A Plausible Mechanism for Gene Correction by Chimeric Oligonucleotides†

Howard B. Gamper, Jr.,*,‡ Allyson Cole-Strauss,‡ Richard Metz,§ Hetal Parekh,‡ Ramesh Kumar,§ and Eric B. Kmiec*,‡

Department of Biological Sciences, University of Delaware, Newark, Delaware 19716, and Kimeragen, Inc., Newtown, Pennsylvania 18940

Received September 21, 1999; Revised Manuscript Received February 10, 2000

ABSTRACT: Self-complementary chimeric oligonucleotides that consist of DNA and 2′-*O*-methyl RNA nucleotides arranged in a double-hairpin configuration can elicit a point mutation when targeted to a gene sequence. We have used a series of structurally diverse chimeric oligonucleotides to correct a mutant neomycin phosphotransferase gene in a human cell-free extract. Analysis of structure−activity relationships demonstrates that the DNA strand of the chimeric oligonucleotide acts as a template for high-fidelity gene correction when one of its bases is mismatched to the targeted gene. By contrast, the chimeric strand of the oligonucleotide does not function as a template for gene repair. Instead, it appears to augment the frequency of gene correction by facilitating complex formation with the target. In the presence of RecA protein, each strand of a chimeric oligonucleotide can hybridize with double-stranded DNA to form a complement-stabilized D-loop. This reaction, which may take place by reciprocal four-strand exchange, is not observed with oligonucleotides that lack 2′-*O*-methyl RNA segments. Preliminary sequencing data suggest that complement-stabilized D-loops may be weakly mutagenic. If so, a low level of random mutagenesis in the vicinity of the chimera binding site may accompany gene repair.

A self-complementary oligonucleotide that folds into a double-hairpin configuration can act as a template for gene repair of a point mutation when synthesized with a chimeric DNA and 2′-*O*-methyl RNA backbone of sufficient length (see Figure 1 for representative chimeras). Correction of genetic mutations by chimeric oligonucleotides was first described in 1996 (*1*−*3*), and recent work (*4*, *5*) has confirmed that the corrected sequence is genetically stable. Yet to-date, the frequency of gene correction using these agents remains variable, ranging from nondetectable to as high as 40% (*6*−*10*). This variability and the lack of a clearly defined mechanistic pathway have engendered controversy regarding the use of chimeric oligonucleotides in human gene therapy. In a previous study (*11*), we showed that a chimeric DNA/2′-*O*-methyl RNA oligonucleotide (Figure 1A) corrected a point mutation in the neomycin phosphotransferase (*neo*) gene of pK^{s}m4021 with a frequency of 0.1%. Correction was achieved in a cell-free extract (*12*) prepared from human hepatoma HuH-7 cells after which recovered plasmid was electroporated into *E. coli* for detection of kanamycin-resistant (kan$^+$) colonies. Here we have designed and tested additional chimeras in the cell-free extract. Our results indicate that gene correction may involve a complement-stabilized D-loop, a four-stranded joint molecule in which each strand of the chimera is hybridized to a complementary strand of the target DNA. Sequence analysis of kan$^+$ clones suggests that gene repair is accompanied by a low level of

FIGURE 1: Structure and sequences of chimeric oligonucleotides used in this study. (A) Chimera **1** and its target site in the *neo* gene of pK^{s}m4021. With the exception of the mutator bases, the targeting region of this and related chimeras (**1**−**19**) was homologous to a specific site in the *neo* gene that codes for premature termination of translation. In the presence of cell-free extract, introduction of the targeted G-C to C-G mutation at position 4021 restored expression by converting an amber (TAG) codon to a tyrosine (TAC) codon. DNA residues are capitalized, and 2′-*O*-methyl RNA residues are in lower case letters. Nonhomologous bases are boxed. (B) Chimera **23** has been used to correct the sickle mutation in the β-globin gene of human lymphoblastoid cells (*2*). The targeting region of this and related chimeras (**20**−**25**) was homologous to a portion of the wild-type β-globin gene. Some of these chimeras formed stable joint molecules with a homologous 46 bp DNA.

† This work was supported by NIH Grant RO1 HL58563-01Al (E.B.K.).

* Address correspondence to either of these authors at the Department of Biological Sciences, 122 Wolf Hall, University of Delaware, Newark, DE 19716. Telephone: 302-831-3221. Fax: 302-831-8786. E-mail: ekmiec@udel.edu.

‡ University of Delaware.

§ Kimeragen, Inc.

10.1021/bi9921891 CCC: $19.00
Published on Web 04/18/2000

nonspecific mutagenesis probably attributable to the joint itself.

Pairing of a chimera with a dsDNA[1] target would probably require recombinase-mediated catalysis. Without the assistance of other enzymes such as helicases or nucleases, such a reaction is problematic. Chimeras are highly stable double hairpins that contain one strand composed primarily of 2′-*O*-methyl RNA nucleotides. While it is known that *E. coli* RecA protein can catalyze the formation of RNA−DNA hybrid from complementary single strands (*13*), there are no reports that it catalyzes strand exchange between a single-stranded RNA and a double-stranded DNA. Equally unproven is the ability of RecA protein to catalyze pairing and strand exchange of two double-stranded homologous DNA molecules. While RecA protein can bind and align two homologous duplexes (*14*−*19*), initiation of strand exchange appears to be a three-strand event requiring that one of the molecules have a single-stranded overhang or gap homologous to the end of the other duplex (*20*). Furthermore, double-strand break repair by RecA has recently been inferred to occur by three-strand exchange (*21*). We investigated whether RecA protein could utilize a chimeric oligonucleotide in a strand exchange assay with dsDNA. Strand exchange occurred and formation of a joint was dependent upon the presence of 2′-*O*-methyl RNA nucleotides in one strand of the chimera. We suggest that chimeric oligonucleotides can search out and find a target gene by homologous recombination. Elucidation of the mechanism of strand exchange may reveal new aspects of recombinase action.

MATERIALS AND METHODS

Nucleic Acid Substrates. pK^{s}m4021 was derived from pWE15 (Stratagene) by introducing a T→G change at position 4021 to inactivate the *neo* gene. Chimeras were synthesized on 1000 Å controlled pore glass supports using DNA and 2′-*O*-methyl RNA phosphoramidite nucleoside monomers. Deprotected chimeras were purified to 98% homogeneity by HPLC and stored in water at −20 °C. Single-stranded oligonucleotides were purified by denaturing PAGE. Double-stranded 46 bp substrate for the strand exchange reactions was prepared by annealing two complementary 46-mers, one of which had been 5′ end-labeled by T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP. The resulting duplex was electrophoresed in a nondenaturing 12% polyacrylamide gel at 8 °C, released into solution by shaking the excised band overnight at 8 °C in Tris−EDTA buffer, and concentrated by passage through a Nensorb-20 column (Dupont). After drying in a Speed-Vac (Savant), the DNA was dissolved in water and stored at −20 °C.

