Role of *ERCC1* in removal of long non-homologous tails during targeted homologous recombination

Gerald M. Adair¹, Rhonda L. Rolig,
Dana Moore-Faver, Marina Zabelshansky,
John H. Wilson² and Rodney S. Nairn

Department of Carcinogenesis, The University of Texas M.D. Anderson Cancer Center, Science Park–Research Division, Smithville, TX 78957 and ²The Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030, USA

¹Corresponding author
e-mail: adair@odin.mdacc.tmc.edu

The XpF/Ercc1 structure-specific endonuclease performs the 5' incision in nucleotide excision repair and is the apparent mammalian counterpart of the Rad1/Rad10 endonuclease from *Saccharomyces cerevisiae*. In yeast, Rad1/Rad10 endonuclease also functions in mitotic recombination. To determine whether XpF/Ercc1 endonuclease has a similar role in mitotic recombination, we targeted the *APRT* locus in Chinese hamster ovary *ERCC1*⁺ and *ERCC1*[−] cell lines with insertion vectors having long or short terminal non-homologies flanking each side of a double-strand break. No substantial differences were evident in overall recombination frequencies, in contrast to results from targeting experiments in yeast. However, profound differences were observed in types of *APRT*⁺ recombinants recovered from *ERCC1*[−] cells using targeting vectors with long terminal non-homologies—almost complete ablation of gap repair and single-reciprocal exchange events, and generation of a new class of aberrant insertion/deletion recombinants absent in *ERCC1*⁺ cells. These results represent the first demonstration of a requirement for *ERCC1* in targeted homologous recombination in mammalian cells, specifically in removal of long non-homologous tails from invading homologous strands.

Keywords: *ERCC1*/gene targeting/homologous recombination/terminal non-homology/XpF–Ercc1 endonuclease

Introduction

Nucleotide excision repair (NER) is responsible for processing DNA lesions that cause large distortions in DNA helical structure, such as UV photoproducts and bulky covalent chemical adducts (de Laat *et al.*, 1999). In eukaryotes, there is striking conservation of the genes involved in NER; the biochemical steps constituting this repair pathway are essentially the same in yeast (*Saccharomyces cerevisiae*) and mammalian cells (Aboussekhra and Wood, 1994; de Laat *et al.*, 1999). After damage recognition, lesion demarcation and formation of a pre-incision complex, a dual incision step

catalyzes release of a single-stranded oligonucleotide containing the damage site, allowing repair synthesis and ligation (Aboussekhra *et al.*, 1995; Mu *et al.*, 1996; de Laat *et al.*, 1999). The proteins responsible for dual incision are structure-specific endonucleases that recognize transitional DNA duplex/single-stranded regions. In mammalian cells, the XpG protein performs the initial incision 2–9 nucleotides (nt) 3' to DNA damage and the XpF/Ercc1 complex makes the second incision 16–25 nt 5' to the damage (Matsunaga *et al.*, 1996; Bessho *et al.*, 1997b; Evans *et al.*, 1997; de Laat *et al.*, 1998). The precise location of the DNA incisions is dependent on the nature of the DNA lesion and results in excision of a 24–32 nt oligonucleotide fragment (Huang *et al.*, 1992; Moggs *et al.*, 1996).

In *S. cerevisiae*, the *RAD1* and *RAD10* gene products also form a heterodimeric complex, which incises DNA specifically at 5'-double-strand (ds)–3'-single-strand (ss) junctions (Tomkinson *et al.*, 1993; Bardwell *et al.*, 1994) and the role of the Rad1/Rad10 endonuclease in NER in yeast is analogous to XpF/Ercc1 function in the dual incision step of NER in mammalian cells (Davies *et al.*, 1995; de Laat *et al.*, 1999). Rad1/Rad10 endonuclease is capable of recognizing and incising 'bubble substrates', as well as transitional structures with single-stranded flaps or unpaired 'tails' positioned 3' to duplex structures (Bardwell *et al.*, 1994; Davies *et al.*, 1995; Rodriguez *et al.*, 1996). Purified XpF/Ercc1 complex has been shown to incise similar structures, with a polarity of incision consistent with nicking in duplex DNA on the 5' side of ds–ss junctions (Park *et al.*, 1995; Matsunaga *et al.*, 1996; Bessho *et al.*, 1997b; de Laat *et al.*, 1998). Rad1 and Rad10 proteins also exhibit significant amino acid conservation with the mammalian XpF and Ercc1 proteins, respectively (van Duin *et al.*, 1986; Brookman *et al.*, 1996; Sijbers *et al.*, 1996). Mutation or gene disruption of *RAD1* or *RAD10* in yeast and *XPF* or *ERCC1* in mammalian cells results in pronounced UV hypersensitivity and NER deficiency.

RAD1 and *RAD10* are required for certain specialized pathways of mitotic recombination as well as NER (for review see Paques and Haber, 1999). Genetic studies have shown that intrachromosomal recombination between direct or inverted repeats involves distinct *RAD1/RAD10* and *RAD52* recombination pathways (Klein, 1988, 1995; Schiestl and Prakash, 1988, 1990; Prado and Aguilera, 1995) and that *RAD1* or *RAD10* knockouts reduce the frequency of targeted homologous integration of a plasmid into a chromosomal locus (Schiestl and Prakash, 1988, 1990; Saffran *et al.*, 1994). *RAD1* and *RAD10* are required both for removal of long, 3'-ended non-homologous tails from single-strand annealing (SSA) recombination intermediates during double-strand break (DSB)-induced recombination between direct repeats, and removal of

end-blocking non-homologous sequences from invading homologous strands during DSB-induced recombination between plasmid inverted repeats (Fishman-Lobell and Haber, 1992; Ivanov and Haber, 1995; Prado and Aquilera, 1995; Saparbaev *et al.*, 1996; Paques and Haber, 1997, 1999; Sugawara *et al.*, 1997; Colaiacovo *et al.*, 1999). Among the genes in the *RAD3* NER epistasis group, only *RAD1* and *RAD10* appear to be directly involved in these mitotic recombination pathways (Ivanov and Haber, 1995). The above studies also demonstrated that although the Rad1/Rad10 endonuclease is required for efficient removal of end-blocking non-homologies of >20 nt, shorter terminal non-homologies can be removed by *RAD1/RAD10*-independent pathways.

In addition to the characteristic UV-hypersensitive phenotypes displayed by all rodent NER complementation groups, the CHO *ERCC1* and *ERCC4* (*XPF*) complementation groups are uniquely defined by extreme hypersensitivity to agents that cause DNA interstrand crosslinks (Hoy *et al.*, 1985; Collins, 1993; Busch *et al.*, 1997). This phenotype is consistent with involvement of these two NER genes in recombination, since DNA interstrand crosslinks may be subject to repair by recombinational mechanisms (Bessho *et al.*, 1997a). Because of the structural homology of the Rad10 and Ercc1 proteins, the requirement for *RAD10* in specific pathways of homologous recombination in yeast (Ivanov and Haber, 1995; Paques and Haber, 1999), and the *ERCC1*-deficient phenotype of extreme hypersensitivity to DNA crosslinks, it has been postulated that *ERCC1* is a dual-function repair/recombination gene in mammalian cells (Weeda *et al.*, 1997). However, neither *XPF* nor *ERCC1* has yet been shown directly to be required for any specific recombination pathway in mammalian cells; their presumed function in recombination is based primarily on their DNA crosslink-sensitive phenotypes, conserved structural homologies with Rad1 and Rad10 polypeptides, and the observations that the XpF/Ercc1 and Rad1/Rad10 complexes perform analogous functions in NER.

