MOLECULAR AND CELLULAR BIOLOGY, Dec. 1987, p. 4441–4452 Vol. 7, No. 12
0270-7306/87/124441-12$02.00/0
Copyright © 1987, American Society for Microbiology

Functional Domains of *SIR4*, a Gene Required for Position Effect Regulation in *Saccharomyces cerevisiae*

MARK MARSHALL,† DEBBIE MAHONEY,[1] ALAN ROSE,[1] JAMES B. HICKS,[2] AND JAMES R. BROACH[1]*

Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544,[1] and Scripps Research Clinic and Foundation, La Jolla, California 92037[2]

Received 5 March 1987/Accepted 1 September 1987

The product of the *Saccharomyces cerevisiae SIR4* gene, in conjunction with at least three other gene products, prevents expression of mating-type genes resident at loci at either end of chromosome III, but not of the same genes resident at the *MAT* locus in the middle of the chromosome. To address the mechanism of this novel position effect regulation, we have conducted a structural and genetic analysis of the *SIR4* gene. We have determined the nucleotide sequence of the gene and found that it encodes a lysine-rich, serine-rich protein of 152 kilodaltons. Expression of the carboxy half of the protein complements a chromosomal nonsense mutation of *sir4* but not a complete deletion of the gene. These results suggest that *SIR4* protein activity resides in two portions of the molecule, but that these domains need not be covalently linked to execute their biological function. We also found that high-level expression of the carboxy domain of the protein yields dominant derepression of the silent loci. This anti-Sir activity can be reversed by increased expression of the *SIR3* gene, whose product is normally also required for maintaining repression of the silent loci. These results are consistent with the hypothesis that *SIR3* and *SIR4* proteins physically associate to form a multicomponent complex required for repression of the silent mating-type loci.

Control of mating-type loci in *Saccharomyces cerevisiae* presents a striking example of position effect regulation or altered gene expression as a function of chromosomal position. Genes encoding regulatory proteins that determine the mating type of the cell are present at three different locations on chromosome III (2, 29, 45) (Fig. 1). The set of genes resident at the *MAT* locus are transcribed, and their products establish the cell's mating type (7, 10, 20, 27, 28, 40). The same genes resident at either end of the chromosome, at loci designated *HML* and *HMR*, are not transcribed and do not contribute to the establishment or maintenance of cell type (18, 27, 30). These silent loci serve solely as repositories of mating-type information that normally can be activated only by transposition of a copy of the genes into the *MAT* locus (11, 14, 15, 17, 22, 29, 45).

Differential expression of *MAT* versus *HML* and *HMR* results from repression of expression of the genes residing at the silent loci *HML* and *HMR*. The products of three genes, *SIR2* (or *MAR1*) through *SIR4*, unlinked to each other or to the mating-type loci, are required to maintain the silent loci in a transcriptionally inactive state (8, 16, 31, 32). In the absence of any one of these products, both silent loci are expressed at a level equivalent to that of *MAT*. A fourth gene, *SIR1*, is required for complete repression of the silent loci, but its inactivation yields only partial expression of the otherwise silent genes (13, 31).

Repression of the genes resident at *HML* and at *HMR* also requires the integrity of *cis*-acting sites adjacent to each locus. Deletion of a site designated E, which lies to the left of each locus as they are conventionally represented, yields full expression of the genes in the adjacent locus (1, 5). Both E sites are less than 250 base pairs (bp) in length and can function in *cis* in an orientation-independent manner to repress expression from the promoters of the mating-type genes lying up to several kilobases away (3). In addition, placement of an E site adjacent to a gene other than those normally present at *HML* and *HMR* yields *SIR*-dependent repression of expression from that gene (3). This is true of genes transcribed by polymerase III as well as those transcribed by polymerase II (35). Several studies have implicated a second site at each locus whose integrity is required for complete repression of the contiguous genes (1, 5, 12). This site is designated I and is located on the side opposite E at each locus (1, 5). As a working model, we assume that the products of the *SIR* genes act through the E and I sites to render the DNA in its vicinity refractory to the activity of RNA polymerases.

To address the mechanism by which the *SIR* products might render DNA refractory to transcription, we have undertaken a genetic and molecular genetic analysis of the *SIR4* gene. We present here the complete nucleotide sequence of the *SIR4* gene as well as an analysis of its predicted protein structure. We also provide genetic evidence demonstrating that the protein contains two separate and distinct domains that, although normally linked as a single polypeptide, can act in *trans* to effect repression of the silent loci. We also show that one of these domains apparently encompasses a site for direct interaction with the *SIR3* protein, an interaction essential for establishing and maintaining repression. These properties of *SIR4* are considered in terms of possible mechanisms for *SIR*-mediated repression of the silent loci.

MATERIALS AND METHODS

Strains and media. *Escherichia coli* strains used in this study were C600 (*thr pro leuB hsdR hsdM*), HB101 (*hsdR hsdM recA13 supE44 lacZ4 leuB proA2 thi1*), GM33 (*dam-3*), and JM101 [Δ(*lac-pro*) *thi1 strA endA sbcB15 hsdR4 supE* F′ *traD36 proAB lacI*^{q} *Z*Δ*M15*]. Luria-Bertani medium, with or without 100 μg of ampicillin per ml, and M9 medium with Casamino Acids were prepared as described by Maniatis et al. (24). *S. cerevisiae* strains used in this study are listed with

* Corresponding author.
† Present address: Merck, Sharp, and Dohme Research, West Point, PA 19486.

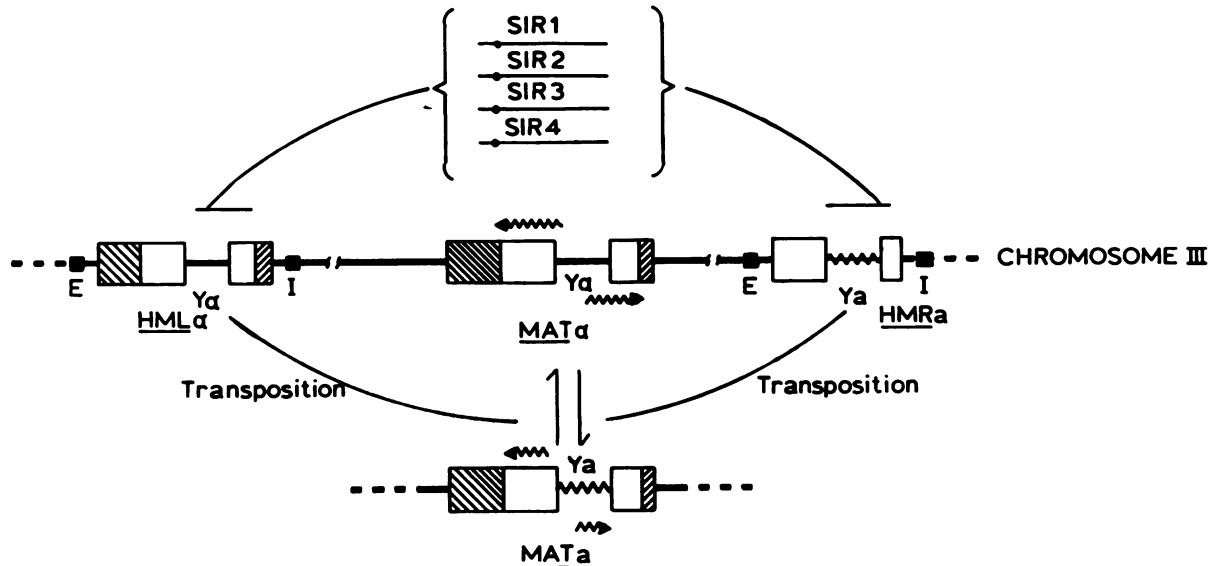

FIG. 1. Mating-type genes of *S. cerevisiae*. Schematic diagram of chromosome III of *S. cerevisiae*, highlighting the organization and structure of the three mating-type loci, *MAT*, *HML*, and *HMR*. Segments common to all three loci (open boxes) flank allele-specific sequences, which consist of either α mating information (solid line) or a mating information (jagged line). *MAT* and *HML* share additional homology (hatched boxes) not present at *HMR*. Flanking the regions of homology at *HML* and *HMR* are sites, E and I, though which the *SIR* gene products act to maintain these loci in a transcriptionally silent state. Under the influence of the *HO* gene product, allele-specific sequences resident at *HML* or *HMR* can be transposed to *MAT*, where they are transcribed to yield regulatory elements that establish the mating type of the cell.

their genotypes and sources in Table 1. Strains Y558 and Y554 were constructed by transforming strain S150-2B to histidine prototrophy with *Sph*I-plus-*Cla*I-digested pMM7.1 or pMM10.7 plasmid DNA, respectively (see below), as described by Rothstein (33). Similarly, strain Y801 was constructed by transforming strain S150-2B to uracil prototrophy with pAR59 plasmid DNA, digested with *Bgl*II plus *Pvu*II. Strain JR281B-7C is a segregant from a cross between strains H027B and S150-2B. Rich medium (YEP) and minimal media (SC), supplemented with either 2% glucose or 2% galactose as carbon source, were prepared as described by Sherman et al. (36).

Plasmids. Genomic *SIR3* and *SIR4* clones, designated pKAN63 and pKAN59, were obtained from K. Nasmyth (37). Their structures are described in Ivy et al. (13). Plasmids pJH3A and pJH5.1A are also described in Ivy et al. (13). Vectors YEp51, YEp52, and YEp13 have been previously described (4).

Random insertions of *Xho*I octanucleotide linkers into plasmid pJH5.1A were generated essentially as described by Shortle and Botstein (38). The approximate position of the linker insertion in individual plasmids was determined by restriction analysis after their recovery by transformation of *E. coli* to ampicillin resistance. The precise positions of a number of linker insertions were determined by sequence analysis as described below.

Plasmids pJHG610 and pJHG745 were constructed by inserting the *Xho*I to *Hin*dIII fragment spanning the carboxy end of *SIR4* from linker insertion derivatives 610 and 745 of plasmid pJH5.1A into *Sal*I-plus-*Hin*dIII-digested YEp51 vector DNA. Similarly, plasmid pSIR4.3 was constructed by inserting the *Sma*I to *Hin*dIII fragment from plasmid pJH3.1A into plasmid YEp51 DNA that had been digested with *Sal*I, treated with DNA polymerase I Klenow fragment plus deoxynucleotide triphosphates, and then digested with *Hin*dIII.

