6674–6687 *Nucleic Acids Research, 2003, Vol. 31, No. 22*
DOI: 10.1093/nar/gkg844

Identification of factors influencing strand bias in oligonucleotide-mediated recombination in *Escherichia coli*

Xin-tian Li[1,2], Nina Costantino[3], Lin-yu Lu[1], De-pei Liu[2], Rory M. Watt[1], Kathryn S. E. Cheah[1], Donald L. Court[3] and Jian-Dong Huang[1,*]

[1]Department of Biochemistry, The University of Hong Kong, 3/F Laboratory Block, Faculty of Medicine Building, 21 Sassoon Road, Pokfulam, Hong Kong SAR, P. R. China, [2]National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences (CAMS) and Peking Union Medical College (PUMC), Beijing 100005, P. R. China and [3]Gene Regulation and Chromosome Biology Laboratory, Division of Basic Science, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, MD 21702, USA

Received May 2, 2003; Revised July 31, 2003; Accepted September 22, 2003

ABSTRACT

Recombinogenic engineering methodology, also known as recombineering, utilizes homologous recombination to create targeted changes in cellular DNA with great specificity and flexibility. In *Escherichia coli*, the Red recombination system from bacteriophage lambda has been used successfully to modify both plasmid and chromosomal DNA in a highly efficient manner, using either a linear double-stranded DNA fragment or a synthetic single-stranded oligonucleotide (SSO). The current model for Red/SSO-mediated recombination involves the SSO first annealing to a transient, single-stranded region of DNA before being incorporated into the chromosome or plasmid target. It has been observed previously, in both eukaryotes and prokaryotes, that mutations in the two strands of the DNA double helix are 'corrected' by complementary SSOs with differing efficiencies. Here we investigate further the factors that influence the strand bias as well as the overall efficiency of Red/SSO-mediated recombination in *E.coli*. We show that the direction of DNA replication and the nature of the SSO-encoded mismatch are the main factors dictating the recombinational strand bias. However, the influence that the SSO-encoded mismatch exerts upon the recombinational strand bias is abolished in *E.coli* strains that are defective in mismatch repair (MMR). This reflects the fact that different base–base mispairs are corrected by the *mutS/H/L*-dependent MMR pathway with differing efficiencies. Furthermore, our data indicate that transcription has negligible influence on the strand bias. These results demonstrate for the first time that the interplay between DNA replication and MMR has a major effect on the efficiency and strand bias of Red/SSO-mediated recombination in *E.coli*.

INTRODUCTION

Homologous recombination-based DNA engineering technology, also known as recombinogenic engineering or recombineering, allows chromosomal or plasmid DNA to be modified in an exquisitely precise manner, and eliminates the need for *in vitro* manipulations using restriction enzymes or DNA ligases [reviewed in (1,2)]. In *Escherichia coli*, PCR-generated targeting 'cassettes' have been used to generate deletions, insertions and substitutions in efficient procedures controlled by the *rac*-encoded RecET system, or the bacteriophage-λ Red recombination system (3–7). Using the Red system, it has been shown recently that synthetic single-stranded oligonucleotides (SSOs) may also be used in place of linear double-stranded DNA fragments to create sequence-specific mutations with significantly increased efficiency. When a selection of point mutations were created at various locations throughout the chromosome, >1% of the SSO-treated cells were found to be recombinant, making it possible to screen without the use of a selectable marker (8–10). In a similar manner, SSOs have also been used to introduce genetic modifications in yeast (11–13) and in mammalian cells (10,14,15).

At the present time the exact mechanism(s) by which SSOs mediate these genetic changes remains to be established. Ellis *et al.* (8) have proposed a model where the SSO first anneals to a single-stranded region near the replication fork, before DNA polymerase and ligase complete the recombination process by joining the annealed SSO to the chromosome or plasmid DNA. Consistent with this 'annealing-integration' model, the authors (8) demonstrated that two (complementary) recombinogenic SSOs targeted to different strands at the same chromosomal location had different recombination efficiencies, with the lagging strand SSO (Lag-SSO, which

*To whom correspondence should be addressed. Tel: + 1 852 2819 2810; Fax: +1 852 2855 1254; Email: jdhuang@hkucc.hku.hk

has the same sequence as the Okazaki fragment in DNA replication) being significantly more competent for recombination than the other. This strand bias may be explained by the presence of more single-stranded regions during lagging strand synthesis, which the SSO may access, compared with during leading strand synthesis.

It is possible that other biological processes such as transcription and/or DNA damage repair may also be involved in generating the strand bias observed in SSO-mediated recombination. In fact, in contrast to the situation in *E.coli*, where the strand bias has thus far only been associated with the direction of DNA replication, there are some data to suggest that in yeast and mammalian cells, it is also influenced by the direction of transcription of the target genes (12,15–17). A mutational strand bias has also been observed for SSO-mediated recombination in a number of other eukaryotic systems (12,13,15,16,18). However, the experimental findings and conclusions drawn in these reports differ substantially. In some cases, it was found that recombination occurred at a higher frequency when Template-SSOs (T-SSOs, where the oligonucleotide sequence was defined as being complementary to the mRNA) were used (12,13,15,16), while the converse was observed in other cases (18,19). Working with yeast, the Sherman laboratory concluded that the SSO-mediated mutagenesis of the CYC1 gene was influenced more by the direction of replication, than the direction of transcription.

There are several possible explanations as to how transcription may impose a strand bias upon SSO-mediated recombination. It may result from a difference in steric hindrance that the SSOs experience while accessing the template versus the non-template strands during transcription, or it may be a result of the SSOs being preferentially displaced from the template strand by RNA polymerase (12).

The mismatch repair pathway (MMR) is another possible source of the recombinational strand bias [reviewed in (20)]. MMR is essential for the faithful duplication of the genome, and improves the fidelity of DNA replication by up to three orders of magnitude, through the correction of replication errors introduced by the DNA polymerase holoenzyme. A mismatch existing in the form of a base–base mispair, or a small insertion/deletion loop in double-stranded DNA, is first recognized, then bound by a MutS homodimer (21–23). This is joined by multiple MutS homodimers, forming a higher-order multimeric-complex (24), which 'grows' along the DNA double helix until it eventually encounters a hemi-methylated GATC site (24). The MutL and MutH proteins are then recruited to nick the (unmethylated) newly synthesized strand, which is digested through the mismatch by an exonuclease, assisted by DNA helicase II (UvrD) and SSB. The Polymerase III holoenzyme then 'fills-in' the gap and DNA ligase repairs the nick. In nick-directed MMR, a mismatch in an Okazaki fragment is readily recognized and repaired due to the availability of nearby nicks. Direct interactions between the MMR proteins and the DNA replication machinery have also been observed (25–27), suggesting an additional mechanism by which MMR may be directed to the newly synthesized strand (20). MMR may also contribute to the Red/SSO-mediated recombination strand bias as a result of correcting different base–base mismatches with differing efficiencies. Mispairs [T/G, C/A and G/G], [A/A] and [G/A, A/G, T/T, C/C, C/T and T/C] are repaired at high, intermediate and low frequencies, respectively (28), suggesting that the sequence of the SSO may play a role in the generation of strand bias.

Performing Red/SSO-mediated recombination experiments with sets of complementary oligos at six locations spread throughout the *E.coli* chromosome, Ellis *et al.* (8) demonstrated that the observed strand bias was consistent with the recombination events occurring at (or near) replication forks. However, large discrepancies in the magnitude of the strand bias, and in the overall number of recombinants formed, suggests that other as yet undetermined factors may also be involved in the recombination process in *E.coli*. Although Ellis *et al.* suggested that transcription did not play a role in generating strand bias, they did not test genes that would be expected to be highly transcribed during growth in rich medium. In this paper we further investigate the factors that influence Red/SSO-mediated recombination in *E.coli*. Our results confirm that DNA replication plays a major role in determining the strand bias associated with Red/SSO-mediated recombination with both plasmid and chromosomal targets, and that *recA*-mediated recombination, transcription, nucleotide excision repair (NER) and transcription-coupled repair (TCR) exert little, if any influence on this process. We show that in certain cases, the sequence of the correction-SSO (i.e. the nature of the encoded mismatch) may be the dominant factor in determining the efficiency of the Red/SSO-mediated recombination event, capable of overriding any replication-imposed strand bias. Furthermore, our results demonstrate that the mismatch-dictated recombinational strand bias is mediated through the MMR pathway.

MATERIALS AND METHODS

Media, chemicals and reagents

Tetracycline (tet), chloramphenicol (cm), kanamycin (kan) and ampicillin (amp) were added to LB-agar media plates at 10, 12.5, 15 and 50 μg/ml, respectively. IPTG was added to induce *mkan* expression at 800 μg/LB-agar plate (~50 μg/ml agar), and to 1 mM in liquid LB-media, respectively. Restriction enzymes were purchased from New England Biolabs or Amersham Biosciences; T4 DNA ligase was from Gibco BRL®; Alkaline Phosphatase was from Promega; Taq polymerase was from Roche; Expand High-fidelity polymerase was purchased from Boehringer Mannheim; oligonucleotides were purchased from Proligo (Singapore) in a desalted form without further purification; PCR products and DNA digest products were purified using QIAquick gel-extraction kits (Qiagen) after agarose gel electrophoresis. CBT medium (200 ml) consisted of: M9 buffer (5×, 40 ml); MM1 buffer (0.2 ml); LB-media (2 ml); $MgSO_4$ (0.1 M, 2 ml); $CaCl_2$ (0.5 M, 0.2 ml); glucose (40%, 2 ml); leucine (0.79%, 2 ml); valine (0.79%, 2 ml); isoleucine (0.79%, 2 ml); vitamin B1 (0.05%, 0.4 ml); water added to 200 ml.

Oligonucleotides

See Table 1.

