

Assays for DNA double-strand break repair by microhomology-based end-joining repair mechanisms

Kaja Kostyrko and Nicolas Mermod*

Institute of Biotechnology, University of Lausanne, and Center for Biotechnology UNIL-EPFL, Lausanne, Switzerland

Received May 27, 2015; Revised November 17, 2015; Accepted November 19, 2015

ABSTRACT

DNA double stranded breaks (DSBs) are one of the most deleterious types of DNA lesions. The main pathways responsible for repairing these breaks in eukaryotic cells are homologous recombination (HR) and non-homologous end-joining (NHEJ). However, a third group of still poorly characterized DSB repair pathways, collectively termed microhomology-mediated end-joining (MMEJ), relies on short homologies for the end-joining process. Here, we constructed GFP reporter assays to characterize and distinguish MMEJ variant pathways, namely the simple MMEJ and the DNA synthesis-dependent (SD)-MMEJ mechanisms. Transfection of these assay vectors in Chinese hamster ovary (CHO) cells and characterization of the repaired DNA sequences indicated that while simple MMEJ is able to mediate relatively efficient DSB repair if longer microhomologies are present, the majority of DSBs were repaired using the highly error-prone SD-MMEJ pathway. To validate the involvement of DNA synthesis in the repair process, siRNA knock-down of different genes proposed to play a role in MMEJ were performed, revealing that the knock-down of DNA polymerase θ inhibited DNA end resection and repair through simple MMEJ, thus favoring the other repair pathway. Overall, we conclude that this approach provides a convenient assay to study MMEJ-related DNA repair pathways.

INTRODUCTION

During their lifetime cells constantly face DNA damage that may result from the by-products of normal metabolic processes, or from exogenous factors such as chemical agents or ionizing radiation (IR). One of the most serious types of DNA damage are double-stranded breaks (DSBs), which can block replication and transcription, or lead to the loss of chromosome fragments. Eukaryotic cells thus possess various mechanisms that sense and repair DSBs. The two major mechanisms responsible for DSB repair in

normal cells consist of the non-homologous end-joining (NHEJ) and homologous recombination (HR) pathways (1–3). These two mechanisms can compete for broken DNA ends in the cell, and the choice between them is made depending on the type of the DSB and the phase of the cell cycle. NHEJ, the main pathway used in higher eukaryotes, is active throughout the cell cycle. It is a fast process, which very efficiently repairs easily ligatable DSBs (4). In contrast, HR is a much more complex mechanism, active mainly in the late S and G2 phase of the cell cycle. HR is considered to be a relatively error-free pathway, as it can repair very complicated DSBs with high fidelity. However, it requires extensive DNA end processing and a homologous DNA molecule as a template, hence its more prominent use following chromosome replication.

In recent years it became apparent that a third mechanism of DSB repair also exists in eukaryotic cells (5,6). This pathway, which may be masked by the main repair processes in normal cells, was described under various names, including alternative end-joining (alt-EJ), alternative or backup NHEJ (alt-NHEJ, a-NHEJ, B-NHEJ), or microhomology-mediated end-joining (MMEJ) (5–11). It is still unclear whether these alternate repair mechanisms consist of one or several distinct pathways. A common feature of these end-joining mechanisms is the use of short 2–25 nt homologies to join broken DNA ends.

Many fundamental findings on the functioning of the DSB repair pathways have been made using *in vivo* plasmid end-joining assays. These assays are most commonly based on the reconstitution of a functional reporter gene by one of the DSB repair mechanisms after the induction of a break in a non-functional substrate. Many such assays were constructed to investigate NHEJ and HR in various types of cells (4,12–19), and recently also MMEJ (20–22). However, with a growing number of new studies proposing the occurrence of mechanistically distinct MMEJ pathways, the need arises for more specific assays that may allow to better distinguish these mechanisms.

Here, we sought to design an assay to measure synthesis-dependent (SD)-MMEJ, a recently proposed variant of MMEJ (23). Both MMEJ and SD-MMEJ may start with a 5'–3' end resection (21,23), similarly to HR, but diverge in later steps. Repair may be carried out by MMEJ if the re-

*To whom correspondence should be addressed. Tel: +41 21 693 61 51; Fax: +41 21 693 76 10; Email: nicolas.mermod@unil.ch

sulting ssDNA overhangs contain short regions of homology, which can pair together to mediate alignment of the two sides of the DSB. However, if the resection fails to expose any microhomologies, the MMEJ machinery may be unable to rejoin the two ends. Moreover, the presence of long ssDNA overhangs precludes the use of the NHEJ pathway, which cannot process these types of substrates (24,25). The SD-MMEJ model offers a solution in such situations. In this mechanism, a non-processive DNA polymerase may copy a sequence from the DNA up- or downstream of the break, which subsequently serves to align the two sides of the break enabling the continuation of the MMEJ pathway. The SD-MMEJ model offers a solution in such situations. In this mechanism, a non-processive DNA polymerase may copy a sequence from up- or downstream of the break, which subsequently serves as a microhomology used to align the two broken DNA ends enabling the continuation of the MMEJ pathway (23). This can be accompanied by the amplification of a sequence near the DSB and in its insertion between the joined DNA extremities, which is termed a templated insertion.

In this study, we attempted to construct a specific SD-MMEJ assay and compared it with a previously published reporter designed to measure simple MMEJ, where the rejoining of a functional Green Fluorescent protein (GFP) sequence by either mechanism can be followed by flow cytometry (22). The use of these assays in Chinese Hamster Ovary (CHO) cells and the analysis of junction DNA sequences revealed that DNA repair resulted from both types of mechanisms, although SD-MMEJ was more frequently used. This suggested that SD-MMEJ, while error-prone, is a very robust mechanism able to repair incompatible DSBs without any need for pre-existing homology. The MMEJ assay was validated by the siRNA knock-down of genes involved in alternative end-joining pathways, which indicated that the depletion of polymerase θ decreased the efficiency of simple MMEJ in favor of NHEJ and SD-MMEJ, suggesting that this enzyme is important for the former pathway. Overall, we conclude that this assay can be used to decipher the molecular mechanisms of multiple MMEJ-related DSB repair pathways.

MATERIALS AND METHODS

Cell culture

Adherent CHO DG44 cells (26) were cultivated in DMEM/F12+GlutaMAXTM supplemented with 1x HT and 10% fetal bovine serum (Gibco, Invitrogen), and with the antibiotic-antimycotic solution (Sigma-Aldrich, #A5955).

Construction of the recombination assays

The MMEJ and SD-MMEJ reporter cassettes were constructed by inserting restriction sites in the GFP coding sequence of the pSV40-GFP plasmid described previously (27). The MMEJ vector was based on a previously described reporter (22). Briefly, a naturally occurring 9-bp sequence (CGCGCCGAG) was duplicated and an 18-bp I-SceI recognition site was inserted in between the two copies of the sequence (Figures 1A and 2A). Two in frame stop

codons present in the inserted sequence prevent the expression of a functional GFP from the intact vector. Digestion with I-SceI linearizes the vector and creates a DSB with 3' overhangs.