Plasmids containing insertion/deletion mutations of *SIR4* were constructed as follows. Plasmid pKAN59 was digested with *Eco*RI, ligated at low DNA concentration, and transformed into *E. coli* C600. The plasmid recovered by this procedure, pBR-SIR4, consists of plasmid pBR322 containing the genomic fragment spanning *SIR4* formerly present on pKAN59. A *Bam*HI fragment spanning the yeast *HIS3* gene was inserted into *Bam*HI-plus-*Bcl*I-digested pBR-SIR4 DNA that had been propagated in and isolated from *E. coli* GM33 (see Fig. 6). The resulting plasmid was designated pMM7.1, and the allele obtained by its transplacement into yeast cells was designated *sir4Δ1*::*HIS3*. Allele *sir4Δ2*::*HIS3* was obtained by digesting plasmid pMM7.1 with *Sna*BI, which cuts

TABLE 1. Yeast strains

Strain	Genotype	Source
DC14	*MAT*a *his1*	CSHL[a]
DC17	*MAT*α *his1*	CSHL
HO27B	*HML*α *mata1 HMR*α *ade8 his4 gal1 leu2-3,112*	CSHL
IV16-17A	*MAT*α *his4 leu2-3,112 trp1 ura3-52 can1 sir4-351*	CSHL
S150-2B	*MAT*a *his3Δ leu2-3,112 trp1-289 ura3-52 gal2*	CSHL
FVY1-1C	*MAT*α *ade his3 leu2-3,112*	F. Volkert
Y558	*MAT*a *his3Δ leu2-3,112 trp1-289 ura3-52 sir4Δ-1*::*HIS3*	This study
Y554	*MAT*a *his3Δ leu2-3,112 trp1-289 ura3-52 sir4Δ-2*::*HIS3*	This study
Y801	*MAT*a *his3Δ leu2-3,112 trp1-289 ura3-52 sir4Δ-3*::*URA3*	This study
JK-1A	*MAT*a *leu2-3,112 ura3-52 ade2 trp1*	This study
MS-43α	*MAT*α *leu2-3,112 ura3-52 ade2 trp1*	This study
JR281B-7C	*HML*α *mata1 HMR*α *leu2-3,112 ura3-52 his3Δ trp1*	This study

[a] CSHL, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

uniquely within the 5′ portion of the *SIR4* gene, and *Xho*I, which cuts within the *HIS3 Bam*HI fragment 3′ to the *HIS3* gene. The digested DNA was incubated with DNA polymerase I Klenow fragment plus deoxynucleotide triphosphates and then ligated at low DNA concentration. The plasmid recovered by this procedure was designated pMM10.7. Plasmid pAR59 consists of vector pSP64 (Promega Biotech) into which is inserted the *Hin*dIII to *Nde*I fragment spanning the *URA3* gene, bracketed by the 445-bp *Bgl*II to *Hin*dIII fragment from the 5′ side of *SIR4* and the *Nde*I to *Hin*dIII fragment from the 3′ side of *SIR4* (cf. Fig. 2 and 6).

Plasmids pSIR4.6 and pAR42 were designed for galactose-inducible expression and high-copy propagation of the yeast *SIR4* and *SIR3* genes, respectively. Plasmid pSIR4.6 consists of genomic DNA extending from the *Hin*dIII site immediately upstream of the 5′ end of the *SIR4* coding region to the first *Hin*dIII site 3′ of the coding region, inserted immediately downstream of the *GAL10* promoter in the yeast expression vector YEp52 (4). It was constructed by cloning fragments from a partial *Hin*dIII digestion of plasmid pKAN59 into *Hin*dIII-digested vector DNA. To construct plasmid pAR42, a *Sal*I site was introduced immediately 5′ of the *SIR3* coding region by treating an appropriate subclone of pKAN63 with Bal31 and reclosing in the presence of synthetic *Sal*I linkers. The modified *SIR3* gene was then cloned immediately downstream from the *GAL10* promoter in the expression vector YEp53, a derivative of YEp51 in which the yeast *URA3* gene is substituted for *LEU2* as the selectable marker for the plasmid (A. Rose and J. Broach, unpublished results).

Quantitative mating assays. Tester strains (DC14 and DC17) and the strain to be tested were grown to mid-log phase (1×10^7 to 2×10^7 cells per ml) in liquid culture and then briefly sonicated to disperse clumps. Viable cell count was determined by serial dilution and plating on YEPD plates or, for plasmid-bearing strains, on plates selective for the plasmid-borne marker. Equal numbers (10^4 to 10^5) of test cells and tester cells were mixed and filtered onto a sterile 25-mm nitrocellulose disk. The disk was transferred to an appropriate agar plate and incubated at 30°C for 6 h. After incubation, the disk was transferred to 5 ml of sterile water. Cells were dispersed by vortexing and then plated at appropriate dilutions on medium selective for diploids. Mating efficiency was scored as the number of diploids formed per viable test cell (or in the case of plasmid-bearing strains, the number of diploids per plasmid-bearing test cell).

Sequence determination of *SIR4*. *SIR4* sequence was determined by using the chain termination method of Sanger et al. (34). Templates and primers for sequencing were obtained by one of four procedures. The particular procedure used for each segment is indicated in Fig. 2.

RESULTS

DNA sequence of the *SIR4* gene. We have determined the complete nucleotide sequence of the *SIR4* gene present on plasmid pKAN59, using the method of Sanger et al. (34), by the strategy diagrammed in Fig. 2. Nasmyth previously isolated plasmid pKAN59, consisting of a 5.5-kilobase segment of yeast genomic DNA inserted into the yeast vector YEp13, on the basis of its ability to complement a *sir4* strain. Ivy et al. demonstrated that the genomic DNA fragment carried on this plasmid contains the *SIR4* gene by its ability to complement the *sir4-351* ochre mutation and its ability to direct integration into the *SIR4* locus by homologous recombination (13). Within this segment we found only one extended open coding region, the complete nucleotide sequence of which is shown in Fig. 3.

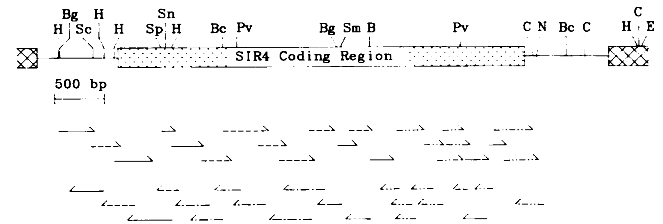

FIG. 2. Strategy for sequencing *SIR4*. The region of plasmid pKAN59 encompassing the yeast genomic fragment spanning *SIR4* is diagrammed, showing the positions of various restriction sites, the location of the extended open coding region (dotted region), and the flanking vector sequences (hatched region). Arrows beneath the diagram represent the extent and orientation of sequence information derived from individual sequencing reactions. The particular protocol used to generate template and primer for each reaction is indicated by the form of the arrow. (Solid arrows) Templates were obtained by subcloning restriction fragments from pKAN59 onto phages M13mp11 and M13mp12. Primer was the M13-specific pentadecamer 5′d(TCCCAGTCACGACGT). (Dashed arrows) Templates were derived from specific linker insertion derivatives of plasmid pJH5.1A, inserted into M13mp12. (Dot-dashed arrows) Templates were obtained as a set of unidirectional deletions essentially as described by Henikoff (9), starting with phage M13mp11, into which had been cloned the *Hin*dIII fragment spanning most of the *SIR4* gene. (Double-dot-dashed arrows) Template was heat-denatured pBR-SIR4 duplex DNA. Specific primers were synthesized as dictated by initial sequence information, using phosphoramidite chemistry with an Applied Biosystems model 380A DNA synthesizer. Restriction site abbreviations: B, *Bam*HI; Bc, *Bcl*I; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; N, *Nde*I; Pv, *Pvu*II; Sc, *Sca*I; Sm, *Sma*I; Sn, *Sna*BI; Sp, *Sph*I.

Several lines of evidence indicate that the extended open coding region corresponds to the *SIR4* gene. First, this region is transcribed. Ivy et al. showed that the fragment completely encompassed a 4.4-kilobase poly(A) transcript found in yeast cells and that the orientation of this transcript is the same as that of the extended open coding region (13). In addition, we have mapped the 5′ end of a poly(A)-containing transcript isolated from yeast cells to a position approximately 48 bp upstream from the first ATG codon of the open reading frame (M. Marshall, Ph.D. thesis, Princeton University, Princeton, N.J., 1985). Second, we constructed a frameshift mutation within this open coding region by filling in the *Bgl*II site in the middle of the gene (see Fig. 5). This mutant plasmid present in a strain containing the *sir4-351* ochre mutation fails to yield a Sir^+ phenotype, even though the parent plasmid yields full complementation. Finally, transplacement into a SIR^+ strain of a number of different constructs carrying interruptions of this coding region yields $sir4^-$ strains (see below).

The *SIR4* DNA sequence shares no obvious homology with either the *SIR2* or the *SIR3* DNA sequences. A perfect TACTAAC sequence, essential for mRNA splicing in yeasts, is present at nucleotide 1484. However, no exact homologies to GTATGT, the 5′ donor site sequence, or any interruptions in the open reading frame requiring splicing are evident within the sequence. In addition, S1 nuclease analysis with a probe overlapping the TACTAAC sequence showed that none of the *SIR4* transcript is spliced at this site (Marshall, Ph.D. thesis).

