Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 6851–6856, June 1997
Genetics

Gene targeting by linear duplex DNA frequently occurs by assimilation of a single strand that is subject to preferential mismatch correction

W.-Y. LEUNG, A. MALKOVA, AND J. E. HABER*

Rosenstiel Basic Medical Sciences Research Center and Department of Biology, Brandeis University, Waltham, MA 02254-9110

Communicated by Charles M. Radding, Yale University School of Medicine, New Haven, CT, April 15, 1997 (received for review December 19, 1996)

ABSTRACT **To study targeted recombination, a single linear 2-kb fragment of *LEU2* DNA was liberated from a chromosomal site within the nucleus of *Saccharomyces cerevisiae*, by expression of the site-specific HO endonuclease. Gene targeting was scored by gene conversion of a chromosomal *leu2* mutant allele by the liberated *LEU2* fragment. This occurred at a frequency of only 2×10^{-4}, despite the fact that nearly all cells successfully repaired, by single-strand annealing, the chromosome break created by liberating the fragment. The frequency of Leu^+ recombinants was 6- to 25-fold higher in *pms1* strains lacking mismatch repair. In 70% of these cases, the colony was sectored for Leu^+/Leu^-. Similar results were obtained when a 4.1-kb fragment containing adjacent *LEU2* and *ADE1* genes was liberated, to convert adjacent *leu2* and *ade1* mutations on the chromosome. These results suggest that a linear fragment is not assimilated into the recipient chromosome by two crossovers each close to the end of the fragment; rather, heteroduplex DNA between the fragment and the chromosome is apparently formed over the entire region, by the assimilation of one of the two strands of the linear duplex DNA. Moreover, the recovery of Leu^+ transformants is frequently defeated by the cell's mismatch repair machinery; more than 85% of mismatches in heteroduplex DNA are corrected in favor of the resident, unbroken (mutant) strand.**

Despite the fact that *Saccharomyces* is celebrated for its ability to carry out homologous recombination, the frequency of gene replacement by transformation of a linear exogenous DNA is surprisingly low (1–3). This could be due to several factors. For example, "naked" DNA, not yet associated with histones or other nuclear proteins, might be rapidly degraded before it has a chance to undergo recombination. Alternatively, the replacement of homologous DNA sequences may be an inherently inefficient process. Previously we showed that the frequency of gene replacement can be increased by providing additional copies of the target sequence (3), a result that suggests that the search for homology is a rate-limiting step in the process. In these and previous experiments it was also not known if there was a special subpopulation of cells that were especially adept at this process, either because they were more proficient in recombination or because they were able to take up more copies of the transforming DNA.

To avoid some of the uncertainties inherent in transformation, we have devised a way to liberate a single linearized fragment of DNA from a chromosome within the nucleus and to examine its capacity to be "captured" by homologous recombination. The chromosomal region illustrated in Fig. 1*A* contains a *LEU2* gene flanked by HO endonuclease recognition sites. When a galactose-inducible HO endonuclease is expressed, cleavage occurs at these sites, liberating a 2.0-kb *LEU2* fragment and leaving a broken chromosome. The chromosomal break itself can be very efficiently repaired by the process of single-strand annealing (4–6), in which extensive 5′-to-3′ exonuclease digestion exposes complementary regions of partially duplicated *HIS4* genes and produces a deletion restoring *HIS4* function. During this process, any DNA repair functions that are induced by the appearance of a double-strand break (DSB) should be synthesized. We then demand that the liberated *LEU2* fragment replace a *leu2* mutation located elsewhere on the same chromosome, to produce a His^+ Leu^+ cell. Surprisingly, this process is no more frequent than when DNA is introduced by conventional transformation. We demonstrate that one significant reason for this inefficiency is the biased action of the cell's mismatch repair machinery, to remove the invading *LEU2* allele in favor of the resident *leu2* mutant allele. This result prompts us to suggest that gene replacements, at least those involving the replacement of small heterologies or mismatches, occur by the formation of extensive heteroduplex DNA across the target region rather than by crossings-over limited to the ends of the fragment.

MATERIALS AND METHODS

Strains. The relevant genotypes of strains G304 and its derivatives are shown in Table 1. Strain G304 (*MATα-inc ade1 leu2–3,112 ura3–52 trp1*) has been described previously (6). The structure of the *his4*Δ*–URA3*–cs::*LEU2*::cs–Δ*his4* construct (6) is shown in Fig. 1*A*. Strain G324 lacks the *GAL*::*HO*-containing *TRP1 CEN4* plasmid pFH800 (7). Strains G304, G366, G487, G491, G523, G543, AM48, and AM62 all carry the plasmid pFH800. Strain G366 is identical to G304 except that it has undergone a deletion of the *PMS1* gene (8). Strains G487 and G491 are wild type and *pms1*Δ strains, isogenic with G304, but carry a 4-bp fill-in of the *Asp*718 site in *leu2*. Strains G523 and G543 are wild type and *pms1*Δ derivatives, isogenic to strain G304, but carrying an *ADE1* gene inserted into the *Ase*I site adjacent to *LEU2* (which is surrounded by HO cleavage sites) and having the same *ADE1* gene inserted at the chromosomal *leu2–3,112* locus (Fig. 1*B*). In the latter case, the *ADE1* gene was then mutated to *ade1-E* by the integration and excision of plasmid pRHB134 carrying the *ade1-E* mutation (a 4-bp fill-in mutation of the *Esp*I site that creates a *Pst*I site) on a YIp5 (*URA3*-containing) plasmid (Fig. 1*B*). Strain AM48 is similar to strain G304, but in a different genetic background, and exhibits 3-fold higher *LEU2* capture. Strain AM62 is identical to AM48 except that it contains a *leu2-K* sequence flanked at each end by half of the HO endonuclease cleavage site, all integrated on chromosome I at the *ade1* locus. Here, the HO-liberated *LEU2* fragment has perfect homology to the ends of the target sequence.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*To whom reprint requests should be addressed. e-mail: haber@hydra.rose.brandeis.edu.

