THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 277, No. 14, Issue of April 5, pp. 11756–11764, 2002
*Printed in U.S.A.*

# Repair of Sequence-specific $^{125}$I-induced Double-strand Breaks by Nonhomologous DNA End Joining in Mammalian Cell-free Extracts*

Received for publication, November 27, 2001
Published, JBC Papers in Press, January 30, 2002, DOI 10.1074/jbc.M111304200

**Andrea Odersky‡, Irina V. Panyutin§, Igor G. Panyutin§, Christian Schunck‡¶, Elke Feldmann‡, Wolfgang Goedecke‡, Ronald D. Neumann§, Günter Obe‡, and Petra Pfeiffer‡‖**

*From the ‡Institut für Genetik FB9, Universität Essen, Universitätsstrasse 5, D-45117 Essen, Germany and the §Department of Nuclear Medicine, Warren G. Magnussen Clinical Center, National Institutes of Health, Bethesda, Maryland 20854*

**In mammalian cells, nonhomologous DNA end joining (NHEJ) is considered the major pathway of double-strand break (DSB) repair. Rejoining of DSB produced by decay of $^{125}$I positioned against a specific target site in plasmid DNA via a triplex-forming oligonucleotide (TFO) was investigated in cell-free extracts from Chinese hamster ovary cells. The efficiency and quality of NHEJ of the "complex" DSB induced by the $^{125}$I-TFO was compared with that of "simple" DSB induced by restriction enzymes. We demonstrate that the extracts are indeed able to rejoin $^{125}$I-TFO-induced DSB, although at approximately 10-fold decreased efficiency compared with restriction enzyme-induced DSB. The resulting spectrum of junctions is highly heterogeneous exhibiting deletions (1–30 bp), base pair substitutions, and insertions and reflects the heterogeneity of DSB induced by the $^{125}$I-TFO within its target site. We show that NHEJ of $^{125}$I-TFO-induced DSB is not a random process that solely depends on the position of the DSB but is driven by the availability of microhomology patches in the target sequence. The similarity of the junctions obtained with the ones found *in vivo* after $^{125}$I-TFO-mediated radiodamage indicates that our *in vitro* system may be a useful tool to elucidate the mechanisms of ionizing radiation-induced mutagenesis and repair.**

Mammalian genomes constantly suffer a variety of types of damage, of which double-strand breaks (DSB)[1] are considered the most dangerous. DSB may arise spontaneously in the cell or may be induced by exogenous agents, such as ionizing radiation. The estimation that mammalian cells suffer at least 10 spontaneous DSB/day suggests that efficient repair of DSB is critical for cell survival (1). Failure to do so can result in deleterious genomic rearrangements, cell cycle arrest, or cell death.

Recent studies have revealed that DSB in the genomes of higher eukaryotes can be repaired by at least three different pathways (2): (i) Homologous recombination repair, the most accurate process, is able to restore the original sequence at the break. Because of its strict dependence on extensive sequence homology, this mechanism is suggested to be active mainly during the S and $G_2$ phases of the cell cycle (3, 4). (ii) Single-stranded annealing is another homology-dependent but less accurate process that can repair DSB between direct repeats and thereby produces mainly interstitial deletions (4). (iii) Nonhomologous DNA end joining (NHEJ) comprises at least two different processes (5). The major and best investigated NHEJ pathway depends on the Ku70/80 heterodimer, the catalytic subunit of the DNA-dependent protein kinase, DNA ligase IV, and its essential co-factor XRCC4 (6, 7). In contrast to homologous recombination repair and single-stranded annealing, NHEJ can operate in the absence of sequence homology (although short sequence homologies, so-called microhomologies, may facilitate the process) and is able to rejoin broken ends directly (2). This process is supposed to occur mainly in the $G_0$ and $G_1$ phases of the cell cycle and is considered to be the major pathway of DSB repair in mammalian cells, although it is typically accompanied by loss or gain of a few nucleotides. The regulation of these different pathways and their relative contributions to mammalian DSB repair have yet to be comprehended (1).

To elucidate the mechanisms of NHEJ, many studies have made use of restriction endonucleases (RE) to introduce defined DSB in the genomic DNA of cultured mammalian cells (8–13) or in plasmids to be offered as DSB substrates in transfection assays (14–16) or cell-free extracts (17–22). The fact that RE-induced DSB are exactly defined with respect to their structure (depending on the enzyme used: 5′- or 3′-overhangs or blunt ends; always 3′-hydroxyl and 5′-phosphate) and position within a given DNA sequence has greatly facilitated study of the efficiency and fidelity of DSB repair mechanisms in the above-mentioned systems by comparing the original DSB termini and the resulting repair site (junction). As opposed to such "clean" DSB, which are repaired very efficiently because they are accepted substrates of DNA-modifying enzymes, DSB generated by ionizing radiation or certain chemical agents are more complex and may, for instance, contain damaged sugar and base moieties and 5′-hydroxyl and 3′-phosphate groups. In addition, the investigation of the repair of such complex DSB on the molecular level is aggravated by the fact that these "dirty" DSB are usually randomly distributed and not positioned within a specific DNA sequence. Experimental approaches comprise the

* This work was supported by Grant 96.053.2 from the Wilhelm-Sander-Stiftung für Krebsforschung (to P. P.) and by a fellowship of the Heisenberg-program of the Deutsche Forschungsgemeinschaft (to P. P.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Present address: MetaSystems GmbH, Robert-Bosch-Str. 6, D-68804 Altlussheim, Germany.

‖ To whom correspondence should be addressed: Institut für Genetik FB9 (S05 T04 B26), Universität Essen, Universitätsstr. 5, D-45117 Essen, Germany. E-mail: petra.pfeiffer@uni-essen.de.

[1] The abbreviations used are: DSB, double-strand break(s); bl., blunt; coh. cohesive; ccc, covalently closed circle; NHEJ, nonhomologous DNA end joining; oc, open circle; Pu, purine; Py, pyrimidine; RE, restriction enzyme; SSB, single-strand break(s); TFO, triplex-forming oligo; CHO, Chinese hamster ovary; MOPSO, 3-(*N*-morpholino)-2-hydroxypropanesulfonic acid.



This paper is available on line at http://www.jbc.org

This is an Open Access article under the CC BY license.

analysis of the mutational spectra generated by ionizing radiation or chemicals in selectable cellular genes (23) and the use of oligonucleotides with unusual terminal structures in cell-free extracts (24) and plasmids carrying at their ends oligonucleotides damaged by bleomycin (25–27).

A novel approach called gene-targeted radiotherapy has recently opened the possibility to target the radiodamage produced by Auger electron emitters such as $^{125}$I to a specific DNA sequence (as opposed to random targeting of total genomic DNA in traditional radiotherapy) (28). Auger electron emitters are a large group of radioisotopes that decay by electron capture and/or conversion emitting a cascade of low energy electrons that produces a highly charged daughter atom. The combined effect of low energy electrons and positively charged daughter atoms results in highly localized damage to the molecular structures within a short range from the decay site (Auger effect). Decay of $^{125}$I results in emission of, on average, 21 electrons and produces a correspondingly positively charged tellurium atom. Incorporated into DNA, the decay of $^{125}$I produces DSB localized mostly within one turn of the double-helix around the decay site (10 bp) with an efficiency of 0.8 DSB/decay. This extremely short range of radiodamage produced by $^{125}$I led to the idea of targeting this Auger electron emitter to specific genes within genomic or plasmid DNA (29).

Sequence-specific delivery of $^{125}$I-induced radiodamage is achieved by the use of triplex-forming oligonucleotides (TFO), short single-stranded oligonucleotides capable of forming triple helixes (triplexes) with polypurine:polypyrimidine sequences. In such triplexes, the TFO occupies the major groove of the target double-helix and forms Hoogsteen hydrogen bonds with the purines of the Watson-Crick base pairs. The specificity of sequence recognition is comparable with that provided by complementary Watson-Crick base pairing (30–32).

To investigate the repair of site-specific $^{125}$I-induced DSB, a TFO labeled on its 3′-end with $^{125}$I ($^{125}$I-TFO) was used to introduce DSB within its target sequence on plasmid pUC19-MDR1 (33). The linearized plasmid was incubated with cell-free extracts from CHO cells capable of performing efficient NHEJ (5). We show that the repair of the $^{125}$I-induced DSB is about a factor of 10 less efficient than the repair of RE-induced DSB. The resulting spectrum of junctions shows deletions of varying sizes resembling the ones found in selectable genes after irradiation of mammalian cells with ionizing radiation. Our study may contribute to the understanding of how the damage produced by Auger electron emitters is repaired by mechanisms of NHEJ, which is important for their application in gene-targeted radiotherapy.

