Mechanisms of Double-Strand-Break Repair During Gene Targeting in Mammalian Cells

Philip Ng* and Mark D. Baker*²

*Department of Molecular Biology and Genetics and ²Department of Pathobiology, University of Guelph, Guelph, Ontario, Canada N1G 2W1

Manuscript received July 24, 1998
Accepted for publication November 10, 1998

ABSTRACT

In the present study, the mechanism of double-strand-break (DSB) repair during gene targeting at the chromosomal immunoglobulin \( \mu \)-locus in a murine hybridoma was examined. The gene-targeting assay utilized specially designed insertion vectors genetically marked in the region of homology to the chromosomal \( \mu \)-locus by six diagnostic restriction enzyme site markers. The restriction enzyme markers permitted the contribution of vector-borne and chromosomal \( \mu \)-sequences in the recombinant product to be determined. The use of the insertion vectors in conjunction with a plating procedure in which individual integrative homologous recombination events were retained for analysis revealed several important features about the mammalian DSB repair process:

1. The presence of the markers within the region of shared homology did not affect the efficiency of gene targeting.
2. In the majority of recombinants, the vector-borne marker proximal to the DSB was absent, being replaced with the corresponding chromosomal restriction enzyme site. This result is consistent with either formation and repair of a vector-borne gap or an “end” bias in mismatch repair of heteroduplex DNA (hDNA) that favored the chromosomal sequence.
3. Formation of hDNA was frequently associated with gene targeting and, in most cases, began ~645 bp from the DSB and could encompass a distance of at least 1469 bp.
4. The hDNA was efficiently repaired prior to DNA replication.
5. The repair of adjacent mismatches in hDNA occurred predominantly on the same strand, suggesting the involvement of a long-patch repair mechanism.

Gene targeting is a powerful technology for precisely altering chromosomal genes in the study of gene structure and function, creating animal models for human genetic diseases and, perhaps ultimately, human gene therapy (Waldman 1992, 1995; Bertling 1995; Guénet 1995; Vega 1995). Unfortunately, in mammalian cells, this power is greatly tempered by the inefficiency of the process. That is, the absolute frequency of gene targeting is generally ~10⁻²/cell for most loci in a number of different cell lines studied. To complicate matters further, random integration of the vector typically outnumbers gene targeting by about 1000:1. Clearly, a better understanding of the mechanism of mammalian gene targeting will help to improve this technology for precise and efficient genome modification.

In yeast, transformation studies have led to the double-strand-break repair (DSBR) model as the proposed mechanism of gene targeting (Orr-Weaver et al. 1981, 1988; Orr-Weaver and Szostak 1983; Szostak et al. 1983). According to the canonical yeast DSBR model, recombination is initiated by a double-strand break (DSB) that is enlarged to form a gap with 3’-OH single-stranded DNA tails by exonuclease activity. One 3’-OH end invades a homologous duplex, anneals with its complementatory sequence to generate a region of asymmetric heteroduplex DNA (hDNA), and primes DNA repair synthesis displacing a D-loop. The D-loop is enlarged and anneals to the second 3’-OH end, forming another region of asymmetric hDNA from which repair synthesis initiates, finally forming two Holliday junctions. Branch migration of the Holliday junctions can generate regions of symmetric hDNA. Resolution of the two Holliday junctions in the opposite planes results in crossover products, while resolution in the same plane results in noncrossover products. A revised version of the DSBR model was later proposed (Sun et al. 1991) to account for observations in yeast that most gene conversion results from mismatch repair (MMR) of hDNA rather than from gap repair (Petes et al. 1991). In the revised DSBR model, the DSB is processed to 3’-OH single-stranded tails (rather than a gap), which then undergo strand invasion with the homologous chromosomal template, and recombination proceeds as above.

Gene targeting in mammalian cells has also been interpreted within the framework of the yeast DSBR model. For example, a DSB in the vector, within the region of
homology to the chromosomal target, greatly enhanced the efficiency of gene targeting (Jasin and Berg 1988; Valancius and Smithies 1991; Hasty et al. 1992) suggesting that homologous DNA ends flanking the DSB were recombinogenic and interacted directly with homologous chromosomal sequences. Insertion vectors bearing a preformed gap within a region homologous to the chromosomal target were repaired using the chromosome as template (Jasin and Berg 1988; Valancius and Smithies 1991; Smithies and Kim 1994). Vector-borne heterologies in close proximity on one or both sides of the DSB were frequently lost following targeted vector integration, suggesting the possibility that the DSB might have been enlarged to form a gap or that formation and repair of hDNA might be associated with vector integration (Hasty et al. 1991, 1995; Pennington and Wilson 1991; Valancius and Smithies 1991; Deng et al. 1993; Donoho et al. 1998; Ng and Baker 1998). Nevertheless, there are many mechanistic details of the gene-targeting process relating to the processing of the vector-borne DSB, the extent and location of hDNA, the efficiency of repair of DNA mismatches, and whether or not there exists a strand bias in repair that we know little about.

To address these questions, we have investigated the mechanism of gene targeting at the chromosomal immunoglobulin μ-locus in a murine hybridoma using specially designed insertion vectors. Insertion vectors are useful in the study of mammalian gene targeting because the site of initiation of homologous recombination (site of the vector-borne DSB) is known and, following vector integration, both participating DNA sequences are available for analysis at the recombinant locus. To exploit the power of this approach, we modified the insertion vector using site-directed mutagenesis to create six diagnostic restriction enzyme markers in the μ-gene constant (Cμ) region of homology to the chromosomal μ-locus. Thus, the vector-borne Cμ region was genetically distinguishable from the endogenous restriction enzyme site that each insertion destroys is indicated in italics, and the novel restriction enzyme site that bears a single copy of the trinitrophenyl (TNP)-specific immunoglobulin μ-gene and synthesizes cytolytic, TNP-specific IgM (Köhler and Shulman 1980; Köhler et al. 1982). The deficiency in igm482 is due to the presence of a 2-bp deletion in the Cμ3 exon that renders the μ-chain nonfunctional. Unlike the wild-type Sp6 hybridoma, the TNP-specific IgM produced by the mutant igm482 hybridoma is nontylocytic; it cannot activate complement-dependent lysis of TNP-coupled sheep red cells. Thus, whereas the wild-type Sp6 hybridoma can be detected as a plaque-forming cell (PFC) in a sensitive TNP-specific plaque assay (ε.o.p. ~0.8 TNP-specific PFC/cell), the mutant igm482 hybridoma cannot (ε.o.p. ~10⁻⁸ TNP-specific PFC/cell; Köhler et al. 1982; Baumann et al. 1985; Baker et al. 1988). The igm482 mutation destroys a wild-type XmnI restriction enzyme site but creates a Xbal site. With the exception of the 2-bp deletion, the mutant igm482 and wild-type Sp6 chromosomal μ-genes are isogenic.

