MOLECULAR AND CELLULAR BIOLOGY, Jan. 1994, p. 170–180
0270-7306/94/$04.00+0

Vol. 14, No. 1

Characterization of DNA End Joining in a Mammalian Cell Nuclear Extract: Junction Formation Is Accompanied by Nucleotide Loss, Which Is Limited and Uniform but Not Site Specific

ANDREA L. NICOLÁS AND C. S. H. YOUNG*

Department of Microbiology, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Received 23 July 1993/Returned for modification 8 September 1993/Accepted 22 September 1993

Mammalian cells have a marked capacity to repair double-strand breaks in DNA, but the molecular and biochemical mechanisms underlying this process are largely unknown. A previous report has described an activity from mammalian cell nuclei that is capable of multimerizing blunt-ended DNA substrates (R. Fishel, M. K. Derbyshire, S. P. Moore, and C. S. H. Young, Biochimie 73:257–267, 1991). In this report, we show that nuclear extracts from HeLa cells contain activities which preferentially join linear plasmid substrates in either a head-to-head or tail-to-tail configuration, that the joining reaction is covalent, and that the joining is accompanied by loss of sequence at the junction. Sequencing revealed that there was a loss of a uniform number of nucleotides from junctions formed from any one type of substrate. The loss was not determined by any simple site-specific mechanism, but the number of nucleotides lost was affected by the precise terminal sequence. There was no major effect on the efficiency or outcome of the joining reaction with substrates containing blunt ends or 3′ or 5′ protruding ends. Using a pair of plasmid molecules with distinguishable restriction enzyme sites, we also observed that blunt-ended DNA substrates could join with those containing protruding 3′ ends. As with the junctions formed between molecules with identical ends, there was a uniform loss of nucleotides. Taken together, the data are consistent with two models for the joining reaction in which molecules are aligned either throughout most of their length or by using small sequence homologies located toward their ends. Although either model can explain the preferential formation of head-to-head and tail-to-tail products, the latter predicts the precise loss of nucleotides observed. These activities are found in all cell lines examined so far and most likely represent an important repair activity of the mammalian cell.

The genetic integrity of all organisms depends on the ability to repair damage to DNA caused by endogenous and exogenous agents or mechanisms. Thus, in those organisms in which it has been examined, it is not surprising that a considerable amount of genetic information is dedicated to encoding the gene products involved in the repair of DNA damage. These gene products probably operate in multiple, and perhaps functionally redundant, pathways.

Among the most dangerous types of damage is the formation of a double-strand break (DSB), because a whole section of a chromosome could be irrevocably lost to the daughter cell if the covalent bonds in the DNA are not re-formed. Evidence for the lethality of DSBs was established by early work on the response of organisms to ionizing radiation, but the first genetic evidence in eukaryotes was obtained in the yeast *Saccharomyces cerevisiae*, in which *rad* mutations defined functions involved in DSB repair (reviewed by Friedberg [6]). The most convincing demonstration of this is in mutations of the *RAD52* gene. When yeast cells undergo mating-type switching, a DSB is introduced at the *MAT* locus. In mutant cells lacking the *RAD52* gene product, the break cannot be repaired, and the cells die (18, 40). Recent work of Bennett et al. (1) showed that an induced DSB in a dispensable plasmid led to the complete cessation of cell division by causing a blockage at the G_2 phase. These results imply a very strong link between the control of the cell cycle and the integrity of the genome. Although such clear genetic evidence is not available in mammals, there are several inherited human repair deficiencies and artificially created rodent cell lines which are very sensitive to the action of agents which can cause DSBs (reviewed by Jeggo [11]). Some of these deficiencies are probably caused by the lack of an ability to repair such breaks, although definitive proof of this is still lacking.

In general, a DSB in DNA can be repaired in two ways. The chromosome can be reconstructed exactly, using information present on a homologous chromatid or chromosome, as in the DSB repair model for yeast recombination (35). However, an alternative mechanism(s) employs nonhomologous joining processes. Such mechanisms could have several possible outcomes. The correct ends could be rejoined faithfully with no loss of sequence, the ends could be rejoined unfaithfully with loss or insertion of nucleotides at the site of the join, or they could be joined to give translocations or inversions. Anything other than faithful rejoining is potentially mutagenic. However, unlike the situation in unicellular organisms, an imperfect joining mechanism is not necessarily a serious drawback for a somatic cell in a complex eukaryote. Any mutation introduced could well be in an intron or in a transcriptionally repressed area of the genome. Somatic mammalian cells are also diploid, so an-

* Corresponding author. Mailing address: Department of Microbiology, College of Physicians and Surgeons, Columbia University, 701 West 168th Street, New York, NY 10032.

other copy of an undamaged gene would be available in most cases. For these reasons, the repair of a DSB by nonhomologous means in a mammalian cell is completely plausible; indeed, experimental evidence favors nonhomologous mechanisms as the major pathway for the repair of DSBs in mammalian cells (5).

The ability of mammalian cells to join the ends of linear DNA transformed into a cell has been clearly demonstrated (5, 41). Once linear DNA has been joined end to end, it can be maintained in the cell in a large extrachromosomal array (27). When mammalian cells integrate DNA, they do so mainly nonhomologously (30), probably at transient DSBs (22, 23). Similarly, in transfection assays using linearized viral DNA, end joining of the fragments is a prominent outcome in the recovered viable progeny (reviewed by Roth and Wilson [31]). The terminal sequences present in the junctions have undergone limited alterations (or none), even if the two ends are not complementary, and from the pattern of nucleotide loss (or gain), models for the mechanism have been developed (32). Despite the sophistication of these analyses, however, a complete understanding of the joining processes requires a detailed biochemical analysis of the reaction in vitro.

The repair of DSBs by nonhomologous means will undoubtedly employ many of the common enzymes of DNA metabolism. DNA polymerases and nucleases could be involved at the site of the break to restore or remove terminal nucleotides. These activities alone are not sufficient to heal a DSB, because any joining process must involve some form of ligation activity. Although there are three known ligases in mammalian cells (reviewed by Lindahl and Barnes [17]), all of them operate most efficiently on substrates with staggered complementary ends, and all require 3′ hydroxyl and 5′ phosphate groups to complete the ligation reaction. The efficient joining of the ends of DNA transfected into mammalian cells cannot be explained by any of these ligases alone. Junctions can be formed directly between 3′ and blunt ends, an activity not performed by any known ligase at high efficiency. In addition, a naturally occurring DSB would not necessarily have complementary ends or even 3′ hydroxyls and 5′ phosphates, the termini may be blocked by bulky adducts, or the ends could have drifted apart. Any plausible joining mechanism would have to be able to accommodate these possibilities, and ligase alone is not capable of providing these activities. For these reasons, we and others (2a, 3, 4, 24) have begun the characterization and purification of DNA end-joining activities from mammalian cells.

In this report, we describe the characteristics of DNA end joining in an in vitro system, using nuclear extracts and linear plasmid substrates. We have characterized the sequences at the junctions resulting from the in vitro end-joining process and have observed the behavior of different DNA ends in this system. The results show that linear molecules are efficiently joined, predominantly in head-to-head (H:H) and tail-to-tail (T:T) configurations, that nucleotides are lost concomitantly with the joining reaction, and that the pattern of nucleotide loss is determined, in part, by the terminal sequences.