## EXPERIMENTAL PROCEDURES

### *Cell Culture*

The two wild-type Chinese hamster ovary cell lines, CHO-K1 and AA8, were grown at 37 °C in a humidified 5% $CO_2$ atmosphere in Ham's F-12 medium enriched with 10% fetal calf serum, 2 mM L-glutamine, 100 units/ml penicillin, and 100 μg/ml streptomycin.

### *Cell-free Extracts*

Whole cell extracts from CHO-K1 and AA8 cells were prepared exactly as described previously (5, 17). In each preparation, ~$5 \times 10^8$ cells of each cell line were used to yield 0.5–1 ml of extract with a protein concentration ranging between 6–10 mg/ml. The extracts were stored in 50-μl aliquots in liquid nitrogen and remained active for 6–12 months. Directly prior to use in the NHEJ reaction, the extract aliquot was dialyzed against freshly prepared M buffer (50 mM MOPSO-NaOH, pH 7.5, 40 mM KCl, 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol) on microdialysis filters (0.025-μm pore diameter; catalog number VSWPO2500; Millipore) for 30 min at 4 °C.

### *DNA Substrates*

*$^{125}$I-TFO-induced DSB*—Labeling of the TFO with $^{125}$I-dC was performed by extension of the 3′-end of a primer in the presence of $^{125}$I-dCTP (PerkinElmer Life Sciences) and Klenow fragment of DNA polymerase I as described previously (33). To form a triplex, topoisomerase-relaxed pUC19-MDR1, a 2727-bp derivative of pUC19 containing a 32-bp polypurine-polypyrimidine fragment from the MDR1 gene as TFO-target sequence (see Fig. 1 and Refs. 33 and 34) was mixed with purified $^{125}$I-TFO in 30 mM NaAc buffer, pH 5.0, and heated to 70 °C for 3 min followed by slow cooling to room temperature. For the accumulation of $^{125}$I decays, the sample was stored at −70 °C. After a period of 60 days (the half-life of $^{125}$I), about 50% of total covalently closed circular (ccc) pUC19-MDR1 was converted to open circle (oc), and about 20% was converted to linear DNA indicative of double-strand breakage of the plasmid as estimated by separation of the products in 1.5% agarose gels containing ethidium bromide. To remove contaminating oc and ccc DNA, the linear form of pUC19-MDR1 was purified twice over 1.5% preparative low melting point NuSieve agarose (BioProducts FMC) gels in TAE buffer (40 mM Tris-HAc, pH 7.4, 12 mM NaAc, 0.1 mM EDTA) containing 0.5 μg/ml ethidium bromide. Electrophoresis was performed at 2 V/cm for 24 h with continuously recirculated TAE buffer containing 0.5 μg/ml ethidium bromide, and separation of DNA in oc, linear, and ccc forms was visualized under UV light. Linear DNA was purified using the Agar ACE™ agarose-digesting enzyme (Promega) according to the manufacturer's instructions. The samples were purified further by two extractions with phenol and phenol:chloroform: isoamyl alcohol (25:24:1; Invitrogen) and precipitated with ethanol. After resuspension in 50 μl of TE (10 mM Tris-HCl, pH 7.6, 0.1 mM EDTA) the samples were finally purified by gel filtration through G-50 Microspin columns (Amersham Biosciences). The resulting linearized pUC19-MDR1 substrate used in extract joining assays was found to contain on the average less than 5% of contaminating oc DNA and no ccc DNA at all.

*RE-induced DSB*—The three substrates for cohesive (coh.) and blunt (bl.) end ligation were derived from pUC19-MDR1 by digestion with a single restriction enzyme (*Bam*HI: 5′-coh.; *Pst*I: 3′-coh.; *Hin*cII: bl.). The five substrates containing noncomplementary ends were derived from a 4-kb modified pUC19-MDR1-λ construct harboring a 1.25-kb fragment of λ-DNA between the restriction sites used for substrate preparation. Generation of substrates containing two noncomplementary ends was controlled by quantitative excision of the λ insert. Each substrate was named after the pair of RE used in its preparation (*Eco*/*Asp*, 5′/5′; *Sac*/*Kpn*, 3′/3′; *Eco*/*Sma*, 5′/bl.; *Sac*/*Sma*, 3′/bl.; *Eco*/*Kpn*, 5′/3′). All RE-linearized substrates were gel-purified using a gel extraction kit (Qiagen).

### *Assay for NHEJ and Analysis of Products*

In standard reactions, 10 ng of $^{125}$I-TFO- or RE-linearized plasmid substrate, respectively, were incubated for up to 360 min at 25 °C in a total volume of 10 μl containing 6–8 μg/μl of extract protein in M buffer supplemented with 1 mM ATP, pH 7.5, and 200 μM dNTPs (50 μM each) and 50 ng/μl bovine serum albumin. The reactions were terminated by adjustment to 20 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1% SDS and incubation at 65 °C for 5 min. After digestion for 30 min at 37 °C with 2 mg/ml proteinase K, equivalents of 2 ng of substrate DNA were electrophoresed in 1% agarose gels in the presence of 1 μg/ml ethidium bromide to separate oc from ccc products and visualized by *in situ* gel hybridization (35) using a pUC19-specific probe labeled with [$^{32}$P]α-dCTP by random priming. Reaction products were quantified in a phosphorimaging facility (Packard Bioscience) as percentages of the total radioactivity/lane. Circular joined products were cloned by transformation of 4-ng equivalents of substrate DNA of each NHEJ sample in *Escherichia coli* strain DH5α to yield single clones that were purified by miniscale extraction. In the case of $^{125}$I-TFO-linearized pUC19-MDR1, the samples were digested with *Bgl*II prior to transformation to remove oc contaminants originating from substrate preparation that could yield false positives. Clones from $^{125}$I-TFO-linearized pUC19-MDR1 were subjected again to cleavage with *Bgl*II, and only *Bgl*II-resistant clones were analyzed by sequencing (Seqlab). The clones from ligation products (*Bam*, *Pst*, and *Sma*) were subjected to cleavage with the original RE to check for accurate ligation. The clones from NHEJ products (*Eco*/*Asp*, *Sac*/*Kpn*, *Eco*/*Sma*, *Sac*/*Sma*, and *Eco*/*Kpn*) were analyzed directly by sequencing (ABI Prism 377 DNA Sequencer; PerkinElmer Life Sciences).

For the analysis of dimer products from $^{125}$I-TFO-linearized pUC19-MDR1, the dimer band was gel-purified using a gel extraction kit (Qiagen). Dimer junctions were amplified by PCR with 2.5 units of *Taq* polymerase in *Taq* buffer (MBI Fermentas) in a total volume of 50 μl containing 1 ng of dimer product, 20 pmol of each primer (pUC19-MDR1-For, 5′-GGGGCCTCTTCGCTATTACG; pUC19-MDR1-Rev, 5′-

AGGCACCCCAGGCTTTACACTTTA), 2.5 mM $MgCl_2$, and 200 $\mu$M of each dNTP. PCR was performed in a thermocycler (PerkinElmer Life Sciences) for 30 cycles (30 s 95 °C; 30 s 54 °C; 1 min 72 °C). The resulting 300-bp PCR product was digested with *Bgl*II to remove PCR products possibly originating from oc contaminants. *Bgl*II-resistant PCR product was gel-purified and subcloned using a cloning kit (Invitrogen). The resulting clones were purified by miniscale extraction and subjected again to cleavage with *Bgl*II, and only *Bgl*II-resistant clones were analyzed by sequencing.

### *Calculations for Fig. 8*

For the diagrams in Fig. 8 (*A* and *B*), the following calculations were performed.

*Distribution of DSB*—The distribution of breaks around the $^{125}$I decay site had been measured previously as single-strand breaks (SSB) occurring in the Pu-rich and Py-rich strand, respectively (33) (see *bars* in Fig. 1) and is given here as the average probability ((Pu + Py)/2) of all types of DSB (*gray bars* in Fig. 8, *A* and *B*; see "Discussion" for details) to occur at a given base pair position.