The vectors PCMEN16 and PCMEN15 used in this study are identical to the enhancer-trap and enhancer-positive vectors PCMEN and PCMEN⁺, respectively, described previously (Ng and Baker 1998) except for the following modifications: site-directed mutagenesis (Deng and Nickoloff 1992) was used to alter the 5.8-kb Cμ region segment from the wild-type Sp6 hybridoma by introducing a 4-bp insertion into the endogenous restriction enzyme sites Avall, Sad, AluI, EarI, and Xhel at the noncoding nucleotide positions 239, 645, 1199, 1677, and 2114, respectively (Xbal site of vector linearization defined as Cμ nucleotide position 0). In the destruction of each of the endogenous restriction enzyme sites, a novel site in the vector-borne Cμ region is created (Figure 1A). In addition, the vector-borne Cμ region differs from the mutant igm482 Cμ region (Figure 1B) by the presence of the wild-type sequence in the Cμ3 exon denoted by the XmnI site (Cμ nucleotide position 1586). Therefore, the vector-borne Cμ region bears six diagnostic markers that distinguish it from the corresponding endogenous sites in the recipient mutant igm482 chromosomal Cμ region. The oligonucleotides used in site-directed mutagenesis are listed below. For each oligonucleotide, the particular 4-bp insertion is indicated in boldface, the endogenous restriction enzyme site that each insertion destroys is indicated in italics, and the novel restriction enzyme site that is created is underlined: (1) 5′-CCCTCTCT TGTTGGTACCTCACAGAGTCACTC-3′ (Avall → KpnI); (2) 5′-GCTTGGCCAAAGCTATATGCCATCACCTTGGACC-3′ (Sad → EcoRI); (3) 5′-CCAGGATCTTCCCTTTAAAGTATGGTTGGGGG-3′ (AluI → Dral); (4) 5′-GGCAGGTTCTCCTTACGAGCCACCTATCC-3′ (Earl → AaII); and (5) 5′-GC TATTGACCATGCTTAGTACTGCCTACACCGGC-3′ (Nhel → SacI). Following site-directed mutagenesis, DNA sequencing (Mobix, McMaster University, Hamilton, Ontario, Canada) was performed to verify that only the desired 4-bp insertion was introduced. The vectors PCMEN15 and PCMEN16 were identical except that in PCMEN15, the 372-bp NsiI/NdeI fragment encompassing the SV40 early region enhancer sequence responsible for neo gene expression was deleted from the pSV2neo vector backbone whereas it was retained in PCMEN16. Thus, whereas PCMEN15 is 13.4 kb, PCMEN16 is 13.7 kb. The enhancer-trap feature of PCMEN15 was exploited in recombinant isolation procedure 2 (see below). With the exception of the introduced Cμ region markers, linearization of PCMEN16 and PCMEN15 at the unique Xbal site within the Cμ region provides 1.5 kb and 4.3 kb of overall homology to the chromosomal Cμ region on the 5′ and 3′ sides of the vector cut site, respectively.

**Vector transfer and transformant isolation:** To effect gene targeting, 8.7 pmol of either PCMEN15 or PCMEN16 was linearized within the Cμ region at the unique Xbal site (Figure 1A)
and transferred to $2 \times 10^7$ recipient mutant igm482 hybridomas by electroporation as described (Baker et al. 1988). Following electroporation, the hybridoma culture was resuspended in DMEM (Dulbecco's modified Eagle medium containing 13% bovine calf serum and 5.3 $\times$ 10$^{-4}$ m 2-mercaptoethanol). An average of 30% of the hybridomas survived electroporation as determined by Trypan blue staining. Screening of the surviving hybridomas for targeted recombinants was accomplished by two different procedures.

**Procedure 1:** This method of identifying targeted recombinants was described previously (Baker et al. 1988). It is based upon the ability of the transfer vector bearing the wild-type $C_{\mu}$ region to undergo homologous recombination with the mutant igm482 chromosomal $C_{\mu}$ region, correcting the 2bp mutant igm482 $C_{\mu}$ deletion (mutant igm482 $T\mu$ site is changed to the wild-type $Xmn$ site). Restoration of the wild-type $C_{\mu}3$ exon in the expressed 5' $C_{\mu}$ region restores cytoytic, TNP-specific IgM production in the hybridoma. Targeted hybridomas are identified as TNP-specific PFC in a plaque assay (Baker et al. 1988).

Two electroporations were conducted with the vector pC-En$\mu$ (denoted A and B) while three electroporations were performed using the vector pC-En$\mu$ (denoted I, II, and III). For each hybridoma culture, 2 days post-electroporation, the absolute frequency of TNP-specific PFC in the nonselected recipient population was determined by plaque assay. To determine the frequency of PFC among the G418$^R$ transformants, a portion of each electroporated culture was selected for resistance to G418 (600 $\mu$g of active G418/ml) and subjected to the plaque assay. Transformation frequency was determined by distributing the electroporated culture at densities of 10$^4$, 10$^5$, and 10$^8$ hybridomas/ well in multiwell-tissue-culture plates in DMEM containing G418. Following G418 selection, the number of growth-positive wells was enumerated and the mean frequency of G418$^R$ transformants/hybridoma determined by the Poisson distribution.

In electroporations involving pC-En$\mu$ and pC-En$\mu$, PFCs making TNP-specific IgM were recovered directly from plaque assay plates used to determine the absolute frequency of gene targeting, as described by Baker and Read (1992), and then deposited into DMEM containing 600 $\mu$g/ml of G418 to select for those that were G418$^R$. As shown in a previous study (Ng and Baker 1998), an enhancer-trap vector similar to pC-En$\mu$ provided an ~15-fold enrichment in the frequency of targeted vector integration at the chromosomal locus as a result of expression of the enhancerless neo gene in the integrated vector. Therefore, this feature permitted recombinant hybridomas making normal, TNP-specific IgM also to be isolated directly from the transformed G418$^R$ population in those electroporations involving pC-En$\mu$. Thus, in the plating experiments described above to enumerate the frequency of G418$^R$ transformants, culture supernatant from each growth-positive well was tested for the presence of normal, TNP-specific IgM by spot test (Köhler and Shulman 1980; Baker 1989). Hybridomas in culture wells that reacted positively were cloned at limiting dilution and retested and those making normal, TNP-specific IgM were saved for DNA analysis to verify the gene-targeting event.

**Procedure 2:** Procedure 2 involved a modification of plating procedure 1 described above for use with the enhancer-trap, pC-En$\mu$ vector. The modification permitted the product(s) of an individual integrative gene targeting event to be retained for analysis within a single culture well. As for procedure 1, 8.7 pmol of XbaI-linearized pC-En$\mu$ vector was introduced into $2 \times 10^7$ recipient mutant igm482 hybridomas by electroporation. Two electroporations (denoted IV and V) were performed. To isolate individual gene-targeting events, immediately following electroporation, the hybridoma culture was resuspended in 1188 ml of DMEM and 0.1-ml aliquots of the culture segregated in 96-well microtiter plates. Two days later, each culture well received 0.1 ml of DMEM plus G418 yielding a final concentration of 600 $\mu$g/ml. Following outgrowth of G418$^R$ colonies, the number of growth-positive wells was enumerated. From the Poisson distribution, it was determined that each growth-positive well originated from a single G418$^R$ transformant. To identify targeted recombinants at the chromosomal $\mu$- locus, genomic DNA from each G418$^R$ transformant was analyzed by Southern blot analysis. As in procedure 1, a spot test was used to identify those recombinants making normal, TNP-specific IgM (Köhler and Shulman 1980; Baker 1989).