In studies examining the potential role of *ERCC1* in recombination in mammalian cells, no substantial differences were observed between *ERCC1*⁺ and *ERCC1*⁻ cells in overall frequencies of homologous recombination between extrachromosomal substrates (Nairn *et al.*, 1991), plasmid-chromosome targeting in *ERCC1* nullizygous mouse embryonic stem cells (Melton *et al.*, 1998), or intrachromosomal recombination between direct repeats in CHO *ERCC1*⁺ and *ERCC1*⁻ cell lines (Sargent *et al.*, 1997). However, cells from *ERCC1* knockout mice show increased genomic instability (Melton *et al.*, 1998) and recombination-dependent deletions and rearrangements involving duplicated sequences at the Chinese hamster ovary (CHO) *APRT* locus appear to be suppressed in *ERCC1*⁺ cells relative to *ERCC1*⁻ cells (Sargent *et al.*, 1997). These phenotypes are suggestive of a subtle, rather than general, recombination defect in *ERCC1*⁻ cells. Furthermore, although the functions of *RAD1* and *RAD10* in removal of long 3'-ended, non-homologous single-stranded tails from SSA recombination intermediates, or from invading homologous strands during DSB-induced recombination between direct or inverted repeats in yeast, have been elaborately described (for review see Paques and Haber, 1999), it is possible that these genes

may have other roles in recombination that have not been recognized.

To demonstrate directly a role for *ERCC1* in homologous recombination, we designed gene targeting vectors that would allow us to examine how the presence of terminal non-homology flanking both sides of a DSB would affect targeted homologous recombination in CHO *ERCC1*⁺ and *ERCC1*⁻ cells. For these targeting experiments, illustrated in Figure 1, we used 'ends-in' insertion-type *APRT* targeting vectors in which both arms of targeting homology were blocked with either long or short terminal non-homologies, assuming that the efficiency of these targeting vectors in generating certain classes of recombinants would be dependent on successful removal of the non-homologous sequences blocking the 3'-OH ends of invading homologous strands. As targeting recipients, we used *ERCC1*⁺ and *ERCC1*⁻ mutant or knockout cell lines containing the same endogenous 3 bp deletion at the hemizygous *APRT* target locus. Recombinants generated in these experiments were characterized by Southern analysis to determine the relative frequencies and class distributions of target gene conversions, targeted insertions and targeting vector corrections. Our results, which constitute the first direct demonstration of a requirement for *ERCC1* in a homologous recombination pathway in mammalian cells, indicate that *ERCC1*: (i) is required for the removal of long non-homologous tails from the 3'-OH ends of invading DNA strands during targeted homologous recombination and (ii) may suppress formation of deletions arising from recombination intermediates that require removal of such 3'-OH-end-blocking non-homologous sequences for normal resolution by homologous recombination pathways.

Results

Effects of terminal non-homology on the frequency of targeted recombination in *ERCC1*⁺ and *ERCC1*⁻ cell lines

The first question we addressed was whether *ERCC1* knockout or mutation would have a significant effect on the overall frequency of targeted recombination when both arms of targeting homology in an ends-in insertion vector were blocked by terminal non-homology, flanking each side of a DSB. In yeast plasmid-chromosome recombination experiments in which a DSB was introduced into a region of targeting homology on a non-replicating plasmid (creating an insertion vector with targeting homology flanking each side of a DSB), *RAD1* or *RAD10* knockouts reduced the frequencies of targeted integration into the homologous chromosomal target gene locus by 5- to 9-fold compared with *RAD*⁺ cells, in the absence of any blocking terminal non-homology (Schiestl and Prakash, 1988, 1990). In the present study, we carried out a series of *APRT* targeting experiments in which CHO *ERCC1*⁺, *ERCC1*⁻ mutant cells and *ERCC1*⁻ knockout cells were electroporated either with pAG6 (Figure 1, top left), an ends-in insertion-type vector with very short (18 and 11 nt) terminal non-homologies blocking each arm of *APRT* targeting homology, or with the pAG6ins0.9 insertion vector (Figure 1, top right), which has much longer (271 and 670 nt) terminal non-homologies. *APRT*⁺ targeted

Fig. 1. Targeted recombination at the CHO *APRT* locus. The pAG6 and pAG6ins0.9 targeting vectors carry a full-length Chinese hamster *APRT* gene, disrupted by insertion of 33 or 945 bp of heterologous sequence into the exon 3 *XhoI* site. Linearization of pAG6 at a unique *HindIII* site within the heterologous insert creates an ends-in insertion vector configuration in which 18 and 11 terminal non-homologies flank each side of the DSB. Linearization of pAG6ins0.9 at a unique *SalI* site within the larger heterologous insert generates an insertion vector in which 271 and 670 nt terminal non-homologies flank each side of the DSB. The *ERCC1*⁺ (ATS49tg) and *ERCC1*⁻ mutant (U9S50tg) or knockout (E1KO7-5) cell lines used in this study all have the same 3 bp deletion in exon 5 of the *APRT* target gene locus, resulting in loss of the exon 5 *MboII* (Mb*) site. Electroporation of these cell lines with the pAG6 or pAG6ins0.9 targeting vectors, followed by selection for ALASA-resistant clones, results in recovery of *APRT*⁺ recombinants with the recombinant class structures shown. B, *BamHI*; Bg, *BglII*; E, *EcoRI*; H, *HindIII*; Mb, *MboII*; P, *PstI*; S, *SacI*; X, *XhoI*; Xb, *XbaI*.

recombinants were selected by plating into ALASA (alanosine/azaserine/adenine medium), as described in Materials and methods. Results from these experiments are shown in Table I.

In *APRT* targeting experiments using the pAG6 (short terminal non-homology) vector, the overall frequencies of *APRT*⁺ recombinants obtained in *ERCC1*⁺ and *ERCC1*⁻ mutant or knockout cell lines were quite similar. Surprisingly, even in *APRT* targeting experiments using the pAG6ins0.9 (long terminal non-homology) vector, the overall frequencies of *APRT*⁺ recombinants obtained in both *ERCC1*⁻ cell lines were only slightly lower than the frequencies in *ERCC1*⁺ ATS49tg cells.

