 Nucleic Acids Research, 1999, Vol. 27, No. 12 **2521–2528**

# Palindromy is eliminated through a structure-specific recombination process in rodent cells

**Susanna M. Lewis***

Program in Genetics and Genomic Biology, Hospital for Sick Children Research Institute and Department of Immunology, University of Toronto, Toronto, Canada

Received January 29, 1999; Revised and Accepted April 20, 1999

## ABSTRACT

**Higher eukaryotes are proficient at remodeling palindromic DNA. As shown here, a fully palindromic 15.4 kb circular DNA can be introduced into rodent cells with the novel result that the molecule is reproducibly and site-specifically converted to a monomeric circle. The dimer-to-monomer conversion has not been described previously, and in particular is undetectable in *Escherichia coli*. Comparative DNA sequence analyses of the new junctions found within the monomer circles suggest that the resolution process involves formation of hairpin DNA structures followed by the introduction of single-strand nicks near their termini. By extension, hairpin nicking combined with non-homologous end-joining may be important as a general mechanism for the maintenance of genomic stability in mammalian cells. It is suggested that the absence of a comparable strategy for coping with palindrome-induced structure in *E.coli* and other unicellular organisms underlies a fundamental difference in DNA sequence organization in the genomes of prokaryotes versus higher eukaryotes.**

## INTRODUCTION

Palindromic DNA is biologically compromised. Anywhere a particular sequence is contiguous to an inverted copy of itself, there is the potential for intrastrand base pairing. Because the self-complementary nature of palindromic DNA allows it to adopt a non-B-form structure under certain circumstances, palindromy can interfere with genetic function in both prokaryotes and eukaryotes (reviewed in 1,2). The presence of palindromic sequences presents a special threat to genome stability and might be considered an exceptional type of DNA damage. In many situations, DNA containing a palindromic sequence is indistinguishable from native B-form DNA. It is only when DNA has the opportunity to form intrastrand base-pairs that a latent ability to obstruct replication and cause rearrangements becomes manifest.

The high instability of palindromic sequences, particularly in bacteria, makes them difficult, and in some cases impossible, to study by standard molecular genetic techniques. As a consequence, the impact that a palindromic DNA sequence may have on genome maintenance and transmission is not widely appreciated. Nevertheless, several lines of evidence have suggested that the problem of palindromy is significant and that different organisms may have developed specific strategies for coping with DNA sequences that comprise a perfect, uninterrupted inverted repeat (3,4).

Here, both qualitative and quantitative differences between *Escherichia coli* and vertebrate cells in their ability to convert a fully palindromic circular DNA to a non-palindromic form have been uncovered. It is possible to show that hamster or mouse cells reproducibly reduce a palindromic dimer circle to a simple monomeric circle through illegitimate recombination. *Escherichia coli*, in contrast, is incapable of either maintaining or resolving a palindromic plasmid at any detectable frequency. An analysis of the ability of either cell type to manipulate various forms of identical DNA is presented below.

These studies indicate that the repair of genomes possessing palindromy can be accomplished by fundamentally different approaches. Acquisition of hairpin structure is thought to be a key step in palindrome processing in either rodent or *E.coli* cells (5 and references cited therein). Here, evidence indicates that differences arise in what occurs after hairpin structure forms. It is suggested that different repair paradigms may well account for the marked difference in sequence organization between bacteria and higher eukaryotes.

## MATERIALS AND METHODS

### Preparation of transfected DNA: monomer circles, open linear DNA, hairpin linear DNA and palindromic dimer circles

The polyoma-based shuttle plasmid pJH298 (6) was the source plasmid for all DNA in this study. Palindromic dimer circles were prepared by incubating a gel-purified 7.8 kb fragment generated by *Sal*I and *Bam*HI co-digestion, at a concentration of 35 μg/ml, with T4 DNA ligase. The 15.6 kb palindrome dimer circles were then purified on an agarose gel. Hairpin linear and open linear DNA was prepared from this same pJH298 fragment, as described (7).

*Correspondence should be addressed to: Immunology Department, Medical Sciences Building, 1 King's College Circle, University of Toronto, Toronto, ON M5S 1A8, Canada. Tel: +1 416 978 4230; Fax: +1 416 978 1938; Email: susanna.lewis@utoronto.ca

### Cell lines

204-1-8 and S41 are Abelson murine leukemia virus (A-MuLV)-transformed pre-B type cells derived from a wild-type mouse and a mouse with severe combined immunodeficiency (*scid*) defect, respectively. These cells, along with NIH 3T3 cells, were obtained from sources described previously (7). CBP-9 and SCD-9 are wild-type and *scid* mouse embryonic fibroblast cell lines, respectively, and were a generous gift of Dr Gillian Wu (Ontario Cancer Institute). CHO-K1 and a derivative, CHO-K1$^{d}$, are Chinese hamster ovary cell lines, the original source of which was Dr Penny Jeggo (University of Sussex).

### Transfections

204-1-8 and S41 cells were transfected by the method given previously (8), with DNA in the amounts specified here. All other cell lines were transfected using Protocol 1 described in Sambrook *et al*. (9): cells were treated for 6 min with a solution containing 1 mg/ml DEAE–dextran and DNA in the specified quantities.

### Transformations

*Escherichia coli* strain DH10B was used [*F*$^{-}$ mcrA Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80d*lac*ZΔM15 Δ*lac*X74 *end*A1 *rec*A1 deoR Δ(*ara*, *leu*)7697 *ara*D139 *gal*U galK *nup*G *rps*L λ$^{-}$]. Competent bacteria were either purchased (Bethesda Research Laboratories) or prepared following protocols provided by the supplier. Standard transformation procedures were used; details are specified in Lewis (7). Selective media contained 10 μg/ml chloramphenicol and/or 100 μg/ml ampicillin.

