Published online 4 January 2017

Nucleic Acids Research, 2017, Vol. 45, No. 8 ***e63***
doi: 10.1093/nar/gkw1339

A 'Semi-Protected Oligonucleotide Recombination' Assay for DNA Mismatch Repair *in vivo* Suggests Different Modes of Repair for Lagging Strand Mismatches

Eric A. Josephs* **and Piotr E. Marszalek***

Department of Mechanical Engineering and Materials Science, Edmund T. Pratt, Jr. School of Engineering, Duke University, Durham, NC, USA

Received November 23, 2016; Revised December 16, 2016; Editorial Decision December 19, 2016; Accepted December 20, 2016

ABSTRACT

In *Escherichia coli*, a DNA mismatch repair (MMR) pathway corrects errors that occur during DNA replication by coordinating the excision and re-synthesis of a long tract of the newly-replicated DNA between an epigenetic signal (a hemi-methylated d(GATC) site or a single-stranded nick) and the replication error after the error is identified by protein MutS. Recent observations suggest that this 'long-patch repair' between these sites is coordinated in the same direction of replication by the replisome. Here, we have developed a new assay that uniquely allows us to introduce targeted 'mismatches' directly into the replication fork via oligonucleotide recombination, examine the directionality of MMR, and quantify the nucleotide-dependence, sequence context-dependence, and strand-dependence of their repair *in vivo*—something otherwise nearly impossible to achieve. We find that repair of genomic lagging strand mismatches occurs bi-directionally in *E. coli* and that, while all MutS-recognized mismatches had been thought to be repaired in a consistent manner, the directional bias of repair and the effects of mutations in MutS are dependent on the molecular species of the mismatch. Because oligonucleotide recombination is routinely performed in both prokaryotic and eukaryotic cells, we expect this assay will be broadly applicable for investigating mechanisms of MMR *in vivo*.

INTRODUCTION

DNA mismatch repair (MMR) helps to ensure genomic stability by repairing incorrectly paired nucleotides (such as a G to T or A to C) or tracts of inadvertent nucleotide insertions/deletions that occur during replication (1,2). This repair is orchestrated through a pathway that increases replication fidelity 100-fold and whose components are highly conserved from *Escherichia coli* through humans. In the methyl-directed mismatch repair pathway in *E. coli* (3,4), these replication errors are identified by a homodimer of MutS which, with protein MutL, activates a latent nicking endonuclease MutH. As the *E. coli* genome is methylated at d(GATC) sites by dam methylase, there is a brief window of time immediately after replication (before these sites are fully-methylated) where MutH can initiate repair by nicking the DNA at the nearest d(GATC) site (5) on the un-methylated and, hence newly-replicated, strand (6,7)—although a pre-formed single-stranded break in the DNA has also been found to be sufficient to initiate repair in the absence of MutH (8). At the site of the nick or single-strand break, helicase UvrD is loaded by MutL back toward the replication error and, with the appropriate 5′-to-3′ or 3′-to-5′ exonucleases, the strand of DNA between the nick and the error is digested and re-synthesized. This excision/re-synthesis that occurs in MMR is termed 'long-patch repair,' as the distance between the d(GATC) and the mismatch can be separated by hundreds of base-pairs while still promoting efficient repair (7).

While the key biochemical components of MMR have long since been identified, there remains significant dispute over the mechanisms by which (i) the epigenetic strand-discrimination signal (a hemi-methylated d(GATC) site or a pre-formed nick) can be rapidly found after an error has been identified and (ii) how repair can then be coordinated back between the two sites over potentially large stretches of DNA (9). A confounding factor in elucidating these mechanisms has been the diversity of behaviors observed when MutS binds to mismatched sites. ADP-bound MutS dimers have been observed to undergo an ADP-ATP exchange after binding to a mismatch (10). A few seconds after this ex-

*To whom correspondence should be addressed. Tel: +19196605474; Email: eaj20@duke.edu
Correspondence may also be addressed to Piotr E. Marszalek. Tel: +19196605381; Fax: +1916608963; Email: pemar@duke.edu

change, MutS undergoes a conformational transition to a 'sliding clamp,' a long-lived structure that diffuses randomly along the DNA (11–13) and which itself can recruit and form transient complexes with rapidly-moving MutL(H) sliding clamps that can diffuse on both sides of the MutS (14). It remains unclear, during these cascading diffusion events, how the relative location of the mismatch is retained so that excision may be efficiency directed back toward that site. Alternatively, a tetrameric form of MutS (15), its predominate state in solution, has been observed to form loops in heteroduplex DNA molecules (16,17), with complexes of MutSL(H) recently being observed having a propensity to form intra-strand loops that directly bridge mismatches and the sites of hemi-methylated d(GATC) or pre-formed nicks (18). Such a looping mechanism would allow for both the mismatch and strand-discrimination signal to be simultaneously bound by MutS dimers to confine DNA excision and re-synthesis effectively between the two sites, although a precise mechanism remains elusive. However, while truncation of the MutS C-terminal domain, which contains the residues required for tetramerization, imparted significant repair defects *in vitro* (15), Mendillo *et al.* (19) found that MutS tetramers were non-essential for repair *in vivo* and that mutation of the MutS tetramerization domain resulted only in a moderate mutator phenotype during a spontaneous rifampicin resistance assay (Rif assay). Lastly, Hasan and Leach (20) recently used an unstable trinucleotide repeat (TNR) array on the *E. coli* genome to measure 'single-unit instability,' a quantitative measure of the ability of the MMR system to correct three-nucleotide insertion/deletion loops which frequently occur during replication of long CTG·CAG tracts. They found that frequency of single-unit instability was inversely correlated with the distance of the nearest genomic d(GATC) in the direction the replication fork moves during DNA replication (away from the origin of replication), but uncorrelated with the distance of the nearest d(GATC) site on the opposite side. This result is suggestive that MMR protein complexes identify d(GATC) sites through its association with the replisome during replication, although the DNA replication machinery is often not present in *in vitro* experiments and not necessary for MMR to occur *in vitro* (21).

As can be seen above, there remain major difficulties in relating the biochemistry of MMR to *in vivo* experimental systems. These challenges stem from the fact that, in general, cellular assays to deconstruct MMR *in vivo* must rely on rare, (approximately) random errors that occur during replication, or that MMR efficiency must often be evaluated indirectly through spontaneous appearance of a reporter phenotype able to survive a screening process. These assays, however reproducible, also tend to be semi-quantitative at best. Here, we have developed an assay (Figure 1) that allows us to directly evaluate and quantify MMR efficiency *in vivo* in a nucleotide-, sequence context-, strand-, direction-, and chromosomal context- / orientation- dependent manner. The assay is based on a variation of the genomic engineering technique known as 'oligonucleotide recombination' (22) and allows us to introduce targeted 'mismatches' into the genome in a process that is known to interact directly with the MMR pathway at the replication fork in both prokaryotic (23,24) and eukaryotic cells (25,26). This assay reveals a number of new insights which would be nearly impossible to resolve with any other method. First, in *E. coli* we find that repair of lagging strand G-T mismatches and T-T mismatches differs in both the directional bias of repair and the effect of MutS mutations on this directional bias: while lagging strand T–T mismatches are repaired weakly but almost exclusively from its 3′- end (the direction of replication), lagging strand G–T mismatches are effectively repaired bi-directionally, from both its 5′- and 3′- ends. In strains which possess a mutation in MutS that impairs its ability to tetramerize but not its ability to dimerize (27), long-patch repair directed from the 3′- end is substantially reduced, with no significant effect on 5′- coordinated repair. While further validation with other mismatches will be necessary, the heretofore unobserved differences between repair of lagging strand G–T and T–T mismatches in directional origin of repair and the effects of mutations on MMR proteins suggests that lagging strand MMR may be coordinated *in vivo* by different modes that depend on the molecular species of the mismatch. Furthermore, using this assay to probe the subtler effects of sequence context of those mismatches we find that G-T mismatches are repaired slightly but significantly more efficiently *in vivo* when the mismatch is flanked by purine nucleotides than when flanked by pyrimidine nucleotides, which is consistent with *in vitro* studies of human MutS homologue (hMSH) activity with DNA mismatches in analogous sequence contexts (28). As can be seen from these demonstrations, this assay provides a newfound ability to directly quantify and probe MMR *in vivo* in a profoundly more controlled way than spontaneous phenotypic reporter assays have in the past, and represents a powerful new way to deconstruct the mechanistic aspects of the complex MMR pathway that have so far remained elusive.

MATERIALS AND METHODS

Materials

Escherichia coli strains SIMD50 (W3110 $galK_{tyr145UAG}$ *ΔlacU169* [λ cI857 *Δ(cro-bioA) (int-cIII<>bet)*]) and SIMD90 (SIMD50 *mutS<>cat*) were obtained as a generous gift of the laboratory of Don Court (National Cancer Institute, Frederick, MD, USA). M9 minimal salts (5x) were obtained from Sigma-Aldrich Co. M63 galactose-selective media (3% KH_2PO_4 w/w, 7% K_2HPO_4 w/w, 2% $(NH_4)SO_4$ w/w, 2% D-galactose w/w, 1 mM $MgSO_4$, 0.5 mg/L $FeSO_4$, 1 mg/L D-biotin) was prepared as previously described (29). Taq 2X MasterMix was obtained by New England Biolabs (Ipswich, MA, USA) and used for all PCR reactions. Gene Pulser(R)/MicroPulser(tm) Electroporation Cuvettes, 0.1 cm gap were obtained from Bio-Rad Laboratories. Oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA, USA) with standard desalting and used without further purification.