Escherichia coli transformation

Recombination-competent cells were prepared according to Yu *et al.* (4). Briefly, overnight cultures inoculated from a

Table 1. Oligonucleotides

SSO	Sequence	Use/function
mKan-1	GGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTAGGACTG	G introduces a stop codon into the kanamycin (*kan*) resistance gene
mKan-2	TATTCGGCAAGCAGGCATCG	Amplification of the *mkan* resistance gene
6A	GCTATTCGGCTA***C***GACTGGGCACAA	Non-template SSO to correct the nonsense mutation in the *mkan* gene, ***C*** indicates the mismatch
6B	TTGTGCCCAGTC***G***TAGCCGAATAGC	Template SSO to correct the nonsense mutation in *mkan*, ***G*** indicates the mismatch
8A	TTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA***C***GACTGGGCA-CAACAGACAATCGGCTGCTCTGATGCC	Non-template SSO to correct the nonsense mutation in *mkan*, ***C*** indicates the mismatch
8B	GGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTC***G***TAGCCGAAT-AGCCTCTCCACCCAAGCGGCCGGAGAA	Template SSO to correct the nonsense mutation in *mkan*, ***G*** indicates the mismatch
12A	GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA***C***-GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC	Non-template SSO to correct the nonsense mutation in the *mkan* gene, ***C*** indicates the mismatch
12B	GAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTC-***G***TAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGC	Template SSO to correct the nonsense mutation in *mkan*, ***G*** indicates the mismatch
12A-TAT	GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA-***T***GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC	Non-template SSO to correct the nonsense mutation in *mkan*, ***T*** indicates the mismatch
12B-ATA	GAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTC-***A***TAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGC	Template SSO to correct the nonsense mutation in *mkan*, ***A*** indicates the mismatch
DY380(+)-1	GGCGCTGCAAAAATTCTTTGTCGAACAGGGTGTCTGGATCTAATGC-GCCGCTACAGGGCGCGTAA	Amplification of the *mkan* and *amp*^{r} genes, for construction of the DY380(+) and DY330(+) strains
DY380(+)-2	GGCGCTGCAAAAATTCTTTGTCGAACAGGGTGTCTGGATCTAATGC-GCCGCTACAGGGCGCGTAA	Amplification of the *mkan* and *amp*^{r} genes, for construction of DY380(+) and DY330(+)
DY380(+)-detect	ACGTTAAATCTATCACCGCAAGGG	Confirms *mkan* and *amp*^{r} insertion into DY380 and DY330 chromosomes
PGK-frt-2	TATTCGGCAAGCAGGCATCG	Confirms *mkan* insertion in DY380(+) and DY330(+) chromosomes
DY380(–)-1	TGGCGGTGATAATGGTTGCATGTACTAAGGAGGTTGTATGTAATGC-GCCGCTACAGGGCGCGTAA	Amplification of the *mkan* and *amp*^{r} genes, for construction of DY380(–) and DY330(–)
DY380(–)-2	GGCGCTGCAAAAATTCTTTGTCGAACAGGGTGTCTGGATCGCAGCC-CGGTGGACAAGAAGCGAAC	Amplification of the *mkan* and *amp*^{r} genes, for construction of DY380(–) and DY330(–)
PGK-frt-3	GATCGTTTCGCATGATT	Confirms *mkan* insertion in DY380(–) and DY330(–) chromosomes
Plac-Kan-F	GCCAGTTCCGCTGGCGGCATTTTAACTTTCTTTATCACACAGGAAA-CAGCTATGATTGAACAAGATGGATTGCAC	Amplification of *mkan* gene for construction of pLac-mKan
LacA-Kan-F	CGCTGAACTTGTAGGCCTGATAAGCGCAGCGTATCAGGCAATTTTT-ATAATTCAGAAGAACTCGTCAAGAAGGCG	Amplification of *mkan* gene for construction of pLac-mKan
CM>MutS1	ACAACGAAAATAAAAACCATCACACCCCATTTAATATCAGGGAAC-CGGACCGGGGATCCTCTAGAGTCGACCTGC	Amplification of *cm*^{r} gene to replace *mutS*
CM>MutS2	CGCGCCGTCAGCCAGCGCGCGCCACTCATCCAGCTCTTCGTTATAG-CCCGAAACGCTACACAGGCTCCTGAGATT	Amplification of *cm*^{r} gene to replace *mutS*
CM>MutS-detect1	GCGGGCAGTGAGCGCAACG	Confirms *mutS*< >*cm*^{r} replacement
CM>MutS-detect2	ATCAAGAACTCGCCGTTCGCTTCTT	Confirms *mutS*< >*cm*^{r} replacement
CM>MutL1	CGATGCAGGTGCGACGCGTATCGATATTGATATCGAACGCGGTGGG-GCGACGGGGATCCTCTAGAGTCGACCTGC	Amplification of *cm*^{r} gene to replace *mutL*
CM>MutL2	GTCAGATGCACAGCATGTGACCATCAGGGCAGTGCCTTTACCCTTA-CGCCAAACGCTACACAGGCTCCTGAGATT	Amplification of *cm*^{r} gene to replace *mutL*
MutL-detect	TCAGTACTGGTCGAAACCGG	Confirms *mutL*< >*cm*^{r} replacement
CM>MutH-1	GCTCTCTCCTCCCGAAACTGAAGAACAGTTGTTAGCGCAAGCACAG-CAACCGGGGATCCTCTAGAGTCGACCTGC	Amplification of *cm*^{r} gene to replace *mutH*
CM>MutH-2	TTCTGACGCTGCCACGCGGCTTTTATTTGAAGAAGAATTTCACCAGT-GCAAAACGCACACAGGCTCCTGAGATT	Amplification of *cm*^{r} gene to replace *mutH*
MutH-detect	ATTGAACAACGCATGTGG	Confirms *mutH*< >*cm*^{r} replacement
CM>Mfd-1	TAACCGGCGCTGTCCAGTTGGGTTCGTAATGCATCTCGTGACAGGC-GCTGAAACGCTACACAGGCTCCTGAGATT	Amplification of *cm*^{r} gene to replace *mfd*
CM>Mfd-2	ATCAGGACATTATCTCCTCGCGCCTTTCCACCCTTTACCAGCTACCG-ACGCGGGGATCCTCTAGAGTCGACCTGC	Amplification of *cm*^{r} gene to replace *mfd*
Mfd-detect	CCGGTGGTACTCATTGCACC	Confirms *mfd*< >*cm*^{r} replacement
CM>UvrB-1	CGGAAGCCTATGTACCGAGTTCCGACACTTTCATTGAGAAAGATGC-CTCGCGGGGATCCTCTAGAGTCGACCTGC	Amplification of *cm*^{r} gene to replace *uvrB*
CM>UvrB-2	TCGCCACGAACGCGGAAAGTACCACGCTGGAATGCTTGATCATTAC-GAGCAAACGCTACACAGGCTCCTGAGATT	Amplification of *cm*^{r} gene to replace *uvrB*
UvrB-detect	GTTCAAACTGAATTCCGCTT	Confirms *uvrB*< >*cm*^{r} replacement
100	AAGTCGCGGTCGGAACCGTATTGCAGCAGCTTTA***T***CATCTGCCGCT-GGACGGCGCACAAATCGCGCTTAA	Non-template SSO to correct *galKam*
101	TTAAGCGCGATTTGTGCGCCGTCCAGCGGCAGATG***A***TAAAGCTGCT-GCAATACGGTTCCGACCGCGACTT	Template SSO to correct *galKam*
144	AAGTCGCGGTCGGAACCGTATTGCAGCAGCTTTA***C***CATCTGCCGCT-GGACGGCGCACAAATCGCGCTTAA	Non-template SSO to correct *galKam*
145	TTAAGCGCGATTTGTGCGCCGTCCAGCGGCAGATG***G***TAAAGCTGCT-GCAATACGGTTCCGACCGCGACTT	Template SSO to correct *galKam*
NC219	AAGTCGCGGTCGGAACCGTATTGCAGCAGCTTTA***G***CATCTGCCGCT-GGACGGCGCACAAATCGCGCTTAA	Creating a TAG nonsense codon and converting *galK*$^+$ to *galK*$^-$

single colony were diluted 50-fold in LB-medium (10 ml) and grown to OD_{600} = 0.6, before λ-Red expression was induced by shifting cultures to 42°C (15 min). After chilling on ice (20 min), cells were harvested by centrifugation and washed with ice-cold sterile water (3× 1.5 ml). Cell pellets were resuspended in sterile water (50 µl) and transformed with DNA [100 ng of SSO, and 10 ng plasmid (when applicable), unless otherwise stated] by electroporation at 1.75 kV, 25 mF using the *E.coli* Pulser (Bio-Rad). LB-media (1 ml) was quickly added, then the cells were incubated at 32°C for 1.5 h with shaking, before being spread onto the appropriate selective LB-agar, or minimal galactose plates [see Ellis *et al.* (8)].

PCR

PCR was performed on a PTC-200 thermal cycler (Peltier Thermal Cycler Company). Each reaction included: Expand™ High Fidelity polymerase (1 U); polymerase buffer (10×, 5 µl, containing 10 mM $MgCl_2$); dNTPs (10 mM of each, 1 µl); primers (5 µM, 1 µl of each); ddH_2O added to 50 µl. Thermal cycling included an initial denaturation step of 94°C (5 min) followed by 35 cycles of: 50 s at 94°C, 50 s at 55°C and 50s at 72°C, followed by cooling to 4°C.