The two SD-MMEJ reporter cassettes were designed using microhomologies already present in the GFP sequence. In the first assay (SD-MMEJ-1), the two 5-bp microhomologies (CGAGG) are 7 bp apart (Figures 1B and 2B). In the second assay (SD-MMEJ-2) the two 7-bp microhomologies (CCACCCT) are 5 bp apart (Supplementary Figure S1). To enable the formation of DSBs with 5' incompatible overhangs and prevent re-ligation upon digestion, two restriction sites (separated by 3 bp) were introduced into each vector inside one of the microhomologies. SpeI and AflII recognition sites were used in SD-MMEJ-1, and AflII and EcoRI in SD-MMEJ-2. In both cases, in-frame stop codons are present inside the restriction sites to prevent GFP expression from the intact vectors.

Transfection and FACS analysis

MMEJ and SD-MMEJ plasmids were digested with the appropriate restriction enzymes (New England Biolabs) for 5 h and purified by ethanol precipitation. Aliquots were analyzed by gel electrophoresis to confirm complete digestion. A total of 2×10^5 CHO cells were seeded into each well of a 12 well plate (TPP) in 1 ml of complete medium. On the following day each well was co-transfected with 900 ng of linearized GFP reporter plasmid and with 100 ng of undigested pGL3-CMV-dsRed construct to normalize for transfection efficiency using Fugene 6, according to manufacturer's instructions (Promega). The pSV40-GFP vector (pGFP) was transfected in parallel as a positive control of GFP expression. Expression of GFP and dsRed was monitored by fluorescence microscopy (Carl Zeiss Microscope Axio Observer.A1) and flow cytometry. For flow cytometry, cells were harvested 24 h following transfection and resuspended in 0.5 ml of PBS with 2% FBS (Gibco, Invitrogen). Data were acquired using the CyAn analyzer (Beckman Coulter) and analyzed using the FlowJo software (Tree Star). GFP repair efficiency was calculated as the ratio of GFP-positive cells over the number of dsRed-positive cells.

Junction sequence analysis

For the analysis of junction sequences, CHO cells were transfected as before with the MMEJ or SD-MMEJ vectors, but without the pGL3-CMV-dsRed plasmid. After 24 h, cells were harvested and GFP-positive cells were sorted by FACS (MoFlo Astrios Cell Sorter, Beckman Coulter), to enrich the cell population in transfected cells. Total DNA was isolated from the sorted cells using the DNeasy Blood & Tissue Kit (Qiagen). Next, one of the two following methods was used to obtain the sequences of the repaired vectors. In the first method the repaired GFP sequences were amplified by PCR using primers GFP-NcoI-F1 (ATTCCGGTACTGTTGGTAAAGCCACCA) and GFPp2-Rev (GTATCTTATCATGTCTGCT). The PCR products were cleaved with NcoI and XbaI and ligated into the NcoI and XbaI-cleaved pSV40 vector prior to transformation into recombination deficient *Escherichia coli* cells (XL10-Gold UI-

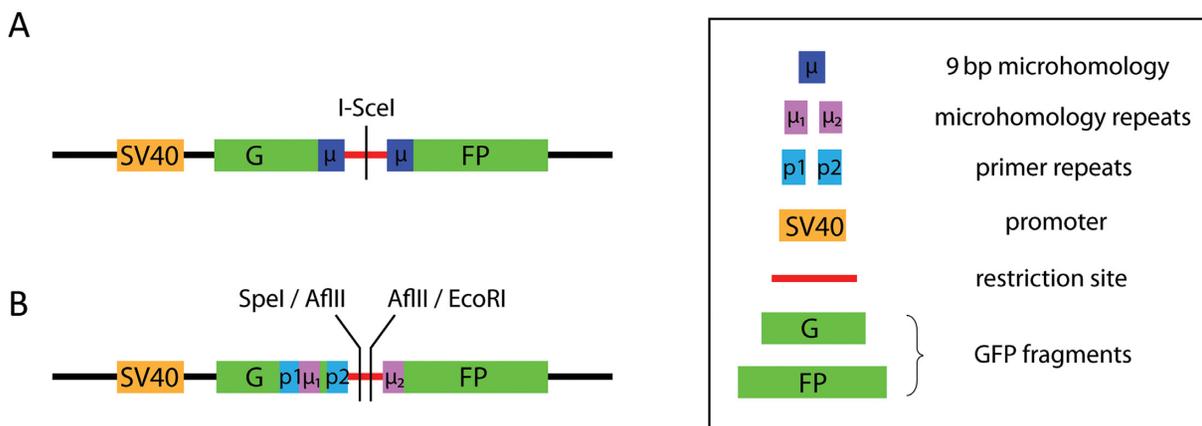


Figure 1. GFP-reconstitution reporter cassettes aimed at detecting the regular microhomology-mediated end-joining (MMEJ) and the synthesis-dependent MMEJ (SD-MMEJ) DNA repair pathways. (A) MMEJ reporter. (B) SD-MMEJ reporter. Description and abbreviations are listed in the text.