Southern analysis of *APRT*⁺ recombinants in *ERCC1*⁺ cells

In previous CHO *APRT* targeting experiments, employing conventional ends-in insertion vectors, with two unblocked arms of *APRT* targeting homology flanking a DSB (Adair *et al.*, 1989, 1998), we have consistently recovered three major classes of *APRT*⁺ recombinants: (i) target gene conversions, in which the 3 bp target gene deletion has been corrected by unidirectional transfer of sequence information from the targeting vector (Figure 1A); (ii) targeted insertions at the chromosomal *APRT* gene locus, produced by single-reciprocal exchange/crossover events (Figure 1B); and (iii) targeting vector

Table 1. Frequencies of targeted homologous recombination in *ERCC1*⁺ and *ERCC1*⁻ cell lines after electroporation with pAG6 or pAG6ins0.9 targeting vectors

Targeting vector/cell line	Total No. of cells electroporated (No. of cuvettes)	Frequency of APRT ⁺ recombinants ^a	Frequency of GPT ⁺ transfectants ^b
<i>Hind</i> III-cut pAG6			
ATS49tg (<i>ERCC1</i> ⁺)	8.5 × 10 ⁷ (5)	5.5 ± 0.7 (5.5)	13.7 ± 2.9 (13.2)
U9S50tg (<i>ERCC1</i> ⁻)	6.8 × 10 ⁷ (4)	4.7 ± 0.2 (4.6)	10.6 ± 0.5 (10.5)
E1KO7-5 (<i>ERCC1</i> ^{ko})	3.4 × 10 ⁷ (2)	4.3 ± 1.4 (3.9)	12.6 ± 4.2 (11.2)
combined <i>ERCC1</i> ⁻	10.2 × 10 ⁷ (6)	4.6 ± 0.4 (4.2)	11.2 ± 1.2 (10.8)
<i>Sal</i> I-cut pAG6ins0.9			
ATS49tg (<i>ERCC1</i> ⁺)	23.8 × 10 ⁷ (14)	5.1 ± 0.9 (5.1)	10.9 ± 1.0 (9.6)
U9S50tg (<i>ERCC1</i> ⁻)	6.8 × 10 ⁷ (4)	3.4 ± 0.1 (3.4)	7.2 ± 0.1 (7.2)
E1KO7-5 (<i>ERCC1</i> ^{ko})	17.0 × 10 ⁷ (10)	3.0 ± 0.7 (2.2)	8.8 ± 1.7 (5.9)
combined <i>ERCC1</i> ⁻	23.8 × 10 ⁷ (14)	3.1 ± 0.5 (3.0)	8.4 ± 1.2 (8.4)

^aAPRT⁺ recombinants per 10⁶ viable cells plated; calculated either as the mean frequency ± SEM for all individual electroporations, or (in parentheses) by dividing the total number of ALASA-resistant colonies obtained from all electroporations by the total number of viable cells plated.

^bGPT⁺ transfectants per 10⁴ viable cells plated; calculated either as the mean frequency ± SEM for all individual electroporations, or (in parentheses) by dividing the total number of HAT-resistant colonies obtained from all electroporations by the total number of viable cells plated.

corrections, which reflect DSB/gap repair of the vector *APRT* gene by unidirectional transfer of sequence information from the chromosomal gene (Figure 1C). These three classes of recombinants are readily distinguishable on the basis of diagnostic restriction fragment patterns.

Target gene conversions (Figure 1A) will show correction of the 3 bp target gene deletion and restoration of the exon 5 *Mbo*II restriction site, reflected by loss of the 2.0 kb *Mbo*II fragment and appearance of 1.5 and 0.5 kb fragments characteristic of a wild-type *APRT* gene (as shown for ATS49tg recombinant 9.10-32, in Figure 2A, lane 3), but will retain normal restriction fragment patterns characteristic of the endogenous *APRT* locus for all other restriction enzymes (lane 3 of Figure 2B and C).

Targeted integration of the insertion vector by single reciprocal exchange/crossover will result in a duplication of *APRT* gene sequences at the target gene locus (Figure 1B), generating a wild-type copy of the *APRT* gene at one end (upstream) of the integrated vector sequences, and either an *Mbo*II site-deleted or wild-type copy of the *APRT* gene at the other end, depending on whether the insertion was accompanied by conversion of the downstream *Mbo*II site deletion. An example of an ATS-49tg *Mbo*II+/Δ targeted insertion (recombinant 99.14-12) is shown in Figure 2A, lane 4; an example of an *Mbo*II+/+ targeted insertion with conversion of the downstream *Mbo*II deletion site (recombinant 99.10R-5) is shown in Figure 2A, lane 5. Targeted insertion recombinants show diagnostic restriction fragment patterns following double digestion with *Hind*III and *Sac*I (13.1 and 4.0 kb fragments; Figure 2B, lanes 4 and 5) or with *Bgl*II and *Xho*I (2.2, 3.85, 4.6 and 2.4 kb fragments; Figure 2C, lanes 4 and 5). Note that both the upstream and downstream *APRT* gene copies contain a wild-type exon 3 *Xho*I site (Figure 2C, lanes 4 and 5); this is the site that is disrupted by non-homologous inserts in the pAG6 and pAG6ins0.9 targeting vectors.

Vector correction recombinants retain an uncorrected *APRT* target allele (with 3 bp exon 5 *Mbo*II site deletion) at the original chromosomal locus, but have also acquired an ectopically integrated, wild-type *APRT* gene, generated by recombinational correction of the vector *APRT* sequence.

Fig. 2. Southern analysis of APRT⁺ recombinants obtained from ATS49tg pAG6ins0.9 targeting experiments. Lane 1, APRT⁺ CHO-AT3-2 cells; lane 2, APRT⁻ ATS49tg cells; lane 3, a target gene conversion; lane 4, an *Mbo*II+/Δ targeted insertion; lane 5, an *Mbo*II+/+ targeted insertion; lane 6, a targeting vector correction. (A) *Mbo*II digests, 1.4 kb *Eco*RI-*Xba*I *APRT* probe. (B) *Hind*III-*Sac*I double digests, 3.9 kb *Bam*HI *APRT* probe. (C) *Bgl*II-*Xho*I double digests, 3.9 kb *APRT* probe.

Thus, vector corrections will have wild-type 1.5 and 0.5 kb *APRT* *Mbo*II fragments, as well as the 2.0 kb *Mbo*II fragment characteristic of the original target gene locus (as shown for recombinant 99.14-28, in Figure 2A, lane 6). The 0.8 kb *Mbo*II fragment seen in clone 99.14-28 represents the 3' junction-fragment produced by ectopic integration of the corrected *APRT* targeting vector sequence. After digestions with other restriction enzymes, vector corrections will show hybridizing fragments characteristic of the original *APRT* target gene locus, plus novel fragments that have been produced by ectopic integration of the corrected, vector-derived *APRT* gene (lane 6 of Figure 2C and D).

Effects of terminal non-homology on the class distributions of APRT⁺ targeted recombinants in *ERCC1*⁺ and *ERCC1*⁻ cell lines

A total of 289 independent APRT⁺ targeted recombinants, generated by electroporation of *ERCC1*⁺ and *ERCC1*⁻ mutant or knockout cell lines with the pAG6 or pAG6ins0.9 targeting vector, were characterized by Southern analysis and classified by recombinant type, as

Table II. Class distributions of targeted recombination events in *ERCCI*⁺ and *ERCCI*⁻ cells after electroporation with pAG6 or pAG6ins0.9 targeting vectors

Targeting vector/ cell line	APRT ⁺ recombinants					Targeting vector correction
	Total analysed	Target gene conversions	Targeted insertions			
			<i>Mbo</i> II+/Δ ^a	<i>Mbo</i> II+/+ ^b	aberrant ^c	
<i>Hind</i> III-cut pAG6						
ATS49tg (<i>ERCCI</i> ⁺)	49	17 (35%)	10 (20%)	5 (10%)	0	17 (35%)
U9S50tg (<i>ERCCI</i> ⁻)	28	15 (54%)	3 (11%)	4 (14%)	0	6 (21%)
E1KO7-5 (<i>ERCCI</i> ^{ko})	29	12 (41%)	3 (10%)	6 (21%)	0	8 (28%)
<i>Sa</i> I-cut pAG6ins0.9						
ATS49tg (<i>ERCCI</i> ⁺)	90	41 (46%)	27 (30%)	8 (9%)	0	14 (16%)
U9S50tg (<i>ERCCI</i> ⁻)	29	24 (83%)	1 (3%)	0	4 (14%)	0
E1KO7-5 (<i>ERCCI</i> ^{ko})	67	55 (82%)	0	1 (2%)	11 (16%)	0

^aTargeted insertions reflecting single reciprocal exchange/crossover without conversion of the *Mbo*II site deletion in the downstream copy of the *APRT* gene.