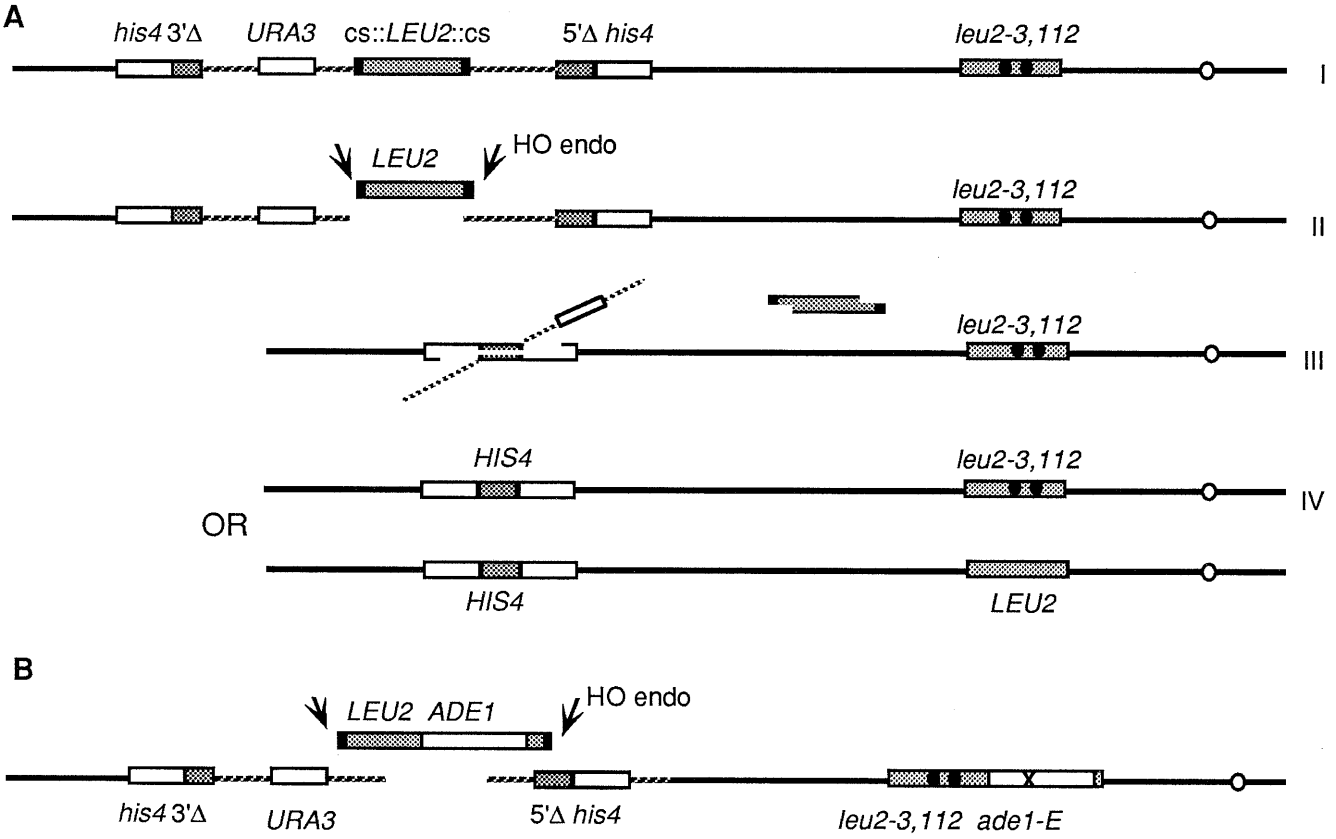

FIG. 1. Liberation of a *LEU2* fragment and subsequent repair of the broken chromosome. (*A*) pJH825 is pBR322 plasmid containing *URA3* inserted at the *Hin*dIII site, *LEU2* surrounded by two 117-bp HO endonuclease recognition sites (cs), all inserted at the *Bam*HI site, and a 1-kb internal fragment of the *HIS4* gene inserted at the *Eco*RI site. When this plasmid is integrated at *HIS4*, it creates a His^- duplication of part of the *HIS4* gene (I). Expression of the HO endonuclease from a galactose-inducible promoter liberates the *LEU2* fragment (II). The broken chromosome III is repaired by single-strand annealing to produce a His^+ recombinant (III). The liberated *LEU2* fragment may recombine with the *leu2–3,112* allele at its normal chromosomal location approximately 22 kb away from *HIS4* (IV). (*B*) Liberation and capture of a *LEU2* fragment containing a 1.9-kb insertion of the *ADE1* gene. Integration of this entire fragment into a wild-type *LEU2* target produces a Leu^+ Ade^+ recombinant.

The *leu2–3,112* double mutation contains two 1-bp insertions, each an extra G in a run of 4 or 3 Gs, respectively at positions 907 and 1449 of the 2.23-kb *LEU2* region sequence (9) (J. K. Moore, L. Berenfeld, and J.E.H., unpublished work).