Analysis of the predicted amino acid sequence of *SIR4*. The open reading frame of *SIR4* codes for a protein of 1,358 amino acids with a predicted molecular weight of 151,902.

```
        10         20         30         40         50         60         70         80         90        100
GCTTTAGTTC AAAGATCTGA TGACAATGCA CACGCCTTGA AGAAGAGATT AGCTGCTTAC CATGCTCAAA CCGAACCAAT TGTTGACTTT TACAAAAAGA

       110        120        130        140        150        160        170        180        190        200
CCGGTATCTG GCCTGGTCTT GATGCTTCCC AACCTCCTCC TACTGTTTGG GCTGACATCT TGAACAAGCT AGGTAAGGAT TAATGACCTT GCGCAATGTC

       210        220        230        240        250        260        270        280        290        300
TTCTAAATAT CTTTTCTTTT TTTTTAAATT AATGAAAATA TAATAAATTT TTGTATGTGA GTACATATAT CCGCAGTATA CTATAAATGA AATAATAACA

       310        320        330        340        350        360        370        380        390        400
TAACTGTTTG TAACTGGCTA TAACGTCCTT AAACATGTGC ACTGCCATTA AGTACTTAGT CAATATTTTT ATCGTTGAGA ACGAACGCAA ATCATATTAC

       410        420        430        440        450        460        470        480        490        500
TAAATAGATC AACAAAAGCG CATCACATAA ACAAAAACGG TATAACAGTA ATAGTGAAGC TTTTTATAAA CCTATAACGA AATAAGCGCT TTTCAGGTTT

       510        520        530        540        550   →    560        570        580        590        600
TATAATTCAG TCAGAGTTCT AACTGGACAT CGTTTTGCAG GGGATAAAAA AAAAAAGGAA GCTTCAACCC ACAATACCAA AAAAGCGAAG AAAACAGCCA

            612             624             636             648             660             672             684
ATG CCA AAT GAC AAT AAG ACA CCC AAT AGG TCC AGC ACT CCC AAG TTT ACT AAA AAA CCA GTA ACC CCG AAT GAT AAA ATT CCT
 M   P   N   D   N   K   T   P   N   R   S   S   T   P   K   F   T   K   K   P   V   T   P   N   D   K   I   P

            696             708             720             732             744             756             768
GAA AGA GAA GAA AAA TCC AAT GAA GTG AAG ACA CCT AAA ATT CCA TTA TTC ACG TTT GCC AAA AGC AAA AAC TAT TCA AGG CCG
 E   R   E   E   K   S   N   E   V   K   T   P   K   I   P   L   F   T   F   A   K   S   K   N   Y   S   R   P

            780             792             804             816             828             840             852
TCA ACC GCA ATT CAT ACC TCA CCT CAT CAA CCA AGT GAT GTA AAA CCG ACT TCC CAT AAA CAG TTG CAA CAG CCA AAA TCC TCA
 S   T   A   I   H   T   S   P   H   Q   P   S   D   V   K   P   T   S   H   K   Q   L   Q   Q   P   K   S   S

            864             876             888             900             912             924             936
CCA CTG AAA AAA AAT AAC TAT AAT TCT TTT CCT CAC TCA AAT CTG GAA AAA ATA AGC AAC AGC AAA CTA CTC TCC CTT CTT CGG
 P   L   K   K   N   N   Y   N   S   F   P   H   S   N   L   E   K   I   S   N   S   K   L   L   S   L   L   R

            948             960             972             984             996            1008            1020
TCC AAA ACG TCA GCA GGA AGA ATT GAA TCA AAC AAT CCT TCA CAT GAT GCT AGC AGG TCT CTA GCA AGT TTT GAA CAA ACA GCA
 S   K   T   S   A   G   R   I   E   S   N   N   P   S   H   D   A   S   R   S   L   A   S   F   E   Q   T   A

           1032            1044            1056            1068            1080            1092            1104
TTT TCC CGG CAT GCG CAG CAA CAA ACT TCT ACC TTC AAT TCA AAG CCT GTA CGT ACC ATA GTA CCG ATA TCA ACA TCA CAA ACC
 F   S   R   H   A   Q   Q   Q   T   S   T   F   N   S   K   P   V   R   T   I   V   P   I   S   T   S   Q   T

           1116            1128            1140            1152            1164            1176            1188
AAT AAC TCA TTT TTA TCA GGA GTA AAA AGC TTA CTA AGT GAA GAG AAA ATA AGG GAT TAC TCT AAA GAG ATT CTA GGC ATA AAC
 N   N   S   F   L   S   G   V   K   S   L   L   S   E   E   K   I   R   D   Y   S   K   E   I   L   G   I   N

           1200            1212            1224            1236            1248            1260            1272
TTG GCA AAC GAA CAG CCT GTT TTA GAG AAG CCA CTT AAA AAA GGA TCA GCA GAC ATC GGC GCC TCT GTG ATT TCT TTG ACC AAG
 L   A   N   E   Q   P   V   L   E   K   P   L   K   K   G   S   A   D   I   G   A   S   V   I   S   L   T   K

           1284            1296            1308            1320            1332            1344            1356
GAC AAA TCT ATC AGG AAA GAT ACC GTA GAA GAA AAA AAA GAA GAA AAG TTG AAT ATA GGT AAA AAC TTT GCC CAT TCT GAT TCA
 D   K   S   I   R   K   D   T   V   E   E   K   K   E   E   K   L   N   I   G   K   N   F   A   H   S   D   S

           1368            1380            1392            1404            1416            1428            1440
CTA TCA GTT CCG AAG GTT AGT GCC GGT GAC AGT GGT ATA TCA CCG GAG GAG AGC AAG GCA AGA AGT CCT GGA ATT GCA AAA CCG
 L   S   V   P   K   V   S   A   G   D   S   G   I   S   P   E   E   S   K   A   R   S   P   G   I   A   K   P

           1452            1464            1476            1488            1500            1512            1524
AAT GCC ATA CAA ACA GAA GTG TAT GGA ATA AAT GAG GAA TCT ACT AAC GAG CGT TTA GAA ATA AAT CAA GAA AAA CCA GTA AAA
 N   A   I   Q   T   E   V   Y   G   I   N   E   E   S   T   N   E   R   L   E   I   N   Q   E   K   P   V   K

           1536            1548            1560            1572            1584            1596            1608
TTA GAT GAG AAT AGT GCA AAT AGT ACG GTC GCA TCG GCC TTA GAT ACC AAT GGG ACG TCA GCC ACG ACA GAA ACT CTA ACA TCA
 L   D   E   N   S   A   N   S   T   V   A   S   A   L   D   T   N   G   T   S   A   T   T   E   T   L   T   S

           1620            1632            1644            1656            1668            1680            1692
AAG AAG ATC GTT CCA TCT CCA AAA AAA GTC GCC ATT GAT CAA GAT AAA ATA ACA CTG CAC GAT GAG AAA ACA CTT GCA CCT TCG
 K   K   I   V   P   S   P   K   K   V   A   I   D   Q   D   K   I   T   L   H   D   E   K   T   L   A   P   S

           1704            1716            1728            1740            1752            1764            1776
AAG CAT CAG CCT ATA ACA TCT GAA CAA AAA ATG AAG GAA GAC GCC GAC CTC AAG AGG ATG GAA ATC TTA AAG TCA CCT CAT TTG
 K   H   Q   P   I   T   S   E   Q   K   M   K   E   D   A   D   L   K   R   M   E   I   L   K   S   P   H   L

           1788            1800            1812            1824            1836            1848            1860
TCG AAA AGT CCA GCT GAC AGA CCT CAA GGG CGT AGA AAC AGC CGA AAT TTT TCC ACT AGA GAT GAA GAA ACT ACA AAA CTT GCT
 S   K   S   P   A   D   R   P   Q   G   R   R   N   S   R   N   F   S   T   R   D   E   E   T   T   K   L   A

           1872            1884            1896            1908            1920            1932            1944
TTT CTT GTT GAA TAT GAA GGC CAA GAA AAC AAC TAT AAC TCT ACT TCT CGA AGC ACA GAA AAG AAA AAT GAT ATG AAC ACT TCT
 F   L   V   E   Y   E   G   Q   E   N   N   Y   N   S   T   S   R   S   T   E   K   K   N   D   M   N   T   S

           1956            1968            1980            1992            2004            2016            2028
GCA AAG AAT AAA AAT GGT GAA AAC AAG AAA ATT GGT AAG AGG CCG CCT GAA ATC ATG AGC ACT GAA GCT CAC GTA AAC AAA GTA
 A   K   N   K   N   G   E   N   K   K   I   G   K   R   P   P   E   I   M   S   T   E   A   H   V   N   K   V

           2040            2052            2064            2076            2088            2100            2112
ACC GAA GAA ACC ACA AAG CAG ATA CAG AGT GTA CGA ATA GAT GGT CGA AAA GTG CTT CAA AAA GTA CAA GGA GAA TCC CAC ATT
 T   E   E   T   T   K   Q   I   Q   S   V   R   I   D   G   R   K   V   L   Q   K   V   Q   G   E   S   H   I

           2124            2136            2148            2160            2172            2184            2196
GAT TCG AGA AAC AAT ACC CTG AAT GTT ACA CCA TCA AAG AGG CCC CAG CTA GGA GAA ATA CCG AAT CCT ATG AAA AAG CAT AAG
 D   S   R   N   N   T   L   N   V   T   P   S   K   R   P   Q   L   G   E   I   P   N   P   M   K   K   H   K

           2208            2220            2232            2244            2256            2268            2280
CCT AAT GAA GGG CGA ACT CCA AAT ATC TCA AAC GGT ACA ATA AAC ATC CAA AAG AAA TTA GAG CCT AAG GAA ATT GTG CGA GAT
 P   N   E   G   R   T   P   N   I   S   N   G   T   I   N   I   Q   K   K   L   E   P   K   E   I   V   R   D

           2292            2304            2316            2328            2340            2352            2364
ATT TTG CAT ACG AAA GAA TCA TCA AAT GAG GCT AAG AAA ACT ATT CAA AAC CCT TTA AAT AAA TCA CAA AAC ACT GCT CTT CCT
 I   L   H   T   K   E   S   S   N   E   A   K   K   T   I   Q   N   P   L   N   K   S   Q   N   T   A   L   P

           2376            2388            2400            2412            2424            2436            2448
TCC ACA CAT AAA GTT ACA CAA AAA AAA GAT ATA AAA ATT GGA ACT AAT GAC CTT TTT CAG GTT GAA TCT GCT CCA AAA ATA TCC