*Relative Frequencies of Junction Breakpoints*—The relative frequencies for the occurrence of the breakpoints of junctions 2–34 (see Fig. 5) at a particular nucleotide (*black bars* in Fig. 8*A*) were calculated as follows: (i) Blunt junctions. The number of a particular junction was normalized to the total number of junctions (64) and divided by two because the breakpoint can be either counted to the left or to the right side of the deletion (*e.g. junction 8* in Fig. 5); because this junction occurred twice, its relative frequency would be 2/64 = 0.0313. Because the breakpoint can be counted either to the A on the left side or to the C on the right side, the relative frequency of this breakpoint at the A and C, respectively, is 0.0156. (ii) Microhomology junctions. The calculation was performed as for blunt junctions with the additional inclusion of a factor for the microhomology (*e.g. junction 23* in Fig. 5) that occurs twice and exhibits a 2-bp homology (AG) with three possible breakpoints. Therefore, the relative frequency of the breakpoints would be $2/64 \times 2 \times 3 = 0.0052$ for any nucleotide within the microhomology and each of the nucleotides flanking the microhomology on the left and right side, respectively (A, A, and G on the left side and the A, G, and T, on the right side). Each *black bar* in the diagram represents the sum of the relative frequencies of all breakpoints occurring at a particular nucleotide of the target sequence.

*$\chi^2$ Test*—The $\chi^2$ test was performed for Fig. 8*A*. Multiplication of the average probability of a DSB at a given base pair by the number of total junctions [(Pu + Py)/2] × 4 yields the expected frequency (*E*) of a junction to occur at this base pair, which was compared with the observed frequency (*O*) of junctions occurring at this position. The $\chi^2$ value was calculated using the formula $(O - E)^2/E$. Because the distribution of observed junctions spans a larger sequence region (27 bp) than the distribution of DSB (19 bp; 5′-GAAG. . . . GAGT), only the junctions falling into this 19-bp region were taken into account resulting in 19 categories yielding a degree of freedom of 18. The estimated $\chi^2$ value is 49.54 ($\Sigma(O - E)^2/E$) and significantly larger than 28.87, the value for the 5% interval for the degree of freedom of 18. Therefore, the hypothesis that junction formation is a random process that follows the distribution of the $^{125}$I-TFO-induced DSB has to be rejected (see "Discussion").

*Distribution of Deleted Nucleotides*—The distribution of nucleotides deleted around the decay site (see *black bars* in Fig. 8*B*) was calculated as follows. In the 64 junctions (see Fig. 4, *junctions 2–34*), a total of 422 nucleotides were deleted (*e.g.* the G* in the target sequence that was lost unambiguously in 41 cases and was part of a microhomology in 14 cases; see the *dots* in the sequences of Fig. 5). Because it was unknown from which of the two DSB ends the G in the corresponding microhomology originated, 14 was divided by two (14/2 = 7) so that the relative frequency at which the G* is lost in all 64 junctions is (41+7)/422 = 0.1137. Each *black bar* in the diagram represents the relative frequency with which a particular nucleotide was deleted from the target sequence.

## RESULTS

*Experimental System*—Annealing the $^{125}$I-TFO to its target sequence within pUC19-MDR1 and subsequent incubation for 60 days at −70 °C yield sequence-specific DSB within a short region of about 10 bp in each direction opposite the $^{125}$I-dC within the unique *Bgl*II site (A/**G**ATCT) of the plasmid (Fig. 1). The distribution and relative frequencies of breaks had been determined previously by analysis of the SSB occurring in the Pu- and Py-rich strand, respectively, which is indicated schematically in Fig. 1 (33). The slightly asymmetric distribution of SSB in the two strands reflects the structure of the Py motif triple helix.


FIG. 1. **Sequence of the TFO (*shaded*) and location of its target sequence (*bold*) within the polylinker of pUC19-MDR1 (33).** The *enlarged C* at the 3′-end of the TFO indicates the $^{125}$I-dC whose decay induces sequence-specific breaks that were measured previously as intensities of SSB at individual bases in the Pu- and Py-rich strand, respectively, in relative units (*bars*). The maxima occur opposite of the $^{125}$I-dC in both strands and coincide with the unique *Bgl*II site (*underlined*) of pUC19-MDR1 (33).

Gel-purified $^{125}$I-TFO-linearized pUC19-MDR was subjected to DNA end joining in cell-free extracts from CHO-K1 and AA8 cells as described under "Experimental Procedures." For comparison, extract joining reactions were also carried out with pUC19-MDR1 linearized by restriction endonucleases. NHEJ reaction products were separated in agarose gels, and the corresponding repair sites (junctions) were cloned in *E. coli* for subsequent sequence analysis.

*Efficiency of NHEJ of $^{125}$I-TFO-induced DSB Compared with RE-induced DSB*—To determine the efficiency of NHEJ of the $^{125}$I-TFO-linearized substrate, we used different RE-linearized substrates for comparison. Substrates generated by cleavage with a single RE have compatible ends that allow measurement of the efficiency of ligation of cohesive 5′- (*Bam*) or 3′-ends (*Pst*), respectively, or blunt ends (*Hinc*II). Substrates generated by cleavage with two different RE have noncomplementary DNA ends (*Eco/Asp*, 5′/5′; *Sac/Kpn*, 3′/3′; *Eco/Sma*, 5′/bl.; *Sac/Sma*, 3′/bl.; *Eco/Kpn*, 5′/3′) that allow measurement of the efficiency of genuine nonhomologous end joining. This type of end joining is more complex and requires more factors than "simple" cohesive or blunt end ligation because the ends must be converted first into a ligatable form by DNA fill-in synthesis and/or exonucleolytic removal of nonmatching bases (36) (Fig. 2 and below). Rejoining of $^{125}$I-TFO-induced DSB is expected to be even more complex because these dirty breaks may contain damaged sugar and base moieties, 5′-hydroxyl and 3′-phosphate groups that are not substrates for DNA-modifying enzymes such as DNA ligase or DNA polymerase and therefore must be removed prior to NHEJ (25, 27). In addition to that, it is important to note that each RE substrate contains only a single type of DSB with ends exactly defined in structure and sequence. In contrast, the $^{125}$I-TFO substrate represents a mixture of molecules containing many different types of DSB because of the fact that the $^{125}$I-TFO induces multiple breaks distributed along a 19-bp region (see also "Discussion" and Fig. 8). Therefore, the term "complex DSB" used below not only includes the presumptive dirty DSB but also a large variety of DSB ends differing in structure and sequence.

The extract-mediated NHEJ reaction converts all three different substrate types into monomeric oc reaction intermediates, ccc products, and linear multimers (mostly dimers), which are readily separated in agarose gels. In standard reactions, about 30–50% of the RE substrate input are converted into ccc and dimer products and the ratio of ccc:dimer product is ~2:1 (but may vary with the batch of extract used and other factors

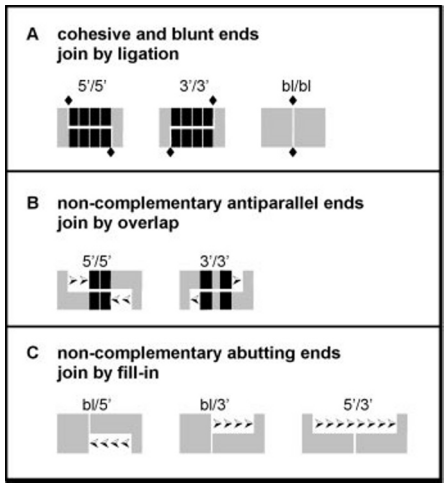


FIG. 2. **Major pathways of accurate NHEJ as observed *in vitro* with RE substrates in *Xenopus* eggs (18) and mammalian cells (5, 17) that involve the Ku70/80 heterodimer, the catalytic subunit of the DNA-dependent protein kinase, DNA ligase IV, and the associated XRCC4 protein (5).** *A*, cohesive (5′/5′; 3′/3′; *black boxes* represent complementary bases in cohesive overhangs) and bl. ends are joined by ligation (*black diamonds)* to restore the original restriction site. *B*, DNA ends with noncomplementary anti-parallel 5′/5′- or 3′/3′-overhangs form short mismatched overlaps at positions of complementary bases (*black boxes*) that determine the patterns of DNA fill-in synthesis (*arrowheads*) (37, 38). *C*, sequences of 5′- and 3′-overhangs in abutting terminus configurations are preserved by fill-in synthesis (*arrowheads*) (35). Although fill-in of a 5′-overhang can be primed at the recessed 3′-OH group of the same end, fill-in of a 3′-overhang can be primed only at the 3′-OH of the abutting terminus, which may be a blunt end or 5′-overhang.