Growth and Induction of HO Endonuclease. Cells were grown at 30°C. Prior to HO induction, by the addition of 2% galactose to the medium, cells containing the *GAL*::*HO* gene on a *TRP1* plasmid pFH800 (7) were grown on synthetic medium minus tryptophan to stationary phase and then grown in yeast extract/peptone (YEP)-lactate, as described previously (4–6). HO induction was terminated by washing cells with sterile water and plating them on appropriate nutritional plates.

Transformation Conditions. Cotransformation of approximately 3 μg of gel-purified *Hpa*I–*Sal*I fragment of *LEU2* and 3 μg of circular *URA3*-containing plasmid YCp50, with 2.8×10^8 cells in 1 ml, was carried out according to the method of Chen *et al.* (10).

DNA Analysis. Southern blot analysis was carried out as described previously (4–6).

RESULTS

Efficient Single-Strand Annealing Is Accompanied by Inefficient Gene Conversion. Strain G304 carries the *his4*Δ–*URA3*–cs::*LEU2*::cs–Δ*his4* construct illustrated in Fig. 1*A*. The strain carries a galactose-inducible HO gene (7) that cleaves the two HO cleavage sites (cs) flanking *LEU2*. After a 2- to 3-hr induction of HO (see *Materials and Methods*), at least 75% of cells became His^+ (Table 1, part A). Formation of deletions by single-strand annealing can be followed on Southern blots of DNA taken at intervals after HO induction (Fig. 2). HO cleavage is highly efficient, so that most of the cells yield the expected liberated 2.0-kb *LEU2* fragment. There are only very faint bands of 7.4 and 5.0 kb that are the products of HO cleavage at only one of the two sites (these bands are marked by dots in the 0.5-hr lane of Fig. 2). The *LEU2* fragment is not rapidly degraded, as it is easily seen even 2 hr after it is liberated. It disappears with about the same kinetics as the cut ends of the chromosomal DNA that are engaged in single-strand annealing. Fig. 2*B* shows the efficient repair of the broken chromosome, which occurs even though the ends of the two chromosomal segments that reanneal are initially separated by 2 kb.

Despite efficient reannealing/recombination of the 0.9-kb *his4* segments flanking the double-strand break, the liberated *LEU2* fragment recombines poorly with its *leu2–3,112* target sequence located about 22 kb away on the same chromosome (Table 1, part A). The *LEU2* and *leu2–3,112* sequences are identical except for the two frameshift mutations. Among His^+ recombinants in strain G304, only 0.012% of the cells were also Leu^+. The low frequency of Leu^+ recombinants is found despite the fact that any DNA damage-inducible genes necessary for the recombinational repair of the chromosomal double-strand break must have been induced. The frequency of HO-stimulated events (9.1×10^{-5}) represents an approximately 250-fold increase over the spontaneous frequency of His^+ Leu^+ events (3.6×10^{-7}).

By crossing with appropriate tester strains, we also analyzed 80 His^+ Leu^- derivatives to determine if perhaps one of the

Table 1. Gene conversion of *leu2* alleles by a *LEU2* fragment liberated in the nucleus or introduced by transformation

Strain	Genotype difference	His^+ total, %	His^+ Leu^+/His^+, %
A. HO-induced gene conversion of *leu2-3,112* by *LEU2*			
G324	No GAL::HO plasmid	0.06 ± 0.04	0.06 ± 0.05
G304		75.8 ± 11.6	0.012 ± 0.004
G366	*pms1*Δ	100 ± 11.6	0.07 ± 0.008
B. HO-induced gene conversion of *leu2-A* by *LEU2*			
G487		100 ± 1	0.017 ± 0.002
G491	*pms1*Δ	87.5 ± 12	0.42 ± 0.042
C. HO-induced gene conversion of *leu2-3,112*::*ade1-E* by *LEU2*::*ADE1*			
G523		89.6 ± 10.1	0.036 ± 0.003
G543	*pms1*Δ	97.7 ± 14.8	0.25 ± 0.047
D. HO-induced gene conversion of *leu2-K* with homologous DNA ends			
AM48		70.8 ± 16.5	0.030 ± 0.02
AM62	Plus hcs::*leu2-K*::*hcs** inserted at *ADE1*	88.4 ± 11.5	0.047 ± 0.02
E. Transformation of linear *LEU2* DNA and circular *URA3*-marked plasmid			Leu^+/Ura^+, %
G358			0.076
G364	*pms1*Δ		0.26

The relevant genotype of strains G304 and its derivatives is shown in Fig. 1A and are more fully described in Materials and Methods. *hcs = half of the HO endonuclease cleavage site (cs) homologous to the end of the HO-liberated *LEU2* fragment.

two *leu2–3* or *leu2–112* alleles had been converted more frequently. In all 80 cases, both *leu2* alleles were still present, suggesting that there was not a special problem in gene converting one of the two sites. As shown below, similar frequencies of His^+ Leu^+ recombinants were also obtained when the liberated *LEU2* fragment recombined with another allele, *leu2-K*.