 S   T   H   K   V   T   Q   K   K   D   I   K   I   G   T   N   D   L   F   Q   V   E   S   A   P   K   I   S

           2460            2472            2484            2496            2508            2520            2532
TCA GAG ATT GAC CGG GAG AAT GTT AAA TCA AAG GAT GAA CCG GTT TCC AAG GCT GTT GAA AGC AAA TCT TTA TTA AAT TTG TTT
 S   E   I   D   R   E   N   V   K   S   K   D   E   P   V   S   K   A   V   E   S   K   S   L   L   N   L   F

           2544            2556            2568            2580            2592            2604            2616
TCA AAC GTA CTC AAG GCT CCT TTC ATT AAA AGT GAA AGC AAA CCT TTT TCT AGT GAT GCT CTG TCA AAA GAA AAA GCC AAT TTT
 S   N   V   L   K   A   P   F   I   K   S   E   S   K   P   F   S   S   D   A   L   S   K   E   K   A   N   F

           2628            2640            2652            2664            2676            2688            2700
TTG GAA ACT ATC GCT TCC ACT GAA AAG CCA GAA AAT AAG ACT GAT AAG GTG TCT CTA TCT CAG CCA GTT AGT GCA AGT AAG CAT
 L   E   T   I   A   S   T   E   K   P   E   N   K   T   D   K   V   S   L   S   Q   P   V   S   A   S   K   H
```

```
        2712        2724        2736        2748        2760        2772        2784
GAG TAT AGC GAT AAT TTT CCA GTT TCT CTA TCT CAA CCT TCA AAG AAA TCT TTC GCA AAT CAT ACA GAG GAT GAG CAA ATT GAA
E   Y   S   D   N   F   P   V   S   L   S   Q   P   S   K   K   S   F   A   N   H   T   E   D   E   Q   I   E

        2796        2808        2820        2832        2844        2856        2868
AAA AAG AAG ATC TGC CGT GGG AGA ATG AAT ACG ATA ATA ACT CAC CCG GGA AAA ATG GAG CTG GTA TAC GTG TCC GAC TCA GAC
K   K   K   I   C   R   G   R   M   N   T   I   I   T   H   P   G   K   M   E   L   V   Y   V   S   D   S   D

        2880        2892        2904        2916        2928        2940        2952
GAT TCT TCT TCA GAT AAT GAT AGC CTA ACT GAC TTG GAA AGT TTA AGC TCT GGT GAA TCA AAT GAA ATC AAA GTA ACT AAT GAT
D   S   S   S   D   N   D   S   L   T   D   L   E   S   L   S   S   G   E   S   N   E   I   K   V   T   N   D

        2964        2976        2988        3000        3012        3024        3036
TTA GAT ACA ACT GCT GAA AAG GAC CAA ATT CAG GCA GGC AAA TGG TTT GAT CCT GTA TTG GAT TGG CGA AAA TCT GAT CGT GAA
L   D   T   S   A   E   K   D   Q   I   Q   A   G   K   W   F   D   P   V   L   D   W   R   K   S   D   R   E

        3048        3060        3072        3084        3096        3108        3120
TTG ACG AAA AAC ATT CTT TGG AGG ATA GCT GAT AAA ACG ACA TAC GAT AAA GAA ACA ATA ACT GAC CTT ATT GAA CAA GGG ATC
L   T   K   N   I   L   W   R   I   A   D   K   T   T   Y   D   K   E   T   I   T   D   L   I   E   Q   G   I

        3132        3144        3156        3168        3180        3192        3204
CCA AAA CAT AGT TAT TTA AGT GGC AAT CCA TTA ACT AGT GTG ACT AAC GAC ATT TGC TCT GTT GAA AAC TAT GAA ACA TCA AGT
P   K   H   S   Y   L   S   G   N   P   L   T   S   V   T   N   D   I   C   S   V   E   N   Y   E   T   S   S

        3216        3228        3240        3252        3264        3276        3288
GCT TTC TTT TAC CAA CAA GTG CAC AAG AAG GAC AGA TTA CAA TAT TTG CCA TTA TAT GCA GTT TCT ACA TTT GAA AAC ACA AAT
A   F   F   Y   Q   Q   V   H   K   K   D   R   L   Q   Y   L   P   L   Y   A   V   S   T   F   E   N   T   N

        3300        3312        3324        3336        3348        3360        3372
AAC ACT GAA AAA AAT GAT GTG ACC AAT AAA AAT ATC AAT ATA GGG AAA CAT AGC CAA GAA CAA AAT TCT TCT TCC GCT AAA CCA
N   T   E   K   N   D   V   T   N   K   N   I   N   I   G   K   H   S   Q   E   Q   N   S   S   S   A   K   P

        3384        3396        3408        3420        3432        3444        3456
TCC CAA ATT CCA ACC GTG TCT TCT CCA TTA GGA TTC GAA GAA ACA AAG CTA AGT ACC ACG CCT ACT AAA AGC AAT AGA AGA GTG
S   Q   I   P   T   V   S   S   P   L   G   F   E   E   T   K   L   S   T   T   P   T   K   S   N   R   R   V

        3468        3480        3492        3504        3516        3528        3540
TCG CAT AGT GAT ACT AAT TCT AGC AAA CCC AAA AAC ACG AAG GAG AAC CTT TCA AAA AGC TCT TGG AGG CAA GAA TGG CTT GCA
S   H   S   D   T   N   S   S   K   P   K   N   T   K   E   N   L   S   K   S   S   W   R   Q   E   W   L   A

        3552        3564        3576        3588        3600        3612        3624
AAT TTG AAA CTT ATT TCC GTT TCG TTG GTT GAT GAG TTC CCT TCG GAG CTT TCC GAC AGT GAT AGA CAA ATA ATT AAC GAA AAA
N   L   K   L   I   S   V   S   L   V   D   E   F   P   S   E   L   S   D   S   D   R   Q   I   I   N   E   K

        3636        3648        3660        3672        3684        3696        3708
ATG CAG TTA CTT AAA GAT ATA TTT GCT AAC AAC CTT AAA TCA GCA ATT TCC AAT AAT TTT AGA GAG AGT GAC ATC ATT ATA CTG
M   Q   L   L   K   D   I   F   A   N   N   L   K   S   A   I   S   N   N   F   R   E   S   D   I   I   I   L

        3720        3732        3744        3756        3768        3780        3792
AAA GGT GAA ATA GAA GAT TAC CCA ATG AGT TCT GAA ATT AAG ATT TAC TAC AAC GAA TTA CAG AAC AAG CCT GAT GCA AAA AAA
K   G   E   I   E   D   Y   P   M   S   S   E   I   K   I   Y   Y   N   E   L   Q   N   K   P   D   A   K   K

        3804        3816        3828        3840        3852        3864        3876
GCC AGA TTT TGG TCA TTT ATG AAG ACT CAG AGA TTT GTT TCC AAC ATG GGA TTT GAT ATT CAG AAG TCC TGT GAA CCT GTT TCT
A   R   F   W   S   F   M   K   T   Q   R   F   V   S   N   M   G   F   D   I   Q   K   S   C   E   P   V   S

        3888        3900        3912        3924        3936        3948        3960
ATA TCC ACG TCG GTC AAA CCA CAT GTG GTC GAG CCC GAG CAT ATG GCA GAT GCC AAA ATT ATG CCT AAG GAT ATA CTG CAA ATT
I   S   T   S   V   K   P   H   V   V   E   P   E   H   M   A   D   A   K   I   M   P   K   D   I   L   Q   I

        3972        3984        3996        4008        4020        4032        4044
ACA AAA AAA CCG TTA ATG GTT AAA AAT GTG AAG CCT TCT TCT CCG CCA GAT GTG AAG TCT TTG GTG CAG CTG AGC ACA ATG GAA
T   K   K   P   L   M   V   K   N   V   K   P   S   S   P   P   D   V   K   S   L   V   Q   L   S   T   M   E

        4056        4068        4080        4092        4104        4116        4128
ACG AAA ACC CTA CCA GAA AAG AAG CAA TTC GAC AGT ATT TTT AAT TCT AAC AAA GCA AAA ATA ATC CCT GGA AAT GGG AAG CAC
T   K   T   L   P   E   K   K   Q   F   D   S   I   F   N   S   N   K   A   K   I   I   P   G   N   G   K   H

        4140        4152        4164        4176        4188        4200        4212
GCA TCA GAA AAC ATC TCA CTC TCT TTC TCA AGA CCT GCG TCC TAC GGC TAT TTT TCT GTT GGA AAA AGG GTT CCA ATC GTT GAA
A   S   E   N   I   S   L   S   F   S   R   P   A   S   Y   G   Y   F   S   V   G   K   R   V   P   I   V   E

        4224        4236        4248        4260        4272        4284        4296
GAT CGT CGA GTG AAA CAA CTC GAC GAT ATA ACA GAC AGT AAT ACA ACA GAA ATT TTA ACT AGT GTT GAC GTT TTA GGA ACA CAT
D   R   R   V   K   Q   L   D   D   I   T   D   S   N   T   T   E   I   L   T   S   V   D   V   L   G   T   H

        4308        4320        4332        4344        4356        4368        4380
TCA CAA ACT GGT ACT CAA CAA TCA AAT ATG TAC ACA TCA ACC CAA AAA ACA GAA CTT GAA ATT GAT AAT AAG GAT AGT GTC ACC
S   Q   T   G   T   Q   Q   S   N   M   Y   T   S   T   Q   K   T   E   L   E   I   D   N   K   D   S   V   T

        4392        4404        4416        4428        4440        4452        4464
GAA TGT TCG AAG GAC ATG AAA GAA GAT GGT CTT TCC TTT GTG GAT ATC GTT TTG TCG AAA GCA GCA TCG GCG CTG GAT GAA AAG
E   C   S   K   D   M   K   E   D   G   L   S   F   V   D   I   V   L   S   K   A   A   S   A   L   D   E   K

        4476        4488        4500        4512        4524        4536        4548
GAA AAA CAA TTG GCT GTT GCA AAT GAA ATT ATT CGG TCT TTG TCA GAT GAA GTT ATG AGG AAT GAA ATT AGA ATA ACT TCA CTT
E   K   Q   L   A   V   A   N   E   I   I   R   S   L   S   D   E   V   M   R   N   E   I   R   I   T   S   L

        4560        4572        4584        4596        4608        4620        4632
CAA GGT GAT TTA ACT TTT ACA AAG AAA TGT CTT GAA AAT GCG AGA AGT CAA ATA TCT GAA AAA GAT GCT AAA ATT AAC AAA TTG
Q   G   D   L   T   F   T   K   K   C   L   E   N   A   R   S   Q   I   S   E   K   D   A   K   I   N   K   L

        4644        4656        4668
ATG GAA AAA GAT TTT CAA GTG AAT AAG GAG ATA AAA CCG TAT TGA
M   E   K   D   F   Q   V   N   K   E   I   K   P   Y   *

      4687       4697       4707       4717       4727       4737       4747       4757       4767       4777
CTTTTTATCA TTCTACAAAA GACCAGTAAC GAAGTGTACC CTGTTTTTCT TTGTCGTTTT CTTCTATAAT GAACAATTTA CTTTACGGTC TTTAATCGAT

      4787       4797       4807       4817       4827       4837       4847
TTTAATGATT AGTTGGGCTT TTGGCTACAT ATGTAAAATT ACATATAATA TAATGGACAC ATAAGCAAAA
```