MATERIALS AND METHODS

Preparation of nuclear extracts. Nuclear extracts were prepared from cells grown either in suspension or as monolayers in 15-cm-diameter tissue culture dishes (Falcon), using a protocol adapted from that of Peterson et al. (28). Extracts were routinely prepared from 10 to 12 monolayer plates or from 1 to 3 liters of suspension culture. All procedures were performed at 0 to 4°C. Cell monolayers were scraped from plates into 5 ml of phosphate-buffered saline (PBS) per plate and centrifuged at 1,000 rpm for 5 min in an IEC-CR 6000 with a small rotor (no. 269). The pellets prepared from both types of culture were washed once with 10 to 20 ml of PBS and then resuspended in 5 to 15 ml of 50 mM Tris (pH 7.5)–50 mM KCl–15 mM $MgCl_2$–0.5% Nonidet P-40–5 mM phenylmethylsulfonyl fluoride (PMSF)–25 mM β-glycerophosphate–0.5 mM dithiothreitol (DTT)–5 μg each of aprotinin and leupeptin per ml. The cells were then disrupted with a Wheaton Dounce homogenizer (B pestle) until at least 85% cell rupture was obtained, usually 10 to 20 strokes. At this point, reticulocyte standard buffer (0.1 M Tris [pH 7.5], 10 mM NaCl, 5 mM $MgCl_2$) with 5 mM PMSF, 25 mM β-glycerophosphate, 0.5 mM DTT, and 5 μg each of aprotinin and leupeptin per ml was added at one-half the volume of the original suspending buffer, and the nuclei were centrifuged as described above. The nuclear pellet was resuspended in a volume of extraction buffer A (50 mM Tris [pH 7.5], 0.2 mM EDTA, 0.2 mM EGTA, 20% glycerol, 200 mM NaCl, 0.1 mM DTT, 2 mM $MgCl_2$, 5 mM PMSF, 25 mM β-glycerophosphate, 5 μg each of aprotinin and leupeptin per ml) which was equal to the volume of the nuclei, and then an equal volume of extraction buffer B (extraction buffer A at 600 mM NaCl) was added. Nuclei were extracted for 30 min with gentle rotation at 4°C. The nuclear extract was clarified at 100,000 × *g* for 60 min in a Beckman SW50.1 rotor, and the protein concentration of the supernatant was determined by the Bradford colorimetric assay (Bio-Rad). The clarified extract was then distributed in 100-μl aliquots and stored at −80°C.

Preparation of substrate DNA for the in vitro reaction. All substrates used in the end-joining reactions were made from plasmid DNA, prepared from 500-ml cultures, and processed by column chromatography (Qiagen Inc.). The plasmid was exhaustively digested with the restriction enzyme appropriate for each experiment until no nicked circular DNA was observable on an ethidium bromide-stained gel. The linear substrate was phenol-chloroform extracted, ethanol precipitated, and resuspended to a concentration of 1 mg/ml.

The plasmids were either pUC18 or derivatives thereof. The plasmid with a unique *Ssp*I site in the polylinker was created in two steps. First, the *Ssp*I site, normally present at bp 2501, was removed by the insertion of an 8-bp *Kpn*I linker. This derivative was then modified by inserting an 8-bp *Ssp*I linker (AAATATTT) into the *Hin*cII site of pUC18. The two plasmids used in the 3′-to-blunt end-joining experiment were created by the insertion of a 13-bp *Sfi*I linker, with a nested *Sma*I site, at the *Hin*cII site. One plasmid contains the original pUC18 *Ssp*I site at bp 2501, while the other contains a single novel site located at bp 183, the original site of restriction by *Nde*I.

For convenience, each substrate is named after the restriction enzyme used in its preparation.

In vitro reaction conditions. To clone junctional fragments, 20 μg of linearized substrate was incubated with nuclear extract, containing 1 to 2 mg of protein, for 2 h at 25°C in 50 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.5)–0.5 mM DTT–0.2 mM ATP–10 mM $MgCl_2$ in a total volume of 200 μl (Fig. 1A). After incubation, proteinase K and sodium dodecyl sulfate were added to 100 μg/ml and 0.5%, respectively, and the mixture was incubated at 37°C for 1 h. The end-joined DNA was then extracted with phenol-chloroform, ethanol precipitated,

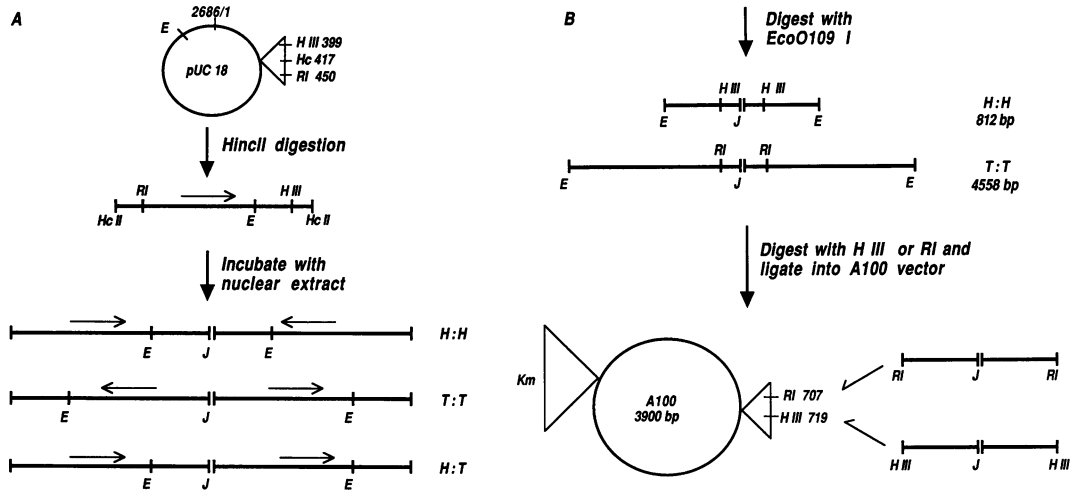

FIG. 1. Structure of H:H and T:T junctional fragments and the strategy used in their cloning. HIII, *Hin*dIII; HcII, *Hin*cII; RI, *Eco*RI; E, *Eco*O109I; J, site of the junction. (A) The creation of *Hin*cII substrate from pUC18 and the products which result after incubation with the nuclear extract are diagrammed. All three possible orientations, H:H, T:T, and H:T, are shown. The head of the linear *Hin*cII substrate is defined as that *Hin*cII half-site which is closest to the pUC18 origin. The orientation of each linear half of the multimer in relationship to the original substrate is shown by an arrow. (B) The H:H and T:T fragments are digested with *Eco*O109I, isolated, and cloned as shown and as described in Materials and Methods. Km, kanamycin resistance gene.

washed with 70% ethanol, and resuspended to 1 mg/ml in Tris-EDTA (TE). All subsequent DNA manipulations were carried out by the same procedures.

Cloning of junction fragments. The cloning strategy is illustrated in Fig. 1B. The total DNA product from the end-joining reaction was digested for 1 h with *Eco*O109I and loaded on a 1% standard agarose gel in Tris-borate-EDTA, with ethidium bromide at 0.5 μg/ml. The DNA was electrophoresed for 1 h at 80 V to separate the H:H and T:T bands. These were then excised from the gel, and the DNA was electroeluted from the gel slices at 4°C into 8 M ammonium acetate on an IBI electroeluter. Electroelution at 4°C was essential for subsequent successful cloning of the junctional fragments. The eluted junctional fragments were precipitated by addition of glycogen and ethanol, washed with 70% ethanol, and then resuspended in 10 μl of TE. The H:H junctional fragments were digested with *Hin*dIII, and the T:T fragments were digested with *Eco*RI. After 1 h of digestion, the DNA was extracted and resuspended in 8 μl of TE.

These small junctional fragments were then ligated overnight, using T4 DNA ligase, with 250 to 400 ng of prepared A100 vector (see below). An equal amount of prepared vector was ligated without insert as a control. The entire ligation mixture was electroporated into *Escherichia coli* XLI-Blue cells on a Bio-Rad Gene Pulser set to 25 μF, 2.25 kV, and 200 Ω. Immediately after pulsing, 0.8 ml of SOC medium (Bethesda Research Laboratories) was added, and the cells were chilled for 15 min on ice. The electroporation mixture was divided into four tubes of 0.2 ml each and incubated at 37°C for 1 h. Each tube was plated on one LB plate containing kanamycin at 50 μg/ml and tetracycline at 30 μg/ml. The former maintains selection for the pA100 vector, and the latter selects for maintenance of the *lacI*-containing plasmid in XLI-Blue, as a bacterial contamination control.

Preparation of the pA100 vector. pA100 was derived from pBluescriptII KS$^+$ (Stratagene) by inserting the kanamycin resistance gene of Tn*903* at the *Sca*I site. The insertion of the kanamycin resistance gene at this site interrupts the ampicillin resistance gene of the original plasmid. DNA preparations of pA100 were exhaustively digested with either *Hin*dIII or *Eco*RI, treated with calf alkaline phosphatase (Boehringer Mannheim) or shrimp alkaline phosphatase (U.S. Biochemical) as recommended by the supplier, and extracted, precipitated, and resuspended at 0.5 μg/μl. Preparations for use in cloning were stored at −20°C.

Sequencing of cloned junctions. Individual colonies were inoculated into 5 ml of LB with kanamycin (250 μg/ml) and tetracycline (62.5 μg/ml). DNA was prepared for sequencing by the Speedprep method (8), one half of the final yield of DNA was denatured, and T3 primer was annealed. Sequencing was performed by using the Sequenase kit (U.S. Biochemical) with dITP termination solutions.

