High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides

Hilary M. Ellis, Daiguan Yu*, Tina DiTizio, and Donald L. Court†

Gene Regulation and Chromosome Biology Laboratory, Division of Basic Sciences, National Cancer Institute at Frederick, P.O. Box B, Frederick, MD 21702

Communicated by Sankar Adhya, National Institutes of Health, Bethesda, MD, April 3, 2001 (received for review February 16, 2001)

Homologous DNA recombination is a fundamental, regenerative process within living organisms. However, in most organisms, homologous recombination is a rare event, requiring a complex set of reactions and extensive homology. We demonstrate in this paper that Beta protein of phage λ generates recombinants in chromosomal DNA by using synthetic single-stranded DNAs (ssDNA) as short as 30 bases long. This ssDNA recombination can be used to mutagenize or repair the chromosome with efficiencies that generate up to 6% recombinants among treated cells. Mechanistically, it appears that Beta protein, a Rad52-like protein, binds and anneals the ssDNA donor to a complementary single-strand that generates recombinants in vivo. ssDNA overhangs to which Beta protein can use short (50-bp) homologous recombination functions each in a RecA-like fashion. We note, however, that oligonucleotides obtained from some other sources were not as active as the ones we used here. The sequences of the most active of the two oligos used for curing each Tn10 insertion are indicated below: 

**Materials.** Oligonucleotides (oligos) were provided by Life Technologies (Frederick, MD) as salt free but were not purified in any special way. We note, however, that oligonucleotides obtained from some other sources were not as active as the ones we used here. The sequences of the most active of the two oligos used for curing each Tn10 insertion are indicated below: 

**Construction of HME31, a Genetic Insertion.** The cat-sacB cassette, carrying genes for chloramphenicol resistance and sucrose sensitivity, was amplified by PCR using oligos that contain homologies with the galK gene. The oligos are GACCGTTAGCAGCGATCCCGCTGCGGTTCCAGAATATCGGCAACACG; and GACCGTTAGCAGCGATCCCGCTGCGGTTCCAGAATATCGGCAACACG. The galK homology on the 5' end and the cat-sacB primers on the 3' end of each oligo. The PCR product was used for homologous recombination in the strain HME6. Recombinants that had inserted the cassette into galK were selected as chloramphenicol-resistant colonies and tested as described (4).

**Construction of HME41, a Genetic Inversion.** To create an inversion of the gal operon, the operon was first deleted by recombination with the gal<del>-specific oligonucleotide (see below) into...
Results

Homologous Recombination with ssDNA. Activities of the λ Red proteins suggested that cells expressing the Red functions might be able to use a simple single-stranded oligo as a substrate for recombination. To test this idea, we asked whether the λ Red system could use a 70-base single-stranded oligo to correct the amber mutation galK_{A124U5} (4) located on the bacterial chromosome. For this purpose, we synthesized an oligo corresponding to the non-template DNA strand of galK (see Fig. 1). Strains HME6 and HME9 carrying the amber mutation were induced with the gal operator, followed by 40 bases of sequence downstream of galM. The 81-base gal2>del sequence is 5’-GGGATTTAATCTTTGATGTAAATTCTTGTGTAAACGA- TTCC/GGCTTTATCTTCAAATGACACGCAAAACCAGA-A and was selected as temperature-resistant survivors because the slash represents the deletion junction. To reinsert the gal operon in the inverted orientation, we created a PCR product of the gal operon in which homology segments found upstream (44 bases) and downstream (43 bases) flanked the complete operon, which was in inverted orientation with respect to these flanking chromosomal homology regions. PCR primers were 5’ATTGATTTTACGTGTGGATTGTGATTTGAAACAATGAGTAC, in which the slash represents the deletion junction. To reinvert the gal operon in the inverted orientation, we created a PCR product of the gal operon in which homology segments found upstream (44 bases) and downstream (43 bases) flanked the complete operon, which was in inverted orientation with respect to these flanking chromosomal homology regions. PCR primers were 5’CATCATAAACCTCCTGTATATCTTTAATCAGGACGATATAAA and the chromosome. This PCR product was crossed with the 70-base oligo, which gave 2.2 × 10^6 recombinants per 10^8 viable cells, the 60-, 50-, and 40-base oligos recombined less efficiently, yielding 4.5 × 10^5 recombinants per 10^8 cells. Recombination efficiency dropped to 5.2 × 10^5 recombinants per 10^8 viable cells with the 30-base oligo, and Gal+ colonies were near background levels with a 20-base oligo. A similar requirement for homology length was observed for recombination of linear dsDNA using the λ Red system (4). The drop in recombination efficiency from 10^2 to 10^3 is consistent with the observation that Beta binds poorly to oligos shorter than 36 bases (17).