### Calculation of recovery

The recoveries of products from mammalian cell transfections were determined based on colony counts after *E.coli* transformation. The number of colonies expected for 100% recovery was the number of pJH298 (or related) DNA molecules transfected into the rodent cell culture, corrected for the size of the aliquot of harvested material that was transformed into *E.coli* and normalized to the per molecule recovery obtained by a parallel transformation of *E.coli* with 10 pg pUC19. For dimer DNA, the recovery was calculated according to an expected two resolution circles per input dimer. By these calculations, the values in Tables 2 and 3 provide a direct comparison of monomer circle recoveries for each different species of DNA (open linear, hairpin, inverted dimer and monomer circle) as transfected into NIH 3T3 or other rodent cells; variations in *E.coli* transformation from one experiment to the next have been factored out.

### DNA analysis

DNA sequence determination was as described in Lewis (7). All insertions from this study, as well as those previously defined (7), were analyzed for this study by searching the GenBank database maintained by the NCBI with the Blastn algorithm for identities (10). Insertions were also analyzed against the pJH298 sequence itself. Vector identities of greater than nine nucleotides are indicated by the single asterisks in Figure 2.

## RESULTS

### A novel palindrome resolution function is exhibited by rodent cells

The fact that *E.coli* is not transformed by DNA that is composed of a head-to-head and tail-to-tail dimer of a plasmid vector illustrates the degree to which palindromy can be incompatible with replicon maintenance (11–13). The prohibition against inverted repeat structure is so profound in the case of extrachromosomal elements that molecular biologists commonly exploit the situation in 'directional' or 'forced' cloning. Here a DNA fragment with heterologous ends is ligated to a correspondingly cut vector DNA. The reason that it is possible to specifically recover the desired product is because palindromic vector-to-vector ligation products, though formed, are without biological function in *E.coli*.

To investigate whether a palindromic DNA circle suffers a similar fate in mammalian cells it was necessary to construct the test DNA *in vitro* on a preparative scale. A circular, 15.6 kb DNA was prepared through self-ligation of a 7.8 kb linear fragment (Materials and Methods). The linear DNA was obtained by *Bam*HI and *Sal*I digestion of a plasmid DNA so that any two molecules must ligate to one another in inverted orientation. The parental plasmid and the palindromic dimer are shown in Figure 1 (left column, top and bottom diagrams, respectively). The palindromic DNA circle was constructed from a shuttle vector (6), so that the same DNA preparation could be tested in either rodent or bacterial cells.

Consistent with earlier observations (11–13), the palindromic dimer DNA was not maintained in *E.coli*. The palindromic DNA preparation, introduced into *E.coli* strain DH10B by either of two standard methods, gave a transformation efficiency of $<2 \times 10^{3}$ ampicillin-resistant colonies/μg (Table 1). These few transformants (of which a small number were also chloramphenicol resistant) proved to be due to a ligation artifact. DNA obtained from a large number of randomly selected colonies was analyzed and found to consist of a single copy of the vector into which a random *Sal*I/*Bam*HI-cut DNA fragment originating from the *E.coli* chromosome had been cloned (data not shown). The artifactual transformants were evidently generated during palindromic dimer preparation as a consequence of trace amounts of *E.coli* chromosomal DNA in the original plasmid DNA. (The artifactual transformants did not greatly confound the analysis because they were easily identified on the basis of a *Bam*HI and *Sal*I co-digestion.)

To verify that the extremely low transformation efficiency of the palindromic dimer DNA preparations was an attribute of DNA itself, mixing experiments were performed. Co-transformation with pUC19 DNA confirmed that palindromic dimer DNA preparations were free of factors that would interfere with bacterial transformation (Table 1, experiment 2).

After testing in bacteria, palindromic dimer DNA was transfected into mouse or hamster cell lines. Recipient cells represented different tissue types, including two mouse pre-B-like cell lines (204-1-8 and S41), several mouse fibroblastoid cell lines (NIH 3T3, SCD-9 and CBP-9) and an epithelial-like hamster cell line (CHO-K1). Two days after transfection with the palindromic dimer DNA, material was recovered from the cells by an alkaline lysis procedure and used to transform *E.coli*. Ampicillin-resistant colonies arose and, notably, a

considerable fraction of these (up to 90%) were also chloramphenicol resistant, suggesting that the input DNA had become rearranged. Colony numbers were not reduced by *Dpn*I treatment of the DNA prior to *E.coli* transformation (8). The *Dpn*I resistance of the transforming DNA indicated that the detected products were replicated in a eukaryotic cell and thus preexisted the transformation step itself (data not shown).

**Table 1.** Comparison of transformation efficiency of circular, linear and hairpin DNA and palindromic dimers introduced into *E.coli* DH10B

| Experiment | DNA (amount) | Colonies (per μg)[a] |
|---|---|---|
| 1 ($CaCl_2$) | pUC19 (10 pg) | $5.0 \times 10^8$ |
| | Hairpin (5 ng) | $<2 \times 10^2$ [b] |
| | Dimer (60 ng) | <17[b] |
| 2 (electroporation) | pUC19 (10 pg) | $2.2 \times 10^9$ |
| | Hairpin (5 ng) | $<2 \times 10^2$ [b] |
| | Dimer (60 ng) | $1.9 \times 10^3$ |
| | Hairpin[c] + pUC19[d] | $1.7 \times 10^9$ [e] |
| | Dimer[f]+ pUC19[d] | $1.8 \times 10^9$ [e] |
| 3 (electroporation) | pJH298 (37.5ng) | $3.2 \times 10^7$ |
| | Linear (20 ng) | $2.1 \times 10^4$ |
| | Hairpin (5 ng) | $<2 \times 10^2$ [b] |
| | Dimer (60 ng) | $6.7 \times 10^2$ |
| 4 (electroporation) | pUC19 (10 pg) | $1.8 \times 10^9$ |
| | Dimer (60 ng) | $1.2 \times 10^3$ |
| 5 (electroporation) | pUC19 (10 pg) | $2.9 \times 10^8$ |
| | Linear (100 ng) | $1.7 \times 10^2$ |
| | Dimer (100 ng) | $2.1 \times 10^2$ |

[a]Ampicillin-resistant colonies; numbers include the artifactual products discussed in the text.
[b]No transformants detected.
[c]An aliquot of 5 ng of hairpin DNA was used.
[d]An aliquot of 10 pg of pUC19 was used.
[e]Number given is per μg pUC19 (other DNA was non-transforming).
[f]An aliquot of 60 ng of dimer DNA was used.