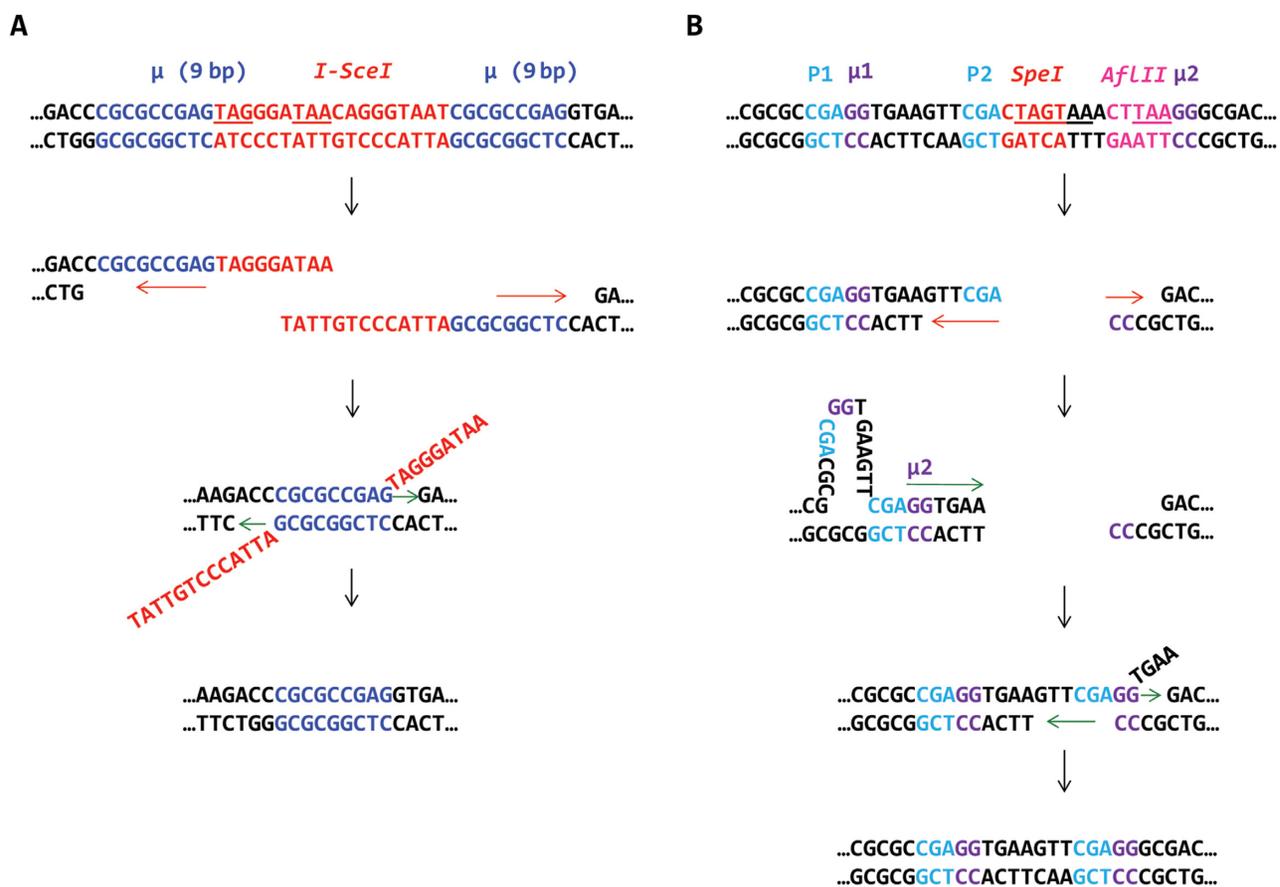


Figure 2. Anticipated mechanism of repair of the MMEJ and SD-MMEJ-1 reporters. Stop codons are underlined. Red arrows indicate the direction of the 5'-3' resection. Green arrows indicate the direction of DNA synthesis. (A) MMEJ reporter. Adapted from (22). (B) SD-MMEJ-1 reporter. P1/P2 stand for primer repeats, and μ_1/μ_2 for microhomology repeats. The SD-MMEJ repair mechanism is adapted from (23).

tracompetent Cells, Stratagene) and plating on ampicillin-containing Luria Broth plates. Colonies were picked and analyzed by colony PCR amplification for the presence of the insert using primers GFP-NcoI-F1 (ATTCCGGTAC TGTTGGTAAAGCCACCA) and SV40.lateR (TCCAAA CTCATCAATGTATC). Plasmids isolated from positive clones were analyzed by Sanger sequencing. In the second

method, recombination deficient *E. coli* cells (XL10-Gold Ultracompetent Cells, Stratagene) were transformed with 6–12 μ l of total DNA and plated on ampicillin-containing Luria Broth plates. Plasmid DNA was isolated from individual colonies and analyzed by Sanger sequencing. Half of all the sequenced junctions were obtained using the first method and another half using the second method. The

frequencies of particular junction patterns was not altered when comparing the data obtained by either method.

siRNA and transfections

Small interfering RNA duplexes were specifically designed to target the Chinese hamster homologs of DNA-PKcs, Ku70, Ligase I, Ligase III, polymerase θ and Pold3. The siRNAs were designed and provided by Microsynth AG (Balgach, Switzerland). Three RNA duplexes were designed per mRNA to increase the probability of successful knock-down. Three negative (non-targeting) siRNAs were also designed as controls. For siRNA-mediated knock-down, 1.5×10^5 CHO-DG44 cells were seeded into each well of a 12 well plate (TPP) in 1 ml of complete medium, and transfected with equimolar amounts of three targeting or non-targeting siRNA duplexes at a final concentration of 50 nM using Lipofectamine RNAiMAX (Invitrogen), with the reverse transfection protocol, according to the manufacturer's instructions. After two days, the siRNA-treated cells were re-transfected with 1 μ g of the digested MMEJ assay vector using Fugene 6 (Promega). GFP-positive cells were sorted and the junction sequences were amplified and sequenced as described above.

RESULTS

CHO cells restore GFP expression more efficiently from the MMEJ than the SD-MMEJ reporters

To measure the efficiency of extrachromosomal MMEJ and SD-MMEJ in mammalian cells, we constructed GFP reporter assays aimed at specifically detecting these pathways. The MMEJ repair substrate contains two 9-bp microhomologies flanking the I-SceI restriction site (Figures 1A and 2A). After I-SceI digestion and end-resection, annealing at these microhomologies should enable the restoration of a functional GFP coding sequence (Figure 2A). Two versions of the SD-MMEJ reporter were constructed, differing by the size of the microhomologies and by their distance. They contain a GFP sequence interrupted by two tandem restriction sites, which serve to create a DSB with non-complementary 5' overhangs (Figures 1B and 2B, Supplementary Figure S1). Since there are no extended microhomologies in the sequence surrounding the break, microhomologous sequences should be amplified from another fragment of the vector by a DNA polymerase, as illustrated by the examples of a SD-MMEJ mechanism shown in Figure 2B and Supplementary Figure S1. In this mechanism, the microhomology (μ 1) sequence following P1 will be copied to follow P2, acting as a primer, to provide on each side of the DSB the two microhomologous μ 1 and μ 2 sequences needed to align the two broken ends. The use of these direct repeats should enable the restoration of a functional GFP sequence. A simple re-ligation of the cleaved DNA ends prevents the expression of a functional GFP, due to the presence of in-frame stop codons, while the use of microhomologies other than the pre-set ones may lead to deletions in the open reading frame and/or frame shifts.