End-labeling reaction. Small junctional fragments bounded by *Eco*RI or *Hin*dIII sites were prepared as described above. After digestion, the junctional fragments were precipitated and resuspended in 3 μl of TE. The recessed termini were end labeled by using the Klenow fragment of DNA polymerase and 20 μCi of [α-^{32}P]dATP with 8 μM unlabeled deoxynucleoside triphosphates at room temperature for 30 min. The reaction was stopped by the addition of 1 μl of 0.5 mM EDTA. After ethanol precipitation, the DNA was resuspended in 4 μl of TE and 1 μl of loading dye and electrophoresed on an 8% denaturing acrylamide sequencing gel. The markers for this experiment were junctional fragments of known sequence that had been subjected to the same end-labeling procedures. They included T4 ligase-produced H:H and T:T fragments and fragments that had been previously cloned from the products of the in vitro reaction and contained 12 and 24 nucleotides lost from the site of the junction.

RESULTS

Nuclear extracts from mammalian cells join DNA molecules end to end. To characterize the biochemical components which allow the mammalian cell to join DNA molecules efficiently, it is necessary to develop simple in vitro systems.

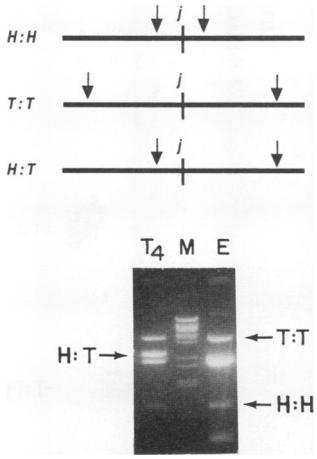

FIG. 2. Relative proportions of the end-joined products. The structure of each possible orientation of joined product is diagrammed. The arrow represents the head of each linear molecule, and J represents the junction site. Lane M contains 1 μg of *Hin*cII substrate incubated with the extract and cut with *Eco*O109I, to distinguish each of the three orientations. Lane T_4 contains 1 μg *Hin*cII substrate incubated with 400 U of T4 ligase. Lane M contains *Bst*EII-digested lambda DNA. Each orientation of join is indicated with the appropriate abbreviation. The samples were electrophoresed on a standard 1% agarose gel containing ethidium bromide.

As a first step, we elected to extract components from isolated nuclei, with a protocol based on a well-characterized method used previously for the isolation of mammalian cell transcription factors (28). It seemed likely that such an extraction procedure would allow recovery of many proteins capable of binding to DNA, including those that promote efficient end joining. A previous publication has presented preliminary characterization of the activities found in this extract (4). Briefly, the end-joining activity was found to join blunt linear DNA into linear multimers. The activity was optimal at 25°C and 200 mM NaCl, it required $MgCl_2$, and the reaction was linear for at least 60 min. The addition of ATP to the reaction stimulated the activity, but a residual level of activity remained if no ATP was added. Even though linear multimers were efficiently produced, there were no detectable circular forms. The present report describes the nature of the products and of the junctions created by the end-joining process.

Linearized blunt-ended DNA was reacted under standard conditions, and the DNA products were analyzed by further restriction enzyme digestion. The enzymes used were either the one originally used to make the unit-length linear substrate or others whose sites are located some distance from the junction. Multimerized products from an incubation with substrate prepared by using *Hin*cII were redigested with *Hin*cII. Unlike the case with joining by T4 ligase, where the large majority of the joins were faithful, the junctions formed in the nuclear extract could not be recut by the enzyme used to make the substrate (data not shown), indicating that some nucleotides were lost or gained prior to or during the joining process. The suggestion that the joining is not faithful will be confirmed by sequence analysis below. More remarkably, the products of a digestion with an enzyme which cuts asymmetrically (*Eco*O109I) were predominantly in H:H and T:T configurations (Fig. 2, lane E), even though the control with T4 ligase showed the expected predominance of head-to-tail (H:T) product (Fig. 2, lane T_4).

In the reaction shown in Fig. 2, approximately half of the input linear substrate was converted to the H:H and T:T forms. The residual linear material, when cut with *Eco*O109I, generated two bands, the larger of which migrated slightly faster than the predicted position for the H:T product, but it is clearly distinguishable in the T4 reaction. In the analysis of the reaction products with a variety of different substrates, the H:T species was either undetectable or present in much lower quantities than expected from random joining.

In summary, the nuclear extract joins blunt linear DNA to produce linear multimers, which are primarily in the H:H and T:T configurations, and the majority of joins are not faithful.

The nucleotide sequences at H:H and T:T junctions formed with a blunt-ended substrate. The results discussed above suggest either that the joining reaction is intrinsically nonconservative or that the substrate is partially digested by nucleases in the extract prior to joining. We note, however, that there is little evidence for extensive nuclease activity in the extract, because the linear input DNA is remarkably stable during the incubation period (as can be deduced from the sharpness of the individual bands in Fig. 2, lane E, and other data not shown). However, to gain a better understanding of the sequences lost during the incubation, it was necessary to clone and sequence many individual junction sites. The particular pattern of nucleotide loss, with a variety of different substrates, may yield important clues to the mechanism underlying end joining. To examine the junctions, we developed a system to isolate and clone the predominant H:H and T:T junctional fragments from a mixture of DNA products arising from incubation in the extract (Fig. 1A). The junction fragments were cloned into the pA100 vector as described in detail in Fig. 1B. They were then sequenced without prior screening of the plasmid DNA to determine the exact nature of nucleotide loss. The pattern of nucleotide loss was examined for a number of substrates with different ends. These substrates were all prepared from pUC18 DNA or its derivatives, and the linearization site was always in the polylinker. The results with a *Hin*cII blunt-ended substrate will be described first.

Remarkably, all of the H:H and T:T junctions formed from the *Hin*cII substrate showed a loss of 12 nucleotides, with one exception that had lost 24. The specific number of nucleotides lost is incompatible with a random exonucleolytic activity prior to end joining. The sequences of 37 of the 38 junctions which had lost 12 nucleotides are shown in Fig. 3 and are organized in two different configurations merely to illustrate the pattern of base loss, not to imply any specific mechanism. At the top of each panel in Fig. 3, the reacting linear molecules are shown paired along their length. Because of the nature of the primary nucleotide sequence at the terminus of the substrates, it was not possible to determine the numbers of nucleotides lost from each partner. However, the pairs of arrows indicate the two extremes between which the deletion must lie, and the dotted line indicates symmetrical loss of nucleotides. The lower part of each panel is arranged to show the duplex sequence that would be created by faithful end joining of the two linear molecules. The bold sequence in the duplex corresponds to the symmetric loss of nucleotides from each partner; below, two of the three single-stranded examples illustrate asymmetric loss. One junction of the 38 analyzed displayed an unambiguous sequence. In this case, the loss was definitely asymmetric, involving 11 nucleotides from one partner and a single nucleotide from the other. Note, however, that even at the extremes, nucleotides must have

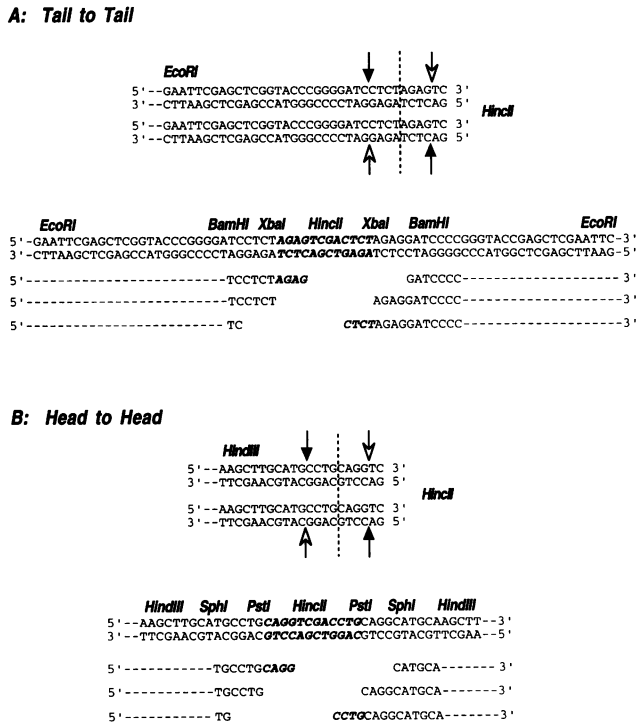

FIG. 3. Pattern of nucleotide loss for the *Hinc*II substrates. Two linear *Hinc*II substrate molecules are shown paired in the T:T (A) or H:H (B) orientation. Nucleotides are lost from the junction site on a sliding scale between the two sets of arrows; the filled-in arrows are at one extreme, while the open arrows represent the other extreme. The dashed line shows the loss resulting from a symmetrical deletion. The duplex sequence below shows the joined product, with the deleted nucleotides shown in boldface. The first line below shows the pattern of nucleotide loss which results when the deletion is made at the open arrows; an appropriate space has been left for each deleted nucleotide. The second line shows the pattern of loss when the deletion is made at the dashed line, while the third line shows the results of a deletion at the filled in arrows. The final sequence from these three schemes and any other on the sliding scale between the two sets of arrows is the same, as can be seen when comparing the sequence across these three lines. Two other types of junctions occurred with very low frequency and are not depicted above. One T:T junction was obtained which had lost 12 nucleotides, but in the pattern of 1 nucleotide from one duplex and 11 from the other; this sequence is unambiguous. The other was a T:T junction which had lost 24 nucleotides in a pattern which doubled the extent of the nucleotide loss on both duplexes.