^bTargeted insertions reflecting single reciprocal exchange/crossover with conversion of the *Mbo*II site deletion in the downstream copy of the *APRT* gene.

^cAberrant targeted insertions in which a variable portion of the integrated targeting vector sequence appears to have been deleted or retains heterologous insert sequences.

described above. Analysis of 106 APRT⁺ recombinants from targeting experiments employing the pAG6 vector demonstrated that each of the three expected APRT⁺ recombinant classes shown in Figure 1 (target gene conversions, targeted insertions and targeting vector corrections) was recovered at comparable frequencies in *ERCCI*⁺ and *ERCCI*⁻ cell lines (Table II).

Analysis of 90 APRT⁺ recombinant clones from ATS49tg/pAG6ins0.9 targeting experiments demonstrated that in *ERCCI*⁺ cells, all three classes of APRT⁺ targeted recombinants were obtained even when both arms of *APRT* targeting homology were blocked by substantial terminal non-homologies, with class frequency distributions that were not statistically different from those for pAG6 targeting experiments (Table II). However, a strikingly different class distribution of APRT⁺-targeted recombinants was observed in pAG6ins0.9 targeting experiments using *ERCCI*⁻ mutant or knockout cell lines. Analysis of 96 APRT⁺ recombinants obtained from *ERCCI*⁻/pAG6ins0.9 targeting experiments revealed that target gene conversions represented >80% of all recombinants recovered from *ERCCI*⁻ cell lines, with no vector corrections and virtually no normal targeted insertions (Table II).

Recombinant structures of aberrant targeted insertion/deletion events, seen only in *ERCCI*⁻ cells

Fourteen of the APRT⁺-targeted recombinants recovered from pAG6ins0.9 targeting experiments using *ERCCI*⁻ cell lines represented a novel class of aberrant targeted insertions in which a variable portion of the integrated vector sequence appears to have been deleted (Figure 1D); another aberrant targeted insertion event produced a unique target gene triplication in which the middle *APRT* gene copy retains the non-homologous insert sequences originally present in the vector. The recovery of such aberrant insertion/deletion events in these targeting experiments was quite unexpected; targeted homologous integration of an ends-in, insertion-type vector

would normally be expected to result in precise, target gene-templated repair of the DSB/gap in the vector sequence, and conservative integration of the entire vector into the chromosomal target gene locus (Szostak *et al.*, 1983; Adair *et al.*, 1989, 1998; Pennington and Wilson, 1991; Valancius and Smithies, 1991; Hasty *et al.*, 1992, 1995; Stahl, 1996; Osman and Subramani, 1998). Southern analyses of six *ERCCI*⁻ targeted/aberrant insertion recombinants are shown in Figure 3 and their deduced structures in Figure 4. APRT⁺ recombinant 9U50/6ins0.9 S-17 (lane 2 of Figure 3A–D) was the only apparently normal *Mbo*II+/Δ targeted insertion recovered from *ERCCI*⁻/pAG6ins0.9 targeting experiments. *Bgl*II–*Xho*I double digests of this recombinant yielded a 2.2, 3.85, 4.6, 2.4 kb restriction fragment pattern characteristic of a normal targeted insertion, in which both copies of the *APRT* gene duplication produced at the target gene locus have a normal wild-type exon 3 *Xho*I restriction site (Figure 3C, lane 2). *Bgl*II–*Xho*I double digests of recombinant 9U50/6ins0.9 S-23 (Figure 3C, lane 3) revealed a 2.2, 3.85, 9.4, 4.6, 2.4 kb restriction fragment pattern, suggesting an aberrant insertion/triplication structure in which both the upstream and downstream *APRT* gene copies contain a wild-type exon 3 *Xho*I restriction site, but the middle copy still retains the non-homologous insert that disrupts this site in the pAG6ins0.9 vector (Figure 4B). Restriction fragment patterns for *Mbo*II (1.5, 0.5, 1.55, 0.9, 0.5, 1.5 and 0.5 kb; Figure 3A, lane 3), *Pst*I (3.9, 9.4, 8.45 and 2.2 kb; Figure 3B, lane 3), *Hind*III–*Eco*RI (5.9, 7.1, 1.35, 1.45, 6.65, 1.35 and 2.65 kb; Figure 3D, lane 3) and other digests (not shown) are all consistent with the novel aberrant insertion/triplication structure shown in Figure 4B.

Southern analyses of four representative *ERCCI*⁻ aberrant insertion/deletion recombinants (9U50/6ins S-3 and E1KO7-5 99.12-14, 99.12-39 and 96.6-2), with variable-sized deletions of the integrated targeting vector sequence ranging from ~0.6 to 7.8 kb, are shown in lanes 4–7 of Figure 3A–D. The deduced structures of these recombinants are depicted in Figure 4C–F. Analyses of *Mbo*II

Fig. 3. Southern analysis of aberrant insertion recombinants obtained from *ERCC1*⁻ cell pAG6ins0.9 targeting experiments. Lane 1, APRT⁻ E1KO7-5 cells; lanes 2–4, one normal targeted insertion (9U50/6ins S-17) and two aberrant insertion recombinants (9U50/6ins S-23, 9U50/6ins S-3) from U9S50tg targeting experiments; lanes 5–7, three aberrant insertion recombinants (99.12-14, 99.12-39 and 96.6.2-11) from E1KO7-5 targeting experiments. (A) *Mbo*II digests, 1.4 kb *Eco*RI–*Xba*I APRT probe. (B) *Pst*I digests, 3.9 kb *Bam*HI APRT probe. (C) *Bgl*II–*Xho*I double digests, 3.9 kb *Bam*HI APRT probe. (D) *Hind*III–*Eco*RI double digests, 3.9 kb *Bam*HI APRT probe.

(Figure 3A), *Pst*I (Figure 3B), *Bgl*II–*Xho*I (Figure 3C), *Hind*III–*Eco*RI (Figure 3D) and other restriction fragment patterns allowed determination of the approximate size and extent of the deleted region in each of these recombinants.

Discussion

In *S.cerevisiae*, the Rad1/Rad10 endonuclease appears to be required not only for removal of long, 3'-ended, non-homologous single-stranded tails from SSA recombination intermediates during DSB-induced recombination between direct repeats, but also for the removal of such blocking non-homologies from the 3'-OH ends of invading homologous strands in other pathways of DSB-induced homologous recombination (Fishman-Lobell and Haber, 1992; Ivanov and Haber, 1995; Prado and Aquilera, 1995; Saparbaev *et al.*, 1996; Sugawara *et al.*, 1997; Paques and Haber, 1997, 1999; Colaiacovo *et al.*, 1999). Although two previous studies in mammalian cells failed to demonstrate a general effect of *ERCC1* deficiency on either the rate of spontaneous intrachromosomal homologous recombination between direct repeats (Sargent *et al.*, 1997) or the

overall frequency of gene targeting (Melton *et al.*, 1998), neither study was designed to investigate the role of *ERCC1* in recombination events that would specifically require the removal of terminal non-homologies. To examine this most likely role of *ERCC1* directly, we conducted gene targeting experiments employing two insertion-type vectors (pAG6 and pAG6ins0.9), in which both arms of *APRT* targeting homology were blocked by very short or substantially longer terminal non-homologies flanking each side of a DSB.

