MOLECULAR AND CELLULAR BIOLOGY, Dec. 1984, p. 2735–2744
0270-7306/84/122735-10$02.00/0

Vol. 4, No. 12

Primary Structure of the *RAD52* Gene in *Saccharomyces cerevisiae*

KENJI ADZUMA, TOMOKO OGAWA, AND HIDEYUKI OGAWA*

Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan

Received 30 May 1984/Accepted 10 September 1984

The *RAD52* gene of *Saccharomyces cerevisiae*, which is involved in genetic recombination and DNA repair, was cloned by transformation of *rad52-1* mutant cells to methyl methanesulfonate resistance with *Bam*HI fragments of Rad$^+$ genomic DNA inserted into the *Escherichia coli-S. cerevisiae* shuttle vector YRp7. A plasmid carrying a 2.0-kilobase *Bam*HI fragment was found to partially complement methyl methanesulfonate sensitivity of the *rad52-1* mutant. By using this fragment as a hybridization probe, a plasmid that fully complemented the methyl methanesulfonate sensitivity of the mutant was isolated, which carries a 3.3-kilobase *Sal*I fragment containing most of the 2.0-kilobase *Bam*HI fragment. Analysis of the nucleotide sequence of the *Sal*I fragment revealed the presence of a large open reading frame of 1,512 nucleotides. The *rad52-1* mutant DNA has a single-base change in this reading frame, which leads to an amino acid substitution. Analysis of mRNA synthesized in yeast by the S1 mapping technique disclosed possible transcription initiation and termination points of the *RAD52* gene and suggested formation of the gene product without splicing of the transcript.

Our understanding of the molecular mechanism of genetic recombination in *Escherichia coli* has been advanced significantly by in vitro studies with purified proteins (10). On the other hand, lack of an in vitro method limits our knowledge of eucaryotic recombination. It is therefore appealing to study the mechanism of eucaryotic recombination by the use of a protein(s) whose function in recombination is suggested by in vivo studies.

Among eucaryotic organisms that are available for such an approach, *Saccharomyces cerevisiae* has several advantages. Genetic studies on its chromosomes have been carried out extensively (22), *S. cerevisiae-E. coli* shuttle vectors have been developed (1, 36, 37), and many mutations (about 60) have been identified which affect genetic recombination (16).

Among the mutations producing defects in recombination, *rad52-1* on chromosome XIII (28) has been investigated most extensively. Originally the *rad52-1* mutant was isolated as a strain which was extremely sensitive to ionizing radiations and to methyl methanesulfonate (MMS) (11, 28). Cells carrying the *rad52-1* mutation have defects in repair of double-strand breaks (15, 30), completion of mating-type interconversion (14, 18), transformation by linear DNA (24), and recombination during meiosis and mitosis (12, 27). Furthermore, *rad52-1/rad52-1* diploid cells are known to be defective in production of viable spores (11, 27) and accumulate single-strand breaks in their chromosomes during meiosis (29).

Although the *RAD52* gene is believed to code for a protein involved in recombination, the gene product has not been identified yet. As a first step toward understanding its role in recombination in molecular terms, we cloned the *RAD52* gene and analyzed its structure by determining the DNA sequence and the location and orientation of the transcripts by S1 nuclease protection experiments.

The cloning of the *RAD52* gene has been independently reported by Schild et al. (32).

* Corresponding author.

MATERIALS AND METHODS

Strains and media. *S. cerevisiae* strains g160-2b (*MAT***a** *rad52-1 ade2-1 arg4 arg9 trp1 his5 lys1-1 ilv3 leu2 pet*), XS2470-7C (*MAT*α *rad52-1 trp1 leu1-1 his3-532*), and a standard wild-type strain, S288C (*MAT*α *gal2 SUC2 mal*), were obtained from S. Nakai (National Institute of Radiological Sciences, Chiba City, Japan). Strain D13-1A (*MAT***a** *trp1 his3-532 gal2 SUC2 mal*) (36) was obtained from K. Murakami (Kanazawa University, Kanazawa City, Japan). The *E. coli* strains used were N23-53 (*thr tyr phe trpC galT-6 str recA41*) (23) and HMS174 (r$^-$ m$^+$ Rif^r *recA1*) (4).

Synthetic complete medium (SC) consisted of 0.67% yeast nitrogen base without amino acids (Difco Laboratories) and 2% glucose. For nutrient medium, YPD (1% yeast extract, 2% peptone, 2% glucose) or MYPD (0.3% malt extract, 0.3% yeast extract, 0.5% peptone, 1% glucose) was used. Regeneration agar for yeast transformation consisted of SC containing 1 M sorbitol and 2% Bacto-agar (Difco).

***S. cerevisiae* transformation.** *S. cerevisiae* transformation was carried out by the method of Struhl et al. (37) with the following modifications. *S. cerevisiae* cells harvested at early-log phase (10^7 per ml) were treated with a solution of 20 mM EDTA and 30 mM 2-mercaptoethanol at 30°C for 10 min as suggested by H. Tanaka, K. Murakami, and H. Yoshikawa (personal communications) before treatment with Helicase (Pharmindustrie, Clichy, France) or Zymolyase 60000 (Seikagaku Kogyo, Tokyo, Japan). The optimal conditions for transformations with respect to generation of spheroplasts with these enzymes were determined for each strain. For strains g160-2b and XS2470-7C, the cells were treated with 3 mg of Helicase and 5 μg of Zymolyase 60000 per ml, respectively, at 30°C for 1 h. Under these conditions, transformation frequencies were about 1×10^3 for strain g160-2b and 5×10^4 for strain XS2470-7C with respect to Trp$^+$ per μg of YRp7 DNA.

Preparation of DNA and determination of nucleotide sequences. *S. cerevisiae* DNA was prepared by the method of Cryer et al. (8). Covalently closed circular plasmid DNA from *E. coli* was prepared by the procedure of Clewell and

Helinski (7). Single-stranded viral DNAs from recombinant M13mp7 phages were prepared according to the manual provided by Bethesda Research Laboratories. Nucleotide sequences of DNA were determined as described by Maxam and Gilbert (19). For purine-specific modifications, we used acetic acid by the method of Som and Tomizawa (33).

Hybridization and preparation of DNA probes. DNA blotting to a nitrocellulose filter and the subsequent DNA hybridization were carried out as described previously (31, 34), and in situ colony hybridization was performed by the method of Grunstein and Wallis (13). ^{32}P-labeled AK28 or SL1 probe was prepared by the terminal labeling method by using [γ-^{32}P]ATP (3,000 Ci/mmol) and T4 polynucleotide kinase after *Hin*fI digestion (specific activity, 0.3×10^7 to 1.2×10^7 cpm/μg).

Preparation of RNA from *S. cerevisiae* and S1 mapping of the transcripts. *S. cerevisiae* total RNA was prepared from the spheroplasts of S288C by the method of Sripati and Warner (35) with a slight modification. RNA containing polyadenylate tracts, polyadenylated [poly(A)$^+$] RNA, was isolated by polyuridine-Sepharose-4B (Pharmacia) chromatography as described previously (39). Typical yields of RNA from 1.0 g (wet weight) of spheroplasts were roughly 200 optical density units at 260 nm for the total RNA and 1.5 to 3.0 optical density units at 260 nm for the poly(A)$^+$ RNA.