FIG. 3. Nucleotide sequence of *SIR4*. The nucleotide sequence of yeast genomic DNA extending from a *Hin*dIII site 600 bp upstream of the initial ATG of the *SIR4* coding region to a site 170 bp downstream of the TGA codon terminating the coding region is shown. Beneath the sequence of the coding region itself is indicated the predicted amino acid sequence of the *SIR4* protein. In addition, the location of the probable cap site of the *SIR4* mRNA is marked by the arrow above the sequence.

TABLE 2. Proteins exhibiting homology to the yeast *SIR4* protein

Protein[a]	Extent of overlap[b]	% Identity in overlap	z score[c]
Myosin heavy chain I, nematode	424[d]	15	9.2
50S ribosomal protein L23, *E. coli*	87	20	6.8
H-2 class II histocompatibility antigen AαC	76	21	6.3 (10.9)
cAMP-dependent protein kinase, bovine	95[d]	19	4.8
Coat protein, tobacco mosaic virus	89[d]	17	4.3
Gene IV protein, M13, fd, and f1 phages	92[d]	19	4.2
MATα2 protein, *S. cerevisiae*	87	21	3.9 (8.4)
*MAT***a***1* protein, *S. cerevisiae*	96	24	3.5 (5.8)
80S ribosomal protein eL12, shrimp	87	14	3.8
72K DNA-binding protein, adenovirus 2	196	21	3.4 (5.5)
50S ribosomal protein L1, *Xenopus*	74	31	3.3
Cloacin DF13 protein, *E. coli* plasmid C1	241	16	3.0 (5.2)
Laminin B1, mouse	108	22	3.0 (3.7)
β-Fibrinogen, bovine/human	162	17	2.7 (9.8)

[a] Protein homologies were identified by comparison of the predicted amino acid sequence of *SIR4* protein with 3,477 polypeptides in the Dayhoff data bank, using the algorithm described by Lipman and Pearson (23) at a setting of Ktup = 2, unless otherwise noted. Polypeptides with an optimized homology score of >60 were subjected to random shuffle analysis to determine the statistical significance of the identified homology. Only those peptides whose homologies are at least possibly significant are shown.

[b] Number of amino acid residues encompassed by the homology. The location within *SIR4* protein of each of the homologous domains is shown in Fig. 4.

[c] Significance of homologies are determined by comparison of the similarity score of the actual sequence with that of sequences derived by a random shuffle of the residues within the homologous domain. This is represented as a *z* value, which is calculated as follows: (similarity score − mean of random scores)/standard deviation of random scores. *z* values of >3 are interpreted to be possibly significant; those >6, as probably significant; and those >10 as significant. Statistical analysis of each protein was conducted at the same Ktup setting as that at which the protein was originally identified. Parenthetical value is the *z* value based on a comparison of the identified protein to a truncated version of the *SIR4* protein, encompassing that third of the *SIR4* protein spanning the region of homology.

[d] These proteins were identified in the Dayhoff bank, using Ktup = 1 in the Lipman and Pearson algorithm.

We determined the codon bias to be 0.033, which is virtually random with respect to codon preference. This is consistent with the observation that *SIR4* is normally transcribed at a very low level (13). The protein is unusually rich in serine (12.8%) and lysine (11.6%) and on the average is very hydrophilic. *SIR4* has a predicted pI of 9.7.

We compared the predicted protein sequence of *SIR4* for similarities with 3,477 polypeptide sequences (780,111 residues) in the Dayhoff amino acid bank. A number of proteins were identified that exhibited some sequence similarity to *SIR4*. The statistical significance of the sequence similarity of each of these proteins was determined by a random shuffle analysis, as described by Lipman and Pearson (23). The level of significance is expressed in terms of a "*z* score." *z* scores of >3 signify possible statistical significance. Proteins exhibiting significant homology to *SIR4* are listed in Table 2, and the region of *SIR4* to which each of these proteins is homologous is indicated in Fig. 4. As is evident from this list, the diversity of function of these proteins makes it difficult to draw conclusions about the biological significance of these homologies.

SIR4 mutations exhibit unusual intragenic _trans_-complementation. We have observed an unusual intragenic complementation pattern for various *sir4* mutant alleles. Ivy et al. previously showed that plasmid pJH5.1A (see Fig. 5) could complement the *sir4-351* ochre allele (13). From the *SIR4* DNA sequence, we noted that this plasmid encodes only the carboxy-terminal 46% of the *SIR4* protein. This implied that either *SIR4* activity is encompassed entirely within the carboxy half of the protein or the carboxy terminus of *SIR4* protein can yield wild-type activity by acting in *trans* with a truncated protein produced by the chromosomal *sir4* gene containing the ochre mutation. To determine which of these possibilities is true, we examined the ability of various plasmid-borne subfragments of *SIR4* to complement different mutant alleles of *SIR4* in the chromosome. We examined three chromosomal *sir4* alleles in addition to *sir4-351*: *sir4Δ1::HIS3*, a deletion/insertion mutation in which the middle third of the gene is replaced by a fragment containing the yeast *HIS3* gene (see Fig. 6); *sir4Δ2::HIS3*, a deletion/insertion mutant in which almost all the amino portion of the gene is replaced by the same *HIS3* fragment in the same orientation as that in *sir4Δ1::HIS3*; and *sir4Δ3:URA3*, a complete deletion of the *SIR4* coding region, replaced by the yeast *URA3* gene. We made these null alleles of *SIR4* in vitro and used them to replace the wild-type copy of *SIR4* in vivo in strain S150-2B by homologous recombination.

We measured *SIR4* function by quantitative mating assays. Loss of *SIR4* activity results in expression of α mating-type genes, resident at *HML* in our test strain, in addition to **a** mating type genes resident at *MAT* and *HMR*. Concurrent expression of both **a** and α mating-type genes in a cell yields a sterile phenotype typical of **a**/α diploid cells. Thus, the degree to which the test strain regains its capacity to mate as an **a** cell is a direct measure of the degree to which *HML* is repressed and, accordingly, a measure of the level of functional *SIR4* activity.

Each of the chromosomal *sir4* mutations reduced the mating of strain S150-2B to between $1/10^4$ and $1/10^5$ that of the SIR^+ parent (Table 3). We introduced all or parts of the *SIR4* gene into these $sir4^-$ strains on a high-copy *LEU2* plasmid and then scored *SIR4* function by quantitative mating assays. The results are summarized in Table 3 and Fig. 5. Plasmid pKAN59, carrying the intact *SIR4* gene, restores mating competence to that of wild type for all four

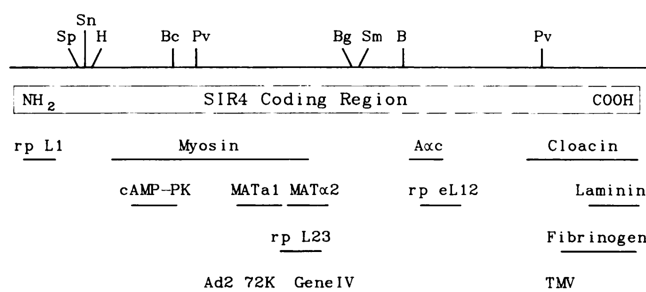

FIG. 4. Proteins homologous to *SIR4*. Beneath a diagram of the *SIR4* coding region are shown the locations and extents of the regions of *SIR4* protein exhibiting statistically significant homologies to each of those proteins listed in Table 2. Restriction site abbreviations are as indicated in the legend to Fig. 2. Ad2, Adenovirus type 2; TMV, tobacco mosaic virus.

TABLE 3. Complementation of *sir4* mutations[a]

Plasmid	Plasmid-borne gene	No. of diploids formed/haploid plasmid-bearing test cell with given allele:			
		sir4-351	*sir4Δ1::HIS3*	*sir4Δ2::HIS3*	*sir4Δ3::URA3*
None	None	0.00015	0.00003	0.00001	0.00001
pKAN59	*SIR4*	0.5	0.6	0.6	0.8
pJH5.1A	*sir4*[b]	0.8	ND[c]	ND	0.009
pSIR4.3[d]	*sir4*[b]	0.4	0.3	0.007	0.007
pJH3A	*sir4*[b]	0.0002	0.00001	ND	ND
pKAN69	*SIR3*	0.03	0.00001	ND	ND

[a] Strain S150-2B containing the indicated chromosomal *sir4* allele and harboring the designated plasmid was assayed for restoration of Sir activity by the quantitative mating assay described in Materials and Methods, using strain DC17 (*MATα his1*) as mating tester. Under these assay conditions, strain S150-2B exhibits a mating efficiency of 0.7.

[b] Plasmids pJH5.1A, pSIR4.3, and pJH3A encode and expresses the carboxy 46, 45, and 12% of *SIR4*, respectively.

[c] ND, Not determined.

[d] Mating assays were performed on glucose-containing media, which permitted only low-level expression from the *GAL10-SIR4* fusion resident on the plasmid.

mutant strains. In addition, as previously reported, plasmid pJH5.1A, a derivative of pKAN59 in which the promoter and amino-terminal 54% of the *SIR4* gene is deleted, yields wild-type levels of *SIR* activity in a *sir4-351* background. Plasmid pSIR4.3 carries a construct consisting of the *GAL10* promoter fused to the carboxy 45% of the *SIR4* gene (at the *Sma*I site at position 2831 of the sequence in Fig. 3). We presume that translation of this truncated gene initiates at the first available AUG codon within the fused coding region, at residue 747 of the protein. Both the *sir4-351*- and the *sir4Δ1::HIS3*-containing strains were complemented by

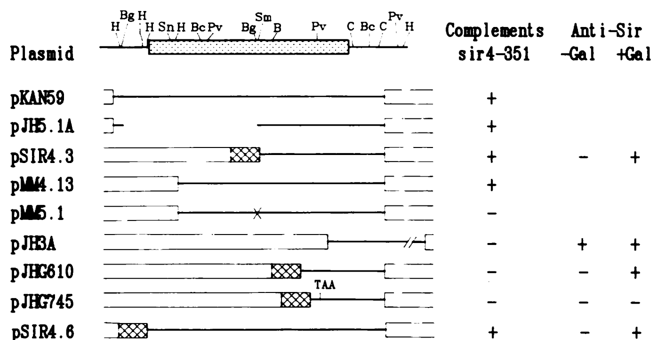

FIG. 5. Biological activity of plasmid-borne fragments of the *SIR4* gene. Beneath a diagram of the genomic region spanning the *SIR4* gene (dotted box) are shown the specific subcloned segments (heavy lines) carried on each of the plasmids listed to the left. Cross-hatched boxes represent the *GAL10* promoter, oriented to yield galactose-inducible transcription of the cloned *SIR4* segment. For plasmids pJHG610 and pJHG745 the promoter fragment includes the initial ATG of the *GAL10* coding region, which is in-frame with the *SIR4* coding region in the former case and out-of-frame in the latter case (noted by the TAA codon that terminates the nonsense peptide expected to be produced from that plasmid). Plasmid pMM5.1 is identical to plasmid pMM4.13 except that the *Bgl*II site within the coding region was filled in to yield a frameshift mutation at the site of the "X." The ability of each plasmid to restore normal mating to a *sir4-351* strain, summarizing data presented in Table 3, is indicated to the right of the diagram of each plasmid. These complementation tests were performed on glucose-containing media. As such, for those plasmids carrying *GAL10-SIR4* fusions, the assay results represent the complementing activity of the *SIR4* construct in its transcriptionally repressed state. Anti-Sir activity of individual plasmids, summarizing results shown in Fig. 7, is indicated. Anti-Sir activity was determined by the mating pattern of HO27B harboring the indicated plasmid as measured on media promoting induction (+Gal) or repression (−Gal) of the *GAL10* promoter. Restriction site abbreviations are as in the legend to Fig. 2.

plasmid pSIR4.3 to levels near that obtained with the intact gene (Table 3). We observe complementation even though the strains were grown in glucose medium, which yields very low-level expression from the *GAL10* promoter (42). In contrast, plasmid pSIR4.3 fails to complement *sir4Δ2::HIS3* or *sir4Δ3::URA3* under similar conditions. Thus, *SIR4* function does not reside solely in the carboxy half of the gene. Rather, these results suggest that two domains of the protein, one at the amino end and one at the carboxy end, are required for *SIR4* repression of the silent mating-type loci but that these two domains need not be covalently linked to establish and maintain repression.