Generation of E. coli *strain variants*

Strain SIMD50, which expresses the single-stranded DNA recombinase Beta from the λ phage in a heat-inducible manner at 42°C (30), was transformed using Red-mediated oligo-mediated recombination and screened

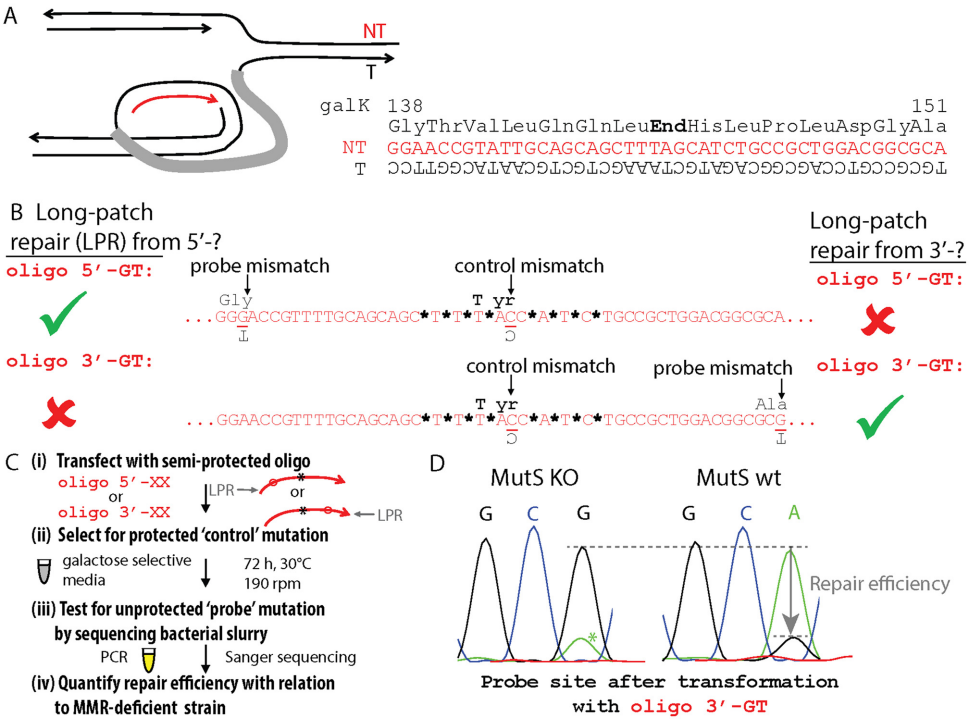


Figure 1. A 'semi-protected oligonucleotide recombination' (SPORE) assay to quantify mismatch repair (MMR) efficiency *in vivo* in a nucleotide-, sequence-context-, strand-, direction- and chromosomal context-/orientation-dependent manner. (**A**) (left) In the SPORE assay presented here, a synthetic oligonucleotide (oligo, red) with significant homology to non-template strand (NT) of galactose kinase gene *galK* is designed to hybridize with the lagging strand during replication (22). (right) In the *E. coli* strains used, the oligo is designed to target the region surrounding an amber mutation. (**B**) Example segments of two of the 70-nucleotide-long synthetic oligos used in the SPORE assay. See text for details. Oligos are designed to possess (i) MMR-inactive 'control' mismatch designed to correct the amber mutation after the oligo is incorporated into the genome at the replication fork and (ii) a MMR-reactive 'probe' mismatch to one side of the control mismatch that introduces a silent mutation. Phosphorothioate bonds (*), which flank the control mismatch, block long-patch repair of the probe mismatch from the opposite end. (**C**) Simplified protocol of the SPORE assay. See text and Experimental Procedures for details. (**D**) Quantification of repair efficiencies is obtained by comparing the decrease in the sequencing signal at the probe mutation site relative to that of a SPORE assay using a MMR-deficient (MutS KO) strain, after selecting for the control mutation by ability to metabolize galactose. See also Supplementary Figure S1 for example chromatograms.

according to the standard protocol (31,32) using oligos: 5′-CTGCGGTCGAAGCTCTGGAAAATCTTGATCCCCGGTCACTCACCCCGCGTCAGGCGCTGG-3′ (underline indicates mismatched nucleotides), and 5′-CGACGCCCATACGCCCATGATGGCTGCTTATGCTGCTCTGAAAGCCCAGCATCCCGAGAT-3′, for MutS D835R and MutS ^{15}AAYAAL^{20} mutations, respectively. Briefly, using sterile technique, bacterial colonies of SIMD50 grown on Luria broth (LB) agar plates were picked and grown in 5 mL LB overnight at 30°C with shaking (190 rpm). 0.5 mL of the growth solution was then added to 17 mL of LB in 50 mL centrifuge tubes and grown for 2 h at 30°C with shaking. The tubes were heat shocked at 42°C in a water bath for 15 min with agitation then immediately cooled in ice water for 5 min. The tubes were then spun in a centrifuge at 6500×g for 7 min at 4°C and the LB gently decanted. Bacterial pellets were re-suspended in 1 mL of pure water followed by an additional 30 mL of water, then spun again at 6500xg for 7 min at 4°C. The tubes were immediately removed and the supernatant gently removed with a pipette, and the bacterial pellets were then re-suspended in 1 mL of water and spun for 30 s at 13 500×g in a chilled 1.5 mL falcon tube using a desktop centrifuge at 4°C. The supernatant was removed by pipetting and the pellets re-suspended into 1200 μL of 15% glycerol and stored at −80°C in 300 μL aliquots until use or used fresh by resuspension in pure ice-cold water. 50 μL of electrocompetent bacteria were thawed on ice and gently mixed with 2 μL of 100 μM in H_2O of one of the oligos described above. The mixtures were electroporated at 1.8 kV using a GenePulser Xcell electroporation system (Bio-Rad), then immediately mixed with 1 mL room-temperature LB and grown for 30 min at 30°C with shaking. After 30 min, the 100 μL was plated on LB–agar plates and grown overnight at 30°C.

Bacterial colonies were picked, spotted on a labeled LB-agar plate, and screened for MutS ^{15}AAYAAL^{20} mutations by ability to initiate a PCR reaction using primers 5′-GCTGCTTATGCTGCT-3′ and 5′-AAACCTTTGCTGTCCTG-3′, or screened for MutS D835R by testing a PCR product enriched using primers 5′-AGCCACATATTGCCATC-3′ and 5′-ATAACGCCACCGAATAC-3′ for ability to be digested by restriction endonuclease BglII (New England Biolabs). Successfully screened colonies were spread again on LB-agar plates, grown overnight at 30°C, screened a second time to obtain an isogenic colony. Mutations were confirmed by Sanger sequencing a PCR-amplified segment of DNA coding of the N- or C-terminus of *mutS*.

'*Semi-protected oligonucleotide recombination*' *(SPORE) assay* (Figure 1C)

Designs of synthetic oligonucleotides for SPORE assays were derived from oligo 144 in (24) to target the lagging strand at the *galK* gene (Figure 1). Synthetic oligos were designed to contain a (MMR-inactive) 'control' C–C mismatch which corrects an amber mutation in the *galK* gene of SIMD50 (and derivatives) and allows successful transformants to metabolize galactose, and a (MMR-active) 'probe' mismatch that introduces a silent mutation into the *galK* gene and is located approximate 20 nt away from the 'control' mismatch on either its 5′- or 3′- side. See 'Results' for extended discussion of the design of oligonucleotides for SPORE. The 'control' mismatch is designed to be flanked by phosphorothioate bonds that block exonuclease activity (33) (and hence, long patch repair).

Escherichia coli strains were made electrocompetent and Beta was induced following the procedure described above. Although induction of Beta was not strictly necessary for the SPORE assay (data not shown), we observed that it increased transformation efficiency of the oligonucleotides and hence the ratio of 'signal' to 'noise' from the background, untransformed population; see discussion. 50 μL of electrocompetent bacteria were thawed on ice and gently mixed with 2 μL of 100 μM in H_2O of one of the following oligos to investigate the effects of molecular species of mismatch (mismatched nucleotide underlined):

oligo 5′-GT: 5′-AGTTCTTCCGCTTCACTGGAAGTCGCGGTCGGGACCGTATTGCAGCAG*C*T*T*T*TACCA*T*C*T*G*CCGCTGGACG-3′, where N* indicates a phosphorothioated DNA base

oligo 3′-GT: 5′-GTCGCGGTCGGAACCGTATTGCAGCAG*C*T*T*TACCA*T*C*T*GCCGCTGGACGGCGCGCAAATCGCGCTTAACG-3′

oligo 5′-TT: 5′-AGTTCTTCCGCTTCACTGGAAGTCGCGGTCGGTACCGTATTGCAGCAG*C*T*T*T*TACCA*T*C*T*G*CCGCTGGACG-3′

oligo 3′-TT: 5′-GTCGCGGTCGGAACCGTATTGCAGCAG*C*T*T*TACCA*T*C*T*GCCGCTGGACGGCGCTCAAATCGCGCTTAACG-3′

To investigate the effects of sequence context of the mismatched nucleotides, oligo 5′-GT (with its mismatched G flanked by purines) and oligo 3′-GT (with its mismatched G flanked by pyrimidines) were compared with:

oligo 5′-GT2: 5′-AGTTCTTCCGCTTCACTGGAAGTCGCGGTCGGAACCGTGTTGCAGCAG*C*T*T*TACCA*T*C*T*G*CCGCTGGACG-3′

oligo 3′-GT2: 5′-GTCGCGGTCGGAACCGTATTGCAGCAG*C*T*T*TACCA*T*C*T*GCCGCTGGACGGCGCACAGATCGCGCTTAACG-3′

The mixtures were electroporated at 1.8 kV using a GenePulser Xcell electroporation system, then immediately mixed with 1 mL room-temperature LB and grown for 30 min at 30°C with shaking. After 30 min, the bacteria were spun down at 13 500×g for 15 s, the medium was decanted, then the bacteria washed in 1 mL of M9 minimal media and spun down again. After decanting, the bacteria were re-suspended in 1000 μL of M63 media, divided into two samples of 500 μL in 1.5 mL centrifuge tubes, and incubated at 30°C with shaking for 72 h. The bacteria were then spun down at 13 500×g for 3 min, and re-suspended in 20 μL 25% glycerol and stored at −80°C. 2 μL of thawed bacterial stocks were then used directly to PCR a segment of the *galK* gene using Taq polymerase in 40 μL reactions using primers 5′-ACAATCTCTGTTTGCCAACG-3′ and 5′-GGCTGGCTGCTGGAAG-3′. The reaction mixture was then sent for purification and Sanger sequencing by Eton Biosciences at its North Carolina branch (Durham, NC) using sequencing primer 5′-ACAATCTCTGTTTGCCAACG-3′.