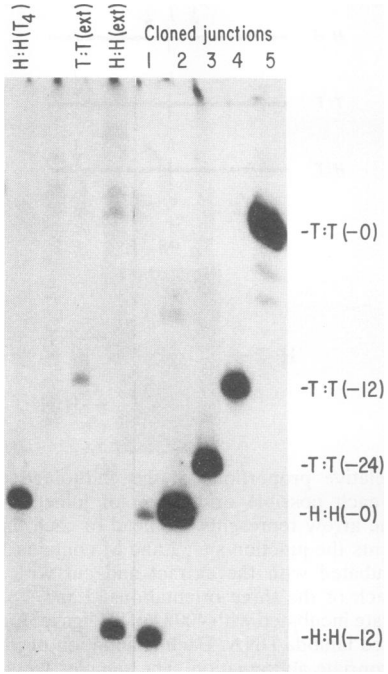

FIG. 4. End labeling of junction fragments produced in the nuclear extract. *Hinc*II substrate was incubated with either T4 DNA ligase [lane H:H(T_4)] or the nuclear extract [lanes H:H(ext) and T:T(ext)] and then digested with *Eco*O109I. The bands for the H:H and the T:T orientation were isolated and digested with *Hin*dIII and *Eco*RI, respectively. The resulting fragments were end labeled and electrophoresed on a standard 8% denaturing acrylamide gel. The size markers for this experiment are fragments from junctions which had been previously cloned and sequenced. These fragments were removed from the vector with *Hin*dIII, in the case of the H:H, or with *Eco*RI, in the case of the T:T, and they were treated exactly as described above. The orientation of the junctional fragments and the number of nucleotides lost are shown at the right. Junction fragments were derived as follows: lane 1, H:H from extract; lane 2, H:H from T4; lane 3, an unusual T:T fragment, with the loss of 24 nucleotides, obtained from extract; lane 4, T:T from extract; lane 5, T:T from T4.

been lost from both partners in the creation of all 38 junctions. These sequencing data also explain why the bulk population of multimerized product could not be redigested with *Hinc*II but could be digested with either *Xba*I or *Bam*HI (data not shown).

Objections could be raised that the specificity of nucleotide loss is an artifact of the cloning step in *E. coli*. Because the H:H and T:T fragments are palindromic, there could be selection against completely faithful junctions at either the ligation or transformation steps of the cloning procedure. Several different experiments were performed to demonstrate that the observed loss of nucleotides is intrinsic to the incubation conditions and not an artifact of the cloning procedure.

The most obvious test of the fidelity of the procedure was to clone H:H and T:T junction fragments formed by incubation of *Hinc*II substrate with T4 DNA ligase. Twelve junctions were sequenced, and all showed faithful reconstruction of the *Hinc*II site. To ensure that the transformation did not select preexisting deleted junction fragments preferentially, products from reactions with T4 DNA ligase were mixed in an equal proportion with those from the nuclear extract after the preparation of the clonable fragments. The mixed material was then ligated into the pA100 vector, and the ligation mixtures were electroporated. Upon sequencing of 20 of the resulting colonies picked at random, it was found that both the perfect T4 junctions and the imprecise deleted junctions were recovered with similar efficiencies.

As a final test, we examined a population of junctional fragments directly. For this experiment, the small junctional fragments were isolated exactly as in the cloning procedure (Fig. 1B), except that after the *Eco*RI or *Hin*dIII digestion step, the material was end labeled, and displayed directly on an 8% denaturing acrylamide gel (Fig. 4). The size markers for this experiment (Fig. 4, lanes 1 to 5) were junctional fragments, derived from plasmids containing previously cloned and sequenced junctions, obtained by incubation of the *Hinc*II substrate with either T4 DNA ligase or nuclear

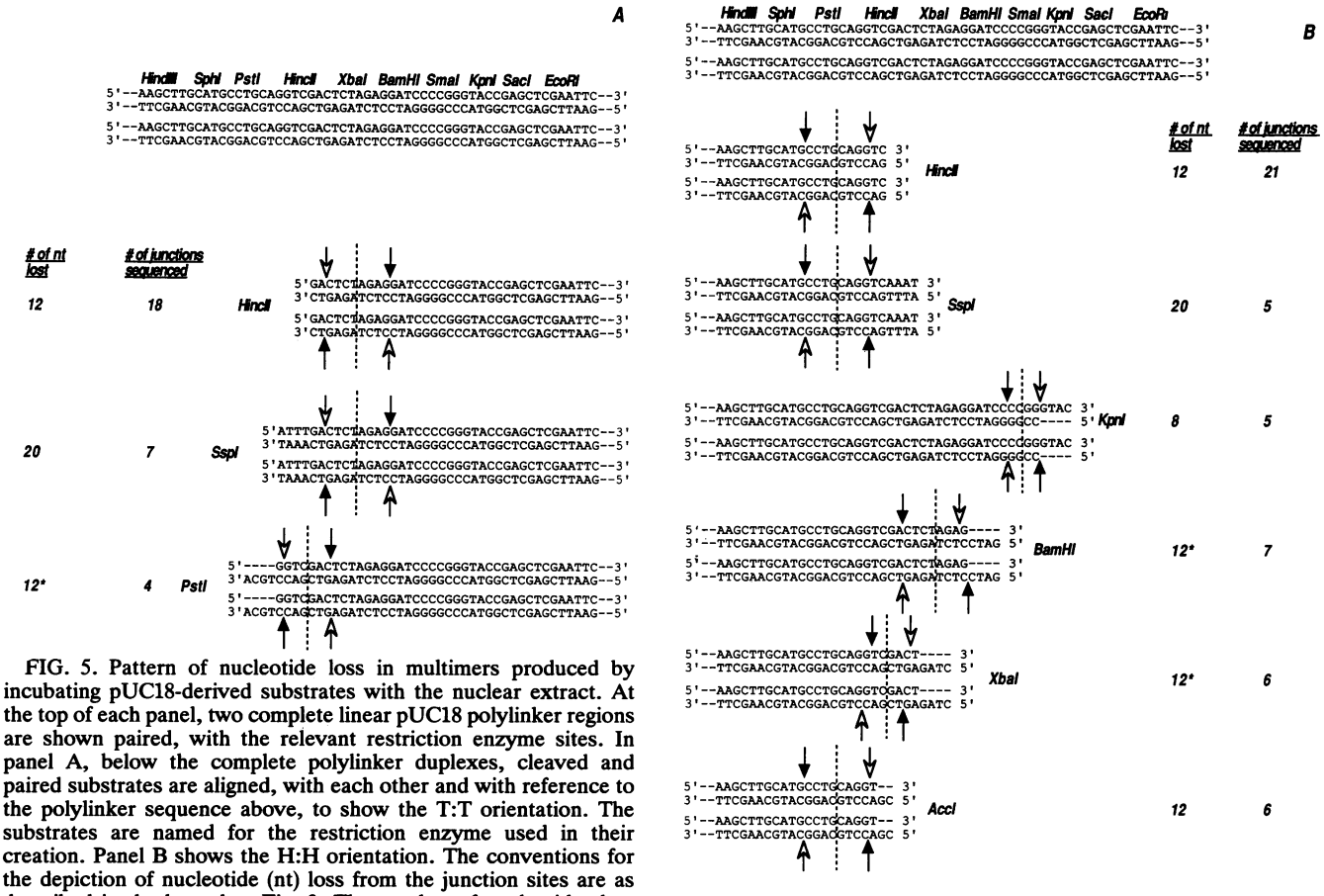

FIG. 5. Pattern of nucleotide loss in multimers produced by incubating pUC18-derived substrates with the nuclear extract. At the top of each panel, two complete linear pUC18 polylinker regions are shown paired, with the relevant restriction enzyme sites. In panel A, below the complete polylinker duplexes, cleaved and paired substrates are aligned, with each other and with reference to the polylinker sequence above, to show the T:T orientation. The substrates are named for the restriction enzyme used in their creation. Panel B shows the H:H orientation. The conventions for the depiction of nucleotide (nt) loss from the junction sites are as described in the legend to Fig. 3. The number of nucleotides lost includes the overhanging extension, where present. The number of observations indicates the number of independent junctions sequenced for each substrate. Faithfully joined junctions were also observed.