To analyze the events by which the palindromic dimer DNA had become reconfigured in rodent cells, chloramphenicol-resistant colonies were selected at random and DNA preparations were mapped by restriction enzyme digestion. The plasmids were all monomeric in size and, with one exception, possessed a simple structure consistent with a resolution event (Fig. 1, bottom right box). Recombination had apparently occurred very close to the two symmetry axes of the original palindromic dimer, although the breakpoints in the monomer products were variable. This was indicated by the fact that many of the monomeric circles could be linearized with *Bam*HI or *Sal*I. In the input dimer, one axis of symmetry contains a unique cleavage site for *Bam*HI, while at the other, a unique cleavage site for *Sal*I is present (Fig. 1, bottom right box).

In summary, palindromic DNA was not transforming when introduced directly into *E.coli*. Nor, to the limits of the assay, were monomeric resolution products recovered after direct

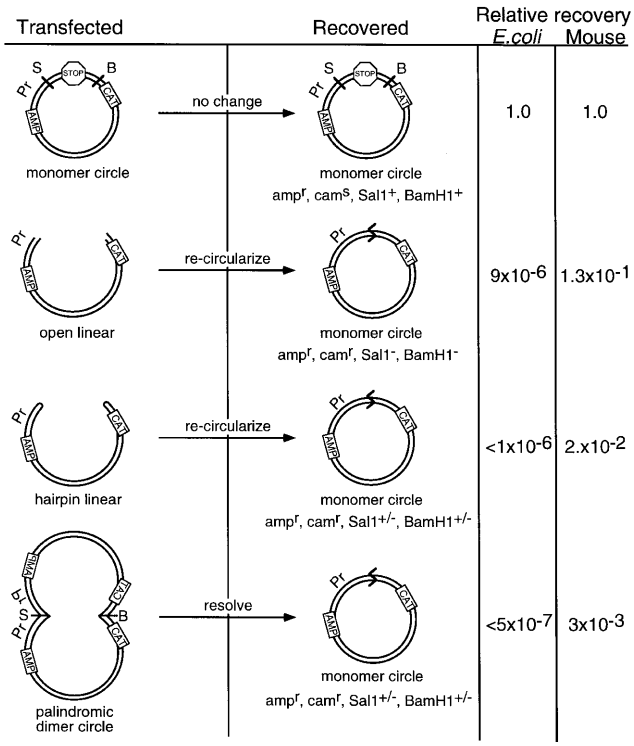


**Figure 1.** (Left) Diagram of the DNAs used in this study. The top diagram shows the parental pJH298 plasmid (6). B, *Bam*HI (not shown are two sites located 50 bp apart); S, *Sal*I; Pr, a promoter; CAT, structural gene for chloramphenicol acetyltransferase; STOP, a transcriptional terminator. The plasmid and its derivatives also contain a functional β-lactamase gene (AMP), as well as polyoma and OriC replication origins (not indicated). Linear, hairpin and palindromic dimer circle molecules are shown in descending order. (Middle) Monomer circles produced by illegitimate recombination. The characteristics of each, with respect to drug resistance and the two axis-specific restriction sites are given for each. Recovery of monomer circles scored for each form of DNA relative to an input monomer. For *E.coli*, values are from Table 1, experiment 3, and exclude molecules determined to be an *in vitro* ligation artifact (see text). For mouse, values given are averages from Table 2.

transformation. In contrast, rodent cells reproducibly converted a completely palindromic dimer circle to monomeric form through recombination at or very near sites specified by the palindromic axes of symmetry. This novel function has not been previously reported for any other organism.

## DNA sequence analysis of the monomer junctions suggests that a hairpin nicking step is involved in dimer resolution

The presence of P (for palindromic) insertions is a distinctive feature of the junctions that arise when hairpin-terminated DNA becomes joined. For example, P insertions are observed when DNA possessing preformed hairpin ends is recircularized in mouse cells (7). P insertions are also seen in joints arising from site-specific V(D)J recombination in lymphoid cells (14,15), a process that requires connection of hairpin-terminated cleavage products (16,17). For these two established examples of hairpin end-joining, the presence of P insertions is evidence that, prior to ligation, the covalently closed hairpin termini were opened by the introduction of a single-stranded nick positioned near, but not exactly at, the tip (see 7,15,18 for further discussion).

The DNA sequence of the palindromic dimer resolution products obtained in this study gave indications that here too, hairpin nicking was involved. As shown in Figure 2A (bold typeface), small amounts of inverted-repeat sequence remained at many of the resolution junctions. The residual palindrome was always quite short, never greater than 9 bp. Interestingly, the residual palindromy showed cell type-specific differences, where lymphoid cell products possessed rather longer P nucleotide stretches than did those arising from transfection of non-lymphoid cells. This trend mirrored differences in the P nucleotides seen in hairpin linear recircularization experiments (as summarized in table 2 of 7). The presence of P nucleotides within the junctions formed by palindromic dimer resolution and the differences in P nucleotides associated with different cell types suggested that dimer resolution and hairpin end-joining are mechanistically related processes (7; Fig. 2A).