The frequency of *LEU2* gene conversion after HO induction is very similar to that obtained in a cotransformation experiment. Approximately 3 μg of an analogous linearized gel-purified *Bam*HI fragment, containing the *Hpa*I–*Sal*I *LEU2* fragment bounded by 117-bp HO cut site segments, was transformed into yeast along with an equivalent concentration of circular *URA3*-marked centromeric plasmid (Table 1, part E). The frequency of Leu^+ Ura^+ colonies, compared with Ura^+ colonies, was 7.6×10^{-4}. Thus, a principal limitation in transformation appears to be the "capture" of a fragment of DNA by gene conversion rather than any problems of getting transforming DNA to the nucleus.

The limitation on *LEU2* capture is not attributable to the approximately 60 bp of HO cut site sequences at the ends of the fragment that are not homologous to the target. Strain AM62 was created carrying both *leu2–3,112* and a *leu2-K* mutant flanked by sequences perfectly homologous to ends of the HO-liberated *LEU2* fragment, integrated at *ADE1*. When this strain was induced by addition of galactose, *LEU2* capture was only about twice that of strain AM48, with only the *leu2–3,112* target (Table 1, part C), as expected if there were two equally attractive targets (3). Southern blot analysis showed that both sites were used equally (data not shown).

Deletion of the Mismatch Repair Gene *PMS1* Improves Gene Replacement. The apparent failure of *LEU2* to recombine could be explained if the *LEU2* fragment formed heteroduplex DNA with its target, but the mismatch correction system repaired the heteroduplex DNA in favor of the resident, mutant, information. Thus a recombinational intermediate would be formed, but cells would not become Leu^+. Both bacterial and vertebrate mismatch correction systems preferentially correct a mismatch on the same strand as a nick or gap such as would be found in recombination intermediates (11–

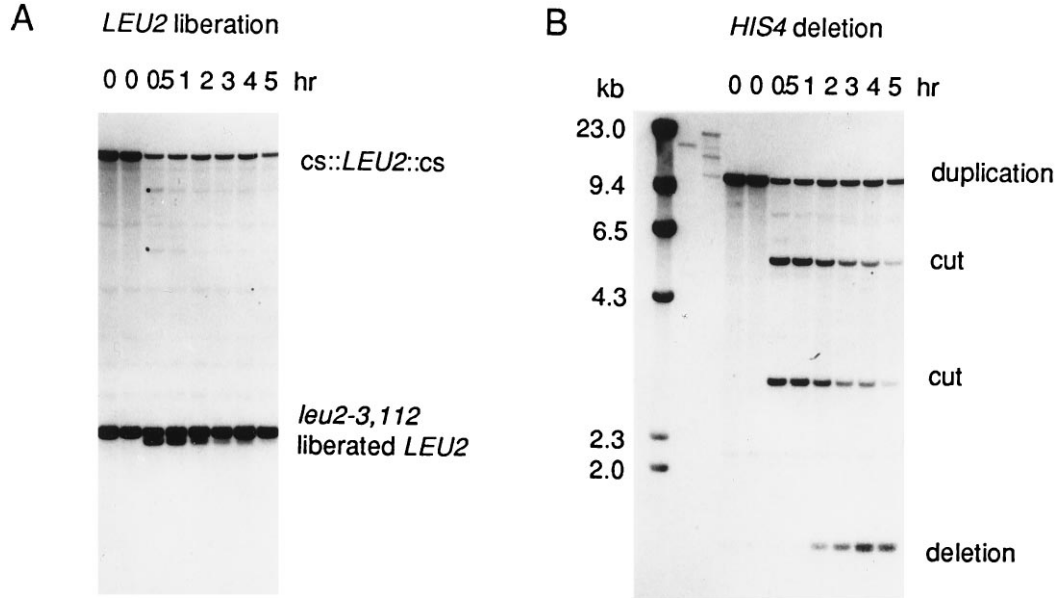

FIG. 2. Physical monitoring of the liberation of a *LEU2* fragment and the repair of the broken chromosome by single-strand annealing. Strain G304, carrying the *his4′–URA3*–cs::*LEU2*::cs–*′his4* insertion illustrated in Fig. 1 and the *TRP1*, *GAL1*::*HO* centromere plasmid pFH800 was grown in YEP-lactate to which 2% galactose (final concentration) was added at $t = 0$ hr. DNA was extracted at intervals, purified, and cleaved with restriction enzymes for Southern blot analysis. (*A*) The liberation of the *LEU2* fragment is seen on a *Hpa*I–*Sna*BI digest of DNA, probed with a *LEU2*-specific probe. Virtually all of the DNA that is cleaved by HO endonuclease is cut at both flanking cut sites to produce a 2.1-kb fragment. Two very faint bands of 7.4 kb and 5.0 kb, marked by dots in the 0.5-hr lane, are the restriction fragments expected for cleavage at only one site. The liberated *LEU2* fragment is visible for more than 2 hr before it is apparently degraded. (*B*) The HO endonuclease-cleaved chromosomal DNA efficiently recombines to form a His^+ recombinant. *Hpa*I–*Sna*BI-digested DNA, probed with a *HIS4*-specific probe, initially reveals a 10.2-kb restriction fragment containing the entire *his4′–URA3*–cs::*LEU2*::cs–*′his4* region. HO endonuclease cleavage produces two fragments of 5.3 and 2.9 kb, respectively. After approximately 1 hr, the final His^+ recombinant band of 1.4 kb is visible.