extract. The majority of H:H and T:T junctional sequences, produced in the mammalian nuclear extract [H:H(ext) and T:T(ext) in Fig. 4], migrated with those markers missing 12 nucleotides at the site of the join. The population of junctional fragments did not contain any appreciable amount of sequence arising from faithfully joined substrate. Thus, the individually sequenced junctions are truly representative of the majority of junctions formed in the extract in that both show a loss of 12 nucleotides.

Sequence loss at junctions formed with other substrates. The uniformity of nucleotide loss at the junctions could result from several different mechanisms. The joining reaction might be site specific, and the join might occur only at the specific sequences internal to the *Hin*cII site. An argument against this is that the sequences surrounding the H:H and T:T junctions, formed from the *Hin*cII substrate (Fig. 3), are not the same although, fortuitously, each sequence might contain such a joining site at precisely the same linear position. Alternatively, the joining reaction might not be site specific but might always occur a defined number of nucleotides from the terminus. In addition, the specific number of nucleotides lost might be affected by the precise nature of the terminal sequence or by the presence or absence of overhanging single strands. Finally, it is not ruled out that the loss is the result of prior activity of a nuclease which digests a precise number of nucleotides from the end.

To test these possibilities, reactions were performed with different substrates prepared from pUC18 and its derivatives, and the resulting junctions were cloned and sequenced as described above. The results for T:T and H:H junctions are shown in Fig. 5A and B, respectively. As in Fig. 3, the diagrams are organized to illustrate the numbers of nucleotides lost and their relative positions in the substrate and are not meant to imply that the interacting molecules are paired along their length. The first comparison depicted is between junctions formed with either the *Hin*cII or *Ssp*I substrate (Fig. 5). The latter has four extra A · T base pairs added to the *Hin*cII half-site. Junctions formed from the *Ssp*I substrate contained a uniform deletion, but in this case the total loss was 20 nucleotides. This result eliminates a simple end-counting mechanism as the sole criterion for the determination of the location of the end-joining site, but the result is still compatible with a site-specific mechanism, because the junctional sequences are identical in the products of both *Hin*cII and *Ssp*I substrates. A site-specific mechanism, however, is not supported by the observations with the H:H junctions formed from the *Kpn*I substrate (Fig. 5B), which terminates with a 3′ overhanging end. In this case, individual junctional sequences again demonstrated uniform loss but of only eight nucleotides, inclusive of the single-strand extension. Because the site of joining is displaced several nucleotides from the region in which *Hin*cII substrates are joined, this result strongly argues against a site-specific mechanism.

The chemical nature of the terminus is not the cause of the lower number of nucleotides lost with this 3′ overhanging end, because a *Pst*I substrate was joined with loss of 12 nucleotides (Fig. 5A). A variety of substrates with 5′ overhanging ends was also examined (Fig. 5B), and the junctions always showed a loss of 12 nucleotides. In at least two cases (the *Xba*I and *Bam*HI substrates), the junctions were formed in sequences displaced from those used in the *Hin*cII substrate reaction. Again, this result argues against a simple site-specific mechanism. It should be pointed out that several junctions formed from the *Pst*I, *Xba*I, and *Bam*HI substrates were faithfully joined. This is not surprising, as the crude nuclear extract contains mammalian ligase activity, which could join overhanging ends faithfully.

Taken together, these results with substrates containing a variety of terminal structures and sequences suggest that the end-joining mechanism is neither site specific nor restricted to the loss of a precise and constant number of nucleotides from the end. In the discussion, a model which can explain the precise nature of the nucleotide loss will be addressed.

The nuclear extract can join 3′ ends to blunt ends. One of the most surprising results to come from previous transfection experiments was the observation that 3′ ends could join to blunt ends efficiently and with no loss of sequence (33), a reaction that no known ligase can perform. We also examined the nuclear extract to test for this unusual 3′ to blunt end-joining reaction. The substrates for the reaction (Fig. 6A) were two plasmids, derived from pUC18, which contain a 13-bp *Sfi*I linker with a nested *Sma*I site. The two plasmid DNAs were restricted either with *Sfi*I to give a 3′ overhang or with *Sma*I to give a blunt end. Furthermore, the two plasmids could be distinguished by an external *Ssp*I restriction site difference, allowing us to discriminate between products formed by 3′-to-3′, blunt-to-blunt, or 3′-to-blunt joining.

After incubation with the nuclear extract and restriction of the products with *Ssp*I, the H:H junctional fragment formed between the *Sfi*I- and *Sma*I-restricted DNA substrates could be detected as easily as those formed by 3′-to-3′ and blunt-to-blunt pairs of molecules (data not shown). The 3′-to-blunt junctional fragments were cloned as described above, and three types of joins were observed (Fig. 6B). There was a loss of eight nucleotides in two distinct and nonoverlapping patterns; in the third type, there was a loss of nine nucleotides at the site of the join. Once again, because of the repetitive nature of the polylinker primary sequence, the exact location of the breakpoint could not be determined, but the nucleotide loss was restricted to distinct regions in the three junctional types, as indicated in Fig. 6. Despite the ambiguities inherent in the sequence of the substrates, the majority of the junctions must have been formed with at least one nucleotide lost from the blunt end and at least two nucleotides lost from the 3′ end. We cannot determine whether the junction is normally formed by removal of the 3′ overhang in its entirety and several nucleotides from the blunt end. However, in the junctions with a deletion of eight nucleotides, the loss of sequence is clearly asymmetric and of more limited extent than most of those shown in Fig. 5.

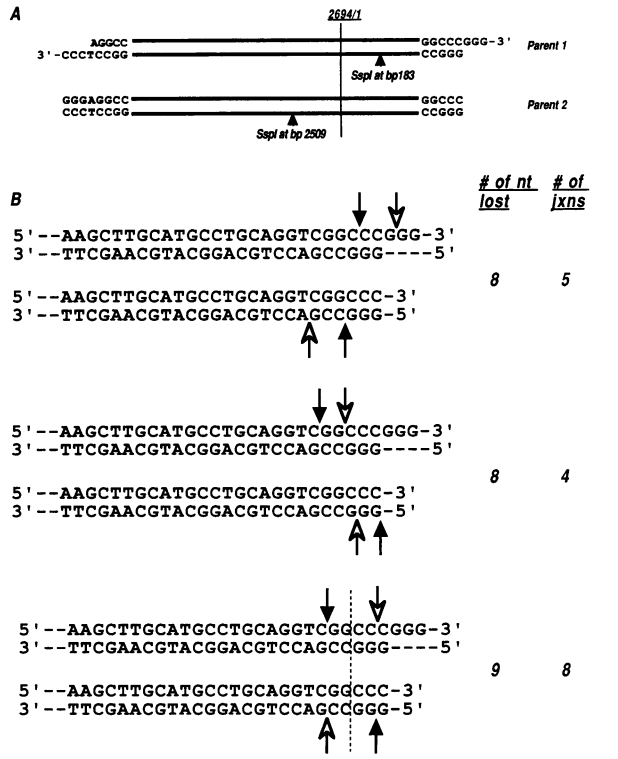

FIG. 6. Joining of 3′ ends to blunt ends in the extract. The substrates used for this experiment are shown in panel A. Parents 1 and 2 are identical except for the location of the *Ssp*I site. Parent 1 was digested with *Sfi*I to give the 3′ end, while parent 2 was digested with *Sma*I to give the blunt end. After incubation with the extract, the heterodimeric H:H band was isolated from an *Ssp*I digestion of the reaction products, cloned, and sequenced. In panel B, the sequence of each individual type of junction and the pattern of nucleotide (nt) loss are shown. The conventions for describing the patterns of nucleotide loss are given in the legends to Fig. 3 and 5. The total number of nucleotides lost and the total number of observations for each junction type are shown next to the sequence information.

