

Strand bias influences the mechanism of gene editing directed by single-stranded DNA oligonucleotides

Kerry Falgowski, Carly Falgowski, Cassie York-Vickers and Eric B. Kmiec*

Marshall Institute for Interdisciplinary Research, Marshall University, Robert C. Byrd Biotechnology Science Center, 1700 Third Avenue, Suite 220, Huntington, WV 25755, USA

Received June 29, 2010; Revised January 19, 2011; Accepted January 21, 2011

ABSTRACT

Gene editing directed by modified single-stranded DNA oligonucleotides has been used to alter a single base pair in a variety of biological systems. It is likely that gene editing is facilitated by the direct incorporation of the oligonucleotides via replication and/or by direct conversion, most likely through the DNA mismatch repair pathway. The phenomenon of strand bias, however, as well as its importance to the gene editing reaction itself, has yet to be elucidated in terms of mechanism. We have taken a reductionist approach by using a genetic readout in *Escherichia coli* and a plasmid-based selectable system to evaluate the influence of strand bias on the mechanism of gene editing. We show that oligonucleotides (ODNs) designed to anneal to the lagging strand generate 100-fold greater ‘editing’ efficiency than ‘those that anneal to’ the leading strand. The majority of editing events (~70%) occur by the incorporation of the ODN during replication within the lagging strand. Conversely, ODNs that anneal to the leading strand generate fewer editing events although this event may follow either the incorporation or direct conversion pathway. In general, the influence of DNA replication is independent of which ODN is used suggesting that the importance of strand bias is a reflection of the underlying mechanism used to carry out gene editing.

INTRODUCTION

Gene editing using single-stranded DNA oligonucleotides (ODNs) is being developed as a molecular treatment of inherited disorders. Applications of gene editing in mouse cells range from gene alteration in embryonic stem cells

(1–4) to reversal of single base errors in genes in somatic cells (5–8). In parallel to the development of therapeutic strategies, many laboratories have been elucidating the mechanism and regulation of gene editing (9). Several years ago, our laboratory proposed a three-branched pathway model for nucleotide conversion directed by ODNs as a mechanistic framework (10). While all three paths ultimately lead to a converted base pair, two of them have gained the largest amount of experimental support. The first involves the ‘editing of a mutant’ base pair by the cell’s inherent DNA repair activity, most likely operating through the DNA mismatch repair system. Proponents of this path cite the fact that correction can take place on genes that are neither being transcribed nor replicated (11,12). Alternatively, cells that are being replicated most likely use a mode of gene editing that involves the incorporation of the ODN into a replicating region that covers the target site (13–16). Acting in this manner, the ODN incorporates and serves as part of an Okazaki fragment by homologously aligning at the target and priming extension of newly synthesized DNA. If this occurs, then technically gene editing takes place in the absence of mismatch repair since a second round of replication could generate a different (edited) base pair (9,10). Strong evidence for this model was provided by the work of Radecke *et al.* (17) in studies where an incorporated ODN was cross-linked and, subsequently, isolated directly from the target site. These data strengthened the replication-dependent branch of the original model (10).

Another observation related to the mechanism of gene editing centers on a phenomenon known as strand bias (18). Many groups have reported that while either strand, transcribed or non-transcribed (lagging or leading), can be targeted for nucleotide exchange, there appears to be bias toward one or the other, dependent on the host cell type. This phenomenon was recognized in early studies of gene editing (19) but its relevance and importance to mechanism of action is still not known. In addition, it is not known what role, if any, it plays in

*To whom correspondence should be addressed. Tel: +1 304 696 3830; Fax: +1 304 696 3839; Email: kmiec@marshall.edu

directing a particular path of nucleotide exchange. It is also important to note that strands designated as 'T or NT' can also serve as either the lagging or leading strands during DNA replication. As such, the strand bias effect may also be reflected in whether the cells are able to replicate or not during a gene editing event.

We decided to examine the relationship between strand bias and mechanism by asking whether the targeting of a particular strand could, in fact, influence which mode of editing is chosen for resolution. To do so, we simplified the reaction conditions by utilizing a selectable system that relies on phenotypic and genetic readout in *Escherichia coli* to assess gene editing activity. This approach eliminates several reaction parameters that could influence the interpretation of the data. In this case, our results indicate that ODNs that anneal to the transcribed strand or the lagging strand of replication (in a system that selects for the acquisition of antibiotic resistance) direct more editing events. But, in most cases, this activity results in a mixed population of corrected and non-corrected templates or molecules. Annealing to the strand designated as nontranscribed (or leading strand of DNA replication in our system) leads to fewer modified cells, but the vast majority of them contain a homogenous population of converted genes. The replication mode is the more prevalent form of gene editing when ODNs annealing to the lagging strand are used, while the direct conversion and replication via incorporation modes of editing appear at an equal level when the leading strand is the target.

MATERIALS AND METHODS

Plasmid and ODNs

The pK^Sm4021 plasmid containing the targeted mutant kanamycin resistance gene was constructed from pWE15 by converting a T to a G residue at position 4021. Single-stranded ODNs listed below, were synthesized and obtained from Integrated DNA Technologies, Inc. (Coralville, IA, USA);

Kan49NT 5' G*C*C*GATTGTCTGTTGTGCCAGT
CGTAGCCGAATAGCCTCTCCACC*C*A*A 3'
Kan49T 5' T*T*G*GGTGGAGAGGCTATTCGGCTA
CGACTGGGCACAACAGACAATC*G*G*C 3'
Kan49NT-PM 5' G*C*C*GATTGTCTGTTGTGCCA
GTCCTAGCCGAATAGCCTCTCCACC*C*A*A 3'
Kan49T-PM 5' T*T*G*GGTGGAGAGGCTATTCGG
CTAGGACTGGGCACAACAGACAATC*G*G*C 3'
Kan49NT/3'InvdT 5' G*C*C*GATTGTCTGTTGTGC
CCAGTCGTAGCCGAATAGCCTCTCC
ACC*C*A*A/InvdT 3'
Kan49T/3'InvdT 5' T*T*G*GGTGGAGAGGCTATTC
GGCTACGACTGGGCACAACAGACA
ATC*G*G*C/InvdT 3'

The Kan49T ODN is complementary to the transcribed strand of the plasmid, which is also the template for the lagging strand synthesis. The Kan49NT ODN is complementary to the non-transcribed strand of the plasmid, which is also the template for the leading strand synthesis.

Electroporation, plating and selection

Electroporation mixtures contained 1 µg of the pK^Sm4021 plasmid, 1.05 µg of ODNs in dH₂O up to 50 µl. Five microliters (100 ng of plasmid, 105 ng of ODN) of the mixture was immediately used to transform 20 µl of competent *E. coli* DH10B (Invitrogen) by electroporation using a Cell-Porator apparatus (Life Technologies, Inc.). The transformed cells were then allowed to recover in 5 ml of super optimal broth with catabolite repression (SOC) media at 37°C for the indicated incubation recovery time (1 h, 2 h, 3 h, 4 h). A 100 µl of a 5 × 10⁻⁴ dilution aliquot from the 5 ml culture was plated onto an ampicillin (100 µg/ml) plate for each sample. The remainder of the culture treated with the Kan49NT ODN was pelleted by centrifugation, re-suspended in 100 µl of Luria Broth (LB) broth, then plated onto a kanamycin (50 µg/ml) plate. For cultures treated with the Kan49T ODN, a 1 ml aliquot was pelleted by centrifugation, re-suspended in 100 µl of LB broth, then plated onto a kanamycin (50 µg/ml) plate. Colony counts were determined after a 24 h incubation period at 37°C by an Accu-count 1000 plate reader (Biologics). At least three independent experiments were performed for each reaction condition. To account for variability of electroporation efficiencies, kanamycin colony counts were normalized to 5 × 10⁵-ampicillin colonies. Therefore, correction efficiency of the kan^S to kan^R conversion was measured as the number of kanamycin colonies per 5 × 10⁵-ampicillin colonies.