Plasmid construction

p(+)mKan and p(–)mKan. DNA inserts for plasmid constructs were amplified by PCR with Expand High Fidelity polymerase using primers containing restriction enzyme sites suitable for cloning. All plasmids were sequenced to check for possible mutations introduced by PCR. Primers mKan-1 and mKan-2 were used to amplify a 559 bp fragment from plasmid pGK-frt (Courtesy of Dr F. Stewart), which was digested with EagI–NcoI, then ligated with the 4287 bp EagI–NcoI fragment from pGK-frt to form pmKan. To construct p(+)mKan and p(–)mKan, the purified EcoRI–ClaI fragment from pmKan was ligated to the 2942 bp EcoRI–ClaI fragment of pBSKS and pBSSK, to form p(+)mKan and p(–)mKan, respectively.

p(+)Lac-mKan and p(–)Lac-mKan. Primers Plac-Kan-F and LacA-Kan-F were used to amplify a 894 bp DNA fragment from pGK-frt by PCR, which was inserted into pALA-1172 via Red-mediated homologous recombination in DY380 cells as described previously (4,8) to form the plasmid pALA1172-Lac-Kan. The purified 133bp BamHI–EagI, 1103 bp NcoI–SacII and 1762 bp BamHI–SacII fragments from pALA1172-Lac-Kan were ligated to the 526 bp EagI–NcoI fragment from pmKan to form pALA1172-Lac-mKan. The EcoRI–SphI DNA fragments from pALA1172-Lac-mKan (1219 bp) and pmKan (3325bp) were purified and ligated to form pLacI. The 1145 bp EcoRV–EagI fragment from pALA1172-Lac-mKan and the 3982 bp EagI–ClaI fragment from pmKan were ligated to form pLac-mKan. Finally, the 1193 bp EcoRI fragment from pLac-mKan was inserted into the EcoRI site of pLacI in both directions to form p(+)Lac-mKan and p(–)Lac-mKan.

Construction of mutants

To construct *E.coli* strain DY380(+), PCR was used to amplify a 2397 bp fragment from p(+)mKan using the primers DY380(+)-1 and DY380(+)-2. The purified DNA fragment (which contained the *mkan* gene and an *amp*^{r} selectable marker flanked by two regions homologous to the chromosomal target) was inserted into DY380 between the λ prophage *cro* gene and the *bioA* gene adjacent to the prophage [see Lee *et al.* (29)] via Red-mediated homologous recombination as described previously (4,8). *amp*^{r} colonies were selected, then screened by PCR using the primers DY380(+)-detect and PGK-frt-2 to positively identify those that contained the correctly inserted cassettes. Other mutant strains were similarly generated by Red-mediated recombination using appropriate PCR-generated cassettes (see Table 2). The construction of *E.coli* HME6 and HME41, which contain the normal and inverted *gal* operon containing a *galKam* mutation, respectively, have been described by Ellis *et al.* (8).

RESULTS

Recombinational strand bias is influenced by the SSO sequence

Two reporter plasmids, p(+)mKan and p(–)mKan (Fig. 1A and B) were constructed to determine whether the relative direction of transcription or DNA replication exerted any influence upon the strand bias of Red/SSO-mediated recombination. These plasmids contained a ColE1 replication origin, an ampicillin (*amp*) resistance gene and a mutated kanamycin (*kan*) resistance gene, under the control of a constitutive promoter (*Tn5*). Analogous to the reporter system established by Metz *et al*. (30), the *kan* gene had been engineered to contain a nonsense mutation (Y22Stop) in its coding sequence. This nonsense mutation meant that cells transformed with p(+)mKan or p(–)mKan would be resistant to ampicillin, but remain sensitive to kanamycin. The two plasmids were identical, apart from the *kan* reporter gene cassettes, which were orientated in opposing directions relative to the direction of DNA replication. This meant that while the sequence of the template and non-template strands remained the same in the two constructs, the replication lagging and leading strands were reversed.

Oligo 12A, a NT-SSO (a Non-Template-SSO, i.e. an oligo with a sequence that corresponded to the non-template strand in transcription, or to that of the mRNA) and oligo 12B, a T-SSO (a T-SSO, i.e. an oligo with a sequence that corresponded to the template strand in transcription, or complementary to that of the mRNA), were designed to repair this nonsense mutation and concomitantly create a silent mutation, with respect to the wild-type *kan*^{r} gene, by changing the TA**G** codon to TA**C** (i.e. stop to tyrosine; see Fig. 1G). Recombination-competent DY380 cells were co-transformed by electroporation with plasmid (10 ng) and SSO (100 ng unless otherwise stated). Recombination between the SSO and the target DNA bearing the nonsense mutation was indicated by the appearance of kanamycin-resistant (*kan*^{r}) colonies, with the efficiency of the recombination calculated by dividing the number of *kan*^{r} colonies by the number of the *amp*^{r} colonies. Sequencing of plasmid DNA isolated from *kan*^{r} colonies confirmed that only the intended changes were introduced. When expression of the recombination-promoting proteins (i.e. the Red proteins: Exo, Beta and Gam) was not induced, or when a homologous correction-SSO was not present, there was no growth of *kan*^{r} colonies. Similarly, transformation with an SSO of unrelated sequence did not generate *kan*^{r} colonies.

Table 2. Bacterial strains

Strain	Genotype
DH10B	F⁻ *mcrA Δ (mrr–hsdRMS–mcrBC) ϕ80dlacZΔ M15 Δ lacX74 deoR recA1 endA1 araΔ139* Δ *(ara, leu) 7649 galU galK rspL nupG*
DY330	W3310 ΔlacU169 gal490 λcI857 Δ(*cro-bioA*)
DY380	DH10B [λ*cI857* (*cro–bioA*)< >*tet*][a]
DY330(+)	DY330 [(*cro–bioA*)< >(*mkan-amp*^{r})][b]
DY330(–)	DY330 [(*cro–bioA*)< >(*amp*^{r}*-mkan*)][b]
DY380(+)	DY380 [(*cro–bioA*)< >(*mkan-amp*^{r})][b]
DY380(–)	DY380 [(*cro–bioA*)< >(*amp*^{r}*-mkan*)][b]
DY380Δ*mutS*	DY380 [*mutS*< >*cm*^{r}]
DY380(+)Δ*mutS*	DY380(+) [*mutS*< >*cm*^{r}]
DY380(–)Δ*mutS*	DY380(–) [*mutS*< >*cm*^{r}]
DY380(+)Δ*mutL*	DY380(+) [*mutL*< >*cm*^{r}]
DY380(–)Δ*mutL*	DY380(–) [*mutL*< >*cm*^{r}]
DY380(+)Δ*mutH*	DY380(+) [*mutH*< >*cm*^{r}]
DY380(–)Δ*mutH*	DY380(–) [*mutH*< >*cm*^{r}]
DY380(+)Δ*mfd*	DY380(+) [*mfd*< >*cm*^{r}]
DY380(–)Δ*mfd*	DY380(–) [*mfd*< >*cm*^{r}]
DY380(+)Δ*uvrB*	DY380(+) [*uvrB*< >*cm*^{r}]
DY380(–)Δ*uvrB*	DY380(–) [*uvrB*< >*cm*^{r}]
HME6	W3310 *galK*$_{tyr145UAG}$ λ*cI857* Δ*(cro-bioA)*
HME41	W3310 IN*gal*< >(*galM*$^+$ *K*$_{tyr145UAG}$ *T*$^+$*E*$^+$) λ*cI857* Δ*(cro-bioA)*[c]

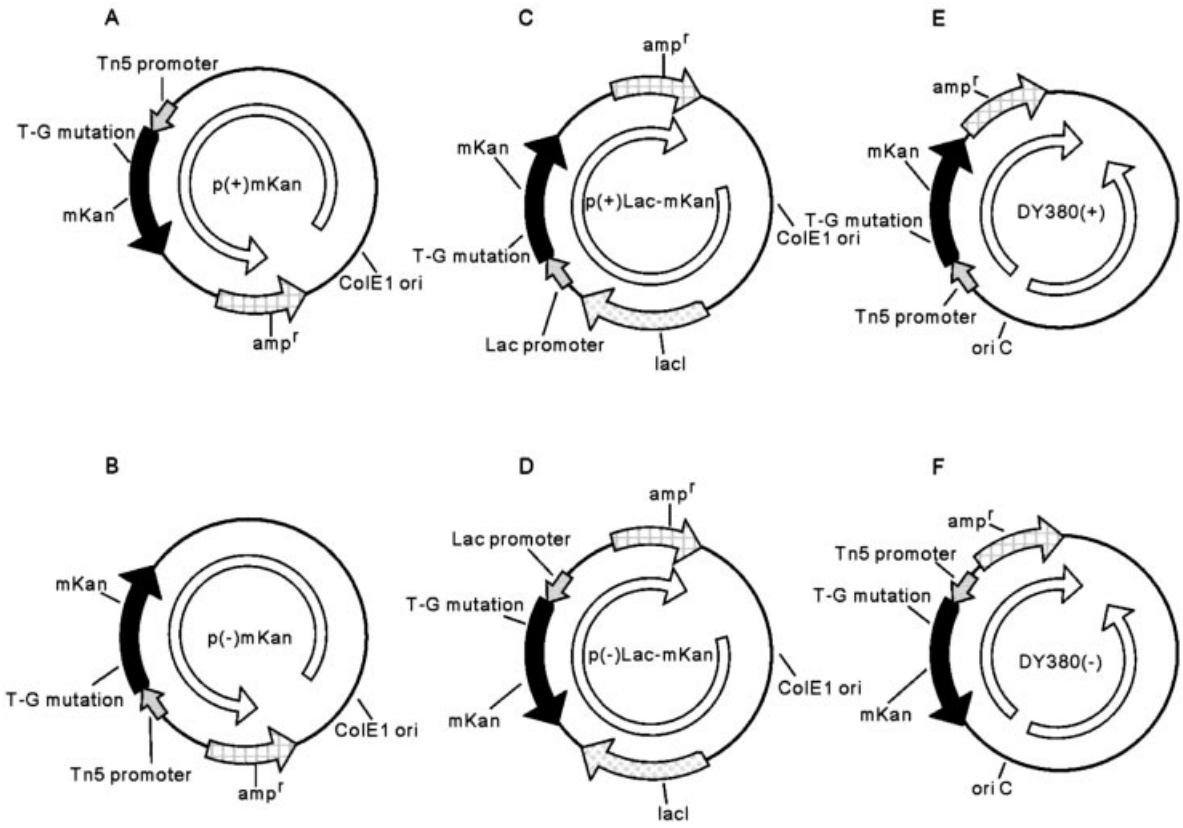
[a]The symbol < > indicates a replacement generated by recombineering technology. For example, *(cro–bioA)*< >*tet* indicates the substitution of *cro–bio*A with *tet.*
[b]*mkan-amp*^{r} and *amp*^{r}*-mkan* indicates that the *mkan* gene has been inserted in opposite directions in the two DY380 or DY330 strains.
[c]IN*gal* indicates that the *gal* gene has been inserted in opposite directions as in the HME6 strain.