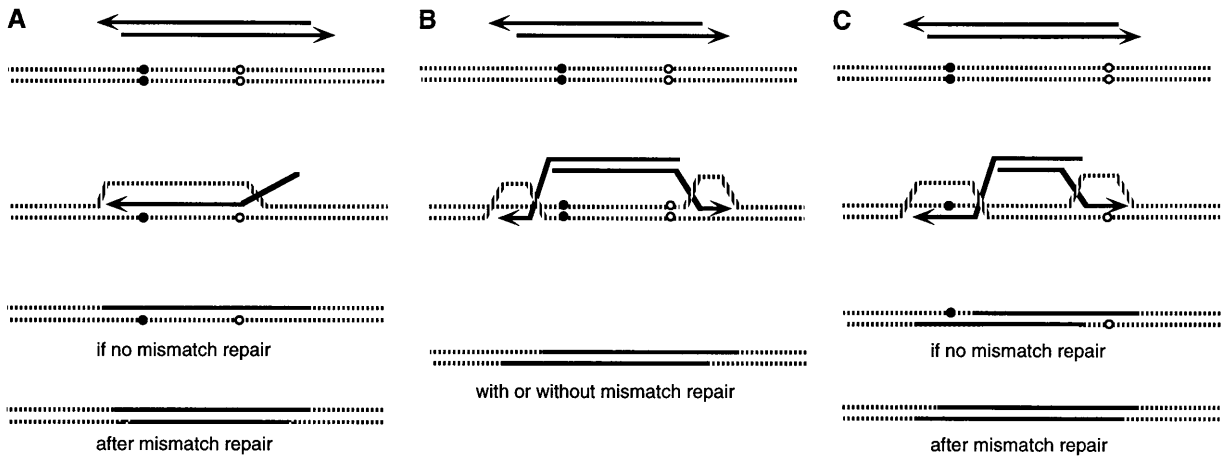

FIG. 3. Alternative mechanisms to incorporate transforming DNA into the chromosome. (*A*) Assimilation of a single strand of DNA to form heteroduplex DNA over most of its length. As shown in successive vertical drawings, heteroduplex DNA can then be mismatch corrected to yield a wild-type transformant. If mismatch correction is blocked, a sectored colony should be produced, with markers at opposite ends of the transforming fragment assimilated into one strand. Prior steps of forming and removing a displacement loop (D-loop) involving the complementary resident strand are not shown. (*B*) Assimilation of a double-stranded fragment of DNA by two crossover events confined to the ends of the recombining fragment. In this case, the central part of the region is replaced with two strands of donor DNA and mismatch correction plays no role in the process. (*C*) A hybrid model in which there are longer regions of heteroduplex DNA associated with two crossing-over events. In this case the markers at opposite ends of the transforming DNA will form heteroduplex on two different DNA strands of the recipient.

13). Such biased mismatch repair has also been documented during mating-type gene (*MAT*) switching in yeast (6, 14), where a deletion of the mismatch repair gene *PMS1* led to sectored (mutant/wild type) colonies, similar to postmeiotic segregation events in meiotic cells (15). The *pms1* mutant does not completely prevent mismatch repair of 1- to 4-bp insertions or base pair substitutions; about 15% of repair events are still observed (6, 16, 17).

To test if mismatch repair was a major factor in reducing *LEU2* capture, we measured the frequency of His^+ Leu^+ recombinants in a strain where mismatch correction was prevented by a *pms1* deletion mutation (strain G366). The *leu2–3,112* double mutation contains two 1-bp insertions separated by 542 bp, lying 666 and 781 bp from the left and right ends of the 2.0-kb liberated *LEU2* fragment. The absence of mismatch repair increased capture of the *LEU2* fragment approximately 6-fold (Table 1, part A). Thus, preventing mismatch repair of heteroduplex DNA formed between the *LEU2* fragment and the resident *leu2–3,112* locus leads to an increased recovery of at least one DNA strand that is Leu^+. This conclusion is supported by a direct analysis of His^+ colonies arising after HO induction. Among 11,123 His^+ *pms1* colonies plated at a density of approximately 60 colonies per plate, there were 11 that were also Leu^+ (0.1%, in good agreement with the results in Table 1, part A). Eight of these 11 colonies were sectored (Leu^+/Leu^-), as expected if most Leu^+ recombinants escaped mismatch correction of heteroduplex DNA (*LEU2/leu2–3,112*). There were no sectored colonies among an equivalent sample of His^+ colonies from the *PMS1* strain G304.

Similar results were also obtained by liberating the *LEU2* fragment by HO endonuclease in strains isogenic to G304 in which *leu2–3,112* was replaced by a 4-bp insertion mutation, *leu2-A*, resulting from filling-in the *Asp*718 site of *LEU2* (Table 1, part B). Here, the frequency of Leu^+ transformants increased 25-fold in a *pms1*Δ derivative.

The effect of deleting *PMS1* is also seen with conventional DNA transformation. When a *LEU2* fragment and a *URA3*-marked plasmid were cotransformed, there was a 3.6-fold increase in Leu^+ Ura^+ transformants in the *pms1*Δ *leu2–3,112* strain compared with wild type (Table 1, part E).