DISCUSSION

The repair of DSBs in DNA is an activity which is fundamental to the integrity of the genome. In mammalian cells, the primary method of study has been the analysis of the fate of transfected linear DNA (reviewed by Roth and Wilson [31]). Perhaps the most surprising finding from the in vivo experiments has been that the efficiency of broken-end joining is not significantly affected by the precise chemical nature of the ends (33), implying that the cell has mechanisms for joining any end created by endogenous and exogenous agents. The transfection studies have allowed the elaboration of molecular models to account for the efficient repair of DSBs, but a complete mechanistic description requires the development of in vitro biochemical assays. Accordingly, we and others (3, 4, 24) have chosen to approach this problem in mammalian cells by testing whole cell or nuclear extracts for representative repair activities.

In this report, we describe an end-joining activity in mammalian cell nuclear extracts that joins linear DNA into linear multimers. The efficiency of the reaction is relatively unaffected by the structure of the termini, yielding a ladder of multimers from both blunt-ended and overhanging substrates. When cut with an enzyme outside of the junction site, these multimers are seen to consist mainly of molecules joined in H:H and T:T configurations (Fig. 2), with very little H:T product formed. This is very different from the random ligation activity of T4 DNA ligase, in which the H:T product is expected to be formed at least twice as frequently as either

of the other two forms. Another very striking characteristic of the end-joining activity is the nonconservative nature of joint formation, which was demonstrated both by the inability to recut the reaction products with the enzymes used to create the various substrates (data not shown) and also by extensive sequence analysis (see below). The overall characteristics are therefore different from those of the end-joining activity found in an extract from *Xenopus* eggs (29, 37), in which H:T circular molecules were formed with high efficiency and with little loss of terminal sequence. Similarly, the activities outlined in this report are distinguishable from those described by others using mammalian cell extracts. Thus, the activity obtained by Thacker and colleagues (3, 24) promoted joining of 5′-ended substrates much more efficiently than joining of those with blunt ends, and the majority of junctions could be recut with the enzyme used to create the original substrate. Circular DNA molecules were also formed (3), although their frequency declined with changing extraction conditions and/or cell type (24). Another point of difference concerns the relative lack of faithfulness of joining in ataxia cell extracts compared with those from wild-type cells (3), a distinction not observed by us (see below).

The joining activity is not confined to HeLa cell nuclear extracts. All cell lines tested so far, from a variety of mammalian species and from different tissue origins, contain this activity. Cell lines tested included those derived from SCID mice, recently shown to be deficient in the repair of DSBs (2, 7, 9), or the simian virus 40-transformed derivative of the AT5BIVA strain of human ataxia telangiectasia. Although the direct cause of the deficiency in ataxia telangiectasia is unknown, ataxia cell lines are deficient in various aspects of DSB repair (25, 26, 36). In nuclear extracts from both cell types, linear multimers were formed in the nuclear extract, predominantly in H:H and T:T configurations, and the junctions could not be recut with the enzyme used to make the substrate. Junctional sequences, formed by using a *Hin*cII substrate and an extract from ataxia cells, were identical to those found in products formed by the HeLa extract. These observations support the idea that the activity that we have examined is a widespread and important part of the cellular machinery for the repair of broken DNA.

The most striking novel features of the activity described in this report are the preferential formation of H:H and T:T junctions and the loss of sequence either prior to or during the reaction. The loss of sequence could be an important step in the reaction or could arise from limited nuclease digestion unconnected mechanistically to subsequent joining. Before an attempt could be made to distinguish between these two possibilities, it was necessary to characterize the junctions in detail at the sequence level. Accordingly, a strategy for cloning and sequencing individual junctions was developed (Fig. 1). Using this system, we were able to examine nucleotide loss in sequences formed from the joining of blunt, 3′, or 5′ ends and of 3′ to blunt ends. In all cases, nucleotides were lost from the majority of junctions (Fig. 5 and 6), but although the number of nucleotides lost was uniform for any given substrate, the precise number varied with the substrate. This uniformity and specificity of nucleotide loss is not easily reconciled with random double-strand nuclease activity occurring before, and unrelated to, the joining reaction itself. An alternative possibility would be that an exonucleolytic activity stalls at certain precise sequences. However, some of the sequencing data suggest that the loss of nucleotides is asymmetric between the two partners. This is true both for a single sample from the T:T *Hin*cII junctions (legend to Fig. 3) and for those 3′-to-blunt joins in which eight nucleotides were lost (Fig. 6). In the former case, it would be necessary to postulate different degrees of digestion of the two individual molecules. This seems unlikely, because both partners are identical in sequence. In the latter, the two types of junction would require that the prior digestion from the 3′ extension be either more or less extensive than that of the blunt end, depending on the type of junction formed. Again, this seems unlikely for a double-strand exonucleolytic activity in the extract. We conclude that random double-strand nuclease activity is unlikely to be the cause of the uniform and specific nucleotide loss observed but do not preclude single-strand nuclease activity as a necessary component of the joining reaction itself (see below).

The sequencing data show that the joining reaction is not site specific. Two lines of evidence support this conclusion. First, the blunt-ended *Hin*cII substrate was joined with approximately equal efficiencies in both H:H and T:T orientations (Fig. 2), yet the sequences surrounding the junctions are clearly not identical (Fig. 3). Second, different substrates yielded junctions from different regions of the polylinker sequence (Fig. 5). Thus, it is quite clear that there is no single sequence which directs the joining. Similarly, although the nucleotide loss is most frequently of 12 nucleotides, others show the loss of 8, 9, or 20 nucleotides (Fig. 5 and 6). Thus, the end-joining mechanism does not simply count a given number of nucleotides in from the end and cause a join to be made. These negative conclusions serve to reinforce the idea that any plausible model for end joining must account for the uniformity and specificity of nucleotide loss.

We were particularly interested in whether the nuclear extract could join 3′ to blunt ends. This is the most surprising of all the activities revealed in the in vivo analyses performed previously (33), as no known DNA ligase is capable of joining single-stranded ends to each other, or to blunt ends, with high efficiency. Previous results with the *Xenopus* egg extract demonstrated that such an activity could indeed be found in vitro (29, 37). In the present study, we designed two distinguishable plasmids in which the products formed by 3′-to-blunt joining could be differentiated from those derived from blunt-to-blunt and 3′-to-3′ joins. Such a pair is illustrated in Fig. 6, and evidence for such joining was obtained. However, in contrast to the situation in transfection with simian virus 40 molecules (33), all junctions showed evidence for nucleotide loss, either of eight or of nine nucleotides (Fig. 6). Although the exact site of exchange was ambiguous, it is clear that sequence must have been lost from both partners in the majority of samples. In no case was there complete retention of the terminal sequence of both partners, unlike the situation in vivo. This reaction therefore can be considered yet another example of the joining reactions, described above, for substrates with identical ends. One difference, however, should be noted. All sequences were obtained from the H:H product, because the T:T product was not observed in several independent experiments, either in the ethidium bromide-stained gel or after Southern blotting (data not shown). The significance of this observation is unknown, but it may reveal an aspect of the normal joining mechanism, as discussed below.

The end-joining activity is unusual in that there is preferential formation of H:H and T:T products, and an absolute or relative lack of H:T product, for all of the substrates described in this report (Fig. 2 and data not shown). The lack of H:T product and other data (not shown) which demon-

strate the necessity for terminal homology for efficient multimer formation are consistent with the involvement of an alignment protein, which pairs the linear molecules along their entire homology, prior to the joining process. This model, which is discussed in detail in the accompanying article by Derbyshire et al. (2a), is supported by the observation that the human homologous pairing protein, HPP1, stimulates a similar kind of end joining in more purified extracts from lymphocyte lines (4) (but see below). The necessary involvement of an alignment activity would explain the lack of H:T junctions and would obviously preclude the formation of circular covalently closed monomer molecules, in agreement with our physical and biological observations (data not shown). This interpretation of the in vitro observations is reminiscent of that previously proposed to explain the results of transformation of *S. cerevisiae* with linear DNA incapable of homologous recombination with the chromosome. It was shown that such DNA formed inverted dimer circles almost exclusively (13, 14). A point of difference is that the number of base pairs lost from the yeast junctions was always much greater than that observed in the mammalian examples. However, despite the attractiveness of the full-length alignment model and the apparent biological precedent, both from yeast studies (13, 14) and from the cytological evidence of the fate of broken chromosome ends in *Zea mays* (19), the model does not readily explain the specificity and uniformity of nucleotide loss.