ERCC1 mutation or knockout has no significant effect on the overall frequency of targeted homologous recombination

Previous *APRT* gene targeting experiments using the pAG6 insertion vector (Pennington and Wilson, 1991) suggested that ATS49tg cells were capable of carrying out conservative homologous recombination events requiring removal of short terminal non-homologies from the ends of invading homologous strands of the vector. In our present targeting experiments using the pAG6 vector, *ERCC1* mutation or knockout appeared to have no significant effect on the overall frequency of targeted homologous recombination. These results suggest that *ERCC1* is not generally required for targeted homologous recombination and demonstrate that the short terminal non-homologies blocking both arms of targeting homology in the pAG6 vector can be effectively removed by *ERCC1*-independent pathways. While our results appear to support the conclusion of Melton *et al.* (1998) that *ERCC1* is not required for the general recombination pathways involved in gene targeting, they are difficult to reconcile with findings by Schiestl and Prakash (1988, 1990) that in yeast, *RAD1* or *RAD10* knockouts markedly reduce the frequencies of targeted integration by linearized insertion vectors with no terminal non-homology.

We anticipated that XpF/Ercc1 endonuclease activity might be required for removal of the long terminal non-homologies blocking both arms of *APRT* targeting homology in the pAG6ins0.9 vector. We therefore expected *ERCC1*⁻ cells to show lower frequencies of homologous recombination than *ERCC1*⁺ cells in targeting experiments using this vector. However, even in pAG6ins0.9 targeting experiments, *ERCC1* mutation or knockout appeared to have relatively little effect on overall frequencies of targeted homologous recombination. Our analyses of the recombinant class distributions obtained from these experiments, however, revealed a more subtle recombination-deficient phenotype in *ERCC1*⁻ cells, which was only evident in experiments using the pAG6ins0.9 vector.

ERCC1⁻ cells are specifically deficient in carrying out homologous recombination events requiring the removal of long terminal non-homologies

In previous targeting experiments employing insertion-type vectors with two unblocked arms of *APRT* targeting homology flanking a DSB (Adair *et al.*, 1989, 1998), we recovered three major classes of APRT⁺ recombinants: target gene conversions, targeted insertions and vector corrections. Our ATS49tg/pAG6 targeting experiments yielded recombinant class distributions virtually identical to those reported by Pennington and Wilson (1991), with

Fig. 4. Structures of normal and aberrant targeted insertion recombinants. Deduced structures of the six APRT⁺ recombinants shown in Figure 3. Abbreviations as in Figure 1.

nearly equal representation of each class (conversions, targeted insertions and vector corrections) among the 49 APRT⁺ recombinants analyzed. In our pAG6 experiments, recombinant class distributions in both *ERCC1*⁻ cell lines differed slightly, but not significantly, from distributions observed in the *ERCC1*⁺ ATS49tg cell line. These results suggest that XpF/Ercc1 endonuclease is not required for generation of any of these classes of recombinants when the targeting vector contains only short terminal non-homologies.

Although *ERCC1* mutation or knockout seemed to have relatively little effect on the overall frequency of targeted homologous recombination, even in pAG6ins0.9 targeting experiments, our analyses of APRT⁺ recombinants recovered from such experiments revealed that the recombinant class distributions and specific types of recombination events recovered from *ERCC1*⁻ mutant or knockout cell lines were profoundly different from those from *ERCC1*⁺ ATS49tg cells. Among 96 APRT⁺ recombinants from pAG6ins0.9/*ERCC1*⁻ cell targeting experiments, we recovered no targeting vector correction recombinants, and only two normal-looking targeted insertions (one from each *ERCC1*⁻ cell line). These results indicate a specific requirement for *ERCC1* in the

recombination pathway(s) leading to these two recombination products, but only when both arms of targeting vector homology are blocked by long terminal non-homologies.

In current models of DSB-induced homologous recombination, resection of 5' ends at a DSB by a 5'→3' exonuclease creates long 3'-ended single-stranded tails, which can invade a homologous DNA duplex, base-pair and provide an extendible 3'-OH primer-end for new DNA synthesis (Belmaaza and Chartrand, 1994; Stahl, 1996; Osman and Subramani, 1998; Paques and Haber, 1999). Initial interaction and alignment of invading strand-target DNA duplex homology is presumed to involve localized unwinding of the duplex and protein-assisted base-pairing to create a relatively unstable paranemic joint, in which the invading strand is paired, but not intertwined with its complementary partner (Bianchi *et al.*, 1983; Paques and Haber, 1997; Wong *et al.*, 1998). Conversion of a paranemic joint to a more stable plectonemic joint, by strand transfer and intertwining of paired strands, is constrained unless one strand has a free end. Pairing of a 3'-ended homologous invading strand would allow rapid stabilization of the paranemic joint by primed DNA synthesis, but the presence of terminal non-homology

would preclude 3'-OH-end extension. Since the invading strand is generally presumed to be the recipient of genetic information, its 3'-OH-end extension using the complementary strand of the target duplex as template should result in unidirectional transfer of sequence information from chromosomal duplex to invading strand.

In plasmid-chromosome recombination experiments employing ends-in, insertion-type vectors, targeted integrations are presumed to arise by single reciprocal exchange/crossover recombination events that result in precise, target gene-templated repair of the DSB/gap in the vector *APRT* sequence, and conservative integration of the entire vector into the chromosomal target gene locus (Stahl, 1996; Osman and Subramani, 1998; Paques and Haber, 1999). Vector correction events also require precise DSB/gap repair of the vector *APRT* sequence by unidirectional transfer of sequence information from the chromosomal target locus. These two classes of recombinant could be generated by either conventional DSB repair or synthesis-dependent strand annealing (SDSA) recombination pathways; both pathways would require complete removal of terminal non-homology from the 3' ends of invading strands to create a free 3'-OH end, which can prime and be extended by target gene-templated DNA synthesis. In contrast, target gene conversions can arise by several pathways that would not necessarily require removal of targeting vector terminal non-homologies. The near-complete ablation of normal targeted insertions and vector corrections in *ERCC1*-pAG6ins0.9 targeting experiments suggests that *ERCC1*⁻ cells are specifically deficient in their ability to carry out recombination events requiring removal of long 3'-end-blocking non-homologous tails from invading strands.

Although short terminal non-homologies can be efficiently removed by *ERCC1*-independent pathways, CHO cells appear to be incapable of effectively removing long blocking non-homologies from the 3' ends of invading homologous strands in the absence of functional XpF/Ercc1 endonuclease. Based on the numbers of targeted insertions and target vector corrections obtained from *ERCC1*⁺ cell targeting experiments and from *ERCC1*⁻ cells in pAG6 targeting experiments, one would expect ~70% of the *APRT*⁺ recombinants in *ERCC1*⁻ targeting experiments to be either targeted insertions or vector corrections if *ERCC1* were not required for the removal of long terminal non-homologies. Therefore, our recovery of no vector corrections and only two normal targeted insertions among 96 *APRT*⁺ recombinants isolated from pAG6ins0.9 targeting experiments in *ERCC1*⁻ mutant or knockout cell lines suggests that *ERCC1*⁻ cells are profoundly deficient (~30-fold reduced) in their ability to remove long non-homologous tails precisely from the 3' ends of invading homologous strands during recombination.