Intragenic complementation obtained with different *sir4* alleles is independent of the mating type of the cell. We crossed strain Y558(pSIR4.3), which carries the *sir4Δ1::HIS3* allele and plasmid pSIR4.3, to a *MATα* strain, FVY1-1C. Upon scoring of the progeny spores from this diploid, we found that plasmid pSIR4.3 could complement the *sir4Δ1::HIS3* disruption in both *MAT***a** and *MATα* backgrounds (data not shown). In a similar fashion, we found that pSIR4.3 failed to complement *sir4Δ2::HIS3* in either a *MAT***a** or a *MATα* background. Thus, intragenic complementation affects expression of *HMR* in exactly the same fashion as it does that of *HML*.

Expression of smaller domains of the carboxy end of *SIR4* does not yield complementation of *sir4-351*. Plasmids pJH3A and pJHG610, which express the carboxy 12 and 25% of the *SIR4* gene, respectively (see below), do not restore mating competence to any of the strains harboring chromosomal *sir4* mutations (Fig. 5).

Extragenic complementation of *sir4* mutants by excess *SIR3* activity. Ivy et al. (13) observed that the *SIR3* gene on a high-copy-number vector (plasmid pKAN63) could complement strains containing the *sir4-351* ochre mutation, although not as efficiently as could the intact *SIR4* clone (plasmid pKAN59). To determine whether *SIR3* complementation of *sir4-351* is allele specific or results from bypass suppression, we introduced the multicopy *SIR3* plasmid into a strain bearing the null allele *sir4Δ1::HIS3* and tested for complementation of the mating defect (Fig. 6). We found that *SIR3* did not complement the *sir4Δ1::HIS3* deletion/insertion mutation (Table 3). Thus, hyper-expression of *SIR3* does not abrogate the need for some *SIR4* activity to maintain repression of the silent loci.

Overexpression of the carboxy end of *SIR4* disrupts normal *SIR* repression. We have determined that expression of the carboxy end of the *SIR4* protein causes disruption of *SIR* repression in vivo. Ivy et al. (13) previously noted that plasmid pJH3A, which carries a yeast genomic fragment

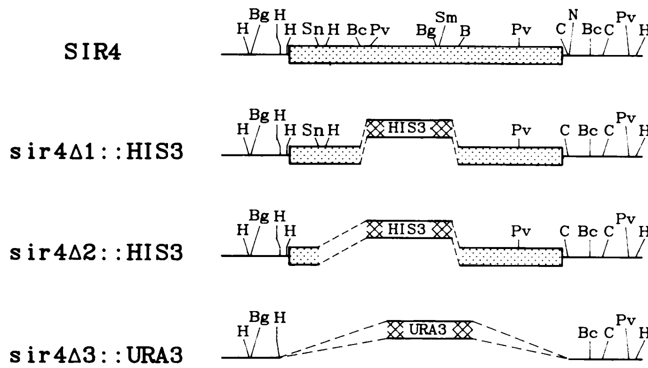

FIG. 6. Deletion alleles of *sir4*. The structure of three deletion alleles of *sir4*, constructed as described in Materials and Methods, are shown beneath a restriction map of the region spanning the *SIR4* gene (dotted box). Restriction sites are as indicated in the legend to Fig. 2.

spanning the carboxy end of *SIR4* and the adjacent transcription unit, induced a dominant Sir$^-$ phenotype. From analysis of the sequence of *SIR4*, we find that plasmid pJH3A yields a translational fusion of the carboxy 12% of *SIR4* to the amino portion of a reading frame within the pBR322 moiety of the plasmid. Thus, we presume that plasmid pJH3A expresses at some unknown level in yeasts a fusion protein consisting in part of the carboxy 12% of the *SIR4* protein.

To determine whether anti-*SIR* activity of plasmid pJH3A is a consequence of the expression of the carboxy end of *SIR4*, we fused the *GAL10* promoter, including the initial ATG of the *GAL10* gene, to the *SIR4* coding region at sites of two random linker insertions, designated 610 and 745. For linker 610, this yields an in-frame fusion of the *GAL10* coding region to that of *SIR4*, allowing galactose-inducible expression of the carboxy 25% of the *SIR4* protein. For linker 745, the fusion is out of frame. We examined anti-Sir activity of these plasmids in strain HO27B (*HMLα mata1 HMRα SIR*$^+$). The null *mata1* allele in strain HO27B does not contribute functional mating-type information to the cell (40, 43). Strain HO27B normally mates as an **a**, since strains lacking functional mating-type information behave as **a** cells (43). However, if the silent α mating-type genes of strain HO27B are expressed, the strain mates as an α cell, since no functional **a** mating-type information is present to cause a sterile **a**/α phenotype. Thus, strain HO27B mates as an α cell when the silent loci are expressed but as an **a** cell when the silent loci are repressed.

Results of this analysis are shown in Fig. 7 and summarized in Fig. 5. As previously reported, we find that the presence of plasmid pJH3A completely converts the mating type of strain HO27B from **a** to α. We found that strain HO27B carrying the pJHG610 fusion mates as an **a** strain on galactose-free media and as an α strain on galactose-containing medium. On the other hand, strain HO27B carrying the pJHG745 fusion mates as an **a** on either medium. Thus, the presence of the carboxy end of the *SIR4* protein in the cell disrupts normal repression of *HML* and *HMR* by the *SIR* products. Consistent with this, we find that plasmid pSIR4.3, which carries the carboxy 45% of *SIR4* under control of the *GAL10* promoter and complements *sir4-351* at low-level expression (see above), has anti-*SIR* activity in strain HO27B when the strain is grown on galactose to induce high-level expression (Fig. 7). Finally, we have found that overexpression of the entire *SIR4* gene yields disruption of *SIR* repression. That is, strain HO27B containing the *GAL10-SIR4* plasmid pSIR4.6 mates as an **a** when grown on

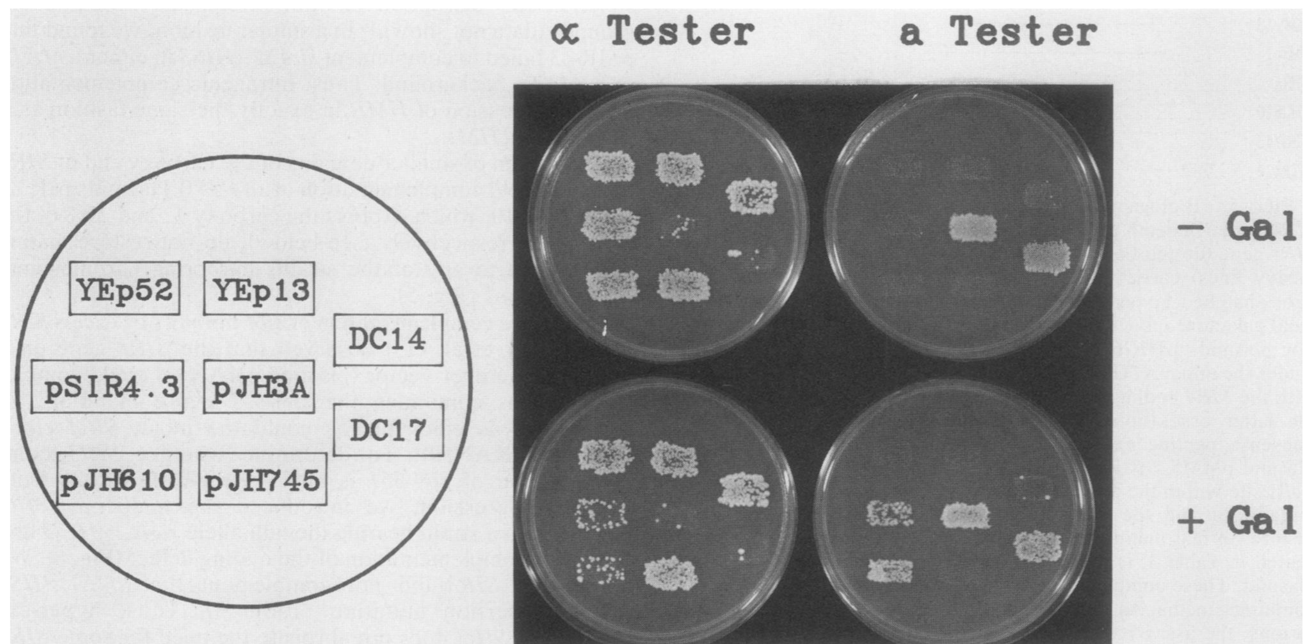

FIG. 7. Anti-Sir activity of *SIR4* plasmids. Strains 14**a**, 17α, and HO27B containing the indicated plasmids were patched in the designated pattern on SC minus leucine plates that contained 2% glycerol plus 2% ethanol as carbon source and either contained (lower plates) or lacked (upper plates) 1% galactose. After incubation for 2 days at 30°C, the master plates were each replicated to two synthetic minimal plus glucose plates spread with a lawn of either an **a** *leu2* (JK-1A; right) or an α *leu2* (KAY5-3B; left) tester strain. The galactose-containing master plate was replicated to the lower two plates and the galactose-free plate was replicated to the two upper plates. The photograph shows the mating tester plates after 3 days of incubation at 30°C.