Quantitative Sanger sequencing analysis of SPORE assays

Raw chromatogram data from the sequencing reads (Supplementary Figure S1) were imported into MATLAB (MathWorks, Inc; Natick, MA), where peak heights for each nucleotide signal were algorithmically extracted, the sequence determined from the maximum signal at each 'peak,' and the second highest 'peak' of the second strongest signal within those called peaks also extracted. In the SPORE assay, we wish to identify the fraction of the population of 'probe' mismatches that were repaired by MMR in the population of cells which was successfully transformed by the oligonucleotide (which we verify by the presence of the 'control' mismatch, allowing the bacteria to metabolize galactose in this case). This is performed by quantifying the relative drop of the 'probe' mutation when the SPORE assay is performed in an experimental strain vs. when performed in a MutS KO strain (SIMD90) (Figure 1D).

However, to quantify the fraction of the population with probe mismatches that were repaired from the Sanger sequencing chromatograms, a normalization procedure is necessary not only because the signal strength of each nucleotide varies slightly according to a normal distribution, but also because the signal of at one nucleotide position may affect the relative signal strength of nearby nucleotides (34). This correlation especially presents a challenge for the assay, where we introduce a targeted 'probe' mutation that we will expect to have different signal strengths as a result of varying repair efficiencies. To normalize the experimental chromatogram data, we first normalized the raw chromatograms of the MutS KO experiments with respect to the signal strength of the 'C' signal strength at the 'control' mutation site. The most 'stable' peak (peak with the smallest variance across the MutS KO samples) located >20 nt outside of the locations of either probe site was identified (stds. of <0.01 for G–T and T–T mismatches, respectively). This distant site is not expected to have any correlation with any changes in the probe signal and be robust across all the data sets. The signal strength at that stable site was used to normalize all other raw experimental chromatograms by dividing by their signal strengths by the strength of the signal at those stable sites.

MutS KO chromatograms were only used in the initial normalization if they satisfied signal-to-noise and positive-selection criteria: (i) that the G/C signal strength ratio at the control mutation site was <0.1 (spurious G signals at the site of the control mutation indicate a background, untransformed population that can artifactually increase the apparent repair efficiency); (ii) mean signal strength of the

second strongest peak in each nucleotide read between the two probe mutation sites was <0.2 and (iii) that the G signal strength at the control mutation was less than the mean background signal strength of (ii). This normalization process resulted in highly robust data to compare the efficiencies of repair at the probe sites. Repair efficiencies were determined as follows: first we derived the mutational efficiency of the oligo $ME = P_{X,ex}/<P_{X,KO}>$ where is $P_{X,ex}$ the normalized signal strength of the probe mutation site of the channel of the mutation nucleotide (X = G or T) for each experimental run, and $<P_{X,KO}>$ is the mean signal strength of the probe mutation site of the channel of the mutation nucleotide for the MutS KO runs. Repair efficiency RE was defined as $RE = (1 - ME - <RE_{KO}>)/(1 - <RE_{KO}>)$—essentially as $RE \approx 1 - ME$, with the remaining terms as minor corrections ($<RE_{KO}>$ as the mean repair efficiency of the relevant MutS KO experiments) to account for any offset of the mean MutS KO results and set the mean apparent MMR efficiency of the MutS KO strain to 0%.

Experimental chromatogram data were then used provided they satisfied signal-to-noise and positive-selection criteria and a fourth criterion: (iv) the apparent 'repair efficiency' at the probe site that was not tested (i.e. the 5′-probe site when reviewing data after an oligo transfection targeting a 3′-probe site) was within 20% of 1 (i.e. 'full repair'). Data were compared for statistically significant differences and 95% confidence in their effect sizes using two-sided *t*-tests. Data and statistical tests for all experiments which passed the signal-to-noise and positive-selection criteria are in the Supplemental Information.

RESULTS

A 'semi-protected oligonucleotide recombination' (SPORE) assay quantifies long-patch repair efficiency in a nucleotide-, strand- and directionality-dependent manner

In our assay, which we term a 'semi-protected oligonucleotide recombination' (SPORE) assay (Figure 1A–C), we introduce targeted 'replication-errors' directly into the *E. coli* chromosome during replication by building on traditional oligonucleotide recombination techniques (26,35,36). Oligonucleotide recombination is a genomic engineering technique where synthetic single-stranded oligonucleotides (oligos), which contain 50–90 nt nucleotides (nt) that are complementary to a segment of chromosomal DNA but also are designed to flank one or more mismatched nt, are transfected into a cell. These oligos can then become incorporated into the genome at a low frequency ($\sim 10^{-7}$ to 10^{-5} per transfected cell) (36,37). Red-mediated oligonucleotide recombination (22), one of the best studied oligonucleotide recombination techniques where transfection of the oligo is accompanied by the expression of the single-stranded DNA recombinase Beta from the λ phage, increases the frequency of oligonucleotide incorporation to approximately 10^{-5} to 10^{-3} per transfected cell. There is substantial evidence that, during Red-mediated recombination, synthetic oligos are incorporated at the replication fork and interact with the replisome (22,38): for example, incorporation rates are dependent on whether the oligo is designed to bind to the lagging or leading strand template (22,39), studies where plasmid DNA has been targeted have shown that the plasmids must be actively replicating for oligonucleotide recombination to occur (23), and the extreme 5′- and 3′-ends of the oligos themselves are subject to digestion by the exonucleases associated with DNA polymerases (40). Furthermore, there is also substantial evidence that when bound to their chromosomal template these oligos are subject to proofreading by the DNA mismatch repair proteins: disruption of MutS, MutL, MutH, dam methylase and UvrD each enhance the probability that a cell will be successfully transformed (24,41,42); and incorporation of C–C mismatches or insertion/deletion bulges >3 nts, each of which is not repaired by MMR (43), as well as MMR-inactive artificial nucleotides that are weakly recognized by MutS (37,44), are incorporated with the same probability *in vivo* with or without a functional MMR system. There is similar evidence that oligonucleotide recombination occurs primarily at the replication fork (25,45–48) in a MMR-dependent manner (26,49–53) in eukaryotic systems as well. Therefore, oligonucleotide recombination techniques can be used to gain valuable insights into MMR *in vivo*.

In a SPORE assay (Figure 1C), a synthetic oligonucleotide is designed to contain a chemically-protected, MMR-inactive 'control' mismatch (MM) and an unprotected, MMR-active 'probe' mismatch (Figure 1B). The control mismatch is designed to produce a selectable mutation—here it introduces a C–C mismatch (which is recognized extremely weakly by MutS and not subject to MMR) that corrects an amber mutation in the galactose kinase gene *galK*, allowing successful transformants to metabolize galactose (22). To protect it from exonuclease digestion during repair, the mismatched C is flanked by phosphorothioate bonds (33), which blocks DNA excision that occurs during long-patch repair originating from the opposite side as the probe mismatch and protects the mismatched C from excision tracts that 'overshoot' the probe mismatch (54). In this study, the probe mismatch is located 20 nt away and either 5′- or 3′- of the control mismatch, respectively, and introduces a silent mutation in the *galK* gene if left unrepaired. After bacteria are transfected with the synthetic oligonucleotide and grown in selective media, PCR-enriched genomic DNA obtained directly from aliquots of the bacterial media are sequenced and the strength of the probe mutation signal compared to those obtained from a SPORE assay using a MMR-defective strain (Figure 1D). Hence, the presence of the control mismatch in the SPORE assay serves the dual role of allowing us to separate oligonucleotide recombination efficiency from mismatch repair efficiency of the probe mismatch, and allowing us to individually probe the efficiency of long-patch repair originating from each direction (5′- or 3′-).

We performed the SPORE assay using oligos that targeted the lagging strand template during replication (22,24) (Figure 1A and B). oligo 5′-GT and oligo 3′-GT (see Materials and Methods) were designed to introduced G–T mismatches, one of the most efficiently repaired mismatches that interacts very strongly with MutS (55–57), at a probe site located either 5′- or 3′- of the 'control' mismatch, respectively, while oligo 5′-TT and oligo 3′-TT instead introduced T–T mismatches, one of the least effectively repaired mismatches at the same 'probe' sites. These experiments were

performed on strains with wild-type MutS (MutS wt) and with MutS knocked-out (MutS KO). We note that there are no d(GATC) sites in the synthetic oligo, and once incorporated into the genome, the 'probe' mismatch finds itself located 63 or 58 nt away from the nearest 5′- or 3′- chromosomal d(GATC) sites, respectively. The 'probe' mismatches were located either 32 nt away from the 5′-end or 16 nt away from the 3′-end of the synthetic oligo, respectively.

