Efficient Incorporation of Large (>2 kb) Heterologies Into Heteroduplex DNA: 
Pms1/Msh2-Dependent and -Independent Large Loop Mismatch Repair in 
Saccharomyces cerevisiae

Jennifer A. Clikeman, Sarah L. Wheeler and Jac A. Nickoloff

Department of Molecular Genetics and Microbiology, University of New Mexico School of Medicine, Albuquerque, New Mexico 87131

Manuscript received November 10, 2000
Accepted for publication January 2, 2001

ABSTRACT

DNA double-strand break (DSB) repair in yeast is effected primarily by gene conversion. Conversion can conceivably result from gap repair or from mismatch repair of heteroduplex DNA (hDNA) in recombination intermediates. Mismatch repair is normally very efficient, but un repaired mismatches segregate in the next cell division, producing sectored colonies. Conversion of small heterologies (single-base differences or insertions <15 bp) in meiosis and mitosis involves mismatch repair of hDNA. The repair of larger loop mismatches in plasmid substrates or arising by replication slippage is inefficient and/or independent of Pms1p/Msh2p-dependent mismatch repair. However, large insertions convert readily (without sectoring) during meiotic recombination, raising the question of whether large insertions convert by repair of large loop mismatches or by gap repair. We show that insertions of 2.2 and 2.6 kbp convert efficiently during DSB-induced mitotic recombination, primarily by Msh2p- and Pms1p-dependent repair of large loop mismatches. These results support models in which Rad51p readily incorporates large heterologies into hDNA. We also show that large heterologies convert more frequently than small heterologies located the same distance from an initiating DSB and propose that this reflects Msh2-independent large loop-specific mismatch repair biased toward loop loss.

In Saccharomyces cerevisiae, most DNA double-strand breaks (DSBs) are repaired by recombination, principally gene conversion, with or without an associated crossover (Paques and Haber 1999). Gene conversion also plays a significant role in the repair of chromosomal DSBs in mammalian cells (Taghian and Nickoloff 1997; Liang et al. 1998) and conversion from pseudogene donors has been implicated in human diseases (e.g., Watnick et al. 1998). Alleles suffering a DSB usually receive information from unbroken alleles, and most conversion tracts are continuous (Petes et al. 1991; Nickoloff and Hoekstra 1998). Although these features can be explained by a model in which conversion occurs in a double-strand gap (Szostak et al. 1983), most meiotic and mitotic conversion in yeast involves mismatch repair of heteroduplex DNA (hDNA; Petes et al. 1991; Nickoloff and Hoekstra 1998; Weng and Nickoloff 1998; Nickoloff et al. 1999). When mismatches escape repair, markers in hDNA segregate in the next cell division, producing sectored colonies; this is termed postmeiotic segregation (PMS) for meiotic events; analogous events can occur in mitotic cells. Gene conversion can be described in four steps: initiation, end-processing, hDNA formation, and resolution of intermediates. DSBs are potent initiators of gene conversion, and recent studies have clarified how broken ends are processed into 3’ single-stranded tails (reviewed in Paques and Haber 1999), but the subsequent hDNA formation step remains unclear. hDNA can form as a result of strand invasion (synapsis) and by branch migration of Holliday junctions. In yeast, strand exchange is presumed to be mediated by Rad51p, a homolog of Escherichia coli RecA (Shinohara et al. 1992; Ogawa et al. 1993; Sung and Stratton 1996). RecA and Rad51p bind to processed ends (3’ single-stranded tails), forming nucleoprotein filaments, and both display DNA-dependent ATPase activity and can pair or transfer complementary DNA strands in vitro.

Resolution of recombination intermediates includes mismatch repair of hDNA and/or resolution of Holliday junctions. Mismatch repair proteins are conserved from bacteria to higher eukaryotes. In E. coli, mutHLS mediates the dominant mismatch repair mechanism that involves excision and new synthesis of long DNA tracts that can extend >1 kb (Modrich 1991; Rasmussen et al. 1998). MutS functions in mismatch recognition and MutL couples the mismatch-bound MutS to proteins involved in later steps. Yeast have several mutS homologs, including MSH2, MSH3, MSH6, and several mutL homologs, including PMS1 and MLH1 (Crouse 1998). Msh2p in complex with Msh6p or Msh3p binds to single-base or loop mismatches, respectively. Despite this conservation at the protein level, repair efficiencies of various types of mismatches differ markedly among different
organisms. In *E. coli* most single-base mismatches and small loops (<4 bases) are repaired efficiently, but GC and larger loop mismatches are repaired only as part of a tract initiated by another mismatch (Carraway and Marinus 1993). In yeast most single-base mismatches are repaired efficiently; GC and palindromic loop mismatches that form stable stem-loop structures are poorly repaired unless a well-repaired mismatch is nearby (Nag et al. 1989; Nag and Petes 1991; Weng and Nickoloff 1998). Short loop mismatches (1–15 bases) are well repaired in yeast (Bishop and Kolodner 1986; Bishop et al. 1989; Kramer et al. 1989), but studies of larger loop mismatches have led to conflicting conclusions. In transformation assays with artificial hDNA plasmid substrates, a 38-base loop was repaired with low efficiency (30–50%; Kramer et al. 1989; Luhr et al. 1998). Two reports indicate that loops >15 bases formed by DNA polymerase slippage are not subject to repair (Tran et al. 1996; Sia et al. 1997), but another report indicates a role for Msh3p in the repair of 94-base loops (Harfe and Jinks-Robertson 1999; Harfe et al. 2000). Repair of relatively large loop mismatches (16–216 bases) in plasmid hDNA substrates has been demonstrated *in vitro* with yeast nuclear extracts; this activity is independent of MSH2, MSH3, MLH1, and PMS1 (Corrette-Bennett et al. 1999). In contrast, even very large insertion mutations (>1 kbp) are converted during meiosis (reviewed in Petes et al. 1991); if these insertions convert as a result of inclusion in hDNA it would suggest that very large loop mismatches are efficiently repaired. It has been suggested that conversions of large insertions might be mediated by gap repair (Szostak et al. 1983; Tran et al. 1996) rather than by mismatch repair.