Most models of recombination invoke Watson-Crick base pairing as a means to stabilize the intermediate formed between the recombining partners, and nonhomologous models are no exception, but with the added proviso that the sequence homologies may be quite small and not energetically sufficient by themselves to maintain the structure (see, for example, reference 33). Although these small homologies are not necessary for nonhomologous recombination, there are many examples of their use (21, 33). It is apparent from inspection of the duplex sequences presented in the bottom half of each panel of Fig. 3 that because of the presence of the polylinker in the substrate DNA, there will be homologies of at least 6 bp between interacting molecules in either the H:H or T:T configuration. These small homologies could be used to direct alignment in an overlapping arrangement, as opposed to the complete side-by-side alignment described above. Furthermore, if the homologies nearest to the termini are used preferentially, the alignment would lead, precisely, to the observed loss of nucleotides in each product. This is shown for the *Hin*cII substrate (Fig. 7A) and can be deduced by inspection for all other sequences. The rare exception, namely, the loss of 24 nucleotides from a single T:T junction, would occur with use of the penultimate terminal homology. In contrast, the H:T configuration does not lend itself to much sequence alignment close to the termini (Fig. 7A) and, if the alignment requirement for nonhomologous recombination is strong, could explain the relative lack of H:T junction formation observed with all the substrates described in this report. It might also explain the lack of T:T junctions in the 3′-to-blunt end-joining reaction, in which this configuration shows much less overlapping homology than the H:H configuration. Assuming that alignment of small terminal sequences is an important part of the mechanism, it is then possible to outline a detailed set of steps in the end-joining reaction, based on those proposed for the single-strand annealing model of homologous recombination (16) (Fig. 7B). A very similar model has been proposed recently to account for the formation of small deletional junctions in the recircularization of linear plasmids in in vitro extracts (36a).

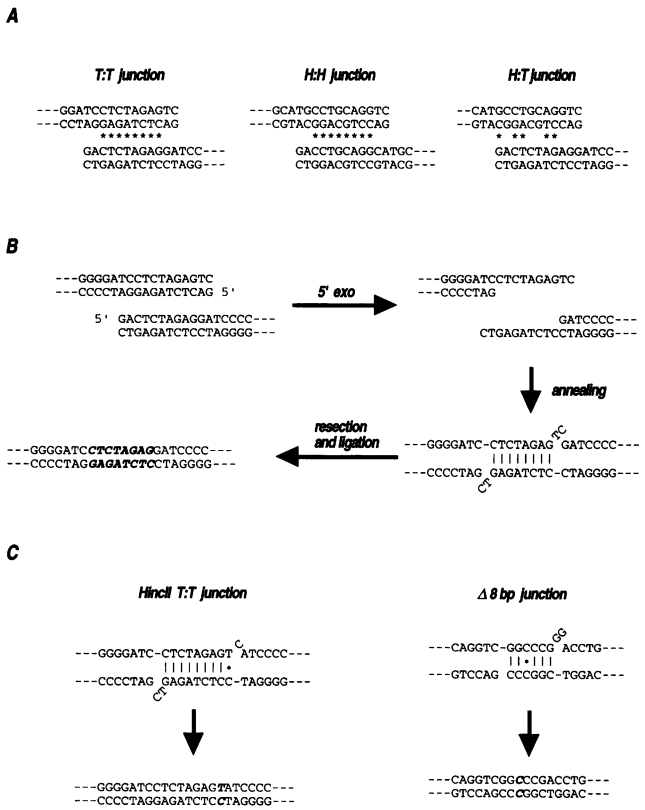

FIG. 7. Use of sequence alignment in the creation of junctions. Panel A shows duplexes from the *Hin*cII substrate aligned to give maximum homology (*) in the T:T and H:H configurations. The dashed lines indicate the rest of the plasmid sequence. An example of a potential alignment in the H:T configuration is shown; it is the one with the greatest relative sequence homology (5-of-11 match), with the exception of a three-nucleotide sequence at the very terminus, which contain a 2-of-3 match. In panel B, a modified version of the single-strand-annealing model is depicted. For convenience, the 5′-3′ exonuclease is shown to have digested 10 nucleotides from each partner. This exposes complementary sequences, which can anneal (||), leaving short unpaired single-strand extensions. These are then resected, and the nick is sealed by ligase. The region of hybrid DNA, in which the top strand comes from the leftward partner and the bottom strand comes from the rightward partner, is shown in boldface. Panel C illustrates the formation of either a flanking or an internal mismatched base pair (•), based on the model described in panel B.

For simplicity of representation, the first step involves degradation by an exonuclease from the 5′ end to uncover regions of homology between the two partners, but a helicase activity would be equally plausible (39). Following annealing of the exposed complementary sequences, resection of unpaired single strands, and ligation, the final recombinant product would contain a region of hybrid DNA. The hybrid structure helps to explain several of the sequencing observations, as illustrated in Fig. 7C. The single asymmetric T:T junction from the *Hin*cII substrate can be explained if the hybrid DNA included, in this case, a flanking mismatched base pair. Incorporation of the upper-strand T from the leftward partner would generate the observed unique sequence. Aligning the terminal sequences of the substrates used in the 3′-to-blunt joining reaction, to give the observed loss of eight nucleotides, creates an internal mismatch and

should lead to the segregation of two distinguishable sequences at similar frequencies. This is indeed the case (Fig. 6). Thus, this alignment model, based on small terminal homologies, not only can explain the uniformity and specificity of nucleotide loss but also leads to specific predictions of mismatch formation that are borne out by the sequencing data themselves. It is also consistent with the predominance of H:H and T:T over H:T junctions observed with the specific substrates used here. Further confirmation of this last point could be obtained by designing substrates in which the H:T configuration possessed terminal homology, while the H:H and T:T configurations did not. The prediction would be that the former would predominate in the products of the reaction. This result would also eliminate the alternative full-length alignment model described above.

Although detailed speculation as to the proteins involved in the mechanism of joining is premature, it seems likely either that some form of single-strand nucleolytic activity is an intrinsic part of the joining process or that helicases unwind the duplex termini prior to single-strand annealing. Exonucleolytic activity seems to be an intrinsic property of several strand-exchange proteins, such as SEP1 (12) and HPP1 (20), in which resection occurs in 5′-to-3′ and 3′-to-5′ directions, respectively. This resection may be necessary for strand exchange to be initiated in homologous recombination, but results of others (3a) also suggest that the 3′-to-5′ activity of HPP1 may be required for end joining. Recent evidence from reactions involving extracts of both *Xenopus* oocytes and yeast nuclei also shows that exonuclease digestion is an important component in the homologous recombination of linear DNA molecules with overlapping homology (10, 15). Alternatively, it is not implausible that the opening up of the termini to allow single-strand annealing between the partners might be accomplished by helicases. Some repair-deficient syndromes are associated with loss of helicase activity (34, 38), and it will be informative to test nuclear extracts from them for end-joining activity.

Our current efforts are directed toward the purification of the components of this potent joining activity and to a further examination of cells that are deficient in DNA repair. To date, the activity from HeLa cells can be recovered from several ion-exchange chromatography steps and maintains the specific characteristics of the unfractionated extract, namely, the predominance of H:H and T:T product and the nonconservation of nucleotides at the junctions. Purification of individual components and reconstitution of the complete joining reaction will allow an unambiguous answer to the question of whether the loss of nucleotides is an intrinsic part of this novel activity.

ACKNOWLEDGMENTS

This work was supported by NIGMS grant RO1 GM31452 and by a minority supplement awarded to A.L.N. Core facilities were provided in part by grant CA13696 from the NCI, to the Columbia Comprehensive Cancer Center.

We thank Patricia Munz for constructing the *Ssp*I and *Sfi*I linker derivatives of pUC18, and for help with cell culture, to which Tony Wild also contributed. Larry Wiater donated the parent plasmid from which pA100 was derived. We are grateful to Lorraine Symington for valuable advice and initial criticism of the manuscript and to John Wilso, who pointed out a simpler and more compelling explanation than the one we initially favored for the loss of nucleotides at the junctions. This work has benefited from the continuing informed criticism of Richard Fishel.