The results (Figure 2 and Supplementary Figure S1) were remarkably robust—the average standard deviation of experimental repair efficiencies observed for each of the 10 SPORE assays described in that figure was 5.09% (±0.44%, SEM; see also Supplementary Tables S1–S5) and as low as 1.7% in some cases—allowing us to identify and quantify even the weak repair efficiency of T–T mismatches using Sanger sequencing. Next-generation sequencing methods will likely increase the resolution of this technique further. We also note that, while not affecting the results, we consistently observed a small signal of a 'repaired' nucleotide at the probe sites in the MutS KO strain (asterisk in Figure 1D)—this is likely the result of digestion of the extreme 5′- and 3′-ends of the oligo by exonucleases associated with the DNA polymerases. This digestion is expected to occur in 10–20% of oligos given the distance of the probe sites from the oligos' 3′-ends and less frequently for the 5′-ends (40), and indicates a direct interaction with the replisome. Note that all extracted data from the SPORE assays and statistical tests comparing experiments are located in the Supplemental Information (Supplementary Tables S1–S5; Figures S2 and S3).

A SPORE assay reveals differences in directional bias and effects of MutS mutations during lagging strand mismatch repair of G–T versus T–T mismatches

The results of the assay show that G–T mismatches are efficiently repaired by wild-type MutS from both 5′- and 3′-originated long-patch repair (Figure 2, right, and Supplementary Figure S1A and B), with repair efficiencies of 87.4% (±3.5%; 95% confidence) and 78.6% (±2.7%; 95% confidence) relative to average MutS KO repair, respectively. T–T mismatches (Figure 2, left), conversely, were repaired almost exclusively from the 3′-end (repair efficiency of 10.9 ± 2.8%; 95% confidence), and strength of the probe mutation signals for a 5′-T–T probe (4.8 ± 4.2%; 95% confidence) was not statistically different than that of the MutS KO strain ($P = 0.11$; two-sided t-test; 95% confidence; Supplementary Figure S3). We subsequently repeated this assay using a strain possessing a destabilizing mutation in *mutS* (MutS ^{15}AAYAAL^{20}), which is known to decrease intracellular MutS concentration 20–50% and decrease MutH activation by MutSL *in vitro* by 65% (59). This mutation disrupted G–T repair by ∼85% from both 5′- and 3′-ends and was sufficient to drop the repair efficiency of the T–T mismatch to below the level of our experimental sensitivity (Figures 2 and S1C). This result demonstrates that the SPORE assay is appropriate for robustly identifying and quantifying a variety of MutS mutants *in vivo*.

We sought to identify any potential role of tetrameric MutS in lagging strand repair *in vivo*, and repeated the experiment (Figure 2 and S1D) using a strain possessing a mutation in *mutS* which abolishes the ability of MutS to tetramerize but not its ability to dimerize (MutS D835R) (27). This mutation substantially disrupted repair of the G–T mismatch from the 3′-end (repair efficiency of 26.4 ± 4.6%; 95% confidence) and was sufficient to abolish repair of the T–T dimers. However, this mutation had a markedly weaker effect on repair of the G–T mismatches from the 5′-end (from 87.4% to 80.7% (±6.4%; 95% confidence)) as compared to MutS wt, which was not a statistically significant difference ($P = 0.13$; Supplementary Figure S2). Interestingly, a destabilized, tetramerization-null double mutant (MutS ^{15}AAYAAL^{20} D835R) was sufficient to completely abolish all mismatch repair of the lagging strand (Figure 2 and Supplementary Figure S1E).

A SPORE assay reveals subtle differences in repair efficiencies of lagging strand G–T mismatches based on sequence context of the mismatch

We considered the possibility that the nucleotides which immediately flank the mismatch may affect their repair efficiency *in vivo* (28,56,60). Sequence context-dependent effects have also been observed to occur during oligonucleotide recombination (37,43,49), but in the absence of an 'control' mismatch on the oligo it is difficult to de-convolve differences in oligonucleotide incorporation (which can be affected by oligonucleotide uptake (43), melting temperature, or secondary structure (61), for example) from those arising from 'true' MMR-related differences in repair. Using the SPORE assay we can probe these subtle effects directly.

We were able to identify one alternative 'probe' mismatch site at each 5′- and 3′-end of the 'control' mismatch site on the SPORE oligonucleotide that would result in a silent mutation in the *galK* gene (oligo 5′-GT2 and oligo 3′-GT2, respectively). These new designs had the effect of switching the sequence contexts of the mismatched 'probe' site from the mismatched G being flanked by purine (R) nucleotides to being flanked by pyrimidines (Y) and vice versa (from G[G]A to T[G]T for 5′-'probe' mismatches, and from C[G]C to A[G]A for 3′-'probe' mismatches, where [G] is the site of the mismatched G), with the purine or pyrimidine identity of the flanking bases being a useful heuristic identified by *in vitro* studies to describe the effects of sequence contexts on MMR efficiency (28). In that referenced report, Mazurek *et al*. measured the kinetic efficiency of the hMSH ATPase activity after incubating a hMSH with a series of DNA duplexes where the mismatch site was flanked by every possible combination of nucleotides. While they used human homologues of *E. coli* MutS, from their work we would expect a similar effect of moving the 'probe' mismatch site from being flanked by purines to pyrimidines as would decrease hMSH ATPase activity from 66.3×10^{-4} M^{-1} min^{-1} to 58.7×10^{-4} M^{-1} min^{-1} for those specific 5′-sites and from 63.3×10^{-4} M^{-1} min^{-1} to 58.3×10^{-4} M^{-1} min^{-1} for those 3′-sites.

We tested these alternative SPORE oligonucleotides *in vivo* (Figure 3). While there was no statistically significant ($P > 0.1$) differences between 5′- and 3′-repair of G–T mismatches in the same context (Y[G]Y or R[G]R), we did find a slight decrease in repair efficiency in G–T mismatches

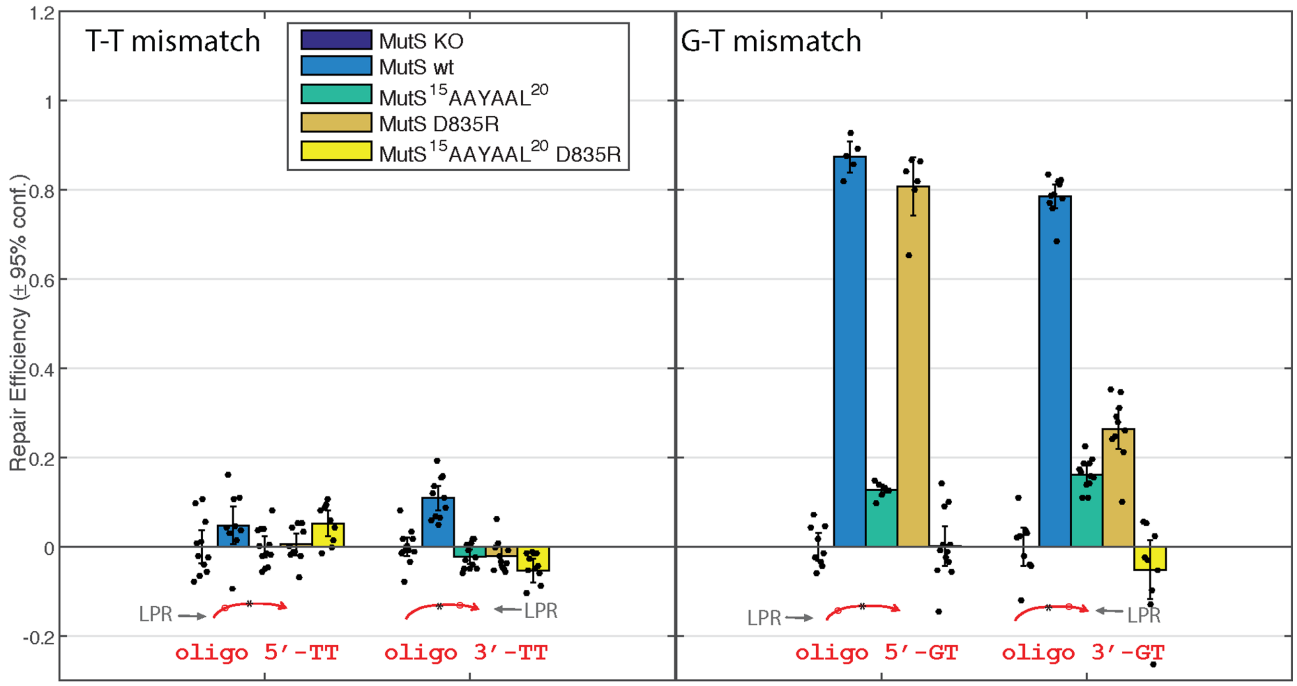


Figure 2. Repair efficiencies obtained by semi-protected oligonucleotide recombination (SPORE) assay for lagging strand repair of T–T (left panel) and G–T (right panel) mismatches. See Text and Experimental Procedures for details, with oligos showing direction of long-patch repair (LPR) allowed below. Beeswarm plot of individual experimental data points overlaid over bar graphs for repair of each strain (see text for details) tested for 5′-directed long patch repair (using oligo 5′-XT, where X is G or T; see Figure 1) or 3′-long patch repair (using oligo 3′-XT). Error bars are 95% confidence around mean (bar height) repair, and repair efficiency of 0% is defined as the mean repair efficiency of the mismatch repair defective (MutS KO) strain for each oligo. See Supplementary Tables S1–S5. MutS KO, a MMR-defective, MutS knock-out strain; MutS wt, strain with functional MMR pathway; MutS ^{15}AAYAAL^{20}, a strain with a mutation in the *mutS* gene which destabilizes the protein (58); MutS D835R, a strain with a mutation in the *mutS* gene which disrupts the ability of MutS to tetramerize but does not affect its ability to dimerize (27); MutS ^{15}AAYAAL^{20} D835R, a strain with a double mutation in the *mutS* gene.