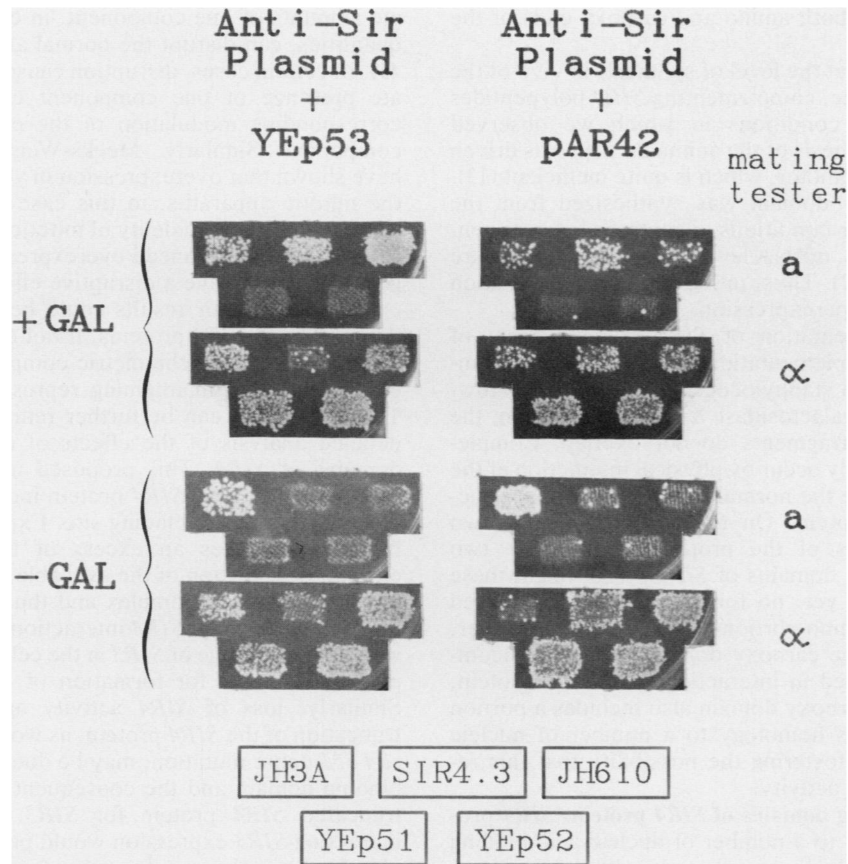

FIG. 8. Reversal of *SIR4* anti-Sir activity by overexpression of *SIR3*. Strain JR281B-7C (*HMLα mata1 HMRα*) was sequentially transformed, first with one of the *LEU2-sir4* plasmids possessing anti-Sir activity or the control vector YEp51 or YEp52 and then with the *URA3-GAL10-SIR3* plasmid pAR42 or the control vector YEp53. Patches of these doubly transformed strains were grown on SC plus 2% glycerol–2% ethanol plates lacking uracil and leucine and either containing (+GAL) or lacking (−GAL) 2% galactose. After 2 days of growth, these master plates were replicated to similar plates lacking also adenine and histidine and spread with a lawn of either an **a** (JK-1A) or an α (MS-43) mating-type tester strain. Shown are portions of these replica plates after an additional 2 days of incubation at 30°C. The particular anti-Sir plasmid present in each strain is indicated in the key at the bottom.

glucose but mates predominantly as an α when grown on galactose (data not shown). Thus, overexpression of the carboxy end of *SIR4*, whether as a partial fragment or as the intact protein, interferes with *SIR* repression.

***SIR3* overexpression reverses the anti-*Sir* activity of *SIR4* fragments.** The observation that the *sir4-351* allele is complemented by overexpression of *SIR3* while a *sir4* gene disruption is not suggests that *SIR3* and *SIR4* might interact to repress *HML* and *HMR*. If such an interaction does occur, it is possible that the anti-*SIR* effect of the carboxy terminus of *SIR4* is due to disruption of such an interaction. To test this hypothesis, we examined the mating pattern of an *HMLα mata1 HMRα* strain containing both an anti-*SIR* plasmid and a galactose-inducible *SIR3* gene. These results are shown in Fig. 8. Strain JR281B-7C harboring an anti-Sir plasmid (either plasmid pJH3A or the galactose-inducible plasmid pSIR4.3 or pJHG610) and plasmid pAR42 (YEp53-*SIR3*) or, as control, YEp53, was tested for mating after growth in the presence or absence of galactose. As described in the preceding section, an *HMLα mata1 HMRα* strain harboring plasmid JH3A mates as an α cell and such a strain harboring plasmid pSIR4.3 or pJHG610 mate as **a** cells when grown in the absence of galactose but as α cells when grown in its presence. The concurrent presence of vector YEp53 in the strain does not alter these mating patterns (Fig. 8). In contrast, α mating of galactose-grown JR281B-7C containing either of the anti-*SIR* plasmids is substantially reduced when plasmid pAR42 is also present in the strain (top line of Fig. 8, comparing α mating of anti-Sir, YEp53 strains to that of anti-Sir, pAR42 strains). Thus, although the presence of the carboxy end of *SIR4* disrupts *SIR* repression to allow phenotypic expression of α information from *HML* and *HMR* in strain JR281B-7C, such disruption is blocked and repression of the silent loci is restored by coexpression of *SIR3* at high levels.

DISCUSSION

Intragenic complementation of *sir4* mutations. We have provided evidence that *SIR4* activity resides in two separate domains of the *SIR4* protein and that these two domains need not be physically linked to effect repression. We found that a subclone expressing only the carboxy-terminal 46% of the protein could restore near-wild-type *SIR4* activity in strains carrying a *sir4-351* mutation (an ochre mutation located at an undetermined site within the gene) as well as strains containing the *sir4Δ1::HIS3* allele, which expresses only the amino 26% of the *SIR4* gene. Since the same subclone fails to complement strains more extensively deleted for *SIR4*, we conclude that *SIR4* activity requires

protein domains from both amino and carboxy ends of the normal *SIR4* protein.

It should be noted that the level of synthesis in vivo of the each of the two separate, complementing *SIR4* polypeptides is quite low. Under conditions in which we observed complementation, synthesis of the amino domain was driven by the normal *SIR4* promoter, which is quite inefficient (13). Similarly, the carboxy domain was synthesized from the *GAL10* promoter under conditions of repressed expression. Under these conditions, mRNA levels from the promoter are <0.01 copy per cell (42). Thus, intragenic complementation is not an artifact of hyperexpression.

Intragenic complementation of *SIR4* is reminiscent of β-galactosidase α-complementation and of fragment complementation seen with staphylococcal nuclease (46). However, in contrast to β-galactosidase α-complementation, the complementing *SIR4* fragments do not overlap. Complementation could possibly occur by physical interaction of the two domains to restore the normal three-dimensional structure of the native protein. On the other hand, the two complementing regions of the protein could define two independent functional domains of *SIR4*. What might these two functions be? As yet, no function can be associated specifically with the amino portion of the protein. However, as described below, the carboxy domain probably encompasses a region involved in interaction with *SIR3* protein. The complementing carboxy domain also includes a portion of the protein that has homology to a number of nucleic acid-binding proteins, fostering the possibility that this region also possess such activity.

Possible DNA-binding domains of *SIR4* protein. *SIR4* protein exhibits homology to a number of nucleic acid-binding proteins. These homologies are loosely clustered in two domains: one near the center of the molecule and the second spanning the carboxy third of the protein. We have found that a fragment of the *SIR4* protein encompassing the carboxy 45% of molecule and produced from plasmid pSIR4.3 is retained on a single-strand DNA cellulose column (M. Marshall and J. Broach, unpublished observations). Thus, we suspect that this portion of the molecule, which includes the second domain of homology, does in fact possess DNA-binding activity. We have as yet been unable to define specific double-strand DNA-binding activity of the protein, but such studies are in progress.

Model for *SIR3-SIR4* interaction in formation of a repressor complex. Our results suggest that *SIR* repression of the silent mating-type loci requires formation of a multicomponent repressor complex with stringent stoichiometric constraints. First, we find that overexpression of either all of the *SIR4* protein or only the carboxy end of the protein disrupts repression. This overexpression disruption can be reversed by coincident overexpression of *SIR3* protein. In addition, overexpression of *SIR3* yields allele-specific suppression of *sir4* mutants. That is, overexpression of *SIR3* will not bypass the need for *SIR4* activity but it will yield a partial Sir^+ phenotype in a strain carrying the *sir4-351* ochre allele. Thus, excess *SIR3* protein will both restore function to a mutant *SIR4* protein lacking its carboxy end and mitigate the loss of Sir repression caused by overexpression of the carboxy end of *SIR4*.

Precedent for interpreting these observations emerges from investigations of bacteriophage morphogenesis and of suprastructure assembly in yeasts. For example, assembly of bacteriophage T4 requires interaction of a number of molecular components, many of which must be present in reasonably precise stoichiometric amounts. Disproportionate amounts of one component, in either excess or limiting quantities, can disrupt the normal assembly process (6, 39, 41). In certain cases, disruption caused by the disproportionate presence of one component can be alleviated by a corresponding modulation of the expression of a second component. Similarly, Meeks-Wagner and Hartwell (25) have shown that overexpression in yeasts of a component of the mitotic apparatus, in this case a histone protein, can adversely affect the fidelity of mitotic chromosome transmission. However, balanced overexpression of paired histones genes does not have a disruptive effect.

In this light, our results could be interpreted to suggest that *SIR3* and *SIR4* proteins, if not *SIR1* and *SIR2* proteins as well, form a stoichiometric complex that is required for establishing and maintaining repression of the silent loci. This hypothesis can be further refined on the basis of our detailed analysis of the effects of overexpression of subdomains of *SIR4*. This proposed model predicts that the carboxy 12% of the *SIR4* protein includes at least a portion of the *SIR3* protein-binding site. Expression of the anti-*SIR* fragment provides an excess of the *SIR3* binding site, entrapping a portion of the available *SIR3* protein in the cell in a nonfunctional complex and thus making it unavailable for productive *SIR3-SIR4* interaction. Increasing the expression or gene dosage of *SIR3* in the cell makes additional *SIR3* protein available for formation of a productive complex. Similarly, loss of *SIR4* activity arising from mutational truncation of the *SIR4* protein, as would be the case with the *sir4-351* ochre mutation, may be due to deletion of the *SIR3* binding domain and the consequent reduced affinity of the truncated *SIR4* protein for *SIR3* protein. In this case, increasing *SIR3* expression would promote productive complex formation by compensating for this reduced affinity for *SIR3* protein by providing a substantially increased concentration of *SIR3* protein in the cell.

An alternative hypothesis to the structural model presented above is that *SIR3* protein catalyzes some requisite posttranslational modification of *SIR4* protein and that hyperexpression of the carboxy end of the *SIR4* protein effectively limits available *SIR3* activity. Consistent with this hypothesis is the observation that *SIR4* protein is multiply phosphorylated in vivo (R. Schnell, W. Kimmerly, and J. Rine, personal communication). However, the primary structure of *SIR3* does not resemble that of known protein kinases, so it is unlikely that *SIR3* functions in this capacity. In addition, suppression of the *sir4* ochre allele by overexpression of *SIR3* is difficult to accommodate in this, or similar, models. Thus, we favor a model postulating formation of a multicomponent complex mediating *SIR* repression as the most economical explanation of our observations.