REFERENCES

1. **Bennett, C. B., A. L. Lewis, K. K. Baldwin, and M. A. Resnick.** 1993. Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. Proc. Natl. Acad. Sci. USA **90**:5613–5617.
2. **Biedermann, K. A., J. Sun, A. J. Giaccia, L. M. Tosto, and J. M. Brown.** 1991. *scid* mutation in mice confers hypersensitivity to ionizing radiation and a deficiency in DNA double-strand break repair. Proc. Natl. Acad. Sci. USA **88**:1394–1397.

2a. **Derbyshire, M. K., L. H. Epstein, C. S. H. Young, P. L. Munz, and R. Fishel.** 1994. Nonhomologous recombination in human cells. Mol. Cell. Biol. **14**:156–169.

3. **Fairman, M. P., A. P. Johnson, and J. Thacker.** 1992. Multiple components are involved in the efficient joining of double stranded DNA breaks in human cell extracts. Nucleic Acids Res. **20**:4145–4152.
4. **Fishel, R., M. K. Derbyshire, S. P. Moore, and C. S. H. Young.** 1991. Biochemical studies of homologous and nonhomologous recombination in human cells. Biochimie **73**:257–267.
5. **Folger, K. R., E. A. Wong, G. Wahl, and M. R. Capecchi.** 1982. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. Mol. Cell. Biol. **2**:1372–1387.
6. **Friedberg, E. C.** 1988. Deoxyribonucleic acid repair in the yeast *Saccharomyces cerevisiae*. Microbiol. Rev. **52**:70–102.
7. **Fulop, G. M., and R. A. Phillips.** 1990. The *scid* mutation in mice causes a general defect in DNA repair. Nature (London) **347**:479–482.
8. **Goode, B. L., and S. C. Feinstein.** 1992. "Speedprep" purification of template for double-stranded DNA sequencing. BioTechniques **12**:148–149.
9. **Hendrickson, E. A., X.-Q. Qin, E. A. Bump, D. G. Schatz, M. Oettinger, and D. T. Weaver.** 1991. A link between double-strand break-related repair and V(D)J recombination: the *scid* mutation. Proc. Natl. Acad. Sci. USA **88**:4061–4065.
10. **Huang, K. N., and L. S. Symington.** 1993. A 5′-3′ exonuclease from *Saccharomyces cerevisiae* is required for in vitro recombination between linear DNA molecules with overlapping homology. Mol. Cell. Biol. **13**:3125–3134.
11. **Jeggo, P. A.** 1990. Studies on mammalian mutants defective in rejoining double-strand breaks in DNA. Mutat. Res. **239**:1–16.
12. **Johnson, A. W., and R. D. Kolodner.** 1991. Strand exchange protein 1 from *Saccharomyces cerevisiae*. A novel multifunctional protein that contains DNA strand exchange and exonuclease activities. J. Biol. Chem. **266**:14046–14054.
13. **Kunes, S., D. Botstein, and M. S. Fox.** 1985. Transformation of yeast with linearized plasmid DNA. Formation of inverted dimers and recombinant plasmid products. J. Mol. Biol. **184**: 375–387.
14. **Kunes, S., D. Botstein, and M. S. Fox.** 1990. Synapsis-mediated fusion of free DNA ends forms inverted dimer plasmids in yeast. Genetics **124**:67–80.
15. **Lehman, C. W., and D. Carroll.** 1991. Homologous recombination catalyzed by a nuclear extract from *Xenopus* oocytes. Proc. Natl. Acad. Sci. USA **88**:10840–10844.
16. **Lin, F.-L., K. Sperle, and N. Sternberg.** 1984. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. Mol. Cell. Biol. **4**:1020–1034.
17. **Lindahl, T., and D. E. Barnes.** 1992. Mammalian DNA ligases. Annu. Rev. Biochem. **61**:251–281.
18. **Malone, R. E., and R. E. Eposito.** 1980. The *RAD 52* gene is required for homothallic interconversion of mating types and spontaneous mitotic recombination in yeast. Proc. Natl. Acad. Sci. USA **77**:503–507.
19. **McClintock, B.** 1939. The behavior in successive nuclear divisions of a chromosome broken at meiosis. Proc. Natl. Acad. Sci. USA **25**:405–416.
20. **Moore, S. P., and R. Fishel.** 1990. Purification and characterization of a protein from human cells which promotes homologous pairing of DNA. J. Biol. Chem. **265**:11108–11117.
21. **Munz, P. L., and C. S. H. Young.** 1987. The creation of adenovirus genomes with viable, stable, internal redundancies centered about the E2b region. Virology **158**:52–60.
22. **Murnane, J. P., M. J. Yezzi, and B. R. Young.** 1990. Recombi-

nation events during integration of transfected DNA into normal human cells. Nucleic Acids Res. **18**:2733–2738.
23. **Nalbantoglu, J., C. Miles, and M. Meuth.** 1988. Insertion of unique and repetitive DNA fragments into the *aprt* locus of hamster cells. J. Mol. Biol. **200**:449–459.
24. **North, P., A. Ganesh, and J. Thacker.** 1990. The rejoining of double-strand breaks in DNA by human cell extracts. Nucleic Acids Res. **18**:6205–6210.
25. **Pandita, T. K., and W. N. Hittelman.** 1992. Initial chromosome damage but not DNA damage is greater in ataxia telangiectasia cells. Radiat. Res. **130**:94–103.
26. **Pandita, T. K., and W. N. Hittelman.** 1992. The contribution of DNA and chromosome repair deficiencies to the radiosensitivity of ataxia-telangiectasia. Radiat. Res. **131**:214–223.
27. **Perucho, M., D. Hanahan, and M. Wigler.** 1980. Genetic and physical linkage of exogenous sequences in transformed cells. Cell **22**:309–317.
28. **Peterson, C. L., K. Orth, and K. L. Calame.** 1986. Binding in vitro of multiple cellular proteins to immunoglobulin heavy-chain enhancer DNA. Mol. Cell. Biol. **6**:4168–4178.
29. **Pfeiffer, P., and W. Vielmetter.** 1988. Joining of nonhomologous DNA double strand breaks *in vitro*. Nucleic Acids Res. **16**:907–924.
30. **Robins, D. M., S. Ripley, A. S. Henderson, and R. Axel.** 1981. Transforming DNA integrates into the host chromosome. Cell **23**:29–39.
31. **Roth, D., and J. Wilson.** 1988. Illegitimate recombination in mammalian cells, p. 621–653. *In* R. Kucherlapati and G. R. Smith (ed.), Genetic recombination. American Society for Microbiology, Washington, D.C.
32. **Roth, D. B., T. N. Porter, and J. H. Wilson.** 1985. Mechanisms of nonhomologous recombination in mammalian cells. Mol. Cell. Biol. **5**:2599–2607.
33. **Roth, D. B., and J. H. Wilson.** 1986. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. Mol. Cell. Biol. **6**:4295–4304.
34. **Schaeffer, L., R. Roy, S. Humbert, V. Moncollin, W. Vermeulen, J. H. J. Hoeijmakers, P. Chambon, and J.-M. Egly.** 1993. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science **260**:58–63.
35. **Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl.** 1983. The double-strand-break repair model for recombination. Cell **33**:25–35.
36. **Thacker, J.** 1989. The use of integrating DNA vectors to analyse the molecular defects in ionising radiation-sensitive mutants of mammalian cells including ataxia telangiectasia. Mutat. Res. **220**:187–204.
36a. **Thacker, J., J. Chalk, A. Ganesh, and P. North.** 1992. A mechanism for deletion formation in DNA by human cell extracts: the involvement of short sequence repeats. Nucleic Acids Res. **20**:6183–6188.
37. **Thode, S., A. Schäfer, P. Pfeiffer, and W. Vielmetter.** 1990. A novel pathway of DNA end-to-end joining. Cell **60**:921–928.
38. **Troelstra, C., A. Van Gool, J. De Wit, W. Vermeulen, D. Bootsma, and J. H. J. Hoeijmakers.** 1992. *ERCC6*, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. Cell **71**:939–953.
39. **Wake, C. T., F. Vernaleone, and J. H. Wilson.** 1985. Topological requirements for homologous recombination among DNA molecules transfected into mammalian cells. Mol. Cell. Biol. **5**:2080–2089.
40. **Weiffenbach, B., and J. E. Haber.** 1981. Homothallic mating type switching generates lethal chromosome breaks in *rad52* strains of *Saccharomyces cerevisiae*. Mol. Cell. Biol. **1**:522–534.
41. **Wilson, J. H., P. B. Berget, and J. M. Pipas.** 1982. Somatic cells efficiently join unrelated DNA segments end-to-end. Mol. Cell. Biol. **2**:1258–1269.