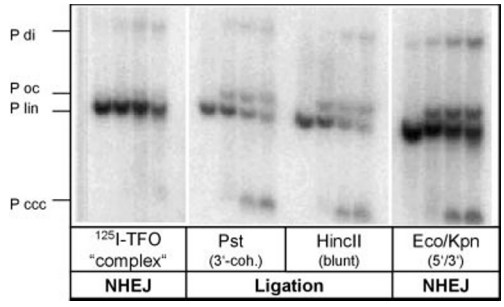


FIG. 3. **Examples of kinetics of ligation and NHEJ in the CHO-K1 extract.** The indicated $^{125}$I-TFO- and RE-linearized pUC19-MDR1 substrates were incubated with CHO-K1 extract at 25 °C for different times (each *block* of *four lanes* corresponds to 0, 15, 75, and 360 min). The reaction products were separated by agarose gel electrophoresis and visualized by *in situ* hybridization with a pUC19-specific $^{32}$P-labeled probe and subsequent exposure to an x-ray film. All of the samples shown originate from the same experiment but different gels. Band designations are linear plasmid (*P lin*) for the input substrate, open circular (*P oc*), and covalently closed circular (*P ccc*) products and linear dimer (*P di*) product.

like protein concentration and DNA concentration). We did not find any quantitative or qualitative differences between the CHO-K1 extract and the AA8 extract. A representative example of the reaction kinetics of three of the eight RE substrates and the $^{125}$I-TFO substrate is given in Fig. 3. As reflected by the levels of ccc product formation after 6 h at 25 °C, the reaction is most efficient with the ligation of cohesive (*Pst*) and blunt ends (*Hin*cII) that converts on the average 37% of the input substrate into ccc product (and 12% into dimers). Rejoining of noncomplementary RE ends (*Eco/Kpn*) is somewhat less efficient and converts on the average 29% of the linear input into ccc product (and 13% into dimers). For the $^{125}$I-TFO-linearized substrate, however, ccc product formation is drastically decreased to 2.3% (6.8% dimers) and reaches only about one-tenth of the efficiency obtained with RE-induced DSB. This decrease in efficiency is consistent with the assumption that complex DSB require more extensive modifications to be converted into a form that is accepted by the DNA-modifying enzymes participating in the NHEJ reaction (*e.g.* DNA ligase IV).

*Analysis of Junctions*—Isolation of single NHEJ events for sequence analysis of the junctions was achieved by two different strategies: (i) transfection of total reaction products in *E. coli*, which results in preferential cloning of the junctions in circular products (with decreasing efficiency for ccc > oc ≫ lin) and (ii) PCR amplification of junctions of gel-purified linear dimers and subsequent subcloning in *E. coli* to produce single clones suitable for sequence analysis.

Because the $^{125}$I-TFO substrate represents a mixture of plasmid molecules containing a large variety of different DSB, it can be expected that the spectrum of junctions obtained from this substrate is more heterogeneous than the spectra of junctions obtained from the different RE substrates. In addition, the presence of dirty DSB may reduce the fidelity of NHEJ. We therefore investigated the sequences of 96 junctions derived from the RE substrates (12 junctions for each of the 8 different substrates; Fig. 4) and 71 junctions derived from the $^{125}$I-TFO substrate (Fig. 5).

*RE-induced DSB Are Rejoined with High Accuracy*—To investigate the fidelity of the NHEJ reaction using different substrates, it is important to define the term "accurate NHEJ" (Fig. 2). Although it is obvious that "accurate ligation" of complementary cohesive or blunt restriction ends restores the original restriction site used to create the DSB (Fig. 2*A*), the definition of accurate NHEJ is not self-evident because joining of noncomplementary restriction ends necessarily causes a change in the original sequence. Still, general rules were established for NHEJ of noncomplementary ends because extracts from *Xenopus* eggs (18) and mammalian cells (5, 17) generate highly reproducible spectra of junctions using two main pathways: the "overlap" and "fill-in" pathways (Fig. 2, *B* and *C*). The pathway used is determined by the structure of the ends being joined; although the overlap pathway typically joins DNA ends containing 5′- or 3′-anti-parallel single-stranded overhangs (5′/5′; 3′/3′), the fill-in pathway joins abutting DNA ends (5′/bl.; 3′/bl.; 5′/3′). In the first case, the ends form incompletely matched overlaps by pairing of single fortuitously complementary bases, and the overlap structure determines the patterns of subsequent repair reactions (Fig. 2*B*) (38). In the second case, the sequences of participating 5′- or 3′-overhangs are preserved fully by fill-in DNA synthesis in a process in which the ends are transiently held together (presumably by the Ku70/80 heterodimer) (5) so that the 3′-hydroxyl group of the 5′-overhang or blunt end can serve as a primer to direct repair synthesis of the 3′-overhang (Fig. 2*C*) (35).

Cloning of single joining events was achieved by transformation of circular products in *E. coli*. Here, we have analyzed 36 cloned junctions derived from the three RE substrates containing complementary ends (Fig. 4, *Ia*, *Ib*, and *Ic*), and 60 from the five RE substrates containing noncomplementary ends (24 overlap junctions (Fig. 4, *IIa* and *IIb*) and 36 fill-in junctions (Fig. 4, *IIIa*, *IIIb*, and *IIIc*)).

The spectra of ligation junctions (Fig. 4, *Ia*, *Ib*, and *Ic*) show that the accuracy of ligation is high and reaches 100% for 5′-cohesive and blunt ends and 92% for 3′-cohesive ends. The accuracy of NHEJ is slightly decreased when compared with ligation but still high with 50 and 66%, respectively, for the overlap junctions (Fig. 4, *IIa* and *IIb*) and 83, 25, and 67% for the fill-in junctions (Fig. 4, *IIIa*, *IIIb*, and *IIIc*). These results are consistent with previous studies (5) and show that the NHEJ reaction is a highly accurate process, at least on substrates generated by restriction endonucleases producing clean

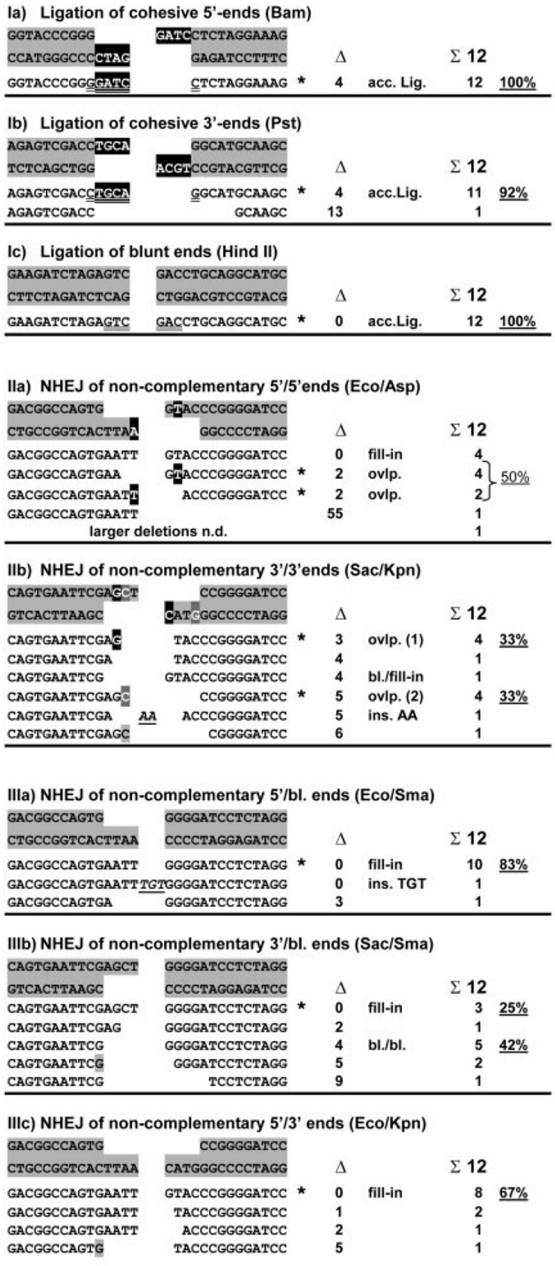


FIG. 4. **Spectrum of junctions generated by NHEJ in the CHO-K1 extract.** *Ia*, *Ib*, and *Ic*, ligation of cohesive and blunt ends. *IIa* and *IIb*, NHEJ of anti-parallel ends by overlap formation (*ovlp.*). *IIIa*, *IIIb*, and *IIIc*, NHEJ of abutting ends by fill-in. Terminus configurations (*shaded*) including the flanking double-stranded sequences are shown at the *top* of each panel with complementary bases used for overlap formation between anti-parallel ends highlighted by *white letters* on a *black background*. The junctions are listed below as top strand sequences with accurate junctions marked by *asterisks*. *Double-underlining* indicates a restored restriction site by accurate ligation (*acc. lig.*) used for quicker analysis in ligation assays. Overlaps formed by *Eco/Asp* via the T:A match contain a G:A mismatch that segregates in *E. coli* into either G:C or T:A (37). *Sac/Kpn* can form two different overlaps marked by *(1)* and *(2)*. *(1)* uses the G:C match marked in *black*, and *(2)* uses the C:G match in *gray*. The total numbers of clones analyzed are listed next to Σ, and the total numbers of bases lost from both ends are listed below Δ. Microhomologies at junction breakpoints not originating from overlap formation between single-stranded overhangs are marked in *black letters* on a *gray background*. Untemplated bases inserted (*ins.*) at junction breakpoints are printed in *underlined bold italic letters*.

ends that are accepted substrates of DNA-modifying enzymes.