To determine the location and orientation of transcripts from the *RAD52* region, the S1 mapping procedure of Berk and Sharp (2) was used. As probes we used the unlabeled single-stranded viral DNAs of recombinant M13mp7 phages which carried various portions of the SL1 or AK28 fragment. The length of DNA protected from S1 digestion was analyzed by Southern blotting analysis by using ^{32}P-labeled SL1 fragment as probe. A 20- to 30-μg poly(A)$^+$ RNA sample prepared from S288C was mixed with 0.5 to 1.0 μg of single-stranded viral DNA of the recombinant M13mp7 and precipitated with ethanol. The pellet was suspended to 30 μl hybridization buffer [40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid), pH 6.4, 1 mM EDTA, 400 mM NaCl, 80% formamide], heated at 80°C for 10 min, and then incubated at 49°C for 3 h. The mixture was diluted 10-fold with precooled S1 buffer (50 mM sodium acetate, pH 4.6, 280 mM NaCl, 4.5 mM $ZnSO_4$) containing 100 to 1,200 U of S1 nuclease (Sigma Chemical Co.) and incubated at 37°C for 30 min. After S1 digestion, the sample was extracted once with phenol-chloroform-isoamyl alcohol (25:24:1) and precipitated with ethanol together with 10 μg of carrier tRNA. The resulting precipitate was suspended in 40 μl of loading buffer (10 mM Tris-hydrochloride, pH 7.6, 1 mM EDTA, 0.1% each of xylene cyanol-FF and bromophenol blue, 50% glycerol). Half of the sample was electrophoresed on a neutral agarose gel in 89 mM Tris-borate containing 2 mM EDTA. To the remaining half, 5 μl of 250 mM NaOH was added, and the mixture was heated at 90°C for 1 min and run on an alkaline agarose gel in 30 mM NaOH containing 2 mM EDTA (20). After electrophoresis, the DNA in the gels was blotted to nitrocellulose filters and subjected to hybridization with ^{32}P-labeled SL1 fragment as probe.

Construction of a *S. cerevisiae* gene bank. *S. cerevisiae* genomic DNA was prepared from a Rad$^+$ strain, D13-1A. Fifty micrograms of D13-1A DNA was digested with *Bam*HI and ligated to 6 μg of *Bam*HI-digested DNA of the *E. coli-S. cerevisiae* shuttle vector YRp7. This vector carries a 1,453-base-pair *Eco*RI fragment containing the *S. cerevisiae TRP1* gene and *S. cerevisiae* chromosomal replication origin *ARS1* (autonomously replicating sequence) in the *Eco*RI site of pBR322 (36, 40). For transformation of *E. coli* with this ligation mixture, N23-53 cells were grown to 6×10^8 per ml in 400 ml of Luria broth (LB) and treated twice with 50 mM $CaCl_2$ for 5 min at 0°C. The cells were mixed with 1/10 of the ligation mixture and kept at 0°C for 15 min. The mixture was then diluted with 1.2 liters of LB prewarmed to 37°C and incubated at 37°C for 90 min. Tetracycline was added to a final concentration of 4 μg/ml, and the shaking was continued for 30 min at 37°C. At this point, D-cycloserine was added to a final concentration of 100 μg/ml, and shaking was continued for an additional 90 min. The cells were then

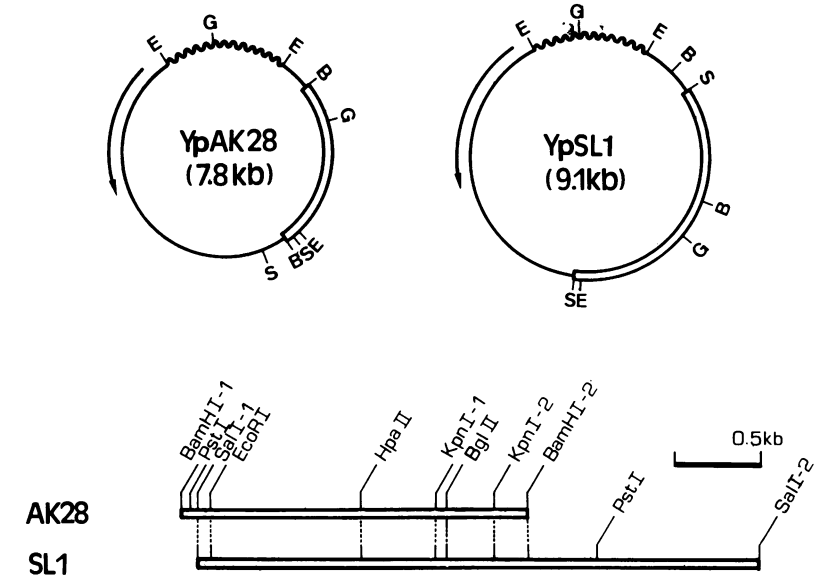

FIG. 1. Restriction maps of plasmids YpAK28 and YpSL1. Thin line, pBR322 sequence; wavy line, *TRP1-ARS1* sequence; open bar, AK28 or SL1. Arrow indicates the location and orientation of *bla* (Amp^r). Restriction sites shown are *Eco*RI (E), *Bam*HI (B), *Bgl*II (G), and *Sal*I (S). Detailed maps of the cloned *Sal*I 3.3-kb fragment (SL1) in YpSL1 and of the *Bam*HI 2.0-kb fragment (AK28) in YpAK28 are shown below the maps of the plasmids.

harvested and plated on 200 LB plates containing 50 μg of ampicillin per ml. The resulting ampicillin-resistant (Amp^r) colonies (ca. 3,000) were examined for tetracycline sensitivity, and the sensitive (Tet^s) colonies (which constituted 30 to 40% of the total number of Amp^r colonies) were retained. By repeating the above transformation eight times, we obtained 8,400 Amp^r Tet^s transformants harboring hybrid plasmids. They were divided randomly into 28 groups, each consisting of 300 independent clones. From each group, plasmid DNAs were prepared as a mixture of 300 clones and pooled separately (DNA pools of *S. cerevisiae* gene bank).

Strategy for selection of a hybrid plasmid complementing the MMS sensitivity of a *rad52-1* mutant from the *S. cerevisiae* gene bank. Each sample (ca. 10 μg of DNA) of 28 DNA pools of our *S. cerevisiae* gene bank was used to transform a *S. cerevisiae rad52-1 trp1* mutant, g160-2b. The spheroplasts (0.25 ml) generated from 5×10^8 cells grown in 50 ml of YPD by Helicase treatment were mixed with the DNA and diluted 10-fold with a solution containing 20% polyethylene glycol 4000, 10 mM Tris-hydrochloride, pH 7.6, and 10 mM $CaCl_2$. The resulting spheroplast-DNA mixture (2.5 ml) was combined with 22.5 ml of melted regeneration agar lacking tryptophan at 50°C, and 5-ml aliquots were poured onto five plates of the regeneration agar (25 ml). When the plates were solidified, 5-ml aliquots of the regeneration agar were overlaid onto each plate to bury the spheroplasts into the sandwich of the regeneration agars. After incubation for 24 h at 30°C, 5 ml of the regeneration agar containing 0.08% MMS was overlaid onto four out of five plates, and incubation was continued for 3 to 4 days at 30°C. The final concentration of MMS in the plate was expected to be 0.01%. The number of Trp^+ transformants appearing on the control plate (without MMS) was about 1×10^4 to 5×10^4. In the presence of MMS, g160-2b spheroplasts carrying YRp7 survived at a frequency of 10^{-4}.