As shown in Figure 2A and B, the NT-SSO (12A) was more efficiently recombined with the target gene at each concentration tested, regardless of its orientation in the plasmid. The strand bias ratio, which was defined as the recombination efficiency of the NT-SSO (12A) divided by that of the T-SSO (12B) targeted to the same site, was calculated for each set of recombination experiments. As may be seen in Figure 2, the strand bias ratio was >1 for all cases tested, indicating that it was determined predominantly by either the direction of transcription, or by an intrinsic sequence property of the NT- or T-SSO used. The recombinational strand bias was also strongly influenced by the direction of DNA replication. This was demonstrated by when a correction-SSO was both a NT- and Lag-SSO (i.e. an oligo with a sequence that corresponded to the lagging strand in replication, or to the Okazaki fragment), the resulting strand bias ratio was significantly larger than when an analogous NT-SSO was at the same time a Lead-SSO (i.e. an oligo with a sequence that corresponded to the Leading-strand in replication; compare Fig. 2B with A, or D with C).

To investigate whether this apparent recombination strand bias varied with the size of the SSO used, an analogous set of targeting experiments was conducted using oligos of 25 (6A and 6B), 73 (8A and 8B) and 91 nt (12A and 12B) in length. As was found previously, the NT-SSOs (6A, 8A and 12A) had a higher recombination efficiency than the T-SSOs (6B, 8B and 12B), regardless of length (Fig. 2C and D). To confirm that the observed mutational strand bias was not specific to the ColE1 replication origin, the same set of targeting experiments were conducted on an analogous plasmid containing a different replication origin (pSC101). As before, the NT-SSOs had a significantly higher recombination efficiency in all cases tested (data not shown).

A chromosome-based reporter gene system was constructed to discount the possibility that NT-SSOs recombined more efficiently than T-SSOs only when plasmids were used as recombination targets. The mutated kanamycin (*mkan*) resistance reporter gene cassette (under the control of the *Tn5* promoter) was inserted in both directions into the chromosome of *E.coli* strain DY380, concomitantly deleting the *tet* marker and part of the *bio* operon near 17 min 40 s on the chromosome, to create DY380(+) and DY380(–) (see Fig. 1E and F). In DY380(+), oligo 12A is both a NT- and a Lead-SSO (with respect to the direction of chromosomal replication), whilst in DY380(–) oligo 12A is a NT- and a Lag-SSO. Correspondingly, oligo 12B acts as a T- and Lag-SSO in DY380(+), and as a T- and Lead-SSO in DY380(–). As shown in Figure 3, the NT-SSO (oligo 12A) had the higher recombination efficiency in each case (i.e. a strand bias ratio >1), regardless of the orientation of the *mkan* gene on chromosome, or the length or concentration of SSO used. As was found for the plasmid-based reporter gene, the largest strand bias ratio was observed when the NT-SSO was also a Lag-SSO, i.e. the recombination efficiency was highest when the sequence of the correction-oligo corresponded concomitantly to the Okazaki fragment in replication and to the transcriptional non-template strand (Fig. 3B and D). In the most extreme case, a strand bias ratio of 450 was calculated (Fig. 3B, using 1000 ng of SSO). These data correlated well with that obtained from the plasmid-based reporter system, again indicating that DNA replication together with transcription and/or the sequence of the correction-SSO plays a major role in determining which strand is targeted more effectively.

An additional set of Red/SSO-mediated recombination assays were conducted in DY330, an *E.coli* W3110 *recA+* strain that harbors a defective lambda prophage (29), to confirm that the observed phenomenon was not strain specific. The *Tn5-mkan* reporter gene cassette was inserted into the DY330 chromosome in both directions, to create DY330(+) and DY330(–), which are analogous to DY380(+) and



G

Wild Type *kan*
(NT) GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA**T**GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC
(T) CGTGCGTCCAAGAGGCCGGCGAACCCACCTCTCCGATAAGCCGAT**A**CTGACCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAG
H A G S P A A W V E R L F G Y D W A Q Q T I G C S D A A V F

mkan
(NT) GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA**G**GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC
(T) CGTGCGTCCAAGAGGCCGGCGAACCCACCTCTCCGATAAGCCGAT**C**CTGACCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAG
H A G S P A A W V E R L F G *

12A NT 5′-GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA**C**GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC

12B T 3′-CGTGCGTCCAAGAGGCCGGCGAACCCACCTCTCCGATAAGCCGAT**G**CTGACCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAG

Corrected *kan*
(NT) GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA**C**GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC
(T) CGTGCGTCCAAGAGGCCGGCGAACCCACCTCTCCGATAAGCCGAT**G**CTGACCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAG
H A G S P A A W V E R L F G Y D W A Q Q T I G C S D A A V F

12A-TAT NT 5′-GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA**T**GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC

12B-ATA T 3′-CGTGCGTCCAAGAGGCCGGCGAACCCACCTCTCCGATAAGCCGAT**A**CTGACCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAG

Corrected *kan*
(NT) GCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA**T**GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC
(T) CGTGCGTCCAAGAGGCCGGCGAACCCACCTCTCCGATAAGCCGAT**A**CTGACCCGTGTTGTCTGTTAGCCGACGAGACTACGGCGGCACAAG
H A G S P A A W V E R L F G Y D W A Q Q T I G C S D A A V F

Figure 1. Key features of the plasmids and chromosomes containing the *mkan* reporter gene. The solid black arrows indicate the orientation of the genes, while the solid grey arrows specify the direction of transcription (under the control of the *Tn5* or *Lac* promoters). The open arrows inside the plasmid (**A**–**D**) or chromosome (**E** and **F**) denote the direction of replication from the *ori.* In (A), (C) and (E), the direction of transcription of the reporter genes is the same as that of DNA replication. In these reporter genes, the non-template strands (NT) are the same as the leading strands in replication. In (B), (D) and (F), the direction of transcription of the reporter gene is opposite to the direction of DNA replication, hence the NT strands are also the lagging strands in replication. In (A) and (B), and in (E) and (F), the mutated kanamycin (*mkan*) gene is under the control of the *Tn5* promoter, whilst in (C) and (D) it is under the control of the inducible *Lac* promoter. *oriC* is the *E.coli* chromosomal replication origin. (**G**) Shows that the wild type *kan* gene sequence (TAT), the amber mutant (TAG), the sequences of the 12A, 12B, 12A-TAT, 12B-ATA oligonucleotides and the corresponding sequences (TAC) or (TAT) after recombination. Drawings are not to scale.

