Nucleic Acids Research, Vol. 19, No. 25 7201–7205

Oligonucleotide capture during end joining in mammalian cells

David B.Roth[+], G.Neal Proctor[§], Lisa K.Stewart[†] and John H.Wilson*
Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, TX 77030, USA

ABSTRACT

Extra nucleotides (termed filler DNA) are found at about 10% of the junctions of the genetic rearrangements that arise by illegitimate recombination in mammalian cells. Such filler DNAs could arise by the joining of oligonucleotide fragments to broken ends prior to end joining. We tested this possibility by microinjecting mixtures of defined oligonucleotides with SV40 genomes that were linearized in the intron for T antigen, a site where incorporation of extra nucleotides does not impair viability. Using an injection ratio of 1000 oligonucleotides per DNA end, we screened viable genomes for incorporation of single-stranded and double-stranded oligonucleotides with varying degrees of complementarity to the ends of the linear SV40 molecules. Genomes from 510 independent plaques were screened by restriction digestion to identify those that had picked up a restriction site unique to the injected oligonucleotides. Double-stranded oligonucleotides that were fully complementary to the SV40 ends were readily incorporated, but uptake of other oligonucleotides was not detected by restriction analysis. Nucleotide sequences of junctions from 12 genomes derived from co-injection of noncomplementary oligonucleotides revealed two with filler DNA, but neither could be assigned unambiguously to the injected oligonucleotides.

INTRODUCTION

A wide variety of illegitimate recombination events occur in mammalian cells (1). These include chromosome translocations, rearrangements in transfected DNA, and the assembly of immunoglobulin and T-cell receptor genes by V(D)J recombination. One feature common to all such rearrangements is the frequent presence of extra nucleotides, termed filler DNA, at the recombination junctions (2). Filler DNA is found at approximately 10% of junctions generated by circularization of linear DNA molecules transfected into fibroblasts, and at about half of reciprocal chromosomal translocations (2). Extra nucleotides can have profound biological consequences; for example, addition of nucleotides to junctions formed by V(D)J recombination serves to amplify antigen receptor diversity.

At least two specialized mechanisms may operate in lymphoid cells to generate extra nucleotides during V(D)J recombination: the incorporation of non-templated nucleotides by terminal deoxynucleotidyl transferase (TdT) (3,4), and a mysterious process that adds one to two nucleotides that are palindromic with respect to the termini they adjoin (so-called P nucleotides) (5). Extra nucleotides are also present at V(D)J junctions produced in fibroblasts rendered capable of V(D)J recombination by cotransfection with recombination activating genes, RAG1 and RAG2 (J.E. Hesse, personal communication): a procedure that does not result in the activation of TdT expression (6). These observations suggest that processes responsible for producing filler DNA in fibroblasts may participate in the generation of junctional diversity during antigen receptor gene rearrangement (1,2).

As a first approach to characterizing such activities, we surveyed a large number of junctions with filler DNAs that were produced by illegitimate recombination in fibroblasts (2). Many contained more than 10 nucleotides of filler DNA, and several contained 20 to 40 nucleotides. Half of the filler DNAs longer than 10 bp contained nucleotide sequences that appeared to have been derived from other DNA molecules in the cell. Since mammalian cells are capable of joining virtually any two DNA ends by highly efficient mechanisms (7), we suggested that double- or single-stranded oligonucleotides might be added to broken DNA ends prior to their repair by end joining (2). We set out to test this possibility by asking whether defined oligonucleotides can be captured during the joining of DNA ends.

As in previous studies of end joining, we used SV40 genomes that were linearized in the intron in the gene for T antigen because the intron is largely nonessential and no particular sequence must be recreated by the end-joining event (2,7–10). This feature allows the majority of end-joining products to be recovered as viable viruses. To test whether oligonucleotides could be captured during end joining, we chose to microinject defined mixtures of SV40 genomes and oligonucleotides. Microinjection ensures that oligonucleotides and SV40 genomes arrive in the nucleus at the same time in a defined ratio regardless of the type of oligonucleotide tested. Linear SV40 genomes with

* To whom correspondence should be addressed

Present addresses: [+]Laboratory of Molecular Biology, NIDDK, Bethesda, MD 20892, [§]Transkaryotic Therapies, Inc., 195 Albany Street, Cambridge, MA 02139 and [†]Department of Neuro-Oncology, MD Anderson Cancer Center, Houston, TX 77030, USA

complementary or blunt restriction ends were microinjected in various combinations with single-stranded or double-stranded oligonucleotides that were 16 nucleotides long. The oligonucleotides were selected to allow us to assess whether oligonucleotides, if captured, were added to ends according to the same set of rules that govern the end joining of longer molecules (7).