Perhaps even more striking was the observation that similarity between the products of hairpin end-joining and palindromic dimer resolution extended to actual sequence identity in a number of cases. Isolates arising through palindrome dimer resolution were found that were indistinguishable from those that arose through hairpin end-joining. Recombinants presented in Figure 2B provide examples from hairpin end-joining experiments (7) that are seen to exactly match one or more of the independently derived dimer resolution products listed in Figure 2A. Accordingly, the dimer-to-monomer conversion is highly likely to involve some or all of the same biochemical operations that take place in hairpin end-joining.

## Monomer junctions contain characteristic, but atypical insertions

DNA sequence analysis revealed junctional insertions in somewhat less than half of the monomer circle junctions (Fig. 2, central column). These were distinct from palindromic nucleotides and fell into two general categories: small insertions that were comprised of fewer than ~25 bp of apparently random sequence; larger insertions, among which vector-derived patches could be detected.

Small, random insertions are associated with vector recircularization in rodent cell lines, when either open or hairpin linear DNA is transfected (7; unpublished observations). The

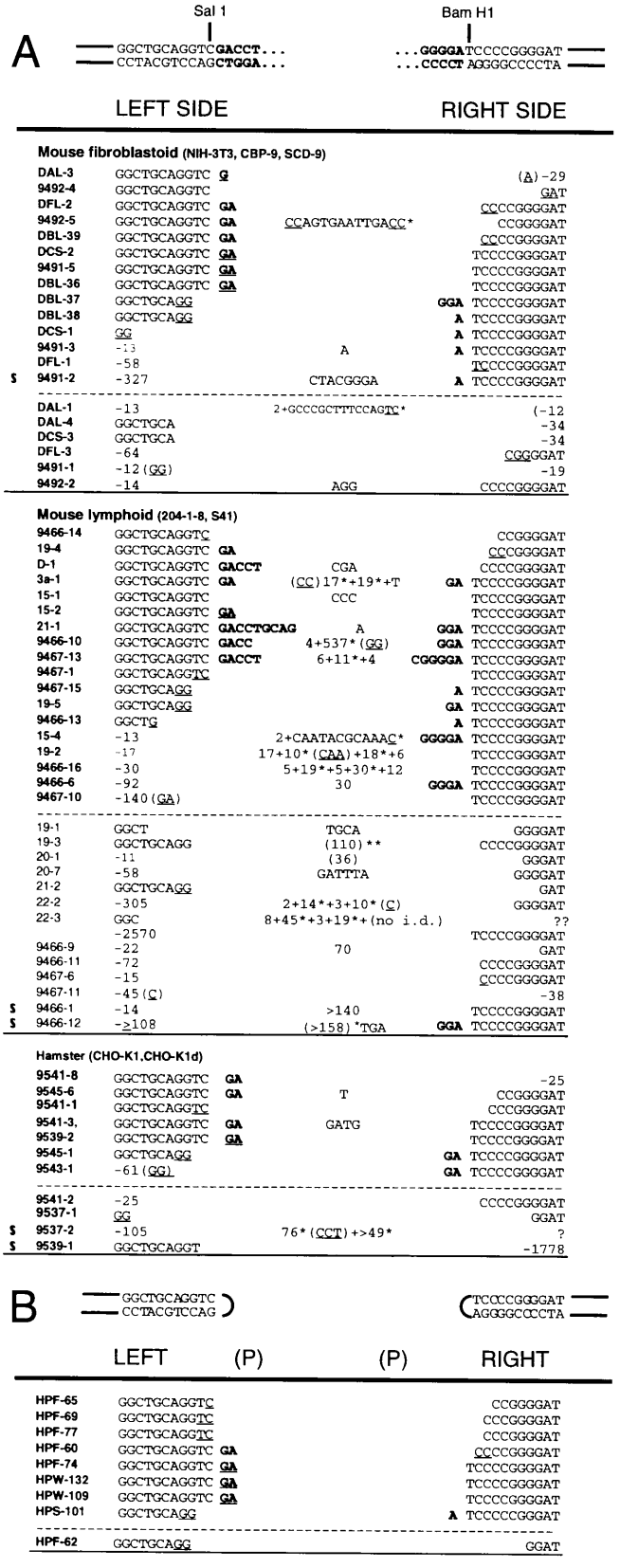


**Figure 2.** DNA sequence analysis of resolution junctions. All repeat junctions shown were derived in separate transfection experiments. Each junction sequence reads continuously from left to right, but is displayed in a gapped format in order to indicate the origin of every nucleotide (see header). Nucleotides for which the assignment is ambiguous (indicating that recombination took place in a region of microhomology) are underlined. Inserts are shown in the middle column. Single asterisks denote patches of sequence appearing as insertions that are vector derived. As necessary, the length rather than the DNA sequence of a patch is given. Where sequence information has not been provided, any microhomology is indicated by the underlined nucleotides in parentheses. In the left-most column, s designates isolates that were chloramphenicol sensitive. (**A**) DNA sequences of resolution junctions (as in Fig. 1, bottom row) resulting from palindromic dimer DNA transfection. The header gives the sequence at the *Sal*I symmetry axis (left) and at the *Bam*HI symmetry axis (right). Nucleotides just past the symmetry axes are analogous to P nucleotides (see text) and are shown in bold. Joints are grouped according to the cell type from which they were isolated. Within each section, the dashed line separates recombinants where the sequence from neither symmetry axis was retained (grouped below the dashed line) from those in which at least one axis is still present (above the dashed line). The minimum number of P nucleotides is indicated because every ambiguous assignment was made to the more deleted end. Missing nucleotides, as scored from the *Sal*I symmetry center (left) or *Bam*HI symmetry center (right), are indicated by the numbers. (**B**) Sequences of a subset of junctions arising from hairpin recircularization (7). Each junction shown is an exact match to one or more of the junctions given in (A).

remaining insertions, however, were qualitatively different from those that had previously been encountered. Patches of sequence that were derived from either the vector itself or the genome were apparent (Fig. 2, central column, asterisks). Also, insertions that were not obviously vector-derived were atypical with respect to size. In one junction (marked by two asterisks in Fig. 2) the insert matched a segment of mouse genomic DNA at 38 of 43 positions (Materials and Methods; 10).