Isolation of DNA segments containing the complete *RAD52* gene (SL1) and the *rad52-1* mutation (RX1). Genomic DNA of a Rad^+ strain, D13-1A, was digested with *Sal*I, fractionated by 0.6% agarose gel electrophoresis, and transferred to a nitrocellulose filter. The DNA filter was subjected to hybridization with a ^{32}P-labeled AK28 fragment. The resulting autoradiogram demonstrated that only one DNA band of ca. 3.3 kb in length was hybridized with the probe. A 100-μg portion of the *Sal*I digest of D13-1A DNA was electrophoresed on 0.6% agarose gel, and a gel slice containing 3.0- to 3.5-kb DNA fragments was cut out. The DNA extracted from the slice (ca. 1 μg) was ligated to 1 μg of YRp7 DNA at the *Sal*I site. The ligation mixture was used to transform *E. coli* HMS174 to Amp^r Tet^s. Three hundred Amp^r Tet^s colonies of the transformants were immobilized on a nitrocellulose filter and subjected to hybridization with a ^{32}P-labeled AK28 probe. Three colonies showed a positive response. The plasmids recovered from the colonies were found to have an insertion of the identical 3.3-kb *Sal*I fragment (SL1) at the *Sal*I site of YRp7.

The 2.0-kb *Bam*HI fragment (1.5 μg) purified from the genomic DNA (85 μg) of the *rad52-1* mutant g160-2b was inserted to YRp7 (4 μg) at the *Bam*HI site. The resulting DNA was used for transformation of *E. coli* HMS174 followed by the colony hybridization procedure as described above. Among 550 Amp^r Tet^s colonies, 3 were hybridized with the probe. The plasmids recovered from these clones had an insertion (RX1) which was structurally identical with AK28 in size and in restriction pattern.

Measurement of MMS sensitivity. Yeast cells in log phase (1×10^7 per ml) or in stationary phase (2×10^8 per ml) in SC liquid medium supplemented with 0.2% Casamino Acids (Difco) and histidine were harvested, washed twice with 0.1 M phosphate buffer, (pH 7.0), and suspended with the same buffer. After addition of MMS to a final concentration of 0.3%, the cells were incubated at 30°C. At appropriate time intervals, 0.1-ml portions were withdrawn and immediately diluted 10-fold with 10% sodium thiosulfate. After appropriate dilutions, the cells were plated on SC agar medium supplemented with 0.1% Casamino Acids and histidine. The colonies formed were counted after 5 to 7 days of incubation at 29°C.

To assay sensitivity with MMS-containing plates, cells were grown in SC liquid medium supplemented with 0.2% Casamino Acids and histidine and plated on SC agar medium containing 0.1% Casamino Acids, histidine, and various concentrations of MMS. The plates were used immediately after preparation. Surviving colonies were counted after 3 to 6 days of incubation at 29°C.

RESULTS

Complementation of MMS sensitivity of the *rad52-1* mutant with cloned DNA segments. A *S. cerevisiae rad52-1 trp1* mutant, g160-2b, was used for selection of the *RAD52* gene from the *S. cerevisiae* gene library, which was constructed by joining quasi-random *Bam*HI fragments of Rad^+ genomic DNA to plasmid YRp7. This plasmid was one of the *S. cerevisiae-E. coli* shuttle vectors and carried the *S. cerevisiae TRP1* gene and the *ARS1* segment, which allowed autonomous plasmid replication in *S. cerevisiae*. Since *rad52-1* mutants are extremely sensitive to MMS, MMS-resistant (MMS^r) Trp^+ transformants were selected on a plate lacking tryptophan and containing 0.01% MMS. Transformations with 4 of 28 DNA pools of the *S. cerevisiae* gene library yielded 20 to 400 Trp^+ MMS^r transformants, whereas none of the other DNA pools did. A single colony of the Trp^+ MMS^r transformants was inoculated into YPD, incubated overnight, and plated on YPD medium. Most of the colonies formed were found to lose the MMS^r and Trp^+ phenotypes simultaneously. DNA prepared from the Trp^+ MMS^r transformants was used to transform *E. coli* HMS174 to Amp^r. A plasmid obtained from an Amp^r transformant allowed the *rad52-1* mutant to grow on an MMS-containing plate. This plasmid, YpAK28, had an insertion of the 2.0-kb *Bam*HI fragment of *S. cerevisiae* DNA (denoted AK28) at the *Bam*HI site of YRp7 (Fig. 1). YpAK28 was also capable of complementing the MMS sensitivity of another *rad52-1* strain, XS2470-7C. However, when this *Bam*HI fragment (AK28) was inserted in the opposite orientation at the same *Bam*HI site of YRp7, this reconstructed plasmid (YpAK28-i) could not complement the MMS sensitivity of either *rad52-1* mutant strain. This result suggested that AK28 did not contain a fully functional *RAD52* gene. Therefore, we attempted to isolate a larger fragment containing AK28 that could complement the MMS sensitivity when inserted in either orientation into YRp7.

DNA fragments generated by *Sal*I digestion of genomic DNA of a Rad^+ strain, D13-1A, were fractionated by agarose gel electrophoresis, transferred to a nitrocellulose filter, and subjected to hybridization with a ^{32}P-labeled AK28 fragment. DNA of ca. 3.3 kb in length was hybridized with the probe (data not shown). The DNA extracted from the gel slice corresponding to this 3.3kb band was inserted into YRp7 at the *Sal*I site and used for transformation of *E. coli* HMS174. The transformants obtained were subjected to

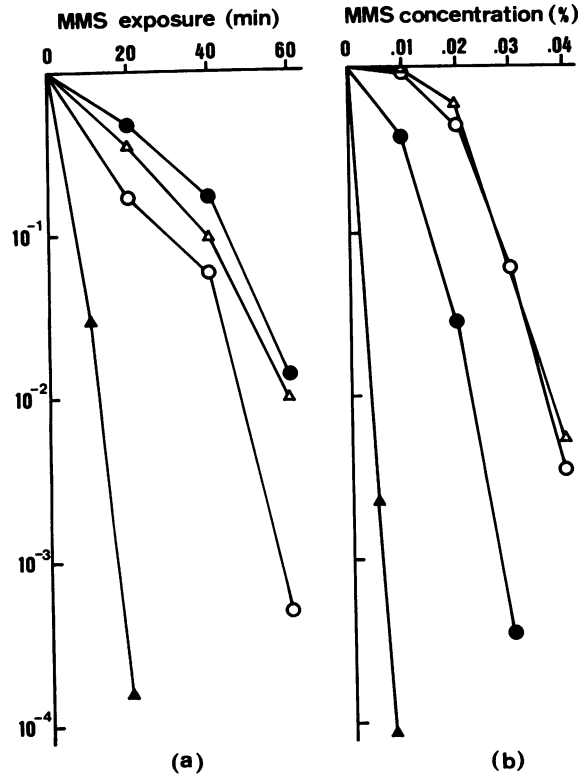

FIG. 2. MMS sensitivities of YpAK28 and YpSL1 transformants. (a) Survival fractions after increasing time of MMS treatment in a liquid culture. Log-phase cells grown in SC lacking tryptophan were treated with 0.3% MMS in 0.1 M phosphate buffer (pH 7.0), and the fractions of survivors were measured by plating a portion to SC plates lacking tryptophan. (b) Survival fractions after incubation on MMS plates. Cells in stationary phase in SC lacking tryptophan were plated on SC plates lacking tryptophan and containing various concentration of MMS. Results shown are for *rad52-1* mutant cells, XS2470-7C, carrying YpSL1 (○), YpAK28 (●), or YRp7 (▲); and Rad$^+$ cells, S288C (△).

in situ colony hybridization with ^{32}P-labeled AK28. Among 300 colonies, 3 gave a positive autoradiographic result and were found to harbor recombinant plasmids containing an identical 3.3-kb *Sal*I fragment (denoted SL1) at the *Sal*I site. In contrast to the plasmids carrying AK28, plasmids carrying SL1 in both possible orientations (YpSL1 and YpSL1-i; Fig. 1) conferred MMS resistance on both *rad52-1* mutants, XS2470-7C and g160-2b. Therefore, the SL1 fragment was most likely to contain the complete *RAD52* gene.