MATERIALS AND METHODS

Cell Lines, Viruses and DNAs

The established monkey kidney cell line, CV1, was grown according to standard procedures (9). The substitution mutant, su1910, whose construction was described previously (7), is a derivative of SV40 that contains a polylinker inserted into the intron of the large-T antigen gene. SV40 genomes were linearized at the BglII site or the SmaI site in the polylinker before microinjection. SV40 DNA preparations, plaque assays, and viral infections were carried out according to established procedures as described previously (9,11).

Oligonucleotides

Three 16-nucleotide long oligonucleotides were chemically synthesized. Oligonucleotide 1 (5'GATCCGAAGGGGTTCG) and oligonucleotide 2 (3'GCTTCCCCAAGCCTAG) were obtained from New England Biolabs in both phosphorylated and nonphosphorylated forms; they pair in a central 12-nucleotide stretch leaving single-strand protrusions that form BamHI ends. Oligonucleotide 3 (3'CTAGGCTTCCCCAAGC) was synthesized at Baylor College of Medicine in a nonphosphorylated form; it pairs with oligonucleotide 1 to give a double-stranded oligonucleotide with blunt ends.

Microinjection

Linear SV40 genomes with or without added oligonucleotides were microinjected under continuous flow conditions into nuclei of CV1 cells in slightly subconfluent monolayers on 100-mm plates. Injection were done under TS buffer after washing the cell monolayers twice with TS buffer. After injection the monolayers were washed twice with TS buffer and then complete medium was added. Injected cells were allowed to recover overnight in the incubator and then the monolayers were overlaid with agar and treated as for a standard plaque assay. All microinjections were carried out in buffer (58 mM NaCl, 38 mM KCl, 5.8 mM Tris pH 7.5, and 0.58 mM EDTA) at an SV40 DNA concentration of 1 μg/ml, which we estimate introduces about 5 to 10 genomes per cell. Unless otherwise noted oligonucleotides were present at a 2000-fold molar excess. Microinjection sites were arranged in a pattern and at each injection site four adjacent cells were microinjected.

Virus were isolated from each injection site that yielded a plaque. Restriction analysis (BglII or SmaI digestion, as described in Results) of viral DNA from these primary plaques showed that about half contained two or more distinct genomes. Individual genomes were isolated from mixed primary plaques by plaque assay at high dilution. Only one viral isolate was characterized from each injection site, ensuring that all analyzed genomes were generated by independent events.

Miniwell Analysis and DNA Sequencing

Analysis of individual plaques was determined a miniwell labelling procedure followed by restriction digestion as described previously (12). To prepare viral DNA for sequencing, confluent monolayers of CV1 cells in 100 mm plates were infected with picked plaque suspensions. Viral DNA was harvested from Hirt supernatants (13) and purified on CsCl density gradients as described previously (12). Viral DNAs were subjected to double-strand DNA sequence analysis (14), using minor modifications that have been described previously (7).

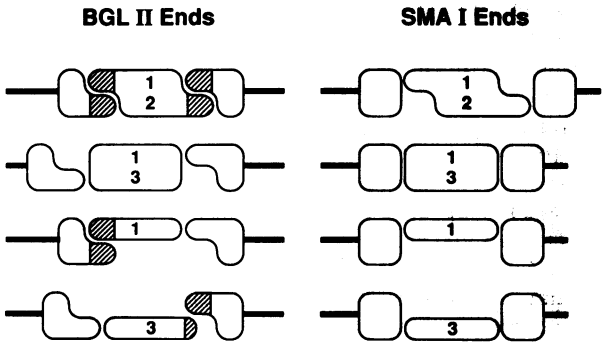

Figure 1. The variety of oligonucleotides and SV40 DNA ends tested by microinjection. Schematic representations of the restriction ends on the SV40 genomes are shown flanking different oligonucleotides. Numbers in the oligonucleotides refer to specific single-stranded oligonucleotides whose sequences are shown in Materials and Methods. Shaded areas represent complementary bases.