A related question concerns whether large loop mismatches form *in vivo* i.e., can Rad51p incorporate very large insertions into hDNA? Studies of RecA provide insight into this question. RecA-mediated strand transfer *in vitro* is impeded by point mutations and blocked by a 2-kbp heterology (DasGupta and Radding 1982). However, heterologies >1 kbp are readily incorporated into hDNA when RecA is augmented by single-strand binding protein (SSB) and an ATP regeneration system (Bianchi and Radding 1983), and 1.3-kbp loop mismatches were detected *in vivo* during recA-dependent λ recombination (Lichten and Fox 1984). RecA mediates strand transfer between homologous DNAs without ATP hydrolysis, but incorporation of heterologies into hDNA requires ATP hydrolysis (Rosselli and Stasiak 1991; Kim et al. 1992); this reaction is also facilitated by RuvA and RuvB (Iype et al. 1994; Adams and West 1996). In yeast, 32-base loop mismatches were detected during meiosis (Nag and Petes 1993), but a similar analysis of larger loops has not been reported. Yeast Rad51p shares many biochemical properties with RecA, including strand transfer activity that is facilitated by the yeast SSB homolog, replication protein A (Sung 1997). Like RecA, Rad51p promotes strand transfer between homologous DNAs without ATP hydrolysis *in vitro* (Sung and Stratton 1996).

In this study we analyzed mitotic gene conversion in diploid yeast in which events were initiated at a defined DSB created by HO nuclease. We demonstrate that a heterology of 2.6 kbp is converted even more often than an equidistant small heterology and that this large heterology frequently segregates in *pms1* and *msh2* mutants. These results indicate that large heterologies are readily incorporated into hDNA, that the resulting large loop mismatches are efficiently repaired, and that the majority of this repair involves Pms1p and Msh2p.

**MATERIALS AND METHODS**

**Plasmid DNA, plasmid rescue, PCR, and yeast strains:** Plasmid manipulation, plasmid rescue, PCR, restriction fragment length polymorphism (RFLP) mapping strategies, yeast culture, and strain construction were described previously (Sweetser et al. 1994; Nickoloff et al. 1999). Strain genotypes are given in Table 1. Recombinate substrate structures (Figure 1) were confirmed by Southern hybridization and by restriction mapping of rescued plasmids and/or PCR products. All strains carry an integrated copy of *GAL1* promoter-driven HO nuclease (*GALHO*) and a copy of *ura3* (recipient allele) with a 24-bp HO recognition sequence at position 432 (HO432) into which DSBs are introduced upon growth in medium with galactose. In some strains the second (donor) copy of *ura3* was inactivated by the nonrevertible frameshift mutation X764 (Sweetser et al. 1994); in other strains the donor copy was wild type (*ura3*). *ura3* alleles flanked by pUC19 and LEU2 were constructed by using derivatives of the RecR1 transplacement vector; in some strains HIS3 was inserted ~8 kbp from the telomere on the same arm as *ura3* (Nickoloff et al. 1999). *pms1* and *msh2* knockout vectors were kindly provided by R. Kolodner and E. Alani. The msh2 knockout vector, pEAl99, replaces nearly all of the *MSH2* sequence with a related question concerns whether large loop mismatches form *in vivo* i.e., can Rad51p incorporate very large insertions into hDNA? Studies of RecA provide insight into this question. RecA-mediated strand transfer *in vitro* is impeded by point mutations and blocked by a 2-kbp heterology (DasGupta and Radding 1982). However, heterologies >1 kbp are readily incorporated into hDNA when RecA is augmented by single-strand binding protein (SSB) and an ATP regeneration system (Bianchi and Radding 1983), and 1.3-kbp loop mismatches were detected *in vivo* during recA-dependent λ recombination (Lichten and Fox 1984). RecA mediates strand transfer between homologous DNAs without ATP hydrolysis, but incorporation of heterologies into hDNA requires ATP hydrolysis (Rosselli and Stasiak 1991; Kim et al. 1992); this reaction is also facilitated by Ruva and RuVB (Iype et al. 1994; Adams and West 1996). In yeast, 32-base loop mismatches were detected during meiosis (Nag and Petes 1993), but a similar analysis of larger loops has not been reported. Yeast Rad51p shares many biochemical properties with RecA, including strand transfer activity that is facilitated by the yeast SSB homolog, replication protein A (Sung 1997). Like RecA, Rad51p promotes strand transfer between homologous DNAs without ATP hydrolysis *in vitro* (Sung and Stratton 1996).