Leung *et al.* (1997) have presented evidence that gene targeting-mediated conversion can occur by assimilation of a single strand that may normally be subject to preferential mismatch correction in favor of the resident unbroken strand. Our finding that target gene conversions represent >80% of all *APRT*⁺ recombinants recovered from pAG6ins0.9 targeting experiments in *ERCC1*⁻ mutant or knockout cell lines suggests that the inability of *ERCC1*⁻ cells to generate targeted insertion or vector

correction recombinants, by removal of terminal non-homologies and extension of the 3'-OH ends of invading vector strands, may favor such an alternative target gene conversion pathway involving assimilation of invading strands into the chromosomal duplex. Consequently, many of the invading vector strands that would normally be extended by 3'-OH-end DNA synthesis and resolved as either targeted insertions or vector conversion events in *ERCC1*⁺ cells, may instead lead to target gene conversions in *ERCC1*⁻ cells.

Role of *ERCC1* in maintaining genomic stability

A novel class of aberrant targeted integration events, not seen in *ERCC1*⁺ cells, represented >15% of the *APRT*⁺ recombinants recovered from pAG6ins0.9 targeting experiments in *ERCC1*⁻ mutant or knockout cell lines. All but one of these aberrant insertion recombinants appear to have sustained substantial deletions of the integrated targeting vector and downstream *APRT* gene sequences, retaining only one intact copy of the *APRT* gene at the target gene locus, with partial to very extensive deletions of the integrated pAG6ins0.9 vector sequence. We believe that these recombinants reflect deletogenic events produced by aberrant resolution of, or attempted DNA replication through, arrested recombination intermediates containing unremoved non-homologous single-strand tails. Aberrant deletion events have also been observed during spontaneous recombination between direct chromosomal repeats in *ERCC1*⁻ cells (Sargent *et al.*, 1997). The inability of *ERCC1*⁻ cells to remove long terminal non-homologies from the 3'-OH ends of DNA strands during SSA- or SDSA-mediated recombination between chromosomal repeat sequences could lead to DSBs during DNA replication. These DSBs could in turn promote end-joining by non-homologous/illegitimate recombination pathways, leading to increased frequencies of deleterious deletions or chromosomal rearrangements that might contribute to the genomic instability observed in *ERCC1*⁻ knockout mice (McWhir *et al.*, 1993; Weeda *et al.*, 1997; Melton *et al.*, 1998).

The results of this study demonstrate, for the first time, that *ERCC1* is required for removal of long non-homologous tails from the 3'-OH ends of invading strands during targeted homologous recombination, and suggest that the primary role of XpF/Ercc1 endonuclease in homologous recombination in mammalian cells is analogous to that of Rad1/Rad10 endonuclease in yeast. Our findings further suggest that the function of XpF/Ercc1 endonuclease in the removal of end-blocking non-homologous tails from recombining DNAs may ensure normal resolution of homologous recombination intermediates, which otherwise might be diverted into and processed by non-homologous end-joining pathways that could lead to deleterious deletions or rearrangements.

Materials and methods

Cell lines and culture conditions

The cell lines used for these experiments are all derived from CHO-AT3-2, a CHO cell line that is hemizygous for the *APRT* locus (Adair *et al.*, 1983). ATS49tg, a spontaneous 6-thioguanine-resistant (HGPRT⁻), 8-azaadenine-resistant (*APRT*⁻) mutant of CHO-AT3-2, contains a single, mutationally inactivated copy of the *APRT* gene, with a 3 bp deletion that has eliminated the exon 5 *Mbo*II restriction site (Smith and

Adair, 1996). U9S50tg, which contains the same spontaneous 3 bp *APRT* deletion (Smith and Adair, 1996), is an HGPRT⁻, APRT⁻ derivative of UVL9, an *ERCC1*⁻ mutant that was isolated from CHO-AT3-2 after ethylmethane sulfonate mutagenesis (Clarkson *et al.*, 1983; Rolig *et al.*, 1998). E1KO7-5, a CHO *ERCC1*⁻ knockout cell line generated by targeted disruption of the hemizygous *ERCC1* gene locus in ATS49tg, was obtained by gene targeting/positive-negative selection, exactly as described in Rolig *et al.* (1997) for targeted disruption of the *ERCC1* gene in CHO-K1 cells. Diagnostic restriction fragment patterns for E1KO7-5 at the disrupted *ERCC1* locus were identical to those exhibited by *ERCC1* knockout cell line CHO-7-27 (Rolig *et al.*, 1997). The repair phenotypes of the *ERCC1*⁻ cell lines used in the present study (U9S50tg and E1KO7-5) were essentially the same as those measured independently in several studies of *ERCC1* mutants (Thompson *et al.*, 1981; Collins, 1993; Busch *et al.*, 1997; Rolig *et al.*, 1997), exhibiting UV hypersensitivity and extreme DNA crosslink hypersensitivity; in addition, northern and western blot analyses indicated that neither *ERCC1* mRNA nor protein was expressed in U9S50tg (Rolig *et al.*, 1998) or E1KO7-5 (data not shown).

Cells were maintained as subconfluent, exponentially growing monolayer cultures in alpha-modified minimal essential medium, containing 2 mM L-glutamine, penicillin (50 U/ml), streptomycin (50 µg/ml) and 10% fetal bovine serum, in a 37°C incubator (5% CO₂/95% air). APRT⁺ recombinants were selected in ALASA selection medium. GPT⁺ transfectants were selected in hypoxanthine/methotrexate/thymidine (HAT) selection medium. The composition of these selection media and conditions for selection have been described in detail elsewhere (Adair *et al.*, 1998; Nairn and Adair, 1999). Alanosine (NSC-529469) was obtained from the Drug Synthesis and Chemistry Branch of the National Cancer Institute.

APRT targeting vectors and electroporation conditions

The pAG6 targeting vector has been described previously (Pennington and Wilson, 1991). The 3.9 kb *Bam*HI fragment of this pSV2gpt-derived vector carries a full-length Chinese hamster *APRT* gene that has been disrupted by a 33 bp polylinker insertion into the exon 3 *Xho*I site. The pAG6ins0.9 targeting vector was derived from pAG6 by insertion of a non-functional 0.9 kb *Nae*I fragment of the herpes simplex virus thymidine kinase gene into the *Nae*I site of the pAG6 polylinker insert, creating a 945 bp heterologous insertion at the exon 3 *Xho*I site. Linearization of pAG6 by cleavage at a unique *Hind*III restriction site within the polylinker insertion sequence creates an ends-in insertion vector configuration in which 18 and 11 nt terminal non-homologies flank each side of the DSB. Linearization of pAG6ins0.9 by cleavage at a unique *Sa*I site within its larger heterologous insert generates an insertion vector with 271 and 670 nt terminal non-homologies flanking each side of the DSB. The structures of both *APRT* targeting vectors, as well as the structure of the *APRT* target locus in CHO-ATS49tg, U9S50tg and E1KO7-5 cells, are shown in Figure 1.

Methods for electroporation of cells with *APRT* targeting vectors and selection of APRT⁺ recombinants have been described in detail elsewhere (Adair *et al.*, 1998; Nairn and Adair, 1999). To ensure the independence of all analyzed recombination events, only one ALASA-resistant clone from each original plate, seeded immediately after electroporation, was isolated for Southern analysis.