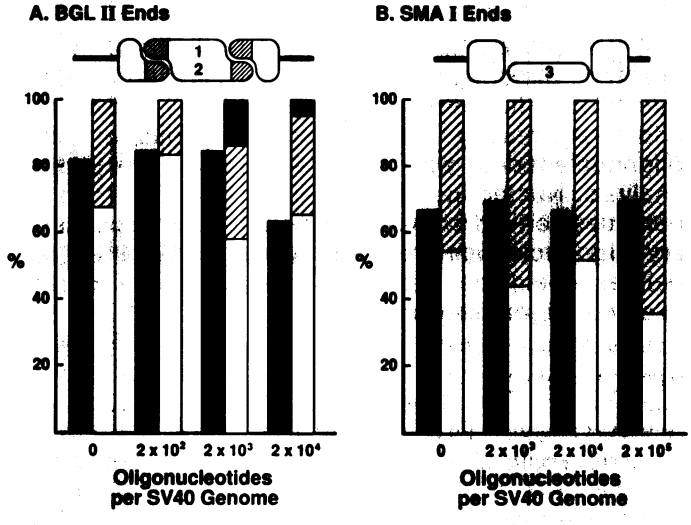

Figure 2. Efficiency of plaque formation and distribution of genome types as a function of different ratios of oligonucleotides to SV40 genomes. (A) Results with double-stranded oligonucleotides complementary to the BglII ends on the SV40 genomes. (B) Results with single-stranded oligonucleotides and SmaI-cut SV40 genomes. Oligonucleotides were injected at a variety of ratios, as indicated along the bottom axis. The percentage of injection sites (each with four injected cells, Materials and Methods) that yielded plaques is shown by the gray bars. The numbers of injection sites in (A) were 33 at 0:1, 72 at 200:1, 66 at 2000:1, and 73 at 20,000:1; in (B) they were 12 at 0:1, 30 at 2000:1, 18 at 20,000:1, and 10 at 200,000:1. The distribution of genome types is shown as a percentage of those tested. Genomes that were BglII$^+$ or SmaI$^+$ are shown as white bars; genomes that were BglII$^-$XmnI$^-$ or SmaI$^-$XmnI$^-$ are shown as hatched bars; genomes that were BglII$^-$XmnI$^+$ are shown as solid bars. Numbers of genomes tested in (A) were 31 at 0:1, 55 at 200:1, 101 at 2000:1, and 44 at 20,000:1; in (B) they were 11 at 0:1, 25 at 2000:1, 42 at 20,000:1, and 16 at 200,000:1.

RESULTS

Experimental Design

To test whether oligonucleotides could be captured during end joining, we chose to microinject mixtures of linear SV40 genomes and defined oligonucleotides. For these studies we used a viable mutant of SV40 that contains a polylinker in the intron of the T-antigen gene. These genomes were linearized in the polylinker at the unique BglII site to give ends with single-strand protrusions or at the unique SmaI site to give blunt ends. Linearization provides a selection for cell-mediated events, since the genome must be circularized before lytic infection is initiated. In previous studies we have shown that many sequence alterations, including insertions, are tolerated in this intron, allowing viable virus to be isolated as plaques (2,8,9,15).

Microinjection was chosen in preference to DEAE dextran-mediated transfection, which we have used previously, because it guarantees that the ratio of oligonucleotides to SV40 molecules we set in the mixture is the same ratio that is present initially in the nucleus. In addition, microinjection introduces oligonucleotides and SV40 genomes into the nucleus at the same time, which would be difficult to ensure by transfection.

Three 16-nucleotide long oligonucleotides were used to create a variety of single-stranded and double-stranded oligonucleotides for microinjection (Figure 1). The 16-nucleotide length was chosen arbitrarily, but only a few filler DNAs in rearrangements from nonlymphoid cells are longer (2). To facilitate identification of genomes that picked up oligonucleotides, the particular sequences (see Methods) were chosen so that capture of an oligonucleotide destroys the restriction site (BglII or SmaI) at which the SV40 molecules were linearized. Moreover, each oligonucleotide carried a unique internal XmnI site to allow positive identification of captured sequences. The XmnI site was chosen because it has four unspecified bases in its recognition sequence (5'GAANNNNTTC). Thus by sequence analysis we can unambiguously identify an XmnI$^+$ genome as deriving from an injected oligonucleotide.