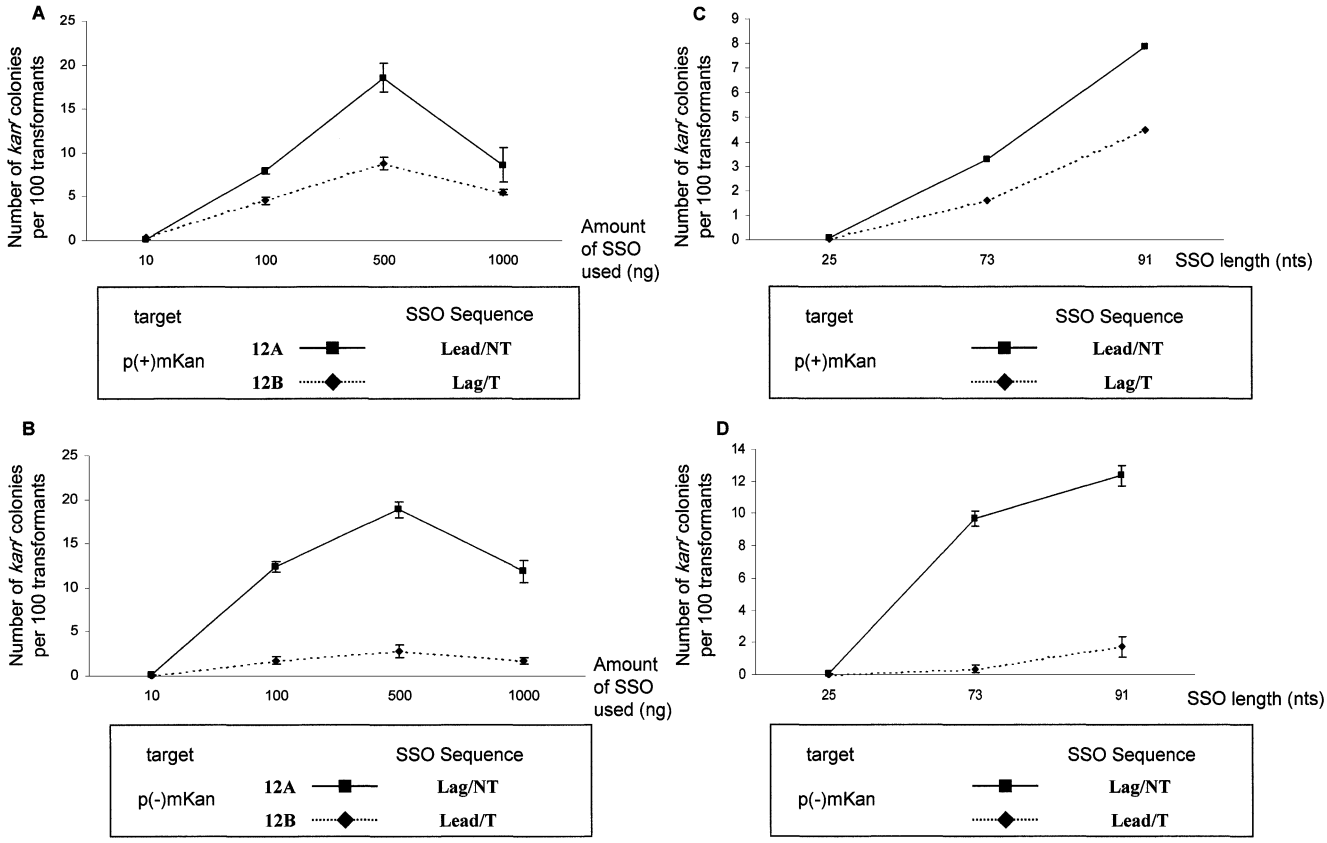


Figure 2. The strand bias observed in SSO-mediated recombination with plasmid DNA. Plasmid (10 ng) and SSO DNA were co-transformed into recombination-competent DY380. Aliquots of the plasmid/SSO-cotransformed cells were plated onto LB + Amp or LB + Amp + Kan plates, and after incubation overnight at 32°C, the number of colonies was counted. The plasmid and SSO used in each experiment are noted at the bottom of each panel under 'target' and 'SSO Sequence', respectively. T (template) or NT (non-template) indicates that the SSO has the same sequence as the transcriptional template, or the transcript, respectively. Lag or Lead indicates that the SSO has the same sequence as the newly synthesized Okazaki fragment or the leading strand in replication, respectively. The *y*-axis represents the number of *kan*^{r} colonies per 100 *amp*^{r} colonies after co-transformation of SSO and reporter plasmid. At least six independent experiments were carried out for each set of conditions. In (**A**) and (**B**) the length of SSO was kept constant (91mer), but the concentration was varied from 10 to 1000 ng (per reaction). In (**C**) and (**D**) three lengths of oligonucleotide were used: 25 (6A and 6B), 73 (8A and 8B) and 91 nt (12A and 12B), with the concentration maintained at 100 ng per reaction. All SSOs in this figure introduce TA**G** to TA**C** or **C**TA to **G**TA substitutions, with this mismatch located at the centre of the oligo.

DY380(–), respectively. DY380, derived from DH10B, is *recA* defective (29) and has several restriction and methylation defects not present in DY330, which is much more closely related to wild-type *E.coli* K-12. Moreover, DY380 is *deo*R$^-$, which facilitates the uptake of larger DNA fragments. Despite these numerous genetic differences, the NT-SSOs were again found to be substantially more recombinogenic than the T-SSOs (Fig. 3E and F). As was the case in the corresponding DY380 strains, the recombination efficiencies of the NT-SSOs were augmented substantially when their sequences also corresponded with that of the replicative lagging strand (i.e. they were also Lag-SSOs; see Fig. 3F).

To determine whether the higher recombination frequencies associated with the NT-SSOs were exerted through a transcription-related process, or were due to intrinsic sequence differences between the NT- and T-SSOs, a further two targeting oligos were designed and tested (see Fig. 1G). Oligo 12A-TAT (a NT-SSO) is identical to 12A except that it encodes for a TA**G** to TA**T** substitution, whereas oligo 12A changes the TA**G** codon to TA**C**. Similarly, oligo 12B-ATA (a T-SSO) creates a **C**TA to **A**TA substitution, whilst 12B changes the **C**TA to **G**TA. As shown in Figure 4A (Nos 1–4), the T-SSO (12B-ATA) underwent recombination with the target gene with substantially higher efficiency than the NT-SSO (12A-TAT), regardless of whether it was a Lag-SSO or a Lead-SSO. In contrast to what was found for SSOs 12A and 12B, the recombination efficiencies of SSOs 12A-TAT and 12B-ATA were highest when they were both a Lag- and T-SSO. These results suggested that the sequence of the correction-SSO, specifically the nature of the nucleotide mismatch it encoded for, strongly influenced the recombinational efficiency and strand bias.

Correction oligonucleotides were also targeted to the *galKam* gene in *E.coli* HME6 and HME41 (8) [which contain the normal and inverted *gal* operon, respectively, as well as an amber stop codon in the middle of the *galK* gene (Tyr145UAG)] to test whether the observed SSO sequence-effect was gene or location-specific. Regardless of being targeted to the replicative leading or lagging strand, oligo 144, which corrects the TA**G** codon to TA**C** was recombined with a significantly higher efficiency than oligo 145, which converts **C**TA to **G**TA (Table 3), indicating a sequence-specific bias

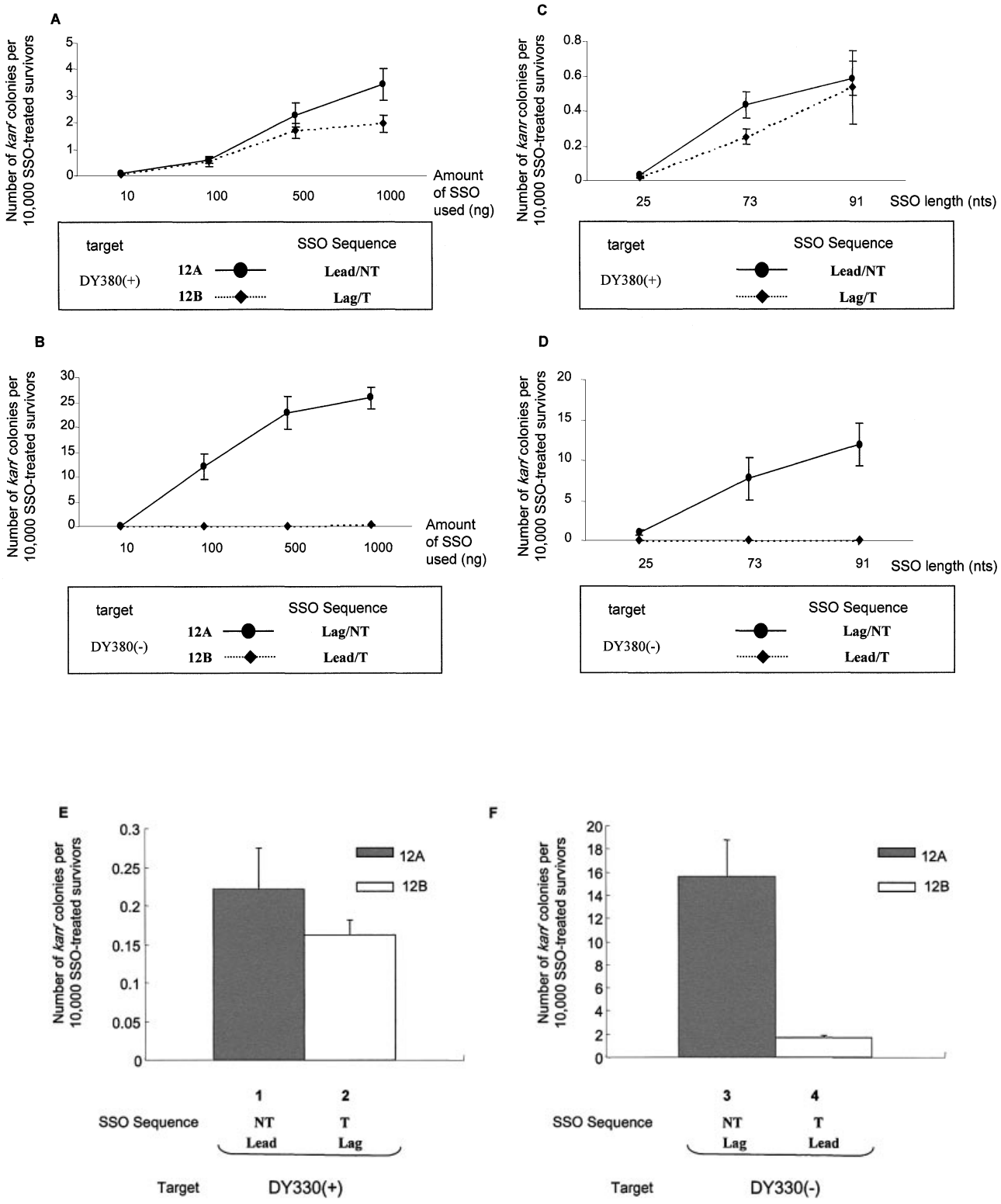
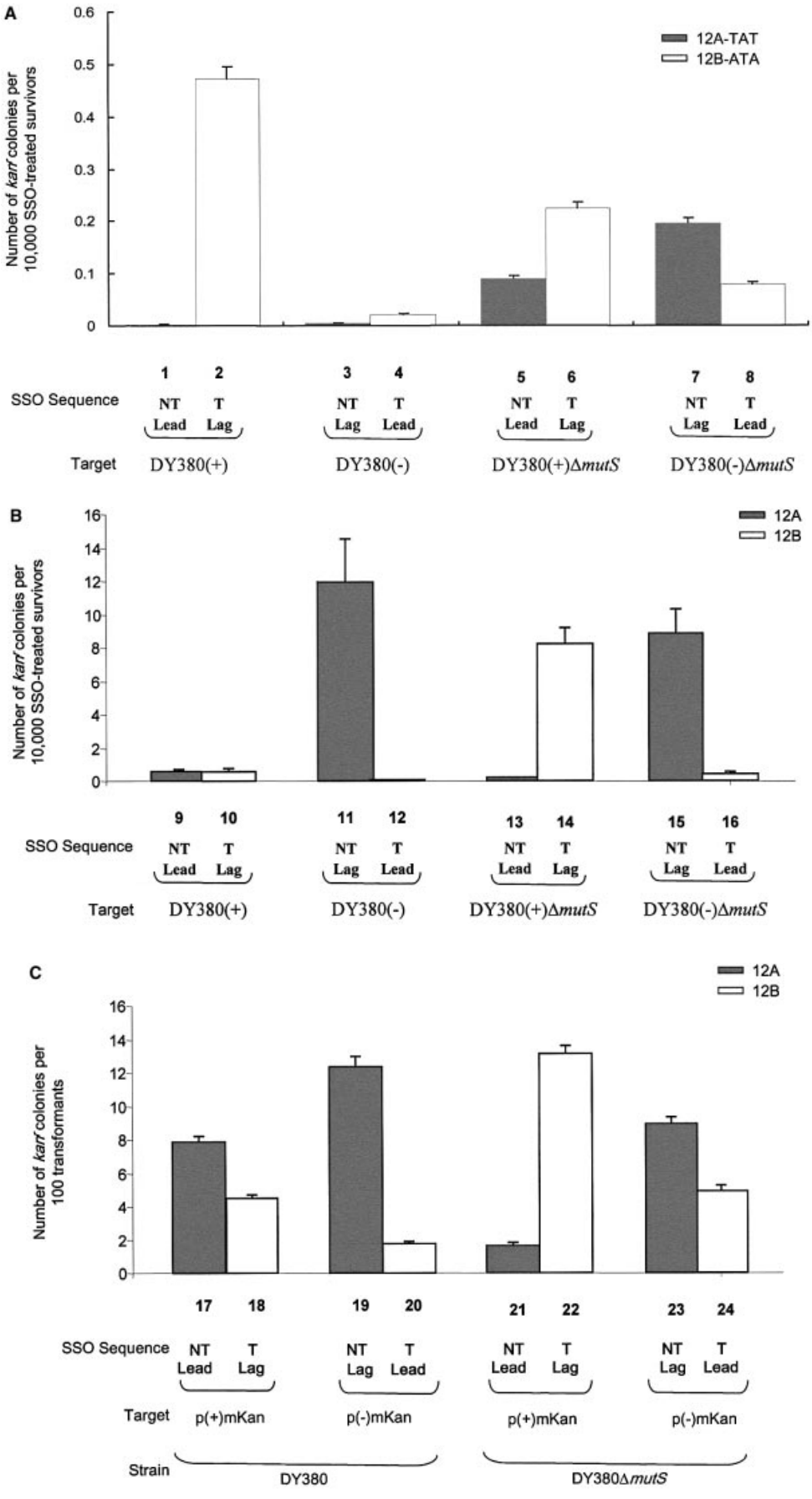