**Analysis of $\mu$-gene structure:** Hybridoma genomic DNA was prepared by the SDS-proteinase K procedure of Gross-Bel-lard et al. (1973). Restriction enzymes were purchased from Bethesda Research Laboratories (Gaithersburg, MD), New England Biolabs (Beverly, MA), and Pharmacia (Piscataway, NJ) and used in accordance with the manufacturers' specifications. Gel electrophoresis, transfer of DNA onto nitrocellulose membrane, 32P-labelled probe preparation, and hybridization were all performed according to standard procedures (Sambrook et al. 1989).

**PCR analysis:** PCR was used to specifically amplify the 5' and 3' $C_{\mu}$ regions in targeted G418$^R$ recombinants. Oligonucleotide primers (synthesized at Mobix, McMaster University) AB9703 (5'-CTACTTGGAGAAGCCAGGATCTAGG-3') and AB9745 (5'-ACCGGATCTTTACGCTGGTGAG-3') were used to specifically amplify the 5' $C_{\mu}$ region, while primers AB9703 and AB9438 (5'-GGACCATCAGTTCAGCTTTCCA-3') were used for the 3' $C_{\mu}$ region.

The PCR reaction contained the following (per 50$\mu$l volume): 1 $\mu$l of target DNA (1 $\mu$g/ml), 20 mm TrisHCl (pH 8.0), 50 mm KCl, 0.2 mm dNTP, 0.24 $\mu$m of each of the primer pair, and 1.5 mm MgCl2 with a 50-$\mu$l light mineral oil overlay. The reagents were mixed in a 0.5-ml microfuge tube and the solution was heated to 94°C for 2 min in a Perkin Elmer (Norwalk, CT) Thermocycler 680 and then maintained at 80°C, during which time 0.5 $\mu$l of Taq DNA polymerase (5 units/ $\mu$l) was added. PCR was performed for 45 cycles. The parameters for each cycle were as follows: 1 min at 94°C (denaturing), 1 min at 63°C (annealing), and 2 min at 72°C (extension).

A final extension for 5 min at 72°C was included after the last cycle. A total of 2 to 4 $\mu$l of the PCR product was digested with the various restriction enzymes and visualized by ethidium bromide staining following agarose gel electrophoresis.

**RESULTS**

**Experimental system:** Our gene-targeting assay utilized the mutant igm482 hybridoma as a recipient for transfer of the enhancer-trap and enhancer-positive vectors pC-En$\mu$ and pC-En$\mu$, respectively, in which the $C_{\mu}$ region of homology to the chromosomal locus was modified by inclusion of six diagnostic restriction enzyme markers. The basic scheme was as follows: vector DNA (8.7 pmol) was linearized within the $C_{\mu}$ region of homology at the unique XbaI site and transferred to $2 \times 10^7$ recipient mutant igm482 hybridomas by electroporation (Baker et al. 1988). Hybridoma survival averaged ~30%. Targeted recombinants derived by homologous vector integration were isolated by two procedures, following which they were examined to determine the fate of each $C_{\mu}$ region marker. From this analysis, infor-
mation regarding mammalian gene-targeting reaction was obtained.

Isolation of targeted recombinants: Procedure 1: As detailed in materials and methods, isolation of targeted recombinants by procedure 1 was based on the ability of the gene-targeting reaction to correct the 2-bp mutant igm482 C_μ3 deletion in the expressed 5' C_μ region of the hybridoma, restoring production of normal, TNP-specific IgM and permitting its detection as a PFC (Baker et al. 1988). As shown in Table 1, three electroporations with the enhancer-trap pC_μEn_{916} vector (denoted I, II, and III) and two electroporations with the enhancer-positive vector pC_μEn_{916} (denoted A and B) were performed. The absolute frequencies of gene targeting with the enhancer-trap vector pC_μEn_{916} and enhancer-positive vector pC_μEn_{916} were 1.34 × 10^{-3} cell and 1.71 × 10^{-6} cell, respectively, and were similar to the absolute frequencies of gene targeting of 1.74 × 10^{-6} cell and 1.82 × 10^{-9} cell, respectively, that were obtained with the enhancer-trap vector pC_μEn^- and enhancer-positive vector pC_μEn^+, which were used in a previous study (Ng and Baker 1998) and in which the C_μ region bore the completely wild-type sequence. Thus, the presence of the six diagnostic C_μ region markers in pC_μEn_{916} and pC_μEn_{916} had no effect on the absolute frequency of gene targeting at the chromosomal C_μ locus. In contrast, the 28-fold reduction in the frequency of G418^a transformants obtained with the enhancer-trap vector pC_μEn_{916} compared to the enhancer-positive vector pC_μEn_{916} led to a corresponding 21-fold increase in the frequency of PFC among the G418^a transformants. These results are consistent with the enrichment in gene targeting obtained with the enhancer-trap vector pC_μEn^- in our previous study (Ng and Baker 1998). From the electroporations with vectors pC_μEn_{916} and pC_μEn_{916} summarized in Table 1, a total of 23 and 5 G418^a were isolated, respectively.

The chromosomal C_μ gene structure in the G418^a PFC was analyzed to verify homologous vector integration. As shown in Figure 1B, the haploid recipient mutant igm482 C_μ region is present on a 12.5-kb EcoRI fragment. Targeted integration of one copy of the vector by a single reciprocal crossover between the vector-borne and chromosomal C_μ regions will duplicate the C_μ region, converting the endogenous 12.5-kb EcoRI fragment into two EcoRI fragments, one bearing the 5' and the other, the 3' C_μ region (Figure 1C). For vector integration involving the enhancer-trap vector pC_μEn_{916}, the EcoRI fragment bearing the 5' C_μ region is expected to be 16.2 kb, whereas for the enhancer-positive vector pC_μEn_{916}, it is expected to be 16.5 kb (in parentheses). For both vectors, the 3' C_μ region is expected on a 9.6-kb EcoRI fragment. The results of the Southern blot analysis revealed that the G418^a PFC could be placed into three classes according to their C_μ gene structures. Two of the 5 and 14 of the 23 G418^a PFC generated with pC_μEn_{916} and pC_μEn_{916}, respectively, were of the class I type, bearing a recombinant C_μ gene structure depicted in Figure 1C for single copy targeted vector integration. Class II recombinants composed 2/5 and 2/23 of the G418^a PFC generated with pC_μEn_{916} and pC_μEn_{916}, respectively. These recombinants bore, in addition to the two recombinant EcoRI C_μ gene fragments depicted in Figure 1C, EcoRI C_μ gene fragment(s) that corresponded to the unit size of the vector (13.4 kb for pC_μEn_{916} and 13.7 kb for

<table>
<thead>
<tr>
<th>Transfer vector</th>
<th>Electroporation</th>
<th>Frequency of G418^a transformants</th>
<th>Frequency of PFC^c</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>2.46 × 10^{-5}</td>
<td>0.69 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>5.72 × 10^{-5}</td>
<td>1.99 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>6.08 × 10^{-5}</td>
<td>2.45 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>4.75 × 10^{-5}</td>
<td>1.71 × 10^{-6}</td>
</tr>
<tr>
<td>pC_μEn_{916}</td>
<td>A</td>
<td>1.32 × 10^{-3}</td>
<td>1.16 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1.36 × 10^{-3}</td>
<td>1.56 × 10^{-6}</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>1.34 × 10^{-3}</td>
<td>1.36 × 10^{-6}</td>
</tr>
</tbody>
</table>

^a Values are from two and three electroporations of pC_μEn_{916} and pC_μEn_{916}, respectively, into the recipient igm482 hybridoma.