Southern analysis

Genomic DNA was isolated from APRT⁺ recombinants by conventional methods (Adair *et al.*, 1998; Nairn and Adair, 1999). Restriction endonucleases were used according to suppliers' directions (Amersham, Boehringer-Mannheim, New England Biolabs). Methods of electrophoresis in agarose gels, capillary blotting, hybridization and autoradiography were as previously described (Adair *et al.*, 1998; Nairn and Adair, 1999). *APRT* probes used (Figure 1) were: (i) a 3.9 kb *Bam*HI fragment containing the entire hamster *APRT* gene and (ii) a 1.4 kb fragment of the *APRT* gene extending from the intron 2 *Eco*RI site in intron to an *Xba*I site just downstream of the *APRT* polyadenylation signal (Adair *et al.*, 1998).

Acknowledgements

We thank Chaline Brown, Megan Lowery, April Orbison and Angela Bolt for technical assistance and Judy Ing for artwork. This work was supported by USPHS grants from the National Institutes of Health to G.M.A. (CA28711), R.S.N. (CA36361) and J.H.W. (GM38219) and Pilot Project and core support from NIEHS Center grant ES07784 to G.M.A.

and R.S.N. R.L.R. was the recipient of a Rosalie B.Hite pre-doctoral fellowship.

References

- Aboussekhra, A. and Wood, R.D. (1994) Repair of UV-damaged DNA by mammalian cells and *Saccharomyces cerevisiae*. *Curr. Opin. Genet. Dev.*, **4**, 212–220.
- Aboussekhra, A. *et al.* (1995) Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell*, **80**, 859–868.
- Adair, G.M., Stallings, R.L., Nairn, R.S. and Siciliano, M.J. (1983) High-frequency structural gene deletion as the basis for functional hemizygosity of the adenine phosphoribosyltransferase locus in Chinese hamster ovary cells. *Proc. Natl Acad. Sci. USA*, **80**, 5961–5964.
- Adair, G.M., Nairn, R.S., Wilson, J.H., Seidman, M.M., Brothman, K.A., MacKinnon, C. and Scheerer, J.B. (1989) Targeted homologous recombination at the endogenous adenine phosphoribosyltransferase locus in Chinese hamster cells. *Proc. Natl Acad. Sci. USA*, **86**, 4574–4578.
- Adair, G.M., Scheerer, J.B., Brothman, A., McConville, S., Wilson, J.H. and Nairn, R.S. (1998) Targeted recombination at the Chinese hamster *APRT* locus using insertion versus replacement vectors. *Somat. Cell Mol. Genet.*, **24**, 91–105.
- Bardwell, A.J., Bardwell, L., Tomkinson, A.E. and Friedberg, E.C. (1994) Specific cleavage of model recombination and repair intermediates by the yeast Rad1–Rad10 DNA endonuclease. *Science*, **265**, 2082–2085.
- Belmaaza, A. and Chartrand, P. (1994) One-sided invasion events in homologous recombination at double-strand breaks. *Mutat. Res.*, **314**, 199–208.
- Bessho, T., Mu, D. and Sancar, A. (1997a) Initiation of DNA interstrand cross-link repair in humans: the nucleotide excision repair system makes dual incisions 5' to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand. *Mol. Cell. Biol.*, **17**, 6822–6830.
- Bessho, T., Sancar, A., Thompson, L.H. and Thelen, M.P. (1997b) Reconstitution of human excision nuclease with recombinant XPF–ERCC1 complex. *J. Biol. Chem.*, **272**, 3833–3837.
- Bianchi, M., DasGupta, C. and Radding, C.M. (1983) Synapsis and the formation of paranemic joints by *E.coli* RecA protein. *Cell*, **34**, 931–939.
- Brookman, K.W. *et al.* (1996) *ERCC4* (XPF) encodes a human nucleotide excision repair protein with eukaryotic recombination homologs. *Mol. Cell. Biol.*, **16**, 6553–6562.
- Busch, D.B. *et al.* (1997) Phenotypic heterogeneity in nucleotide excision repair mutants of rodent complementation groups 1 and 4. *Mutat. Res.*, **383**, 91–106.
- Clarkson, J.M., Mitchell, D.L. and Adair, G.M. (1983) The use of an immunological probe to measure the kinetics of DNA repair in normal and UV-sensitive mammalian cell lines. *Mutat. Res.*, **112**, 287–299.
- Colaiacono, M.P., Paques, F. and Haber, J.E. (1999) Removal of one nonhomologous DNA end during gene conversion by a *RAD1*- and *MSH2*-independent pathway. *Genetics*, **151**, 1409–1423.
- Collins, A.R. (1993) Mutant rodent cell lines sensitive to ultraviolet light, ionizing radiation and cross-linking agents: a comprehensive survey of genetic and biochemical characteristics. *Mutat. Res.*, **293**, 99–118.
- Davies, A.A., Friedberg, E.C., Tomkinson, A.E., Wood, R.D. and West, S.C. (1995) Role of the Rad1 and Rad10 proteins in nucleotide excision repair and recombination. *J. Biol. Chem.*, **270**, 24638–24641.
- de Laat, W.L., Appeldoorn, E., Jaspers, N.G.J. and Hoeijmakers, J.H.J. (1998) DNA structural elements required for ERCC1-XPF endonuclease activity. *J. Biol. Chem.*, **273**, 7835–7842.
- de Laat, W.L., Jaspers, N.G. and Hoeijmakers, J.H. (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev.*, **13**, 768–785.
- Evans, E., Moggs, J.G., Hwang, J.R., Egly, J.M. and Wood, R.D. (1997) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J.*, **16**, 6559–6573.
- Fishman-Lobell, J. and Haber, J.E. (1992) Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene RAD1. *Science*, **258**, 480–484.
- Hasty, P., Rivera-Perez, J. and Bradley, A. (1992) The role and fate of DNA ends for homologous recombination in embryonic stem cells. *Mol. Cell. Biol.*, **12**, 2464–2474.
- Hasty, P., Rivera-Perez, J. and Bradley, A. (1995) Gene conversion during