Our demonstration that recombination with ssDNA could change a single base on the chromosome at galK led us to test whether deletions could also be engineered by using ssDNA. To compare the efficiency of these two types of changes using our protocol, we first inserted a 3.3-kbp chloramphenicol resistance cassette at the galK amber site in HME6, creating strain HME31. HME6 and HME31 were induced for Red function and electroporated with the 70-base oligo described above. Gal+ recombinants are generated with the same frequency in both strains, demonstrating that single strand-mediated recombination can create a 3.3-kbp deletion as efficiently as a single base change.

ssDNA Recombination Requires Only Beta. To determine which λ functions are required for recombination with ssDNA, we created a series of replacements of λ genes by a cat cassette (Fig. 2). The physical structure of each of these replacements was verified by PCR. We then tested each of these mutant prophage strains for efficiency of gal recombination with the 70-base oligo. The results (Fig. 2) indicate that only λ Beta is absolutely required for ssDNA recombination with the chromosome. Deletion of exo had no effect, and deletion of gam caused a minor 5-fold reduction in efficiency, suggesting that Gam, although not required, enhances recombination efficiency. To confirm that Beta was the only λ function required for ssDNA recom-
of transcription at galK or some other asymmetry, an inversion of the entire gal operon containing the galK amber mutation was created (see Materials and Methods). If the bias results from the asymmetry of transcription, the non-template strand oligo should still recombine more efficiently than the template strand oligo within the gal inversion. Instead, in this case, recombination of the template strand oligo in the gal inversion strain was 14-fold more efficient than recombination of the non-template strand oligo (Table 1). Thus, the strand bias observed does not result from the asymmetry of transcription. The strand bias is also not a consequence of simple sequence differences between the two oligos because either one may be more recombinogenic depending on orientation of the operon within the chromosome. As shown below, our results are more consistent with a bias caused by the direction of replication through the region containing the gal operon.

**Strand Bias Correlates with Direction of Replication Through the Recombination Site.** To determine whether strand bias is a general phenomenon or specific to galK, recombination at five additional genes on the chromosome was examined (Fig. 3). Each gene tested contained a Tn10 insertion causing an auxotrophic defect

---

**Table 1. Strand bias of recombination at galK**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Oligo</th>
<th>Gal+ recombinants†</th>
</tr>
</thead>
<tbody>
<tr>
<td>HME6</td>
<td>NT</td>
<td>1.5 x 10⁵</td>
</tr>
<tr>
<td>HME6</td>
<td>T</td>
<td>0.3 x 10⁵</td>
</tr>
<tr>
<td>HME41</td>
<td>NT</td>
<td>0.1 x 10⁵</td>
</tr>
<tr>
<td>HME41</td>
<td>T</td>
<td>1.4 x 10⁵</td>
</tr>
</tbody>
</table>

*Cells were induced for 15 min at 42°C and electroporated with 200 ng of either the non-template (NT) or template (T) oligo. HME41 is HME6 with the gal operon inverted. By convention, the non-template DNA has the same sequence as the transcript.

†Total recombinant colonies on minimal galactose medium per 10⁸ viable cells following electroporation.

---

**Fig. 2.** The defective λ prophage. The prophage contains λ genes from cl to int (4). Some λ genes are not shown for clarity. A deletion (Δ) removes cro through attR from the right side of the prophage, ρ, and the arrow indicate the early left promoter and transcript used to express Gam, Beta, and Exo. The C857 repressor prevents ρ expression at 32°C but is inactive at 42°C. attL and attR indicate the left and right junction of the prophage. Gene replacements with the cat cassette and a simple deletion Δ(gam-N) are shown below the prophage map and are located to the left of the strain name they represent. For the experiment, cells were induced for 15 min at 42°C and electroporated with 200 ng of the 70-base oligo of Fig. 1. The Gal+ recombinants were selected on minimal galactose medium, and the numbers are normalized per 10⁸ cells surviving each electroporation.

---

**A Strand Bias in Recombination Levels When Comparing Two Complementary ssDNAs.** Although Beta can mediate strand exchange between ssDNA and a target DNA with single-stranded ends (13), it could not be shown to mediate strand invasion of a fully double-stranded DNA (22). Because there is no requirement for RecA, recombination of ssDNAs may not involve strand invasion. Instead, we believe that the recombination events described here involve pairing of the Beta-bound oligo at transiently single-stranded regions of the chromosome. Single-stranded regions are formed during replication of the chromosome, repair of DNA, and transcription-induced supercoiling.