Assay for Gene Correction in a Cell-Free Extract. Chimera (1.5 μg for a 68-mer) was incubated 45 min at 37 °C with pK^{s}m4021 (1.0 μg; molar ratio 360 to 1) and HuH-7 cell-free extract (30 μg of protein; for preparation, see *11*) in 100 μL of 20 mM Tris, pH 7.4, 15 mM $MgCl_2$, 0.4 mM DTT, and 10 mM ATP. The reaction was terminated by phenol extraction, and the plasmid DNA was recovered by ethanol precipitation. Analysis of the gene repair event relied on a bacterial readout system. Plasmid (100 ng) recovered from the cell-free extract was transfected into competent *E. coli* cells (strain DH10B) by electroporation (400 V, 300 μF, 4 kΩ) in a Cell-Porator apparatus (Life Technologies). After a 60 min recovery period in SOC medium, the bacteria were serially diluted, and 100−200 μL aliquots were spread on 10 cm plates containing solid media with ampicillin (100 μg/mL) or kanamycin (50 μg/mL). Usually, 100 and 200 μL aliquots of the transfected bacteria were added to the kan plates, while 100 μL aliquots of 2×10^3, 4×10^3, and 8×10^3 dilutions of this same suspension were added to the amp plates. By using these relative volume additions, the plates contained less than 1000 antibiotic-resistant colonies following overnight growth at 37 °C. Colonies were counted electronically using an Accucount 1000 instrument (Biologics). Gene correction frequency for each chimeric oligonucleotide was determined from at least three independent experiments by calculating the average ratio of kan$^+$ to amp$^+$ colonies obtained from a set amount of bacteria. In some experiments, individual kan$^+$ *E. coli* colonies were cultured overnight under amp selection. Plasmid DNA was purified, and the sequence of the *neo* gene was determined using a Perkin-Elmer 310 Genetic Analyzer.

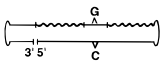
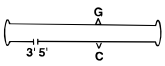
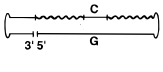
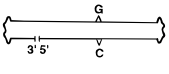
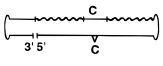
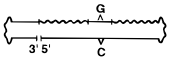
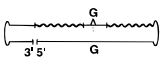
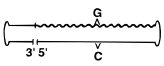
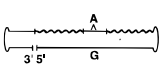
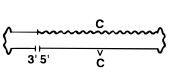
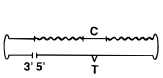
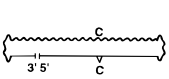
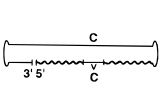
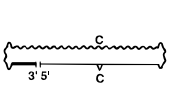
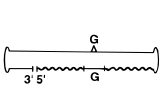
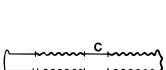
Strand Exchange Reaction. Chimeric oligonucleotide (80 nM) was incubated for 10 min at 37 °C with RecA protein (6.8 μM; nuclease-free; U. S. Biochemical) in 25 mM Tris−OAc, pH 6.8, 1 mM DTT, 1 mM $Mg(OAc)_2$, and 1 mM ATPγS. After raising the [$Mg(OAc)_2$] to 5 mM, a gel-purified, ^{32}P-end-labeled 46 bp duplex was added at 20 nM to the reaction mixture, and incubation (final volume 10 μL) was continued for 10 min. Reactions were stopped with 1% SDS and stored at −20 °C prior to analysis. Aliquots were electrophoresed 2 h at 8 °C in a nondenaturing 12% polyacrylamide gel containing 1 mM $MgCl_2$. Joint molecules were detected by autoradiography of the dried gel and quantitated using a phosphorimager. The structure of the joint molecule was verified by co-electrophoresis with a stable Y-arm joint formed by hybridization of non-self-complementary oligonucleotides. In separate reactions, RecA protein was incubated 20 min at 37 °C in strand exchange buffer with radiolabeled single-stranded 46-mer, double-stranded 46-mer, or chimeric oligonucleotide **20**. Analysis of the nucleic acid on a sequencing gel failed to detect any loss of label or degradation.

RESULTS

Gene Correction Assay. We have previously described a human cell-free extract that catalyzes gene repair by chimeric oligonucleotides (*11*). The extract was prepared from HuH-7 cells by a modification of the protocol described by Li and Kelly (*12*). Extracts similar to this have been used to study replication, recombination, and repair of DNA (*12, 22, 23*). Chimeric oligonucleotides used in this assay were designed to correct an amber point mutation previously introduced into the *neo* gene of pWE15, a plasmid normally conferring both ampicillin and kanamycin resistance. Correction was monitored by electroporating plasmid recovered from the cell-free extract into *recA*$^-$ *E. coli* and scoring for kanamycin-resistant colonies. Bacterial strains deficient in homologous recombination or mismatch repair must be used for readout

[1] Abbreviations: dsDNA, double-stranded DNA; ssDNA, single-stranded DNA; HPLC, high-performance liquid chromatography; PAGE, polyacrylamide gel electrophoresis; DTT, dithiothreitol; ATPγS, adenosine-5′-*O*-(3-thiotriphosphate); SDS, sodium dodecyl sulfate; bp, base pair(s); amp, ampicillin; kan, kanamycin.