Ratio of Oligonucleotides to SV40 Genomes

To establish an appropriate ratio of oligonucleotides to SV40 genomes to use in microinjection experiments, we tested two mixtures at a variety of molar ratios (Figure 2). A fully complementary double-stranded oligonucleotide (nonphosphorylated) was mixed with BglII linearized SV40 molecules at molar ratios of 200:1, 2000:1, and 20,000:1 (Figure 2A), and a single-stranded oligonucleotide (nonphosphorylated) was mixed with SmaI linearized SV40 molecules at molar ratios of 2000:1, 20,000:1, and 200,000:1 (Figure 2B).

To assess whether oligonucleotides interfered with the ability of linear SV40 genomes to form plaques—a measure of the efficiency of end joining, we determined the percentage of injection sites that yielded plaques in the presence and absence of oligonucleotides (Figure 2). With the possible exception of the highest ratio of complementary double-stranded oligonucleotides, there was no significant effect of oligonucleotides on the ability of SV40 genomes to form plaques.

To measure the effect of added oligonucleotides on the frequency with which the SV40 restriction site was regenerated during circularization, we prepared SV40 DNAs from individual plaques and digested them with BglII or SmaI. Once again there was no dramatic difference in the fraction that had regenerated the restriction site in the presence and absence of oligonucleotides (Figure 2). Plaques that did not regenerate the restriction site were tested for the presence of an XmnI site. XmnI$^+$ genomes were found at the two higher ratios of complementary double-stranded oligonucleotides (Figure 2A).

Injection of a Variety of Oligonucleotides

Using a molar ratio of 2000:1, we tested the mixtures of oligonucleotides (nonphosphorylated) and SV40 genomes indicated in Figure 1. As shown in Table 1, all mixtures of oligonucleotides yielded BglII$^-$ or SmaI$^-$ virus at about the same frequency as the corresponding injections in Figure 2. None of the new combinations, however, gave XmnI$^+$ virus.

The absence of 5' phosphates at the ends of SV40 genomes does not interfere with the efficiency with which they are joined (16), nor does it change the frequency with which a restriction site is regenerated during circularization (Roth and Wilson, unpublished). Nevertheless, we repeated a few of the injections using oligonucleotides with 5' phosphates. Once again, double-stranded oligonucleotides with ends complementary to the SV40 genome were the only mixture that yielded XmnI$^+$ virus.

Nucleotide Sequences of Junctions

To confirm that the XmnI$^+$ virus arose from the injected oligonucleotides, the sequences around the XmnI sites in seven viral genomes were determined. As expected, all seven contained the diagnostic GGGG sequence that characterized the injected XmnI site. Five were exact joinings of the oligonucleotide to the BglII sites in the SV40 genome; two were exact at one end, but contained a deletion at the other end (Figure 3).

Table 1. Restriction analysis of viral genomes.

Mixture	Oligo	Nonphosphorylated BglII$^-$ BglII$^+$	XmnI$^-$	XmnI$^+$	Phosphorylated BglII$^-$ BglII$^+$	XmnI$^-$	XmnI$^+$	Total
A. BGLII ENDS								
1.	none	21	10	0	15	6	0	52
2.	1 + 2	120	50	15	30	12	2	229
3.	1 + 3	14	11	0				25
4.	1	12	6	0	8	0	0	26
5.	3	19	7	0				26
B. SMAI ENDS								
Mixture	Oligo	Nonphosphorylated SmaI$^-$ SmaI$^+$	XmnI$^-$	XmnI$^+$	Phosphorylated SmaI$^-$ SmaI$^+$	XmnI$^-$	XmnI$^+$	Total
6.	none	6	5	0	29	6	0	46
7.	1 + 3	8	13	0	42	7	0	70
8.	1 + 2	17	14	0				31
9.	1				18	2	0	20
10.	3	39	44	0				83