The 70-base oligo corresponding in sequence to the non-template strand for transcription of galK generated ~5-fold more Gal+ recombinants than the complementary oligo corresponding to the template strand (Table 1). A strand bias of 4- to 10-fold was seen in each of four independent experiments. This strand bias might result from differences in oligo sequence, or from inherently asymmetric processes in the cell, such as transcription. To test whether strand bias resulted from the direction
Table 2. Strand bias in recombination at several locations

<table>
<thead>
<tr>
<th>Strain and (relevant allele)</th>
<th>Oligo†</th>
<th>Recombinants‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>HME52(thrA::Tn10)</td>
<td>cc</td>
<td>3.8 x 10⁵</td>
</tr>
<tr>
<td></td>
<td>cw</td>
<td>1.8 x 10⁵</td>
</tr>
<tr>
<td>DY374(nadA::Tn10)</td>
<td>cc</td>
<td>6.9 x 10⁶</td>
</tr>
<tr>
<td></td>
<td>cw</td>
<td>3.8 x 10⁶</td>
</tr>
<tr>
<td>HME53(tpcC::Tn10)</td>
<td>cc</td>
<td>1.0 x 10⁵</td>
</tr>
<tr>
<td></td>
<td>cw</td>
<td>0.3 x 10⁵</td>
</tr>
<tr>
<td>HME54(cysI::Tn10)</td>
<td>cc</td>
<td>0.1 x 10⁵</td>
</tr>
<tr>
<td></td>
<td>cw</td>
<td>5.1 x 10⁵</td>
</tr>
<tr>
<td>HME55(metC::Tn10)</td>
<td>cc</td>
<td>0.4 x 10⁵</td>
</tr>
<tr>
<td></td>
<td>cw</td>
<td>1.3 x 10⁵</td>
</tr>
</tbody>
</table>

*Cells were cultured for 15 min at 42°C and electroporated with 200 ng of the appropriate oligo containing 35 bases of wild-type sequence to each side of the indicated Tn10 element (16).
†Oligos are labeled as clockwise (cw) referring to the DNA strand whose 5' to 3' sequence proceeds in the clockwise direction with respect to the genetic map of the E. coli chromosome, and counterclockwise (cc) for the opposite strand (see Fig. 3).
‡Total prototrophic recombinants per 10⁸ viable cells following electroporation.

(16). For each mutant, two complementary 70-base oligos that included 35 bases of wild-type sequence on each side of the Tn10 insertion point were tested for efficiency of recombination (see Materials and Methods). We detected a strand bias at each of these sites, ranging from a 2- to 50-fold difference in recombination efficiency between the two complementary oligos used (Table 2). The strand bias was reproduced at each locus in three independent experiments (data not shown). As was the case for recombination at galK, the preferred strand at each of these genes had no correlation with the direction of transcription. Instead, strand preference correlated in every case with the direction of DNA replication through the gene being tested. Replication forks elongate in both directions from the origin (ori; Fig. 3), and the oligonucleotide that recombines more readily corresponds to the leading strand DNA (Table 2 and Fig. 3).

Discussion

Strand bias resulting from DNA replication suggests that ssDNA recombination may occur near the replication fork. The process of replication results in transient regions of ssDNA that may be accessible to Beta-mediated annealing of the introduced oligo. The increased recombination efficiency of the “lagging strand” oligos may reflect the increased frequency of single-stranded regions during lagging versus leading strand synthesis (23). DNA polymerase and DNA ligase may complete joining of the annealed strand to the chromosome. We observed rather substantial differences between the loci in both the degree of strand bias and the absolute number of recombinants (see Table 2). Our data do not provide an explanation of these differences; however, they may reflect changes in the structure or rate of replication of the chromosomal regions.

The result demonstrating that a single base change in galK is repaired at the same rate as a large insertion at the same position raises another important question that has not been resolved. How does the 70 base ssDNA used to repair these two mutations find its complementary sequence in these two examples? One might expect a gap created by discontinuous synthesis to reveal a 70-base segment to which Beta could anneal the complementary 70-base oligo. However, in the insertion mutant, the 35-base homology segments that flank the insertion DNA would be 3.3 kbp apart. How does Beta anneal to these separated 35-base targets at the same efficiency as it does to one 70-base target? Because the efficiency is the same, we must assume the two targets are not independent annealing events. Therefore, we assume either that the gap is very long, including the entire 3.3 kbp to allow direct annealing, or that there are multiple gaps spanning this distance and that the limiting event is finding the first complementary sequence. Most studies would support the latter of these two possibilities (23). Alternatively one can imagine that Beta does not depend on primer and can stimulate strand invasion of the ssDNA into dsDNA. This model, however, ignores the in vitro result that strand invasion (or displacement) occurs only after the ssDNA has been first annealed to an adjacent gapped region (16, 23).