Table 1: Structure−Activity Relationships Observed for Gene Correction by Chimeric Oligonucleotides in a Human Cell-Free Extract[a]

DNA Strand Directs Correction			RNA Strand Augments Correction		
	Chimera	Relative Activity		Chimera	Relative Activity
1		0.70±0.06	9		0.34±0.01
2		~0	10		0.34±0.06
3		1.00±0.04	11		0.73±0.05
4		0.014±0.008	12		1.39±0.01
5		~0	13		1.67±0.09
6		1.00±0.17	14		1.66±0.13
7		0.019±0.009	15		1.90±0.20
8		2.03±0.33	16		0.05±0.02

[a] In the schematic drawings, DNA is represented by a smooth line and 2′-*O*-methyl RNA by a wavy line. Length is proportional to the number of bases. Mutator bases that are mismatched to the target sequence are denoted by a carat. All chimeras possessed the standard clamp depicted in Figure 1A except for chimera **15**, which contained a hairpin clamp the stem of which (partially denoted by a boldface line) was homologous in sequence to the target. Hairpin loops consisted of T_4 or U_4. Gene repair reactions in the cell-free extract were conducted in triplicate. Colony counts (kan^+/amp^+) obtained in the bacterial readout assay were 232/747 (chimera **1**), 0/642 (**2**), 426/970 (**3**), 3.9/624 (**4**), 0/798 (**5**), 423/962 (**6**), 2.7/326 (**7**), 386/434 (**8**), 118/790 (**9**), 125/836 (**10**), 223/696 (**11**), 329/540 (**12**), 422/570 (**13**), 204/280 (**14**), 373/444 (**15**), and 5.7/260 (**16**). In all experiments, 10^3 more bacteria were plated on kan plates than on amp plates. The gene repair activity of each double hairpin was normalized relative to chimera **3**, which elicited 440 kan^+ colonies per 10^6 amp^+ colonies.

since these bacteria do not catalyze targeted nucleotide substitution by chimeric oligonucleotides (*11*). In this assay, the frequency of gene correction was given by the ratio of kan^+ to amp^+ colonies.

DNA Strand of the Chimera Directs Gene Correction. Structures of the various chimeric oligonucleotides tested are shown in Table 1, together with their relative correction frequencies observed in the cell-free extract. While they shared the same homologous targeting region to the mutant *neo* gene in pK^sm4021, they varied with respect to identity of the mutator bases and distribution of 2′-*O*-methyl RNA in the backbone. The first set of eight demonstrates that a mismatch between the mutator base on the DNA strand of the chimera and the complementary base of the target sequence is necessary and sufficient for targeted mutagenesis. Hence, chimera **1**, with both mutator bases mismatched to the target (to form C-C and G-G mismatches), directed correction while chimera **2**, with no mismatches to the target, lacked activity. The mutator bases in chimeras **3** and **4** were chosen so that only one was mismatched to the target. Consequently, only the DNA strand in **3** (forming a C-C mismatch with the target) and only the chimeric strand in **4** (forming a G-G mismatch with the target) could specify a change in gene sequence. Of the two, only number **3** had significant gene repair activity. Chimeras **5** and **6** substantiated the dominant role played by the DNA strand. Both of these oligonucleotides were designed to restore activity to the *neo* gene by converting the amber (TAG) codon to a tyrosine (TAT) codon. The template for this correction was the chimeric strand in **5** (forming an A-G mismatch with the target) and the DNA strand in **6** (forming a T-C mismatch with the target). Of this pair, only number **6** exhibited activity. Finally, chimeras **7** and **8** were synthesized with the 2′-*O*-methyl RNA segments present in the lower strand instead of the top strand. Only the chimeric strand in **7** (forming a C-C mismatch with the target) and only the DNA strand in **8** (forming a G-G mismatch with the target) could mediate gene correction. Of this pair, only number **8** exhibited significant activity.

Initiation of gene correction probably involves heteroduplex formation between the DNA strand of the chimera and the complementary strand of the *neo* gene. Use of the DNA strand as a template for mismatch repair could then lead to a base pair change at position 4021 through the process of mismatch repair, a pathway previously implicated in chimera-mediated mutagenesis (*11*, *24*). Although mismatch repair can correct all possible mismatches, some are repaired more efficiently than others (*23*, *25*, *26*). For example, a C-C mismatch is a relatively poor substrate compared to a G-G mismatch; hence, chimera **8** may be twice as effective as chimera **3** for this reason.

Chimeric Strand of the Chimera Augments the Frequency of Gene Correction. The role of 2′-*O*-methyl RNA in chimera-mediated correction was investigated by testing oligonucleotides with varying lengths of this nucleic acid. An all-DNA construct was half as active (compare **9** with **1**), a chimera with a continuous RNA sequence in the upper strand was twice as active (compare **12** with **1**), and a chimera with RNA segments in both strands was 20-fold less active (compare **16** with **3**). Chimeric oligonucleotides containing additional 2′-*O*-methyl RNA in the hairpin loops or in the upper portion of the clamp did not exhibit enhanced correction frequencies (compare **10** with **9**, **11** with **1**, and **14** with **13**). Based on these comparisons, we conclude that the RNA content of the homologous targeting region modulates the frequency of gene correction. It is conceivable that homologous pairing and strand exchange of the chimera with target DNA might generate a complement-stabilized D-loop in which both strands of the chimera are hybridized to the target and that 2′-*O*-methyl RNA might stabilize such a structure. As long as one arm of the joint is DNA, gene correction can proceed by mismatch repair. In the context of this model, the enhanced activity of chimeras **3** and **6** over **1** might be due to the absence of a mismatch in the upper 2′-*O*-methylated RNA containing arm of the complement-stabilized D-loop, thereby enhancing stability of the joint.

Clamp Placement Does Not Greatly Alter the Frequency of Gene Correction. In the original design of the chimera, a

Table 2: Chimeric Oligonucleotides with Differing Placements of the Nick and Clamp Correct the *neo* Gene in a Cell-Free Extract[a]

Chimera	Relative Activity
3	1.00±0.04
17	1.10±0.01
18	0.82±0.02
19	1.00±0.06

[a] Table 1 describes the chimera representations. Reactions were conducted in triplicate. Colony counts (kan^+/amp^+) for each chimera were 426/970 (chimera **3**), 154/320 (**17**), 210/584 (**18**), and 187/424 (**19**).