Restriction- enzyme digest and DNA sequencing

Individual colonies grown on either kanamycin or ampicillin plates were grown at 37°C for 8 h in LB broth with the same selective antibiotic (kanamycin 50 µg/ml, ampicillin 100 µg/ml). Plasmid populations were isolated from each individual culture using the Qiagen QIAprep Spin Miniprep Kit (Valencia, CA, USA). Then, 500 ng of recovered plasmid for each sample was digested with 1 U of the BfaI enzyme for 2 h in a total reaction volume of 20 µl. Ten microliter of the digest reaction was loaded onto a 0.8% agarose gel containing ethidium bromide. Densitometry of Restriction Fragment Length Polymorphism (RFLP) bands were analyzed with Alpha Innotech AlphaView[®] Software version 3.0.3.0. DNA sequencing analysis was performed on the individual purified-plasmid populations by cycle sequencing at the Molecular Biology Core Facility at the A.I. duPont Hospital for Children (Wilmington, DE, USA).

RESULTS

Strand bias with respect to correction efficiency

The plasmid pK^Sm4021 was constructed from pWE15 by converting a T to a G residue at position 4021 (Figure 1). This change introduces a stop codon in the coding region of the kanamycin resistance gene and creates a BfaI restriction site. The inherent ampicillin-resistance gene remains intact. The mutation is targeted by an ODN, which is designed to be complementary to either the T or NT DNA strand, with the exception of a centrally

located mismatched base. The mismatched base pairing induces a targeted nucleotide exchange to occur, which restores kanamycin resistance. ODNs are designed to direct a nucleotide-exchange reaction by converting the mutant TAG codon to TAC. This maintains the same amino acid sequence for kanamycin resistance, however, by converting the codon to TAC instead of back to the wild-type codon TAT it allows for the ability to distinguish by DNA sequence analysis conversion events that were ODN-directed and not from contamination of

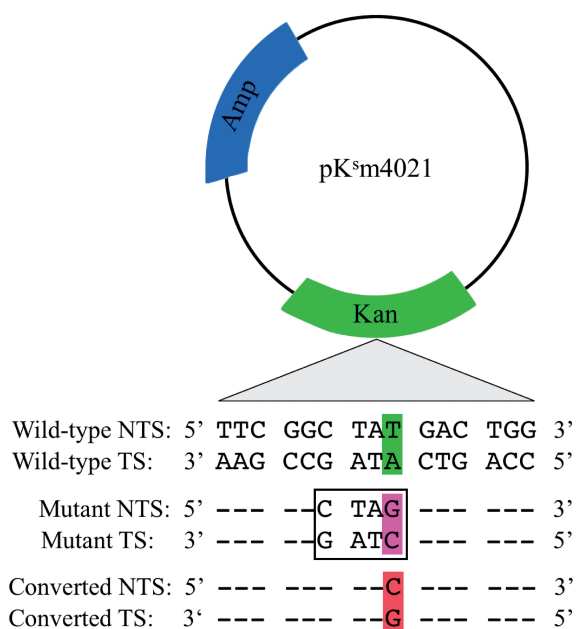


Figure 1. Model system for gene repair utilizing the pK^Sm4021 plasmid. The mutant kan^S plasmid has a TAT to TAG mutation early within the open reading frame of the kanamycin resistance gene, which creates a premature stop codon and a truncated, non-functional protein product. The mutation is targeted by an ODN, which is designed to be complementary to either the T- or NT-DNA strand, with the exception of a centrally located mismatched base. The mismatched base pairing induces a targeted nucleotide-exchange reaction to occur, which restores kanamycin resistance. ODNs are designed to direct a nucleotide exchange reaction by converting the mutant TAG codon, to TAC. This maintains the same amino acid sequence for kanamycin resistance, however, by converting the codon to TAC instead of back to the wild-type codon TAT it allows for the ability to distinguish by DNA sequence analysis conversion events that were ODN-directed and not from contamination of wild-type plasmid. The boxed four bases of the mutant-plasmid sequences are recognized and cut by the BfaI restriction enzyme.

wild-type plasmid. The boxed four bases of the mutant plasmid sequences are recognized and cut by the BfaI restriction enzyme (Figure 1). The pK^Sm4021 plasmid has a low-copy ColE1 replication origin and is therefore maintained at a copy number of 10–20.

The gene editing reaction was begun by mixing the ODN with plasmid, pK^Sm4021, immediately prior to electroporation into *E. coli* strains DH10B. The cells were allowed to recover at 37 for 1 h, 2 h, 3 h or 4 h prior to plating on both ampicillin and kanamycin plates to observe possible changes to correction efficiency over time (see ‘Methods and Materials’ section). Antibiotic-resistant colonies were determined after 24 h of incubation at 37°C. The ODNs used in these experiments were 49 bases in length and annealed to either the T or replication lagging strand (Kan49T) or NT or replication leading (Kan49NT) strand of the mutant kanamycin resistance gene in plasmid pK^Sm4021 (Table 1). Average colony counts for each recovery time-point of cells treated with either Kan49T or Kan49NT are shown in Figure 2A and B. Cells treated with the Kan49T, as opposed to the Kan49NT, had about 100-fold more kanamycin-resistant colonies at each time-point. Figure 2C shows ampicillin colony counts of Kan49T- and Kan49NT-treated cells at a 5×10^{-4} dilution for each recovery point. Ampicillin colonies for both ODNs steadily increase over time and are within standard deviation of each other.

The correction efficiency defined as the number of kanamycin colonies per 5×10^5 -ampicillin colonies was calculated for each time point during the recovery period prior to antibiotic selection is given in Figure 2D. When only the mutant plasmid was transformed into the cells as a negative control, a correction efficiency of 0.17 ± 0.06 was observed. As additional controls, we determined correction efficiency for ODNs Kan49T-PM and Kan49NT-PM, which have sequences that bear perfect complementarity to the target site. The correction efficiencies for Kan49T-PM- and Kan49NT-PM-treated cells were 0.15 ± 0.05 ($P = 0.09$) and 0.11 ± 0.02 ($P = 0.3$), respectively. The two perfectly matched ODNs do not have the capacity to direct the conversion of the target base, as no significant correction efficiency defined as $P < 0.01$ was observed above background. Alternatively, when we transfected cells with the targeting ODNs a significant correction efficiency above background was observed. The average correction efficiency for Kan49NT-treated cells was 0.52 ± 0.06 ($P = 0.0003$),

Table 1. List of ODN used to investigate correction efficiency with respect to strand targeting

Kan49NT	5'	G*C*C*GATTGCTGTGTGCCAGTC <u>G</u> TAGCCGAATAGCCTCTCCACC*C*A*A 3'
Kan49T	5'	T*T*G*GGTGGAGAGGCTATTCGGCTA <u>C</u> GACTGGGCACAACAGACAATC*G*G*C 3'
Kan49NT-PM	5'	G*C*C*GATTGCTGTGTGCCAGTC <u>C</u> TAGCCGAATAGCCTCTCCACC*C*A*A 3'
Kan49T-PM	5'	T*T*G*GGTGGAGAGGCTATTCGGCTA <u>G</u> GACTGGGCACAACAGACAATC*G*G*C 3'
Kan49NT/3'InvdT	5'	G*C*C*GATTGCTGTGTGTGCCAGTC <u>G</u> TAGCCGAATAGCCTCTCCACC*C*A*A/InvdT 3'
Kan49T/3'InvdT	5'	T*T*G*GGTGGAGAGGCTATTCGGCTA <u>C</u> GACTGGGCACAACAGACAATC*G*G*C/InvdT 3'

NT-ODNs target the NT-DNA strand, while T-ODNs target the T DNA strand. PM-ODNs are perfectly complementary to the DNA strand; therefore, it is not designed to elicit a correction event. Kan49NT/3'InvdT and Kan49T/3'InvdT ODNs have an inverted 3'-end nucleotide (/InvdT) which blocks the extension of DNA polymerases. ODNs were synthesized with three phosphorothioate linkages (*) to protect from degradation from Integrated DNA Technologies, Inc. The centrally located bold and underlined base directs the targeted nucleotide exchange event.