- vector insertion in embryonic stem cells. *Nucleic Acids Res.*, **23**, 2058–2064.
- Hoy,C.A., Thompson,L.H., Mooney,C.L. and Salazar,E.P. (1985) Defective DNA cross-link removal in Chinese hamster cell mutants hypersensitive to bifunctional alkylating agents. *Cancer Res.*, **45**, 1737–1743.
- Huang,J.C., Svoboda,D.L., Reardon,J.T. and Sancar,A. (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc. Natl Acad. Sci. USA*, **89**, 3664–3668.
- Ivanov,E.L. and Haber,J.E. (1995) *RAD1* and *RAD10*, but not other excision repair genes, are required for double-strand break-induced recombination in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **15**, 2245–2251.
- Klein,H.L. (1988) Different types of recombination events are controlled by the *RAD1* and *RAD52* genes of *Saccharomyces cerevisiae*. *Genetics*, **120**, 367–377.
- Klein,H.L. (1995) Genetic control of intrachromosomal recombination. *BioEssays*, **17**, 149–159.
- Leung,W., Malkova,A. and Haber,J.E. (1997) Gene targeting by linear duplex DNA frequently occurs by assimilation of a single strand that is subject to preferential mismatch correction. *Proc. Natl Acad. Sci. USA*, **94**, 6851–6856.
- Matsunaga,T., Park,C.H., Bessho,T., Mu,D. and Sancar,A. (1996) Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. *J. Biol. Chem.*, **271**, 11047–11050.
- McWhir,J., Selfridge,J., Harrison,D.J., Squires,S. and Melton,D.W. (1993) Mice with DNA repair gene (ERCC-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. *Nature Genet.*, **5**, 217–224.
- Melton,D., Ketchen,A.M., Nunez,F., Bonatti-Abbondandolo,S., Abbondandolo,A., Squires,S. and Johnson,R. (1998) Cells from *ERCC1*-deficient mice show increased genome instability and a reduced frequency of S-phase-dependent illegitimate chromosome exchange but a normal frequency of homologous recombination. *J. Cell Sci.*, **111**, 395–404.
- Moggs,J.G., Yarema,K.J., Essigmann,J.M. and Wood,R.D. (1996) Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *J. Biol. Chem.*, **271**, 7177–7186.
- Mu,D., Hsu,D.S. and Sancar,A. (1996) Reaction mechanism of human DNA repair excision nuclease. *J. Biol. Chem.*, **271**, 8285–8294.
- Nairn,R.S. and Adair,G.M. (1999) Use of gene targeting to study recombination in mammalian DNA repair mutants. *Methods Mol. Biol.*, **113**, 499–517.
- Nairn,R.S., Adair,G.M., Christmann,C.B. and Humphrey,R.M. (1991) Ultraviolet stimulation of intermolecular homologous recombination in Chinese hamster ovary cells. *Mol. Carcinogen.*, **4**, 519–526.
- Osman,F. and Subramani,S. (1998) Double-strand break-induced recombination in eukaryotes. *Prog. Nucleic Acid Res. Mol. Biol.*, **58**, 263–299.
- Paques,F. and Haber,J.E. (1997) Two pathways for removal of nonhomologous DNA ends during double-strand break repair in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **17**, 6765–6771.
- Paques,F. and Haber,J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol. Biol. Rev.*, **63**, 349–404.
- Park,C.H., Bessho,T., Matsunaga,T. and Sancar,A. (1995) Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *J. Biol. Chem.*, **270**, 22657–22660.
- Pennington,S.L. and Wilson,J.H. (1991) Gene targeting in Chinese hamster ovary cells is conservative. *Proc. Natl Acad. Sci. USA*, **88**, 9498–9502.
- Prado,F. and Aguilera,A. (1995) Role of reciprocal exchange, one-ended invasion crossover and single-strand annealing on inverted and direct repeat recombination in yeast: different requirements for the *RAD1*, *RAD10* and *RAD52* genes. *Genetics*, **139**, 109–123.
- Rodriguez,K., Wang,Z., Friedberg,E.C. and Tomkinson,A.E. (1996) Identification of functional domains within the *RAD1*-*RAD10* repair and recombination endonuclease of *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **271**, 20551–20558.
- Rolig,R.L., Layher,S.K., Santi,B., Adair,G.M., Gu,F., Rainbow,A.J. and Nairn,R.S. (1997) Survival, mutagenesis and host cell reactivation in a Chinese hamster ovary cell *ERCC1* knock-out mutant. *Mutagenesis*, **12**, 277–283.
- Rolig,R.L., Lowery,M.P., Adair,G.M. and Nairn,R.S. (1998) Characterization and analysis of Chinese hamster ovary cell *ERCC1* mutant alleles. *Mutagenesis*, **13**, 357–365.
- Saffran,W.A., Greenberg,R.B., Thaler-Scheer,M.S. and Jones,M.M. (1994) Single strand and double strand DNA damage-induced reciprocal recombination in yeast. Dependence on nucleotide excision repair and *RAD1* recombination. *Nucleic Acids Res.*, **22**, 2823–2829.
- Saparbaev,M., Prakash,L. and Prakash,S. (1996) Requirement of mismatch repair genes *MSH2* and *MSH3* in the *RAD1*-*RAD10* pathway of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics*, **142**, 727–736.
- Sargent,R.G., Rolig,R.L., Kilburn,A.E., Adair,G.M., Wilson,J.H. and Nairn,R.S. (1997) Recombination-dependent deletion formation in mammalian cells deficient in the nucleotide excision repair gene *ERCC1*. *Proc. Natl Acad. Sci. USA*, **94**, 13122–13127.
- Schiestl,R.H. and Prakash,S. (1988) *RAD1*, an excision repair gene of *Saccharomyces cerevisiae*, is also involved in recombination. *Mol. Cell. Biol.*, **8**, 3619–3626.
- Schiestl,R.H. and Prakash,S. (1990) *RAD10*, an excision repair gene of *Saccharomyces cerevisiae*, is involved in the *RAD1* pathway of mitotic recombination. *Mol. Cell. Biol.*, **10**, 2485–2491.
- Selfridge,J., Pow,A.M., McWhir,J., Magin,T.M. and Melton,D.W. (1992) Gene targeting using a mouse HPRT minigene/HPRT-deficient embryonic stem cell system: inactivation of the mouse *ERCC-1* gene. *Somat. Cell Mol. Genet.*, **18**, 325–336.
- Sijbers,A.M. et al. (1996) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell*, **86**, 811–822.
- Smith,D.G. and Adair,G.M. (1996) Characterization of an apparent hotspot for spontaneous mutation in exon 5 of the Chinese hamster *APRT* gene. *Mutat. Res.*, **352**, 87–96.
- Stahl,F. (1996) Meiotic recombination in yeast: coronation of the double-strand-break repair model. *Cell*, **87**, 965–968.
- Sugawara,N., Paques,F., Colaiacovo,M. and Haber,J.E. (1997) Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc. Natl Acad. Sci. USA*, **94**, 9214–9219.
- Szostak,J.W., Orr-Weaver,T.L., Rothstein,R.J. and Stahl,F.W. (1983) The double-strand-break repair model for recombination. *Cell*, **33**, 25–35.
- Thompson,L.H., Busch,D.B., Brookman,K., Mooney,C.L. and Glaser,D.A. (1981) Genetic diversity of UV-sensitive DNA repair mutants of Chinese hamster ovary cells. *Proc. Natl Acad. Sci. USA*, **78**, 3734–3737.
- Tomkinson,A.E., Bardwell,A.J., Bardwell,L., Tappe,N.J. and Friedberg,E.C. (1993) Yeast DNA repair and recombination proteins Rad1 and Rad10 constitute a single-stranded-DNA endonuclease. *Nature*, **362**, 860–862.
- Valancius,V. and Smithies,O. (1991) Double-strand gap repair in a mammalian gene targeting reaction. *Mol. Cell. Biol.*, **11**, 4389–4397.
- van Duin,M., de Wit,J., Odijk,H., Westerveld,A., Yasui,A., Koken,H.M., Hoeijmakers,J.H. and Bootsma,D. (1986) Molecular characterization of the human excision repair gene *ERCC-1*: cDNA cloning and amino acid homology with the yeast DNA repair gene *RAD10*. *Cell*, **44**, 913–923.
- Weeda,G. et al. (1997) Disruption of mouse *ERCC1* results in a novel repair syndrome with growth failure, nuclear abnormalities and senescence. *Curr. Biol.*, **7**, 427–439.
- Wong,B.C., Chiu,S.K. and Chow,S.A. (1998) The role of negative superhelicity and length of homology in the formation of paranemic joints promoted by RecA protein. *J. Biol. Chem.*, **273**, 12120–12127.

Received May 24, 2000; revised August 17, 2000;
accepted August 18, 2000