In this study we analyzed mitotic gene conversion in diploid yeast in which events were initiated at a defined DSB created by HO nuclease. We demonstrate that a heterology of 2.6 kbp is converted even more often than an equidistant small heterology and that this large heterology frequently segregates in *pms1* and *msh2* mutants. These results indicate that large heterologies are readily incorporated into hDNA, that the resulting large loop mismatches are efficiently repaired, and that the majority of this repair involves Pms1p and Msh2p.

**Materials and Methods**

**Plasmid DNA, plasmid rescue, PCR, and yeast strains:** Plasmid manipulation, plasmid rescue, PCR, restriction fragment length polymorphism (RFLP) mapping strategies, yeast culture, and strain construction were described previously (Sweetser et al. 1994; Nickoloff et al. 1999). Strain genotypes are given in Table 1. Recombination substrate structures (Figure 1) were confirmed by Southern hybridization and by restriction mapping of rescued plasmids and/or PCR products. All strains carry an integrated copy of *GAL1* promoter-driven HO nuclease (*GALHO*) and a copy of *ura3* (recipient allele) with a 24-bp HO recognition sequence at position 432 (HO432) into which DSBs are introduced upon growth in medium with galactose. In some strains the second (donor) copy of *ura3* was inactivated by the nonrevertible frameshift mutation X764 (Sweetser et al. 1994); in other strains the donor copy was wild type (*ura3*). *ura3* alleles flanked by pUC19 and LEU2 were constructed by using derivatives of the RecR1 transplacement vector; in some strains HIS3 was inserted ~8 kbp from the telomere on the same arm as *ura3* (Nickoloff et al. 1999). *pms1* and *msh2* knockout vectors were kindly provided by R. Kolodner and E. Alani. The msh2 knockout vector, pEAl99, replaces nearly all of the *MSH2* sequence with a
et al. 1996). All recombinant products were independent since only one product of a particular phenotype was isolated from a parent population. Recombinants that lost or retained pUC19 were distinguished by Southern hybridization. ura3 alleles in recombinants that retained pUC19 were rescued as described (Nickoloff et al. 1999). ura3 alleles in Leu+ recombinants that lacked pUC19 were PCR amplified by using primers specific for sequences upstream of ura3 and in LEU2. In either case, conversion tracts were characterized by mapping silent RFLPs with established procedures (Sweetser et al. 1994). Leu+ products that had lost pUC19 were assumed to have converted all markers in the recipient allele and were not characterized further.

In strains carrying a URA3 donor, recombination was induced and cells were seeded to YPD plates as above, but the resulting YPD colonies were replica plated only to leucine omission medium to score Leu+ recombinants. Because the Ura phenotype was not informative for recombination in strains carrying a URA3 donor, we determined whether the Leu+ half of Leu+/− colonies was parental (retained HO432) or recombinant (lost HO432) in 30 Leu+/− products from each strain (wild type, pms1, msh2) as follows. Leu+ cells from Leu+/− colonies were dispersed to leucine omission plates, incubated for 2 days, and the resulting colonies were replica plated to YPGal to reinduce GALHO for 24 hr. Cells from YPGal colonies were then dispersed to semiselective leucine medium on which Leu+ colonies appear pink and Leu− colonies appear white in an ade2 background (Myers and Nickoloff 1999). If the Leu+ half of Leu+/− colonies was parental, reinduction of GALHO stimulated recombination and many Leu− colonies arose, whereas GALHO induction does not stimulate recombination in Leu+ recombinants, and few or no Leu− colonies arose. Leu+/− colonies with parental sectors were presumed to arise during G2 and were scored as nonsectored in calculations of sector rates.

In patch assays, pms1 and msh2 mutants show higher frequencies of spontaneous Ura+ products than wild type (our unpublished results). To avoid problems associated with jackpots, recombination assays in these mutants (and in wild-type controls done in parallel) were not performed with individual parent colonies. Instead, frozen stocks of subclones with low background levels of Ura− cells were identified by spreading 1-cm2 patches on YPD plates, incubating for 2 days, and replica plating to uracil omission medium. Approximately 5–10× 105 cells from each of several different areas of patches that exhibited a low level of Ura+ papillae were transferred to tubes with 1.5 ml of YPGly medium and treated as above, except that GALHO was induced using YPGal with 5% galactose for only 2 hr. The shorter induction period minimizes cell division prior to plating, providing a more accurate estimate of marker segregation rates. Statistical analyses were performed by using t-tests unless otherwise specified.