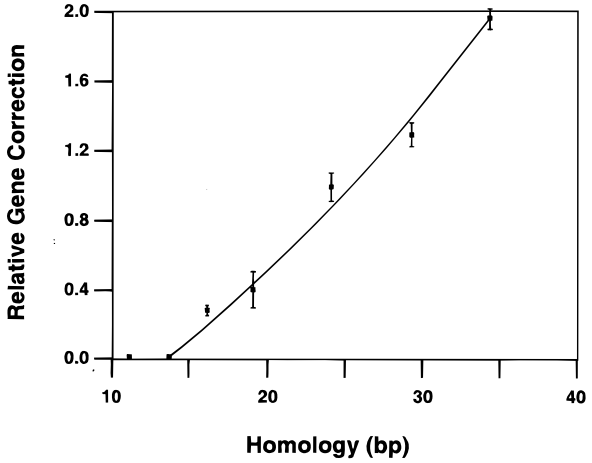


FIGURE 2: The frequency of gene correction by chimeric oligonucleotides is dependent on the length of shared homology with the target. A series of chimeric oligonucleotides derived from chimera **1** by altering the length of the homologous targeting region was evaluated for correction of the *neo* gene in pK^sm4021 using cytoplasmic HuH-7 extract. As the length of the targeting region was increased, the ratio of DNA to 2′-*O*-methyl RNA bases in the chimeric strand was kept constant. Colony counts (kan^+/amp^+) for each chimera were 2.9/824 (11 bp homology), 2.4/553 (13 bp), 62/680 (16 bp), 88/624 (19 bp), 232/747 (24 bp), 228/601 (29 bp), and 339/566 (34 bp).

nick was embedded in the duplex by extending the 3′ end of the chimeric strand of the oligonucleotide to form a 5 bp long G-C clamp bridged by four T's at the end. The nick was provided as a swivel to permit plectonemic coiling of the chimera with dsDNA target. It was recessed within the duplex to enhance thermal stability and nuclease resistance of the molecule. Because the G-C clamp does not bear homology to the target gene, its placement at either end of the homologous targeting region might be allowed. To investigate this possibility and determine whether the clamp could extend from either strand of the chimera, we tested gene repair activity of the chimeras shown in Table 2. In the cell-free extract, all four chimeras elicited similar levels of gene repair. These results support our supposition that the enhanced activity of chimera **8** is attributable to the type of mismatch formed with target DNA.

Length of the Homologous Targeting Region Influences the Frequency of Gene Correction. Figure 2 shows the relationship between the level of gene repair in the cell-free extract and the length of the homologous targeting region in the chimeric oligonucleotide. Detectable repair activity required at least 16 bp of homology, a length somewhat greater than that needed by RecA protein to promote homologous alignment of a ssDNA with a dsDNA (*27*). Chimera **1**, which shared 24 bp of homology with the target, was half as active as a longer chimera with 34 bp of homology. Synthesis of a chimera with a 6 bp clamp homologous to the target increased activity modestly relative to a similar chimera with a 5 bp G-C clamp (compare **15** with **14** in Table 1).

Table 3: Fidelity of Chimera-Mediated Gene Correction[a]

Chimera	Sequence of Codon 51 TAC	TAT	TTG	TGG	AGG[a]
3	5	-	-	-	-
4	1	1	1	1	-
6	-	5	-	-	-
7	-	1	1	-	1
8	5	-	-	-	-
9	5	-	-	-	-
14	5	-	-	-	-
15	5	-	-	-	-
16	3	1	1	-	-

[a] Two frame shift mutations generated this sequence:

CTA - TTC - GGC - TAG +G GAC ⟶ CTA - TCG - GCT - AGG - GAC (-T)

[a] Kan^+ colonies generated by treatment of pK^sm4021 with different chimeras were grown overnight in liquid culture containing ampicillin. Plasmid DNA was isolated, and a portion of the *neo* gene was sequenced. Codon 51 sequences generated by the chimeras are listed. TAC and TAT = Tyr; TTG = Leu; TGG = Trp; AGG = Arg.

Fidelity of Gene Correction. For some of the chimeras, plasmid DNA from up to five kan^+ colonies was sequenced to determine the fidelity of gene correction. The results of this analysis are summarized in Table 3. They show that chimeras **3**, **8**, **9**, **14**, and **15** directed a G-C → C-G conversion at position 4021 and that chimera **6** directed a G-C → T-A conversion at the same site. These conversions were specified by the mutator base on the DNA strand of the chimera and indicate a high degree of fidelity for the correction event. By contrast, the few kan^+ colonies generated by chimeric oligonucleotides **4**, **7**, and **16** displayed a mixture of mutations, most of which were not specified by the chimera. These mutations were chimera-dependent since the spontaneous mutation frequency in the gene repair assay (i.e., 0.28 kan^+ colony per 10^6 amp^+ colonies) was at least

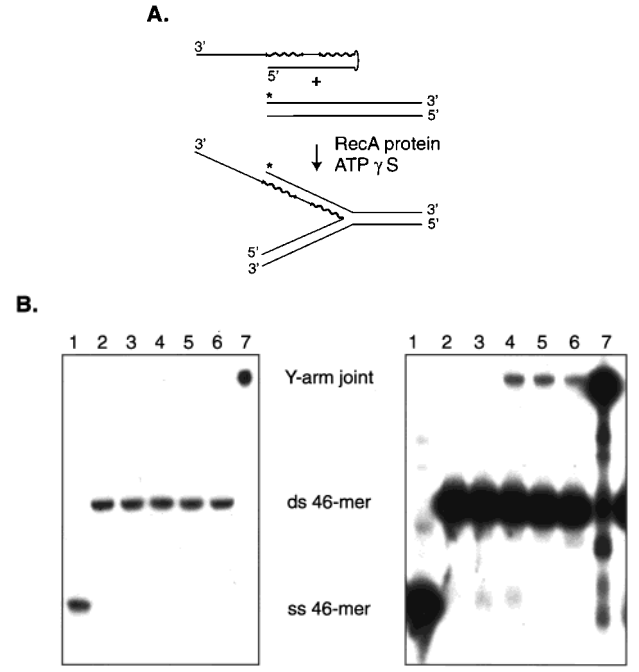


FIGURE 3: RecA-mediated alignment and strand exchange of chimeric oligonucleotide **20** with a homologous 46 bp DNA yields a complement-stabilized Y-arm joint. (A) Reaction scheme for strand exchange of chimera **20** with a ^{32}P end-labeled 46 bp duplex. One end of the 46 bp duplex was homologous to the 25 bp long targeting region of the chimera. (B) Y-arm joint formation: lane 1, single-strand 46-mer standard; lane 2, no chimera control; lane 3, no ATPγS control; lane 4, complete reaction terminated with SDS; lane 5, complete reaction terminated with SDS/$\mathbf{C}_{25}$; lane 6, complete reaction terminated with SDS/$\mathbf{W}_{25}$; lane 7, immobile Y-arm standard. $\mathbf{W}_{25}$ and $\mathbf{C}_{25}$ are single-stranded 25-mers complementary to the two strands that make up the stem of chimera **20**.

22-fold lower than that observed with any one of these oligonucleotides (see Table 1). We suggest that while these chimeras may be able to form complexes with target DNA, their structure is not compatible with the high-fidelity repair pathway used by the other chimeras. Instead, these complexes enter into a different mutagenic pathway that is both less efficient and template-independent. The extent of our data is insufficient to determine whether chimeric oligonucleotides with optimal structure enter into this pathway.