The linearized vectors were transiently transfected into CHO cells, and the appearance of GFP-positive cells was analyzed by flow cytometry. Over 50% of cells transfected

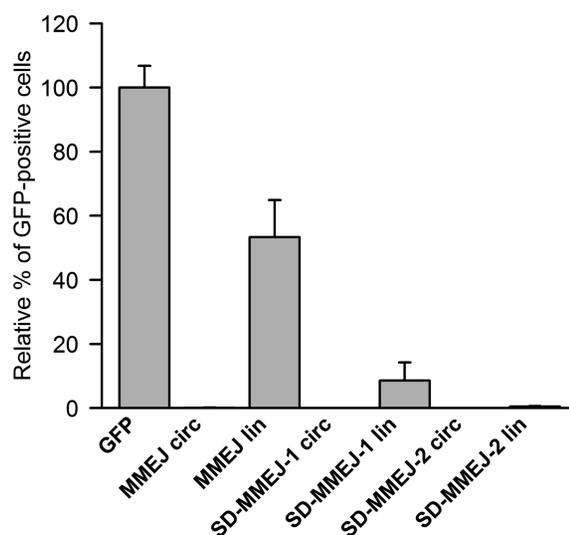


Figure 3. Frequency of GFP coding sequence reconstituted from the transient transfection of linearized MMEJ and SD-MMEJ reporter constructs in CHO cells. Cells were transfected with a control circular GFP expression plasmid (GFP) to assess transfection efficiency, or with the non-digested (circ) or linearized (lin) MMEJ or SD-MMEJ reporters. The percentages of GFP-positive cells are shown with respect to cells transfected with the control plasmid (set to 100%). Mean of three experiments, error bars represent standard error of the mean (SEM) values.

with the MMEJ reporter were able to reconstitute a functional GFP coding sequence (Figure 3). In contrast, only ~9% of the cells transfected with one of the SD-MMEJ assays (SD-MMEJ-1) successfully repaired the GFP gene. The second SD-MMEJ reporter (SD-MMEJ-2) yielded only a few GFP-positive cells (~0.5%). The difference in efficiency between the two SD-MMEJ constructs could result from the choice of the restriction enzymes used to generate the break, and/or from the spacing of elements and variable flanking sequences (Figure 2B and Supplementary Figure S1).

These results suggested that the MMEJ mechanism may more efficiently reconstitute a functional GFP coding sequence than SD-MMEJ pathway in CHO cells, possibly because of the presence of the pre-existing and relatively long (9 bp) microhomologies, which may provide a greater chance of successful GFP reconstitution. In contrast, the lack of long homology in the SD-MMEJ assays may force the repair machinery to copy the microhomology from a more distant region of the plasmid, which may not necessarily lead to the reconstitution of a functional GFP. To assess how the reporter cassettes were repaired, and to unambiguously identify the mechanisms used, we sequenced the re-joined plasmids isolated from the GFP-expressing CHO cells (Supplementary Figure S2).

MMEJ and SD-MMEJ reporters are more frequently repaired by the SD-MMEJ pathway

The sequences of 12 junctions were amplified and cloned from cells transfected with the MMEJ vector, whereas 13 junctions were analyzed from the cells transfected with the SD-MMEJ-1 or SD-MMEJ-2 vector. The DNA sequences

of the junctions were determined and analyzed for the presence of potential microhomologies, deletions and/or primer/microhomology pairs in the vicinity (i.e. within 300 bp) of the junction. The presence of limited sequence loss (0–6 bp) and a lack of microhomology was interpreted as the result of the NHEJ mechanism. Junctions showing evidence of long resections (>6 bp), with at least 2 bp of pre-existing microhomology, were classified as simple MMEJ products. If these short homologous sequences were absent, we searched for the occurrence of ≥ 2 bp of the SD-MMEJ microhomology P sequences followed by ≥ 2 bp of μ sequences, either as direct or inverted repeats up- or downstream of the junction, as well as for the presence of a templated inserts, which are all hallmarks of the SD-MMEJ pathway (23).

Analysis of all the sequenced repair products revealed that the majority (67–75%) of the junctions from cells transfected with either type of reporters must have resulted from the SD-MMEJ mechanism (Table 1, Supplementary Tables S1–S3). The MMEJ mechanism seemed to account for only 25–33% of all the junctions analyzed, despite the presence of the 9 bp microhomology region in the vicinity of the cleavage site for the MMEJ reporter. Interestingly, we did not observe any repair products that could be unambiguously attributed to the NHEJ pathway. These results indicated not only that the SD-MMEJ pathway may be more efficient than simple MMEJ, but also that the appearance of GFP fluorescence in these types of assays should not be used as a sole determinant of repair efficiency by a given mechanism. Consequently, sequencing of the repaired junctions should always be performed to reliably estimate the true efficiency of the end-joining mechanisms in cells.

Out of the 7 MMEJ-like junctions obtained from all assays, 3 occurred at the 9 bp microhomologies (Table 1, Supplementary Tables S1–S3). In the 3 remaining cases, fortuitous 2, 3 and 4 bp microhomologous sequences were used for alignment. This indicated that the MMEJ mechanism may show a preference for relatively longer microhomologies. Strikingly, out of all the sequences obtained from cells transfected with the MMEJ reporter, only one contained a GFP sequence successfully reconstituted using the pre-set 9 bp microhomology (Supplementary Table S1, junction number 4). This implies that precise, error-free repair events relying on cleavage-proximal microhomologous sequences correspond to rarely occurring events.

In contrast, none of the SD-MMEJ constructs was repaired using the designed microhomologies, even though in 8/18 of SD-MMEJ-like junctions, the SD-MMEJ mechanism relied on 5 bp (2 bp + 3 bp) microhomology sequences, as designed in the SD-MMEJ-1 reporter (Table 1, Supplementary Tables S2 and S3). In most cases (11/18), the ‘primer’ sequence was 2 bp long and the amplified microhomology used for bridging the breaks was 2 bp (8/18) or 3 bp (7/18) long. This suggested that even very short sequences can serve as starters for the SD-MMEJ DNA polymerase, making this process very robust in repairing breaks completely devoid of extended microhomologies. This, however, likely entails poor repair fidelity, consistently with the low frequency of successful GFP reconstitution events from the SD-MMEJ reporters.

The repair of the MMEJ cassette was accompanied by slightly larger deletions than those obtained from the SD-MMEJ assays (Table 1, Supplementary Figure S3). This difference likely reflects the difference between the 5′ protruding ends in the SD-MMEJ vectors and the recessive 5′ ends in the MMEJ vector ($\Delta = 16$ bp), suggesting that the extent of the 5′–3′ end resection was similar for all three constructs. We concluded that these reporter constructs can all be used to study both types of repair mechanisms, and that the type of repair mechanism that is used is not predefined by the type of microhomologies that is present at the DNA extremities.

Taken together, these results indicated that, while the simple MMEJ pathway is potentially more precise if larger homologies are present, the SD-MMEJ pathway seems to be much more frequently used in these cells, irrespective of the presence or not of extended microhomologies. We thus hypothesized that the SD-MMEJ mechanism plays the role of a salvage repair pathway when other mechanisms have failed.

Depletion of polymerase θ inhibits DSB repair through MMEJ in favor of other end-joining pathways

We next used this strategy to assess the relative efficiency of MMEJ and SD-MMEJ pathways in cells depleted of factors thought to be implicated in such repair processes. Short interfering RNA (siRNA) were used to knock-down the expression of Ligase I, Ligase III, DNA polymerase θ and DNA polymerase δ subunit 3 (Pold3), all of which were previously proposed to play a role in alternative DSB repair pathways (11,20,28–32). We also used siRNAs against two NHEJ genes—Ku70 and DNA-PKcs, and non-targeting control siRNAs. The decrease in the target mRNA levels upon siRNA treatment was verified by real-time PCR (Supplementary Figure S4). Since the reporter vectors did not significantly differ in the relative frequency of the MMEJ and SD-MMEJ mechanisms, we used only the MMEJ reporter, as it allows for higher recovery of re-joined vectors. Repaired plasmids were obtained from the cells as before, and 20–24 junction sequences were characterized for each condition.