### RESULTS

**Experimental design:** We previously studied DSB-induced allelic recombination in multiply marked copies of ura3 (Nickoloff et al. 1999). In that system, one copy of ura3 was inactivated by insertion of a 24-bp HO site (HO432) and the second copy by a +1 frameshift mutation (X764). There were nine additional phenotypically silent RFLP mutations present at ~100-bp intervals in ura3 to allow high resolution mapping of conversion tracts initiated by DSBs in HO432 following galactose induction of GALHO (Figure 1A). This system provided information about gene conversion tract lengths, directionality, and symmetry relative to a defined DSB. Prod-
products showing loss of heterozygosity at all markers centromere-distal to HO432 could result from gene conversion, break-induced replication (BIR), or G2 crossovers (Figure 2). Products showing loss of heterozygosity at all markers could have resulted from these three processes, as well as by chromosome loss. By using a telomere-proximal HIS3 gene (HIS3-telV) linked to ura3 (110 kbp from HO432), we showed that BIR and chromosome loss were rare, with most products resulting from gene conversion. Associated crossovers were seen in ~20% of conversions, and, as expected, ~25% of crossovers (5% of total events) led to loss of HIS3-telV.

Recombination was induced in liquid medium with galactose as described in MATERIALS AND METHODS, and cells were seeded to nonselective solid medium (YPD), which supports growth of parental cells and all types of recombinant products. DSBs in HO432 lead to conversion (loss) of HO432; if the conversion tract is short (i.e., does not also encompass X764), the product will be Ura+ and these were identified by replica plating YPD colonies to medium lacking uracil. Recombinants with longer conversion tracts that encompass X764 have the same Ura− phenotype as parents; Ura− recombinants were distinguished from Ura− parents by using a reinduction assay as described in MATERIALS AND METHODS. Because nearly all tracts are continuous, tracts in Ura+ recombinants usually terminate before X764, or from independent events in G2. We believe the majority of Ura+ products provides an estimate of tract lengths. Another class of products is sectored Ura+/−, which can arise from segregation of unrepaired hDNA that encompasses X764, or from independent events in G2. We believe the majority of Ura+/− sectors reflect independent events in G2 because X764 produces a 4/5 bubble mismatch that displayed low segregation rates in several studies of direct repeat and plasmid × chromosome recombination (Sweetser et al. 1994; Cho et al. 1998; Weng and Nickoloff 1998; Nickoloff et al. 1999). To generate complete gene conversion tract spectra, we analyze tracts in sets of Ura+ and Ura− products, then combine the data in proportion to the measured Ura+ and Ura− frequencies; this produces spectra that are not biased by selection of a particular phenotype. Adjusting Ura+ and Ura− frequencies by adding half of the Ura+/− products to the Ura+ class and half to the Ura− class does not significantly alter product spectra (data not shown).

In the prior system (Nickoloff et al. 1999), all heterologies were small, ranging from single-base substitutions to linker insertions (Figure 1A). In the present study we analyzed conversion of large heterologies flanking the 1.2-kbp fragment carrying ura3, including a 2.2-kbp heterology consisting of most of pUC19, and a 2.6-kbp heterology consisting of a 2.2-kbp LEU2 frag-
ment and the remaining 0.4 kbp of pUC19. Beyond these large heterologies is essentially unlimited homology (>100 kbp; Figure 1, B and C). As described above, some strains were marked with HIS3-telV to monitor G2 crossovers and chromosome loss.

**Large heterologies do not influence recombination frequencies, gene conversion tract spectra, or rates of chromosome loss:** Total induced recombination frequencies (including Ura+, Ura-, and sectored Ura+/− products) were similar for substrates with small heterologies or small and large heterologies (Table 2, experiments 1 and 2), indicating that large heterologies do not inhibit DSB-induced recombination. Total induced recombination was also similar in strains carrying the distant HIS3-telV marker (Table 2, experiments 3 and 4). In fact, none of these recombination frequencies are significantly different (all \( P > 0.15 \)).

We characterized conversion tracts in 91 independent recombinant strains SW3476 (small and large heterologies). As expected, continuous tracts were predominant. Of the 91 products, 4 were Ura+ Leu−; i.e., they did not convert X764 but did convert the more distal LEU2 marker and therefore had discontinuous tracts; these were not characterized further. The conversion tract spectrum generated from the remaining 87 products shares several features with the spectrum obtained previously with strain DY3515-15 (small heterologies). In both cases, 79% of tracts were bidirectional, most tracts were long (the most common product converted all markers), and no crossovers were observed without an associated gene conversion (data not shown). Both spectra also displayed a bias toward conversion of markers promoter-proximal (5′) to the DSB. In the presence or absence of large heterologies, 5′ unidirectional tracts comprised 19 and 20% of products, respectively, but no 3′ unidirectional tracts were recovered (Table 3).