RecA Protein-Mediated Strand Exchange. To ascertain the feasibility of complement-stabilized D-loop formation in the absence of other reactions, we investigated whether RecA protein could catalyze strand exchange when presented with a chimeric oligonucleotide homologous to one end of a gel-purified 46 bp DNA duplex (Figure 3A). Chimera **20** had a 21 base long single-stranded 3′ overhang instead of a hairpin clamp to facilitate loading of RecA protein onto the molecule (*28*). The sequence of the overhang bore no homology to the β-globin sequence of the target, thus limiting strand exchange to the homologous targeting region. A two-step reaction protocol was employed. In the first step, chimera was incubated with excess RecA protein in the presence of 1 mM ATPγS and 1 mM $Mg(OAc)_2$ to form a presynaptic filament. Gel mobility shift assay verified the formation of a complex between RecA protein and chimera **20** (data not shown). In the second step, strand exchange conditions were established by adding radiolabeled 46 bp duplex and increasing the $Mg(OAc)_2$ concentration to 5 mM. Under these conditions, the RecA filament has been shown to bind two duplexes (*19*). If homologous alignment and strand exchange occurred within the filament, we expected the two substrates to form a complement-stabilized Y-arm joint as diagrammed in Figure 3A. After RecA protein was denatured by addition of SDS, the reactions were quick-frozen in dry ice and stored at −20 °C until analyzed by nondenaturing PAGE. By rapidly cooling the reaction mixtures and conducting electrophoresis at 8 °C in the presence of $MgCl_2$, we intended to reduce dissociation of the joint by branch migration.

Figure 3B presents two exposures of a representative set of strand exchange reactions analyzed by nondenaturing PAGE. In the left-hand panel, joint molecule formation was too low to detect following short exposure of the dried gel to X-ray film. However, in the right-hand panel, after a longer exposure joint molecule formation was readily seen. This joint was identified as a Y-arm structure due to its comigration with an immobile Y-arm standard (Figure 3B, lane 7) formed by hybridization of three non-self-complementary oligonucleotides. Consistent with a Y-arm structure, the deproteinized joint did not hybridize with 25-mers complementary to the upper or lower strands of the chimera (Figure 3B, lanes 5 and 6). Strand exchange was dependent upon the presence of chimeric oligonucleotide and ATPγS (Figure 3, lanes 2 and 3). Other controls showed that elimination of RecA protein or Mg^{2+} from the reaction also abolished joint molecule formation. Lowering of the pH from 6.8 to 6.2 did not significantly enhance strand exchange whereas replacing ATPγS with ATP eliminated the strand exchange product.

Phosphorimage analysis of the gel in Figure 3B indicated that only 2% of the labeled 46 bp target was incorporated into a Y-arm joint. Given the low efficiency of strand exchange, we considered alternative explanations for how the Y-arm joint might be formed. Despite having been gel-purified, the target duplex was contaminated with <1% single-stranded 46-mers. These oligonucleotides could form presynaptic filaments with RecA protein and initiate strand exchange with the chimeric hairpin. The resulting intermediate could hypothetically enter into subsequent strand exchange or hybridization reactions to generate a Y-arm joint. While we cannot rule out this pathway, there was no obvious relationship between the amount of Y-arm joint formed and the amount of single-stranded 46-mer remaining in the reaction mixture (in Figure 3B compare lanes 2 and 3 with lane 4).

Alternatively, hybridization of partially degraded oligonucleotides could give rise to joint molecules if, for example, the RecA protein was contaminated with a small amount of exonuclease. To investigate this possibility, we prepared three Y-arm joints in which the 5′ radiolabel was on the chimeric hairpin or on either strand of the 46 bp duplex. These joints were gel-purified and re-run on a sequencing gel adjacent to radiolabeled standards (Figure 4). The two strands of the 46 bp duplex that were recovered from the joint showed no signs of degradation. Analysis of the chimeric hairpin was somewhat complicated by its tendency to snap back to full-length or partial hairpin structures that ran ahead of the fully denatured molecule. Due to this phenomenon, only 60% of the counts in lanes 6 and 8 were found in the band corresponding to full-length single-stranded material. Overexposure of the standard chimera (see lane 7) showed a small amount of low molecular weight material, some of which

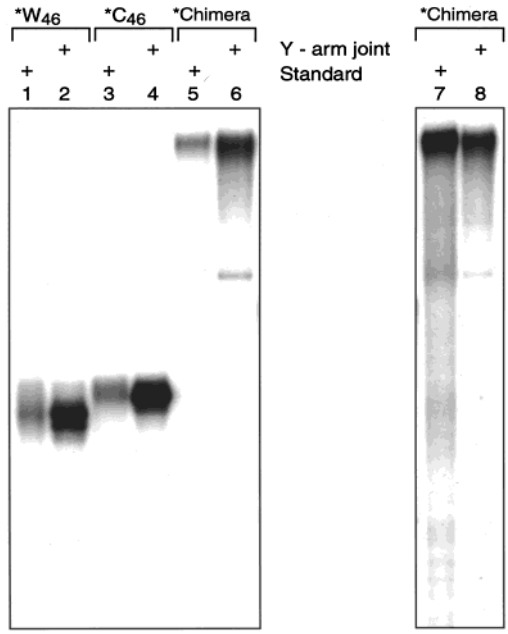


FIGURE 4: Y-arm joints are composed of three full-length oligonucleotides. Chimera **20** was paired with the 46 bp duplex (composed of $\mathbf{W_{46}}$ and $\mathbf{C_{46}}$ strands) in the presence of RecA protein. Three Y-arm joints were formed, each containing a different 5′-^{32}P end-labeled oligonucleotide ($\mathbf{W_{46}}$, $\mathbf{C_{46}}$, or chimera **20**). The joints were extracted from a nondenaturing 12% polyacrylamide gel, denatured in 90% formamide at 90 °C for 2 min, and run on a 7 M urea/12% polyacrylamide gel adjacent to end-labeled standards. Recovered oligonucleotides, which were purposely overexposed, comigrated with the standards. Note that both control and recovered $\mathbf{W_{46}}$ (lanes 1 and 2) were contaminated with $n - 1$ product. In lane 6, partially and completely reannealed chimera can be seen running ahead of the fully denatured molecule. When further exposed, the standard chimera showed the same bands superimposed upon lower molecular weight fragments (lane 7).