*Rejoining of $^{125}$I-TFO-induced DSB Produces a Highly Heterogeneous Spectrum of Junctions*—Radiodamage delivered to

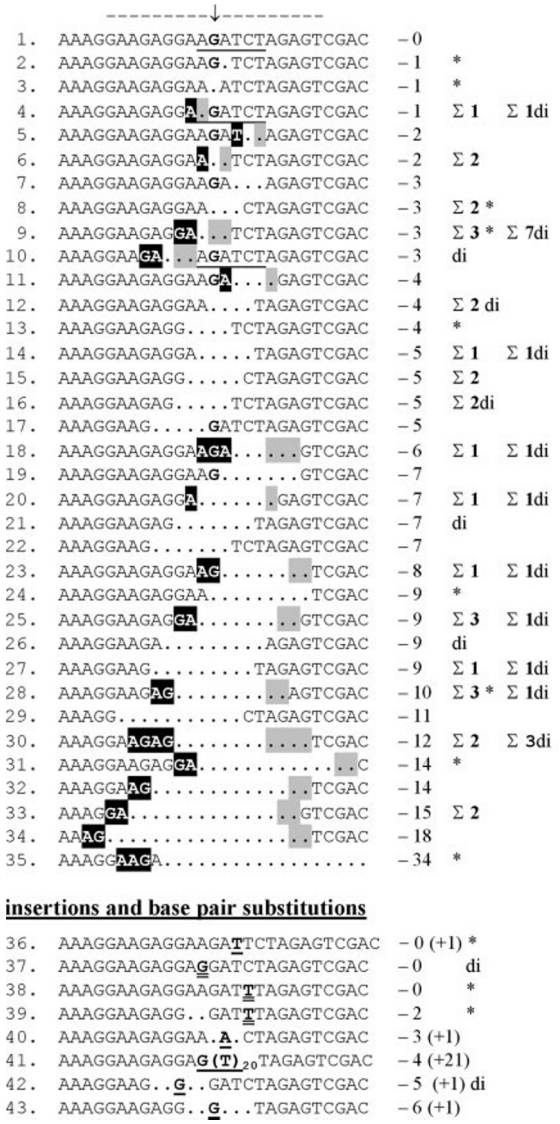


FIG. 5. **Sequences of the junctions in ccc and dimer (*di*) products formed in cell-free extracts from CHO-K1 and AA8 (*) cells.** *Sequence 1* represents the original sequence with intact *Bgl*II site (*underlined*) of the undamaged plasmid; the *bold G* marked by a *vertical arrow* indicates the position of the $^{125}$I in the TFO, and the *dashes* mark the region in which SSB occur in the Pu- and Py-rich strand, respectively (see Fig. 1). Bases deleted in junctions are indicated by *dots*, and the size of the corresponding deletion is given as a *negative numeral* on the *right*. Microhomology patches at junction breakpoints are marked in *white letters* on a *black background* on the *left side*; the corresponding matching bases are in *gray* on the *right side* (note that the microhomology patches were arbitrarily attributed to the left side, although it is impossible to determine which of the nucleotides participated in match formation). For each junction, the total number of clones is given behind the Σ. Σ*di*, total numbers of junctions derived from dimer products. The *asterisks* mark single sequences derived from $^{125}$I-TFO substrate treated with AA8 extract (*e.g.* with Σ3*, two clones were derived from the CHO-K1 extract and one from the AA8 extract). Junctions harboring additional nontemplated nucleotides (*underlined*) or altered bases (*doubly underlined*) are listed under "*insertions and base pair substitutions*."

pUC19-MDR1 by the $^{125}$I-TFO accumulates within a short region of 19 bp around the G in the single *Bgl*II site (A*G*ATCT). As mentioned under "Experimental Procedures," the linear substrate contained up to 5% of oc contaminant. The oc molecules are intermediates that arise during the decay process of the $^{125}$I and contain multiple SSB. Only molecules receiving two closely spaced SSB (one in each strand separated by less than 10 bp) will give rise to linear molecules. Staggered SSB in

opposite strands located further apart will probably not produce linear molecules because long single-stranded tails will melt only upon heating and reanneal instantaneously after cooling so that these molecules will exist most likely in oc form. Furthermore, recent analysis of purified linear $^{125}$I-TFO substrate revealed that in addition to highly localized breaks around the TFO binding site, 25% of the DSB occur outside of a 90-bp fragment containing the TFO-binding motif (33). This out-of-target damage is probably caused by (i) higher energy electrons produced by decay of $^{125}$I and/or (ii) the Auger effect itself if segments of the same molecule or other molecules come close to $^{125}$I because of condensation of DNA in solution. The presence of DSB outside of the target site and the presence of oc-contaminants led us to use a selection procedure to avoid sequencing of large fractions of clones not damaged in the relevant region.

Because the maximal frequency of DSB occurs within and around the single *Bgl*II site and NHEJ of a radiation-induced DSB within the *Bgl*II site is, *a priori*, not expected to restore the site, we have used resistance to cleavage with *Bgl*II as a marker for successful rejoining of the $^{125}$I-TFO-linearized substrate. Therefore, joining products were digested with *Bgl*II prior to transfection in *E. coli* to remove the bulk of oc contaminants (which would also give rise to clones), and the resulting clones were again checked for cleavage with *Bgl*II. A total of 44 *Bgl*II-resistant clones were subjected to sequence analysis, and the junctions are shown in Fig. 5. To obtain a more reliable picture of the NHEJ mechanism that rejoins complex DSB, we also analyzed the junctions arising in the dimer fraction. For this, gel-purified dimers were subjected to PCR, which amplifies exclusively molecules in head-to-tail orientation (equivalent to circular products; because of their palindromic nature, the simultaneously arising head-to-head and tail-to-tail molecules cannot be analyzed). After cleavage of the resulting PCR products with *Bgl*II, the *Bgl*II-resistant material was subcloned in *E. coli* and a total of 25 *Bgl*II-resistant clones were sequenced. Their junctions are also displayed in Fig. 5. Although the selection for *Bgl*II resistance helps to avoid analyzing false positives possibly arising by transfection of oc contaminants and products resulting from plasmids damaged out-of-target, it must be kept in mind that all events are lost which arose by rejoining of DSB that do not affect the *Bgl*II site. Likewise, all events are lost in which the *Bgl*II site is regenerated by chance by use of microhomology patches present in the repetitive TFO target motif (Fig. 1). This issue was verified by sequencing of 17 *Bgl*II-sensitive clones, and we found, as expected, a high proportion of wild-type sequences (76%) and three clones in which the *Bgl*II site had been regenerated by chance (Fig. 5, *junctions 4* and *10*).

Unlike the spectra obtained from RE substrates, which produced only few different junctions per substrate, the spectrum from the $^{125}$I-TFO substrate appears much more heterogeneous as reflected by a total of 43 different junctions. With the exception of three junctions (junctions 36–38), all junctions have lost one or several (up to 34) bases (larger deletions of up to several hundreds of base pairs also existed but were not further analyzed because of loss of the primer binding site for sequencing). Because we did not detect any major differences between the sequences derived from ccc products and those from dimer products, no further distinction was made between these two product forms.