Analysis of individual marker conversion rates revealed a second 5′ conversion bias: markers 5′ of the DSB converted at significantly higher rates than equidistant 3′ markers (Figure 3). Note that these biases are generally independent since individual marker conversion rates are derived from all products, while only ~20% of products have unidirectional tracts. The 5′ conversion bias may reflect a transcriptional effect (WENG et al. 2000).

All markers 5′ of HO432 were lost in 99% of SW3476 products; these may have resulted from gene conversion, BIR, or G2 crossovers. In 79% of products, all markers were lost and these may have resulted from these processes, as well as from chromosome loss. To distinguish these possibilities we monitored recombination with large and small heterologies at ura3 in the presence of the HIS3-telV marker. Retention of HIS3-telV rules out chromosome loss and BIR for products showing partial or complete marker loss at ura3. We found that 90.8% of products retained HIS3-telV; this is similar to the level obtained with small heterologies (Table 4). As seen previously with small heterologies, substantial fractions of products that lost HIS3-telV retained other markers at ura3, or were expected to result from G2 crossovers (estimated in measurements of HIS3-telV gain among His+/− products), with at most 4% of products resulting from chromosome loss (data not shown). Together, these results indicate that large heterologies do not affect G2 crossover frequencies, nor do they enhance BIR or chromosome loss in lieu of gene conversion.

**Large heterologies are readily incorporated into hDNA, and they convert more efficiently than equidistant small heterologies:** Conversion of individual markers decreases with distance from an initiating DSB. In SW3476, the large pUC19 and LEU2 heterologies are located 432 and 738 bp from HO432, respectively. These distances are essentially the same as the small heterologies Ase20 and B3′ in DY3515-13. Interestingly, the large LEU2 heterology converted at higher rates than the small equidistant B3′ heterology; a similar trend is apparent for pUC19 and Ase20 (see Discussion). In contrast, shared small heterologies converted at similar frequencies in the presence or absence of large heterologies (Figure 3). Thus, rather than being refractory to conversion, large heterologies converted as often, or more often, than equidistant small heterologies. This result might be easy to explain if conversion occurred by gap repair (SZOSTAK et al. 1983), but several studies have shown that DSB-induced gene conversion of small heterologies is mediated by mismatch repair of hDNA (PETES et al. 1991; RAY et al. 1991; WENG and NICKOLOFF 1998; NICKOLOFF et al. 1999).

To determine whether conversion of large heterologies is mediated by mismatch repair of hDNA, we examined recombination in mismatch repair (MMR)-defective strains with small and large heterologies. For these experiments, run in parallel with the wild-type strain, GALHO induction was limited to 2 hr to minimize cell division prior to plating, since cell division in liquid reduces the number of observable segregation events (cells divide once per 3 hr in galactose medium; data not shown). DSB-induced recombination frequencies were very similar in wild-type and pms1 mutant cells (Table 2, experiments 5 and 6). DSB-induced recombination was slightly lower in the msh2 mutant (Table 2, experiments 5 and 7), but this difference was not statistically significant (\( P = 0.06 \)). Msh2p (but not Pms1p) plays a role with Rad1p/Rad10p in processing long nonhomologous tails during HO site conversion (SUGAWARA et al. 1997). In the present crosses, the HO432 insertion is 39 bp in length (a 24-bp HO site plus an EcoRI linker); hence, each nonhomologous tail is ~30 nucleotides (nt) in length. The similar recombination frequencies in wild-type and msh2 strains are consistent with data indicating that Msh2p has little or no role in processing short (<30 nt) nonhomologous tails (SUGAWARA et al. 1997). Interestingly, tract lengths
TABLE 2
Recombination frequencies