MMS sensitivity of *rad52-1* mutants carrying YpSL1 and YpAK28. The MMS sensitivity of the *rad52-1* mutant, XS2470-7C, carrying YpSL1 or YpAK28 was examined by two different methods: (i) temporary exposure of the cells to 0.3% MMS in a liquid culture for increasing periods of time, and (ii) spreading the cells on SC plates containing various concentrations of MMS. In the liquid culture method, the sensitivity of the *rad52-1* mutant carrying YpSL1 or YpAK28 was similar to that of the wild-type strain, S288C (Fig. 2a). In the MMS plate method, on the other hand, the mutant carrying YpAK28 was much more sensitive than the YpSL1-carrying mutant; the latter showed the same sensitivity as the wild-type strain (Fig. 2b).

Location of the *rad52-1* mutation in the DNA segment corresponding to AK28. To determine whether the SL1 or AK28 fragment contained the *RAD52* gene, or at least a part of it, rather than an unknown suppressor, we next isolated a DNA segment hybridizable to AK28 from a *rad52-1* mutant strain, g160-2b, to verify the presence of the *rad52-1* mutation in the segment. The *Bam*HI 2.0-kb DNA fragments were prepared from the *Bam*HI digests of g160-2b DNA by agarose gel electrophoresis and joined to YRp7 at the *Bam*HI site. The recombinant DNA was used to transform *E. coli* HMS174. The resulting Amp^r Tet^s colonies were subjected to colony hybridization with ^{32}P-labeled AK28 probe. Three of 550 colonies gave a positive autoradiographic result and harbored a recombinant plasmid carrying a 2.0-kb *Bam*HI fragment (denoted RX1). The restriction map of RX1 was identical to that of AK28. However, the plasmid YpRX1 that carried the RX1 fragment in the same orientation as that of AK28 in YpAK28 could not complement the MMS sensitivity of the *rad52-1* mutants at all. Therefore, we concluded that SL1 and AK28 contain most or all of the *RAD52* gene rather than an unknown suppressor.

Location and orientation of the *RAD52* gene in SL1. Various restriction fragments of SL1 were recloned on YRp7 to examine the ability to complement the MMS sensitivity of *rad52-1* mutant XS2470-7C. It was found that the region responsible for complementation of the MMS sensitivity had the *Hpa*II, *Kpn*I-1, *Bgl*II, and *Bam*HI-2 sites shown in Fig. 1. The 2.2-kb fragment from the *Hpa*II to *Sal*I-2 site could not confer MMS resistance. However, when this fragment was joined to plasmid pAM82 (21) at a site downstream of the *S. cerevisiae* acid phosphatase promoter to allow transcription to proceed in the direction from the *Hpa*II toward the *Sal*I-2 site, the resulting plasmid could confer MMS resistance.

The portable *lacZ* gene of *E. coli*, in which its own promoter and initiation codon are eliminated, was excised from pMC1871 (obtained from M. Casadaban) (5) with *Bam*HI and joined to YpAK28 at the *Bam*HI-2 site. The plasmid thus constructed expressed high β-galactosidase activity in *S. cerevisiae* as well as in *E. coli*, but it lost the ability to complement the MMS sensitivity of the *rad52-1* mutants. From these results, we predicted that the *RAD52* gene contained the *Hpa*II, *Kpn*I-1, *Bgl*II, and *Bam*HI-2 sites and was transcribed in the rightward direction (from *Hpa*II to *Bam*HI-2) on SL1 (Fig. 1).

Nucleotide sequence of the SL1 fragment and the position of the *rad52-1* mutation. The nucleotide sequence of 3,302 base pairs (bp) of the SL1 fragment was determined by the method of Maxam and Gilbert (Fig. 3). The positions of all potential initiation (ATG, GTG) and termination (TAG, TGA, TAA) codons are presented in Fig. 4. The nucleotide T in the left *Sal*I site (*Sal*I-1) is numbered as 1. There is a large open reading frame composed of 1,512 nucleotides from nucleotide 968 to 2479 in phase 3. Another large reading frame includes 807 nucleotides of the SL1 fragment (from 807 to 1) and extends beyond the left border of the fragment (phase 4). All the others were composed of less than 438 nucleotides.

We also determined the nucleotide sequence of the RX1 fragment carrying the *rad52-1* mutation. The region corresponding to most of the largest open reading frame (from 998 to 1949) was sequenced. The only difference found between RX1 and SL1 was at position 1236, at which cytosine was converted to thymine in RX1 as indicated in Fig. 3.

Transcription map of the *RAD52* region. To identify mRNA transcribed from the *RAD52* region, mRNA prepared from a