^b The frequency of transformed G418^a cells.

^c The frequency of PFC in the nonselected hybridoma population following electroporation and in the hybridoma population selected for G418 R. The value in parentheses represents a different measurement of the enrichment in gene targeting obtained with the vector (13.4 kb for pC_μEn_{916} and 13.7 kb for pC_μEn_{916}).
Figure 1—Gene targeting at the chromosomal \(\mu\)-locus. (A) As detailed in materials and methods, the enhancer-positive vector \(pC_\mu En_{\mu16}\) and the enhancer-trap vector \(pC_\mu En_{\mu16}\) are identical except that \(pC_\mu En_{\mu16}\) lacks the SV40 early region enhancer. In each vector, site-directed mutagenesis was used to create the indicated vector-borne restriction enzyme sites (solid gray circles), which replace the endogenous restriction enzyme sites in the mutant \(igm482\) chromosomal \(C_\mu\) region (open circles in B). The nucleotide position of the various diagnostic restriction enzymes with respect to the XbaI site of vector linearization is indicated (the dashed line is nucleotide position 0). The wild-type \(XmnI\) site in the vector-borne \(C_\mu\) exon is denoted by the solid gray square. Although not relevant to the present study, each vector bears an enhancerless Herpes Simplex Virus-1 thymidine kinase (tk) gene. (B) The structure of the haploid, recipient mutant \(igm482\) chromosomal \(\mu\)-gene. The positions of the endogenous restriction enzyme sites are indicated by the open circles. (C) The structure of the chromosomal recombinant \(\mu\)-gene following targeted integration of a single copy of the \(pC_\mu En_{\mu16}\) vector. The diagram depicts the positions of the PCR primers and the sizes of the amplified products they produce. The positions of the diagnostic \(C_\mu\) region markers are not indicated because each gene-targeting reaction has the potential to generate a different marker pattern in each recombinant. In both B and C, the fragment sizes that the indicated restriction enzymes generate, along with the position of the various DNA probe fragments, are shown. Probe fragment N consists of adjacent 475-bp and 495-bp \(NheI\) fragments, probe F is an 870-bp XbaI/BamHI \(C_\mu\) specific fragment, and probe G (not shown) is a 762-bp \(PvuII\) fragment from the \(neo\) gene. E, \(EcoRI\); H, \(HaeIII\); VHTNP, TNP-specific heavy-chain variable region; \(C_\mu\), \(\mu\)-gene constant region; \(neo\), neomycin phosphotransferase gene; tk, HSV-1 thymidine kinase gene. The figures are not drawn to scale.

\(pC_\mu En_{\mu16}\) consistent with targeted integration of \(>1\) copy of the vector in tandem. The remaining \(G418^8\) PFC were of class III, having an unexpected chromosomal \(\mu\)-gene structure suggesting that homologous recombination did not occur by the expected mechanism of targeted vector integration. Figure 2 presents the Southern blot analysis of \(EcoRI\)-digested genomic DNA analyzed with \(C_\mu\) probe F for representative class I and class II \(G418^8\) PFC.

Procedure 2: As described in materials and methods, procedure 2 consisted of a slight modification to procedure 1 that permitted the products of individual vector integration events involving the enhancer-trap \(pC_\mu En_{\mu16}\) vector to be retained in a single culture well. \(G418^8\) recombinants were isolated from two electroporation IV and V. Of the 11,616 wells plated following electroporation IV, 450 generated \(G418^8\) colonies whereas, for electroporation V, 11,712 wells were plated and, of these, 389 generated \(G418^8\) colonies. From the number of growth-negative wells and the Poisson distribution, electroporations IV and V yielded mean values of 0.04 and 0.03 \(G418^8\) cells/well, respectively. Thus, each \(G418^8\) growth-positive well originated from a single \(G418^8\) transformant. Genomic DNA was prepared from each of the 839 \(G418^8\) colonies, digested with \(EcoRI\), and screened by Southern blot analysis. Recombinants arising from targeted vector integration were identified according to the diagnostic \(EcoRI\) fragments described in procedure 1 above.

Of the 839 independent \(G418^8\) transformants ana-
Figure 2.—Analysis of μ-gene structure in targeted recombinants. Genomic DNA from the representative hybridomas was digested with EcoRI, electrophoresed through a 0.7% agarose gel, and hybridized with 32P-labelled probe F. The blot presents representative class I recombinants in which a single vector has integrated by homologous recombination into the mutant igm482 chromosomal μ-locus. In addition, a representative class II recombinant (29-1) is also shown in which >1 copy of the transfer vector has integrated in tandem into the mutant igm482 chromosomal μ-locus. The positions of bands of interest are presented on the left of the blot while the positions of relevant marker bands are indicated on the right.

Figure 3.—Restoration of the XbaI site of vector linearization in the targeted recombinants. Genomic DNA from the indicated hybridomas was digested with the combination HaeII/XbaI, electrophoresed through a 0.7% agarose gel, blotted to nitrocellulose, and hybridized with 32P-labelled μ-specific probe N. The blot presents representative class I recombinants. The positions of bands of interest are on the left of the blot and the positions of relevant marker bands are on the right.

lyzed, 21 arose by homologous recombination between the transfer vector and the chromosomal immunoglobulin μ-locus: 17 recombinants were of class I (Figure 1C) and 4 recombinants were of class II. Five G418R recombinants belonging to class III were also recovered. Random vector integration(s) into the hybridoma genome accounted for the remaining 813 G418R transformants (results not shown). The presence and proportion of the three classes of targeted recombinants is consistent with the results obtained with procedure 1 above as well as with the results of our previous gene-targeting study at the chromosomal immunoglobulin μ-locus (Ng and Baker 1998). To study the mechanism of gene targeting, class I recombinants from isolation procedures 1 and 2 were analyzed further because of their relatively simple recombinant μ-gene structure, which facilitated both ease of analysis and interpretation.

Class I recombinants do not contain random vector integrations: A common feature of all the class I recombinants was that only the fragment sizes expected from the recombinant μ-locus were present with no extraneous band(s) being detected by Southern blot analysis (Figure 2; see also Figure 3). This was confirmed more directly by Southern blot analysis of EcoRI-digested class I recombinant genomic DNA electrophoresed under conditions in which DNA fragments ≥300 bp were retained for transfer onto nitrocellulose. Again, through the use of both chromosome and vector-specific probe fragments (i.e., μ-specific-probe N and neo probe G, respectively; data not shown), no fragments other than those from the recombinant μ-locus (Figure 1C) were visible. This suggested that gene targeting at the immunoglobulin μ-locus was not accompanied by random vector integration, a result consistent with our previous studies (Baker et al. 1988; Ng and Baker 1998) and those of others at different loci in other cell lines (Bollag et al. 1989; Waldman 1992, 1995; Bertling 1995).