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Strain</th>
<th>n</th>
<th>Induced (hr)</th>
<th>Heterologies</th>
<th>MMR</th>
<th>Total</th>
<th>Glucose</th>
<th>Galactose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ura⁺</td>
<td>Ura⁻</td>
<td>Ura⁺⁻⁻⁻⁻</td>
</tr>
<tr>
<td>1</td>
<td>DY3515-13</td>
<td>3</td>
<td>6</td>
<td>Small</td>
<td>WT</td>
<td>65 ± 13</td>
<td>200 ± 36</td>
<td>980 ± 67</td>
</tr>
<tr>
<td>2</td>
<td>SW3476</td>
<td>4</td>
<td>6</td>
<td>Small + Large</td>
<td>WT</td>
<td>28 ± 33</td>
<td>229 ± 49</td>
<td>1195 ± 133</td>
</tr>
<tr>
<td>3</td>
<td>JC3517-13</td>
<td>3</td>
<td>6</td>
<td>Small</td>
<td>WT</td>
<td>65 ± 26</td>
<td>205 ± 24</td>
<td>952 ± 151</td>
</tr>
<tr>
<td>4</td>
<td>JC3525</td>
<td>4</td>
<td>6</td>
<td>Small + Large</td>
<td>WT</td>
<td>44 ± 25</td>
<td>180 ± 21</td>
<td>1028 ± 48</td>
</tr>
<tr>
<td>5</td>
<td>SW3476</td>
<td>4</td>
<td>2</td>
<td>Small + Large</td>
<td>WT</td>
<td>18 ± 12</td>
<td>105 ± 35</td>
<td>512 ± 91</td>
</tr>
<tr>
<td>6</td>
<td>JC3528</td>
<td>4</td>
<td>2</td>
<td>Small + Large</td>
<td>mms1</td>
<td>13 ± 8</td>
<td>179 ± 25</td>
<td>347 ± 68</td>
</tr>
<tr>
<td>7</td>
<td>JC3531</td>
<td>4</td>
<td>2</td>
<td>Small + Large</td>
<td>msh2</td>
<td>129 ± 58</td>
<td>151 ± 12</td>
<td>145 ± 70</td>
</tr>
</tbody>
</table>

* Each determination in experiments 1–4 started with individual colonies of parent strains; experiments 5–7 started with cells from regions of patches that displayed low Ura⁺ backgrounds.

† Number of independent determinations.

‡ Heterologies within and nearby ura3. JC3517-13 and JC3525 also have the large HIS3-telV marker 110 kbp upstream of ura3.

§ Mismatch repair status.

‖ Average uninduced frequencies ± SD. Includes Ura⁺, Ura⁻, and Ura⁺⁻⁻⁻⁻ recombinants. Because background recombination levels were measured by scoring 350–600 colonies per determination, the fivefold higher background in msh2 resulted from a difference of ~1 vs. ~5 recombinants per determination. Although experiments 5–7 were started with patches with low Ura⁺ backgrounds, Ura⁻ background is not detectable and approximately two-thirds of the msh2 background was due to Ura⁺ products.

¶ Average induced frequencies ± SD. For each determination an average of 400–1500 colonies were scored.
TABLE 3

Conversion tract directionality

<table>
<thead>
<tr>
<th>Strain</th>
<th>Heterologies</th>
<th>( n^a )</th>
<th>Bidirectional</th>
<th>HO only</th>
<th>5'</th>
<th>3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>DY3515-13</td>
<td>Small</td>
<td>57</td>
<td>89</td>
<td>0</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>SW3476</td>
<td>Small + large</td>
<td>87</td>
<td>90</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\) Number of products with continuous tracts.

\(^b\) Unidirectional tracts include those <53 bp in length (conversion of HO site only) plus those extending either 5' or 3' from the HO site. Data for DY3515-13 are from Nickoloff et al. (1999).

Figure 3.—Percentage conversion of individual markers among recombinant products as a function of distance from the DSB. Conversion tract spectra from products with continuous tracts (not shown) were generated from 57 products of DY3515-13 with small heterologies (data from Nickoloff et al. 1999) and 87 products of SW3476 with large heterologies (this study). The two large heterologies are shown by shaded symbols. \( P \) values are given for small and large heterologies at comparable distances from the DSB (Fisher exact tests).

TABLE 4

Percentage of HIS3-telV retention, loss, and sectoring

<table>
<thead>
<tr>
<th>Strain</th>
<th>Heterologies</th>
<th>% of products</th>
</tr>
</thead>
<tbody>
<tr>
<td>JC3517-13</td>
<td>Small</td>
<td>93.4</td>
</tr>
<tr>
<td>JC3525</td>
<td>Small + large</td>
<td>90.8</td>
</tr>
</tbody>
</table>
most of these products were induced by DSBs.) Notably, Leu\(^{+/-}\) sector rates were essentially the same in the presence and absence of small heterologies, indicating that the Pms1p- and Msh2p-dependent repair of large loop mismatches does not depend on corepair with small heterologies.

