MOLECULAR AND CELLULAR BIOLOGY, Mar. 1987, p. 1300–1303
0270-7306/87/031300-04$02.00/0

Vol. 7, No. 3

Effect of Deletion and Insertion on Double-Strand-Break Repair in *Saccharomyces cerevisiae*

KEVIN STRUHL

Department of Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02115

Received 4 August 1986/Accepted 11 December 1986

I investigated double-strand-break repair in *Saccharomyces cerevisiae* cells by measuring the frequencies and types of integration events at the *PET56-HIS3-DED1* chromosomal region associated with the introduction of linearized plasmid DNAs containing homologous sequences. In general, the integration frequencies observed in strains containing a wild-type region, a 1-kilobase (kb) deletion, or a 5-kb insertion were similar, provided that the cleavage site in the plasmid DNA was present in the host genome. Cleavage at a plasmid DNA site corresponding to a region deleted in the chromosome caused a 10-fold reduction in the integration frequency even when the site was close to regions of homology. However, although the integration frequency was normal even when cleavage occurred only 25 base pairs (bp) outside the deletion breakpoint, 98% of the events were associated not with the usual heterogenote structure, but instead with a homogenote structure containing two copies of the deletion allele separated by vector sequences. Similarly, when cleavage occurred 80 bp outside the 5-kb substitution breakpoint, 40% of the integration events were associated with homogenote structures. From these observations, I suggest that (i) exonuclease and polymerase activities are not rate-limiting steps in double-strand-break repair, (ii) exonuclease activity is coupled to the initiation step, (iii) the integration frequency is strongly influenced by the amount of homology near the recombinogenic ends, (iv) both ends of a linear DNA molecule might interact with the host chomosome before significant exonuclease or polymerase action, and (v) the average repair tract is about 600 bp.

Transformation of *Saccharomyces cerevisiae* cells by recombinant plasmids containing cloned yeast DNA sequences is associated with autonomous replication or chromosomal integration of the introduced DNA (2–4, 13). If the plasmids are incapable of autonomous replication in yeast cells, transformation is inefficient because it depends on homologous recombination between plasmid and genomic sequences (4, 13). The recombination event usually results in the precise integration of a single copy of the entire plasmid DNA into the genome. This generates a heterogenote structure (6) in which the homologous sequences derived from the host and introduced DNA are separated by the plasmid vector sequences.

The transformation frequency can be increased by a factor of 100 to 1,000 if the plasmid DNA is cleaved within the homologous yeast sequences before being introduced into cells (3, 7). This dramatic stimulation has been attributed to the recombinogenic nature of the ends of the linear molecules, because the resulting transformants contain one or more integrated copies of the introduced DNA (7). Moreover, these integration events occur even when the plasmid DNA is cleaved at two positions to produce a double-strand gap. In this situation, the gap in the plasmid DNA is repaired with information from the host genome. These events depend on *RAD52*, and they are associated with roughly equal frequencies of gene conversion and reciprocal recombination near the site of the double-strand break (8). Thus, it has been suggested that transformation by linear plasmid DNAs provides a model for the repair of double-strand breaks (7) and that normal gene conversion may be initiated by such breaks (14).

A molecular model for the chromosomal integration of linear plasmid DNAs has been proposed (14). In this model, exonucleases enlarge the double-strand break into a gap flanked by 3′ single strands. Then, one of the two 3′ ends invades the host chromosome at the homologous site, thus displacing a D-loop, and the invading end is extended by DNA polymerase until the other 3′ end can anneal to the single-stranded region of the D-loop. The second 3′ end is extended to complete the repair of the gap, and branch migration results in the formation of two Holliday junctions. Finally, this structure is resolved to yield gene conversion or reciprocal recombination.