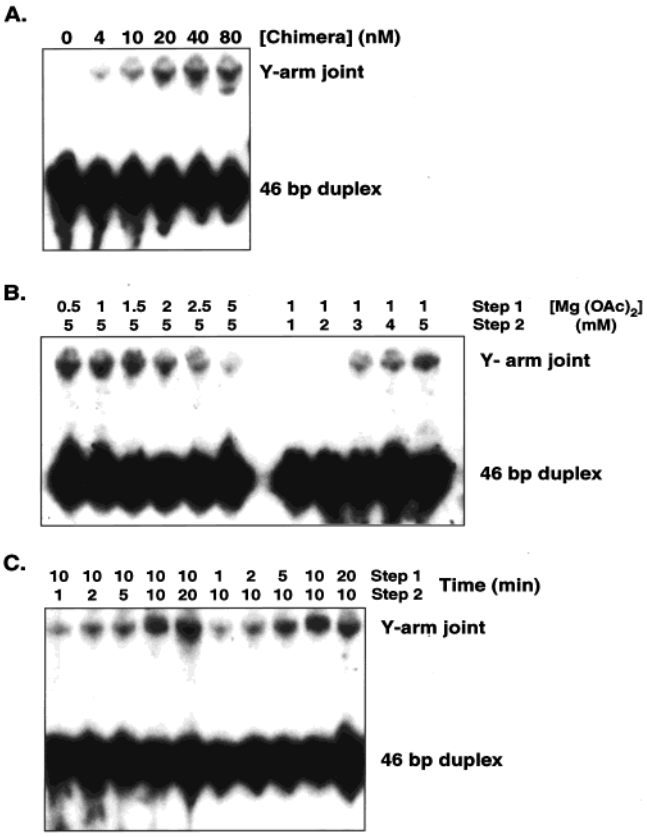


FIGURE 5: Characterization of the strand exchange reaction between chimera **20** and the 46 bp duplex. (A) Dependence of joint molecule formation on [chimera] when using 2 nM target. (B) Dependence of joint molecule formation on [$Mg(OAc)_2$] in steps 1 and 2. (C) Dependence of joint molecule formation on time of incubation during steps 1 and 2.

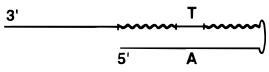
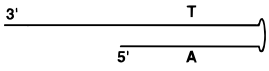
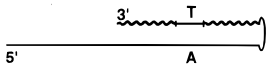
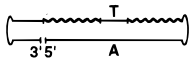
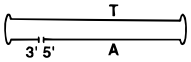
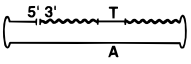
might have participated in joint molecule formation. In any case, the results presented here do not support the hypothesis that recombination was mediated by contaminating exonucleases. Indeed, the majority of Y-arm joints contained three full-length oligonucleotides.

Several parameters of the strand exchange reaction were investigated, and the associated autoradiograms are presented in Figure 5. For example, the yield of the Y-arm joint was proportional to the concentration of chimeric oligonucleotide (Figure 5A). As anticipated from earlier three-strand exchange studies (e.g., *29*), the binding and exchange steps were sensitive to [$Mg(OAc)_2$]. Presynaptic filament formation was optimal at Mg^{2+} concentrations <1.5 mM while strand exchange proceeded best at Mg^{2+} concentrations >5 mM (Figure 5B). The rates of presynapsis and synapsis were similar, with each reaction approaching completion after 10 min incubation at 37 °C (Figure 5C).

In later experiments we found that the yield of Y-arm joint could be significantly improved by cooling the strand exchange reaction to 4 °C prior to adding SDS. This modification of the protocol may stabilize nascent protein-free joint molecules released from synaptic complexes by SDS. Using the modified protocol, we tested several hairpin oligonucleotides for strand exchange activity (Table 4). These oligonucleotides shared the same homologous targeting region but differed with respect to whether 2′-*O*-methyl RNA segments were present in the upper strand and whether a single-stranded overhang or a hairpin clamp was placed at the 5′ or 3′ terminus of the molecule. When using the 46 bp duplex as a target for strand exchange, only the 25 bp long targeting region of these chimeras could interact homologously with the target.

Chimera **20**, with a chimeric upper strand and a 3′ overhang, was the most active in forming joint molecules. Chimera **23**, which possessed a chimeric upper strand and a 3′ hairpin clamp, was nearly as good. The equivalency of chimeras with a terminal overhang or hairpin clamp indicates that either structure is compatible with presynaptic filament formation. By switching the overhang or hairpin clamp to the 5′ end of the molecule, strand exchange was reduced by 60−70% (chimeras **22** and **25**). This differential activity is probably related to the polar nature of many RecA-mediated reactions. DNA constructs of chimeras **20** and **23** (numbers **21** and **24**, respectively) were inactive in the strand exchange assay, whereas a chimera with an upper strand composed entirely of 2′-*O*-methyl RNA exhibited slightly reduced activity in the same type of assay (data not shown). It thus appears that a limited amount of 2′-*O*-methyl RNA facilitates joint molecule formation by chimeric oligonucleotides. Previous studies have shown that 2′-*O*-methyl RNA residues also modulate the efficiency of exchange between single-stranded and double-stranded substrates (*30, 31*). In the strand exchange reactions reported here, the chimeras were completely homologous to the target. When the chimera and target molecules were mismatched at the mutator site as

Table 4: Structure−Activity Relationships for Complement-Stabilized Y-Arm Joint Formation[a]

	Chimera	Y - Arm Joint (% Yield)
20		11.6
21		0
22		4.3
23		10.3
24		0
25		3.1

[a] Strand exchange reactions between hairpin oligonucleotides and the 46 bp duplex were cooled to 4 °C before addition of SDS. The single-stranded overhang and hairpin clamp extensions on these molecules were heterologous to the target duplex. Table 1 describes the chimera representations.