The insertions observed in resolution products suggest that DNA ends created when the palindromic dimer is broken are available for transactions that do not necessarily occur during hairpin end-joining. More generally, however, these data indicate that processing of palindromic DNA is highly related to hairpin end-joining, likely including identical steps. This is because, whether or not unusual insertions were present, other features of the junctions (P nucleotide length and frequency and the extent of deletion at the joint) were consistent with the products of hairpin end-joining. Extra steps indicated by the presence of unusual insertions may be relevant to the resolution of palindromes in a genomic context (see Discussion).

### The different proficiency with respect to palindromic dimer resolution between *E.coli* and mice can be explained by differences in hairpin DNA processing

Palindromic dimer DNA is reproducibly reduced to monomeric form in rodent cells, whereas there is no evidence for a similar capability in *E.coli*. To explore the basis for this difference, the recovery of monomer circles upon introduction of linear, hairpin linear and palindromic dimer DNA preparations into either *E.coli* or NIH 3T3 cells was quantified.

As shown in Figure 1, monomeric circles can, in theory, be produced by illegitimate recombination from each of three different input forms of a vector DNA: 'open' linear and hairpin linear DNA and palindromic dimers. Because these may represent related steps in palindrome processing, it was of interest to compare the recovery of monomer circles from either *E.coli* or NIH 3T3 cells when challenged with each of these various DNAs. As tested here, the open linear, hairpin linear and palindromic dimer DNA possessed the identical primary sequence (see Materials and Methods).

The recovery of monomeric circles from NIH 3T3 cells was quantified by harvesting transfected DNA and counting colonies obtained after transformation of *E.coli*. Ampicillin-resistant colonies were scored for experiments with the parental plasmid, whereas doubly chloramphenicol- and ampicillin-resistant colonies were scored for all other forms. Colony counts were converted to per molecule recoveries as described in Materials and Methods. This approach, where the transfected DNA was introduced into bacteria for analysis, was possible only because every type of DNA apart from the circular, monomer plasmid was poorly transforming for *E.coli* (Table 1). In the case of the hairpin linear DNA experiments, data have been reported previously (7) and raw numbers have been recalculated for consistency (details in Materials and Methods).

As shown in Table 2, and summarized in Figure 1, transfection of NIH 3T3 cells with monomer circular, open-ended linear or hairpin linear DNA or palindromic dimers gave a relative per molecule recovery that was roughly 300:40:7:1. The per molecule recoveries provide some reflection of actual resolution efficiency, as seen by the fact that transfection with either 10 or 120 ng of palindromic dimer DNA gave comparable results, on a per molecule basis (Table 2). The rank order for the various transfected DNAs with respect to the final yield of monomer circles indicated that end-joining (as required for recircularization of open linear DNA) was not limiting for either the hairpin linear DNA recircularization or for palindromic dimer resolution. Likewise, hairpin opening [which is required for recircularization of the hairpin linear DNA in addition to non-homologous end-joining (NHEJ)] was not limiting for palindromic dimer resolution.

**Table 2.** Comparative recovery of monomer circles from circular, linear and hairpin DNA and palindromic dimers transfected into NIH 3T3 cells

| Experiment | DNA (amount) | Recovery (per molecule)[a] |
|---|---|---|
| 1 | Circle (10 ng) | 0.06 |
| 2 | Circle (10 ng) | 0.10 |
| | Circle (10 ng) | 0.18 |
| | Dimer (120 ng) | $7.7 \times 10^{-5}$ |
| | Dimer (120 ng) | $4.9 \times 10^{-4}$ |
| 3 | Circle (10 ng) | 0.10 |
| | Circle (10 ng) | 0.04 |
| | Linear (10 ng) | 0.02 |
| | Linear (10 ng) | $5.4 \times 10^{-3}$ |
| | Hairpin (10 ng) | $1.6 \times 10^{-3}$ |
| | Hairpin (10 ng) | $2.9 \times 10^{-3}$ |
| | Dimer (10 ng) | $<1.6 \times 10^{-4}$ |
| | Dimer (10 ng) | $7.9 \times 10^{-4}$ |

[a]Monomeric circles recovered per circle, linear, hairpin or dimer molecule transfected.

A less extensive analysis was carried out with additional mouse and hamster cell lines (Table 3). Palindromic dimer resolution could be detected in each case, and the various transfected DNAs showed similar trends with respect to the relative recovery of product monomer circles.

Results of the comparable test in *E.coli* are shown in Table 1, which reports relative recoveries of ampicillin-resistant colonies from each different type of DNA. In no experiment was even a single colony recovered upon transformation of hairpin linear DNA; however, to obtain true values for dimer DNA transformations as well as open linear DNA it was necessary to analyze every transformant by restriction digestion and, where appropriate, DNA sequence analysis. This was because, as discussed above, an artifactual ligation product contaminated the palindromic dimer DNA preparations and, in addition, a fraction of the colonies arising from preparations of linear DNA contained low levels of contaminating uncut pJH298. For experiment 5 (Table 1) it was determined that no resolution products existed among the transformants scored after transformation with palindromic dimer DNA (resolved monomers recovered at <10/μg of dimer DNA). In contrast, for experiment 5, recircularized linear molecules (as verified by DNA sequence analysis) were recovered at a frequency of 90/μg of DNA. This analysis showed, as have other studies (19), that

*E.coli* is far less proficient in NHEJ than are mammalian cells. Further, these data suggest that NHEJ, as assessed by recircularization of open linear DNA, may not be limiting in the rejoining of hairpin DNA ends. Recircularization of open linear DNA, though low, was detectable, whereas recircularization of hairpin ends or the resolution of palindromic dimer DNA was not.