Figure 3. The strand bias observed in SSO-mediated recombination with chromosomal DNA. SSOs encoding for TA**G** to TA**C** or **C**TA to **G**TA substitutions were transformed into different DY380 and DY330 derived strains, as indicated at the bottom of each panel. Aliquots of the SSO-treated cells were spread onto LB + Amp or LB + Amp + Kan plates, with the number of colonies that appeared after overnight incubation at 32°C counted. The *y*-axis shows the number of *kan*^{r} colonies per 10 000 colonies after transformation of SSO (i.e. a measure of the number of recombinants among cells that 'survived' electroporation with SSO). In (**A**) and (**B**) the length of the correction-SSO was kept constant (91mer), but the concentration was varied from 10 to 1000 ng. In (**C**) and (**D**) three lengths of oligo were used: 25 (6A and 6B), 73 (8A and 8B) and 91 nt (12A and 12B), with the concentration maintained at 100 ng per reaction. In (**E**) and (**F**), oligos 12A and 12B (100 ng per reaction) were transformed into DY330(+) and DY330(–) as indicated. For (A) through to (F), a minimum of six independent experiments were performed for each set of conditions.



associated with TA**G** to TA**C** codon change. However, when oligos 100 and 101, which encoded for TA**G** to TA**T** and **C**TA to **A**TA alterations, respectively, were used to target the same reporter gene, it was the direction of replication that dictated the recombination efficiency, i.e. the Lag-SSO was more potent in recombination independent of whether it was concomitantly a NT- or T-SSO (Table 3). This was in agreement with the results obtained previously by Ellis *et al.* (8).

Mutations in the MMR pathway abolish the SSO sequence-associated strand bias

A set of *E.coli* reporter strains defective in MMR were constructed to investigate whether MMR influenced the Red/SSO-recombination efficiency, and to further probe the relationship between SSO-encoded mismatch and the mutational strand bias. These were generated by Red-mediated homologous recombination, through the singular replacement of the *mutS*, *mutH* or *mutL* genes with a chloramphenicol-resistance (cm^r) gene cassette, in both DY380(+) and DY380(–). Recombination experiments, identical to those conducted on the DY380(+) and DY380(–) strains, were performed on each of these new strains (Table 2) using correction-SSOs 12A, 12B, 12A-TAT and 12B-ATA, with representative results summarized in Figure 4 (Nos 5–24).

These experiments revealed that inactivation of the MMR pathway (through the deletion of either *mutS*, *mutL* or *mutH*) completely abolished the relationship between the recombinational correction efficiency and the sequence of the SSO-encoded mismatch. This left the strand bias dictated predominantly by the direction of DNA replication, so that the Lag-SSO had a higher recombination efficiency for each case tested, independent of its orientation relative to the direction of transcription. This held true for both chromosomal (Fig. 4A and B) and plasmid targets (Fig. 4C), for both the *mkan* and *galKam* (data not shown) reporter genes tested.

To investigate whether other DNA-repair processes, which transiently generate regions of single-stranded DNA, can affect the recombinational efficiency and strand bias, we performed an analogous set of targeting experiments with strains defective in TCR (31) and NER (32). These were prepared by replacing the *mfd* and *uvrB* genes with a cm^r cassette, in the *E.coli* DY380(+) and DY380(–) reporter strains. However, these mutations led to only minor changes in the overall recombination efficiency and in the strand bias ratio, compared with the corresponding wild type strains (data not shown), indicating that MMR, but neither TCR nor NER, plays an important role in modulating the efficiency of Red/SSO-recombination.

Transcription and Red/SSO-mediated recombination

In order to test whether transcription could directly influence the Red/SSO-mediated recombination, two new reporter plasmids were constructed: p(+)Lac-mKan and p(–)Lac-mKan (Fig. 1C and D). These were similar to the p(+)mKan and p(–)mKan plasmids, respectively, except that the constitutive *Tn5* promoter controlling the transcription of the *mkan* reporter gene was replaced by an inducible *Lac* promoter. Consequently, transcription of *mkan* could be regulated by the presence or absence of an inducer, IPTG. The stringency with which transcription was controlled in these reporter plasmids was first tested by substituting *mkan* with a wild type *kan* gene creating p(+)Lac-Kan and p(–)Lac-Kan. Without the addition of IPTG, p(+)Lac-Kan and p(–)Lac-Kan could not support the growth of kan^r colonies on LB-media plates containing kanamycin (15 μg/ml). However, when IPTG was added, both plasmids generated numerous and equivalent numbers of kan^r colonies. Furthermore, cells cultured in liquid LB-media with or without IPTG gave rise to near identical numbers of colonies when plated on to IPTG and kanamycin-containing LB-media plates, indicating that the transcriptional status of the *kan* gene prior to plating did not affect the survival rate. As the only difference between p(±)Lac-mKan and p(±)Lac-Kan was a point mutation in the coding region of the *kan* gene, this inferred that the transcription of *mkan* could be readily controlled by the addition/omission of IPTG.

For the recombination experiments, p(+)Lac-mKan and p(–)Lac-mKan were first transformed into DY380 to establish stable episomes. Cultures of DY380 cells containing these plasmids were grown in LB-media with IPTG, both before being made recombination-competent, and after electroporation with SSO. As controls, DY380 cultures containing these plasmids were similarly cultured and transformed with SSO in the absence of IPTG. Both the IPTG-treated and control cells were recovered in 20 ml of LB-media with or without IPTG, respectively, for 2 h before being spread onto LB-agar plates containing IPTG and kanamycin. The recombinational strand bias ratio was then calculated for the IPTG-treated cells and the controls, from the number of recombinant colonies that appeared. The addition of IPTG (i.e. induction of transcription) led to no significant changes in the strand bias measured for either the 12A/12B (Table 4) or the 12A-TAT/12B-ATA pairs of oligos (data not shown). This suggests that in Red/SSO-mediated recombination, the process of transcription is not directly involved in the determination of the inter-strand recombinational bias. However, a notable increase in the recombination efficiency was observed.

To examine the influence transcription exerts upon Red/SSO-mediated recombination with a chromosomal gene, oligo NC219 (creating a TAG nonsense mutation) was used to convert wild type *galK* gene to $galK^-$. The *gal* operon in *E.coli* is naturally repressed; however, the addition of fucose (which is a gratuitous inducer) stimulates *gal* transcription dramatically. Thus, recombination experiments could be conducted on the *galK* gene in a transcriptionally active or inactive state,

Figure 4. Recombinational strand bias in the absence of MMR. (**A**) SSOs (12A-TAT or 12B-ATA, 100 ng) and (**B**) SSOs (12A or 12B, 100 ng) were transformed into four different strains of recombination-competent *E.coli*, as indicated underneath each panel. Aliquots of the treated cells were spread onto LB + Amp or LB + Amp + Kan plates and after incubation overnight incubation at 32°C, the number of colonies that appeared was counted. The *y*-axis indicates the number of kan^r colonies per 10 000 survived colonies after transformation of SSO. (**C**) Plasmids (10 ng) and SSOs (12A or 12B, 100 ng) were co-transformed into the indicated recombination-competent *E.coli* strains. Aliquots were spread onto LB + Amp or LB + Amp + Kan plates, and after overnight incubation at 32°C, the number of colonies that appeared was counted. The *y*-axis shows the number of kan^r colonies per 100 amp^r colonies after co-transformation of SSO and reporter plasmid.