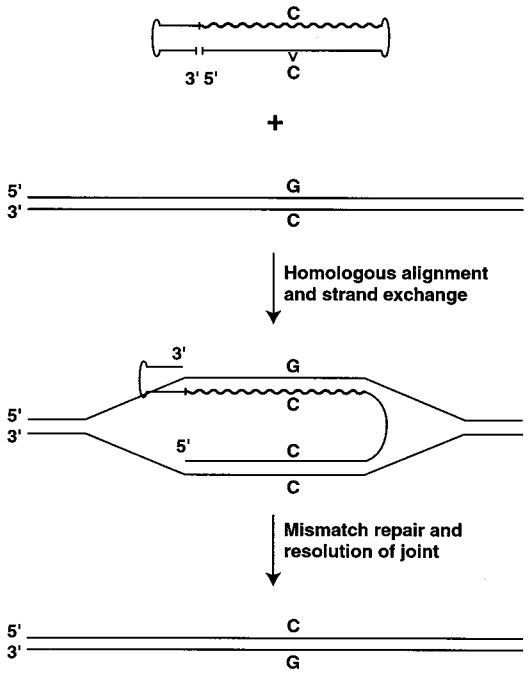


FIGURE 6: A complement-stabilized displacement loop may be a key intermediate in targeted mutagenesis by chimeric oligonucleotides. Scheme for gene correction of the mutant *neo* gene by a chimera with optimized structure.

would be the case in a gene correction experiment, strand exchange was not greatly reduced (data not shown).

DISCUSSION

A complement-stabilized D-loop appears to be a central intermediate in chimera-mediated gene repair (Figure 6). The ability of RecA protein to mediate joint molecule formation between a chimera and a short DNA duplex corroborates genetic evidence (*24*) that homologous recombination is an early step in chimera-mediated gene correction in bacteria and suggests that this reaction might also occur in eukaryotic cells. A complement-stabilized D-loop is significantly more stable than a simple D-loop (*32*) and is more likely to survive in the nucleus as a substrate for DNA repair. This intermediate provides a framework for understanding the structure−activity relationships obtained in the cell-free extract. Thus while the DNA strand of the chimera acts as a template for gene repair, the chimeric strand stabilizes an otherwise labile intermediate. In this context, a chimera with only the DNA strand mismatched to the target should be more efficacious than one in which both strands are mismatched, a result observed here. Aberrant mismatch repair of the joint is believed to mediate the mutation process since a human cell-free extract lacking hMSH2 protein exhibited a reduced level of chimera-mediated gene correction (*11*). We note that aspects of this pathway mimic the ability of eukaryotic cells to carry out homeologous recombination and mismatch repair of the resulting heteroduplex joint (*33*).

Sequence analysis of corrected plasmids calls into question the fidelity of gene repair by chimeric oligonucleotides. While the more active chimeras produced the expected mutation, those with low repair activity generated a high proportion of non-template-directed mutations. Although the ratio of repair activity for the two classes of chimeras was approximately 100 to 1, the absolute frequency of non-template-directed mutation might have been much higher since only those changes which restored activity to the *neo* gene would have been detected. Assuming that the latter mutations arose by an alternative pathway for processing the complement-stabilized D-loop, it is possible that chimeras with optimal structure for gene repair might enter into this pathway and produce unintended changes in the genome. Precedent exists for the induction of mutations by unusual joint structures in DNA (e.g., *34*). Further evaluation of the specificity of chimera-mediated mutagenesis will help determine their suitability as gene therapy agents.

The requirement for homology in the pairing step will probably limit this technique to the introduction of single base pair substitutions, additions, or deletions. In the extract, for example, insertion of a base pair is at least 10-fold less efficient than a substitution event (*11*). Pairing is also likely to be the rate-limiting step in the pathway since mismatch repair is known to be highly active in cell-free extracts (*23*, *25*). In the cell-free system used here, correction of the *neo* gene was 0.05% efficient. While this is orders of magnitude better than traditional gene targeting, it is still at least 100 times lower than the frequencies reported for chimera-mediated point mutation of genes in hepatocytes (*6*−*8*). The discrepancy could be explained by an alternative repair pathway not functional in the extract. If so, insights provided by the cell-free extract might not apply to gene repair taking place in cells.

RecA protein has not been shown to initiate strand exchange between two dsDNAs. While this recombinase can homologously align the two DNAs, initiation of strand exchange takes place between a ssDNA and a dsDNA (*16*−*19*, *20*, *35*). Once strand exchange has started, RecA protein can propagate the heteroduplex into adjoining duplexes via a Holliday junction. Previously, RecA protein has been

reported to catalyze formation of complement-stabilized D-loops when presented with long denatured restriction fragments homologous to an even longer duplex target (*36*, *37*). Here, however, both substrates were double-stranded in the region of shared homology, one due to a hairpin structure and the other due to prior gel purification.

Two mechanisms could account for joint molecule formation by the chimeric oligonucleotides. A three-strand exchange mechanism could be operative if RecA protein can strand-separate either the chimera or the 46 bp duplex. Once a single strand becomes part of a presynaptic filament, it could strand-exchange with a homologous duplex. Secondary hybridization or strand exchange reactions could convert the initial heteroduplex into a Y-arm joint. In 1985, Bianchi et al. (*38*) described strand separation of duplexes less than 30 bp in length by RecA protein. The reaction was dependent upon ATP and required that the duplex have a single-stranded overhang. Substitution of ATPγS for ATP reduced the extent of strand separation by 90%. In light of these results, a chimera with a single-stranded overhang or hairpin loop might be unwound by RecA protein as the first step in a three-strand exchange with the 46 bp duplex.

Contaminating single-stranded 46-mer could also initiate a three-strand exchange reaction with the chimeric hairpin acting as a target. This possibility is considered because nondenaturing PAGE analysis of the gel-purified target duplex showed that it was usually contaminated with 1−2% of single-stranded 46-mer. The single strand was generated during purification of the duplex and could not be avoided. Although we cannot exclude this pathway, there was no clear substrate−product relationship between the single-stranded 46-mer and the Y-arm joint. Following strand exchange, the amount of single-stranded 46-mer usually remained unchanged. In recent experiments that will be described elsewhere, we have observed that RecA protein can catalyze up to 75% joint molecule formation between two hairpin oligonucleotides. In these model reactions, one hairpin had a DNA backbone while the other had a chimeric DNA/RNA backbone with a single-stranded overhang. Homology between the two hairpins was restricted to the double-stranded stems. When using these substrates, no single-stranded DNA or RNA complementary to the stem sequences could be detected before or after strand exchange.

An alternative mechanism for joint formation would be a four-strand exchange between two duplexes homologously aligned in a RecA filament. Zaitsev and Kowalczykoski (*19*) have shown that two duplexes can be accommodated in a RecA filament in the presence of 1 mM ATPγS and $\geq$5 mM $MgCl_2$. Underwinding of supercoiled DNA by homologous dsDNA indicates that such duplexes can be homologously aligned by RecA protein (*16*−*18*). An analysis of strand exchange between partially single-stranded duplex and fully duplex DNA molecules supports a functional role for this duplex-by-duplex pairing (*14*, *15*). The short length of the chimera and the juxtaposition of ends may facilitate the strand exchange observed in this study. Dissociation of RecA protein from the synaptic complex by addition of SDS might yield a highly unwound heteroduplex that could collapse into a Y-arm joint or liberate free chimera together with ss or ds 46-mer. The first pathway might be favored by the presence of 2′-*O*-methyl RNA in the chimera. Ongoing experiments should differentiate between the two mechanisms for strand exchange.

