nucleic Acids Res.* **12**, 5647 (1984); G. F. Carle, M. Frank, M. V. Olson, *Science* **232**, 65 (1986); S. M. Clark, E. Lai, B. W. Birren, L. Hood, *ibid.* **241**, 1203 (1988); B. W. Birren, E. Lai, S. M. Clark, L. Hood, M. I. Simon, *Nucleic Acids Res.* **16**, 7563 (1988).
24. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
25. R. K. Saiki *et al.*, *Science* **230**, 1350 (1985); R. K. Saiki *et al.*, *ibid.* **239**, 487 (1988).
26. Chromosome III is approximately 140 centimorgans (cM) in length with an average ratio of physical size to genetic map distance of 2.4 kb/cM. The *LEU2* locus is 45 ± 5 and 95 ± 5 cM from the short and long arm telomere, respectively. R. K. Mortimer and D. Schild, *Microbiol. Rev.* **49**, 181 (1985); G. F. Carle and M. V. Olson, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3756 (1985).
27. G. B. Dreyer and P. B. Dervan, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 968 (1985).
28. The cleavage reaction conditions were as follows: yeast chromosomal DNA embedded in 0.5% low melting point agarose (~30-μl blocks) was washed extensively with doubly distilled water to remove exogenous EDTA and equilibrated for 12 hours in 100 mM NaCl, 25 mM tris-acetate, pH 7.2 to 7.8, and 1.0 mM spermine·$(HCl)_4$ (NTS buffer). The agarose blocks were transferred to 2.0-ml flat-bottom eppendorf tubes and overlaid with 120 μl of NTS buffer; 1.5 μl each of 100 μM oligonucleotide-$(EDTA)_2$ and 200 μM $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ (freshly prepared) were preequilibrated for 5 min, added to the solution surrounding the agarose block, and incubated for 3 hours. After diffusion and triple-helix formation, the reaction was initiated by addition of 3 μl of 200 mM DTT and allowed to proceed for 8 hours. The inactivated cleavage reagents were removed by three 30-min washes with 1-ml aliquots of 100 mM tris-HCl (pH 8.5), after which the agarose block was reequilibrated in NTS buffer and a second reaction cycle performed with fresh oligonucleotide-$(EDTA)_2$ and DTT. All manipulations were performed at room temperature (22°C). After two cleavage cycles, the chromosomal blocks were washed in 1 ml of 10 mM tris-HCl (pH 7.5), 1 mM EDTA, and loaded on a 1% agarose (0.5 × TBE) (*19*) gel for characterization of the cleavage products.
29. J. K. Keesey, R. Bigelis, G. R. Fink, *J. Biol. Chem.* **254**, 7427 (1979); R. J. Deshaies and R. Schekman, *J. Cell Biol.* **105**, 633 (1987).
30. M. J. Dobson *et al.*, *Nucleic Acids Res.* **10**, 2625 (1982); M. F. Tuite *et al.*, *EMBO J.* **1**, 603 (1982).
31. Secondary cleavage bands are more intense in the control (− target site) yeast strain because of the higher chromosomal DNA concentration in the sample. Thus band intensities can be compared within, but not between, yeast strains.
32. We thank S. Emr for yeast strain SEY6210, the shuttle vector YEp13, and assistance in the design of the yeast transformation; R. Schekman for the *HIS4* marker in YCp503; J. Campbell for the *PGK1* promoter in pMA91; B. Birren for yeast strain YNN295 and assistance with pulsed-field gel electrophoresis; and the Howard Hughes Institute for a predoctoral fellowship to S.A.S. Supported by NIH grant GM 42966.

24 January 1990; accepted 1 May 1990