Surprisingly, in cells transfected with the control siRNA, the frequency of MMEJ-attributed repair (~50%) was higher than that of SD-MMEJ (30%) (Figure 4A, Supplementary Table S4), which contrasted with the results obtained from untreated cells. This indicated that the siRNA transfection alone can alter the frequency with which the two pathways are used. We also observed few junction sequences that could only be attributed to the NHEJ mechanism (~20%). Similar repair patterns were observed in cells depleted in DNA-PKcs and Ligase I, implying that these factors may not be essential for the repair of extrachromosomal DSBs. However, it should be noted that the knock-down of DNA-PKcs and Ku70 decreased the frequency of NHEJ. It also increased the number of GFP-positive cells about 2-fold (data not shown), in line with previous observations that alternative end-joining pathways are more active when NHEJ is disabled (8,33–35). Interestingly, the knock-down of polymerase θ significantly decreased the frequency of simple MMEJ to approximately 20%, in favor of

Table 1. The analysis of the sequenced repair products.

Reporter	Repair mechanism Number (% of junctions)		Deletion size ¹ [bp]	Pre-existing microhomology [bp] (number of junctions)	Primer repeat [bp] + microhomology [bp] (number of junctions)
	MMEJ	SD-MMEJ			
MMEJ	3 (25%)	9 (75%)	17, 21, 27, 53, 54, 54, 64, 73, 81, 91, 109, 117	9 (2), 3 (1)	2 + 2 (2), 3 + 2 (2), 2 + 3 (3), 3 + 3 (1), 2 + 4 (1)
SD-MMEJ-1	3 (33%)	6 (67%)	10, 11, 20, 23, 27, 30, 71, 103, 1163	9 (1), 4 (1), 2(1)	2 + 2 (2), 2 + 3 (2), 3 + 3 (1), 4 + 2 (1)
SD-MMEJ-2	1 (25%)	3 (75%)	24, 48, 52, 76	3 (1)	3 + 2 (1), 2 + 4 (1), 3 + 4 (1)

¹Size of the sequence missing from the reporter construct after repair. Deletions are ordered from smallest to largest. A deletion of 27 bp would be expected from the MMEJ mechanism using the preset 9 bp microhomology on its cognate assay plasmid.

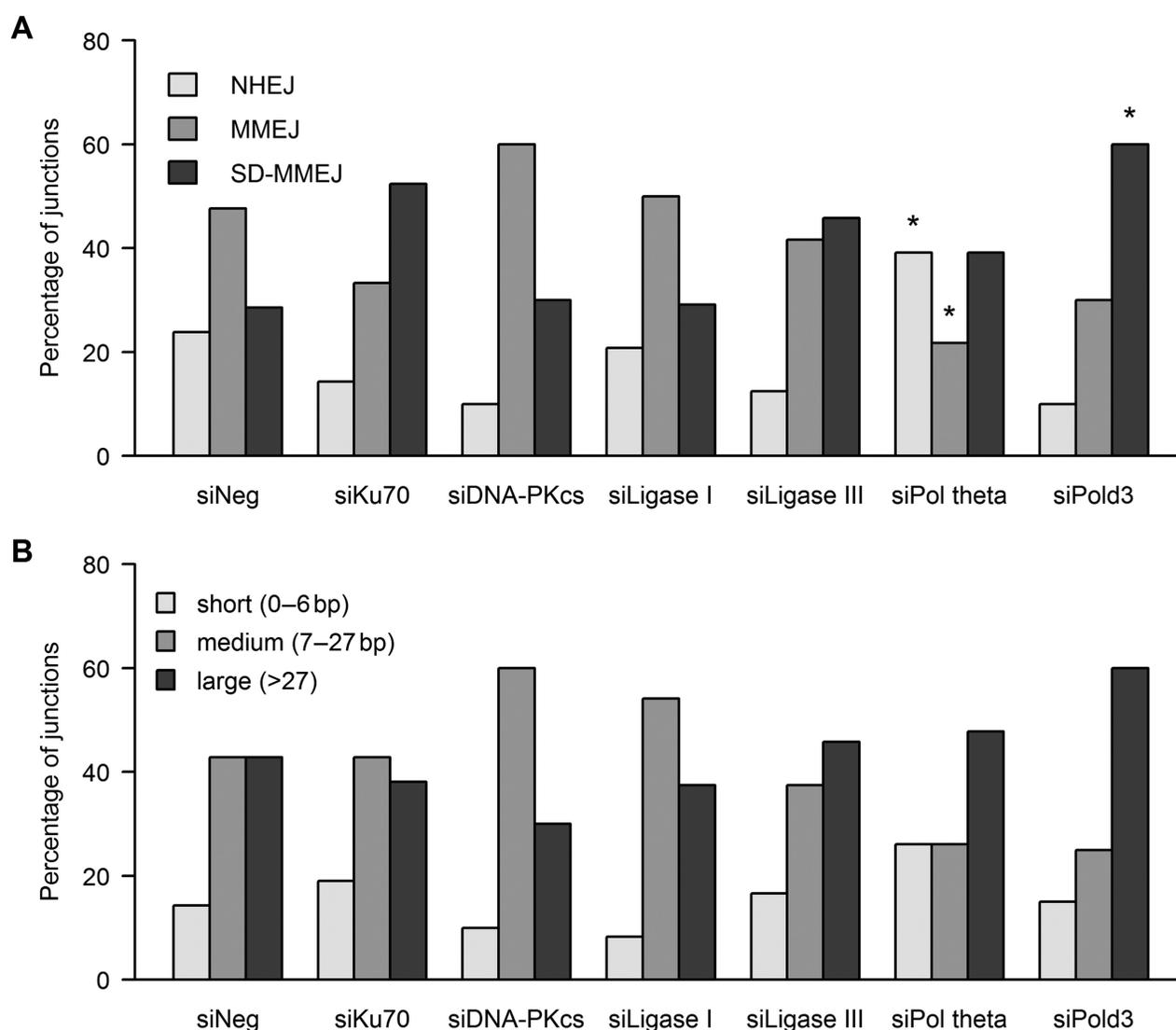


Figure 4. Frequency of DSB repair mechanisms and deletion sizes in cells transfected with MMEJ and SD-MMEJ reporters. (A) The frequency of DSB repair of the MMEJ reporter attributed to the NHEJ, MMEJ and SD-MMEJ mechanisms in cells depleted of NHEJ or MMEJ activities. Statistical significance was calculated using the exact binomial test, and asterisks indicate significant differences between the siRNA treated sample and the control treated with a non-targeting siRNA (siNeg), with a significance level of $P < 0.05$. (B) The sizes of junctional deletions from siRNA-transfected cells were determined as for Figure 3. Deletions in the sequenced MMEJ reporter junctions were classified as short (0–6 bp), medium (7–27 bp) or large (>27 bp).

both NHEJ (to ~40%) and SD-MMEJ (to ~40%), suggesting that it plays an important role in the first pathway. An increase in SD-MMEJ repair was also noted in cells depleted of Pold3 and Ligase III, although in this case at the expense of both MMEJ and NHEJ pathways.