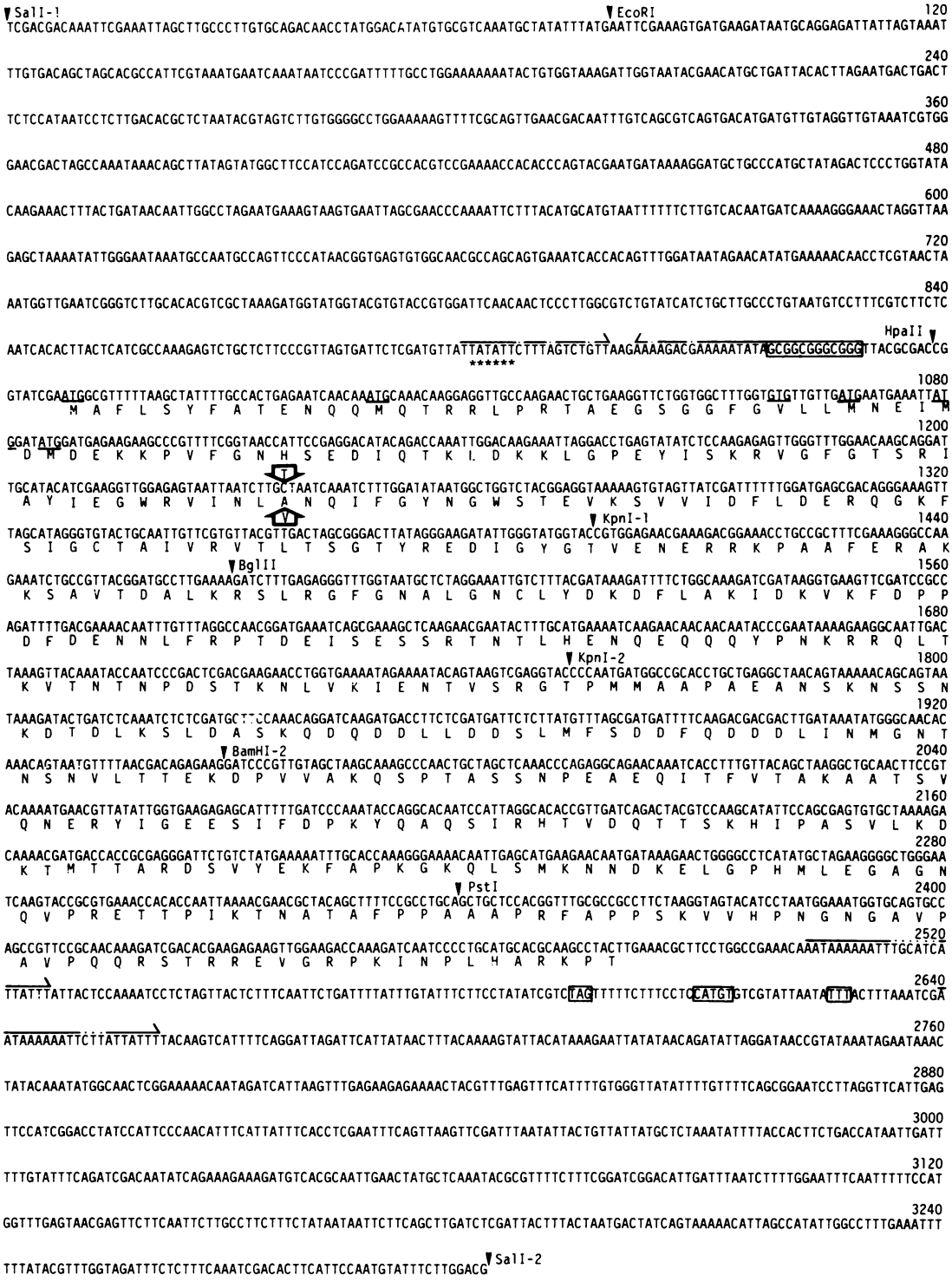

FIG. 3. Nucleotide sequence of SL1. The sequence of a region that makes up ca. 62% of SL1 (ca. 77% of the putative coding region of the *RAD52* gene) was determined in both orientations, and the remaining 38% of the sequence was determined in the same orientation at least twice. The sequence is shown only for the *RAD52* mRNA identical strand (i.e., the l-strand described in the text). Some restriction enzyme sites are indicated. The nucleotides are numbered from the nucleotide T in the *Sal*I-1 site. A TATA box-like sequence for the *RAD52* gene is indicated by asterisks. The inverted repeat sequence found in the 5′-flanking region and the direct repeat sequence downstream of the termination codon are shown by arrows over the sequences. The guanine-plus-cytosine stretch following the inverted repeat and the structure postulated to be involved in efficient transcription termination and polyadenylation by Zaret and Sherman (41) are shown in boxes. The possible initiation codons, five ATG and one GTG, are underlined. The putative amino acid sequence of the *RAD52* protein in its maximal size is shown by using the single-letter code. The *rad52-1* mutation is present at the nucleotide in an open arrow.

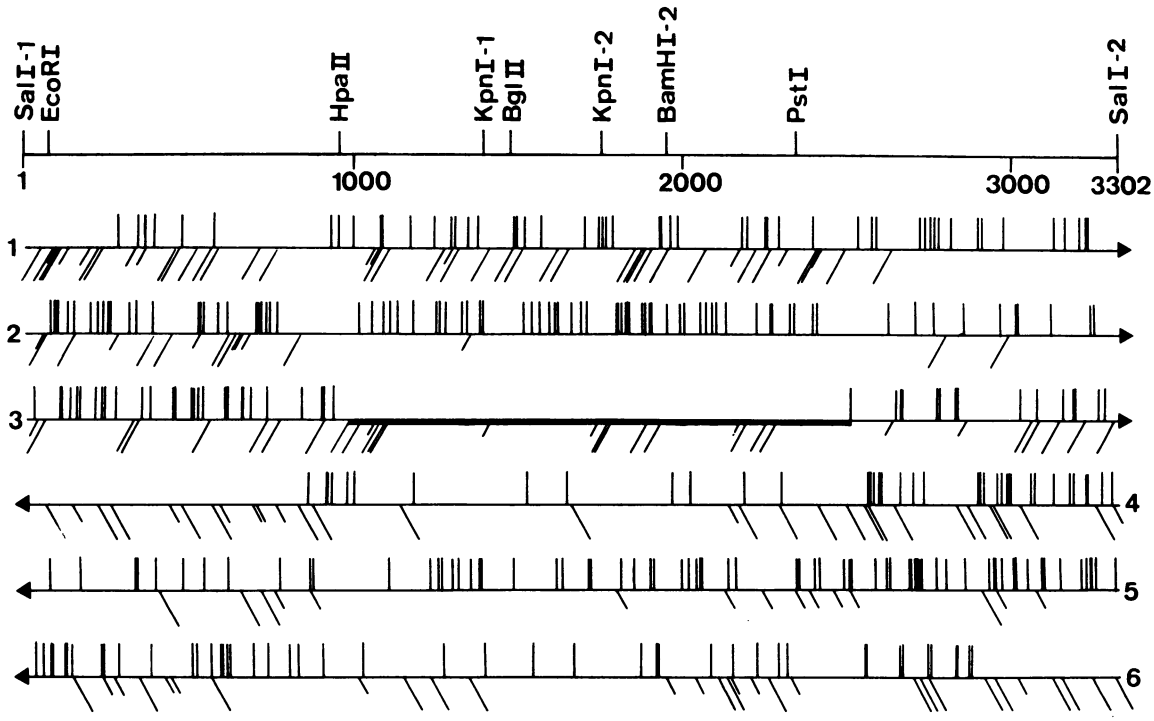

FIG. 4. Potential initiation and termination codons in SL1. A vertical line represents a termination codon (TAA, TGA, and TAG) and long and short slanting lines indicate the initiation codons ATG and GTG, respectively. The orientation is from left to right in phases 1 to 3 (l-strand), and phases 4 to 6 show the opposite orientation (r-strand). The putative *RAD52* coding frame is indicated by the thick line.

Rad$^+$ strain, S288C, was analyzed by the S1 mapping technique with unlabeled single-stranded viral DNAs of recombinant M13mp7 phages carrying different portions of the SL1 or AK28 fragment (Fig. 5). The lengths of the DNA segments protected from S1 nuclease digestion were deduced from their mobilities in neutral and alkaline agarose gel electrophoreses, which were detected by hybridization with ^{32}P-labeled SL1 probe after transfer to nitrocellulose filters.

Using the r-strand of SL1 (R1), we detected two bands different in size (1.65 and 1.78 kilobases [kb]) in both neutral and alkaline agarose gel electrophoreses (Fig. 6), indicating the synthesis of two rightward transcripts from the genomic *RAD52* region corresponding to the cloned SL1 fragment. To

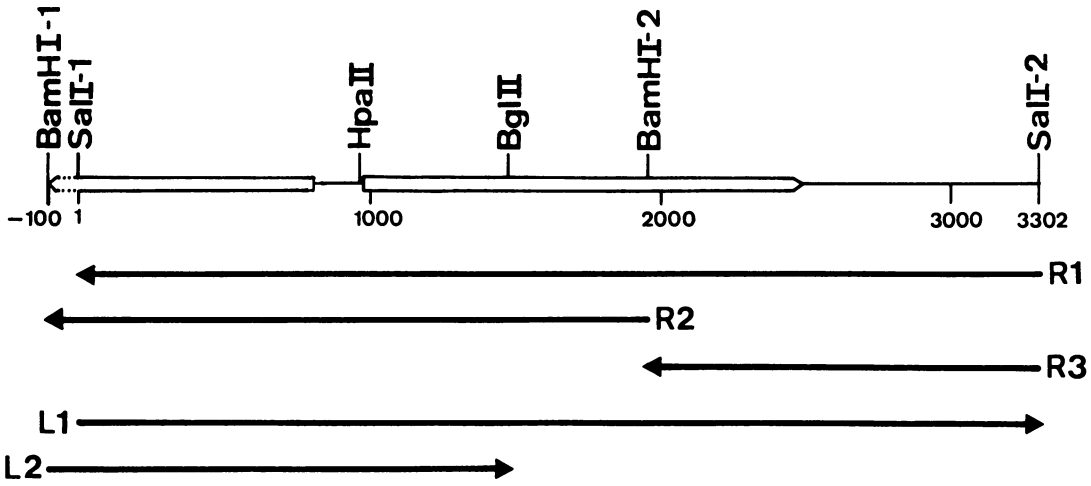

FIG. 5. Strategies to identify transcripts from the *RAD52* region. Open arrows represent the locations and orientations of the largest (phase 3 in Fig. 4) and the next largest (phase 4) open reading frames in SL1. Single-stranded DNAs of recombinant M13mp7 phages used as probes for S1 mapping are depicted by arrows: R1, R2, and R3 represent r-strands of SL1, AK28, and the fragment between the *Sal*I-2 and *Bam*HI-2 sites, respectively; and L1 and L2 were l-strands of SL1 and the fragment between the *Bam*HI-1 and *Bgl*II sites, respectively.