The site of vector linearization is restored in class I recombinants: Another feature common to all the class I recombinants was that the XbaI site used for vector linearization was restored following targeted vector integration as depicted in Figure 1C. This was revealed by Southern blot analysis in which restoration of the XbaI site was expected to generate a 4.4-kb and a 2.8-kb HaeII-XbaI fragment from the 5’ and 3’ Cμ regions, respectively (Figure 1C). The results of this analysis for representative recombinants are presented in Figure 3.

Determination of restriction enzyme marker patterns in class I recombinants: The identity of the restriction enzyme marker at each of the six positions in both the 5’ and 3’ Cμ regions of the class I recombinants was examined to determine whether it was derived from the
transfer vector or the chromosome. This was accomplished by using PCR to specifically amplify a 4.8-kb product from the 5' C_\mu_ region and a 4.6-kb product from the 3' C_\mu_ region (Figure 1C). Following amplification, the 5'- and 3'-C_\mu_ -region PCR products were digested separately with each of the six pairs of diagnostic restriction enzymes and the products analyzed by gel electrophoresis. Restriction maps of the 5'- and 3'-C_\mu_ regions, together with the diagnostic fragment sizes expected following digestion with each of the vector (V) and chromosome (E) specific enzymes, are presented in Figure 4, A and B, respectively.

As an example, Figure 5, A and B, present the results of this analysis with the enzyme pair NheI/ Scal for the 5' C_\mu_ region of G418^-targeted recombinants generated from electroporation V. As shown in Figure 5A, the PCR product amplified from recombinants 3-1, 31-1, and 83-2 was completely sensitive to NheI, generating the diagnostic C_\mu_ fragments of 2.6 kb and 2.2 kb expected from Figure 4A. Thus, in these recombinants, the chromosomal NheI marker was present in the 5' C_\mu_ region. This was confirmed following digestion with Scal where the presence of the full-length 4.8-kb PCR product in these recombinants was indicative of complete resistance to cutting (Figure 5B). For recombinants 102-2, 48-4, and 114-1, complete resistance to digestion with NheI was indicated by the full-length 4.8-kb PCR product (Figure 5A), suggesting that the Scal marker was present in the 5' C_\mu_ region of these recombinants. Indeed, complete sensitivity to Scal digestion was revealed in these recombinants by the presence of the diagnostic 2.6- and 2.2-kb C_\mu_ region fragments (Figure 5B). A different outcome was observed for recombinant 27-1, where digestion of the 4.8-kb PCR product from the 5'
as did digestion with ScaI, ensured that all recombinants were independent isolates. As shown in Table 2, for all recombinants except 27-1, the restriction enzyme site marker in each Cμ position was derived from either the vector or the chromosome. All recombinants that contained the wild-type XmnI site in the Cμ3 exon of the expressed 5′ Cμ region produced cytolytic TNP-specific IgM. However, this analysis revealed that four recombinants, 19/2, (25/4), (60/7), and 83-2, produced cytolytic TNP-specific IgM even though they retained the 2-bp mutant Igm482 deletion (TfI site) in the 5′ Cμ region. The importance of this observation will be discussed below.

With respect to recombinant 27-1, in addition to the partial cutting observed at nucleotide position 2114 in the 5′-Cμ-region PCR product (presented above), nucleotide position 1199 in the 3′-Cμ-region PCR product also revealed partial susceptibility to digestion with the diagnostic restriction enzyme pair AflII/DraI (data not shown). These results indicated that recombinant 27-1 did not consist of a pure population. To investigate this further, recombinant 27-1 was cloned and 19 subclones, each originating from a single cell, were isolated. Southern blot analysis of EcoRI-digested genomic DNA with Cμ-specific probe fragment F confirmed that each of the 19 subclones retained the Cμ region duplicate of the parental 27-1 hybridoma (Figure 1C; data not shown). For each 27-1 subclone, the 5′-Cμ-region PCR product was digested separately with NheI and Scal while the 3′-Cμ-region PCR product was digested separately with AflII and DraI and the products analyzed by gel electrophoresis. The results (data not shown) revealed that the parental 27-1 recombinant was in fact composed of two subpopulations in approximately equal proportion: 7/19 subclones fell into subpopulation A (Table 2, electroporation V) possessing the vector-borne Scal marker in the 5′ Cμ region and the vector-borne Scal marker in the 3′ Cμ region, while the remaining 12/19 subclones fell into subpopulation B (Table 2, electroporation V) in which the chromosomal NheI marker was present in the 5′ Cμ region and the chromosomal AflII marker, in the 3′ Cμ region.

Possible outcomes of the gene-targeting reaction: Before considering the results in Table 2, we first present four possible outcomes of a gene-targeting reaction in which participating Cμ regions bear different markers at homologous positions (vector-borne marker is gray; chromosome marker is white) together with the likely mechanism(s) involved in generating the pattern. In Figure 6A, the vector-borne marker is present in the 5′ Cμ region while the chromosomal marker is in the 3′ Cμ region. This pattern is expected as a result of a single, reciprocal crossover between the Cμ regions 5′ of the marker. However, it might also be generated as a consequence of MMR of hDNA formed during vector integration. In Figure 6B, the chromosomal marker is present in both Cμ regions. This situation is predicted as a consequence of the repairing of a gap that had removed the vector-borne marker but might also arise
TABLE 2
Analysis of restriction enzyme markers in the 5' and 3' Cµ regions of recombinant hybridomas

<table>
<thead>
<tr>
<th>Electroporation</th>
<th>Recombinant</th>
<th>5' Cµ. nucleotide position (bp)</th>
<th>3' Cµ. nucleotide position (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>239</td>
<td>1199</td>
</tr>
<tr>
<td>I</td>
<td>19/2</td>
<td>645</td>
<td>1199</td>
</tr>
<tr>
<td>(25/4,60/7)</td>
<td>AvaI</td>
<td>Sad</td>
<td>AflII</td>
</tr>
<tr>
<td>19/13</td>
<td>AvaI</td>
<td>Sad</td>
<td>AflII</td>
</tr>
<tr>
<td>A2/10</td>
<td>AvaI</td>
<td>BioI</td>
<td>EcoRII</td>
</tr>
<tr>
<td>(A3/1, A5/4</td>
<td>AvaI</td>
<td>EcoRII</td>
<td>XmnI</td>
</tr>
<tr>
<td>A6/4)</td>
<td>AvaI</td>
<td>EcoRII</td>
<td>XmnI</td>
</tr>
<tr>
<td>II</td>
<td>1/16</td>
<td>38/1</td>
<td>49/9</td>
</tr>
<tr>
<td>(38/1, 49/9)</td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td>33/6</td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td>III</td>
<td>5/5-10</td>
<td>(5/7-8)</td>
<td>(5/7-8)</td>
</tr>
<tr>
<td>A</td>
<td>5/1</td>
<td>20±1</td>
<td>AflII</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
<td>58±2</td>
<td>AflII</td>
</tr>
<tr>
<td>IV</td>
<td>58-2</td>
<td>105-3</td>
<td>AflII</td>
</tr>
<tr>
<td></td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td></td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td></td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td></td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td>V</td>
<td>3-1</td>
<td>25-3</td>
<td>AflII</td>
</tr>
<tr>
<td></td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td></td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
<tr>
<td></td>
<td>AvaI</td>
<td>BioI</td>
<td>OhiI</td>
</tr>
</tbody>
</table>