The analysis of deletion sizes in the recovered junctions demonstrated that loss of polymerase θ led to a decrease in medium-sized deletions in favor of very short ones (Figure 4B, Supplementary Figure S5), consistent with the increased usage of the NHEJ pathway. This suggested that the lack of this polymerase may limit repair mechanisms involving medium-size DNA end resection. The proportion of large deletions was decreased in favor of medium-sized ones in the absence of Ligase I, suggesting that the lack of this ligase may also lead to the repair processes that limit resection, but to a lesser extent. Conversely, Pold3 knock-down resulted in an increased number of long deletions, potentially indicating that DNA polymerase δ may inhibit long-range resection. Overall, use of this DSB repair assay indicated that DNA polymerase θ may mediate DSB repair by a MMEJ mechanism involving extensive DNA end resection, whereas the lack of Ligase I or polymerase δ may conversely oppose mechanisms mediating long resections.

DISCUSSION

DSBs are potentially genotoxic DNA lesions, which need to be efficiently repaired by the cells to prevent chromosomal aberrations or serious DNA damage, leading to carcinogenesis or even cell death. The two major pathways responsible for DSB repair in eukaryotes are generally considered to be HR and NHEJ. However, in recent years it became apparent that these mechanisms are assisted by a family of as yet still poorly characterized alternative DSB repair pathways collectively termed MMEJ. These mechanisms are thought to come into play when the main repair pathways are insufficient to repair all the DNA breaks that arise in cells. This is often the case in cancer cells, which suffer from high levels of oxidative stress (36,37). Consistently, many reports have shown that such alternative MMEJ-like DSB repair pathways are more active in tumor cells (38,39).

Studies of these alternative repair mechanisms advanced largely due to the development of plasmid recombination assays, which allow to measure the efficiency of end-joining in the cells. These assays are often based on transiently transfected or genome-integrated reporter substrates, in which DSBs are induced by restriction enzymes. Restoration of reporter gene expression serves to assess the efficiency of DSB repair, which was complemented by the sequencing of the repaired DNA products in some studies. However, a systematic study of various types of MMEJ-related mechanisms and junction sequences by dedicated reporter plasmids was not reported.

Here, we describe various GFP reporter substrates that can be used to study a recently proposed sub-pathway of MMEJ, termed SD-MMEJ (23). This study also describes a MMEJ reporter transient transfection assay that yields high numbers of GFP-positive cells, possibly because of the availability of longer microhomologies. Use of this assay and sequencing of the repaired junctions revealed that the SD-MMEJ mechanism is frequently used by the assessed

tumor cell line, even if less accurate and less likely to yield a functional GFP coding sequence. Thus, we hypothesize that the availability and length of microhomology at the break site does not strictly determine the pathway used for repair. Nevertheless, it should be noted that the vectors used in this study were linearized before transfection and hence prone to the activity of cellular exonucleases. This may have resulted in the removal of DNA sequences near the break site, preventing their use by the intended repair process. This may be less likely to happen in a chromosomal context. In addition, we propose that if the DSB contains blunt or compatible cohesive ends, an apparently rare situation, they may be easily ligated by the NHEJ machinery (Figure 5). However, if the breaks are not compatible, as assessed in this study, extended 5'-3' end resection may create long ssDNA overhangs, which can in turn anneal at pre-existing microhomologies to enable repair by MMEJ. However, even in the presence of microhomology at the break site, the SD-MMEJ mechanism often amplifies alternate microhomologous sequences, as needed for pairing with a compatible single stranded sequence on the other DNA molecule to be rejoined. In most of the observed cases (i.e. 90%), these microhomologies were found less than 100 bp from the break site. However, here we propose that they can also be copied from more distanced regions of the repaired molecule, up to 300 bp, which corresponds to the dynamic persistence length of double stranded DNA, i.e. the minimal length of DNA required to close a loop (40-42). Overall, we conclude that the SD-MMEJ pathway may be a robust repair mechanism that can be preferentially used to repair incompatible DSBs, even when other end-joining pathways might be used, as allowed by the assay conditions used in this study.

Analysis of the frequency of different DNA end-joining pathways was also performed in CHO cells depleted of important NHEJ and MMEJ factors. Out of the siRNAs tested, the knock-down of polymerase θ very significantly affected DNA end-joining in this transient assay. Interestingly, depletion of this enzyme resulted in a decrease of MMEJ repair in favor of NHEJ and, to a lesser extent, SD-MMEJ, indicating that polymerase θ may be required for the former pathway. These results are in line with the proposed role of this DNA polymerase in MMEJ (43-46). While the increase in SD-MMEJ is inconsistent with previous reports assigning polymerase θ to this pathway (23,28), it may be explained by the participation of other DNA polymerases in this process, e.g. translesion synthesis (TLS) or replication enzymes. Indeed, yeast homologs of TLS polymerases η and ζ (Rad30 and Rev3) as well as Pold3 (Pol32) were also reported to play a role in alternative DSB repair (30). Consistently, a recent large scale study of the integration of plasmids in the genome of CHO cells has also implicated the existence of multiple SD-MMEJ pathways relying on distinct DNA polymerases (Kostyrko *et al.*, manuscript in preparation).

In the present work, we describe reporters and assays that may be used to detect repair events mediated by SD-MMEJ, a sub-pathway of MMEJ. Although the construction of a SD-MMEJ-specific reporter proved to be more challenging than initially anticipated, we report that a combination of GFP fluorescence analysis and sequencing can be used to successfully measure the contribution of the different end-