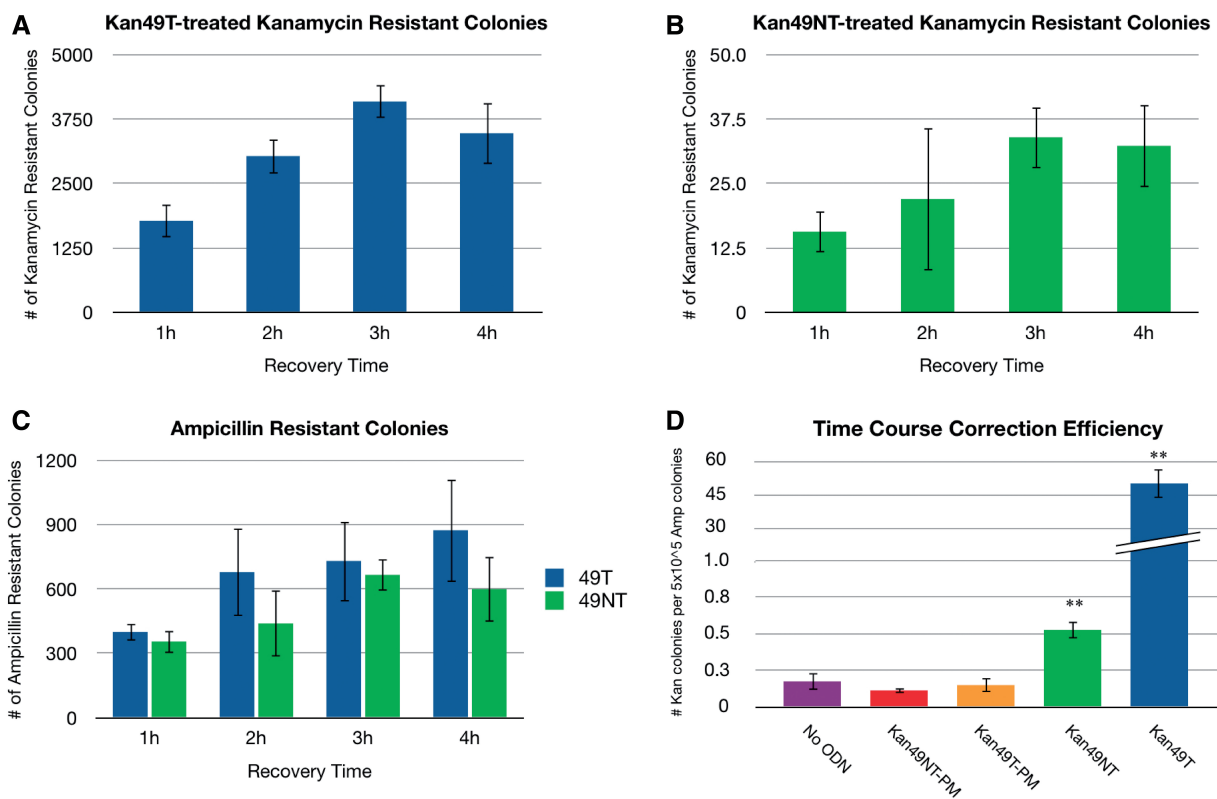


Figure 2. Average kanamycin-resistant colony counts for each recovery time point resulting from Kan49T-treated (A) and Kan49NT-treated (B) cells. (C) Average ampicillin-resistant colony counts of Kan49T- and Kan49NT-treated cells at a 5×10^{-4} dilution calculated from data from each recovery time point. (D) Average correction efficiencies calculated from data from all recovery time points for cells treated with no ODN, Kan49NT-PM, Kan49T-PM, Kan49NT and Kan49T. Correction efficiency is defined as the number of kanamycin-resistant colonies per 5×10^5 ampicillin-resistant colonies. ODN molarity was maintained at a 350:1 ODN to plasmid ratio for each experiment. For each reaction condition at least three independent experiments were conducted. $**P < 0.001$.

which is about a 4-fold increase over background. Alternatively, when cells were treated with the Kan49T ODN a correction efficiency of 50.1 ± 6.7 ($P = 0.00003$) was observed, which is ~100-fold greater than the correction efficiency for Kan49NT-treated cells and about a 300-fold increase over background levels.

RFLP analysis

Twenty random kanamycin-resistant colonies were picked from the kanamycin plates and grown at 37°C for 8 h in LB broth under kanamycin selection. Plasmid populations were isolated from each individual culture and then digested with BfaI. In Figure 3A, a representative map of expected DNA bands resulting from a restriction digest of a mutant plasmid, converted plasmid and mixed plasmid population is presented. The BfaI enzyme recognizes the mutant base present at the target site, resulting in the appearance of a 1301-bp band. The conversion of the target base will result in the loss of a cut site in the 1301-bp band, thereby producing a 1534 band.

Figure 3B shows the RFLP patterns of the 20 kanamycin-resistant plasmid population isolates obtained from Kan49T- or Kan49NT-treated cells for the 1 h-recovery time point (kan^R49T #1-20 and kan^R49NT #1-20). The lane represented by kan^S, is the RFLP from a mutant plasmid population isolated from an amp^R/kan^S

colony from either the Kan49T- or Kan49NT-treated cells, and shows only the mutant band (1301 bp). Analyses of kan^R isolate RFLP results reveals four different banding patterns (A–D). Three of the patterns have the predicted mixed-plasmid-population RFLP. For Pattern (A) there is more of the mutant band (1301 bp) than the converted band (1534 bp), Pattern (B) has an equal mutant band to converted band and for Pattern (C) there is more converted than mutant band. The fourth Pattern (D) has the predicted converted plasmid RFLP in which only the converted band is present.

The results of the RFLP analysis prompted the need to confirm our predictions with sequence analysis of the plasmid populations. Figure 3C depicts representative data of the NT (leading) and T (lagging) strand sequences obtained from each of the RFLP patterns. As expected, the kan^S sequence shows only the mutant G–C base pairs at the target site. For the (A) RFLP, the sequence shows a mixed peak at the target site for the NT strand with the mutant G peak being more dominant than the small converted C peak present. The T strand shows the mutant C peak with a very basal G peak present slightly above background. Sequence analysis for Pattern (B) shows the predicted equal mutant to converted peaks present on both the NT and T strands at the target site. Pattern (C) sequence depicts a greater converted C peak on the NT