The total spectrum can be subdivided in three major groups: (i) junctions that are free of microhomology (blunt junctions: junctions 2, 3, 7, 8, 12–17, 19, 21, 22, 24, 26, 27, and 29; note that the term "blunt junction" does not imply that these junctions arose necessarily by blunt end ligation but that they can also arise by the fill-in mechanism mentioned above; Fig. 2*C*); (ii) junctions that display patches of microhomology of 1–4 bp at their breakpoints (microhomology junctions: junctions 4, 5, 6, 9, 10, 11, 18, 20, 23, 25, 28, and 30–35); and (iii) junctions containing single base substitutions or additional (untemplated) bases not present in the original sequence (insertion junctions: junctions 36–43). The heterogeneity of this spectrum is consistent with the expected heterogeneity of DSB present in the $^{125}$I-TFO substrate and possibly a decreased fidelity of the NHEJ reaction of dirty DSB. A detailed interpretation of the junctions will be presented under "Discussion."

## DISCUSSION

The use of the TFO labeled with $^{125}$I-dCTP allowed us to take advantage of the highly localized energy spectrum produced by Auger electron emitter decay to induce site-specific DSB within a limited region of ~20 bp around the single *Bgl*II site of pUC-MDR1. The nature of the process by which Auger emitters decay and the similarity of the biological effects to those of high linear energy transfer radiation suggest that the majority of such DSB should be of a complex type and thus highly mutagenic. As such, Auger emitting radionuclides fulfill the criteria for a mutagenic agent that induces complex DNA lesions including the destructive loss of nucleotides at the damaged site.

Decay of $^{125}$I is a stochastic process in which one decay may produce, for example, 30 Auger electrons, whereas another decay produces only five (39). Therefore, some decays may result in severe damage, *i.e.* multiple SSB, base and sugar lesions, base loss, or even multiple DSB, whereas others may produce only simple SSB or base damage that in turn results in SSB in aqueous solution. The complexity of the $^{125}$I-TFO-induced lesions is reflected by the fact that the efficiency of the cell-free NHEJ reaction is reduced by a factor of about 10 when compared with clean RE-induced DSB, which indicates that only a small proportion of the $^{125}$I-TFO-damaged plasmid is repaired. This proportion could represent the molecules containing the least damage, *i.e.* the "simplest breaks" resembling the ones induced by RE. On the other hand, DSB containing damaged sugar and base moieties would have to be converted into structures accepted by the enzymes involved in NHEJ (*e.g.* DNA ligase IV). We have shown previously that our extracts are capable, although at reduced efficiency, of rejoining other complex DSB that had been induced by bleomycin and contain 3′-phosphoglycolate termini (25). Still, we do not know at present whether the reduced NHEJ efficiency for $^{125}$I-TFO-induced DSB reflects the *in vivo* situation or simply is due to the lack in our extracts of some components necessary to remove damaged DNA moieties prior to NHEJ. To clarify this issue, transfection experiments similar to the ones described previously (40) would have to be performed to compare the joining capacity of $^{125}$I-TFO- and RE-linearized plasmids *in vivo*. It is also worth mentioning that $^{125}$I-TFO-induced DSB reduce ccc product formation to a greater degree than dimer formation in comparison with RE-induced DSB. This may be explained in part by the fact that dimers can exist in three possible orientations where different degrees of homology are available at the termini. The tail-to-tail orientation especially exposes the redundant TFO motif, which offers ample microhomology patches (see below). This does not apply for the RE substrates where the DSB termini are located outside of the TFO motif. However, because of the palindromy, these tail-to-tail and head-to-head products are not accessible to cloning and sequence analysis.

The majority (64%) of the 65 sequences derived from the rejoining of the $^{125}$I-TFO substrate (Fig. 5, *junctions 2–35*) show small patches of sequence homology at the junction, indicating that microhomologies play a role in junction formation. In contrast to the blunt junctions (36%), which have

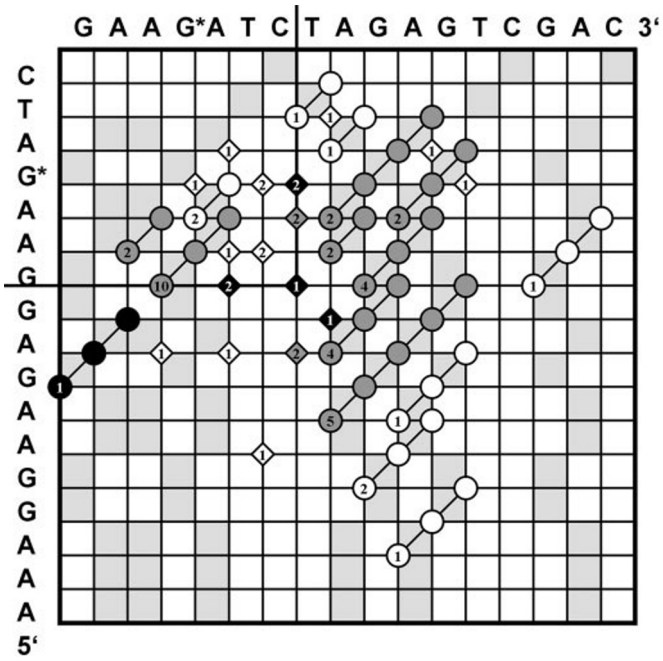


FIG. 6. **Distribution of the breakpoints of *junctions 2–34* (see Fig. 5; because of the large size of its deletion, *junction 35* was not included in this diagram) (41).** The nucleotide sequence of the Pu-rich strand is shown along the axes in 5′-3′-direction from *bottom* to *top* and *left* to *right*. *G** indicates the position of the $^{125}$I in the TFO (note that the motif GAAG*ATC between the *bold lines*, although present only once, is shown in both sequences because this allows inclusion of all junction breakpoints in one diagram). The *grid lines* represent phosphodiester bonds between the bases. The *gray squares* mark base homologies between the *vertical* and *horizontal strands*. The junctions are indicated by *diamonds* or *circles* containing *numerals* that indicate the number of junctions found for this particular sequence. *Open symbols* represent junctions derived from ccc products, *black symbols* represent junctions derived from dimers, and *gray symbols* represent junctions derived from both ccc and dimer products. Blunt junctions are drawn as *diamonds* at intersections of the *grid lines*. Their nucleotide sequences can be determined by reading the *vertical strand* from *bottom* to *top* until the *horizontal line* indicated by the *diamond* is reached, then following the *vertical line* from the *diamond* to the corresponding nucleotide in the *horizontal strand*, and finally continuing to the *right* with the sequence of the *horizontal strand*. The junctions containing patches of microhomology at their breakpoints are denoted by *circles*. Because the homology makes it impossible to determine the precise location of the junction breakpoint, each *circle* in a row connected by a *diagonal* represents a possible breakpoint within the microhomology patch. For example, the sequence of junction 9 (see Fig. 5), which contains a 2-bp homology (. . . . GAGG*ATCT . . .) is given in the diagram by the *diagonal row* of *three gray circles* marked by the *number 10* (*left half* in the *top quarter*). Note that the homology leads to a 2-bp ambiguity with respect to the position of the breakpoint, which can lie between G and G*, G* and A, or A and T.

always precisely defined breakpoints, the breakpoints of microhomology junctions are ambiguous because it is unknown from which of the two DSB ends a nucleotide of the homology originated and where exactly the breakpoint is located within the homology patch. This feature is the hallmark of all microhomology junctions and becomes clearer in Fig. 6 where the 64 junction sequences 2–34 (Fig. 5) are displayed in a two-dimensional diagram as blunt junctions (Fig. 6, *diamonds*) and microhomology junctions (Fig. 6, *circles*), respectively (41, 42). A comparison of the distribution of chance homologies between the vertical and horizontal strand (*gray squares*) and the distribution of the two junction types shows that 65% of the blunt junctions accumulate within a region that is free of microhomology patches (see TCT 3′ of the G*) but only 35% occur in regions exhibiting microhomology. The resulting over-representation of microhomology junctions *versus* blunt junctions in regions containing microhomology indicates that the NHEJ process prefers the use of small homologies whenever available. In our case, especially the highly redundant AG and GA motifs of the TFO-binding site (Fig. 6, see *vertical strand*) and the adjacent *Bgl*II and *Xba*I site (Fig. 6, see AG*A and AGAG in the *horizontal strand*) contribute to the high proportion of microhomology junctions.

The importance of microhomologies in the process of junction formation is further underscored by the fact that the observed frequency of a microhomology exceeds the expected probability of this microhomology to occur by chance at a breakpoint in a DNA duplex of unbiased sequence composition (Fig. 7) (41). Interestingly, the observed numbers of breakpoints that coincide with a microhomology increase with the increasing size of the microhomology, which is inversely proportional to the expected values. This result strongly indicates that microhomologies are important for the process of junction formation from radiation-induced DSB.