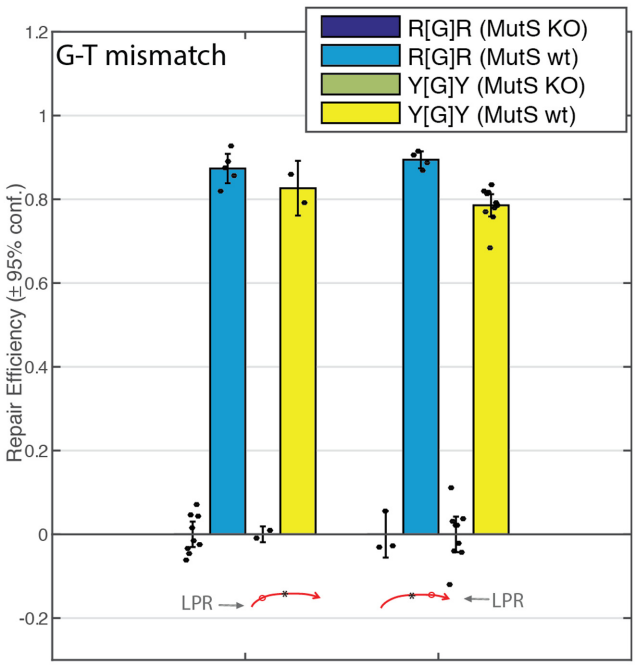


Figure 3. Repair efficiencies, with oligos showing direction of long-patch repair (LPR) allowed below as in Figure 2, obtained by semi-protected oligonucleotide recombination (SPORE) assay of G–T mismatches where the mismatched G ([G]) is flanked by either pYrimidines (Y[G]Y) or puRines (R[G]R).

flanked by pyrimidines (82.64% (±15.0%; 95% confidence) and 78.6% (±2.7%; 95% confidence) for 5′- and 3′-repair, respectively) compared with those flanked by purines (87.4% (±3.5%; 95% confidence) and 89.4% (±6.9%; 95% confidence) for 5′- and 3′-repair, respectively) that was statistically significant ($P = 0.00045$ for 3′-repair). Interestingly, the ∼10% decrease in repair efficiencies of G–T mismatches flanked by pyrimidines compared with those flanked by purines almost exactly mirrors the decrease in hMSH ATPase reported by Mazurek *et al.* (28), and shows that SPORE is capable of discerning even these subtle effects via Sanger sequencing.

DISCUSSION

The results presented demonstrate that the SPORE assay is a remarkably versatile molecular tool that can be used to simultaneously probe how of the molecular species of a mismatch, the sequence context of a mismatch, and any mutation in MMR proteins all affect the mechanistic components of MMR—such as repair efficiency and direction of epigenetic signal used to coordinate long patch repair—*in vivo* and with quantitative sensitivity. The SPORE assay does so in an extremely specific and targeted manner that is beyond the capabilities of spontaneous phenotypic reporter assays: for example, spontaneous rifampicin resistance may arise from mutation in the *rpoB* gene at any one of several distinct sites that can accommodate, overall, all

six possible transitions or traversions (62) as well as with several in-frame deletions (63). Different mutations in *rpoB* that result in rifampicin resistance can also affect cellular fitness differently (64), and the SPORE assay removes this added complication. The SPORE assay also does not require whole genome sequencing (65), e.g., to identify rare, spontaneous mutations, and only requires the sequencing at a single site. Thus, we expect that future work using next generation 'deep' sequencing techniques (66) at the targeted site rather than Sanger sequencing will further allow for extremely high sensitivity quantifying MMR efficiency as well as simultaneous characterization of multiple 'probe' mismatches in single experiment for different mutational strains, allowing for a streamlined deconstruction of the MMR pathway.

However, we must consider the extent to which the repair of 'replication errors' generated by oligonucleotide recombination is reflective of the native MMR and the extent to which blockage of exonuclease digestion represents a 'true' measure of MMR directionality. As mentioned, based on previous work characterizing oligonucleotide recombination as a genomic engineering technique, the mechanistic overlap between these two is likely to be substantial. There is a large body of work from multiple laboratories which have showed that the frequency of successful recombination events improves in lockstep with any cellular changes that inhibit or overload the DNA MMR pathway (24,42,44,67), which has also been found to be true for oligonucleotide recombination in organisms other than *E. coli* (45,51,52,68). As to the second point, a concern would be that long-patch repair and digestion up to the phosphorothioate bonds can destabilize the synthetic oligo on the genomic DNA, which may cause it to melt off the chromosomal DNA after long-patch repair but prior to incorporation of the 'control' mutation. Depending on the side of the phosphoriate bonds which have been digested, this may result in false positives or false negatives. However, based on the designs of the oligos, this does not appear likely: conservatively, the melting temperatures of segments of the oligos that flank the phosphorothioate bonds are $\gg$30°C (growth temperature), with the lowest melting temperature of one of those flanking sites (3′-end of oligo 5′-GT and oligo 5′-TT) being 57.9°C, at $[Na^+] = 120$ mM and $[Mg^{2+}] = 5$ mM (4,17). Hence, even without considering the nucleotides flanked by phosphorothioate bonds, we would still only expect aberrant melting to produce a false negative rate of <3% in this case (with $\ll$ 0.1% for other long-patch repair events of these oligos; http://unafold.rna.albany.edu/) which can still be improved and controlled for in the future through design of different oligonucleotides. Lastly, we will note that a challenge in the present study is the limited number of possible 'silent' mutation sites that can be used as 'probe' mismatch sites around the amber mutation in *galK* found in the SIMD50/SIMD90 strains. In future work, a number of alternative genes can be used to introduce selectable 'control' mutations—such as *rpoB* for rifampicin resistance, *rpsL* for streptomycin resistance, *malK* for maltose metabolism, or *tolC* for resistance to colcin E1, to name a few (69)—that can be used to vary the chromosomal location or wider sequence context for SPORE assays.

Acknowledging these limitations, the SPORE assay has already yielded new insights that previously would have been very difficult to observe otherwise. For one matter, the SPORE assay is able to reveal that MMR of lagging strand mismatches can occur bi-directionally, particularly in the case of G-T mismatches where repair efficiency is approximately equal from either 5′- or 3′-direction when controlling for sequence context. This finding is contrary to the conclusions of a recent report by Hasan and Leach that MMR efficiency at an unstable trinucleotide repeat (TNR) array was inversely correlated with the distance to the nearest origin-distal d(GATC) site but not the distance to the origin-proximal d(GATC) site (20). The authors had suggested that this finding implied MMR was coordinated by an interaction with the replisome during replication. We will note, however, that our finding of bidirectional repair *in vivo* is still perfectly compatible with their observed inverse correlation if repair of some mismatches (like T–T mismatches) are repaired in a directionally-biased manner in the direction of replication while others (like G–T mismatches) appear to be repair equally from both sides. Because TNR arrays may expand or contract on either strand during replication, this inverse correlation will also hold if repair of leading-strand mismatches are preferentially coordinated by the replisome in the direction of replication while lagging strand repair may be mixed (bi- and uni-directional, as we have seen). This may also occur if repair of lagging strand mismatches can be coordinated by non-d(GATC) features such as the natural breaks in the Okazaki fragments or the single-stranded DNA / double-stranded DNA (ssDNA-dsDNA) junction at the 5′-end of the replicating oligo (70). These last arguments are particularly compelling since Hasan and Leach found that even a 2 kb separation between the mismatch and the closest remaining d(GATC) site, twice as far as has ever been observed *in vitro* (54), did not abolish MMR and there is recent evidence that MMR can occur independently of hemi-methylated d(GATC) sites in some cases *in vivo* (71). Since oligonucleotide recombination can be efficiently performed on the leading strand as well as the lagging strand (22), we are currently testing these hypotheses using the SPORE assay.