**DISCUSSION**

This study provides new insight into the efficiency of incorporation of large heterologies into hDNA, the efficiency of large loop mismatch repair, and the proteins involved in this repair. Our analysis was performed in mitotic cells, but our conclusions also apply to meiotic conversion (see below). Numerous studies support the notion that conversion of small heterologies (single-base differences and insertions <40 bp in length) in both meiosis and mitosis is mediated by mismatch repair of hDNA (Petes et al. 1991; Nickoloff and Hoekstra 1998; Weng and Nickoloff 1998; Nickoloff et al. 1999). Furthermore, heterologies >1 kbp convert readily with little PMS in meiosis. Thus, one might conclude that meiotic cells are capable of incorporating large heterologies into hDNA and efficiently repairing large loop mismatches. However, this conclusion was inconsistent with plasmid transformation assays that indicated that small loop (1–12 bases) and most single-base mismatches are repaired with >90% efficiency, but 38-base loops were repaired with only 30–50% efficiency (Bishop and Kolodner 1986; Bishop et al. 1987; Kramer et al. 1989; Luhr et al. 1998). Furthermore, the repair of single-base mismatches and loop mismatches ≤12 bases involved Pms1p and Msh2p, among other proteins, yet the low efficiency repair of a 38-base loop mismatch was independent of Pms1p and showed limited dependence on Msh2p (Bishov et al. 1987; Kramer et al. 1989; Luhr et al. 1998). Similarly, the repair of 1- and 7-base loop mismatches arising by replication slippage was dependent on Pms1p and Msh2p (Tran et al. 1996), whereas repair of loop mismatches >15 bases was independent of these proteins (Tran et al. 1996; Sia et al. 1997; Harfe and Jinks-Robertson 1999). The apparent lack or poor repair of loop mismatches >15 bases in length and the lack of involvement of key mismatch repair proteins led to the suggestion that meiotic conversion of very large insertions is not mediated by mismatch repair of hDNA, but instead by gap repair (Szostak et al. 1983; Tran et al. 1996). The high rates of segregation of the large LEU2 heterology in msh2 and pms1 mutants (Figure 4) indicate that most mitotic conversion of large heterologies involves mismatch repair of hDNA and that that this repair is largely dependent on Msh2p and, to a lesser extent, on Pms1p. This conclusion has recently been extended to meiotic conversion: 1.1-kbp heterologies convert during meiosis without PMS in wild-type cells, but PMS is increased in msh2 (but not pms1) mutants (H. Kearney and T. Petes, personal communication). Our data also indicate that heterologies as large as 2.6 kbp are readily incorporated into hDNA during mitotic gene conversion, at least when present in the recipient allele. Although the efficiency of incorporation may differ for large heterologies present in recipient and donor alleles, we believe that this is unlikely because meiotic conversion shows parity; i.e., conversion involving marker gain occurs at a similar rate to marker loss (reviewed in Petes et al. 1991).

Recent results suggest overlap among mismatch repair and nucleotide excision repair pathways. In addition to its role in nucleotide excision repair, Rad1p/Rad10p endonuclease cleaves nonhomologous 3' tails during single-strand annealing and during DSB-induced recombination when invading ends have 3' nonhomologous tails (Paques and Haber 1999), as in the present system. In meiotic cells, about one-half of repair events at a 26-base loop mismatch involve Msh2p, Rad1p, and Rad10p; the remaining repair events are independent of these proteins, or less likely, involve gap repair (Kirkpatrick and Petes 1997). Thus, Rad1p/Rad10p is a structure-specific endonuclease that recognizes both 3' flaps and relatively large loop mismatches. It was recently shown that a 1.1-kbp heterology displayed even higher PMS in rad1 and rad10 than in msh2 (H. Kearney and T. Petes, personal communication). It is likely that Rad1p/Rad10p is also involved in large loop repair in mitosis; we could not test this because these proteins are required to cleave 3' nonhomologous tails.
produced by HO nuclease. Because Msh3p and Msh6p are involved in loop and single-base mismatch repair, respectively, it is also likely that large loops will segregate frequently in msh3, but not msh6, mutants.

Why are large loop mismatches repaired poorly in transformed plasmid substrates, with residual repair being largely independent of Pms1p and Msh2p, while these proteins have significant roles in the efficient repair of large loop mismatches during DSB-induced chromosomal gene conversion? This difference might reflect the substrate context (plasmid vs. chromosomal), although this explanation is inadequate because chromosomal loop mismatches >16 bases produced by replication slippage are not processed by Pms1p, Msh2p, or Msh6p (Tran et al. 1996; Sia et al. 1997; Harfe and Jinks-Robertson 1999). We propose that these disparate results reflect the differential accessibility of one or more components of the Msh2-dependent repair system to mismatches present in artificial (plasmid) hDNA, arising by replication slippage, or arising by recombination. Different mismatch repair complexes may be active during these processes, some of which may or may not contain Msh2p and Msh2p-interacting proteins. Alternatively, there may be a single large loop mismatch repair complex, but its recognition or repair activities may be modulated by interactions with specific factors present at replication forks, such as proliferating cell nuclear antigen (PCNA), or at sites of DSB repair, such as Rad51p. In this regard it is interesting that human MSH2 and MLH1 were recently found in complex with a large number of proteins involved in recombinational repair and replication-associated repair, including BRCA1, ATM, BLM, the RAD50-MRE11-NBS1 complex, and PCNA (Wang et al. 2000). Notably, BRCA1 associates with BRCA2, which in turn associates with RAD51 (Wong et al. 1997; Chen et al. 1998; Moynahan et al. 1999).