**Table 3.** Recovery of monomer circles from linear and hairpin DNA and palindromic dimers transfected into mouse or hamster cell lines

| Cell line | DNA ($n$)[a] | Recovery (per molecule)[b] |
|---|---|---|
| Mouse fibroblastoid | | |
| NIH 3T3 | Circle ($n$ = 5) | 0.10 |
| | Linear ($n$ = 2) | $1.3 \times 10^{-2}$ |
| | Hairpin ($n$ = 2) | $2.3 \times 10^{-3}$ |
| | Dimer ($n$ = 4) | $3.3 \times 10^{-4}$ |
| CBP9 | Dimer ($n$ = 1) | $6.9 \times 10^{-4}$ |
| SCD9 | Dimer ($n$ = 1) | $1.8 \times 10^{-4}$ |
| Mouse lymphoid | | |
| 204-1-8 | Hairpin ($n$ = 2) | $3.5 \times 10^{-5}$ |
| | Dimer ($n$ = 5) | $1.1 \times 10^{-7}$ |
| S-41 | Hairpin ($n$ = 2) | $8.4 \times 10^{-5}$ |
| | Dimer ($n$ = 3) | $1.3 \times 10^{-6}$ |
| Hamster epithelioid | | |
| CHO-K1[d] | Circle ($n$ = 2) | $7.6 \times 10^{-2}$ |
| | Dimer ($n$ = 2) | $5.8 \times 10^{-5}$ |
| CHO-K1 | Circle ($n$ = 2) | $3.9 \times 10^{-2}$ |
| | Dimer ($n$ = 2) | $6.5 \times 10^{-7}$ |

[a]$n$, number of transfections.
[b]Number of monomeric circles recovered per circular, linear, hairpin or dimer molecule transfected. Number is the average of the indicated number of transfections.

One can deduce from these observations that differences between palindrome processing in *E.coli* and NIH 3T3 cells may not be entirely attributable to differences in NHEJ. Another factor, related to hairpin opening, may also contribute. Possibilities to consider are that efficient hairpin nicking may not occur in bacteria, a hairpin nicking function may be present that creates ends that are refractory to ligation, or perhaps hairpin nicking activities are present but are not effectively integrated into the NHEJ pathway for double-strand break repair.

## DISCUSSION

The notion that palindromic DNA, due to its potential for self-pairing into an aberrantly folded structure, is a form of DNA damage has received increased attention in the last several years (reviewed in 1,2). Evidence that there exists a mechanism to alleviate the effect of large secondary structures that form at palindromes has been most fully developed for *E.coli*, (3). The extent to which palindromes also create problems in mammalian cells and whether sequences bearing palindromes may be subject to specific repair processes is only beginning to be explored (20). The present study reveals that mammalian cells contend with palindromy in a quite efficient and reproducible manner. Here, it can be seen that rodent cells are capable of converting a fully palindromic plasmid to a monomeric circle. No similar activity is detected in *E.coli*, suggesting that bacteria may employ fundamentally different tactics when confronted with a problematic palindromic DNA.

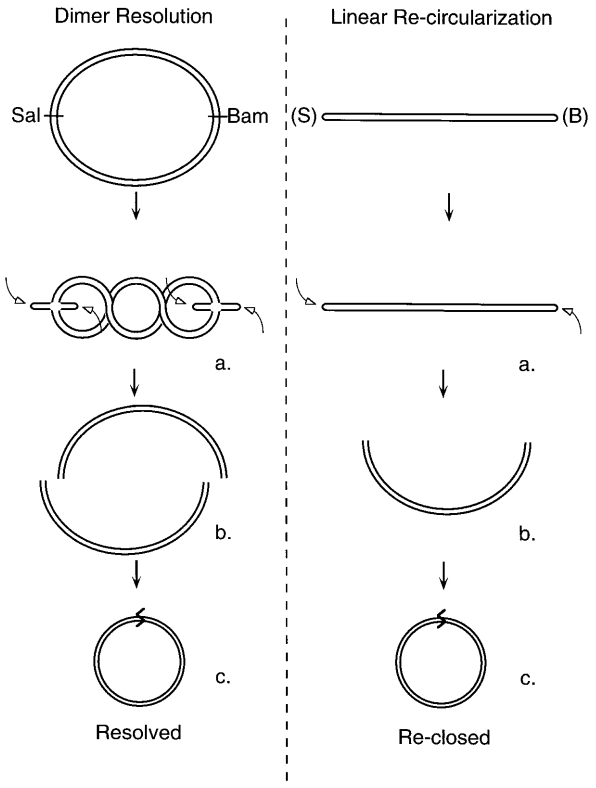


**Figure 3.** (Left) Mechanism of palindromic dimer resolution. Axis-specific recombination of the palindromic dimer may follow cruciform extrusion via the following steps: (a) hairpins are nicked at a near tip position; (b) nicked cruciforms are resorbed, leading to formation of open linear DNA; (c) linear molecules are circularized by end-joining. (Right) Steps in hairpin linear recircularization: (a) hairpins are nicked at a near tip position; (b) this leads to the formation of open linear DNA; (c) linear molecules are circularized by end-joining.

### Mechanism of inverted dimer resolution

As is apparent from the analysis presented here, the two symmetry centers in the circular palindrome are hotspots for illegitimate recombination. The symmetry centers would seem to be demarcated somehow and an obvious inference is that intrastrand base pairing creates a structural alteration at the recombination sites. A model for the resolution of palindromic dimers in mammalian cells is depicted in Figure 3 (left). Through cruciform extrusion at the two axes of symmetry, four hairpin structures arise. This creates targets for a hairpin nicking activity that introduces a single-strand break very near the tips of each structure. The net result is that the multiply nicked dimer DNA has acquired two axis-specific double-strand

breaks. Unit length linear molecules, liberated by branch migration, then become circularized through the action of NHEJ functions.