determine the regions that specify these two transcripts, the same mRNA preparation was used for hybridization with the r-strand of shorter fragments, R2 and R3 (Fig. 5). The analysis with the R2 strand showed the presence of only one band of 0.94 kb in length after treatment with an optimum concentration of S1 nuclease (333 U). With the R3 strand, two bands (0.69 and 0.83 kb) were detected. These results indicate that the rightward transcripts are initiated from a single point at position ca. 1010 (0.94 kb upstream from the *Bam*HI-2 site at position 1949) and terminate at different points at position ca. 2640 or ca. 2780 (0.69 and 0.83 kb downstream from the *Bam*HI-2 site, respectively). Therefore, the total lengths of the respective two transcripts were calculated to be 1.63 and 1.77 kb. These lengths were compatible with those found by hybridization with the R1 strand described above. There seemed to be no splicing point(s) in the transcripts, because if splicing had occurred the alkaline electrophoresis would have resulted in the bands migrating faster than in neutral electrophoresis, each corresponding to the respective exon(s). The locations and orientations of these two transcripts coincided well with those of the region specifying the large open reading frame (968 to 2479) in phase 3 (Fig. 4).

When the l-strand of SL1 (L1) was used for hybridization with the mRNA, S1 digestion yielded one band of 0.85 kb in length on both neutral and alkaline agarose electrophoreses. Analysis with the l-strand of the *Bgl*II-to-*Bam*HI-1 fragment (L2), which carries additionally ca. 0.1 kb at the left end of SL1 but lacks the right half of SL1 from the *Bgl*II site, showed that the resulting band was about 0.1 kb longer than that with the L1 strand. These results indicated that the leftward transcript started at position ca. 850, proceeded through the *Sal*I-1 and *Bam*HI-1 sites, and was not spliced. The location and orientation of this transcript corresponded to the open reading frame extending beyond the *Sal*I-1 site (from 807 to 1 and beyond; phase 4 in Fig. 4). The deduced transcription map of the *RAD52* region is summarized in Fig. 7.

DISCUSSION

Coding region of the *RAD52* gene. The 3.3-kb *Sal*I fragment (SL1) isolated from *S. cerevisiae* Rad$^+$ DNA complemented the MMS sensitivity of the *rad52-1* strains, whereas the corresponding DNA segment from the *rad52-1* mutant DNA (RX1) did not. Schild and co-workers have independently isolated a DNA fragment which complements X-ray sensitivity of the *rad52-1* mutant and showed that a plasmid carrying this fragment integrated at the chromosomal *RAD52* locus (32). Their fragment and SL1 share a common restriction map. Analysis of the nucleotide sequence of SL1 showed the presence of a large open reading frame composed of 1,512 nucleotides (position 968 to 2479; Fig. 4). We believe that this open reading frame or at least a part of it specifies the *RAD52* protein, as deduced from the following results.

The results of MMS complementation with fragments subcloned from SL1 into YRp7 indicated that the *RAD52* gene includes the *Hpa*II, *Kpn*I-1, *Bgl*II, and *Bam*HI-2 sites. Cleavage at the *Hpa*II site seems to cause a defect in transcription initiation of the *RAD52* gene, because connecting the *S. cerevisiae* acid phosphatase promoter to this site led to the restoration of the *RAD52* activity to confer the MMS resistance. This large open reading frame, but none of the other frames, includes *Kpn*I-1 (position 1936), *Bgl*II (position 1470), and *Bam*HI-2 (position 1949) sites and is in very close (10 bp) proximity to the *Hpa*II (position 959) site. A DNA segment (RX1) cloned from the *rad52-1* mutant DNA contained a single base change, cytosine to thymine, at position 1236 (Fig. 3). This change leads to the amino acid substitution of alanine by valine in this frame, which is consistent with the observation that the *rad52-1* mutation is not nonsense suppressible (28). Our S1 mapping experiments showed that this open reading frame is transcribed and gives rise to two mRNAs with two different termination points and that these mRNAs are not spliced. All of the sequenced *S. cerevisiae* nuclear genes containing an intron are known to have the sequence $\frac{G}{T}GTA\frac{T}{C}GT$ at the 5′ border of each intron and the sequence TACTAAC within the introns, which are postulated to be involved in the selection of the splicing sites (17, 26). These sequences are not present in SL1. When the portable *lacZ* gene of *E. coli* of Casadaban was joined to YpAK28 at the *Bam*HI-2 site such that the above reading frame continued into the reading frame of the *lacZ* gene, the resulting plasmid expressed high β-galactosidase activity in *S. cerevisiae* as well as in *E. coli*. Therefore, this frame is translated in *S. cerevisiae* to produce a protein. In addition, this plasmid was found to lose the *RAD52* activity to confer MMS resistance, showing that the protein specified by this frame most probably has the *RAD52* activity.

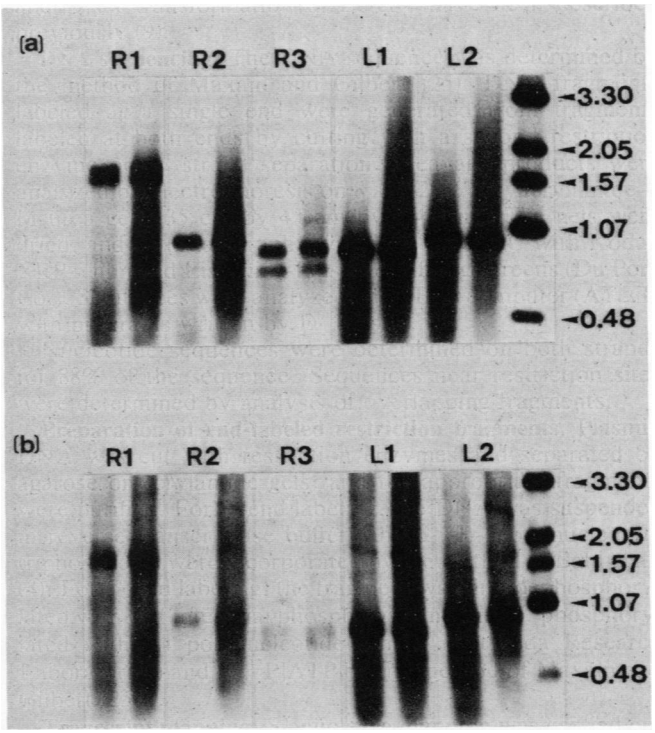

FIG. 6. S1 mapping autoradiogram. Polyadenylated RNA (30 μg) prepared from S288C was hybridized to 1.0 μg each of the indicated recombinant M13mp7 DNA, digested with S1 nuclease, fractionated on 1.0% neutral (a) or 1.0% alkaline (b) agarose gel electrophoresis, and subjected to Southern blotting analysis with ^{32}P-labeled SL1 fragment (10^7 cpm/μg) as probe. The amounts of S1 nuclease used here were 333 U for the left lane in each column and 111 U for the right lane. Single-stranded DNAs used as probes are indicated at the top of each column, using the same designations used in the text and in Fig. 5. Lengths of the bands of the DNA and the DNA-RNA hybrids were calibrated by relative mobilities to coelectrophoresed size marker DNAs consisting of 250 pg each of the restriction fragments in SL1 and AK28 (3.30, 2.05, 1.57, 1.07, and 0.48 kb in length). Autoradiography was performed at −80°C for 5 days with an intensifying screen.