In this report, I examine several aspects of double-strand-break repair. The basic experiment is to create double-strand breaks or gaps at various positions within yeast DNA sequences present on recombinant plasmids and then to introduce these molecules into strains containing wild-type, deletion, or substitution alleles at the homologous chromosomal region. The relevant structures of the recombinant plasmids and host chromosomal regions are shown in Fig. 1. The recombinant plasmids are derived from YIp5-Sc2812, which contains a 6.1-kilobase (kb) yeast DNA fragment encoding the entire *PET56-HIS3-DED1* gene region (12) cloned into the *URA3* integrating vector YIp5 (13). YIp5-Sc3320 contains a mutated *Hin*dIII site at position +328 (with respect to the *his3* mRNA intitiation site) (12), whereas YIp5-Sc3309 contains a 400-base pair (bp) segment of the *GAL1,10* promoter region inserted between positions −173 and −137 (11). *S. cerevisiae* KY114 contains a wild-type *PET56-HIS3-DED1* chromosomal region (12); strain KY117 contains *his3-Δ200*, a deletion between −181 and +883 (12); and strain KY174 contains *his3-CYH2*, a 5-kb *CYH2* fragment inserted between an artificial *Eco*RI site at −35 and the *Bgl*II site at +442 (10). All of these strains contain the *ura3-52* mutation and hence are unable to grow in the absence of uracil.

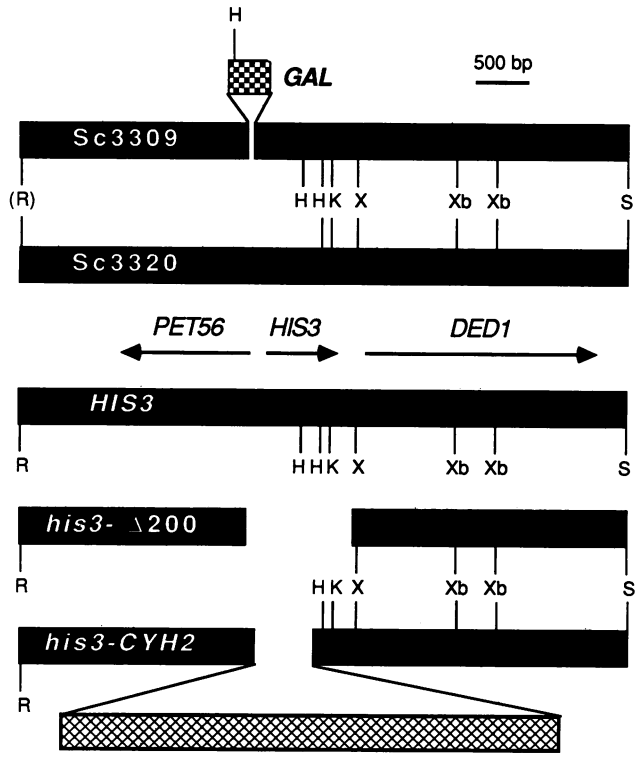

FIG. 1. Structures of transforming DNAs and host chromosomes. The black bar in the middle (marked *HIS3*) represents the 6.1-kb *Eco*RI-*Sal*I fragment from the *S. cerevisiae PET56-HIS3-DED1* chromosomal region from the wild-type strain KY114. The locations and orientations of the RNA transcripts are shown above the bar, and various restriction endonuclease cleavage sites are shown below the bar. Abbreviations: R, *Eco*RI; H, *Hin*dIII; K, *Kpn*I, X, *Xho*I; Xb, *Xba*I; S, *Sal*I. The two black bars below the wild type depict the structures of the *his3-Δ200* and *his3-CYH2* alleles that are contained in strains KY117 and KY174, respectively. The *his3-Δ200* allele contains a 1-kb deletion (blank space), and the *his3-CYH2* allele contains a 5-kb substitution that includes the entire *CYH2* gene ( ). The two black bars above the wild-type diagram the structures of the transforming DNAs (excluding the YIp5 vector sequences). Sc3309 substitutes approximately 400 bp of DNA (including 365 bp of the *GAL1,10* regulatory region, polylinker sequences from plasmids pUC8, and the 30-bp region between the *Eco*RI and *Hin*dIII sites of plasmid pBR322;) for a 40-bp region between the *HIS3* and *PET56* genes (11). The *Hin*dIII site represents the boundary between the *GAL* and pBR322 regions and is located about 30 bp from the boundary of the black bar (nucleotide −173); hence, it is located about 40 bp from the deletion endpoint of *his3-Δ200* (nucleotide −181). Sc3320 is simply a 4-bp insertion that destroys one of the two *Hin*dIII sites in the *HIS3* structural gene. Both Sc3309 and Sc3320 contain a mutated *Eco*RI site [indicated by (R)].