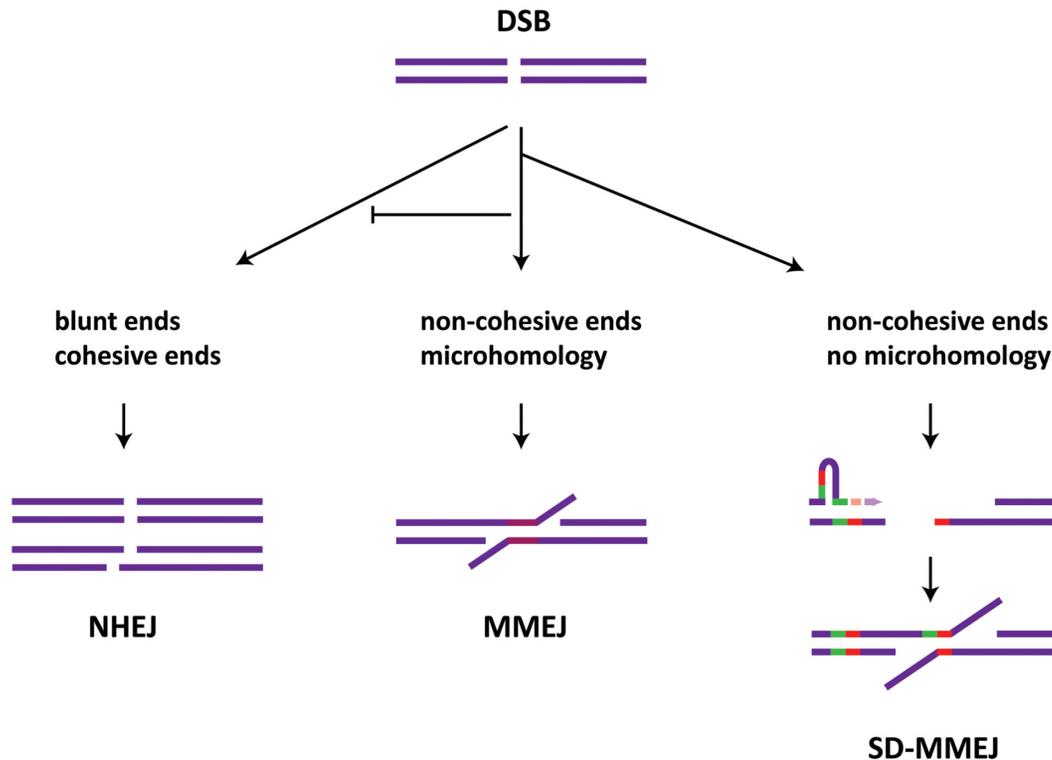


Figure 5. Proposed model for the end-joining repair of extrachromosomal DSBs in eukaryotic cells. The structure of the break and the microhomologies used for the particular repair mechanism are shown as red or green lines.

joining pathways to the repair of extrachromosomal DSBs, using rapid transient transfection assays. This may provide a favorable approach to study the recombination of free DNA in a cellular environment, such as the end-joining of plasmids or viral vectors delivered into the cells, or during viral infection. When compared with a genome-integrated reporter counterpart, it may also be used to assess if extrachromosomal DNA breaks may be treated differently than genomic breaks, as suggested by previous studies (47,48).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Ioannis Xenarios and other members of the Swiss Institute of Bioinformatics group, and especially Thomas Junier, for overall support during this project.

FUNDING

Swiss Government Commission for Technology and Innovation (grant number 13939.1 PFLS-LS), Selexis SA, University of Lausanne. Funding for open access charge: University of Lausanne.

Conflict of interest statement. N.M. is the co-founder and shareholder of Selexis SA. K.K. is jointly funded by the Swiss Government and Selexis SA. Selexis SA had no role and commercial interest in the current study.

REFERENCES

- Jackson, S.P. (2002) Sensing and repairing DNA double-strand breaks. *Carcinogenesis*, **23**, 687–696.
- Chapman, J.R., Taylor, M.R.G. and Boulton, S.J. (2012) Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell*, **47**, 497–510.
- Hiom, K. (2010) Coping with DNA double strand breaks. *DNA Repair (Amst)*, **9**, 1256–1263.
- Mao, Z., Bozzella, M., Seluanov, A. and Gorbunova, V. (2008) Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA Repair (Amst)*, **7**, 1765–1771.
- Bennardo, N., Cheng, A., Huang, N. and Stark, J.M. (2008) Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet.*, **4**, e1000110.
- Wang, H., Perrault, A.R., Takeda, Y., Qin, W. and Iliakis, G. (2003) Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res.*, **31**, 5377–5388.
- Audebert, M., Salles, B. and Calsou, P. (2004) Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.*, **279**, 55117–55126.
- Boboila, C., Yan, C., Wesemann, D.R., Jankovic, M., Wang, J.H., Manis, J., Nussenzweig, A., Nussenzweig, M. and Alt, F.W. (2010) Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4. *J. Exp. Med.*, **207**, 417–427.
- Della-Maria, J., Zhou, Y., Tsai, M.-S., Kuhnlein, J., Carney, J.P., Paull, T.T. and Tomkinson, A.E. (2011) Human Mre11/human Rad50/Nbs1 and DNA ligase IIIalpha/XRCC1 protein complexes act together in an alternative nonhomologous end joining pathway. *J. Biol. Chem.*, **286**, 33845–33853.
- Gigi, V., Lewis, S., Shestova, O., Mijušković, M., Deriano, L., Meng, W., Luning Prak, E.T. and Roth, D.B. (2014) RAG2 mutants alter DSB repair pathway choice in vivo and illuminate the nature of 'alternative NHEJ'. *Nucleic Acids Res.*, **42**, 6352–6364.
- Paul, K., Wang, M., Mladenov, E., Bencsik-Theilen, A., Bednar, T., Wu, W., Arakawa, H. and Iliakis, G. (2013) DNA ligases I and III