Red-mediated ssDNA recombination provides a rapid and highly efficient approach for generating recombinant DNA molecules. Because this process is highly efficient, one can directly screen for mutations in the absence of selection (24). The examples here involve chromosomal genes; however, it is also possible to create mutations in genes cloned on plasmids. Swaminathan et al. (24) have used this Red-mediated recombination system to create mutations within the murine Bcr2 gene contained on a bacterial artificial chromosome (BAC). In this case, the high rate of recombination made it possible to use a PCR-based screen to identify recombinants.

Recombination with oligos can also be used to create deletions. Although we have not made extensive comparisons, deletion of an insertion cassette at galK by oligo recombination occurred with the same high efficiency as a single base change. Note that deletion of the different Tn10 insertion elements described here is also very efficient and again is high enough to permit screening for desired deletions in the absence of selection. For the nadA::Tn10 insertion, recombinants that deleted the insertion made up 6% of cells that survive electroporation. The ability to delete individual genes or chromosome segments efficiently and precisely makes Red-mediated recombination with oligos a very useful tool for systematic deletion analysis of the E. coli genome. This system is useful for analysis of other bacterial genomes. Red recombination has already been adapted for use in Salmonella (N. P. Higgins, unpublished observations), and it seems reasonable to propose that expression of the Red functions, or related proteins, may mediate recombination by oligos in other bacteria. Note that, because bet is the only gene required for ssDNA recombination, appropriate shuttle vectors containing the bet gene could be used to export this recombination system to other species.

Thirty years ago the discovery of restriction enzymes led to revolutionary changes in molecular biology and to the advent of genetic engineering. However, homologous recombination is a more precise, efficient, and versatile means of engineering DNA molecules (4, 25). The precision by which recombinant DNA molecules can be created using restriction enzymes and DNA ligase is limited by the availability of unique and correctly positioned restriction sites. In vivo recombination can generate recombinant molecules without the requirements for unique or special sites. Moreover, the products engineered by recombination are generated directly in the chromosome or episomal plasmid without intermediate cloning steps. Engineering recombinant DNA molecules by in vivo homologous recombination, a method we will call “recombinering,” has been confined to yeast (25), and only recently has it been fully exploited in E. coli (4, 6). Here, we have described an extension of recombinering using ssDNA that we suggest can be adapted to modify the genomes of diverse organisms.

Recombination of oligos is not unique to E. coli. Short single-stranded oligos have been shown to recombine in yeast with the CYC1 gene. As we have observed for E. coli, oligo recombination in yeast demonstrated a strand bias that appeared unrelated to transcription (26, 27). Although the direction of replication through CYC1 was unknown, the authors speculated that this bias might reflect preferential annealing of the oligo to either the leading or lagging strand during replication. The yeast
functions required for this recombination could not be determined. Although yeast Rad52 has been suggested to be functionally and structurally similar to λ Beta (28), Rad52 function was not required for oligo recombination (29), possibly because a second Beta-like function is present.

Oligo-mediated recombination occurs in both E. coli and yeast, suggesting that recombination via annealing of ssDNAs may occur in a wide range of organisms. Thus, recombineering with ssDNA may be applicable to higher eukaryotes as well, perhaps by a mechanism as simple as overexpressing Beta or a functionally similar ssDNA annealing protein, or more directly by introducing the protein itself into the eukaryotic cell during electroporation. We believe that homologous recombination with ssDNA will provide valuable technology for studying bacterial pathogens, and may open new avenues to engineer mutations or markers in the chromosomes of eukaryotic cells. Because oligos are so efficiently incorporated in vivo, the potential also exists to cross special chemical adducts and modifications from the oligos directly into the chromosome, thus providing special “tags” in the DNA of living cells.

We thank N. Costantino, E.-C. Lee, H. Peters, J. Sawitzke, K. Sergueev, S. K. Sharan, and H. Wilson for discussions and advice. We also thank T. Baker for technical help and M. Mills for preparing the manuscript and editorial comments. T. DiTizio is presently an undergraduate at Hood College in Frederick, MD.