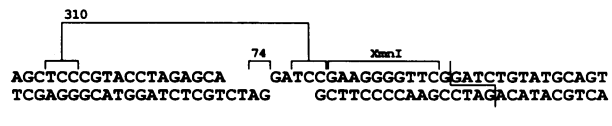

Figure 3. Two junctions at which the oligonucleotide was incoporated imprecisely. One end of the input oligonucleotide is shown at the gap; the other end is shown at the site where it joined to the SV40 end precisely (jagged line). The XmnI site is shown above the oligonucleotide. The junctions at the other side of the oligonucleotide are indicated by connecting lines above the sequence. For plaque isolate 310 the junction involved three nucleotides of homology (5'CTCCG), as shown by the bracket. The junction in plaque isolate 74 involved no homology.

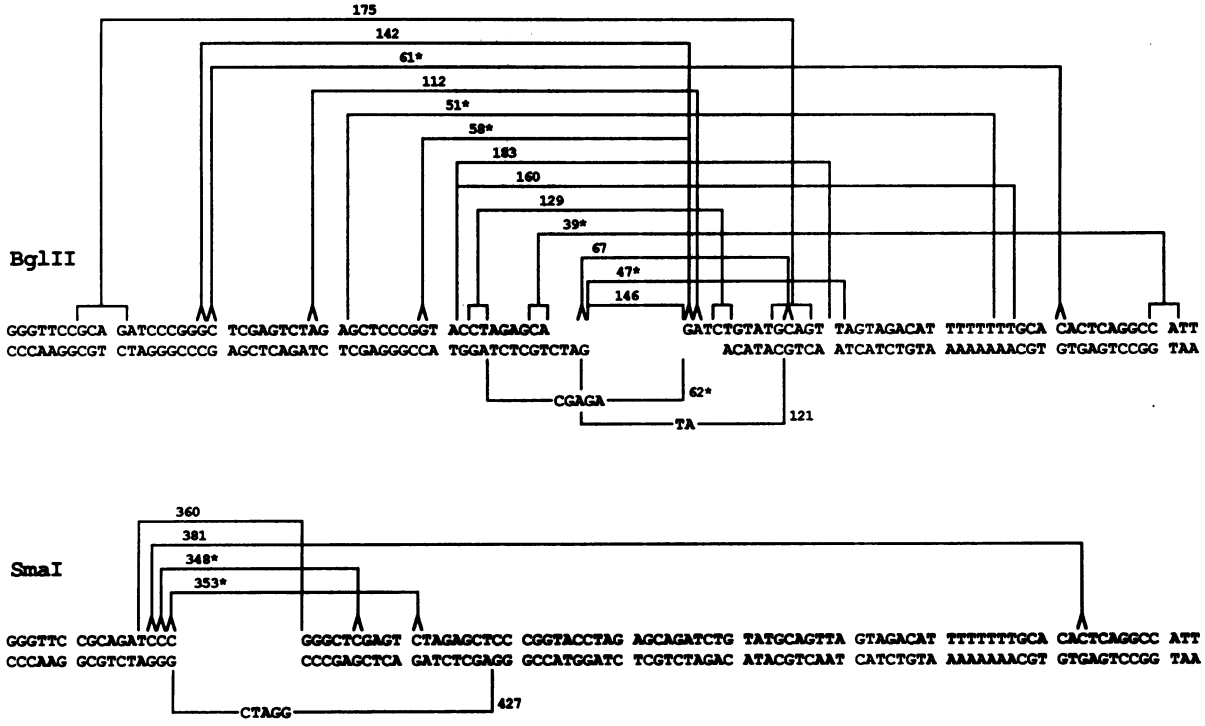

Figure 4. Junctions from 20 viruses that were XmnI$^-$ and BglII$^-$ or SmaI$^-$. Numbers identify different plaque isolates, which were all from experiments using nonphosphorylated oligonucleotides. Lines above the sequences indicate junctions that had no extra nucleotides; lines below the sequences indicate junctions that contained extra nucleotides, which are shown in a 3′ to 5′ orientation. Junctions that arose in injections with no added oligonucleotides are indicated with an asterisk. Brackets at the ends of some of the connecting lines indicate homologous nucleotides at the junctions. Isolates 67 and 112 are from mixture 2 (Table 1); isolates 121, 129, 142, 146 are from mixture 3; isolates 175 and 183 are from mixture 4; isolate 160 is from mixture 5; isolate 360 is from mixture 7; isolate 381 is from mixture 8; and isolate 427 is from mixture 10.