Sir regulation of silent mating-type loci. We have analyzed the *SIR4* gene with the hope of providing some insight into the mechanism of Sir regulation of the silent mating-type loci. As described above, products of the *SIR2*, *SIR3*, and *SIR4* genes, and to a lesser extent that of the *SIR1* gene, act through specific sites flanking *HML* and *HMR* to render the DNA across these loci refractory to transcription by RNA polymerase II or III (1, 3, 5, 8, 13, 16, 32, 35). In addition, this system prevents cleavage by HO endonuclease at the HO cleavage site within each locus (19, 21, 44). Thus, by several criteria Sir repression acts by making DNA in the vicinity of the E sites inaccessible to enzymes that normally should be able to interact with sequences resident there. The only known activity to which DNA at the silent loci is accessible is the process of double-strand break-mediated

recombination in the course of directed transposition of mating-type sequences to the *MAT* locus.

The mechanism by which the *SIR* gene products render DNA inaccessible is not known, although several characteristics of the process have been defined. Chromatin structure across *HML* or *HMR*, as measured by the location and intensity of DNase I and micrococcal nuclease-sensitive sites of genomic DNA in isolated nuclei, is only minimally altered in Sir^+ versus Sir^- strains or in comparison with *MAT* (27). The only striking difference is the appearance of a strong DNase I-hypersensitive site at the HO cleavage site at *HML* and *HMR* in conjunction with loss of SIR repression. In addition, Miller and Nasmyth (26) have shown that the continued presence of the *SIR* products is required to maintain repression. Shifting strains carrying temperature-sensitive mutations in *sir3* or *sir4* to the nonpermissive temperature yielded immediate activation of the silent loci, even in α-factor-arrested cells. However, the converse was not true. Repression could be reestablished only if cells passed through S phase at the permissive temperature. Shifting α-factor-arrested cells to the permissive temperature did not yield repression. Thus, it appears that *SIR* repression can only be established during the process of DNA replication.

These observations in conjunction with the results presented in this report suggest that repression of the silent mating-type loci involves a rigorous organization of the components of the repression machinery. Not only may the protein components need to be present in precise stoichiometic amounts, but their interaction with the target site itself requires its assuming a specific structure, one that occurs only during DNA replication. One possibility is that the repressor complex nucleates around a single-stranded version of the E sites, a possibility consistent with the similarity of *SIR4* protein to proteins that interact with RNA and with our observation that a *SIR4* fragment binds single-stranded DNA. Further analysis of the properties of *SIR4* protein should help to clarify the molecular basis of this novel repression system.

ACKNOWLEDGMENTS

We thank Steve Pilder and Kevin Ryder for participation in this project.

This work was supported by Public Health Service grants GM34545 (to J.R.B.) and GM25624 (to J.B.H.) from the National Institutes of Health. J.R.B. is an Established Investigator of the American Heart Association.

LITERATURE CITED

1. **Abraham, J., K. A. Nasmyth, J. N. Strathern, A. J. S. Klar, and J. B. Hicks.** 1984. Regulation of mating-type information in yeast: negative control requiring sequences both 5′ and 3′ to the regulated region. J. Mol. Biol. **176**:307–331.
2. **Astell, C. R., L. Ahlstrom-Jonasson, M. Smith, K. Tatchell, K. A. Nasmyth, and B. D. Hall.** 1981. The sequence of the DNAs coding for the mating-type loci of Saccharomyces cerevisiae. Cell **27**:15–23.
3. **Brand, A. H., L. Breeden, J. Abraham, R. Sternglanz, and K. Nasmyth.** 1985. Characterization of a silencer in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. Cell **41**:41–48.
4. **Broach, J. R., Y. Y. Li, L. C.-C. Wu, and M. Jayaram.** 1983. Vectors for high-level, inducible expression of cloned genes in yeast, p. 83–117. *In* M. Inouye (ed.), Experimental manipulation of gene expression. Academic Press, Inc., New York.
5. **Feldman, J. B., J. B. Hicks, and J. R. Broach.** 1984. Identification of the sites required for repression of a silent mating type locus in yeast. J. Mol. Biol. **178**:815–834.
6. **Floor, E.** 1970. Interaction of morphogenetic genes of bacteriophage T4. J. Mol. Biol. **47**:293–306.
7. **Haber, J. E.** 1983. Mating type genes of Saccharomyces cerevisiae, p. 559–619. *In* J. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
8. **Haber, J. E., and J. P. Geroge.** 1979. A mutation that permits the expression of normally silent copies of mating-type information in Saccharomyces cerevisiae. Genetics **93**:13–35.
9. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**: 351–360.
10. **Herskowitz, I., and Y. Oshima.** 1981. Control of cell type in yeast, p. 181–209. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), The molecular biology of the yeast Saccharomyces. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Hicks, J. B., J. N. Strathern, and A. J. S. Klar.** 1979. Transposable mating type genes in Saccharomyces cerevisiae. Nature (London) **282**:478–483.
12. **Hicks, J. B., J. N. Strathern, A. J. S. Klar, S. Ismail, and J. R. Broach.** 1984. Structure of the SAD mutation and the location of control sites at silent mating type genes in yeast. Mol. Cell. Biol. **4**:1278–1285.
13. **Ivy, J. M., A. J. S. Klar, and J. B. Hicks.** 1986. Cloning and characterization of four *SIR* genes of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6**:688–702.
14. **Klar, A. J. S.** 1980. Interconversion of yeast cell types by transposable genes. Genetics **95**:631–648.
15. **Klar, A. J. S., and S. Fogel.** 1979. Activation of mating type genes by transposition in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA **76**:4539–4543.
16. **Klar, A. J. S., S. Fogel, and K. MacLeod.** 1979. MAR1—a regulator of HMa and HM(alpha) loci in Saccharomyces cerevisiae. Genetics **93**:37–50.
17. **Klar, A. J. S., J. B. Hicks, and J. N. Strathern.** 1982. Directionality of yeast mating-type interconversion. Cell **28**: 551–561.
18. **Klar, A. J. S., J. N. Strathern, J. R. Broach, and J. B. Hicks.** 1981. Regulation of transcription in expressed and unexpressed mating type cassettes of yeast. Nature (London) **289**:239–244.
19. **Klar, A. J. S., J. N. Strathern, and J. B. Hicks.** 1981. A position-effect control for gene transposition: state of expression of yeast mating type genes affects their ability to switch. Cell **25**:517–524.
20. **Klar, A. J. S., J. N. Strathern, and J. B. Hicks.** 1984. Developmental pathways in yeast, p. 151–195. *In* R. Losick and L. Shapiro (ed.), Microbial development. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Kostriken, R., J. N. Strathern, A. J. S. Klar, J. B. Hicks, and F. Heffron.** 1983. A site specific endonuclease essential for mating type switching in Saccharomyces cerevisiae. Cell **35**:167–174.
22. **Kushner, P. J., L. C. Blair, and I. Herskowitz.** 1979. Control of yeast cell types by mobile genes: a test. Proc. Natl. Acad. Sci. USA **76**:5264–5268.
23. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Meeks-Wagner, D., and L. H. Hartwell.** 1986. Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell **44**:43–52.
26. **Miller, A. M., and K. A. Nasmyth.** 1984. Role of DNA replication in the repression of silent mating type loci in yeast. Nature (London) **312**:247–251.
27. **Nasmyth, K. A.** 1982. The regulation of yeast mating-type chromatin structure by SIR: an action at a distance affecting both transcription and transposition. Cell **30**:567–578.
28. **Nasmyth, K. A.** 1982. Molecular genetics of yeast mating type.

Annu. Rev. Genet. **16**:439–500.
29. **Nasmyth, K. A., and K. Tatchell.** 1980. The structure of transposable yeast mating type loci. Cell **19**:753–764.
30. **Nasmyth, K. A., K. Tatchell, B. D. Hall, C. R. Astell, and M. Smith.** 1981. A position effect in the control of transcription at yeast mating type loci. Nature (London) **289**:244–250.
31. **Rine, J., and I. Herskowitz.** 1987. Four genes responsible for a position effect on expression from *HML* and *HMR* in *Saccharomyces cerevisiae*. Genetics **116**:9–22.
32. **Rine, J. D., J. N. Strathern, J. B. Hicks, and I. Herskowitz.** 1979. A suppressor of mating-type locus mutations in Saccharomyces cerevisiae: evidence for and identification of cryptic mating type loci. Genetics **93**:877–901.
33. **Rothstein, R. J.** 1983. One-step gene disruption in yeast. Methods Enzymol. **101**:202–211.
34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
35. **Schnell, R., and J. Rine.** 1986. A position effect on the expression of a tRNA gene mediated by the *SIR* genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6**:494–501.
36. **Sherman, F., G. R. Fink, and J. B. Hicks.** 1979. Methods in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Shore, D., M. Squire, and K. Nasmyth.** 1984. Characterization of two genes required for the position-effect control of yeast mating-type genes. EMBO J. **3**:2817–2823.
38. **Shortle, D., and D. Botstein.** 1983. Directed mutagenesis with sodium bisulfite. Methods Enzymol. **100**:457–468.
39. **Showe, M., and L. Onorato.** 1978. Kinetic factors and form determination of the head of bacteriophage T4. Proc. Natl. Acad. Sci. USA **75**:4165–4169.
40. **Sprague, G. F., Jr., L. C. Blair, and J. Thorner.** 1983. Cell interactions and regulation of cell type in the yeast Saccharomyces cerevisiae. Annu. Rev. Microbiol. **37**:623–660.
41. **Sternberg, N.** 1976. A genetic analysis of bacteriophage (lambda) head assembly. Virology **71**:686–692.
42. **St. Johns, T. P., and R. W. Davis.** 1981. The organization and transcription of the galactose gene cluster of *Saccharomyces*. J. Mol. Biol. **152**:285–315.
43. **Strathern, J. N., J. B. Hicks, and I. Herskowitz.** 1981. Control of cell type in yeast by the mating type locus: the (alpha)1-(alpha)2 hypothesis. J. Mol. Biol. **147**:357–372.
44. **Strathern, J. N., A. J. S. Klar, J. B. Hicks, J. A. Abraham, J. M. Ivy, K. A. Nasmyth, and C. McGill.** 1982. Homothallic switching of yeast mating type cassettes is initiated by a double-strand cut in the MAT locus. Cell **31**:183–192.
45. **Strathern, J. N., E. Spatola, C. McGill, and J. B. Hicks.** 1980. Structure and organization of transposable mating type cassettes in Saccharomyces yeast. Proc. Natl. Acad. Sci. USA **77**:2839–2843.
46. **Zabin, I.** 1982. Beta-galactosidase (alpha)-complementation: a model of protein-protein interaction. Mol. Cell. Biochem. **49**:87–96.