In at least 70% of all events recovered from the *pms1*Δ strain, both the *leu2–3* and *leu2–112* alleles were apparently included in heteroduplex DNA, as the His^+ colonies were sectored for Leu^+/Leu^-. These two alleles lie 666 and 781 bp from the left and right ends of the 2.0-kb *LEU2* fragment. Thus, the central 542 bp of the recombining *LEU2* fragment appears to have been in heteroduplex DNA. This suggests that so-called "ends-out" recombination† does not occur primarily by two local crossover events, each confined to a few hundred base pairs at the end of the fragment (Fig. 3*B*), because such a process would swap Leu^+ double-stranded DNA for *leu2–3,112* DNA so that there would be no heteroduplex DNA involving these markers. To examine this point further, we constructed a strain carrying a 4.1-kb fragment bounded by HO cut sites, carrying both the *LEU2* and *ADE1* genes, and in which the chromosomal target region contained the same two genes, but with *leu2–3,112* and *ade1-E* (Fig. 1*B*). As before, HO-cutting of chromosome III leads to formation of His^+ colonies, of which 0.036% were also Leu^+. Deletion of *PMS1* in this strain led to a 7-fold increase in the frequency of colonies that were also Leu^+ (0.25%), of which 24/38 (63%) were sectored Leu^+/Leu^- (Table 2). Half of these Leu^+ colonies were also coconverted or sectored for the adjacent *ADE1* gene. Leu^+ Ade^- colonies may have arisen if only DNA covering the *leu2* alleles was assimilated or if the *ADE1* allele was mismatch corrected back to *ade1-E*. Among the colonies sectored for both markers, if a single strand were assimilated into heteroduplex, such colonies should have the cis configuration Leu^+ Ade^+/Leu^- Ade^-, whereas if there were long heteroduplexes at either end of a double crossover, as in Fig. 3*C*, one might expect some or all Leu^+ and Ade^+ sectors to be in trans. Of seven doubly sectored colonies, all were in the cis configuration. As discussed above, no such sectors were seen in an equivalent number of *PMS1* colonies.

†The terms "ends-out" and "ends-in" integrative recombination were defined by Thaler and Stahl (18). "Ends-out" recombination is exemplified by one-step gene replacement (1) in which the ends of the integrating DNA are pointed away from each other. The two DNA ends thus invade noncontiguous regions of homology (19, 20). "Ends-in" recombination is exemplified by the integration of a gapped plasmid molecule (21) in which the invading ends are pointed toward each other and the two ends define a single break or gap in a linear region of homology (18, 19).

Table 2. *LEU2* and *ADE1* capture in *pms1*Δ strain G543

Colony type	No. of His^+ colonies
Leu^+ Ade^+/Leu^+ Ade^+	8
Leu^+ Ade^+/Leu^+ Ade^-	0
Leu^+ Ade^+/Leu^- Ade^+	4
Leu^+ Ade^+/Leu^- Ade^-	7
Leu^+ Ade^-/Leu^- Ade^+	0
Leu^+ Ade^-/Leu^+ Ade^-	6
Leu^+ Ade^-/Leu^- Ade^-	13
Leu^- Ade^+/Leu^- Ade^+	11
Leu^- Ade^+/Leu^- Ade^-	16

An HO-liberated *LEU2*::*ADE1* fragment could gene convert a *leu2-3,112*::*ade1-E* target locus to produce His^+ colonies that were also Leu^+ and/or Ade^+ (Fig. 1*B*). In the experiment shown, 14,380 His^+ colonies were examined. Each colony type is shown as two half-sectors. Those that are identical to the left and right of the slash were not sectored, while those that are nonidentical are colonies that were actually sectored.

DISCUSSION

Gene Replacement by Single-Strand Assimilation. Our data suggest that conversion of *leu2* alleles by a linearized 2.0-kb *LEU2* fragment does not occur by two localized crossovers at each end of the DNA fragment (Fig. 3*B*). Such a process should replace both DNA strands carrying the *leu2* alleles in the middle of the fragment with two strands of *LEU2* DNA, so that there would be no heteroduplex DNA covering these sites. Our results show that heteroduplex DNA frequently covers both *leu2–3* and *leu2–112*, located 542 bp apart and more than 650 bp from each end of the recombining DNA. We therefore propose that frequently only one end of the transforming DNA invades the duplex and is assimilated, by the agency of a strand exchange protein and possibly a helicase that unwinds the parental duplex (Fig. 3*A*). The process can be facilitated if a single-stranded exonuclease removes the complementary invading strand as its partner is wound into the chromosome (22). This will produce a long heteroduplex region of DNA that will be acted upon by the mismatch repair system. The assimilation of a single strand of a 2- to 4-kb fragment would also be the expected result if strand invasion is inherently inefficient, where it would be unlikely that both ends of a linear fragment would independently engage in strand invasion (Fig. 3*B*) before strand assimilation from one end covers the entire region (Fig. 3*A*).

Undoubtedly, some "ends-out" transformation events are the consequence of two independent crossing-over events at the ends (1, 19, 20), such as those creating large chromosomal deletions, using a fragment containing two very distant regions of homology flanking a selectable marker (20). However, at least when the DNAs are 2 to 4 kb in length and differ by only one or two heterologies, our results suggest that most events involve the assimilation of only one DNA strand.

Recently, Negritto *et al.* (23) independently developed a similar strategy for the study of gene replacement, in which a chromosomal *sam2*::*HIS3* disruption is converted to *SAM2* by HO endonuclease-induced liberation of a homologous (*SAM2*) or a homeologous (*SAM1*) fragment. Their results differ from ours in several respects. First, the frequency with which the disrupted gene is converted to wild type by a completely homologous HO-derived linear fragment is only 15-fold above the spontaneous background, compared with at least a 250-fold stimulation in our case. This may be explained by the fact that Negritto *et al.* (23) require the replacement of a large insertion, whereas our experiments involve only frameshift insertions of 1 or 4 bp. Heteroduplex formation across an entire assimilated strand may often be impeded by a large heterology.