Chromosomal markers from the endogenous mutant igm482 Cµ region are indicated in normal type face while those in boldface indicate vector-borne markers from the Cµ region of the transfer vector. The indicated Cµ nucleotide position defines the location of the various markers relative to the XbaI site of vector linearization. For isolation procedure 1, recombinants A2/10 and (A3/1, A5/4, A6/4) from electroporation I and recombinants 5/1 and 1/3 from electroporations A and B, respectively, were isolated directly from plaque assay plates as TNP-specific PFC (Baker et al. 1988). The remaining recombinants isolated by procedure 1 were identified as a result of screening culture supernatant for TNP-specific IgM by spot test (Köhler and Shulman 1980; Baker 1989). For isolation procedure 2, all recombinants were identified as described in the materials and methods.
Figure 6.—(A-D) Analysis of gene-targeting mechanisms. This figure presents four possible outcomes of a gene-targeting reaction in which the participating regions of homology bear different markers at homologous positions. Vector-borne markers, gray circles; chromosomal markers, open circles.

Figure 7.—Summary of the estimated minimum hDNA tract length in targeted recombinants. This figure presents the positions of the various Cμ region markers in relation to the XbaI site of vector linearization (nucleotide position 0 is in bold) and the estimated minimum extent of hDNA that was observed across the Cμ region in the various recombinants indicated.

as a result of MMR of hDNA. Figure 6C presents the situation in which the vector-borne marker is present in both Cμ regions, a pattern consistent with MMR of hDNA. In Figure 6D, the chromosomal marker is present in the 5′ Cμ region and the vector-borne marker is present in the 3′ Cμ region. This marker pattern is consistent with MMR of hDNA. However, according to the DSBR model, it would also be consistent with rightward migration of the leftward Holliday junction prior to MMR. Of the 33 recombinants analyzed (Table 2), only 2, 25-3 and 114-1, contained the complete set of vector-borne markers in the 5′ Cμ region and the corresponding chromosomal markers in the 3′ Cμ region. This marker pattern can be explained simply on the basis of a single, reciprocal crossover at, or near, the site of vector linearization, 5′ of the marker proximal to the DSB (at nucleotide position 239). The remaining 31 recombinants bore different Cμ marker patterns and are considered in the next sections in relation to the mechanism of gene targeting.

Frequent loss of the vector-borne marker proximal to the DSB: A striking feature common to the majority (31/33) of the recombinants was that in both the 5′ and 3′ Cμ regions, the marker closest to the XbaI site of vector linearization was the chromosomal Avall restriction enzyme site. This indicated that the vector-borne KpnI marker was frequently lost as a consequence of gene conversion during the DSB repair event that resulted in vector integration.

Formation of hDNA during mammalian gene targeting: As indicated above, 2/33 recombinants (25-3 and 114-1) bore a marker pattern in which the 5′ Cμ region contained vector-borne markers, while the 3′ Cμ region contained chromosomal markers consistent with a crossover occurring at or near the site of the vector-borne DSB. From Table 2, a further eight recombinants [1/16, (38/1, 49/9), 45-5, 17-1, 84-6, 20-1, and 79-1] bore a similar marker pattern. However, in some (45-5, 17-1, and 84-6), vector-borne sites positioned proximal to the DSB (including KpnI) were replaced by the corresponding chromosomal sites, suggesting that gene conversion of markers had also occurred prior to the crossover event that integrated the transfer vector (Figure 6A).

Evidence in support of the idea that the gene conversion resulted from MMR of hDNA formed during the gene-targeting reaction was obtained from 19 recombinants (Table 2, 19/2, (25/4, 60/7), 19/13, 83-2, 102-2, 48-4, and 27-1). In these recombinants, either the same vector-borne marker was present in equivalent positions in both Cμ regions, as depicted in Figure 6C, or a chromosomal marker was present in the 5′ Cμ region, while the corresponding vector-borne marker was present in the equivalent position in the 3′ Cμ region as illustrated in Figure 6D. From the number of markers exhibiting these patterns in each recombinant, an estimate of the minimum length of hDNA that might have been formed across the Cμ region during gene targeting was derived. As shown in Figure 7, in most recombinants, evidence for hDNA was observed beginning at the second marker position, 645 bp from the site of DSB and in some cases extended for at least 1469 bp on the one side of the DSB examined. However, both the number of recombinants with hDNA intermediates as well as the measurement of hDNA length likely represent underestimates given the difficulty in quantifying hDNA in the company of a functional MMR system and gap repair. For example, recombinants 58-2, 105-3, 3-1, and 31-1 bear a Cμ region marker pattern that might be interpreted on the basis of MMR of hDNA and/or gap repair involving the entire Cμ region.

As indicated in Table 2, an interesting marker pattern
was observed in recombinant 27-1. This recombinant was actually composed of two cell populations in approximately equal proportion that differed for markers in two Cμ region positions. The mitotic sectoring observed in recombinant 27-1 can be explained on the basis of a failure to completely repair the hDNA intermediate prior to DNA replication and cell division. As explained earlier, procedure 2 ensured that all products of integrative recombination were retained for analysis in individual culture wells. Therefore, mitotic sectoring as observed in the case of recombinant 27-1 could only be detected in recombinants isolated by this procedure.

With the exception of recombinant 27-1, all other recombinants isolated by procedure 2 in which evidence for hDNA was obtained were pure cultures in which a given Cμ region marker position contained either a vector-borne or a chromosomal restriction enzyme site. This suggested that the majority of hDNA generated during mammalian gene targeting was efficiently repaired prior to DNA replication. For any given marker position within hDNA, repair toward both vector and chromosomal sequences was observed, suggesting an absence of bias in MMR.

For several recombinants in Table 2 [19/13, A2/10, (A3/1, A5/4, A6/4), 33/6, 5/5-10, (5/7-8), 5/1, 1/3, 111-2, 118-2, 102-2, and 27-1], the presence of the same vector-borne marker in equivalent positions in both the 5' and 3' Cμ regions suggests that hDNA formation during mammalian gene targeting was frequently symmetrical across the Cμ region.

Adjacent markers within hDNA were repaired predominately in the same direction: From Table 2 and Figure 7, the results suggested that in 18 recombinants [19/2, (25/4, 60/7), 19/13, A2/10, (A3/1, A5/4, A6/4), 33/6, 5/5-10, (5/7-8), 5/1, 1/3, 111-2, 118-2, 102-2, and 27-1] hDNA had spanned at least two markers. Therefore, in these recombinants, the question of whether adjacent markers were repaired in either the same or opposite directions could be addressed. Examination of the marker patterns revealed, in general, continuous tracts of either the vector-borne or chromosomal markers, suggesting that adjacent sites were repaired predominately in the same direction. However, repair of adjacent markers in opposite directions was also apparent, albeit less frequently. This was evident by the punctuation of a continuous tract of vector-borne markers with a chromosomal marker or vice versa, as seen in ten recombinants [A2/10, (A3/1, A5/4, A6/4), 33/6, 5/1, 111-2, 118-2, 48-4, and 27-1].