- cooperate in alternative non-homologous end-joining in vertebrates. *PLoS One*, **8**, e59505.
12. Bärtsch, S., Dücker, K., Würzler, F.E. and Sengstag, C. (1997) Ectopic mitotic recombination in *Drosophila* probed with bacterial beta-galactosidase gene-based reporter transgenes. *Nucleic Acids Res.*, **25**, 3917–3924.
 13. Hamilton, A.A. and Thacker, J. (1987) Gene recombination in X-ray-sensitive hamster cells. *Mol. Cell. Biol.*, **7**, 1409–1414.
 14. Kucherlapati, R.S., Eves, E.M., Song, K.Y., Morse, B.S. and Smithies, O. (1984) Homologous recombination between plasmids in mammalian cells can be enhanced by treatment of input DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **81**, 3153–3157.
 15. Puchta, H., Kocher, S. and Hohn, B. (1992) Extrachromosomal homologous DNA recombination in plant cells is fast and is not affected by CpG methylation. *Mol. Cell. Biol.*, **12**, 3372–3379.
 16. Rubnitz, J. and Subramani, S. (1984) The minimum amount of homology required for homologous recombination in mammalian cells. *Mol. Cell. Biol.*, **4**, 2253–2258.
 17. Seidman, M.M. (1987) Intermolecular homologous recombination between transfected sequences in mammalian cells is primarily nonconservative. *Mol. Cell. Biol.*, **7**, 3561–3565.
 18. Seluanov, A., Mittelman, D., Pereira-Smith, O.M., Wilson, J.H. and Gorbunova, V. (2004) DNA end joining becomes less efficient and more error-prone during cellular senescence. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 7624–7629.
 19. Zhang, X. and Paull, T.T. (2005) The Mre11/Rad50/Xrs2 complex and non-homologous end-joining of incompatible ends in *S. cerevisiae*. *DNA Repair (Amst.)*, **4**, 1281–1294.
 20. Oh, S., Harvey, A., Zimbric, J., Wang, Y., Nguyen, T., Jackson, P.J. and Hendrickson, E.A. (2014) DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells. *DNA Repair (Amst.)*, **21**, 97–110.
 21. Truong, L.N., Li, Y., Shi, L.Z., Hwang, P.Y.-H., He, J., Wang, H. and Hendrickson, E.A. (2013) Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *PNAS*, **110**, 7720–7725.
 22. Wang, H., Shao, Z., Shi, L.Z., Hwang, P.Y.-H., Truong, L.N., Berns, M.W., Chen, D.J. and Wu, X. (2012) CtIP protein dimerization is critical for its recruitment to chromosomal DNA double-stranded breaks. *J. Biol. Chem.*, **287**, 21471–21480.
 23. Yu, A.M. and McVey, M. (2010) Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.*, **38**, 5706–5717.
 24. Muñoz-Galván, S., López-Saavedra, A., Jackson, S.P., Huertas, P., Cortés-Ledesma, F. and Aguilera, A. (2013) Competing roles of DNA end resection and non-homologous end joining functions in the repair of replication-born double-strand breaks by sister-chromatid recombination. *Nucleic Acids Res.*, **41**, 1669–1683.
 25. Zhou, Y., Caron, P., Legube, G. and Paull, T.T. (2014) Quantitation of DNA double-strand break resection intermediates in human cells. *Nucleic Acids Res.*, **42**, e19.
 26. Urlaub, G. and Chasin, L.A. (1980) Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 4216–4220.
 27. Grandjean, M., Girod, P.-A., Calabrese, D., Kostyrko, K., Wicht, M., Yerly, F., Mazza, C., Beckmann, J.S., Martinet, D. and Mermod, N. (2011) High-level transgene expression by homologous recombination-mediated gene transfer. *Nucleic Acids Res.*, **39**, e104.
 28. Chan, S.H., Yu, A.M. and McVey, M. (2010) Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in *Drosophila*. *PLoS Genet.*, **6**, 1–16.
 29. Costantino, L., Sotiriou, S.K., Rantala, J.K., Magin, S., Mladenov, E., Helleday, T., Haber, J.E., Iliakis, G., Kallioniemi, O.P. and Halazonetis, T.D. (2014) Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science*, **343**, 88–91.
 30. Lee, K. and Lee, S.E. (2007) *Saccharomyces cerevisiae* Sae2- and Tel1-dependent single-strand DNA formation at DNA break promotes microhomology-mediated end joining. *Genetics*, **176**, 2003–2014.
 31. Liang, L., Deng, L., Nguyen, S.C., Zhao, X., Maulion, C.D., Shao, C. and Tischfield, J.A. (2008) Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Res.*, **36**, 3297–310.
 32. Yousefzadeh, M.J., Wyatt, D.W., Takata, K.-I., Mu, Y., Hensley, S.C., Tomida, J., Bylund, G.O., Doublé, S., Johanson, E., Ramsden, D. et al. (2014) Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ. *PLoS Genet.*, **10**, e1004654.
 33. Decottignies, A. (2007) Microhomology-mediated end joining in fission yeast is repressed by pku70 and relies on genes involved in homologous recombination. *Genetics*, **176**, 1403–1415.
 34. Kabotyanski, E.B., Gomelsky, L., Han, J.-O., Stamato, T.D. and Roth, D.B. (1998) Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.*, **26**, 27–31.
 35. Liang, F., Romanienko, P.J., Weaver, D.T., Jeggo, P.A. and Jasin, M. (1996) Chromosomal double-strand break repair in Ku80-deficient cells. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 8929–8933.
 36. Liou, G.-Y. and Peter, S. (2010) Reactive oxygen species in cancer. *Free Radic. Res.*, **44**, 479–496.
 37. Szatrowski, T.P. and Nathan, C.F. (1991) Production of large amounts of hydrogen peroxide by human tumor cells. *Cancer Res.*, **51**, 794–798.
 38. Bentley, J., Diggle, C.P., Harnden, P., Knowles, M.A. and Kiltie, A.E. (2004) DNA double strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. *Nucleic Acids Res.*, **32**, 5249–5259.
 39. Tobin, L.A., Robert, C., Nagaria, P., Chumsri, S., Twaddell, W., Ioffe, O.B., Greco, G.E., Brodie, A.H., Tomkinson, A.E. and Rassool, F.V. (2012) Targeting abnormal DNA repair in therapy-resistant breast cancers. *Mol. Cancer Res.*, **10**, 96–107.
 40. Bednar, J., Furrer, P., Katritch, V., Stasiak, a.Z., Dubochet, J. and Stasiak, A. (1995) Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.*, **254**, 579–594.
 41. Hays, J.B., Magar, M.E. and Zimm, B.H. (1969) Persistence length of DNA. *Biopolymers*, **8**, 531–536.
 42. Lu, Y., Weers, B. and Stellwagen, N.C. (2002) DNA persistence length revisited. *Biopolymers*, **61**, 261–275.
 43. Ceccaldi, R., Liu, J.C., Amunugama, R., Hajdu, I., Primack, B., Petalcorin, M.I.R., O'Connor, K.W., Konstantinopoulos, P.A., Elledge, S.J., Boulton, S.J. et al. (2015) Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. *Nature*, **518**, 258–262.
 44. Kent, T., Chandramouly, G., Mcdevitt, S.M., Ozdemir, A.Y. and Pomerantz, R.T. (2015) Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase θ. *Nat. Struct. Mol. Biol.*, **22**, 230–237.
 45. Koole, W., van Schendel, R., Karambelas, A.E., van Heteren, J.T., Okihara, K.L. and Tijsterman, M. (2014) A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat. Commun.*, **5**, 3216.
 46. Mateos-Gomez, P.A., Gong, F., Nair, N., Miller, K.M., Lazzarini-Denchi, E. and Sfeir, A. (2015) Mammalian polymerase θ promotes alternative NHEJ and suppresses recombination. *Nature*, **518**, 254–257.
 47. Schulte-Uentrop, L., El-Awady, R.A., Schliecker, L., Willers, H. and Dahm-Daphi, J. (2008) Distinct roles of XRCC4 and Ku80 in non-homologous end-joining of endonuclease- and ionizing radiation-induced DNA double-strand breaks. *Nucleic Acids Res.*, **36**, 2561–2569.
 48. Willers, H., McCarthy, E.E., Hubbe, P., Dahm-Daphi, J. and Powell, S.N. (2001) Homologous recombination in extrachromosomal plasmid substrates is not suppressed by p53. *Carcinogenesis*, **22**, 1757–1763.