The relative activity of the chimeras is different in the gene repair and strand exchange assays. This probably reflects differences between a complicated, multistep eukaryotic reaction (gene repair) and a simpler prokaryotic reaction (strand exchange). Although the DNA hairpins failed to yield joint molecules with the 46 bp target, recent experiments have shown that these molecules (e.g., **21** and **24**) can form joints with other substrates such as plasmid DNA. Eventually, we hope to use RecA protein as a tool to form complement-stabilized D-loops in plasmid DNA. These joints could then be used as substrates to dissect the pathway of mutagenesis in cell-free extracts or with purified enzymes. In parallel, we intend to investigate whether eukaryotic recombinases such as Rad51 can catalyze joint molecule formation by chimeric oligonucleotides as well as the role of 2′-*O*-methyl RNA in modulating the extent of strand exchange.

ACKNOWLEDGMENT

We thank Drs. K. Adzuma, P. Glazer, Y.-M. Hou, and S. Kowalczykowski for their comments on an early version of this paper.

REFERENCES

1. Kmiec, E. B. (1996) *Semin. Oncol. 23*, 188−193.
2. Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Gryn, J., Holloman, W. K., and Kmiec, E. B. (1996) *Science 273*, 1386−1389.
3. Yoon, K., Cole-Strauss, A., and Kmiec, E. B. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 2071−2076.
4. Alexeev, V., and Yoon, K. (1998) *Nat. Biotechnol. 16*, 1343−1346.
5. Santana, E., Peritz, A. E., Iyer, S., Uitto, J., and Yoon, K. (1998) *J. Invest. Dermatol. 111*, 1172−1177.
6. Kren, B. T., Cole-Strauss, A., Kmiec, E. B., and Steer, C. J. (1997) *Hepatology 25*, 1462−1468.
7. Kren, B. T., Bandyopadhyay, P., and Steer, C. J. (1998) *Nat. Med. 4*, 285−290.
8. Kren, B. T., Parashar, B., Bandyopadhyay, P., Chowdhury, N. R., Chowdhury, J. R., and Steer, C. J. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 10349−10354.
9. Stephenson, J. (1999) *JAMA 281*, 119−120.
10. Zhang, Z., Eriksson, M., Falk, G., Graff, C., Presnell, S. C., Read, M. S., Nichols, T. C., Blomback, M., and Anvret, M. (1998) *Antisense Nucleic Acid Drug Dev. 8*, 531−536.
11. Cole-Strauss, A., Gamper, H., Holloman, W. K., Munoz, M., Cheng, N., and Kmiec, E. B. (1999) *Nucleic Acids Res. 27*, 1323−1330.
12. Li, J. J., and Kelly, T. J. (1985) *Mol. Cell. Biol. 5*, 1238−1246.
13. Kirkpatrick, D. P., Rao, B. J., and Radding, C. M. (1992) *Nucleic Acids Res. 20*, 4339−4346.
14. Chiu, S.-K., Wong, B. C., and Chow, S. A. (1990) *J. Biol. Chem. 265*, 21262−21268.
15. Chow, S. A., Chiu, S. K., and Wong, B. C. (1991) *Biochimie 73*, 157−161.
16. Conley, E. C., and West, S. C. (1989) *Cell 56*, 987−995.
17. Conley, E. C., and West, S. C. (1990) *J. Biol. Chem. 265*, 10156−10163.
18. Lindsley, J. E., and Cox, M. M. (1990) *J. Biol. Chem. 265*, 10164−10171.
19. Zaitsev, E. N., and Kowalczykowski, S. C. (1999) *J. Mol. Biol. 287*, 21−31.
20. Cox, M. M. (1995) *J. Biol. Chem. 270*, 26021−26024.
21. Shan, Q., and Cox, M. M. (1998) *Mol. Cell 1*, 309−317.

22. Jessberger, R., and Berg, P. (1991) *Mol. Cell. Biol. 11*, 445−457.
23. Thomas, D. C., Roberts, J. D., and Kunkel, T. A. (1991) *J. Biol. Chem. 266*, 3744−3751.
24. Metz, R., DiCola, M., Avissar, P., Thomson, N., Primerano, J., and Kumar, R. (1998) *Conference Proceedings, 1st Annual Meeting of the American Society of Gene Therapy*, American Society of Gene Therapy, Seattle, WA, p 164a.
25. Holmes, J., Clark, S., and Modrich, P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5837−5841.
26. Muster-Nassal, C., and Kolodner, R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7618−7622.
27. Hsieh, P., Camerini-Otero, C. S., and Camerini-Otero, R. D. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6492−6496.
28. Pugh, B. F., and Cox, M. M. (1987) *J. Biol. Chem. 262*, 1326−1336.
29. Rosselli, W., and Stasiak, A. (1990) *J. Mol. Biol. 216*, 335−352.
30. Kotani, H. Germann, M. W., Andrus, A., Vinayak, R., Mulla, B., and Kmiec, E. B. (1996) *Mol. Gen. Genet. 250*, 626−634 (1996).
31. Havre, P. A., and Kmiec, E. B. (1998) *Mol. Gen. Genet. 258*, 580−586.
32. Kutyavin, I. V., Rhinehart, R. L., Lukhtanov, E. A., Gorn, V. V., Meyer, R. B., and Gamper, H. B. (1996) *Biochemistry 35*, 11170−11176.
33. Leung, W.-Y., Malkova, A., and Haber, J. E. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 6851−6856.
34. Wang, G., Seidman, M. M., and Glazer, P. M. (1996) *Science 271*, 802−805.
35. Kim, J.-I., Cox, M. M., and Inman, R. B. (1992) *J. Biol. Chem. 267*, 16444−16449.
36. Jayasena, V. K., and Johnston, B. H. (1993) *J. Mol. Biol. 230*, 1015−1024.
37. Sena, E. P., and Zarling, D. A. (1993) *Nat. Genet. 3*, 365−372.
38. Bianchi, M., Riboli, B., and Magni, G. (1985) *EMBO J. 4*, 3025−3030.

BI9921891