Table 3. SSO-sequence associated strand bias (chromosomal target)

Target	SSO[a]	gal$^+$/10^8 survivors[b]	Strand bias ratio[c]
HME6[a]	Oligo 144 (NT/Lag)	8.0×10^6	1000
HME6[a]	Oligo 145 (T/Lead)	8.0×10^3	
HME41[a]	Oligo 144 (NT/Lead)	4.0×10^5	8
HME41[a]	Oligo 145 (T/Lag)	5.0×10^4	
HME6[a]	Oligo 100 (NT/Lag)	3.0×10^5	8.6
HME6[a]	Oligo 101 (T/Lead)	3.5×10^4	
HME41[a]	Oligo 100 (NT/Lead)	2.0×10^4	0.033
HME41[a]	Oligo 101 (T/Lag)	6.0×10^5	

[a]NT, T, Lag and Lead indicates that the SSO sequence corresponds to the non-template, template, lagging or leading strand, respectively in the HME6 or HME41 strains.
[b]The columns show the number of *gal*$^+$ recombinants for every 10^8 cells that survived electroporation and transformation with SSO. Every experiment was repeated at least three times, with the numbers quoted being an average of each count.
[c]Strand bias ratio = number of *gal*$^+$ colonies (NT-SSO plate)/number of *gal*$^+$ colonies (T-SSO plate).

Table 4. Transcription and strand bias in Red/SSO-mediated recombination

Target	SSO	IPTG[b]	*kan*^{r}/10^4 survivors[c]	Strand bias ratio[d]
DY380/p(+)*mkan*[a]	12A (NT/Lead)		17	1.7
DY380/p(+)*mkan*[a]	12B (T/Lag)		10	
DY380/p(+)Lac-*mkan*[a]	12A (NT/Lead)	–	0.48	1.5
DY380/p(+)Lac-*mkan*[a]	12B (T/Lag)	–	0.31	
DY380/p(+)Lac-*mkan*[a]	12A (NT/Lead)	+	9.0	1.7
DY380/p(+)Lac-*mkan*[a]	12B (T/Lag)	+	5.3	
DY380/p(–)Lac-*mkan*[a]	12A (NT/Lag)	–	4.3	2.0
DY380/p(–)Lac-*mkan*[a]	12B (T/Lead)	–	2.2	
DY380/p(–)Lac-*mkan*[a]	12A (NT/Lag)	+	13	2.4
DY380/p(–)Lac-*mkan*[a]	12B (T/Lead)	+	5.5	

[a]Plasmids were established as stable episomes in the host strain before transformation with correction-SSO (100 ng).
[b]The (+) or (–) sign indicates whether IPTG was added to or omitted from the culture media prior to plating, respectively.
[c]The columns show the number of *kan*^{r} recombinants for every 10^4 cells that survived electroporation and transformation with SSO.
[d]Strand bias ratio = number of *kan*^{r} colonies (NT-SSO plate)/number of *kan*^{r} colonies (T-SSO plate).

through the addition or omission of fucose from the growth medium. We found that fucose neither increased nor reduced *galK*$^-$ recombinants (data not shown).

DISCUSSION

Mutational studies in *E.coli* have demonstrated previously that of the three Red proteins from bacteriophage lambda, Beta is the only one that is required for SSO-mediated recombination activity (8). Exo is only required when double-stranded DNA targeting cassettes are used, whilst Gam potentiates the efficiency of the recombination process through the inhibition of intracellular exonucleases. However, the exact mechanism by which Beta mediates oligonucleotide-directed recombination, and the factors that affect its activity and efficiency remain to be firmly established.

According to the current working model (2,8,10), the first step in the recombination process involves the Beta-coated SSO hybridizing to a region of single-stranded DNA transiently generated during DNA replication. From *in vitro* studies, and with comparison to its 'natural' λ-Exo generated substrate (33), the Beta-SSO nucleoprotein filament may promote annealing with the target ssDNA locus, preferentially with 3′ to 5′ polarity, driving the hybridization through mismatches to form a Beta-bound duplex moiety. Thus, any cellular processes generating single-stranded DNA regions may provide entry points for Beta-coated SSO. It has been postulated that the Beta-coated SSO may access the areas of single-stranded DNA during DNA replication, transcription, DNA supercoiling or during DNA damage repair (2,8). If these processes are asymmetric, such asymmetry may be reflected in the strand bias in Red/SSO-mediated recombination. During replication, the higher recombinogenic efficiency of the SSO that primes the lagging strand, may reflect the increased amount of single-stranded DNA present on the lagging strand template relative to that present on the leading strand template that is available for annealing (8,10,34,35). As transcription and some DNA damage repair processes operate with strand-specificity (i.e. are strand asymmetrical), any associated Beta/SSO recombination events may be directed preferentially to a particular strand. This may possibly be due to steric crowding; specific protein–protein and/or protein–DNA interactions; or through a more favorable orientation of the bases on one strand relative to the other, resulting in a quantifiable targeting strand bias.

In order to identify external factors and cellular processes that contributed to the strand bias observed in Red/SSO recombination, we constructed a number of different reporter systems, in which the reporter gene had a defined orientation relative to the direction of replication and transcription, and where the transcriptional state could be controlled. With these reporter systems, we then tested the properties of SSOs and several cellular processes, which generate transient single-stranded DNA regions, for their influences on the efficiency and strand bias of Red/SSO-mediated recombination.

The sequence of the SSO appears to be one major factor determining the recombinational strand bias in the targeted correction of point mutations in both chromosomal and plasmid-based reporter genes. In *E.coli* strains with wild type MMR activities, non-template SSOs that encoded for a TA**G** to TA**C** mutation were found to be consistently more potent in recombination than the complementary T-SSOs that created a **C**TA to **G**TA substitution. This was independent of the orientation of the target gene relative to the direction of replication, and held true for both the *mkan* and *galKam* reporter systems tested. Over the range of different oligonucleotide lengths and concentrations tested, the *mkan* reporter gene, whether located chromosomally or on a plasmid, was consistently mutated most efficiently with a non-template oligo that encoded for a (TA**G** to TA**C**) alteration (Fig. 3). When targeted to chromosome and plasmid-based reporter genes, this type of correction-oligo (TA**G** to TA**C**) had the highest recombinational efficiency when its sequence corresponded to that of the lagging strand in replication (Fig. 4B, compare columns 9 and 11). In the most extreme case, the NT-SSO (12A) changed TA**G** to TA**C** 450 times more efficiently than the corresponding T-SSO (12B).

Conversely, when oligos that encoded for TA**G** to TA**T** (12A-TAT), or **C**TA to **A**TA (12B-ATA) substitutions were

targeted to the same genetic loci of *mkan*, a higher recombination efficiency was always found to be associated with the template oligo (**C**TA to **A**TA). When oligos 100 (TA**G** to TA**T**) and 101 (**C**TA to **A**TA) were similarly targeted to the chromosomal *galKam* reporter, we again confirmed that the strand bias was now dictated by the direction of replication as reported previously (8), i.e. the Lag-SSO was always more potent in recombination, regardless of the direction of transcription. The results obtained with *mkan* and *galKam* reporter systems indicate that the polarity and magnitude of the strand bias is determined both by the direction of DNA replication and the nature of the SSO-encoded mismatch, with the outcome at any specific locus determined by the relative contributions from these two factors.

To investigate whether the MMR process influenced the efficiency and strand bias of Red/SSO-recombination in *E.coli*, a selection of experiments were conducted in DY380 strains where the *mutL*, *mutH* or *mutS* genes were singularly deleted, using both episomal and chromosomal reporter systems. Deletion of any one of these three genes generated very similar Red/SSO-recombination mutational efficiencies (data not shown), with only the Δ*mutS* data included with this report for the sake of brevity (see Fig. 4). When *galKam* was used as the reporter, the recombination efficiency increased dramatically in *mutS*, *mutH* and *mutL* mutants, compared with the corresponding wild type strains HME6 and HME41 (N. Costantino and D. Court, unpublished results). This increase in recombination efficiency was also found for the *kan* reporter gene in some cases (Fig. 4B, compare 10 versus 14). This observation is consistent with results obtained previously by Dekker *et al.* (36), who found that MMR strongly suppressed the efficiency of gene modification by SSOs in murine embryonic stem cells. As DY380 is a *recA1* strain (derived from DH10B), our results indicate that the Red-mediated recombination of ssDNA oligonucleotides can occur in *recA* and *mutS* double mutant strains. This is markedly different to the findings of Metz *et al.* (30), who reported that both *recA* and *mutS* were required in *E.coli*, when RNA/DNA chimeric oligonucleotides (RDO) were used to correct a similarly designed point mutation in the *kan* gene. These contrasting results suggest that RDO- and Red/SSO-mediated recombination either follow different mechanistic pathways, or that the Red proteins can somehow substitute for both RecA and MutS activities at the same time.

In *E.coli* strains with wild type MMR activities, the strand bias was influenced by the nature of the mismatch encoded by the SSO, and by the direction of DNA replication as discussed above. A NT-SSO creating a TA**G** to TA**C** mutation underwent recombination more efficiently than its corresponding T-SSO, regardless of the direction of replication. However, in *E.coli* strains defective in MMR, the nature of the mismatch encoded by the oligo had no effect upon the strand bias, leaving it to be determined mainly by the direction of replication, i.e. the Lag-SSO was always the more recombinogenic. This was true for all SSO pairs tested, and for both *mkan* and *galKam* reporter genes. These results are best explained by the MMR system recognizing and repairing mismatches at (or near) the replication fork. Comparing the two SSOs that create TA**G** to TA**C** (12A) and **C**TA to **G**TA (12B) substitutions, oligo 12A will form a C/C mispair with the plasmid or chromosome DNA, while 12B will form a G/G mispair. The C/C mismatch is neither detected nor eliminated by the MMR machinery as efficiently as the G/G mismatch (28). Consequently, when MMR is functional, the oligo that introduces a C/C mismatch (NT-SSO 12A) should have a higher recombination efficiency than the oligo that forms a G/G mismatch (T-SSO 12B). This rationale was supported by our experimental results, for both plasmid and chromosomal-based targets (Fig. 4B and C; compare 9 and 10, 11 and 12, 17 and 18, 19 and 20), indicating that the sequence of the correction-SSO may override any replication-imposed strand bias, when the full complement of MMR proteins are present. Consistent with these results, SSOs creating T/T and C/C mispairs in yeast exhibited the highest recombination efficiency (37).