The plasmid DNAs were cleaved with various restriction endonucleases, and the resulting products were electrophoretically separated on a 0.7% low-temperature gelling-melting agarose gel (10). After melting of the appropriate gel slices at 73°C, the desired DNA segments were purified by phenol extraction and ethanol precipitation. Cultures from all three yeast strains were grown to early log phase (A_{600} = 1.0), and the cells were made competent for transformation by incubation in lithium chloride (5). Equal amounts of DNA (about 1 μg for each data point) were mixed with equal numbers of cells, and Ura$^+$ transformants were selected on plates lacking uracil. To minimize the variability in transformation efficiency, all of the individual reactions were performed in parallel on two occasions. The average number of Ura$^+$ transformants for each combination of linear plasmid DNA and host strain is shown in Table 1; the error is estimated to be ±30%.

Plasmid DNAs containing a simple double-strand break (experiments 1, 2, 9, 10, and 12 as listed in Table 1) produced about 40 Ura$^+$ colonies in all nine cases for which the site of cleavage was present in the host strain. These transformants are due to high-frequency integration of linear DNA because the number of colonies is only 10-fold below that obtained with an equal amount of circular YCp50, a DNA molecule capable of autonomous replication. Genomic hybridization analysis (data not shown) of 24 colonies from experiment 2A indicates that 22 colonies contain a single copy integrated at the *his3* locus; the remaining 2 colonies contained two tandem copies each. Multiple integration events occurred much less frequently than described previously (7), presumably because the transformation procedure avoids the necessity for making spheroplasts. Thus, the integration frequency is relatively unaffected by the position of the cleavage site or by the protruding 5′ or 3′ ends generated by the restriction endonuclease.

High-frequency integration was observed even when the cleavage site was located only 25 or 80 bp away from nonhomologous sequences (experiments 2B and 12C, respectively). In contrast, only four Ura$^+$ colonies were observed in experiment 1B, in which *Kpn*I cleaved YIp5-Sc3309 DNA was introduced into the *his3-Δ200* strain which deletes the *Kpn*I site. Thus, high-frequency integration requires homology between sequences at or near the double-strand break site and the host genome.

Linear molecules containing double-strand gaps ranging in size from 150 to 600 bp (experiments 3, 4, 11, and 13) yielded high integration frequencies in all seven cases in which both cleavage sites were present in the host genome. Similar frequencies were observed in the three "short-gapped" situations in which only one site was present in the genome (experiments 6A, 4B, and 7B). In contrast, the integration

TABLE 1. Integration frequences of plasmid DNA-yeast strain combinations

No. of expt.	Source of DNA	Enzyme	Avg no. of Ura$^+$ colonies obtained in strain containing:		
			HIS3 (A)[a]	*his3-Δ200* (B)[b]	*his3-CYH2* (C)
1	YIp5-Sc3309	*Kpn*I	44	4	
2	YIp5-Sc3309	*Xho*I	77	38	
3	YIp5-Sc3309	*Xba*I	30	59	
4	YIp5-Sc3309	*Kpn*I + *Xho*I	78	34	
5	YIp5-Sc3309	*Kpn*I + *Xba*I	29	4	
6	YIp5-Sc3309	*Hin*dIII	35	6	
7	YIp5-Sc3309	*Hin*dIII + *Xho*I	6	45	
8	YIp5-Sc3309	*Hin*dIII + *Xba*I	6	13	
9	YIp5-Sc3320	*Kpn*I	50		39
10	YIp5-Sc3320	*Xho*I	45		49
11	YIp5-Sc3320	*Xba*I	33		75
12	YIp5-Sc3320	*Hin*dIII	76		49
13	YIp5-Sc3320	*Hin*dIII + *Xho*I	67		30
14	YIp5-Sc3320	*Hin*dIII + *Xba*I	14		9
15	YCp50	Uncut	400	375	550

[a] Wild type.

[b] For column B, all Ura$^+$ colonies were His$^-$ with the following exceptions: experiment 2B, 4 His$^+$ colonies of 204 examined; experiment 3B, 49 His$^+$ colonies of 59 examined.

frequency was reduced by a factor of 5 to 10 in experiment 6B, a situation in which both cleavage sites were absent in the host. Thus, high-frequency integration occurs, provided that at least one of the two ends is homologous to the genome.