**DISCUSSION**

In this study, we examined the mechanism of gene targeting in a murine hybridoma utilizing insertion vectors in which the Cμ region of homology bore six novel restriction enzyme markers that allowed it to be distinguished from the haploid, chromosomal immunoglobulin μ-gene target locus. The insertion sites for the various diagnostic restriction enzyme markers were planned so as to minimize disruption of the otherwise perfect sequence homology shared between the vector-borne and chromosomal Cμ region. This was deemed important because at least 132 bp of continuous sequence homology is required for intrachromosomal homologous recombination in mammalian cells (Waldman and Liskay 1988). As was evident from our results, the Cμ region markers in pCμEn16 and pCμEn16 had no effect on the absolute frequency of gene targeting at the chromosomal μ-locus.

With the exception of the Cμ region markers, both insertion vectors bore 1.5 kb and 4.3 kb of overall sequence homology to the chromosomal μ-gene on the 5’ and 3’ sides of the DSB, respectively. The μ-gene structure and Cμ region marker patterns in the class I recombinants indicated that this amount of bilateral homology was sufficient to mediate correct repair of the DSB. Thus, the generation of the class I recombinants can be explained by the two-sided invasion, DSBR model for homologous recombination (Orr-Weaver et al. 1981; Szostak et al. 1983; Szostak et al. 1983) rather than recombination models invoking one-sided invasion, which can result in loss or gain of sequence information (Ellis and Bernstein 1989; Aratani et al. 1992; Bernstein et al. 1992; Belmaaza and Charrand 1994; Ferguson and Holloman 1996).

The examination of markers in the 5’ and 3’ Cμ regions of every recombinant revealed important information about the mammalian gene-targeting reaction, including the processing of the vector-borne DSB, hDNA formation, and mismatch repair (MMR), which are discussed below. However, first, an apparent difference between the Cμ region marker patterns in recombinants isolated by procedures 1 and 2 deserves comment. Unlike procedure 1, procedure 2 involved segregating the transformants immediately following electroporation, thus enabling the products of individual integrative recombination events to be retained for analysis in individual culture wells. Also, whereas procedure 1 recombinants were identified as PFC synthesizing cytolytic, TNP-specific IgM, procedure 2 identified recombinants on the basis of an alteration in the structure of the haploid, recipient chromosomal μ-gene identified through Southern blot screening. This explains why procedure 2 recovered recombinants that retained the mutant igm482 Tfi site in the 5’ Cμ3 exon, whereas procedure 1 did not. However, it does not explain why a predominance of vector-borne markers exists in both the 5’ and 3’ Cμ regions of PFC isolated by procedure 1, while in those recombinants isolated by procedure 2 that were subsequently identified as being PFC (17-1, 84-6, 111-2, 118-2, 20-1, 79-1, 25-3, 83-2, 102-2, and 114-1) the 5’ Cμ region bears predominantly vector-borne markers and the 3’ Cμ region bears chromosomal markers. Nevertheless, the following results suggest that
these differences are more apparent than real. First, the TNP-specific spot test (Köhler and Shulman 1980; Baker 1989) was used in both procedures 1 and 2 to aid in identification of recombinants producing cytolytic, TNP-specific IgM. Therefore, different Cμ marker patterns in the recombinants cannot be attributed to differences in their methods of identification. Second, in an earlier gene-targeting study (Ng and Baker 1998) in which recombinants derived by homologous integration of the wild-type vector-borne Cμ region into the mutant igm482 chromosomal Cμ region were isolated as TNP-specific PFC according to procedure 1, no bias for the wild-type or mutant igm482 sequence in the Cμ3 exon in either the 5′ or the 3′ Cμ region was observed. Therefore, in the present study, we conclude that isolation procedures 1 and 2 do not influence the outcome of homologous recombination.

In the majority of the recombinants in Figure 6 (31/33), the vector-borne KpnI site closest to the DSB was lost in favor of the chromosomal Avall site. This result is consistent with two possibilities related to the processing of the vector-borne DSB: (i) the chromosomal Avall site was introduced as a result of MMR of hDNA with an end-bias toward the sequence of the unbroken (chromosomal) strand (Gilbertson and Stahl 1996; Weng and Nickoloff 1998) or (ii) the DSB was enlarged to form a gap that encompassed the KpnI site, following which gap repair, using the homologous chromosomal sequence as template resulted in the introduction of the Avall site as predicted by the canonical DSBR model of homologous recombination (Orr-Weaver et al. 1981; Szostak et al. 1983). With regard to the former possibility, our data is not consistent with a single MMR tract favoring the chromosomal strand as, for more distant markers, MMR toward vector-borne sequences was also observed (see below). Thus, if an end-bias in MMR exists, it would have to decrease with distance. While we cannot rule out MMR of hDNA as a possible mechanism, our data would appear to be more consistent with processing of the vector-borne DSB to at least a small gap, probably not extending much further past the KpnI site. Support for the lack of extensive gap formation in the class I recombinants was also obtained from our Southern blot analyses, as any gap that had formed on the 5′ side of the DSB and extended into tk gene sequences would have no homologous template available for repair and would be expected to generate a μ-gene deletion following vector integration, an event that was not observed. In addition, previous studies in mammalian cells suggest that extensive end degradation of transferred vector DNA does not occur in either targeted or random transformants (Shulman et al. 1990; Hasty et al. 1992; Jiang et al. 1992; Pfleiffer et al. 1994; Richard et al. 1997). The above mechanisms could also explain removal of vector-borne heterologies at, or close to, the DSB in other mammalian gene-targeting studies (Jasin and Berg 1988; Pennington and Wilson 1991; Valancius and Smithies 1991; Hasty et al. 1992, 1995; Deng et al. 1993; Smithies and Kim 1994).

The marker analysis suggested that hDNA was formed during mammalian gene targeting (Figure 7). In most cases, hDNA was estimated to first begin at least 645 bp from the DSB and to encompass a distance of at least 1469 bp on the one side of the DSB that was examined. The presence of the same vector-borne marker at equivalent positions in both the 5′ and 3′ Cμ regions in several of the recombinants suggested that hDNA frequently formed symmetrically adjacent to the DSB. The formation of symmetric hDNA is expected if Holliday junction branch migration extended into the adjacent homologous Cμ region DNA. In yeast gene targeting (Orr-Weaver et al. 1988), hDNA tracts of at least several hundred base pairs were formed adjacent to the DSB. However, of the 15 recombinants analyzed in that study, the formation of hDNA between the vector and chromosome resulted in transfer of only the chromosomal marker to the vector. Evidence for transfer of the vector-borne marker to the chromosome was not obtained. With the assumption that MMR was unbiased, the authors interpreted their results to mean that hDNA was largely asymmetric and that extensive symmetric hDNA was not formed. In contrast, as indicated above, our results suggested that symmetric hDNA was formed frequently during mammalian gene targeting. Because the marker used in the yeast study was located very close to the vector-borne DSB (only 308 bp away), it might often have resided within the region of asymmetric hDNA formed as a result of pairing between an overhanging plasmid end with the chromosome. Thus, symmetric DNA, if formed, may not have been detected. In contrast, most of the Cμ region markers utilized in our study were located further from the DSB, affording a better opportunity for detection of symmetric hDNA. Another possible explanation for the apparent absence of symmetric hDNA observed by Orr-Weaver et al. (1988) may be that, for the markers used, their assumption that MMR repair was unbiased was incorrect. In addition, it is also possible that, while the mechanism of gene targeting in yeast and mammalian cells shares an overall similarity, there may be subtle differences in detail.