The small group of insertion junctions (11% of the total of 72 junctions) comprises sequences containing base pair substitutions or additional untemplated bases not present in the original sequence. The base pair substitutions (Fig. 5, *junctions 37–39*) are not necessarily linked to a DSB rejoining event but could be explained by the repair of single bases that have been damaged by radiation. On the other hand, insertion of untemplated nucleotides is often observed at junctions (5, 17, 43). The insertion of one or a few nucleotides (Fig. 5, *junctions 40*, *42*, and *43*) can be explained by the action of the DNA polymerase that fills gaps in the junctions and sometimes adds single nucleotides to the 3′-hydroxyl of a DSB end (see also Fig. 4, *IIb* and *IIIa*) (44–47). The addition of a longer stretch of nucleotides (Fig. 5, *junction 41*) could also be the result of polymerase action or alternatively reflect the capture of an oligonucleotide (48) possibly originating from residual fragments of mitochondrial or nuclear DNA still present in our whole cell extract preparations (49).

Although the $^{125}$I-TFO-induced DSB are defined with respect to their location within a 19-bp region of the target sequence, the analysis of the underlying joining mechanisms is still complicated by the fact that the structure of the ultimate DSB participating in the formation of a particular junction is unknown. In principle, a DSB can result from two SSB that are located precisely opposite of each other (blunt) or are separated by one or several bases (5′- or 3′-staggered). Only closely spaced SSB (<10bp) in opposite strands are likely to give rise to DSB because of the expected high stability of the intervening duplex (50). As mentioned above, the ends of the linear $^{125}$I-TFO substrate represent a mixture of blunt and staggered DSB at all possible positions. The probability of a certain type of DSB to occur at a given sequence position can be calculated by multiplication of the probabilities of the corresponding SSB to occur at the corresponding bases, which had been determined previously (for details see Fig. 1 and "Experimental Procedures") (33). Thus, the probability of a blunt DSB occuring at a certain position is given by multiplication of the probability of the SSB at this base in one strand with the probability of the SSB at the corresponding base in the opposite strand. Likewise, the probability of a staggered DSB is given by multiplication of the probability of the SSB at a particular base in one strand with the probability of the SSB at any other base in the opposite strand. The sum of all these probabilities reflects the probability of this particular base to be found at an end in any type of DSB.

The distribution of the *gray bars* in Fig. 8*A* shows a fairly symmetrical distribution of DSB around the decay site with the maximum at the central G. If the process of junction formation were a random process and solely determined by the distribu-

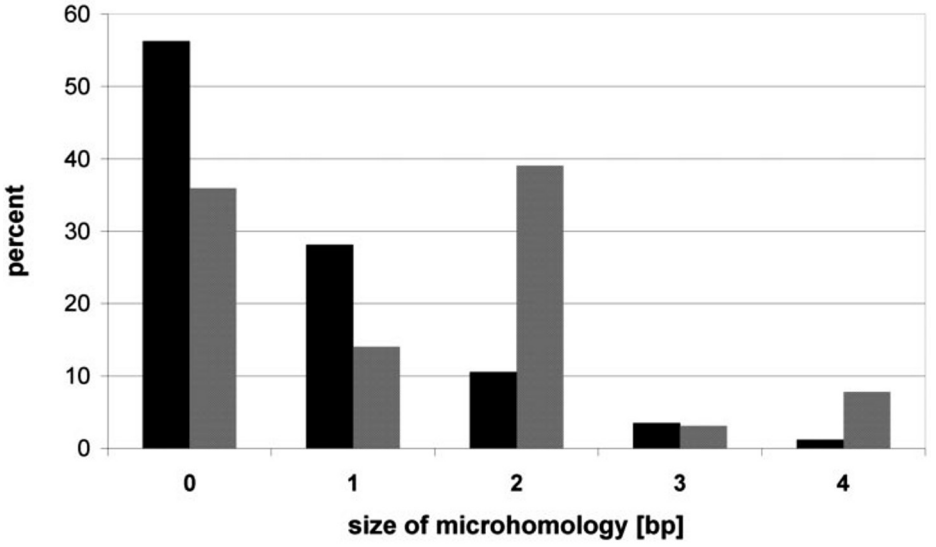


FIG. 7. **Analysis of the frequency of microhomologies at junctions.** The expected probability of a homology of $x$ nucleotides homology to occur by chance at a breakpoint in a DNA duplex of unbiased sequence composition is given by the equation $P(x) = (x + 1)(1/4)^x(3/4)^2$ with $(x + 1)$ being the number of different ways that chance identities could yield the specified homology, $(1/4)^x$ being the probability that $x$ nucleotides match, and $(3/4)^2$ being the probability that nucleotides flanking the matching nucleotides do not match (42). This allows one to calculate the percentage of breakpoints expected (*black bars*) to be located within a given microhomology and compare them with the observed numbers (*gray bars*) as derived from the 64 sequences shown in Fig. 5.

tion of the $^{125}$I-TFO-induced DSB, the distribution of junctions resulting from this substrate should resemble the distribution of DSB. As seen by the *black bars* in Fig. 8*A* and confirmed in a $\chi^2$ test (for details see "Experimental Procedures"), the distribution of junctions is significantly different from the expected distribution. Therefore, other parameters, like the availability of microhomologies (*e.g.* over-representation of junctions at the AGAG motifs left and right of the central G) and the chemical complexity of the original DSB, are likely to contribute to the process of junction formation.

In contrast, the distribution of deleted nucleotides follows nearly precisely the distribution of breaks (Fig. 8*B*). Almost all junctions (with the exception of *junctions 36–38* in Fig. 5) have lost one or several bases. As seen in the figure, bases are most frequently lost around the central G, the site of most efficient DSB induction and thus parallels directly the distribution of DSB. It remains, however, unclear whether this loss of bases is the direct result of the original DSB lesion, which was possibly accompanied by the loss of one or several bases, the result of the NHEJ reaction, which had to remove damaged bases to provide structures that can be processed by DNA-modifying enzymes, or the result of both.

In addition to the simple blunt or staggered DSB discussed so far, the possibility of DSB that comprise multiple SSB in one or both strands and thus are effectively accompanied by the deletion of several bases has to be considered as well. Such lesions can be regarded as double-stranded gaps and therefore have the potential to create larger deletions. As is seen in the spectra derived from RE substrates, deletions are rarely formed by NHEJ in our cell-free system (Fig. 4). If occurring at all, they are mostly small and usually range between 1 and 5 bp. Only 4% of the junctions contain larger deletions (6–55 bp), indicating that the NHEJ process tends to preserve the sequence information at the DSB without extensive nucleotide loss. Fig. 8*C* shows that 50% of the $^{125}$I-TFO deletions observed are small, too, and range between 1 and 5 bp with a pronounced maximum at 3 bp. The other 50% range between 6 and 18 bp. This fraction is considerably bigger than the corresponding fraction from the RE junctions. Therefore, it cannot be excluded that a significant fraction of the $^{125}$I-TFO substrate molecules contain double-stranded gaps that subsequently result in the observed high fraction of larger deletions.

In conclusion, we have established an *in vitro* system that allows us to investigate the repair of a single radiodamaged site on a sequence level. With respect to the presence of deletions, base pair substitutions, and insertions, the spectrum of junc-

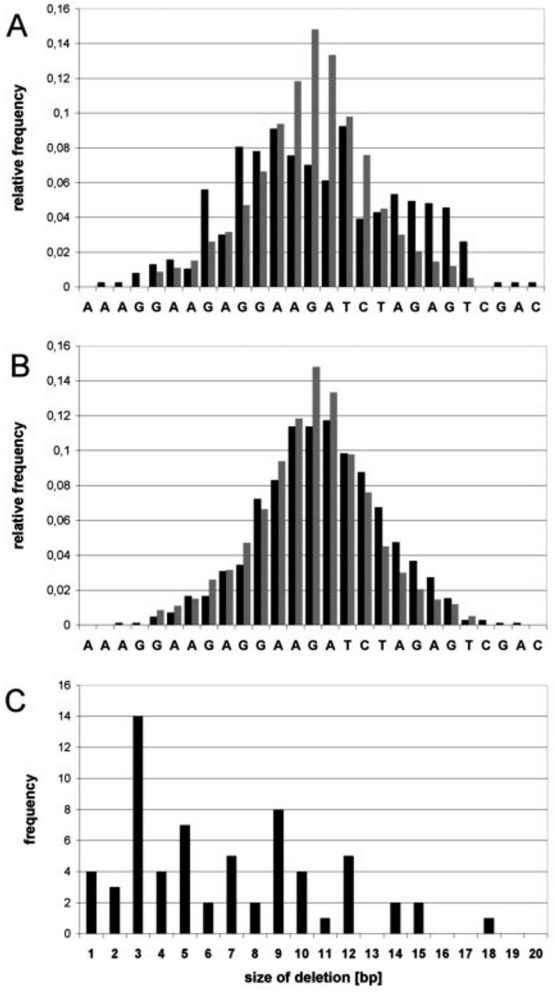


FIG. 8. **Analysis of the distribution of junction breakpoints and delections.** *A*, distribution of junction breakpoints (*black bars*) *versus* distribution of breaks (*gray bars*) around the decay site. Each *black bar* represents the sum of the relative frequencies of all breakpoints occurring at a particular nucleotide of the target sequence (for details see "Experimental Procedures"). *B*, distribution of nucleotides deleted around the decay site (*black bars*) *versus* distribution of breaks (*gray bars* are the same as in *A*). Each *black bar* represents the relative frequency how often a particular nucleotide was deleted from the target sequence (for details see "Experimental Procedures"). *C*, frequencies of the different deletion sizes (in bp) in the 64 junctions (Fig. 5: junctions *2–34*).