Additional questions raised by our present study relate to the mechanism and efficiency of incorporation of large heterologies into hDNA. In vitro and in vivo strand transfer catalyzed by RecA on loop mismatches >1 kb in length (Bianchi and Radding 1983; Lichter and Fox 1984), but incorporation of small or large heterologies requires ATP hydrolysis (Rosselli and Stasiak 1991; Kim et al. 1992). Our results suggest that Rad51p readily incorporates large heterologies into hDNA. Cells with a mutant Rad51p that cannot bind ATP show the same hypersensitivity to the radiomimetic agent methylmethane sulfonate as null rad51 mutants. In contrast, cells with Rad51p that can bind, but cannot hydrolyze, ATP (ATPase mutant), are as resistant to methylmethane sulfonate as wild type (Sung and Stratton 1996). We recently found that a Rad51p ATPase mutant is capable of incorporating heterologies into hDNA in vivo, albeit at reduced efficiency (J. Clikeman and J. Nickoloff, unpublished results). How are large heterologies incorporated into hDNA?

One model suggests that incorporation occurs in a processive manner, with homology tested along the entire length of the heterology. There is in vitro evidence that RecA searches processively (Gonda and Radding 1983). We can envision an alternative (“leap-frog”) model in which large heterologies are bypassed in a single step (Figure 5). In this view, homology is identified discontinuously; once a heterology is encountered, Rad51p promotes a second synapsis event on the other side of the heterology. One can also imagine a hybrid model in which homology is tested at intervals. If heterologies are incorporated in a processive manner, a reduction in conversion may only be detectable once a certain threshold length of heterology is reached. If such a threshold exists, our data suggest that it is >2.6 kbp.

We found that wild-type cells converted the 2.6-kbp LEU2 heterology more frequently than a small heterology at the same locus. This result could reflect either more efficient incorporation of large heterologies into hDNA or differential repair processing of these distinct mismatches. We and others have shown that additional heterologies increase gene conversion tract lengths (Schultes and Szostak 1990; Nickoloff et al. 1999; Vedel and Nicolas 1999), but we interpret these marker effects as a reflection of altered mismatch repair processing rather than more extensive hDNA formation. Large heterologies appear to be incorporated into hDNA at least as efficiently as small heterologies, but is it possible that large heterologies are actually incorporated more efficiently? Indirect evidence in favor of this idea comes from an in vitro study of RecA pairing activity showing increasing pairing between a 151-bp segment and its complement with increasing length of a linked heterologous segment (Gonda and Radding 1983). However, we favor an alternative explanation that is based on the large loop-specific Msh2p-independent repair system identified in yeast nuclear extracts; this system is nick independent and leads to preferential loss of loops (Corrette-Bennett et al. 1999). In our recombination assay, loop loss is scored as conversion and large loop-specific repair would enhance conver-
sion of large heterologies. If the large loop-specific repair system operates independently of Msh2p in vitro, as it does in vitro (Corrette-Bennett et al. 1999), our msh2 results would suggest that 70% of conversions of the 2.6-kbp LEU2 heterology are mediated by the Msh2p system and 30% by the large loop-specific system. It is worth noting that this 30% value is similar to the difference in conversion frequencies of the large LEU2 heterology and small B3′ heterology (23%; Figure 3). Although the conversion frequencies of the 2.2-kbp pUC19 and single-base Ase20 heterologies were not significantly different (P = 0.08; Fisher’s exact test), a trend toward increased conversion of large heterologies is apparent. It is important to note that there is reduced sensitivity for detecting increased conversion of a large heterology located 5′ of the DSB because even small heterologies at this position convert in >90% of products. Thus, at these distances from the DSB it is not possible for large loop-specific repair to enhance conversion as much for a 5′ heterology as for a 3′ heterology.

In conclusion, yeast has evolved systems that allow efficient incorporation of large heterologies into hDNA, and the resulting large loop mismatches appear to be processed by Msh2p/Pms1p-dependent and -independent repair systems. Loop-specific repair systems have also been identified in mammalian cells (Taghian and Nickoloff 1998; Littman et al. 1999); in the case of palindromic loops, repair favors loop retention (Taghian and Nickoloff 1998), but nonpalindromic loops are preferentially lost in plasmid substrates (Weiss and Wilson 1987) and in recombination intermediates (Bill et al. 2001). Why would cells require a large loop-specific mismatch repair system, and why would it evolve with a bias toward loop removal? A reasonable hypothesis is that it functions to eliminate insertion mutations produced by integration of stray DNA fragments via illegitimate recombination, or, perhaps more importantly, to remove integrated viral DNA. A large loop-specific repair system that is not 100% effective may create a balance between removal of potentially harmful DNA insertions and retention of foreign DNA to provide raw material to drive evolution.

We thank E. Alani and R. Kolodner for kind gifts of plasmids, T. Petes and H. Kearney for communicating results prior to publication, T. Petes for helpful comments, and K. Spitz for expert technical assistance. This work was supported by National Institutes of Health grant CA55302 to J.A.N.

LITERATURE CITED