However, since MMR is initiated bi-directionally *in vitro* with approximately equal rates (5,54), the observed bias in the direction from which repair is initiated in the repair of T–T mismatches (3′- $\gg$ 5′-) and the asymmetric effects of MutS D835R mutations on the repair of G–T mismatches would still suggest, regardless, that MMR is coordinated in some way through an interaction between MMR and the replisome (20). Alternatively, this asymmetry in repair may arise from an asymmetry in the replicated DNA on the lagging strand itself. That MutS D835R mutations do not appear to affect repair of the G-T mismatches from the 5′- would suggest that only MutS dimers are necessary to coordinate repair from this side, while repair of 3′- repair is significantly enhanced by the ability of MutS to tetramerize (or form loops in the heteroduplex DNA). One could consider that upon forming a sliding clamp at the mismatch, the MutS and MutL(H) sliding clamps can diffuse randomly in either direction (Figure 4): if it encounters the 5′- ssDNA–dsDNA junction at the end of the Okazaki fragment (Figure 4i), the asymmetry of the junction is sufficient information

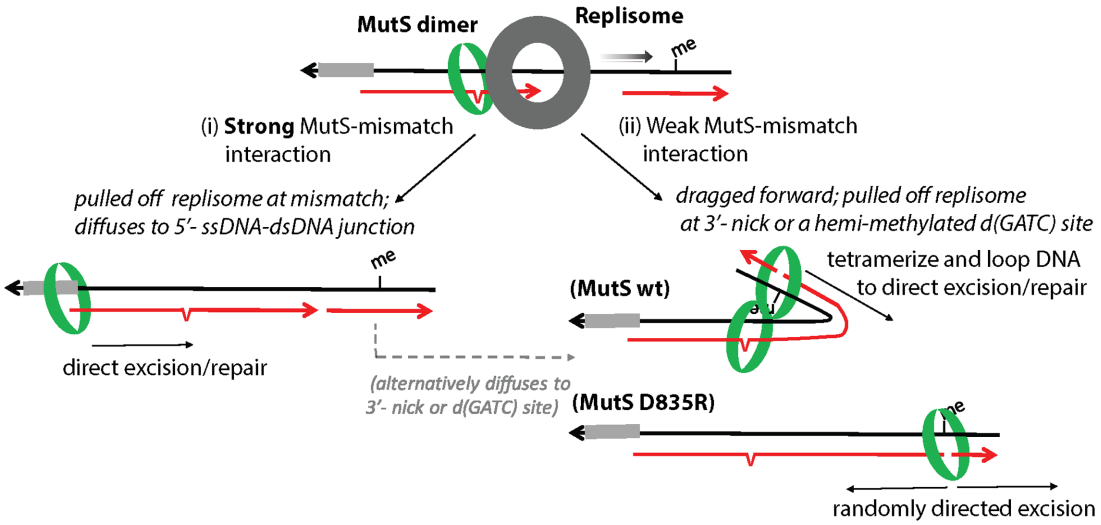


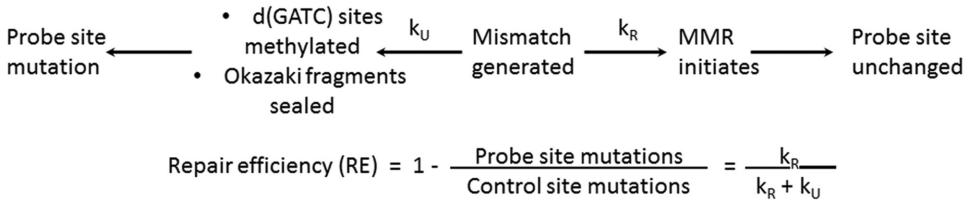
Figure 4. Asymmetry in the bi-directional repair of lagging strand mismatches. (i) An 'activated' MutS dimer detaches from the replisome at a mismatch site and can initiate and direct repair by diffusing either to the 5′-single-stranded DNA/double-stranded DNA (ssDNA-dsDNA) junction (with MutL), or (ii) by looping the DNA as a tetramer after diffusing to and nicking a 3′-hemi-methylated d(GATC) site with MutLH. If MutS is unable to tetramerize, repair directed from hemi-methylated d(GATC) sites becomes uncoordinated.

for the MutL sliding clamps, loaded by activated MutS (14), to itself load UvrD back toward the mismatched site and direct excision/resynthesis (72). A similar mechanism has recently been proposed for the anti-homeologous recombination activity of MutS (73). If the activated MutS diffuses away from the mismatch (to the 3′- end) and instead encounters a hemi-methylated d(GATC) site first (Figure 4ii), with MutL it can direct MutH to produce a single stranded nick but from a single-stranded nick alone does not possess sufficient information to properly direct excision. In this case, we would propose that looping of the DNA by the MutS tetramers (18), or direct contact between the MutS at the d(GATC) site and a separate MutS dimer which has bound the mismatched site after the initial MutS dimer has departed, could direct efficient excision between the two sites. Otherwise, since MMR has already initiated, excision is apparently directed in a random direction, resulting in the ∼50% defect in 3′- MMR that we observe.

One consideration is that 5′- and 3′-repair are both possible during an MMR event that occurs naturally, while in a SPORE experiment we chemically block long-patch repair from one direction. However, by constructing a simple mathematical model, we can use the information obtained in the series of SPORE experiments to estimate the relative importance for lagging strand MMR of repair originating from each of the directions. Such modeling can provide a useful means to compare the relative effects of different MMR protein mutations, different lesions, and different sequence contexts, and to compare the results of SPORE assays with the substantial biochemical and spontaneous reporter assay literature. We do this by considering that there is a finite time window during which repair can be coordinated after replication, and to observe an a wild-type 'probe' site, repair must have initiated before this time window closes. Therefore, our mathematical model behaves according to Scheme 1.

If we assume, for simplicity, that to first approximation the probability that repair initiates (with rate k_R) and the probability that the epigenetic signals required for repair are removed (with rate k_U) each follow first-order kinetics, then in this model RE can be described as: $RE = k_R / (k_R + k_U)$ (74). Furthermore, the ∼ 2–4 s needed to re-methylate lagging strand d(GATC) sites located 3′- of the probe sites (75) is approximately the same as the time required for the next Okazaki fragment to fill in the 5′-ssDNA–dsDNA junction (76,77), so here we also assume that ${k_U}^{5'} \approx {k_U}^{3'}$. From this simplified model, we estimate that, while MutS tetramerization increases the 3′-repair rates 10-fold from 0.359 k_U to 3.66 k_U, the 5′-repair rate in the absence of MutS tetramerization is 4.18 k_U, which may help to explain why MutS D835R exhibited only a moderate mutator defect in Rif assays but show significant defect in their 3′-directed repair in SPORE assays (19). Furthermore, from the data shown here we can estimate the overall repair rate (${k_R}^{5'} + {k_R}^{3'}$) of G-T mismatches (10.6 k_U) is about two orders of magnitude greater that of T–T mismatches (0.122 k_U).

Intriguingly, while the repair efficiencies of different mismatches have long been known to depend on the molecular species of the mismatch (55–57,78) the SPORE assay has also revealed that the directional bias and effect of MutS mutations on the repair of T–T mismatches appear to differ from those of the repair of G–T mismatches. This is unexpected because crystal structures showing MutS bound to different types of mismatches revealed a common binding mode across these various lesions (79,80), and so it has been generally accepted that, after recognition by MutS, MMR proceeds according to a common mechanism for all replication errors (78). Our finding raises the possibility that the initiation of MMR may be coordinated differently in a manner that depends on the molecular species of the mismatch. How could such differences in directional bias of the repair of G–T versus T–T mismatches be explained within in this context? One possibility (Figure 4) is that the repair com-



Scheme 1. Simplified model of repair efficiencies from SPORE assays.

plex traveling with the replisome (81) may detach from the replisome at mismatches with which MutS interacts more strongly (like G–T), enabling it to diffuse randomly as a sliding clamp to identify a strand discrimination signal, while the repair complex must either identify the mismatch without respect to the replisome (58) or be dragged 3′- by the replisome from sites with which MutS only weakly interacts (like T–T). In this model, we would therefore expect mismatches that interact strongly with MutS (57) to be repaired bi-directionally, while mismatches poorly recognized by MutS will have a bias in the direction of repair initiation from the direction of replication (3′- end), as we observe in the limited cases presented here. A more rigorous validation of all possible mismatches and sequence contexts will be required, and at present we are performing an exhaustive characterization of repair efficiencies and mechanisms of all twelve possible mismatches (in addition to insertion/deletion bulges) in different sequence contexts using the SPORE assay.

Because oligonucleotide recombination is routinely performed in eukaryotic cells as well as prokaryotic cells, we would expect the SPORE assay to be readily adaptable to deconstructing MMR in these systems. As mentioned above, foundational studies have linked MMR (26,49–53), replication (25,45–48), and DNA damage response (45,48,82–85) directly to oligonucleotide recombination events in eukaryotic cells. In fact, many of the necessary adaptations necessary to perform the SPORE assay into eukaryotic systems, such as consideration of cell cycle progression (46,86) and toxicity of the chemically-modified synthetic oligonucleotides (87,88), have already been described. The use of synthetic 'locked nucleic acids' (LNAs) appears to be particularly promising in this regard (37). Furthermore, we will also note that our observations with regards to repair of lagging strand G-T mismatches in *E. coli* do exhibit similarities to that of 5′- and 3′-directed MMR in humans of G–T heteroduplexes (89). In eukaryotic mismatch repair, 5′-directed repair does not require a nick to direct excision while 3′-repair does, and these results may suggest a conservation of mechanism with respect to how repair is coordinated directionally. Use of the SPORE assay in human cells to identify any mechanistic differences in the mismatch repair of, for example, chemotherapeutically-derived DNA damage sites recognized by different MutS homologues versus naturally-occurring replication errors may further help in the identification of new pharmaceutical targets for synthetic lethality (90).

In conclusion, an optimized ability to directly insert / observe targeted genomic lesions using the SPORE assay presented here opens up new avenues to directly test the biochemical and mechanistic hypotheses of MMR directly within living cells in a highly specific and controlled manner.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We wish to thank the laboratory of Donald Court (National Cancer Institute, Frederick, MD) for their generous gift of *E. coli* strains SIMD50 and SIMD90 as well as Paul Modrich (Duke University Medical Center, Durham, NC; Howard Hughes Medical Institute) for helpful discussions.