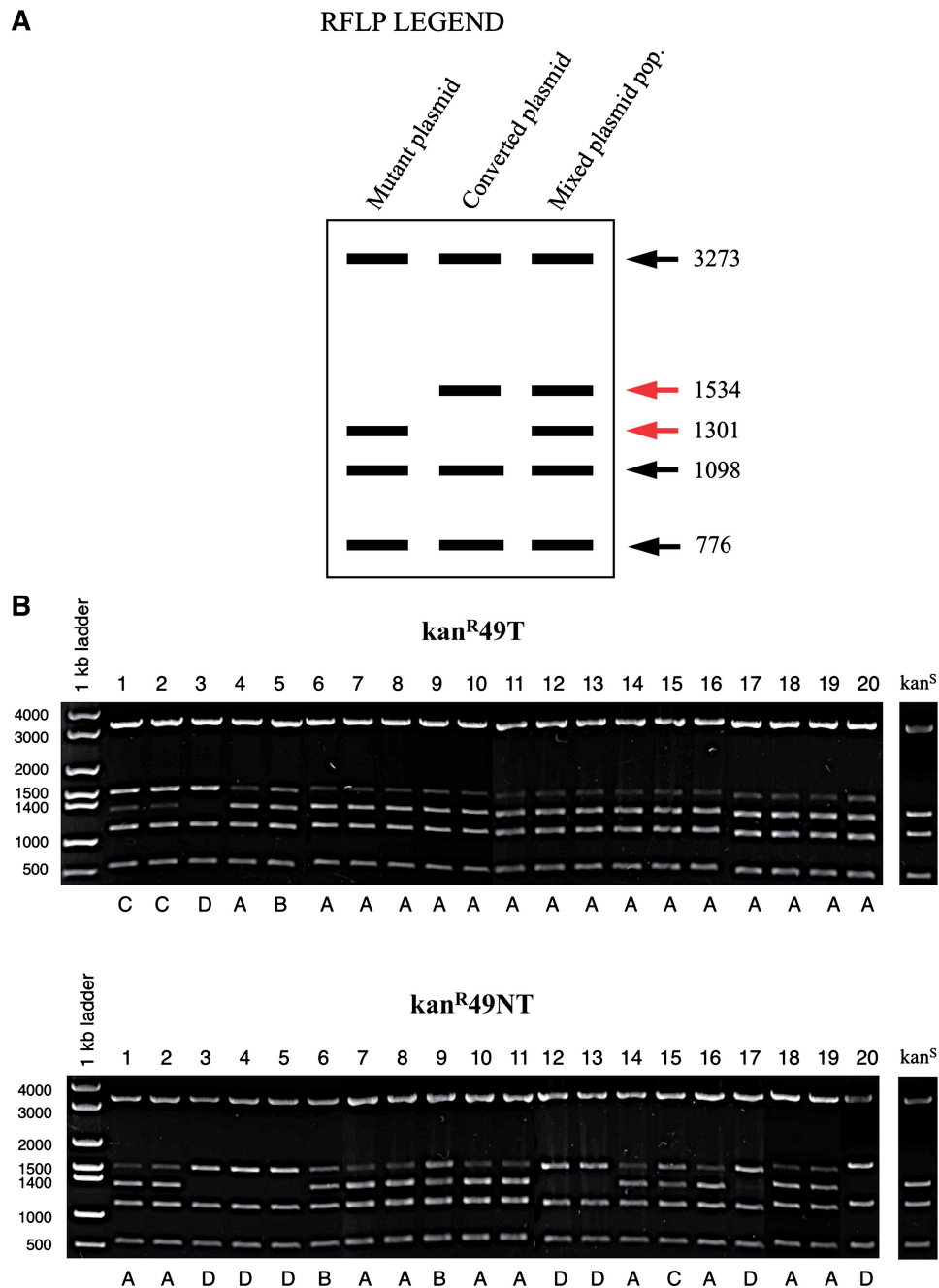


Figure 3. (A) Representative map of expected DNA bands resulting from a restriction digest of a mutant plasmid, converted plasmid and mixed plasmid population is presented. The Bfal enzyme recognizes the mutant base present at the target site, resulting in the appearance of a 1301 bp band. The conversion of the target base will result in the loss of a cut site in the 1301 bp band thereby producing a 1534 band. (B) RFLP patterns of the 20 kanamycin-resistant plasmid population isolates resulting from Kan49T- or Kan49NT-treated cells for the 1 h recovery time point (kan^R49T #1-20 and kan^R49NT #1-20). Isolate kan^S is the RFLP from a mutant-plasmid population isolated from an amp^R/kan^S colony from the respective Kan49T- and Kan49NT-treated cells. Four different RFLP patterns are shown, A–D. (C) Representative data of the NT and T strand sequences obtained from each of the RFLP patterns. The wild-type base pairing at the target site is T–A, the mutant base pairing is G–C, while the converted base pairing is C–G.

strand than the mutant G peak at the target site and a mixed peak on the T strand in which the converted G peak is slightly larger than the mutant C peak. Finally, Pattern (D) shows only the presence of the converted C–G base pairs at the target site. Therefore, the RFLP analyses of plasmid populations generated by the respective ODNs are consistent with the DNA sequence-level analyses.

For each recovery time-point, 20 kan^R colonies of Kan49T- and Kan49NT-treated cells were processed as a population sampling to analyze any RFLP preferences with regard to strand bias and time of recovery. The percentages for each type of RFLP present in the 20 sample pool are shown in Figure 4A and B; for Kan49T- and Kan49NT-treated cells. The average of each RFLP

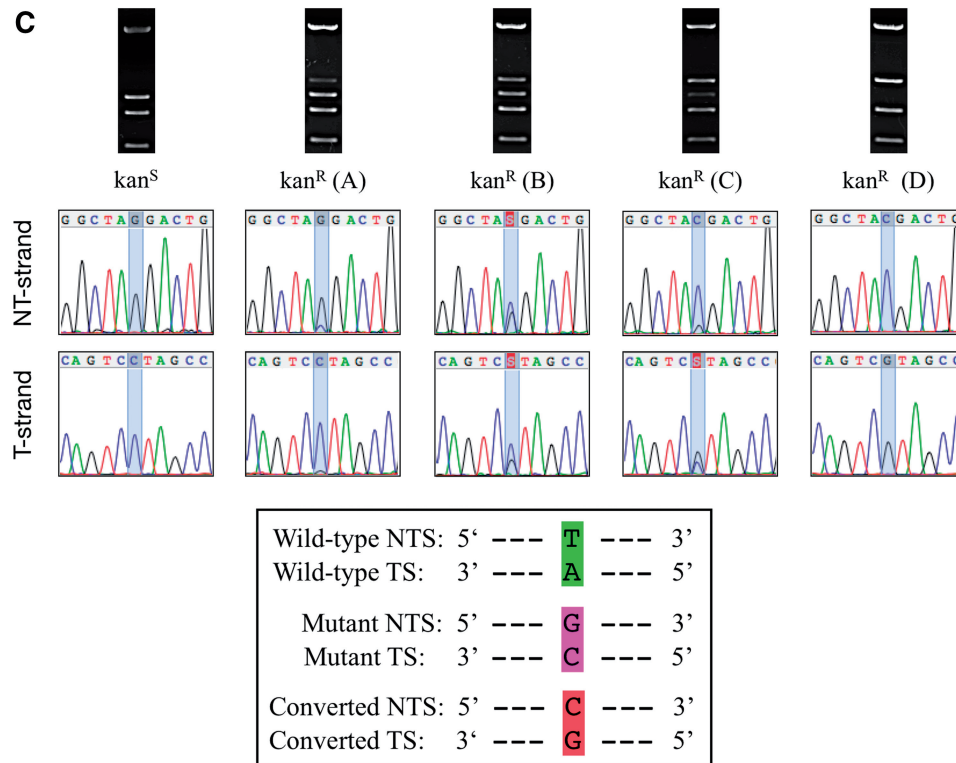


Figure 3. Continued.

banding Pattern (A–D) was calculated for the entire time course (Figure 4C) since there are too few events generated during the 1 h time point to make valid interpretations of the experimental results. The majority of kanamycin-resistant colonies resulting from Kan49T-treated cells (kan^R49T) had the (A) RFLP pattern with a percentage of $68.8 \pm 12.5\%$. Pattern (D) was next most common with an average of $21.3 \pm 12.5\%$ ($P = 0.0009$). Cells targeted by the Kan49NT ODN did not display this bias since Patterns (A) and (D) were present at the same percentages with respective averages of $48.8 \pm 7.5\%$ and $42.5 \pm 8.7\%$ ($P = 0.16$). For both targeting ODNs, Patterns (B) and (C) were very rare in the sampling population. These data indicate that there is an approximately equal probability of (D) or (A) events taking place when gene editing is initiated on the NT or leading strand of replication.

ODNs with 3' inverted dT blocks DNA polymerase extension

The data above reveal four RFLP categories, yet the reaction mechanism responsible for producing each banding pattern had not been fully elucidated. The current thinking is that modified single-stranded ODNs induce nucleotide exchange using two pathways; direct repair or incorporation during replication, with the latter route supported by the most published data (14–17, 20, 21). Hence, we decided to inhibit the incorporation pathway by utilizing ODNs with a 3'-end inverted nucleotide to block extension by DNA polymerase. A change in any of the frequencies of the RFLP categories, when

the inverted ODNs were used, would likely affect the pathway of incorporation more than the direct repair pathway.