Nucleotide sequences were determined around the junctions in 12 BglII$^-$ or SmaI$^-$ viral genomes that arose in injections with nonphosphorylated oligonucleotides (Figure 4) along with 8 similar junctions in viruses from injections without oligonucleotides (marked with an asterisk in Figure 4). Among these 20 junctions, three were found that carried extra nucleotides. A five-nucleotide insert was present in isolate 62, which arose in the absence of injected oligonucleotides. Two-nucleotide and five-nucleotide inserts were present in isolates 121 and 427, which arose in mixtures 3 and 10, respectively (see Table 1). The two-nucleotide insert is too short to make an assignment, but the five-nucleotide insert exactly matches the terminal five nucleotides of the oligonucleotide used in the injection.

DISCUSSION

Our goals in this study were (i) to assess whether short defined oligonucleotides could be captured during the joining of DNA ends in mammalian cells, and (ii) if they were captured, to determine whether oligonucleotide addition was governed by the same rules that apply to end joining itself. Microinjection was chosen as the method for delivering DNA to the nucleus so that the ratio of oligonucleotides to SV40 genomes could be controlled. Microinjected oligonucleotides in quantities approaching 0.1% to 1% of the nuclear DNA mass did not kill cells directly, nor did they interfere detectably with SV40 infection (Figure 2). Thus, oligonucleotides, even in large amounts, do not seem to be toxic to cells.

We analyzed a total of 608 independent SV40 genomes from microinjected cells: 510 from injections with oligonucleotides and 98 from injections without oligonucleotides. Among the 510 viruses that arose from injections of SV40 genomes and oligonucleotides, we found unambiguous evidence for capture of double-stranded oligonucleotides with ends that were complementary to the BglII ends on the SV40 genomes. With this mixture oligonucleotides were found in 7% of total virus (or 22% of BglII$^-$ virus). Capture of oligonucleotides with fully complementary ends would require a DNA ligase but none of the rest of the machinery that has been hypothesized to promote efficient end joining in cells (1,7,17,18).

Among 281 junctions in viruses from injections of oligonucleotides and SV40 genomes that did not have complementary double-stranded ends, none (less than 0.4%) showed unambiguous evidence for a captured oligonucleotide. The sequenced junction in isolate 427 contained an insert that exactly matched the terminal five nucleotides of the injected single-stranded oligonucleotide, which is unlikely to occur by chance. Nevertheless, assignment of the filler DNA to the injected oligonucleotide is ambiguous in this case because the insert is part of a six-nucleotide direct repeat (3′CTAGGGCTAGGG) that spans the junction (Figure 4). This is a common motif among inserts (2,19,20) and has been hypothesized to arise by slipped mispairing near ends (8).

Among twenty sequenced junctions, we found extra nucleotides at three (15%), which is near the 10% value observed for all rearragements in nonlymphoid cells (2). The seventeen sequenced

junctions that did not contain extra nucleotides are similiar to the junctions we observed after DEAE dextran-mediated transfection (7). The deletion endpoints are not clustered and homology at the junctions ranged from 0 to 4 nucleotides (6 genomes had 0 nucleotides of homology; 8 genomes had 1 nucleotide; 2 had 2 nucleotides; and 1 had 4 nucleotides). Among these junctions there may be a few more nucleotides removed from the ends prior to end joining than was observed after DEAE dextran-mediated transfection. In our previous experiments, which analyzed the joining of mismatched ends, 97% of all junctions occurred within 15 nucleotides of the input end (7); in these studies, which examined the joining of matched ends, 90% of the junctions were within 15 nucleotides of the input end.

The data presented here demonstrate that, under the conditions tested, double-stranded oligonucleotides with ends fully complementary to the target DNA can be captured by end joining, whereas oligonucleotides with noncomplementary termini or blunt ends, as well as single-stranded oligonucleotides, are not substrates for efficient end joining. One trivial explanation for these results might be degradation of the input oligonucleotides. This could explain the lack of capture of single-stranded oligonucleotides; however, it is not obvious why double-stranded oligonucleotides with blunt ends should be less stable than oligonucleotides with complementary termini. The simplest interpretation is that capture of oligonucleotides by mismatched end joining is at best rare, and may not occur with high enough efficiency to participate significantly in the formation of filler DNA. Thus, the class of filler DNAs seemingly derived from other DNA molecules in the cell may originate from other mechanisms (for example, a replicative process) rather than end joining.