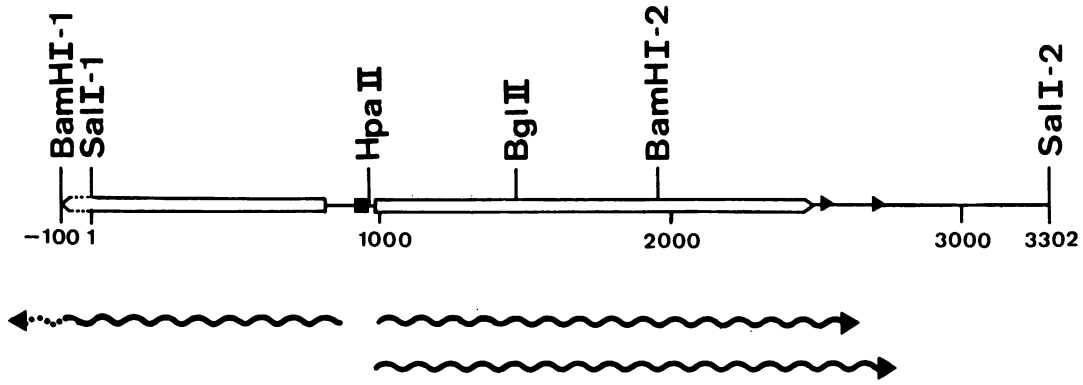

FIG. 7. Transcription map in the *RAD52* region. Locations and orientations of the transcripts verified from the results in Fig. 6 are shown by wavy lines. The largest and the next largest open reading frames are indicated by open arrows. Inverted repeat and direct repeat sequences found in the flanking regions of the *RAD52* gene are depicted by a closed box and by two triangles, respectively.

The results of the subcloning experiments and sequence determination of the *rad52-1* mutation, together with those of the fusion experiment with the acid phosphatase promoter, indicate that the initiation codon of the *RAD52* gene is most likely located within the region between *Hpa*II (position 959) and the *rad52-1* mutation site (position 1236). In this region, there are five ATG triplets (at positions 968, 1007, 1067, 1079, and 1085) and one GTG (position 1058) that are candidates for a translation initiation codon. Since there is no direct evidence to favor one initiation codon to the others, we tentatively assume that the structural gene starts at the first ATG (position 968) determined in this paper (Fig. 3). We are attempting to determine the precise location of the 5′ end of the *RAD52* mRNA. Preliminary results indicate that a major transcript starts 10 bp downstream from the second ATG triplet in the *RAD52* frame. This suggests that the third, fourth, or fifth ATG triplet is used as translation initiation codon of an authentic *RAD52* product, indicating that its molecular weight may not be larger than 52,400. Many, but not all, *S. cerevisiae* genes have been known to have a PuXXATGPuXT sequence at the translational start site (9). The fourth and fifth ATG triplets have the surrounding sequences ATTATGGAT (position 1076 to 1084) and GATATGGAT (position 1082 to 1090), respectively, which fit the canonical sequence among the above six candidates.

Chow and Resnick (6) have isolated DNase from *S. cerevisiae* by using an antiserum raised against a purified single-stranded DNA-binding endonuclease from *Neurospora crassa* whose activity was not found in the *rad52-1* mutant. The purified DNase has a molecular weight of 70,000. Since our sequencing of the *RAD52* gene indicated that the molecular weight of the *RAD52* protein should be smaller than 56,064 even if the translation starts at the first ATG triplet, this nuclease does not seem to be the *RAD52* protein.

5′- and 3′-flanking regions of the *RAD52* gene. According to present views, a sequence related to TATA$^{A}_{T}$A (TATA box) is usually encountered about 30 bp upstream from the transcription initiation point(s) of numerous eucaryotic genes and has been postulated to be involved in promoter recognition by eucaryotic RNA polymerase II (3). The TATA box is known to have highly conserved sequence, TATA, and to rarely contain guanine (3). In this regard, in the region preceding the probable transcription initiation point at position ca. 1010, a sequence TATATT (at position 900 to 905; Fig. 3) seems to most fit the canonical hexanucleotide. Another feature worth noting in this region is the presence of a 17-bp inverted repeat followed by a 12-bp guanine-plus-cytosine stretch (at position 899 to 937; Fig. 3). This inverted repeat might be involved in regulation of the expression of the *RAD52* gene or the adjacent gene or both. We should recall again that the elimination of the region upstream from the *Hpa*II site (at position 959) resulted in inactivation of the *RAD52* gene function, probably due to prevention of transcription initiation of the gene. Therefore, the sequence concerned with transcription initiation must be located upstream from the *Hpa*II site.

Our S1 mapping results indicated the possibility that transcription termination of the *RAD52* gene occurs at two different points. The 3′-flanking region of the *RAD52* structural gene contains a direct repeat AATAAAAAATT----ATTATTT at positions 2503 and 2640 (21 and 158 bp downstream from the termination codon TGA, respectively; Fig. 3). These sequences are followed by about 0.15 kb to the respective transcription termination points. Zaret and Sherman have suggested that the entire structure TAG-$X_{1\text{-}14}$-TA(T)GT-(A+T-rich)-TTT in the 3′-noncoding region of most *S. cerevisiae* genes may have a role in efficient transcription termination and polyadenylation (41). The 3′-flanking region of the *RAD52* structural gene seems to have a related structure, TAG-X_{13}-CATGT-(9/12 A+T)-TTT which appears from 111 bp downstream of the termination codon TGA (Fig. 3).

Complementation ability of the truncated *RAD52* gene in YpAK28. Plasmid YpAK28 carries the truncated *RAD52* gene instead of the complete one. It can, however, partially complement MMS sensitivity of the *rad52-1* mutants. Since all of the Trp$^+$ *rad52-1* mutants transformed by this plasmid exhibited MMS resistance and simultaneously lost both Trp$^+$ and MMS^r characters at a high frequency similar to that of parental vector YRp7, the partial MMS resistance shown by transformants carrying YpAK28 is probably conferred by the YpAK28 plasmid in its extrachromosomal state. Therefore, the truncated *RAD52* product whose C-terminal end is encoded from the vector DNA sequence presumably has adequate *RAD52* activity to complement MMS sensitivity of the *rad52-1* mutants. Taking into consideration the pBR322 sequence (25, 38), a short peptide comprising 17 amino acids is expected to be fused to the *RAD52* polypeptide truncated at the *Bam*HI-2 site. The

truncated *RAD52* product is probably not so active and requires more time to repair MMS damage completely. That is, when temporarily exposed to MMS in a liquid culture, even the truncated product can repair MMS damage, although it cannot overcome the successive production of damage by continuous exposure on an MMS plate.

In contrast to YpAK28, YpAK28-i does not confer MMS resistance. It is possibly because this plasmid encodes the truncated *RAD52* product fused with 87 amino acids, and the fusion of the 87 amino acids is more deleterious than the fusion of 17 amino acids in the product of YpAK28. The lack of complementation activity by the plasmid carrying the *RAD52-lacZ* fusion gene, in which the truncated *RAD52* polypeptide is fused with β-galactosidase (1,013 amino acids, from its eighth residue), may be explained similarly. Until the corresponding sequences are integrated as single-copy sequences into the chromosome, however, the other explanations, including copy numbers of the respective plasmids, are also possible.