Cells were transfected with ODNs (Kan49NT/3'InvdT and Kan49T/3'InvdT) identical to the Kan49NT and Kan49T ODNs with the exception of a 3'-end inverted nucleotide, then allowed to recover for 1 h, 2 h, 3 h or 4 h prior to plating for antibiotic selection. Correction efficiencies (number of kanamycin resistant colonies per 5×10^5 ampicillin-resistant colonies) were calculated from triplicate experiments for Kan49NT/3'InvdT and Kan49T/3'InvdT-treated cells to be 0.35 ± 0.30 and 6.2 ± 4.5 , respectively, and once again presented as an average (Figures 4C and 5A). When the Kan49NT/3'InvdT correction efficiency was compared to the Kan49NT correction efficiency of 0.52 ± 0.06 ($P = 0.15$), there was only a modest change based mostly on the low numbers achieved using the NT ODN. Alternatively, the blocked 3'-end of the Kan49T/3'InvdT caused almost an 8-fold reduction in the correction efficiency when compared the Kan49T ODN correction efficiency of 50.1 ± 6.7 ($P = 0.0002$).

Average RFLP percentages were also calculated for the 3' inverted ODNs. Ten to twenty kan^R colonies of Kan49NT/3'InvdT- and Kan49T/3'InvdT-treated cells were processed as a population sampling to analyze RFLP percentages each recovery time-point (1 h, 2 h, 3 h, 4 h). Average number of kanamycin-resistant colonies resulting from each of the four RFLP categories were calculated from the RFLP percentages and the total number of kanamycin-resistant colonies produced from

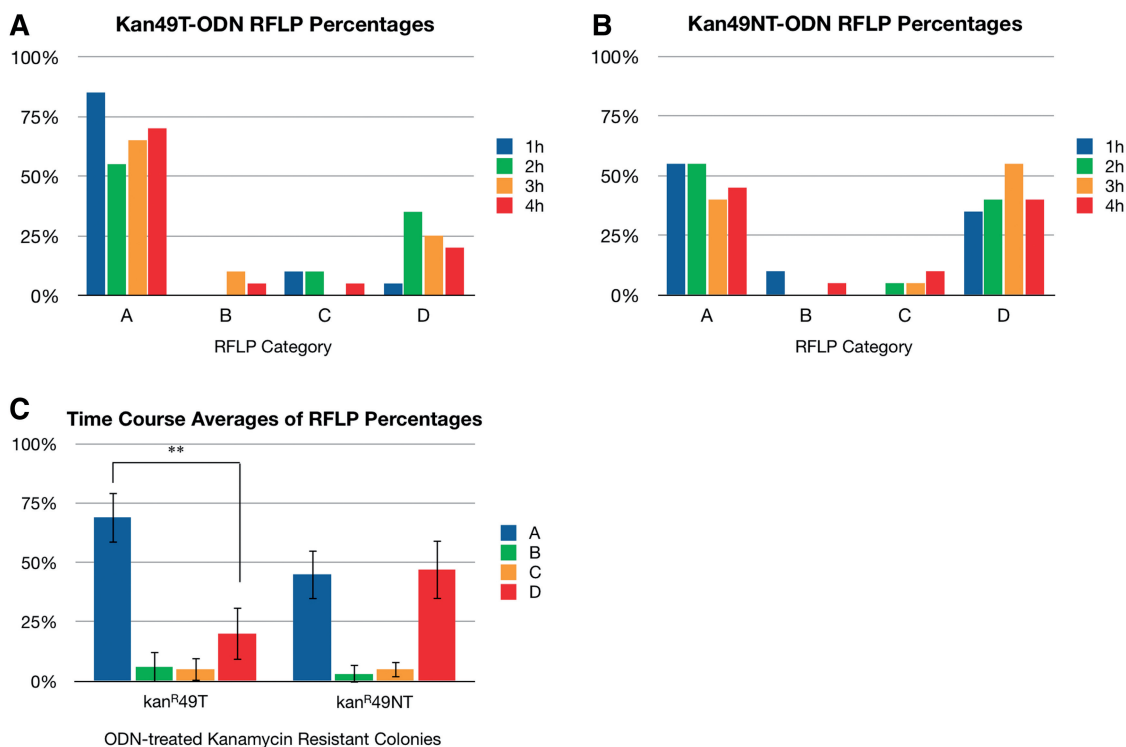


Figure 4. The recovery time course (1 h, 2 h, 3 h and 4 h) percentages for each of the four RFLPs (A, B, C and D) present in the 20 sample pool for kanamycin-resistant colonies arising from (A) Kan49T-(kan^R49T) and (B) Kan49NT-(kan^R49NT) treated cells. (C) Time course averages with respective standard deviations were calculated for each of the four RFLPs resulting from kan^R49T and kan^R49NT kanamycin-resistant colonies. ** $P < 0.001$.

Kan49NT/3'InvdT- and Kan49T/3'InvdT-treated cells. The average number of kanamycin-resistant colonies arising from each of the four RFLP categories treated the 3'-end inverted ODNs was compared to kanamycin-resistant colonies resulting from Kan49NT- and Kan49T-treated cells (Figure 5B and C). Kanamycin-resistant colonies resulting from cells targeted by the Kan49NT ODN (kan^R49NT) had relatively equal percentages of the (A) and (D) patterns which predicted out of the total number of kanamycin-resistant colonies 12 ± 3 colonies had the (A) pattern and 11 ± 6 colonies had the (D) pattern (Figure 5B). The (B) and (C) RFLP banding patterns were rare in the sampling populations. The kan^R colonies resulting from cells targeted by the Kan49NT/3'InvdT ODN showed a dramatic decrease in the (A) RFLP present in the kanamycin-resistant population. The predicted number of kan^R colonies resulting from Kan49NT/3'InvdT-treated cells with the (A) pattern in the total kanamycin-resistant population was 1 ± 1 ($P = 0.0001$). There were no significant changes ($P < 0.01$) in the number of kanamycin-resistant colonies with (B), (C) or (D) patterns when comparing the kan^R49NT and kan^R49NT/3'InvdT populations. Similarly, kan^R colonies produced from cells transfected with the Kan49T/3'InvdT ODN had a similar shift in RFLP frequencies as the Kan49NT/3'InvdT kan^R colonies. The results in Figure 5C demonstrate the significant decrease observed in the (A) pattern numbers from 2066 ± 565 kan^R 49T colonies to 23 ± 46 ($P = 0.0002$),

kan^R 49T/3'InvdT colonies in sample populations. There were no significant changes ($P < 0.01$) in the number of kanamycin-resistant colonies with (B) or (C) patterns when comparing the kan^R49T and kan^R49T/3'InvdT populations. The results of annealing Kan49T to the lagging strand, however, did produce an interesting result as there is some reduction in the number of editing events of the D-type when the blocked 3' ODN is used. These data suggest involvement of replication in the generation of D-type events albeit lower levels than the impact on A-type events. Overall, the data consistently support the notion that ODNs annealing to the lagging strand direct gene editing using a replication mode rather than the direct conversion mode. Stated in a different way, the incorporation mode is more prevalent in gene editing when ODNs designed to anneal to the lagging strand are used.

DISCUSSION

Modified single-stranded ODNs can direct the exchange of nucleotides at specific sites in extra-chromosomal and chromosomal environments in a process known as gene editing. The mechanism of the reaction is not fully elucidated but likely involves DNA repair and replication acting in a concerted fashion to enable sequence exchange. Overall, the field has witnessed an increase in the amount of data validating gene editing. Many papers that focus on the mechanism of action suggest that there is a preference for editing either the NT or the T strand of the target gene