Using various chimeric homologous–homeologous fragments, Negritto *et al.* (23) suggest that gene conversion occurred primarily by crossings-over at the ends of the linear fragment, while our results argue to the contrary. They draw this conclusion from the observation that the severely inhibiting effects of providing homeologous (15% diverged) *SAM1* sequences for gene conversion of *sam2*::*HIS3* were seen only when the homeology was at the ends of the fragment. But these homeologous sequences may be incapable of initiating the formation of heteroduplex DNA at either DNA end, thus effectively preventing any strand assimilation, while such divergence may be better tolerated in the interior of the fragment, once a region of homologous base pairing is established. Moreover, gene conversion of *sam2*::*HIS3* by identical *SAM2* sequences was stimulated 40-fold by a mutation in the mismatch repair gene, *msh2* (23). Such a stimulation is inconsistent with the idea that recombination occurs only at the two (perfectly homologous) ends of a DNA fragment but is consistent with our observations for *pms1*.

Differences Between "Ends-in" and "Ends-out" Recombination. The data we have presented here argue that during "ends-out" gene replacement by a linear DNA fragment, only one of the two strands of DNA is assimilated, creating heteroduplex DNA over the entire length. This process may be different from the apparently more limited formation of heteroduplex DNA at the ends of linear DNA fragments that undergo "ends-in" recombination (gap repair), which appears usually to be limited to a few hundred base pairs (24, 25). However, other studies of "ends-in" recombination show that a marker on a linearized plasmid, several kilobases from the DNA end, can be transferred to the intact chromosome in cells in which gap repair has also occurred (26). In these transformation experiments, it is possible that donation of a marker to the chromosome occurred by an "ends-out" process analogous to that we have studied, while gap repair involved another DNA molecule.

We believe that "ends-in" recombination and "ends-out" recombination often proceed by different mechanisms. For example, the efficiency of these two types of events seems to be quite different. We have compared the efficiency of "ends-out" gene conversion involving an HO-liberated *LEU2* fragment with the "ends-in" repair of an HO-induced double-strand break within the same *LEU2* sequences on a centromeric plasmid that recombines with a *leu2* chromosomal sequence. In both cases, physical monitoring showed that nearly all of the DNA was cleaved by HO endonuclease. While "ends-out" *LEU2* capture occurred at a frequency of less than 5×10^{-3} (even when mismatch correction is eliminated), "ends-in" gap repair occurred in more than 10% of cells (F. Pâques, W.-Y.L., and J.E.H., unpublished results). The 20-fold difference we see between "end-in" and "ends-out" recombination is substantially higher than the 3-fold difference noted by Hastings *et al.* (19) in a well controlled transformation experiment. Possibly the discrepancy between these results reflects the difference between HO-endonuclease liberation of a single linearized DNA molecule, already in chromatin, within the nucleus versus the introduction of one or more copies of naked DNA by transformation.

Even though biased mismatch repair of the invading DNA strand represents a significant barrier to efficient "ends-out" recombination, there must be other limitations that cause gene replacement to be 100-fold less efficient than an equivalent "ends-in" event. Physical analysis shows that the linearized fragment is not rapidly degraded. One distinctive difference is the direction of new DNA synthesis that could be primed by an invading 3′ end. An increasingly popular mechanism to account for gene conversions without an associated crossing-over is termed "synthesis-dependent strand annealing" (27–30), in which a newly synthesized strand is displaced as a single strand from a migrating replication "bubble" (31). When the synthe-

sized strand extends far enough to anneal with the second DNA end (which may never have invaded the donor), the complementary strand can be synthesized and gap repair is accomplished. If one envisions a similar scenario in an "ends-out" event, bubble migration progresses *away* from the opposite DNA end and there is no way to terminate this process or to engage the other end. No stable heteroduplex DNA is formed and many of these events might simply be lost. Indeed, from this point of view, the successful events that we have detected may have come from invasion of a 5′ end, where no new DNA synthesis would be initiated and the strand could be entirely assimilated.

Mismatch Repair Is Biased During Transformation. A second major conclusion of our work is that mismatch correction strongly favors restoration of heteroduplex DNA containing small heterologies to the genotype of the resident DNA strand. If heteroduplex DNA were formed, but there were no bias in mismatch repair, then we would have expected at best a 2-fold increase in Leu^+ transformants by deleting *PMS1* instead of the 7- to 25-fold increase we observed. This strong strand preference is expected if the mismatch correction system selects which strand to repair on the basis of a nearby nick or gap in the DNA (8, 11–14).

These results are reminiscent of the effects of the mismatch correction system in *Pneumococcus*, where the Pms1 homologue, HexB, has been shown to prevent efficient transformation by DNA containing a single mismatch (32). This result has often been interpreted as evidence that HexB acts as an anti-recombinator by provoking the removal of mismatch-containing heteroduplex DNA. Moreover, the *Escherichia coli* homologue of Pms1, MutL, acts to prevent recombination between DNA substrates that have many heterologous sites (33). However these studies do not distinguish between two alternative interpretations: (*i*) the mismatch system actually prevents stable heteroduplex DNA formation or (*ii*) highly preferential mismatch correction, in favor of the resident strand, excises and eliminates the transforming DNA after heteroduplex DNA is formed. We believe that the latter interpretation is more likely. Indeed, when transforming pneumococcal DNA contains not only a single mismatch but also other types of (repairable) DNA modifications, the efficiency of the transformation of the low-efficiency marker increases, as would be expected if heteroduplex DNA were formed but other DNA repair activities altered the pattern of mismatch repair (34).