In experiment 5A, the DNA contained a 1.6-kb gap with respect to the wild-type genome, yet the integration frequency was similar to or only slightly lower than the normal level. However, such large gaps can significantly lower the integration frequency in situations in which one of the two ends either is not homologous to the genome or maps close to a region of nonhomology. For example, a 700-bp gap with one homologous end produced 35 Ura^+ colonies (experiment 6A), whereas only 6 Ura^+ colonies were observed when the gap is increased to 1.1 (experiment 7A) or 2.6 kb (experiment 8A). Similarly, direct comparisons between experiments 4B and 5B, 7A and 7B, 7B and 8B, or 11C and 14C show that the integration frequency decreased about four- to eightfold when the gap is enlarged by 1 kb.

As mentioned above, the integration frequency was high even when the transforming DNA was cleaved at a position mapping only 25 bp from a deletion breakpoint (line 2B). However, all 38 of the transformants listed in Table 1 were His^-, even though the transforming DNA contained a functional *his3* allele. Analysis of 204 more transformants revealed that 200 (98%) were His^-, whereas only 4 (2%) were His^+. Genomic hybridization analysis (data not shown) indicated that all 16 colonies examined did not contain the usual heterogenote structure. Instead, 15 of these strains had a homogenote structure containing two copies of the *his3-Δ200* allele separated by a single copy of the Ura^+ vector sequence; the exceptional case was probably the result of multiple integration events. Equivalent homogenotes were also observed in experiment 3B, in which the transforming DNA was cleaved to generate a 600-bp gap with endpoints located approximately 0.9 and 1.5 kb away from the *his3-Δ200* breakpoint. Of 59 colonies examined, 49 (83%) were His^+ heterogenotes and 10 (17%) were His^- homogenotes.

To account for the homogenote structures, it appears that 1.3 kb of DNA containing the entire *HIS3* structural gene and the *GAL* segment of YIp5-Sc3309 have been degraded by exonucleases and possibly also recombined out by a reciprocal exchange. The high integration frequency suggests that this extensive activity is not the rate-limiting step for chromosomal integration. Moreover, this extensive degradation is not an artifact associated with the introduction of purified DNA into yeast cells. If this were the case, high integration frequencies would be observed even when DNA was cleaved at sites corresponding to sequences deleted in *his3-Δ200*, because exonucleolytic action would generate recombinogenic 3′ ends that are homologous to the host genome. However, the low integration frequencies in experiment 1B and 6B are inconsistent with this possibility. Moreover, the low integration frequency in experiment 6B, in which the *Hin*dIII site upstream of the *his3* gene in YIp5-Sc3309 is located only 40 bp from homologous sequences in *his3-Δ200*, suggests that newly introduced DNA is subject to minimal exonuclease activity. This is supported by the observation that cohesive ends generated by restriction endonuclease treatment are readily ligated upon introduction into yeast cells (9). Thus, these observations strongly suggest that exonuclease activity is coupled to the initiation event.

These observations also suggest that the integration frequency is strongly influenced by the amount of homology near the recombinogenic ends. In principle, both ends of a linear molecule can initiate the repair process if they are homologous to chromsomal sequences. In experiment 2B, both *Xho*I ends were homologous to the genome but the upstream end contained only 25 bp of homology before the *Δ200* interruption whereas the downstream end contained 2.5 kb of uninterrupted homology. Assuming that an invading 3′ end does not suffer further exonuclease action once it invades a duplex, one would expect His^+ colonies if the *his3-Δ200* chromosome were to be invaded by the upstream end. Nevertheless, only 4% of the colonies were His^+, far below the 50% that would be expected if both ends were equally capable of initiating the recombination event.

Homogenote structures were also observed when the transforming DNA was cleaved only 80 bp outside the point of a 5-kb insertion. Genomic hybridization analysis (data not shown) of 15 colonies from experiment 12C and 8 colonies from experiment 13C indicate that only 60% are associated with the normal heterogenote structure. The remaining 40% are homogenotes that contain two copies of the *his3-CYH2* allele. In these cases, the entire 5-kb *CYH2* substitution was copied during the integration. The relatively high frequency of these events strongly suggests that extensive DNA polymerase activity can occur during double-strand-break repair and that this polymerization is not the rate-limiting step in the process.