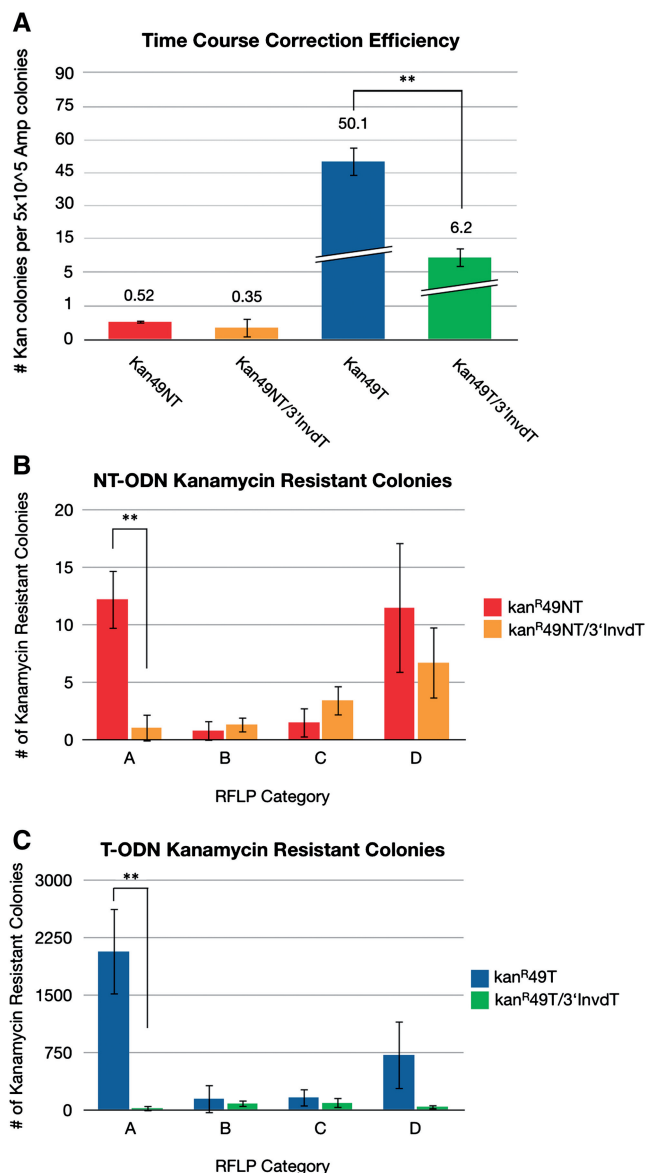


Figure 5. (A) Average correction efficiencies for the entire time course of cells treated with Kan49NT, Kan49T, Kan49NT/3'InvdT and Kan49T/3'InvdT ODNs. Correction efficiency is defined as the number of kanamycin resistant colonies per 5×10^5 ampicillin-resistant colonies. ODN molarity was maintained at a 350:1 ODN to plasmid ratio for each experiment. For each reaction condition at least three independent experiments were conducted. (B and C) Average number of kanamycin-resistant colonies resulting from each of the four RFLP categories was calculated from the RFLP percentages and the total number of kanamycin-resistant colonies of the time course (1 h, 2 h, 3 h and 4 h) from kan^R49NT, kan^R49T, kan^R49NT/3'InvdT and kan^R49T/3'InvdT ODN-treated resistant colonies. Average RFLP percentages were calculated from 10–20 plasmid isolates resulting from each ODN-treated kanamycin-resistant colony populations. Three independent experiments were performed to determine average number of kanamycin-resistant colonies for each recovery time point. $**P < 0.001$.

(19,22,23). In this work, we examine the influence of strand bias on the pathway of correction chosen by the cell to resolve the mutation. By using a simplistic model system, coupled to a genetic selectable readout, we provide some rationale for the phenomenon seen in the gene

editing approach. By utilizing ODNs designed to be complementary to either the T or NT strand (by our definition of the orientation of the gene in the plasmid), we were able to discern differences in correction efficiency and the pathway used to direct the editing mode, i.e. repair, replication or a combination of the two.

Our results lead us to project two roadmaps for gene editing based on which strand is selected; they have both common and uniquely distinctive features. It is important to note that Brachman and Kmiec (18) suggested that the designations 'T' and 'NT', respectively, can be misleading since each strand can also serve as either the 'lagging' or 'leading' strand of DNA synthesis (at the target site). Further, it was found that lagging-strand activity and structure is dominant in determining which of the two strands exhibits a higher degree of targeting success. In the current system, the T strand is also the lagging strand of DNA synthesis and thus the higher number of kan^R colonies (100-fold greater) from plasmid targeted with the T ODN is predictable. In addition, annealing to the T strand can also induce higher numbers of kan^R colonies because the editing of the T strand enables immediate kanamycin resistance as oppose to targeting the NT strand. In the latter case, a second event must take place to generate the corrected gene phenotype and kan^R colonies.

Under the reaction conditions we use in this study, cells are allowed to undergo DNA replication by incubation in a large volume of media; there is a reduced level of stress in the environment and thus replication and cell division are enabled. As was seen in the upper panel of Figure 3A, the vast majority of plasmid populations per isolated bacterial clone gave rise predominantly to a single RFLP Pattern (A). We now present Figure 6A as a model to explain the types and ratios of RFLP patterns seen in Figures 3 and 4 in cells targeted on the T strand. The model illustrates the initial pathways involved in a single-bacterial cell prior to cell division, therefore only reflects plasmid replication. The magnitude and RFLP pattern of resistant clones generated by gene editing directed by the strand most likely follows the pathway that generates the population of plasmids illustrated in route (A). This pathway is dominated by incorporation of the T-ODN into the newly synthesized strand during plasmid replication, resulting in one mutant plasmid and one plasmid with a C–C mismatch. This intermediate will be susceptible to kanamycin selection since neither strand has the converted base to 'enable' antibiotic resistance. After another round(s) of plasmid replication the C–C mismatch plasmid will produce a C–G converted plasmid, which will give the cell kanamycin-resistant phenotype to produce a resistant colony. These events result in a mixed population with more mutant plasmids than converted plasmids in the plasmid population exhibiting the (A) RFLP. The other RFLP Patterns (B, C and D) are represented but in much lower numbers. Pattern D must arise from two early conversion events prior to plasmid replication since all members of that population contain fully converted strands. This is clearly true when the NT-ODN is used, however, there may be a role for DNA replication in generating (D)-type colonies when the T-ODN is used (see below).