FUNDING

National Science Foundation [MCB-1244297 to P.E.M.]; National Institutes of Health [F32GM112502 to E.A.J.]. Funding for open access charge: National Science Foundation.
Conflict of interest statement. None declared.

REFERENCES

1. Kunkel,T.A. and Erie,D.A. (2005) DNA mismatch repair. *Annu. Rev. Biochem.*, **74**, 681–710.
2. Jiricny,J. (2013) Postreplicative mismatch repair. *Cold Spring Harbor Perspect. Biol.*, **5**.
3. Au,K.G., Welsh,K. and Modrich,P. (1992) Initiation of methyl-directed mismatch repair. *J. Biol. Chem.*, **267**, 12142–12148.
4. Lahue,R.S., Au,K.G. and Modrich,P. (1989) DNA mismatch correction in a defined system. *Science*, **245**, 160–164.
5. Cooper,D.L., Lahue,R.S. and Modrich,P. (1993) Methyl-directed mismatch repair is bidirectional. *J. Biol. Chem.*, **268**, 11823–11829.
6. Lahue,R.S., Su,S.S. and Modrich,P. (1987) Requirement for d(GATC) sequences in *Escherichia coli mutHLS* mismatch correction. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 1482–1486.
7. Bruni,R., Martin,D. and Jiricny,J. (1988) d(GATC) sequences influence Escherichia coli mismatch repair in a distance-dependent manner from positions both upstream and downstream of the mismatch. *Nucleic Acids Res.*, **16**, 4875–4890.
8. Längle-Rouault,F., Maenhaut,M.G. and Radman,M. (1987) GATC sequences, DNA nicks and the MutH function in *Escherichia coli* mismatch repair. *EMBO J.*, **6**, 1121–1127.
9. Kolodner,R.D., Mendillo,M.L. and Putnam,C.D. (2007) Coupling distant sites in DNA during DNA mismatch repair. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 12953–12954.
10. Blackwell,L.J., Bjornson,K.P., Allen,D.J. and Modrich,P. (2001) Distinct MutS DNA-binding modes that are differentially modulated by ATP binding and hydrolysis. *J. Biol. Chem.*, **276**, 34339–34347.
11. Acharya,S., Foster,P.L., Brooks,P. and Fishel,R. (2003) The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Mol. Cell*, **12**, 233–246.
12. Qiu,R., DeRocco,V.C., Harris,C., Sharma,A., Hingorani,M.M., Erie,D.A. and Weninger,K.R. (2012) Large conformational changes

in MutS during DNA scanning, mismatch recognition and repair signalling. *EMBO J.*, **31**, 2528–2540.
13. Jeong,C., Cho,W.-K., Song,K.-M., Cook,C., Yoon,T.-Y., Ban,C., Fishel,R. and Lee,J.-B. (2011) MutS switches between two fundamentally distinct clamps during mismatch repair. *Nat. Struct. Mol. Biol.*, **18**, 379–385.
14. Liu,J., Hanne,J., Britton,B.M., Bennett,J., Kim,D., Lee,J.-B. and Fishel,R. (2016) Cascading MutS and MutL sliding clamps control DNA diffusion to activate mismatch repair. *Nature*, **539**, 583–587.
15. Bjornson,K.P., Blackwell,L.J., Sage,H., Baitinger,C., Allen,D. and Modrich,P. (2003) Assembly and molecular activities of the MutS tetramer. *J. Biol. Chem.*, **278**, 34667–34673.
16. Allen,D.J., Makhov,A., Grilley,M., Taylor,J., Thresher,R., Modrich,P. and Griffith,J.D. (1997) MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.*, **16**, 4467–4476.
17. Jiang,Y. and Marszalek,P.E. (2011) Atomic force microscopy captures MutS tetramers initiating DNA mismatch repair. *EMBO J.*, **30**, 2881–2893.
18. Josephs,E.A., Zheng,T. and Marszalek,P.E. (2015) Atomic force microscopy captures the initiation of methyl-directed DNA mismatch repair. *DNA Repair*, **35**, 71–84.
19. Mendillo,M.L., Putnam,C.D. and Kolodner,R.D. (2007) Escherichia coli MutS tetramerization domain structure reveals that stable dimers but not tetramers are essential for DNA mismatch repair in vivo. *J. Biol. Chem.*, **282**, 16345–16354.
20. Hasan,A.M. and Leach,D.R. (2015) Chromosomal directionality of DNA mismatch repair in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 9388–9393.
21. Junop,M.S., Yang,W., Funchain,P., Clendenin,W. and Miller,J.H. (2003) In vitro and in vivo studies of MutS, MutL and MutH mutants: correlation of mismatch repair and DNA recombination. *DNA Repair (Amst.)*, **2**, 387–405.
22. Ellis,H.M., Yu,D., DiTizio,T. and Court,D.L. (2001) High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 6742–6746.
23. Thomason,L.C., Costantino,N. and Court,D.L. (2016) Examining a DNA replication requirement for bacteriophage lambda Red- and Rac prophage RecET-promoted recombination in Escherichia coli. *mBio*, **7**, 10.1128/mBio.01443-01416.
24. Costantino,N. and Court,D.L. (2003) Enhanced levels of lambda Red-mediated recombinants in mismatch repair mutants. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 15748–15753.
25. Aarts,M. and te Riele,H. (2010) Parameters of oligonucleotide-mediated gene modification in mouse ES cells. *J.Cell. Mol. Med.*, **14**, 1657–1667.
26. Dekker,M., Brouwers,C. and te Riele,H. (2003) Targeted gene modification in mismatch-repair-deficient embryonic stem cells by single-stranded DNA oligonucleotides. *Nucleic Acids Res.*, **31**, e27.
27. Manelyte,L., Urbanke,C., Giron-Monzon,L. and Friedhoff,P. (2006) Structural and functional analysis of the MutS C-terminal tetramerization domain. *Nucleic Acids Res.*, **34**, 5270–5279.
28. Mazurek,A., Johnson,C.N., Germann,M.W. and Fishel,R. (2009) Sequence context effect for hMSH2-hMSH6 mismatch-dependent activation. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 4177–4182.
29. Warming,S., Costantino,N., Court,D.L., Jenkins,N.A. and Copeland,N.G. (2005) Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.*, **33**, e36.
30. Datta,S., Costantino,N., Zhou,X. and Court,D.L. (2008) Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 1626–1631.
31. Sawitzke,J.A., Thomason,L.C., Costantino,N., Bubunenko,M., Datta,S. and Court,D.L. (2007) Recombineering: in vivo genetic engineering in E. coli, S. enterica, and beyond. *Methods Enzymol.*, **421**, 171–199.
32. Sharan,S.K., Thomason,L.C., Kuznetsov,S.G. and Court,D.L. (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protoc.*, **4**, 206–223.
33. Mosberg,J.A., Gregg,C.J., Lajoie,M.J., Wang,H.H. and Church,G.M. (2012) Improving lambda red genome engineering in Escherichia coli via rational removal of endogenous nucleases. *PLoS One*, **7**, e44638.
34. Carr,I.M., Robinson,J.I., Dimitriou,R., Markham,A.F., Morgan,A.W. and Bonthron,D.T. (2009) Inferring relative proportions of DNA variants from sequencing electropherograms. *Bioinformatics (Oxford, England)*, **25**, 3244–3250.
35. Parker,B.O. and Marinus,M.G. (1992) Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 1730–1734.
36. Swingle,B., Markel,E., Costantino,N., Bubunenko,M.G., Cartinhour,S. and Court,D.L. (2010) Oligonucleotide recombination in Gram-negative bacteria. *Mol. Microbiol.*, **75**, 138–148.
37. van Ravesteyn,T.W., Dekker,M., Fish,A., Sixma,T.K., Wolters,A., Dekker,R.J. and Te Riele,H.P. (2016) LNA modification of single-stranded DNA oligonucleotides allows subtle gene modification in mismatch-repair-proficient cells. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 4122–4127.
38. Lajoie,M.J., Gregg,C.J., Mosberg,J.A., Washington,G.C. and Church,G.M. (2012) Manipulating replisome dynamics to enhance lambda Red-mediated multiplex genome engineering. *Nucleic Acids Res.*, **40**, e170.
39. Li,X.T., Costantino,N., Lu,L.Y., Liu,D.P., Watt,R.M., Cheah,K.S., Court,D.L. and Huang,J.D. (2003) Identification of factors influencing strand bias in oligonucleotide-mediated recombination in Escherichia coli. *Nucleic Acids Res.*, **31**, 6674–6687.
40. Li,X.t., Thomason,L.C., Sawitzke,J.A., Costantino,N. and Court,D.L. (2013) Bacterial DNA polymerases participate in oligonucleotide recombination. *Mol. Microbiol.*, **88**, 906–920.
41. Ryu,Y.S., Biswas,R.K., Shin,K., Parisutham,V., Kim,S.M. and Lee,S.K. (2014) A simple and effective method for construction of Escherichia coli strains proficient for genome engineering. *PLoS One*, **9**, e94266.
42. Nyerges,A., Csorgo,B., Nagy,I., Latinovics,D., Szamecz,B., Posfai,G. and Pal,C. (2014) Conditional DNA repair mutants enable highly precise genome engineering. *Nucleic Acids Res.*, **42**, e62.
43. Sawitzke,J.A., Costantino,N., Li,X.-T., Thomason,L.C., Bubunenko,M., Court,C. and Court,D.L. (2011) Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. *J. Mol. Biol.*, **407**, 45–59.
44. Wang,H.H., Xu,G., Vonner,A.J. and Church,G. (2011) Modified bases enable high-efficiency oligonucleotide-mediated allelic replacement via mismatch repair evasion. *Nucleic Acids Res.*, **39**, 7336–7347.
45. Ferrara,L. and Kmiec,E.B. (2006) Targeted gene repair activates Chk1 and Chk2 and stalls replication in corrected cells. *DNA Repair*, **5**, 422–431.
46. 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.
47. Kow,Y.W., Bao,G., Reeves,J.W., Jinks-Robertson,S. and Crouse,G.F. (2007) Oligonucleotide transformation of yeast reveals mismatch repair complexes to be differentially active on DNA replication strands. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 11352–11357.
48. Aarts,M. and te Riele,H. (2010) Subtle gene modification in mouse ES cells: evidence for incorporation of unmodified oligonucleotides without induction of DNA damage. *Nucleic Acids Res.*, **38**, 6956–6967.
49. Dekker,M., de Vries,S., Aarts,M., Dekker,R., Brouwers,C., Wiebenga,O., de Wind,N., Cantelli,E., Tonelli,R. and te Riele,H. (2011) Transient suppression of MLH1 allows effective single-nucleotide substitution by single-stranded DNA oligonucleotides. *Mut. Res./Fundam. Mol. Mech. Mutagen.*, **715**, 52–60.
50. 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.
51. Engstrom,J.U., Suzuki,T. and Kmiec,E.B. (2009) Regulation of targeted gene repair by intrinsic cellular processes. *Bioessays*, **31**, 159–168.
52. Ferrara,L., Parekh-Olmedo,H. and Kmiec,E.B. (2004) Enhanced oligonucleotide-directed gene targeting in mammalian cells following treatment with DNA damaging agents. *Exp. Cell Res.*, **300**, 170–179.
53. Rodriguez,G.P., Romanova,N.V., Bao,G., Rouf,N.C., Kow,Y.W. and Crouse,G.F. (2012) Mismatch repair-dependent mutagenesis in nondividing cells. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 6153–6158.