These results are valuable in understanding transformation in higher eukaryotes. The preferential repair of heteroduplex DNA in favor of resident information may also explain why the recovery of transformants in mammalian cells is inefficient and why targeted transformation is greatly increased by the use of donor DNA that is derived from the same cells used as recipients (35) or by using cells defective in mismatch repair (36). When donor DNA contains a number of base pair substitutions and small heterologies, mismatch repair may eliminate the selectable marker by its cocorrection along with adjacent mismatches, in favor of the resident strand.

We thank Rhona Borts for the kind gift of a plasmid to create the *ade1-E* allele. Michael Lichten, Susan Lovett, and members of the Haber laboratory offered trenchant comments on the manuscript. This work was supported by National Institutes of Health Grant GM20056.

1. Rothstein, R. (1983) *Methods Enzymol.* **103,** 202–211.
2. Schiestl, R. H. & Geitz, D. H. (1989) *Curr. Genet.* **16,** 337–346.
3. Wilson, J. H., Leung, W.-Y., Bosco, G., Dieu, D. & Haber, J. E. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 177–181.
4. Fishman-Lobell, J., Rudin, N. & Haber, J. E. (1992) *Mol. Cell. Biol.* **12,** 1292–1303.
5. Sugawara, N. & Haber, J. E. (1992) *Mol. Cell. Biol.* **12,** 563–575.
6. Haber, J. E. & Leung, W.-Y. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 13949–13954.
7. Nickoloff, J. A., Singer, J. D., Hoeckstra, M. F. & Heffron, F. (1989) *J. Mol. Biol.* **207,** 527–541.
8. Ray, B. L., White, C. I. & Haber, J. E. (1991) *Mol. Cell. Biol.* **11,** 5372–5380.
9. Andreadis, H., Hsu, Y. P., Kohlhaw, G. B. & Schimmel, P. (1982) *Cell* **31,** 319–325.
10. Chen, D.-Z., Yang, B.-C. & Kuo, T.-K. (1992) *Curr. Genet.* **21,** 83–84.
11. Modrich, P. (1991) *Annu. Rev. Genet.* **25,** 229–253.
12. Holmes, J., Clark, S. & Modrich, P. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5837–5841.
13. Varlet, I., Canard, B., Brooks, P., Cerovic, G. & Radman, M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 10156–10161.
14. Haber, J. E., Ray, B. L., Kolb, J. M. & White, C. I. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 3363–3367.
15. Williamson, M., Game, J. C. & Fogel, S. (1985) *Genetics* **110,** 609–646.
16. Kramer, W., Kramer, B., Williamson, M. S. & Fogel, S. (1989) *J. Bacteriol.* **171,** 5339–5346.
17. Bishop, D. & Kolodner, R. D. (1986) *Mol. Cell. Biol.* **6,** 3401–3409.
18. Thaler, D. & Stahl, F. W. (1988) *Annu. Rev. Genet.* **22,** 169–197.
19. Hastings, P. J., McGill, C., Schafer, B. & Strathern, J. N. (1993) *Genetics* **135,** 973–980.
20. Surosky, R. T. & Tye, B. K. (1987) *Methods Enzymol.* **153,** 243–253.
21. Orr-Weaver, T., Szostak, J. W. & Rothstein, R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6354–6358.
22. Corrette-Bennett, S. & Lovett, S. T. (1995) *J. Biol. Chem.* **270,** 6881–6885.
23. Negritto, M. T., Wu, X., Kuo, T., Chu, S. & Bailis, A. M. (1997) *Mol. Cell. Biol.* **17,** 278–286.
24. Orr-Weaver, T., Nicolas, A. & Szostak, J. W. (1988) *Mol. Cel. Biol.* **8,** 5292–5298.
25. Sweetser, D. B., Hough, H., Whelden, J. F., Arbuckle, M. & Nickoloff, J. A. (1994) *Mol. Cell. Biol.* **14,** 3863–3875.
26. Roitgrund, C., Steinlauf, R. & Kupiec, M. (1993) *Curr. Genet.* **23,** 414–422.
27. Hastings, P. W. (1988) *BioEssays* **9,** 61–64.
28. Nassif, N., Penney, J., Pal, S., Engels, W. R. & Gloor, G. B. (1994) *Mol. Cell. Biol.* **14,** 1613–1625.
29. Nelson, H. H., Sweetser, D. B. & Nickoloff, J. A. (1996) *Mol. Cell. Biol.* **16,** 2951–2957.
30. Mueller, J. E., Clyman, J., Huang, Y. J., Parker, M. M. & Belfort, M. (1996) *Genes Dev.* **10,** 351–364.
31. Formosa, T. & Alberts, B. M. (1986) *Cell* **47,** 793–806.
32. Claverys, J. P. & Lacks, S. A. (1986) *Microbiol. Rev.* **50,** 133–165.
33. Rayssinguier, C., Thaler, D. S. & Radman, M. (1989) *Nature (London)* **342,** 396–401.
34. Mejean, V., Devejiian, J. C., Rives, I., Alloing, G. & Claverys, J. P. (1991) *Nucleic Acids Res.* **19,** 5525–5531.
35. te Riele, H., Maandag, E. R. & Berns, A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 5128–5132.
36. de Wind, N., Dekker, M., Berns, A., Radman, M. & te Riele, H. (1995) *Cell* **82,** 321–330.