For oligos that create TA**G** to TA**T** and **C**TA to **A**TA substitutions, T/C and A/G mispairs are encoded by the NT-SSO and T-SSO, respectively. These two types of mis-pairs are treated relatively equally by the MMR system, leaving the strand bias dictated predominantly by the direction of replication. For the *galKam* reporter, this is illustrated by oligos 100 and 101 being more recombinogenic when acting as lagging strand-oligos (data not shown). Therefore, MMR only contributes to the strand bias when the two SSO-encoded mis-pairs are corrected with differing efficiencies. However, with the *mkan* reporter system, the T-SSO (12B-ATA) was always more potent than the NT-SSO (12A-TAT) (Fig. 4A, compare 1 and 2, 3 and 4), suggesting that the sequence adjacent to the mismatched base pair may also influence the mis-match recognition and repair processes. In MMR defective strains, the Lag-SSO always has a higher efficiency, no matter what mismatch it encodes for. It will be interesting to see whether a similar relationship between MMR activity, strand bias and the sequence of correction-SSOs exists in eukaryotic systems.

To evaluate the contribution of other single-stranded DNA-generating processes to the overall efficiency and strand bias of Red/SSO-mediated recombination, we generated mutations that inactivated the TCR or NER machinery, through deletion of the *mfd* or *uvrB* genes. These mutations resulted in a recombination efficiency, and strand bias that was essentially identical to that found for the corresponding wild type system (data not shown). We notice that this observation does not agree with a model proposed by Igoucheva *et al.* that SSO-mediated recombination in mammalian cells involves TCR (17). However, our results indicate (perhaps not surprisingly) that the Beta/SSO-filament cannot opportunistically anneal with the short, transient regions of ssDNA generated during the nucleotide excision process.

Targeting genes under conditions known to have a large effect on transcription, the transcriptional state was found to have negligible influence on the strand bias with both plasmid (Table 4) and chromosomal reporters. This is in sharp contrast to the data published by Kmiec and co-workers (12) who found that transcription affected the strand bias of SSO-mediated recombination in yeast. They argued that a NT-SSO bound to the template strand is preferentially displaced by RNA polymerase during transcription, whilst a T-SSO binding to the non-template strand is affected to a lesser extent. While this may be true in some cases, it cannot explain our observations in *E.coli*. Their model also failed to explain the existence of a strand bias in other transcriptionally inactive

genes in yeast (19). We observed a notable increase in the overall efficiency of Red/SSO-mediated recombination with the *mkan* reporter gene upon transcriptional activation, but not with the *gal* reporter gene. We are currently investigating the mechanistic basis for this apparent discrepancy. However, a great deal of work will be required to fully establish the relationship between transcription and SSO-mediated recombination within these, and other species.

In summary, we have shown that MMR exerts an important influence on the Red/SSO-mediated recombination process in *E.coli*, even though it is not essential for activity. MMR limits the overall recombination efficiency and may impose a strand bias by preferentially repairing some, but not all types of oligoncleotide-introduced mismatches. The direction of replication strongly influences the recombinational strand bias; however, our results suggest that transcription and other cellular processes transiently generating asymmetric single-stranded regions (such as TCR) do not play a major role. This study provides a number of key insights into the cellular processes involved in Red/SSO-mediated oligonucleotide recombination in *E.coli*, and should act as a guide to future research into this fascinating area.

ACKNOWLEDGEMENTS

We thank Dr David Smith for editing the Abstract and Dr Francis Stewart for the pGK-frt plasmid. This work was supported by a joint grant from the National Science Foundation of China and the Research Grant Council of Hong Kong (NSFC/HKU 17) to J.D.H. and D.L., and by a CRCG Seed Funding grant (University of Hong Kong) to J.D.H.

REFERENCES

1. Copeland,N.G., Jenkins,N.A. and Court,D.L. (2001) Recombineering: a powerful new tool for mouse functional genomics. *Nature Rev. Genet.*, **2**, 769–779.
2. Court,D.L., Sawitzke,J.A. and Thomason,L.C. (2002) Genetic engineering using homologous recombination. *Annu. Rev. Genet.*, **36**, 361–388.
3. Zhang,Y., Buchholz,F., Muyrers,J.P.P. and Stewart,A.F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genet.*, **20**, 123–128.
4. Yu,D., Ellis,H.M., Lee,E.C., Jenkins,N.A., Copeland,N.G. and Court,D.L. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **97**, 5978–5983.
5. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA*, **97**, 6640–6645.
6. Murphy,K.C. (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J. Bacteriol.*, **180**, 2063–2071.
7. Murphy,K.C., Campellone,K.G. and Poteete,A.R. (2000) PCR-mediated gene replacement in *Escherichia coli*. *Genetics*, **246**, 321–330.
8. 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. USA*, **98**, 6742–6746.
9. Swaminathan,S., Ellis,H.M., Waters,L.S., Yu,D., Lee,E.C., Court,D.L. and Sharan,S.K. (2001) Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis*, **29**, 14–21.
10. Zhang,Y., Muyrers,J.P., Rientjes,J. and Stewart,A.F. (2003) Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. *BMC Mol. Biol.*, **4**, 1.
11. Moerschell,R.P., Tsunasawa,S. and Sherman,F. (1988) Transformation of yeast with synthetic oligonucleotides. *Proc. Natl Acad. Sci. USA*, **85**, 524–528.
12. 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.
13. Liu,L., Rice,M.C. and Kmiec,E.B. (2001) *In vivo* gene repair of point and frameshift mutations directed by chimeric RNA/DNA oligonucleotides and modified single-stranded oligonucleotides. *Nucleic Acids Res.*, **29**, 4238–4250.
14. Gamper,H.B., Parekh,H., Rice,M.C., Bruner,M., Youkey,H. and Kmiec,E.B. (2000) The DNA strand of chimeric RNA/DNA oligonucleotides can direct gene repair/conversion activity in mammalian and plant cell-free extracts. *Nucleic Acids Res.*, **28**, 4332–4339.
15. Igoucheva,O., Alexeev,V. and Yoon,K. (2001) Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther.*, **8**, 391–399.
16. 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.
17. 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.
18. Yamamoto,T., Moerschell,R.P., Wakem,L.P., Komar-Panicucci,S. and Sherman,F. (1992) Strand-specificity in the transformation of yeast with synthetic oligonucleotides. *Genetics*, **131**, 811–819.
19. Yamamoto,T., Moerschell,R.P., Wakem,L.P., Ferguson,D. and Sherman,F. (1992) Parameters affecting the frequencies of transformation and co-transformation with synthetic oligonucleotides in yeast. *Yeast*, **8**, 935–948.
20. Jiricny,J. (1998) Replication errors: challenging the genome. *EMBO J.*, **17**, 6427–6436.
21. Su,S.S. and Modrich,P. (1986) *Escherichia coli* mutS-encoded protein binds to mismatched DNA base pairs. *Proc. Natl Acad. Sci. USA*, **83**, 5057–5061.
22. 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.
23. Parker,B.O. and Marinus,M.G. (1992) Repair of DNA heteroduplexes containing small heterologous sequences in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **89**, 1730–1734.
24. Joshi,A. and Rao,B.J. (2002) ATP hydrolysis induces expansion of MutS contacts on heteroduplex: a case for MutS treadmilling? *Biochemistry*, **41**, 3654–3666.
25. Lopez de Saro,F.J. and O'Donnell,M. (2001) Interaction of the beta sliding clamp with MutS, ligase and DNA polymerase I. *Proc. Natl Acad. Sci. USA*, **98**, 8376–8380.
26. Umar,A., Buermeyer,A.B., Simon,J.A., Thomas,D.C., Clark,A.B., Liskay,R.M. and Kunkel,T.A. (1996) Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell*, **87**, 65–73.
27. Gu,L., Hong,Y., McCulloch,S., Watanabe,H. and Li,G.M. (1998) ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. *Nucleic Acids Res.*, **26**, 1173–1178.
28. Kramer,B., Kramer,W. and Fritz,H.J. (1984) Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*. *Cell*, **38**, 879–887.
29. Lee,E.C., Yu,D., Martinez de Velasco,J., Tessarollo,L., Swing,D.A., Court,D.L., Jenkins,N.A. and Copeland,N.G. (2001) A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics*, **73**, 56–65.
30. Metz,R., DiCola,M., Kurihara,T., Bailey,A., Frank,B., Roecklein,B. and Blaese,M. (2002) Mode of action of RNA/DNA oligonucleotides: progress in the development of gene repair as a therapy for alpha(1)-antitrypsin deficiency. *Chest*, **121**, 91S–97S.
31. Svejstrup,J.Q. (2002) Mechanisms of transcription-coupled DNA repair. *Nature Rev. Mol. Cell. Biol.*, **3**, 21–29.
32. VanHouten,B. (1990) Nucleotide excision repair in *Escherichia coli*. *Microbiol. Rev.*, **54**, 18–51.
33. Li,Z., Karakousis,G., Chiu,S.K., Reddy,G. and Radding,C.M. (1998) The beta protein of phage lambda promotes strand exchange. *J. Mol. Biol.*, **276**, 733–744.

34. Kim,S., Dallmann,H.G., McHenry,C.S. and Marians,K.J. (1996) Coupling of a replicative polymerase and helicase: a tau-DnaB interaction mediates rapid replication fork movement. *Cell*, **84**, 643–650.
35. Marians,K. (1996) In Neidhardt.F.C., Curtis,R.,III, Ingraham,J.L., Lin,E.E.C., Low,R.B., Magasanik,B., Reznikoff,W.S., Riley,M., Schaechter,M. and Umbarger,H.E. (eds), *Escherichia coli* and *Salmonella*. ASM Press, Washington DC, Vol. 1, pp. 749–763.
36. 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.
37. Brachman,E.E. and Kmiec,E.B. (2003) Targeted nucleotide repair of cyc1 mutations in *Saccharomyces cerevisiae* directed by modified single-stranded DNA oligonucleotides. *Genetics*, **163**, 527–538.