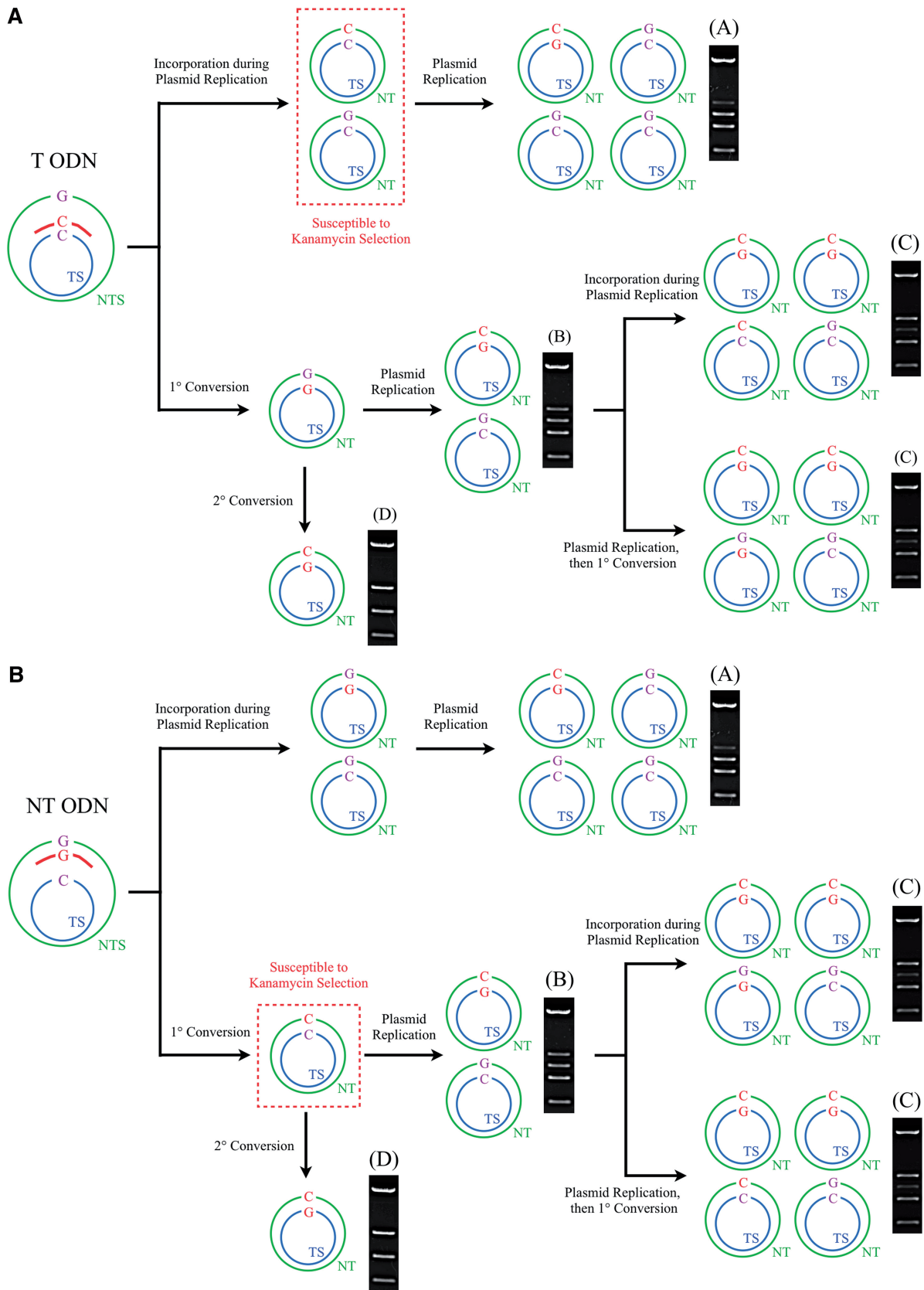


Figure 6. Proposed pathway models of resolving the gene repair reaction to produce each of the four RFLP patterns. **(A)** Pathway for resolving gene repair in Kan49T-treated cells. **(B)** Pathway for resolving gene repair in Kan49NT-treated cells. Each grouping represents the plasmid populations within a single-bacteria cell. Only plasmid replication is depicted, not bacterial replication. (See ‘Discussion’ section for detailed explanation).

Pattern B arises from the replication of a dual population of plasmids that are generated by plasmid replication after a primary conversion event and Pattern C arises from an additional primary conversion or incorporation via replication event of one of the mutant plasmids present in the population diagramed in (B). Thus, under these conditions, the repair or non-incorporation pathway of gene editing is in fact operative when T strand annealing takes place, albeit at much lower levels.

When we developed model pathways for corrected plasmids generated from NT strand gene editing (Figure 6B), there were additional experimental observations to consider. First, the number of kan^R colonies produced by reactions involving the NT strand is much lower (100-fold). Second, the NT strand serves as the leading strand of replication in the target plasmid used in our systems. Thus, our results remain consistent with our own work and others (18,20,21,24–26) in that the lagging strand is more amenable to gene-editing activity than the leading strand of replication. We observe very low correction efficiency despite the fact that our model predicts that the first intermediate in the incorporation via plasmid replication pathway resulting in an (A) RFLP should acquire antibiotic resistance from the presence of the converted T strand. While the population of plasmids with RFLP patterns depicted in (B) and (C) are present, the most obvious result is that plasmid populations bearing the (A) and (D) RFLP patterns appear at approximately equal frequencies. This result differs from the (A)/(D) ratio observed when ODNs that anneal to the T strand are used (Figure 4C), which is slanted to the (A) RFLP type. Thus, it appears that the leading strand, in this case also the NT strand, engages gene editing events via the replication repair pathway are somewhat impeded. The data confirm that primary and secondary conversion events are fully operational. Hence, strand bias not only reflects the extent of gene editing activity, but also reflects the pathway used to achieve the altered phenotype. Our results suggest that plasmid replication occurred prior to bacterial replication and there was equal plasmid segregation in subsequent daughter cells, since there were not significant changes to RFLP frequencies over the time course.

Once our model was proposed, we challenged our hypothesis by manipulating the pathways of gene repair. One of the proposed pathways that has recently gained strong support as a mechanism for gene repair is the incorporation pathway, by which the ODN becomes part of the newly synthesized DNA strand. In support for this pathway, Radecke *et al.* (17) has demonstrated the incorporation of a labeled ODN with extensible 3' hydroxyl ends into the genome of transfected and gene converted cells. In addition, cell cycle manipulation correlated to ODN-mediated gene repair events, suggest that cells progressing through S-phase have the highest levels of gene repair events (14–16,20,21,32). Therefore, we utilized ODNs with a 3'-end inverted nucleotide to block the DNA polymerase from extending into the 3'-end of the ODN, thereby incorporating the ODN into the DNA. Our model shows that the majority of the kanamycin-resistant colonies resulting from cells treated with the Kan49T

ODN are generated by incorporation during replication. The model would therefore predict that the inhibition of the replication pathway via addition of the Kan49T/3'InvdT ODN would lead to a dramatic reduction in the correction efficiency when compared to the Kan49T ODN. This result was confirmed since addition of the 3'-end inverted nucleotide to the Kan49T/3'InvdT ODN resulted in a correction efficiencies of 6.2 ± 4.5 , an 8-fold decrease from the Kan49T correction efficiency of 50.1 ± 6.7 ($P = 0.00002$) (Figure 5A).

Our model predicts a change in the (A) RFLP frequencies observed in the kanamycin-resistant colony populations resulting from cells treated with the Kan49NT, Kan49T ODN and the 3'-end inverted NT and T ODNs. Since the (A) pattern results from a gene editing event via the replication pathway, our model would predict a decrease in the frequency of the (A) RFLP with the transfection of the 3'-end inverted ODNs. This prediction was tested with the addition of the 3'-end inverted nucleotide to both the T and NT ODNs, insofar that there was a reduction in the predicted number of kanamycin-resistant colonies in the population with an (A) pattern from 2066 ± 565 kan^R49T colonies to 23 ± 46 kan^R49T/3'InvdT colonies (Figure 5C). This trend was also observed when employing the NT ODNs. Although there was no reduction in the correction efficiency with the addition of the 3'-end inverted nucleotide to the NT ODN, there was a decrease in the frequency of the (A) pattern in the kan^R49NT/3'InvdT colony population from 12 ± 3 kan^R49NT colonies to 1 ± 1 kan^R49NT/3'InvdT colonies. 'With regard to the action of the NT-ODN' there were no observed significant differences ($P < 0.01$) in the (B), (C), and (D) RFLP kan^R-type colonies within the total kanamycin-resistant population after the addition of the 3'-end inverted nucleotide ODNs. These results may suggest that only the replication pathway was inhibited and there was no increase in the preference for the conversion pathway for the NT-ODN. The results presented in Figure 5C are intriguing in that the population of (D)-type colonies are also affected to some degree by kan^R 49 T/3'InvdT. These data suggested that when the ODN that anneals to the lagging strand is used, DNA replication activity may also impact the creation of the (D)-type colonies.

In *E. coli*, gene editing has been accomplished using a technique known as 'recombineering' (27–29). Single- or double-stranded linear fragments (or ODNs) are introduced into the cell and direct alterations in the chromosome. When ssODNs are used, the frequency is enhanced when λ - β protein is expressed in the cell (30). Consistent with our results herein, the impact strand bias on the mechanism of mutagenesis has been observed. The lagging strand targets at higher levels where regions between Okazaki fragments provide a site for ssODN hybridization. These observations are also consistent with work done in eukaryotic cells wherein a recent model suggests a three branched pathway for possible routes of correction (10). The major route is replication-based incorporation of the ssODN into the new synthesized strand; several groups have now provided strong evidence to support this model (17,22,31). What often

escapes examination is the observation that targeting the leading strand (in our work, NT) engages an alternative pathway of conversion most likely directed by mismatch repair. Our work here demonstrates that the conversion pathway used in gene editing is, at least in part, dependent on which of the two strands is identified for change prior to plasmid amplification.

Our results also raise a cautionary note for workers employing selectable systems to evaluate mechanisms of ssODN mutagenesis or gene editing. The time at which selection is applied and correction evaluated could skew results. Independent of defining the leading or lagging strands of replication, it is important to take into consideration which ODN targets the T strand. Correction or editing of the T strand will obviously generate larger numbers of antibiotic resistant cells since the expression of the resistance factor is dependent on the conversion of the T-strand. In contrast, selecting the NT strand will initially generate a much smaller antibiotic-resistant population, yet approximately half of those colonies will have corrected homozygous genotypes. Thus, taken together, the ‘strand bias’ effect of gene editing appears to have downstream implications for measurements of repair frequencies and can impact the pathway used for the editing event.