ACKNOWLEDGMENTS

We thank S. Nakai for his generous gift of *S. cerevisiae* strains; T. Ohnishi and K. Yamamoto for help with the construction of *S. cerevisiae* gene library; and J. Tomizawa, Y. Sakakibara, R. Wickner, and I. Herskowitz for critical readings of the manuscript.

This work was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Science, and Culture of Japan and by a Naito Foundation research grant for 1983 to T. Ogawa.

LITERATURE CITED

1. **Beggs, J. D.** 1978. Transformation of yeast by a replicating hybrid plasmid. Nature (London) **275**:104–109.
2. **Berk, A., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**:1274–1278.
3. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
4. **Campbell, J. L., C. C. Richardson, and F. W. Studier.** 1978. Genetic recombination and complementation between bacteriophage T7 and cloned fragments of T7 DNA. Proc. Natl. Acad. Sci. U.S.A. **75**:2276–2280.
5. **Casadaban, M. J., A. Martinez-Arias, S. K. Sharina, and J. Chou.** 1983. β-Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. Methods Enzymol. **100**:293–308.
6. **Chow, T. Y.-K., and M. A. Resnick.** 1983. The identification of a deoxyribonuclease controlled by the *RAD52* gene of *Saccharomyces cerevisiae*, p. 447–455. *In* E C. Friedberg and B. A. Bridges (ed.), Cellular responses to DNA damage. Alan R. Liss, New York.
7. **Clewell, D. B., and D. R. Helinski.** 1969. Supercoiled circular DNA protein complex in *Escherichia coli*: induced conversion to open circular DNA form. Proc. Natl. Acad. Sci. U.S.A. **62**:1159–1166.
8. **Cryer, D. R., R. Eccleshall, and J. Marmur.** 1975. Isolation of yeast DNA. Methods Cell Biol. **12**:39–44.
9. **Dobson, M. J., M. F. Tuite, N. A. Roberts, A. J. Kingsman, and S. M. Kimgsman.** 1982. Conservation of high efficiency promoter sequences in *Saccharomyces cerevisiae*. Nucleic Acids Res. **10**:2625–2637.
10. **Dressler, D., and H. Potter.** 1982. Molecular mechanisms in genetic recombination. Annu. Rev. Biochem. **51**:727–761.
11. **Game, J. C., and R. K. Mortimer.** 1974. A genetic study of X-ray sensitive mutants in yeast. Mutation Res. **24**:281–292.
12. **Game, J. C., T. J. Zamb, R. J. Braun, M. Resnick, and R. M. Roth.** 1980. The role of radiation (*rad*) genes of meiotic recombination in yeast. Genetics **94**:51–68.
13. **Grunstein, M., and J. Wallis.** 1979. Colony hybridisation. Methods Enzymol. **68**:379–389.
14. **Haber, J. E., B. Weiffenbach, D. T. Rogers, J. McCusker, and J. B. Rowe.** 1981. Chromosomal rearrangements accompanying yeast mating-type switching: evidence for a gene conversion model. Cold Spring Harbor Symp. Quant. Biol. **45**:991–1002.
15. **Ho, K. S. Y.** 1975. Induction of DNA double strand breaks by X-rays in a radiation sensitive strain of the yeast *Saccharomyces cerevisiae*. Mutation Res. **30**:327–334.
16. **Kunz, B. A., and R. H. Haynes.** 1981. Phenomenology and genetic control of mitotic recombination in yeast. Annu. Rev. Genet. **15**:57–89.
17. **Langford, C. J., and D. Gallwitz.** 1983. Evidence for an intron contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell **33**:519–527.
18. **Malone, R., and R. E. Esposito.** 1980. The *RAD52* gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast. Proc. Natl. Acad. Sci. U.S.A. **77**:503–507.
19. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNa with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
20. **McDonell, M. W., M. M. Simon, and F. W. Studier.** 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. **110**:119–146.
21. **Miyanohara, A., A. Toh-e, C. Nozaki, F. Hamada, N. Ohtomo, and K. Matsubara.** 1983. Expression of hepatitis B surface antigen gene in yeast. Proc. Natl. Acad. Sci. U.S.A. **80**:1–5.
22. **Mortimer, R. K., and D. Schild.** 1980. Genetic map of *Saccharomyces cerevisiae*. Microbiol. Rev. **44**:519–571.
23. **Ogawa, H., K. Shimada, and J. Tomizawa.** 1968. Studies on radiation sensitive mutants of *Escherichia coli*. I. Mutants defective in the repair synthesis. Mol. Gen. Genet. **101**:227–244.
24. **Orr-Weaver, T. L., J. W. Szostak, and R. J. Rothstein.** 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. U.S.A. **78**:6354–6358.
25. **Peden, K. W. C.** 1983. Revised sequence of the tetracycline resistance gene of pBR322. Gene **22**:277–280.
26. **Pikiely, C. W., J. L. Teem, and M. Robash.** 1983. Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan mRNA splicing. Cell **34**:395–403.
27. **Prakash, S., L. Prakash, W. Burke, and B. A. Montelone.** 1980. Effects of the *RAD52* gene on recombination in *Saccharomyces cerevisiae*. Genetics **94**:31–50.
28. **Resnick, M. A.** 1969. Genetic control of radiation sensitivity in *Saccharomyces cerevisiae*. Genetics **62**:519–531.
29. **Resnick, M. A., J. N. Kasimos, J. C. Game, R. J. Braun, and R. M. Roth.** 1981. Changes in DNA during meiosis in a repair-deficient mutant (*rad52*) of yeast. Science **212**:543–545.
30. **Resnick, M. A., and P. Martin.** 1976. The repair of double-strand breaks in the nuclear DNA of *Saccharomyces cerevisiae* and its genetic control. Mol. Gen. Genet. **143**:119–129.
31. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
32. **Schild, D., B. Konforti, C. Perez, W. Gish, and R. K. Mortimer.** 1983. Isolation and characterization of yeast DNA repair genes. I. Cloning of the *RAD52* gene. Curr. Genet. **7**:85–92.
33. **Som, T., and J. Tomizawa.** 1982. Origin of replication of *Escherichia coli* plasmid RSF1030. Mol. Gen. Genet. **187**:375–383.
34. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
35. **Sripati, C. E., and J. R. Warner.** 1978. Isolation, characterization, and translation of mRNA from yeast. Methods Cell Biol. **20**:61–81.
36. **Stinchcomb, D. T., K. Struhl, and R. W. Davis.** 1979. Isolation and characterisation of a yeast chromosomal replicator. Nature (London) **282**:39–43.
37. **Struhl, K., D. T. Stinchcomb, S. Scherer, and R. W. Davis.** 1979. High frequency transformation of yeast: autonomous replica-

tion of hybrid DNA molecules. Proc. Natl. Acad. Sci. U.S.A. **76**:1035–1039.
38. **Sutcliffe, J. G.** 1978. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43**:77–90.
39. **Taylor, J. M., and T. P. H. Tse.** 1976. Isolation of rat liver albumin messenger RNA. J. Biol. Chem. **251**:7461–7467.
40. **Tschumper, G., and J. Carbon.** 1980. Sequence of yeast DNA fragment containing a chromosomal replicator and the *TRP1* gene. Gene **10**:157–166.
41. **Zaret, K. S., and F. Sherman.** 1982. DNA sequence required for efficient transcription termination in yeast. Cell **28**:563–573.