tions described here resembles closely the one obtained previously *in vivo* by transfection of a $^{125}$I-TFO-linearized plasmid in mammalian cells (40). This indicates that our *in vitro* system yields reliable results. In future experiments, it will be interesting to dissect the contributions of the different DSB repair mechanisms by using cell-free extracts from mutant CHO cell lines with defined defects in these pathways.

*Acknowledgment*—We thank George Poy for help with sequencing of the RE junctions.

REFERENCES

1. Haber, J. E. (2000) *Trends Genet.* **16,** 259–264
2. Pfeiffer, P., Goedecke, W., and Obe, G. (2000) *Mutagenesis* **15,** 289–302
3. Haber, J. E. (1999) *Trends Biochem. Sci.* **24,** 271–275
4. Haber, J. E. (2000) *Curr. Opin. Cell Biol.* **12,**286–292
5. Feldmann, E., Schmiemann, V., Goedecke, W., Reichenberger, S., and Pfeiffer, P. (2000) *Nucleic Acids Res.* **28,** 2585–2596
6. Critchlow, S. E., and Jackson, S. P.(1998) *Trends Biochem. Sci.* **23,** 394–398
7. Featherstone, C., and Jackson, S. P. (1999) *Curr. Biol.* **9**, R759–R761
8. Bryant, P. E. (1985) *Int. J. Radiat. Biol. Relat Stud. Phys. Chem. Med.* **48,** 55–60
9. Bryant, P. E. (1984) *Int. J. Radiat. Biol. Relat Stud. Phys. Chem. Med.* **46,** 57–65
10. Natarajan, A. T., and Obe, G. (1984) *Chromosoma* **90,** 120–127
11. Obe, G., Palitti, F., Tanzarella, C., Degrassi, F., and De Salvia, R. (1985) *Mutat. Res.* **150,** 359–368
12. Rouet, P., Smih, F., and Jasin, M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 6064–6068
13. Rouet, P., Smih, F., and Jasin, M. (1994) *Mol. Cell. Biol.* **14,** 8096–8106
14. Roth, D. B., and Wilson, J. H. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 3355–3359
15. Lutze, L. H., Cleaver, J. E., Morgan, W. F., and Winegar, R. A. (1993) *Mutat. Res.* **299,** 225–232
16. King, J. S., Valcarcel, E. R., Rufer, J. T., Phillips, J. W., and Morgan, W. F. (1993) *Nucleic Acids Res.* **21,** 1055–1059
17. Daza, P., Reichenberger, S., Göttlich, B., Hagmann, M., Feldmann, E., and Pfeiffer, P. (1996) *Biol. Chem. Hoppe-Seyler* **377,** 775–786
18. Pfeiffer, P., and Vielmetter, W. (1988) *Nucleic Acids Res.* **16,** 907–924
19. North, P., Ganesh, A., and Thacker, J. (1990) *Nucleic Acids Res.* **18,** 6205–6210
20. Nicolas, A. L., Munz, P. L., and Young, C. S. (1995) *Nucleic Acids Res.* **23,** 1036–1043
21. Baumann, P., and West, S. C. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 14066–14070
22. Labhart, P. (1999) *Eur. J. Biochem.* **265,** 849–861
23. Phillips, J. W., and Morgan, W. F. (1994) *Mol. Cell. Biol.* **14,** 5794–5803
24. Beyert, N., Reichenberger, S., Peters, M., Hartung, M., Gottlich, B., Goedecke, W., Vielmetter, W., and Pfeiffer, P. (1994) *Nucleic Acids Res.* **22,** 1643–1650
25. Chen, S., Inamdar, K. V., Pfeiffer, P., Feldmann, E., Hannah, M. F., Yu, Y., Lee, J. W., Zhou, T., Lees-Miller, S. P., and Povirk, L. F. (2001) *J. Biol. Chem.* **276,** 24323–24330
26. Bennett, R. A., Gu, X. Y., and Povirk, L. F. (1996) *Int. J. Radiat. Biol.* **70,** 623–636
27. Gu, X. Y., Bennett, R. A., and Povirk, L. F. (1996) *J. Biol. Chem.* **271,** 19660–19663
28. Panyutin, I. G., Winters, T. A., Feinendegen, L. E., and Neumann, R. D. (2000) *Q. J. Nucl. Med.* **44,** 256–267
29. Sedelnikova, O. A., Luu, A. N., Karamychev, V. N., Panyutin, I. G., and Neumann, R. D. (2001) *Int. J. Radiat. Oncol. Biol. Phys.* **49,** 391–396
30. Panyutin, I. G., and Neumann, R. D. (1997) *Nucleic Acids Res.* **25,** 883–887
31. Panyutin, I. G., and Neumann, R. D. (1996) *Acta Oncol.* **35,** 817–823
32. Panyutin, I. G., and Neumann, R. D. (1994) *Nucleic Acids Res.* **22,** 4979–4982
33. Panyutin, I. V., Luu, A. N., Panyutin, I. G., and Neumann, R. D. (2001) *Radiat. Res.* **156,** 158–166
34. Sedelnikova, O. A., Panyutin, I. G., Luu, A. N., Reed, M. W., Licht, T., Gottesman, M. M., and Neumann, R. D. (2000) *Antisense Nucleic Acid Drug Dev.* **10,** 443–452
35. Thode, S., Schäfer, A., Pfeiffer, P., and Vielmetter, W. (1990) *Cell* **60,** 921–928
36. Pfeiffer, P. (1998) *Toxicol. Lett.* **96–97,**119–129
37. Pfeiffer, P., Thode, S., Hancke, J., Keohavong, P., and Thilly, W. G. (1994) *Mutagenesis* **9,** 527–535
38. Pfeiffer, P., Thode, S., Hancke, J., and Vielmetter, W. (1994) *Mol. Cell. Biol.* **14,** 888–895
39. Nikjoo, H., Laughton, C. A., Terrissol, M., Panyutin, I. G., and Goodhead, D. T. (2000) *Int. J. Radiat. Biol.* **76,** 1607–1615
40. Mezhevaya, K., Winters, T. A., and Neumann, R. D. (1999) *Nucleic Acids Res.* **27,** 4282–4290
41. Roth, D. B., and Wilson, J. H. (1986) *Mol. Cell. Biol.* **6,** 4295–4304
42. Roth, D. B., Porter, T. N., and Wilson, J. H. (1985) *Mol. Cell. Biol.* **5,** 2599–2607
43. Roth, D. B., Chang, X. B., and Wilson, J. H. (1989) *Mol. Cell. Biol.* **9,** 3049–3057
44. King, J. S., Fairley, C. F., and Morgan, W. F. (1994) *J. Biol. Chem.* **269,** 13061–13064
45. King, J. S., Fairley, C. F., and Morgan, W. F. (1996) *J. Biol. Chem.* **271,** 20450–20457
46. Clark, J. M. (1988) *Nucleic Acids Res.* **16,** 9677–9686
47. Reichenberger, S., and Pfeiffer, P. (1998) *Eur. J. Biochem.* **251,** 81–90
48. Roth, D. B., Proctor, G. N., Stewart, L. K., and Wilson, J. H. (1991) *Nucleic Acids Res.* **19,** 7201–7205
49. Lin, Y., and Waldman, A. S. (2001) *Genetics* **158,** 1665–1674
50. Tijssen, P. (1993) in *Laboratory Techniques in Biochemistry and Molecular Biology* (van der Vliet, P. C., ed) Elsevier Science Publishers B.V., Amsterdam