In this model, palindromic dimer resolution involves many of the same operations by which a hairpin linear DNA is recircularized (Fig. 3, right; 7). The notion that most steps in dimer resolution and hairpin end-joining are shared is experimentally supported by the observed similarities in the monomeric circles produced in either case (Fig. 2). Indeed, as noted in Results, some of the circles generated in rodent cells are identical.

There is an incomplete overlap between dimer resolution and hairpin end-joining nonetheless. Vector-derived and genome-derived insertions are frequently included in the resolution junctions (Fig. 2; 7), although similar insertions are rare when the corresponding hairpin linear DNA is transfected into the same cells (7). To gain insight into what sort of transactions might give rise to the palindrome-associated inserts, it is informative to examine their properties in some detail. Vector-derived patches, where found, occurred in a variety of complicated arrangements. They could be seen as a single short stretch or as conglomerates of short segments pieced together in arbitrary orientations and interspersed with segments of non-vector origin (Fig. 2). In some cases a long continuous stretch of vector sequence comprised the junctional insertion. For these identified inserts, it was possible to determine that many exhibited a 'micro-homology' of one or two nucleotides at the point where they were connected to adjacent sequences (microhomologies are underlined in Fig. 2A). Though not proved in every case, it was clear for some recombinants (e.g. 9466-10) that, according to restriction site analysis, the vector-derived insertions were copies of sequences that remained at their normal location in the plasmid backbone. Thus the junctional insertion was either copied from the vector or had been incorporated as a fragment derived from a separate molecule, i.e. it was not translocated from one spot in the vector to another.

Although other junctional insertions had no discernible match to the plasmid, these may have arisen from events that were similar to those generating the vector-derived insertions. This was suggested when it was discovered by searching the available public database (Materials and Methods; 10) that one insertion contained a segment with a 38 bp identity to a mouse genomic sequence (marked by two asterisks in Fig. 2). Possibly, other long insertions lacking vector identities might likewise include genomic DNA (presumably derived from a still unsequenced region).

The vector-derived and genome-derived insertions in palindrome resolution junctions indicate that 3'-ends are available for extension or ligation before NHEJ takes place. This could occur if the four hairpin tips created by cruciform extrusion (Fig. 3) are not always cut simultaneously. According to both theoretical and experimental considerations a cruciform will not undergo an instantaneous collapse if single-strand breaks are introduced within the extruded domain (12,21). Thus there is no reason to suppose that in the model shown in Figure 3, hairpin nicking should be synchronous. For dimer resolution, between steps (a) and (b) (Fig. 3) a liberated 3'-end might participate in transactions other than direct end-joining. Borrowing from the literature on illegitimate recombination in plants, the filler sequences could be acquired when a free 3'-end becomes annealed at a fortuitous microhomology in either the genome or the vector itself, elongated and released prior to end-joining (22,23). As discussed, the nature of the insertions suggests that repeated dissociation of newly synthesized DNA from its template can evidently occur, indicating that multiple rounds of invasion/extension are a central feature of the process. The non-processive, synthesis-associated variation in the NHEJ pathway is not as evident in mammalian cells as in plants, when end-joining is assessed with extrachromosomal substrates (22); however, the present study indicates that the possibilities for interim invasion/extension of DNA ends prior to ligation may be enhanced during palindrome processing.

Evidence that similar insertions accumulate in junctions arising from rearrangements of chromosomally located palindromes has been reported (20). Insertions, comprised of sequences that lie near, but non-adjacent, to the symmetry center of a large palindrome were observed for a large transgenic inverted repeat (20). Thus the variant aspect of end-joining revealed by the accumulation of insertions in palindrome rearrangements by the extrachromosomal assay may also be significant in a physiological context.

### Palindrome resolution, genome stability and evolution

For *E.coli*, it has been proposed that if a palindromic sequence adopts a hairpin structure during replication and thereby halts fork progression, the SbcCD nuclease can cut near the tip of the hairpin to initiate RecA-dependent post-replication repair. The SbcCD-mediated pathway is thought to fully reconstitute the palindrome within the replicated genome (3,5). If repair is successful, however, the problematic sequence remains and, overall, a genome containing the palindrome will be disadvantaged due to underlying replication difficulties. In the absence of repair, palindromic sequences that have a high probability of intrastrand base pairing during replication will either not become replicated or may be fully deleted due to replication by-pass (reviewed in 1; see 24 for a recent discussion).

Problems with the replication of palindromic DNA, manifested as palindrome instability, are observed not only in *E.coli*, but also in other organisms, including yeast and humans (25,26; reviewed in 1,2). However, several studies, including this one, indicate that for mammalian cells palindrome remodeling need not involve either homology-dependent post-replication repair or replication by-pass (4,20). By breaking a palindrome at the symmetry center and rejoining the ends through NHEJ, the mammalian strategy creates changes that specifically interrupt the perfect symmetry of the inverted repeat (4,27). Because central asymmetries greatly reduce the likelihood that an inverted repeat can subsequently adopt a problematic, self-annealed structure (reviewed in 28–30), this operation specifically ensures that the palindrome-bearing sequence is altered to allow proper replication without rearrangement. In mouse cells, a minimal sequence change at the center of a large palindrome is sufficient to achieve stable transmission and this can be acquired in one generation (4,20; L.Cunningham, M.Jasin and S.Lewis, unpublished observations).