Strong support for extensive gene conversion adjacent to the DSB in a number of recombinants was evident from examination of the Cμ region marker patterns in Table 2. The results suggested that gene conversion resulted from MMR of a hDNA intermediate. With the exception of recombinant 27-1, which consisted of a sectored colony that likely resulted from replication of an incompletely repaired hDNA intermediate, it was evident that mismatches present in the hDNA of other recombinants were all efficiently repaired prior to DNA replication. Although the pathways of recombination may differ, this result agrees with MMR of hDNA generated during extrachromosomal
recombination in Chinese hamster ovary cells (Deng and Nickoloff 1994; Miller et al. 1997).

With the possible exception of the replaced vector-borne KpnI site, the Cµ region marker pattern in many recombinants revealed that MMR of hDNA could occur toward either vector-borne or chromosomal sequences and that it frequently occurred continuously in the same direction. Thus, during mammalian gene targeting, a long-patch MMR mechanism is involved in the repair of DNA mismatches as reported previously in both eukaryotic and prokaryotic cells (Modrich 1993; Modrich and Lahue 1996). Nevertheless, a long-patch MMR mechanism was not absolute, as we also observed occasionally that continuous conversion tracts could be interrupted by a marker from the opposite strand. Discontinuity in MMR has also been reported in other mammalian gene-targeting studies (Brinster et al. 1989; Steeg et al. 1990), although, in contrast to our results, repair of closely linked mismatches in opposite directions appeared to predominate.

Our finding of extensive conversion tracts contrasts with a previous gene-targeting study (Elliot et al. 1998) in which a mutant neo gene stably integrated in the genome of embryonic stem (ES) cells was corrected by recombination with a transferred neo gene fragment modified to contain 5 to 8 silent markers. In that study, marker analysis revealed relatively short conversion tracts in the G418+ recombinants with 80% being ≤58 bp. In fact, the longest tract (511 bp) was seen in only 2.5% of the recombinants. One main difference between the studies, which may account for the variance in tract length, is that targeted recombinants in our study resulted from reciprocal exchange between the vector and the chromosome whereas, in the study by Elliot et al. (1998), the recombinants resulted exclusively from nonreciprocal exchange between the vector and chromosome. Indeed, studies of intrachromosomal homologous recombination in mammalian cells (Godwin and Liskay 1994) as well as mitotic recombination in yeast (Carpenter 1984; Ahn and Livingston 1986; Aguilera and Klein 1989a,b; Vincent and Petes 1989) suggested that longer conversion tracts were often resolved as crossovers but short conversion tracts were not.

A number of recombinants bore the same vector-borne marker at equivalent positions in both the 5' and 3' Cµ regions, suggesting MMR of symmetric hDNA. This result differed from previous studies of gene targeting at the hprt locus in murine ES cells where loss of a chromosomal marker and gain of a vector-borne marker was either not observed (Valancius and Smithies 1991; Deng et al. 1993) or observed at low frequency (Hasty et al. 1995). Differences in the location of vector-borne markers with respect to the DSB can likely explain this discrepancy. In our study, the various markers were spaced at regular intervals across the Cµ region enabling the pattern of hDNA repair to be accurately assessed while, in the latter studies, the markers were located either within close proximity to the DSB or at a great distance. As indicated above, vector-borne markers in close proximity to the DSB may frequently be lost by gap repair or may be encompassed in asymmetric hDNA while markers too distant from the DSB may fail to be encompassed within hDNA. It might also be that there are subtle differences in the mechanistic details of gene targeting at the two target loci, the nature of the cell lines, or the mismatches studied.

As revealed in Table 2, all Cµ region marker positions that were not under selective pressure contained either the chromosomal or vector-borne restriction site. Furthermore, the unselected XbaI site of vector linearization was restored in all recombinants. At first glance, this might suggest that the mammalian gene-targeting reaction occurred with fidelity. However, this conclusion becomes tenuous when one considers that the sum of DNA in these sites represents <1% of the total Cµ region DNA potentially involved in homologous recombination. Evidence for the mammalian gene-targeting reaction being subject to error has been suggested previously (Thomas and Capecchi 1986; Doetschman et al. 1988; Brinster et al. 1989; Baker and Read 1993). In this study, four recombinants [19/2, 25/4, 60/7, and 83-2] were found to produce cytolytic TNP-specific IgM even though the mutant igm482 TfiI site was present in the expressed 5' Cµ3 exon. In these clones, a mutation introduced by gene targeting may have restored the reading frame in the 5' Cµ region. Alternatively, the mutant igm482 TfiI site in the 5' Cµ region may have been corrected to the wild-type XmnI site present in the 3' Cµ region by intrachromosomal gene conversion (Baker 1989) but only in a subpopulation of cells that was below the level of detection by the PCR assay.

The duplicate Cµ region structure and marker patterns in the class I recombinants in Table 2 have been interpreted on the basis of a single vector integration event. However, the possibility exists that in some class I recombinants, the initial vector integration event might have been followed by an additional vector-directed gene replacement event. To account for the pure recombinant clones that were obtained, any hDNA generated by such a putative gene replacement event would have to be repaired before DNA replication and cell division. What is the likelihood that vector integration in the class I recombinants was accompanied by vector-directed gene replacement? In this study, 8 class II recombinants were recovered. Our analysis of the class II recombinants (Ng and Baker 1999, accompanying article) suggested that they contained two tandem vector copies integrated at the chromosomal immunoglobulin µ locus as a result of two vector integration events. If, during gene targeting, resolution results in the production with equal frequency of both crossover and noncrossover products, then it might be expected that
~8 additional class I recombinants were also generated by two vector interactions with the chromosome, one a vector integration and the other a gene replacement, event. We can estimate the fraction (F) of the total number of independent recombinants that were G418°C recombinants derived by vector integration using the absolute frequencies of gene targeting for procedure 2 (Table 1) as $F = (\text{frequency of G418° transforms})/ (\text{absolute frequency of G418° PFC})$. Accordingly, crossover and noncrossover products of gene targeting at the chromosomal $\mu$-locus do not occur with equal frequency, a finding consistent with the results of other mammalian gene targeting studies (Thomas et al. 1986; Song et al. 1987; Thomas and Capecci 1987; Adair et al. 1989). From the above information, it is possible to estimate that, of the total of 33 class I recombinants, ~1 may have been generated by both vector integration and vector-directed gene replacement events. Therefore, we conclude that the $C_{\mu}$ region marker patterns in the majority of class I recombinants can be interpreted on the basis of a single vector integration event.

We thank Joan Hamilton and Gary Smith for help with the illustrations. This work was supported by an operating grant from the Medical Research Council of Canada to M.D.B.

LITERATURE CITED


Miller, E. M., H. L. Hough, J. W. Cho and J. A. Nickoloff, 1997 Mismatch repair by efficient nick-directed, and less efficient mis-
Mechanisms of Double-Strand-Break Repair

match-specific, mechanisms in homologous recombination intermediates in Chinese hamster ovary cells. Genetics 147: 743-753.


Communicating editor: C. Kozak