FUNDING

Funding for open access charge: NIH 7R01CA089325-08.

Conflict of interest statement. None declared.

REFERENCES

- Aarts,M., Dekker,M., de Vries,S., va der wal,A. and Te Riele,H. (2006) Gene of a mouse mutant by oligonucleotide-mediated gene modification in ES cells. *Nucleic Acids Res.*, **34**, e147.
- Dekker,M., Brouwers,C., Aarts,M., van der Torre,J., de Vries,S., van de Vrugt,H. and te Riele,H. (2006) Effective oligonucleotide-mediated gene disruption in ES cells lacking the mismatch repair protein MSH3. *Gene Ther.*, **13**, 686–694.
- Morozov,V. and Wawrousek,E.F. (2008) Single-strand DNA-mediated targeted mutagenesis of genomic DNA in early mouse embryos is stimulated by Rad51/54 and by Ku70/86 inhibition. *Gene Ther.*, **15**, 468–472.
- Pierce,E.A., Liu,Q., Igoucheva,O., Omarrudin,R., Ma,H., Diamond,S.L. and Yoon,K. (2003) Oligonucleotide-directed single-base DNA alterations in mouse embryonic stem cells. *Gene Ther.*, **10**, 24–33.
- Parekh-Olmedo,H., Ferrara,L., Brachman,E. and Kmiec,E.B. (2005) Gene therapy progress and prospects: targeted gene repair. *Gene Ther.*, **12**, 639–646.
- Bertoni,C. (2005) Oligonucleotide-mediated gene editing for neuromuscular disorders. *Acta Myol.*, **3**, 194–201.
- Bertoni,C., Morris,G.E. and Rando,T.A. (2005) Strand bias in oligonucleotide-mediated dystrophin gene editing. *Hum. Mol. Genet.*, **14**, 221–233.
- Andrieu-Soler,C., Halhal,M., Boatright,J.H., Pandove,S.A., Nickerson,J.M., Stodulkova,E., Stewart,R.E., Ciavatta,V.T., Doat,M., Jeanny,J.C. *et al.* (2007) Single-stranded oligonucleotide-mediated in vivo gene repair in rd1 retina. *Mol. Vis.*, **13**, 692–706.
- Engstrom,J.U., Suzuki,T. and Kmiec,E.B. (2009) Regulation of targeted gene repair by intrinsic cellular processes. *BioEssays*, **31**, 159–168.
- Parekh-Olmedo,H. and Kmiec,E.B. (2007) Progress and prospects: targeted gene alteration (TGA). *Gene Ther.*, **14**, 1675–1680.
- Igoucheva,O., Alexeev,V., Scharer,O. and Yoon,K. (2006) Involvement of ERCC1/XPF and XPG in oligodeoxynucleotide-directed gene modification. *Oligonucleotides*, **16**, 94–104.
- Igoucheva,O., Alexeev,V., Pryce,M. and Yoon,K. (2003) Transcription affects formation and processing of intermediates in oligonucleotide-mediated gene alteration. *Nucleic Acids Res.*, **31**, 2659–2670.
- Wu,X.S., Xin,L., Yin,W.X., Shang,X.Y., Lu,L., Watt,R.M., Cheah,K.S., Huang,J.D., Liu,D.P. and Liang,C.C. (2005) Increase efficiency of oligonucleotide-mediated gene repair through slow replication fork progression. *Proc. Natl Acad. Sci. USA*, **102**, 2508–2513.
- Olsen,P.A., Randol,M. and Krauss,S. (2005) Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides. *Gene Ther.*, **12**, 546–551.
- Engstrom,J.U. and Kmiec,E.B. (2008) DNA Replication, cell cycle progression and the targeted gene repair reaction. *Cell Cycle*, **7**, 1402–1414.
- Brachman,E.E. and Kmiec,E.B. (2005) Gene repair in mammalian cells is stimulated by the elongation of S phase and transient stalling of replication forks. *DNA Repair*, **4**, 445–457.
- Radecke,S., Radecke,F., Peter,I. and Schwarz,K. (2006) Physical incorporation of a single-stranded oligodeoxynucleotide during targeted repair of a human chromosomal locus. *J. Gene Med.*, **8**, 217–228.
- Brachman,E.E. and Kmiec,E.B. (2004) DNA replication and transcription direct a DNA strand bias in the process of targeted gene repair in mammalian cells. *J. Cell Sci.*, **117**, 3867–3874.
- Brachman,E.E. and Kmiec,E.B. (2002) The ‘biased’ evolution of targeted gene repair. *Curr. Opin. Mol. Ther.*, **4**, 171–176.
- Olsen,P.A., Randol,M., Luna,L., Brown,T. and Krauss,S. (2005) Genomic sequence correction by single-stranded DNA oligonucleotides role of DNA synthesis and chemical modifications of the oligonucleotide ends. *J. Gene Med.*, **7**, 1534–1544.
- Wu,X.S., Xin,L., Yin,W.X., Shang,X.Y., Lu,L., Watt,R.M., Kathryn,S.E.C., Huang,J.D., Liu,D.P. and Liang,C.C. (2005) Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. *Proc. Natl Acad. Sci. USA*, **102**, 2508–2513.
- Sorensen,C.B., Krogsdam,A.M., Andersen,M.S., Kristiansen,K., Bolund,L. and Jensen,T.G. (2005) Site-specific strand bias in gene correction using single-stranded oligonucleotides. *J. Mol. Med.*, **83**, 39–49.
- Liu,L., Rice,M.C., Drury,M., Cheng,S., Gamper,H. and Kmiec,E.B. (2002) Strand bias in targeted gene repair is influenced by transcriptional activity. *Mol. Cell Biol.*, **22**, 3852–3863.
- Papaloannou,I., Disterer,P. and Owen,J. (2009) Use of internally nuclease-protected single-strand DNA oligonucleotides and silencing of the mismatch repair protein, MSH₂ enhances the replication of corrected cells following gene editing. *J. Gene Med.*, **11**, 267–274.
- Olsen,P.A., McKeen,C. and Krauss,S. (2003) Branched oligonucleotides induce in vivo gene conversion of a mutated EGFP reporter. *Gene Ther.*, **10**, 1830–1840.
- Andrieu-Soler,C., Casas,M., Faussat,A.M., Gandolphe,C., Doat,M., Tempé,D., Giovannangeli,C., Cohen,F.B. and Concordet,J.P. (2005) Stable transmission of targeted gene modification using single-stranded oligonucleotides with flanking LNAs. *Nucleic Acids Res.*, **33**, 3733–3742.
- Sharan,S.K., Thomason,L.C., Kuznetkov,S.G. and Court,D.L. (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.*, **4**, 206–223.
- Sawitzke,J.A., Thomason,L.C., Costantino,N., Bubunenkov,M., Datta,S. and Court,D.L. (2007) Recombineering in vivo genetic engineering in E. coli, S. enteric, and beyond. *Method. Enzymol.*, **421**, 171–199.
- Court,D.L., Sawitzke,J.A. and Thomason,L.C. (2002) Genetic engineering using homologous recombination. *Annu. Rev. Genet.*, **36**, 361–388.

30. Murphy, B.R., Moayedpardazi, H.S., Gewirtz, A.M., Diamond, S.L. and Pierce, E.A. (2008) Delivery and mechanistic considerations for the production of knock-in mice by single-stranded oligonucleotide gene targeting. *Gene Ther.*, **14**, 304–315.
31. Huen, M.S., Li, X.T., Lu, L.Y., Watt, R.M., Liu, D.P. and Huang, J.D. (2006) The involvement of replication in single-stranded oligonucleotide-mediated gene repair. *Nucleic Acids Res.*, **34**, 6183–6194.
32. Ferrara, L. and Kmiec, E.B. (2004) Camptothecin enhances the frequency of oligonucleotide-directed gene repair in mammalian cells by inducing DNA damage and activating homologous recombination. *Nucleic Acids Res.*, **32**, 5239–5248.