54. Grilley,M., Griffith,J. and Modrich,P. (1993) Bidirectional excision in methyl-directed mismatch repair. *J. Biol. Chem.*, **268**, 11830–11837.
55. Su,S.-S., Lahue,R.S., Au,K.G. and Modrich,P. (1988) Mispair specificity of methyl-directed DNA mismatch correction in vitro. *J. Biol. Chem.*, **263**, 6829–6835.
56. Jones,M., Wagner,R. and Radman,M. (1987) Repair of a mismatch is influenced by the base composition of the surrounding nucleotide sequence. *Genetics*, **115**, 605–610.
57. Groothuizen,F.S., Fish,A., Petoukhov,M.V., Reumer,A., Manelyte,L., Winterwerp,H.H., Marinus,M.G., Lebbink,J.H., Svergun,D.I., Friedhoff,P. *et al.* (2013) Using stable MutS dimers and tetramers to quantitatively analyze DNA mismatch recognition and sliding clamp formation. *Nucleic Acids Res.*, **41**, 8166–8181.
58. Pluciennik,A., Burdett,V., Lukianova,O., O'Donnell,M. and Modrich,P. (2009) Involvement of the β clamp in methyl-directed mismatch repair in vitro. *J. Biol. Chem.*, **284**, 32782–32791.
59. Pluciennik,A., Burdett,V., Lukianova,O., O'Donnell,M. and Modrich,P. (2009) Involvement of the beta clamp in methyl-directed mismatch repair in vitro. *J. Biol. Chem.*, **284**, 32782–32791.
60. Macpherson,P., Humbert,O. and Karran,P. (1998) Frameshift mismatch recognition by the human MutSα complex. *Mutat. Res./DNA Repair*, **408**, 55–66.
61. Bonde,M.T., Klausen,M.S., Anderson,M.V., Wallin,A.I., Wang,H.H. and Sommer,M.O. (2014) MODEST: a web-based design tool for oligonucleotide-mediated genome engineering and recombineering. *Nucleic Acids Res.*, **42**, W408–W415.
62. Garibyan,L., Huang,T., Kim,M., Wolff,E., Nguyen,A., Nguyen,T., Diep,A., Hu,K., Iverson,A. and Yang,H. (2003) Use of the rpoB gene to determine the specificity of base substitution mutations on the Escherichia coli chromosome. *DNA Repair*, **2**, 593–608.
63. Severinov,K., Soushko,M., Goldfarb,A. and Nikiforov,V. (1993) Rifampicin region revisited. New rifampicin-resistant and streptolydigin-resistant mutants in the beta subunit of Escherichia coli RNA polymerase. *J. Biol. Chem.*, **268**, 14820–14825.
64. Reynolds,M.G. (2000) Compensatory evolution in rifampin-resistant Escherichia coli. *Genetics*, **156**, 1471–1481.
65. Lee,H., Popodi,E., Tang,H. and Foster,P.L. (2012) Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, E2774–2783.
66. Schmitt,M.W., Kennedy,S.R., Salk,J.J., Fox,E.J., Hiatt,J.B. and Loeb,L.A. (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 14508–14513.
67. Lennen,R.M., Nilsson Wallin,A.I., Pedersen,M., Bonde,M., Luo,H., Herrgård,M.J. and Sommer,M.O.A. (2016) Transient overexpression of DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects. *Nucleic Acids Res.*, **44**, e36–e36.
68. DiCarlo,J.E., Conley,A.J., Penttilä,M., Jäntti,J., Wang,H.H. and Church,G.M. (2013) Yeast oligo-mediated genome engineering (YOGE). *ACS Synth. Biol.*, **2**, 741–749.
69. Wang,H.H., Kim,H., Cong,L., Jeong,J., Bang,D. and Church,G.M. (2012) Genome-scale promoter engineering by coselection MAGE. *Nat. Methods*, **9**, 591–593.
70. Tham,K.-C., Hermans,N., Winterwerp,H.H.K., Cox,M.M., Wyman,C., Kanaar,R. and Lebbink,J.H.G. (2013) Mismatch repair inhibits homeologous recombination via coordinated directional unwinding of trapped DNA structures. *Mol. Cell*, **51**, 326–337.
71. Sneppen,K. and Semsey,S. (2014) Mismatch repair at stop codons is directed independent of GATC methylation on the Escherichia coli chromosome. *Sci. Rep.*, **4**.
72. Mechanic,L.E., Frankel,B.A. and Matson,S.W. (2000) Escherichia coli MutL loads DNA helicase II onto DNA. *J. Biol. Chem.*, **275**, 38337–38346.
73. Tham,K.-C., Kanaar,R. and Lebbink,J.H.G. (2016) Mismatch repair and homeologous recombination. *DNA Repair*, **38**, 75–83.
74. Colquhoun,D. and Hawkes,A.G. (1995), *Single-Channel Recording*. Springer, pp. 397–482.
75. Stancheva,I., Koller,T. and Sogo,J.M. (1999) Asymmetry of Dam remethylation on the leading and lagging arms of plasmid replicative intermediates. *EMBO J.*, **18**, 6542–6551.
76. Georgescu,R.E., Kurth,I., Yao,N.Y., Stewart,J., Yurieva,O. and O'Donnell,M. (2009) Mechanism of polymerase collision release from sliding clamps on the lagging strand. *EMBO J.*, **28**, 2981–2991.
77. Breier,A.M., Weier,H.U. and Cozzarelli,N.R. (2005) Independence of replisomes in Escherichia coli chromosomal replication. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 3942–3947.
78. DeRocco,V.C., Sass,L.E., Qiu,R., Weninger,K.R. and Erie,D.A. (2014) Dynamics of MutS–mismatched DNA complexes are predictive of their repair phenotypes. *Biochemistry*, **53**, 2043–2052.
79. Lamers,M.H., Perrakis,A., Enzlin,J.H., Winterwerp,H.H., de Wind,N. and Sixma,T.K. (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G-T mismatch. *Nature*, **407**, 711–717.
80. Natrajan,G., Lamers,M.H., Enzlin,J.H., Winterwerp,H.H., Perrakis,A. and Sixma,T.K. (2003) Structures of escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates. *Nucleic Acids Res.*, **31**, 4814–4821.
81. Liao,Y., Schroeder,J.W., Gao,B., Simmons,L.A. and Biteen,J.S. (2015) Single-molecule motions and interactions in live cells reveal target search dynamics in mismatch repair. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E6898–E6906.
82. 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.
83. Crouse,G.F. (2016) Non-canonical actions of mismatch repair. *DNA Repair*, **38**, 102–109.
84. Rodriguez,G.P., Song,J.B. and Crouse,G.F. (2013) In vivo bypass of 8-oxodG. *PLoS Genet.*, **9**, e1003682.
85. Rodriguez,G.P., Song,J.B. and Crouse,G.F. (2012) Transformation with oligonucleotides creating clustered changes in the yeast genome. *PLoS One*, **7**, e42905.
86. 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.
87. Papaioannou,I., Disterer,P. and Owen,J.S. (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.
88. Olsen,P.A., Randøl,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.
89. Constantin,N., Dzantiev,L., Kadyrov,F.A. and Modrich,P. (2005) Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *J. Biol. Chem.*, **280**, 39752–39761.
90. Begum,R. and Martin,S.A. (2016) Targeting mismatch repair defects: a novel strategy for personalized cancer treatment. *DNA Repair*, **38**, 135–139.