Whereas a disparity between *E.coli* and higher eukaryotes in resolution of palindromic dimer DNA is demonstrated in the present study, there is additional evidence that this major difference between genome maintenance strategies arose long ago. As first revealed by reassociation kinetics, bacterial and yeast genomes largely lack sequences that are arranged as

inverted repeats. In contrast, a zero time 'snap-back' fraction constitutes fully several percent of the non-satellite DNA of higher eukaryotes (31–36). The fact that inverted repeats (many of which can be shown to be interrupted by a central asymmetry) are common in higher eukaryotes and conspicuously rare in bacteria has never been assigned any functional significance. According to the present work, it seems plausible that the interrupted inverted repeats that occur throughout higher eukaryotic genomes may be the result of center-specific DNA repair events. If palindrome-containing sequences have consistently been stabilized by creating central alterations in higher eukaryotes but not in yeast or bacteria, the observed accumulation of 'repaired' palindromes over evolutionary time (i.e. interrupted inverted repeats) is expected in the one case and not in the other.

## ACKNOWLEDGEMENTS

The author gratefully acknowledges the technical assistance of Linda Czycyk and Susie Suh, the helpful insights on junctional insertions provided by Dr Avraham Levy and Dr Angel Islas and valuable comments on the manuscript by Dr David Pulleyblank, Dr Marc Shulman and Dr Howard Lipshitz. The CBP9 and SCD-9 cell lines were a generous gift of Dr Gillian Wu. This work was funded by the American Cancer Society and by the Medical Research Council of Canada, grant no. MT13219. The author is a Research Scientist of the National Cancer Institute of Canada supported with funds provided by the Canadian Cancer Society.

## REFERENCES

1. Leach,D.R.F. (1994) *Bioessays*, **16**, 893–900.
2. Leach,D.R.F. (1996) *Genet. Eng.*, **18**, 1–11.
3. Leach,D.R., Okely,E.A. and Pinder,D.J. (1997) *Mol. Microbiol.*, **26**, 597–606.
4. Akgün,E., Zahn,J., Baumes,S., Brown,G., Liang,F., Romanienko,T.J., Lewis,S. and Jasin,M. (1997) *Mol. Cell. Biol.*, **17**, 5559–5570.
5. Connelly,J.C., Kirkham,L.A. and Leach,D.R.F. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 7969–7974.
6. Hesse,J.E., Lieber,M.R., Mizuuchi,K. and Gellert,M. (1989) *Genes Dev.*, **3**, 1053–1061.
7. Lewis,S.M. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 1332–1336.
8. Lieber,M.R., Hesse,J.E., Mizuuchi,K. and Gellert,M. (1987) *Genes Dev.*, **1**, 751–761.
9. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. **3**, pp. 16.42–16.44.
10. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) *J. Mol. Biol.*, **215**, 403–410.
11. Hagan,C.E. and Warren,G.J. (1982) *Gene*, **19**, 147–151.
12. Mizuuchi,K., Mizuuchi,M. and Gellert,M. (1982) *J. Mol. Biol.*, **156**, 229–243.
13. Collins,J. (1981) *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 409–416.
14. Lafaille,J.J., DeCloux,A., Bonneville,M., Takagaki,Y. and Tonegawa,S. (1989) *Cell*, **59**, 859–870.
15. Meier,J.T. and Lewis,S.M. (1993) *Mol. Cell. Biol.*, **13**, 1078–1092.
16. McBlane,J.F., van Gent,D.C., Ramsden,D.A., Romeo,C., Cuomo,C.A., Gellert,M. and Oettinger,M.A. (1995) *Cell*, **83**, 387–395.
17. Ramsden,D.A. and Gellert,M. (1995) *Genes Dev.*, **9**, 2409–2420.
18. Lewis,S.M. (1994) *Semin. Immunol.*, **6**, 131–141.
19. King,J.S., Valcarcel,E.R., Rufer,J.T., Phillips,J.W. and Morgan,W.F. (1993) *Nucleic Acids Res.*, **21**, 1055–1059.
20. Collick,A., Drew,J., Penberth,J., Bois,P., Luckett,J., Scaerou,F., Jeffreys,A. and Reik,W. (1996) *EMBO J.*, **15**, 1163–1171.
21. Haniford,D.B. and Pulleyblank,D.E. (1985) *Nucleic Acids Res.*, **13**, 4343–4363.
22. Gorbunova,V. and Levy,A.A. (1997) *Nucleic Acids Res.*, **25**, 4650–4657.
23. Rubin,E. and Levy,A.A. (1997) *Mol. Cell. Biol.*, **17**, 6294–302.
24. Pinder,D.J., Blake,C.E., Lindsey,J.C. and Leach,D.R. (1998) *Mol. Microbiol.*, **28**, 719–727.
25. Henderson,S.T. and Petes,T.D. (1993) *Genetics*, **133**, 57–62.
26. Rosche,W.A., Trinh,T.Q. and Sinden,R.R. (1995) *J. Bacteriol.*, **177**, 4385–4391.
27. Lewis,S.M., Akgün,E. and Jasin,M. (1999) *Ann. NY Acad. Sci.*, **870**, 45–57.
28. Murchie,A.I.H. and Lilley,D.M.J. (1992) *Methods Enzymol.*, **211**, 158–180.
29. Chalker,A.F., Okely,E.A., Davison,A. and Leach,D.R.F. (1993) *Genetics*, **133**, 143–148.
30. Davison,A. and Leach,D.R.F. (1994) *Genetics*, **137**, 361–368.
31. Bonner,T.I. (1970) *Carnegie Inst. Yearb.*, **68**, 1079–1088.
32. Britten,R.J. and Kohne,D.E. (1968) *Science*, **161**, 529–540.
33. Cech,T.R. and Hearst,J.E. (1975) *Cell*, **5**, 429–446.
34. Davidson,E.H., Hough,B.R., Amenson,C.S. and Britten,R.J. (1973) *J. Mol. Biol.*, **77**, 1–23.
35. Klein,H.L. and Welch,S.K. (1980) *Nucleic Acids Res.*, **8**, 4651–4669.
36. Wilson,D.A. and Thomas,C.A.,Jr (1974) *J. Mol. Biol.*, **84**, 115–144.