The generation of homogenotes containing two copies of the *his3-CYH2* allele seems to involve the repair of a 5-kb gap in plasmid DNA with chromosomal information. However, the frequency of this event was considerably higher than that observed with much smaller gaps. Once possible explanation for this apparent anomaly is that the large-gap effect occurs only when the recombinogenic ends are homologous to sequences that are far apart on the host genome. In this view, *his3-CYH2* homogenotes from experiment 12C and 13C would be generated easily because, although a large gap must ultimately be repaired, the original ends of the plasmid DNAs could interact with genomic sequences that are close together. In contrast, the low integration frequency in experiment 14C would be due to a true 1.8-kb gap between the *Hin*dIII and *Xba*I sites. One implication of this idea is that both ends of the plasmid DNA would interact with the host chromosome at a stage of the repair process before extensive exonuclease or polymerase action.

On the assumption that the lengths of repair tracts fit a Poisson distribution, two results in this report suggest that the average repair tract is approximately 600 bp. First, the heterogenotes in experiment 2B represent repair tracts less then 25 bp in length, because larger tracts can be resolved only by homogenote formation. Assuming an average length of 600 bp, the predicted frequency of repair tracts less than 25 bp long is 4%, a value in excellent accord with that observed (2%). Second, the homogenotes in experiment 3B represent repair tracts greater than 1,500 bp in length; the observed value (17%) agrees well with the predicted value (10%). An average repair tract of 600 bp may explain the decrease in integration frequency that is observed in large-gapped molecules. Other workers (1; T. L. Orr-Weaver, Ph.D. thesis, Harvard University, Boston, Mass., 1983) have determined a similar average repair length by measuring the conversion frequency of heteroalleles during double-strand-break repair involving complete homology between the transforming DNA and host chromosome. However, although the average repair tract appears to be about 600 bp, more extensive repair can occur in certain circumstances without significant reduction in the integration frequency.

I thank Andrew Murray, Terry Orr-Weaver, and Jack Szostak for fruitful discussion.

This work was supported by Public Health Service grant GM 30186 from the National Institutes of Health.

LITERATURE CITED

1. **Ahn, B.-Y., and D. M. Livingston.** 1986. Mitotic gene conversion lengths, coconversion patterns, and the incidence of reciprocal recombination in a *Saccharomyces cerevisiae* plasmid system. Mol. Cell. Biol. **6**:3685–3693.
2. **Beggs, J.** 1978. Transformation of yeast by a replicating hybrid plasmid. Nature (London) **275**:104–109.
3. **Hicks, J. B., A. Hinnen, and G. R. Fink.** 1979. Properties of yeast transformation. Cold Spring Harbor Symp. Quant. Biol. **43**:1305–1313.
4. **Hinnen, A., J. B. Hicks, and G. R. Fink.** 1978. Transformation of yeast. Proc. Natl. Acad. Sci. USA **75**:1929–1933.
5. **Ito, H., Y. Fukuda, K. Murata, and A. Kimura.** 1983. Transformation of intact yeast cells treated with alkali cations. J. Bacteriol. **153**:163–168.
6. **Morse, M. L., E. M. Lederberg, and J. Lederberg.** 1956. Transductional heterogenotes in *Escherichia coli*. Genetics **41**:758–775.
7. **Orr-Weaver, T. L., R. J. Rothstein, and J. W. Szostak.** 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA **78**:6354–6358.
8. **Orr-Weaver, T. L., and J. W. Szostak.** 1983. Yeast recombination: the association between double-strand gap repair and crossing over. Proc. Natl. Acad. Sci. USA **80**:4417–4421.
9. **Stinchcomb, D. T., K. Struhl, and R. W. Davis.** 1979. Isolation and characterization of a yeast chromosomal replicator. Nature (London) **282**:39–43.
10. **Struhl, K.** 1983. Direct selection for gene replacement events in yeast. Gene **26**:231–241.
11. **Struhl, K.** 1984. Genetic properties and chromatin structure of the yeast *gal* regulatory element: an enhancer-like sequence. Proc. Natl. Acad. Sci. USA **81**:7865–7869.
12. **Struhl, K.** 1985. Nucleotide sequence and transcriptional mapping of the yeast *pet56-his3-ded1* gene region. Nucleic Acids Res. **13**:8587–8601.
13. **Struhl, K., D. T. Stinchcomb, S. Scherer, and R. W. Davis.** 1979. High frequency transformation of yeast: autonomous replication of hybrid DNA molecules. Proc. Natl. Acad. Sci. USA **76**:1035–1039.
14. **Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. Stahl.** 1983. The double-strand-break repair model for recombination. Cell **33**:25–35.