The lack of efficient capture of most of the tested oligonucleotides may reflect a property of the end-joining machinery. Previous results obtained using this cell line (CV1) demonstrated that complementary, blunt, or mismatched ends are joined with equally high efficiency, and suggested that one mechanism for joining might involve direct ligation of single strands (7). Recent studies using Xenopus eggs suggested that an end-pairing activity is involved in mismatched end joining (18); this activity apparently requires double-stranded DNA substrates at least 30 base pairs in length (P. Pfeiffer, personal communication). Perhaps a similar activity is necessary for efficient joining in mammalian cells; it is also possible that this activity does not recognize short oligonucleotides.

ACKNOWLEDGEMENTS

We thank De Dieu and Kathleen Marburger for excellent technical assistance. We thank Petra Pfeiffer for valuable discussions and the Molecular Biology Group in the Department of Biochemistry for helpful criticism. We thank Sandra Pennington for providing organizational help with the manuscript. This work was supported by grants from the National Institutes of Health (GM38219) and the American Cancer Society (CD-420).

REFERENCES

1. Roth,D.B. and Wilson,J.H. (1988) In Kucherlapati,R. and Smith,G.R. (eds.), *Genetic Recombination*. American Society for Microbiology, Washington, pp. 621–653.
2. Roth,D.B., Chang,X.-B. and Wilson,J.H. (1989) Mol. Cell. Biol. 9, 3049–3057.
3. Alt,F.W. and Baltimore,D. (1982) Proc. Natl. Acad. Sci. USA 79, 4118–4122.
4. Landau,N.R., Schatz,D.G., Rosa,M. and Baltimore,D. (1987) Mol. Cell. Biol. 7, 3237–3243.
5. Lafaille,J.J., DeCloux,A., Bonneville,M., Takagaki,Y. and Tonegawa,S. (1989) Cell 59, 859–870.
6. Schatz,D.G., Oettinger,M.A. and Baltimore,D. (1989) Cell 59, 1035–1048.
7. Roth,D.B. and Wilson,J.H. (1986) Mol. Cell. Biol. 6, 4295–4304.
8. Roth,D.B., Porter,T. and Wilson,J.H. (1985) Mol. Cell. Biol. 5, 2599–2607.
9. Wilson,J.H., Berget,P. and Pipas,J.M. (1982) Mol. Cell. Biol. 2, 1258–1269.
10. Roth,D.B. and Wilson,J.H. (1985) Proc. Natl. Acad. Sci. USA 82, 3355–3359.
11. Wilson,J.H., DePamphilis,M. and Berg,P. (1976) J. Virol. 20, 391–399.
12. Wake,C.T., Gudewicz,T., Porter,T., White,A. and Wilson,J.H. (1984) Mol. Cell. Biol. 4, 387–398.
13. Hirt,B. (1967) J. Mol. Biol. 26, 365–369.
14. Zagursky,R.J., Baumeister,K., Lomax,N. and Berman,M.L. (1985) Gene Anal. Techn. 2, 89–94.
15. Porter,T., Pennington,S.L., Adair,G.M., Nairn,R.S. and Wilson,J.H. (1990) Nucleic Acids Res. 18, 5173–5180.
16. Chang,X-B. and Wilson,J.H. (1987) Proc. Natl. Acad. Sci. USA 84, 4959–4963.
17. Pfeiffer,P. and Vielmetter,W. (1988) Nucleic Acids Res. 16, 907–924.
18. Thode,S., Schafer,A., Pfeiffer,P. and Vielmetter,W. (1990) Cell 60, 921–928.
19. Wessler,S., Tarpley,A., Purugganan,M.S. and Okagaki,R. (1990) Proc. Natl. Acad. Sci. USA 87, 8731–8735.
20. Gheysen,G., Villaroel,R. and Montagu,M.V. (1991) Genes Dev. 5, 287–297.